SUPPLEMENTAL MATERIAL

SUPPLEMENTAL METHODS

Treatment of ASF mice by emulsifiers

Six weeks old gnotobiotic ASF C57BL/6 mice were exposed to sodium carboxymethylcellulose (CMC, average M_W ~250,000) (Sigma-Aldrich, St. Louis, MO, USA) or polysorbate-80 (P80) (Sigma- Aldrich) diluted in drinking water (1.0% or indicated concentration) (not blinded). The same water (reverse-osmosis treated Atlanta city water) was used for the water-treated (control) group. These solutions were autoclaved and changed every week. Fresh feces were collected every other week for downstream analysis. After 11 weeks of emulsifier treatment, mice were fasted for 5-h at which time blood was collected by retrobulbar intraorbital capillary plexus. Hemolysis-free serum was generated by centrifugation of blood using serum separator tubes (Becton Dickinson, Franklin Lakes, NJ, USA). Mice were then euthanized, and colon length, colon weight, spleen weight and adipose weight were measured. Organs were collected for downstream analysis.

Quantification of fecal lipocalin-2 (Lcn-2) by ELISA

For quantification of fecal Lcn-2 by ELISA, frozen fecal samples were reconstituted in PBS containing 0.1% Tween 20 to a final concentration of 100 mg/mL and vortexed for 20 min to get a homogenous fecal suspension [1]. These samples were then centrifuged for 10 min at 14 000 g and 4°C. Clear supernatants were collected and stored at −20°C until analysis. Lcn-2 levels were estimated in the supernatants using Duoset murine Lcn-2 ELISA kit (R&D Systems, Minneapolis, MN, USA) using the colorimetric peroxidase substrate tetramethylbenzidine, and

optical density (OD) was read at 450 nm (Versamax microplate reader).

Fecal flagellin and lipopolysaccharide load quantification

Levels of fecal bioactive flagellin and lipopolysaccharide (LPS) were quantified as previously described [2] using human embryonic kidney (HEK)-Blue-mTLR5 and HEK-BluemTLR4 cells, respectively (Invivogen, San Diego, CA, USA) [2]. Fecal material was resuspended in PBS to a final concentration of 100 mg/mL and homogenized for 10 s using a Mini-Beadbeater-24 without the addition of beads to avoid bacteria disruption. Samples were then centrifuged at 8000 g for 2 min and the resulting supernatant was serially diluted and applied on mammalian cells. Purified *E. coli* flagellin and LPS (Sigma-Aldrich) were used for standard curve determination using HEK-Blue-mTLR5 and HEK-Blue-mTLR4 cells, respectively. After 24 h of stimulation, the cell culture supernatant was applied to QUANTI-Blue medium (Invivogen) and the alkaline phosphatase activity was measured at 620 nm after 30 min.

Bacterial quantification by q-PCR

For quantification of total fecal bacterial load and ASF composition analysis, total bacterial DNA was isolated from weighted feces using QIAamp DNA Stool Mini Kit (Qiagen, Hilden, Germany). DNA was then subjected to quantitative PCR using QuantiFast SYBR Green PCR kit (Biorad, Hercules, CA, USA) with universal 16S rRNA primers or with primers specific for each member of the ASF defined microbiota (**Table 1**) [3]. Results are expressed as bacteria number per mg of stool, using a standard curve, or as relative values. Principal coordinate analysis was performed using euclidean distance with the 8 ASF members.

Short-chain fatty acids (SCFA) quantification

The SCFA in the lumen samples of the M-SHIME were extracted with diethyl ether and analyzed using a gas chromatograph as previously described [4].

