Supporting Information

Claesson et al. 10.1073/pnas.1000097107

SI Materials and Methods

Subject Recruitment. Subjects were clinically examined at either of two local hospitals that serve a population base of ~481,000 in the Cork city and county region. The subjects were defined as (*i*) community-dwelling (control healthy), (*ii*) attending an outpatient day hospital (out-patient), (*iii*) in short-term rehabilitation hospital care (Rehabilitation; under 6-wk stay), or (*iv*) in long-term institutionalized care (Long stay; more than 6 wk). The mean age of the subjects was 78 (± 7) y, with a range of 65–96 y. A full-analysis proforma was also applied to all subjects returning after 3 mo to allow for the identification of changes in lifestyle or health status or major dietary changes during this time.

Sequence-Based Microbiota Analysis and Phylogeny. The 16S rRNA V4 amplicons were sequenced on a 454 Genome Sequencer FLX platform (Roche Diagnostics) according to 454 protocols. Raw sequencing reads were quality trimmed according to published recommendations (1) using a locally installed version of the Ribosomal Database Project (RDP) Pyrosequencing Pipeline (2) applying the following criteria: (*i*) exact matches to primer sequences and barcode tags, (*ii*) no ambiguous bases (Ns), and (*iii*) read-lengths no shorter than 150 bp. For large-scale assignments into the new Bergey bacterial taxonomy (3), we used the Naïve Bayesian Classifier (RDP classifier), which provides rapid taxonomic classifications from domain to genus of both partial and full-length rRNA gene sequences along with bootstrap-like confidence estimates (4). Our choice of 50% as confidence-value threshold was based on what was found suitable for V4 amplicons from the human gut environment (5). Trimmed sequences with their classifications were imported into a MySQL database for efficient storage and advanced querying. Pyrosequencing reads were aligned using Infernal (6) and associated covariance models obtained from the Ribosomal Database Project Group. These were based on secondary structural information from full-length 16S rRNA genes sequences of 508 fully sequenced genomes and were further trimmed to encompass only the V4 region to increase alignment speeds.

Statistical Analysis. The average proportion of total reads assigned to the different phyla and genera was compared according to antibiotic use for the 161 baseline samples using the nonparametric Mann–Whitney test. Although the data are more suitable to the nonparametric methods, independent-sample *t* tests were also conducted in tandem with the Mann–Whitney tests and provided the same conclusions, unless indicated otherwise in *Results*.

The average UniFrac distances within the 26 individuals from time 0 to time 3 mo were compared with the average between individual UniFrac distances at both time 0 and time 3 mo using robust *t* tests to allow for dependence between observations. These comparisons were made for both the weighted and unweighted UniFrac distances.

The change in the proportion of reads assigned to the different phyla and genera from time 0 to time 3 mo was compared according to antibiotic use for the 26 subjects who provided samples at both time points using the nonparametric Mann–Whitney test. Again, independent sample *t* tests were also conducted in tandem and provided the same conclusions as the Mann–Whitney test, unless indicated otherwise in *Results*.

1. Huse SM, Huber JA, Morrison HG, Sogin ML, Welch DM (2007) Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol* 8:R143.
2. Cole JR, et al. (2009) The Ribosomal Database Project: Improved alignments and new tools for rRNA analysis. *Nucleic Acids Res* 37:D141–D145.
3. Lilburn TG, Garrity GM (2004) Exploring prokaryotic taxonomy. *Int J Syst Evol Microbiol* 54:7–13.
4. 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.
5. Claesson MJ, et al. (2009) Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS ONE* 4:e6669.
6. Nawrocki EP, Eddy SR (2007) Query-dependent banding (QDB) for faster RNA similarity searches. *PLOS Comput Biol* 3:e56.

