SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Collection of fecal samples from human microbiota donors

Human fecal specimens were collected, de-identified, and stored in a biorepository in accordance with protocols approved by the Washington University Human Research Protection Office and local IRBs for now completed observational studies conducted in the United States, Bangladesh, Malawi, and Venezuela (Yatsunenko et al., 2012; Ridaura et al., 2013; Subramanian et al., 2014; Mondal et al., 2012; Smith et al., 2013; Turnbaugh et al., 2009). Specimens were maintained at -80°C until use. Except for the fecal samples acquired from 10 raw primal dieters for the present study, all previously collected samples had been analyzed by 16S rRNA sequencing and the results described in previous reports (Yatsunenko et al., 2012; Ridaura et al., 2013; Subramanian et al., 2014).

A subset of samples from healthy adult microbiota donors representing each geographic region and culinary tradition were selected from our biorepository using the following criteria: (i) the donor was in his/her third-to-fifth decade of life; (ii) if female, the donor was not known to be pregnant or was greater than 6 months post-partum; (iii) the body mass index of the donor was greater than 21 kg/m^2; (iv) there was no reported past medical history of systemic or gastrointestinal diseases; (v) the donor was on his/her current diet for at least five years; (vi) the donor was not known to have consumed antibiotics or probiotics for at least three months prior to sampling; and (vii) the bacterial phylogenetic configuration of the sample was representative of samples collected from the geographic/culinary cohort (i.e., near the centroid of their cohort in principal coordinates space based on unweighted UniFrac analysis of fecal bacterial 16S rRNA datasets). "Primal" dieters were defined as group that had consumed a "raw primal" diet (i.e., raw meats and vegetables) for 8 to 11 years. In addition to adult microbiota donors, a fecal microbiota sample was obtained from a 24-month old Bangladeshi child from the Mirpur thana of Dhaka city (23.8042°N 90.3667°E) who had a healthy growth phenotype as defined by serial anthropometry (Subramanian et al., 2014). Microbiota donor features are reported in **Table S1a**.

Production of gnotobiotic mouse diets representative of those consumed by the human microbiota donors

Diets used in the current study are described in **Figure S2** and **Table S1b**. All ingredients were purchased at local grocery stores (Whole Foods Market, Global Foods Market, and Schnuck's Markets, all in St. Louis, MO). Organic products were purchased whenever possible to minimize any effects of residual fertilizers and pesticides. Fruits, vegetables, and meats were de-pitted or de-boned as needed and then chopped manually. Any requisite cooking was performed at this point. All ingredients were ultimately ground and mixed for 10 minutes in an industrial food processor (Robot Coupe Model R23, Jackson, MS), then further homogenized in an industrial mixer (Globe Equipment Company, Port Washington, NY).

The *Bangladeshi diet* was based on dietary analyses performed by the Institute of Nutrition and Food Science at the University of Dhaka (Islam et al., 2010). Rice was cooked in a standard commercial rice cooker. Tilapia fish was simmered on a hot plate (Corning Life Sciences, Tewksbury, MA; heat setting of 2) for 30 minutes with minimal water added to prevent burning prior to the addition of the vegetables (potato, spinach, tomato, cauliflower, okra, and Indian cucumber). Yellow lentils were cooked separately in a glass beaker on a hot plate for 90 minutes, with occasional stirring and minimal water periodically added to prevent burning. Ingredients were then mixed as above. For experiments involving turmeric, the Bangladeshi diet was split into two portions immediately after it was cooked. One portion was set aside as the "unsupplemented" diet. Turmeric was heated in vegetable oil (at a 1:1 volume-to-weight ratio, e.g. 1.0 mL oil to 1.0 g turmeric) in a glass beaker on a hot plate for 10 min with occasional stirring, then added to the Bangladeshi diet and mixed in the industrial mixer for an additional 10 minutes. The amount of turmeric added to the Bangladesh diet (0.8 mg/g diet) was equivalent to 20 times the median daily consumption in South Asia (Ferrucci et al., 2010). The *raw primal diet* was based on an analysis of the most commonly consumed foods and overall nutritional intake gleaned from the food diaries of seven individuals living in the USA. Ingredients were not cooked in keeping with the food preparation habits of these people. An

