Maternal fecal microbiome predicts gestational age, birth weight and neonatal growth in rural Zimbabwe.

Ethan K. Gough[1,*], Thaddeus J. Edens[2,*], Hyun Min Geum[3], Iman Baharmand[3], Sandeep K. Gill[3], Ruairi C. Robertson[4], Kuda Mutasa[5], Robert Ntozini[5], Laura E Smith[5,6], Bernard Chasekwa[5], Florence D. Majo[5], Naume V. Tavengwa[5], Batsirai Mutasa[5], Freddy Francis[7], Lynnea Carr[8], Joice Tome[5], Rebecca J. Stoltzfus[10], Lawrence H. Moulton[1], Andrew J. Prendergast[1,4,5], Jean H. Humphrey[1,4], Amee R. Manges[3,9], SHINE Trial Team[11]

[1] Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
[2] Devil's Staircase Consulting, West Vancouver, British Columbia, Canada
[3] School of Population and Public Health, University of British Columbia, Vancouver, Canada
[4] Blizard Institute, Queen Mary University of London, London, UK
[5] Zvitambo Institute for Maternal and Child Health Research, Harare, Zimbabwe
[6] Department of Population Medicine and Diagnostics, Cornell University, Ithaca, NY, USA
[7] Department of Experimental Medicine, University of British Columbia, Canada
[8] Department of Microbiology and Immunology, University of British Columbia, Canada
[9] British Columbia Centre for Disease Control, Vancouver, British Columbia, Canada
[10] Goshen College, Goshen, Indiana, USA
[11] Members of the SHINE Trial team who are not named authors are listed in https://academic.oup.com/cid/article/61/suppl_7/S685/358186

ARTICLE INFO

ABSTRACT

Background: Preterm birth and low birth weight (LBW) affect one in ten and one in seven livebirths, respectively, primarily in low-income and middle-income countries (LMIC) and are major predictors of poor child health outcomes. However, both have been recalcitrant to public health intervention. The maternal intestinal microbiome may undergo substantial changes during pregnancy and may influence fetal and neonatal health in LMIC populations.

Methods: Within a subgroup of 207 mothers and infants enrolled in the SHINE trial in rural Zimbabwe, we performed shotgun metagenomics on 351 fecal specimens provided during pregnancy and at 1-month postpartum to investigate the relationship between the pregnancy gut microbiome and infant gestational age, birth weight, 1-month length-, and weight-for-age z-scores using extreme gradient boosting machines.

Findings: Pregnancy gut microbiome taxa and metabolic functions predicted birth weight and WAZ at 1 month more accurately than gestational age and LAZ. Blastocystis sp, Brachyspira sp and Treponeme carriage were high compared to Western populations. Resistant starch-degraders were important predictors of birth outcomes. Microbiome capacity for environmental sensing, vitamin B metabolism, and signalling predicted increased infant birth weight and neonatal growth; while functions involved in biofilm formation in response to nutrient starvation predicted reduced birth weight and growth.

Interpretation: The pregnancy gut microbiome in rural Zimbabwe is characterized by resistant starch-degraders and may be an important metabolic target to improve birth weight.

Funding: Bill and Melinda Gates Foundation, UK Department for International Development, Wellcome Trust, Swiss Agency for Development and Cooperation, US National Institutes of Health, and UNICEF.

Introduction

Children weighing <2,500 grams (g) at birth (low birth weight [LBW]), due to intrauterine growth retardation or preterm birth (<37 weeks completed gestation), are at high risk of morbidity and mortality during infancy and childhood, and of adverse effects over the life

Hyun Min Geum is currently affiliated with the University of British Columbia, School of Medicine.

* Corresponding author: Amee R. Manges, MPH, PhD, School of Population and Public Health, University of British Columbia, 137-2206 East Mall, Vancouver, BC Canada
E-mail address: amee.manges@ubc.ca (E.K. Gough).

Research in context

Evidence before this study

Preterm birth and low birth weight are a persistent public health challenge in low-income and middle-income countries (LMIC) that predict poor child health. There is need to identify more effective targets for intervention. A limited number of small studies, predominantly from high-income western populations, suggest that the maternal gut microbiome undergoes important metabolic changes during pregnancy that may play an underappreciated role in preterm, low birth weight and subsequent neonatal health. However, studies of the pregnancy gut microbiome in LMICs, where preterm and low birth weight are most prevalent, are extremely sparse; and metagenomic studies to date severely underrepresent populations outside of North America and Europe.

Added value of this study

We used shotgun metagenomics to investigate the relationship of pregnancy gut microbiome taxa and metabolic functions with infant gestational age, birth weight, and neonatal growth in a subsample of women and their infants enrolled in the Sanitation, Hygiene, Infant Nutrition Efficacy trial in rural Zimbabwe. In this population, resistant-starch degrading bacteria were the predominant gut taxa and were important predictors of birth weight and neonatal growth. Microbiome functions involved in signaling, starch, vitamin B and energy metabolism were also important predictors. Gut microbiome predictors of birth weight and neonatal growth largely reflected the low-diversity, resistant-starch dominated diet of the study population. This is the first study to identify gut microbiome markers of birth weight and early infant growth that likely reflect the dietary patterns of the mother, in a sub-Saharan African population where there is a very high prevalence of preterm birth and low birth weight.

Implications of all the available evidence

The pregnancy gut microbiome of rural Zimbabwean mothers, primarily the abundance of resistant-starch degraders, is an important contributor to birth weight and neonatal growth. Microbiome functions that predicted these outcomes suggested that bacterial capacity to respond to nutrient availability in the maternal diet may be important to promoting improved birth weight and early infant growth in populations with a monotonous starch-rich diet.

course [1]. Globally, LBW and preterm birth affect one in seven [2] and one in ten [3] livebirths respectively, primarily in low-income and middle-income countries (LMICs). However, nutritional interventions to promote optimal gestation, fetal and neonatal growth have not had a significant impact, or their effects have been considerably heterogeneous [4]. There is a need to better understand the underlying nutritional and environmental determinants of LBW and preterm birth so more effective targets for intervention can be identified.

Recent evidence suggests that the composition and function of the maternal gut microbiome before and during pregnancy may affect birth weight and duration of gestation [5−8]. Changes in maternal gut microbiome composition [9−11] and metabolic activity [12] have been observed during pregnancy; for example, increased carbohydrate metabolism and butyrate production [12]. These alterations in the maternal microbiome may contribute directly to fetal growth, or indirectly through influences on maternal nutritional status and gestational weight gain. In contrast, other studies have reported no differences in gut microbiome composition during pregnancy [9] or have reported that the effect of other host characteristics such as pre-pregnancy weight, gestational weight gain, fasting blood glucose, and place of residence explain 3-5 fold more variance in microbiome composition than stage of pregnancy [13].

The gut microbiome can directly suppress [14,15] or promote [16] intestinal inflammation, and has been associated with biomarkers of environmental enteric dysfunction [17,18]. Biomarkers of intestinal damage and microbial translocation that indicate enteric dysfunction, in turn, have been associated with shorter gestation [19−21] and smaller birth size [21]. Enteric pathogens also promote intestinal inflammation and are associated with alterations in microbiome composition [22]. Inadequate sanitation and hygiene have been linked to preterm birth in LMICs [23] where exposure to enteric pathogens is prevalent [24].

Nutrition during pregnancy can also influence fetal growth and development [25,26]. Dietary influences on the gut microbiome are well established [27,28], and the intestinal microbiome plays an important role in nutrient harvesting [29]. The gut microbiome can regulate blood pressure through the production of short chain fatty acids (SCFAs) [30] from dietary fiber [31]. Production of SCFAs by the gut microbiome during pregnancy has been associated with lower blood pressure in observational studies [8] and pregnancy microbiome composition has been associated with pre-eclampsia [32], which is an important risk factor for preterm birth [33]. In addition, the gut microbiome may contribute to prevention of micronutrient deficiencies, such as folate deficiency [34,35]. Plasma folate has been related to LBW when deficient [36], and higher carriage of intestinal *Bifidobacteria sp* was positively associated with plasma folic acid [37].

Indirectly, the gut microbiome may affect pregnancy outcomes through an impact on maternal nutritional status and gestational weight gain during pregnancy, which influence fetal growth. In adults, gut bacteria-derived SCFAs regulate food intake, body weight, energy expenditure, and satiety [38,39], and oral probiotic administration has been shown to regulate central adiposity during and after pregnancy [40]. In observational studies, mothers who were overweight prior to pregnancy or who gained more weight during pregnancy showed greater gut *Bacteroides, Clostridium, Staphylococcus* and *Enterobacteriaceae* (*Escherichia coli*) abundance [37,41].

Furthermore, the maternal gut is a predominant source of bacteria to colonize the infant gastrointestinal tract at birth [42,43]. Childhood microbiome immaturity (delayed acquisition of specific taxa with child age) is related to early-life growth [44−46]. A dysbiotic maternal gut microbiome can, in part, be directly transferred to the newborn [47], regardless of whether the dysbiosis results from dietary deficiencies or pathogen exposure. Growth during the neonatal period may, therefore, also be impacted via acquisition of the maternal pregnancy gut microbiome by the infant at birth.

However, the evidence for a relationship between the pregnancy gut microbiome and adverse birth outcomes comes predominantly from high-income settings, and the results of these small studies have been inconsistent [48]. Moreover, human microbiome species from settings outside of high-income European and North American populations are severely underrepresented, which highlights the need for evidence from more diverse populations and environments [49]. We sought to better understand the role of the maternal microbiome during pregnancy on gestational age, birth weight, and neonatal growth. Our hypothesis is that pathogen carriage, and bacterial species and functions in the fecal microbiome of pregnant mothers that are related to diet, water, sanitation, and hygiene will predict gestational age, birth weight and neonatal growth. We investigated our hypothesis in a subsample of mothers and infants participating in the Sanitation Hygiene Infant Nutrition Efficacy (SHINE) Trial [50]. SHINE was a 2 × 2 factorial cluster-randomized trial designed to test the impact of improved household water quality, sanitation, and

hygiene (WASH) and improved infant and young child feeding (IYCF) on linear growth and anaemia at 18 months (mo) of infant age in rural Zimbabwe [50].