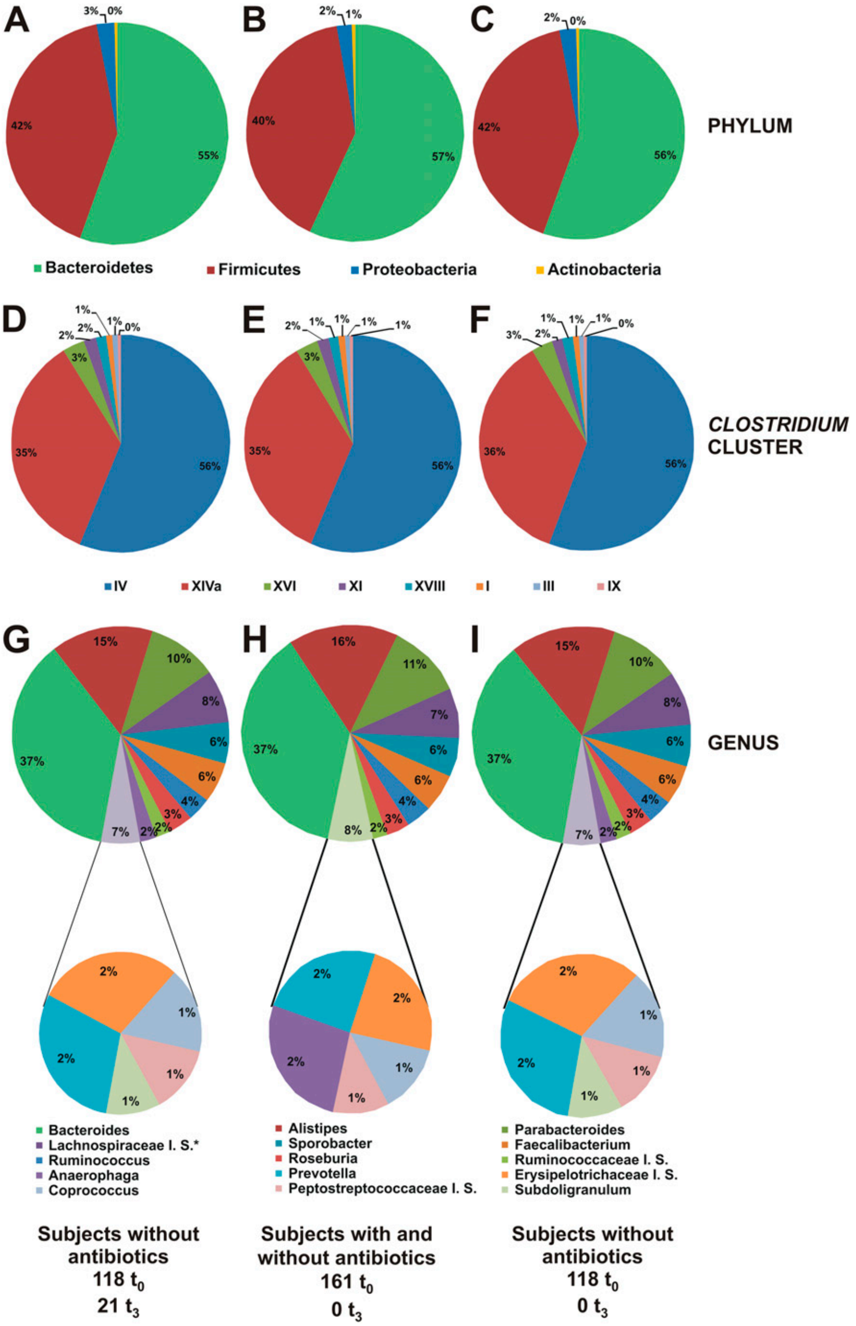

Fig. S1. Aggregate microbiota composition at the level of phylum, *Clostridium* cluster level, and genus, as indicated. Only the 4 most populated phyla (*A*–*C*), 8 most populated *Clostridium* clusters (*D*–*F*), and 15 largest genera (*G*–*I*) are shown.

