ARTICLES

An obesity-associated gut microbiome with increased capacity for energy harvest

Peter J. Turnbaugh[1], Ruth E. Ley[1], Michael A. Mahowald[1], Vincent Magrini[2], Elaine R. Mardis[1,2] & Jeffrey I. Gordon[1]

The worldwide obesity epidemic is stimulating efforts to identify host and environmental factors that affect energy balance. Comparisons of the distal gut microbiota of genetically obese mice and their lean littermates, as well as those of obese and lean human volunteers have revealed that obesity is associated with changes in the relative abundance of the two dominant bacterial divisions, the Bacteroidetes and the Firmicutes. Here we demonstrate through metagenomic and biochemical analyses that these changes affect the metabolic potential of the mouse gut microbiota. Our results indicate that the obese microbiome has an increased capacity to harvest energy from the diet. Furthermore, this trait is transmissible: colonization of germ-free mice with an 'obese microbiota' results in a significantly greater increase in total body fat than colonization with a 'lean microbiota'. These results identify the gut microbiota as an additional contributing factor to the pathophysiology of obesity.

The human 'metagenome' is a composite of *Homo sapiens* genes and genes present in the genomes of the trillions of microbes that colonize our adult bodies. The latter genes are thought to outnumber the former by several orders of magnitude[1]. 'Our' microbial genomes (the microbiome) encode metabolic capacities that we have not had to evolve wholly on our own[2,3], but remain largely unexplored. These include degradation of otherwise indigestible components of our diet[4], and therefore may have an impact on our energy balance.

Colonization of adult germ-free mice with a distal gut microbial community harvested from conventionally raised mice produces a dramatic increase in body fat within 10–14 days, despite an associated decrease in food consumption[5]. This change involves several linked mechanisms: microbial fermentation of dietary polysaccharides that cannot be digested by the host; subsequent intestinal absorption of monosaccharides and short-chain fatty acids; their conversion to more complex lipids in the liver; and microbial regulation of host genes that promote deposition of the lipids in adipocytes[5]. These findings have led us to propose that the microbiota of obese individuals may be more efficient at extracting energy from a given diet than the microbiota of lean individuals[2,5].

In a previous study, we performed a comparative 16S-rRNA-gene-sequence-based survey of the distal gut microbiota of adult C57BL/6J mice homozygous for a mutation in the leptin gene (Lep^{ob}) that produces obesity, as well as the microbiota of their lean (*ob*/+ and +/+) littermates[6]. Members of two of the 70 known divisions of Bacteria[7,8], the Bacteroidetes and the Firmicutes, consisted of more than 90% of all phylogenetic types in both groups of mice, just as they do in humans[6,9,10]. However, the relative abundance of the Bacteroidetes in *ob/ob* mice was lower by 50%, whereas the Firmicutes were higher by a corresponding degree[6]. These differences were division-wide, and not attributable to differences in food consumption (a runted *ob/ob* mouse weighed less than his *ob/ob* littermates owing to reduced chow consumption, but still exhibited a markedly greater per cent body fat and ratio of Firmicutes to Bacteroidetes)[6].

We have observed analogous differences in the distal gut microbiota of obese versus lean humans; the relative abundance of Bacteroidetes increases as obese individuals lose weight on either a fat- or a carbohydrate-restricted low-calorie diet. Moreover, the increase in Bacteroidetes was significantly correlated to weight loss but not to total caloric intake[9].

To determine if microbial community gene content correlates with, and is a potential contributing factor to obesity, we characterized the distal gut microbiomes of *ob/ob*, *ob*/+, and +/+ littermates by random shotgun sequencing of their caecal microbial DNA. Mice were used for these comparative metagenomics studies to eliminate many of the confounding variables (environment, diet and genotype) that would make such a proof of principle experiment more difficult to perform and interpret in humans. The caecum was chosen as the gut habitat for sampling because it is an anatomically distinct structure, located between the distal small intestine and colon, that is colonized with sufficient quantities of a readily harvested microbiota for metagenomic analysis. The predicted increased capacity for dietary energy harvest by the *ob/ob* microbiome was subsequently validated using biochemical assays and by transplantation of lean and obese caecal microbiotas into germ-free wild-type mouse recipients. These transplantation experiments illustrate the power of marrying metagenomics to gnotobiotics to discover how microbial communities encode traits that markedly affect host biology.