We performed whole metagenome shotgun sequencing of maternal fecal specimens collected during pregnancy and at 1mo post-partum to identify maternal microbes and metabolic functions that might influence infant gestational age, size at birth, or growth during the neonatal period. We also investigated the impact of the SHINE WASH intervention, baseline sanitation and hygiene-related factors, and maternal characteristics on the maternal fecal microbiome.

Methods

Study design and participants

The study design and methods for the SHINE trial and for microbiome analyses have been reported previously [51,52]. In brief, SHINE was a cluster-randomized trial of the independent and combined effects of improved IYCF and WASH in two rural Zimbabwean districts with 15% antenatal HIV prevalence. The WASH intervention was designed to limit exposure to human and animal feces, was initiated during pregnancy, and included, at the household level: construction of a ventilated improved pit latrine, installation of two hand-washing stations plus monthly delivery of liquid soap and water chlorination solution, provision of a play space for the infant, and hygiene counseling. Infants had birth weight measured in institutions by personnel who had been provided with Seca scales and had been trained in their use. At 1mo of infant age, a home visit was conducted to measure weight and length, as previously described [50]. The SHINE trial is registered at ClinicalTrials.gov (NCT01824940).

In a pre-specified substudy of SHINE, 1,656 mother-child pairs were selected for biological specimen collection from the mother during pregnancy then at 1mo post-partum, and intensive collection from the infant at 1, 3, 6, 12 and 18 months of follow-up. This substudy sample was enriched for HIV-positive mothers because we were also interested in investigating outcomes associated with fetal HIV-exposure. Of these substudy pairs, two hundred seven mother-child pairs with nearly complete longitudinal fecal sample collection were investigated as part of the fecal microbiome study; these are the focus of our analyses in this manuscript.

Fecal specimens

Mothers collected fecal specimens prior to a home visit by a research nurse during pregnancy and at the 1mo post-partum study visit. Fecal specimens were immediately placed in a cold box and transported to the regional laboratory, where they were stored at -80°C until transfer to the central laboratory in Harare for long-term archiving at -80°C. Fecal specimens were transferred via private courier on dry ice from Harare, Zimbabwe to Vancouver, British Columbia. The Qiagen DNeasy PowerSoil Kit was used to extract total DNA from 200mg of feces, according to manufacturer's instructions.

Whole metagenome library preparation and sequencing

Paired-end libraries were constructed using the Illumina TruSeq kit and using New England Biosystem TruSeq compatible library preparation reagents. Libraries were sequenced at the British Columbia Genome Sciences Centre using the Illumina HiSeq 2500 platform. Forty-eight libraries were pooled and included per sequencing lane. Sixteen negative process controls, consisting of sterile water subjected to DNA extraction, library preparation, and sequencing, were included to capture microbial contamination of laboratory reagents.

Bioinformatics

Sequenced reads were trimmed of adapters and filtered to remove low-quality, short ($<$60 base-pairs), and duplicate reads, as well as those of human, other animal or plant origin using KneadData with default settings [53]. Species composition was determined by identifying clade-specific markers from reads using MetaPhlAn3 with default settings [54]. Relative abundance estimates were obtained from known assigned reads, and unknown read proportions were estimated from total, assigned and unassigned, reads. Percent human DNA was estimated from KneadData output, using the proportion of quality-filtered reads that align to the human genome. DNA extraction from fecal samples did not include enrichment for viral DNA, nor was cDNA synthesized. Given the smaller viral genome sizes, sequencing depth, and limitations of MetaPhlAn3 for virus identification, we did not include viruses in our current analyses. We applied a minimum relative abundance threshold of 0.1% for taxa and included taxa meeting this threshold in $\geq$5% of specimens in all downstream analyses.

Functional gene and metabolic pathway composition was determined using HUMAnN3 with default settings against the UniRef90 database [55]. Functionally annotated reads were further classified into level-4 enzyme commission (EC) categories using provided scripts. Enzyme family and pathway abundance estimates were normalized using reads per kilobase per million mapped reads (RPKM) and then re-normalized to relative abundance. We applied a minimum relative abundance threshold of 3×10^{-7}% for EC and pathway features, at $\geq$5% prevalence, in all downstream analyses.

Statistical analyses

All analyses were conducted in R version 3.3.2 [56]. Infant weight and length at 1mo of age were converted to Z-scores using the WHO growth standard [57]. The association between epidemiologic variables and birth weight or infant growth was assessed using multivariable linear regression. Epidemiologic variables associated with infant growth were selected for inclusion in XGBoost models.

Maternal fecal microbiome composition was investigated using descriptive approaches (e.g. boxplots) and α- and β-diversity metrics. α-diversity was assessed using the Shannon, Simpson, inverse Simpson, evenness and richness metrics; while β-diversity was evaluated using the Bray$-$Curtis dissimilarity index and visualized using principal coordinate analyses (PCoA) performed in the package *vegan* [58]. Permanova was performed to assess significant differences in β-diversity by visit, HIV-status, season of birth, and WASH assignment using the *adonis* function [59] in *vegan*.

The bivariate relationship between epidemiologic variables and maternal fecal taxa was examined using zero-inflated beta regression fitted by generalized additive models for location, scale and shape (GAMLSS), with a log-link, in the R package *gamlss* [60]. A separate model was fit for each microbiome taxon, yielding a relative abundance ratio (RAR). Epidemiologic variables were assessed individually and included baseline household characteristics (number of occupants, diet diversity, food insecurity, wealth index, improved floor, improved latrine, time to water, water treatment), maternal characteristics (age, height, mid-upper-arm circumference (MUAC), education, religion, HIV-status, anti-retroviral therapy, depression score, parity), mode and location of delivery, season of fecal specimen collection, diet composition on a normal day, randomized WASH allocation, and percent human DNA. To assess diet composition on a normal day, mothers responded to questions about her diet and her family's diet (meals and snacks) consumed 'yesterday', unless yesterday was a feast day or celebratory day, then she was asked to respond for the day before or her last 'normal' diet day. False discovery rate (FDR) adjustment was used to calculate q-values from p-values [61]. Results are reported when the q-value was less than 0.05.

Independent associations between the maternal pregnancy microbiome and gestational age, birth weight, WAZ or LAZ at 1mo was evaluated using extreme gradient boosting machines (XGBoost). Microbiome data are very complex, consisting of hundreds of species or pathways, and thousands of enzyme commissions (EC), and comprise of highly right-skewed distributions with many zeros. Analytic methods that can address these challenges and select the most important subset of features for a given outcome are limited. XGBoost builds an optimized predictive model by creating an ensemble from a series of weakly predictive models. In aggregate, each additional model that is fitted is parameterized to improve the overall prediction accuracy of the ensemble when the new model is combined with previous models. Only features that improve model accuracy are retained in the process. XGBoost is also non-parametric, can capture non-linear relationships, and can accommodate high-dimensional data [62]. XGBoost or other decision-tree based methods (eg. Random Forests) have been applied in other microbiome analyses [45,63,64].

The XGBoost models were developed using microbiome relative abundances, diversity measures, percent human DNA reads, percent unknown reads, and selected epidemiologic variables. Epidemiologic variables were selected for inclusion in the XGBoost model based on: (i) their association with birth weight, LAZ or WAZ at 1mo in multiple linear regression models (Supplemental Table 4); (ii) their association with microbiome diversity indices in simple linear regression models ($p<0.05$); and (ii) their association with microbiome taxon abundance in GAMLSS models ($q<0.05$). XGBoost model selection was performed in 3 stages. In stage one, the *BayesianOptimization* function of the *rBayesianOptimization* package was used with 10-fold cross-validation to select model hyperparameters (Supplementary Table 1) by minimizing the mean squared error (MSE). Models with the lowest MSE (in the 5th percentile) were retained, and from these models the variables that contributed to the top 95% of variable importance by proportion were retained. In stage two, all epidemiologic variables were included with the microbiome variables obtained in stage one. *BayesianOptimization* was run as described in stage one but using leave-one-out cross-validation. Microbiome variables that contributed to the top 95% of variable importance by proportion were retained. In stage three, all epidemiologic variables, microbiome features, and hyperparameters selected in stage two were used to fit our final models. Final models were fit using leave-one-out cross-validation to minimize the MSE. This process was implemented separately for each outcome (gestational age, birth weight and neonatal growth measures), and separately for microbiome taxa, pathways, and enzymes. Finally, all models were re-run excluding epidemiological variables to assess the contribution of the maternal pregnancy microbiome alone to the accuracy of our final models. XGBoost models were fitted using the H20.ai engine and *h2o* R package interface with the *XGBoost* package.

XGBoost model performance was evaluated using three estimates: (i) pseudo-R-squared (pseudo-R^2) between the final XGBoost-predicted outcome value and the observed outcome value, where pseudo-R^2=1·0 is perfect prediction; (ii) the mean absolute error (MAE), which is defined as the average absolute difference between observed and predicted outcome values estimated from cross-validation and summarizes model performance in the actual units of the outcome variable (kilograms for birth weight, weeks for gestational age, standard deviations for LAZ and WAZ at 1mo); and (iii) the root mean squared error, which is defined as the average squared difference between observed and predicted outcome values estimated from cross-validation. We used the scaled relative importance for each variable in a model to identify the twenty most informative variables for further interpretation, where the most important variable is ranked first, and the importance of subsequent variables are relative to the first variable.

The marginal relationships between the twenty most important epidemiologic variables, microbiome species, pathway or enzyme relative abundances and each outcome were visualized for interpretation [65] using accumulated local effects plots (ALE). ALE plots can be interpreted as showing a marginal effect, adjusted for all covariates retained in the final model. That is, the plots show the expected change in the outcome variable per increment in a feature, either epidemiologic or microbiome, adjusted for the variables retained in the model. The range of the feature across all observations is partitioned into intervals such that each interval contains roughly the same number of observations. The corresponding expected changes in the outcome values per increment, or effect sizes, are averaged. The resulting effect sizes are plotted cumulatively and centered about the average effect size [66]. ALEs were generated using the *ALEplot* package and were plotted using *ggplot2*. Standard deviations were calculated per increment and were used to calculate and plot increment-wise 95% confidence intervals.

Ethics statement

All SHINE mothers provided written informed consent. The Medical Research Council of Zimbabwe (MRCZ/A/1675), Johns Hopkins Bloomberg School of Public Health (JHU IRB # 4205.), and the University of British Columbia (H15-03074) approved the study protocol, including the microbiome analyses.