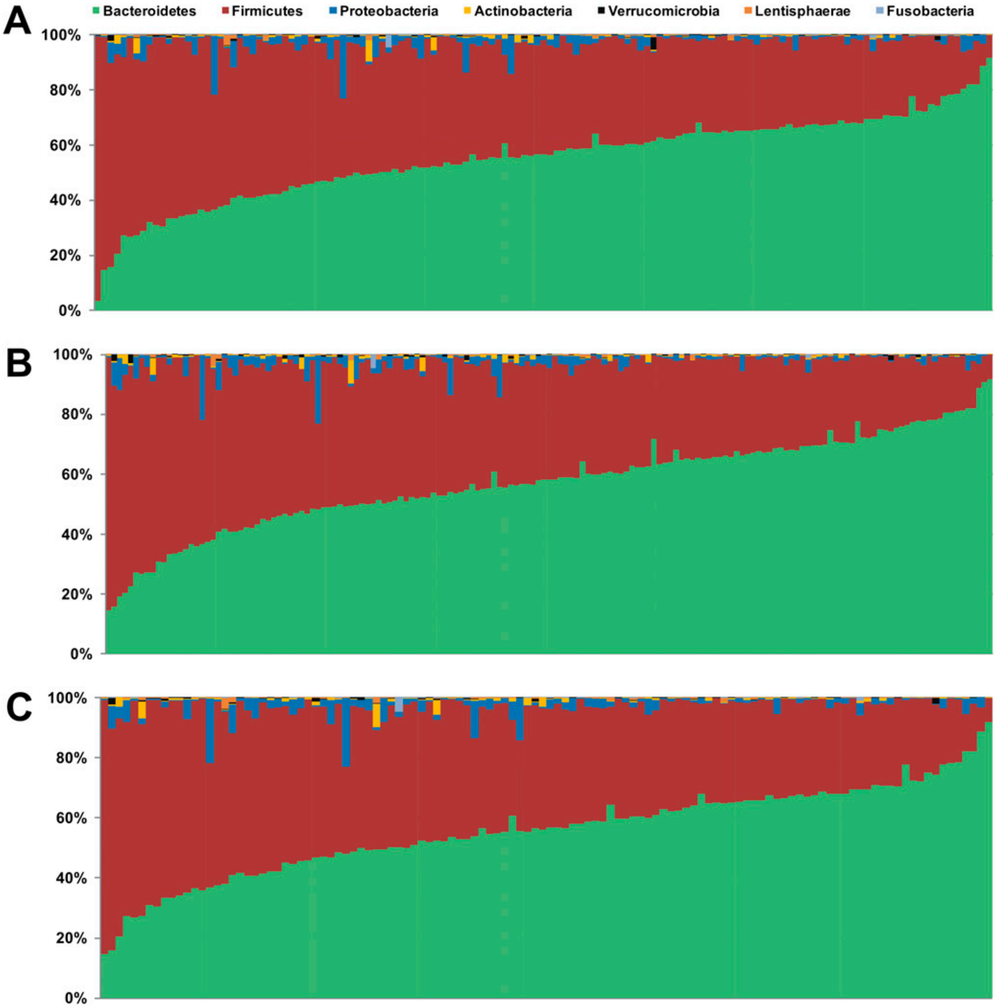

Fig. S2. Microbiota composition, at phylum level, in 139 (118 time 0 and 21 time 3 mo) fecal samples from elderly Irish subjects who had not been treated with antibiotics (*A*), 161 time-0 samples from subjects either treated or not treated with antibiotics (*B*), and 118 time-0 samples from subjects not treated with antibiotics (*C*). Only the seven largest phyla are shown. Samples were ordered by the proportion of all reads assigned to phylum *Bacteroidetes* to illustrate the dramatic interindividual variation.

