CellPress
REVIEWS

Review

Impact of the Microbiome on the Human Genome

Thiviya Jeyakumar,[1,2] Nicole Beauchemin,[1,3] and Philippe Gros[1,2,3,*]

Humans live in a microbial world that includes pathogenic bacteria, viruses, and fungi that cause lethal infections. In addition, a large number of microbial communities inhabit mucosal surfaces where they provide key metabolic activities, facilitating adaptation to changing environments. New genome technologies enable both sequencing of the human genome and sequence-based cataloging of microbial communities inhabiting human mucosal surfaces. These have revealed intricate two-way relationships between the microbiome and the genome, including strong effects of human genotypes on the composition and activity of the microbiome. Likewise, the microbiome plays an important role in training and regulating the immune system, and acts to modify expression of human genetic risk for debilitating chronic inflammatory and immune conditions. These studies are suggesting a new role of the microbiome in human health and disease.

Genetic Analysis of Host–Microbiome Interaction

A genetic approach to considering the possible impact of the microbial world on the human genome and its evolution relies on the prediction that such human–microbe interactions have in fact left recognizable genetic fingerprints over time that can be detected by comparative genome sequencing.

McFall-Ngai and colleagues first reported that 37% of the ~23 000 human genes have homologs in the Bacteria and Archaea, and another 28% originated in unicellular eukaryotes [1], highlighting the contribution of the microbial world as a provider of building blocks for the human genome. Subsequent microbial effects on human genotypes and associated phenotypes that significantly increase the capacity of humans to adapt, thrive, and reproduce in changing environments can also be detected by genome sequencing [2]. Indeed, selective pressure from microbial pathogens that result in retention or elimination of certain human genotypes is well documented [3,4]. This selection may take place slowly over time in populations living in areas of endemic disease, or may happen abruptly following an epidemic [3,4].

Humans also live in contact with a large community of nonpathogenic commensals termed 'the **microbiome**' (see Glossary). These microbes live in symbiosis with humans on the skin, gut, and other mucosal surfaces, and they do not generally cause life-threatening diseases except in immunocompromised hosts. However, recent evidence indicates that this microbiome plays a critical role in training the host immune system during early years and can also modulate its activity throughout the life course [5,6]. Hence, variations in microbiome-directed regulation of immune functions may impact the susceptibility of humans to pathogenic microbes, or their response to chronic exposure to microbial products. Thus, the microbiome and its products are associated with susceptibility to chronic inflammatory and autoimmune diseases that carry significant morbidity, such as **inflammatory bowel disease (IBD)**, obesity, asthma, Parkinson's disease, and even cancer (reviewed by [7,8]). The onset and progression of such diseases have a clear impact on human fitness and longevity, and ultimately on vertical transmission of specific vulnerability and resistance genotypes over generations. The analysis of such genomic fingerprints is complicated by several factors: the horizontal transmission of microbiomes between individuals, the expected modest generational effect (both penetrance and expressivity), the complex nature of the human genes involved, and their further modulation by environmental factors, including diet, lifestyle, and chemical insults. Although a significant body of association data on three-way interactions between humans, the microbiome, and the environment is emerging, formal cause-to-effect relationships have relied on studies in animal models.

Highlights

While the selective pressure exerted by pathogens on the human genome is clear, commensal microbes may have also helped humans adapt to changing environments, also leaving marks in the human genome.

Affordable genomic sequencing, and cataloging of microbial populations living on humans, is revealing an important human/microbiome interplay in health and disease.

The human microbiome plays a critical role in training and regulating the immune system to promote health and fitness. Certain human genetic variants that cause changes in microbial composition are also associated with diseases, suggesting that the microbiome may impact expression of disease-associated genetic variants in humans.

The microbiome may thus have influenced the retention or elimination of genetic variants affecting fitness, and may have contributed to sculpting the human genome.

[1]Department of Biochemistry, McGill University, Montreal, Canada

[2]McGill Center for the Study of Complex Traits, McGill University, Montreal, Canada

[3]Goodman Cancer Research Center, McGill University, Montreal, Canada

*Correspondence:
philippe.gros@mcgill.ca

The focus of this review is on host–microbiome interactions in health and disease, with an emphasis on humans and on mouse models of human diseases. The reader is referred to recent articles that review host–microbiome interactions in non-human primates [9], insects [10], and other invertebrates and early-branching animal Phyla [11]. It has also been proposed that the microbiome may have benefited human evolution in other ways, such as adaptation to changing climates and nutritional environments. This topic will not be reviewed in detail herein, but the reader is directed towards superb reviews on this subject [2,12].

Genomic Fingerprints Left by Microbial Pathogens

Major advances in genome technologies permit deep genomic sampling of multiple individuals from the same or different populations to investigate genome–environment interactions (environmental niche, diet, epidemics, etc.), including the impact of different microbiomes. These analyses have identified direct cause-to-effect relationships between specific human genotypes and unique microbes, while in other cases such evidence is less strident, indirect, and lacking direct validation. Different types of genomic adaptation that increase fitness in response to selective pressure from changing environments have been described. Classical examples include skin pigmentation (*SLC24A5*, skin color) [13], hair morphology (*EDAR*, ectodysplasin A; thermo regulation) [14], hypoxic conditions (*EPAS1*; high altitude) [15], and lactose tolerance (*LCT*, glycoside hydrolase; lactose tolerance) [16] to name a few. The innate and adaptive immune systems represent a clear illustration of gene amplification for adaptation to changing environment, in this case sensing different microbial molecules, to increase diversity and functional redundancy in gene families such as immunoglobulins, T cell receptors, and pattern-recognition molecules [3,17]. These genetic variants are either under positive or negative selection, or balancing selection (for review, see [4,18]). Variants under positive selection provide a selective advantage under certain environmental conditions and are often found as homozygotes, while those under balancing selection can preserve genetic diversity by driving advantages in heterozygotes against select pressures, but being deleterious to host fitness in homozygotes. Several pathogenic microbes have positively selected certain allelic variants in human genes over time, leaving obvious fingerprints on the genome. Box 1 provides an illustration of some of the best characterized cases of positive or balancing selection by pathogens.