Microbiota analysis by 16S rRNA gene sequencing using Illumina technology

16S rRNA gene amplification and sequencing were done using the Illumina MiSeq technology following the protocol of Earth Microbiome Project with their modifications to the MOBIO PowerSoil DNA Isolation Kit procedure for extracting DNA (www.earthmicrobiome.org/emp-standard-protocols) [5, 6]. Bulk DNA was extracted from frozen M-SHIME suspension or feces using a PowerSoil-htp kit from MoBio Laboratories (Carlsbad, CA, USA) with mechanical disruption (bead-beating). The 16S rRNA genes, region V4, were PCR amplified from each sample using a composite forward primer and a reverse primer containing a unique 12-base barcode, designed using the Golay error-correcting scheme, which was used to tag PCR products from respective samples [6]. We used the forward primer 515F 5'-*AATGATACGGCGACCACCGAGATCTACAC***TATGGTAATT*GT***GTGCCAGCMGCCGCGGTAA-3': the italicized sequence is the 5' Illumina adapter B, the bold sequence is the primer pad, the italicized and bold sequence is the primer linker and the underlined sequence is the conserved bacterial primer 515F. The reverse primer 806R used was 5'-*CAAGCAGAAGACGGCATACGAGAT* XXXXXXXXXXXX **AGTCAGTCAG *CC*** *GGACTACHVGGGTWTCTAAT*-3': the italicized sequence is the 3' reverse complement sequence of Illumina adapter, the 12 X sequence is the golay barcode, the bold sequence is the primer pad, the italicized and bold sequence is the primer linker and the underlined sequence is the conserved bacterial primer 806R. PCR reactions consisted of Hot Master PCR mix

(Quantabio, Beverly, MA, USA), 0.2 μM of each primer, 10-100 ng template, and reaction conditions were 3 min at 95°C, followed by 30 cycles of 45 s at 95°C, 60s at 50°C and 90 s at 72°C on a Biorad thermocycler. PCRs products were purified with Ampure magnetic purification beads (Agencourt, Brea, CA, USA), and visualized by gel electrophoresis. Products were then quantified (BIOTEK Fluorescence Spectrophotometer) using Quant-iT PicoGreen dsDNA assay. A master DNA pool was generated from the purified products in equimolar ratios. The pooled products were quantified using Quant-iT PicoGreen dsDNA assay and then sequenced using an Illumina MiSeq sequencer (paired-end reads, 2 x 250 bp) at Cornell University, Ithaca.

16S rRNA gene sequence analysis

Forward and reverse Illumina reads were joined using the fastq-join method [7, 8], sequences were demultiplexed, quality filtered using Quantitative Insights Into Microbial Ecology (QIIME, version 1.8.0) software package [9]. QIIME default parameters were used for quality filtering (reads truncated at first low-quality base and excluded if: (1) there were more than three consecutive low quality base calls (2), less than 75% of read length was consecutive high quality base calls (3), at least one uncalled base was present (4), more than 1.5 errors were present in the bar code (5), any Phred qualities were below 20, or (6) the length was less than 75 bases). Sequences were assigned to operational taxonomic units (OTUs) using UCLUST algorithm [10] with a 97% threshold of pairwise identity (without the creation of new clusters with sequences that do not match the reference sequences), and classified taxonomically using the Greengenes reference database 13_8 [11]. A single representative sequence for each OTU was aligned and a phylogenetic tree was built using FastTree [12]. The phylogenetic tree was used for computing the unweighted UniFrac distances between samples [13, 14], rarefaction were performed and used to

compare abundances of OTUs across samples. Principal coordinates analysis (PCoA) plots were used to assess the variation between experimental group (beta diversity). Alpha diversity curves were determined for all samples using the determination of the number of observed species. LEfSE (LDA Effect Size) was used to investigate bacterial members that drive differences between groups [15]. Unprocessed sequencing data are deposited in the European Nucleotide Archive under accession numbers PRJEB19279 (M-SHIME suspension) and PRJEB19272 (transplant of M-SHIME suspension to germfree recipient mice).