American diet was based on analyses of food journals, obtained from four individuals living in the USA with no self-imposed dietary restrictions, detailing 31 cumulative days of food intake (eight days each from three subjects and seven days from another). A dietician analyzed these journals with the Nutrition Data System for Research (NDS-R; version 4.03_31) to quantify the macronutrients and micronutrients consumed by each subject. The means of daily consumption of dietary energy (kilocalories), carbohydrates, proteins and fats were calculated. These mean values served as the benchmarks against which *in silico* menus were evaluated. Randomly generated menus were created by sampling 41 of the journal entries. Thousands of randomized menus were produced; the representative diet used for our study was the one whose calculated values was closest to the means. All *in silico* diet design was conducted in software R version 3.1.2 (R Development Core Team, 2013). The *Malawian diet* was based on staple foods consumed by individuals living in rural southern Malawi, and cooked using procedures described in an earlier publication (Smith et al., 2013). The *Amerindian diet* was based on foods consumed by individuals from the Guahibo Amerindian families living in villages near Puerto Ayacucho in the Amazonas State of Venezuela (Yatsunenko et al., 2012). Catfish and eggs were simmered on a hot plate for 20 minutes with minimal water added to prevent burning, prior to addition of the remaining ingredients (corn flour, cassava flour, and powdered whole milk). Minimal modifications to the recipes were made as needed to ensure that all diets satisfied the recommended daily intake of macro- and micronutrients for mice (National Research Council (US) Subcommittee on Laboratory Animal Nutrition, 1995).

Aliquots (250 g or 500 g) of each diet type were double-sealed (vacuum packed) and double-bagged before sterilization by irradiation (dose 20-50 kGy; Steris Isomedix Corporation, Chicago, IL). Sterility was confirmed by (i) spore strips (North American Science Associates Inc, Northwood, OH) that were placed in each box of bagged food prior to irradiation and (ii) culturing aliquots of a representative bag from each irradiated box for 7 days at 35°C under aerobic conditions in BHI broth (BD Diagnostics, Sparks, MD), nutrient broth (BD Diagnostics, Sparks, MD), and Sabouraud dextrose broth (BD Diagnostics, Sparks, MD) prior to plating on

blood agar plates. The nutritional content of diets was confirmed by direct analyses of micronutrients and macronutrients (N-P Analytic Laboratories, Nestlé Purina Petcare, St. Louis, MO; see **Table S1b)**. Irradiated, vacuum-packed diets were stored at 4°C until use.

Preparation of human fecal microbiota for transplantation into gnotobiotic mice

Clarified fecal microbiota suspensions were prepared from previously frozen human fecal samples in an anaerobic chamber (Coy Lab Products, Grass Lake, MI; atmosphere composed of 75% N_2, 20% CO_2, and 5% H_2) using a protocol detailed in an earlier report (Ridaura et al., 2013) that was modified as follows: in place of pulverization with a mortar and pestle, samples (500 mg) were diluted in 15 mL of reduced PBS (PBS supplemented with 0.1% Resazurin (w/v) and 0.05% L-cysteine-HCl), homogenized in a blender for 1 minute, and passed through a 40 μm pore diameter nylon cell strainer (BD Falcon, Franklin Lakes, NJ). Clarified fecal microbiota suspensions were stored in reduced PBS with 15% glycerol at -80°C until use.

Animal husbandry

All gnotobiotic mouse experiments were performed using protocols approved by the Washington University Animal Studies Committee. Conventionally raised C57BL/6 *Ret* +/- mice were re-derived as germ-free by embryo transfer (Faith et al., 2011) and maintained in sterile, flexible, plastic gnotobiotic isolators (Class Biologically Clean Ltd., Madison, WI) under a strict 12-hour cycle. All mice used in our experiments were 8-12 weeks old.