Many genes of the immune system are evolving under purifying selection from pathogenic microbes. Indeed, mutations in certain host defenses and immune pathways cause life-threatening susceptibility to infections with bacterial, parasitic, fungal, and viral pathogens. There are over 350 such immunodeficiencies (inborn errors of immunity) with a known genetic etiology [19]. While a few of these are 'syndromic', affecting multiple organs, many are primary immunodeficiencies (PIDs) that affect only the immune system and host defense genes. Although some of these mutations cause very severe clinical presentation and broad susceptibility to infections, such as severe combined immunodeficiencies (SCID, mutations in *JAK3*), and chronic granulomatous disease (CGD, mutations in subunits of the NADPH oxidase), others cause a very narrow infection susceptibility phenotype [20,21]. For example, mutations in *TLR3* only cause susceptibility to herpes simplex encephalitis, while mutations in *IL17* and *CARD9* only cause susceptibility to fungal pathogens (chronic mucocutaneous candidiasis) [22–24]. Mendelian susceptibility to mycobacterial disease (MSMD) manifests as disseminated or recurrent mycobacterial infections following perinatal vaccination with bacillus Calmette–Guérin (BCG) [25]. Mutations causing MSMD affect ontogeny of immunocytes (*IRF8*), expression of key activating cytokines (*IL12*, *IFNG*) and their receptors (*IL12RB1*, *IFNGR1/2*), and the intracellular signaling pathways (*TYK2*, *STAT1*, *NEMO*) that are linked to these receptors [25]. Hence, it is clear that variants causing such severe primary immunodeficiencies have been quickly eliminated from the population. Negative selection can also arise for genes that have no detrimental effects on host immune systems per se, but predispose to novel selective pressures such as the effect of deadly European pathogens on the Native American populations. For example, prior to the arrival of Europeans, there was stable positive selection for *HLA-DQA1* in ancient indigenous populations of British Columbia, Canada (near 100% of the population). However, in modern populations, and following European contact,

Glossary

Allele: corresponds to one of two or more alternative forms of a gene that arise by mutation within or next to a gene, and that can be distinguished by different nucleotide sequences. The presence of different alleles of the same gene may result in different observable phenotypic traits in humans, such as differences in pigmentation or susceptibility to certain diseases.
Dysbiosis: microbial imbalance or maladaptation within a particular environment on or inside the body, such as the intestinal microbiome, that shows changes in the composition or metabolic activity of its bacterial communities which confers novel properties, sometimes causing diseases in humans.
Genome-wide association studies (GWAS): a type of population study in humans that searches for the statistical association of discrete nucleotide variants at millions of locations in the genome (near or within genes) with specific traits.
Inflammatory bowel disease (IBD): a group of chronic diseases of mild or severe inflammation of the intestinal mucosa that includes ulcerative colitis and Crohn's disease. In these conditions, the normal intestinal tissue barriers are disrupted, leading to continuous exposure of the underlying tissues to gut microbes.
Inflammatory diseases: conditions such as inflammatory bowel disease, asthma, and multiple sclerosis, where sustained activation of the immune system and associated inflammatory pathways lead to chronic or acute pathologies.
Microbiome: the microbial populations living on and within the human body (i.e., gut, skin, lungs) that, unlike microbial pathogens, are not traditionally considered to cause disease but that behave as commensal microbes.

Box 1. Pathogens Have Left Genetic Footprints in Human Populations

Allelic variants associated with protection against certain life-threatening infections have been selected for in populations where the disease is endemic or has been prevalent in the past. The clearest example of such selection is the *Plasmodium* malarial parasite. Variants in erythrocyte-specific proteins affecting the capacity of *Plasmodium* to invade/replicate in these cells are enriched in malaria-endemic areas. The erythrocyte cell-surface 'Duffy antigen' (*DARC*) is used by *Plasmodium vivax* for invasion, and absence of DARC completely protects against *P. vivax* malaria; *P. vivax* has driven fixation of Duffy negativity in sub-Saharan Africa [88,89]. Deletion of the erythrocyte Cl^-/HCO_3^- anion exchanger (*SLC4A1*) causes Southeast Asian ovalocytosis, a condition common in Southeast Asia that is associated with reduced malaria incidence [90–92]. A case of balancing selection is the sickle-cell hemoglobin variant (HbS) prevalent in Africa, the Middle East, and India (5–40% carrier frequency). Although HbS homozygosity causes severe anemia, heterozygotes are asymptomatic and show strong protection against malaria [93,94]. In West Africa, high rates of resistance to Lassa hemorrhagic fever suggests selective pressure for *LARGE* allelic variants, a gene implicated in virus infectivity [95]. The *CCR5* gene (C-C chemokine receptor type 5) codes for the cell-surface HIV-1 receptor on T lymphocytes. A deletion in *CCR5* protects against HIV-1 infection [96,97]. While this variant (frequent in European populations) may contribute to relatively low HIV levels in Europe [98], an unknown historical viral disease may have triggered its initial selection in Europeans. Blood type O individuals are more susceptible to cholera (*Vibrio cholerae*) than other blood groups, and in flood-heavy Bangladesh, where cholera outbreaks often happen, the frequency of type O blood is among the lowest in the world [99]. Furthermore, the proapoptotic protease caspase-12 modulates immune responses against bacterial infections, and *caspase-12*-active populations are more susceptible to sepsis [100]; the inactive *caspase-12* isoform is spreading in human populations, suggesting selective pressure for more effective antibacterial immune responses [101]. Finally, while the origin of the microbial pathogen exerting selective pressure is sometimes unknown, the genetic fingerprints left in the human genomes are clear; for example, a bacterial pathogen is likely to have driven the prevalence of pathological *CFTR* variants (dF508) in Northern European populations [26].

this frequency has decreased to ~35% over the past 150–200 years, a change that could not be explained simply by changes in demographics, European admixture, or balancing selection. This suggests that the same protective genetic variant(s) positively selected by local pathogens in ancient environments had experienced strong negative selection following contact with novel European pathogens [26].

In simple terms, balancing selection implies that certain genetic variants present in human populations may have been preserved because they provide a fitness advantage in one set of environmental conditions, while decreasing fitness under other circumstances. Except for a few cases (e.g., the sickle-cell hemoglobin variant; HbS), examples of balancing selection are difficult to pinpoint in humans. One instance that has yet to reach scientific consensus is the mutations in *CFTR* in northern Europeans (Caucasian). Homozygosity for mutations in *CFTR* causes cystic fibrosis, a deadly disease that causes chronic infections of the lung. Because of the high frequency of mutant *CFTR* **alleles** in Europeans, it has been theorized that unaffected heterozygotes have a protective advantage against a yet to be identified infectious agent during historic epidemics, with cholera, typhoid, and tuberculosis being lead candidates [27].

Balancing selection of certain protein variants has been studied in mouse models of experimental infections and inflammation, providing insight into how fluctuating selective forces can alter the effects of genetic variants on host response to different pathogens. For example, mutations in immune genes such as *STAT1*, *IRF1*, *IRF8*, *IFNG*, *THEMIS*, and several others [28] cause high susceptibility to pulmonary tuberculosis. However, inactivation of such 'infection-protective' genes and associated pathways completely protects mutant mice against acute neuroinflammation in cerebral malaria [29–33]. RNA and chromatin sequencing studies show that the same pathways are activated (by proinflammatory transcription factors *IRF1/8* and *STAT1*) in the lungs infected with *Mycobacterium tuberculosis*, and the brains of *Plasmodium berghei*-infected mice, suggesting that immune pathways required for protection against pulmonary tuberculosis drive pathological neuroinflammation in cerebral malaria. Likewise, mutations in the complement pathway (C5a) have opposite effects on

vulnerability of mice to infection with the fungal pathogen *Candida albicans*, and the malarial parasite *P. berghei* [34,35]. Interestingly, circulating levels of C5a have been associated with different severity of pregnancy-associated malaria [36].

