SUPPLEMENTARY MATERIAL

SUPPLEMENTARY MATERIALS AND METHODS

Microbial genome sequencing

The following bacterial strains are incorporated into the commercially available FMP: *Bifidobacterium animalis* subsp. *lactis* (strain CNCM I-2494); *Lactobacillus delbrueckii* subsp. *bulgaricus* (strains CNCM I-1632, CNCM I-1519), *Lactococcus lactis* subsp. *cremoris* (strain CNCM I-1631), and *Streptococcus thermophilus* (strain CNCM I-1630). We performed shotgun 454 FLX pyrosequencing of both *L. delbrueckii* subsp. *bulgaricus* strains (CNCM I-1632, CNCM I-1519), plus the *L. lactis* subsp. *cremoris* strain (33-, 32- and 51-fold coverage, respectively). Using the Newbler assembler (v1.1, 454 Life Sciences) and already sequenced strains of these species, we obtained draft genome assemblies with N50 contig sizes of 69,466 and 55,626 and 55,851 bp, respectively. The total sizes of the assembled *L. delbrueckii* genomes were 1,767,897 bp (CNCM I-1632) and 1,795,730 bp (CNCM I-1519), while the *L. lactis* assembly had an aggregate genome size of 2,511,332 bp. A finished genome sequence of the *B. animalis* subsp. *lactis* genome and a deep draft assembly of the *S. thermophilis* genome were previously generated by Integrated Genomics (see **Table S3** for a summary of genome metrics).

Annotation and comparative genomic analysis

The predicted proteins of all the sequenced bacterial species used in this study were annotated using BLASTP searches (*E*-value $<10^{-5}$) against version 54 (v54) of the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (*S1-3*) and the Carbohydrate Active Enzyme (CAZy) database (*S4*). BLAST results were parsed into a lookup table for each genome, and each fecal microbiome, using a Perl script (lookup_KEGG_for_genes_BLAST.pl) that assigned to each gene the KEGG Orthology (KO) identifier associated with its top BLAST hit (*E*-value $<10^{-5}$). In cases where the query protein's top match to the reference database was ambiguous (i.e., when there were multiple top matches with equally good scores), the gene was annotated with the KOs of all the top-matching database entries.

These lookup tables were then used to calculate the 'coverage ratio' of each KEGG metabolic pathway for each bacterial genome in this study using the Perl script kegg_key.pl. For each edge (enzymatic reaction) of each KEGG metabolic pathway, this script checks for the presence of a gene whose assigned KO corresponds with that reaction in (i) a given microbial genome, (ii) defined collections of microbial genomes, or (iii) larger, incompletely sequenced microbiomes. KEGG pathway coverage ratios for several combinations of genomes, as well as several published gut microbiomes, were calculated and used to perform average linkage hierarchical clustering in Cluster 3.0 (*S5*) using an un-centered correlation similarity metric. A heatmap visualization of this clustered data was generated using the Java Treeview application (jtreeview.sourceforge.net).

Culturing of *B. animalis* subsp. *lactis*

A frozen stock of *B. animalis* subsp. *lactis* (strain CNCM I-2494) was streaked out on MRS-agar plates (BD/Difco) and transferred to a Coy chamber for overnight growth at 37^{o}C under an atmosphere of 5% H_2, 20% CO_2 and 75% N_2. Single colonies were picked and inoculated into 10 ml of pre-reduced MRS broth (BD/Difco) that had been stored for 24 h in an anaerobic chamber. The medium was not supplemented with cysteine. Each culture was passaged four times to stationary phase, and growth curves were recorded to discern growth kinetics under these conditions. 100 μl of an overnight culture was used to inoculate 10 ml of fresh anaerobic MRS in 27 ml Balch tubes with rubber stoppers and aluminum crimp tops. The initial headspace of the tubes was composed of 5% H_2, 20% CO_2, and 75% N_2 at ambient pressure. Tubes were incubated at 37^{o}C, and 4 ml aliquots were collected at mid-log phase (9 h post-inoculation, OD_{600}=0.2) and during late stationary phase (27 h post-inoculation, OD_{600}=2.6). Each aliquot from each culture (n=2) was immediately combined with 8 ml of RNAprotect Bacteria Reagent (Qiagen), incubated for 5 min at room temperature, then centrifuged (3,220 x *g*; 15 min at 25^{o}C). The pellets were snap-frozen in liquid nitrogen and stored at -80^{o}C, and total cellular RNA was subsequently isolated as described previously (*S6*).

Human studies

Production and distribution of the FMP to study participants – The FMP used for this study was produced in Danone's pilot plant located in Fort Worth, TX. Batches were shipped directly to subjects by an independent delivery service so that the names of study participants would remain unknown to all but those in the MOAFTS study group. Each subject received 4 shipments of FMP, spaced at 2-week intervals. Each shipment was composed of sufficient numbers of cups (pots) so that study participants could consume one 4 oz serving twice a day (each serving consisted of a single pot). Each co-twin chose her flavors (strawberry, vanilla, and/or peach). The same flavor selection was shipped each time. Each subject was allowed to vary the sequence of selected flavors according to her wishes.

Analysis of the effects of FMP consumption on stool parameters – Stool consistency, difficulty of passage, and frequency were assessed using a daily stool diary in which participants recorded the time of day for each bowel movement. Participants rated the stool consistency using the seven point Bristol Stool Form Scale (*S7*) and the difficulty of passage using a five-point scale (no difficulty to extreme difficulty).

Quantitative PCR (qPCR) analysis of fecal levels of FMP strains – qPCR was used to define the levels of selected FMP strains in fecal samples obtained from MZ co-twins and gnotobiotic mice (*S8-12*). The PCR primer sets targeting each strain's 16S rRNA gene or CRISPR locus are described in **Table S14**, as are the amplification conditions. Samples were analyzed on an Applied Biosystems 7900HT instrument using SYBR Green chemistry. Standard curves were constructed using genomic DNA prepared from a known number of bacterial cells harvested from monocultures grown to stationary phase (cells were DAPI stained and counted by microscopy); C_t values for each reaction could, therefore, be expressed in terms of cell equivalents (CE).

To calculate the concentration of a given bacterial strain in each fecal sample, three serial dilutions of extracted fecal DNA (10 ng, 1 ng, 0.1 ng) were assayed in at least two independent qPCR reactions. C_t values falling within the linear range of the assay were referenced to the standard curves, while those outside the linear range were excluded from the analysis. For human samples, data were log-transformed and normalized to fecal mass ($\log_{10}$ CE/g feces). For mouse samples, data were log-transformed and normalized to mass of template DNA ($\log_{10}$ CE/μg DNA).

Multiplex pyrosequencing of 16S rRNA genes in fecal samples and the FMP – A total of 126 fecal samples (9 samples per individual) were collected over the course of 4 months according to the schedule shown in **Fig. 1**. All fecal samples were frozen at -20^{o}C within 30 min after they were produced. The samples were maintained at this temperature for the period of time required to ship them to a biospecimen repository (less than 24 h). As soon as samples were received, they were de-identified and stored at -80°C. DNA was extracted from frozen, pulverized fecal samples by bead beating followed by phenol-chloroform extraction, as described previously (*S13*). Methods for generating and performing multiplex pyrosequencing of amplicons from variable region 2 (V2) of bacterial 16S rRNA genes are described in (*S13*). Bacterial V2 16S rRNA gene sequencing data were pre-processed to remove sequences with low quality scores, sequences with ambiguous characters, and sequences outside the length bounds (200-300 nt). All subsequent data processing and analyses were done using QIIME software (*S14*). Pyrosequencing 'noise' was removed with an algorithm implemented in QIIME. 16S rRNA reads were binned according to their sample-specific, error-correcting barcode incorporated into the reverse primer. Similar sequences were binned into phylotypes using CD-HIT with minimum pairwise identity of 97% (*S15*).

Aliquots of freshly produced as well as 30 day-old FMP from 6 batches of each flavor were sent directly from the pilot production plant to one of our labs using the same shipping protocol that was used to deliver the FMP to study participants. DNA was extracted and amplicons from the V2 region of bacterial 16S rRNA were generated and sequenced using the protocols described above. 49,959 high quality reads were obtained from a total of 33 FMP samples (1,332±187 reads/sample (mean ± S.D)): 43,729 reads of these were classified using the GreenGenes database (http://greengenes.lbl.gov/cgi-bin/nph-index.cgi) and belonged to the genera Streptococccus, Lactococcus, Lactobacillus or Bifidobacterium.

Co-occurrence analysis – To determine whether there were statistically significant associations between the presence of *B. animalis* subsp. *lactis* and the occurrence of resident gut bacterial species-level phylotypes in human fecal samples, a co-occurrence analysis was performed using software tools present in QIIME under the script otu_category_significance.py. We used this script to employ an

ANOVA test to search our fecal 16S rRNA datasets for phylotypes whose relative abundances were higher in samples in which *B. animalis* subsp. *lactis* was present versus samples in which this strain was absent, as determined using qPCR. To avoid biases that might be introduced by differences in sample sequencing depth, we randomly selected an even number of sequences/sample (1,644 sequences) prior to performing the analysis. The raw *p*-values were corrected for multiple tests using the false discovery rate (fdr) correction (*S16*). We also performed the analysis at the genus and family levels by binning all operational taxonomic units (OTUs) that mapped to the same family or genus based on classification with the RDP classifier, using the summarize_taxa.py script in QIIME.

Shotgun sequencing of fecal microbiomes – Forty-eight fecal samples from four twin pairs were selected for multiplex shotgun pyrosequencing of total community DNA (454 FLX chemistry). For each individual, two fecal samples were analyzed before initiation of FMP consumption, two samples during the period when FMP was being consumed, and two samples after consumption had ceased.

Each fecal community DNA sample was randomly fragmented to an average length of 500 bp by nebulization, and then labeled with a distinct multiplex identifier (MID; Roche) using the manufacturer's protocol. Equivalent amounts of 12 MID-labeled samples from each family were pooled prior to each pyrosequencing run. Shotgun reads were subsequently filtered using publicly available software (*S17*) to remove (i) all reads less than 60 nt, (ii) LR70 reads with at least one degenerate base (N) or reads with two continuous and/or three total degenerate bases, (iii) all duplicates, defined as sequences whose initial 20 nt were identical and shared an overall identity of >97% throughout the length of the shortest read, and (iv) all sequences with significant similarity to human reference genomes (i.e., having a BLASTN match with *E*-value $\leq 10^{-5}$, bitscore ≥ 50, and percent identity $\geq 75\%$) to ensure continued de-identification of samples.

Reads obtained from shotgun sequencing of the twins' fecal microbiomes were used to query v54 of the KEGG GENES database using BLASTX (*E*-value $<10^{-5}$, bitscore >50, and %identity >50). A comparable annotation was performed for published fecal microbiomes that had been generated from 124 deeply sampled unrelated adult Europeans (*S18*), and from a pair of obese adult MOAFTS twins (*S19*).

Microbial RNA-Seq analysis of human fecal metatranscriptomes – Methods used to prepare fecal RNA, generate RNA-Seq data, and perform computational analyses of these datasets are described in detail below and in one of our previously published studies (*S6*, **Fig. S4B**). The only difference in the analysis pipeline used for human versus mouse RNA-Seq datasets was the choice of reference genomes. To map RNA-Seq reads from human fecal samples to genes, we used a collection of 131 bacterial genomes encompassing much of the diversity documented in human gut communities, plus the genomes of the five FMP strains; these genomes are listed in **Table S15**.

