Originally published 1 February 2013; corrected 25 April 2013

www.sciencemag.org/cgi/content/full/339/6119/548/DC1

Supplementary Materials for

Gut Microbiomes of Malawian Twin Pairs Discordant for Kwashiorkor

Michelle I. Smith, Tanya Yatsunenko, Mark J. Manary, Indi Trehan, Rajhab Mkakosya, Jiye Cheng, Andrew L. Kau, Stephen S. Rich, Patrick Concannon, Josyf C. Mychaleckyj, Jie Liu, Eric Houpt, Jia V. Li, Elaine Holmes, Jeremy Nicholson, Dan Knights, Luke K. Ursell, Rob Knight, Jeffrey I. Gordon*

*To whom correspondence should be addressed. E-mail: jgordon@wustl.edu

Published 1 February 2013, *Science* **339**, 548 (2013)
DOI: 10.1126/science.1229000

This PDF file includes:

Materials and Methods
Supplementary Text
Figs. S1 to S15
Full Reference List

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencemag.org/cgi/content/full/339/6119/548/DC1)

Tables S1 to S10 (as Excel files)

Corrections: Three typos in two of the supplementary tables have been corrected. These typos do not affect any of the results described in the paper or their interpretation. The corrections are as follows:

1) Table S1A: In the row titled “Different gender,” column “Total number of twin pairs,” the percentage in the brackets was changed from 3.7% to 33.7%. The raw numbers are provided in the table to calculate the percentage.

2) Table S2A: The column headers “HAZ” and “WAZ” have been switched. In accordance with standard practice in the field, the classification of moderate acute malnutrition was based on WHZ scores, not HAZ or WAZ scores. The

classification of severe acute malnutrition was also based on WHZ scores (marasmus) and/or the presence of bilateral pedal edema (kwashiorkor). HAZ and WAZ scores were simply provided in this table to give the reader additional anthropometric information about the twins. Note that the average values for Z scores presented in table S1C, as well as the scores shown in fig. S5, were reported correctly in the paper.

3) Table S2B: The information about sample h196B.M2.D07.h8.F has been removed. This sample was not used in our analyses.

MATERIALS AND METHODS

Studies in humans

Enrollment, clinical data and sample collection – Subjects were recruited for the present study using procedures approved by the College of Medicine Research Ethics Committee of the University of Malawi, and by the Human Research Protection Office of Washington University in St. Louis. Twin pairs were recruited through health centers located in the Malawian villages of Makhwira, Mitondo, M'biza, Chamba, and Mayaka. A team of one USA pediatrician and a minimum of two trained local personnel visited each site every month where the weight and height of each infant or child was measured in three replicates, and each child was checked for bilateral pitting edema. On average, twin pairs where one or both co-twins developed kwashiorkor received 9±8 (mean ± SD) weeks of RUTF treatment, while those with marasmus were treated for 11±8 weeks. RUTF was produced in Malawi (*38*).

During each visit for a scheduled fecal sample collection, each child wore a commercial disposable diaper lined with plastic. Each fecal specimen was flash frozen in a cryogenic storage container, filled with liquid nitrogen, within 10 min after it was produced. All samples were subsequently stored at -80^{o}C prior to analyses. For twin pairs who remained concordant for a well nourished phenotype, fecal samples were collected on average every 3 months, yielding 5±1 (mean ± SD) samples/individual. In the case of twins who became discordant for kwashiorkor, sampling was increased to every 2 weeks during the period of RUTF treatment, producing 8±3 samples per child.

Zygosity tests – Buccal smears were collected using kits from Oragene. We selected 48 autosomal SNPs with high heterozygosity in the Yoruban HapMap sample (n=90) from the Illumina DNA Test Panel of 360 SNP loci that have been optimized for the BeadExpress. The SNP minor allele frequencies in Yorubans ranged from 0.28-0.43, and each autosome was sampled by at least one marker (with the exception of 7,19 and 21), with a minimum marker separation of 3Mb. We ran PREST [Pedigree Relationship Statistical Test (*39*)] to estimate the identity-by-descent sharing and kinship between the relative pairs, and compared reported relationships versus the inferred relationships from the genetic data.

Criteria for selection of 9 well nourished twin pairs and 13 twin pairs discordant for kwashiorkor – Selection of twins in the well nourished group was based on the following criteria: (i) both co-twins in a twin pair had to remain well nourished (a number of twin pairs had to be removed from the original selection due to disease occurring at later time points); (ii) their ages spanned the period from birth through three years; and (iii) the availability of at least four samples from each child, representing one year of the child's life. Among the 19 same gender twin pairs discordant for kwashiorkor, one of the co-twins died during the study in each of three families. In the case of another three families, samples became available much later in the study and could not be included in the analysis. Therefore, we focused on 13 twin pairs where both co-twins survived and enough fecal samples were available for our analyses. Well nourished children were in the same age group as twin pairs discordant for kwashiorkor: the first group included children sampled between three weeks and 27 months of age, whereas the latter group was sampled between three weeks and 30 months.

Isolation and sequencing of fecal microbial community DNA – Fecal samples from humans were pulverized with a mortar and pestle in liquid nitrogen. DNA was extracted from aliquots of the pulverized material (*40*). Amplicons, generated by PCR of variable region 4 (V4) of bacterial 16S rRNA genes present in the fecal DNA samples were subjected to multiplex sequencing with an Illumina HiSeq instrument using procedures described in a previous publication (*41*). Multiplex shotgun 454 Titanium FLX pyrosequencing of each fecal community DNA sample was conducted according to ref. (*10*).

Analysis of metagenomic datasets

16S rRNA – We used QIIME (*42*) to cluster V4-16S rRNA Illumina reads at 97% nucleotide sequence identity (97 %ID). A closed-reference OTU (operational taxonomic unit) picking protocol with uclust against the Greengenes database was employed. Reads that did not match the database were excluded from further analyses. Of the 110,819,831 Illumina reads that passed QIIME's quality filters, 93.5% (103,583,462) hit a reference sequence at greater than or equal to 97% ID. For all downstream analysis, 16S rRNA datasets were rarefied to 116,440 V4-16S rRNA reads/sample. Taxonomy was assigned using RDP classifier 2.4. Summarize_taxa.py script in QIIME was used to calculate the relative abundance of phylum-, class-, order-, genus-, species-level taxa. UniFrac distances between samples were calculated using a *de novo* tree constructed from representative sequences defining each OTU. A table of OTU counts per sample was generated and used in combination with the tree to calculate beta diversity.

Shotgun sequences from fecal microbiomes – Shotgun reads produced by 454 Titanium shotgun sequencing were filtered as described in ref. (*10)*. In addition, searches against a database of the 462 human gut bacterial genomes (listed in the **table S3**) were conducted with megablast. A sequence read was annotated as the best hit to this database if the E-value was $\leq 10^{-20}$, and at least 90% identical between query and subject. The relative abundances of reads that mapped to each of the 462 genomes were adjusted to genome sizes. For sequences that mapped to more than one genome, counts were adjusted according to unique matches using a previously described procedure (*43*): for example, in the case of a non-unique read that equally matched genome A and genome B, and where two unique reads matched genome A and eight unique reads matched genome B, 0.2 counts of the non-unique read would be assigned to genome A and 0.8 counts to genome B.

Searches against protein-coding component of the KEGG database (version 58) were conducted with BLASTX. Counts were normalized to the mapped reads. 26±4% of the reads mapped to KEGG ECs and 55±14% to the 462 reference human gut genomes. Unmapped reads were excluded from the analyses shown, although repeating the analyses including these reads had little effect on the results. To quantify the differences in KEGG EC profiles among fecal microbiomes, evenly rarefied matrices of EC counts were created with all samples, and Hellinger distances were calculated using QIIME. t-test and linear mixed-effects regression analyses were conducted in R statistical software using package NLME (*44*) for the latter analysis.

Studies in gnotobiotic mice

Colonization of germ-free animals – All experiments involving mice were performed using protocols approved by the Washington University Animal Studies Committee. Germ-free 8-week-old, male C57BL/6J mice were maintained in flexible plastic gnotobiotic isolators (Class Biologically Clean Ltd., Madison, WI) under a strict 12-hour light cycle (lights on at 0600h) and fed *ad libitum*. Transplantation of human fecal microbiota was performed by first pulverizing the frozen fecal sample with mortar and pestle in liquid nitrogen. Approximately 200 µg of the pulverized sample was diluted in 4 mL of sterile PBS/2mM DTT at room temperature in an anaerobic chamber (Coy Lab Products, Grass Lake, MI; atmosphere composed of 75% N_2/20% CO_2/5% H_2). The fecal material was suspended by vortexing (3 min at room temperature) and the resulting suspension was allowed to settle by gravity (2 min at room temperature in the anaerobic chamber). The supernatant was placed in a sterile polypropylene tube that was capped and placed in a Balch tube. The Balch tube was then capped, taken out of the Coy chamber and transported, at room temperature, to our gnotobiotic mouse facility (~10 minute trip). The polypropylene tube was then removed from the protective Balch tube, placed in a transfer sleeve attached to the gnotobiotic isolator and its external surface was sterilized during a 20 min exposure to 2% chlorine dioxide. Once introduced into the gnotobiotic isolator, the tube was opened, and a ~200 □L aliquot was gavaged, using a flexible plastic tube attached to a sterile syringe, directly into the stomach of each germ-free recipient mouse (n=5-10 mice per donor microbiota preparation).

All mice that received a fecal microbiota sample from a given human donor were co-housed (n=5/cage) in a given gnotobiotic isolator. Different isolators were used for different donors. The sterility of germ-free mice and their isolators was established by routine culture-based surveys of fecal samples as well as by PCR of fecal DNA using universal bacterial 16S rRNA primers.

Preparation of human diets given to gnotobiotic mice – Ingredients for constructing a representative Malawian diet were purchased from vendors in the USA. Meseca® instant corn flour was obtained from Restaurant Depot (College Point, NY). Organic mustard greens, yellow onions and vine-ripened tomatoes were purchased fresh from Whole Foods Supermarkets. Ground peanuts were from East Wind Nut Butters (Tecumseh, MO).

Food was prepared in 18 kg batches. Four kg of corn flour were added to 12.3 L of freshly autoclaved water and blended for 5 min using an industrial mixer (Globe SP30P 30-quart pizza mixer; gear speed 1; Globe Food Equipment Company, Dayton, OH). This step was followed by addition of 1.63 kg of a relish. This relish was produced by adding 770 g of organic mustard greens to 154 g of organic yellow peeled onions, 385 g of organic vine-ripened tomatoes and 321 mL of water. The mixture was then pureed for 10 min in a vertical cutter mixer (Robot Coupe Model R23, Jackson, MS) and simmered for 30 min on a Corning stirrer/hot plate (high setting) prior to addition to the corn-water solution. Following addition of the relish, food was blended in the mixer for 5 min. Two to three batches of food were prepared in succession over the course of a day, with each batch being temporarily stored in covered plastic containers. At the end of the day, subsets of each batch were mixed together so that there was a homogeneous product. The finished product was then vacuumed packed in 500 g aliquots in FDA/USDA-compliant poly-nylon vacuum pouches, double-bagged, and sterilized by irradiation (20-

50 kGy) within 24 h of its production (Steris Co; Chicago, IL). The macro- and micronutrient content of the cooked and irradiated Malawian diet was defined by N.P. Analytical Laboratories (St Louis, MO) (**table S6**). Food was stored at 4°C for up to six months. Sterility was verified using the same culture methods employed to survey gnotobiotic isolators (see above).

Peanut-based RUTF was produced by Harlan Laboratories (Madison, WI) using a recipe described in a previous publication (*38*) and ingredients purchased in the USA. RUTF was packaged and sterilized as above.

Plastic bags containing 500 g of a given diet were introduced into each gnotobiotic isolator after sterilizing the external surface with 2% chlorine dioxide. A small incision was placed at one corner of each bag so that the food, which had the consistency of paste, could be extruded into sterile glass petri dishes placed in two locations in each cage. Food was refreshed twice per day, allowing for *ad libitum* feeding.

Control groups of mice harboring a kwashiorkor or healthy co-twin microbiota were fed an autoclaved mouse chow diet (B&K Universal, East Yorkshire, U.K; diet 7378000) *ad libitum*.

Sample collection – Fecal samples collected from mice were frozen and stored at 80°C within 30 min of their production. Fecal and urine samples collected for metabolomic analyses were frozen in liquid nitrogen within 2-10 min after they were produced and subsequently stored at -80°C. Just prior to sacrifice, blood was obtained by retro-orbital phlebotomy. Cecal samples were collected and frozen in liquid nitrogen immediately after sacrifice.