In each of the experiments described this report, feeding of the initial diet was started three days prior to colonization with uncultured fecal microbiota samples or with cultured bacterial consortia. Age-matched male germ-free C57BL/6 mice within a given isolator were gavaged with the same bacterial community, or maintained as germ-free. Mice were co-housed as described below. Autoclaved Aspen hardwood lab bedding (NEPCO, Warrensburg, NY) was replaced with new bedding at the start of each new diet phase to limit carryover of ingredients from the prior diet phase. Total intestinal transit times were measured at the end of each diet

phase by gavaging mice with red carmine dye and recording the time from gavage to first appearance of the dye in their feces (see below for details)

Six-phase travel experiment. Six groups of mice (*n*=6 mice per group) were colonized with uncultured fecal microbiota samples obtained from six healthy adults representing different cultural/culinary traditions and geographic locations. Each group of mice designated to receive the same donor's microbiota was placed in a separate gnotobiotic isolator prior to gavage; within each isolator, mice were triply housed (*n*=3 mice per cage). Mice were fed a sequence of sterilized diets formulated to represent those consumed by the microbiota donors (ingredients and preparation detailed above). The starting and ending 'home' diets were given for 14 and 8 days, respectively, while each of the four intermediate 'travel' diets was administered for 8 days.

Three-phase travel experiment. Six groups of mice (*n*=5 mice per group) were colonized with uncultured fecal microbiota samples obtained from healthy adults living in Bangladesh (*n*=3 donors) and the USA (specifically, $USA_{unrestricted}$; *n*=3 donors). Each group of mice resided in a separate flexible film gnotobiotic isolator where they were co-housed in two cages (*n*=2-3 mice per cage). Mice were fed a sequence of sterilized Bangladeshi and primal diets with the starting and ending diet representing a 'local' diet of the microbiota donor: mice colonized with a $USA_{unrestricted}$ microbiota were fed the primal diet, while mice colonized with a Bangladeshi microbiota were fed the Bangladeshi diet. The starting and ending phase diets were given for 14 and 8 days, respectively, while the second (intermediate) phase 'non-native' diet was administered for 8 days.

Assessing the effects of turmeric in germ-free mice versus those colonized with the clonally arrayed bacterial culture collection. Two groups of mice (*n*=6 mice per group) were either maintained as germ-free or colonized with the clonally arrayed culture collection derived from the fecal microbiota of the Bangladeshi child with a healthy growth phenotype. Each group of mice resided in a separate gnotobiotic isolator (*n*=3 mice per cage). Mice were fed a sequence of sterilized diets formulated to represent Bangladeshi diet, with or without turmeric. Each diet

phase was 10 days in length. The starting and ending diets were unsupplemented (i.e., lacked turmeric), while the intermediate phase diet contained turmeric.

Comparing the effects of BSH_{hi} and BSH_{lo} consortia with and without turmeric in wild-type Ret+/+ and heterozygous Ret+/- C57BL6/J mice. Three groups of mice (*n*=5 mice per group) were colonized with (i) the entire clonally arrayed culture collection, (ii) the 7-member BSH_{hi} consortium, or (iii) the 7-member BSH_{lo} consortium. Each group of mice was housed in a separate gnotobiotic isolator containing two cages (*n*=2-3 mice per cage). Mice were fed a sequence of sterilized diets formulated to represent the Bangladeshi diet, with or without turmeric. Each diet phase was 8 days in length. The starting and ending diets were unsupplemented while the intermediate diet contained turmeric. Two additional groups of *Ret+/-* mice (*n*=5 mice per group) were colonized with either the BSH_{hi} or BSH_{lo} consortium. Recipient mice containing a given consortium were housed in the same isolators (but in separate cages) as wild-type mice harboring the same consortium (*n*=2-3 mice per cage).

Effects of turmeric on host gene expression in mice harboring the BSH_{lo} consortium. Two groups of wild-type mice (*n*=3-4 mice per group; 1 cage per group) were colonized with the "BSH_{lo}" consortium. Mice were fed the Bangladeshi diet, with or without turmeric, over the course of a single diet phase that was 10 days long. Distal ileum and liver samples collected at the time of sacrifice were used for mouse RNA-Seq, performed as described below.