Shotgun sequencing of microbiomes

Bulk DNA was prepared from the caecal contents of two *ob/ob* and +/+ littermate pairs. A lean *ob*/+ mouse from one of the litters was also studied. All caecal microbial community DNA samples were analysed using a 3730xl capillary sequencer (10,500 ± 431 (s.e.m.) unidirectional reads per data set; 752 ± 13.8 (s.e.m.) nucleotides per read; 39.5 Mb from all five plasmid libraries). Material from one of the two obese and lean sibling pairs was also analysed using a highly parallel 454 Life Sciences GS20 pyrosequencer[11]: three runs for the +/+ mouse (known as lean1), and two runs for its *ob/ob* littermate (ob1) produced a total of 160 Mb of sequence (345,000 ± 23,500 (s.e.m.) unidirectional reads per run; 93.1 ± 1.56 (s.e.m.) nucleotides per read) (Supplementary Tables 1–3). Both sequencing platforms have unique advantages and limitations: capillary sequencing allows more confident gene calling (Supplementary Fig. 1) but is affected by cloning bias, whereas pyrosequencing can achieve higher sequence coverage with no cloning bias, but produces shorter reads

[1]Center for Genome Sciences, and [2]Genome Sequencing Center, Washington University, St. Louis, Missouri 63108, USA.

©2006 Nature Publishing Group

(Supplementary Table 2). The three pyrosequencer runs of the lean1 caecal microbiome (94.9 Mb) yielded 0.44× coverage (on the basis of PROmer sequence alignments[12]) of the 3730xl-derived sequences obtained from the same sample (8.23 Mb), whereas the two pyrosequencer runs of the microbiome of its *ob/ob* littermate (ob1; 65.4 Mb) produced 0.32× coverage of the corresponding 3730xl sequences (8.19 Mb).

Taxonomic analysis of microbiomes

Environmental gene tags (EGTs) are defined as sequencer reads assigned to the NCBI non-redundant, Clusters of Orthologous Groups[13] (COG), or Kyoto Encyclopedia of Genes and Genomes[14] (KEGG) databases (Fig. 1a; Supplementary Fig. 2; Supplementary Table 4). Averaging results from all data sets, 94% of the EGTs assigned to the non-redundant database were bacterial, 3.6% were eukaryotic (0.29% *Mus musculus*; 0.36% fungal), 1.5% were archaeal (1.4% Euryarcheota; 0.07% Crenarcheota), and 0.61% were viral (0.57% double stranded DNA viruses) (Supplementary Table 5). The relative abundance of the eight bacterial divisions identified from EGTs and 16S rRNA gene fragments was comparable to our previous PCR-derived, 16S-rRNA-gene-sequence-based surveys of these caecal samples, including the increased ratio of Firmicutes to Bacteroidetes in obese versus lean littermates (Supplementary Fig. 2). In addition, comparisons of the lean1 and ob1 reads obtained with the pyrosequencer against the finished genome of *Bacteroides thetaiotaomicron* ATCC29148[1], and a deep draft genome assembly of *Eubacterium rectale* ATCC33656 (50% of total contig bases present in contigs ≥75.9 kb; http://gordonlab.wustl.edu/supplemental/Turnbaugh/obob/) provided independent confirmation of the greater relative abundance of Firmicutes in the *ob/ob* microbiota. These organisms were selected for comparison because both are prominently represented in the normal human distal gut microbiota[10] and species related to *B. thetaiotaomicron* (Bacteroidetes division) and *E. rectale* (Firmicutes division) are members of the normal mouse distal gut microbiota[6]. The ratio of sequences homologous to the *E. rectale* versus *B. thetaiotaomicron* genome was 7.3 in the ob1 caecal microbiome compared with 1.5 in the lean1 microbiome.