Isolation and sequencing of mouse fecal DNA – Isolation of fecal DNA, bacterial V4-16S rRNA amplicon and whole community shotgun sequencing, and analyses of the resulting datasets were conducted as described above for human samples.

PCR-Luminex assays for enteropathogens

DNA isolated from human donor fecal samples and mouse fecal pellets was tested for 22 enteropathogens using custom, multiplex PCR-Luminex assays (*15-18*). These assays were designed to detect Adenovirus, *Ancylostoma duodenale*, *Ascaris lumbricoides*, *Campylobacter jejuni/coli*, *Cryptosporidium* spp., *Cyclospora cayetanensis*, *Cystoisospora belli*, *Enterocytozoon bieneusi*, *Encelphalitozoon intestinalis*, enteroaggregative *E. coli* (EAEC), enterohemorrhagic *E. coli* (EHEC), enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC), enteroinvasive *E. coli* (EIEC)/*Shigella* spp., *Entamoeba histolytica*, *Giardia lamblia*, *Necator americanus*, *Salmonella* spp., *Strongyloides stercoralis*, *Trichuris trichiura*, *Vibrio cholera/parahaemolyticus* and *Yersinia* spp. Atypical EPEC was inferred if *eae* was present without *bfpa* (characteristic of typical EPEC) or *stx1/stx2* (present with EHEC). Our assay for the heat-stable enterotoxin of ETEC detects *STh*, while *ipaH* detects but does not discriminate between enteroinvasive *E. coli* (EIEC) and *Shigella* spp.

Metabolomic analyses

Targeted gas chromatography-mass spectrometry of short chain fatty acids (SCFA) – Fecal pellets and cecal samples were weighed in 2 mL polytetrafluoroethylene (PFTE) screw cap vials. A mix of internal standards (20 □M of acetic acid-$^{13}C_2$,d_4,

propionic acid-d_6, butyric acid-$^{13}C_4$, lactic acid-3,3,3-d_3 and succinic acid-$^{13}C_4$ in a volume of 10 µL) was added to each vial, followed by 20 µL of 33% HCl. Diethyl ether (1 mL) was introduced and the mixture vortexed vigorously for 10 min. The two phases were separated by centrifugation (4,000 x g for 5 min at room temperature). The upper organic layer was transferred into another clear vial and a second diethyl ether extraction was performed. After combining the upper organic layer from the two ether extractions, 60 µL of the ether extract and 20 µL of *N-tert*-butyldimethylsilyl-*N*-methyltrifluoroacetamide (MTBSTFA) were mixed in a 100 µL glass insert in a GC auto-sampler vial and incubated for 2 h at room temperature. Samples were analyzed in a randomized order.

Derivatized samples (1 µL) were injected with 15:1 split into an Agilent 7890A gas chromatography system coupled with 5975C mass spectrometer detector (Agilent, CA). Analyses were carried on a HP-5MS capillary column (30 m x 0.25 mm, 0.25 µm film thickness, Agilent J&W Scientific, Folsom, CA) using electronic impact (70 eV) as ionization mode. Helium was used as a carrier gas at a constant flow rate of 1.26 mL/min and the solvent delay time was set to 3.5 min. The column head-pressure was 10 p.s.i.. The temperatures of injector, transfer line, and quadrupole were 270°C, 280°C and 150°C, respectively. The GC oven was programmed as follows: 45°C hold for 2.25 min; increase to 200°C at a rate of 20°C/min; increase to 300°C at a rate of 100°C/min; hold for 3 min.

Quantification of SCFAs was performed by isotope dilution GC–MS using selected ion monitoring (SIM). For SIM analysis, the m/z for native and labeled molecular peaks for SCFA quantified were: 117 and 122 (acetate), 131 and 136 (propionate), 145 and 149 (butyrate), 261 and 264 (lactate) and 289 and 293 (succinate), respectively. Various concentrations of standards were spiked into control samples to prepare calibration curves for quantification.

Non-targeted GC-MS – Fecal pellets and cecal samples were weighed and 20 volumes of HPLC grade water were added. Homogenization was performed using a bead beater without beads (mini-beadbeater-8; BioSpec Products; Bartlesville, OK; set on homogenize for 2 min at room temperature). After centrifugation (20,000 x g for 10 min at 4°C), a 200 µL aliquot of the supernatant was transferred to a clean tube and 400 µL of ice-cold methanol was added. The mixture was vortexed vigorously and subsequently centrifuged (20,000 x g for 10 min at 4°C). A 500 µL aliquot of the resulting supernatant was combined with 10 µL of lysine-$^{13}C_6$,$^{15}N_2$ (2 mM) and the mixture dried in a speed vacuum.

Derivatization of all dried supernatants followed a method adapted with modifications from (*43*). Briefly, 80 µL of methoxylamine solution (15 mg/mL in pyridine) was added to methoximate reactive carbonyls and the mixture incubated at 37ºC for 16 h. This step was followed by replacement of exchangeable protons with trimethylsilyl groups using *N*-methyl-*N*-(trimethylsilyl) trifluoroacetamide (MSTFA) with a 1% v/v catalytic admixture of trimethylchlorosilane (Thermo-Fisher Scientific, Rockford, IL) (incubated at 70ºC for 1 h). Heptane (160 µL) was added and an aliquot of derivatized samples (1 µL) were injected without split into the Agilent 7890A gas chromatography system and coupled 5975C mass spectrometer detector described above. Analyses were carried on a HP-5MS capillary column (see above) with the following modifications of the protocol described for targeted GC-MS: helium was used at a constant flow rate of 1 mL/min; the solvent delay time was 5.5 min; the column head-

pressure was 8.23 p.s.i; the temperatures of the injector, transfer line, and source were 250°C, 290°C and 230°C, respectively; and the GC oven was programmed for a 60 °C hold for 2 min, followed by increases to (i) 140°C at 10°C/min, (ii) 240°C at 4°C/min, and (iii) 300°C at 10°C/min and finally 300°C for 8 min. Metabolites were identified by co-characterization of standards.

Analysis of GC-MS datasets – Data in instrument-specific format (.D) were converted to common data format (.cdf) files using MSD ChemStation (E02.01, Agilent, CA). CDF files were extracted using the Bioinformatics Toolbox in the MATLAB 7.1 (The MathWorks, Inc., Natick, MA), along with previously utilized custom scripts for alignment of the data in the time domain, and automatic integration and extraction of peak intensities (*45,46*). The resulting dataset included sample information, peak retention time and peak intensities. Data were subsequently mean centered and unit variance was scaled for multivariate analysis.

Quality control of GC-MS data – As noted above, samples were analyzed in a randomized order and metabolite identification was done by co-characterization of standards. Pooled quality control (QC) samples were prepared from 20 µL of each sample and analyzed together with the other samples. The QC samples were also inserted and analyzed every 10 samples. To exclude false positives, raw data for statistically significant metabolites were re-evaluated in MSD ChemStation (E02.01, Agilent, CA).

^{1}H-NMR spectroscopic analyses of urine and plasma – Urinary and plasma samples were analyzed using a Bruker 600 MHz spectrometer (Bruker; Rheinstetten, Germany) and a standard set of protocols (*26*). A total of 25 µL of each sample was well mixed with 30 µL of 0.2 M sodium phosphate buffer [20% deuterium oxide, 0.01% 3-(trimethylsilyl)-[2,2,3,3-$^{2}H_4$]-propionic acid sodium salt (TSP) and 3 mM sodium azide]. 50 µL of the resulting mixture was transferred into a micro-NMR tube with an outer diameter of 1.7 mm. A standard one-dimensional solvent suppression NMR pulse sequence (recycle delay [RD]-90^{o}-t_1-90^{o}-t_m-90^{o}-acquisition) was used to acquire 1-dimensional (1-D) ^{1}H NMR spectral data for both urine and plasma. Additionally, a Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence [RD-90°-(t-180°-t)$_n$-acquisition] was applied to plasma samples to allow visualization of small molecular components without substantial interference of macromolecular signals. A total of 512 scans were collected into 64k data points (1k equals 1024 data points) with a spectral width of 20 ppm.

Multivariate data analysis of ^{1}H-NMR spectral data – ^{1}H-NMR spectra obtained from urine and plasma were automatically phased, referenced and baseline-corrected. The resulting NMR spectra (δ0-10) were imported to MATLAB software and digitized into 20,000 data points with resolution of 0.0005 ppm. The water peak region in urinary spectra (δ 4.7-5.05) and plasma CPMG spectra (δ 4.43-5.15) were removed in order to minimize the effect of the disordered baseline. The urea peak (δ 5.58-6.17) in urinary spectra was also removed, followed by peak alignment (*47*) and normalization to total remaining NMR spectral areas in order to perform further analyses, including principal component analysis (PCA) with a log 2 transformed dataset, orthogonal signal correction-projection to latent structures-discriminant analysis (O-PLS-DA) using unit variance scaling method, and correlation analysis (SIMCA v13.0 and MATLAB software).

SUPPLEMENTARY ONLINE TEXT

Additional information about subject characteristics

Three hundred seventeen twin pairs living in five villages (Makhwira, Mitondo, M'biza, Chamba, Mayaka) in the southern region of Malawi were enrolled in this study. The average age at enrollment was 9±6 months (mean ± SD). We did not have information about the order in which each child in a twin pair was born. The percentage of same-gender twin pairs in the cohort was significantly higher than opposite gender pairs [$p<10^{-8}$, Binomial test, probability of success = 0.66, Confidence Interval = (0.61, 0.71), **table S1A**]. All but 19 twin pairs were breastfeeding at the time of enrollment (age at cessation of breastfeeding, 18±5 months).

The average age at presentation with marasmus was 11±4 months; the corresponding ages for kwashiorkor and MAM were 16±7 and 14±7 months, respectively. Children with marasmus suffered significantly more episodes of diarrhea than did those with kwashiorkor or MAM ($p<0.05$; **table S1C**). Among 244 mothers tested for HIV, 14 were positive; of the 147 children tested, only 1 was positive.

Co-twins with kwashiorkor did not receive antibiotics more frequently than their healthy co-twins. Antibiotic usage is quite common in this rural population based on Integrated Management of Childhood Illnesses (IMCI) guidelines for diarrhea, fever, and respiratory distress (*48*). During the course of this study, we did not provide antibiotics to children with kwashiorkor and did not refer them for medical care or antibiotics unless they were markedly ill at the time of clinical examination: i.e., we used a higher threshold than what IMCI would dictate since we had a very senior and experienced nurse and pediatrician who evaluated all of the children in the study at each visit.

Twin pairs discordant for marasmus had a higher death rate, compared to those discordant for kwashiorkor or MAM ($p=0.0027$, Chi-square test): a co-twin died during the study in 43% of the twin pairs who were discordant for marasmus compared to 9% in twin pairs discordant for kwashiorkor and 14% in those discordant for MAM (**table S1A**). The majority of deaths were attributed by verbal autopsy to diarrhea, malaria or pneumonia. In discordant twin pairs, the proportion of the deaths was not significantly higher in co-twins with malnutrition than in their healthy siblings (table S1A). In 120 of the 317 (37.9%) sampled families, at least one child (not a twin) had died at some point before our study; these deaths were not significantly more frequent in families of twins with SAM.

In the population included in this study, children with kwashiorkor have the highest rates of nutritional recovery and survival among the different types of SAM – whether SAM is subdivided into two groups (kwashiorkor and marasmus) or into three groups (kwashiorkor, marasmic kwashiorkor, marasmus).

Searching for taxonomic signatures of kwashiorkor

As noted in the main text, we selected nine same gender twin pairs who remained well nourished in our study cohort and 13 same gender twin pairs who became discordant for kwashiorkor for metagenomic analyses of their microbiomes.

Amplicons were generated from variable region 4 (V4) of bacterial 16S rRNA genes present in the pre-treatment fecal samples of co-twins in discordant pairs and subjected to sequencing with an Illumina HiSeq instrument (403,603±97,345

reads/sample; **table S2A**). We constructed every pairwise combination of taxa at the phylum-, class-, order-, family-, genus-, and species (97%ID OTU) levels and found that no pairwise combination of taxa has a significantly different proportional representation across all twin pairs. Additionally, we calculated the generalization error of the Random Forests classifier (*49*) to distinguish between healthy and kwashiorkor samples for all pairwise combinations of taxa. The results revealed that across all taxonomic levels no combination provided more than a 1% improvement in predictive ability.

