APPLICATIONS OF NEXT-GENERATION SEQUENCING

The human microbiome: at the interface of health and disease

Ilseung Cho[1,2] *and Martin J. Blaser*[1,2,3,4]

Abstract | Interest in the role of the microbiome in human health has burgeoned over the past decade with the advent of new technologies for interrogating complex microbial communities. The large-scale dynamics of the microbiome can be described by many of the tools and observations used in the study of population ecology. Deciphering the metagenome and its aggregate genetic information can also be used to understand the functional properties of the microbial community. Both the microbiome and metagenome probably have important functions in health and disease; their exploration is a frontier in human genetics.

Microbiota
The microbial organisms that constitute the microbiome. The composition of the microbiota in a community can vary substantially between environmental sites, among host niches and between health and disease.

[1]*Department of Medicine, NYU Langone Medical Center, New York, New York 10016, USA.*
[2]*New York Harbor Department of Veterans Affairs Medical Center (Manhattan), New York, New York 10010, USA.*
[3]*Department of Microbiology, NYU Langone Medical Center, New York, New York 10016, USA.*
[4]*Department of Biology, New York University, New York, New York 10003, USA.*
Correspondence to I.C.
e-mail:
Ilseung.Cho@nyumc.org
doi:10.1038/nrg3182
Published online
13 March 2012

Until recently, the properties of the microbiota of humans (formerly called 'the normal flora') were largely a black box. Cultivation *in vitro*, which has been the cornerstone of microbiology since the nineteenth century, cannot be applied to many of the most densely populated microbial communities[1]. However, DNA-based analyses have expanded our horizon by generating enormous new data sets that can be mined for information on the composition and functional properties of vastly greater numbers of microbial communities. For example, the Human Microbiome Project (HMP)[2] by the US National Institutes of Health has produced a 2.3 terabyte 16S ribosomal RNA metagenomic data set of over 35 billion reads taken from 690 samples from 300 US subjects, across 15 body sites. Large-scale endeavours (for example, the HMP and also the European project, MetaHIT[3]) are already providing a preliminary understanding of the biology and medical significance of the human microbiome and its collective genes (the metagenome).

The aim of these projects, particularly the HMP, is to characterize the compositional range of the 'normal' microbiome of healthy individuals. Important questions concerning the commonalities and differences among healthy individuals in both microbial taxa and functional pathways are being addressed. The presence of major clustering patterns at body sites such as the vagina[4] and the gastrointestinal tract[5] provide new ways to classify individuals and possibly their disease risks. Substantial progress has been made in developing the tools for inquiry and in defining the overarching concepts that advance the field. However, the subject is vast, and the implications for human health and disease are wide-ranging. The study of humans and model animal systems with strong phenotypes is essential for making progress in this field of applied genetics. Although a focus on bacteria is important, inquiries aimed at archaea, viruses and retroviruses are also needed.

The purpose of this Review is to develop the theoretical basis for investigating how microbiome composition and function affect human health. We provide examples of applying this knowledge to better understand human health, and we discuss how microbiome changes could alter host–microbiome interactions to mitigate disease. We also consider the next steps in the development of this field, particularly regarding the need to focus on the inheritance of the microbiome and on its involvement in modulating complex traits.

Characterizing the microbiome

Animals have had residential microbes carrying out metabolic functions for at least 500 million years, at a conservative estimate[6,7]. Extensive congruent phylogenies of animal hosts and their microbiota, involving both individual organisms and whole microbial populations[1,8,9], suggest the existence of specific selection based on co-adaptation. Cooperative interactions between microbes and their hosts typically involve microbial participation in host functions such as defence, metabolism and reproduction[10]. For example, comparing germ-free and normal mice indicates that microbiota are responsible for most of the metabolites that are detected in plasma[11].

Below we describe the efforts to categorize the composition and complex dynamics of our microbiota. Functional variation of host microbiota can be mediated by the introduction or extinction of particular microbial groups or by a change in population structure[12–14]. Such

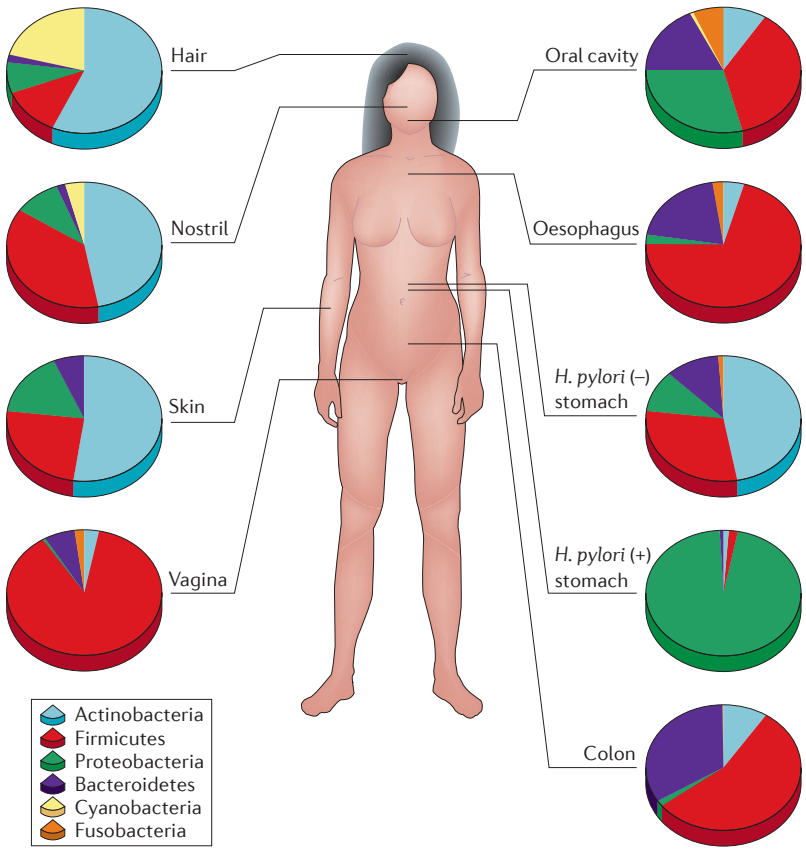

Figure 1 | **Compositional differences in the microbiome by anatomical site.** High-throughput sequencing has revealed substantial intra-individual microbiome variation at different anatomical sites, and inter-individual variation at the same anatomical sites[4,5,25,52,89,93]. However, higher-level (for example, at the level of phyla) taxonomic features display temporal (longitudinal) stability in individuals at specific anatomical sites. Such site-specific differences and the observed conservation between human hosts provide an important framework to determine the biological and pathological significance of a particular microbiome composition. The figure indicates the relative proportion of sequences determined at the taxonomic phylum level at eight anatomical sites. Certain features, such as the presence (+) or absence (–) of *Helicobacter pylori*, can lead to permanent and marked perturbations in community composition[93].

alterations can in turn be induced through selection by environmental factors[10,15], such as dietary changes or exposure to antibiotics[10,15].

Tools for studying the metagenome. The taxonomic diversity that is inherent in complex environmental communities and the task of identifying specific associations with host traits create unique challenges. One approach to metagenome analysis involves assigning unassembled sequences generated by shotgun high-throughput sequencing (HTS)[16] to the NCBI non-redundant Clusters of Orthologous Groups (COG) or the Kyoto Encyclopedia of Genes and Genomes (KEGG) databases[17]. This method facilitates the assessment of interactions that occur within the microbiome, and potentially between a microbiome and its host[18]. However, because a substantial fraction of the metagenome (~33%) is not well-represented by reference genomes, this strategy provides only a limited understanding of the functional potential of the microbiota. An alternative approach is to use catalogues of known genes to identify functional clusters in a sample; such clusters could correspond to the proposed taxonomic enterotypes[5]. A catalogue of the microbial genes present in the human gut, for example, is being generated using several approaches including sequencing, assembling and characterizing non-redundant microbial genes from faecal samples[19], and whole-genome sequencing of reference microbial species[20].

As technologies for sequencing and bioinformatics continue to evolve (see REF. 21 for a review of the state-of-the-art technologies), scientific priorities will include elucidating the 'core' metagenome that occupies a specific human niche and discerning the differences between normal and diseased hosts. As an example of the latter goal, Greenblum *et al.*[22] applied new tools to understand interhost metagenomic variation in relation to phenotypes such as obesity and inflammatory bowel disease (IBD). By categorizing metagenomic sequences based on gene function, they constructed community-level metabolic networks varying in gene abundance, and examined the topological features of these networks in relation to host phenotype. Their analysis identified specific network topologies related to obesity and IBD; skewed topologies chiefly differ in genes related to host interactivity, particularly metabolic functions. Such topological tools can now be applied to explore differences in other host disease states.

Taxonomic variation. The composition of the microbiome varies by anatomical site (FIG. 1). The primary determinant of community composition is anatomical location: interpersonal variation is substantial[23,24] and is higher than the temporal variability seen at most sites in a single individual[25]. The temporal stability observed at an anatomical site suggests that individuals can be grouped according to the major enterotypes present in the colon[5] or the vagina[4]. However, minor perturbations such as dietary changes can rapidly cause substantial intestinal metagenomic changes, and enterotypes are known to cluster based on the dietary abundance of animal protein relative to carbohydrate[26]. Similarly, nasopharyngeal microbiota in young children varies seasonally[24], and vaginal microbiota varies with menses[4].