16S-based metagenome prediction

PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States) was used to predict the metagenome based on microbiota composition analysis [16]. A closed-OTU picking strategy was using against the Greengenes reference database 13_8 [11], and metagenomes were predicted. Predicted metagenomes were categorized at level 2 and 3 of the Kyoto Encyclopedia of genes and genomes (KEGG) pathways. Gene-E was used for heatmap representation of pathways with an altered abundance following emulsifier consumption compared with water-treated control group.

Intraperitoneal injection of M-SHIME suspension to Rag$^{-/-}$ mice and quantification of sera IL-6 by ELISA

Luminal M-SHIME suspensions were used for intraperitoneal injection to Rag$^{-/-}$ mice (100 μL). Two hours post-injection, mice were bled retro-orbitaly, and IL-6 levels were estimated in the serum using Duoset murine ELISA kit (R&D Systems, Minneapolis, MN, USA), as previously reported [2].

Immunostaining of mucins and localization of bacteria by FISH

Mucus immunostaining was paired with fluorescent *in situ* hybridization (FISH), as previously described [17], in order to analyze bacteria localization at the surface of the intestinal mucosa [18, 19]. Briefly, colonic tissues (proximal colon, 2nd cm from the cecum) containing fecal material were placed in methanol-Carnoy's fixative solution (60% methanol, 30% chloroform, 10% glacial acetic acid) for a minimum of 3 h at room temperature. Tissues were then washed in methanol 2 x 30 min, ethanol 2 x 15 min, ethanol/xylene (1:1) 15 min and xylene 2 x 15 min, followed by embedding in Paraffin with a vertical orientation. Five µm sections were performed and dewax by preheating at 60°C for 10 min, followed by xylene 60°C for 10 min, xylene for 10 min and 99.5% ethanol for 10 minutes. Hybridization step was performed at 50°C overnight with EUB338 probe (5'-GCTGCCTCCCGTAGGAGT-3', with a 5' labeling using Alexa 647) diluted to a final concentration of 10 µg/mL in hybridization buffer (20 mM Tris–HCl, pH 7.4, 0.9 M NaCl, 0.1% SDS, 20% formamide). After washing 10 min in wash buffer (20 mM Tris–HCl, pH 7.4, 0.9 M NaCl) and 3 x 10 min in PBS, PAP pen (Sigma-Aldrich) was used to mark around the section and block solution (5% fetal bovine serum in PBS) was added for 30 min at 4°C. Mucin-2 primary antibody (rabbit H-300, Santa Cruz Biotechnology, Dallas, TX, USA) was diluted 1:1500 in block solution and apply overnight at 4°C. After washing 3 x 10 min in PBS, block solution containing anti-rabbit Alexa 488 secondary antibody diluted 1:1500, Phalloidin-Tetramethylrhodamine B isothiocyanate (Sigma-Aldrich) at 1µg/mL and Hoechst 33258 (Sigma-Aldrich) at 10µg/mL was applied to the section for 2h. After washing 3 x 10 min in PBS slides were mounted using Prolong anti-fade mounting media (Life Technologies, Carlsbad, CA, USA). Observations were performed with a Zeiss LSM 700 confocal microscope with software

Zen 2011 version 7.1. This software was used to determine the distance between bacteria and epithelial cell monolayer, as well as the mucus thickness.

SUPPLEMENTAL FIGURE LEGENDS

Figure S1: Microbiota stabilization in the mucosal-M-SHIME system. A 8-vessel M-SHIME system was set up and inoculated with fresh human feces at day -7. Luminal microbiota composition was analyzed at days -7, -6, -3, -2, 1, 4, 6, 8, 11 and 13 using Illumina sequencing of the V4 region of 16S rRNA genes. (**A**) Principal coordinates analysis (PCoA) of the weighted UniFrac distance matrix. (**B**) Microbiota richness (number of OTUs observed) at days -6, 1 and 11. Data are the means +/- S.E.M (*N*=2). Each treatment group was compared with the control group (day -6) and statistical significance was determined using two-way ANOVA with Bonferroni's multiple comparisons test.