Sample collection from gnotobiotic mice

Before colonization and throughout each experiment, fecal pellets were collected from mice at pre-established time points, including on the days of transit time measurements, with care taken to collect pellets prior to the passage of carmine dye. Samples were snap frozen in liquid nitrogen within 30 minutes of their collection and stored at -80°C until use. At the conclusion of each experiment (>48 hours after the last carmine dye assay), the following biospecimens were procured after weighing and sacrificing each animal: liver, small intestine, cecal contents, colon, and feces. The small intestine was partitioned into 8 equal-sized segments.

Segment 8 was operationally defined as 'terminal ileum'. Samples were immediately snap frozen in liquid nitrogen and stored at -80°C until use.

Multiplex sequencing and analysis of PCR amplicons generated from bacterial 16S rRNA genes

Genomic DNA was extracted from mouse fecal pellets as described previously (Turnbaugh et al., 2009). Primers 515F and 806R with sample-specific barcodes were used to generate PCR amplicons spanning variable region 4 (V4) of bacterial 16S rRNA genes present in the fecal samples. Libraries of these V4-16S rRNA amplicons were subjected to multiplex sequencing on the Illumina MiSeq platform (250 nt paired-end reads; see **Table S4** for details regarding multiplexing and sequencing depth). Paired-end reads were trimmed to 200 base pairs to retain the highest-quality sequences, and aligned using the flash aligner (Magoč and Salzberg, 2011). 16S rRNA reads were grouped into operational taxonomic units based on whether they shared ≥97% nucleotide sequence identity (97%ID OTUs). We employed an open-reference OTU-picking strategy (QIIME; version 1.5.0) (Caporaso et al., 2010), where reads were clustered against the Greengenes reference database (DeSantis et al., 2006) using the UCLUST algorithm (Edgar, 2010). Any reads that did not match entries in the reference dataset were clustered *de novo*. OTUs were picked using the combined V4-16S rRNA sequence datasets from all mouse experiments, so that OTU IDs and associated taxonomic annotations would be uniform across the corresponding analyses. OTUs with proportional abundances of at least 0.1% in at least 1% of all samples were retained for downstream analysis. Taxonomy was assigned to OTUs using the Ribosomal Database Project (RDP) version 2.4 classifier, which enabled classifications at the phylum, class, order, family, genus, and species level (Ridaura et al., 2013). Samples were rarefied to a depth of 4080 reads per sample. Unweighted and weighted UniFrac was used to calculate beta diversity indices; principal coordinates analysis (PCoA) plots were generated using these data.

Community profiling by sequencing (COPRO-Seq)

Genomic DNA was extracted from mouse fecal pellets (Turnbaugh et al., 2009) that had been collected throughout experiments involving mice colonized with either the complete culture collection, or the BSH_{hi} or BSH_{lo} consortia. Multiplexed DNA libraries were prepared from genomic DNA samples as described (Faith et al., 2011; McNulty et al., 2011) and subjected to shotgun sequencing using an Illumina MiSeq platform to generate 75 nt unidirectional reads (see **Table S4** for details regarding multiplexing and sequencing depth). The analytic pipeline described in a previous publication (McNulty et al., 2011), and available at https://github.com/nmcnulty/COPRO-Seq, was used to map reads to the sequenced genomes of members of the culture collection and calculate their relative abundances in the fecal microbiota of gnotobiotic mice.

Profiling the fecal meta-transcriptomes of gnotobiotic mice using microbial RNA-Seq

Bacterial community transcriptional responses to turmeric were profiled in fecal samples collected from eight male age-matched wild-type C57BL/6 mice, colonized with the 7-member BSH_{hi} or BSH_{lo} 7- consortia, at two time points (8 and 16 days after gavage, which corresponded to periods where animals were consuming the Bangladeshi diet lacking and containing turmeric, respectively). Multiplexed cDNA libraries were prepared as described previously (Faith et al., 2011; McNulty et al., 2011; Ridaura et al., 2013) with the following modification: after removal of 5S rRNA and tRNA with a MEGAclear column (Life Technologies, Carlsbad, CA) remaining rRNA was depleted using the Ribo-Zero™ rRNA Removal Kit for Bacteria (Epicentre, Madison, WI). Multiplexed libraries were sequenced using a Illumina NextSeq instrument to generate 75 nt unidirectional reads (see **Table S4** for details regarding multiplexing and sequencing depth). Analysis was performed as previously described (Ridaura et al., 2013) with the custom reference database consisting of the 14 genomes of the 14 strains that comprised the two BSH consortia. Following gene assignments, pathway analysis was performed using the KEGG reference database (release 72.1) (Kanehisa and Goto, 2000; Kanehisa et al., 2014).