Intriguingly, there were more EGTs that matched Archaea (Euryarchaeota and Crenarchaeota) in the caecal microbiome of *ob/ob* mice compared with their lean *ob*/+ and +/+ littermates (binomial test of pooled obese versus pooled lean capillary-sequencing-derived microbiomes, $P < 0.001$; Supplementary Table 5).

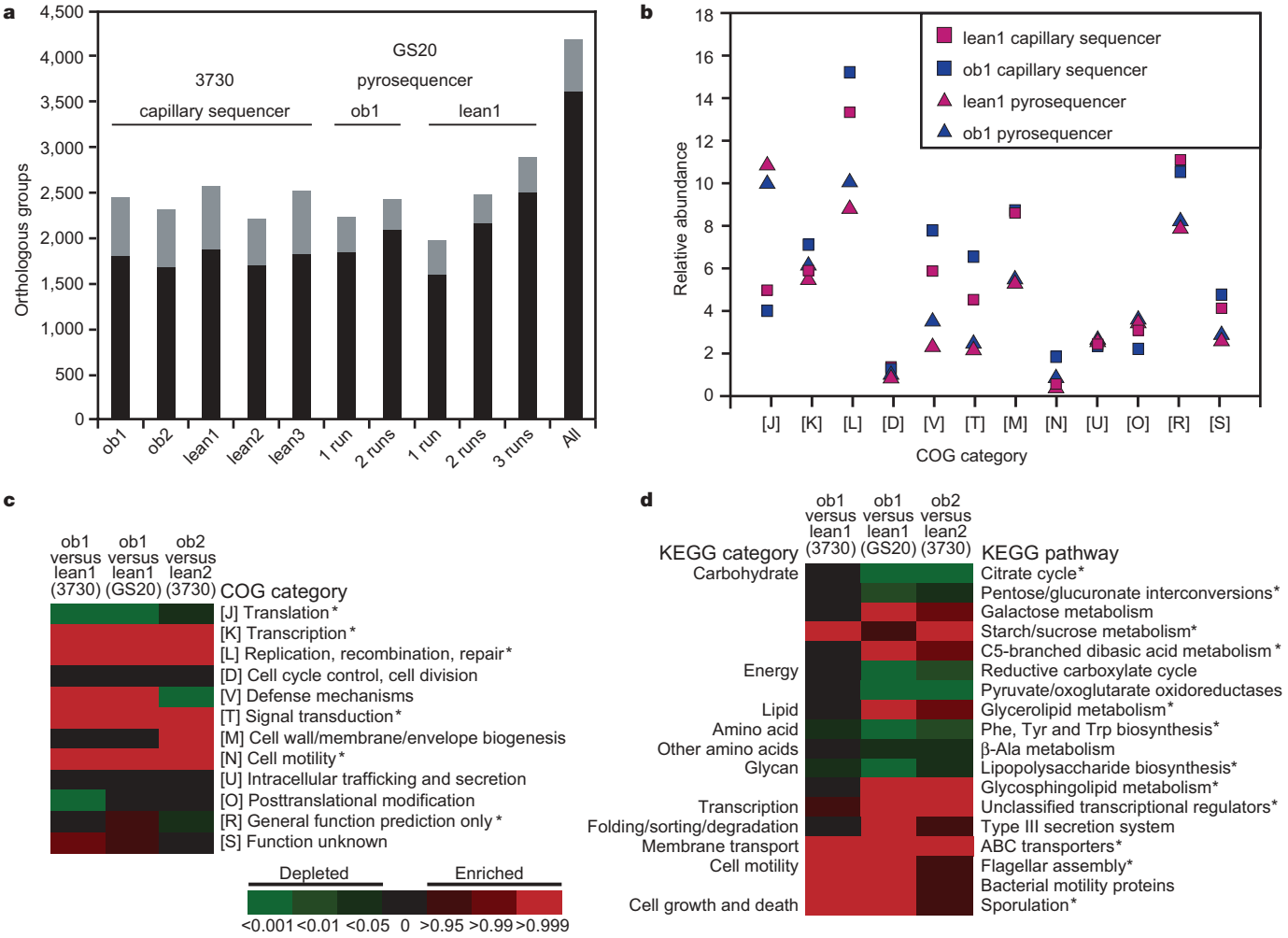

Figure 1 | Comparison of data sets obtained from the caecal microbiomes of obese and lean littermates. **a**, Number of observed orthologous groups in each caecal microbiome. Black indicates the number of observed groups. Grey indicates the number of predicted missed groups. **b**, Relative abundance of a subset of COG categories (BLASTX, *e*-value $< 10^{-5}$) in the lean1 (red) and ob1 (blue) caecal microbiome, characterized by capillary- and pyro-sequencers (squares and triangles, respectively). **c**, **d**, A subset of COG categories (**c**) and all KEGG pathways (**d**) consistently enriched or depleted in the caecal microbiomes of both obese mice compared with their lean littermates. Red denotes enrichment and green indicates depletion on the basis of a cumulative binomial test (brightness indicates the level of significance). Black indicates pathways whose representation is not significantly different. Asterisks indicate groups that were consistently enriched or depleted between both sibling pairs using a more stringent EGT assignment strategy (*e*-value $< 10^{-8}$). For additional details see Supplementary Discussion; Supplementary Figs 5 and 6, and Supplementary Tables 6, 8 and 9.

©2006 Nature Publishing Group

Methanogenic archaea increase the efficiency of bacterial fermentation by removing one of its end products, H_2. Our recent studies of gnotobiotic normal mice colonized with the principal methanogenic archaeon in the human gut, *Methanobrevibacter smithii*, and/or *B. thetaiotaomicron* revealed that co-colonization not only increases the efficiency, but also changes the specificity of bacterial polysaccharide fermentation, leading to a significant increase in adiposity compared with mice colonized with either organism alone[15].

Comparative metagenomic analysis

Using reciprocal TBLASTX comparisons, we found that the Firmicutes-enriched microbiomes from *ob/ob* hosts clustered together, as did lean microbiomes with low Firmicutes to Bacteroidetes ratios (Fig. 2a). Likewise, Principal Component Analysis of EGT assignments to KEGG pathways revealed a correlation between host genotype and the gene content of the microbiome (Fig. 2b).