Analysis of enteropathogens in fecal microbiota

We used established PCR-Luminex assays (*15-18*) to test for 22 common bacterial, parasitic, and viral enteropathogens in the input human microbiota as well as in recipient mouse fecal samples. We have previously reported that the PCR-Luminex corrected median fluorescence correlates semi-quantitatively with gene copy numbers for these assays. Details are described elsewhere (*15-18*); in brief, the relationship displays a linear correlation with an R^2 of > 0.8 for most targets.

Human fecal samples from families 196 and 57 harbored protozoan (*Giardia* and *Cryptosporidium*) as well as bacterial pathogens [*Campylobacter*, atypical enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC), and enteroinvasive *E. coli* (EIEC)/*Shigella*]. This diversity and multiplicity of enteropathogens is not surprising in children from low-income countries (*50*). The presence of several enteropathogens appeared to cluster within the families (e.g., *Giardia*, atypical EPEC, ETEC, EIEC/*Shigella*, and *Campylobacter*); however, numbers and/or levels of pathogens were not significantly different between the healthy and kwashiorkor co-twins (two-way ANOVA with Bonferroni post-test; **fig. S8A**)

Atypical EPEC and ETEC, detectable in the fecal microbiota of both co-twins in family 196, were transferred to recipient gnotobiotic mice but cleared rapidly (**fig. S8B**). *Cryptosporidium* DNA was present at high levels in the kwashiorkor co-twin from family 196 but was only detected at two time points in one of the gnotobiotic recipients that did not lose more weight than the other mice in the same recipient group (**fig. S8C**). We concluded that the markedly discordant weight loss phenotype in recipients of these microbiota was not due to the transfer and subsistence of these pathogens.

Shared microbial responses between mice receiving transplants from discordant pairs 196 and 57

As in the case of discordant twins belonging to family 196, transplantation of fecal microbiota from discordant twin pair 57 was highly efficient. Only 4 of 69 and 2 of 76 species-level taxa detected in the input community were not detected in the output fecal microbiota of healthy and kwashiorkor co-twin microbiota transplant recipients, respectively (**table S7C,D**). Moreover, 87.4-92.3% of the 835 ECs present in the input communities were detected in the recipient mice, and the proportional representation of these ECs in the input and output fecal samples was highly correlated (R^2=0.784-0.859; **fig. S7C,D**).

Thirty-three species-level taxa were identified whose proportional representation was significantly different in the fecal microbiota of the two groups of family 57 transplant recipients when they were consuming a Malawian diet (**table S8D**). Six of these taxa showed corresponding significant differences in recipients of the discordant

twin pair 196 microbiota (noted in bold font in **table S8D**). Taxa that showed higher proportional representation in both sets of mice with the kwashiorkor microbiota included *Bilophila wadsworthia*, *Dorea formicigenerans*, a producer of formate used by sulfur-reducing bacteria (*51*), *Coprococcus comes*, a Firmicute associated with Crohn's disease (*52*), plus an unclassified Bacteroidales taxon. Examining 16S rRNA data generated from the human fecal samples prior to transplantation, we found that as with family 196, *B. wadsworthia*, as well as members of the order Clostridiales (*Coprococcus comes*, *Dorea formicigenerans*), and the order Bacteroidales were overrepresented in the microbiota of the kwashiorkor co-twin in family 57 compared to the well nourished sibling (**table S8B,D**).

Like family 196 recipient mice, switching from the Malawian diet to RUTF produced a rapid change in microbiota configuration, with the family 57 kwashiorkor microbiota recipients exhibiting a more dramatic change than recipients of the healthy sibling's microbiota (**fig. S9C,D**). Moreover, the transplanted kwashiorkor gut community underwent many of the same changes seen with the transplanted family 196 kwashiorkor microbiota, including changes in *Bifidobacterium* (*B. longum*, *B. bifidum* and another unclassified taxon), *R. torques*, and *Bacteroidales* (*B. uniformis*, *Parabacteroides distasonis* and an unidentified *Parabacteroides* taxon) (**table S7D** and **table S8C**). **Table S9** compares common changes observed in mice and in twins in families 196 and 57 to changes noted in the 13 discordant twin pairs as a function of health status and diet; e.g., *Collinsella aerofaciens, Coprococcus comes, Ruminococcus obeum, Dorea formicigenerans* and *Bifidobacterium longum* showed similar patterns in most twin pairs.

GC-MS of mouse fecal samples, collected on the last day of each diet period, revealed a very similar metabolic profile to that seen in mice who had received transplants from discordant pair 196, with essential and non-essential amino acids, fatty acids, and products of nucleotide metabolism manifesting statistically significant differences in their levels when mice were consuming RUTF versus the Malawian diet (**fig. S14,** $p<0.05$; unpaired Student's t-test).

The initial global metabolite profiles from samples obtained during the first phase of the Malawian diet period were significantly different between mice transplanted with the kwashiorkor versus the healthy co-twin microbiota. For discordant pair 196, ^{1}H-NMR analysis of urine disclosed that the kwashiorkor microbiota was associated with higher levels of allantoin (an end product of purine metabolism in non-primates, equivalent to uric acid in humans), creatine and creatinine (markers of muscle turnover) and dimethylamine (linked to gut microbial metabolism of choline via trimethylamine), and lower concentrations of 2-oxoadipate (product of lysine metabolism) and trimethylamine (from choline) (**Table 1**). Mice harboring transplanted microbiota from kwashiorkor co-twin 57 also showed increased urinary excretion of allantoin, creatine and dimethylamine (**table S10**), phenylacetylglycine (microbial-host co-metabolite from colonic phenylalanine metabolism), trimethylamine-*N*-oxide (alternative metabolic product of trimethylamine derived from choline), and acetate; these latter differences were not significant in recipients of the discordant family 196 twin microbiota, although there was a trend towards increased trimethylamine-*N*-oxide excretion in mice harboring the kwashiorkor community.

SUPPLEMENTARY FIGURES

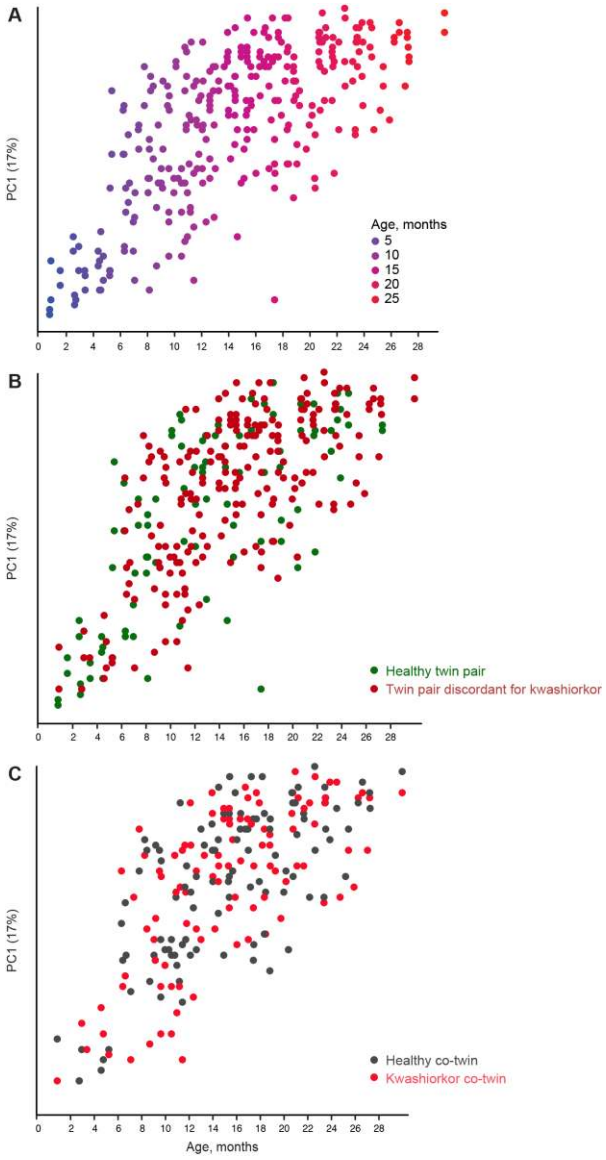


fig. S1. PCoA of Hellinger distances generated from microbiome KEGG EC profiles of Malawian twin pairs who remained concordant for healthy status and twin pairs who became discordant for kwashiorkor. (**A**) Taken from **Fig. 1A** showing PC1 coordinates of all 308 sequenced co-twin fecal microbiomes plotted against age. Each sphere represents a microbiome colored by the age of the human donor. (**B**) Same as panel **A**, but colored by the health status of a twin pair (green, concordant for healthy; red, discordant for kwashiorkor). (**C**) PC1 coordinates of fecal microbiomes from twin pairs discordant for kwashiorkor (n=215 microbiomes) plotted against age, colored by the health status of each co-twin (those who developed kwashiorkor over the course of the study are colored in red before and after onset of the disease; healthy co-twin samples are colored black). Similar results were obtained using other distance metrics (Bray Curtis, Euclidian, and Kulzyncki; data not shown).

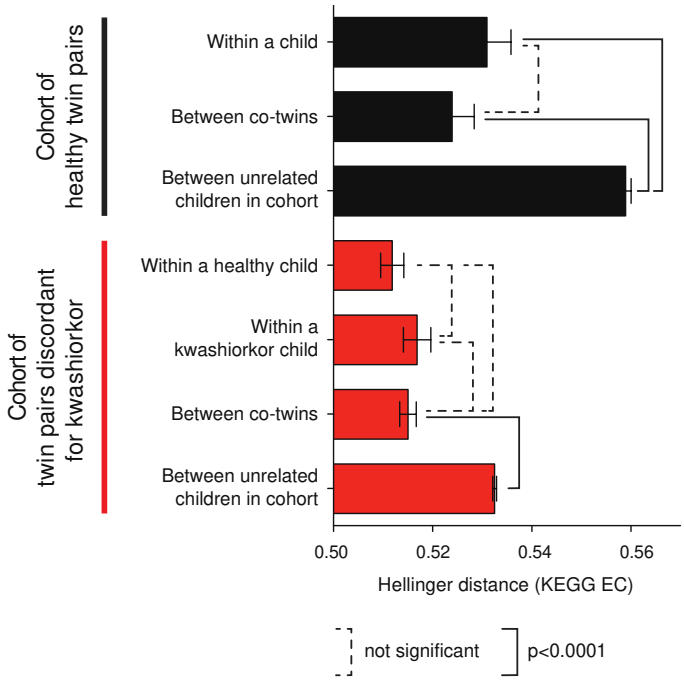


fig. S2. Hellinger distances within and between the fecal microbiomes of Malawian twin pairs who remained healthy, and twin pairs who became discordant for kwashiorkor. Hellinger distances were derived from the KEGG EC profiles of healthy twin pairs (black bars) and twin pairs discordant for kwashiorkor (red bars). Distances between all microbiomes that originate from a given child, and distances between microbiomes sampled from a given twin pair, are plotted next to the average distance between microbiomes from unrelated children. $p<0.0001$, unpaired Student's t-test with Bonferroni correction. Similar results were obtained using other distance metrics (Bray Curtis, Euclidian, and Kulzyncki; data not shown).