Data sharing

All relevant data are within the paper and its Supporting Information files except for the raw data which the trial team will begin loading as individual participant data with an accompanying data dictionary at http://ClinEpiDB.org in mid-2021. Prior to that time, the data are housed on the ClinEpiDB platform at the Zvitambo Institute for Maternal and Child Health Research and available upon request from Ms. Virginia Sauramba (vsauramba@zvitambo.co.zw).

Role of funding source

The funders had no role in study design, data collection, data analyses, interpretation, or writing of the manuscript.

Results

Study population

Two hundred seven SHINE mothers were included in these analyses (Figure 1). These mothers gave birth to 215 infants (199 singletons and 8 sets of twins). A comparison of mothers included in the microbiome substudy and the entire SHINE cohort is provided in Table 1. There were 97 mothers living with HIV in the substudy, with median (IQR) CD4 count of 444 cells/mm³ (300,652). The majority (93%) were receiving antiretroviral therapy (ART) and 46% were receiving cotrimoxazole prophylaxis at baseline (Table 1). In the microbiome substudy compared to the overall SHINE cohort, there were slightly fewer mothers allocated to the WASH intervention arm and fewer mothers in the lowest wealth quintile, more HIV-positive mothers with a handwashing station, and more HIV-negative mothers with livestock in the home at baseline. Approximately 45% of households met minimum dietary diversity standards, and few households experienced severe food insecurity (Table 1). Mothers and infants included in these analyses otherwise resembled the overall SHINE population (Table 1). Importantly, 100% of mothers consumed cereals on a typical day (Supplementary Table 2). In Zimbabwe, the most commonly consumed cereal is maize, a major source of resistant-starch.
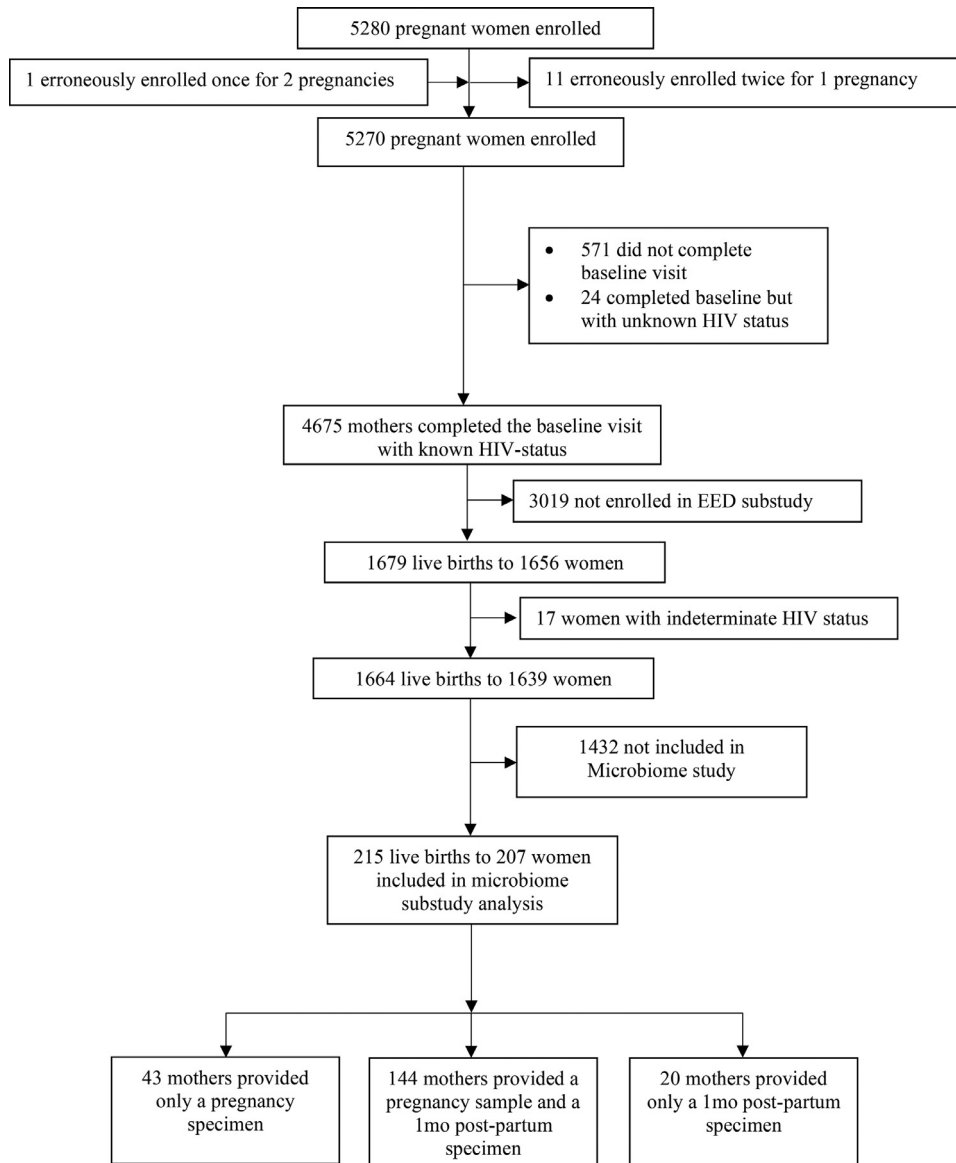
Figure 1. Flow of participants through the SHINE microbiome substudy.

A total of 351 whole metagenome sequencing datasets were produced. 144 mothers provided samples from the gestational visit (between 9- and 37-weeks gestation) and at 1mo post-partum (between 12 and 90 days after birth); an additional 43 provided only a pregnancy sample and 20 only a postpartum sample. There were median 1(IQR: 1,2) and 2(IQR: 1,4) datasets per randomization cluster at each visit, respectively. Since there was little opportunity for expression of intra-cluster correlation, it was not adjusted for in any of our analyses. There was very little variability in sequencing depth across specimens (mean 11·4±3·0 million quality-filtered read pairs for all datasets). Negative controls revealed negligible levels of microbial contamination of reagents (mean 734±3,462 quality-filtered reads) relative to the sequencing depth achieved across specimens. The median proportion of human reads detected was 0·01% but ranged widely from 0·0004% to 19·3% (Supplementary Figure 1), indicating high human DNA enrichment in some fecal specimens.

After applying our relative abundance and prevalence threshold criteria, 144 unique taxa at the species level or above, 424 microbiome pathways, and 1,956 enzyme commission features were included in downstream analyses from the pregnancy and 1mo post-partum metagenome datasets. Of these, 39 taxa (28%) were defined by MetaPhlAn3 as co-abundance gene groups (CAGs) [67]. CAGs are putative species genomes derived from assembled sequences in publicly available metagenomes, and for which no bacterial culture-derived representative exists [49,67]. These represent novel bacteria that are prevalent in rural Zimbabwean mothers and that are missing from reference databases comprised of bacteria isolated or sequenced primarily from North American and European populations. Eight of these CAGs were defined at the phylum (Firmicutes) level only, meaning no genetically closer relative exists in reference databases at finer taxonomic levels.

Baseline variables influencing maternal microbiome composition during pregnancy and 1mo post-partum

In PCoA analyses of microbiome β-diversity by epidemiologic variables (using Bray-Curtis distances), maternal microbiome composition differed modestly by HIV-status (Permanova p-value = 0·012, goodness of fit=0·19) and randomized WASH assignment (p-value = 0·022, goodness of fit = 0·23) (Supplementary Figure 2). However, microbiome composition did not vary by trimester of pregnancy (p-value = 0.343) (Supplemental Figure 3).

Table 1
Baseline characteristics of the microbiome sub-study versus all SHINE mothers and their infants