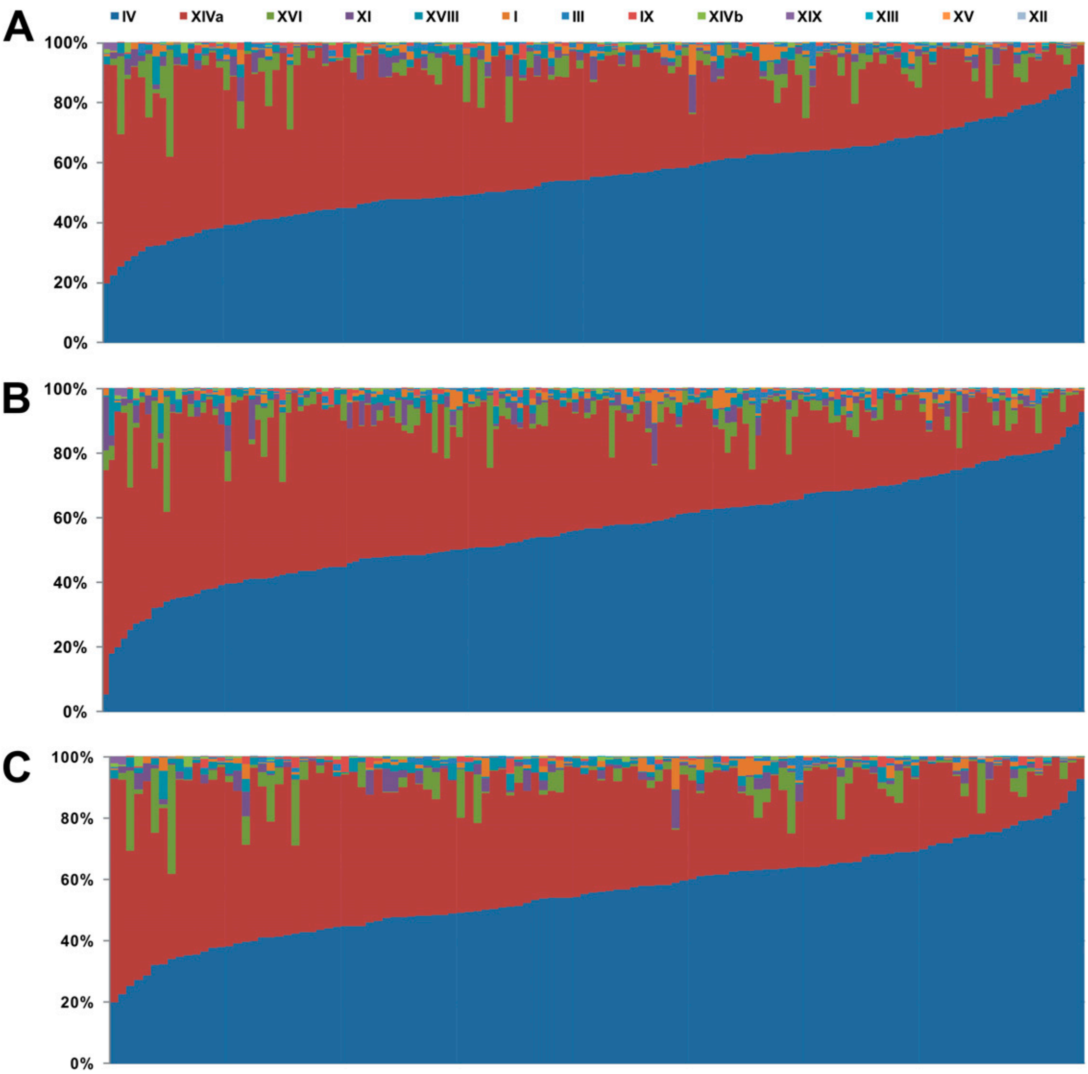

Fig. S3. Interindividual variation in the proportions of *Clostridium* clusters in 139 (118 time 0 and 21 time 3 mo) fecal samples from elderly Irish subjects who had not been treated with antibiotics (*A*), 161 time-0 samples from subjects either treated or not treated with antibiotics (*B*), and 118 time-0 samples from subjects not treated with antibiotics (*C*). Samples were ordered by the proportion of all reads assigned to *Clostridium* cluster IV to illustrate the dramatic interindividual variation.

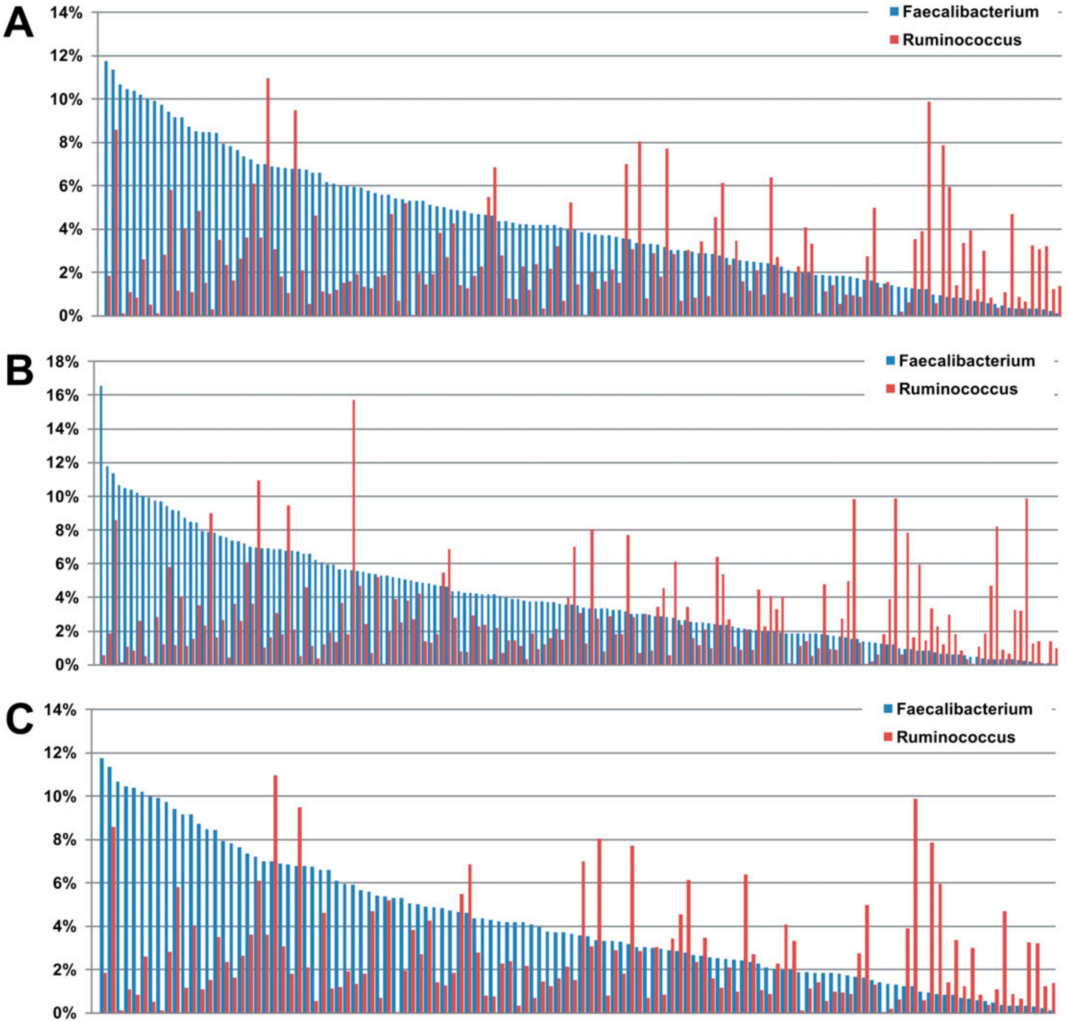

Fig. S4. Interindividual variation in the proportions of the genera *Faecalibacterium* and *Ruminococcus* in 139 (118 time 0 and 21 time 3 mo) fecal samples from elderly Irish subjects who had not been treated with antibiotics (*A*), 161 time-0 samples from subjects either treated or not treated with antibiotics (*B*), and 118 time-0 samples from subjects not treated with antibiotics (*C*). Samples are ordered by the proportion of all reads assigned to *Faecalibacterium* to illustrate the dramatic interindividual variation.

