



















Interconnected pathways link faecal microbiota plasma lipids and brain activity to childhood malnutrition related cognition

Received: 21 March 2024

Accepted: 23 December 2024

Published online: 08 January 2025

 Check for updates

T. Portlock ^{1,10}, T. Shama ^{2,10}, S. H. Kakon ^{2,10}, B. Hartjen ^{3,4,10}, C. Pook ^{1,10}, B. C. Wilson ^{1,10}, A. Bhutto³, D. Ho ¹, I. Shannon ¹, A. M. Engelstad ^{3,5}, R. Di Lorenzo ^{3,4}, G. Greaves³, N. Rahman ³, C. Kelsey^{3,4}, P. D. Gluckman ^{1,4}, J. M. O'Sullivan ^{1,6,7,8} , R. Haque ^{2,3}, T. Forrester ⁹ & C. A. Nelson ^{3,4,5} 

Malnutrition affects over 30 million children annually and has profound immediate and enduring repercussions. Survivors often suffer lasting neuro-cognitive consequences that impact academic performance and socio-economic outcomes. Mechanistic understanding of the emergence of these consequences is poorly understood. Using multi-system SHAP interpreted random forest models and network analysis, we show that Moderate Acute Malnutrition (MAM) associates with enrichment of faecal *Rothia mucilaginosa*, *Streptococcus salivarius* and depletion of *Bacteroides fragilis* in a cohort of one-year-old children in Dhaka, Bangladesh. These microbiome changes form interconnected pathways that involve reduced plasma odd-chain fatty acid levels, decreased gamma and beta electroencephalogram power in temporal and frontal brain regions, and reduced vocalization. These findings support the hypothesis that prolonged colonization by oral commensal species delay gut microbiome and brain development. While causal links require empirical validation, this study provides insights to improve interventions targeting MAM-associated neurodevelopmental deficits.

Malnutrition is a significant global health issue responsible for an estimated 45% of all child deaths worldwide, making it one of the leading causes of mortality among children under the age of 5¹. Moderate acute malnutrition (MAM) is the most common and treatable form of malnutrition and is characterized by delayed growth, proportionate reductions in mass of most organs and tissues, and alterations in tissue architecture². MAM is also associated with neurocognitive impairments thought to result from structural and

functional changes to the brain³. Children who survive MAM are likely to suffer long-term consequences including impaired neurocognitive development, leading to long-term deficits in cognition and behaviour³. This leads to poor school performance and economic prospects as an adult⁴. Although much is known about the health, social, and economic ramifications of MAM, significant gaps in our knowledge remain. Nutritionally wasted children show significant brain atrophy on MRI⁵. Although re-feeding reverses this atrophy,

¹The Liggins Institute, University of Auckland, Auckland, New Zealand. ²Infectious Diseases Division, International Centre for Diarrheal Disease Research, Dhaka, Bangladesh. ³Division of Developmental Medicine, Department of Pediatrics, Boston Children's Hospital, Boston, MA, USA. ⁴Department of Pediatrics, Harvard Medical School, Boston, MA, USA. ⁵Harvard Graduate School of Education, Cambridge, MA, USA. ⁶Singapore Institute for Clinical Sciences, Agency for Science Technology and Research, Singapore, Singapore. ⁷The Maurice Wilkins Centre, The University of Auckland, Auckland, New Zealand. ⁸MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton, UK. ⁹Faculty of Medical Sciences, University of the West Indies (UWI), Kingston, Jamaica. ¹⁰These authors contributed equally: T. Portlock, T. Shama, S. H. Kakon, B. Hartjen, C. Pook, B. C. Wilson.

 e-mail: justin.osullivan@auckland.ac.nz; Charles.Nelson@childrens.harvard.edu

functional and microstructural deficits persist. One crucial gap is: how does MAM influence host physiology and cognitive development?

Many aspects of MAM, including host nutritional status, dietary intake, antibiotic administration, and infections impact the diversity, composition, and functionality of the microbiome^{6,7}. To this end, several studies in low- and middle-income countries (LMICs) have shown differences in faecal microbiome profiles between malnourished and well-nourished children^{8,9}. For example, a study in Bangladesh found that malnourished children, compared to well-nourished children, had enrichment of faecal *Bifidobacterium* and *Escherichia* species¹⁰. The development of the infant gut microbiome was shown to be slowed by malnutrition with this effect persisting after refeeding¹¹. Work using mouse models points to a possible causal role of the faecal microbiome in growth and weight gain, as mice colonized using faecal microbial transplantation with samples from malnourished children but not well-nourished controls showed impairments in weight gain and growth¹². Critically, perturbations of the faecal microbiome associated with MAM may have downstream consequences for brain and cognitive development^{13,14}. Given the rapid development that is occurring early in infancy and the fact that this overlaps a dynamic period of change in the microbiome, it is attractive to speculate that MAM-induced alterations to the gut microbiome impact brain structure and behaviours, directly or indirectly through the gut–brain axis¹⁵. Studies of well-nourished children living in upper-middle-income countries have shown that the faecal microbiome is associated with cognitive and brain development, although the direction of cause and effect remains unclear, with both increased and decreased faecal microbiota alpha diversity being linked to positive cognitive outcomes and neural development¹⁶. Alterations in the faecal microbiome may contribute to negative neurological outcomes observed in malnourished children, potentially through the disruption of nutrient absorption or the accumulation of toxic metabolites¹⁷. However, few studies have examined the link between the faecal microbiome and cognition in malnourished children^{11,18}. Therefore, more work is needed to understand how the faecal microbiome mediates the association between MAM and cognitive development.

Brain and behavioural development may be impacted by MAM indirectly through the circulating plasma lipidome^{19,20}. Several circulating plasma lipids, including cholesterol, phosphatidylcholines, phosphatidylethanolamine, and sphingolipids comprise 50% of the dry weight of the brain and have unique roles in neurological structure and function²¹. The brain relies upon nutrients circulating in the blood for its supply of resources. MAM impairs the blood brain barrier, which regulates accessibility of the brain to circulating lipids²². Microbially derived products (e.g. microbially encoded plasma lipids and e.g. bile acid derivatives²³), have been hypothesized to be key components in plasma lipidome-mediated communication between the faecal microbiome and the brain²⁴. This is supported by observations that lipids transformed and synthesized by the gut microbiome can influence structuring and signalling within host cells and tissues^{24–26}. For example, by influencing the functionality of membrane-embedded proteins through effects on membrane structure and function²⁷. Plasma lipid profiles have been shown to alter in children with MAM however, the host, microbial, and dietary contributions to these changes remain unresolved.

Given the range of impacts of MAM that are potentially mediated through the gut–brain axis, there has been increasing interest in understanding how its alterations are associated with impacts on infant neurocognitive development. The present discovery-based study examines the impact of MAM on the composition of the infant faecal microbiome, plasma lipidome, neural activity, and cognitive outcomes in a cohort of well-nourished and MAM 12-month-old Bangladeshi children. Interpreted Random Forest classifiers and co-abundance network analysis were used to understand multi-modal interactions that provide putative mechanistic insights into MAM-

associated developmental delays. Overall, this study provides important information about faecal–blood–brain–behaviour associations in children impacted by MAM. This information has the potential to contribute to the design of improved therapeutic methods and trials that test the efficiency of current therapies.

Results

Study population characteristics

Dhaka is the second most densely populated city in the world. Dhaka is the capital of Bangladesh; a country with the highest rate of childhood malnutrition globally²⁸. The Mirpur region was chosen as the site of our study to assess the impact of early-life malnutrition. Children with MAM ($n = 159$) and well-nourished controls ($n = 75$) at 12 ± 1 months of age with no history of chronic medical conditions, no known congenital anomalies, and no antibiotic use within the past month were recruited from the Mirpur region as part of a community-based, open-label, randomized clinical trial on the early emergence of executive function and development in Bangladeshi children using nutritional and psychosocial intervention (Fig. 1a)²⁹. MAM was defined according to WHO guidelines, using a threshold between two and three standard deviations below the mean z-score for weight-for-length/weight-for-height (WLZ/WHZ)³⁰. The MAM was associated significantly (MWU, p val < 0.05) with sociodemographic factors that included the principal toilet system used (septic-tank/toilet), vaginal delivery mode, and water treatment method (boiling) (Table 1, Supplementary Data 1, Supplementary Fig. 1).

MAM is associated with reduced Shannon diversity and enriched *Rothia mucilaginosa* and *Streptococcus salivarius*

Given the dominant role nutrition plays in driving faecal microbiome composition³¹, we sought to characterize and compare the faecal microbiome of MAM and well-nourished children to identify microbiome signatures associated with MAM. Stool metagenomes were extracted, sequenced (40.53 ± 8.5 million reads with no significant difference in read counts between MAM and well-nourished (MWU, $p = 0.71$), and profiled according to their species and functional compositions. Across all samples, 4 kingdoms, 18 phyla, 32 classes, 55 orders, 118 families, 319 genera, 1015 species, 611 functional pathways, and 2,828,874 gene families were detected.

For all children, there was a mean species richness of 56.1 ± 17.7 per sample and mean Shannon diversity of 2.96 ± 0.72 . MAM was associated with a lower Shannon diversity (MWU, $p = 0.025$) and Pielou's evenness (MWU, $p = 0.009$) than their well-nourished counterparts (Fig. 1e, Supplementary Data 2); consistent with previous observations in malnourished children³². The observed differences in alpha diversity were underscored by a significant compositional difference in the Bray–Curtis dissimilarity between the nutritional groups (PERMANOVA, $R^2 = 2.22$, $p = 0.008$, Fig. 1d). Significance for this difference was not observed with other measures of dissimilarity (Jaccard, Weighted-UniFrac, Unweighted-UniFrac, and Bray–Curtis) despite the significant correlation between all pairwise comparisons between them (Mantel $q < 0.05$, Supplementary Fig. 2).

Redundancy analysis (RDA) was used to investigate the significance of MAM in the context of other biologically relevant covariates and measure the extent to which these covariates constrain the compositional variance (Supplementary Fig. 3). MAM, delivery mode, biological sex, and duration of exclusive breast-feeding explained 2.27% of the total variance collectively with MAM status as the only factor that significantly explained the constrained variance ($p = 0.022$).