Figure S2: Microbiota stabilization in the mucosal-M-SHIME system. A 8-vessel M-SHIME system was set up and inoculated with fresh human feces at day -7. Following a 7-day stabilization period, vessels were treated with either water (*N*=2), CMC 1% (w/v, *N*=3) or P80 1% (v/v, *N*=3) for 13 days. Luminal microbiota composition was analyzed using Illumina sequencing of the V4 region of 16S rRNA genes. (**A-C**) Microbiota richness (number of OTUs observed) at days -6 (**A**), 1 (**B**) and 11 (**C**). Data are the means +/- S.E.M (*N*=2-3). Each treatment group was compared with the control group (water-treated) and statistical significance was determined using two-way ANOVA with Bonferroni's multiple comparisons test.

Figure S3: P80 alters mucus associated mucosal-M-SHIME microbiota composition. A 8-vessel M-SHIME system was set up and inoculated with fresh human feces at day -7. Following a 7-day stabilization period, vessels were treated with either water (*N*=2), CMC 1% (w/v, *N*=3) or P80 1% (v/v, *N*=3) for 13 days. Mucus-associated microbiota composition was analyzed using

Illumina sequencing of the V4 region of 16S rRNA genes. Principal coordinates analysis (PCoA) of the weighted UniFrac distance matrix at days -6 and 11 is represented (*N*=2-3).

Figure S4: Figure 5: Both CMC and P80 alter microbial gene expression in the mucosal-M-SHIME model. A 8-vessel M-SHIME system was set up and inoculated with fresh human feces at day -7. Following a 7-day stabilization period, vessels were treated with either water (*N*=2), CMC 1% (w/v, *N*=3) or P80 1% (v/v, *N*=3) for 13 days. M-SHIME-associated luminal metatranscriptome was analyzed by Illumina sequencing. Following functional classification of the reads to the subsystems orthology using MG-RAST, data were plotted as follow: *i*) the difference in abundance between the two compared groups is indicated in log2 fold change; *ii*) X-axis=relative expression in water-treated samples at day *x* versus water-treated samples at day *y*; *iii*) Y-axis=relative expression in emulsifier-treated samples at day *x* versus emulsifier-treated samples at day *y*. All dot above the diagonal (draw in **A**) represent gene with an increased expression in emulsifier -treated M-SHIME vessels versus water-treated M-SHIME system. All dot below the diagonal (draw in **A**) represent gene with a decreased expression in emulsifier-treated M-SHIME vessels versus water-treated M-SHIME system. Due to the log2 fold change representation, genes expressed in one condition only are represented as presented in **A** panel. (**B**) Day 1 versus day -6 expression in water and CMC-treated M-SHIME suspensions. (**C**) Day 11 versus day -6 expression in water and CMC-treated M-SHIME suspensions. (**D**) Day 11 versus day 1 expression in water and CMC-treated M-SHIME suspensions. (**E**) Day 1 versus day -6 expression in water and P80-treated M-SHIME suspensions. (**F**) Day 11 versus day -6 expression in water and P80-treated M-SHIME suspensions. (**G**) Day 11 versus day 1 expression in water and P80-treated M-SHIME suspensions. (*N*=2-3).

Figure S5: Both CMC and P80 alter microbial gene expression in the mucosal-M-SHIME model at various range of functional depths. A 8-vessel M-SHIME system was set up and inoculated with fresh human feces at day -7. Following a 7-day stabilization period, vessels were treated with either water (*N*=2), CMC 1% (w/v, *N*=3) or P80 1% (v/v, *N*=3) for 13 days. M-SHIME-associated luminal metatranscriptome was analyzed by Illumina sequencing. Following functional classification of the reads to Kyoto Encyclopedia of genes and genomes (KEGG) orthology classification at levels 2, 3 and 4 using MG-RAST, days -6, 1 and 11 metatranscriptomes were visualized using area plots (*N*=2-3).