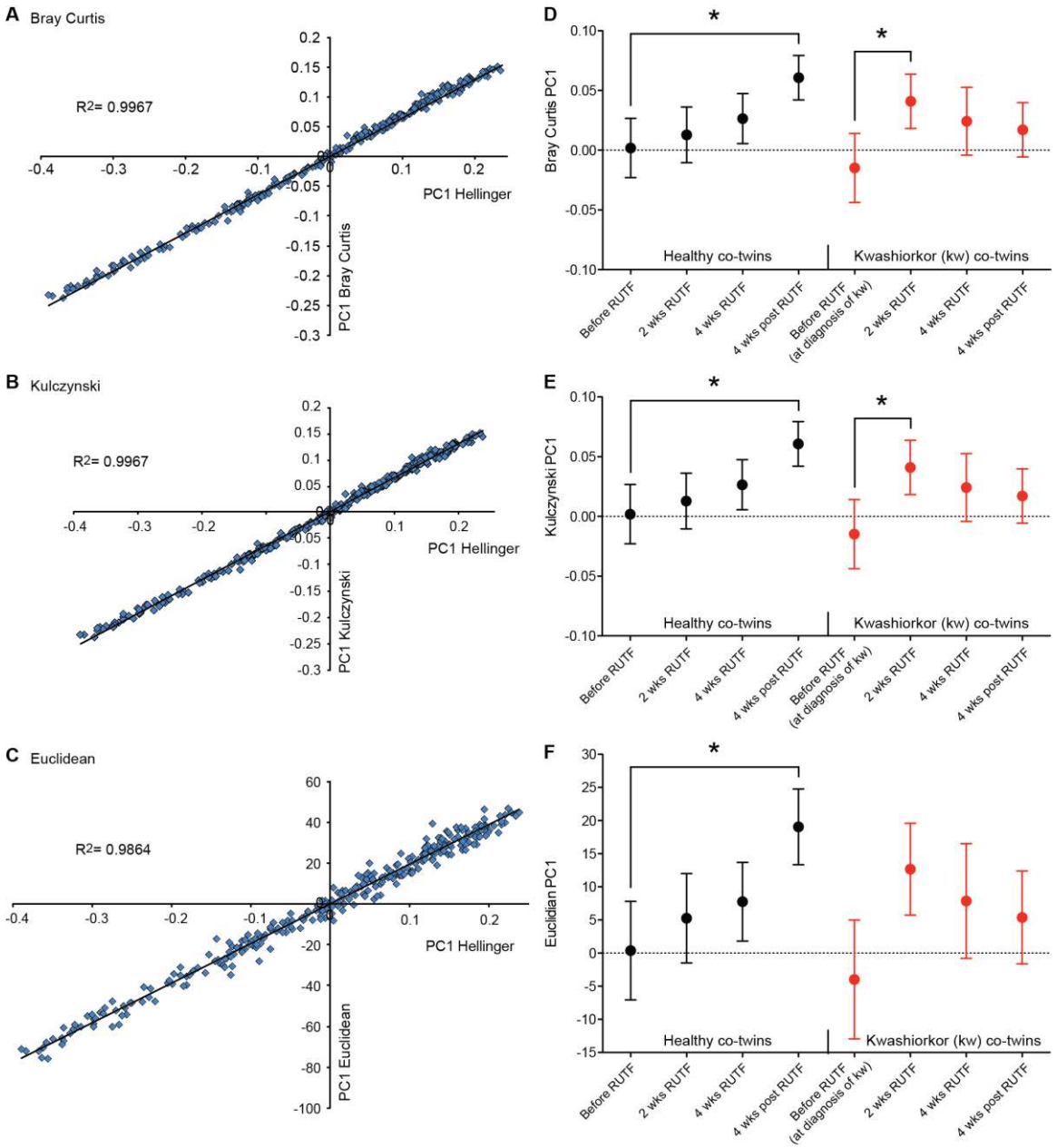


fig S3. Comparison of four distance metrics used to differentiate the fecal microbiomes of Malawian twin pairs who became discordant for kwashiorkor. The analysis is based on KEGG EC profiles. (**A-C**) Position of 308 microbiomes along Principal Coordinate 1 (PC1) defined from PCoA of Hellinger distances on the x-axis plotted against PC1 generated from PCoA of Bray Curtis (panel A), Kulczynski (panel B) and Euclidian (panel C) distances on the y-axis. R^2 values for the linear regression are shown. (**D-F**) Average ± SEM of the PC1 coordinate obtained from panels A-C for gut microbiomes sampled before, during and after RUTF treatment in co-twins discordant for kwashiorkor; panel D, Bray Curtis; panel E, Kulczynski; panel F**,** Euclidian distances. *p<0.05, Friedman test with Dunn's post-hoc test applied to the data shown in panels D-F.

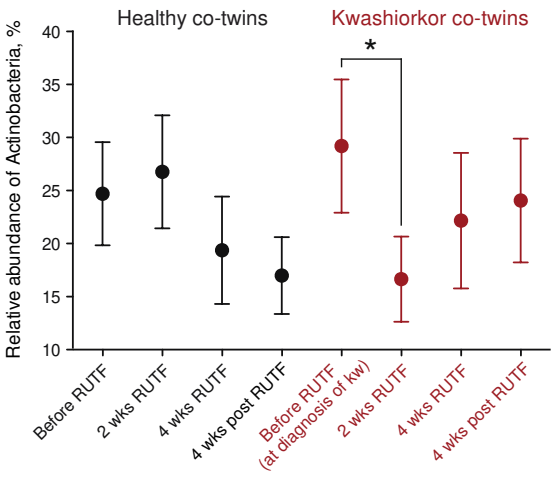


fig. S4. Relative abundance of Actinobacteria in the fecal microbiomes of healthy twin pairs and twin pairs discordant for kwashiorkor. Percent relative abundance of Actinobacteria in the fecal microbiome decreases significantly with 2 weeks of RUTF treatment in co-twins with kwashiorkor but not in their healthy co-twins; *, $p<0.05$ Friedman test with Dunn's post hoc test and after false discovery rate correction for comparison of other bacterial phyla. Data were generated by blasting shotgun reads against a database of 462 sequenced human gut microbial genomes.

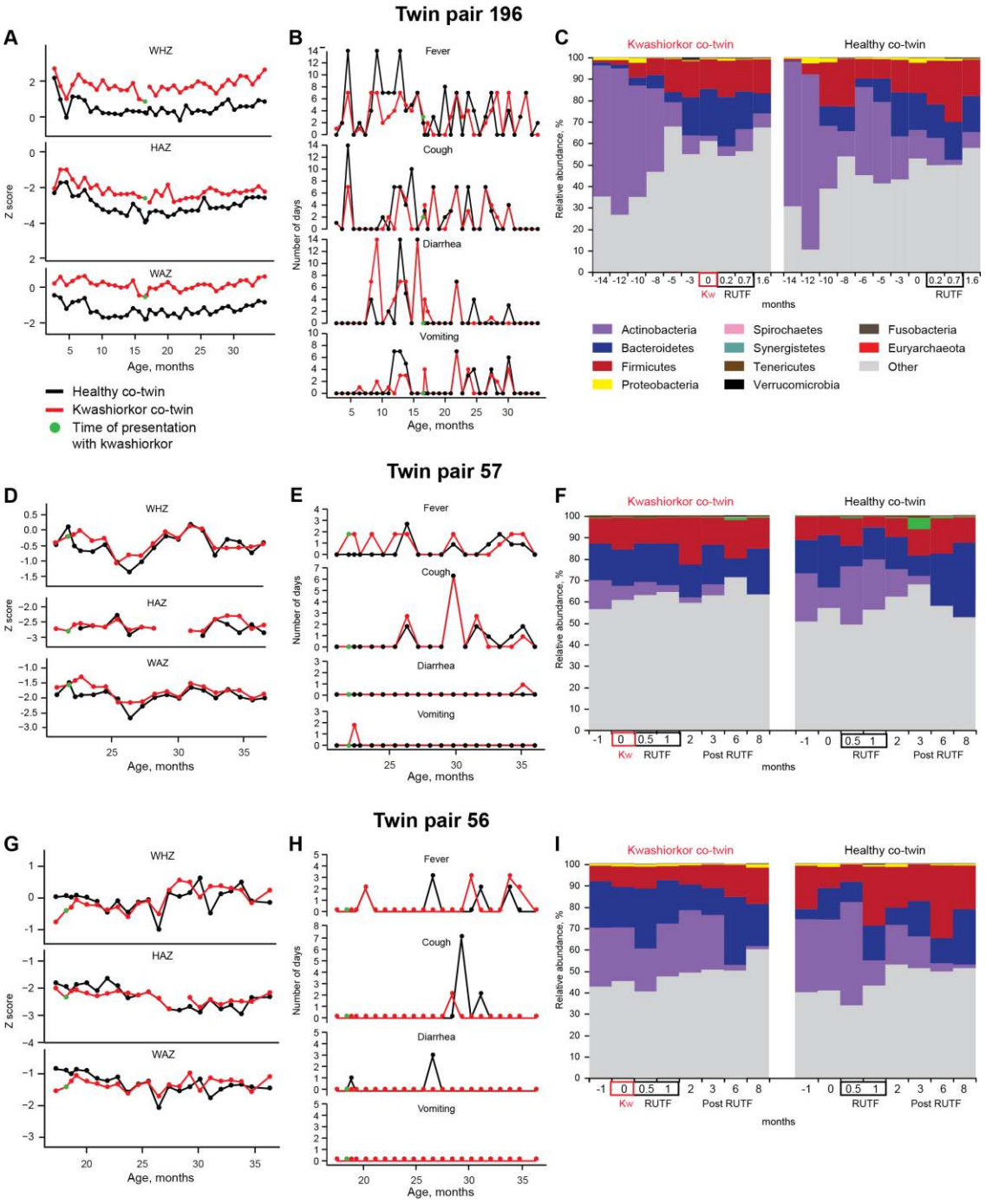


fig. S5. Anthropometric measurements, health status, and functional composition of the fecal microbiomes of twin pairs from families 196, 57, and 56. Shown are anthropometric Z scores, episodes of diarrhea, fever, cough and vomiting, and relative abundance of major bacteria taxa based on the BLAST analysis to 462 bacterial genomes for twin pair 196 (**A-C**), twin pair 57 (**D-F**), and twin pair 56 (**G-I**). The DZ male co-twins from family 196 were enrolled from the village of Chamba at 2.7-months of age when both were healthy. Their last visit took place when they were 34.7-months-old. There were no symptoms of cough, fever, diarrhea or vomiting at the time when the co-twin with kwashiorkor first presented with his disease. The MZ female co-twins from

family 57 were enrolled from the village Mbiza at 20.6-months of age when both were healthy. Their last visit took place when they were 36.2 months old. At the time of presentation of kwashiorkor, the affected female presented with two days of fever preceding the clinic visit. The MZ female co-twins from family 56 were enrolled from the village Mbiza at 17.2 months of age when both were healthy. Their last visit took place when they were 36.3 months old. The number of episodes of fever, cough, diarrhea and vomiting were similar between the co-twins and was not higher in the child who developed kwashiorkor, either before or at the time of presentation with the disease.

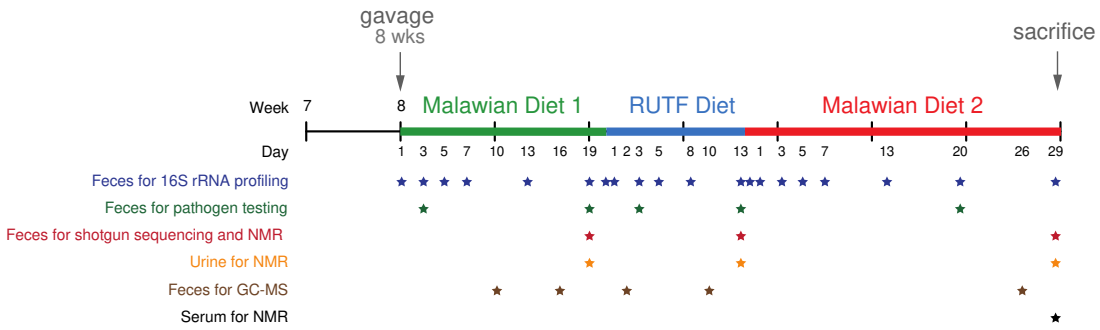
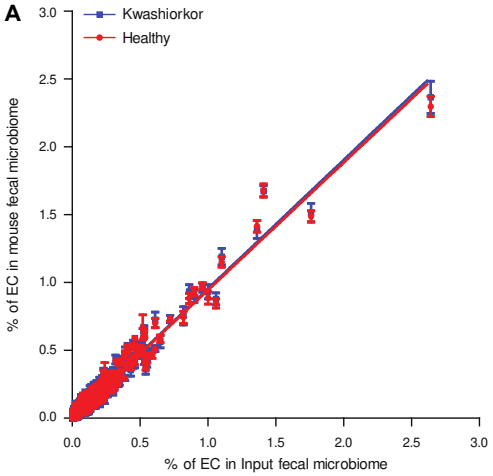


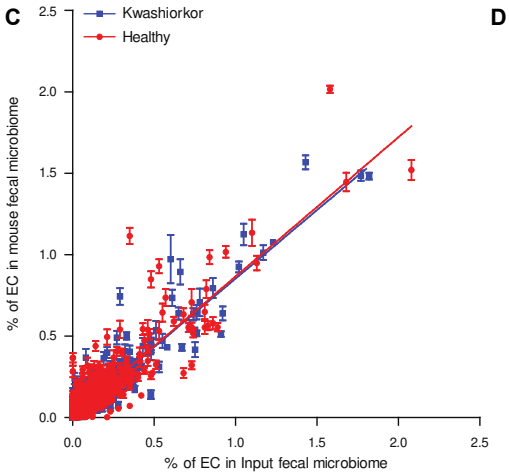
fig. S6. Sampling scheme for gnotobiotic mouse recipients of fecal microbiota transplants from discordant twin pair 196. A star designates the time point at which a sample was used for the indicated analysis. Amplicons were generated from the V4 region of bacterial 16S rRNA genes present in the input human samples, as well as in the fecal microbiota of all mouse recipients, and sequenced (204,585±181,417 reads/sample; 10 recipients/microbiota; up to 8 time points surveyed/diet/mouse; n=379 samples; **table S2B**). DNA isolated from fecal samples collected at the end of each diet period was also subjected to shotgun pyrosequencing (n=54 samples; **table S2C**).