Profiling the host transcriptome using RNA-Seq

We used RNA-Seq to characterize gene expression in the livers and ileums of seven age-matched wild-type C57BL/6 mice colonized with the BSH_{10} consortium and fed the Bangladeshi diet with or without turmeric. For each sample, 20-30 mg of tissue was placed in pre-cooled RNAlater®-ICE (Life Technologies, Carlsbad, CA) and stored at -20°C for at least 16 hours. Total RNA was extracted using the RNeasy Mini Kit (Qiagen, Venlo, Netherlands) after tissue disruption and homogenization (TissueLyser, Qiagen). 5S rRNA and tRNA were depleted from the sample (MEGAclear column purification) followed by treatment with RNase-free TURBO™ DNase (Life Technologies). Cytoplasmic and mitochondrial rRNA were then depleted using the Ribo-Zero™ rRNA Removal Kit for Human/Mouse/Rat (Epicentre, Madison, WI). (This approach was used rather than poly(A) enrichment to obtain more even coverage since Ribo-Zero™ does not have a 3′ end bias). RNA was precipitated with ethanol, followed by a second DNase treatment (Baseline-ZERO™ DNase; Epicentre, Madison, WI). First strand cDNA synthesis was accomplished using SuperScript II Reverse Transcriptase (Life Technologies), followed by second strand cDNA synthesis with *E. coli* DNA polymerase, *E. coli* DNA ligase, and RNase H (all from New England Biolabs, Ipswich, Massachusetts). Double-stranded cDNA was purified using QIAquick purification columns (Qiagen) before shearing (BioRuptor XL sonicator; Diagenode, Denville, NJ). ~250 bp fragments were size-selected by extraction from 2% agarose gels. Multiplex libraries were prepared and subjected to sequencing on the Illumina NextSeq platform to generate 75 nt unidirectional reads.

Analysis was limited to reads that contained exact 5′ sequence matches to barcodes utilized for multiplexing and no more than 1 expected error, as calculated by Q scores via usearch (version 8.0.1517) (Edgar, 2010). (See **Table S4** for details regarding multiplexing, sequencing depth, filtering, and mapping of the resulting high-quality sequences). Reads were mapped to the *Mus musculus* C57BL/6J strain genome (UCSC mm10) using Bowtie2 (version 2.2.4) (Langmead and Salzberg, 2012), with subsequent alignment and splice junction mapping (TopHat2, version 2.0.13) (Kim et al., 2013). 81.8%- 89.0% of input sequences could be mapped. These data were imported into R (version 3.1.2) (R Development Core Team, 2013) for analysis

using the *edgeR* package (version 3.8.5) (Robinson et al., 2010). Separate analyses were conducted for liver and distal ileum RNA-Seq datasets. Low-expression genes were filtered out: only those with an abundance of at least 1 count per million reads in more than one sample and at least 10 counts per million reads in one sample were retained. Trimmed mean of M-values (TMM) normalization was applied. The coefficient of biological variation (the square root of dispersion under the negative binomial model) was estimated across the entire dataset for each tissue and for each gene. Differentially expressed genes were then identified using the exact negative binomial test, which is appropriate since a single variable (turmeric consumption) was being investigated. Significance was corrected for multiple comparisons using the method of Benjamini-Hochberg (Benjamini and Hochberg, 1995), as reported in **Table S11**.