Reads were then assigned to COGs and KOs (KEGG orthology terms) by BLASTX comparisons against the STRING-extended COG database[13], and the KEGG Genes database[14] (version 37). We tallied the number of EGTs assigned to each COG or KEGG category, and used the cumulative binomial distribution[3], and a bootstrap analysis[16,17], to identify functional categories with statistically significant differences in their representation in both sets of obese and lean littermates. As noted above, capillary sequencing requires cloned DNA fragments; the pyrosequencer does not, but produces relatively short read lengths. These differences are a likely cause of the shift in relative abundance of several COG categories obtained using the two sequencing methods for the same sample (Fig. 1b). Nonetheless, comparisons of the caecal microbiomes of lean versus obese littermates sequenced with either method revealed similar differences in their functional profiles (Fig. 1c).

The *ob/ob* microbiome is enriched for EGTs encoding many enzymes involved in the initial steps in breaking down otherwise indigestible dietary polysaccharides, including KEGG pathways for starch/sucrose metabolism, galactose metabolism and butanoate metabolism (Fig. 1d; Supplementary Fig. 3 and Supplementary Table 6). EGTs representing these enzymes were grouped according to their functional classifications in the Carbohydrate Active Enzymes (CAZy) database (http://afmb.cnrs-mrs.fr/CAZY/). The *ob/ob* microbiome is enriched ($P < 0.05$) for eight glycoside hydrolase families capable of degrading dietary polysaccharides including starch (CAZy families 2, 4, 27, 31, 35, 36, 42 and 68, which contain α-glucosidases, α-galactosidases and β-galactosidases). Finished genome sequences of prominent human gut Firmicutes have not been reported. However, our analysis of the draft genome of *E. rectale* has revealed 44 glycoside hydrolases, including a significant enrichment for glycoside hydrolases involved in the degradation of dietary starches (CAZy families 13 and 77, which contain α-amylases and amylomaltases; $P < 0.05$ on the basis of a binomial test of *E. rectale* versus the finished genomes of Bacteroidetes—*Bacteroides thetaiotaomicron* ATCC29148, *B. fragilis* NCTC9343, *B. vulgatus* ATCC8482 and *B. distasonis* ATCC8503).

