Diet Drives Convergence in Gut Microbiome Functions Across Mammalian Phylogeny and Within Humans
Brian D. Muegge, et al.
Science 332, 970 (2011);
DOI: 10.1126/science.1198719

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 9, 2011):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/332/6032/970.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2011/05/18/332.6032.970.DC1.html
http://www.sciencemag.org/content/suppl/2011/05/19/332.6032.970.DC2.html

A list of selected additional articles on the Science Web sites related to this article can be found at:
http://www.sciencemag.org/content/332/6032/970.full.html#related

This article cites 29 articles, 16 of which can be accessed free:
http://www.sciencemag.org/content/332/6032/970.full.html#ref-list-1

This article has been cited by 2 articles hosted by HighWire Press; see:
http://www.sciencemag.org/content/332/6032/970.full.html#related-urls

This article appears in the following subject collections:
Microbiology
http://www.sciencemag.org/cgi/collection/microbio
Diet Drives Convergence in Gut Microbiome Functions Across Mammalian Phylogeny and Within Humans

Brian D. Muegge,1 Justin Kuczynski,2 Dan Knights,3 Jose C. Clemente,4 Antonio González,5 Luigi Fontana,6 Bernard Henriot,6 Rob Knight,2,7 Jeffrey I. Gordon1,*

Coevolution of mammals and their gut microbiota has profoundly affected their radiation into myriad habitats. We used shotgun sequencing of microbial community DNA and targeted sequencing of bacterial 16S ribosomal RNA genes to gain an understanding of how microbial communities adapt to extremes of diet. We sampled fecal DNA from 33 mammalian species and 18 humans who kept detailed diet records, and we found that the adaptation of the microbiota to diet is similar across different mammalian lineages. Functional repertoires of microbiome genes, such as those encoding carbohydrate-active enzymes and proteases, can be predicted from bacterial species assemblages. These results illustrate the value of characterizing vertebrate gut microbiomes to understand host evolutionary histories at a supraorganismal level.

2,163,286 reads [mean = 55,469 ± 28,724 (SD) per sample; 261 ± 83 nt per read] (table S3) (2). Shotgun reads were functionally annotated using three databases: KEGG (for KEGG Orthology (KO) groups and Enzyme Commission (E.C.) numbers), CAZY (for carbohydrate-active enzymes), and MEROPS (for peptidases) (3–5). When shotgun reads were assigned to phylogenetic bins using the program MEGAN (6), the results revealed that fecal microbiomes were dominated by members of Bacteria, had low levels of Eukarya (0.15 to 5.35% of identifiable reads), and archaeons were variably represented (0 to 1.77% of assignable reads, with none detected in any carnivore microbe). Seventeen samples had reads assigned to known viruses (table S4) (2).

Procrustes analysis (least-squares orthogonal mapping) was used to test whether the functional properties of a microbiome can be predicted from the bacterial species that compose it (2). Procrustes analysis attempts to stretch and rotate the points in one matrix, such as points obtained by principal coordinates analysis (PCoA), to be as close as possible to points in the other matrix, thus preserving the relative distances between points within each matrix (7, 8) (Fig. 1A). We first took the 16S rRNA data set and used the UniFrac metric to compare the overlap between each pair of communities in terms of their evolutionary distance (9). The similarity in functional profiles was then determined using the Bray-Curtis distance metric applied to KO groups, E.C.s, CAZyomes, or peptidases. Principal-coordinates reduction was performed separately on the 16S rRNA and annotated shotgun (microbiome) data sets, and the point clouds were aligned using Procrustes. For each comparison, the goodness of fit, or $M^2$ value, of the transformed data sets was measured over the first three dimensions. The statistical significance of the goodness of the fit was measured by a Monte Carlo label permutation approach (2).

The agreement between phylogenetic and functional measurements was remarkable for all mammals, regardless of their diet, host lineage, or gut physiology. Figure 1, B to E, shows how the goodness of fit was robust to different functional databases. The analysis was also robust to taxonomy or phylogenetic-based species classification, weighted or unweighted metrics, and whether one or more
member of each mammalian species was considered (fig. S1). For both bacterial 16S rRNA and whole community gene data sets, the PCoA plots separated carnivores and omnivores from herbivores, emphasizing the importance of diet in differentiating gut microbial communities (P < 0.05) (2). Our previous study using full-length 16S rRNA sequences revealed that the fecal microbiota of conspecics were significantly more similar than the communities of different host species (7). The V2 16S rRNA data generated in this study confirmed this result, using both weighted and unweighted UniFrac distances (P < 0.05 by 1000 Monte Carlo permutations) (2).