Taxonomic differences between MAM and well-nourished infants were calculated using Multivariate Association with Linear Models (MaAsLin2) whilst controlling for those covariates as fixed effects. A differential abundance of 5/286 taxa (2%) was identified (Fig. 1f, Supplementary Data 3). Specifically, the faecal microbiomes from MAM children had a greater prevalence and abundance of *Rothia*

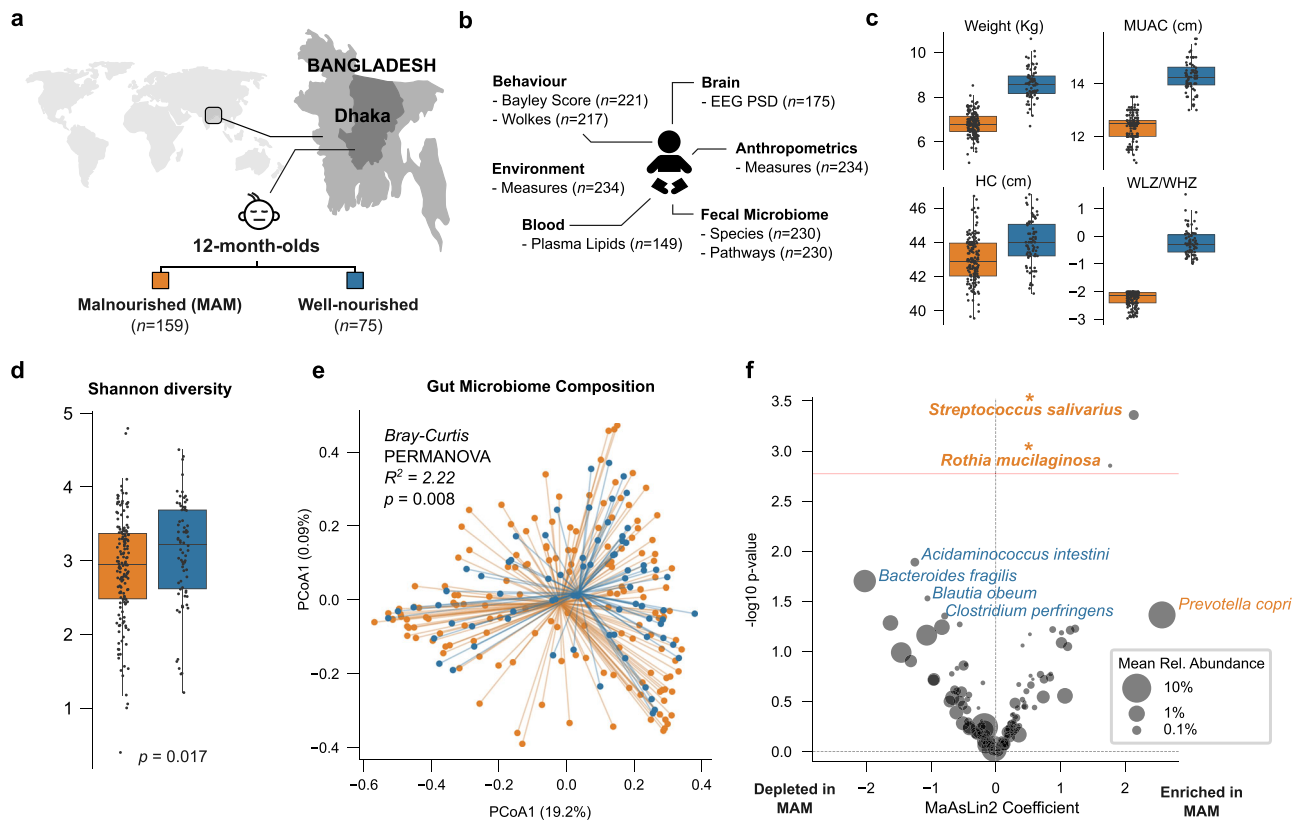


Fig. 1 | MAM impacts the 12-month-old faecal microbiome. **a** Schematic of study design. **b** Summary of data collected. Boxplots of significant differences (MMU, $p < 0.05$) in **c** anthropometric measurements between MAM ($n = 156$) and well-nourished controls ($n = 74$). **d** Boxplots of significant difference (MWU, $p = 0.01$) in alpha diversity between MAM ($n = 159$) and well-nourished controls ($n = 75$). Box plots show the median, the 25th–75th percentiles (IQR), and whiskers extending up to $1.5 \times$ IQR. **e** PCoA Scatterplot of Bray–Curtis beta diversities of samples (each marker is a single child's sample). **f** Volcano plot of differences in species relative

abundance (CPM) that are associated with MAM ($n = 159$) versus well-nourished controls ($n = 75$) measured using MaAsLin2 (two-sided and corrected for multiple comparisons). Red horizontal line signifies a significance threshold of $q = 0.25$, with point size representing the mean relative abundance across all 12-month-old children and *Streptococcus salivarius* ($q = 0.23$) and *Rothia mucilaginosa* ($q = 0.23$) being the only statistically significant results as signified by an asterisk. All subplots are coloured orange and blue representing MAM and well-nourished 12-month-olds, respectively.

mucilaginosa (MaAsLin2, coef = 1.77, $q = 0.23$), and *Streptococcus salivarius* (MaAsLin2, coef = 2.14, $q = 0.23$), when compared to well-nourished controls (Fig. 1f). These species were lowly abundant (mean relative abundance of $0.14 \pm 1.1\%$ and $0.004 \pm 0.019\%$ for *S. salivarius* and *R. mucilaginosa*, respectively) but prevalent (82.2% and 40.4%) of across all samples (Supplementary Fig. 4). Of note, whilst *Prevotella* and *Bacteroides* genera individually did not meet significance threshold, the *Prevotella/Bacteroides* ratio was elevated in MAM children (MAM/Well-nourished = 14.4, MWU $p = 0.020$). Of the other covariates, *Bifidobacterium bifidum* was observed to be enriched in those children that were born by vaginal delivery mode (MaAsLin2, coef = 1.77, $q = 0.23$).

Functional pathway analyses revealed no significant differences in the composition of the overall functionome between MAM and well-nourished controls (PERMANOVA, $R^2 = 8.76$, $p = 0.365$, Supplementary Fig. 5, Supplementary Fig. 7). There were no significant differences in the relative abundances (Supplementary Data 4) or completeness of individual pathways. MelonnPan was employed to predict the abundance of 81 faecal metabolites based on these functional gene abundance profiles. Subsequent statistical analysis using MaAsLin2 adjusted for confounders revealed no significant differences in the predicted metabolite concentrations.

MAM impacts brain activity and communication

Having identified the MAM-associated differences in the faecal microbiome community, the next phase of the investigation was to

look at the other half of the gut-brain axis. Resting-state electroencephalography (EEG) assessments of participants were performed to investigate the impacts of MAM on brain activity. Comparisons of EEG power spectral density (PSD) between children with MAM and well-nourished controls identified significant differences in the beta (12–30 Hz) and gamma (30–45 Hz) frequency bands distributed across temporal and frontal regions (Fig. 2a, b). These bands are associated with concentration, alertness, and higher mental activity and were observed to have higher amplitudes in the well-nourished children compared to those with MAM (Supplementary Data 5)^{33–35}.

Malnourished children often present with long-term impairments in neural and cognitive development³⁶. The Bayley Scales of Infant and Toddler Development Fourth Edition (BSID-IV) was used to assess development in cohort³⁷. When compared to well-nourished children and controlling for confounders, there was a significant reduction in Expressive Communication, Fine Motor, and Gross Motor Scores in the MAM children (MaAsLin2, mean difference (MAM-well-nourished) = -2.02 , -1.68 , -2.69 , $q = 0.021$, 0.021 , 0.047 , respectively; Fig. 2c, Supplementary Data 6).

To complement this method of assessing development, Wolke scoring was performed for the cohort (Fig. 2d, Supplementary Data 7). The Wolke Behavioural Rating Scale assesses children's socio-emotional development across five dimensions³⁸. As with the Bayley scoring, vocalization scores were disproportionately reduced in MAM children (MaAsLin2, mean difference (MAM-well-nourished) = -1.47 ,

Table 1 | Baseline child characteristics

	MAM (n = 159)	Well-nourished (n = 75)	pval
WLZ/WHZ	-2.24 ± 0.26	-0.23 ± 0.48	5.30e-35
MUAC (cm)	12.4 ± 0.49	14.26 ± 0.6	1.23e-34
Weight (kg)	6.8 ± 0.52	8.58 ± 0.69	1.78e-32
HC (cm)	42.98 ± 1.33	43.97 ± 1.35	9.98e-07
Principal type of toilet facility used by household members—septic tank or toilet	106/159 (66.7%)	68/75 (90.7%)	5.01e-05
Water treatment method—boil	70/159 (44.0%)	51/75 (68.0%)	0.001
Father's occupation—businessman	0/159 (0.0%)	6/75 (8.0%)	0.001
Principal type of toilet facility used by household members—pit latrine	40/159 (25.2%)	6/75 (8.0%)	0.002
Delivery Mode—vaginal	104/159 (65.4%)	35/75 (46.7%)	0.007
Number of years lived in current household	5.39 ± 6.28	3.83 ± 5.22	0.011
Toilet facility shared with other households	129/159 (81.1%)	49/75 (65.3%)	0.013
Mother's income taka	1425.79 ± 3108.0	600.0 ± 2046.75	0.016
Years of father's education	4.96 ± 3.67	6.45 ± 4.29	0.019
Father's occupation—daily labourer	38/159 (23.9%)	8/75 (10.7%)	0.021
Water treatment method—water filter	6/159 (3.8%)	9/75 (12.0%)	0.023
Type of cooking fuel—wood	22/159 (13.8%)	3/75 (4.0%)	0.023
Mother's occupation—housewife	117/159 (73.6%)	65/75 (86.7%)	0.028
Monthly total expenditure taka	14,854.4 ± 6388.11	18,469.33 ± 10,791.88	0.032
Language—Bengali	146/159 (91.8%)	74/75 (98.7%)	0.041
Language—Urdu	13/159 (8.2%)	1/75 (1.3%)	0.041
Principal type of toilet facility used by household members—water-sealed or slab latrine	13/159 (8.2%)	1/75 (1.3%)	0.041
Uses social media—false	109/159 (68.6%)	41/75 (54.7%)	0.042
Sex of child—male	82/159 (51.6%)	47/75 (62.7%)	0.123
Months of exclusive breastfeeding	5.3 ± 1.44	5.12 ± 1.71	0.336

Plus, minus values are mean values ± SD from continuous variables (±). *p*-values (*pval*) were calculated using two-sided MWU. All other variables are categorical (True/False) with their *p*-values calculated using Fisher's Exact test.

WLZ/WHZ weight-for-length/weight-for-height, MUAC mid-upper arm circumference, HC head circumference.

$q = 1.44e-9$) in addition to corresponding reductions to activity and approach scores.

