Ecological and Evolutionary Forces Shaping Microbial Diversity in the Human Intestine

Ruth E. Ley,1 Daniel A. Peterson,1 and Jeffrey I. Gordon1,*
1 Center for Genome Sciences, Washington University School of Medicine, St. Louis, MO 63108, USA
*Contact: jgordon@molecool.wustl.edu
DOI 10.1016/j.cell.2006.02.017

The human gut is populated with as many as 100 trillion cells, whose collective genome, the microbiome, is a reflection of evolutionary selection pressures acting at the level of the host and at the level of the microbial cell. The ecological rules that govern the shape of microbial diversity in the gut apply to mutualists and pathogens alike.

Our intestinal tract is a nutrient-rich environment packed with up to 100 trillion \((10^{14})\) microbes. The vast majority reside in our colon where densities approach \(10^{11} \text{--} 10^{12}\) cells/ml, the highest recorded for any microbial habitat (Whitman et al., 1998). Today, there are 6.5 billion humans living on Earth. Together, we represent a gut reservoir of \(10^{23} \text{--} 10^{24}\) microbial cells. This number is just five orders of magnitude less than the world’s oceans, which contain an estimated \(10^{29}\) cells (Whitman et al., 1998). Therefore, together with other mammals, the human gut constitutes a substantial microbial habitat in our biosphere.

Because we are born germ free, the microbes that populate our intestinal tract must come from the outside. Microbial diversity on our planet is vast: although 55 divisions (deep evolutionary lineages) of Bacteria and 13 divisions of Archaea have been described to date (DeLong and Pace, 2001; Rappe and Giovannoni, 2003; Rondon et al., 1999), much diversity remains unexplored. The intestine is remarkable for its exclusivity: selection pressures whittle down the microbial diversity of the outside world so that the gut microbiota in adults is dominated by members of just two divisions of bacteria—the Bacteroidetes and Firmicutes—and one member of Archaea, *Methanobrevibacter smithii* (Bäckhed et al., 2005; Eckburg et al., 2005). The gut microbial community presumably has strict requirements for membership: an arsenal of enzymes to utilize available nutrients; cell-surface molecular paraphernalia to attach to the “right” habitat, evade bacteriophages, and appease a reaction-ready immune system; genetic gadgetry for mutability to stay well adapted; the ability to grow rapidly to avoid washout; and the stress resistance needed when making the jump to other hosts via a largely dry and toxic “ex-host” environment. Written into the microbiome is a survival guide for microbes who wish to live in the gut environment.

Lederberg (2000) has emphasized the importance of having a broad ecologic view of our relationships with microbes. In this view, we are seen as superorganisms composed of an amalgam of both microbial and *H. sapiens* cells, where the survival of microbe and human is interdependent. The gut is a natural laboratory for studying the “microevolution” of humans. Experimental and computational tools are now in hand to comprehensively define diversity in our gut microbiota, decipher the gene content of the microbiome, and to explore the metabolome encrypted by this collection of microbial genes—a collection that is estimated to exceed the number of our own human genes by at least two orders of magnitude. The results of this exploration promise to reveal the operating principles that underlie beneficial host-microbial and microbial-microbial relationships, as well as new ecologic perspectives and “ecogenomic” views of how pathogens arise and function within our indigenous microbial communities.

In this review, we argue that the microbial diversity of the human gut is the result of coevolution between microbial communities and their hosts. We suggest that the peculiar structure of microbial diversity in the human gut resulted from natural selection operating at two levels. Host level, “top-down” selection on the community favors stable societies with a high degree of functional redundancy. An opposing “bottom-up” force is selection pressure driving microbial cells to become functionally specialized (Figure 1). In addition to the selection pressures shaping microbes in the gut, we discuss factors that constrain diversity. Finally, we present an ecologic view of pathogenic relationships: the challenges potential pathogens face when encountering a microbiota where there is functional redundancy with virtually all niches and habitats filled and the idea that microbial community structure should be considered as a factor that can influence predisposition to specific diseases in certain host contexts.

Who Passes the Gauntlet? Patterns of Microbial Diversity in the Gut

Definitions

Bacteria and Archaea in the gut multiply by binary fission. They can “differentiate” genetically by a number of mechanisms, including lateral gene transfer via phage- and mobile element-mediated insertions, conjugal transfer, or uptake of naked extracellular DNA. Alterations to vertically
inherited genomes can come from mutations (Giraud et al., 2001b) and chromosomal rearrangements. Within this mutable context, the term “individual” is misleading (single cells have been referred to as “dividuals” [Koestler, 1967]), and the term “species” is ambiguous. We use “species” here to refer to named types (e.g., Eubacterium rectale), and “phylogtype” (phylogenetic type) to refer to clusters of related 16S rRNA gene sequences characterized by levels of pairwise sequence identity (≥ 97% ID is the commonly used threshold used for a “species”).