In the absence of marked perturbations the aggregate microbiota of an individual varies rather narrowly within host-specific boundaries; the basis of such boundaries have not been established, but may represent Nash equilibria[13]. Because minor microbial populations have the potential to bloom, the temporal variation observed in a host may be mirrored by the inter-individual variation observed at a single time point[27]; that the system is dynamic suggests that there are greater interpersonal similarities than a snap-shot view indicates. However, large perturbations such as antibiotic exposure[28] or enteric infections (L.A. David, Harvard Society of Fellows, Cambridge, USA, personal communication), can lead to transient disequilibrium[29] or to the development of a new stable state.

16S ribosomal RNA
A component of the 30S small subunit of prokaryotic ribosomes. Sequencing of the 16S rRNA has been used to identify prokaryotic taxonomy in complete environmental samples such as the microbiome.

Microbiome
The totality of microbes, their genetic information and the milieu in which they interact. Microbiomes typically consist of environmental or biological niches containing complex communities of microbes.

Metagenome
The genetic information of a complex population — typically from microbes in an environmental or host niche sample — that is constituted by the genomes of many individual organisms. The metagenome provides information about the functional genetic potential of the aggregate population.

Among all mammals, the microbiota composition is extensively conserved at high taxonomic levels[7], but variation increases at progressively lower taxonomic levels[30]. Consequently, 85% of the sequences obtained from the distal gut of the mouse represent genera that are not detected in humans[31]. Furthermore, intraspecies variability of the microbiota among human populations is substantial[5]. This degree of variation was unanticipated *a priori*. In retrospect, however, extensive taxonomic variation is unsurprising: a human harbours a climax population of ~10^{14} bacterial cells, can host 10^5–10^6 bacterial generations per human generation, is omnivorous and has accumulated genetic and epigenetic diversity as a host species for >1 million years. Indicator organisms such as *Helicobacter pylori*[32] and *Streptococcus mutans*[33] highlight some differences across the microbiota[34] and metagenome[35] among human ethnic groups; however, the extent of ethnic variation in overall metagenomic composition is unknown. The microbiomes of monozygotic twins are more closely related to one another than to those of unrelated individuals[36,37] but not strikingly so, indicating important postnatal influences on composition.

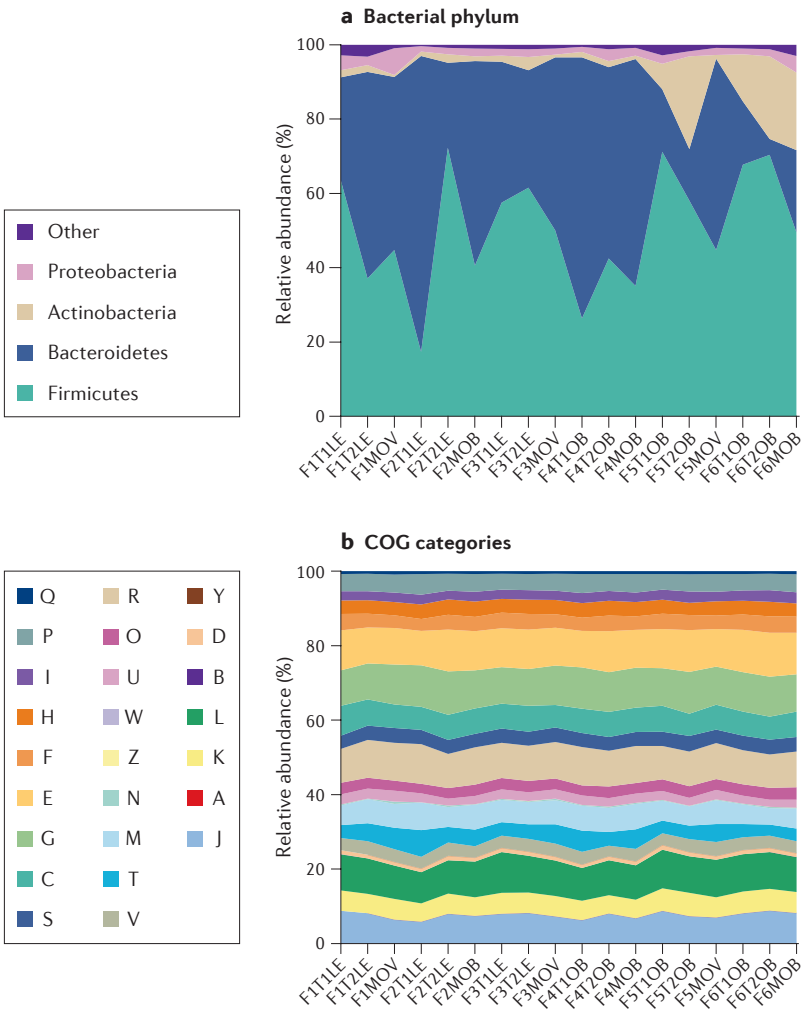

Figure 2 | **Conservation of bacterial genes despite taxonomic variation.** **a** | Turnbaugh *et al.*[37] studied the distal gut microbiome in lean and obese human twins and their mothers. Significant or substantial taxonomic variation was observed among the individuals, although Firmicutes and Bacteroidetes still constituted most of the taxa that were found in the distal gut. **b** | Through metagenomic analyses, the functional characteristics of the microbiome, as identified by Clusters of Orthologous Groups (COG) pathway analysis, are largely conserved, despite the taxonomic variation[37]. The labels on the X axes are individual identifiers. COG pathways are denoted by: A, RNA; B, chromatin; C, energy; D, cell cycle; E, amino acids; F, nucleotides; G, carbohydrates; H, coenzymes; I, lipids; J, translation; K, transcription; L, DNA; M, envelope; N, cell motility; O, protein turnover; P, inorganic; Q, second metabolites; R, general function; S, unknown; T, signal transduction; U, secretion; V, defence; W, extracellular; Z, cytoskeleton. Images are reproduced, with permission, from REF. 37 © (2009) Macmillan Publishers Ltd. All rights reserved.

Functional variation. The extensive lower-level taxonomic variation and large compositional differences observed even between highly related host organisms (for example, mice and humans) is counterbalanced by the substantial conservation of metagenomic core functions (FIG. 2). This reflects the conservation of core bacterial properties involved in nucleic acid and protein synthesis, and in metabolic and structural requirements. Of the >50 known phyla, most of the human microbiota is composed of <10 (and mostly 6) phyla. Bacteria from other phyla, usually of plant origin, may be present in skin, nasopharyngeal or gut samples[24,25,38]; however, these are generally infrequent (representing <0.01% of the sequences) and probably represent transient carriage from food- and air-borne exposures.

Why did the particular restriction of diversity to only a few phyla evolve not only in humans, but also in perhaps all other vertebrates? One possibility is that within the conserved boundaries for the microbiome that are permitted by the human genome, there exists a large range of contingency organisms and contingency genes. According to this hypothesis, the genes may only be active at some moment in the lifespan of the host, or perhaps at a frequency of less than once per lifespan.

The parallel needs of individual bacteria lead to both competition for key substrates and to functional redundancy in the microbiota. Nevertheless, the enormous bacterial biomass also provides many unique or minimally redundant bacterial genes[19].

Resilience and community disturbance. Resilience, the ability to withstand disturbance, is a central concept in ecology. The resilience of the human colonic microbiome is beautifully illustrated by recent studies of twins that were examined before, during and after 7 weeks of ingesting a fermented milk product containing a sample of ~10^8 *Bifidobacterium*, *Lactobacillus*, *Lactococcus* and *Streptococccus* species[36]. Despite the daily oral inoculations, the composition of the microbiota at the 16S rDNA level and the metagenome were essentially unaffected. Although the microbiome of human adults seems to be highly resilient, no comparable studies have been carried out in children; microbial population structures seem to be more dynamic in children[39], and so resilience may be lower. An important natural experiment has been occurring over the past 70 years in which most of the world's population have been exposed to pharmacological doses of antimicrobial

agents. Such usage has been based on the implicit belief that the human microbiome is completely resilient and returns to the pretreated composition after antibiotic-induced perturbation. However, studies of indicator organisms, such as *H. pylori*, show that individual hosts can experience extirpations of bacterial species[40]. Medium- and long-term selection of resistant organisms and the destabilization of the microbiome with new species compositions are also seen, even in the absence of further antibiotic exposure[28,41]. Thus, despite the extensive resilience that is inherent in a complex ecosystem, there may be loss of recovery from continued perturbations[29], which has important implications for human health[42].

Medical scientists are familiar with Koch's postulates, which are used as criteria to determine whether a microbe causes disease[43]. However, when considering the pathogenicity of the microbiome it might be better to focus on community characteristics, which are largely governed by richness, composition and interactions among the constituent members[7,16,44]. Substantial perturbation (community disturbance[45]) tests the resilience of the community, such as its ability to resist invasion by exogenous microbes; stable diverse communities resist pathogens[46]. At present, 16S rRNA analyses focus on taxonomic differences at or above the species level. However, examinations below the subspecies level, relating to strains or even alleles, may be more informative. However, the technology (particularly the informatics tools) are not yet sufficiently developed for these applications.