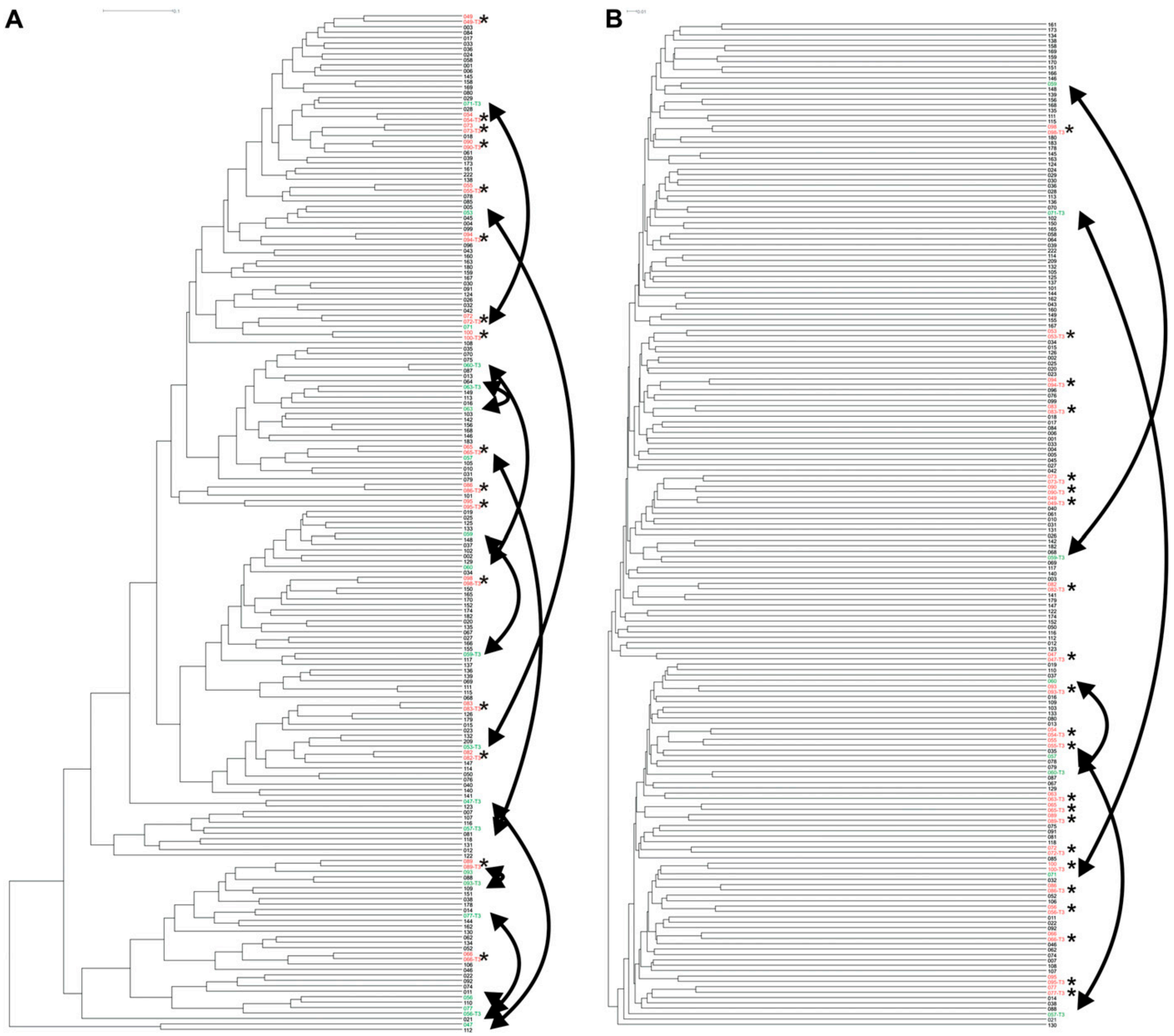

Fig. S5. Weighted (*A*) and unweighted (*B*) UniFrac tree of 161 time-0 fecal microbiota composition datasets and 26 time 3-mo datasets; paired samples that colocalize in the tree are highlighted in red and with asterisks, and unpaired samples are in green and with arrows. Scale bar indicates a weighted UniFrac distance of 0.1 and 0.01 for weighted and unweighted, respectively.