Characteristic	SHINE mothers included in the microbiome sub-study n = 207		All SHINE mothers completing the baseline visit n = 4,675	
	HIV-positive n = 97	HIV-negative n = 110	HIV-positive n = 773	HIV-negative n = 3902
Trial Arm				
Standard of care	23 (23.7%)	32 (29.1%)	173 (22.4%)	911 (23.4%)
IYCF	28 (28.9%)	32 (29.9%)	164 (21.2%)	929 (23.8%)
WASH	20 (20.6%)	20 (18.2%)	214 (27.7%)	996 (25.5%)
WASH + IYCF	26 (26.8%)	26 (23.4%)	222 (28.7%)	1066 (27.3%)
Household Characteristics				
Median Number of Occupants (IQR)	4.0 (3.0,5.0)	5.0 (3.0,6.0)	4.0 (3.0, 6.0)	5.0 (3.0, 6.0)
Wealth Quintile				
1 (lowest)	18 (20.0%)	9 (9.3%)	210 (27.5%)	714 (18.5%)
2	21 (23.3%)	24 (24.7%)	178 (23.3%)	752 (19.5%)
3	21 (23.3%)	19 (19.6%)	150 (19.6%)	775 (20.1%)
4	17 (18.9%)	26 (26.8%)	112 (14.6%)	823 (21.3%)
5 (highest)	12 (13.3%)	19 (19.6%)	115 (15.0%)	801 (20.7%)
Sanitation				
Improved latrine at household	29 (32.2%)	36 (37.1%)	217 (28.7%)	1227 (32.2%)
Water				
Main source of household drinking water improved	63 (70.0%)	60 (61.9%)	450 (59.6%)	2440 (63.7%)
Median one-way walk time to fetch water (IQR), min	10 (5,15)	10 (5,15)	10.0 (5, 20)	10.0 (5, 20)
Median one-way walk time to fetch water (IQR), min	10 (5,15)	10 (5,15)	10.0 (5.0, 20.0)	10.0 (5.0, 20.0)
Treat drinking water to make it safer	15 (15.9%)	15 (14.2%)	84 (12.1%)	449 (12.6%)
Hygiene				
Improved floor at household	44 (47.7%)	69 (62.7%)	361 (47.8%)	2118 (55.6%)
Maternal characteristics				
Mean age (sd), years	31.2 (5.9)	27.2 (6.7)	29.0 (6.3)	25.7 (6.7)
Mean height (sd), cm	161.0 (6.2)	160.6 (5.4)	159.6 (9.0)	159.8 (8.1)
Mean mid-upper-arm circumference (sd), cm	27.0 (2.7)	27.3 (3.5)	26.3 (3.0)	26.4 (3.1)
Mean years of schooling completed (sd)	9.2 (2.2)	9.4 (1.9)	9.2 (2.1)	9.6 (1.8)
Median parity (IQR)	2.0 (1.0,3.0)	2.0 (1.0,3.0)	2.0 (1.0, 3.0)	2.0 (1.0, 3.0)
Married	85 (94.4%)	89 (91.8%)	691 (94,5%)	3550 (95.8%)
Religion				
Apostolic	50 (55.6%)	30 (31.0%)	348 (47.2%)	1732 (46.4%)
Other Christian	35 (38.9%)	63 (64.9%)	313 (42.4%)	1699 (45.5%)
Other	3 (3.3%)	2 (2.1%)	77 (10.4%)	306 (8.2%)
HIV status and therapy				
Mean CD4 count (sd)	444 (300,652)	NA	460.5 (349, 614)	NA
Currently receiving cotrimoxazole	32 (45.7%)	NA	399 (55.7%)	NA
Currently receiving ART	65 (92.6%)	NA	582 (81.2%)	NA
Potential depression ($\geq$ 10 on Edinburgh Postnatal Depression Scale	11 (11.7%)	6 (6%)	103 (14.0%)	350 (9.4%)
Trimester of Specimen collection				
First trimester (0 to $\leq$84 weeks)	13 (15.1%)	21 (22.8%)		
Second trimester (>84 to $\leq$196 weeks)	65 (75.6%)	56 (60.9%)		
Third trimester (>196 weeks)	8 (9.3%)	15 (16.3%)		
Diet quality and food security				
Household meets minimum Diet Diversity Score	35 (43%)	44 (44%)	252 (38.2%)	1364 (40.2%)
Median Coping Strategies Index score (IQR)	0 (0,3)	1 (0,4)	2.0 (0, 10)	1.0 (0, 7)
Mean days of staple food in the household (sd)	144 (119)	176 (120)	114.6 (146)	129.3 (129)
Season of stool collection				
Dry (May to September)	40 (44.4%)	59 (57%)	281 (39.4%)	1566 (42.5%)
Rainy (October to April)	50 (55.6%)	44 (43%)	433 (60.6%)	2120 (57.5%)
Infant Characteristics*				
Female sex	48 (48%)	50 (43%)	363 (50.1%)	1830 (49.1%)
Mean birth weight (sd), kg	3.0 (0.5)	3.0 (0.5)	3.0 (0.5)	3.1 (0.5)
Birthweight < 2500 g	13 (13%)	11 (10%)	84 (11.5%)	306 (9.2%)
Preterm (<37 weeks gestation)	6 (9.8%)	9 (11.5%)	63 (17.3%)	326 (16.3%)
Vaginal delivery	91 (96%)	104 (94%)	63 (17.3%)	326 (16.3%)
Institutional delivery	81 (86%)	107 (96%)	612 (92.5%)	3207 (92.4%)
Exclusive breastfeeding initiation	89 (98%)	103 (95%)	545 (83.6%)	3026 (88.7%)
Mean LAZ at 1-month (sd)	-1.1 (1.3)	-0.9 (1.3)	-1.1 (1.3)	-0.9 (1.4)
Mean WAZ at 1-month (sd)	-0.6 (1.4)	-0.3 (1.5)	-0.8 (1.3)	-0.5 (1.2)

* *n=215 infants born to 207 mothers (100 born to HIV-positive mothers and 115 born to HIV-negative mothers). Accurate gestational age measurement was unavailable for 20 HIV-exposed infants and 13 HIV-unexposed infants; and twins were excluded from gestational age estimation. n, sample size; IYCF; infant and young child feeding; WASH, water sanitation, and hygiene; IQR; inter-quartile range; min, minutes; sd; standard deviation; cm; centimeters; ART, antiretroviral therapy; Kg, kilograms; g, grams; AGA, adequate size for gestational age; SGA small for gestational age; LAZ, length-for-age z-score, WAZ, weight-for-age z-score.

In bivariate analyses of differences in relative abundance of individual taxa by epidemiologic variables using GAMLSS models, only one taxon differed by maternal HIV-status and two taxa differed by WASH intervention, after FDR adjustment, at the post-partum visit only. Higher *Treponema berlinense* (RAR=4·67, q=0·003) was observed in HIV-negative mothers at the post-partum visit; while higher *Brachyspira sp* CAG 700 (RAR = 6·9, q<0·001) and lower *Akkermansia muciniphila* (RAR=0·37, q=0·042) were observed in mothers assigned

to the WASH intervention (Supplementary Table 3) at the post-partum visit. There was no evidence that any of the α-diversity metrics varied by these variables.

Pathogen detection

Overall, based on MetaPhlAn3 results, carriage of pathogens was uncommon. At a minimum relative abundance of 0·1%, *Brachyspira pilosicoli*, a spirochete associated with human zoonotic spirochaetosis, was carried by 20 mothers, while 22 mothers carried *Brachyspira sp* CAG 700. Interestingly, lower wealth quintile, lower primary education and greater maternal height were associated with higher *Brachyspira sp* abundance (Supplementary Table 3). Higher percentage of human DNA was associated with higher relative abundance of *Blastocystis* and *Alistipes shahii* at 1mo postpartum (Supplemental Table 3). Other non-pathogenic spirochetes were identified in many SHINE mothers, including *Tremponema succinifaciens* (n=68) and *Treponema berlinense* (n=17) (Supplementary Figure 4). Spirochetes have been shown to be over-represented in many non-Western populations [68]. Thirty-one mothers were also positive for *Blastocystis sp* subtype 1 during at least one visit. *Giardia intestinalis* or *Cryptosporidium spp* were detected but below the 0·1% relative abundance threshold. *Shigella* was not detected and *Salmonella enterica, Campylobacter jejuni* and *C. coli* were also present at <0·1% relative abundance. *Escherichia coli* was common, but MetaPhlAn3 does not separate by *E. coli* pathotype; therefore, presence of enteroaggregative or enteropathogenic *E. coli* could not be assessed.

Epidemiologic variables, birth weight, and neonatal growth

In multivariable linear regression models, maternal height, MUAC, and gestational age, were associated with infant birth weight, as expected (Supplementary Table 4). While maternal MUAC and gestational age were associated with WAZ and LAZ at 1mo respectively.

Maternal microbiome diversity, birth weight, and neonatal growth

To ensure temporality between maternal microbiome characteristics and pregnancy outcomes, all analyses of the association between maternal microbiome, gestational age, birth weight, and neonatal growth utilized the pregnancy visit metagenome datasets only. In simple linear regression models, microbiome Shannon diversity was not associated with infant birth weight or LAZ at 1mo (Supplementary Figure 5) but was associated with higher WAZ at 1mo (Supplementary Figure 5). Higher pregnancy microbiome taxon evenness was associated with higher infant WAZ at 1mo, and there was evidence for a positive association with the Shannon diversity index (*p*-value = 0·09).

XGBoost model performance

Variables chosen for inclusion in XGBoost model building were infant sex, HIV-status, maternal MUAC, maternal height, birth season, pregnancy trimester and season of maternal specimen collection, microbiome diversity metrics, percent unknown, percent human DNA, and gestational age (for growth models). Performance measures for final models of microbiome species, ECs, or pathways, including or excluding epidemiologic variables, are presented in Supplementary Table 5. For all outcomes, the performance of models that included only microbiome features performed as well as the models that also included epidemiologic variables, indicating that model performance was predominantly driven by microbiome composition. Overall, model performance was best for birth weight, followed by neonatal growth, then gestational age. In terms of microbiome features, pathway relative abundance was more predictive than species abundance, while EC abundance was as good or better than pathway abundance

at predicting outcomes, although this may be driven by the higher overall number of ECs.

Important XGBoost variables for birth weight

ALE plots are shown in Figures 2-3, and Supplementary Figures 6-15 to visually interpret the marginal relationships between each outcome and the individual top 20 variables in the final XGBoost models adjusted for the other epidemiologic variables and microbiome features retained in the model. Microbiome features are plotted using percentiles of the abundance distribution to improve visualization near zero where abundance data are denser. Epidemiologic and microbiome diversity variables are plotted on the original scale.

For infant birth weight, gestational age and maternal height were the most informative variables, predicting a cumulative increase in birth weight of 150g and 200g, respectively, over the range of observed values (Figure 2). Taxa that are associated with resistant-starch degradation, specifically members of the *Ruminococcaceae, Lachnospiraceae*, and *Eubacteriaceae* families [69,70], were the next most important predictors of birth weight. However, the relationships were quite variable, and most effects were only observed in the extremes of the abundance distributions. For example, *E. eligens* predicted a cumulative decrease of 250g from the lowest abundance up to approximately the 40th abundance percentile, while *R. torques* predicted a cumulative decrease of 150g above the 25th abundance percentile. In contrast, *R. intestinalis* and *Butyrivibrio sp CAG 318* predicted a cumulative increase of 200g and 100g, respectively, above the 80th abundance percentile (Figure 2). Maternal MUAC was also among the 20 most important variables, predicting a 75g increase in birth weight.

The bacterial enzyme with the highest-ranking importance was glycogen synthase (~200g cumulative decrease up to the median abundance), exhibiting greater importance in the model than gestational age and maternal height (Supplementary Figure 6). Glycogen synthase is involved in starch metabolism. The remaining top 20 enzymes are involved in metabolic pathways that engage in signalling, energy, and vitamin B metabolism (Supplementary Figure 6). For example, histidine kinase (involved in environmental sensing) predicted a 300g increase in birth weight in the lowest abundance percentile. While thiazole biosynthesis pathway (vitamin B1 metabolism) predicted a 200g increase in the lowest abundance percentile and N10-formyl-tetrahydrofolate biosynthesis (vitamin B9 metabolism) predicted a 75g increased (Supplementary Figure 7).

Similarly, top ranking metabolic pathways are engaged in starch degradation, energy metabolism and signalling processes (Supplementary Figure 7). For example, sucrose degradation II pathway and the superpathway of anaerobic sucrose degradation predicted a 200g increase above the 80th abundance percentile (Supplementary Figure 7). The signalling ppGpp biosynthesis pathway predicted an 85g decrease; while, in contrast, the signalling pathway purine ribonucleosides degradation predicted a ~150g increase (Supplementary Figure 7).