The Procrustes results prompted us to use a nearest-neighbor model to test whether the functional configuration of a microbiome could be predicted from its 16S rRNA sequences. Using a fecal sample’s nearest neighbor, as defined by unweighted or weighted UniFrac, to predict the sample’s functional profile generated a significantly better prediction than a random neighbor; this was true for KOs, E.C.s, CAZymes, and peptidases (P < 0.0001, 106 Monte Carlo permutations) (2).

The concordance of diet and microbiome structure and function raises the question of whether it is caused primarily by coevolution between mammals and their gut microbiota or microbiome, or by the many parallel dietary shifts that have occurred over the course of mammalian evolution (10). We tested which of these hypotheses, which have traditionally been viewed as competing but need not be mutually exclusive, were supported by looking for congruence between mammalian phylogeny and subsets of bacterial species, KOs, CAZymes, peptidases, or other enzymatic activities. Briefly, the mammalian phylogenetic tree defines sets of organisms that are monophyletic; that is, groups containing all and only the descendants of a common ancestor. We reasoned that if bacterial taxa or functions originated rarely, then these taxa or functions should be vertically transmitted during mammalian speciation. Therefore, there should be more cases in which a given taxon or function occurred in all members of a monophyletic mammalian group than chance would predict. Using this analytic approach (2), we found

### Table 1. Overview of mammals in this study. Sample abbreviations used in figures and tables are noted in parentheses. See table S1 for additional details.

<table>
<thead>
<tr>
<th>Herbivores</th>
<th>Carnivores</th>
<th>Omnivores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bighorn sheep 1 (BigHornSD)*</td>
<td>Armadillo (Armadillo)†</td>
<td>Baboon 1 (BaboonSTL)†</td>
</tr>
<tr>
<td>Bighorn sheep 2 (BigHornW)†</td>
<td>Bush dog (BushDog1)†</td>
<td>Baboon 2 (BaboonW)†</td>
</tr>
<tr>
<td>Colobus (Colobus)†</td>
<td>Echidna (Echidna)†</td>
<td>Black bear (BlackBr2)†</td>
</tr>
<tr>
<td>Gazelle (Gazelle3)†</td>
<td>Hyena (Hyena2)†</td>
<td>Black lemur (BlackLemur)†</td>
</tr>
<tr>
<td>Giraffe (Giraffe2)†</td>
<td>Lion 1 (Lion1)†</td>
<td>Callimicos (Callimicos)†</td>
</tr>
<tr>
<td>Rock hyrax 1 (HyraxSD)*</td>
<td>Lion 2 (Lion2)†</td>
<td>Chimpanzee 1 (Chimp1)†</td>
</tr>
<tr>
<td>Rock hyrax 2 (HyraxSTL)†</td>
<td>Polynesian armadillo (PolBr2)†</td>
<td>Chimpanzee 2 (Chimp2)†</td>
</tr>
<tr>
<td>Kangaroo (Kroo3)†</td>
<td></td>
<td>Ringtailed lemur (RTLemur)†</td>
</tr>
<tr>
<td>Okapi 1 (Okapi1)†</td>
<td></td>
<td>Saki (Saki)†</td>
</tr>
<tr>
<td>Okapi 2 (Okapi2)†</td>
<td></td>
<td>Spectacled bear (SpecBr2)†</td>
</tr>
<tr>
<td>Springbok (SpgbkW)†</td>
<td></td>
<td>Squirrel (Squirrel)†</td>
</tr>
<tr>
<td>Transcaspian Ursal sheep (Ursal2)†</td>
<td></td>
<td>Zebra (ZebraSTL1)†</td>
</tr>
<tr>
<td>Visayan warty pig (VWPig)*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*San Diego Zoological Park. †St. Louis Zoo. ‡Wild, free-living.
that the overall distribution of microbial species and microbiome functions in the gut does not mirror mammalian phylogeny. 198 different named bacterial genera were detected in our data set; of these, only 3 were significantly associated with the mammalian phylogenetic tree more than would be expected by chance (Prevotella, Barnesiella, and Bacteroides). No CAZymes or peptidases and only 18 of the 3866 KOs tested were associated with host phylogeny. We repeated the analysis using a more relaxed constraint that a taxon or function occurs in a given monophyletic group more frequently than expected by chance rather than requiring strict presence or absence agreement (2). The relaxed definition gave similar results; only three additional genera and a total of 90 KOs were detected as having a significant association with the mammalian tree. We concluded that bacterial taxa and functions are evolutionarily labile and do not explain the concordance between bacterial communities and microbiome functions.