The gut ecosystem is very dynamic. Within a given intestinal habitat (address), some microbial members function as entrenched “residents” (autochthonous components), while others act more like hitchhikers (allochthonous members) from ingested food, water, and various components of our environment. Relationships between members of the microbiota and humans are frequently described as commensal (one partner benefits while the other seems unaffected) rather than mutualistic (both partners derive benefit). In this review, we address the evolution of the entrenched residents. We favor the term mutualist for the permanent residents of the gut based on the view that selection for stable communities favors beneficial relationships.

**Low Levels of Deep Diversity**

The most comprehensive and least biased enumerations of microbial diversity within the mammalian gut have come from sequencing 16S rRNA genes. These sequences are obtained directly from DNA extracted from gut mucosal biopsies or feces using PCR primers targeted to broad phylogenetic groups. The largest data sets consist of 13,335 16S rRNA sequences from three healthy adult humans (Eckburg et al., 2005) and 5,088 sequences from 19 adult mice belonging to the commonly used C57Bl/6 inbred strain (Ley et al., 2005). In the study of the human intestinal microbiota, bacterial and archaeal 16S rRNA sequences were derived from biopsies taken from six regions of the colon and one stool sample from each individual. In the mouse study, 16S rRNA bacterial sequences were obtained from the cecal contents. Neither study targeted Eukarya, although recent work in mice suggests an unappreciated diversity of fungi (Scupham et al., 2006).

The human intestinal samples contained members of seven divisions of Bacteria (Firmicutes, Bacteroidetes, Actinobacteria, Fusobacteria, Proteobacteria, Verrucomicrobia, Cyanobacteria), which, together with divisions described in previous studies (Spirochaetes, VadinBE97 [Bäckhed et al., 2005]), brought the total number to nine. The mouse survey showed a very similar number and set of divisions, although Fusobacteria was not detected and TM7 was found. The Firmicutes and the Bacteroidetes dominated, together accounting for >98% of all 16S rRNA sequences in each mammalian host.

**Radiation of a Few Lineages**

The human intestinal samples contained members of seven divisions of Bacteria (Firmicutes, Bacteroidetes, Actinobacteria, Fusobacteria, Proteobacteria, Verrucomicrobia, Cyanobacteria), which, together with divisions described in previous studies (Spirochaetes, VadinBE97 [Bäckhed et al., 2005]), brought the total number to nine. The mouse survey showed a very similar number and set of divisions, although Fusobacteria was not detected and TM7 was found. The Firmicutes and the Bacteroidetes dominated, together accounting for >98% of all 16S rRNA sequences in each mammalian host.
Figure 2. Comparison of Microbial Diversity in the Human Colon, Mouse Cecum, Ocean, and Soil

(A) Percent representation of divisions in each environment. For the human gut, 16S rRNA sequence data (n = 11,831) were obtained from the Arb alignment supplied by Eckburg et al. (2005). We used our available Arb alignment for mouse data (n = 5088; Ley et al., 2005); the ocean data set (n = 1082) was obtained from adult conventionally raised adults (Rawls et al., 2004; J.F. Rawls, R.E.L., and J.I.G., unpublished data). Soil data sets from two independent studies were combined (n = 1379; Axelrood et al., 2002; Tringe et al., 2005). Stromatolite data (n = 321; Papineau et al., 2005) were obtained from GenBank. Stromatolite, soil, ocean, and fish sequences were aligned using the nast online Arb alignment tool (http://greengenes.lbl.gov/cgi-bin/nph-NAST_align.cgi).

(B) Phylogenetic architecture of the microbial communities shown in (A). For each habitat, the number of phylotypes per 100 16S rRNA gene sequences is shown for differing thresholds of 16S rRNA gene pairwise sequence identity (%ID). The blue bar highlights the phylogenetic fans were unique to each person. For each habitat, the number of phylotypes per 100 16S rRNA gene sequences is shown for differing thresholds of 16S rRNA gene pairwise sequence identity (%ID). The blue bar highlights the phylogenetic fans were unique to each person.

Mucosal communities sampled at different points along the colon were similar within each of the three human hosts (Eckburg et al., 2005; Zoetendal et al., 2002), and their composition overlapped with the fecal microbiota (the colon has incomplete peristalsis that permits backflow and mixing). Interpersonal differences in microbial community composition are illustrated in Figure 3: in several places in the 16S rRNA tree, regions representing phylotypic fans were unique to each person.