MAM is associated with a reduction in circulating odd-chain fatty acids and ceramides

In addition to the connection between the gut and brain via the vagus nerve, the circulatory system serves as another important pathway for the bidirectional communication system. Adequate nutrition in children is characterized by concentrations of circulating lipids necessary for healthy growth and development³⁹. Uptake, transport, and metabolism of lipids from the diet and the microbiome are crucial processes during the developmental window. Therefore, untargeted LC-MS/MS was used to identify and quantify the levels of 792 plasma lipids in the children of the cohort (Fig. 3a).

MAM was associated with major changes (MaAsLin2: 317/792 (40%) $q < 0.25$, PERMANOVA: $R^2 = 10.31$) to the plasma lipidome after controlling for delivery mode, sex, and duration of exclusive breastfeeding as biologically relevant covariates (Fig. 3a, Supplementary Fig. 5, Supplementary Data 8, Supplementary Data 9). Of these changes, 128 (16%) compounds increased, and 189 (24%) decreased in relative abundance. Depletion in the relative abundance of seven lipids with diverse functions was observed, including those that are known to be specific to neurological development and function, such as ceramides, glycosphingolipids, and fatty amides. Several other plasma lysolipids from the lysophosphatidylcholine (LPC), and lysophosphatidylethanolamine (LPE) classes were depleted in the MAM children, as well as bacterially derived odd-chain fatty acids (OCFA). By contrast, oxidized glycerophospholipids, steroid conjugates, and glycerophosphoserines were observed to increase in relative abundance in MAM children (Fig. 3b).

Multimodal Random Forest classification of MAM reveals cross-mode influences

Having established the existence of changes associated with MAM across the faecal microbiome, brain, behaviour, and plasma lipids, the relative importance of changes in each of these domains for the prediction of MAM was measured. PERMANOVA was used to evaluate the difference in beta diversity with respect to all phenotypical covariates including MAM (Supplementary Fig. 5). Across all covariates, nutritional status alone significantly explained the variance in the greatest number of datasets.

To account for the non-linear associations of the features, separate Random Forest classifiers were trained, using either faecal microbiome taxonomic and functional profiles, neuroimaging (EEG PSD), lipidome and behavioural data (Bayley scores), to distinguish MAM from well-nourished children in the cohort (Fig. 4b). Within the predictors trained on individual feature sets the best predictors of MAM in 12-month-old children were: (1) plasma lipids (AUCROC = 0.93 ± 0.05); (2) brain and behavioural metrics (Wolke score, EEG PSD, Bayley score, AUCROC = 0.73 ± 0.05 , 0.71 ± 0.10 , 0.68 ± 0.07 , respectively) and (3) faecal microbiome taxonomic, functional, and predicted metabolite profiles (AUCROC = 0.56 ± 0.07 , 0.53 ± 0.07 , 0.52 ± 0.06 , respectively).

To account for inter-dataset interactions between features, multimodal models were trained on individually scaled and concatenated data from the faecal microbiome taxonomic and functional profiles, neuroimaging (EEG), plasma lipidome, and Bayley and Wolke scores. After including children that had a full dataset (MAM $n = 52$, well-nourished $n = 50$) and hyperparameter tuning, the models were highly predictive of MAM (AUCROC = 0.82 ± 0.05 , Fig. 4c). Shapley additive value (SHAP) interpretation of the multimodal models was performed to measure the

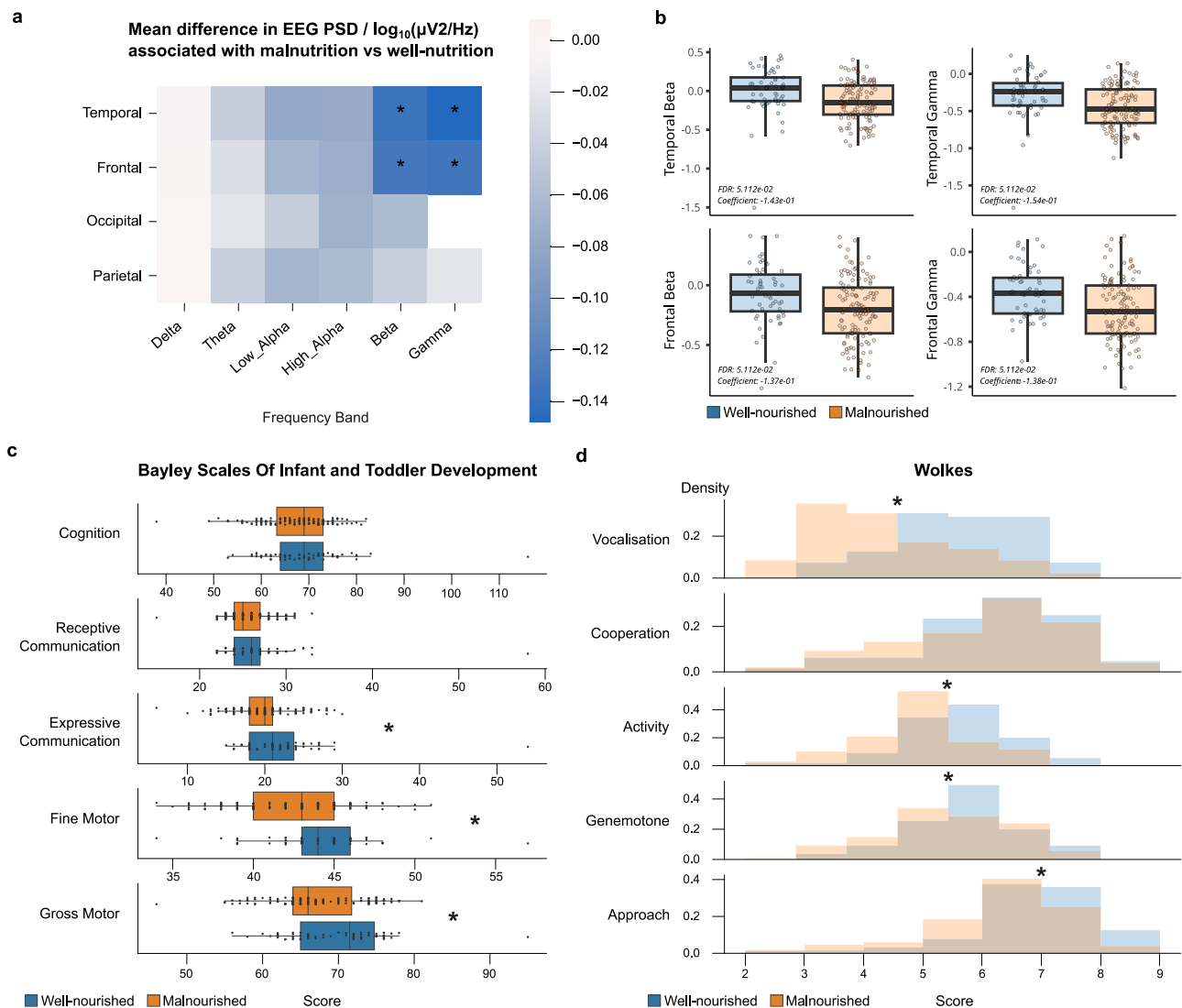


Fig. 2 | Differences in cognitive development of 12-month-old children associate with MAM. **a** Heatmap of lobe and frequency-specific changes in EEG resting state power spectral density (PSD) in MAM ($n = 120$) versus well-nourished ($n = 55$) 12-month-olds. PSD was measured in $\mu\text{Volts-squared per Hz}$ and \log_{10} -transformed for normalization. **b** Boxplots of distributions of significant EEG PSD changes as in subplot A. **c** Bayley score differences in MAM ($n = 151$) versus well-nourished

controls ($n = 70$). **d** Wolke score differences in MAM ($n = 152$) versus well-nourished controls ($n = 65$). $*q < 0.05$. Box plots show the median, the 25th–75th percentiles (IQR), and whiskers extending up to $1.5 \times \text{IQR}$. All differences were measured using MaAsLin2 (two-sided and corrected for multiple comparisons as is the default). All subplots are coloured orange and blue representing MAM and well-nourished 12-month-olds, respectively.

importance of each feature (Fig. 4d, Supplementary Data 10, Supplementary Fig. 8). Prediction importance was highly correlated with the MaAsLin2 significance (i.e. those features that changed significantly were likely to have a higher importance for the model prediction (Spearman correlation between $\text{mean}|\text{SHAP}|$ and $-\log_2(p)$ of $\rho = 0.51$). Interestingly, despite the low predictive performance of the microbiome in predicting MAM, when integrated into the multimodal model, the majority of important features were found to be within the faecal microbiome-derived datasets (taxonomic, functional, and predicted metabolite profiles) (Fig. 4e).

SHAP interaction values are an extension of SHAP values that provide insight into the interactions between pairs of features in a model. SHAP interaction values for each bootstrap of the multimodal models were calculated and averaged. A majority of important interactions ($\text{mean } \Sigma(\text{SHAP interaction}) > 5.017e-5$ (90% quantile)) were observed between those microbiome-derived datasets and the other datasets (Fig. 4f, Supplementary Data 11).

A multimodal predictive network analysis reveals the importance of *Bacteroides fragilis* in infant neurocognitive development

Network analysis is a useful tool to understand complex systems that emerge from interactions between multiple components. To better understand the relationship between the important features (i.e., $\text{mean}|\text{SHAP}| > 0.0016$ or 90% percentile), their co-abundance was calculated, and a relationship network was constructed (Fig. 5a). Spearman correlation of the features that were important in predicting MAM were calculated and filtered by significance ($\rho = 0.51$, $q < 0.05$) (1052/3906 correlations, Supplementary Data 11).