Figure S6: CMC and P80 do not alter short-chain fatty acid production in the mucosal-M-SHIME model. A 8-vessel M-SHIME system was set up and inoculated with fresh human feces at day -7. Following a 7-day stabilization period, vessels were treated with either water (*N*=2), CMC 1% (w/v, *N*=3) or P80 1% (v/v, *N*=3) for 13 days. Short-chain fatty acids (acetate, propionate and butyrate) concentration were determined every day in (**A**) water-treated, (**B**) CMC-treated and (**C**) P80-treated M-SHIME suspensions. Data are the means +/- S.E.M (*N*=2-3).

Figure S7: CMC and P80 do not alter branched short-chain fatty acid production in the mucosal-M-SHIME model. A 8-vessel M-SHIME system was set up and inoculated with fresh human feces at day -7. Following a 7-day stabilization period, vessels were treated with either water (*N*=2), CMC 1% (w/v, *N*=3) or P80 1% (v/v, *N*=3) for 13 days. Branched-chain fatty acids concentration was determined every day in water-treated, CMC-treated and P80-treated M-SHIME suspensions. Data are the means +/- S.E.M (*N*=2-3).

Figure S8: CMC and P80 does not alter LPS levels in the mucosal-M-SHIME model. A 8-vessel M-SHIME system was set up and inoculated with fresh human feces at day -7. Following a 7-day stabilization period, vessels were treated with either water (*N*=2), CMC 1% (w/v, *N*=3) or P80 1% (v/v, *N*=3) for 13 days. (**A**) Luminal M-SHIME suspension LPS levels at days -7, -6, -3, -2, 1, 4, 6, 8, 11 and 13 days. Data are the means +/- S.E.M (*N*=2-3).

Figure S9: Altered microbiota composition in mice receiving CMC- and P80-treated M-SHIME suspensions. A 8-vessel M-SHIME system was set up and inoculated with fresh human feces at day -7. Following a 7-day stabilization period, vessels were treated with either water, CMC 1% or P80 1% for 13 days. Germfree C57BL/6 mice (3-4 weeks old) were removed from isolator and were orally administered with 200µL of luminal M-SHIME suspension (day 11 sample). Transplanted mice were then housed in isolated ventilated cages Isocages (Techniplast). For each condition (water-treated M-SHIME, CMC-treated M-SHIME, P80-treated M-SHIME), 2 cages containing 2-3 animals were used, with each cage receiving suspension from an independent M-SHIME vessel (for each condition, *N*=2 vessel, *N*=2 cages, and *N*=5 animals). Fecal microbiota composition was analyzed using Illumina sequencing of the V4 region of 16S rRNA genes. (**A-B**) Microbiota richness (number of OTUs observed) at days 13 (**A**) and 64 (**B**). (**C-F**) LEfSE (LDA Effect Size) was used to investigate bacterial members that drive differences between groups at days 13 (**C-D**) and 64 (**E-F**). (**C and E**) LDA scores for the differentially altered taxa. Blue, water-enriched taxa; red, CMC-enriched taxa; green, P80-enriched taxa. Only taxa meeting an LDA significant threshold >2.0 are represented. (**D and F**) Taxonomic cladogram obtained from LEfSe analysis of 16S sequences. Blue, water-enriched taxa; red, CMC-enriched taxa; green, P80-enriched taxa. The brightness of each dot is proportional to its

effect size. Data are the means +/- S.E.M (*N*=5). Each treatment group was compared with the control group (water-treated) and statistical significance was determined using two-way ANOVA with Bonferroni's multiple comparisons test.