The current human genome is an admixture of the genomes of archaic and modern humans (*Homo sapiens* × Neanderthals). Genomic regions depleted in archaic ancestry can identify functions specific to modern humans ('Neanderthal desert'; purifying selection of Neanderthal alleles), examples of which include *FOXP2* (language) [37] and *AMY1* (amylase) [38]. Interestingly, innate immunity genes display a greater degree of Neanderthal ancestry than other genes, with gene/loci such as *OAS*, the *TLR6-1-10* cluster, *IRF6*, *IFITM1-3*, *IL17A*, and *IL17F* showing the highest Neanderthal introgression score genome-wide. This suggests that Neanderthal diversity at these loci may have undergone positive selection in humans, supporting better adaptation against pathogens. On the other hand, the *TLR6-1-10* cluster (Neanderthal origin) is associated with increased risk of allergic disease and IBD, suggesting that these variants may have been historically beneficial to resist some microbial assaults, but are now associated with **inflammatory disease** risk [39,40].

Human Genetic Effects on the Microbiome

The symbiotic relationship between microbes and humans is expected to involve two-way interactions that optimize the mutual benefits of cohabitation (Figure 1). Human genetics influences the microenvironment of the gut microflora, both with respect to chemical characteristics (diet, lifestyle) and activity of the immune system (IgA production, antibacterial peptides). Conversely, the microbiome composition and activity may create specific conditions in the host (including initial training of their immune system and punctual modulation of immune function) that may affect susceptibility to inflammatory or infectious diseases which, over time, may alter transmission of disease vulnerability or disease-resistance alleles. Although there is clear evidence of the former, much less is known about the latter.

In humans, genetic variants have been shown to have a strong impact on the microbiome [41,42]. Studies conducted on human twins indicated that the microbiome of monozygotic twins is more similar than that of dizygotic twins [42–44]. In family studies, as many as 20 bacterial taxa were found to show heritability [42]. The systematic identification of host genetic variants affecting microbiome composition is important, as **dysbiosis** of the gut microbiome has been implicated in multiple diseases, including IBD, necrotizing enterocolitis, type 2 diabetes, asthma, and rheumatoid arthritis (for a review see [45]). So-called mGWAS (microbiome **genome-wide association studies**) have shown that these genetic effects are complex, and although many loci have been mapped few have been replicated. One validated association is that of the lactase gene (*LCT*) with abundance of *Bifidobacterium* (a member of the phylum Actinobacteria) [46,47]. In other studies, as few as nine loci have been identified as affecting bacterial taxa [47], while as many as 42 [48] or even 83 loci/SNPs [46] affect interindividual variability in the microbiome; however, even in the best settings (twin pairs), host genetics only accounts for between 5 and 10% of total phenotypic variance of the trait [45]. Additional complicating factors in the mGWAS studies include variability in geography and genetic backgrounds, and methodology and statistical approaches used. Therefore, although such association studies have pointed to an effect of human genetic make-up on the gut microbiome, these studies lack sufficient statistical power and resolution to map major single gene effects.

In humans, dysbiosis of the gut microflora has been associated with IBD [49], a disease which is strongly influenced by human genetic factors; GWAS studies of >60 000 IBD patients have mapped >200 loci impacting disease (reviewed in [50]). Importantly, up to 40 of these IBD loci also affect the composition of the microbiome [51]. Thus, defects in immune or other pathways involved in sensing and appropriately responding to the commensal microbiome may allow for overgrowth of pathobionts. In mouse models, inactivation of genes and proteins involved in sensing microbial ligands (*NOD2*, *TBET*, *ATG16L1*, and others) in gut mucosa causes emergence of a 'colitogenic flora' with a dominant IBD-like phenotype that is transmissible to a naïve host [45]. In humans, *NOD2* is a risk

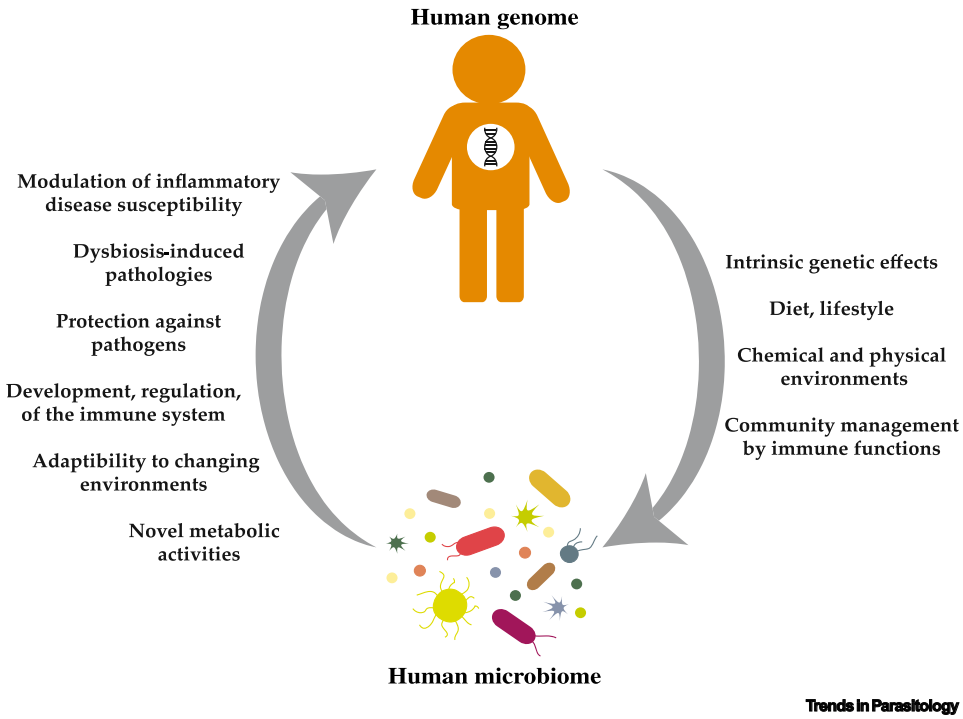

Figure 1. Two-Way Interactions between Humans and Their Microbiome.