Extinctions. The human microbiome represents one or more complete ecosystems. The trophic organization of species-rich communities is similar to other complex network topologies, in that it shows extreme heterogeneity and is dominated by a few highly connected nodes[47]. Such communities may resist random perturbations but if keystone species[48] are lost, effects may cascade, causing secondary extinctions; high biodiversity diminishes this risk[12]. The substantial nonlinear interactions present in complex, co-evolved systems ensure that ecological networks are robust against random removals[49]. However, if a system is repeatedly perturbed, the effects of gene loss can be amplified by downstream effects on co-colonizing microbes and on the host. Because of allelopathy, the effects of extinctions may be magnified[50]. In the short-term, functional redundancy may mask extinction effects but in the longer-term, extinctions lead to losses of contingency responses and cause ecological crashes[49]. Considering the importance of guilds of bacteria that exploit parallel and sequential metabolic pathways, these concepts are relevant to the human metagenome. As a result of modern lifestyles, horizontal microbial transmission has been diminishing, and there has been unprecedented selection against existing, long-present microbes[40]. An example is provided by the loss of a dominant species, *H. pylori*, from the human stomach[51,52], which has led to this body site harbouring alternative, stable states characterized by the presence or absence of *H. pylori*.

In summary, as with other complex ecosystems, the microbiomes that populate specific human anatomical niches are species-rich, but possess particular overall community characteristics at higher organizational levels. All are subject to perturbation in the course of normal development and ageing, and especially with disease. As our knowledge of the fundamental characteristics and biology of the human microbiome grows, so will our ability to understand disease-related variation.

Influences on the microbiota during host life cycles

As described in the previous section, differences in microbiota composition exist across body sites and among individuals. However, changes are also evident across the human lifespan. Important questions in this field involve determining whether such temporal changes are life-stage-specific, and whether they are predetermined by host genetic characteristics or by environmental factors.

Inheritance of microbiota. The congruent phylogenies of mammals and their microbiota[8] provide strong evidence for the inheritance of the microbiota[7]. Although inheritance of the microbiota from the father is presently little studied, increasing evidence supports inheritance from the mother[34,53]. Until the amniotic sac ruptures, a fetus is considered to be sterile, or essentially sterile. Immediately after vaginal delivery, founding microbial populations in the baby closely resemble that of their mother's vagina[54], with lactobacilli predominating. Because lactic-acid-producing bacteria dominate in both the mother's vagina and milk, the initial bloom of lactobacilli in the baby's gastrointestinal tract cannot be considered accidental. Lactobacilli represent the pioneer community in mice[55] and humans[39], in which they prepare the gastrointestinal tract for subsequent microbial successions until microbial maturity is reached.

The repeated opportunities for the microbiota to be transferred from a mother to her baby may be disrupted by modern lifestyles. The availability of delivery by Caesarean section, as opposed to vaginal delivery, is an obvious example of the potential impact of medical practice on microbiota composition; substantial differences in the founding microbiota population[54] can persist for months[56] (FIG. 3). In many host species, paternal contributions to offspring traits have been well documented[57,58]; these observations have been extended to the microbiome, in which paternal contributions to *H. pylori* allele composition in the offspring have been shown[59]. In any event, there is evidence for extensive horizontal transfer of microbial genes within human populations, involving microbes in different functional classes and inhabiting different ecological niches[60], indicating the site-specificity and dynamism of selection on the human microbiome. Even so, microbial inheritance can provide important confirmation of human ancestry[61].

In *Drosophila melanogaster*, microbial influences have an effect on mating preference for >30 generations[62]. Could microbiome composition therefore affect mating

Extinction
The loss of an organism or group of organisms (usually of a species) from an ecosystem.

Enterotype
A recently proposed classification unit of animals that is based on the bacteriological composition of their gut microbiome. There are reported to be at least three distinct enterotypes, which are independent of ethnic background and diet.

Nash equilibria
Concepts from game theory in which players know the strategies of the others, and in which any change from their strategy puts them in a less favourable position.

Resilience
A term in ecology indicating the capacity of a system to absorb disturbance and to reorganize itself while undergoing change, so as to retain essentially the same function, structure and identity.

Extirpations
The loss of species in a locality (for example, an individual host).

Allelopathy
A phenomenon in which a microbe uses chemical means to aid its competition within a group of microbes. Allelopathy may involve manipulation of third parties (for example, the host) to favour competition.

Mating preference
The selection or choice of sexual partners that is often based on traits of a potential mate. Genetic differences between selected and non-selected hosts are a source of selectable variation.

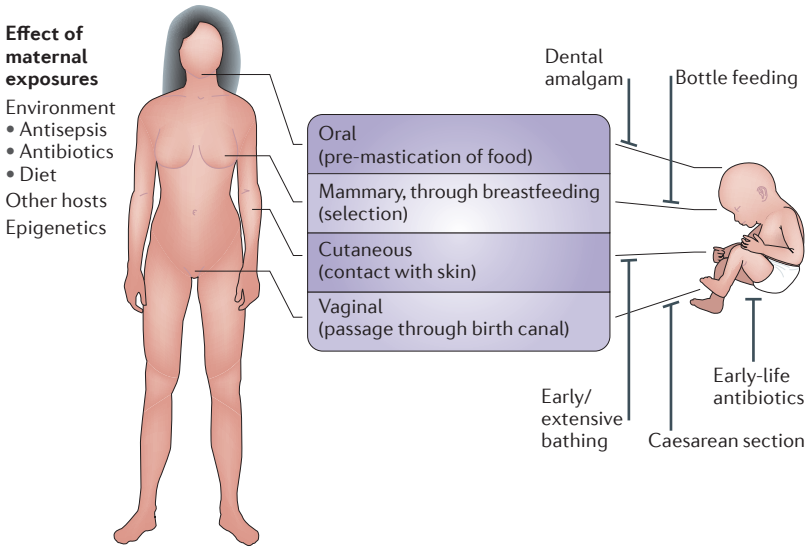

Figure 3 | **Acquisition of the microbiome in early life by vertical transmission, and factors modifying mother-to-child microbial transmission.** Through live-birth, mammals have important opportunities for mother-to-child microbial transmission through direct surface contact. However, many modern practices can reduce organism and gene flow; several examples are illustrated. After initial introductions, there is strong selection by hosts for microbes with specific phenotypes, consistent with the extensive conservation shown in FIG. 1. Acquisition is modified by differences in offspring genetics and epigenetics (with respect to both maternal and paternal genes) that inform the competition for host resources by the vertically transmitted or environmentally acquired microbes. Ancestral organisms that have particular tissue-specific and niche-specific adaptations facilitate tissue tropisms and are selected for, thus explaining the conserved niche-specificity compositions.

in humans? Odour is one means to affect mating preference, and human axillary and oral odours are largely influenced by microbial products, especially mercaptans[63]. In general, the greater the force of mating preference, the more likely that those populations will become sexually isolated[64,65]; this could affect tribal differentiation and other ethnic differences in humans. We speculate that metagenome composition has affected mating preference in humans, representing another phenotype under strong selective pressure.

Postnatal influences on the microbiota. Over a lifetime, each human develops a densely populated microbiome, a process that is recapitulated in every individual and in every generation. The eruption of teeth is responsible for major successions in the oral microbiota[66,67], suggesting that succession may be a general property of microbiome dynamics in humans. In mice, succession clearly occurs in the gastrointestinal tract[68]. Exposure (or not) to environmental microbes is another important, but highly variable, reservoir for the resident microbiota. Antibiotic use in early life produces major shifts in both microbiota characteristics and in host developmental phenotypes, in both farm animals[69] and experimental animals[70,71]. Whether such precedents are applicable to human children is unknown, but it seems likely. If so, then both the timing of microbiome succession and the specific organisms that are present may affect development. The concept of time-dependent compositional variation affecting host immunological, metabolic, cognitive and reproductive development is a potentially important and testable hypothesis. We further speculate that nature orchestrates microbiome development to optimize fecundity, reaching a climax state at or near parturition to maximize success for the next generation of hosts. The noted heterozygote advantage for fecundity[72] may be an analogue of harbouring a genetically diverse microbiota.

Microbiome dynamics in adults. Our knowledge of microbiome dynamics, especially age-related changes during human adulthood, is limited. The older literature (predating the use of HTS), clearly shows that the postmenopausal vaginal microbiota differs substantially from that during the reproductive period[73,74]. Similarly, in the stomach, the age-related progressive development of gastric atrophy (which is enhanced by the presence of *H. pylori*[75,76]) selects for gastric microbiota that are substantially different from those that are found in the stomach of younger, *H. pylori*-negative hosts[77]. Analogous changes may be occurring in other body sites as senescence advances. In the gut, the ratio of Bacteroidete to Firmicute species changes with age[78].

These concepts are particularly relevant to oncogenesis, which is generally age-related. In the multistep Nordling hypothesis of oncogenesis[79], 4–6 somatic cell mutations are needed for cancer development. We propose that shifts in age-related microbiota contribute to this multistep process. Residential microbes can contribute to somatic mutagenesis by causing genotoxicity as a result of inflammation, increased cell proliferation and the production of pro-mutagenic metabolites (for example, butyrate)[80]. Genes may have alternative effects at different life stages, illustrating the idea of antagonistic pleiotropy[81]. We hypothesize that specific human microbiota and their genes that are beneficial early in life may be harmful later in life. The dominant gastric bacterium *H. pylori* provides an example: early in life, inflammatory responses in the host improve the control of infection[82,83] and allergy[84], but later in life promote atrophy and oncogenesis[85]. A related hypothesis is that co-evolved microbiota are adaptive for the human species both by supporting early-in-life host functions and by leading to later-in-life host demise[86].