This pattern of high levels of strain variation but far fewer intermediate and deep lineages was also observed in the mouse gut microbiota (Ley et al., 2005). This type of extreme fan-like phylogenetic architecture may be a signature feature of the gut ecosystem. Shallow phylotype fans have been described in other natural habitats (e.g., the ocean [Acinas et al., 2004]), but these habitats appear to harbor a greater number of intermediate-level phylogenetic groups than the gut (Figure 2B). The pattern of phylotype fans in the gut recalls the patterns in classic adaptive radiations (Schuter, 2000), such as the radiation of terrestrial snail species of French Polynesia (Garrett, 1884) where a few successful early colonists gave rise to a variety of descendants. Shallow, wide radiations result from extreme selection pressure followed by détente. In the case of the snails, the selection pressure came in the form of a wide ocean to cross, but once there, the early snails filled an island’s worth of empty niches (professions) across a range of habitats (addresses). Similarly, the phylogenetic architecture of the gut could have resulted from the diversification of a discrete limited initial community (a population bottleneck) into strains. In addition, the phylogenetic “shallowness” of the intestinal community reflects the relatively short time that the mammalian gut has existed as a habitat (~100 million years for mammals with placentas [Murphy et al., 2001], versus >3.85 billion years for the ocean).

How Do They Get There? 100 Trillion Family Heirlooms

Pathogens that use an oral-fecal route for their transmission (e.g., members of the Proteobacteria such as Vibrio cholera) can exploit environmental reservoirs outside of their hosts to proliferate. However, members of the
Firmicutes and Bacteroidetes detected in the human gut do not appear to grow outside of their host and likely rely on the close contact of parents and offspring for transmission. One testable prediction of this parent-to-offspring transmission hypothesis is that microbial communities will be similar in members of a given family. In experiments with C57Bl/6 mice, we showed that animals inherited their microbial communities from their mothers (Ley et al., 2005), using the recently developed UniFrac metric (Lozupone and Knight, 2005). Related mice shared similar bacterial communities: the effect was evident across multiple generations: the mothers that were sisters shared microbiotas with each other as well as with their offspring. These findings demonstrate that a microbiota can be inherited vertically from mothers and is stable enough over time that kinship relationships are reflected in community composition.

Culture-based studies in humans indicate that babies acquire their initial microbiota from the vagina and feces of their mothers (Mandar and Mikelsaar, 1996). Babies delivered by caesarian section have an altered colonization pattern relative to their vaginally-delivered counterparts (Gronlund et al., 1999). Host genotype represents a confounding factor when attempting to show vertical transmission of microbes from parents to offspring. In a study of adult monzygotic twins, their siblings, and marital partners, Zoetendal et al. (2001), used DNA fingerprinting methods to show greater similarities between the gut microbial communities of monzygotic twins than between monzygotic twins and their unrelated marital partners. The similarities between the monzygotic twins’ gut communities were interpreted as an effect of genotype on microbial diversity. An alternative explanation is that the observed similarities between twins’ communities were due to their colonization from a shared mother. Indeed, the study also included dizygotic twins and siblings: their microbiota were as similar to each other’s as the microbiota of monzygotic twins, despite a lower level of host genetic relatedness (Zoetendal et al., 2001). Similarly, in the only small subunit (SSU) rRNA sequence-based study of kinship effects on the normal human microbiota, Frank et al. (2003) showed that marital partners had different communities in their ear canals but that the same bacterial species was dominant within members of a given family.

In our study of mice, genotype was shown to act on the relative abundance of groups whose specific composition was determined by kinship. The mothers were heterozygotes for a mutation (ob) in the leptin gene. The offspring were produced from matings of ob/+ males and females.
so a subset was obese (ob/ob), while their wild-type (+/+) and heterozygous ob/+ siblings were lean. The composition of the cecal microbiota (i.e., phyotypes present irrespective of their abundance) was dictated by the mother regardless of their offspring’s genotype. However, there was a division-wide 50% reduction in the abundance of Bacteroidetes in ob/ob animals and a proportional division-wide increase in the Firmicutes compared to their lean +/- or ob/+ littermates (Ley et al., 2005).