Humans live in symbiosis with their commensal microbiome, with many complex and intricate two-way interactions, in which human genetic variants may affect the composition and activity of the microbiome, and conversely where alterations in the microbiome can impact human health by modulating penetrance and expressivity of genetically determined human disease traits. The microbiome not only secretes essential nutrients and provides dynamic genetic and biochemical diversity to humans for rapid adaptation to new nutritional and climate environments, but it also educates and modulates the immune system, and can play key roles in modulating disease susceptibility. In turn, alongside host lifestyle, diet, and age, human genetic variation, generally of genes involved in tissue integrity, immunity, and dietary metabolism, can alter the microenvironment these microbial species inhabit, and thus may make niches that are more suitable for colonization by certain microbes over others.

allele for IBD, and *NOD2* loss-of-function mutations are associated with different forms of colitis; interestingly, mGWAS studies indicate that members of the Enterobacteriaceae are enriched in individuals with IBD-associated *NOD2* variants [51]. Likewise, adherent invasive *Escherichia coli* has been associated with IBD and linked to variants at *NOD2*, *ATG16L1*, and *IRGM1* [52,53]. Another interesting case is that of *CARD9*, a known risk factor for IBD in humans that plays a key role in antifungal immune responses [24]. Correspondingly, *Card9* mutant mice show an increased abundance of fungi that have been implicated in IBD-associated dysbiosis [54,55]. Finally, variants at key proinflammatory cytokines, such as interferon (IFN)γ, are associated with an abundance of *Akkermansia muciniphila*, a microbe whose reduced or increased abundance in the human intestinal tract has been linked to several disease states (reviewed in [56]), and IFNγ-deficient mice have increased *A. muciniphila* populations in the gut [57]. Variants in genes affecting the architecture of the intestinal mucosa may also alter the composition of the microbiome towards dysbiosis, and ultimately disease. *FUT2* is a fucosyltransferase that removes sugars from mucin, the major biological barrier between the mucosal cell wall and the microbiota, and *Muc2*$^{-/-}$ mice develop severe colitis due to the absence of this protective barrier [58]. *FUT2* is a risk factor for IBD, and *FUT2* variants have been found to impact gut microbiota, including the abundance of *Faecalibacterium* and Proteobacteria, changes of which are associated with IBD [59,60]. Thus, host genetic variation can alter the microbial microenvironment and disrupt the ability of the host to maintain proper control over commensals. This dysbiotic microbiome can

then be a factor in promoting disease progression, rather than merely being a by-product of disease occurrence.

Similar two-way host genetic–microbiome effects have also been reported in obesity. Obesity can be transferred from humans to germ-free mice by fecal transplant [61]. Also, patterns of obesity-associated microbiome show inheritance in twin studies (monozygotic vs. dizygotic) [44], and obesity has been associated with several heritable microbial taxa [62], including the family Christensenellaceae, *Methanobrevibacter smithii*, the genus *Blautia* and the species *Akkermansia muciniphila* (negative association). Christensenellaceae occurrence in humans is about 40% dependent on host genetic makeup, making it one of the most heritable human bacteria currently identified [42]. *M. smithii* is found enriched in nonobese individuals compared with the obese, while *Blautia* is positively associated with obesity [63,64]. Finally, genetic variants at *SLC9A2*, *ELAVL4*, and *LINGO2* are associated with genetic risk of obesity and they also impact the abundance of *Blautia* [62]. These findings provide remarkable evidence of host genetic variation influencing the human commensal microbiome in a heritable manner, and further contribute to disease and overall host fitness.

Although these observations do not represent direct cause-to-effect relationships, they clearly point to human genetic variants impacting microbiome composition and properties. Direct characterization of the impact of specific genetic variants can be done using mouse models (Figure 2). Overall, these findings strongly suggest that microbiota can drive the incidence and pathogenesis of common human diseases in individuals bearing specific genetic risk. Because these conditions have significant morbidity and impact reproductive fitness, the commensal microbiome may act to modulate transmission of these genetic variants.

Impact of the Microbiome on the Human Genome

Although the capacity of microbial pathogens to influence human genome evolution is obvious, the capacity of commensal microbes to do so is more difficult to demonstrate. Clear pressure of the microbial world on the human genome is visible in pathogen-associated molecular pattern (PAMP) genes, a system that pre-dates evolution of the 'modern' adaptive immune system. This innate immune defense system is comprised of a number of gene families (*TLR*, *NLR*, *NOD*, *IPAF*) expressed on the surface and inside most cells, and that have broad specificity for many microbes or their products, including commensal microbes that inhabit mucosal surfaces [65]. Members of these gene families display relatively rapid evolution, as testified by important interspecies differences in gene copy number and associated sequence conservation. For example, the family of receptors that recognize the major cytomegalovirus (CMV) associated surface protein in humans (*KIR*) have 14 highly polymorphic gene copies [66]. Similar situations exist for other innate receptor gene families [67].

The host microbiota is instrumental in educating and priming the immune system during development, and it also modulates its activity in adult life (Figure 3). For example, germ-free mice show defects in multiple immune cell populations, including compromised T helper type 2 (Th2) skewing of T cells ($CD4^+$) [68], compromised innate lymphoid cell function [69], deficient IgA production [70], and susceptibility to infections [71,72]. More specifically, filamentous segmented bacteria (SFB) direct a T-helper cell type 17 (Th17) response in T cells [73], *Bacteroides* glycolipids inhibit natural killer T (NKT) cells [74], and several groups of bacteria induce $CD4^+Foxp3^+$ suppressor T cells [75]. In a recent systematic study of germ-free mice colonized individually with 53 single bacterial species from the human gut, Geva-Zatorsky analyzed the impact of colonization on the number and function of 18 different immune cell populations. The authors found that most of the strains induced (i) expansion of dendritic cells ($CD103^+$) that play a central role in priming, activating, and inducing tolerance of the immune system, and (ii) contraction of the plasmacytoid dendritic cell (pDC) compartment. A subset of strains induced regulatory T cells (Tregs), and/or modulated the expression of interleukin (IL)-22 by innate lymphoid cells [6]. In another study of the relationship between interindividual variation in the human gut microbial composition and inflammatory