Disease links and health implications

Overall, how does the microbiome affect human health? Current studies focus on describing the variant microbe populations that occur in specific disease states, or the temporal microbial changes that are observed over the course of a disease. For many conditions, the challenge is to discover whether there is a causal link between microbiome variation and pathology. Unfortunately, limitations in the definitions and stratification of clinical syndromes, including irritable bowel syndrome and non-ulcer dyspepsia (NUD), reduce the potential of microbiome studies. Below, we review some recent investigations into specific diseases (TABLE 1); these investigations are preliminary but some observations are promising.

Table 1 | **Examples of associations of human conditions with particular microbiota characteristics**

Disease	Relevant finding	Refs
Psoriasis	Increased ratio of Firmicutes to Actinobacteria	88
Reflux oesophagitis	Oesophageal microbiota dominated by gram-negative anaerobes; gastric microbiota with low or absent *Helicobacter pylori*	75,133
Obesity	Reduced ratio of Bacteroidetes to Firmicutes	17,31
Childhood-onset asthma	Absent gastric *H. pylori* (especially the cytotoxin-associated gene A (*cagA*) genotype)	96,134
Inflammatory bowel disease (colitis)	Larger populations of Enterobacteriaceae	113
Functional bowel diseases	Larger populations of *Veillonella* and *Lactobacillus*	135
Colorectal carcinoma	Larger populations of *Fusobacterium spp.*	101,102
Cardiovascular disease	Gut-microbiota-dependent metabolism of phosphatidylcholine	136

The cutaneous microbiome. The cutaneous microbiome is an obvious target in specific diseases such as psoriasis, a chronic, idiopathic inflammatory dermatological condition[87]. In studies predating HTS, the use of PCR and cloning led to observations that Firmicute species were significantly over-represented and that Actinobacteria were significantly under-represented in psoriatic lesions compared with both unaffected skin in patients with psoriasis and in unaffected controls[88]. Studies to explore these findings using HTS are currently underway[89]. Atopic dermatitis, another chronic inflammatory condition, has increased in incidence approximately threefold over the last 30 years in industrialized countries, suggesting a potential role for microbiome alterations. Classic atopic dermatitis occurs in skin regions, such as the antecubital fossae and the popliteal fossae, that have similar microbial populations[89], suggesting a microbiome role. Similarly, *Propionibacterium acnes* has been implicated in the common dermatological condition, acne. *P. acnes* thrives in the cutaneous pilosebaceous units, secretes enzymes that cause local injury and inflammation, and is widely accepted to have a function in acne development[90]. However, investigations are ongoing to examine the involvement of other microbes in the development of acne. Chronic skin ulcers, which are often secondary to venous stasis or diabetes, lead to substantial morbidity. Cutaneous microbiome shifts have been noted in these conditions, such as an increased abundance of Pseudomonadaceae in patients with chronic ulcers that were treated with antibiotics, and an increased abundance of Streptococcaceae in diabetic ulcers[91]. Such shifts may interact with aberrantly expressed host cutaneous defence response genes[92], thereby increasing disease risk.

The gastric microbiome. The discovery that *H. pylori* was adapted to survive in the acidic gastric environment overturned the dogma that the stomach is sterile. In *H. pylori*-negative individuals, gastric microbiota diversity is high; most of the prominent gastric phylotypes (*Streptococcus*, Actinomyces, *Prevotella* and *Gemella*) also are abundant in the oropharynx of these individuals[93]; this indicates either that many constituents are swallowed from more proximal sites, or that close relatives of the oral microbiota colonize more distally. By contrast, among *H. pylori*-positive individuals, *H. pylori* usually accounts for >90% of sequence reads from the gastric microbiota[93], markedly reducing the overall diversity of this microbiota. The ability of *H. pylori* to dominate the gastric microbiota indicates an evolved fitness for that specialized niche. *H. pylori* is a classical amphibiont; the presence (or absence) of an *H. pylori*-dominated gastric microbiota is strongly associated with particular diseases that show important age-related differences[85]. Its presence increases risks for developing peptic ulcer disease, gastric mucosa-associated lymphoid tissue (MALT) tumours, and gastric adenocarcinomas[94]. Conversely, it is also associated with a decreased risk of reflux oesophagitis[95] and childhood-onset asthma[96], thus demonstrating the complex biological interactions between hosts and microbiota.

The colonic microbiota and colorectal cancer. The colonic microbiota has been suspected for a long time to be involved in the development of colorectal cancers[97], possibly by synthesizing short-chain fatty acids (SCFAs) and other metabolites. SCFAs, in particular butyrate, may induce apoptosis, cell cycle arrest and differentiation, through WNT signalling[98]. Microbes may also be genotoxic to colonic epithelial cells, as demonstrated by the induction of aneuploidy and tetraploidy by *Enterococcus faecalis*[99]. The colonic microbiota might also promote colorectal cancer by eliciting host responses, for example, by stimulating exaggerated immune responses, potentially through T helper 17 (Th17) cells[99].

Further evidence of a link between colonic microbiota and colorectal cancer is suggested by the ability of antibiotic administration to not only alter the composition of the colonic microbiota but also to affect the expression of host genes that are involved in cell cycle regulation, thus reducing epithelial proliferation[100]. Early studies evaluating specific microbes were limited to identifying culture-dependent species, such as *Streptococcus bovis,* but could not adequately assess anaerobic constituents. However, members of the anaerobic genus *Fusobacterium* have recently been associated with colorectal cancer: whole-genome sequences of *Fusobacterium* species were compared between tumour tissue and matched normal colon tissue using both quantitative PCR analysis and HTS[101,102]. *Fusobacterium nucleatum* is a mucosally adherent, pro-inflammatory microbe that was first

Antecubital fossae
The triangular areas on the anterior (flexor) aspects of elbow joints.

Popliteal fossae
The shallow depressions that are found on the flexor aspects of knee joints.

Pilosebaceous units
The anatomic structure around each hair shaft that consists of the hair shaft and follicle, the sebaceous gland and the erector pili muscle.

Amphibiont
An organism (for example, a microbe) that may have a pathogenic or symbiotic relationship with another organism (for example, its host), depending on context. This is a more specific term than commensal.

identified in the mouth[103]. In colorectal cancer samples, *F. nucleatum* sequences were significantly enriched compared with samples obtained from control tissue, while both Bacteroidetes and Firmicutes were depleted relative to other bacteria in *Fusobacterium*-rich malignancies[102]. The enrichment of *Fusobacterium* species (not limited to *F. nucleatum*) was confirmed when evaluating the mucosal microbiome of colorectal cancers compared to adjacent normal tissues in an expanded collection of 99 biopsies[101]. However, the causal direction of the association has not yet been ascertained.

The colon microbiota and inflammatory bowel disease. The microbiome is essential for the activation of host immune responses[104]. For example, Th17 cell differentiation in the mouse lamina propria requires the presence of segmented filamentous bacteria (SFB)[105], and polysaccharide A produced by *Bacteroides fragilis* mediates the conversion of $CD4^+$ T cells into regulatory T cells[106]. The inflammatory bowel diseases have long been considered to reflect interactions between microbes and the host. IBD susceptibility is associated with host polymorphisms in bacterial sensor genes such as nucleotide-binding oligomerization domain-containing 2 (*NOD2*; also known as caspase recruitment domain-containing protein 15 (*CARD15*))[107,108] and Toll-like receptor 4 (*TLR4*)[109], and symptoms in patients with IBD sometimes improve following antibiotic treatment[110]. Early childhood exposure to antibiotics has been associated with a significantly increased risk for Crohn's disease[111], suggesting that gut microbiome perturbations are important for disease risk. Microbial diversity is significantly diminished in Crohn's disease[112], suggesting a decreased gut microbiome resilience that could affect immune interactions. Gut microbiome population structures in patients with ulcerative colitis or Crohn's disease[19] depart from normality, but remain clustered by disease within their characteristic deviated patterns. Specific bacteria of the Enterobacteriaceae family may act together with a disordered microbiome to increase the risk of ulcerative colitis[113]. Between twins that are discordant for ulcerative colitis, those affected had significantly reduced bacterial diversity, but increased proportions of Actinobacteria and Proteobacteria[114]. Patients with Crohn's disease have over-representation of *Enterococcus faecium* and of several Proteobacteria compared with controls[115]. The microbial patterns observed for the conditions described above are preliminary, and their specificity and causal direction have not been established.

The gut microbiota and diseases of the liver. The gut microbiota may be involved in hepatological conditions, including non-alcoholic fatty liver disease (NAFLD)[116], alcoholic steatosis and hepatocellular carcinoma. The liver is the first solid organ to be exposed to the metabolic products that are generated by the gut microbiome, including acetaldehyde, ammonia and phenols. Compared with germ-free mice, the presence of a microbiome in normal mice leads to the suppression of intestinal epithelium angiopoietin-related protein 4, which normally inhibits lipoprotein lipase; this microbiome-mediated effect consequently increases downstream triglyceride accumulation in the hepatic parenchyma and adipocytes[117]. Chronic exposure to ethanol disturbs the gut microbiome[118,119], but roles for the microbiome in steatosis are unresolved. Particular colonic commensals of the mouse (for example, *Helicobacter hepaticus*) promote the development of hepatocellular carcinoma[120]. Patients with cirrhosis have a substantially altered microbiome, including community-wide changes at multiple taxonomic levels, with enrichment of Proteobacteria and Fusobacteria phyla, and of Enterobacteriaceae, Veillonellaceae and Streptococcaceae families[121]. Although many observations suggest links between microbiome composition and liver disease, definitive associations in humans are lacking.