B

Correlation of % EC abundance in input donor sample microbiome versus microbiome of recipient mice consuming a Malawian diet 19 days after gavage

Input Sample	Sample ID	R square	Sy.x
Family 196 - kwashiorkor co-twin	k196A.M1.D19.k3.F	0.9195	0.042
	k196A.M1.D19.k4.F	0.92	0.04088
	k196A.M1.D19.k6.F	0.9351	0.03793
	k196A.M1.D19.k7.F	0.923	0.04123
	k196A.M1.D19.k8.F	0.9358	0.038
Family 196 - healthy co-twin	h196B.M1.D19.h10.F	0.9195	0.042
	h196B.M1.D19.h2.F	0.92	0.04088
	h196B.M1.D19.h3.F	0.9351	0.03793
	h196B.M1.D19.h4.F	0.923	0.04123
	h196B.M1.D19.h5.F	0.9358	0.038
	h196B.M1.D19.h6.F	0.933	0.03751
	h196B.M1.D19.h7.F	0.9236	0.03882
	h196B.M1.D19.h8.F	0.8933	0.04835
	h196B.M1.D19.h9.F	0.9229	0.04068



D

Correlation of % EC abundance in input donor sample microbiome versus microbiome of recipient mice consuming a Malawian diet 19 days after gavage

Input Sample	Sample ID	R square	Sy.x
Family 57 - kwashiorkor co-twin	k57A.M1.D19.k1.F	0.859	0.05032
	k57A.M1.D19.k2.F	0.849	0.05225
	k57A.M1.D19.k3.F	0.8443	0.05271
	k57A.M1.D19.k4.F	0.7992	0.06039
	k57A.M1.D19.k5.F	0.8472	0.05276
Family 57 - healthy co-twin	h57B.M1.D19.h1.F	0.801	0.06463
	h57B.M1.D19.h2.F	0.7922	0.0631
	h57B.M1.D19.h3.F	0.786	0.06609
	h57B.M1.D19.h4.F	0.7838	0.06556
	h57B.M1.D19.h5.F	0.8062	0.06141

fig. S7. Efficient transmission of microbiome functions from human donors to recipient gnotobiotic mice. (**A**) Correlation between the percent relative abundance of KEGG ECs in the input fecal microbiomes of twins belonging to discordant pair 196 and those in the output fecal microbiomes of mice three weeks after transplantation while animals were consuming a Malawian diet (M1 phase; mean values ± SEM shown). (**B**) R-squared values for input fecal microbiomes versus microbiomes of individual recipient mice in panel **A**. (**C**,**D**) Comparable data as in panels A and B but in this case from co-twins belonging to discordant pair 57.

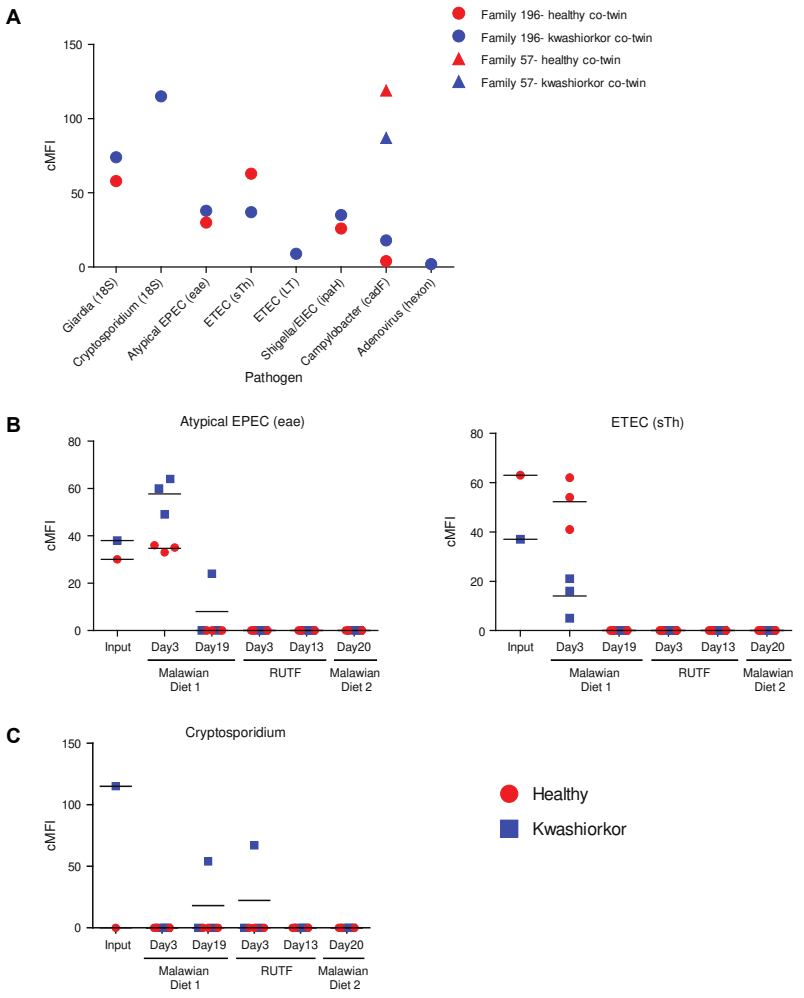


fig. S8. PCR-Luminex assays of enteropathogens in human fecal microbota donor samples from families 196 and 57, and in mouse experiments involving fecal microbiota transplants from discordant twins in family 196. **(A)** Fecal samples collected from family 196 and family 57 donors, at the time that the co-twin was diagnosed with kwashiorkor. **(B)** Assays of atypical enteropathogenic *E. coli* (eae) and enterotoxigenic E. coli (STh) in input human donor fecal samples and fecal samples collected from gnotobiotic mice at the indicated time points on the indicated diets. **(C)** Assay of *Cryptosporidium* in the same samples used in panel B. *Cryptosporidium* DNA appeared in one gnotobiotic mouse recipient of the kwashiorkor co-twin's microbiota 19 days after gavage and became undetectable shortly after treatment with RUTF was initiated. The degree of weight loss in this animal on the Malawian diet was not greater than other members of this treatment group. Corrected median fluorescence intensity (cMFI), a measure of amplicon quantity, is shown on the y-axis. Horizontal lines in panels B and C represent mean values for data on the indicated day for the indicated treatment groups.

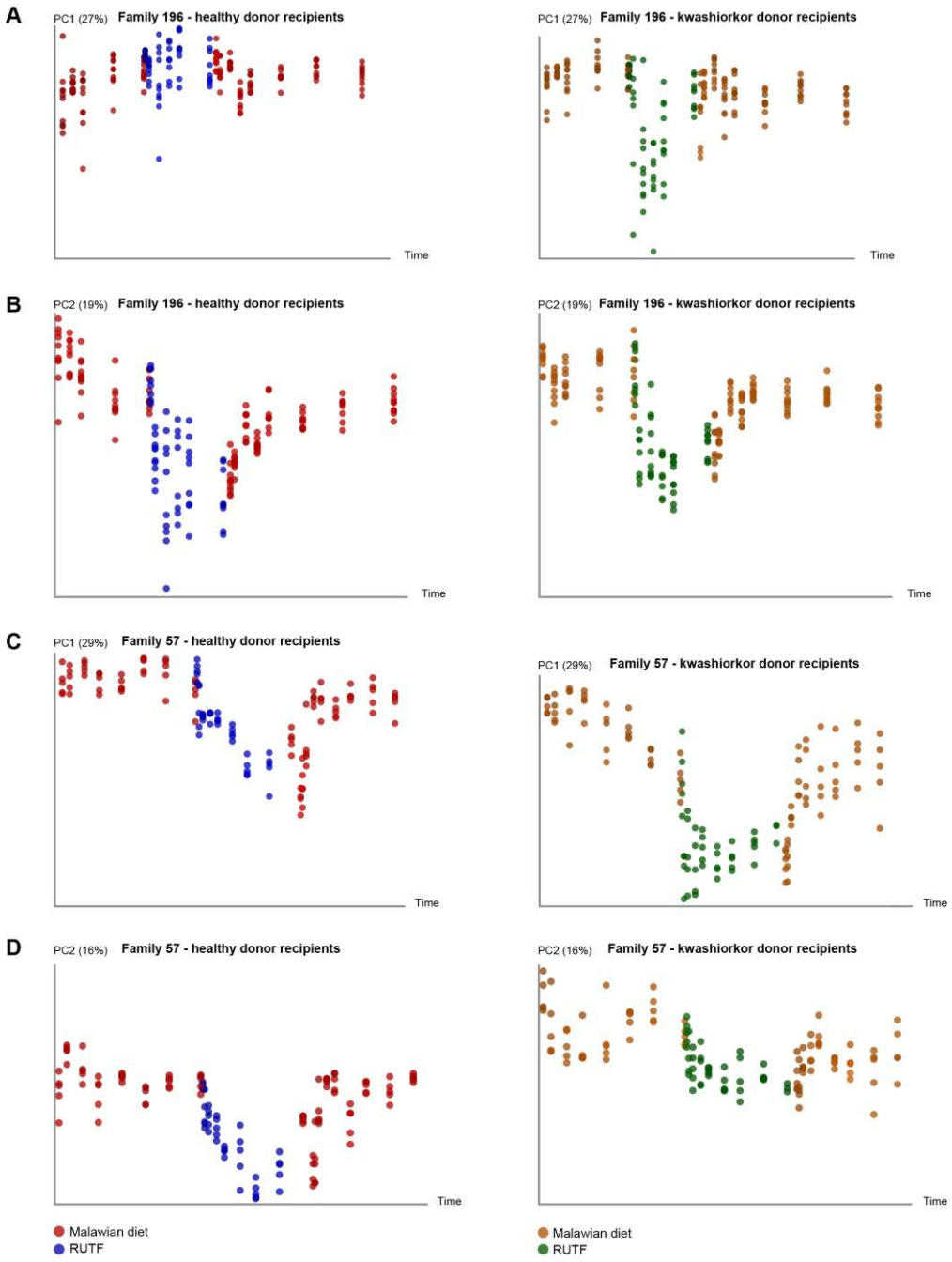


fig. S9. Weighted UniFrac-based PCoA plots of changes in the bacterial phylogenetic configurations of the fecal microbiota of gnotobiotic mice containing transplanted gut microbial communities from discordant pairs 196 (A, B) and 57 (C, D). Each sphere represents a single mouse fecal community, and are colored by donor microbiota and diet: healthy microbiota recipients consuming a Malawian diet (red) or RUTF (blue); kwashiorkor microbiota recipients consuming a Malawian diet (orange) or RUTF (green).