Studies in gnotobiotic mice

Using INSeq to assay the determinants of fitness in a saccharolytic member of the 15-species model human microbiota – INsertion Sequencing (INSeq) is a method based on a mutagenic transposon modified so that discrete fragments of adjacent chromosomal DNA can be captured when the transposon is excised from bacterial genomes by the restriction enzyme MmeI (*S20*). Sequencing fragments excised from a mixed population of tens of thousands of transposon mutants provides information about the location of each transposon in the genome. The number of occurrences of the transposon insertion site sequence mirrors the relative abundance of that mutant in the mixed population. By identifying mutants that significantly decrease in relative abundance after passage through a selective condition, INSeq allows a genome-wide map of *in vivo* fitness determinants to be created.

To determine whether introduction of the FMP strain consortium results in differences in the *in vivo* fitness requirements of a human gut symbiont present in the 15-member community, *B. thetaiotaomicron* strain VPI-5482 was mutagenized with the INSeq transposon (*S20*). A library of 34,544 randomly inserted transposon mutant strains covering 3,435 of the organism's 4,779 genes was introduced, by gavage, together with the other 14 non-mutagenized members of the community into germ-free mice. Fecal samples were subsequently collected from each mouse (n = 10) before (d7, d14), immediately after (d16), and 7 d after (d21) initial introduction of the FMP strains. INSeq libraries were prepared as described (*S20*) and sequenced using an Illumina GA IIx instrument (~1,000,000 36 nt reads/sample; **Table S6**). Resulting sequences were mapped to the *B. thetaiotaomicron* reference genome and quantified

as described (*S20*). We found that insertions in 626 genes showed a significantly decreased relative abundance in the day 14 fecal microbiota (multiple hypothesis testing-corrected q<0.001), reflecting a fitness requirement for these genes in the colonization process (**Table S6**). Analysis of fecal samples collected just prior to, one day after, and 7 days after introduction of the FMP consortium established that exposure to the FMP strains did not impose significant new fitness pressures on specific genes present in this saccharolytic bacterial species.

Animal sampling – Fecal samples were obtained from each animal at the time points indicated in **Fig. 1**. Each fecal sample was collected directly as it emerged from the anus into a 1.7 ml screw cap Eppendorf tube, which was immediately deposited in a stainless steel dewer containing liquid N_2 (the dewer was introduced into the gnotobiotic isolator on the day of collection after it had been sterilized in the isolator's entry port with chlorine dioxide spray (Clidox-S; PRL Pharmacal)). Various subsets of samples were subjected to COPRO-Seq, INSeq and microbial RNA-Seq analyses. Blood samples were collected into lithium heparin tubes (Becton Dickinson), placed immediately on ice, and then centrifuged (2,700 x *g*; 3 min at 4°C). The resulting plasma supernatant was stored at -80^{o}C until assay. Urine was collected directly into Eppendorf tubes and immediately frozen in liquid N_2. Upon sacrifice, ceca were dissected and cecal contents were frozen immediately at -80^{o}C.

Isolation of DNA from cecal contents and feces – Microbial community DNA was prepared in a two-step process consisting of a crude extraction step followed by additional purification and RNAse treatment.

Crude extraction. The sample (typically 25-100 mg of frozen feces or 50-125 mg of frozen cecal contents) was combined with 250 µl of 0.1 mm zirconium beads (BioSpec Products), 500 µl Buffer A (200 mM NaCl, 200 mM Tris, 20 mM EDTA), 210 µl SDS (20% v/v, filter-sterilized), and 500 µl phenol:chloroform:isoamyl alcohol (25:24:1, pH 7.9; Ambion), and the mixture was briefly chilled on ice. Samples were then disrupted using a Mini-BeadBeater-8 (BioSpec) set to 'homogenize' (bead-beating for 2 min at room temperature, followed by placement on ice for 1-2 min, followed by bead-beating for an

additional 2 min). The aqueous phase (~600 µl) was collected after centrifugation (6,800 x *g*; 3 min at 4°C) and combined with an equal volume of phenol:chloroform in 2 ml 'light' phase-lock gel tubes (5Prime) per the manufacturer's protocol. After spinning, the aqueous phase was combined with 1 volume of chilled 100% isopropanol (-20°C) and 1/10 volume sodium acetate (3 M, pH 5.5). Following incubation at -20°C for 1 h, the precipitate was pelleted (20,800 x *g*; 20 min at 4°C), washed in 100% EtOH, dried, and resuspended in 5 µl TE (pH 7.0) per milligram of original sample material.

RNAse treatment and further purification. Aliquots of crude DNA were transferred to a 96-well plate. Buffer PM (Qiagen) was mixed with RNase A (Qiagen) to a final concentration of 1.3 mg/ml. Three volumes of this mixture were added to each well and the reactions were allowed to incubate at room temperature for 2 min. Following RNase digestion, samples were applied to a QIAquick 96 PCR purification plate (Qiagen) and processed according to the manufacturer's instructions using a QIAvac 96 vacuum manifold. DNA was eluted in 100 µl of Buffer EB (Qiagen). DNA quality and purity were verified using a NanoDrop spectrophotometer (model ND-1000).

Preparing DNA libraries for Illumina sequencing and COPRO-Seq analysis in a 96-well format – DNA libraries were prepared for sequencing using a modified version of Illumina's sample preparation protocol for generating libraries from genomic DNA. The six steps of this modified preparation are as follows:

(i) Fragmentation. Two micrograms of each purified DNA sample was suspended in 100 µl Buffer EB and fragmented by sonication in 1.7 ml Eppendorfs using the BioruptorXL multi-sample sonicator (Diagenode) set on 'high'. Samples were sonicated over the course of 20 min using successive cycles of 30 sec 'on' followed by 30 sec 'off '. Sonicated samples were subsequently cleaned up using the MinElute 96 UF PCR Purification Kit (Qiagen) per the manufacturer's instructions. Each sonicated DNA sample in each well of the 96-well plate was eluted with 22 µl Buffer EB.

(ii) 'Add-only' enzymatic modification. Ten microliters of each eluate from step (i) was transferred to a 96-well plate where it was subjected to enzymatic blunting. Each 20 µl reaction contained: 10 µl DNA,

2 µl T4 DNA ligase buffer [10X; New England Biolabs (NEB), catalogue number B0202S], 1 µl dNTPs (1 mM; NEB, N1201AA), 0.5 µl Klenow DNA polymerase (5 U/µl; NEB, M0210S), T4 PNK (10 U/µl; NEB, M0201S), and 6 µl molecular grade water. Blunting reactions were incubated (25^{o}C, 30 min) then heat-inactivated (75^{o}C, 20 min). Residual dNTPs were dephosphorylated by adding 1 µl of shrimp alkaline phosphatase (1 U/µl; Promega, M820A) to each reaction. Reactions were incubated (37^{o}C, 30 min) and heat-inactivated (75^{o}C, 30 min). Adenine tailing reactions were set up in 30 µl reaction volumes that contained 21 µl of the inactivated phosphatase reaction, 6.4 µl T4 DNA ligase buffer (diluted to 1X; NEB, B0202S), 0.6 µl dATP (5 mM), and 2 µl Klenow 3'->5' exo^{-} (5 U/µl; NEB, M0212L). Reactions were incubated (37 ^{o}C, 30 min) and heat-inactivated (75^{o}C, 20 min).

(iii) Ligation. Customized Illumina adapters containing maximally distant 4 bp barcodes described elsewhere (*S6*) were ligated to the polyA-tailed DNA in 50 µl reactions as follows. Thirty microliters of the inactivated A-tailing reaction described in the preceding paragraph was added to 5 µl T4 DNA ligase buffer (10X), 5 µl adapter mix (1 µM final concentration per adapter), and 9 µl water at 4^{o}C. One microliter of T4 DNA ligase (2,000,000 U/µl; NEB M0202M) was subsequently added and reactions were incubated (16^{o}C, 1 h), followed by heat-inactivation (65^{o}C, 10 min). Ligation reactions were cleaned up using the MinElute 96 UF PCR Purification Kit (Qiagen) according to the manufacturer's recommended protocol. DNA was eluted in 22 µl Buffer EB.

(iv) Gel purification. 10 µl of each elution was separated by gel electrophoresis on 2% agarose. DNA migrating at 200 bp was excised and gel slices were purified using a QIAquick 96 PCR Purification Kit (Qiagen).

(v) PCR amplification. Each library was PCR amplified for 19 cycles using Illumina's standard amplification primers with modifications to impart barcode-specificity (*S6*) and Illumina's recommended amplification conditions/reagents. Products were purified using a QIAquick 96 PCR Purification Kit (Qiagen). An aliquot of this purification was also subjected to 2% agarose gel electrophoresis to confirm the absence of any significant adapter-dimer contamination.

(vi) Library pooling and sequencing. The concentration of each purified library was quantified using the Qubit dsDNA HS Assay Kit (Invitrogen). Barcoded libraries were subsequently pooled (typically in groups of 16) at equivalent final concentrations. Sequencing was performed using the standard Illumina GA IIx sequencing protocol, with libraries loaded on the flow cell at a concentration of 2.0-2.5 pM).

A custom software pipeline was written in Perl for performing COPRO-Seq data processing in a computer cluster environment running Sun Grid Engine. These data processing steps are schematized in **Fig. S4A**. Briefly, raw Illumina GA IIx reads from a sequencing pool were first de-plexed by barcode and trimmed to 34 bp (30 bp genome sequence + 4 bp barcode). Trimmed reads were aligned to the genomes of the 20 microbial strains used in this study using Illumina's ELAND aligner. Perfect, unique alignments to the reference genomes were retained, while those mapping less than perfectly or having multiple possible alignments to the reference genomes were filtered out, ensuring that only high-quality, unambiguous reads were used. Hits to each genome were then tallied, after which the summed counts for each genome were normalized by that genome's 'informative genome size' (term defined in *Results*) to adjust for both genome size and uniqueness relative to all other genomes in the experiment. The Perl scripts supporting the COPRO-Seq analytic pipeline can be downloaded from: http://gordonlab.wustl.edu/projects/2011-McNulty_etal.

Characterizing gene expression with microbial RNA-Seq – Following extraction of total nucleic acid with phenol-chloroform, and precipitation with isopropanol, fecal samples were subjected to DNAse digestion (*S6*). Total RNA was then: (i) passed through a MEGAClear column (Ambion) to deplete RNAs <200 nt (removing most 5S rRNA and tRNA species; (*S6*)); (ii) subjected to another round of DNAse digestion; (iii) passed through another MEGAClear column; and (iv) subjected to a hybridization-based pulldown of 16S and 23S rRNA using custom-designed biotinylated oligonucleotides that contain short rDNA sequences conserved across a set of 37 human gut-derived sequenced microbial genomes (*S6*). The depletion protocol, which has been adapted to 96-well format, is described elsewhere (*S6*). PCR (30 cycles) targeting a short (<100 bp) region from the genome of an abundant community member (typically *Bacteroides WH2*) was used to verify the absence of detectable gDNA in each RNA preparation.