Important XGBoost variables for WAZ at 1mo

For WAZ at 1mo, the most important taxa were also predominantly starch-degrading species in the families *Ruminococcaceae, Lachnospiraceae*, and *Eubacteriaceae* (Figure 3). One species outside of these families (*Catenibacterium mitsuokai*), that is also associated with fiber fermentation [71] and consumption of animal fat [72], predicted a ~0.3 sd increase in WAZ. *Treponema berlinense* predicted a 0.6 sd decrease above the 97th abundance percentile. The Simpson diversity index and evenness were also important predictors, indicating a 0.2 sd and 0.7 sd increase through the range of observed diversity values, respectively. However, species richness predicted a 1 sd cumulative decrease in WAZ. As with birth weight, gestational age
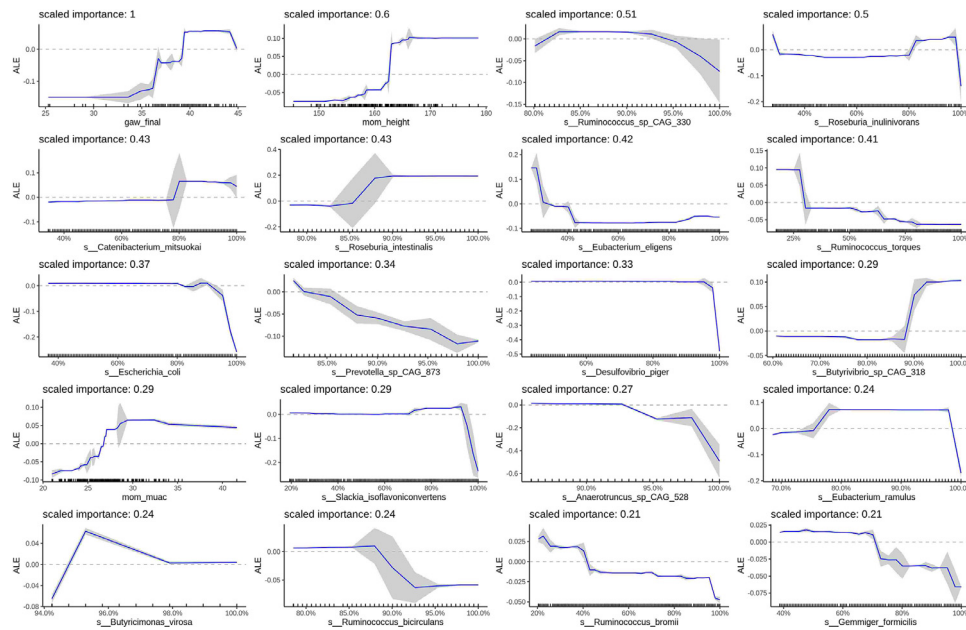
Figure 2. Relationships between infant birth weight in Kg, epidemiologic variables, and maternal gut microbiome species relative abundance. The top 20 predictors of infant birth weight by variable importance score are shown. For microbiome abundances, the x-axis represents the percentile of the abundance distribution. Epidemiologic and microbiome diversity variables are on the original scale. Tick marks on the x-axis are a rug plot of individual feature abundance percentiles. ALEs were generated using the _ALEplot_ package and were plotted using _ggplot2_. Standard deviations (sd) were calculated per increment in microbiome feature and were used to calculate and plot increment-wise 95% confidence intervals as the average change in the outcome $\pm 1.96(\text{sd}/\sqrt{n})$, where n is the number of observed feature values, and sd is the standard deviation of the change in the outcome variable in an interval. gaw_final, gestational age; mom_height, maternal height in centimeters; mom_muac, maternal mid-upper arm circumference in millimeters; pct_human, percent human reads; pct_unknown, percent unknown reads.

and maternal anthropometry were also important predictors of 1mo WAZ (Figure 3).

Important microbiome pathways and ECs to 1mo WAZ are also involved in vitamin B, starch metabolism, and signalling. For example, flavin biosynthesis pathway III (vitamin B2 synthesis) predicted

an increase of 0.5 sd (Supplementary Figure 9). Cellulase abundance predicted ~0.5 sd increase above the 87th abundance percentile. D-fructuronate degradation and sucrose degradation pathway III predicted 1 sd and 0.2 sd increase in WAZ, respectively (Supplementary Figure 9). While enzymes involved in signalling predicted 0.2 sd
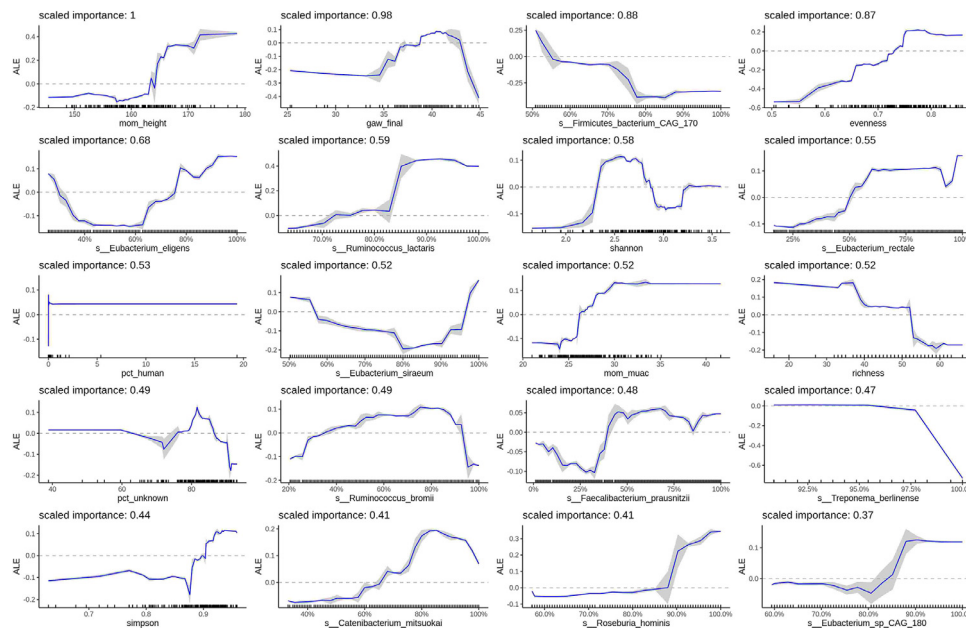

Figure 3. Relationships between infant WAZ at 1-month, epidemiologic variables, and maternal gut microbiome species relative abundance. The top 20 predictors of infant WAZ at 1-month by variable importance score are shown. For microbiome abundances, the x-axis represents the percentile of the abundance distribution. Epidemiologic and microbiome diversity variables are on the original scale. Tick marks on the x-axis are a rug plot of individual feature abundance percentiles. ALEs were generated using the _ALEplot_ package and were plotted using _ggplot2_. Standard deviations (sd) were calculated per increment in microbiome feature and were used to calculate and plot increment-wise 95% confidence intervals as the average change in the outcome $\pm 1.96(\text{sd}/\sqrt{n})$, where n is the number of observed feature values, and sd is the standard deviation of the change in the outcome variable in an interval. gaw_final, gestational age; mom_height, maternal height in centimeters; mom_muac, maternal mid-upper arm circumference in millimeters; pct_human, percent human reads; pct_unknown, percent unknown reads.

increase (e.g cytidine deaminase) or >0.2 sd decrease (e.g. GMP reductase) (Supplementary Figure 8).

Important XGBoost variables for LAZ at 1mo

Gestational age, maternal height and MUAC were also among the most important predictors of LAZ at 1mo; however, *Anaerotruncus sp* CAG 528 exhibited a higher relative importance score than maternal height and MUAC (Supplementary Figure 10). Resistant-starch degraders also figured prominently in the set of taxa important to LAZ, as well as the dietary fiber fermenting *Prevotella copri* [73,74] (Supplementary Figure 10). Some important microbiome functions to 1mo LAZ were also important to WAZ and are also involved in starch metabolism, signalling and vitamin B metabolism (Supplemental Figures 11 and 121). For example, holo- [acyl-carrier-protein] synthase (vitamin B5 and Coenzyme A metabolism) and 4-phosphoerythronate dehydrogenase (vitamin B6 metabolism) predicted a 1 sd and 0.2 sd increase in LAZ, respectively, below the 25th abundance percentile (Supplemental Figure 10).

Important XGBoost variables for gestational age

The most important predictors of gestational age were all microbiome taxa or microbiome diversity metrics, except for trimester of specimen collection and maternal MUAC (Supplemental Figure 13). As with other outcomes, several important taxa are associated with resistant-starch degradation, including members of the *Ruminococcaceae, Lachnospiraceae*, and *Eubacteriaceae* families [69,70] and dietary fiber, including *Prevotella copri* [73,74] which predicted a 0.5 week decrease in gestation (Supplemental Figure 13). Another important taxonomic predictor of gestational age was *Slackia isoflavoniconvertens* (1 week increase in gestational age through the range of observed abundances). Species richness also predicted ~0.5 week shorter gestation in enzyme and pathway models (Supplementary Figure 14 and 15).

Discussion

In this study of 207 mothers and their infants in rural Zimbabwe where there is a high prevalence of preterm birth and LBW, we investigated the relationship between pregnancy microbiome characteristics and gestational age, birth weight or neonatal growth. Taxonomic features of the maternal microbiome alone predicted birth weight (pseudo-R^2=0.23), 1mo WAZ (pseudo-R^2=0.21) and LAZ (pseudo-R^2=0.11) as accurately as models that combined microbiome features with sociodemographic and epidemiologic variables, pointing to the importance of the pregnancy microbiome. Gestational age was predicted less accurately (pseudo-R^2=0.05). Prediction accuracy for all outcomes was improved when functional enzyme relative abundances were used (pseudo-R^2: 0.25-0.37), demonstrating the potential value of whole metagenome shotgun sequencing for investigating the human microbiome and health. There were only modest global differences between mothers allocated to the WASH versus the non-WASH arms of the SHINE trial, as measured by β-diversity, and only few differences in taxon abundances after FDR p-value adjustment. In prior analyses, the WASH intervention also had only a modest impact on pathogen carriage and diarrhea in SHINE infants [75]. Surprisingly, the same was true for maternal HIV-status and pregnancy trimester in this substudy.