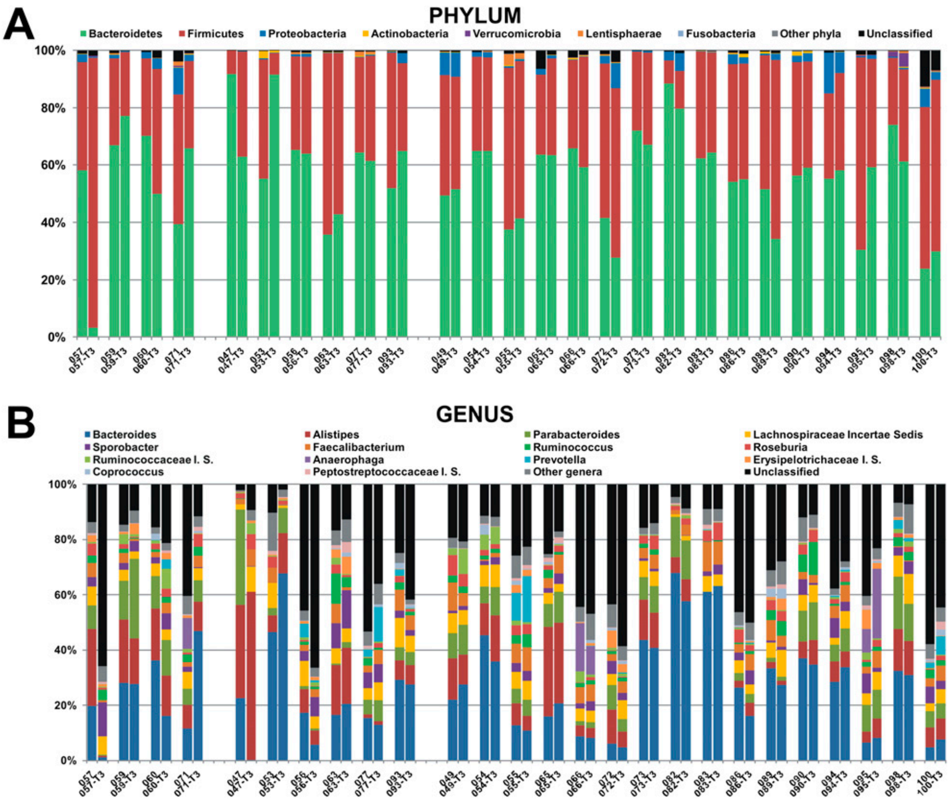

Fig. S6. Variation in the microbiota between time-0 and time 3-mo samples from 26 subjects: phylum assignments (*A*) and genus assignments (*B*). The first of the three sections in both panels contains paired samples that were not closest neighbors in either the unweighted UniFrac tree or the weighted tree. The second section contains samples that only colocated in the unweighted tree. The third section contains samples that colocated in both trees.

Table S1. Distribution of 161 T0 and 26 T3 study participants by antibiotic use and age

	Age (y)		
	Mean (SD)	Median (quartiles 1, 3)	Range (min, max)
All T0 subjects (*n* = 161)	78 (7)	78 (72, 84)	65, 96
T0 subjects – antibiotic users (*n* = 43)	81 (7)	81 (81, 85)	68, 96
T0 subjects – antibiotic-free (*n* = 118)	77 (7)	77 (72, 84)	65, 95
All T3 subjects (*n* = 26)	73 (7)	72 (67, 77)	65, 88
T3 subjects – antibiotic users (*n* = 5)	71 (5)	68 (67, 75)	67, 78
T3 subjects – antibiotic-free (*n* = 21)	73 (7)	72 (67, 77)	65, 88

Table S2. Core microbiota identified in the ELDERMET study and in healthy adults from four previously published studies

	N	Age (mean ± SD years)	Total number of V4 sequences	Number of unique sequences in core	Total number of unique sequences	Subject specificity*	Core proportion: number of unique sequences	Sum of total number of unique sequences	Core proportion: sum of total number of unique sequences
ELDERMET	118	77.3 ± 7.2	3,981,241	406	1,078,207	85.8%	0.04%	1,581,571	1.8%
ELDERMET controls	9	35.0 ± 6	358,096	2,101	125,616	91.4%	1.7%	149,993	9.3%
Tap et al. (1)	17	37.0 ± 7.2	10,364	56	2,781	85.9%	2.0%	4,053	15.6%
Turnbaugh et al. (2)	30	20 twins (21–32 y) and their mothers[†]	7,664	24	2,271	87.1%	1.1%	3,486	13.7%
Zhang et al. (3)	6	36.2 ± 5.8	869	18	345	83.8%	5.2%	429	14.9%
Li et al. (4)	5	38.2 ± 17.4	4,693	64	917	84.1%	7.0%	1,171	20.2%

*Proportion of unique sequences present in only one subject.