Doubled stranded cDNA (dscDNA) was synthesized using random hexanucleotide primers. At the conclusion of the reaction, Illumina adapters containing sample-specific 4 nt barcodes were ligated to the dscDNA. Multiplex sequencing was performed using the Illumina GA IIx instrument. We typically sequenced two barcoded *in vitro* samples/lane of the 8-lane flow cell; *in vivo* samples were not multiplexed (i.e. 1 sample was analyzed per sequencing lane). This allowed us to identify mRNA present at levels representing ≥0.001% of all reads.

The pipeline for processing microbial RNA-Seq data is presented in **Fig. S4B**. The 8-20 million 36 nt cDNA reads from each sequencing lane were separated by barcode, and mapped against the relevant set of genome sequences using the SSAHA2 algorithm (*S21*) to determine the raw unique-match 'counts' (reads) for each gene present in the relevant microbial genome or microbiome. Reads that mapped non-uniquely were added to each gene in proportion to each gene's fraction of unique-match counts (e.g., a non-unique read that maps equally well to gene A with 18 unique reads and gene B with 2 unique reads would be scored as 0.9 of a count to gene A, and 0.1 of a count to gene B; the influence of ties is negligible for RNA-Seq given the small numbers of distinct genomes, but would become more important with more complex communities). Raw counts were then normalized to reads/kb gene length/million mapped reads (RPKM) using one or more gene position file(s) in conjunction with custom Perl scripts.

Data normalization was carried out at two different levels in this study. For our 'top-down' analysis, data were normalized at the level of the entire community metatranscriptome (i.e., raw counts from all species were normalized simultaneously using a single gene position file that included the positions of all genes in the model community metatranscriptome). Data normalized in top-down fashion allowed us to determine, after binning gene expression values by function, how the collective operations of the model community were changing as the result of experimental perturbations. In our 'bottom-up' analysis, data were normalized at the level of individual species (i.e., raw counts from each individual species were normalized separately from one another, in each case using a species-specific gene position file describing the positions of only that species' genes). Data normalized in bottom-up fashion allowed us to

interrogate what statistically significant gene expression changes were occurring within a given species of interest.

Bottom-up normalized transcript data were analyzed by Cyber-T (*S22*) to identify mRNAs that exhibited significant differences in their levels of expression between samples. For each comparison, transcripts were then binned into a list where the magnitude of the difference in their expression was greater than or equal to 4-fold. Binned transcripts were subsequently annotated using the kegg_counting.pl Perl script described above. Each resulting annotated dataset was used to determine the representation of KEGG pathways and categories within these lists.

Further functional comparisons were carried out using ShotgunFunctionalizeR, an R package designed to analyze differences between metagenomic datasets using a Poisson statistical model (*S23*). The RPKM-normalized counts for each gene in the model microbiome were then binned by EC number using each gene's previously assigned annotation. Summed reads in each EC bin were rounded to the nearest integer, and the data were imported into ShotgunFunctionalizeR, which was used to identify ECs that were differentially expressed.

Identification of predictive KEGG categories, pathways and ECs using a Random Forests classifier – To identify KEGG categories, ECs, or pathways that were significantly differentiated across treatment states, we used the Random Forests classifier (*S24*) described in *(S25)*. Mouse samples were divided into 10 pre-treatment samples (experimental day 14) and 20 post-treatment samples (experimental days 15 and 42). To estimate the generalization error of the classifier we used leave-one-out cross-validation, in which each sample's group was predicted by a classifier trained on the other 29 samples. Training was done using default settings for the randomForest package in R (*S24*). Each feature's predictiveness was estimated by calculating the mean increase in estimated generalization error when the values of that feature were permuted at random. Features whose removal caused an average error increase of at least 0.1% were labeled as 'predictive'; those with an increase of at least 1% were labeled as 'highly predictive'.

Non-targeted metabolomics via gas chromatography/mass spectrometry (GC/MS) – Urines were first assayed for creatinine content as measured by a modified Jaffe method using the three microliter "random-urine" routine and CR-S 3000 reagent on the UniCel DxC 600 Synchron Clinical System (Beckman Instruments, Brea, CA). A urine volume equivalent to 0.2 micromoles of creatinine was then aliquoted and spiked with perdeuterated myristic acid (D_{27}-C14:0) as an internal standard for retention-time locking (RTL IS). Following treatment with 7.5 volumes of methanol, the mixture was centrifuged and the supernatant was decanted and dried.

Derivatization of all dried supernatants for GC/MS followed a method adapted with modifications from that of (*S26*). Reagents were from Sigma-Aldrich (St. Louis, MO), unless otherwise noted. Briefly, certain reactive carbonyls were first methoximated at 50ºC with a saturated solution of methoxyamine hydrochloride in dry pyridine, followed by replacement of exchangeable protons with trimethylsilyl (TMS) groups using *N*-methyl-*N*-(trimethylsilyl) trifluoroacetamide with a 1% v/v catalytic admixture of trimethylchlorosilane (Thermo-Fisher Scientific, Rockford, IL) at 50ºC.

GC/MS methods generally followed those of Fiehn (*S27*) and Kind (*S28*), and used a 6890N GC connected to a 5975 Inert single-quadrupole MS (Agilent, Santa Clara, CA). A large-volume, ProSep inlet enabled programmed-temperature vaporization and diversion of heavy contaminants away from the GC and MS, as described below, greatly reducing maintenance time (Apex Technologies, Inc., Independence, KY). The two wall-coated, open-tubular GC columns connected in series were both from J&W/Agilent (part 122-5512), DB5-MS, 15 m in length, 0.25 mm in diameter, with a 0.25 µm luminal film. Prior to each run, initial inlet pressures were empirically adjusted such that the resulting retention time (RT) of the TMS-D27-C14:0 standard was set at ~16.727 min. Under these conditions, derivatized metabolites eluted from the column and reached the electron-ionization (EI) source in the MS at known times (*e.g.,* bis-TMS-lactic acid at ~6.85 min, and TMS-cholesterol at ~27.38 min). A mid-column, microfluidic splitter (Agilent) provided a means for hot back-flushing of the upstream GC column at the end of each run while the oven was held at 325ºC for a terminal "bake-out" (another antifouling and anti-carryover measure analogous to that described in (*S29*)). During this terminal "bake-out," the inlet was also held at 325ºC,

and it was purged with a large flow of the carrier gas, helium. Positive ions generated with conventional EI at 70 eV were scanned broadly from 600 to 50 m/z in the detector throughout the run.

Raw data from Agilent's ChemStation software environment were imported into the freeware, Automatic Mass Spectral Deconvolution and Identification Software (AMDIS) (developed by Drs. Steve Stein, W. Gary Mallard, and their coworkers at National Institute of Standards and Technology (*S30-32*); http://chemdata.nist.gov/mass-spc/amdis/). Deconvoluted spectra were identified, to the extent possible, using several commercial and public spectral libraries. Our primary source was the Fiehn GC/MS Metabolomics RTL Library (a gift from Agilent Technologies, Santa Clara, CA, part number G1676-90000). Additional spectra for comparison were gleaned from the Golm Metabolome Library (courtesy of Dr. Joachim Kopka and coworkers at the Max Planck Institute of Molecular Plant Physiology, Golm, Germany (*S33*); http://csbdb.mpimp-golm.mpg.de/csbdb/gmd/gmd.html), the commercial NIST/EPA/NIH Mass Spectral Library and our own purpose-built spectral library. Where indicated, peak alignment was performed with SpectConnect freeware (courtesy of Dr. Gregory Stephanopoulos, Massachusetts Institute of Technology (*S34*); http://www.spectconnect.mit.edu). Chemometrics were performed with Mass Profiler Professional (a recent descendant of GeneSpring MS, purchased from Agilent), along with our own custom macros, written in Visual Basic for use in the Excel software environment.

The statistical significance of differences in the $\log_2$ spectral abundances of each metabolite in samples obtained at different time points was tested using two approaches. A first-pass, highly-permissive set of pairwise comparisons was calculated between each combination of samples (d0 versus d14, d0 versus d42, d14 versus d42) using a simple two-tailed Student's *t*-test. The resulting *p*-values, which were not corrected for multiple hypothesis testing, are listed in **Table S12**. Given the high ratio of hypotheses tested to samples per group (198 metabolites; 5-8 samples per group), we also guarded against false discovery by performing a more stringent set of calculations that produced a shorter list of metabolites with significant differences in abundance. This latter procedure consisted of first taking metabolite data from all three time points and subjecting them to a one-way ANOVA. The resulting *p*-values were then

adjusted using Benjamini-Hochberg correction, generating *q*-values. The $\log_2$ spectral abundances of all metabolites whose *q*-values were below 0.05 were then subjected to Tukey's HDS (Honestly Significantly Different) post-hoc test to determine which time points were significantly different from one another. All Tukey's HDS *p*-values that were calculated are provided in **Table S12**.

SUPPLEMENTARY RESULTS

Human studies

Analysis of the effects of FMP consumption on stool consistency, difficulty of passage, and frequency – To determine whether there were differences in stool consistency and difficulty of passage of stools between pre-treatment, treatment, and post-treatment study periods, we first constructed a dataset in which the unit of analysis was 'bowel movement'. Using ordinal logistic regression, we analyzed separate models predicting stool consistency and difficulty of stool passage using 'treatment period' as the reference group. We adjusted for clustering of observations using a Huber-White robust variance estimator (STATA 2004). When data from the entire study period were included in the analyses, no significant differences were observed between study periods for either stool consistency or difficulty passing stool. Next, we conducted an alternate analysis in which data from the first two weeks of the treatment and of the post-treatment phases were omitted. We found that women had lower stool consistency scores during the last two weeks of the post-treatment phase compared to the last two weeks of the treatment phase: i.e., stools were softer during the treatment period (OR=0.69; p=0.04). The difference in stool consistency between the pre-treatment phase and the last two weeks of the treatment phase was not significant; however, there was a significant difference between the pre-treatment versus treatment compared to the post-treatment versus treatment odds ratios (p=0.005).

Analyses for stool frequency were conducted similarly to those above with the exception that the unit of analysis was the 'person-day' (i.e., one observation per person per day) and the dependent variable was number of bowel movements per day. We did not find stool frequency to be associated with study period regardless of which study days were included in the analysis.

One participant had a diarrheal illness on three of the FMP treatment days, with a dramatic increase in stool frequency and decrease in stool firmness on these days. She reported taking four 2 mg tablets of loperamide [4-(p-chlorophenyl)-4-hydroxy-N, N-dimethyl- *a, a*-diphenyl-1-piperidinebutyramide monohydrochloride] to relieve her symptoms during this period. Therefore, data from these bowel movements were excluded from the analyses.

Co-occurrence analysis – As noted above, to identify species-level phylotypes that consistently increase or decrease in abundance when *B. animalis* subsp. *lactis* is present in human fecal samples, we performed a co-occurrence analysis using QIIME (see *Supp. Methods*). This analysis indicated that no OTUs present in the pre-treatment microbiota exhibited a statistically significant change in their proportional representation in feces during the period of FMP consumption or during the post-treatment period in any individual after correction for multiple tests. The OTU that most nearly achieved significance was closely related to *Lactococcus lactis* (raw p=0.00067; ANOVA: p>0.05 after FDR correction). A follow-up co-occurrence analysis for all genera also identified the genus Lactococcus as being significantly more abundant when *B. animalis* subsp. *lactis* was present (the latter determined by qPCR). It is reasonable that *L. lactis* would co-occur with *B. animalis* subsp. *lactis* given the presence of both strains in the FMP. A co-occurrence analysis performed at the family level of taxonomy failed to identify any significant differences.