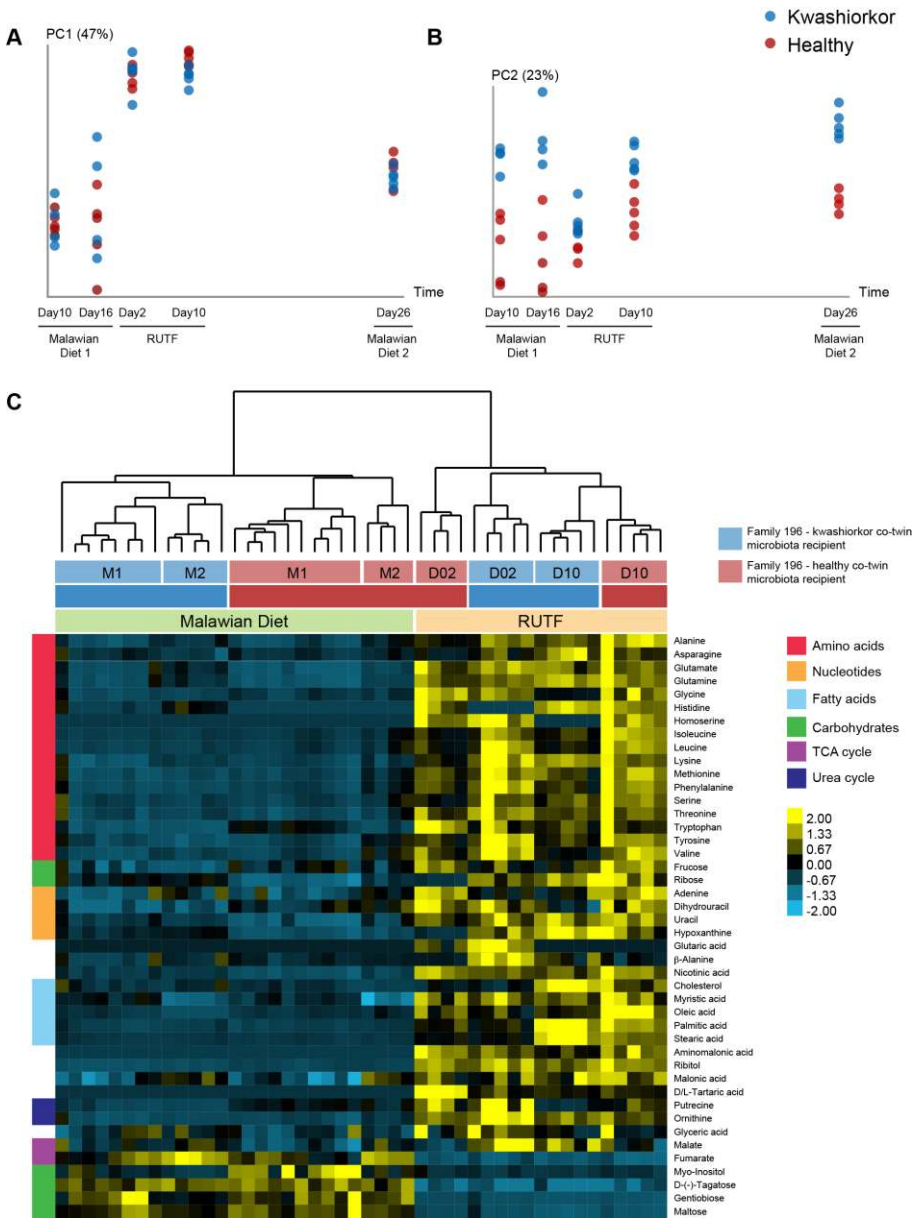


fig. S10. Non-targeted GC-MS analysis of metabolites in fecal samples collected from mice with transplanted microbiota from discordant pair 196. (**A,B**) PCoA plots based on Hellinger distance measurements, showing the effects of diet on the metabolic profile. (**C**) Heatmap showing Z scores of metabolites with statistically significant differences in their levels as a function of diet ($p<0.05$, unpaired t-test). Samples are grouped based on unsupervised hierarchical clustering of metabolite profiles. The results reveal that within a 'donor-diet' cluster, there is distinction between fecal samples obtained during the first versus second exposures to the Malawian diet (M1 and M2) as well as early versus late during the course of RUTF treatment [day 2 (D02) compared to day 10 (D10)]. Color code: blue, fecal microbiota from gnotobiotic recipients of the kwashiorkor co-twin's microbiota; red, recipients of the healthy co-twin's microbiota.

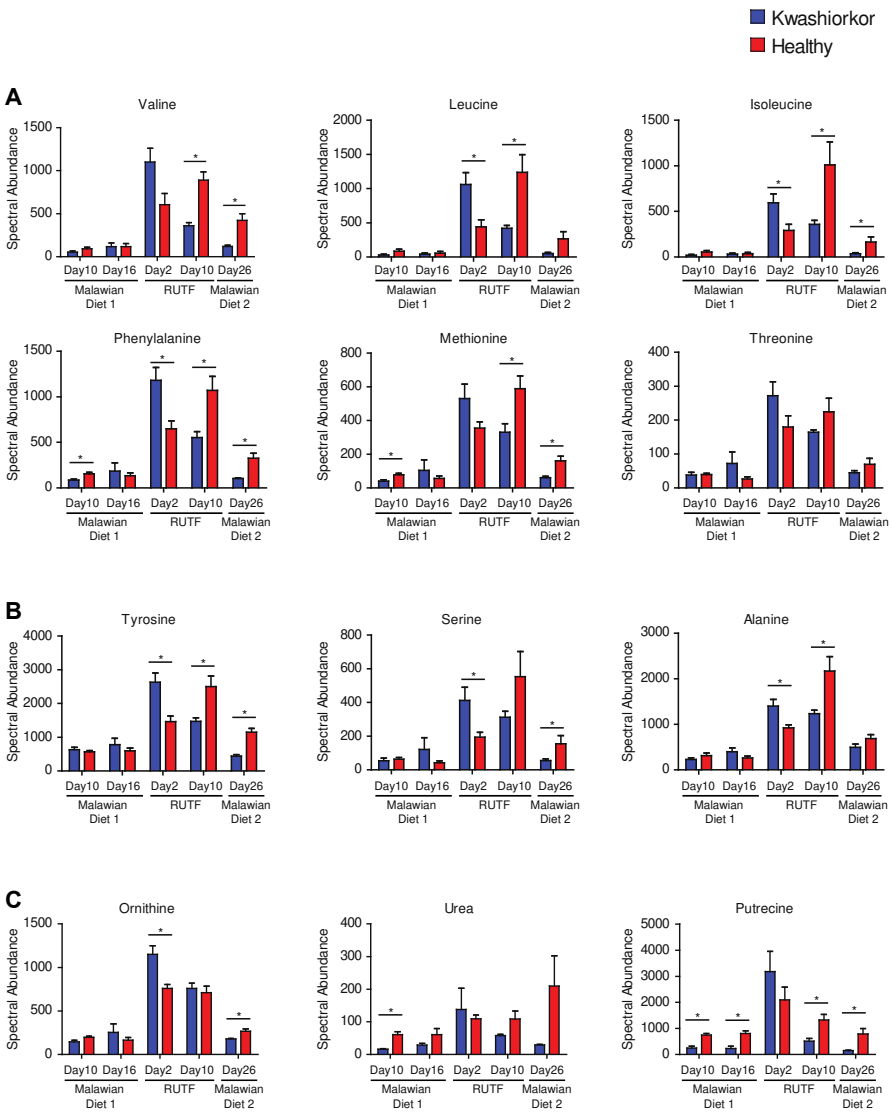


fig. S11. Fecal metabolites differentially affected by diet changes in mice harboring transplanted healthy versus kwashiorkor co-twin fecal microbiota. (**A**) Essential amino acids. (**B**) Non-essential amino acids. (**C**) Urea cycle components. The statistical significance of differences observed in pairwise comparisons between the healthy and kwashiorkor treatment group on a given day of a given diet was evaluated using a two-tailed Student's t-test. Mean values ± SEM are plotted; *, $p<0.05$.

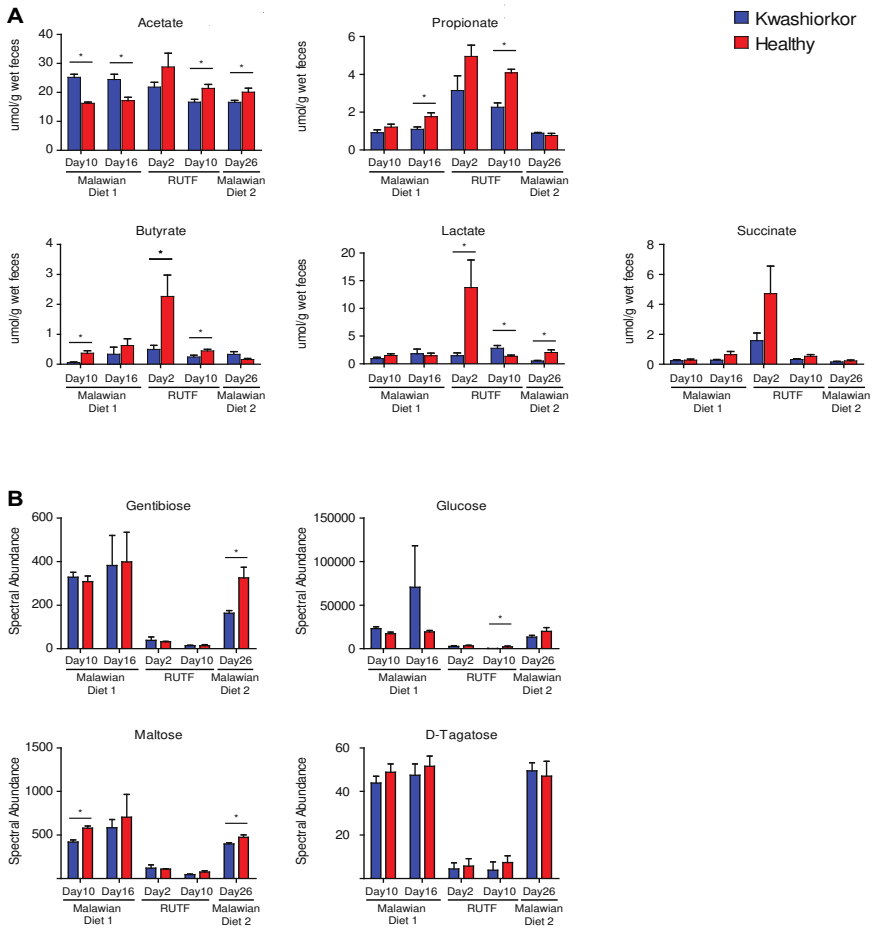


fig. S12. Changes in fecal short chain fatty acid and mono- and disaccharide levels as a function of donor microbiota and diet. (**A**) SCFA. (**B**) Mono- and disaccharides. The statistical significance in pairwise comparisons between the healthy and kwashiorkor treatment group on a given day of a given diet was evaluated using a two-tailed Student's t-test. Mean values ± SEM are plotted; *, $p<0.05$.

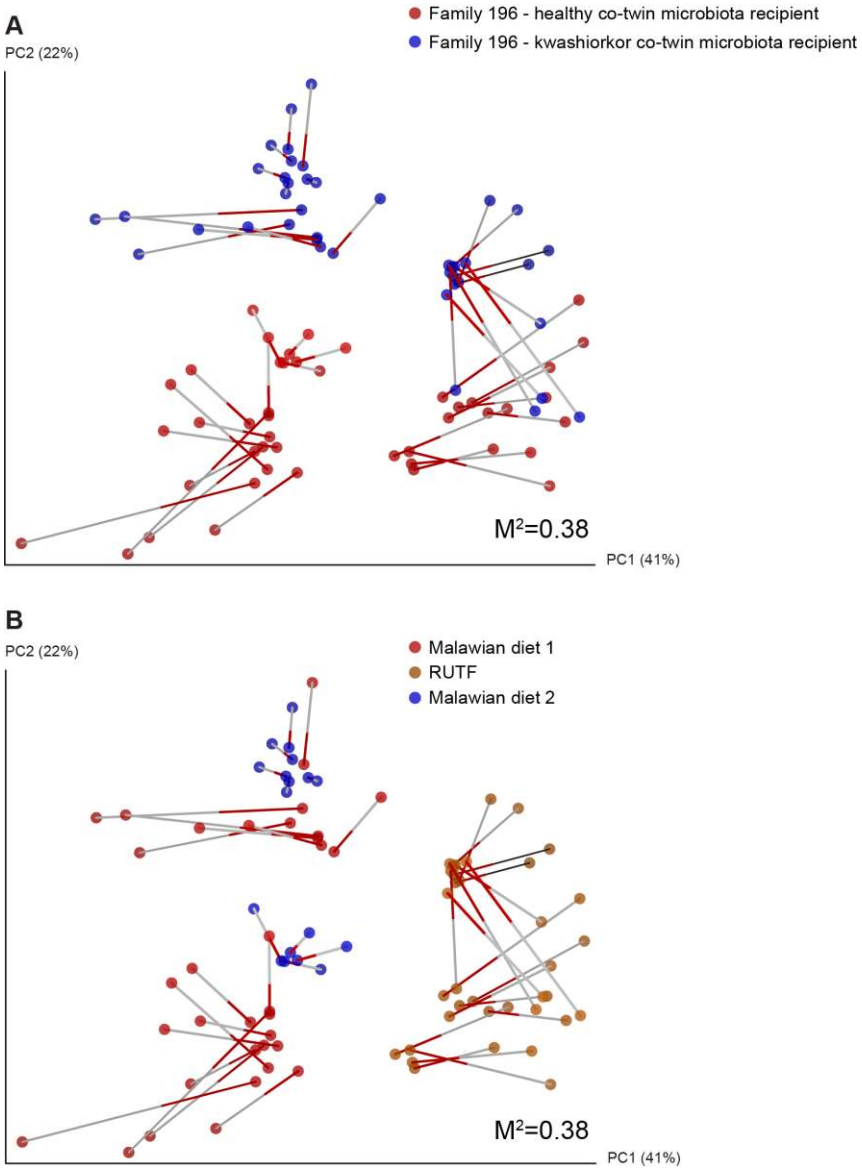


fig. S13. **Procrustes analysis showing correlation between the bacterial phylogenetic configurations and metabolic profiles of fecal samples from mice containing transplanted microbiota from discordant co-twins in family 196.** Each sphere represents a single mouse fecal community. For a given type of data, fecal microbiota were compared using Hellinger distance measurements. The grey end of each line is connected to metabolic data, while the red end of each line is connected to the V4-16S rRNA data. The M^2 value of 0.380 is the fit of the Procrustes transformation over the first three dimensions and is statistically significant ($p<0.0001$, based on 1000 Monte Carlo simulations). (**A**) Spheres are colored based on donor: blue, recipients of the kwashiorkor co-twin's microbiota; red, recipients of the healthy co-twin's microbiota. (**B**) Same as panel A but spheres are colored based on diet: red, first Malawian diet period; orange, RUTF; blue, second Malawian diet period.