Bipartite network analysis provided an additional tool for exploring the interrelationship between host diet, host lineage, gut physiology, and shared and unique bacteria taxa (1). Mammalian hosts and bacterial OTUs were used as nodes in a bipartite graph, with edges connecting OTU nodes to the hosts in which they are found (2). Using 1900 V2 16S rRNA sequences from each mammalian host, the network shows clear separation of fecal communities by host diet (Fig. 2A), mirroring our earlier results based on smaller numbers of full-length 16S rRNA sequences (7).

We reasoned that the bipartite graph approach could also be used to connect mammalian samples to individual microbial gene functions from shotgun reads. The power of the bipartite graph approach is to represent both genes and mammalian species explicitly as nodes, thus visualizing which genes connect with which species. The clear separation by diet disappears when we consider gene functions (Fig. 2B and fig. S2), suggesting that rather than a diet- or physiology-specific set of genes, the relationship among mammalian gut microbiomes is that they share a large core repertoire of functions. We confirmed this result by plotting the frequency of shared taxa in the 39 mammalian fecal samples, and also species- and genus-level OTU bins (2). All of the curves demonstrate an essentially exponential decay as successive samples are added, with no OTUs found in more than 30 samples (Fig. 2C). However, the plot of KO frequency flattens out, with 35 KOs found in all samples. This effect cannot be due to differences in the number of OTUs relative to KOs: There are more OTUs than KOs, and fewer assigned species or assigned genera, yet all the taxonomic curves show the same rapid decay, unlike the KOs.

This result does not imply that there are no differences among the functional configurations of microbiomes of host species having different diets. Rather, it suggests that the differences between microbiomes probably stem from differing abundances of shared functions, such as enzymes that break down chemical substrates in the host diet. We identified 495 E.C.s with significantly different proportional abundance in the 7 carnivorous and 21 herbivorous mammalian microbiomes, using the program Shotgun FunctionalyzeR (adjusted P < 0.001 after multiple hypothesis correction) (table S5) (11). Many of the enzymes distinguishing carnivorous and herbivorous fecal microbiomes are involved in amino acid metabolism. Microbiomes from herbivores were enriched in enzymes that map to biosynthetic reactions for 12 amino acids, whereas no carnivore samples were enriched in amino acid bio-

![OTU network](A) and KO network](B)

**Fig. 2.** Mammalian gut bacterial communities share a functional core. (A and B) Bipartite network diagrams of evenly sampled bacterial 16S rRNA–derived OTUs (A) or KOs (B). Edges connecting mammalian nodes (circles) to species-level OTUs or KOs found in that sample are colored by host diet. Sample labels are removed from the KO diagram for legibility (a high-resolution image of removed labels is presented in fig. S2). (C) Mammalian gut communities share a core suite of KOs. Using evenly subsampled OTU or KO data sets, the distribution of counts is plotted as a function of the number of mammalian host microbiomes where the KO or phylotype was detected. The results demonstrate exponential decay for the 16S rRNA data (OTU, species, and genus), with no OTU or bacterial species found in all samples, although a core set of KOs is detectable in all fecal communities sampled.
enzymes catalyzing OAA production from pyruvate or PEP are significantly increased in the carnivore microbiomes, whereas the reverse reactions are catalyzed by enzymes whose representation is increased in herbivore microbiomes.

Our studies comparing mammalian species revealed a relationship between host diet and gut microbial community structure and function. We next asked whether similar trends could be detected using diet variation within a single free-living host species, namely humans. Quantitative studies of diet in most human populations are complicated by the known inaccuracy of self-reported data (13), so we turned to a group of adults known to keep meticulous records about their daily food composition and consumption. The selected cohort consisted of 18 lean members of the Caloric Restriction Society who typically measure and record all components of their diets on a daily basis with computer software to insure optimal nutrition despite reduced energy intake (14, 15). We collected their dietary records for a 4-day period (conservatively encompassing at least one complete intestinal transit time) before obtaining a single fecal sample, and analyzed macro- and micro-nutrient consumption using a validated protocol (2, 16). An average of 3642 ± 3826 bacterial V2 16S rRNA reads and 54,295 ± 28,086 shotgun reads were obtained per sample (tables S7 to S10).

Procrustes analysis revealed a significant association between the bacterial phylogenetic structure of their fecal communities (16S rRNA) and the functions encoded in their microbiomes \( P < 0.05 \) for KOs, E.C.s, and CAZymes (glucoside hydrolases); not significant for peptidases \( (P = 0.061) \) (fig. S3). These results suggest that the processes that drive the functional differentiation of microorganisms within an individual host species may be fundamentally similar to those that drive their differentiation across mammalian evolution.