[†] Subject ages not published.

1. Tap J, et al. (2009) Towards the human intestinal microbiota phylogenetic core. *Environ Microbiol* 11:2574–2584.
2. Turnbaugh PJ, et al. (2009) A core gut microbiome in obese and lean twins. *Nature* 457:480–484.
3. Zhang H, et al. (2009) Human gut microbiota in obesity and after gastric bypass. *Proc Natl Acad Sci USA* 106:2365–2370.
4. Li M, et al. (2008) Symbiotic gut microbes modulate human metabolic phenotypes. *Proc Natl Acad Sci USA* 105:2117–2122.

Table S3. Breakdown of phylogenetic classifications and top BLAST hits in the RDP database for unique sequences specific to the elderly subjects in the ELDERMET study compared with core microbiota from younger subjects in this as well as the four published studies

Phylum	Class	Genus	*Clostridium* cluster	Best BLAST hit against species in RDP_v10.16	Identity (%)	BLAST E-value
Bacteroidetes	Bacteroidetes	*Rikenella*		*Alistipes indistinctus*	100	3.0E-119
Bacteroidetes	Bacteroidetes	*Alistipes*		*Alistipes finegoldii*	100	2.0E-123
Bacteroidetes	Bacteroidetes	*Alistipes*		*Alistipes massiliensis*	99.38	1.0E-84
Bacteroidetes	Bacteroidetes	*Alistipes*		*Alistipes onderdonkii*	100	4.0E-125
Bacteroidetes	Bacteroidetes	*Alistipes*		*Alistipes putredinis*	96.43	2.0E-108
Bacteroidetes	Bacteroidetes	*Alistipes*		*Alistipes shahii*	99.55	2.0E-123
Bacteroidetes	Bacteroidetes	*Bacteroides*		*Bacteroides caccae*	100	2.0E-123
Bacteroidetes	Bacteroidetes	*Bacteroides*		*Bacteroides dorei*	100	9.0E-123
Bacteroidetes	Bacteroidetes	*Bacteroides*		*Bacteroides fragilis*	100	9.0E-123
Bacteroidetes	Bacteroidetes	*Bacteroides*		*Bacteroides ovatus*	100	2.0E-123
Bacteroidetes	Bacteroidetes	*Bacteroides*		*Bacteroides thetaiotaomicron*	100	2.0E-123
Bacteroidetes	Bacteroidetes	*Bacteroides*		*Bacteroides uniformis*	100	4.0E-125
Bacteroidetes	Bacteroidetes	*Bacteroides*		*Bacteroides vulgatus*	100	2.0E-123
Bacteroidetes	Bacteroidetes	*Bacteroides*		*Bacteroides xylanisolvens*	100	2.0E-123
Bacteroidetes	Bacteroidetes	*Bacteroides*		*Odoribacter splanchnicus*	99.54	2.0E-120
Bacteroidetes	Bacteroidetes	*Parabacteroides*		*Parabacteroides distasonis*	100	3.0E-113
Bacteroidetes	Bacteroidetes	*Parabacteroides*		*Parabacteroides merdae*	100	2.0E-123
Firmicutes	Clostridia	*Anaerostipes*		*Anaerosporobacter mobilis*	94.98	7.0E-96
Firmicutes	Clostridia	*Lachnospiraceae Incertae Sedis*	XIVa	*Bacteroides xylanolyticus*	98.16	9.0E-123
Firmicutes	Clostridia	*Lachnospiraceae Incertae Sedis*	XIVa	*Blautia luti*	99.55	7.0E-102
Firmicutes	Clostridia	*Lachnospiraceae Incertae Sedis*	XIVa	*Blautia schinkii*	95.93	1.0E-81
Firmicutes	Clostridia	*Peptostreptococcaceae Incertae Sedis*	XI	*Clostridium bartlettii*	100	2.0E-92
Firmicutes	Clostridia	*Lachnospiraceae Incertae Sedis*	XIVa	*Clostridium clostridioforme*	97.7	3.0E-113
Firmicutes	Erysipelotrichi	*Catenibacterium*	XVIII	*Clostridium cocleatum*	92.24	1.0E-81
Firmicutes	Clostridia	*Lachnospiraceae Incertae Sedis*	XIVa	*Clostridium indolis*	94.47	9.0E-123
Firmicutes	Clostridia	*Peptostreptococcaceae Incertae Sedis*	XI	*Clostridium ruminantium*	98.61	2.0E-113
Firmicutes	Clostridia	*Coprococcus*		*Coprococcus catus*	100	2.0E-120
Firmicutes	Clostridia	*Dorea*	XIVa	*Coprococcus comes*	100	9.0E-123
Firmicutes	Clostridia	*Dorea*	XIVa	*Dorea longicatena*	99.55	2.0E-123
Firmicutes	Clostridia	*Lachnospiraceae Incertae Sedis*	XIVa	*Eubacterium eligens*	93.55	2.0E-123
Firmicutes	Clostridia	*Papillibacter*	XIVa	*Eubacterium hallii*	98.18	2.0E-123
Firmicutes	Clostridia	*Roseburia*	XIVa	*Eubacterium rectale*	100	2.0E-120
Firmicutes	Clostridia	*Ruminococcaceae Incertae Sedis*	IV	*Eubacterium siraeum*	100	2.0E-89