The gut microbiota and obesity. Genetically obese (*ob*/*ob*) mice have decreased Bacteroidetes/Firmicutes ratios compared with their lean (*ob*/+ and +/+ wild-type) siblings[31]. Transplantation of gut microbiota from the obese (*ob*/*ob*) to germ-free mice conferred an obese phenotype, demonstrating the transmissibility of metabolic phenotypes[17]; the transferred microbiomes had increased capacities for energy harvest. In humans, the relative proportions of members of the Bacteroidete phylum increase with weight loss[122]. In studies of monozygotic and dizygotic twins, obesity was associated with smaller populations of Bacteroidetes, diminished bacterial diversity and enrichment of genes related to lipid and carbohydrate metabolism. Despite substantial taxonomic variation, functional metagenomic differences were minor[37]. Modern lifestyles that change the selection pressures on microbiomes could alter exposures to bacteria during the early lives of hosts and thus may contribute to the development of obesity. Antibiotic use in human infancy (before the age of 6 months) was significantly associated with obesity development[123]. By contrast, perinatal administration of a *Lactobacillus rhamnosus* probiotic decreased excessive weight gain during childhood[124]. These early studies provide support for the concept that perturbations in microbiota could lead to childhood-onset obesity, which might be modifiable. Alterations in the gut microbiome also occur when interventions are used to treat obesity. Roux-en-Y surgery significantly increases levels of *Proteobacteria* and alters specific metabolic markers, such as the production of urinary amines and cresols[125].

The gut microbiota and rheumatoid arthritis. Dysregulation of host responses as a consequence of dysbiosis in the gut lumen could affect distant anatomical sites through the activation of host immune responses. This could be the mechanism that contributes to rheumatoid arthritis, which is another chronic idiopathic inflammatory condition. In mice, the presence of SFBs in the gut microbiome causes the local expansion of Th17 cells[126], which then migrate to peripheral immune compartments and activate B cells into antibody-producing plasma cells. Antibody production leads to the immune-mediated destruction of joints, which occurs in rheumatoid arthritis[127].

Lamina propria
A thin layer of loose connective tissue that lies underneath the epithelium; collectively these tissues constitute the mucosa that line various lumens in the body. The lamina propria is densely populated by immunological and inflammatory cells.

Steatosis
The pathological accumulation and retention of lipids in liver parenchymal cells. Substantial steatosis can compromise cellular functions and is associated with disease processes, including alcoholism, diabetes and hyperlipidaemia.

Commensals
Organisms (for example, microbes) that are involved in a form of symbiosis in which one organism derives a benefit while the other is unaffected.

Probiotic
Living microorganisms that are thought to confer a benefit to the host.

Roux-en-Y surgery
A type of gastric bypass surgery that is primarily used for the treatment of morbid obesity. In Roux-en-Y surgeries, a portion of the small bowel is bypassed to decrease the absorption of nutrients.

Dysbiosis
A condition in which the normal microbiome population structure is disturbed, often through external pressures such as disease states or medications.

Table 2 | **The use of mouse models in microbiome studies**

Model	Advantages	Disadvantages
Inbred mice[137]	• Inexpensive • Often well-characterized • Genetically homogeneous • Allow the study of pathogenetic mechanisms	• Poorly controlled microbial variability • Limited translation potential to humans
Gnotobiotic mice[138]	• Well-controlled microbial variability • Allow for better understanding of specific microbe interactions • Genetically homogeneous • Allow mechanistic studies	• Expensive • Difficult to maintain • Limited translation potential to humans • Physiologically less well-understood than conventional animals
Humanized mice[130]	• More relevant to human disease states • Genetically homogeneous	• Expensive • Difficult to maintain • Physiologically less well-understood than non-chimeric animals
Conventionalized gnotobiotic mice[139]	• Well-controlled microbial variability • Allow for better understanding of specific microbe interactions • Genetically homogeneous • Allow mechanistic studies • Do not require specialized xenobiotic facilities	• Physiologically less well-understood than normal animals • 'Artificial' colonization using known microbiota may not be representative of the real-world situation

Cause or effect?

Microbiome analysis in humans has been largely based on observation; that is, associations of disease phenotypes with particular microbiota constituents. But which is causal: does factor A cause factor B, does factor B cause factor A or does factor C cause both A and B? Hill[128] developed criteria to address the questions: "In what circumstances can we pass from this observed association to a verdict of causation? Upon what basis should we proceed to do so?" The criteria include: the strength of association, including its consistency, specificity, temporality and biological plausibility; whether biological gradients are present; whether experimental support exists; and whether support can be extrapolated from known causal relationships. Although these criteria were advanced largely to unravel epidemiological relationships, they are applicable to genetics and particularly to metagenomics. Sometimes successful treatment trials that achieve the amelioration or cure of a particular condition provide the crucial evidence for a causal relationship. The changed natural history of peptic ulcer disease following the elimination of *H. pylori*[129] demonstrated the pathogenic role of this bacterium.

Model organisms provide an important approach for understanding causation and pathogenesis. Animal models approximate some human diseases (for example, asthma and atherosclerosis), but cannot yet accurately reproduce many diseases (for example, psoriasis). For diseases that can be studied in model organisms, the roles of microbiota can be explored within the constraints of particular model systems (TABLE 2). Standard models of inbred mice are limited by their uncontrolled microbiome diversity. Certain disease states are well-studied in these models, such as the effects of SFBs on Th17 development or the susceptibility to type 1 diabetes in non-obese diabetic (NOD) mice. The use of gnotobiotic mice eliminates the above-mentioned microbiome variability, but the animals are expensive and require specialized facilities and expertise, which limits their widespread use. The recent availability of gnotobiotic animals from commercial sources permits the conventionalization of animals harbouring experimental or control microbiota without needing xenobiotic facilities; such approaches allow for the direct observation of the effects of microbiota on the host. The extension of this concept to humanized model organisms[130] allows better approximation of the effects of the human microbiome on disease processes in tractable animal models.

Perspectives

Inherent complexities in the composition of the microbiome may preclude investigations of microbe-associated diseases using classical approaches such as Koch's postulates. Instead of single organisms being associated with disease, community characteristics (such as composition and metagenomic functionality) may be more relevant. The principles of host interactions with pathogens and commensals share many features, and this may inform the new field; however, the nature of the selection for commensalism is more complex and dynamic than pathogen–host interactions. The scale of the interface suggests that microbiome–host interactions have important bearings on disease susceptibility, and the microbial effects on the balance of host metabolism and immunity[131] provide an excellent model for the broader phenomenon of disease susceptibility. Altering metabolic, immunological or developmental pathways are obvious strategies for modifying disease risk.

Given the ongoing extinction of our ancient commensal organisms, the future of a healthy human microbiome may include the restoration of our ancestral microbial ecology. There are two types of restoration. The first involves restoring ancient organisms (or pathways) in healthy hosts that lack them, as prophylaxis against future imbalances. The second type of

Gnotobiotic
Describes an animal that is colonized solely by known strains of bacteria or other microorganisms. The term also describes germ-free animals, as the status of their microbial communities is known.

Conventionalization
A method in which germ-free animals (particularly mice) are inoculated with gut microbiota to populate the gastrointestinal tract.

Prebiotics
Food ingredients that confer specific changes in the gut microbiome and lead to beneficial effects in the host.

Operational taxonomic unit
(OTU). The smallest phylogenetic unit described by variations in 16S ribosomal RNA sequencing. Dissimilarity of <1% in 16S rRNA sequences has commonly been used to define an OTU but <3% and <5% have also been used.

Non-coexistence
An exclusivity scenario in which the abundance of one species leads to another species being less abundant than would be expected by chance.

Box 1 | Ten areas of microbiome inquiry that should be pursued

- Understanding microbiome characteristics in relation to families: which features are inherited and which are not?*
- Understanding secular trends in microbiome composition: which taxonomic groups have been lost or gained?‡
- For diseases that have changed markedly in incidence in recent decades, do changes in the microbiome have a role? Notable examples include childhood-onset asthma, food allergies, type 1 diabetes, obesity, inflammatory bowel disease and autism.*‡
- Do particular signatures of the metagenome predict risks for specific human cancers and other diseases that are associated with ageing? Can these signatures be pursued to better understand oncogenesis? (Work on *Helicobacter pylori* provides a clear example of this.)*
- How do antibiotics perturb the microbiome, both in the short-term and long-term? Does the route of administration matter?*
- How does the microbiome affect the pharmacology of medications? Can we 'micro-type' people to improve pharmacokinetics and/or reduce toxicity? Can we manipulate the microbiome to improve pharmacokinetic stability?*‡
- Can we harness knowledge of microbiomes to improve diagnostics for disease status and susceptibility?*
- Can we harness the close mechanistic interactions between the microbiome and the host to provide hints for the development of new drugs?‡
- Specifically, can we harness the microbiome to develop new narrow-spectrum antibiotics?‡
- Can we use knowledge of the microbiota to develop true probiotics (and prebiotics)?*‡

*Areas currently under investigation. ‡Proposed areas for investigation.

restoration could be therapeutic, when the aetiological extinctions and imbalances are recognized. This scientific frontier will require an understanding of the biology of re-introductions, as well as developing microbial breeding programmes. In addition to the technical problems that are associated with restoring particular organisms in specific hard-to-reach niches (such as the distal ileum), there also will be substantial biological problems related to understanding how re-introductions affect the population structure of the extant organisms and their interactions with the host.