Our ability to identify *L. lactis* in our co-occurrence analysis was encouraging, but raised the question of why an OTU representing *B. animalis* subsp. *lactis* did not achieve significance, given that nearly every sample collected during the period of FMP consumption was positive for this strain by qPCR. Of the 58 samples deemed positive for *B. animalis* subsp. *lactis* by qPCR, only 7 yielded an OTU in our 16S rRNA dataset with a 100% identity match to *B. animalis* subsp. *lactis*. This OTU was not detected in any of the samples deemed negative by qPCR. Therefore, we concluded that the discrepancy between *B. animalis* subsp. *lactis* being called 'present' by qPCR and by 16S rRNA sequencing was due to inadequate depth of sequencing. Extrapolating, there could be other rare species whose changes in abundance were not detected. To explore this latter possibility, the fecal microbiota of two healthy MZ twin pairs, similar in age and body mass index, but with marked differences in their degree of geographical proximity (**Table S1**), were subjected to deeper sequencing (n=36 samples, yielding an additional 411,177 16S rRNA sequences, resulting in 14,241±2,144 (mean±SD) reads/sample from these individuals). No significant changes at any level of bacterial taxonomy were observed in this small sample dataset.

Studies in gnotobiotic mice

Measurement of adiposity – The body weights and epididymal fat pad weights of mice from both treatment groups were measured at the time of sacrifice. We observed no significant differences between the single and multiple treatment groups in either measurement (p=0.6865, p=0.3516, respectively; two-tailed Student's t-test). Furthermore, all measurements of adiposity and weight were in line with those of mice from other studies that had involved animals from the same inbred strain, who were similarly aged, the same gender, on the same diet, and who harbored comparable defined model human gut microbiota (without FMP strains).

In vitro **studies**

RNA-Seq profiling of **B. animalis** ***subsp.*** **lactis** ***during growth*** **in vitro** – Sequencing of transcripts expressed by *B. animalis* subsp. *lactis* during mid-log phase growth in MRS medium (1.5-2.9 million reads per technical replicate; n=2 independent cultures) detected products from 1,618 of the organism's 1,660 predicted genes, while profiling during late stationary phase indicated that 1,609 of its genes were expressed. The transition from logarithmic to stationary phase was accompanied by significant up- or down-regulation of 98 and 194 genes, respectively, including those involved in various aspects of carbohydrate, amino acid and nucleotide metabolism (see **Table S7A).**

SUPPLEMENTARY FIGURES

McNulty, *et al.*, Fig. S1

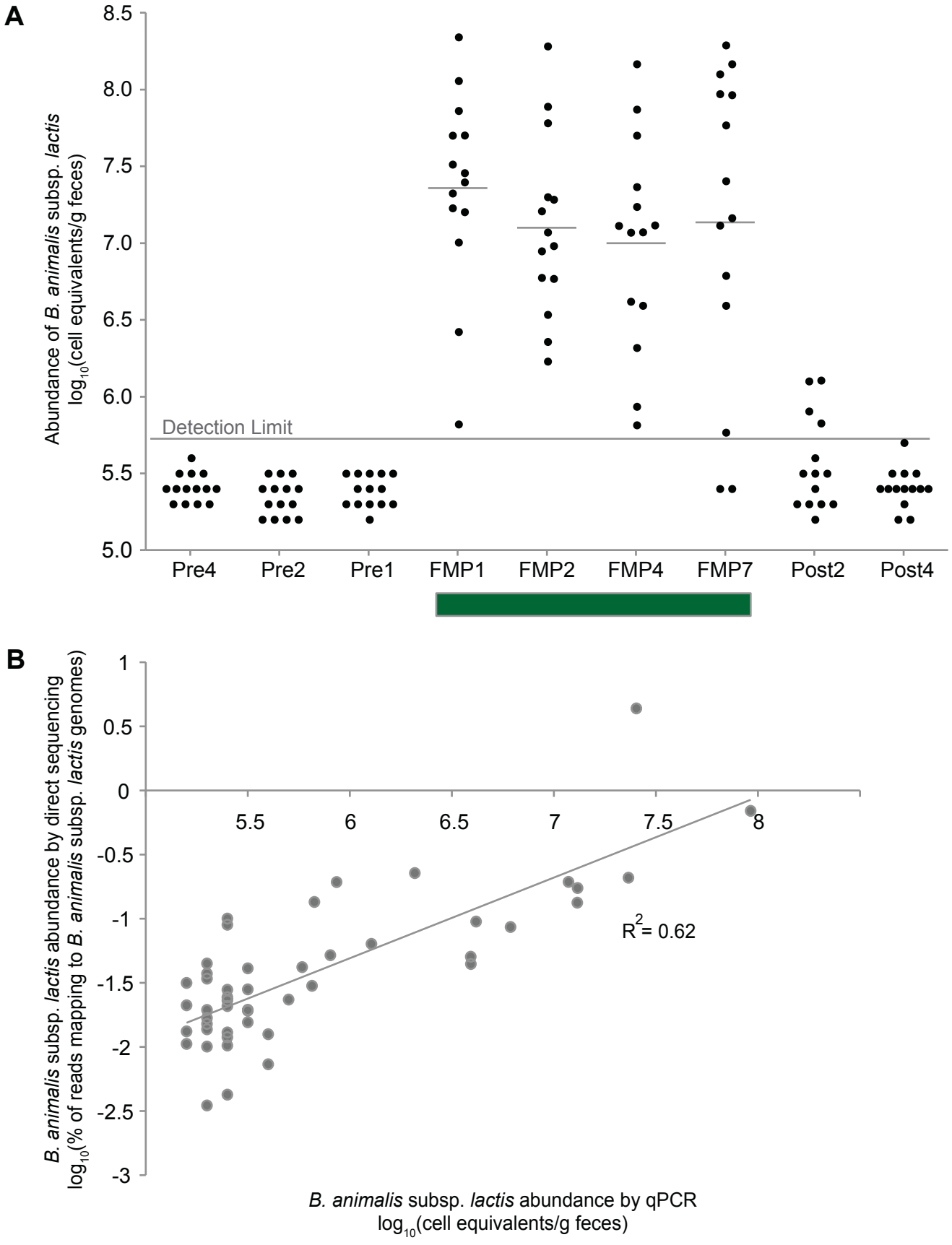

Figure S1. Levels of *B. animalis* subsp. *lactis* (CNCM I-2494) in human fecal samples collected prior to, during, and after consumption of a FMP. (**A**) qPCR assays; each dot represents a sample from a given individual. The green bar denotes the period of FMP consumption. (**B**) Comparison of qPCR results to the number of shotgun reads mapped to the genomes of three *B. animalis* subsp. *lactis* strains. qPCR results are plotted on the X-axis, while the proportional representation of reads that mapped to the *B. animalis* subsp. *lactis* genomes is presented on the Y-axis.

McNulty, *et al.*, Fig. S2

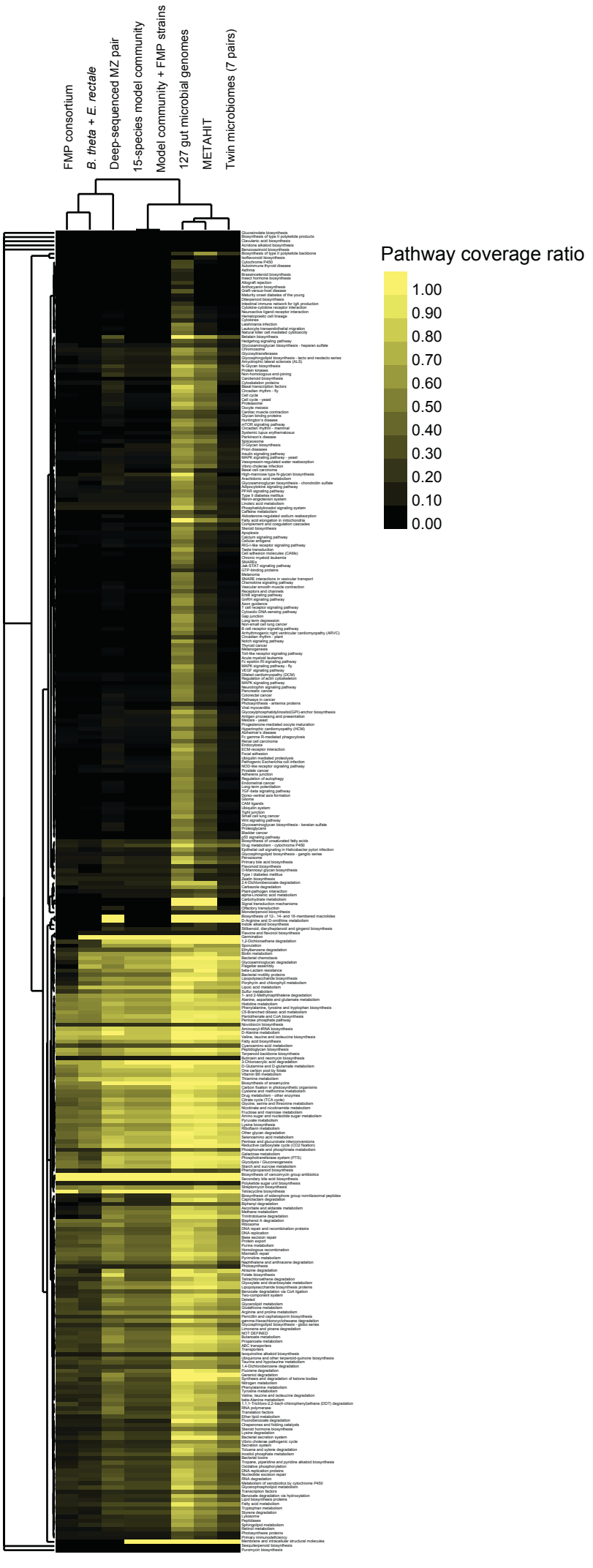

Figure S2. KEGG pathway coverage ratios suggest that the model human gut microbiome encodes many of the functions present in more complex human fecal communities. Genes in (i) the genomes of the five-member FMP strain consortium, (ii) the 15-member model human gut microbiota, (iii) a highly simplified two-member human gut microbiota composed of *B. thetaiotaomicron* and a prominent gut Firmicute (*Eubacterium rectale*) (*S35*), (iv) a reference set of 127 sequenced human gut microbial isolates, (v) the deeply sampled fecal microbiomes of 124 unrelated adult Europeans from a recent METAHIT study (*S18*), (vi) the fecal microbiomes of the 7 twin pairs characterized in the present study, and (vii) the deeply sequenced fecal microbiomes of an obese adult MZ twin pair (*S19*) were annotated using v54 of the KEGG GENES database. The presence/absence of each KO in each KEGG pathway was determined for every set of genes and the pathway coverage ratio (i.e., % of a pathway's components called 'present'; BLASTP *E*-value $<10^{-5}$) was calculated and depicted as the heatmap shown using Cluster 3.0/Treeview.