Important features were more likely to be significantly correlated ($q < 0.05$) with one another (Fig. 5a) than unimportant features ($\text{mean}|\text{SHAP}| < 0.0016$). Leiden cluster analyses revealed that those features that were different between MAM and well-nourished children were positively correlated. A cluster of *B. fragilis*, pyruvate fermentation pathways, plasma ceramides, EEG PSD, and Expressive Communication was identified as being highly correlated with the

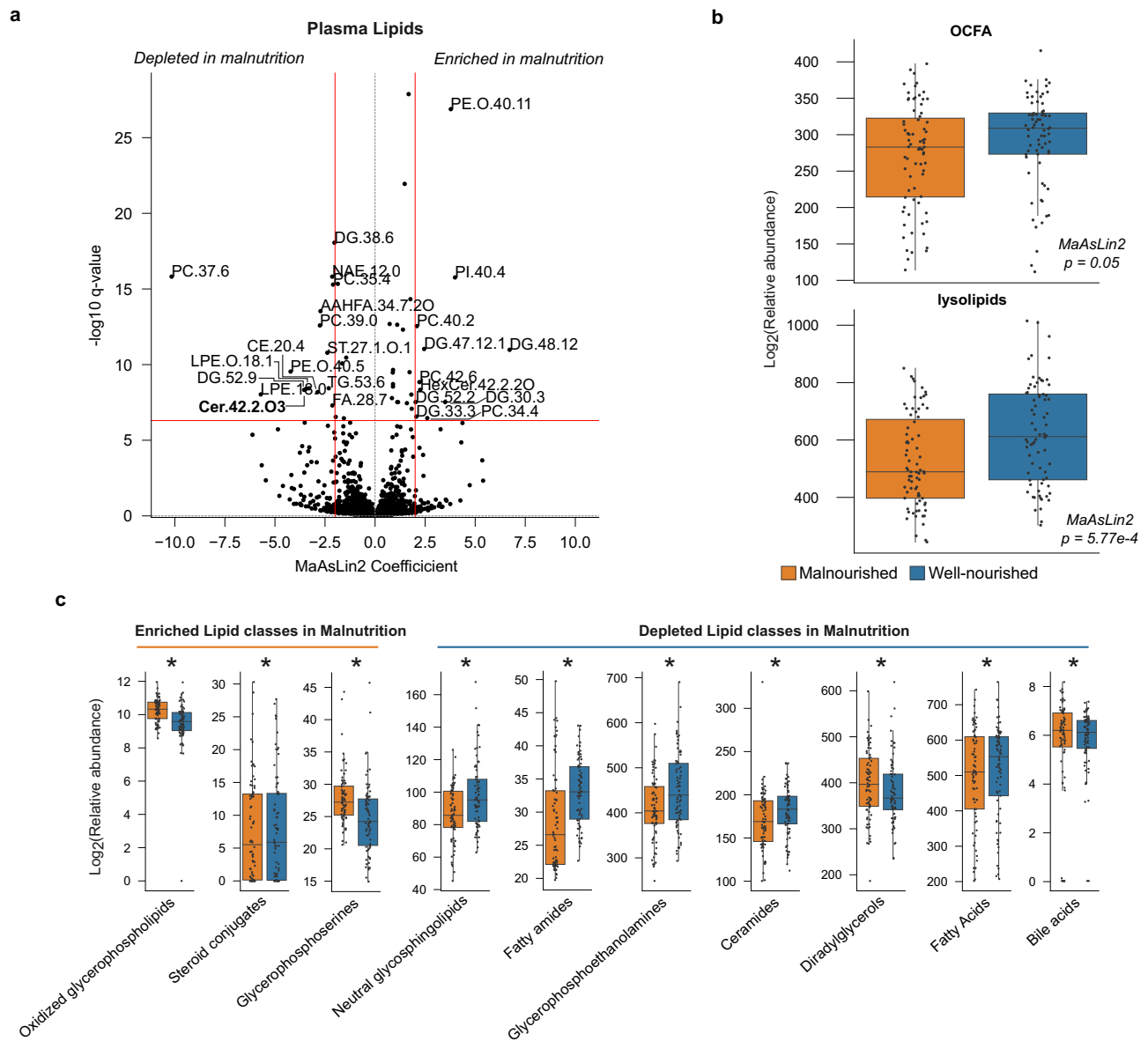


Fig. 3 | MAM results in major, compositional differences in plasma lipids of 12-month-old children. **a** Volcano plot of changes to plasma lipids between MAM ($n=76$) and well-nourished ($n=73$) 12-month-olds. (Upper left and upper right quadrants signify significant changes where the red horizontal line signifies $q < 0.05$ after mixed effect modelling adjusting for covariates. Vertical lines represent the mixed effect model coefficient of -0.1 and 0.1 , respectively). **b** Box plot of changes

in plasma odd-chain fatty acid (OCFA) and lysolipid relative abundance due to MAM. **c** Box plots of changes to lipid classes associated with MAM. * $p < 0.05$. Box plots show the median, the 25th–75th percentiles (IQR), and whiskers extending up to $1.5 \times$ IQR. All differences were measured using MaAsLin2 (two-sided and corrected for multiple comparisons). All subplots are coloured orange and blue representing MAM and well-nourished 12-month-olds, respectively.

well-nourished state (Fig. 5b). Plasma lipids that were depleted (MaAsLin2, $q < 0.05$, coef < 0) from the MAM infant samples were also negatively correlated with EEG PSD high-frequency amplitudes. Notably, EEG measurements were also correlated with bacterial pyruvate fermentation pathways and *B. fragilis* relative abundance. Although not significantly differentially abundant, *B. fragilis* was the 4th most abundant species amongst all children in the cohort (mean RA of 0.0487, 0.0748 in MAM and well-nourished, respectively). A cluster of *P. copri*, *S. salivarius*, *R. mucilaginosus*, glycolysis, peptidoglycan biosynthesis pathways, BCAA pathways, and plasma sphingomyelins was identified as being associated with the MAM condition. The strongest inter-species SHAP interaction effect was observed between *R. mucilaginosus* and *S. salivarius* indicating that their combined presence amplifies the prediction of MAM (mean $\Sigma(|\text{SHAP interaction}|) = 0.00021$).

Discussion

This study employed an interpreted random forest network approach to identify interconnected pathways between the faecal microbiome, plasma lipids, electroencephalogram power spectral density data, and behavioural outcomes from children with MAM and well-nourished controls. MAM was characterized by reduced alpha diversities and enrichment of specific microbial species. Differences in brain electrical activity and expressive communication between MAM and well-nourished children were observed. However, direct correlations between brain EEG and behavioural measures were not significant. Plasma lipidome analysis revealed significant differences in polyunsaturated fatty acids and, notably, ceramides, which are crucial for neural development. Integrated analysis across multimodal models identified non-linear associations between microbially-derived secondary bile acid deoxycholic acid, the microbiome, and behavioural

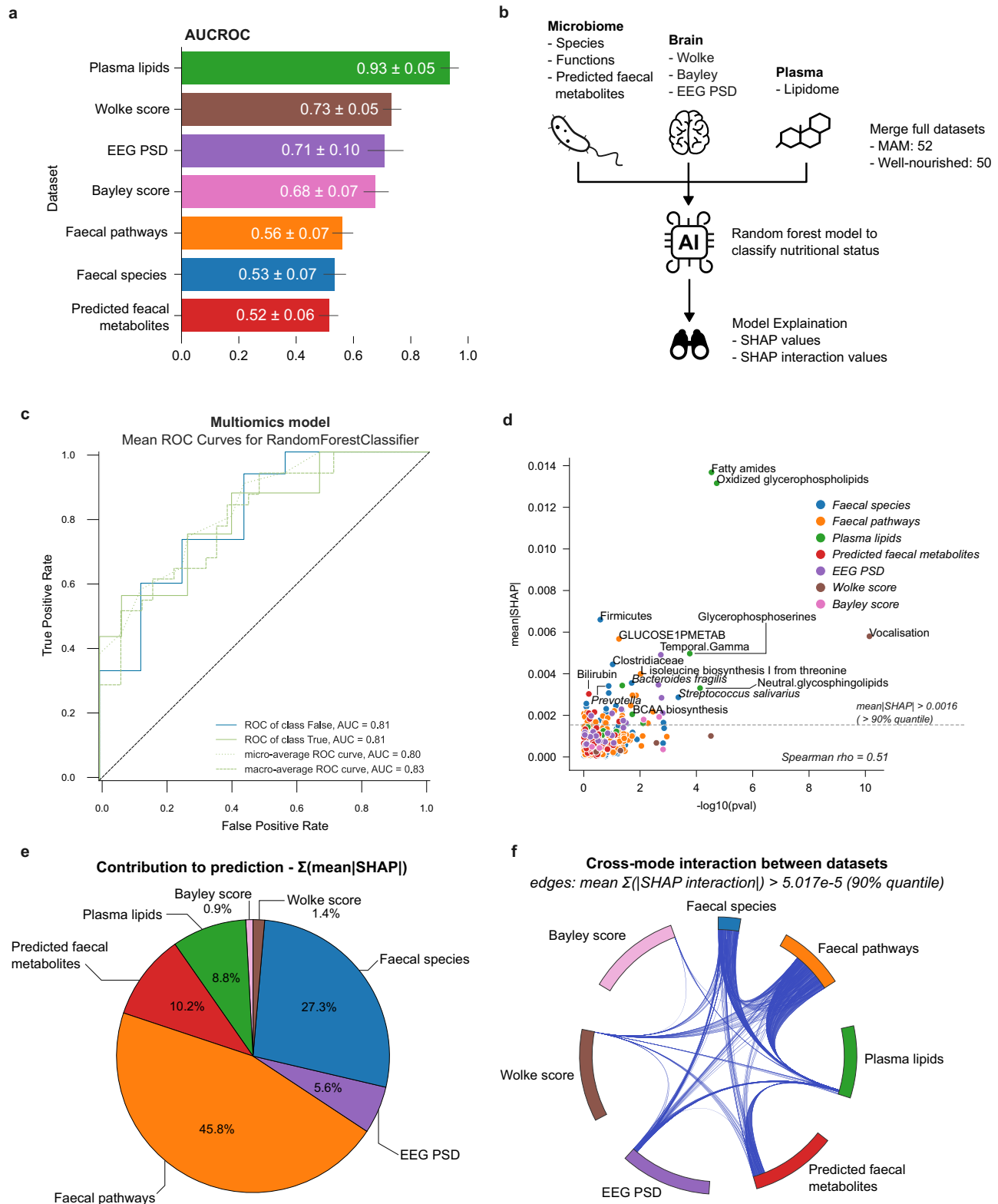


Fig. 4 | Integration of multimodal datasets influences the importance of features of random forest models predicting nutritional status. a Mean AUCROC values of the predictive power of each modal dataset on predicting nutritional status (error bars are standard deviation). **b** Schematic describing interpreted multimodal approach to predict MAM. **c** AUCROC curve of multimodal model prediction of MAM. **d** Scatterplot to show the relationship between mixed effect

model significance as $-\log(pval)$ and mean|SHAP| . **e** Pie chart of proportional contribution of each dataset to multimodal model prediction $\Sigma(\text{mean|SHAP|})$. **f** Circos plot of the contribution of interactions between dataset to MAM prediction. All subplots are coloured orange and blue representing MAM and well-nourished 12-month-olds, respectively.