EGTs encoding proteins that import the products of these glycoside hydrolases (ABC transporters), metabolize them (for example, α- and β-galactosidases KO7406/7 and KO1190, respectively), and generate the major end products of fermentation, butyrate and acetate (pyruvate formate-lyase, KO0656, and other enzymes in the KEGG 'Butanoate metabolism' pathway; and formate-tetrahydrofolate ligase, KO1938, the second enzyme in the homoacetogenesis pathway for converting CO_2 to acetate) are also significantly enriched in the *ob/ob* microbiome (binomial comparison of pyrosequencer-derived ob1 and lean1 data sets, $P < 0.05$) (Fig. 1d; Supplementary Fig. 3 and Supplementary Table 6).

As predicted from our comparative metagenomic analyses, the *ob/ob* caecum has an increased concentration of the major fermentation end-products butyrate and acetate (Fig. 3a). This observation is also consistent with the fact that many Firmicutes are butyrate producers[18–20]. Moreover, bomb calorimetry revealed that *ob/ob* mice have significantly less energy remaining in their faeces relative to their lean littermates (Fig. 3b).

Microbiota transplantation

We performed microbiota transplantation experiments to test directly the notion that the *ob/ob* microbiota has an increased capacity to harvest energy from the diet and to determine whether increased adiposity is a transmissible trait. Adult germ-free C57BL/6J mice were colonized (by gavage) with a microbiota harvested

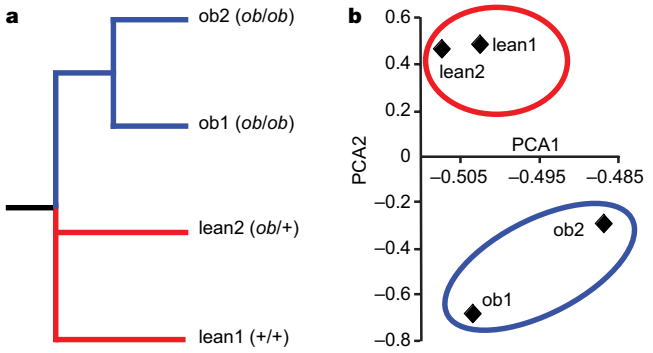

Figure 2 | Microbiomes cluster according to host genotype. a, Clustering of caecal microbiomes of obese and lean sibling pairs based on reciprocal TBLASTX comparisons. All possible reciprocal TBLASTX comparisons of microbiomes (defined by capillary sequencing) were performed from both lean and obese sibling pairs. A distance matrix was then created using the cumulative bitscore for each comparison and the cumulative score for each self–self comparison. Microbiomes were subsequently clustered using NEIGHBOUR (PHYLIP version 3.64). **b**, Principal Component Analysis (PCA) of KEGG pathway assignments. A matrix was constructed containing the number of EGTs assigned to each KEGG pathway in each microbiome (includes KEGG pathways with >0.6% relative abundance in at least two microbiomes, and a standard deviation >0.3 across all microbiomes), PCA was performed using Cluster3.0 (ref. 25), and the results graphed along the first two components.