McNulty, *et al.*, Fig. S3

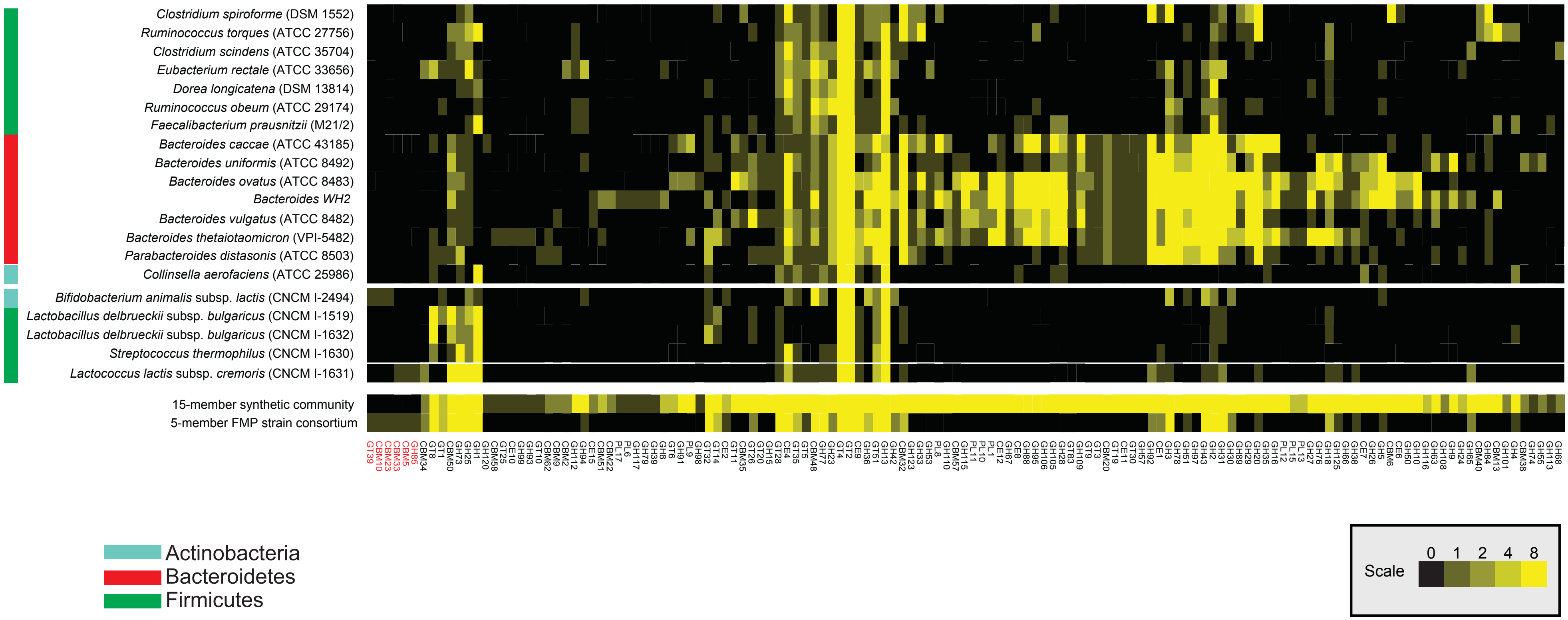

Figure S3. CAZyme profiles of the 20 bacterial strains introduced into gnotobiotic mice. The indicated genomes were annotated for all glycoside hydrolases (GH), glycosyltransferases (GT), carbohydrate binding modules (CBM), and polysaccharide lyases (PL) using the CAZy classification scheme. The Bacteroides possess a larger and more diverse arsenal of CAZymes relative to the Firmicutes/Actinobacteria, though most CAZyme families encoded in the genomes of the FMP strains were also present in defined community members. The small number of FMP strain-specific CAZyme families (CBM5, CBM10, CBM23, CBM33, GH85, and GT39) are highlighted in red. The scale refers to the number of genes in a given CAZy family in a given genome.

McNulty, *et al.*, Fig. S4

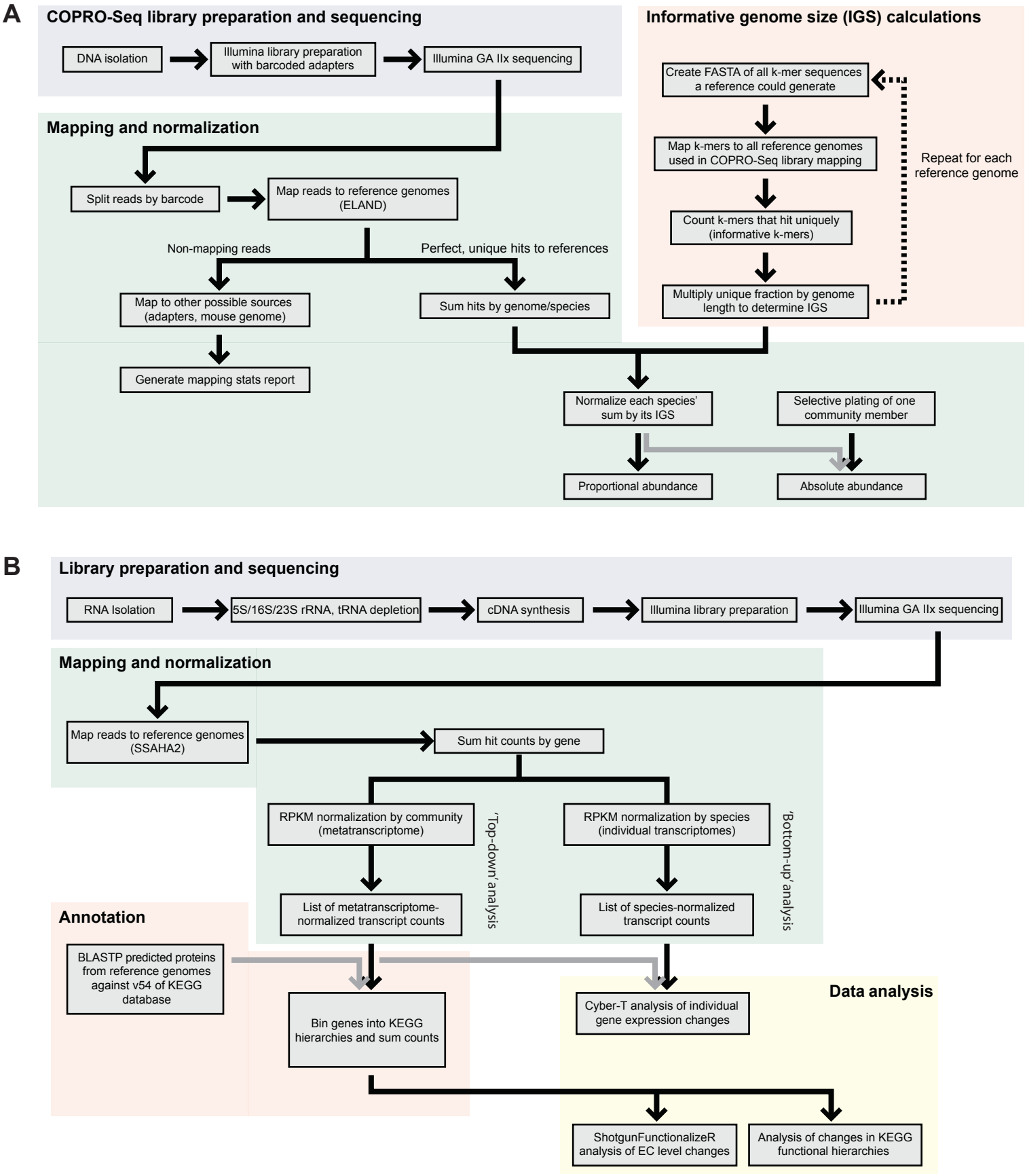

Figure S4. Summary of analysis pipelines utilized in this study. (A) COPRO-Seq. **(B)** RNA-Seq.

McNulty, *et al.*, Fig. S5

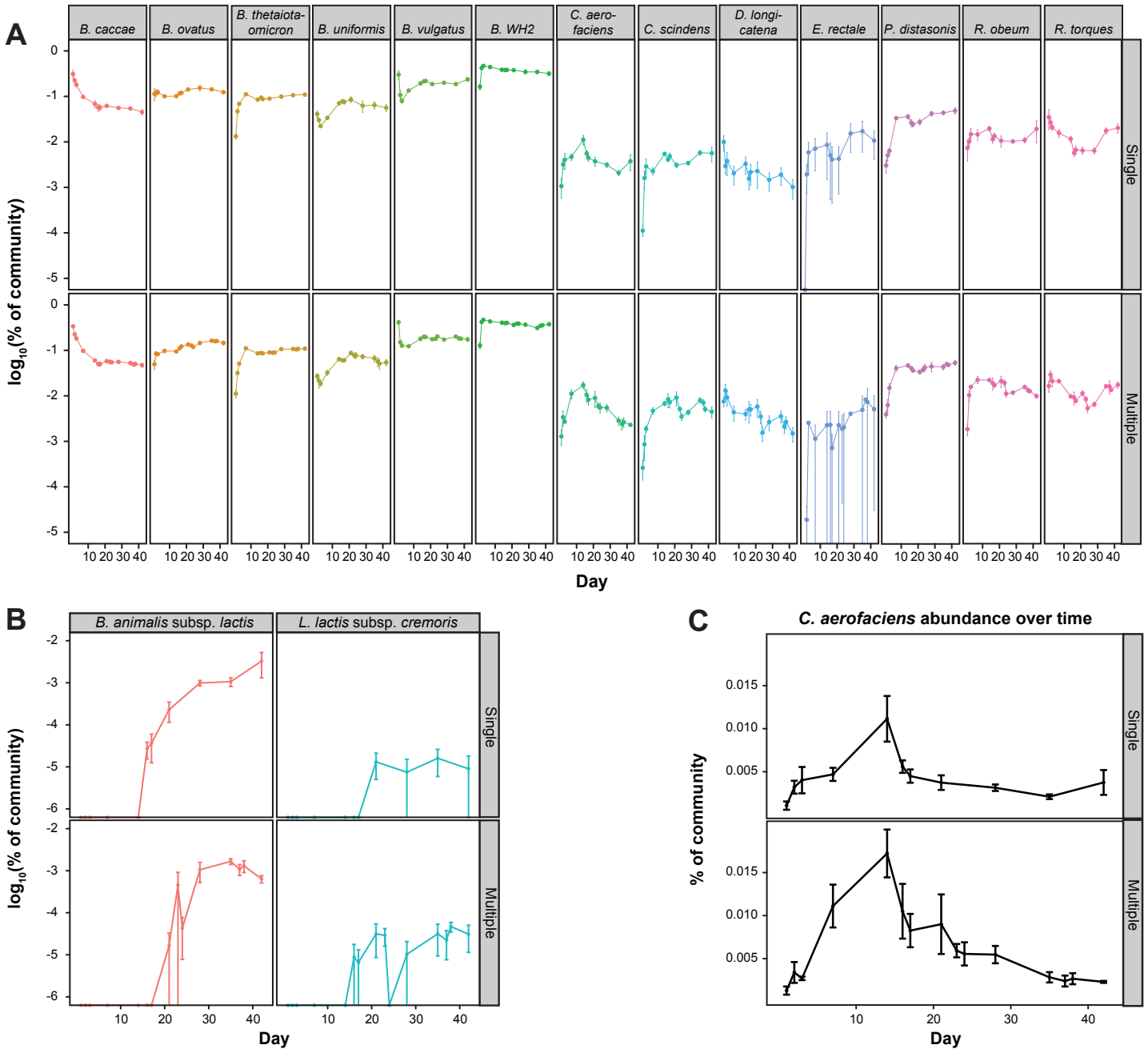

Figure S5. COPRO-Seq-based time series analysis of the abundance of members of the model human microbiota and of the FMP strain consortium in the feces of gnotobiotic mice. Relative abundance, expressed as the $\log_{10}$ of percent representation of all detected community members was calculated over time (d0, time of colonization with the model 15-member community; d14, time of first gavage with the FMP consortium for the single and multiple treatment groups; d21 and d35, times of subsequent gavage with the FMP consortium for the multiple treatment group). For each treatment, animals were gavaged twice over a 24 h period. Mean values ± SEM are plotted (n=5 animals/treatment group; 1 fecal sample/animal/time point; limit of detection = 0.003%). (**A**) COPRO-Seq data for 13 members of the 15 member community (*F. prausnitzii* and *C. spiroforme* were at levels below the limit of detection throughout the study and are not shown). (**B**) Data obtained from the two members of the FMP consortium that persisted at levels above the limit of detection following their introduction into mice. (**C**) Data from panel A representing the response of *C aerofaciens* to introduction of the FMP strain consortium (see text for details).