This mouse study allowed the effects of genotype to be decoupled from effects of kinship because the experimental conditions could be controlled and because analytical methods are now available that take into account composition with or without abundance data (Lozupone and Knight, 2005). These approaches can now be used, in concert, to address some basic unanswered questions about human microbial ecology (see Table 1). For example, we do not understand how resilient (resistant) an individual’s microbiota is to potential colonists/pathogens from contaminated water, fecal material aerosolized by modern toilets (Barker and Jones, 2005), or cohabitation with an unrelated partner (spouse), or to disturbances such as antibiotic treatment, infection with parasites, or drastic shifts in diet. Is mother/caregiver-to-infant transmission a lifelong shaper of an individual’s microbiota? Is there a window of opportunity during infancy when the gut is particularly “open” or susceptible to colonization or, for that matter, to an enduring perturbation of its microbial ecology following antibiotic therapy? Does the early-acquired microbiota persist into adulthood, having gained a foothold by initially “educating” the immune system to help retain a particular suite of organisms? What are the effects of genotype or health status? Controlled studies in mice could tease apart the effects of age, disturbance, exposure, and successional order on the composition and resilience of microbial communities. Such studies should inform and direct sampling designs aimed at answering similar questions in humans. These basic studies will provide more than interesting “natural history”: if people acquire their microbiome and its gene content from their families, it represents another form of genetic heritability—one that has the potential to be reprogrammed.

The Evolutionary Gauntlet: Selection Pressures for Form and Function
The method of gut colonization is one factor that influences diversity. Other factors include the physical and chemical environment and the selection pressures on the host. These selection forces (Figure 1), both direct and indirect, are what mutualistic and pathogenic microbes alike must contend with to be successful.

Chemistry
The gut offers fewer chemical niches (professions) to its microbiota than a habitat such as soil. Gut microbes can derive energy from transferring electrons from organic carbon either to organic carbon (fermentation), to inorganic carbon (methanogenesis), or to sulfate (sulfate reduction), all of which are available in the gut at levels that can sustain populations. In addition, H2 is undoubtedly an important intermicrorganism electron transfer shuttle. Other microbial niches, such as nitrification or photosynthesis, are not viable options in the gut for lack of substrates and light, and therefore entire functional guilds common in other microbial habitats are excluded. Further reduction in the number of niches may result from the peristalsis and churning of the gut. Carbon resource heterogeneity or spatial partitioning of different organic substrates correlates with diversity patterns in soil (Zhou et al., 2002). However, the mixing of gut contents is a homogenizing force that would reduce niche breadth.

Some of the metabolic niches in the gut can be loosely mapped onto its major lineages. Only Archaea are known produce methane. Sulfate reduction is carried out by Delta Proteobacteria (e.g., Desulfotomaculum spp.) and one clade within the Firmicutes (Desulfotomaculum spp.). Fermentation, however, is a phylogenetically widely held skill, and the principal energy pathway for members of the Bacteroidetes and Firmicutes (and thus the dominant energy-producing pathway for the microbiota).

In principle, microbes from any phylogenetic division could acquire the necessary genes to survive as a fermenter in the gut: for example, key glycoside hydrolases that liberate sugars from dietary polysaccharides could be acquired by lateral gene transfer (LGT). Yet the dominance of the Firmicutes and Bacteroidetes shows that acquisition of a gene suite required to colonize the gut has been limited to members of a few phylogenetic divisions. Microbes are more likely to swap genes with microbes from the same environment (Beiko et al., 2005); therefore, dominant groups would monopolize the gene swapping market and newcomers would be rare. However, LGT is a homogenizing evolutionary force (Woese, 2002). If members of the Bacteroidetes and Firmicutes have continually swapped genes over their evolution, their genomic contents would be drawn from a common pool of gut-adapted genes. Whole genome data are needed to make such interdivision comparisons, but the known differences in GC content between members of the two divisions is evidence that LGT has not homogenized their genomes. The codominance of the Bacteroidetes and Firmicutes likely stems from their distinct and complementary metabolic roles within the community. Moreover, these co-evolved, cooperative roles were likely traits of early colonists in mammalian gut evolution and are encoded by a genomic repertoire that is too large and eclectic to be acquired and assembled solely by stochastic processes such as LGT.