The maternal fecal microbiome of rural Zimbabwean mothers was highly enriched for metabolizers of resistant-starch, including *Ruminococcus bromii* and *Faecalibacterium prausnitizii*, most likely driven by the daily consumption of maize by all mothers. These starch-degraders and producers of SCFAs were among the most abundant and most prevalent. Starch-degraders release energy from dietary polysaccharides that are not processed by host enzymes, providing

an important nutrient-harvesting function for the host. Although common to many human intestinal microbiomes, they showed important associations with birth weight and neonatal growth in this study population. *Prevotella copri* and *Eubacterium rectale* were also highly abundant species. *P. copri* tends to be more abundant in the microbiomes of populations living non-Westernized lifestyles [76], and have been associated with high fibre diets. *E. rectale* has also been shown to increase, along with Eubacterium-derived butyrate, in subjects given resistant-starch from maize [77]. As with *R. bromii*, and other resistant starch-degraders, *P. copri* and *E. rectale* are likely more abundant due to dietary selection.

However, the direction of taxon abundance relationships with birth weight and neonatal growth was quite varied. Notably, greater relative abundance of *Butyrivibrio sp* CAG 318 and *Roseburia intestinalis*, predicted larger birth weight, while *Catenibacterium mitsuokai* predicted greater 1mo WAZ. These bacteria are butyrate producers with the capacity to degrade the plant fibers hemicellulose [78,79], mannan [80], and inulin [71,81], respectively. Starch-metabolizing, SCFA-producing microbes can regulate gestational weight gain [82], and may impact fetal growth [83]. However, increasing abundance of several other starch-degrading taxa predicted worse outcomes. Different species of bacteria do appear to prefer different sources of dietary fiber [84], suggesting that the potential benefits of SCFA-producing primary starch-degraders may partly depend on both the dietary-fiber source consumed and the resident gut species. The improvement in outcomes predicted by an increase in some taxa may reflect a response of these particular microbes to specific dietary fibers consumed by some mothers in this population. The abundance of specific taxa may also be a marker of dietary deprivation that could explain decreases in birth weight or neonatal growth. Microbiome evenness was also associated with increasing birth weight and 1mo WAZ. Evenness measures the uniformity of the abundance of taxa present in a microbiome. A less even microbiome may reflect selection for only a few predominant taxa by low dietary diversity, resulting in lower birth weight and neonatal WAZ.

Also, there were remarkably few Bacteroidetes, aside from *Prevotella spp*, that emerged in these analyses. The lack of *Bacteroides spp* specifically could indicate a dietary deficiency, as these are associated with a diet rich in meat or fat [73] and may reflect dietary or functional microbiome differences that contribute to pregnancy outcomes and neonatal growth.

Several pregnancy gut microbiome metabolic pathways and enzymes were also important predictors of birth weight, WAZ and LAZ. These broadly reflected starch metabolism, vitamin B metabolism, signalling and environmental sensing. For example, greater relative abundance of the microbiome enzymes for plant fiber degradation (e.g. cellulase, neopullulanase) and pathways for sucrose and fiber degradation (e.g. D-fructuronate degradation) predicted greater birth weight and neonatal growth. These provide further evidence that a microbiome with an increased capacity to degrade the maize-rich, plant-based Zimbabwean diet, is beneficial for healthier growth. Also, greater abundance of histidine kinase predicted larger birth weight. Histidine kinases are important in bacterial sensing, signal transduction and energy utilization [85]. Bacteria that possess many histidine kinases are generally able to adapt to a variety of environmental stimuli [86]. Histidine kinase abundance may reflect a microbiome that is more robust to unstable nutrient availability. In contrast, glycogen synthase and bacterial functions involved in purine metabolism (e.g. ppGpp synthesis, GMP reductase) predicted decreasing birth weight; while signalling enzymes involved in biofilm dispersal (e.g. cytidine deaminase) predicted increasing WAZ. Glycogen synthesis is an important strategy for bacterial survival and persistence in conditions of fluctuating nutrient availability [87], and facilitates transition into a biofilm state in response to bacterial nutrient starvation [88]. Purine and pyrimidine salvage pathways are also potential markers of biofilm formation and dispersal to a planktonic

state [89–93]. The role of these microbial mechanisms in pregnancy outcomes requires further investigation, but these data suggest that a monotonous, nutrient-poor diet, may induce a starvation response in some maternal gut microbiomes that display characteristics of biofilm formation. Microbiome functions related to vitamin B biosynthesis pathways predicted greater birth weight, WAZ or LAZ. This may reflect varied microbiome capacities for vitamin B metabolism in response to deficiencies in the host diet.

An important taxonomic predictor of longer gestation was *Slackia isoflavoniconvertens*. When present, *S. isoflavoniconvertens* converts dietary isoflavones, largely found in soy [94], to equol [95,96]. Equol has been found to reduce anxiety and depressive behaviour in rodent models [97–99], and humans [99]. This mechanism could influence gestation length [100,101]. In contrast, *Prevotella copri* predicted reduced gestation. *P. copri* has been associated with increased gut inflammation, damage [102] and bacterial translocation [103]. Host inflammation drives increased blood pressure [104–106], which is an important risk factor for preterm birth. The inflammatory contribution of the maternal gut microbiome may contribute to reduced gestation in this sub-Saharan African population. These results should be interpreted with caution as the gestational age outcome model exhibited the weakest performance metrics.

Overall, pathogen carriage was low, aside from *Brachyspira* which is associated with intestinal spirochetosis [92]. However, four spirochetes [107] were carried by many mothers and *Treponema berlinense* predicted lower 1mo WAZ. Treponemes have been found in the microbiota of non-Western populations [108–110] and may represent a critical member of an evolutionarily intact intestinal microbiome, similar to *Helicobacter pylori*, another spirochete, which typically resides in the stomach. The contribution of non-pathogenic *Treponema sp* to intestinal microbiome function is unknown. In a macaque model that tested the impact of a high- versus low-fat diet on maternal and offspring microbiota, the largest change was a significant reduction in non-pathogenic *Treponema spp* (*T. berlinense, T. porcinum* and *T. parvum*), in dams administered a high-fat diet, regardless of obese or lean status. The authors suggested that these *Treponema spp* benefited from a low fat, plant-rich diet (corn and soy bean) [111].

There are notable limitations to our analyses. A large fraction of the sequenced reads was un-assignable. The fact that there was a large proportion of unknown or unidentifiable DNA in these samples is not surprising, as microbiome data from LMICs are under-represented in reference databases. However, the sheer scale of un-assignable DNA reads was an important observation and is rarely reported in microbiome studies. *E. coli* was common and predicted decreasing birth weight, but specific *E. coli* pathotypes could not be distinguished by MetaPhlAn3, limiting our ability to assess the impact of these pathogens. Differences in microbiome characteristics could also be influenced by dietary deprivation. However, wealth quintile was not an important predictor in any models.

In conclusion, our analyses illustrate that the pregnancy fecal microbiome, primarily the abundance of resistant-starch degraders, is an important contributor to birth weight and neonatal growth, and to a lesser extent gestational age, in infants of rural Zimbabwean mothers who consume a diet high in maize. The functional capacity for starch metabolism and environmental sensing may be important microbiome mechanisms during pregnancy; while a nutrient-poor diet dominated by maize may promote microbiome biofilm formation in response to nutrient starvation in the gut, with potentially detrimental consequences for the infant. Future work is warranted to confirm these mechanisms and explore whether interventions to facilitate starch degradation and improve dietary diversity can improve birth weight and neonatal growth in populations with a monotonous starch-rich diet.

Contributors

ARM, LES, RJS, LHM, JHH and AJP conceptualized and designed the study. KM, RN, BC, FDM, NVT, JT, and BM collected data and biospecimens. HMG, IB, SKG, RCR, FF and LC processed biospecimens. TJE performed bioinformatics. ARM, TJE and EKG analyzed and interpreted data. ARM and TJE drafted the manuscript. EKG substantially revised the manuscript. ARM, AJP and JHH verified the underlying data. All authors read and approved the final version of the manuscript.

Declaration of Interests

RCR declares monetary support from Nestle Nutrition Institute for conference attendance (April 2019), outside the submitted work. TJE was paid a scientific consulting fee in relation to the analysis of the data presented here by Zvitambo Instititute for Maternal and Child Health Research. All other authors have no interests to declare.

Declaration of Competing Interest

None.

Data sharing statement

All relevant data are within the paper and its Supporting Information files except for the raw data which the trial team will begin loading as individual participant data with an accompanying data dictionary at http://ClinEpiDB.org by mid-2021. This platform is charged with ensuring that epidemiological studies are fully anonymized by removing all personal identifiers and obfuscating all dates per participant through application of a random number algorithm to comply with the ethical conduct of human subjects research. Researchers must agree to the policies and comply with the mechanism of ClinEpiDB to access data housed on this platform. Prior to that time, the data are housed on the ClinEpiDB platform at the Zvitambo Institute for Maternal and Child Health Research and available upon request from Ms. Virginia Sauramba (vsauramba@zvitambo.co.zw).

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.ebiom.2021.103421.

References

[1] Roseboom TJ, van der Meulen JH, Ravelli AC, Osmond C, Barker DJ, Bleker OP. Effects of prenatal exposure to the Dutch famine on adult disease in later life: an overview. Mol Cell Endocrinol 2001;185:93–8.

[2] Blencowe H, Krasevec J, de Onis M, et al. National, regional, and worldwide estimates of low birthweight in 2015, with trends from 2000: a systematic analysis. Lancet Glob Heal 2019;7:e849–60.

[3] Chawanpaiboon S, Vogel JP, Moller AB, et al. Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. Lancet Glob Heal 2019;7:e37–46.

[4] da Silva Lopes K, Ota E, Shakya P, et al. Effects of nutrition interventions during pregnancy on low birth weight: an overview of systematic reviews. BMJ Glob Heal 2017;2:e000389.

[5] Shiozaki A, Yoneda S, Yoneda N, et al. Intestinal microbiota is different in women with preterm birth: Results from terminal restriction fragment length polymorphism analysis. PLoS One 2014;9:1–7.

[6] Vinturache AE, Gyamfi-Bannerman C, Hwang J, Mysorekar IU, Jacobsson B. Maternal microbiome - A pathway to preterm birth. Semin Fetal Neonatal Med 2016;21:94–9.

[7] Gomez-Arango LF, Barrett HL, McIntyre HD, et al. Connections between the gut microbiome and metabolic hormones in early pregnancy in overweight and obese women. Diabetes 2016;65:2214–23.