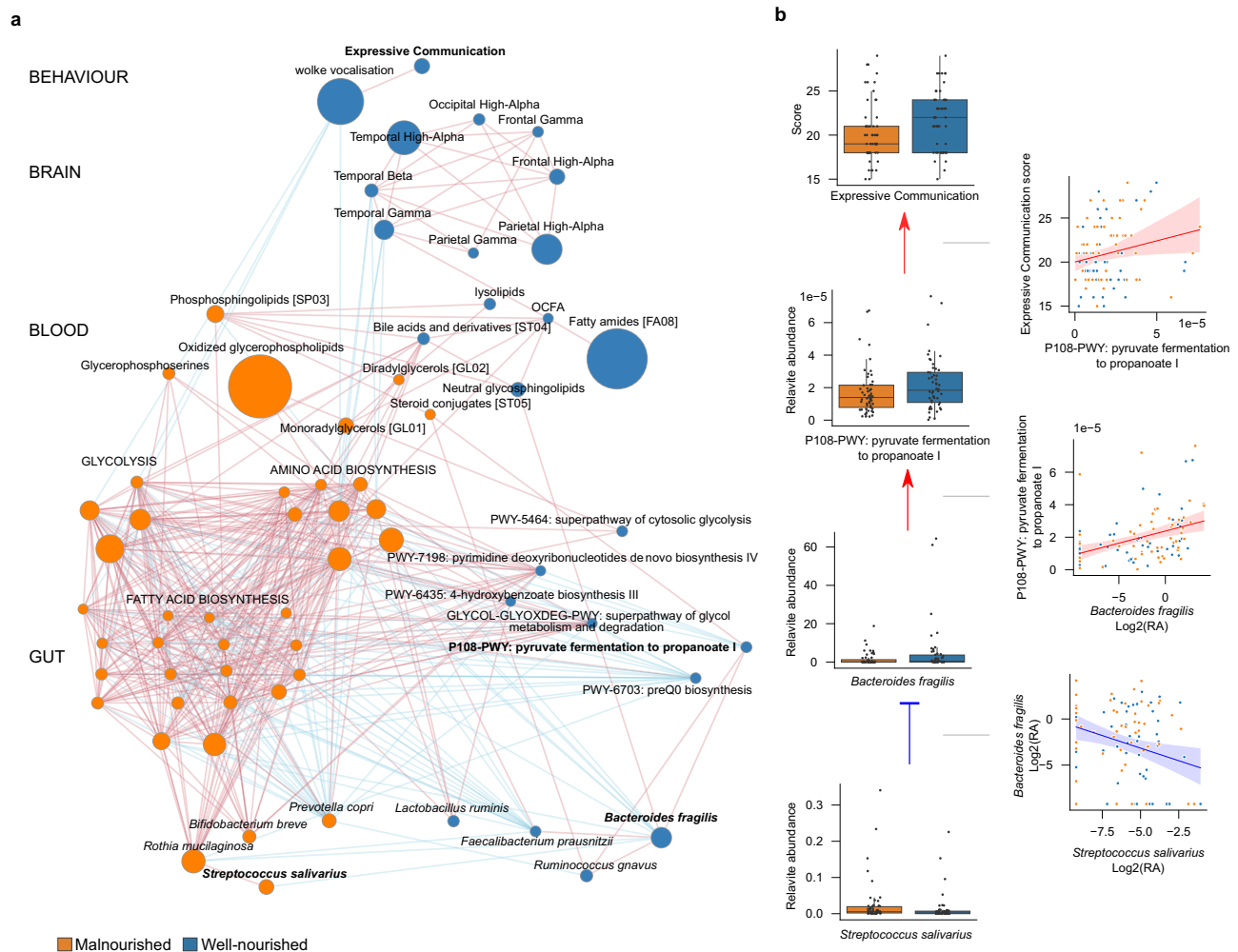


Fig. 5 | *Bacteroides fragilis* forms a network with propanoate synthesis, EEG PSD, and expressive communication that is anti-correlated with a *Rothia mucilaginosa*, *Streptococcus salivarius* focused cluster of features in MAM and well-nourished children. **a** Network illustrating inter-relationships of feature associations that predict MAM. Inclusion in the network requires a feature to be important in the prediction of MAM (mean|SHAP| of >0.0016, 90% quantile) and a significant Spearman correlation ($q < 0.05$). Nodes are features coloured by their enrichment (MaAsLin2) in MAM (orange and blue are enriched (MaAsLin2 coef > 0) and depleted (MaAsLin2 coef < 0 in MAM, respectively). Nodes are scaled proportionately to their mean|SHAP|. Edges

are Spearman correlations coloured red, and blue being positively and negatively correlated respectively. **b** Mechanism of MAM pathogenesis. Scatterplots show a correlation between features, with points coloured orange and blue representing children that are MAM and well-nourished, respectively. The shaded areas represent the 95% confidence intervals for predictions from a linear model. Boxplots are differences in each feature between MAM and well-nourished children. Box plots show the median, the 25th–75th percentiles (IQR), and whiskers extending up to $1.5 \times$ IQR. All subplots are coloured orange and blue representing MAM and well-nourished 12-month-olds, respectively.

outcomes, suggesting microbial signalling pathways impacting behaviour. *Bacteroides fragilis* abundance, linked to fermentation pathways, emerges as a predictive factor for well-nourished children. Causality between faecal microbiome, plasma metabolite changes, and MAM phenotype remains unclear, necessitating further research.

Several key differences between the faecal microbiomes of MAM and well-nourished children were identified. In line with prior research, it was found that MAM children had reduced species alpha diversities. The oral-commensal, Gram-positive, facultative anaerobes *S. salivarius* and *R. mucilaginosa* were enriched in MAM children. This class of species is also found enriched in sufferers of liver cirrhosis, hypothesized to be the result of bile acid dysregulation^{40,41}. In addition, MAM children had a higher P/B ratio. One possible interpretation for the differential *Prevotella* abundance in MAM children is that microbial development is accelerated in MAM children. This hypothesis is in part supported by the elevated abundances of *P. copri* and *Bifidobacterium adolescentis*, which are markers of weaning⁴². Alternatively, selective microbiome community-driven interactions might explain the inverse

correlations that were observed between *P. copri* and *Bifidobacterium* species. The higher P/B ratio also points to a possible depletion in *Bacteroides* which has been observed previously in Bangladeshi children³². Taken together, the findings provide evidence for an accelerated development of the microbiome–faecal–brain axis during childhood MAM, possibly due to mothers having to find alternative food sources for their children rather than relying solely on breastfeeding. However, it should be noted that, while work in adults has shown that the P/B ratio is shaped in part by diet and lifestyle (e.g., hygiene practices, living environment)⁴³ there is a lack of research on *Prevotella* rich microbiomes (previously referred to as the *Prevotella* enterotype) due to their underrepresentation in high-income countries⁴⁴.

Comparisons of the MAM and well-nourished children identified differences both in brain electrical activity and expressive communication that were associated with MAM. Severe acute malnutrition during childhood is generally recognized as having long-term effects on adult cognitive, academic, and behavioural development⁴⁵. Despite

this, our integrated analysis did not identify any significant direct correlations between brain EEG and behavioural measures (Wolke vocalization and Bayley expressive communication) in MAM 12-month-old children. When investigating differences in EEG power spectral density (PSD), disruptions were especially evident for higher frequency canonical frequency bands (high alpha, beta, and gamma) but not the lower frequency bands (delta, theta, or low alpha) in the frontal, temporal and occipital areas. Notably, we did identify connections between the biological markers (faecal microbiome, microbial pathways, and blood metabolome) and both brain (EEG PSD) and behavioural metrics (Bayley, Wolke).

We identified significant differences in plasma lipids between MAM and well-nourished children. These changes included both increases and decreases in polyunsaturated fatty acids (PUFA) with extended chain lengths (>18) or odd chain numbers. The observed changes may be due to the combined effects of reduced dietary intake (e.g. C16 and C18 derivatives) and altered microbial metabolism (e.g. odd chain FA) in the MAM condition. By contrast, the increased very long-chain fatty acids (>C20) likely reflect an alteration in host metabolism. This is consistent with observations from animal models that diets low in protein result in hepatic steatosis, loss of peroxisomes, and mitochondrial dysfunction⁴⁶. Peroxisomes are important for the β -oxidation of very long-chain fatty acids (>C20), branched fatty acids, xenobiotics, and bile acids. The low protein diet peroxisome and mitochondrial dysfunction are reflected in severe metabolic disruption, including to the levels of plasma lipids, consistent with our observations in the MAM individuals. The MAM-associated reduction in 16C and 18C chain FA may reflect the dietary deprivation. Notably, the plasma lipidomes of MAM children also exhibited significant differences in the levels of ceramides and lysolipids (i.e. lipid derivatives in which one or both acyl derivatives have been removed by hydrolysis).

Numerous specific changes in plasma lipid levels stand out as being potentially important for neural development. Firstly, lactosylceramide (hex2cer 34:1) is an essential precursor for the synthesis of all complex glycosphingolipids⁴⁷ that was depleted by approximately 50% in MAM children. Secondly, lysophosphatidylcholine (LPC) and lysophosphatidylethanolamine (LPE) are essential for brain development and growth as they carry fatty acids across the blood-brain barrier, via the major facilitator superfamily domain-containing protein 2A (Mfsd2a)⁴⁸. Phosphatidylcholine (PC) is a precursor to acetylcholine, an essential neurotransmitter for memory and cognitive function. Supplementing neuron differentiation medium with phosphatidylcholine reduces the impact of inflammatory stress and neuronal damage, increasing the number of healthy neurons and modulating neuronal plasticity⁴⁹. By contrast, PI 40:4 increased within the plasma of MAM children. PI 40:4 is a precursor to prostaglandin synthesis that is important for brain development⁵⁰. Previous work has found that microbial fatty acids support myelin biosynthesis and maintenance, and this function may be a possible explanation for the associations found between microbial fatty acid biosynthesis, plasma lipids, and electrical activity in the brain^{51–53}. However, further work is required to test this hypothesis.

Random Forest classification models trained on the fecal microbiome, neuroimaging data, and the plasma lipidome accurately predicted the MAM condition. Integrating the important features of these models and Spearman correlation using network analysis provided a holistic view of the MAM mechanism and highlights the potential importance of a subset of microbes (i.e. *Bacteroides fragilis*) as indicators of childhood neurocognitive and microbiome-faecal-brain axis development in MAM. The integrated analysis identified non-linear associations between circulating bile acids and the microbiome and indirect associations through phosphosphingolipids with behavioural outcomes (i.e. Wolke Vocalization and Bayley Expressive Communication scores). Bile acid concentrations in MAM children were

elevated when compared to well-nourished children. Bile acids are hormone-like molecules that have recognized, albeit poorly described, roles as signalling molecules to the brain⁵⁴. Secondary bile acids are synthesized directly by the microbiome and are detectable in the brain where there are numerous bile acid receptors⁵⁴. Notably, deletion of the Farnesoid X receptor (FXR), involved in bile acid homeostasis was associated with a reduction in depressive and anxiety-like behaviour, but increased motor activity. Collectively, these findings reinforce the hypothesis that there is microbial signalling, through the peripheral circulatory system to the brain that can impact behaviour. However, direct causal effects are yet to be demonstrated for the compounds we have identified⁵⁵.