Cooperation
It remains unclear how much traditional ecological theory can be strictly applied to the microbial world. However, concepts borrowed from ecology can frame our view of microbial interactions in the gut. Ecological theory teaches that head-to-head competition for limiting resources results in competitive exclusion—the loser goes extinct and the community becomes a monoculture of the winner
Table 1. Key Questions and Future Studies

<table>
<thead>
<tr>
<th>Question</th>
<th>Approaches</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. How is the gut microbiota transmitted in humans?</td>
<td>If the transmission is primarily parent-child, we should detect kinship patterns in microbial communities of the hosts. One caveat is that modern standards of hygiene may have altered transmission mechanisms. Parallel studies of families in developed areas where exposure to water and food contaminated with fecal microbes is minimal and families in underdeveloped areas where such exposure is common are required elucidate patterns of transmission.</td>
</tr>
<tr>
<td>2. How distinctly human is our gut microbiota?</td>
<td>If the diversity of the microbiota coevolved with mammalian host species, the relationships of gut communities from different mammals should mirror the evolutionary relationships of the mammals. Furthermore, the microbial communities of distantly related host species should be less functionally interchangeable than the communities of closely related hosts.</td>
</tr>
<tr>
<td>3. How functionally redundant are the members of the human gut microbiota?</td>
<td>Natural selection at the host level will favor functionally stable communities. This “top-down” pressure should produce a community with functional redundancy encoded in its genomes. An opposing “bottom-up,” cell-level selection pressure would favor niche specialization to avoid competition. The sequencing of a large collection of microbial genomes from the gut, from closely and distantly related bacteria and the building-up of “intentional communities” within gnotobiotic mouse models should help answer fundamental questions about how the gut community is assembled and operates.</td>
</tr>
<tr>
<td>4. Are members of our gut microbiota more likely to have exchanged genes with each other than with microbes from other habitats?</td>
<td>The lateral exchange of genes is a fundamental process of microbial genome evolution, one that can confer new characteristics to organisms, allowing them to jump from peak to peak within fitness landscapes. However, microbes are restricted in the gene space they can sample by the suite of potential donors that share their habitat. This would have the effect of homogenizing the genome pool over time within a habitat. The gut is a natural laboratory for studying genome evolution within spatially distinct habitat islands. Moreover, these islands are distributed at a host scale across the globe. The gut is also an alluring model for studies of the biogeographic effects of habitat scale and fractionation on genome evolution in both mutualists and pathogens.</td>
</tr>
<tr>
<td>5. What controls the rate of lateral gene transfer among microbes in the gut?</td>
<td>If lateral gene transfer (LGT) homogenizes genomes over time, genes in any given genome will be drawn randomly from a habitat’s gene pool. The degree to which microbiomes differ between habitats will be a reflection of the degree of mixing between habitats and the rates of LGT within a habitat. Experiments in gnotobiotic models may help quantify the rate of LGT among prominent members of the gut microbiota.</td>
</tr>
<tr>
<td>6. What is the impact and role of Archaea?</td>
<td>The Archaea within the gut appear to follow a different set of rules than the Bacteria. Archaeal diversity appears to be low, although this needs more verification (Rieu-Lesme et al., 2005). By providing the final step in energy extraction from degradation of organic compounds, Archaea can alter the thermodynamics of the whole system, with profound consequences for the host. Are Archaea the power brokers of the gut? Are they pillars of the community and a target for therapeutic manipulation?</td>
</tr>
<tr>
<td>7. Does the immune system respond to gene suites/gene expression rather than to species?</td>
<td>The role of the immune system in shaping the microbiota remains largely enigmatic. The traditional view of bacterial species as entities with discrete phenotypes that the immune system responds to oversimplifies the problem. We need to rethink how an immune system can aid in the retention of a functional suite of microbes whose composition may change over the lifetime of the host. The immune system and the microbiota should be viewed as a coevolved system: a driving force for evolution of the immune system is the need to accommodate (ongoing) diversity in a host microbiota; this in turn allows the host to accommodate environmental (including food) antigens and possibly self-antigens.</td>
</tr>
</tbody>
</table>
plants and animals on islands (Emerson and Kolm, 2005). The process has been invoked to explain high diversity among species (Putman, 1994). Competition is avoided if organisms partition resources and/or cooperate or, alternatively, if another factor, such as the host, limits their growth. Resource partitioning is initiated when an organism finds an empty niche and uses a resource that is not already tapped. E. coli populations have been shown to differentiate into diverse niches when introduced into germ-free mouse guts (Giraud et al., 2001a). Genetic diversification can enable exploitation of new niches by increased mutation rates within a clonal population (Bjedov et al., 2003) and in a mixed community by LGT (Gogarten et al., 2002). Diversification can create positive feedback loops: diversity begets diversity by creating new niches. This process has been invoked to explain high diversity among plants and animals on islands (Emerson and Kolm, 2005).