[8] Gomez-Arango LF, Barrett HL, McIntyre HD, Callaway LK, Morrison M, Dekker Nitert M. Increased Systolic and Diastolic Blood Pressure is Associated with Altered Gut Microbiota Composition and Butyrate Production in Early Pregnancy. Hypertension 2016;68:974–81.

[9] DiGiulio DB, Callahan BJ, McMurdie PJ, et al. Temporal and spatial variation of the human microbiota during pregnancy. Proc Natl Acad Sci 2015;112:11060–5.

[10] Goltsman DSA, Sun CL, Proctor DM, et al. Metagenomic analysis with strain-level resolution reveals fine-scale variation in the human pregnancy microbiome. Genome Res 2018;28:1467–80.

[11] Koren O, Goodrich JK, Cullender TC, et al. Host remodeling of the gut microbiome and metabolic changes during pregnancy. Cell 2012;150:470–80.

[12] Gosalbes MJ, Compte J, Moriano-Gutierrez S, et al. Metabolic adaptation in the human gut microbiota during pregnancy and the first year of life. EBioMedicine 2019;39:497–509.

[13] Yang H, Guo R, Li S, et al. Systematic analysis of gut microbiota in pregnant women and its correlations with individual heterogeneity. npj Biofilms Microbiomes 2020;6:1–12.

[14] Brown EM, Ke X, Hitchcock D, et al. Bacteroides-Derived Sphingolipids Are Critical for Maintaining Intestinal Homeostasis and Symbiosis. Cell Host Microbe 2019;25:668–80 e7.

[15] An D, Na C, Bielawski J, Hannun YA, Kasper DL. Membrane sphingolipids as essential molecular signals for Bacteroides survival in the intestine. Proc Natl Acad Sci 2014;108:4666–71.

[16] Van Den Bogert. B Meijerink M, Zoetendal EG, Wells JM, Kleerebezem M. Immunomodulatory properties of Streptococcus and Veillonella isolates from the human small intestine microbiota. PLoS One 2014;9:1–20.

[17] Terefe Y, Deblais L, Ghanem M, et al. Co-occurrence of Campylobacter Species in Children From Eastern Ethiopia, and Their Association With Environmental Enteric Dysfunction, Diarrhea, and Host Microbiome. Front public Heal 2020;8:99.

[18] Chen RY, Kung VL, Das S, et al. Duodenal Microbiota in Stunted Undernourished Children with Enteropathy. N Engl J Med 2020;383:321–33.

[19] Shivakoti R, Gupte N, Kumar NP, et al. Intestinal Barrier Dysfunction and Microbial Translocation in Human Immunodeficiency Virus–Infected Pregnant Women Are Associated With Preterm Birth. Clin Infect Dis 2018;67:1103–9.

[20] López M, Figueras F, Coll O, et al. Inflammatory Markers Related to Microbial Translocation Among HIV-Infected Pregnant Women: A Risk Factor of Preterm Delivery. J Infect Dis 2016;213:343–50.

[21] Lauer JM, Duggan CP, Ausman LM, et al. Biomarkers of maternal environmental enteric dysfunction are associated with shorter gestation and reduced length in newborn infants in Uganda. Am J Clin Nutr 2018;108:889–96.

[22] David LA, Weil A, Ryan ET, et al. Gut microbial succession follows acute secretory diarrhea in humans. MBio 2015;6.

[23] Padhi BK, Baker KK, Dutta A, et al. Risk of adverse pregnancy outcomes among women practicing poor sanitation in rural India: A population-based prospective cohort study. PLoS Med 2015;12:1–18.

[24] Liu J, Platts-Mills JA, Juma J, et al. Use of quantitative molecular diagnostic methods to identify causes of diarrhoea in children: a reanalysis of the GEMS case-control study. Lancet 2016;388:1291–301.

[25] Poston L, Harthoorn LF, Van Der Beek EM. Contributors to the ILSI Europe Workshop. Obesity in pregnancy: implications for the mother and lifelong health of the child. A consensus statement. Pediatr Res 2011;69:175–80.

[26] Indrio F, Martini S, Francavilla R, et al. Epigenetic Matters: The Link between Early Nutrition, Microbiome, and Long-term Health Development. Front Pediatr 2017;5:1–14.

[27] Johnson AJ, Vangay P, Al-Ghalith GA, et al. Daily Sampling Reveals Personalized Diet-Microbiome Associations in Humans. Cell Host Microbe 2019;25:789–802 e5.

[28] Maier T V, Lucio M, Lee LH, et al. Impact of Dietary Resistant Starch on the Human Gut Microbiome, Metaproteome, and Metabolome. MBio 2017;8.

[29] Schwarzer M, Strigini M, Leulier F. Gut Microbiota and Host Juvenile Growth. Calcif Tissue Int 2017;0:1–19.

[30] Natarajan N, Hori D, Flavahan S, et al. Microbial short chain fatty acid metabolites lower blood pressure via endothelial G protein-coupled receptor 41. Physiol Genomics 2016;48:826–34.

[31] MF Z, Erin N, Po-Yin C, et al. High-Fiber Diet and Acetate Supplementation Change the Gut Microbiota and Prevent the Development of Hypertension and Heart Failure in Hypertensive Mice. Circulation 2017;135:964–77.

[32] Lv L-J, Li S-H, Li S-C, et al. Early-Onset Preeclampsia Is Associated With Gut Microbial Alterations in Antepartum and Postpartum Women. Front Cell Infect Microbiol 2019;9:224.

[33] Hofmeyr GJ, MR Torloni. Calcium supplementation during pregnancy for preventing hypertensive disorders and related problems. Cochrane Database Syst Rev 2018.

[34] Rossi M, Amaretti A, Raimondi S. Folate Production by Probiotic Bacteria. Nutr 2011;3(1):118–34.

[35] Engevik MA, Morra CN, Röth D, et al. Microbial Metabolic Capacity for Intestinal Folate Production and Modulation of Host Folate Receptors. Front Microbiol 2019;10:2305.

[36] Villalpando S. Discussion: Effects of folate and vitamin B12 deficiencies during pregnancy on fetal, infant, and child development. Food Nutr Bull 2008;29:101–11.

[37] Santacruz A, Collado MC, García-Valdés L, et al. Gut microbiota composition is associated with body weight, weight gain and biochemical parameters in pregnant women. Br J Nutr 2010;104:83–92.

[38] Chambers ES, Viardot A, Psichas A, et al. Effects of targeted delivery of propionate to the human colon on appetite regulation, body weight maintenance and adiposity in overweight adults. Gut 2015;64:1744–54.

[39] van der Beek CM, Canfora EE, Lenaerts K, et al. Distal, not proximal, colonic acetate infusions promote fat oxidation and improve metabolic markers in overweight/obese men. Clin Sci 2016;130:2073–82.

[40] Ilmonen J, Isolauri E, Poussa T, Laitinen K. Impact of dietary counselling and probiotic intervention on maternal anthropometric measurements during and after pregnancy: A randomized placebo-controlled trial. Clin Nutr 2011;30:156–64.

[41] Collado MC, Isolauri E, Laitinen K, Salminen S. Distinct composition of gut microbiota during pregnancy in overweight and normal-weight women. Am J Clin Nutr 2008;88:894–9.

[42] Ferretti P, Pasolli E, Tett A, et al. Mother-to-Infant Microbial Transmission from Different Body Sites Shapes the Developing Infant Gut Microbiome. Cell Host Microbe 2018;24:133–45.

[43] Yassour M, Jason E, Hogstrom LJ, et al. Strain-Level Analysis of Mother-to-Child Bacterial Transmission during the First Few Months of Life. Cell Host Microbe 2018;24:146–54 e4.

[44] Smith MI, Yatsunenko T, Manary MJ, et al. Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. Science 2013;339(80):548–54.

[45] Subramanian S, Huq S, Yatsunenko T, et al. Persistent gut microbiota immaturity in malnourished Bangladeshi children. Nature 2014;510:417–21.

[46] Blanton L V, Charbonneau MR, Salih T, et al. Gut bacteria that prevent growth impairments transmitted by microbiota from malnourished children. Science 2016;351.

[47] Bhagavata Srinivasan SP, Raipuria M, Bahari H, Kaakoush NO, Morris MJ. Impacts of Diet and Exercise on Maternal Gut Microbiota Are Transferred to Offspring. Front Endocrinol (Lausanne) 2018;9:716.

[48] Jost T, Lacroix C, Braegger C, Chassard C. Stability of the maternal gut microbiota during late pregnancy and early lactation. Curr Microbiol 2014;68:419–27.

[49] Pasolli E, Asnicar F, Manara S, et al. Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle. Cell 2019;176:649–62.

[50] Humphrey JH, Mbuya MNN, Ntozini R, et al. Independent and combined effects of improved water, sanitation, and hygiene, and improved complementary feeding, on child stunting and anaemia in rural Zimbabwe: a cluster-randomised trial. Lancet Glob Heal 2018;7:e132–47.

[51] Humphrey JH, Jones AD, Manges A, et al. The sanitation hygiene infant nutrition efficacy (SHINE) Trial: Rationale, design, and methods. Clin Infect Dis 2015;61: S685–702.

[52] Gough EK, Prendergast AJ, Mutasa KE, Stoltzfus RJ, Manges AR. Assessing the intestinal microbiota in the SHINE Trial. Clin Infect Dis 2015;61:S738–44.

[53] Huttenhower Lab. GitHub - biobakery/kneaddata: Quality control tool on metagenomic and metatranscriptomic sequencing data, especially data from microbiome experiments. Available at: https://github.com/biobakery/kneaddata. Accessed 8 April 2021.

[54] Truong DT, Franzosa EA, Tickle TL, et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat Methods 2015;12:902–3.

[55] Abubucker S, Segata N, Goll J, et al. Metabolic Reconstruction for Metagenomic Data and Its Application to the Human Microbiome. PLoS Comput Biol 2012;8: e1002358.

[56] RStudio Team. RStudio: Integrated Development for R. Boston, MA: RStudio, PBC; 2020, http://www.rstudio.com/.

[57] De Onis M, Garza C, Onyango A, Martorell R. WHO child growth standards. Acta Paediatr 2006;95:1–104.

[58] Philip D. Computer program review VEGAN, a package of R functions for community ecology. J Veg Sci 2003;14:927–30.

[59] Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral Ecol 2001;26:32–46.

