Supporting Information

Rausch et al. 10.1073/pnas.1106408108

SI Materials and Methods

Human Samples. The biopsy bank of the outpatient clinic of the Department of General Internal Medicine of University Hospital Schleswig-Holstein was screened for individuals with a diagnosis of Crohn disease (CD) ($n = 29$) and for healthy controls ($n = 18$) of Caucasian (northern European) ancestry. CD samples were selected based on the following criteria: (*i*) macro- and microscopically noninflamed sigmoid region biopsies available in biopsy bank and (*ii*) clinical remission with Crohn disease Activity Index (CDAI) score of <150 at time of sampling. The inflammatory activity was independently scored by two investigators. Further patient characteristics are given in Table S1.

All diagnoses were based on standard criteria via radiological and endoscopic examinations. Indications for colonoscopy were monitoring of therapy response and cancer surveillance in CD patients and participation in a volunteer study for the healthy subjects (exclusion of intestinal pathologies before nutritional intervention). All samples (pairs of matching DNA/biopsy) and phenotype information were pseudonymized before the procedure. All procedures related to patients and healthy subjects were approved by the University Hospital Schleswig-Holstein ethics committee (B231/98 and A154/06) and follow the guidelines of the Declaration of Helsinki. All individuals agreed to participation by giving informed consent at least 24 h before the study.

16S rRNA Gene Pyrosequencing. The 16S rRNA gene was amplified by using forward (5′-**CTATGCGCCTTGCCAGCCCGCTCAG***T-C*AGAGTTTGATCCTGGCTCAG-3′) and reverse (5′-**CGTATCGCCTCCCTCGCGCCATCAG**XXXXXXXXXX*CA*TGCTGCCTCCCGTAGGAGT-3′) primers flanking the V1 and V2 hypervariable regions. The 454 Life Sciences primer B (forward) and A (reverse) adapter sequences are denoted in boldface, and the underlined sequences represent the broadly conserved bacterial primers 27F and 338R. A 2-base linker sequence (TC/CA; shown in italics) was added as recommended by Roche (454). A unique 10-base multiplex identifier (designated as XXXXXXXXXX) was added to the reverse primer to tag each PCR product. Template DNA (100 ng) was added to 25-μL PCR reactions performed with Phusion Hot Start DNA Polymerase (Finnzymes). The cycling conditions were as follows: initial denaturation for 30 s at 98 °C; 30 cycles of 9 s at 98 °C, 30 s at 55 °C, and 30 s at 72 °C; and final extension for 10 min at 72 °C. All reactions were performed in duplicate and combined after PCR. PCR products were extracted with the Qiagen MiniElute Gel Extraction Kit and quantified with the Quant-iT dsDNA Broad-Range Assay Kit on a NanoDrop 3300 fluorometer. Equimolar amounts of purified PCR product were pooled and further purified with AMPure beads (Agencourt). A sample of each library was run on an Agilent Bioanalyzer before emulsion PCR and sequencing as recommended by Roche. Amplicon libraries were subsequently sequenced on a 454 Life Sciences GS-FLX using Titanium sequencing chemistry.

Influence of CD Subphenotypes. Because different subphenotypes of CD (i.e., ileal, colonic, ileocolonic) were present in our patient sample, we investigated their influence on the assembly of microbial communities (Table S1). We detected no significant differences in alpha diversity measures between subphenotypes (Shannon H, $F_{2,26} = 0.129$, $P = 0.88$; Shannon evenness (1), $F_{2,26} = 0.132$, $P = 0.877$; Chao1, $F_{2,26} = 0.135$, $P = 0.874$; phylogenetic diversity, $F_{2,26} = 0.041$, $P = 0.960$) nor any significant differences between communities by using measures of beta diversity (*adonis*: Jaccard, $R^2 = 0.065$, $P = 0.889$; Bray-Curtis, $R^2 = 0.06$, $P = 0.913$; unweighted UniFrac, $R^2 = 0.058$, $P = 0.981$).

