Supplementary Information

Materials and Methods

Sequence generation and analysis

All subjects gave written informed consent before participating in this study, which was approved by the Washington University Human Studies Committee.

We studied 12 men and women (21 to 65 years-old; body mass index (BMI) 30 to 43 kg/m^2) who were randomly assigned to one of two low calorie diets: either a fat-restricted (FAT-R; ~30% of calories from fat) or a carbohydrate-restricted (CARB-R; ~25% of calories from carbohydrates). The recommended caloric intake for women on either diet was 1200-1500 kcal/d, and 1500-1800 kcal/d for men. The total fiber content of both diets was similar (~10-15 g/day). A morning stool sample was collected before and at 12, 26 and 52 weeks after starting diet therapy. Stool was also collected at 0 and 52 weeks from two healthy men (aged 32 and 36; BMI 23 kg/m^2). DNA was extracted from morning stool specimens, and bacterial 16S rRNA gene sequences were generated with bacterial primers using protocols described in ref. 1, with the following modifications: (i) replicate PCR reaction mixtures were pooled, concentrated, purified using a Montage PCR cleanup kit (Millipore), and further purified (1% agarose gel electrophoresis) prior to cloning; (ii) three sequence reads were generated per cloned 16S rRNA gene amplicon using vector-specific primers and the internal primer 907R[2].

16S rRNA gene sequences were edited and assembled as outlined in ref. 1. Sequences were aligned using the nast online alignment tool (http://greengenes.lbl.gov/cgi-bin/nph-index.cgi), and checked for chimeras using

1

Bellerophon[3].Non-chimeric sequences >800bp (n=18, 348) were added to an existing Arb alignment using the parsimony insertion tool[4]. Distance matrices, with Olsen correction, were generated in Arb. DOTUR was used (i) to cluster sequences >1kb (n=16,177) into OTUs by % pair-wise identity (%ID, using a furthest-neighbor algorithm and a precision of 0.01), and (ii) to generate Shannon's diversity index[5]. We used UniFrac[6] to cluster the samples based on an Arb-generated neighbor-joining tree. The alignment of the 18,348-sequence dataset is available at http://gordonlab.wustl.edu/microbial_ecology_human_obesity. Sequences have been deposited in GenBank under accession numbers DQ793220-DQ802819, DQ803048, DQ803139-DQ810181, DQ823640-825343.

Statistical analyses

Analysis of variance was conducted using a model comparison approach[7]. The p-value associated with the correlation coefficient describing the relationship between the change in Bacteroidetes and the change in weight was generated by permutation analysis: values were scrambled randomly and a R^2 generated 10,000 times; the distribution of R^2 values was used to assess the probability of obtaining the observed R^2.

Supplementary references

1. Ley, R. E., *et al.* Obesity alters gut microbial ecology. *Proc. Natl. Acad. Sci. USA* **102,** 11070-11075 (2005).

2. Lane, D. J. in *Nucleic acid techniques in bacterial systematics*. (eds. E. Stackebrandt and M. Goodfellow), 115-175 (John Wiley & Sons, Inc., New York, N.Y., 1991).

3. Huber, T., Faulkner, G. & Hugenholtz, P. Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* **20**, 2317-2319 (2004).

4. Ludwig, W., *et al.* ARB: a software environment for sequence data. *Nucleic Acids Res.* **32**, 1363-1371 (2004).

5. Schloss, P. D. & Handelsman, J. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl. Environ. Microbiol.* **71**, 1501-1506 (2005).

6. Lozupone, C. & Knight. R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* **71,** 8228-8235 (2005).

7. Judd, C. M. & McClelland, G. H. Data Analysis: a model-comparison approach (Harcourt, Brace, Jovanovich, 1989).

Table S1: Sequence prefixes by library, and the number of sequences per library (N)

Subject	Sex	Age	Diet Group	0 weeks		12 weeks		26 weeks		52 weeks	
				Library prefix	N	Library prefix	N	Library prefix	N	Library prefix	N
1	F	57	FAT-R	RL178	541	RL240	327	RL197	202	RL310	328
2	F	53	FAT-R	RL182	178	RL242	296	RL205	277	RL305	346
3	F	54	FAT-R	RL187	803	RL251	287	RL200	335	RL385	274
4	F	48	FAT-R	RL188	579	RL241	287	RL201	310	RL311	244
5	M	55	FAT-R	RL180	855	RL244	312	RL198	189	RL307	309
6	M	55	FAT-R	RL184	877	RL243	306	RL239	289	RL308	235
7	F	42	CARB-R	RL176	543	RL246	236	RL199	309		
8	F	30	CARB-R	RL179	767	RL245	215	RL202	271	RL386	294
9	F	42	CARB-R	RL181	539	RL248	302	RL206	325	RL302	337
10	F	49	CARB-R	RL183	481	RL247	309			RL303	254
11	F	35	CARB-R	RL186	865	RL249	227	RL203	304	RL306	331
12	M	54	CARB-R	RL185	831	RL250	284	RL204	290	RL304	300
13	M	32	CONTROL	RL116	100					RL387	252
14	M	36	CONTROL	RL117	93					RL388	303
										TOTAL	18,348