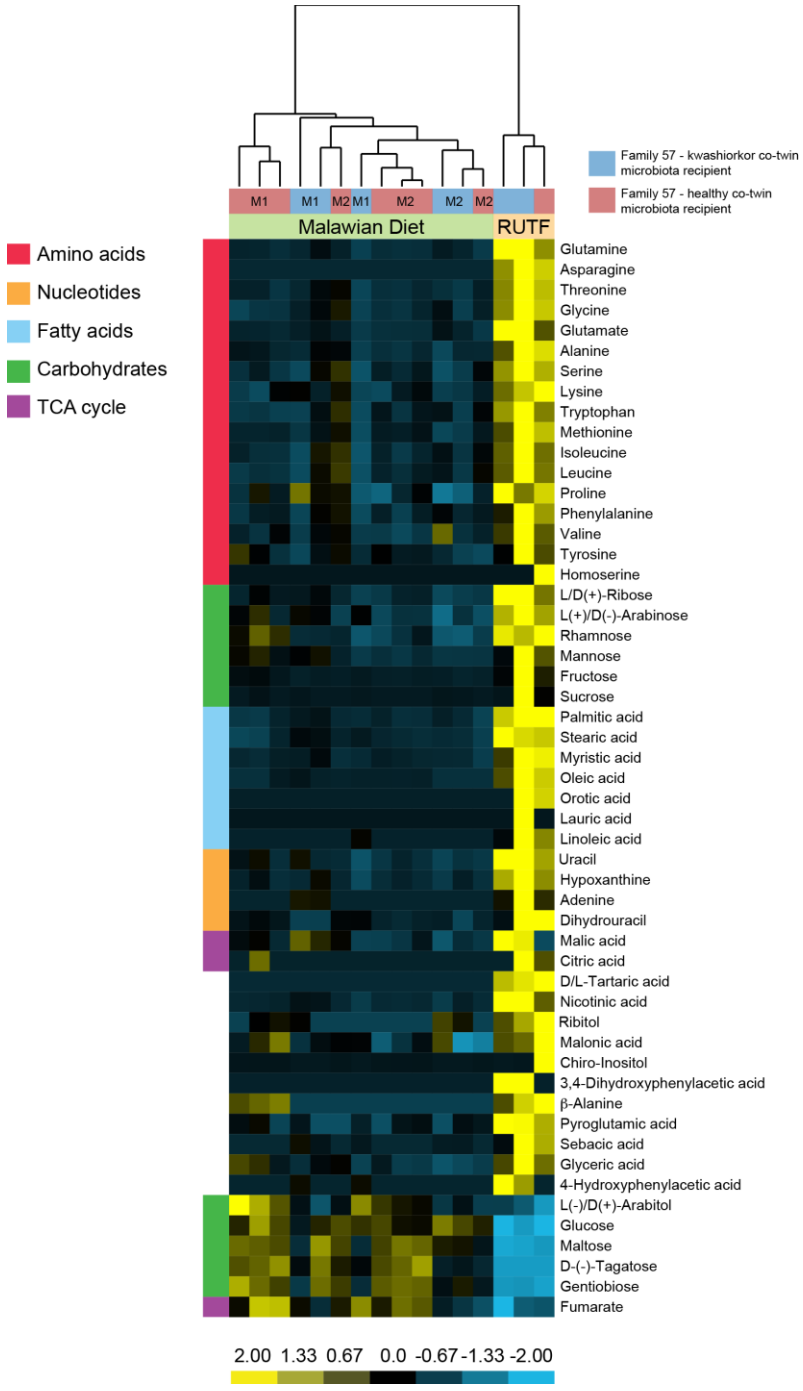


fig. S14. Non-targeted GC-MS analysis of metabolites in fecal samples collected from mice with transplanted microbiota from discordant pair 57. Heatmap of Z scores showing metabolites with significant differences in their levels as a function of diet ($p<0.05$, unpaired t-test). Samples are grouped based on unsupervised hierarchical clustering. The results reveal that within a 'donor-diet' cluster, there is distinction between fecal samples obtained during the first and second exposures to the Malawian diet (M1 and M2).

Smith_Yatsunenko et al. Fig. S15

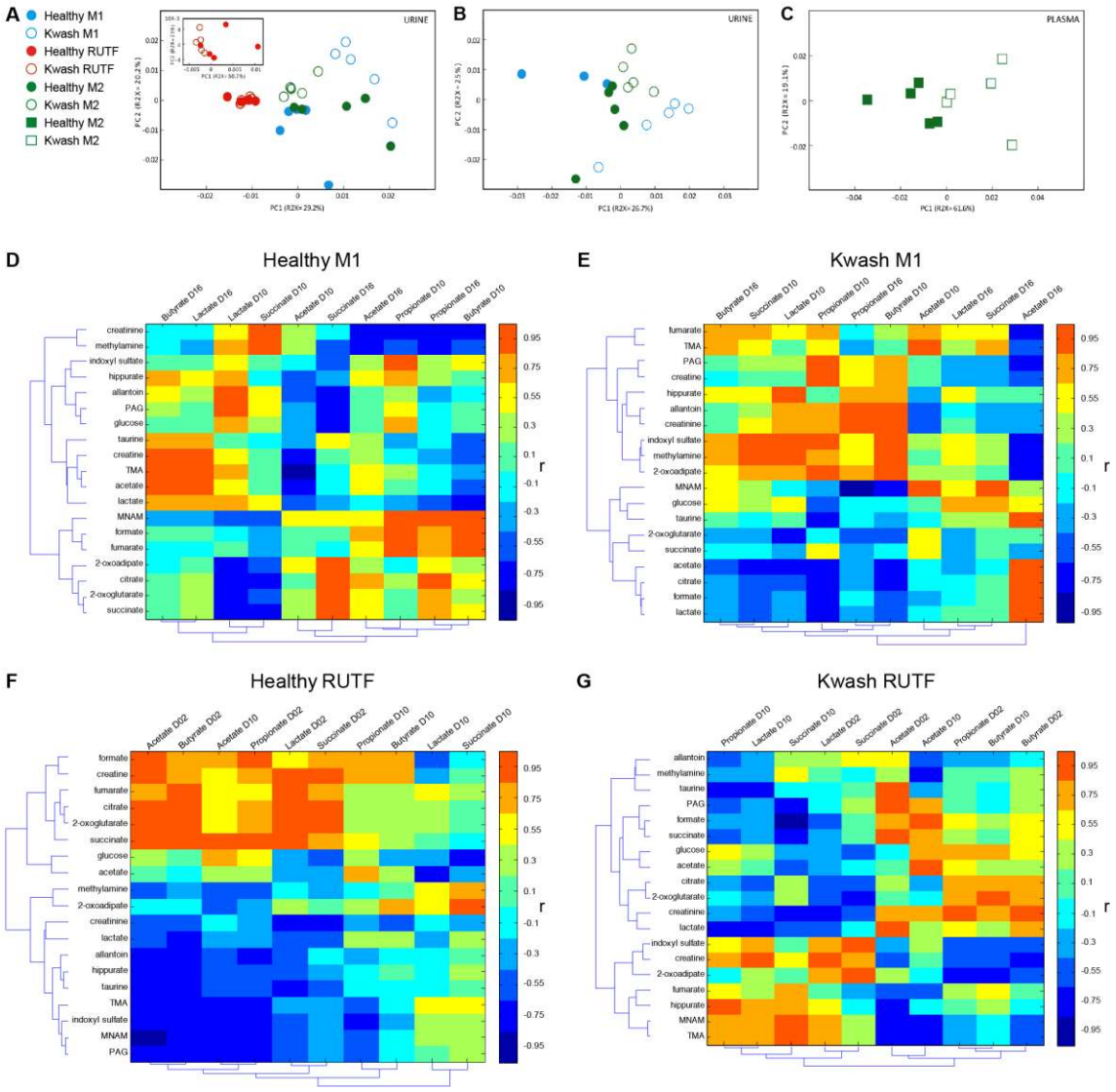


fig. S15. Principal component analysis (PCA) scores plots derived from ^{1}H-NMR analysis of urine obtained from mice with microbiota transplants from discordant twin pair 196 and correlation plots showing correlation structures between urinary metabolites and fecal SCFAs. (**A**) PCA plot of all urine samples from both groups of mice over the three dietary periods emphasizing the impact of RUTF. Inset represents samples collected during the RUTF phase only, emphasizing differences between mice harboring healthy and kwashiorkor co-twin microbiota. (**B**) PCA plot excluding samples collected during the RUTF phase, emphasizing differences between mice harboring healthy and kwashiorkor co-twin microbiota and between the two Malawian diet periods. (**C**) PCA plot of plasma samples collected at the end of the second Malawi diet phase (M2), showing the difference in plasma metabolite profiles between mice with healthy (filled squares) and kwashiorkor (open squares) co-twin

microbiota (higher glucose concentrations and lower levels of citrate, acetate, lactate and tyrosine occur in mice with the healthy co-twin's microbiota; data not shown). All datasets were aligned, normalized to total area, centered and log 2 transformed with an offset of 1.(**D-G**) Correlation plots showing 19 integrated urinary metabolites (y-axis), selected on the basis of group differences as determined by the PCA and PLS-DA models, and clustered according to their correlation with the fecal SCFAs (x-axis). The color represents the correlation between a given urinary metabolite and a fecal SCFA with red colors indicating a positive, and blue a negative correlation value. Panel D, mice with the healthy co-twin microbiota sampled at day 19 on the M1 diet (n=4 mice) correlated with fecal SCFAs at days 10 and 16 after introduction of the M1 diet. Panel E, M1 kwashiorkor (n=4 mice) correlated with five fecal SCFAs at days 10 and 16 on the M1 diet. Panel F, mice with healthy co-twin microbiota sampled on day 13 of RUTF (n=5 mice) correlated with 5 fecal SCFAs at days 2 and 10 after introduction of RTUF. Panel G, mice with kwashiorkor co-twin microbiota on RUTF (n=4 mice) at day 13 correlated with 5 fecal SCFAs sampled on day 2 and day 10 of this diet. Metabolite data have been clustered using a hierarchical clustering algorithm (Euclidian distance; complete linkage) on both axes. Note that for the healthy co-twin microbiota recipients, the TCA cycle intermediates succinate, citrate, 2-oxoglutarate and fumarate are ordered close together in the dendrogram, while in the kwashiorkor twin, fumarate no longer co-clusters with the other TCA intermediates indicating a potential selective inhibition of the TCA cycle. Abbreviations; PAG, phenylacetylglycine; TMA, trimethylamine; MNAM, 1-methylnicotinamide.

SUPPLEMENTARY TABLES

table S1. Characteristics of twins with acute malnutrition.

table S2. Fecal metagenomic datasets obtained from healthy twin pairs, from twin pairs discordant for kwashiorkor and from mice that were the recipients of human fecal microbiota transplants.

table S3. List of 462 sequenced human gut-derived microbial genomes used for analysis of shotgun sequencing datasets.

table S4. Results of the linear mixed-effects regression of PC1, derived from KEGG EC profiles of 308 human microbiomes, with age.

table S5. ECs with significant differences in their representation in the fecal microbiomes of co-twins in each twin pair discordant for kwashiorkor, before diagnosis of kwashiorkor, at time of diagnosis of kwashiorkor, 2 weeks after initiation of RUTF, and 1 month after cessation of RUTF.

table S6. Nutrient analysis of the Malawian and RUTF diets used in this study along with nutritional requirements for mice and 1-3 year-old humans.

table S7. Proportional representation of species-level taxa in human donor microbiota and in fecal samples obtained from gnotobiotic mouse recipients over time during each diet period.

table S8. Species-level taxa whose proportional representation was significantly different between gnotobiotic mouse recipients of a healthy versus kwashiorkor co-twin microbiota transplant, as a function of diet.

table S9. Species-level taxa whose proportional representation changed significantly as a function of diet in gnotobiotic mouse recipients of fecal microbiota transplants from both families 196 and 57 compared to changes observed in the 13 discordant twin pairs.

table S10. Metabolite analysis of urine samples obtained from mice with transplanted healthy or kwashiorkor co-twin microbiota from family 57 at each diet phase.