1. Jost L (2007) Partitioning diversity into independent α and β components. *Ecology* 88: 2427–2439.

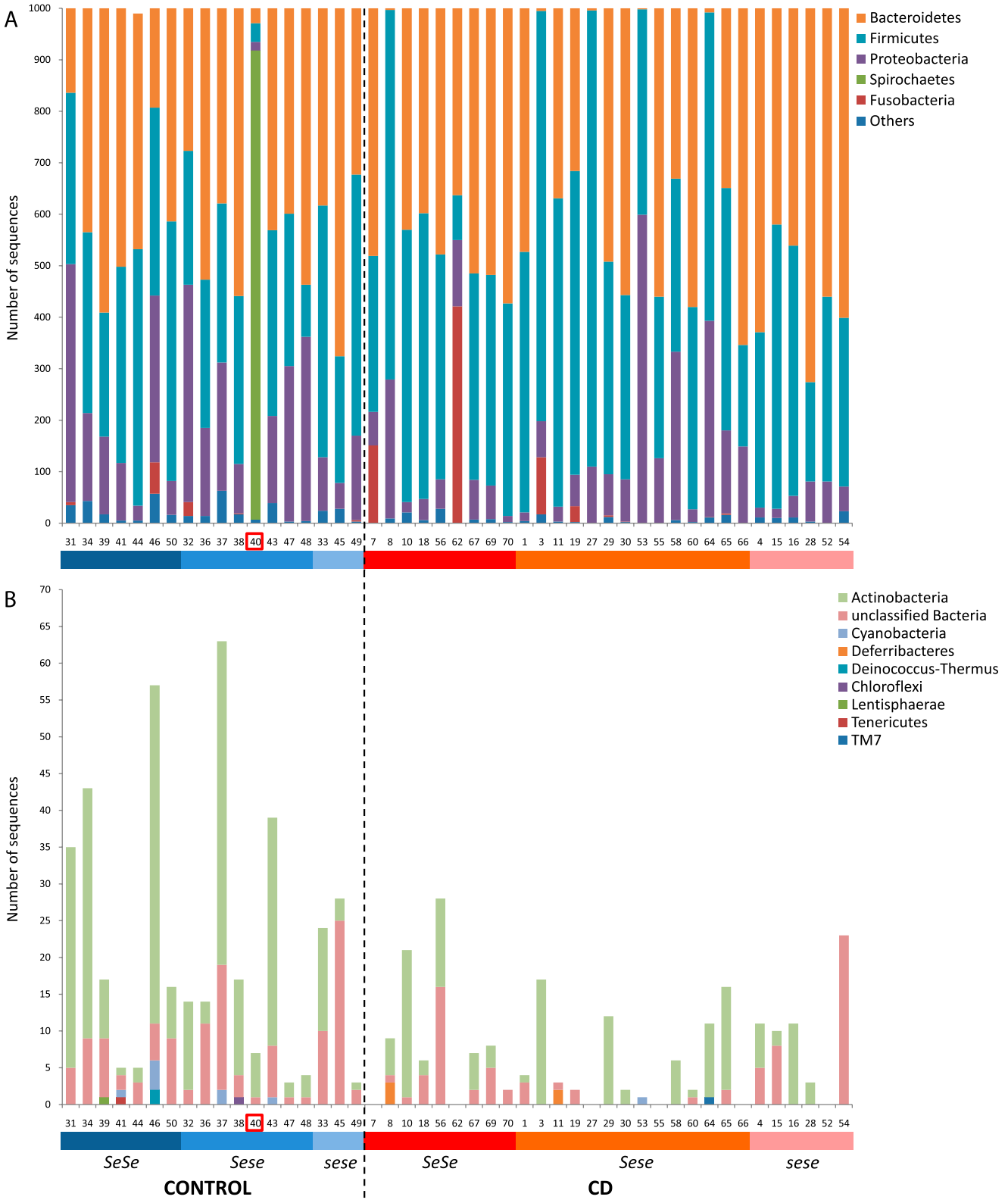

Fig. S1. (*A*) Distribution of read number among the major phyla in sampled individuals. Individual #40 (boxed in red) was excluded from further analysis because of suspected human intestinal spirochaetosis. (*B*) Read number among the rare phyla and unclassified bacteria comprising the "Others" category in *A*.