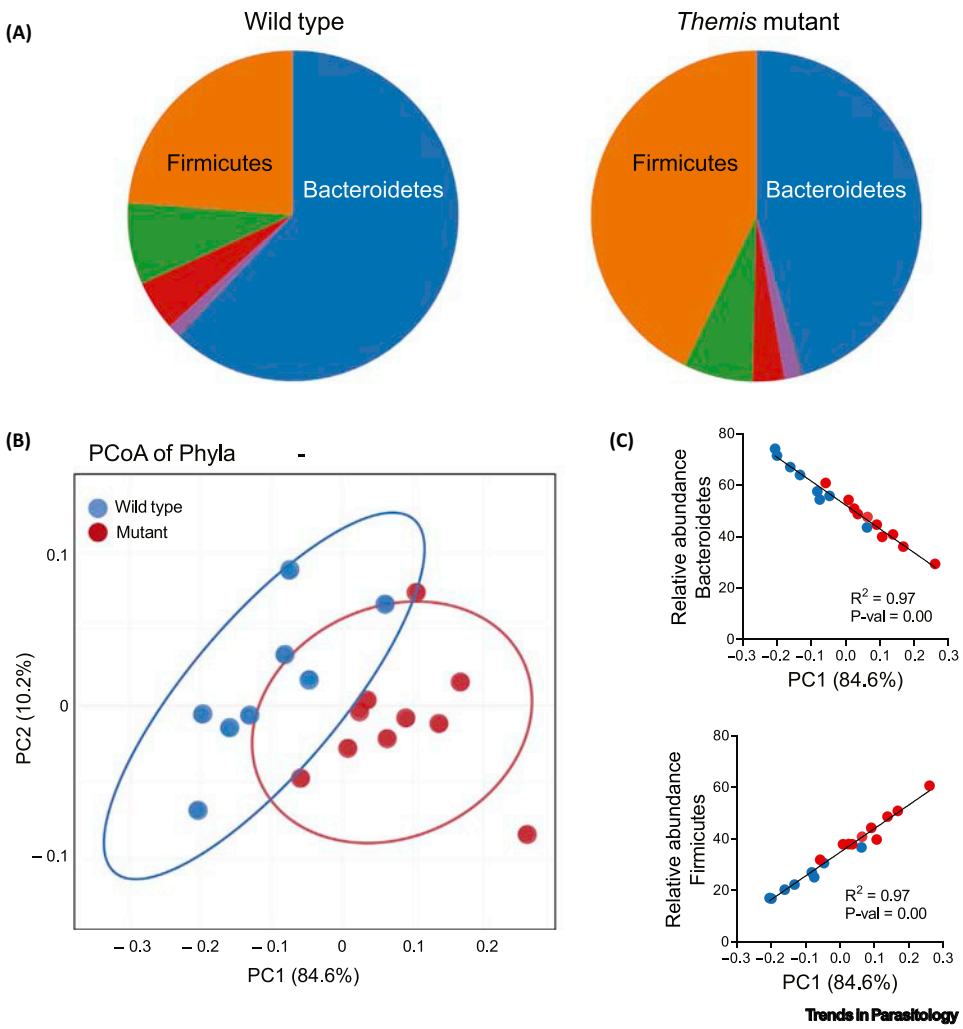

Figure 2. Effect of Host Genetic Variation on the Composition of the Gut Microbiome.

Genetic variants in the human genome may affect the composition and activity of the gut microbiome. These changes may (i) drive the onset and progression of certain diseases *in situ*, including inflammatory bowel disease, and (ii) act at a distance through the production of certain microbial metabolites linked to obesity, asthma, and other diseases that cause significant morbidity and reduced fitness, possibly limiting transmission of susceptibility alleles in permissive hosts. These two-way relationships are difficult to discern in humans, but can be tested more readily in mouse models. In the example below, a loss of function mutation in *Themis*, a gene important for the development and function of the immune system, was introduced in an F2 genetic cross and the effect of the mutation on the composition of the gut microbiome was determined in co-housed wild-type and homozygote mutant mice. The results show that the *Themis* mutation has a very strong effect on the microbial populations of the gut; importantly, human *THEMIS* variants have been associated with a number of chronic inflammatory conditions, including inflammatory bowel disease (IBD) [66]. The gut microbiome of wild type and homozygote *Themis* mutant mice was characterized by 16S V4 rRNA sequencing. (A) Pie chart representing relative abundance of bacterial phyla identifies enrichment of Firmicutes, and reduction in Bacteroidetes in *Themis*-deficient mice. (B) Bray–Curtis dissimilarity-based principal components analysis (PCoA) of microbial community composition show the impact of the *Themis* mutation at the phyla level (with normal confidence ellipses). (C) Relationship between selected phylum and the first or second principal coordinates of the principal component analysis, and identifying string and direct correlation between host genotype and abundance of these two phyla. Abbreviations: PC1, Principal component 1; PC2, Principal Component 2.

cytokine response to microbial stimulation *in vitro*, certain bacterial species were found to predict the cytokine production capacity of their host; these microbiome–cytokine interactions were stimulus-specific, and cytokine-specific. In addition, production of tumor necrosis factor (TNF) and IFNγ was correlated with specific microbial biochemical pathways, including palmitoic acid metabolism and tryptophan degradation, suggesting that microbiome effects on human immune cell function can be operating systemically at a distance from the gut [5]. Finally, germ-free mice can develop healthy immune systems upon colonization with normal mice microbiota, but not microbes from humans or rats [76]. These represent clear examples of the importance of microbial populations that have coevolved with their host in generating dynamic immune systems that enhance host health and fitness.

It is tempting to speculate that the modulatory activity of the gut microbiota on development and function of the immune system may further reveal or exacerbate genetically determined susceptibility to infectious and inflammatory diseases in humans over the life course, hence modulating transmission of these alleles in the population over time. The associated reduced fitness of afflicted individuals may provide one mechanism by which the microbiome may contribute over time to retention or elimination of certain human genotypes.

Indeed, there are some documented host–microbiome interactions in humans and in animal models of diseases that would clearly alter transmission of certain human genotypes. Some clear examples include the modulatory effect of the microbiome on penetrance and expressivity of lethal cancer-causing mutations (leukemia-inducing transgenic *Tet-2* oncogene) [77], ultimately determining survival or death of *Tet-2* transgenic mice; this phenotype was shown to be driven by translocation of normal flora from their normal niche in the gut to other organs. Similarly, in systemic lupus erythematosus prone-mice, translocation of the commensal *Enterococcus gallinarum* from the colon to the liver can induce lupus-like symptoms, while not altering disease susceptibility in healthy nonlupus-prone mice [78]. Additionally, certain human endogenous retroviruses (hERVs), remnants of exogenous retroviral infections that have become fixed in the human genome, have been implicated in the crosstalk between tumor and the microenvironment and worsening prognoses in several cancers [79]. Furthermore, it has been observed that coinfection with other microbes is sometimes required to reveal (or at least hasten) such host gene–pathobiont interactions. The human *ATG16L1* gene is an established risk factor for IBD. However, *Atg16l1*$^{-/-}$ mouse mutants are not susceptible to chemically induced colitis unless they are infected with norovirus (MNV), a virus that alters murine immune responses but which itself does not induce colitis on its own. This 'super-susceptibility' combination causes a phenotype like that seen in Crohn's disease patients. Importantly, this phenotype is driven by (i) the host inflammatory response, and (ii) the presence of a normal gut microflora as it is abrogated by treatment with antibiotics [80].

The gut microflora has also been identified as a potential driver of pathogenesis in neurodegenerative diseases such as Alzheimer's disease (AD) and Parkinson's disease (PD) [81,82]. These diseases are caused by the accumulation of misfolded proteins as ordered polymers with prion-like properties propagating in the human body. Evidence suggests that such pathological proteins could travel from neuroendocrine cells in the gut to the brain (gut–brain axis), by-passing the circulation. In PD, the initial misfolding event of α-synuclein appears to originate from the gut where it would be driven by microbiota: (i) PD patients show altered gut microflora [83], (ii) α-synuclein is deposited in gut neurons early during pathogenesis, and is later found in cell bodies of neurons innervating the gut [84]; (iii) in mouse models of PD, the intestinal microbiota is required for pathogenesis, and intestinal flora from PD patients exacerbates pathogenesis more strongly than microbiota from normal individuals [85]. Similar observations have been made for AD [81]. In both PD and AD, it has been proposed that gut bacteria may produce biophysical mimics of the pathological ordered polymers accumulating in these diseases, and that are normally involved in biofilm formation but that could act to enhance cerebral nucleation of Aβ aggregates [86,87]. Molecular mimicry by microbial products has been proposed to also play a role in autoimmune diseases, although the molecular mechanisms involved (direct, indirect) remain to be elucidated [7].