Recent studies have emphasised the significant role of the faecal microbiome in mediating dietary effects on host physiology, in addition to its influence on the development and function of the nervous system^{56–59}. Multiomic analysis examined associations between childhood MAM, altered brain function, and the microbiome and suggested a mechanism that links the fermentation of pyruvate to butanoate and ceramide biosynthesis to brain function and language development. However, in the absence of causal animal studies, it remains unclear if the faecal microbiome, and plasma metabolite changes are a result of or contribute causally to the wider MAM phenotype. Additional limitations are linked to the structure of the cross-sectional cohort itself and lack of temporal samples, the correlative nature of the data, and the choice of measures (e.g. EEG has poor spatial resolution hampering precise localization of brain signal). Another limitation of this study is the use of relative abundance data, which can introduce compositional bias by distorting changes in microbial taxa and plasma lipids. Although we have applied approaches designed to mitigate compositional bias, such as statistical methods like MaAsLin2, which are widely accepted for handling compositional structures, some biases may remain. Future studies with absolute abundance data could provide more accurate insights.

Notwithstanding these limitations, the integrated dataset and incorporation of age-matched controls, provide evidence that is consistent with the hypothesis that, in infants with MAM, the persistence of oxygen-sensitive oral commensals *Rothia mucilaginosa* and *Streptococcus salivarius* in the gut microbiome hinders gut and neurodevelopmental maturation by out-competing the colonization of *Bacteroides fragilis*, a critical species for production of brain-supportive lipids. Under healthy conditions, *B. fragilis* replaces these oral commensal species as the gut matures, contributing to the synthesis of odd-chain fatty acids essential for brain development, which are associated with enhanced expressive communication. The impact of delays in the sequential colonization sequence could be tested in animal models and should be empirically tested.

Integrative multi-omics analysis highlights inter-connected pathways between *Bacteroides fragilis*, plasma OCFA, temporal/frontal beta/gamma EEG PSD, and vocalization. This primary study provides a testable hypothesis to determine the effects of interventions on MAM-associated behaviour and brain development.

Methods

Ethics

The protocol (registered on ClinicalTrials.gov, study ID number: NCT05629624) was co-developed by researchers from The University of The West Indies, Boston Children's Hospital, USA, International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), and University of Auckland, New Zealand. Ethical approvals were obtained from the Research Review Committee (RRC; August 21, 2021) and Ethical Review Committee (ERC) of icddr,b (protocol no.: PR-21084; September 21, 2021), University of Auckland, New Zealand (approval AH23922; for analyses of collected biological samples) and University of West Indies (CREC-MN.51, 21/22).

Study design and participants

This case-control study was performed on the baseline data from two cohorts of children who were enrolled (between February and December 2022) as part of the M4EFaD intervention, a community-based clinical trial within the Mirpur slum, Dhaka, Bangladesh²⁹. The cohort consisted of a control group of 75 well-nourished children at 12 ± 1 months (WLZ score > -1 SD), an intervention group of 159 children with $WLZ < -2$ and > -3 and/or $MUAC < 12.5$ and > 11.5 cm and having MAM at 12 ± 1 months. Inclusion criteria included an assessment of MAM, no history of chronic medical conditions, known congenital anomalies, and no antibiotic use within the past month.

Recruitment and anthropometric data collection

Enrolment was initiated on February 7, 2022, and will continue until February 2024. Study surveillance workers (SWs) conducted a door-to-door census (~100,000 households) in Mirpur DNCC wards 2, 3, and 5 between January and December 2022. Verbal consent was obtained to participate in the census. The census identified 5736 children aged between 11 and 13 months and 2314 children aged between 34 and 38 months. During the census, if the guardian verbally consented to the study procedure and the babies met the inclusion and exclusion criteria of the study for either the well-nourished control arm or the MAM arm, the SWs proceeded to measure the MUAC of the child. Mothers of babies who were within the MUAC range were invited to visit the icddr,b study clinic for further assessment and enrolment.

Final screening for eligibility and study consent occurred at the icddr,b Mirpur study clinic. Written consent was obtained from the parents or legal guardians of each child included in this study. The consenting process was tailored to each mother's literacy level and involved reviewing the inclusion and exclusion criteria. Comprehension of the study was assessed using scripted points and open-ended questions. For mothers or legal guardians who were illiterate, we involved an impartial witness to assist during the consent process. In such cases, we obtained the mother or the legal guardian's thumbprint along with the witness's signature on the written informed consent.

Following consent, the clinical screening team completed a screening form, capturing the date of enrolment, sex, date of birth (DOB), weight (kg), length/height (cm), head circumference (cm), and MUAC (cm) measurements of the child. The WLZ/WHZ score for each child was calculated using the WHO anthropometric calculator. The child's age was validated using the EPI vaccination card. Bayley scores and EEG data were collected upon enrolment to evaluate brain and cognitive development.

EEG data collection and analysis

Continuous scalp EEG was recorded using NetStation 4.5.4. and 128-channel Hydrocel Geodesic Sensor Nets modified to remove eye electrodes (Electrical Geodesics, Inc. (EGI), Eugene, OR, USA) (Supplementary Fig. 6). Data was sampled at 500 Hz. Impedances were kept under 100 k Ω when possible and measured once at the beginning of the session, and again halfway through. Sessions were conducted in a dimly lit room with the participants sitting on the parent's lap. The participants were separated from the research staff conducting the session by a curtain, but the testing area was not acoustically or electrically shielded. A second research staff member was present in the testing area to help keep the participant engaged. The subsequent (pre-)processing steps were applied to the resting state data where participants watched a 3-min video that featured toys.

EEG data were pre-processed with MatLab (R2021B) using the Harvard Automated Processing Pipeline for Electroencephalography (HAPPE) Version 3^{60,61}. A specified subset of 30 channels was excluded (E1, E8, E14, E17, E21, E25, E32, E38, E43, E44, E48, E49, E56, E63, E68, E73, E81, E88, E94, E99, E107, E113, E114, E119, E120, E121, E125, E126, E127, E128). Data were bandpass filtered (1–100 Hz) and filtered using a 50 Hz cleanline filter for line noise removal. Bad channels were then

automatically identified and rejected, and wavelet thresholding was performed to detect and impute artefacts. Resting-state data were segmented into 2 s epochs; epochs with an amplitude $> [150]$ mV were rejected. Segments were also rejected using segment similarity criteria. Data were then re-referenced to the average of all channels.

EEG outputs from HAPPE were then reformatted and processed using the batch electroencephalography automated processing platform (BEAPP)⁶² to extract power spectra for each participant across the following frequency bands: delta (2–4 Hz), theta (4–6 Hz), low alpha (6–9 Hz), high alpha (9–12 Hz), beta (12–30 Hz), and gamma (30–45 Hz) and the following regions of interest (Supplementary Fig. 6): occipital (E70, E71, E75, E76, E83), temporal (E36, E40, E41, E45, E46, E102, E103, E104, E108, E109), parietal (E52, E53, E59, E60, E85, E86, E91, E92), and frontal (E5, E6, E12, E13, E24, E27, E28, E33, E34, E112, E116, E117, E122, E123, E124). Further, PSD values were normalized by a \log_{10} transform.

Developmental outcomes

The Bayley Scales of Infant and Toddler Development, Fourth Edition (BSID-IV) cognitive, language, and motor subscales were administered to all participants³⁶. Research assistants were trained to research reliability in the administration and scoring of the Bayley-4. Due to cultural differences between Bangladesh and the United States where the assessment was developed, Bangladeshi researchers modified some assessment stimuli to improve cultural responsiveness and relevancy. Pictures for the item naming series and action naming series of the expressive language and receptive language subscales were adapted to include items that Bangladeshi children are more likely to be familiar with. For example, in an item depicting a child in a one-piece pyjama set requiring children to predict what the child was about to do, the pyjama set was replaced by bedtime clothing familiar to Bangladeshi children that would signify that the child was going to sleep. The Wolke ratings are a measure of development that has been validated in previous studies of nutritional impacts on behavioural development in Bangladesh^{38,63–65}.

Biological sample collection

Stool samples were collected from each infant at their home at the baseline visit. Samples were collected in DNA/RNA Shield Faecal Collection Tubes (Zymo Research, #R1101). Peripheral venous blood samples were collected in EDTA Vacutainers, separated into plasma and RBCs, and immediately frozen at -80 °C. Batches of blood and stool samples were air-freighted on dry ice from Bangladesh to the Liggins Institute, New Zealand for processing and analysis.

Microbiome DNA extraction, sequencing, and profiling

DNA was extracted from stool samples using the ZymoBIOMICS MagBead DNA/RNA extraction kit (Zymo Research, #R2136) following the standard protocol. Samples (1 mL) were mechanically lysed in bead-bashing tubes using the MiniG tissue homogenizer prior to extraction of DNA. 200 μ L of the sample was used post-bead bashing for extraction of DNA following the protocol. A volume of 50 L of elute was collected in DNase/RNase Free Water. Samples with a DNA concentration < 14.5 ng/L were re-extracted following the ZymoBIOMICS DNA extraction protocol. DNA concentration measurements were measured using the Nanodrop. Samples were shotgun sequenced (Illumina NovaSeq 150PE reads) to an average sequencing depth of 20 M read-pairs/sample across three batches with ZymoBIOMICS® Microbial Community Standard and water used for positive and negative controls, respectively, for each batch to rule-out batch effect and identify microbial contamination.

Raw sequencing reads were processed using BioBakery3 tools⁶⁶. Contaminant removal was performed with KneadData using paired-end inputs and the human genome reference (hg37). Quality trimming was done with Trimmomatic using the options "HEADCROP:15 SLIDINGWINDOW:4:15 MINLEN:50." Microbial community profiling was

conducted with MetaPhlan3 (Version 3.1, mpa_v31_CHO-COPHlan_201901 database), including viral taxa and unknown estimates via the ‘add_viruses’ and ‘unknown_estimation’ flags. Functional profiling was performed using HUMAnN 3.6 with default parameters, indexing with ‘uniref90_annotated_v201901b’. KneadData-filtered reads and taxonomic profiles were input, with gene families identified using the UniRef90 database and reads mapped to MetaCyc pathways. Functional profiles were renormalized to CPM using the ‘humann2_renorm_table’ function. Predicted stool metabolite profiles were generated from the gene profiles using MelonnPan using default parameters.