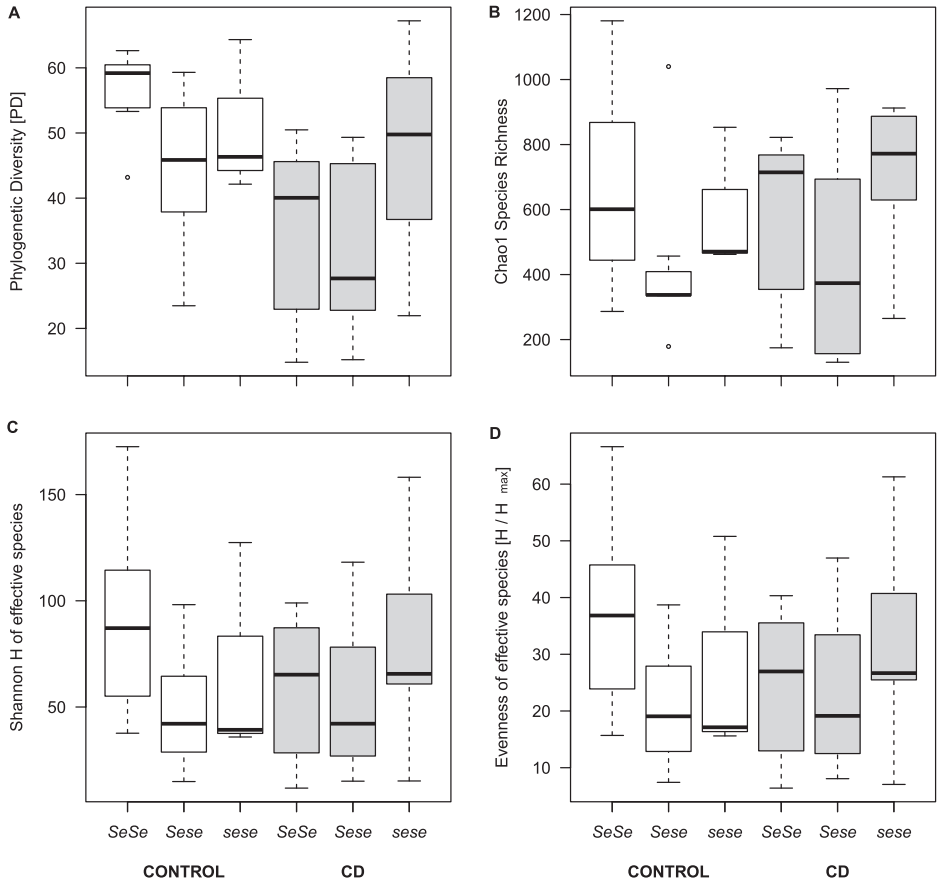

Fig. S2. Mean alpha diversity measures based on species-level operational taxonomic units (OTUs). (*A*) Faith's phylogenetic diversity (1). (*B*) Chao1 species richness metric. (*C*) Shannon H of effective species numbers. (*D*) Shannon evenness of effective species numbers. Error bars indicate SD.