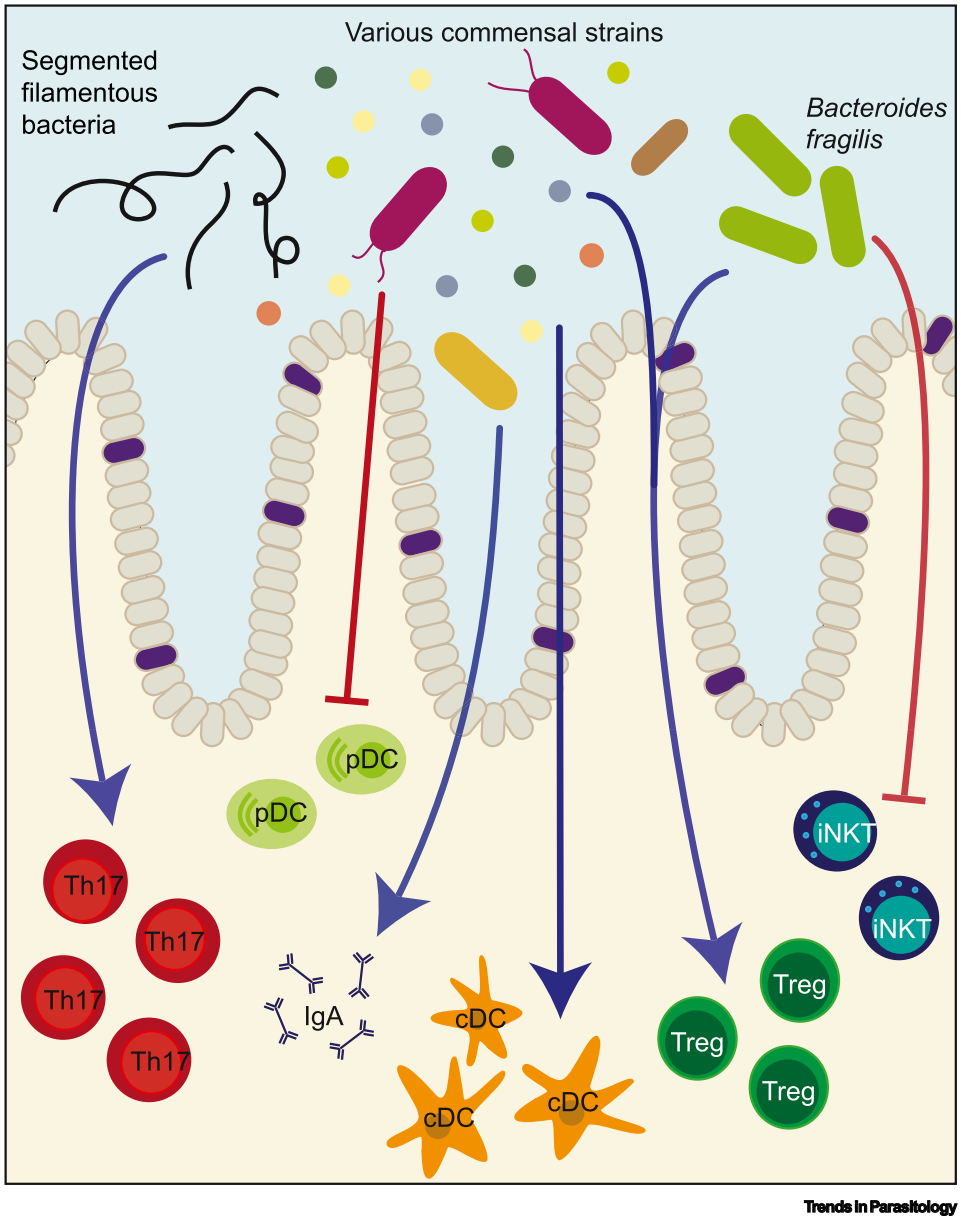

Figure 3. Effects of the Gut Microbiome on Priming and Training the Immune System.
The gut microbiome exerts key immunomodulatory effects on the development and activity of the host's immune system. Clear and replicated examples of host–microbiome immunomodulatory interactions include induction of T-helper cell 17 (Th17) T cells by segmented filamentous bacteria [73], the *Bacteroides fragilis*-mediated inhibition of invariant natural killer T cells (iNKT) differentiation [74], as well as induction of regulatory T cells (Tregs) [75]. Various commensal strains can also activate plasmacytoid dendritic cells (pDCs) (i.e., *Bacteroides vulgatus*) or inhibit their activity (i.e., *Staphylococcus saprophyticus*, *Lactobacillus rhamnosus*), or induce IgA production (i.e., *Fusobacterium varium*), and activate conventional dendritic cells (i.e., *Bacteroides uniformis*) [6]. Furthermore, several additional bacterial species have been identified as potent inducers of Th17 cells (i.e., *Bifidobacterium adolescentis*) and Treg cells (i.e., various *Clostridium* species) in mice [6].

Together, these studies are pointing to emerging paradigms in health and disease that emphasize a critical but previously underestimated three-way interaction between the microbiome, its human host, and the environment. Determining the magnitude of these effects on onset and progression of chronic illnesses in humans promises to be a rich area of investigation, with the potential for new predictive and diagnostic opportunities.

Concluding Remarks

There has been an explosion of research on the microbiome in recent years, with the onset of bulk genome sequencing yielding an unprecedented depth of understanding of the microbial populations living on the mucosal surfaces of humans. The complementary analysis of various metabolites and by-products produced systemically by these microbes has highlighted the impact that these microbial communities can exert at a distance on different organs and biological systems of the human body. Combined with the demonstrated role of the microbiome in educating and modulating the immune systems, this has led to a vast literature on the possible association of changes in the composition and activity of these communities and various disease states. This literature must be considered with caution as associations have not always been replicated, and even when replicated, the strength of some of the associations and the overall effect sizes have been fairly small. Careful validation in model systems is required for formal conclusions. The recognition that microbial populations are themselves subjected to modulatory pressure from environmental factors, such as diet and exposure to various chemical insults, and the fact that disease states in humans are also modulated by intrinsic genetic factors of the host, further complicates establishing cause-to-effect relationships. These three-way interactions need to be better understood for the design and implementation of novel molecular and microbiological tools of clinical value for the prediction, diagnosis, treatment, and resolution of diseases associated with the microbiome (see Outstanding Questions). Nevertheless, research on the microbiome in health and disease over the life course promises to be a rich area of exciting discoveries.

Outstanding Questions

Regarding the association of the microbiome with human disease, how can correlative evidence obtained so far be transformed into cause-to-effect relationships, including the use of appropriate model systems?

How do we untangle three-way relationships between, human genetics, the microbiome, and environment factors to fully understand the causality and progression of clinical disease associated with the microbiome?

How can surrogate markers be identified to inform the design of novel tools and standards of clinical value for the prevention, diagnosis, treatment, and recovery from diseases where the microbiome is thought to play a role?

Acknowledgments

Research in P.G.'s laboratory is supported in part by an award from the Canadian Institute for Advanced Research (Humans and the Microbiome Program), and by the Canadian Institutes of Health Research, and P.G. has salary support from a Distinguished James McGill Professorship from McGill University. T.J. is supported in part by a studentship from the Goodman Cancer Research Center.