At the microbial level, diversity may expand niche space to a degree that is not possible in plant and animal communities. Microbes produce chemical food webs where the product of one microbe becomes the substrate for another, and the removal of waste products (e.g., H₂) can enhance the thermodynamic yield of the first (Thauer et al., 1977). Microbial chemical webs have the potential to be inconceivably complex and changeable as cells adjust their transcriptomes and metabolomes to variations in substrate availability and thermodynamic gradients. The sum of activities of members of the community is an emergent property of the community with consequences for host fitness.

**Selection for Emergent Functionality: Pressure on Hosts**

Each microbial cell is under selection pressure in the gut: this pressure acts on the cell’s phenotype and results in fixation of genes in the genomes of its daughters. Each cell is also part of a community that provides critical services (e.g., nutrient extraction, invasion resistance) for the higher level in the hierarchy—namely the host. Hierarchy theory states that the higher levels impose constraints on the lower levels (Koestler, 1967). Emergent properties of the gut microbial community impact host fitness and consequently the future availability of a gut habitat. Therefore, irrespective of the success of particular suites of bacteria in the gut, if the host does not benefit from their collective behavior, the entire group can be selected against when the host dies without leaving many offspring. Ecological theory suggests that for group selection to take place, the community must occasionally be transferred to a new habitat (Wilson, 1975): this is what happens when new hosts are born and colonized. When the community is transferred, cheaters that benefit from the community without contributing to it are purged (they are present at lower abundance, and the probability that they will be transferred is low; Travisano and Velicer, 2004). Furthermore, transmission to new hosts changes the scale at which competition between members of the microbiota occurs from local (within host) to global (between hosts): this type of change would favor cooperation and altruism over cheating (Griffin et al., 2004).

The diversity of the gut microbiota today may reflect “the ghost of group selection past.” However, the host and his/her gut microbial community constitute a rare biological system: one where group selection could be demonstrated empirically. To test empirically if an emergent trait drives a phenotype that can be selected against at the host level, experiments could be designed in which an intentional community of microbes is seeded into a previously germ-free population of mice and community-level traits (such as the ability to degrade a specific compound requiring a multispecies consortium) used as the basis for artificial selection. Group selection based on emergent properties of the community is standard in engineered systems where a specific product is desired. Functional stability is a property that is always selected for in waste-water treatment systems (by the engineer) and one that may play a critical role in promoting host fitness: after all, the gut must digest food and liberate nutrients and energy in a reliable fashion.

**Functional Redundancy and the Reduced Need for “Keystone Species”**

Our gut microbiota is an efficient and stable natural bioreactor (Sonnenburg et al., 2004): it is resistant and resilient to chaotic blooms of subpopulations (or pathogens) that could be disruptive and reduce host fitness (Bäckhed et al., 2005). Functional redundancy in a microbiota confers stability (also known as the insurance hypothesis [Yachi and Loreau, 1999]) that will be selected for at the host level. Functional redundancy can also obviate the need for “keystone species” (defined as a species with a central role in the system whose loss causes a dramatic change in processes and diversity). If selection pressure at the host level for a reliable stable community is a dominant shaper...
of diversity in the gut, then the resulting functional redundancy should be written into the microbiome as well as scripted in host-microbial interactions.

An extreme test of the functional redundancy of the gut microbiota is replacement of an entire gut microbiota with just one species. Monoassociation of germ-free mice with a single species of Bacteroidetes can recapitulate a number, but certainly not all, host phenotypes observed in “conventionally raised” animals (mice that acquire a complete microbiota beginning at birth). Analysis of the finished genome sequences of three members of Bacteroidetes prominently represented in the human colonic microbiota has revealed a considerable degree of redundancy in their ability to process plant polysaccharides (see below). It is likely that the phylotype fans observed in the gut correspond to “ecotypes” (habitat specialists; Acinas et al., 2004; Palys et al., 1997) with differing levels of functional redundancy. Different species within the family may be functionally redundant, but to coexist they partition the resource base by expressing species-specific substrate preferences and/or “use efficiencies” (Sonnenburg et al., 2005). The alternative strategy, exclusive niche specialization, can result in decreased ability to diversify further (Buckling et al., 2003) and risk to the host that a keystone species is irreplaceably lost in a selective sweep (Cohan, 2002) from, for example, attack by bacteriophages that are abundant in the gut (Breitbart et al., 2003).