1. Faith DP (1992) Conservation evaluation and phylogenetic diversity. *Biol Conserv* 61:1−10.

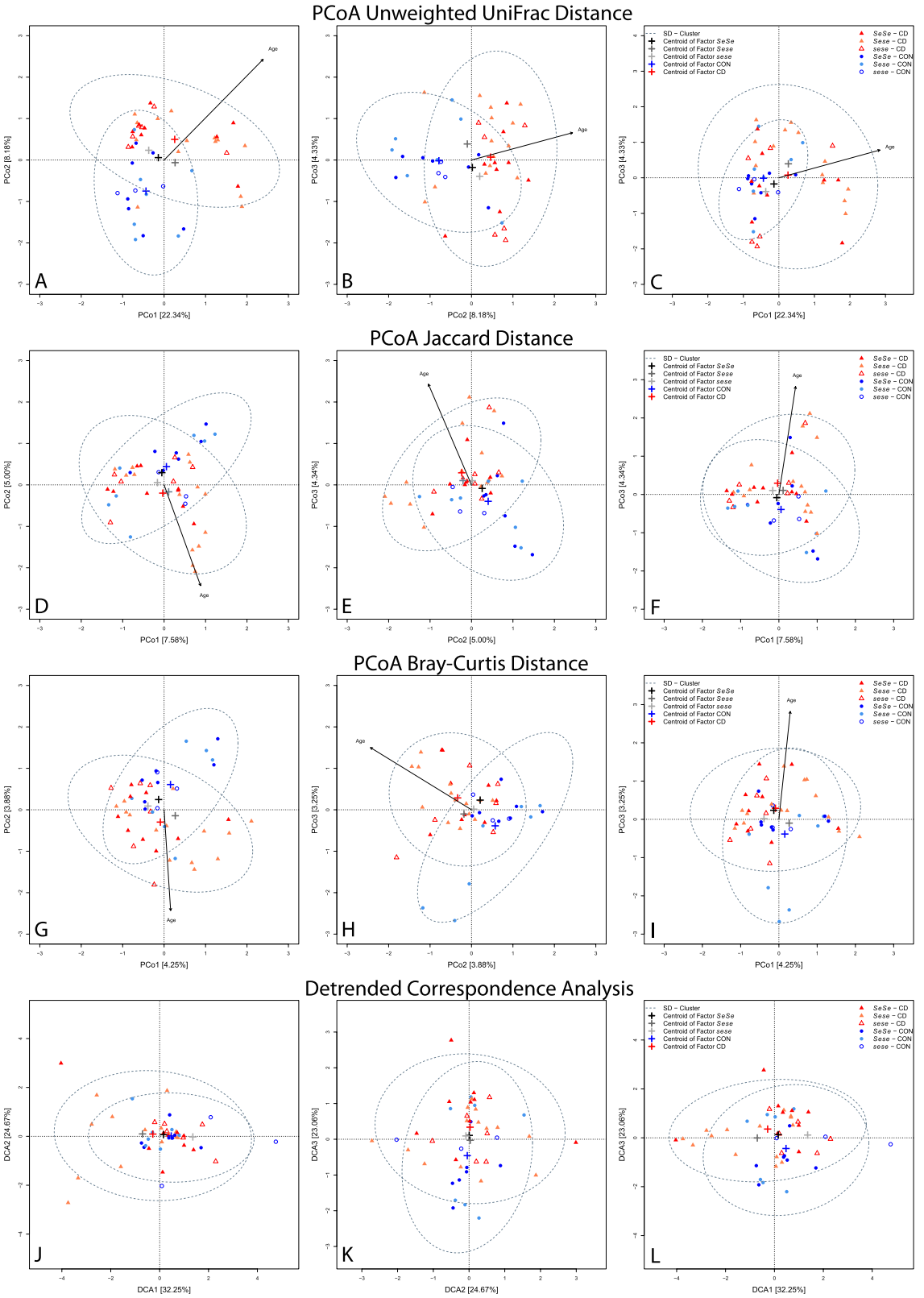

Fig. S3. (*A–C*) Ordination of the unweighted UniFrac distance by principal coordinate analysis (PCoA) in three dimensions (see main text for statistical analysis). (*D–F*) PCoA of the Jaccard index in three dimensions. Centroids of disease status (goodness of fit $R^2 = 0.1048$, $P = 0.002$), genotype within disease status ($R^2 = 0.183$, $P = 0.039$), and age (goodness of fit $R^2 = 0.263$, $P = 0.005$) are correlated with all three axes. Analysis by linear models revealed a correlation of sex with the first axis ($P = 0.047$), whereas the second and third axes show the gradients between disease status ($P = 0.009$) and age ($P = 0.001$), respectively (Table S3). (*G–I*) PCoA of the Bray-Curtis index (data Wisconsin-transformed) in three dimensions. The centroids of disease status (goodness of fit $R^2 = 0.154$, $P < 0.0001$), genotype within disease status (goodness of fit $R^2 = 0.26$, $P = 0.0008$), and age of the subjects (goodness of fit $R^2 = 0.213$, $P = 0.017$) are correlated with

Legend continued on following page

all three axes. Linear models of the site scores revealed a correlation of sex with the first axis ($P = 0.079$), whereas the second and third axes display significant gradients in disease status ($P = 0.0001$) and genotype within disease status ($P = 0.008$), respectively (Table S3). (*J–L*) Detrended correspondence analysis (DCA) visualized in three dimensions. The centroids of disease status (goodness of fit $R^2 = 0.064$, $P = 0.038$), genotype (goodness of fit $R^2 = 0.137$, $P = 0.006$), and genotype–disease status interaction ($R^2 = 0.238$, $P = 0.006$) are highly correlated with all three axes. Analysis by linear models revealed a correlation of genotype with the first axis ($P = 0.003$), whereas the second axis shows no correlation with known factors. The third dimension is significantly influenced by disease status ($P = 0.009$) (Table S3).

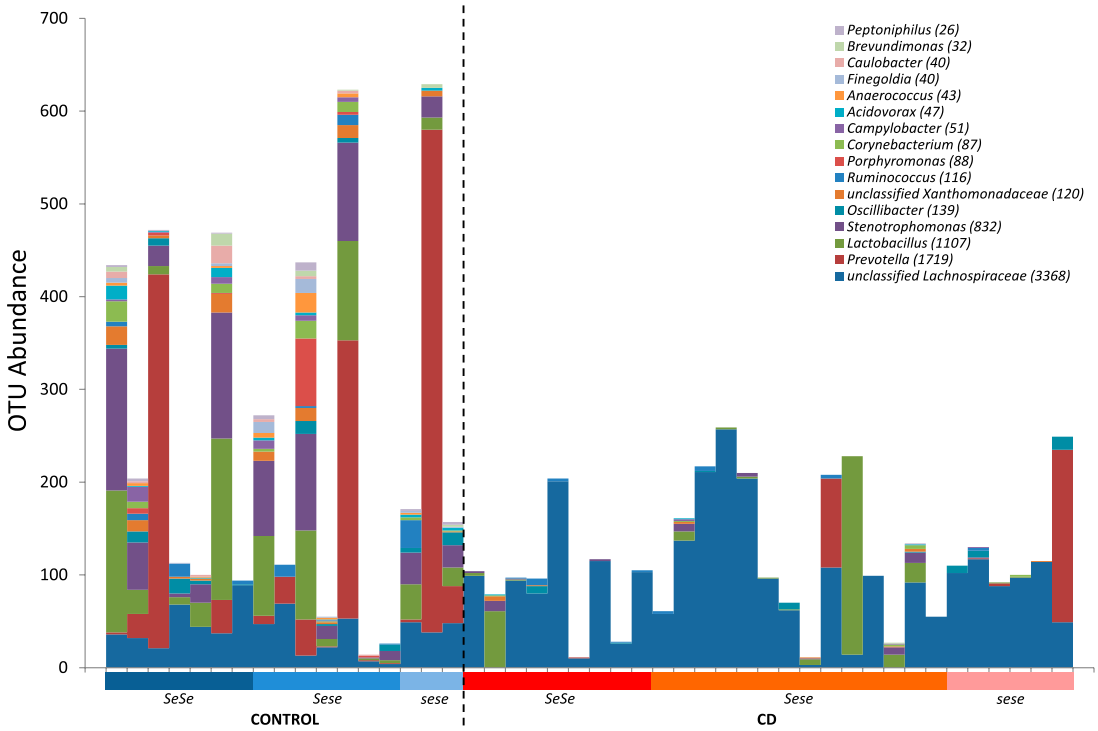

Fig. S4. Abundance of genus-level OTUs identified by indicator species analysis with respect to health status. Numbers in parentheses indicate the total read number in the normalized dataset (1,000 reads per individual).