[60] Stasinopoulos DM, Rigby RA. Generalized additive models for location scale and shape (GAMLSS) in R. J Stat Softw 2007;23:1–46.

[61] Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B 1995;57:289–300.

[62] Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016. p. 785–94.

[63] Zeevi D, Korem T, Zmora N, et al. Personalized Nutrition by Prediction of Glycemic Responses. Cell 2015;163:1079–95.

[64] Ghaemi MS, DiGiulio DB, Contrepois K, et al. Multiomics modeling of the immunome, transcriptome, microbiome, proteome and metabolome adaptations during human pregnancy. Bioinformatics 2018;35:95–103.

[65] Zhao Q, Hastie T. Causal Interpretations of Black-Box Models. J. Bus. Econ. Stat. 2021;39:272–81.

[66] Apley DW, Zhu J. Visualizing the effects of predictor variables in black box supervised learning models. J R Stat Soc Ser B (Statistical Methodol) 2020;82:1059–86.

[67] Nielsen HB, Almeida M, Juncker AS, et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. Nat Biotechnol 2014;32:822–8.

[68] Angelakis E, Bachar D, Yasir M, et al. Treponema species enrich the gut microbiota of traditional rural populations but are absent from urban individuals. New Microbes New Infect 2019;27:14–21.

[69] Zhang L, Ouyang Y, Li H, et al. Metabolic phenotypes and the gut microbiota in response to dietary resistant starch type 2 in normal-weight subjects: a randomized crossover trial. Sci Rep 2019;9:1–11.

[70] Vital M, Howe A, Bergeron N, Krauss RM, Jansson JK, Tiedje JM. Metagenomic insights into the degradation of resistant starch by human gut microbiota. Appl Environ Microbiol 2018;84 e01562-18.

[71] Yan H, Potu R, Lu H, et al. Dietary Fat Content and Fiber Type Modulate Hind Gut Microbial Community and Metabolic Markers in the Pig. PLoS One 2013;8: e59581.

[72] Shin JH, Sim M, Lee JY, Shin DM. Lifestyle and geographic insights into the distinct gut microbiota in elderly women from two different geographic locations. J Physiol Anthropol 2016;35:31.

[73] Wu GD, Chen J, Hoffmann C, et al. Linking long-term dietary patterns with gut microbial enterotypes. Science 2011;334(80):105–8.

[74] De Filippo C, Cavalieri D, Di Paola M, et al. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proc Natl Acad Sci 2010;107:14691–6.

[75] Rogawski McQuade ET, Platts-Mills JA, Gratz J, et al. Impact of water quality, sanitation, handwashing, and nutritional interventions on enteric infections in rural Zimbabwe: the Sanitation Hygiene Infant Nutrition Efficacy (SHINE) Trial. J Infect Dis 2020;221:1379–86.

[76] Tett A, Huang KD, Asnicar F, et al. The Prevotella copri Complex Comprises Four Distinct Clades Underrepresented in Westernized Populations. Cell Host Microbe 2019;26:666–79 e7.

[77] Baxter NT, Schmidt AW, Venkataraman A, Kim KS, Waldron C, Schmidt TM. Dynamics of human gut microbiota and short-chain fatty acids in response to dietary interventions with three fermentable fibers. MBio 2019:10.

[78] Palevich N, Kelly WJ, Leahy SC, et al. Comparative Genomics of Rumen Butyrivibrio spp. Uncovers a Continuum of Polysaccharide-Degrading Capabilities. Appl Environ Microbiol 2019;86.

[79] Emerson EL, Weimer PJ. Fermentation of model hemicelluloses by Prevotella strains and Butyrivibrio fibrisolvens in pure culture and in ruminal enrichment cultures. Appl Microbiol Biotechnol 2017;101:4269–78.

[80] La Rosa SL, Leth ML, Michalak L, et al. The human gut Firmicute Roseburia intestinalis is a primary degrader of dietary β-mannans. Nat Commun 2019;10:1–14.

[81] Hiel S, Gianfrancesco MA, Rodriguez J, et al. Link between gut microbiota and health outcomes in inulin-treated obese patients: Lessons from the Food4Gut multicenter randomized placebo-controlled trial. Clin Nutr 2020;39:3618–28.

[82] Kimura I, Miyamoto J, Ohue-Kitano R, et al. Maternal gut microbiota in pregnancy influences offspring metabolic phenotype in mice. Science 2020;367(80).

[83] Priyadarshini M, Thomas A, Reisetter AC, et al. Maternal short-chain fatty acids are associated with metabolic parameters in mothers and newborns. Transl Res 2014;164:153–7.

[84] Patnode ML, Beller ZW, Han ND, et al. Interspecies Competition Impacts Targeted Manipulation of Human Gut Bacteria by Fiber-Derived Glycans. Cell 2019;179:59–73.

[85] Mascher T, Helmann JD, Unden G. Stimulus Perception in Bacterial Signal-Transducing Histidine Kinases. Microbiol Mol Biol Rev 2006;70:910–38.

[86] Kim D, Forst S. Genomic analysis of the histidine kinase family in bacteria and archaea. Microbiology 2001;147:1197–212.

[87] Wilson WA, Roach PJ, Montero M, et al. Regulation of glycogen metabolism in yeast and bacteria. FEMS Microbiol Rev 2010;34:952–85.

[88] Sekar K, Linker SM, Nguyen J, Grünhagen A, Stocker R, Sauer U. Bacterial Glycogen Provides Short-Term Benefits in Changing Environments. Appl Environ Microbiol 2020;86.

[89] Kim H-M, Davey ME. Synthesis of ppGpp impacts type IX secretion and biofilm matrix formation in Porphyromonas gingivalis. npj Biofilms Microbiomes 2020;6:5.

[90] Yoshioka S, Newell PD. Disruption of de novo purine biosynthesis in Pseudomonas fluorescens Pf0-1 leads to reduced biofilm formation and a reduction in cell size of surface-attached but not planktonic cells. PeerJ 2016;4 e1543−e1543.

[91] Garavaglia M, Rossi E, Landini P. The Pyrimidine Nucleotide Biosynthetic Pathway Modulates Production of Biofilm Determinants in Escherichia coli. PLoS One 2012;7:e31252.

[92] Thorell K, Inganäs L, Backhans A, et al. Isolates from Colonic Spirochetosis in Humans Show High Genomic Divergence and Potential Pathogenic Features but Are Not Detected Using Standard Primers for the Human Microbiota. J Bacteriol 2019;201 e00272-19.

[93] Ha D-G, O'Toole GA. c-di-GMP and its effects on biofilm formation and dispersion: a Pseudomonas aeruginosa review. Microbiol Spectr 2015;3.

[94] Patisaul HB, Jefferson W. The pros and cons of phytoestrogens. Front. Neuroendocrinol. 2010;31:400–19.

[95] Matthies A, Loh G, Blaut M, Braune A. Daidzein and Genistein Are Converted to Equol and 5-Hydroxy-Equol by Human Intestinal Slackia isoflavoniconvertens in Gnotobiotic Rats. J Nutr 2012;142:40–6.

[96] Schröder C, Matthies A, Engst W, Blaut M, Braune A. Identification and expression of genes involved in the conversion of daidzein and genistein by the equol-forming bacterium Slackia isoflavoniconvertens. Appl Environ Microbiol 2013;79:3494–502.

[97] Blake C, Fabick KM, Setchell KDR, Lund TD, Lephart ED. Neuromodulation by soy diets or equol: Anti-depressive & anti-obesity-like influences, age- & hormone-dependent effects. BMC Neurosci 2011;12:28.

[98] Blake C, Fabick KM, Setchell KDR, Lund TD, Lephart ED. Prenatal exposure to equol decrease body weight and depressive-like behaviors in male and female offspring. Curr Top Nutraceuticals Res 2010;8:69–77.

[99] Bax EN, Cochran KE, Mao J, Wiedmeyer CE, Rosenfeld CS. Opposing effects of S-equol supplementation on metabolic and behavioral parameters in mice fed a high-fat diet. Nutr Res 2019;64:39–48.

[100] Dunkel Schetter C, Tanner L. Anxiety, depression and stress in pregnancy: Implications for mothers, children, research, and practice. Curr. Opin. Psychiatry. 2012;25:141–8.

[101] Ding XX, Le Wu Y, Xu SJ, et al. Maternal anxiety during pregnancy and adverse birth outcomes: A systematic review and meta-analysis of prospective cohort studies. J. Affect. Disord. 2014;159:103–10.

[102] Yu C, Zhou B, Xia X, et al. Prevotella copri is associated with carboplatin-induced gut toxicity. Cell Death Dis 2019;10(10):714.

[103] Armstrong AJS, Shaffer M, Nusbacher NM, et al. An exploration of Prevotella-rich microbiomes in HIV and men who have sex with men. Microbiome 2018;6:198.

[104] SE Y, Janice R, Alper Y, Neil GD, GF NE. Novel Role of T Cells and IL-6 (Interleukin-6) in Angiotensin II−Induced Microvascular Dysfunction. Hypertension 2019;73:829–38.

[105] Weiru Z, Wei W, Hong Y, et al. Interleukin 6 Underlies Angiotensin II−Induced Hypertension and Chronic Renal Damage. Hypertension 2012;59:136–44.

[106] Bartolomaeus H, Balogh A, Yakoub M, et al. Short-Chain Fatty Acid Propionate Protects from Hypertensive Cardiovascular Damage. Circulation 2019;139:1407–21.

[107] Hampson DJ. The Spirochete Brachyspira pilosicoli, Enteric Pathogen of Animals and Humans. Clin Microbiol Rev 2017;31:1–32.

[108] Obregon-Tito AJ, Tito RY, Metcalf J, et al. Subsistence strategies in traditional societies distinguish gut microbiomes. Nat Commun 2015;6:1–9.

[109] Schnorr SL, Candela M, Rampelli S, et al. Gut microbiome of the Hadza hunter-gatherers. Nat Commun 2014;5:3654.

[110] Gomez A, Petrzelkova KJ, Burns MB, et al. Gut Microbiome of Coexisting BaAka Pygmies and Bantu Reflects Gradients of Traditional Subsistence Patterns. Cell Rep 2016;14:2142–53.

[111] Ma J, Prince AL, Bader D, et al. High-fat maternal diet during pregnancy persistently alters the offspring microbiome in a primate model. Nat Commun 2014;5:1–11.