Although functional redundancy may be a dominant theme within lineages that occupy the gut, one exceptional organism that appears to be irreversibly adapted to a niche for which there appears to be few if any alternative players is the archaeon Methanobrevibacter smithii. *M. smithii* couples H2 oxidation to CO2 reduction to produce methane (CH4) in one of the least energy-yielding reactions in biology. It competes directly for H2 with sulfate-reducing bacteria but is thought to hold a competitive advantage in vivo (Strochchi et al., 1994). In the comprehensive 16S rRNA sequence-based enumeration study of the colonic microbiota of three healthy individuals, the person with the highest levels of bacterial diversity was the one with the most archaeal (*M. smithii*) sequences (Eckburg et al., 2005) (Figure 3). This observation could be anecdotal, or it could be a hint that archaea have a pronounced effect on the thermodynamics and diversity of the gut microbiota. Interestingly, in engineered systems where functionality is constant, populations of bacteria can fluctuate chaotically, but archaeal populations remain where functionality is constant, populations of bacteria can fluctuate chaotically, but archaeal populations remain.

Pathogenic Species and “Pathogenic Communities”

In the classical view of an enteric pathogenic species, the pathogen is a cheater that benefits from gut microbial community dynamics while imposing a fitness cost to the host and to the community. In this view, pathogenic species are like mutualistic species in that they have a shared interest in their hosts’ survival (Ledberg, 2000). Selection on the host results in selection for pathogens that are not too virulent, and/or a microbial community that can prevent pathogens from building up the necessary population density to cause injury (Czuprynski and Balish, 1981; Zachar and Savage, 1979). Theoretical studies suggest that mutualists typically have a competitive advantage over cheaters (Ferriere et al., 2002). Rather than competing head-to-head with...
mutualists, a successful pathogenic species can use a high infective dose, specialized organelles for attachment/invasion, and/or enterotoxins to induce secretion of water so that it can enter or create habitats that are not occupied by members of the normal microbiota (Nataro and Kaper, 1998). While this behavior results in efficient amplification and return to the environment, it is not compatible with long-term colonization, due to the demise of the host or a robust elicited immune response. As a result, virtually all enteropathogenic bacteria have an environmental reservoir and often a definitive host (e.g., poultry for Salmonella).

In addition to the classic pathogenic species, we propose that another kind of pathogenicity exists in the gut: one in which the whole community is “pathogenic” when its emergent properties contribute to disease. In a “pathogenic community,” no single microbe is pathogenic alone. Instead, the community assemblage is an environmental risk factor that contributes to a disease state. A microbial community will be pathogenic within the context of other risk factors, such as host genotype, diet, and behavior. For instance, the amount of calories available to the host from food is a value modulated to a significant degree by the gut microbiota (Bäckhed et al., 2005). A microbial community whose energy extraction is very efficient could constitute a risk factor for obesity in a person with ready access to food, whereas it might promote health in an individual with more limited access. We have argued that a community whose emergent properties decrease host fitness will be selected against. Diseases such as obesity do not necessarily reduce host fitness today if fitness is only measured as the number of offspring produced. However, they do impose a cost to both the individual and to society. It is up to us as a human society to recognize and select against the microbial communities that may be risk factors for these types of maladies within individuals as part of preventive medicine.

The Human Gut Microbiome Initiative
Comparing multiple genomes, representing 16S rRNA phylotypes with different degrees of relatedness would help answer a number of key questions about the microbiome. Which gene families are widespread among lineages and therefore essential for survival in the gut ecosystem? How much lateral gene transfer occurs between distant versus close relatives in the densely populated distal gut, and how does this relate to the evolution and functional stability of the microbiota’s metabolome? How have mutualists evolved, and how are they continuing to evolve to interact with the immune system? Can features of microbial genome structure and microevolution be used as biomarkers of health or of susceptibility to specific diseases? To what extent are shared elements in mutualists represented in pathogenic microbes that face similar ecological pressures?

To begin to answer these questions, we have turned initially to members of the genus Bacteroides. B. thetaiotaomicron comprises 12% of all Bacteroidetes and 6% of all Bacteria in the 11,831 member human colonic bacterial 16S rRNA data set (Eckburg et al., 2005). B. thetaiotaomicron’s 6.3 Mb genome reflects some of the selective pressures that have defined its habitat and its niche (Xu et al., 2003). It contains an unusually large ensemble of genes involved in acquiring and metabolizing carbohydrates: this arsenal includes 163 outer membrane proteins with homology to two proteins, SusC, SusD, that bind and import starch, 226 predicted glycoside hydrolases, and 15 polysaccharide lyases (Xu et al., 2003). (By contrast, our 2.85 Gb genome contains a relatively “paltry” 99 known or putative glycoside hydrolases and no polysaccharide lyases: it is deficient in enzymes required for degradation of xylan-, pectin-, and arabinose-containing polysaccharides that are common components of dietary fiber [http://afmb.cnrs-mrs.fr/CAZY/]). We have also produced finished genome sequences for B. vulgatus (31% of Bacteroidetes and 15% of Bacteria), and B. distasonis (0.8% and 0.4%) (J. Xu and J.I.G., unpublished data). These Bacteroides species also have highly evolved “glycobiomes” (genes involved the acquisition, breakdown, or synthesis of carbohydrates), that may result from “top-down” selection for functional redundancy.