Table S3. Cont.

Phylum	Class	Genus	*Clostridium* cluster	Best BLAST hit against species in RDP_v10.16	Identity (%)	BLAST E-value
Firmicutes	Clostridia	*Faecalibacterium*	IV	*Faecalibacterium prausnitzii*	100	2.0E-123
Firmicutes	Clostridia	*Sporobacter*	IV	*Oscillospira guilliermondii*	94.66	9.0E-123
Firmicutes	Clostridia	*Roseburia*	XIVa	*Roseburia faecis*	100	4.0E-125
Firmicutes	Clostridia	*Roseburia*	XIVa	*Roseburia hominis*	100	3.0E-119
Firmicutes	Clostridia	*Roseburia*	XIVa	*Roseburia intestinalis*	100	2.0E-123
Firmicutes	Clostridia	*Roseburia*	XIVa	*Roseburia inulinivorans*	100	6.0E-121
Firmicutes	Clostridia	*Ruminococcus*	IV	*Ruminococcus bromii*	100	6.0E-121
Firmicutes	Clostridia	*Lachnospiraceae Incertae Sedis*	XIVa	*Ruminococcus lactaris*	97.24	2.0E-123
Firmicutes	Clostridia	*Lachnospiraceae Incertae Sedis*	XIVa	*Ruminococcus obeum*	100	2.0E-123
Firmicutes	Clostridia	*Lachnospiraceae Incertae Sedis*	XIVa	*Ruminococcus torques*	99.09	1.0E-106
Firmicutes	Clostridia	*Sporobacter*		*Sporobacter termitidis*	93.63	2.0E-123
Proteobacteria	Betaproteobacteria	*Thiobacter*		*Parasutterella excrementihominis*	100	1.0E-118

Table S4. Breakdown of phylogenetic classifications and top BLAST hits in the RDP database for unique sequences specific to younger subjects compared with core microbiota from younger subjects in this as well as the four published studies

Phylum	Class	Genus	*Clostridium* cluster	Best BLAST hit against species in RDP_v10.16	Identity (%)	BLAST E-value
Bacteroidetes	Bacteroidetes	*Bacteroides*		*Bacteroides vulgatus*	100	4.0E-125
Firmicutes	Clostridia	*Anaerostipes*		*Anaerosporobacter mobilis*	95.07	3.0E-98
Firmicutes	Clostridia	*Lachnospiraceae Incertae Sedis*	XIVa	*Blautia luti*	99.56	1.0E-121
Firmicutes	Clostridia	*Lachnospiraceae Incertae Sedis*	XIVa	*Blautia schinkii*	95.98	1.0E-103
Firmicutes	Clostridia	*Lachnospiraceae Incertae Sedis*	XIVa	*Clostridium clostridioforme*	97.77	1.0E-103
Firmicutes	Erysipelotrichi	*Catenibacterium*	XVIII	*Clostridium cocleatum*	92.38	6.0E-84
Firmicutes	Clostridia	*Sporobacter*	III	*Clostridium termitidis*	98.01	4.0E-125
Firmicutes	Clostridia	*Dorea*		*Coprococcus comes*	100	4.0E-125
Firmicutes	Clostridia	*Dorea*	XIVa	*Dorea formicigenerans*	99.55	2.0E-123
Firmicutes	Clostridia	*Dorea*	XIVa	*Dorea longicatena*	99.56	6.0E-121
Firmicutes	Clostridia	*Lachnospiraceae Incertae Sedis*	XIVa	*Eubacterium hallii*	98.21	9.0E-120
Firmicutes	Clostridia	*Roseburia*	XIVa	*Eubacterium rectale*	100	4.0E-125
Firmicutes	Clostridia	*Subdoligranulum*	IV	*Faecalibacterium prausnitzii*	95.67	4.0E-94
Firmicutes	Clostridia	*Lachnospiraceae Incertae Sedis*	XIVa	*Ruminococcus lactaris*	97.32	8.0E-111
Firmicutes	Clostridia	*Lachnospiraceae Incertae Sedis*	XIVa	*Ruminococcus obeum*	100	4.0E-125