McNulty, *et al.*, Fig. S6

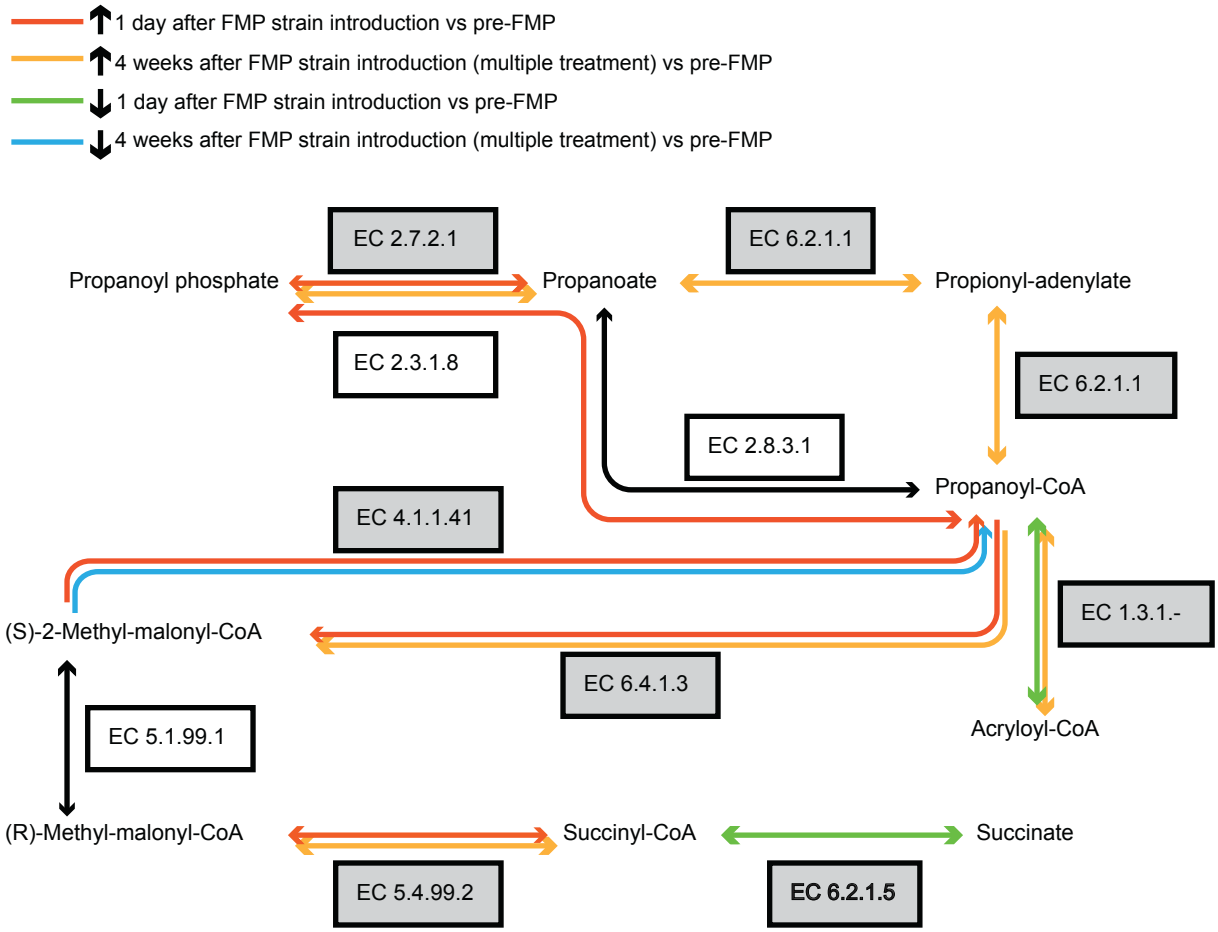

Figure S6. Top-down analysis of the model community's transcriptional response to the FMP strain consortium reveals upregulation of genes involved in interconversion of propionate and succinate. RPKM-normalized RNA-Seq data were binned at the level of EC, and comparisons were made between early responses (d14 versus d15, representing time points just before and 1 day after gavage with the strain consortium) and late responses (d14 versus d42). Boxes and lines are colored according to the key shown above the pathway map and in the legend to panel B of **Fig. 5**.

McNulty, *et al.*, Fig. S7

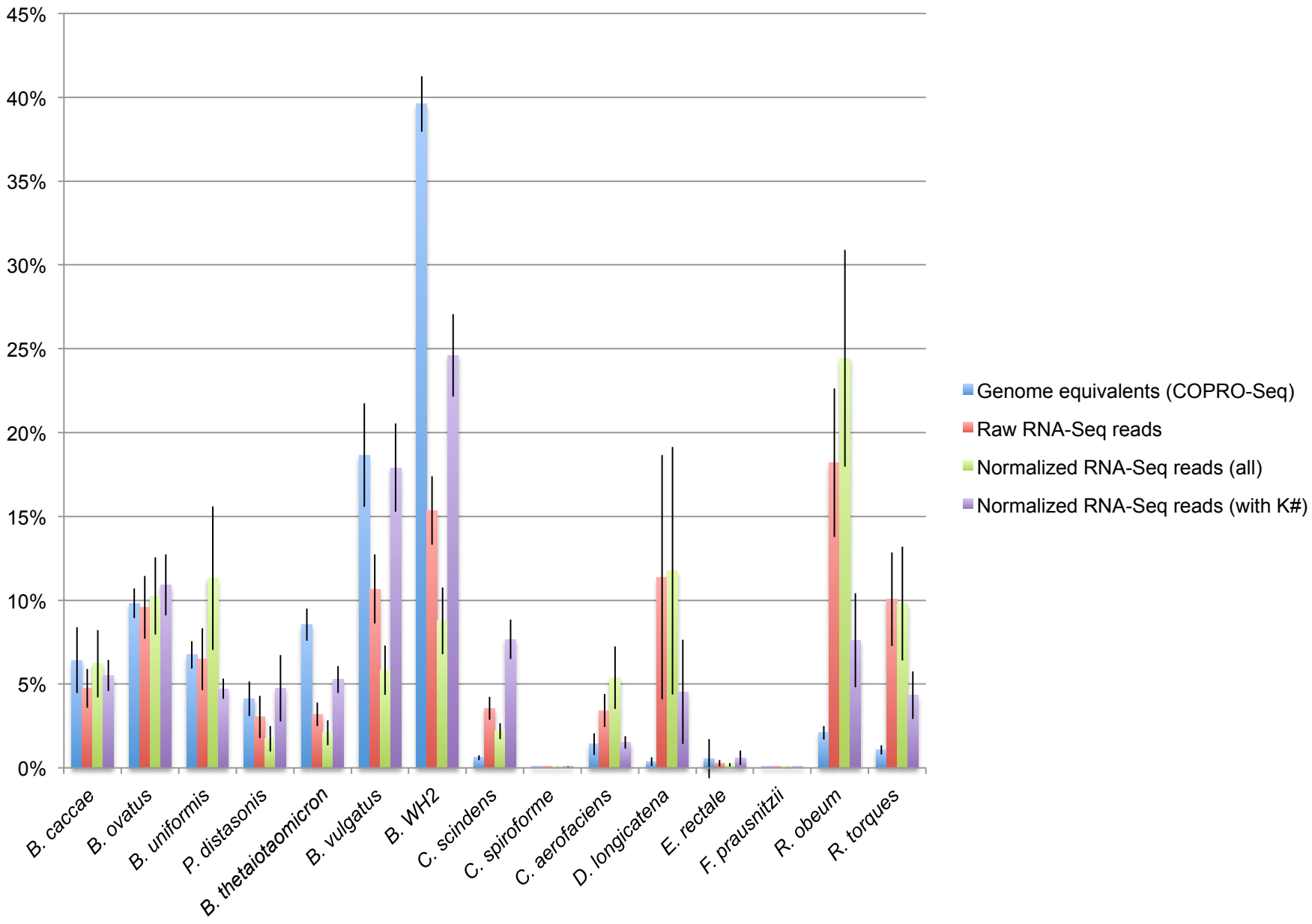

Figure S7. A species' contribution to the metatranscriptome is not necessarily proportional to its abundance in the 15-member community. Microbial RNA-Seq data from d14 of the mouse study were parsed by species to determine the total number of reads that each community member contributed to the total sequenced transcript pool ('metatranscriptome') (both raw and normalized reads as defined in **Fig. S4B**). Data were further broken down into reads that could be mapped to genes with known functions (as defined by KEGG) and those with unknown functions (lacking any KO in v54 of the KEGG GENES database). Mean values ± SD are plotted for each of the four types of data presented. Note, for example, the large number of raw reads attributed to *R. obeum* despite its low proportional abundance in the community. Conversely, *Bacteroides WH2* contributes a far smaller proportion of total raw RNA-Seq reads to the pool than its relative abundance in the community might have suggested.