Table S1. Patient characteristics

	CD				Control			
Parameter	All	*SeSe*	*Sese*	*sese*	All	*SeSe*	*Sese*	*sese*
n	29	8	15	6	18	7	8	3
Sex ratio, female/male	20/9	4/4	13/2	3/3	14/4	6/1	6/2	2/1
Median age, y	39	35.5	43	43	26	28	26.5	25
Age range, y	27–55	29–48	27–55	36–48	23–30	25–30	23–30	23–26
Medication					N/A	N/A	N/A	N/A
Cortisone	5/29	1	3	1				
Azathioprine	1/29	1	0	0				
5-Aminosalicylic acid (5-ASA)	2/29	1	0	1				
Anti-TNFα	2/29	1	0	1				
Disease subtype					N/A	N/A	N/A	N/A
Ileal	8/29	3	4	1				
Ileocolonic	16/29	3	9	4				
Colonic	4/29	2	1	1				
Undetermined	1/29	0	1	0				

N/A, not applicable; *SeSe*, homozygous secretor; *Sese*, heterozygous secretor; *sese*, nonsecretor.

Table S2. Statistical analysis of phyla abundances

		Phyla			
Test	Categories tested	Actinobacteria	Firmicutes	Bacteroidetes	Proteobacteria
ANOVA	Disease status	**0.002**	**0.034**	N/A	**0.020**
	Secretor status	N/A	N/A	**0.036**	N/A
	Genotype	N/A	N/A	N/A	**0.047**
Post hoc Mann–Whitney					
Subsets					
Secretor	Disease status	**0.006***	**0.005***	N/A	**0.012**
Nonsecretor	Disease status	0.603	0.714	N/A	0.095
Genotype	*SeSe:sese*	N/A	N/A	0.121	0.295
	SeSe:Sese	N/A	N/A	0.728	0.154
	Sese:sese	N/A	N/A	**0.044**	**0.017**
Control	*SeSe:sese*	N/A	N/A	N/A	0.5167
	SeSe:Sese	N/A	N/A	N/A	0.3374
	Sese:sese	N/A	N/A	N/A	0.067
CD	*SeSe:sese*	N/A	N/A	N/A	0.596
	SeSe:Sese	N/A	N/A	N/A	0.156
	Sese:sese	N/A	N/A	N/A	0.0622

Significant *P* values are indicated in boldface. N/A, no test applied.
*Significant after sequential Bonferroni correction.

Table S3. Linear model analysis of PCoA and DCA axis scores

Ordination	Transformation	Axis	Model	Akaike information criterion	Factors	Adjusted R^2	*P*
PCoA	None	1, 2, 3	Null	127.205	1		
Unweighted UniFrac	None	1	Best	122.732	Disease status	0.112	0.0134
	None	2	Best	103.740	Disease status	0.412	<0.0001
	None	3	Best	124.938	Genotype	0.087	0.0534
PCoA Jaccard	None	1, 2, 3	Null	115.314	1		
	None	1	Best	113.150	Sex	0.066	0.0471
	None	2	Best	110.158	Disease status	0.125	0.0093
	None	3	Best	106.015	Age	0.200	0.0011
PCoA Bray-Curtis	Wisconsin	1, 2, 3	Null	116,312	1		
	Wisconsin	1	Best	115.060	Sex	0.047	0.0795
	Wisconsin	2	Best	102.922	Disease status	0.268	0.0001
	Wisconsin	3	Best	108.887	Disease status × genotype	0.230	0.0078
DCA	None	1	Null	175.828	1		
	None		Best	167.067	Genotype	0.207	0.0026
	None	2	Null	130.402	1		
	None		Best	> 130.402	N/A	N/A	N/A
	None	3	Null	135.010	1		
	None		Best	129.755	Disease status	0.127	0.0088

N/A, no test applied.