Plasma lipidomics

Plasma samples for lipidomics were thawed on ice in sets of eight bracketed by a procedural blank and quality control and extracted according to a method modified from ref. 67. Briefly, 10 L volume was placed in an amber glass autosampler vial, and 300 L of a mixture of Type 1 water, butanol, methanol, chloroform, and SPLASH Lipidomix in a ratio of 4:15:15:20:1 was added. The mixture was vortexed and sonicated at room temperature before the protein precipitate was removed by centrifugation and an aliquot of the supernatant was transferred to an amber glass autosampler for negative ionization LC–MS/MS. A second aliquot of the supernatant was diluted 5 times with 75% IPA for positive ionization LC–MS/MS. A 5 μ L volume of each sample was injected onto a Phenomenex Kinetex F5 column (100 mm \times 2.1 mm \times 2.6 μ m), and lipids were separated using a ternary gradient of Type 1 water, methanol, and isopropanol containing ammonium acetate. Lipids were quantified and identified with a Q-Exactive mass spectrometer (Thermo Fisher Scientific, Germany) equipped with a heated electrospray ionization HESI source. Data was processed using MS-DIAL v4.92 92⁶⁸. Batches were preceded by six injections of a quality control sample to stabilize the instrument. For full methodological details see the supplementary information. The lipidomics standard initiative report is available (<https://doi.org/10.5281/zenodo.14349715>).

Statistics and reproducibility

Python version 3.9.2 was used to perform all analyses unless otherwise stated⁶⁹. Due to the unequal sample sizes and non-normally distributed data; non-parametric statistical approaches were used for differential analysis. Significance between ordinal-continuous covariates was measured with the Mann–Whitney *U* test (MWU) using the ‘mannwhitneyu’ function from ‘scipy.stats’. Alpha (Gini, Shannon, richness, and Simpson) and Beta (Bray–Curtis Jaccard, weighted-unifrac, unweighted-unifrac) diversities were calculated using BioBakery’s ‘calculate_diversity.R’ on MetaPhlan’s ‘utils’ directory on their GitHub page. Co-linearity between the Beta diversity metrics was measured for significance using the ‘mantel’ function in ‘skbio’ python package. PCoA ordinations (plotted using ‘skbio.stats.ordination.pcoa’ module) was used to visualize the clustering of the beta diversities between samples for each dataset. To quantify the variance explained and constrained by metadata labels monovariate PERMANOVA *p*-values were calculated from those beta diversities using the ‘permanova’ function from the ‘skbio.stats.distance’ module. Associations between features were measured with Spearman correlation (calculated using ‘spearmanr’ function from ‘scipy.stats’ module), where significance was defined as FDR adjusted *p*-values⁶⁵ of <0.05. No blinding was performed for this study.

Associations between categorical variables were evaluated using Fisher’s Exact test (via the ‘fisher_exact’ function from the ‘scipy.stats’ module), with significance defined by *p*-values <0.05. Differential analysis across all datasets was conducted using a mixed modelling approach through MaAsLin2, adjusting for delivery mode, child sex, and duration of exclusive breastfeeding as fixed effects⁷⁰. For the microbiome analysis, total sum scaling (TSS) normalization was used to account for the compositionality of the dataset⁷¹ along with default

parameters of prevalence filtering of 10% and abundance of 0.01%. For all other datasets, no normalization or prevalence/abundance filtering was performed. Significance was defined as a *q*-value of <0.25. This threshold is appropriate for multivariable models, where the inclusion of covariates increases the multiple testing burden^{14,70,72,73}. This allows for the detection of biologically meaningful associations while maintaining an appropriate balance between Type I and Type II errors. Fixed effects were integrated into a formula for Redundancy Analysis (RDA) using the ‘rda’ function in the ‘vegan’ package in the R programming language⁷⁴. *Prevotella/Bacteroides* (P/B) ratio was calculated by dividing the pseudo-value adjusted and log-transformed relative abundance of *Bacteroides* divided by the same transformation of *Prevotella*.

Machine learning

Random Forest classifiers were trained to differentiate MAM from well-nourished controls. Each dataset was split 70/30 percent into training and testing sets. Class imbalance was addressed using the ‘SMOTE’ function from the ‘imblearn’ package for upsampling of the training data. Multimodal models were trained on data that was individually scaled and concatenated. Model hyperparameters, including the number of trees in the forest, maximum tree depth, and minimum sample numbers needed to split internal nodes, were tuned using grid searching with tenfold cross-validation of the test dataset using the ‘pycaret’ python package. The AUCROC of 100 Random Forest models for each prediction was used to measure the performance of each model and identify overfitting. Mean absolute SHapley Additive exPlanations (SHAP) value and the mean sum of the SHAP interaction values for each pair of features were used to interpret the contributions that each feature had on each model’s performance using the ‘shap’ python package⁷⁵. Important features were classified as having mean |SHAP| of above the 90th percentile of all values. Important interactions between features were classified as having mean Σ (SHAP interaction) above the 90th percentile of all values.

Network analysis

Important features (as defined above) in a multimodal model from faecal bacterial species and functional profiles, EEG PSD, Bayley, Wolke scores, and plasma lipids were included in the network analysis. Significant Spearman rho (*q* <0.05) between important features were used as edges. Leiden clustering of the positive correlations was performed using the Leidenalg and igraph python packages. Networks were visualized using Cytoscape (v3.10.2)⁷⁶.

Ethics approval and consent to participate

Ethical approvals were obtained from the Research Review Committee (RRC; August 21, 2021) and Ethical Review Committee (ERC) of icddr,b (protocol no.: PR-21084; September 21, 2021), Institutional Review Board of Boston Children’s Hospital, USA (for analyses of neuropsychological assessments), University of Auckland, New Zealand (approval AH23922; for analyses of collected biological samples) and University of West Indies (CREC-MN.51, 21/22).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The metagenomics data generated in this study have been deposited in the NCBI-SRA database under accession code [PRJNA1087376](https://doi.org/10.1038/s41467-024-55798-3). The lipidomics data generated in this study have been deposited in the Metabolights database under accession code [MTBLS10066](https://doi.org/10.1038/s41467-024-55798-3). The processed EEG, Bayley, Wolkes, Lipidomics spectral data, and anthropometric data are available on figshare (<https://doi.org/10.17608/k6.auckland.25560768>). All differential, feature importance and

correlation data generated in this study are provided in the Supplementary Information/Source Data file.

Code availability

All analysis code is available on the GitHub repository <https://github.com/theoportlock/m4efad> (<https://doi.org/10.5281/zenodo.14428001>). The codebase is organized into scripts, providing a comprehensive framework for replicating the experiments. Detailed documentation and instructions on how to use the code are provided in the repository's README file.

References

- De Onis, M., et al. *WHO Global Database on Child Growth and Malnutrition* (WHO, 1997).
- Martorell, R. & Ho, T. J. Malnutrition, morbidity, and mortality. *Popul. Dev. Rev.* **10**, 49–68 (1984).
- Nyaradi, A., Li, J., Hickling, S., Foster, J. & Oddy, W. H. The role of nutrition in children's neurocognitive development, from pregnancy through childhood. *Front. Hum. Neurosci.* **7**, 97 (2013).
- Jamison, D. T. Child malnutrition and school performance in China. *J. Dev. Econ.* **20**, 299–309 (1986).
- Gunston, G. D., Burkimsher, D., Malan, H. & Sive, A. A. Reversible cerebral shrinkage in kwashiorkor: an MRI study. *Arch. Dis. Child.* **67**, 1030–1032 (1992).
- Morreale, C. et al. Effects of perinatal antibiotic exposure and neonatal gut microbiota. *Antibiotics* **12**, 258 (2023).
- Enav, H., Bckhed, F. & Ley, R. E. The developing infant gut microbiome: a strain-level view. *Cell Host Microbe* **30**, 627–638 (2022).
- Robertson, R. C. et al. The gut microbiome and early-life growth in a population with high prevalence of stunting. *Nat. Commun.* **14**, 654 (2023).
- Fontaine, F., Turjeman, S., Callens, K. & Koren, O. The intersection of undernutrition, microbiome, and child development in the first years of life. *Nat. Commun.* **14**, 1–9 (2023).
- Chen, R. Y. et al. A microbiota-directed food intervention for undernourished children. *N. Engl. J. Med.* **384**, 1517–1528 (2021).
- Subramanian, S. et al. Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature* **510**, 417–421 (2014).
- Blanton, L. V. et al. Gut bacteria that prevent growth impairments transmitted by microbiota from malnourished children. *Science* **351**, aad3311 (2016).
- Kelsey, C., Dreisbach, C., Alhusen, J. & Grossmann, T. A primer on investigating the role of the microbiome in brain and cognitive development. *Dev. Psychobiol.* **61**, 341–349 (2019).
- Kelsey, C. M. et al. Gut microbiota composition is associated with newborn functional brain connectivity and behavioral temperament. *Brain Behav. Immun.* **91**, 472–486 (2021).
- Acua, I. et al. Infant gut microbiota associated with fine motor skills. *Nutrients* **13**, 1673 (2021).
- Carlson, A. L. et al. Infant gut microbiome associated with cognitive development. *Biol. Psychiatry* **83**, 148–159 (2018).
- Goyal, M. S., Venkatesh, S., Milbrandt, J., Gordon, J. I. & Raichle, M. E. Feeding the brain and nurturing the mind: linking nutrition and the gut microbiota to brain development. *Proc. Natl Acad. Sci. USA* **112**, 14105–14112 (2015).
- Vatanen, T. et al. A distinct clade of *Bifidobacterium longum* in the gut of Bangladeshi children thrives during weaning. *Cell* **185**, 4280–4297.e12 (2022).
- Veiga, G. R. S., Ferreira, H. S., Sawaya, A. L., Calado, J. & Florncio, T. M. M. T. Dyslipidaemia and undernutrition in children from impoverished areas of Macei\o, state of Alagoas, Brazil. *Int. J. Environ. Res. Public Health* **7**, 4139–4151 (2010).
- Das, S., Tripathy, B. B., Samal, K. C. & Panda, N. C. Plasma lipids and lipoprotein cholesterol in undernourished diabetic subjects and adults with protein energy malnutrition. *Diabetes Care* **7**, 579–586 (1984).
- Hornemann, T. Mini review: lipids in peripheral nerve disorders. *Neurosci. Lett.* **740**, 135455 (2021).
- De Aquino, C. C. et al. Effect of hypoproteic and high-fat diets on hippocampal blood-brain barrier permeability and oxidative stress. *Front. Nutr.* **5**, 131 (2019).
- Chadaideh, K. S. & Carmody, R. N. Host–microbial interactions in the metabolism of different dietary fats. *Cell Metab.* **33**, 857–872 (2021).
- Lamichhane, S. et al. Linking gut microbiome and lipid metabolism: moving beyond associations. *Metabolites* **11**, 55 (2021).
- An, D. et al. Sphingolipids from a symbiotic microbe regulate homeostasis of host intestinal natural killer T cells. *Cell* **156**, 123–133 (2014).
- Johnson, E. L. et al. Sphingolipids produced by gut bacteria enter host metabolic pathways impacting ceramide levels. *Nat. Commun.* **11**, 2471 (2020).
- Rohr, M. W., Narasimhulu, C. A., Rudeski-Rohr, T. A. & Parthasarathy, S. Negative effects of a high-fat diet on intestinal permeability: a review. *Adv. Nutr.* **11**, 77–91 (2020).
- Ahmed, T. et al. Nutrition of children and women in Bangladesh: trends and directions for the future. *J. Health Popul. Nutr.* **30**, 1 (2012).
- Shama, T. et al. Multidimensional evaluation of the early emergence of executive function and development in Bangladeshi children using nutritional and psychosocial intervention: a randomized controlled trial. *PLoS One.* **19**, e0296529 (2024).
- Lenters, L., Wazny, K. & Bhutta, Z. A. Management of severe and moderate acute malnutrition in children. In *Reproductive, Maternal, Newborn, and Child Health: Disease Control Priorities*. 3rd edn, 205–223 (World Bank, Washington, DC, 2016).
- Wastyk, H. C. et al. Gut-microbiota-targeted diets modulate human immune status. *Cell* **184**, 4137–4153 (2021).
- Roger, K. et al. Impact of early childhood malnutrition on adult brain function: an evoked-related potentials study. *Front. Hum. Neurosci.* **16**, 884251 (2022).
- Mukerji, C. E. et al. Resting frontal gamma power is associated with both expressive language and non-verbal cognitive abilities in young autistic children. *J. Autism Dev. Disord.* **54**, <https://doi.org/10.1007/s10803-024-06308-3> (2024).
- Cantiani, C., Piazza, C., Mornati, G., Molteni, M. & Riva, V. Oscillatory gamma activity mediates the pathway from socioeconomic status to language acquisition in infancy. *Infant Behav. Dev.* **57**, 101384 (2019).
- Bradley, H., Smith, B. A. & Xiao, R. Associations between EEG power and coherence with cognition and early precursors of speech and language development across the first months of life. *PLoS ONE* **19**, e0300382 (2024).
- Martins, V. J. B. et al. Long-lasting effects of undernutrition. *Int. J. Environ. Res. Public Health* **8**, 1817–1846 (2011).
- Nancy, B. & Glen, P. A. *Bayley-4: Scales of Infant and Toddler Development* [Fourth Edition (2019)].
- Wolke, D. Behavioral assessment of the infant. *Child Adolesc. Mental Health* **8**, 101–104 (2003).
- Badaloo, A. V., Forrester, T., Reid, M. & Jahoor, F. Lipid kinetic differences between children with kwashiorkor and those with marasmus. *Am. J. Clin. Nutr.* **83**, 1283–1288 (2006).
- Larabi, A. B., Masson, H. L. P. & Bäumlner, A. J. Bile acids as modulators of gut microbiota composition and function. *Gut Microbes* **15**, 2172671 (2023).
- Qin, N. et al. Alterations of the human gut microbiome in liver cirrhosis. *Nature* **513**, 59–64 (2014).
- Kato, K. et al. Age-related changes in the composition of gut *Bifidobacterium* species. *Curr. Microbiol.* **74**, 987–995 (2017).