Documentation of the weight of each ingredient in each meal consumed by these individuals (table S7) allowed us to perform a follow-up analysis examining the impact on fecal bacterial community structure of three dietary components (total protein, carbohydrate, and insoluble fiber intake). We chose these diet categories because protein intake is markedly different between carnivores and herbivores, and because an extensive literature exists about the impact of ingested polysaccharides and fiber on the gut microbiota (17). Linear regression of the three dietary categories against the position of each individual’s microbiome along principal coordinate 1 of the PCoA plots revealed that total protein intake was significantly associated with KO data [adjusted linear regression coefficient of determination \( (R^2) \) value = 0.307, adjusted \( P \) value = 0.030] (2). In contrast, insoluble dietary fiber was significantly associated with bacterial OTU content (Bray-Curtis metric; adjusted \( R^2 \) value = 0.371; adjusted \( P \) value = 0.013) (table S11). These results confirm that within a single free-living species, both the structure and function of the gut microbiome are significantly associated with dietary intake.

Taken together with our prior work (1), these results teach us that even fecal samples from mammals living in zoos and human samples from a single self-selected population can provide insights into the factors driving the evolution of the gut microbiome. They also compel us, at a time when complete genomes are to be sequenced for 10,000 vertebrates (18), to take the next step and perform systematic studies that rigorously test specific mechanisms that drive the evolution of hosts and their (gut) microbial symbionts. These studies should be guided by experts who can choose taxa that radiated at different points in their evolutionary history, with parallel shifts in their diet, morphology, biogeography, or other key factors known or hypothesized to influence evolution. The results should help address questions such as what functional features in host intestinal environments (including the biochemical characteristics of mucosal surfaces) are related to the representation of specific bacterial taxa and microbiome functions, and how readily microbial populations have been acquired and reacquired during the course of vertebrate evolution. Additionally, our findings emphasize the need to sample humans across the globe with a variety of extreme diets and lifestyles, including relatively ancestral hunter-gatherer lifestyles, in order to provide new insights into the limits of variation within a host species and the possibility that our microbes, in coevolving with our bodies and our cultures, have helped shape our physiological differences and environmental adaptations.

References and Notes
2. See supporting material on Science Online.
18. Genome 10K Community of Scientists, J. Hered. 100, 659 (2009).

Acknowledgments: We thank J. Manchester and S. Wagoner for technical assistance; B. Cantarel, V. Lombard, C. Rancurel, and P. Coutinho for CAZyme annotation; R. Ley and members of the Gordon lab for their suggestions; and S. Bircher, R. Ramey, M. Schlegel, M. Schrenzel, T. Tucker, and P. Tumbaru for past help in procuring mammalian fecal samples. This work was supported by grants from www.sciencemag.org
The Toll-Like Receptor 2 Pathway Establishes Colonization by a Commensal of the Human Microbiota

June L. Round,1,2,10* S. Melanie Lee,3 Jennifer Li,4 Gloria Tran,4 Bana Jabri,4 Talal A. Chatila,3 Sarkis K. Mazmanian1,2

Mucosal surfaces constantly encounter microbes. Toll-like receptors (TLRs) mediate recognition of microbial patterns to eliminate pathogens. By contrast, we demonstrate that the prominent gut commensal Bacteroides fragilis activates the TLR pathway to establish host-microbial symbiosis. TLR2 on CD4+ T cells is required for B. fragilis colonization of a unique mucosal niche in mice during homeostasis. A symbiosis factor (PSA, polysaccharide A) of B. fragilis signals through TLR2 directly on Foxp3+ regulatory T cells to promote immunologic tolerance. B. fragilis lacking PSA is unable to restrain T helper 17 cell responses and is defective in niche-specific mucosal colonization. Therefore, commensal bacteria exploit the TLR pathway to actively suppress immunity. We propose that the immune system can discriminate between pathogens and the microbiota through recognition of symbiotic bacterial molecules in a process that engenders commensal colonization.

Throughout our lives, we continuously encounter microorganisms that range from those essential for health to those causing death (1). Consequently, our immune system is charged with the critical task of distinguishing between beneficial and pathogenic microbes. Toll-like receptors (TLRs) are evolutionarily conserved molecules that promote immune responses, and TLR signaling by innate immune cells is indispensable for proper activation of the immune system during infections. T cells also express functional TLRs (2–4), and TLR signaling has furthermore been shown to restrain immune responses (5). As symbionts and pathogens produce similar molecular patterns that are sensed by TLRs, the mechanisms by which our immune system differentiates between the microbiota and enteric infections remain unknown.