Table S4. List of indicator species and genera and their properties

Taxonomic level	Association	Direction	R^2	*P* (Benjamini and Hochberg-adjusted)*	OTU ID	OTU classification to genus level (no. of reads in normalized dataset)†	Comments
Species	Disease status	Control	0.435	**0.00001 (0.00025)**	1117	*Lactobacillus* spp. (173)	Normal member of the colon microbiome. probiotic characteristics (1, 2)
		Control	0.446	**0.00007 (0.0007)**	2270	*Lactobacillus* spp. (554)	Normal member of the colon microbiome, probiotic characteristics (1, 2)
		Control	0.495	**0.00005 (0.00063)**	2420	*Lactobacillus* spp. (57)	Normal member of the colon microbiome, probiotic characteristics (1, 2)
		Control	0.373	**0.0004 (0.00286)**	2267	*Prevotella* spp. (73)	Mucosa-associated anaerobe, found in the upper and lower gastrointestinal (GI) tract, pathogenic with increasing relevance and resistance (3, 4)
		Control	0.266	0.0146 (0.073)	2165	*Coprobacillus* spp. (377)	Altered abundance in inflammatory bowel disease (IBD) cases and belonging to the core microbiome of the gut (5, 6)
		Control	0.482	**0.00001 (0.00025)**	2460	*Stenotrophomonas* spp. (519)	Ubiquitous, nosocomial pathogen (7)
		Control	0.479	**0.00014 (0.00117)**	2562	*Stenotrophomonas* spp. (224)	Ubiquitous, nosocomial pathogen (7)
		Control	0.514	**0.00004 (0.000625)**	3631	*Stenotrophomonas* spp. (30)	Ubiquitous, nosocomial pathogen (7)
		Control	0.358	**0.00194 (0.0108)**	3844	*Faecalibacterium* spp. (63)	Normal member of the colon microbiome, depletion with potential role in IBD and suspected as probiotic with potential to reduce inflammatory responses (8–11)
		Control	0.33	**0.00157 (0.0098)**	2551	Unclassified Clostridiales spp. (80)	
	Genotype	*sese*	0.545	0.00082 (0.057)	442	Unclassified Lachnospiraceae spp. (17)	Increased in dextran sodium sulfate (DSS) mouse models (12)
		sese	0.504	0.00211 (0.073)	3774	*Coprococcus* spp. (22)	Saccharolytic/cellulolytic/amylolytic, also fucose (13, 14), associated with CD (15)
	Genotype–disease status	*SeSe* and *Sese*-Control	0.576	0.0172 (0.0909)	2420	*Lactobacillus* spp. (57)	Normal member of the colon microbiome, probiotic characteristics (1, 2)

Table S4. Cont.

Taxonomic level	Association	Direction	R^2	*P* (Benjamini and Hochberg-adjusted)*	OTU ID	OTU classification to genus level (no. of reads in normalized dataset)†	Comments
		SeSe and *Sese*-Control	0.557	0.0207 (0.097)	2562	*Stenotrophomonas* spp. (224)	Ubiquitous, nosocomial pathogen (7)
		SeSe and *Sese*-Control	0.598	0.011 (0.065)	3631	*Stenotrophomonas* spp. (30)	Ubiquitous, nosocomial pathogen (7)
		sese-Control	0.69	**0.00417 (0.0339)**	2267	*Prevotella* spp. (73)	Mucosa-associated anaerobe, found in the upper and lower GI tract, pathogenic with increasing relevance and resistance (3, 4)
		sese-Control	0.516	0.01463 (0.0816)	2488	*Brevundimonas* spp. (12)	Pathogen in immunocompromised subjects and distributed in environment (16–18)
		sese-Control	0.536	0.0211 (0.097)	3689	Unclassified Lachnospiraceae spp. (34)	Increased in DSS mouse models (12)
		sese-Control	0.551	**0.00449 (0.0339)**	2193	*Sutterella* spp. (61)	Member of the colon microbiome and associated with GI infections (3, 10, 19, 20)
		sese-Control	0.717	**0.003 (0.0302)**	3844	*Faecalibacterium* spp. (63)	Normal member of the colon microbiome, depletion with potential role in IBD (8–10)
		sese-CD	0.511	0.02053 (0.0972)	4018	*Alistipes* spp. (46)	Normal member of the GI microbiome (6, 21)
		sese-CD	0.759	**0.00073 (0.0302)**	442	Unclassified Lachnospiraceae spp. (17)	Increased in DSS mouse models (12)
		sese-CD	0.608	0.00906 (0.0565)	3774	*Coprococcus* spp. (22)	Saccharolytic/cellulolytic/amylolytic, also fucose (13, 14), associated with CD (15)
Genus	Disease status	Control	0.498	**0.00003 (0.00036)**		Unclassified Xanthomonadaceae (120)	Associated with gut and oral microbiome but majorly environmental organisms with some pathogenicity, especially in plants (7, 22–25)
		Control	0.409	**0.00001 (0.00024)**		*Acidovorax* (47)	Skin microbiome, associated with earthworm development and gut symbiosis (26–28)
		Control	0.540	**0.00001 (0.00024)**		*Stenotrophomonas* (832)	Ubiquitous, nosocomial pathogen (7)
		Control	0.342	**0.00823 (0.03039)**		*Ruminococcus* (116)	Core gut and fecal microbiome, increased in healthy subjects (6, 29, 30)

Table S4. Cont.