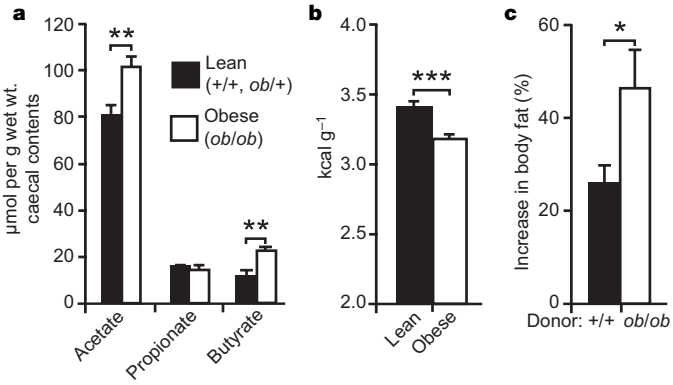

Figure 3 | Biochemical analysis and microbiota transplantation experiments confirm that the *ob/ob* microbiome has an increased capacity for dietary energy harvest. a, Gas-chromatography mass-spectrometry quantification of short-chain fatty acids in the caeca of lean ($n = 4$) and obese ($n = 5$) conventionally raised C57BL/6J mice. **b**, Bomb calorimetry of the faecal gross energy content (kcal g^{-1}) of lean (+/+, *ob*/+; $n = 9$) and obese (*ob/ob*; $n = 13$) conventionally raised C57BL/6J mice. **c**, Colonization of germ-free wild-type C57BL/6J mice with a caecal microbiota harvested from obese donors (*ob/ob*; $n = 9$ recipients) results in a significantly greater percentage increase in total body fat than colonization with a microbiota from lean donors (+/+; $n = 10$ recipients). Total body fat content was measured before and after a two-week colonization, using dual-energy X-ray absorptiometry. Mean values ± s.e.m. are plotted. Asterisks indicate significant differences (two-tailed Student's *t*-test of all datapoints, $^{*}P < 0.05$, $^{**}P \leq 0.01$, $^{***}P < 0.001$).

©2006 Nature Publishing Group

from the caecum of obese (*ob/ob*) or lean (+/+) donors (1 donor and 4–5 germ-free recipients per treatment group per experiment; two independent experiments). 16S-rRNA-gene-sequence-based surveys confirmed that the *ob/ob* donor microbiota had a greater relative abundance of Firmicutes compared with the lean donor microbiota (Supplementary Fig. 4 and Supplementary Table 7). Furthermore, the *ob/ob* recipient microbiota had a significantly higher relative abundance of Firmicutes compared with the lean recipient microbiota ($P < 0.05$, two-tailed Student's *t*-test). UniFrac analysis[21] of 16S rRNA gene sequences obtained from the recipients' caecal microbiotas revealed that they cluster according to the input donor community (Supplementary Fig. 4): that is, the initial colonizing community structure did not exhibit marked changes by the end of the two-week experiment. There was no statistically significant difference in (1) chow consumption over the 14-day period (55.4 ± 2.5 g (*ob/ob*) versus 54.0 ± 1.2 g (+/+); caloric density of chow, 3.7 kcal g^{-1}), (2) initial body fat (2.7 ± 0.2 g for both groups as measured by dual-energy X-ray absorptiometry), or (3) initial weight between the recipients of lean and obese microbiotas. Strikingly, mice colonized with an *ob/ob* microbiota exhibited a significantly greater percentage increase in body fat over two weeks than mice colonized with a +/+ microbiota (Fig. 3c; 47 ± 8.3 versus 27 ± 3.6 percentage increase or 1.3 ± 0.2 versus 0.86 ± 0.1 g fat (dual-energy X-ray absorptiometry): at 9.3 kcal g^{-1} fat, this corresponds to a difference of 4 kcal or 2% of total calories consumed).

Discussion

The primary cause of obesity in the *ob/ob* mouse model is increased food consumption due to leptin deficiency. We have used this model to provide direct experimental evidence that at least one type of obesity-associated gut microbiome has an increased capacity for energy harvest from the diet. This finding provides support for the more general concept that the gut microbiome should be considered as a set of genetic factors that, together with host genotype and lifestyle (energy intake and expenditure), contribute to the pathophysiology of obesity. Yet to be answered are the questions of what mechanisms are responsible for mediating the linkage between the relative abundance of the Bacteroidetes to Firmicutes divisions and adiposity in both mice and humans, and to what extent is this relationship self-perpetuating?

Energy balance is an equilibrium between the amount of energy taken in as food and the amount expended during resting metabolism, as well as the thermic effect of food, physical activity, and loss in the faeces and urine. The alteration in efficiency of energy harvest from the diet produced by changes in gut microbial ecology does not have to be great to contribute to obesity, given that small changes in energy balance, over the course of a year, can result in significant changes in body weight[22]. We are aware of only one report showing that obese humans may have an increased capacity to absorb energy from their diet: in this case, analysis of four lean and four obese individuals given three different diets (high protein/high fat, 'average', and high carbohydrate) revealed that on average, the obese individuals lost less energy to stool compared with their lean counterparts. However, the differences did not achieve statistical significance[23].