Ultra High Performance Liquid Chromatography-Mass Spectrometry (UPLC-MS)

Sample processing - Frozen fecal samples were combined with 20 volumes of ice cold methanol and 1 volume of cholic acid-$^{13}C_1$ (200 mg/ml; Sigma-Aldrich, St. Louis, MO), and then shaken for 2 minutes in a bead beater (Biospec Products; maximal setting; no beads added) before incubation at -20°C for 1 hour. Samples were then centrifuged at 4°C for 10 minutes at 20,800 x g. The supernatant (200 μL) was collected and dried in a SpeedVac at room temperature (requiring a spin time of 2-3 hours). Dried samples were re-suspended in 100 μL of 5% ethanol (in water) by a combination of pipetting, vigorous shaking in a bead beater, and sonication. After centrifugation for 5 minutes at 20,800 x g at 4°C, the supernatant was separated for UPLC-MS. Analyses were performed on a Waters Acquity I Class UPLC system (Waters Corp., Milford, MA) coupled to an LTQ-Orbitrap Discovery (Thermo Fisher Corporation). For the 150 mm x 2.1 mm Waters BEH C18 1.7 mm particle column, injection volume was 5 μL, and the flow rate was 0.3 mL per minute. Mobile phases for positive ion mode were (i) 0.1% formic acid in water and (ii) 0.1% formic acid in acetonitrile, whereas negative ion mode used (i) 5 mM ammonium bicarbonate in water and (ii) 5 mM ammonium bicarbonate in 95/5 acetonitrile/water. The capillary column was maintained at 325°C with a sheath gas flow of 40 (arbitrary units), an aux gas flow of 5 (arbitrary units) and a sweep gas

flow of 3 (arbitrary units) for both positive and negative injections. The spray voltage was 4.5 kV for the positive ion injection and 4 kV for the negative ion injection.

Data pre-processing - Data in instrument specific format (.D) were converted to common data format (.cdf) files using MSD ChemStation (E02.01, Agilent Technologies, Santa Clara, CA). The .cdf files were extracted using Bioinformatics Toolbox in MATLAB 7.1 (The MathWorks, Inc., Natick, MA), along with custom scripts (Cheng et al., 2011) for alignment of data in the time domain, automatic integration, and extraction of peak intensities. The resulting three-dimension data set included sample information, peak retention time, and peak intensities. Data were then mean-centered and unit variance-scaled for multivariate analysis, which was performed in R (version 3.1.2) (R Development Core Team, 2013) using custom scripts.

Quality control of metabolomics data - Pooled quality control (QC) samples were prepared from 8-10 μL aliquots of 10 samples and analyzed together with the other samples. QC samples were inserted and analyzed every 10 samples. Metabolite identification was done by co-characterization of standards.

Statistical analyses

Routine statistical analyses were performed in R (version 3.1.2) (R Development Core Team, 2013) using custom scripts that are available upon request. Linear modeling, including stepwise backward feature selection, statistical correlations and comparisons, and Bonferroni corrections for multiple comparisons were performed using functions within the default *stats* package (version 3.1.2).

Diet-discriminatory OTUs were identified using the Random Forests supervised machine learning algorithm as implemented in the *randomForest* package (version 4.6-10) (Liaw and Wiener, 2002) following 100 replications per diet-microbiota combination (parameters: ntree, 1000; importance scores tracked; otherwise default values). Initial modeling was performed to identify diet-discriminatory taxa in fecal samples obtained from mice colonized with the same donor microbiota. Mean importance scores were then aggregated, and hierarchical clustering was applied to examine whether OTUs were diet-discriminatory across

different microbiota. The accuracy of the model built on the basis of the 6-phase travel experiment results was tested on experimental data from the 3-phase travel using 10,000 replications (parameters: ntree, 1000; importance scores tracked; otherwise default values). Out-of-bag error estimations were determined by iterating through OTUs that had been ranked by mean importance score to determine the performances of the resulting Random Forests models.

Figures were generated using R, employing for several figures *ggplot2* (version 1.0.0) (Wickham, 2009) and *pheatmap* (version 0.7.7) (Kolde, 2013) packages.

REFERENCES

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B Methodol. *57*, 289–300.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat. Methods *7*, 335–336.