These large and similar suites of genes within bacterial genomes could be paralogs resulting from duplication, but they may well be genes acquired by LGT from bacteria that are entrenched members of the microbiota or microorganisms acquired from the environment that are simply passing through our gut (allochthonous members). Indeed, the genomes of the sequenced Bacteroides are peppered with mobile elements that can facilitate LGT. Intriguingly, the capsular polysaccharide synthesis loci of B. thetaiotaomicron contain a number of predicted glycosyltransferases that appear to have been acquired by LGT (Xu et al., 2003), providing a intriguing perspective about the dynamic nature of the interface that exists between our environment, our microbiome, and the surface properties of the microbiota that are seen by our immune system. Lederberg (2000) has suggested that a good pathogen is one whose epitopes stimulate an immune response to competitors: could this (also) be the hallmark of a good mutualist?

For these Bacteroides species to codominate in the gut, they likely have distinct hierarchical substrate preferences would allow resource partitioning and/or metabolic cooperation between congeners (i.e., there would be niche specialization). Our in vivo functional genomic and mass spectrometry-based metabolomic studies of the adaptive foraging behavior of B. thetaiotaomicron in gnotobiotic mice support this view (Sonnenburg et al., 2005). Groups of bacteria assemble on undigested or partially digested food particles and shed elements of the mucus gel layer and/or exfoliated epithelial cells. Attachment to these nutrient reservoirs is directed by nutrient-regulated glycan-specific outer-membrane binding proteins. Attachment helps oppose washout from the gut bioreactor and facilitates harvest of carbohydrates by adaptively expressed glycoside hydrolases. Moreover, when polysaccharide
availability from the diet diminishes, the organism turns to polysaccharides in host mucus (Sonnenburg et al., 2005). Coevolution of glycan structural diversity in the host, and an elaborate collection of nutrient-regulated glycoside hydrolases in members of the microbiota such B. thetaiotaomicron likely ensures that the host and the microbiota can adapt to dietary change and maximize energy harvest, without having to undergo dramatic changes in the representation of this species, or those microbes that use products produced by B. thetaiotaomicron’s “digestive tract” (Sonnenburg et al., 2005).

To obtain a more comprehensive view of the microbiome, we have proposed a human gut microbiome initiative (Gordon et al., 2005; HGMI) that will deliver deep draft whole genome sequences for 100 species representing the bacterial divisions known to comprise our distal gut microbiota. We have identified 86 cultured representatives (22%) of the 395 phylotypes identified in the human colon. The deposited curated genome sequences would herald another phase of the “human” genome sequencing project, provide a key point of reference for interpreting “metagenome” sequencing projects that use total microbial community DNA as starting material for shotgun sequencing, serve as a model for future initiatives that seek to characterize our other extraintestinal microbial communities, and facilitate analysis of how selective pressures and community dynamics have shaped the microbiome in diseased humans and in gut pathogens.

Finally, global warming, loss and homogenization of biological diversity due to species invasions and extinctions (Clavero and Garcia-Berthou, 2005; Mooney and Cleland, 2001), plus human deposition of biologically essential elements such as phosphorus and nitrogen around the planet (Falkowski et al., 2000), impact our biosphere at the microbial level. Defining the gut microbiota and microbiome in people who live in various geographic regions, under various levels of economic development, should provide an opportunity to monitor human “micro-evolution” during this period of profound social, economic, and ecological change, and, hopefully, help us forecast changes in our disease susceptibility.

Supplemental Data
Supplemental Data include Supplemental Experimental Procedures and can be found with this article online at http://www.cell.com/cgi/content/full/124/4/837/DC1/.

ACKNOWLEDGMENTS

We thank Lars Angenent, Justin Sonnenburg, Buck Samuel, Jian Xu, John F. Rawls, and Fredrik Bäckhed for helpful comments. Work from our lab cited in this review has been supported by grants from the NIH (DK30292, DK58529, T32-DK07130 [R.E.L.], T32-CA09547 [D.A.P.]), the NSF (033284), the Ellison Medical Foundation, and the W.M. Keck Foundation.

REFERENCES