McNulty, *et al.*, Fig. S8

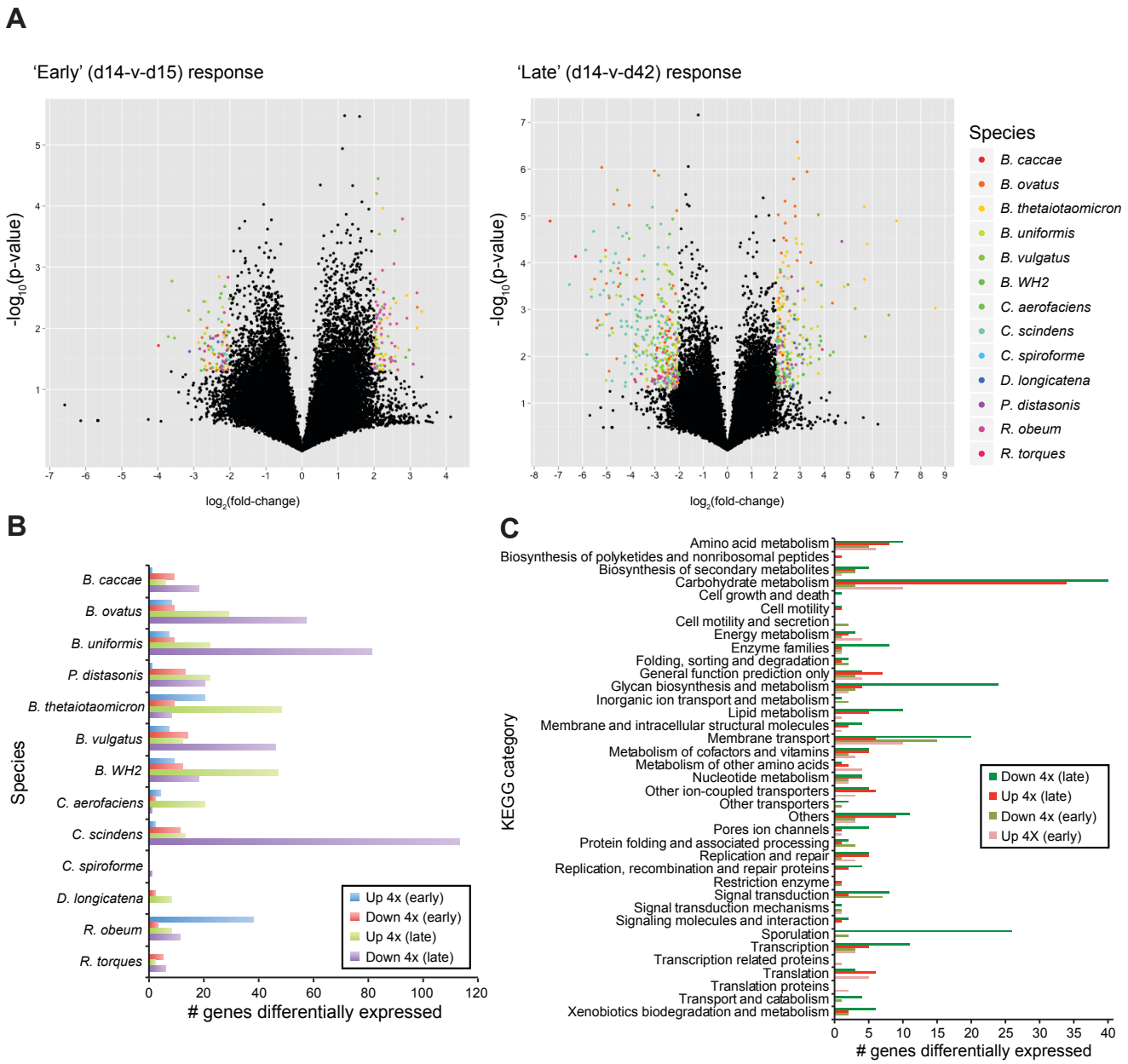

Figure S8. Bottom-up analysis of genes whose expression changes significantly following introduction of the FMP strain consortium. (**A**) Volcano plots of the >48,000 expressed genes detected in at least one fecal RNA sample. Colored points represent genes whose difference in expression following introduction of the FMP strain consortium was (i) ≥4-fold (increase or decrease) relative to the d14 pre-treatment time point and (ii) statistically significant ($p<0.05$; two-tailed Student's *t*-test). Dots are colored according to each gene's species of origin (color key shown to the right of the panel). Black dots represent genes whose change in expression is less than 4-fold at the time points indicated and/or not statistically significant. See **Table S10** for a complete list of all the genes shown as colored dots. (**B**) Breakdown by species of differentially expressed genes. (**C**) Breakdown by KEGG category of differentially expressed genes, showing that late responses are more numerous than immediate ones, and that there is a noticeable bias towards genes involved in carbohydrate and glycan metabolism, particularly in the late response to introduction of the FMP strain consortium.

McNulty, *et al.*, Fig. S9

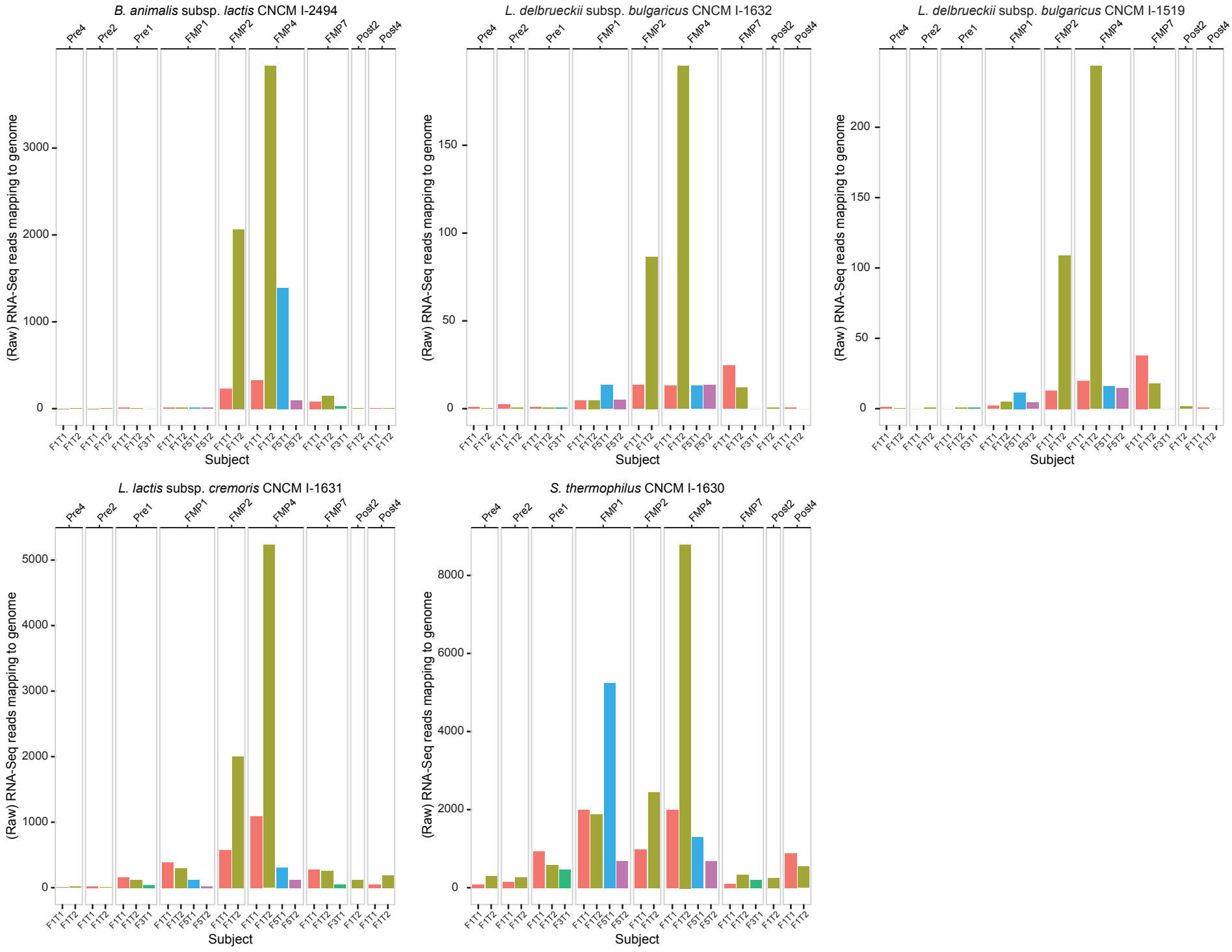

Figure S9. The number of RNA-Seq reads obtained from human fecal samples that map to genomes in the FMP strain consortium peaks shortly after FMP consumption begins. RNA-Seq reads derived from selected human samples were mapped back to the five genomes in the FMP consortium to determine which species' transcripts could be detected over time. *B. animalis* transcripts were detected only during periods of FMP consumption. Reads attributed to *L. lactis*, *S. thermophilus*, and *L. delbrueckii* at time points before consumption began may reflect 'spurious' mapping to related endogenous strains in the gut community. Facet labels located at the top of each bar chart correspond to the sample labels shown in **Fig. 1A** of the main text. Numbers are presented below each bar chart correspond to the code number assigned to each de-identified co-twin (see **Table S1**).

SUPPLEMENTARY TABLE LEGENDS

Table S1. Characteristics of adult female monozygotic (MZ) twins enrolled in study.

Table S2. Summary of human fecal metagenomic datasets. (**A**) Multiplex pyrosequencing of fecal bacterial 16S rRNA V2 amplicons. (**B**) Multiplex shotgun pyrosequencing of total fecal community DNA.

Table S3. Features of the microbial genomes in the 5-member FMP strain consortium and the 15-member model human gut microbiota.

Table S4. Carbohydrate active enzyme (CAZy) annotation data. (A) CAZy summaries by genome. (**B**) CAZy annotations for the 20 bacterial species in this study.

Table S5. COPRO-Seq analysis of bacterial species abundance in mouse fecal samples. (A) Proportional representation of the 20 bacterial species in this study in mouse fecal samples as measured by COPRO-Seq. (**B**) Statistical significance and fold-change of differences in pairwise comparisons of abundance calculated from data in panel A. The group mean for each day/treatment/species combination at time points after the introduction of the FMP strain consortium was compared to the mean for the same treatment/species at d14 (the last time point collected prior to introduction of the FMP strains) using Welch's *t*-test. Values have not been corrected for multiple hypothesis testing. *p*-values <0.05 are highlighted in pink. Fold-changes greater than 2 or less than -2 are highlighted in pink and green, respectively.

Table S6. INSeq analysis. (A) INSeq analysis sequencing statistics. Scale factor corresponds to counts per million normalization; underrepresented samples were re-sequenced and combined

with original data so that all samples were represented by ~1 million reads. (**B**) Genes required by *B. thetaiotaomicron* for survival in the intestines of mice harboring the 15-member model human gut microbiota. The table describes the relative abundance of transposon insertions in each gene (rows) in the input community (average of two independent technical replicates) and in the output communities (fecal samples collected from 10 mice two weeks after introduction of the synthetic model community, immediately prior to introduction of the FMP strain consortium). A *z*-test was used to identify genes whose log-transformed output to input ratios were significantly different from the overall distribution (a uniform value of 1 was added to all counts, and genes with no insertions were removed to allow ratios to be calculated). Resultant *p*-values were corrected for multiple hypothesis testing by *q*-test (*S36*). Genes assigned a *q*-value <0.001 are highlighted in red. Data filtering, normalization, mapping, and statistical analysis were conducted in Perl and Matlab.

Table S7. Differentially expressed *B. animalis* subsp. *lactis* (CNCM I-2494) genes. (A) Log versus stationary growth in MRS medium. (**B**) *In vivo* (mouse) versus *in vitro* (logarithmic phase in MRS) growth.

Table S8. Top-down function-level analysis of the impact of the FMP strain consortium on the model human gut microbiota's metatranscriptome. (**A**) Proportional representation of assignable normalized RNA-Seq counts binned by KEGG category in fecal samples collected over time from singly- and multiply-treated animals. (**B**) Proportional representation of assignable normalized RNA-Seq counts binned by KEGG pathway in fecal samples collected over time from singly- and multiply-treated animals.

Table S9. Model human gut microbiota membrane transport genes demonstrating ≥4-fold increases or decreases in their expression following introduction of the FMP strain consortium.