Figure S10: Altered microbiota metatranscriptome in mice receiving CMC- and P80-treated M-SHIME suspensions. A 8-vessel M-SHIME system was set up and inoculated with fresh human feces at day -7. Following a 7-day stabilization period, vessels were treated with either water, CMC 1% or P80 1% for 13 days. Germfree C57BL/6 mice (3-4 weeks old) were removed from isolator and were orally administered with 200µL of luminal M-SHIME suspension (day 11 sample). Transplanted mice were then housed in isolated ventilated cages Isocages (Techniplast). For each condition (water-treated M-SHIME, CMC-treated M-SHIME, P80-treated M-SHIME), 2 cages containing 2-3 animals were used, with each cage receiving suspension from an independent M-SHIME vessel (for each condition, *N*=2 vessel, *N*=2 cages, and *N*=5 animals). Fecal microbiota composition was analyzed using Illumina sequencing of the V4 region of 16S rRNA genes, and metagenomes were predicted using PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States). (**A-B**) Following functional classification of the reads to Kyoto Encyclopedia of genes and genomes (KEGG) orthology classification at level 4 using PICRUSt, principal coordinates analysis (PCoA) of the euclidean distance matrix were ploted at days 13 (**A**) and 64 (**B**). (**C**) Following functional classification of the reads to Kyoto Encyclopedia of genes and genomes (KEGG) orthology classification at levels 2 and 3 using PICRUSt, days 13 and 64 metatranscriptomes were visualized using area plots (*N*=2-3). For clustering analyzing on principal coordinate plots, categories were compared and statistical significance of clustering were determined using Permanova method

Figure S11: Altered microbiota metatranscriptome in mice receiving CMC- and P80-treated M-SHIME suspensions. A 8-vessel M-SHIME system was set up and inoculated with fresh human feces at day -7. Following a 7-day stabilization period, vessels were treated with either water, CMC 1% or P80 1% for 13 days. Germfree C57BL/6 mice (3-4 weeks old) were removed from isolator and were orally administered with 200µL of luminal M-SHIME suspension (day 11 sample). Transplanted mice were then housed in isolated ventilated cages Isocages (Techniplast). For each condition (water-treated M-SHIME, CMC-treated M-SHIME, P80-treated M-SHIME), 2 cages containing 2-3 animals were used, with each cage receiving suspension from an independent M-SHIME vessel (for each condition, *N*=2 vessel, *N*=2 cages, and *N*=5 animals). Fecal microbiota composition was analyzed using Illumina sequencing of the V4 region of 16S rRNA genes, and metagenomes were predicted using PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States). (**A-H**) Pathways related to flagella synthesis identified using KEGG orthology (KO). Data are the means +/- S.E.M (*N*=2-3). Each treatment group was compared with the control group (water-treated) and statistical significance was determined using Student's *t* tests.

Figure S12: Water-treated M-SHIME suspensions does not promote gross change nor metabolic syndrome when transplanted to germfree recipient mice. A 8-vessels M-SHIME system was set up and inoculated with fresh human feces at day -7. Germfree C57BL/6 mice (3-4 weeks old) were removed from isolator and were orally administered with 200µL of luminal M-SHIME suspension (day 11 sample). Transplanted mice were then housed in isolated ventilated cages Isocages (Techniplast) and compare to germfree animals. (**A**) 5hr fasting blood glucose concentration, (**B**) fat pad weight, (**C**) colon weight, (**D**) colon length, (**E**) colon

weight/length ratio and (**F**) spleen weight. Data are the means +/- S.E.M (*N*=3-5). The data for the water-treated M-SHIME transplanted animals are the same as presented figure 8. Treatment group was compared with the control group (germfree water-treated) and statistical significance was determined using Student's *t* tests.