Taxonomic level	Association	Direction	R^2	*P* (Benjamini and Hochberg-adjusted)*	OTU ID	OTU classification to genus level (no. of reads in normalized dataset)†	Comments
		Control	0.306	0.01572 (0.0503)		*Prevotella* (1,719)	Mucosa-associated pathogen with increasing relevance and resistance (3)
		Control	0.210	**0.00136 (0.00593)**		*Porphyromonas* (88)	Saccharolytic, mucosa-associated, cause of gingivitis and soft tissue infections (3, 14)
		Control	0.406	**0.00025 (0.001714)**		*Peptoniphilus* (26)	Chronic soft tissue infections, diabetic ulcers, periodontitis (31–33)
		Control	0.386	**0.00975 (0.0334)**		*Oscillibacter* (139)	Anaerobic *Oscillibacter* was identified in clamp intestines, unknown ecology but found in healthy human guts (34–36)
		Control	0.331	0.020603 (0.07809)		*Lactobacillus* (1,107)	Normal member of the colon microbiome, probiotic characteristics (1, 2)
		Control	0.360	**0.00012 (0.00096)**		*Finegoldia* (40)	Opportunistic pathogen on mucosa (37), associated with chronic infections (31, 32)
		Control	0.396	**0.00061 (0.00336)**		*Corynebacterium* (87)	Highly diverse group, pathogens and saprophytes (38–40)
		Control	0.343	**0.00009 (0.000864)**		*Caulobacter* (40)	Stomach flora (41)
		Control	0.390	**0.00072 (0.003456)**		*Campylobacter* (51)	Cause of enteritis and increased IBD risk (42)
		Control	0.350	**0.00246 (0.00984)**		*Brevundimonas* (32)	Pathogen in immunocompromised subjects and distributed in environment (16–18)
		Control	0.344	**0.00002 (0.00032)**		*Anaerococcus* (43)	Chronic soft tissue infections, diabetic ulcers, periodontitis (31–33)
		CD	0.486	**0.00063 (0.00336)**		Unclassified Lachnospiraceae (3,368)	Increased in DSS mouse models (12)

*Values in boldface indicate significance after Benjamini and Hochberg adjustment (43).
†Classification obtained on the genus level by Ribosomal Database Project (RDP) classifier at the 80% bootstrap threshold (44, 45).