Table S10. Bottom-up (gene-level) analysis of the impact of the FMP strain consortium on the model community's metatranscriptome. (**A**) Breakdown by microbial species of significantly up/down-regulated genes. (**B**) Breakdown by KEGG category of significantly up/down-regulated genes. (**C**) Model community microbiome genes demonstrating significant increases/decreases in their expression following introduction of the FMP strain consortium.

Table S11. Results of Random Forests supervised classification analysis.

Table S12. Urine metabolites whose levels change significantly in transitions between colonization states. The 'Reverse match score' column contains the AMDIS, dot-product, reverse-match scores (*S32*) which in this case evaluate not only mass-spectral concordance, but also the goodness of fit of chromatographic retention-time index made by comparison to (i) commercial and public target-compound libraries of small metabolites (*S28, 33*), and (ii) our own in-house reference library. Metabolites with match scores less than an arbitrary threshold of 65% were excluded from these results. (**A**) Day 0 (germ-free) versus day 14 (colonized with 15-member model community). (**B**) Day 14 (colonized with 15-member model community) versus day 42 (colonized with 15-member model community plus five-member FMP consortium).

Table S13. ShotgunFunctionalizeR analysis of EC-level changes in the metatranscriptome as a function of FMP strain introduction into mice and humans. The table shows fold-change in mean proportional representation of each EC between groups for each comparison. Values for

nonsignificant EC changes in a comparison (adjusted $p \geq 0.01$, ShotgunFunctionalizeR) are reported as "NS". Note that ECs can have multiple KEGG pathway and category assignments.

Table S14. Primers and amplification conditions used for quantitative PCR assays of FMP consortium strains in fecal DNA. (**A**) Primers used to assay human fecal samples. (**B**) Primers used to assay mouse fecal samples. (**C**) Amplification conditions.

Table S15. List of 136 microbial genomes used to analyze human fecal RNA-Seq data.

SUPPLEMENTARY REFERENCES

S1. M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes, *Nucleic Acids Res* **28**, 27 (2000).

S2. M. Kanehisa, S. Goto, M. Hattori, K. F. Aoki-Kinoshita, M. Itoh, S. Kawashima, T. Katayama, M. Araki, M. Hirakawa, From genomics to chemical genomics: new developments in KEGG, *Nucleic Acids Res* **34**, D354 (2006).

S3. M. Kanehisa, S. Goto, M. Furumichi, M. Tanabe, M. Hirakawa, KEGG for representation and analysis of molecular networks involving diseases and drugs, *Nucleic Acids Res* **38**, D355 (2010).

S4. B. L. Cantarel, P. M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics, *Nucleic Acids Res* **37**, D233 (2009).

S5. M. J. de Hoon, S. Imoto, J. Nolan, S. Miyano, Open source clustering software, *Bioinformatics* **20**, 1453 (2004).

S6. F. E. Rey, J. J. Faith, J. Bain, M. J. Muehlbauer, R. D. Stevens, C. B. Newgard, J. I. Gordon, Dissecting the in vivo metabolic potential of two human gut acetogens, *J Biol Chem* **285**, 22082 (2010).

S7. S. J. Lewis, K. W. Heaton, Stool form scale as a useful guide to intestinal transit time, *Scand J Gastroenterol* **32**, 920 (1997).

S8. O. Firmesse, S. Rabot, L. G. Bermudez-Humaran, G. Corthier, J. P. Furet, Consumption of Camembert cheese stimulates commensal enterococci in healthy human intestinal microbiota, *FEMS Microbiol Lett* **276**, 189 (2007).

S9. J. P. Furet, P. Quenee, P. Tailliez, Molecular quantification of lactic acid bacteria in fermented milk products using real-time quantitative PCR, *Int J Food Microbiol* **97**, 197 (2004).

S10. T. Matsuki, K. Watanabe, J. Fujimoto, T. Takada, R. Tanaka, Use of 16S rRNA gene-targeted group-specific primers for real-time PCR analysis of predominant bacteria in human feces, *Appl Environ Microbiol* **70**, 7220 (2004).

S11. M. A. Nadkarni, F. E. Martin, N. A. Jacques, N. Hunter, Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set, *Microbiology* **148**, 257 (2002).

S12. H. Sokol, B. Pigneur, L. Watterlot, O. Lakhdari, L. G. Bermudez-Humaran, J. J. Gratadoux, S. Blugeon, C. Bridonneau, J. P. Furet, G. Corthier, C. Grangette, N. Vasquez, P. Pochart, G. Trugnan, G. Thomas, H. M. Blottiere, J. Dore, P. Marteau, P. Seksik, P. Langella, 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).

S13. P. J. Turnbaugh, M. Hamady, T. Yatsunenko, B. L. Cantarel, A. Duncan, R. E. Ley, M. L. Sogin, W. J. Jones, B. A. Roe, J. P. Affourtit, M. Egholm, B. Henrissat, A. C. Heath, R. Knight, J. I. Gordon, A core gut microbiome in obese and lean twins, *Nature* **457**, 480 (2009).

S14. J. G. Caporaso, J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Pena, J. K. Goodrich, J. I. Gordon, G. A. Huttley, S. T. Kelley, D. Knights, J. E. Koenig, R. E. Ley, C. A. Lozupone, D. McDonald, B. D. Muegge, M. Pirrung, J. Reeder, J. R. Sevinsky, P. J. Turnbaugh, W. A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld, R. Knight, QIIME allows analysis of high-throughput community sequencing data, *Nat Methods* **7**, 335 (2010).

S15. W. Li, A. Godzik, Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences, *Bioinformatics* **22**, 1658 (2006).

S16. Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing, *J Roy Stat Soc B Met* **57**, 289 (1995).

S17. V. Gomez-Alvarez, T. K. Teal, T. M. Schmidt, Systematic artifacts in metagenomes from complex microbial communities, *ISME J* **3**, 1314 (2009).

S18. J. Qin, R. Li, J. Raes, M. Arumugam, K. S. Burgdorf, C. Manichanh, T. Nielsen, N. Pons, F. Levenez, T. Yamada, D. R. Mende, J. Li, J. Xu, S. Li, D. Li, J. Cao, B. Wang, H. Liang, H. Zheng, Y. Xie, J. Tap, P. Lepage, M. Bertalan, J. M. Batto, T. Hansen, D. Le Paslier, A. Linneberg, H. B. Nielsen, E. Pelletier, P. Renault, T. Sicheritz-Ponten, K. Turner, H. Zhu, C. Yu, M. Jian, Y. Zhou, Y. Li, X. Zhang, N. Qin, H. Yang, J. Wang, S. Brunak, J. Dore, F. Guarner, K. Kristiansen, O. Pedersen, J. Parkhill, J. Weissenbach, P. Bork, S. D. Ehrlich, A human gut microbial gene catalogue established by metagenomic sequencing, *Nature* **464**, 59 (2010).

S19. P. J. Turnbaugh, C. Quince, J. J. Faith, A. C. McHardy, T. Yatsunenko, F. Niazi, J. Affourtit, M. Egholm, B. Henrissat, R. Knight, J. I. Gordon, Organismal, genetic, and transcriptional variation in the deeply sequenced gut microbiomes of identical twins, *Proc Natl Acad Sci U S A* **107**, 7503 (2010).

S20. A. L. Goodman, N. P. McNulty, Y. Zhao, D. Leip, R. D. Mitra, C. A. Lozupone, R. Knight, J. I. Gordon, Identifying genetic determinants needed to establish a human gut symbiont in its habitat, *Cell Host Microbe* **6**, 279 (2009).

S21. Z. Ning, A. J. Cox, J. C. Mullikin, SSAHA: a fast search method for large DNA databases, *Genome Res* **11**, 1725 (2001).

S22. A. D. Long, H. J. Mangalam, B. Y. Chan, L. Tolleri, G. W. Hatfield, P. Baldi, Improved statistical inference from DNA microarray data using analysis of variance and a Bayesian statistical framework. Analysis of global gene expression in Escherichia coli K12, *J Biol Chem* **276**, 19937 (2001).

S23. E. Kristiansson, P. Hugenholtz, D. Dalevi, ShotgunFunctionalizeR: an R-package for functional comparison of metagenomes, *Bioinformatics* **25**, 2737 (2009).

S24. A. Liaw, M. Wiener, Classification and Regression by randomForest, *R News* **2**, 18 (2002).

S25. D. Knights, E. K. Costello, R. Knight, Supervised classification of human microbiota, *FEMS Microbiol Rev* **35**, 343 (2011).

S26. U. Roessner, C. Wagner, J. Kopka, R. N. Trethewey, L. Willmitzer, Technical advance: simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry, *Plant J* **23**, 131 (2000).

S27. O. Fiehn, G. Wohlgemuth, M. Scholz, T. Kind, Y. Lee do, Y. Lu, S. Moon, B. Nikolau, Quality control for plant metabolomics: reporting MSI-compliant studies, *Plant J* **53**, 691 (2008).

S28. T. Kind, G. Wohlgemuth, Y. Lee do, Y. Lu, M. Palazoglu, S. Shahbaz, O. Fiehn, FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry, *Anal Chem* **81**, 10038 (2009).

S29. J. Chen, C. K. Meng, S. B. Narayan, W. Luan, M. J. Bennett, The use of Deconvolution Reporting Software and backflush improves the speed and accuracy of data processing for urinary organic acid analysis, *Clin Chim Acta* **405**, 53 (2009).

S30. W. G. Mallard, J. Reed, U. D. o. C. National Institute of Standards and Technology, Ed. (Gaithersburg, MD, 1997), pp. 58.

S31. J. M. Halket, A. Przyborowska, S. E. Stein, W. G. Mallard, S. Down, R. A. Chalmers, Deconvolution gas chromatography/mass spectrometry of urinary organic acids--potential for pattern recognition and automated identification of metabolic disorders, *Rapid Commun Mass Spectrom* **13**, 279 (1999).

S32. S. E. Stein, An Integrated Method for Spectrum Extraction and Compound Identification from Gas Chromatography/Mass Spectrometry Data, *J Am Soc Mass Spectrom* **10**, 770 (1999).

S33. J. Kopka, N. Schauer, S. Krueger, C. Birkemeyer, B. Usadel, E. Bergmuller, P. Dormann, W. Weckwerth, Y. Gibon, M. Stitt, L. Willmitzer, A. R. Fernie, D. Steinhauser, GMD@CSB.DB: the Golm Metabolome Database, *Bioinformatics* **21**, 1635 (2005).

S34. M. P. Styczynski, J. F. Moxley, L. V. Tong, J. L. Walther, K. L. Jensen, G. N. Stephanopoulos, Systematic identification of conserved metabolites in GC/MS data for metabolomics and biomarker discovery, *Anal Chem* **79**, 966 (2007).

S35. M. A. Mahowald, F. E. Rey, H. Seedorf, P. J. Turnbaugh, R. S. Fulton, A. Wollam, N. Shah, C. Wang, V. Magrini, R. K. Wilson, B. L. Cantarel, P. M. Coutinho, B. Henrissat, L. W. Crock, A. Russell, N. C. Verberkmoes, R. L. Hettich, J. I. Gordon, Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla, *Proc Natl Acad Sci U S A* **106**, 5859 (2009).

S36. J. D. Storey, R. Tibshirani, Statistical significance for genomewide studies, *Proc Natl Acad Sci U S A* **100**, 9440 (2003).