References

1. McFall-Ngai, M. *et al.* (2013) Animals in a bacterial world, a new imperative for the life sciences. *Proc. Natl. Acad. Sci. U. S. A.* 110, 3229–3236
2. Davenport, E.R. *et al.* (2017) The human microbiome in evolution. *BMC Biol.* 15, 127
3. Jacob, J. *et al.* (1991) Intraclonal generation of antibody mutants in germinal centres. *Nature* 354, 389–392
4. Quach, H. and Quintana-Murci, L. (2017) Living in an adaptive world: genomic dissection of the genus *Homo* and its immune response. *J. Exp. Med.* 214, 877–894
5. Schirmer, M. *et al.* (2016) Linking the human gut microbiome to inflammatory cytokine production capacity. *Cell* 167, 1125–1136
6. Geva-Zatorsky, N. *et al.* (2017) Mining the human gut microbiota for immunomodulatory organisms. *Cell* 168, 928–943
7. Cusick, M.F. *et al.* (2012) Molecular mimicry as a mechanism of autoimmune disease. *Clin. Rev. Allergy Immunol.* 42, 102–111
8. Schirmer, M. *et al.* (2019) Microbial genes and pathways in inflammatory bowel disease. *Nat. Rev. Microbiol.* 17, 497–511
9. Nishida, A.H. and Ochman, H. (2019) A great-ape view of the gut microbiome. *Nat. Rev. Genet.* 20, 195–206
10. Raymann, K. and Moran, N.A. (2018) The role of the gut microbiome in health and disease of adult honey bee workers. *Curr. Opin. Insect Sci.* 26, 97–104
11. Bosch, T.C. *et al.* (2019) Evolutionary 'experiments' in symbiosis: the study of model animals provides insights into the mechanisms underlying the diversity of host–microbe interactions. *BioEssays, 1800256*
12. Clayton, J.B. *et al.* (2018) The gut microbiome of nonhuman primates: lessons in ecology and evolution. *Am. J. Primatol.* 80, e22867
13. Lamason, R.L. *et al.* (2005) SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science* 310, 1782–1786
14. Sabeti, P.C. *et al.* (2007) Genome-wide detection and characterization of positive selection in human populations. *Nature* 449, 913–918
15. Simonson, T.S. *et al.* (2010) Genetic evidence for high-altitude adaptation in Tibet. *Science* 329, 72–75

16. Tishkoff, S.A. *et al.* (2007) Convergent adaptation of human lactase persistence in Africa and Europe. *Nat. Genet.* 39, 31–40
17. Hozumi, N. and Tonegawa, S. (1976) Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc. Natl. Acad. Sci. U. S. A.* 73, 3628–3632
18. Deschamps, M. *et al.* (2016) Genomic signatures of selective pressures and introgression from archaic hominins at human innate immunity genes. *Am. J. Hum. Genet.* 98, 5–21
19. Picard, C. *et al.* (2018) International union of immunological societies: 2017 primary immunodeficiency diseases committee report on inborn errors of immunity. *J. Clin. Immunol.* 38, 96–128
20. Russell, S.M. *et al.* (1995) Mutation of Jak3 in a patient with SCID: essential role of Jak3 in lymphoid development. *Science* 270, 797–800
21. McPhail, L.C. *et al.* (1977) Deficiency of NADPH oxidase activity in chronic granulomatous disease. *J. Pediatr.* 90, 213–217
22. Zhang, S.-Y. *et al.* (2007) TLR3 deficiency in patients with herpes simplex encephalitis. *Science* 317, 1522–1527
23. Puel, A. *et al.* (2011) Chronic mucocutaneous candidiasis in humans with inborn errors of interleukin-17 immunity. *Science* 332, 65–68
24. Glocker, E.-O. *et al.* (2009) A homozygous CARD9 mutation in a family with susceptibility to fungal infections. *N. Engl. J. Med.* 361, 1727–1735
25. Rosain, J. *et al.* (2018) Mendelian susceptibility to mycobacterial disease: 2014–2018 update. *Immunol. Cell Biol.* 97, 360–367
26. Lindo, J. *et al.* (2016) A time transect of exomes from a Native American population before and after European contact. *Nat. Commun.* 7, 13175
27. Mowat, A. (2017) Why does cystic fibrosis display the prevalence and distribution observed in human populations? *Curr. Pediatr. Res.* 21, 164–171
28. Torre, S. *et al.* (2018) Genetic analysis of cerebral malaria in the mouse model infected with *Plasmodium berghei*. *Mamm. Genome* 29, 488–506
29. Fehr, T. *et al.* (1997) Crucial role of interferon consensus sequence binding protein, but neither of interferon regulatory factor 1 nor of nitric oxide synthesis for protection against murine listeriosis. *J. Exp. Med.* 185, 921–932
30. Marquis, J.-F. *et al.* (2009) Disseminated and rapidly fatal tuberculosis in mice bearing a defective allele at IFN regulatory factor 8. *J. Immunol.* 182, 3008–3015
31. Berghout, J. *et al.* (2013) Irf8-regulated genomic responses drive pathological inflammation during cerebral malaria. *PLoS Pathog.* 9, e1003491
32. Torre, S. *et al.* (2015) THEMIS is required for pathogenesis of cerebral malaria and protection against pulmonary tuberculosis. *Infect. Immun.* 83, 759–768
33. Langlais, D. *et al.* (2016) The macrophage IRF8/IRF1 regulome is required for protection against infections and is associated with chronic inflammation. *J. Exp. Med.* 213, 585–603
34. Patel, S.N. *et al.* (2008) C5 deficiency and C5a or C5aR blockade protects against cerebral malaria. *J. Exp. Med.* 205, 1133–1143
35. Tuite, A. *et al.* (2004) Genetic analysis of innate immunity in resistance to *Candida albicans*. *Genes Immun.* 5, 576–587
36. Conroy, A.L. *et al.* (2011) Complement activation: a critical mediator of adverse fetal outcomes in placental malaria? *Trends Parasitol.* 27, 294–299
37. Krause, J. *et al.* (2007) The derived FOXP2 variant of modern humans was shared with Neandertals. *Curr. Biol.* 17, 1908–1912
38. Inchley, C.E. *et al.* (2016) Selective sweep on human amylase genes postdates the split with Neanderthals. *Sci. Rep.* 6, 37198
39. Dannemann, M. *et al.* (2016) Introgression of Neandertal- and Denisovan-like haplotypes contributes to adaptive variation in human toll-like receptors. *Am. J. Hum. Genet.* 98, 22–33
40. Pierik, M. *et al.* (2006) Toll-like receptor-1,-2, and-6 polymorphisms influence disease extension in inflammatory bowel diseases. *Inflamm. Bowel Dis.* 12, 1–8
41. Huttenhower, C. *et al.* (2012) Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214
42. Goodrich, J.K. *et al.* (2014) Human genetics shape the gut microbiome. *Cell* 159, 789–799
43. Xie, H. *et al.* (2016) Shotgun metagenomics of 250 adult twins reveals genetic and environmental impacts on the gut microbiome. *Cell Syst.* 3, 572–584
44. Goodrich, J.K. *et al.* (2016) Genetic determinants of the gut microbiome in UK twins. *Cell Host Microbe* 19, 731–743
45. Hall, A.B. *et al.* (2017) Human genetic variation and the gut microbiome in disease. *Nat. Rev. Genet.* 18, 690–699
46. Blekhman, R. *et al.* (2015) Host genetic variation impacts microbiome composition across human body sites. *Genome Biol.* 16, 191
47. Bonder, M.J. *et al.* (2016) The effect of host genetics on the gut microbiome. *Nat. Genet.* 48, 1407–1412
48. Wang, J. *et al.* (2016) Genome-wide association analysis identifies variation in vitamin D receptor and other host factors influencing the gut microbiota. *Nat. Genet.* 48, 1396–1406
49. Yilmaz, B. *et al.* (2019) Microbial network disturbances in relapsing refractory Crohn's disease. *Nat. Med.* 25, 323–336
50. Langlais, D. *et al.* (2017) Genetics of infectious and inflammatory diseases: overlapping discoveries from association and exome-sequencing studies. *Annu. Rev. Immunol.* 35, 1–30
51. Knights, D. *et al.* (2014) Complex host genetics influence the microbiome in inflammatory bowel disease. *Genome Med.* 6, 107
52. Chassaing, B. *et al.* (2011) Crohn disease-associated adherent–invasive *E. coli* bacteria target mouse and human Peyer's patches via long polar fimbriae. *J. Clin. Invest.* 121, 966–975
53. Lapaquette, P. *et al.* (2012) Defects in autophagy favour adherent-invasive *Escherichia coli* persistence within macrophages leading to increased pro-inflammatory response. *Cell. Microbiol.* 14, 791–807
54. Lamas, B. *et al.* (2016) CARD9 impacts colitis by altering gut microbiota metabolism of tryptophan into aryl hydrocarbon receptor ligands. *Nat. Med.* 22, 598–605
55. Lewis, J.D. *et al.* (2015) Inflammation, antibiotics, and diet as environmental stressors of the gut microbiome in pediatric Crohn's disease. *Cell Host Microbe* 18, 489–500