References and Notes

1. United Nations (UN) Inter-Agency Group for Child Mortality Estimation, *Levels & Trends in Child Mortality Report* (2011); www.childinfo.org/files/Child_Mortality_Report_2011.pdf.

2. WHO Multicentre Growth Reference Study Group, WHO Child Growth Standards based on length/height, weight and age. *Acta Paediatr. Suppl.* **450**, 76 (2006). Medline

3. WHO, UN Children's Fund, *WHO Child Growth Standards and the Identification of Severe Acute Malnutrition in Infants and Children* (2009); www.who.int/nutrition/publications/severemalnutrition/9789241598163/en/index.html.

4. C. D. Williams, B. M. Oxon, H. Lond, Kwashiorkor: A nutritional disease of children associated with a maize diet. *Lancet* **226**, 1151 (1935). doi:10.1016/S0140-6736(00)94666-X

5. T. Ahmed, S. Rahman, A. Cravioto, Oedematous malnutrition. *Indian J. Med. Res.* **130**, 651 (2009). Medline

6. C. Gopalan, in *Calorie Deficiencies and Protein Deficiencies: Kwashiorkor and Marasmus: Evolution and Distinguishing Features*, R. A. McCance, E. M. Widdowson, Eds. (Churchill, London, 1968), pp. 48–58.

7. M. H. Golden, Protein deficiency, energy deficiency, and the oedema of malnutrition. *Lancet* **319**, 1261 (1982). doi:10.1016/S0140-6736(82)92839-2 Medline

8. H. Ciliberto *et al*., Antioxidant supplementation for the prevention of kwashiorkor in Malawian children: Randomised, double blind, placebo controlled trial. *BMJ* **330**, 1109 (2005). doi:10.1136/bmj.38427.404259.8F Medline

9. C. A. Lin *et al*., A prospective assessment of food and nutrient intake in a population of Malawian children at risk for kwashiorkor. *J. Pediatr. Gastroenterol. Nutr.* **44**, 487 (2007). doi:10.1097/MPG.0b013e31802c6e57 Medline

10. T. Yatsunenko *et al*., Human gut microbiome viewed across age and geography. *Nature* **486**, 222 (2012). Medline

11. R. E. Black *et al*., Maternal and child undernutrition: Global and regional exposures and health consequences. *Lancet* **371**, 243 (2008). doi:10.1016/S0140-6736(07)61690-0 Medline

12. WHO, World Food Programme, UN System Standing Committee on Nutrition, UN Children's Fund, *Community-Based Management of Severe Acute Malnutrition* (2007); www.who.int/nutrition/topics/statement_commbased_malnutrition/en/index.html.

13. Materials and methods are available as supplementary materials on *Science* Online.

14. L. Lagrone, S. Cole, A. Schondelmeyer, K. Maleta, M. J. Manary, Locally produced ready-to-use supplementary food is an effective treatment of moderate acute malnutrition in an operational setting. *Ann. Trop. Paediatr.* **30**, 103 (2010). doi:10.1179/146532810X12703901870651 Medline

15. J. Liu *et al*., Simultaneous detection of six diarrhea-causing bacterial pathogens with an in-house PCR-luminex assay. *J. Clin. Microbiol.* **50**, 98 (2012). doi:10.1128/JCM.05416-11 Medline

16. M. Taniuchi *et al*., Multiplex polymerase chain reaction method to detect *Cyclospora*, *Cystoisospora*, and Microsporidia in stool samples. *Diagn. Microbiol. Infect. Dis.* **71**, 386 (2011). doi:10.1016/j.diagmicrobio.2011.08.012 Medline

17. M. Taniuchi *et al*., High throughput multiplex PCR and probe-based detection with Luminex beads for seven intestinal parasites. *Am. J. Trop. Med. Hyg.* **84**, 332 (2011). doi:10.4269/ajtmh.2011.10-0461 Medline

18. M. Taniuchi *et al*., Development of a multiplex polymerase chain reaction assay for diarrheagenic *Escherichia coli* and *Shigella* spp. and its evaluation on colonies, culture broths, and stool. *Diagn. Microbiol. Infect. Dis.* **73**, 121 (2012). doi:10.1016/j.diagmicrobio.2012.03.008 Medline

19. S. Devkota *et al*., Dietary-fat-induced taurocholic acid promotes pathobiont expansion and colitis in *Il10*$^{-/-}$ mice. *Nature* **487**, 104 (2012). Medline

20. N. Crum-Cianflone, *Clostridium innocuum* bacteremia in a patient with acquired immunodeficiency syndrome. *Am. J. Med. Sci.* **337**, 480 (2009). doi:10.1097/MAJ.0b013e31819f1e95 Medline

21. T. Itoh, Y. Fujimoto, Y. Kawai, T. Toba, T. Saito, Inhibition of food-borne pathogenic bacteria by bacteriocins from *Lactobacillus gasseri*. *Lett. Appl. Microbiol.* **21**, 137 (1995). doi:10.1111/j.1472-765X.1995.tb01025.x Medline

22. M. F. Fernández, S. Boris, C. Barbés, Probiotic properties of human lactobacilli strains to be used in the gastrointestinal tract. *J. Appl. Microbiol.* **94**, 449 (2003). doi:10.1046/j.1365-2672.2003.01850.x Medline

23. Y. Kato-Mori *et al*., Fermentation metabolites from *Lactobacillus gasseri* and *Propionibacterium freudenreichii* exert bacteriocidal effects in mice. *J. Med. Food* **13**, 1460 (2010). doi:10.1089/jmf.2010.1137 Medline

24. A. A. Salyers, S. E. West, J. R. Vercellotti, T. D. Wilkins, Fermentation of mucins and plant polysaccharides by anaerobic bacteria from the human colon. *Appl. Environ. Microbiol.* **34**, 529 (1977). Medline

25. H. Sokol *et al*., *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 16731 (2008). doi:10.1073/pnas.0804812105 Medline

26. O. Beckonert *et al*., Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat. Protoc.* **2**, 2692 (2007). doi:10.1038/nprot.2007.376 Medline

27. H. Laue, K. Denger, A. M. Cook, Taurine reduction in anaerobic respiration of *Bilophila wadsworthia* RZATAU. *Appl. Environ. Microbiol.* **63**, 2016 (1997). Medline

28. J. C. Edozien, E. J. Phillips, W. R. F. Collis, The free aminoacids of plasma and urine in kwashiorkor. *Lancet* **275**, 615 (1960). doi:10.1016/S0140-6736(60)90502-X Medline

29. R. G. Whitehead, R. F. Dean, Serum amino acids in kwashiorkor. I. Relationship to clinical condition. *Am. J. Clin. Nutr.* **14**, 313 (1964). Medline

30. S. Awwaad, E. A. Eisa, M. El-Essawy, Methionine metabolism in kwashiorkor in Egyptian children. *J. Trop. Med. Hyg.* **65**, 179 (1962). Medline

31. G. Arroyave, D. Wilson, C. De Funes, M. Béhar, The free amino acids in blood plasma of children with kwashiorkor and marasmus. *Am. J. Clin. Nutr.* **11**, 517 (1962).

32. T. R. Ittyerah, S. M. Pereira, M. E. Dumm, Serum amino acids of children on high and low protein intakes. *Am. J. Clin. Nutr.* **17**, 11 (1965). Medline

33. T. R. Ittyerah, Urinary excretion of sulfate in kwashiorkor. *Clin. Chim. Acta* **25**, 365 (1969). doi:10.1016/0009-8981(69)90194-6 Medline

34. D. H. Baker, Utilization of isomers and analogs of amino acids and other sulfur-containing compounds. *Prog. Food Nutr. Sci.* **10**, 133 (1986). Medline

35. N. Orentreich, J. R. Matias, A. DeFelice, J. A. Zimmerman, Low methionine ingestion by rats extends life span. *J. Nutr.* **123**, 269 (1993). Medline

36. M. Mori, S. Manabe, K. Uenishi, S. Sakamoto, Nutritional improvements of soy protein isolate by different levels of methionine supplementation in pregnant rats. *Tokushima J. Exp. Med.* **40**, 35 (1993). Medline

37. J. K. Nicholson, J. A. Timbrell, P. J. Sadler, Proton NMR spectra of urine as indicators of renal damage. Mercury-induced nephrotoxicity in rats. *Mol. Pharmacol.* **27**, 644 (1985). Medline

38. M. J. Manary, Local production and provision of ready-to-use therapeutic food (RUTF) spread for the treatment of severe childhood malnutrition. *Food Nutr. Bull.* **27** (suppl.), S83 (2006). Medline

39. M. S. McPeek, L. Sun, Statistical tests for detection of misspecified relationships by use of genome-screen data. *Am. J. Hum. Genet.* **66**, 1076 (2000). doi:10.1086/302800 Medline

40. P. J. Turnbaugh *et al*., A core gut microbiome in obese and lean twins. *Nature* **457**, 480 (2009). doi:10.1038/nature07540 Medline

41. J. G. Caporaso *et al*., Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **6**, 1621 (2012). doi:10.1038/ismej.2012.8 Medline

42. J. G. Caporaso *et al*., QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335 (2010). doi:10.1038/nmeth.f.303 Medline

43. N. P. McNulty *et al*., The impact of a consortium of fermented milk strains on the gut microbiome of gnotobiotic mice and monozygotic twins. *Sci. Transl. Med.* **3**, 106ra106 (2011). doi:10.1126/scitranslmed.3002701 Medline

44. J. Pinheiro, D. Bates, S. DebRoy, D. Sarkar, R Development Core Team, nlme: Linear and nonlinear mixed effects models. R package version 3.1-104 (2012).

45. M. Chen *et al*., Metabonomic study of aristolochic acid-induced nephrotoxicity in rats. *J. Proteome Res.* **5**, 995 (2006). doi:10.1021/pr050404w Medline

46. J. Cheng, C. Yuan, T. L. Graham, Potential defense-related prenylated isoflavones in lactofen-induced soybean. *Phytochemistry* **72**, 875 (2011). doi:10.1016/j.phytochem.2011.03.010 Medline

47. K. A. Veselkov *et al*., Recursive segment-wise peak alignment of biological ^{1}H NMR spectra for improved metabolic biomarker recovery. *Anal. Chem.* **81**, 56 (2009). doi:10.1021/ac8011544 Medline

48. WHO, UN Children's Fund, *Integrated Management of Childhood Illnesses: Caring for Newborns and Children in the Community* (2011); whqlibdoc.who.int/publications/2011/9789241548045_Manual_eng.pdf

49. L. Breiman, Random forests. *Machine Learning J.* **45**, 5 (2001). doi:10.1023/A:1010933404324

50. M. J. Albert, A. S. Faruque, S. M. Faruque, R. B. Sack, D. Mahalanabis, Case-control study of enteropathogens associated with childhood diarrhea in Dhaka, Bangladesh. *J. Clin. Microbiol.* **37**, 3458 (1999). Medline

51. D. Taras, R. Simmering, M. D. Collins, P. A. Lawson, M. Blaut, Reclassification of *Eubacterium formicigenerans* Holdeman and Moore 1974 as *Dorea formicigenerans* gen. nov., comb. nov., and description of *Dorea longicatena* sp. nov., isolated from human faeces. *Int. J. Syst. Evol. Microbiol.* **52**, 423 (2002). Medline

52. F. Wensinck, J. P. van de Merwe, J. F. Mayberry, An international study of agglutinins to *Eubacterium*, *Peptostreptococcus* and *Coprococcus* species in Crohn's disease, ulcerative colitis and control subjects. *Digestion* **27**, 63 (1983). doi:10.1159/000198931 Medline

53. Subcommittee on Laboratory Animal Nutrition, Committee on Animal Nutrition, Board on Agriculture, National Research Council, *Nutrient Requirements of Laboratory Animals* (The National Academies Press, Washington, DC, ed. 4, 1995).

54. C. M. Chaparro, K. G. Dewey, Use of lipid-based nutrient supplements (LNS) to improve the nutrient adequacy of general food distribution rations for vulnerable sub-groups in emergency settings. *Matern. Child Nutr.* **6** (suppl. 1), 1 (2010). doi:10.1111/j.1740-8709.2009.00224.x Medline