Cheng, J., Xia, Y., Zhou, Y., Guo, F., and Chen, G. (2011). Application of an ultrasound-assisted surfactant-enhanced emulsification microextraction method for the analysis of diethofencarb and pyrimethanil fungicides in water and fruit juice samples. Anal. Chim. Acta *701*, 86–91.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., and Andersen, G.L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl. Environ. Microbiol. *72*, 5069–5072.

Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics *26*, 2460–2461.

Faith, J.J., McNulty, N.P., Rey, F.E., and Gordon, J.I. (2011). Predicting a human gut microbiota's response to diet in gnotobiotic mice. Science *333*, 101–104.

Ferrucci, L.M., Daniel, C.R., Kapur, K., Chadha, P., Shetty, H., Graubard, B.I., George, P.S., Osborne, W., Yurgalevitch, S., Devasenapathy, N., et al. (2010). Measurement of spices and seasonings in India: Opportunities for cancer epidemiology and prevention. Asian Pac. J. Cancer Prev. APJCP *11*, 1621–1629.

Islam, S.N., Khan, M.N.I., and Akhtaruzzaman, M. (2010). A Food Composition Database for Bangladesh with Special reference to Selected Ethnic Foods (INFS-NFPCSP-FAO, Dhaka, Bangladesh).

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. *28*, 27–30.

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2014). Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. *42*, D199–D205.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Kolde, R. (2013). pheatmap: Pretty Heatmaps.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Liaw, A., and Wiener, M. (2002). Classification and Regression by randomForest. R News *2*, 18–22.

Magoč, T., and Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinforma. Oxf. Engl. *27*, 2957–2963.

McNulty, N.P., Yatsunenko, T., Hsiao, A., Faith, J.J., Muegge, B.D., Goodman, A.L., Henrissat, B., Oozeer, R., Cools-Portier, S., Gobert, G., et al. (2011). The Impact of a Consortium of Fermented Milk Strains on the Gut Microbiome of Gnotobiotic Mice and Monozygotic Twins. Sci. Transl. Med. *3*, 106ra106.

Mondal, D., Minak, J., Alam, M., Liu, Y., Dai, J., Korpe, P., Liu, L., Haque, R., and Petri, W.A. (2012). Contribution of enteric infection, altered intestinal barrier function, and maternal malnutrition to infant malnutrition in Bangladesh. Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am. *54*, 185–192.

National Research Council (US) Subcommittee on Laboratory Animal Nutrition (1995). Nutrient Requirements of Laboratory Animals (Washington, D.C.: National Academies Press).

R Development Core Team (2013). R: A language and environment for statistical computing (Vienna, Austria: R Foundation for Statistical Computing).

Ridaura, V.K., Faith, J.J., Rey, F.E., Cheng, J., Duncan, A.E., Kau, A.L., Griffin, N.W., Lombard, V., Henrissat, B., Bain, J.R., et al. (2013). Gut Microbiota from Twins Discordant for Obesity Modulate Metabolism in Mice. Science *341*, 1241214.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinforma. Oxf. Engl. *26*, 139–140.

Smith, M.I., Yatsunenko, T., Manary, M.J., Trehan, I., Mkakosya, R., Cheng, J., Kau, A.L., Rich, S.S., Concannon, P., Mychaleckyj, J.C., et al. (2013). Gut Microbiomes of Malawian Twin Pairs Discordant for Kwashiorkor. Science *339*, 548–554.

Subramanian, S., Huq, S., Yatsunenko, T., Haque, R., Mahfuz, M., Alam, M.A., Benezra, A., DeStefano, J., Meier, M.F., Muegge, B.D., et al. (2014). Persistent gut microbiota immaturity in malnourished Bangladeshi children. Nature *510*, 417–421.

Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., Sogin, M.L., Jones, W.J., Roe, B.A., Affourtit, J.P., et al. (2009). A core gut microbiome in obese and lean twins. Nature *457*, 480–484.

Wickham, H. (2009). ggplot2: elegant graphics for data analysis (Springer New York).

Yatsunenko, T., Rey, F.E., Manary, M.J., Trehan, I., Dominguez-Bello, M.G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R.N., Anokhin, A.P., et al. (2012). Human gut microbiome viewed across age and geography. Nature *486*, 222–227.