Identification of Archaeal and Eukaryotic Members Surveys based on SSU rRNA gene sequencing have largely focused on Bacteria, yet the census of "who's there" in human body habitat–associated communities must also include the other two domains of life: Archaea and Eukarya. Differences in the sequences of archaeal and bacterial 16S rRNA genes, first recognized by Carl Woese in 1977, allowed these two domains of life to be distinguished. The representation of Archaea in human microbial communities is less well defined than that of Bacteria, in part due to the difficulty in optimizing the design of PCR primers that specifically target conserved regions of archaeal (versus bacterial) 16S rRNA genes. Identifying archaeal members is important to our understanding of the functional properties of the microbiota. For example, a major challenge faced by microbial communities when breaking down polysaccharides (the most abundant biologic polymers on Earth) is the maintenance of redox balance in the setting of maximal energy production. Many microbial species have branched fermentation pathways that allow them to dispose of reducing equivalents (e.g., by the production of H_2, which is energetically efficient). However, there is a caveat: the hydrogen must be removed or it will inhibit reoxidation of pyridine nucleotides. Therefore, hydrogen-consuming (*hydrogenotrophic*) species are key to maximizing the energy-extracting capacity of primary fermenters.

In the human gut, hydrogenotrophs include a phylogenetically diverse group of bacterial acetogens, a more limited group of sulfate-reducing bacteria that generate hydrogen sulfide, and methane-producing archaeal organisms (*methanogens*) that can represent up to 10% of the anaerobes present in the feces of some humans. However, the degree of archaeal diversity in the gut microbiota of healthy individuals appears to be low.

Culture-independent surveys of eukaryotic diversity are also confounded by challenges related to the design of PCR primers that target the eukaryotic SSU gene (18S rRNA) as well as the internal transcribed spacer regions of rRNA operons. Metagenomic studies of healthy human adults living in countries with distinct cultural traditions and disparate geographic features and locations have revealed that the degree of eukaryotic diversity is lower than that of bacterial diversity. In the gut, which contains far more microbes than any other body habitat, the representation of fungi is significantly lower in individuals living in Westernized societies than in those living in non-Western societies. The most abundant fungal sequences belong to the phylum-level taxa *Ascomycota* and *Microsporidia*. The phyla *Ascomycota* and *Basidiomycota* appear to be mutually exclusive, and the presence of *Candida* in particular correlates with recent consumption of carbohydrates.

Elucidation of Viral Dynamics Viruses are the most abundant biologic entity on Earth. Viral particles outnumber microbial cells by 10:1 in most environments. Humans are no exception in terms of viral colonization; our feces alone contain 10^8–10^9 viral particles per gram. Despite this abundance, many eukaryotic viral communities remain incompletely characterized, in part because the identification of viruses within metagenomic sequencing datasets is itself very challenging. Characterizing viral diversity requires different approaches: because no single gene is found in all viruses, no universal phylogenetic "barcode of life" equivalent to the SSU rRNA gene exists. One approach has been to selectively purify virus-like particles from community biospecimens, amplify the small amounts of DNA that are recovered, and randomly fragment the DNA and sequence the fragments (*shotgun sequencing*). The resulting sequences can be assembled into larger *contigs* whose function can be computationally predicted from homology to known genes, and the information obtained can be used to populate/expand nonredundant viral databases. These annotated nonredundant databases can then be used for more targeted mining of the rapidly expanding number of shotgun sequencing datasets generated from total-community DNA for known or putative DNA viruses.

Given the dominance of bacteria in the gut microbiota, it is not surprising that *phages* (viruses that infect bacteria) dominate the identifiable components of the gut's DNA virome. *Prophages* are a manifestation of a so-called temperate viral–bacterial host dynamic, in which a phage is integrated into its host bacterium's genome. This temperate dynamic provides a way to constantly refashion the genomes of bacterial species through horizontal gene transfer. Genes encoded by a prophage genome may expand the niche and fitness of their bacterial host, for example, by enabling the metabolism of previously inaccessible nutrient sources. Prophage integration can also protect the host strain from superinfection, "immunizing" the strain against infection by closely related phages. A temperate prophage life cycle allows the virus to expand in a 1:1 ratio with its bacterial host. If the integrated virus conveys increased fitness, the prevalence of the bacterial host and its phage will increase in the microbiota. Induction of a lytic cycle, where the prophage replicates and kills the host, may follow. Lytic cycles can cause high bacterial turnover. Lysis debris (e.g., components of capsules) can be used as nutrient sources by surviving bacteria; this change in the energy dynamic in a community is referred to as a *phage shunt*. A subpopulation of bacteria that undergoes lytic induction may sweep away other sensitive species present in the community, thus increasing the niche space available for survivors (i.e., those bacteria that already have an integrated prophage). Periodic induction of prophages leads to a "constant diversity dynamic" that helps maintain community structure and function.

Interest in viral communities has expanded in recent years, especially given a potentially therapeutic role for phages as an alternative or adjunct to antibiotics. Virome members have evolved elegant survival mechanisms that allow them to evade host defenses, diversify, and establish elaborate and mutually beneficial symbioses with their hosts. A number of recent studies have tried to adapt these mechanisms for therapeutic purposes (e.g., the use of synthetic phages to treat *Pseudomonas aeruginosa* infections in burn patients or in other settings). Phage therapy is not a new idea: Félix d'Herelle, co-discoverer of phages, recognized their potential medical applications nearly a century ago. However, only recently have our technologic capabilities and our knowledge of the human microbiota made phage therapy realistically attainable within our lifetimes.

ECOLOGIC PRINCIPLES AND PARAMETERS FOR COMPARING MICROBIAL COMMUNITIES

At many levels, different people are very much alike: our genomes are >99% identical, and we have similar collections of human cells. However, our microbial communities differ drastically, both between people and between habitats within a single human body. The greatest variation (beta diversity, described below) is between body sites. For example, the difference between the microbial communities residing in a person's mouth versus the same person's gut is comparable to the difference in communities residing in soil versus seawater. Even within a body site, the differences among people are not subtle: gut, skin, and oral communities can all differ by 80–90%, even from the broad, bacterial species–level view. The English poet John Donne said that "no man is an island"; however, from a microbial perspective, each of us consists of not just one isolated island but rather a whole archipelago of distinct habitats that exchange microbes with one another and with the outside environment at some as yet undetermined level. Before we can discuss these differences and understand their relevance to human disease, it is important to understand some basic terms and ecologic principles.

Alpha Diversity *Alpha diversity* is defined as the effective number of species present in a given sample. Communities that are compositionally more diverse (i.e., have more OTUs) or that are phylogenetically more diverse are defined as having greater alpha diversity. Alpha diversity can be measured by plotting the number of different types of SSU rRNA sequences identified at a given phylogenetic level (species, genera, etc.) in a sample as a function of the number of SSU rRNA gene reads collected. The most commonly used metrics of alpha diversity are S_{obs} (the number of species observed in a given number of sequences), Chao1 (a measure based on the number of species observed only once), the Shannon index (a measure of the number of bits of information gained by revealing the identity of a randomly chosen member of the community), and phylogenetic diversity