Figure S13: CMC- and P80-treated M-SHIME suspensions promote low-grade inflammation when transplanted to germfree recipient mice. A 8-vessels M-SHIME system was set up and inoculated with fresh human feces at day -7. Following a 7-day stabilization period, vessels were treated with either water, CMC 1% or P80 1% for 13 days. Germfree C57BL/6 mice (5-10 weeks old) were removed from isolator and were orally administered with 200µL of luminal M-SHIME suspension (day 11 sample). Transplanted mice were then housed in isolated ventilated cages Isocages (Techniplast) and fed irradiated (20-40 kGy) purified high-fat diet D12492-1.5V from Research Diet company. For each condition (water-, CMC-, and P80-treated M-SHIMEs), 2 cages containing 1-2 animals were used, with each cage receiving suspension from an independent M-SHIME vessel (for each condition, N=2 vessel, N=2 cages, and N=3 animals). (**A-B**) Absolute (**A**) and relative (**B**) body weight over time (due to variability in initial body weight (day0), day14 was defined as 100%), (**C**) 5hr fasting blood glucose concentration, (**D**) fat pad weight, (**E**) colon weight, (**F**) colon length, (**G**) colon weight/length ratio and (**H**) spleen weight. Data are the means +/- S.E.M (N=3). Each treatment group was compared with the control group (water-treated) and statistical significance was determined using Student's t tests or two-way ANOVA with Bonferroni's multiple comparisons test.

SUPPLEMENTAL REFERENCES

1 Chassaing B, Srinivasan G, Delgado MA, Young AN, Gewirtz AT, Vijay-Kumar M. Fecal lipocalin 2, a sensitive and broadly dynamic non-invasive biomarker for intestinal inflammation. PLoS One 2012;**7**:e44328.

2 Chassaing B, Koren O, Carvalho FA, Ley RE, Gewirtz AT. AIEC pathobiont instigates chronic colitis in susceptible hosts by altering microbiota composition. Gut 2014;**63**:1069-80.

3 Gomes-Neta JC, Mantz S, Held K, Sinha R, Benson AK, Walter J*, et al.* Evaluation of a new real-time qPCR assay for quantification of the Altered Schaedler Flora microbiota in gnotobiotic mice. 2016.

4 De Weirdt R, Possemiers S, Vermeulen G, Moerdijk-Poortvliet TC, Boschker HT, Verstraete W*, et al.* Human faecal microbiota display variable patterns of glycerol metabolism. FEMS Microbiol Ecol 2010;**74**:601-11.

5 Gilbert JA, Meyer F, Jansson J, Gordon J, Pace N, Tiedje J*, et al.* The Earth Microbiome Project: Meeting report of the "1 EMP meeting on sample selection and acquisition" at Argonne National Laboratory October 6 2010. Stand Genomic Sci 2010;**3**:249-53.

6 Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N*, et al.* Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. The ISME journal 2012;**6**:1621-4.

7 Aronesty E. Comparison of Sequencing Utility Programs. The Open Bioinformatics Journal 2013;**7**:1-8.

8 Aronesty E. Command-line tools for processing biological sequencing data. http://codegooglecom/p/ea-utils 2011.

9 Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK*, et al.* QIIME allows analysis of high-throughput community sequencing data. Nat Med 2010;**7**:335-6.

10 Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 2010;**26**:2460-1.

11 McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A*, et al.* An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J 2012;**6**:610-8.

12 Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol 2009;**26**:1641-50.

13 Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. Appl Environ Microbiol 2005;**71**:8228-35.

14 Lozupone C, Hamady M, Knight R. UniFrac--an online tool for comparing microbial community diversity in a phylogenetic context. BMC Bioinformatics 2006;**7**:371.

15 Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS*, et al.* Metagenomic biomarker discovery and explanation. Genome Biol 2011;**12**:R60.

16 Langille MG, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA*, et al.* Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol 2013;**31**:814-21.

17 Johansson ME, Hansson GC. Preservation of mucus in histological sections, immunostaining of mucins in fixed tissue, and localization of bacteria with FISH. Methods Mol Biol 2012;**842**:229-35.

18 Chassaing B, Koren O, Goodrich JK, Poole AC, Srinivasan S, Ley RE*, et al.* Dietary emulsifiers impact the mouse gut microbiota promoting colitis and metabolic syndrome. Nature 2015;**519**:92-6.

19 Chassaing B, Ley RE, Gewirtz AT. Intestinal epithelial cell toll-like receptor 5 regulates the intestinal microbiota to prevent low-grade inflammation and metabolic syndrome in mice. Gastroenterology 2014;**147**:1363-77 e17.