43. Hjorth, M. F. et al. Prevotella-to-Bacteroides ratio predicts body weight and fat loss success on 24-week diets varying in macronutrient composition and dietary fiber: results from a post-hoc analysis. *Int. J. Obes.* **43**, 149–157 (2019).
44. Tett, A. et al. The *Prevotella copri* complex comprises four distinct clades underrepresented in westernized populations. *Cell Host & Microbe* **26**, 666–679 (2019).
45. Mwene-Batu, P. et al. Long-term effects of severe acute malnutrition during childhood on adult cognitive, academic and behavioural development in African fragile countries: the Lwiro cohort study in Democratic Republic of the Congo. *PLoS ONE* **15**, e0244486 (2020).
46. van Zutphen, T. et al. Malnutrition-associated liver steatosis and ATP depletion is caused by peroxisomal and mitochondrial dysfunction. *J. Hepatol.* **65**, 1198–1208 (2016).
47. D'Angelo, G., Capasso, S., Sticco, L. & Russo, D. Glycosphingolipids: synthesis and functions. *FEBS J.* **280**, 6338–6353 (2013).
48. Tan, S. T., Ramesh, T., Toh, X. R. & Nguyen, L. N. Emerging roles of lysophospholipids in health and disease. *Prog. Lipid Res.* **80**, 101068 (2020).
49. Magaquian, D., Delgado Ocaa, S., Perez, C. & Banchio, C. Phosphatidylcholine restores neuronal plasticity of neural stem cells under inflammatory stress. *Sci. Rep.* **11**, 22891 (2021).
50. Lee, H.-C. et al. LPIAT1 regulates arachidonic acid content in phosphatidylinositol and is required for cortical lamination in mice. *Mol. Biol. Cell* **23**, 4689–4700 (2012).
51. Nunez, P. L., Srinivasan, R. & Fields, R. D. EEG functional connectivity, axon delays and white matter disease. *Clin. Neurophysiol.* **126**, 110–120 (2015).
52. Hoban, A. E. et al. Regulation of prefrontal cortex myelination by the microbiota. *Transl. Psychiatry* **6**, e774 (2016).
53. Albouery, M. et al. Age-related changes in the gut microbiota modify brain lipid composition. *Front. Cell. Infect. Microbiol.* **9**, 444 (2020).
54. Hurley, M. J., Bates, R., Macnaughtan, J. & Schapira, A. H. V. Bile acids and neurological disease. *Pharmacol. Ther.* **240**, 108311 (2022).
55. Huang, F. et al. Deletion of mouse FXR gene disturbs multiple neurotransmitter systems and alters neurobehavior. *Front. Behav. Neurosci.* **9**, 70 (2015).
56. Heiss, C. N. & Olofsson, L. E. The role of the gut microbiota in development, function and disorders of the central nervous system and the enteric nervous system. *J. Neuroendocrinol.* **31**, e12684 (2019).
57. Ceppa, F., Mancini, A. & Tuohy, K. Current evidence linking diet to gut microbiota and brain development and function. *Int. J. Food Sci. Nutr.* **70**, 1–19 (2019).
58. Marques, T. M. et al. Gut microbiota modulation and implications for host health: Dietary strategies to influence the gut-brain axis. *Innov. Food Sci. & Emerg. Technol.* **22**, 239–247 (2014).
59. Fung, T. C., Olson, C. A. & Hsiao, E. Y. Interactions between the microbiota, immune and nervous systems in health and disease. *Nat. Neurosci.* **20**, 145–155 (2017).
60. The MathWorks Inc. *Optimization Toolbox Version: 9.4 (R2022b)* Natick Massachusetts: The MathWorks Inc. (2022).
61. Gabard-Durnam, L. J., Mendez Leal, A. S., Wilkinson, C. L. & Levin, A. R. The Harvard automated processing pipeline for electroencephalography (HAPPE): standardized processing software for developmental and high-artifact data. *Front. Neurosci.* **12**, 97 (2018).
62. Levin, A. R., Mendez Leal, A. S., Gabard-Durnam, L. J. & O'Leary, H. M. BEAPP: the batch electroencephalography automated processing platform. *Front. Neurosci.* **12**, 513 (2018).
63. Wolke, D., Skuse, D. & Mathisen, B. Behavioral style in failure-to-thrive infants: a preliminary communication. *J. Pediatr. Psychol.* **15**, 237–254 (1990).
64. Hamadani, J. D. et al. Randomized controlled trial of the effect of zinc supplementation on the mental development of Bangladeshi infants. *Am. J. Clin. Nutr.* **74**, 381–386 (2001).
65. Tofail, F. et al. Supplementation of fish-oil and soy-oil during pregnancy and psychomotor development of infants. *J. Health Popul. Nutr.* **24**, 48–56 (2006).
66. Beghini, F. et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *elife* **10**, e65088 (2021).
67. Liu, X. et al. Plasma lipidomics reveals potential lipid markers of major depressive disorder. *Anal. Bioanal. Chem.* **408**, 6497–6507 (2016).
68. Tsugawa, H. et al. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat. Methods* **12**, 523–526 (2015).
69. Van Rossum, G. *Python Reference Manual* (CWI (Centre for Mathematics and Computer Science), 1995).
70. Mallick, H. et al. Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput. Biol.* **17**, e1009442 (2021).
71. Gloor, G. B., Wu, J. R., Pawlowsky-Glahn, V. & Egozcue, J. J. It's all relative: analyzing microbiome data as compositions. *Ann. Epidemiol.* **26**, 322–329 (2016).
72. Aatsinki, A.-K. et al. Gut microbiota composition is associated with temperament traits in infants. *Brain Behav. Immun.* **80**, 849–858 (2019).
73. Braun, T. et al. Diet-omics in the study of urban and rural Crohn disease evolution (SOURCE) cohort. *Nat. Commun.* **15**, 3764 (2024).
74. Dixon, P. VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* **14**, 927–930 (2003).
75. Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. In *Adv. Neural Inf. Process. Syst.* 30 (ed. Guyon I. et al.) 4766 (Curran Associates Inc, New York, 2017).
76. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

Acknowledgements

Work on this clinical trial is supported by Wellcome Leap (9942 Culver Blvd Unit 1277 Culver City, CA 90232-4167, United States; www.wellcomeleap.org) to P.D.G., J.M.O., T.F. and C.A.N. as part of the 1kD Programme. We acknowledge our core donors, the Government of Bangladesh, and Canada, for providing unrestricted support and commitment to icddr's research effort. The authors would like to acknowledge the participants in Mirpur, Dhaka, Bangladesh for their contributions to this study. The authors would also like to thank the M4EFaD study team within the Infectious Diseases Division, International Centre for Diarrhoeal Disease Research, Bangladesh, Boston, and Auckland for their work in participant recruitment, sample collection, assessments, and analyses.

Author contributions

T.P., C.K., and J.M.O. drafted and co-wrote the manuscript. T.S., S.H.K., B.C.W., B.H., C.P., A.B., D.H., I.S., A.M.E., N.R., R.D.L., G.G., C.K., P.D.G., R.H., T.F., C.A.N. commented on the manuscript. J.M.O., R.H., T.F., P.D.G., C.A.N. designed the study and analyses. T.S. and S.H.K. performed assessments and obtained samples in Dhaka. R.H. oversaw the Dhaka group. T.P. performed multiomic analyses, B.C.W. and I.S. performed metagenomics, C.P. performed metabolomics, J.M.O. oversaw the Auckland group. B.H. performed EEG analyses, C.A.N. oversaw the Boston group.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-55798-3>.

Correspondence and requests for materials should be addressed to J. M. O'Sullivan or C. A. Nelson.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025