To better understand the implications of microbiota and metagenome variation in human health and disease, the field needs improved informatics tools, including new approaches for understanding the complexity of the metadata[60]. The multidimensionality of the human and microbial phenotypes (and the dynamic, nonlinear interactions) creates challenges for identifying deterministic solutions. For example, in analyses of 16S-rRNA-defined operational taxonomic unit (OTU) populations in mice that received either traditional or Western-type diets, Reshef *et al.*[132] examined top-scoring nonlinear abundance relationships. These often involved 'non-coexistence' that was sometimes related to known factors (for example, diet or host gender) but was often unexplained. Although incomplete, such work leads to new approaches for understanding the underlying complexity.

We also will need new tools to implement the findings of metagenomic analyses that are relevant to human health. Although the known principles of evolution and genetics apply, studies of microbiomes provide new applications, and will lead to new understandings of complex traits. Important questions to pursue are listed in BOX 1. As such, this is a frontier for human preventive medicine and for the medical management of chronic diseases.

1. Baumann, P. & Moran, N. A. Non-cultivable microorganisms from symbiotic associations of insects and other hosts. *Antonie van Leeuwenhoek* **72**, 39–48 (1997).
2. Turnbaugh, P. J. *et al.* The Human Microbiome Project. *Nature* **449**, 804–810 (2007).
3. Ehrlich, S. D. *Metagenomics of the Human Body* (ed. Nelson, K. E.) 307–316 (Springer, 2011).
4. Ravel, J. *et al.* Vaginal microbiome of reproductive-age women. *Proc. Natl Acad. Sci. USA* **108**, 4680–4687 (2011).
This study describes vaginal microbiome differences and similarities in women of reproductive age who vary by ethnicity, and explores factors that are related to bacterial vaginosis.
5. Arumugam, M. *et al.* Enterotypes of the human gut microbiome. *Nature* **473**, 174–180 (2011).
This paper proposes enterotype classifications that are defined by the intrinsic characteristics of the gut microbiome, and that seem to be independent of ethnic or dietary factors.
6. Morris, S. C. & Peel, J. S. The earliest annelids: lower Cambrian polychaetes from the Sirius Passet Lagerstatte, Peary Land, North Greenland. *Acta Palaeontol. Pol.* **53**, 135–146 (2008).
7. Ley, R., Lozupone, C. A., Hamady, M., Knight, R. & Gordon, J. Worlds within worlds: evolution of the vertebrate gut microbiota. *Nature Rev. Microbiol.* **6**, 776–788 (2008).
A review that contrasts the microbial communities in the vertebrate gut with each other and with free-living microbial communities.
8. Ochman, H. *et al.* Evolutionary relationships of wild hominids recapitulated by gut microbial communities. *PLoS Biol.* **8**, e1000546 (2010).
9. Moran, N. A., Munson, M. A., Baumann, P. & Ishikawa, H. A Molecular clock in endosymbiotic bacteria is calibrated using the insect hosts. *Proc. R. Soc. Lond. B* **253**, 167–171 (1993).
10. Benson, A. K. *et al.* Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. *Proc. Natl Acad. Sci. USA* **107**, 18933–18938 (2010).
11. Wikoff, W. R. *et al.* Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc. Natl Acad. Sci. USA* **106**, 3698–3703 (2009).
A comparison of germ-free and normal animals, which shows that the microbiome has substantial effects on host blood metabolites, including on the metabolism of amino acids and organic acids.
12. Petchey, O. L., Eklof, A., Borrvall, C. & Ebenman, B. Trophically unique species are vulnerable to cascading extinction. *Am. Nat.* **171**, 568–579 (2008).
13. Blaser, M. J. & Kirschner, D. The equilibria that allow bacterial persistence in human hosts. *Nature* **449**, 843–849 (2007).
The authors of this paper propose that co-evolved bacteria in human hosts establish homeostases that conform to the principles of Nash equilibria. Understanding such equilibria may provide insight into shifts in microbial communities in health and disease.
14. Maynard Smith, J. *Models in Ecology*. (Cambridge Univ. Press, UK, 1974).
15. Blaser, M. J. Who are we? Indigenous microbes and the ecology of human diseases. *EMBO Rep.* **7**, 956–960 (2006).
16. Tringe, S. G. *et al.* Comparative metagenomics of microbial communities. *Science* **308**, 554–557 (2005).
17. Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027–1031 (2006).
A seminal paper describing the ability of the gut microbiome to extract energy from dietary sources.
18. Warnecke, F. *et al.* Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* **450**, 560–565 (2007).

19. Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).
The authors report the identification of a library of microbial genes that are found in the human gut microbiome using high-throughput metagenomic sequencing.
20. Nelson, K. E. *et al.* A catalog of reference genomes from the human microbiome. *Science* **328**, 994–999 (2010).
21. Kuczynski, J. *et al.* Experimental and analytical tools for studying the human microbiome. *Nature Rev. Genet.* **13**, 47–58 (2012).
22. Greenblum, S., Turnbaugh, P. J. & Borenstein, E. Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. *Proc. Natl Acad. Sci. USA* **109**, 594–599 (2012).
A new method of comparing metagenomic data that involves analysing metabolic networks and their associated genes to describe changes that occur in disease (such as in obesity or IBD).
23. Eckburg, P. B. *et al.* Diversity of the human intestinal microbial flora. *Science* **308**, 1635–1638 (2005).
24. Bogaert, D. *et al.* Variability and diversity of nasopharyngeal microbiota in children: a metagenomic analysis. *PLoS ONE* **6**, e17035 (2011).
25. Costello, E. K. *et al.* Bacterial community variation in human body habitats across space and time. *Science* **326**, 1694–1697 (2009).
This study describes temporal and topographical variations in the human microbiome at various anatomical sites.
26. Wu, G. D. *et al.* Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105–108 (2011).
27. Kuczynski, J. *et al.* Direct sequencing of the human microbiome readily reveals community differences. *Genome Biol.* **11**, 210 (2010).
28. Dethlefsen, L. & Relman, D. A. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc. Natl Acad. Sci. USA* **108**, 4554–4561 (2011).
This paper describes the substantial alterations that occur in the gut microbiome after exposure to antibiotics. It also highlights varied taxonomic changes among individuals.
29. Dethlefsen, L., Huse, S., Sogin, M. L. & Relman, D. A. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol.* **6**, e280 (2008).
30. Huse, S. M. *et al.* Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genet.* **4**, e1000255 (2008).
31. Ley, R. E. et al. Obesity alters gut microbial ecology. *Proc. Natl Acad. Sci. USA* **102**, 11070–11075 (2005).
32. Linz, B. *et al.* An African origin for the intimate association between humans and *Helicobacter pylori*. *Nature* **445**, 915–918 (2007).
33. Douglass, J. M., Li, Y. & Tinanoff, N. Association of mutans streptococci between caregivers and their children. *Pediatr. Dent.* **30**, 375–387 (2008).
34. Li, Y., Ismail, A. I., Ge, Y., Tellez, M. & Sohn, W. Similarity of bacterial populations in saliva from African-American mother-child dyads. *J. Clin. Microbiol.* **45**, 3082–3085 (2007).
35. Li, M. *et al.* Symbiotic gut microbes modulate human metabolic phenotypes. *Proc. Natl Acad. Sci. USA* **105**, 2117–2122 (2008).
36. McNulty, N. P. *et al.* The impact of a consortium of fermented milk strains on the gut microbiome of gnotobiotic mice and monozygotic twins. *Sci. Transl. Med.* **3**, 106ra106 (2011).
37. Turnbaugh, P. J. *et al.* A core gut microbiome in obese and lean twins. *Nature* **457**, 480–484 (2009).
38. Fierer, N., Hamady, M., Lauber, C. L. & Knight, R. The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proc. Natl Acad. Sci. USA* **105**, 17994–17999 (2008).
39. Palmer, C., Bik, E. M., DiGiulio, D. B., Relman, D. A. & Brown, P. O. Development of the human infant intestinal microbiota. *PLoS Biol.* **5**, e177 (2007).
This study describes the taxonomic developments that occur in the infant microbiome and the relationships between these changes and environmental exposures.
40. Blaser, M. J. & Falkow, S. What are the consequences of the disappearing human microbiota? *Nature Rev. Microbiol.* **7**, 887–894 (2009).
This article proposes that our modern lifestyle has led to the extinction of certain microbes, and that their disappearance may have deleterious effects on human health.
41. Sjolund, M., Wreiber, K., Andersson, D. I., Blaser, M. J. & Engstrand, L. Long-term persistence of resistant *Enterococcus* species after antibiotics to eradicate *Helicobacter pylori*. *Ann. Intern. Med.* **139**, 483–487 (2003).
42. Blaser, M. J. Antibiotic overuse: stop the killing of beneficial bacteria. *Nature* **476**, 393–394 (2011).
43. Evans, A. S. Causation and disease: the Henle–Koch postulates revisited. *Yale J. Biol. Med.* **49**, 175–195 (1976).
44. Muegge, B. D. *et al.* Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* **332**, 970–974 (2011).
45. Huston, M. A. *Biological Diversity: The Coexistence Of Species On Changing Landscapes* (Cambridge Univ. Press, UK, 1994).
46. Kennedy, T. A. *et al.* Biodiversity as a barrier to ecological invasion. *Nature* **417**, 636–638 (2002).
47. Strogatz, S. H. Exploring complex networks. *Nature* **410**, 268–276 (2001).
48. Paine, R. T. Food web complexity and species diversity. *Am. Nat.* **100**, 65–75 (1966).
49. Sole, R. V. & Montoya, J. M. Complexity and fragility in ecological networks. *Proc. Biol. Sci.* **268**, 2039–2045 (2001).
50. Borrvall, C. & Ebenman, B. Early onset of secondary extinctions in ecological communities following the loss of top predators. *Ecol. Lett.* **9**, 435–442 (2006).
51. Bik, E. M. *et al.* Molecular analysis of the bacterial microbiota in the human stomach. *Proc. Natl Acad. Sci. USA* **103**, 732–737 (2006).
52. Maldonado-Contreras, A. *et al.* Structure of the human gastric bacterial community in relation to *Helicobacter pylori* status. *ISME J.* **5**, 574–579 (2011).
53. Li, Y., Caufield, P. W., Dasanayake, A. P., Wiener, H. W. & Vermund, S. H. Mode of delivery and other maternal factors influence the acquisition of *Streptococcus mutans* in infants. *J. Dent. Res.* **84**, 806–811 (2005).
54. Dominguez-Bello, M. G. *et al.* Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc. Natl Acad. Sci. USA* **107**, 11971–11975 (2010).
This study shows that infants have largely undifferentiated microbiota across multiple anatomic sites immediately after birth, and that delivery mode determines which types of bacteria of the infant microbiome are the earliest colonizers.
55. Savage, D. C., Dubos, R. & Schaedler, R. W. The gastrointestinal epithelium and its autochthonous bacterial flora. *J. Exp. Med.* **127**, 67–76 (1968).
One of the pioneering studies of the features of the bacterial colonization of the gastrointestinal tract that occurs in early life.
56. Gronlund, M. M., Lehtonen, O. P., Eerola, E. & Kero, P. Fecal microflora in healthy infants born by different methods of delivery: permanent changes in intestinal flora after cesarean delivery. *J. Pediatr. Gastroenterol. Nutr.* **28**, 19–25 (1999).
57. Grant, B. R. & Grat, P. R. Cultural inheritance of song and its role in the evolution of Darwin's finches. *Evolution* **50**, 2471–2487 (1996).
58. Hunt, J. & Simmons, L. W. Maternal and paternal effects on offspring phenotype in the dung beetle *Onthophagus taurus*. *Evolution* **54**, 936–941 (2000).
59. Raymond, J. *et al.* Genetic and transmission analysis of *Helicobacter pylori* strains within a family. *Emerg. Infect. Dis.* **10**, 1816–1821 (2004).
60. Smillie, C. S. *et al.* Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* **480**, 241–244 (2011).
The discovery of a large network of gene exchange that occurs in microbial communities and that allows rapid genetic information transfer to occur in the microbiome. The authors speculate that such networks have roles in specific human diseases.
61. Wirth, T. *et al.* Distinguishing human ethnic groups by means of sequences from *Helicobacter pylori*: lessons from Ladakh. *Proc. Natl Acad. Sci. USA* **101**, 4746–4751 (2004).
62. Sharon, G. *et al.* Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **107**, 20051–20056 (2010).
63. Leyden, J. J., McGinley, K. J., Holzle, E., Labows, J. N. & Kligman, A. M. The microbiology of the human axilla and its relationship to axillary odor. *J. Invest. Dermatol.* **77**, 413–416 (1981).
64. Dobzhansky, T. Further data on the variation of the Y chromosome in *Drosophila pseudoobscura*. *Genetics* **22**, 340–346 (1937).
65. Mayr, E. *Systematics And The Origin Of Species From The Viewpoint Of A Zoologist* (Columbia Univ. Press, New York, USA, 1942).
66. Brailsford, S. R. *et al.* The microflora of the erupting first permanent molar. *Caries Res.* **39**, 78–84 (2005).
67. Cephas, K. D. *et al.* Comparative analysis of salivary bacterial microbiome diversity in edentulous infants and their mothers or primary care givers using pyrosequencing. *PLoS ONE* **6**, e23503 (2011).
68. Schaedler, R. W. The relationshp between the host and its intestinal microflora. *Proc. Nutr. Soc.* **32**, 41–47 (1973).
69. Jukes, T. H. Antibiotics in feeds. *Science* **204**, 8 (1979).
70. Robinson, C. J. & Young, V. B. Antibiotic administration alters the community structure of the gastrointestinal micobiota. *Gut Microbes* **1**, 279–284 (2010).
71. Wlodarska, M. *et al.* Antibiotic treatment alters the colonic mucus layer and predisposes the host to exacerbated *Citrobacter rodentium*-induced colitis. *Infect. Immun.* **79**, 1536–1545 (2011).
72. Gemmell, N. J. & Slate, J. Heterozygote advantage for fecundity. *PLoS ONE* **1**, e125 (2006).
73. Cauci, S. *et al.* Prevalence of bacterial vaginosis and vaginal flora changes in peri- and postmenopausal women. *J. Clin. Microbiol.* **40**, 2147–2152 (2002).
74. Osborne, N. G., Wright, R. C. & Grubin, L. Genital bacteriology: a comparative study of premenopausal women with postmenopausal women. *Am. J. Obstet. Gynecol.* **135**, 195–198 (1979).
75. Peek, R. M. Jr & Blaser, M. J. *Helicobacter pylori* and gastrointestinal tract adenocarcinomas. *Nature Rev. Cancer* **2**, 28–37 (2002).
76. Giannakis, M., Chen, S. L., Karam, S. M., Engstrand, L. & Gordon, J. I. *Helicobacter pylori* evolution during progression from chronic atrophic gastritis to gastric cancer and its impact on gastric stem cells. *Proc. Natl Acad. Sci. USA* **105**, 4358–4363 (2008).
77. Li, X. X. *et al.* Bacterial microbiota profiling in gastritis without *Helicobacter pylori* infection or non-steroidal anti-inflammatory drug use. *PLoS ONE* **4**, e7985 (2009).
78. Mariat, D. *et al.* The Firmicutes/Bacteroidetes ratio of the human microbiota changes with age. *BMC Microbiol.* **9**, 123 (2009).
79. Nordling, C. O. A new theory on cancer-inducing mechanism. *Br. J. Cancer* **7**, 68–72 (1953).
80. Vanhoutvin, S. A. *et al.* Butyrate-induced transcriptional changes in human colonic mucosa. *PLoS ONE* **4**, e6759 (2009).
81. Hamilton, W. D. The moulding of senescence by natural selection. *J. Theor. Biol.* **12**, 12–45 (1966).
A pioneering paper that describes how several key factors (fertility, mortality and age) affect population dynamics.
82. Perry, S. *et al.* Infection with *Helicobacter pylori* is associated with protection against tuberculosis. *PLoS ONE* **5**, e8804 (2010).
83. Higgins, P. D. *et al.* Prior *Helicobacter pylori* infection ameliorates *Salmonella typhimurium*-induced colitis: mucosal crosstalk between stomach and distal intestine. *Inflamm. Bowel Dis.* **17**, 1398–1408 (2011).
84. Arnold, I. C. *et al. Helicobacter pylori* infection prevents allergic asthma in mouse models through the induction of regulatory T cells. *J. Clin. Invest.* **121**, 3088–3093 (2011).
85. Atherton, J. C. & Blaser, M. J. Coadaptation of *Helicobacter pylori* and humans: ancient history, modern implications. *J. Clin. Invest.* **119**, 2475–2487 (2009).
86. Blaser, M. J. & Webb, G. Host demise as a beneficial function of indigenous microbiota in multicellular hosts. in *Am. Soc. Microbiol. Conf. Beneficial Microbes* (Lake Tahoe, Nevada, USA, 2005).
87. Patel, R. V. & Lebwohl, M. Psoriasis. *Ann. Intern. Med.* **155**, ITC2-1 (2011).
88. Gao, Z., Tseng, C. H., Strober, B. E., Pei, Z. & Blaser, M. J. Substantial alterations of the cutaneous bacterial biota in psoriatic lesions. *PLoS ONE* **3**, e2719 (2008).
89. Grice, E. A. & Segre, J. A. The skin microbiome. *Nature Rev. Microbiol.* **9**, 244–253 (2011).
A comprehensive review of the skin microbiome and its connection to several diseases.