Our study in mice demonstrates the feasibility and utility of applying comparative metagenomics to mouse models of human physiologic or pathophysiologic states in order to understand the complex interplay between host genetics, microbial community gene content and the biological properties of the resulting 'superorganism'. As such, it opens the door for comparable investigations of the interactions between humans and their microbial communities, including whether there is a core set of genes associated with the microbiomes of obese versus lean individuals; whether these genes are transmitted from mothers to their offspring; what genetic or behavioural traits of the host can reshape the community[24]; how the microbiome changes as body mass index changes within an individual; the degree to which these changes correlate with energy harvest from their diets; and whether germ-free mice can be used as a bioassay to compare the energy harvesting activities encoded in human gut microbiomes. Our results indicate that if the gut microbiome of obese humans is comparable to that of obese mice, then it may be a biomarker, a mediator and a new therapeutic target for people suffering from this increasingly worldwide disease.

METHODS

DNA was isolated from the caeca of *ob/ob*, *ob/+* and +/+ littermates using a bead beater to mechanically disrupt cells, followed by phenol–chloroform extraction. DNA was sequenced using 3730xl capillary- and GS20 pyro-sequencers: in the case of the latter, DNA was purified further using the Qiaquick gel extraction kit (Qiagen).

Individual reads in the resulting 199.8 Mb data set were directly compared with each other and to reference sequenced gut microbial genomes using MUMmer[12]. Taxonomic assignments were made on the basis of BLASTX searches of the non-redundant database (*e*-value $< 10^{-5}$) and alignment of 16S gene fragments. Reads were also assigned to EGTs (environmental gene tags) by BLASTX searches against the non-redundant database, STRING-extended COG[13], and KEGG[14] (v37) databases. Microbiomes from each animal were clustered according to reciprocal TBLASTX comparisons and their EGT assignments to KEGG pathways. Statistically enriched or depleted COG and KEGG groups were identified using bootstrap[16,17] and cumulative binomial[3] analyses. For the binomial analysis, the probability of observing 'n_1' EGT assignments to a given group in microbiome 1, given 'N_1' EGT assignments to all groups in microbiome 1, was calculated using the cumulative binomial distribution and an expected probability equal to 'n_2/N_2' (the number of EGTs assigned to a given group in microbiome 2 divided by the total number of EGTs assigned to all groups in microbiome 2). Detailed descriptions of these methods and techniques for (1) measuring short-chain fatty acids in caecal samples by gas-chromatography mass-spectrometry, (2) bomb calorimetry of faecal samples, (3) transplanting the caecal microbiota of C57BL/6J *ob/ob* or +/+ donors into 8–9-week-old germ-free +/+ C57BL/6J recipients, (4) measuring the total body fat of transplant recipients, before and after colonization, by dual-energy X-ray absorptiometry, and (5) performing 16S-rRNA-gene-sequence-based surveys of the input (donor) and output (recipient) caecal microbiotas are provided in Supplementary Information.

Received 8 October; accepted 7 November 2006.