Whereas the intestinal microbiota contains hundreds of bacterial species and is integral to human health (6), the mucosal immune system employs an arsenal of responses to control enteric pathogens. Germ-free mice lack proinflammatory T helper 17 (Th17) cells in the gut (7, 8) (Fig. 1A), and only select symbiotic bacteria can induce Th17 cells (9, 10). Most microbes express common TLR ligands (e.g., peptidoglycan, unmethylated CpG, and lipopolysaccharides); therefore, how do symbionts avoid triggering intestinal immunity in their mammalian hosts? We examined the hypothesis that the human gut commensal Bacteroides fragilis evolved molecular mechanisms to suppress Th17 responses during homoeostatic colonization. As predicted previously (7, 10, 11), we found that B. fragilis mono-associated animals did not induce Th17 cell development in the colon compared to germ-free controls (Fig. 1A). The beneficial contributions of B. fragilis require a single immunomodulatory molecule named polysaccharide A (PSA), which prevents and cures lethal enteric disease (12–14). Colonization with B. fragilis in the absence of PSA (B. fragilis/PSA−) however, resulted in significant Th17 cell responses in the gut (Fig. 1, A and B). Colonized mice also produced high amounts of IL-10 from mixed cultures of PSA−CD4+ T cells and PSA+CD4− T cells (Fig. 1A and fig. S5), which indicated that PSA required TLR2 expression on T cells to promote IL-10 production. IL-10 responses to PSA were specific to T cells (fig. S6). Consistent with previous findings (15), proinflammatory IFN-γ production was dependent on TLR2 signaling by DCs (fig. S5); however, IL-10 production was unaffected in cultures containing wild-type CD4+ T cells and T cells deleted in the TLR2 gene (Fig. 2A). Therefore, TLR2 expression by T lymphocytes is necessary for IL-10 production by PSA.

Recent studies have shown that certain gut bacteria can promote regulatory T cell (Treg) induction (16, 17). Treg cells expressing the transcription factor Foxp3 (forkhead box P3) suppress proinflammatory Th17 cell responses. To test whether Treg cells prevent immune responses during B. fragilis colonization, we reconstituted germ-free Rag1−/− mice with bone marrow from Foxp3+ DTR (diphtheria toxin receptor) donors, which allowed for specific ablation of Tregs by administration of diphtheria toxin (DT) (16). Mice were mono-associated with B. fragilis to induce Treg development (Fig. 1F). DT treatment of mice resulted in depletion of Foxp3+ T cells (Fig. 1F), with a comitant increase in Th17 responses (fig. S4), which suggests that Foxp3+ Tregs are required for suppression of Th17 cells during B. fragilis colonization.

PSA is an immunomodulatory bacterial molecule that shapes host immune responses (17). Induction of IL-10 and interferon-γ (IFN-γ) from CD4+ T cells by PSA requires TLR2 signaling (14, 18). We sought to determine the mechanism whereby B. fragilis suppresses Th17 cell responses by testing whether PSA functions through TLR2 signaling by dendritic cells (DCs) and/or CD4+ T cells. PSA elicited a significant increase in IL-10 and IFN-γ production from mixed cultures of wild-type DCs and wild-type CD4+ T cells in vitro (Fig. 2A and fig. S5). When Th2−/− mice were cocultured with wild-type DCs, however, PSA-induced IL-10 production was reduced, whereas IFN-γ expression was not affected (Fig. 2A and fig. S5), which indicated that PSA required TLR2 expression on T cells to promote IL-10 production. IL-10 responses to PSA were specific to T cells (fig. S6). Consistent with previous findings (18), proinflammatory IFN-γ production was dependent on TLR2 signaling by DCs (fig. S5); however, IL-10 production was unaffected in cultures containing wild-type CD4+ T cells and Th2−/− DCs (Fig. 2A). Therefore, TLR2 expression by T lymphocytes is necessary for IL-10 production by PSA.

CD4+ T cells produce IL-10 in response to PSA in the absence of antigen-presenting cells (APCs) (Fig. 2B). Moreover, PSA induces IL-10 expression from purified T cells in a dose-dependent manner, whereas other TLR2 ligands do not (fig. S7). TLR2 can function as either a homodimer or a heterodimer with TLR1 or TLR6 (19). PSA could induce high amounts of IL-10 from wild-type, Th1−/−, and Th6−/− CD4+ T cells; however, IL-10 production was lost only from Th2−/− CD4+ T cells and T cells deleted in the TLR2 adapter molecule MyD88 (Fig. 2B). To determine Treg suppression as a function of cell...