90. McDowell, A. *et al.* A novel multilocus sequence typing scheme for the opportunistic pathogen *Propionibacterium acnes* and characterization of type I cell surface-associated antigens. *Microbiology* **157**, 1990–2003 (2011).
91. Price, L. B. *et al.* Community analysis of chronic wound bacteria using 16S rRNA gene-based pyrosequencing: impact of diabetes and antibiotics on chronic wound microbiota. *PLoS ONE* **4**, e6462 (2009).
92. Grice, E. A. *et al.* Longitudinal shift in diabetic wound microbiota correlates with prolonged skin defense response. *Proc. Natl Acad. Sci. USA* **107**, 14799–1804 (2010).
93. Andersson, A. F. *et al.* Comparative analysis of human gut microbiota by barcoded pyrosequencing. *PLoS ONE* **3**, e2836 (2008).
94. McColl, K. E. *Helicobacter pylori* infection. *N. Engl. J. Med.* **362**, 1597–1604 (2010).
95. el-Serag, H. B. & Sonnenberg, A. Opposing time trends of peptic ulcer and reflux disease. *Gut* **43**, 327–333 (1998).
96. Chen, Y. & Blaser, M. J. Inverse associations of *Helicobacter pylori* with asthma and allergy. *Arch. Intern. Med.* **167**, 821–827 (2007).
97. Plottel, C. S. & Blaser, M. J. Microbiome and malignancy. *Cell Host Microbe* **10**, 324–335 (2011).
98. Lazarova, D. L., Bordonaro, M., Carbone, R. & Sartorelli, A. C. Linear relationship between Wnt activity levels and apoptosis in colorectal carcinoma cells exposed to butyrate. *Int. J. Cancer* **110**, 523–531 (2004).
99. Wu, S. *et al.* A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nature Med.* **15**, 1016–1022 (2009).
100. Reikvam, D. H. *et al.* Depletion of murine intestinal microbiota: effects on gut mucosa and epithelial gene expression. *PLoS ONE* **6**, e17996 (2011).
101. Castellarin, M. *et al. Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome Res.* **22**, 299–306 (2012).
102. Kostic, A. D. *et al.* Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome Res.* **22**, 292–298 (2012).
103. Krisanaprakornkit, S. *et al.* Inducible expression of human β-defensin 2 by *Fusobacterium nucleatum* in oral epithelial cells: multiple signaling pathways and role of commensal bacteria in innate immunity and the epithelial barrier. *Infect. Immun.* **68**, 2907–2915 (2000).
104. Littman, D. R. & Pamer, E. G. Role of the commensal microbiota in normal and pathogenic host immune responses. *Cell Host Microbe* **10**, 311–323 (2011).
105. Ivanov, I. I. *et al.* Specific microbiota direct the differentiation of IL-17-producing T-helper cells in the mucosa of the small intestine. *Cell Host Microbe* **4**, 337–349 (2008).
An important study that describes the immunological interplay between segmented filamentous bacteria and Th17 cells in the distal small bowel.
106. Round, J. L. & Mazmanian, S. K. Inducible Foxp3$^+$ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. *Proc. Natl Acad. Sci. USA* **107**, 12204–12209 (2010).
107. Ogura, Y. *et al.* A frameshift mutation in *NOD2* associated with susceptibility to Crohn's disease. *Nature* **411**, 603–606 (2001).
108. Hugot, J. P. *et al.* Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* **411**, 599–603 (2001).
109. Franchimont, D. *et al.* Deficient host-bacteria interactions in inflammatory bowel disease? The toll-like receptor (TLR)-4 Asp299gly polymorphism is associated with Crohn's disease and ulcerative colitis. *Gut* **53**, 987–992 (2004).
110. Ewaschuk, J. B., Tejpar, Q. Z., Soo, I., Madsen, K. & Fedorak, R. N. The role of antibiotic and probiotic therapies in current and future management of inflammatory bowel disease. *Curr. Gastroenterol. Rep.* **8**, 486–498 (2006).
111. Hviid, A., Svanstrom, H. & Frisch, M. Antibiotic use and inflammatory bowel diseases in childhood. *Gut* **60**, 49–54 (2011).
112. Manichanh, C. *et al.* Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut* **55**, 205–211 (2006).
113. Garrett, W. S. *et al. Enterobacteriaceae* act in concert with the gut microbiota to induce spontaneous and maternally transmitted colitis. *Cell Host Microbe* **8**, 292–300 (2010).
114. Lepage, P. *et al.* Twin study indicates loss of interaction between microbiota and mucosa of patients with ulcerative colitis. *Gastroenterology* **141**, 227–236 (2011).
115. Mondot, S. *et al.* Highlighting new phylogenetic specificities of Crohn's disease microbiota. *Inflamm. Bowel Dis.* **17**, 185–192 (2011).
116. Abu-Shanab, A. & Quigley, E. M. The role of the gut microbiota in nonalcoholic fatty liver disease. *Nature Rev. Gastroenterol. Hepatol.* **7**, 691–701 (2010).
117. Backhed, F., Manchester, J. K., Semenkovich, C. F. & Gordon, J. I. Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *Proc. Natl Acad. Sci. USA* **104**, 979–984 (2007).
118. Mutlu, E. *et al.* Intestinal dysbiosis: a possible mechanism of alcohol-induced endotoxemia and alcoholic steatohepatitis in rats. *Alcohol. Clin. Exp. Res.* **33**, 1836–1846 (2009).
119. Yan, A. W. *et al.* Enteric dysbiosis associated with a mouse model of alcoholic liver disease. *Hepatology* **53**, 96–105 (2011).
120. Fox, J. G. *et al.* Gut microbes define liver cancer risk in mice exposed to chemical and viral transgenic hepatocarcinogens. *Gut* **59**, 88–97 (2010).
121. Chen, Y. *et al.* Characterization of fecal microbial communities in patients with liver cirrhosis. *Hepatology* **54**, 562–572 (2011).
122. Ley, R. E., Turnbaugh, P. J., Klein, S. & Gordon, J. I. Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022–1023 (2006).
123. Ajslev, T. A., Andersen, C. S., Gamborg, M., Sorensen, T. I. & Jess, T. Childhood overweight after establishment of the gut microbiota: the role of delivery mode, pre-pregnancy weight and early administration of antibiotics. *Int. J. Obes.* **35**, 522–529 (2011).
124. Luoto, R., Kalliomaki, M., Laitinen, K. & Isolauri, E. The impact of perinatal probiotic intervention on the development of overweight and obesity: follow-up study from birth to 10 years. *Int. J. Obes.* **34**, 1531–1537 (2010).
125. Li, J. V. *et al.* Metabolic surgery profoundly influences gut microbial-host metabolic cross-talk. *Gut* **60**, 1214–1223 (2011).
126. Ivanov, I. I. *et al.* Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell* **139**, 485–498 (2009).
127. Scher, J. U. & Abramson, S. B. The microbiome and rheumatoid arthritis. *Nature Rev. Rheumatol.* **7**, 569–578 (2011).
128. Hill, A. B. The environment and disease: association or causation? *Proc. R. Soc. Med.* **58**, 295–300 (1965).
129. Hentschel, E. *et al.* Effect of ranitidine and amoxicillin plus metronidazole on the eradication of *Helicobacter pylori* and the recurrence of duodenal ulcer. *N. Engl. J. Med.* **328**, 308–312 (1993).
130. Devoy, A., Bunton-Stasyshyn, R. K., Tybulewicz, V. L., Smith, A. J. & Fisher, E. M. Genomically humanized mice: technologies and promises. *Nature Rev. Genet.* **13**, 14–20 (2012).
131. Shulzhenko, N. *et al.* Crosstalk between B lymphocytes, microbiota and the intestinal epithelium governs immunity versus metabolism in the gut. *Nature Med.* **17**, 1585–1593 (2011).
132. Reshef, D. N. *et al.* Detecting novel associations in large data sets. *Science* **334**, 1518–1524 (2011).
133. Islami, F. & Kamangar, F. *Helicobacter pylori* and esophageal cancer risk: a meta-analysis. *Cancer Prev. Res.* **1**, 329–338 (2008).
134. Blaser, M. J., Chen, Y. & Reibman, J. Does *Helicobacter pylori* protect against asthma and allergy? *Gut* **57**, 561–567 (2008).
135. Tana, C. *et al.* Altered profiles of intestinal microbiota and organic acids may be the origin of symptoms in irritable bowel syndrome. *Neurogastroenterol. Motil.* **22**, 512–519 (2010).
136. Wang, Z. *et al.* Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472**, 57–63 (2011).
137. Larsson, E. *et al.* Analysis of gut microbial regulation of host gene expression along the length of the gut and regulation of gut microbial ecology through MyD88. *Gut* 23 Nov 2011 (doi:10.1136/gutjnl-2011-301104).
138. Backhed, F. *et al.* The gut microbiota as an environmental factor that regulates fat storage. *Proc. Natl Acad. Sci. USA* **101**, 15718–15723 (2004).
139. Turnbaugh, P. J. *et al.* The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci. Transl. Med.* **1**, 6ra14 (2009).

Acknowledgements
This work was supported by NIH grants R01GM63270, R01DK090989, UH2 AR057506, 5 P30 CA016087 and 1UL1RR029893, the Diane Belfer Program for Human Microecology, the Michael Saperstein Medical Scholars Fund and the Levin Fellowship in Gastroenterology.

Competing interests statement
The authors declare no competing financial interests.

FURTHER INFORMATION
Clusters of Orthologous Groups (COG) database: http://clovr.org/docs/clusters-of-orthologous-groups-cogs
Greengenes: http://greengenes.lbl.gov
Human Microbiome Project: http://commonfund.nih.gov/hmp
Kyoto Encyclopedia of Genes and Genomes (KEGG) database: http://www.genome.jp/kegg
Metagenomics of the Human Intestinal Tract (MetaHIT): http://www.metahit.eu
Metagenomics Rapid Annotation using Subsystem Technology (MG-RAST): http://metagenomics.anl.gov
Quantitative Insights Into Microbial Ecology (QIIME): http://www.qiime.org
QIIME database: http://www.microbio.me/qiime
Ribosomal Database Project (RDP): http://rdp.cme.msu.edu
***Nature Reviews Genetics* Series on Applications of Next-Generation Sequencing:** http://www.nature.com/nrg/series/nextgeneration/index.html

ALL LINKS ARE ACTIVE IN THE ONLINE PDF