1. Xu, J. *et al.* A genomic view of the human–*Bacteroides thetaiotaomicron* symbiosis. *Science* **299,** 2074–2076 (2003).
2. Backhed, F., Ley, R. E., Sonnenburg, J. L., Peterson, D. A. & Gordon, J. I. Host–bacterial mutualism in the human intestine. *Science* **307,** 1915–1920 (2005).
3. Gill, S. R. *et al.* Metagenomic analysis of the human distal gut microbiome. *Science* **312,** 1355–1359 (2006).
4. Sonnenburg, J. L. *et al.* Glycan foraging *in vivo* by an intestine-adapted bacterial symbiont. *Science* **307,** 1955–1959 (2005).
5. Backhed, F. *et al.* The gut microbiota as an environmental factor that regulates fat storage. *Proc. Natl Acad. Sci. USA* **101,** 15718–15723 (2004).
6. Ley, R. E. *et al.* Obesity alters gut microbial ecology. *Proc. Natl Acad. Sci. USA* **102,** 11070–11075 (2005).
7. Ley, R. E. *et al.* Unexpected diversity and complexity of the Guerrero Negro hypersaline microbial mat. *Appl. Environ. Microbiol.* **72,** 3685–3695 (2006).
8. Ley, R. E., Peterson, D. A. & Gordon, J. I. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* **124,** 837–848 (2006).
9. Ley, R. E., Turnbaugh, P. J., Klein, S. & Gordon, J. I. Human gut microbes associated with obesity. *Nature* doi:10.1038/nature4441023a (this issue).
10. Eckburg, P. B. *et al.* Diversity of the human intestinal microbial flora. *Science* **308,** 1635–1638 (2005).
11. Margulies, M. *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437,** 376–380 (2005).
12. Kurtz, S. *et al.* Versatile and open software for comparing large genomes. *Genome Biol.* **5,** R12 (2004).
13. von Mering, C. *et al.* STRING: known and predicted protein–protein associations, integrated and transferred across organisms. *Nucleic Acids Res.* **33,** D433–D437 (2005).
14. Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y. & Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **32,** D277–D280 (2004).
15. Samuel, B. S. & Gordon, J. I. A humanized gnotobiotic mouse model of host–archaeal–bacterial mutualism. *Proc. Natl Acad. Sci. USA* **103,** 10011–10016 (2006).
16. DeLong, E. F. *et al.* Community genomics among stratified microbial assemblages in the ocean's interior. *Science* **311,** 496–503 (2006).
17. Rodriguez-Brito, B., Rohwer, F. & Edwards, R. An application of statistics to comparative metagenomics. *BMC Bioinformatics* **7,** 162 (2006).
18. Duncan, S. H., Hold, G. L., Barcenilla, A., Stewart, C. S. & Flint, H. J. *Roseburia intestinalis* sp. nov., a novel saccharolytic, butyrate-producing bacterium from human faeces. *Int. J. Syst. Evol. Microbiol.* **52,** 1615–1620 (2002).

©2006 Nature Publishing Group

19. Barcenilla, A. *et al.* Phylogenetic relationships of butyrate-producing bacteria from the human gut. *Appl. Environ. Microbiol.* **66,** 1654–1661 (2000).
20. Pryde, S. E., Duncan, S. H., Hold, G. L., Stewart, C. S. & Flint, H. J. The microbiology of butyrate formation in the human colon. *FEMS Microbiol. Lett.* **217,** 133–139 (2002).
21. Lozupone, C., Hamady, M. & Knight, R. UniFrac—an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics* **7,** 371 (2006).
22. Flegal, K. M. & Troiano, R. P. Changes in the distribution of body mass index of adults and children in the US population. *Int. J. Obes. Relat. Metab. Disord.* **24,** 807–818 (2000).
23. Webb, P. & Annis, J. F. Adaptation to overeating in lean and overweight men and women. *Hum. Nutr. Clin. Nutr.* **37,** 117–131 (1983).
24. Rawls, J. F., Mahowald, M. A., Ley, R. E. & Gordon, J. I. Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection. *Cell* **127,** 423–433 (2006).
25. de Hoon, M. J., Imoto, S., Nolan, J. & Miyano, S. Open source clustering software. *Bioinformatics* **20,** 1453–1454 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank our colleagues B. Samuel, F. Backhed, D. O'Donnell, M. Karlsson, M. Hickenbotham, K. Haub, L. Fulton, J. Crowley, T. Coleman, C. Semenkovich, V. Markowitz and E. Szeto for their assistance. This work was supported by grants from the NIH and the W.M. Keck Foundation.

Author Information This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the project accession AATA00000000–AATF00000000. The version described in this paper is the first version, AATA01000000–AATF01000000. All 454 GS20 reads have been deposited in the NCBI Trace Archive. PCR-derived 16S rRNA gene sequences are deposited in GenBank under the accession numbers EF95962-100118. Annotated sequences are also available for further analysis in IMG/M (http://img.jgi.doe.gov/m). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.I.G. (jgordon@wustl.edu).

©2006 Nature Publishing Group