1. Lebeer S, Vanderleyden J, De Keersmaecker SCJ (2008) Genes and molecules of lactobacilli supporting probiotic action. *Microbiol Mol Biol Rev* 72:728–764.
2. Zoetendal EG, et al. (2002) Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. *Appl Environ Microbiol* 68:3401–3407.
3. Falagas ME, Siakavellas E (2000) *Bacteroides*, *Prevotella*, and *Porphyromonas* species: A review of antibiotic resistance and therapeutic options. *Int J Antimicrob Agents* 15:1–9.
4. Eckburg PB, et al. (2005) Diversity of the human intestinal microbial flora. *Science* 308:1635–1638.
5. Kassinen A, et al. (2007) The fecal microbiota of irritable bowel syndrome patients differs significantly from that of healthy subjects. *Gastroenterology* 133:24–33.
6. Tap J, et al. (2009) Towards the human intestinal microbiota phylogenetic core. *Environ Microbiol* 11:2574–2584.
7. Hauben L, Vauterin L, Moore ERB, Hoste B, Swings J (1999) Genomic diversity of the genus *Stenotrophomonas*. *Int J Syst Bacteriol* 49:1749–1760.
8. Sokol H, et al. (2009) Low counts of *Faecalibacterium prausnitzii* in colitis microbiota. *Inflamm Bowel Dis* 15:1183–1189.
9. Marteau P (2009) Bacterial flora in inflammatory bowel disease. *Dig Dis* 27(Suppl 1):99–103.
10. Dethlefsen L, Relman DA (2010) Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc Natl Acad Sci USA* 108 (Suppl 1):4554–4561.
11. Sokol H, et al. (2008) *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci USA* 105:16731–16736.
12. Nagalingam NA, Kao JY, Young VB (2011) Microbial ecology of the murine gut associated with the development of dextran sodium sulfate-induced colitis. *Inflamm Bowel Dis* 17: 917–926.
13. Palmonari A, Stevenson DM, Mertens DR, Cruywagen CW, Weimer PJ (2010) pH dynamics and bacterial community composition in the rumen of lactating dairy cows. *J Dairy Sci* 93: 279–287.
14. Duncan SH, Louis P, Flint HJ (2007) Cultivable bacterial diversity from the human colon. *Lett Appl Microbiol* 44:343–350.
15. Wensinck F, van de Merwe JP, Mayberry JF (1983) An international study of agglutinins to *Eubacterium*, *Peptostreptococcus* and *Coprococcus* species in Crohn's disease, ulcerative colitis and control subjects. *Digestion* 27:63–69.
16. Han XY, Andrade RA (2005) *Brevundimonas diminuta* infections and its resistance to fluoroquinolones. *J Antimicrob Chemother* 55:853–859.
17. Egert M, Schmidt I, Bussey K, Breves R (2010) A glimpse under the rim—The composition of microbial biofilm communities in domestic toilets. *J Appl Microbiol* 108:1167–1174.
18. Grossart H-P, Dziallas C, Leunert F, Tang KW (2010) Bacteria dispersal by hitchhiking on zooplankton. *Proc Natl Acad Sci USA* 107:11959–11964.
19. Wexler HM, et al. (1996) *Sutterella wadsworthensis* gen. nov., sp. nov., bile-resistant microaerophilic *Campylobacter gracilis*-like clinical isolates. *Int J Syst Bacteriol* 46:252–258.
20. Sakon H, Nagai F, Morotomi M, Tanaka R (2008) *Sutterella parvirubra* sp. nov. and *Megamonas funiformis* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol* 58:970–975.
21. Hoyles L, McCartney AL (2009) What do we mean when we refer to Bacteroidetes populations in the human gastrointestinal microbiota? *FEMS Microbiol Lett* 299:175–183.
22. Conti S, dos Santos SS, Koga-Ito CY, Jorge AO (2009) Enterobacteriaceae and pseudomonadaceae on the dorsum of the human tongue. *J Appl Oral Sci* 17:375–380.
23. Rajilić-Stojanović M, Smidt H, de Vos WM (2007) Diversity of the human gastrointestinal tract microbiota revisited. *Environ Microbiol* 9:2125–2136.
24. Dewhirst FE, et al. (2010) The human oral microbiome. *J Bacteriol* 192:5002–5017.
25. da Silva ACR, et al. (2002) Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature* 417:459–463.
26. Thakuria D, Schmidt O, Finan D, Egan D, Doohan FM (2010) Gut wall bacteria of earthworms: A natural selection process. *ISME J* 4:357–366.
27. Davidson SK, Stahl DA (2008) Selective recruitment of bacteria during embryogenesis of an earthworm. *ISME J* 2:510–518.
28. Price LB, et al. (2010) The effects of circumcision on the penis microbiome. *PLoS ONE* 5:e8422.
29. Joossens M, et al. (2011) Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. *Gut* 60:631–637.
30. Liu C, Finegold SM, Song Y, Lawson PA (2008) Reclassification of *Clostridium coccoides*, *Ruminococcus hansenii*, *Ruminococcus hydrogenotrophicus*, *Ruminococcus luti*, *Ruminococcus productus* and *Ruminococcus schinkii* as *Blautia coccoides* gen. nov., comb. nov., *Blautia hansenii* comb. nov., *Blautia hydrogenotrophica* comb. nov., *Blautia luti* comb. nov., *Blautia producta* comb. nov., *Blautia schinkii* comb. nov. and description of *Blautia wexlerae* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol* 58:1896–1902.
31. Price LB, et al. (2009) Community analysis of chronic wound bacteria using 16S rRNA gene-based pyrosequencing: Impact of diabetes and antibiotics on chronic wound microbiota. *PLoS ONE* 4:e6462.
32. Dowd SE, et al. (2008) Polymicrobial nature of chronic diabetic foot ulcer biofilm infections determined using bacterial tag encoded FLX amplicon pyrosequencing (bTEFAP). *PLoS ONE* 3:e3326.
33. Kumar PS, Griffen AL, Moeschberger ML, Leys EJ (2005) Identification of candidate periodontal pathogens and beneficial species by quantitative 16S clonal analysis. *J Clin Microbiol* 43:3944–3955.
34. Iino T, Mori K, Tanaka K, Suzuki K-i, Harayama S (2007) *Oscillibacter valericigenes* gen. nov., sp. nov., a valerate-producing anaerobic bacterium isolated from the alimentary canal of a Japanese corbicula clam. *Int J Syst Evol Microbiol* 57:1840–1845.
35. Walker AW, et al. (2011) Dominant and diet-responsive groups of bacteria within the human colonic microbiota. *ISME J* 5:220–230.
36. Mondot S, et al. (2011) Highlighting new phylogenetic specificities of Crohn's disease microbiota. *Inflamm Bowel Dis* 17:185–192.
37. Goto T, et al. (2008) Complete genome sequence of *Finegoldia magna*, an anaerobic opportunistic pathogen. *DNA Res* 15:39–47.
38. Gao Z, Tseng CH, Pei Z, Blaser MJ (2007) Molecular analysis of human forearm superficial skin bacterial biota. *Proc Natl Acad Sci USA* 104:2927–2932.
39. Ventura M, et al. (2007) Genomics of Actinobacteria: Tracing the evolutionary history of an ancient phylum. *Microbiol Mol Biol Rev* 71:495–548.
40. Madigan MT, Martinko JM (2006) *Brock Biology of Microorganisms* (Pearson Prentice Hall, Upper Saddle River, NJ), 11th Ed.
41. Bik EM, et al. (2006) Molecular analysis of the bacterial microbiota in the human stomach. *Proc Natl Acad Sci USA* 103:732–737.
42. Nell S, Suerbaum S, Josenhans C (2010) The impact of the microbiota on the pathogenesis of IBD: Lessons from mouse infection models. *Nat Rev Microbiol* 8:564–577.
43. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc, B* 57:289–300.
44. Cole JR, et al.; Ribosomal Database Project (2003) The Ribosomal Database Project (RDP-II): Previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucleic Acids Res* 31:442–443.
45. Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73: 5261–5267.

Dataset S1. OTU abundances per individual at 97% sequence similarity threshold

Dataset S1