56. Geerlings, S. *et al.* (2018) *Akkermansia muciniphila in* the human gastrointestinal tract: when, where, and how? *Microorganisms* 6, 75
57. Greer, R.L. *et al.* (2016) *Akkermansia muciniphila* mediates negative effects of IFNγ on glucose metabolism. *Nat. Commun.* 7, 13329
58. Johansson, M.E. *et al.* (2008) The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 105, 15064–15069
59. Tong, M. *et al.* (2014) Reprograming of gut microbiome energy metabolism by the FUT2 Crohn's disease risk polymorphism. *ISME J.* 8, 2193–2206
60. Frank, D.N. *et al.* (2007) Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc. Natl. Acad. Sci. U. S. A.* 104, 13780–13785
61. Ridaura, V.K. *et al.* (2013) Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science* 341, 1241214
62. Beaumont, M. *et al.* (2016) Heritable components of the human fecal microbiome are associated with visceral fat. *Genome Biol.* 17, 189
63. Million, M. *et al.* (2013) Correlation between body mass index and gut concentrations of *Lactobacillus reuteri, Bifidobacterium animalis, Methanobrevibacter smithii* and *Escherichia coli*. *Int. J. Obes.* 37, 1460–1466
64. Lahti, L. *et al.* (2013) Associations between the human intestinal microbiota, *Lactobacillus rhamnosus* GG and serum lipids indicated by integrated analysis of high-throughput profiling data. *PeerJ* 1, e32
65. Kawai, T. and Akira, S. (2009) The roles of TLRs, RLRs and NLRs in pathogen recognition. *Int. Immunol.* 21, 317–337
66. Hsu, K.C. *et al.* (2002) The killer cell immunoglobulin-like receptor (KIR) genomic region: gene-order, haplotypes and allelic polymorphism. *Immunol. Rev.* 190, 40–52
67. Fodil, N. *et al.* (2016) Primary immunodeficiencies and inflammatory disease: a growing genetic intersection. *Trends Immunol.* 37, 126–140
68. Mazmanian, S.K. *et al.* (2005) An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* 122, 107–118
69. Gury-BenAri, M. *et al.* (2016) The spectrum and regulatory landscape of intestinal innate lymphoid cells are shaped by the microbiome. *Cell* 166, 1231–1246
70. Hapfelmeier, S. *et al.* (2010) Reversible microbial colonization of germ-free mice reveals the dynamics of IgA immune responses. *Science* 328, 1705–1709
71. Sprinz, H. *et al.* (1961) The response of the germ-free guinea pig to oral bacterial challenge with *Escherichia coli* and *Shigella flexneri*: with special reference to lymphatic tissue and the intestinal tract. *Am. J. Pathol.* 39, 681–695
72. Zachar, Z. and Savage, D.C. (1979) Microbial interference and colonization of the murine gastrointestinal tract by *Listeria monocytogenes*. *Infect. Immun.* 23, 168–174
73. Ivanov, I.I. *et al.* (2009) Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell* 139, 485–498
74. An, D. *et al.* (2014) Sphingolipids from a symbiotic microbe regulate homeostasis of host intestinal natural killer T cells. *Cell* 156, 123–133
75. Round, J.L. and Mazmanian, S.K. (2010) Inducible $Foxp3^{+}$ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. *Proc. Natl. Acad. Sci. U. S. A.* 107, 12204–12209
76. Chung, H. *et al.* (2012) Gut immune maturation depends on colonization with a host-specific microbiota. *Cell* 149, 1578–1593
77. Meisel, M. *et al.* (2018) Microbial signals drive pre-leukaemic myeloproliferation in a Tet2-deficient host. *Nature* 557, 580–584
78. Manfredo Vieira, S. *et al.* (2018) Translocation of a gut pathobiont drives autoimmunity in mice and humans. *Science* 359, 1156–1161
79. Smith, C.C. *et al.* (2018) Endogenous retroviral signatures predict immunotherapy response in clear cell renal cell carcinoma. *J. Clin. Invest.* 128, 4804–4820
80. Cadwell, K. *et al.* (2010) Virus-plus-susceptibility gene interaction determines Crohn's disease gene Atg16L1 phenotypes in intestine. *Cell* 141, 1135–1145
81. Friedland, R.P. and Chapman, M.R. (2017) The role of microbial amyloid in neurodegeneration. *PLoS Pathog.* 13, e1006654
82. Matheoud, D. *et al.* (2019) Intestinal infection triggers Parkinson's disease-like symptoms in $Pink1^{-/-}$ mice. *Nature* 571, 565–569
83. Hill-Burns, E.M. *et al.* (2017) Parkinson's disease and Parkinson's disease medications have distinct signatures of the gut microbiome. *Mov. Disord.* 32, 739–749
84. Kujala, P. *et al.* (2011) Prion uptake in the gut: identification of the first uptake and replication sites. *PLoS Pathog.* 7, e1002449
85. Sampson, T.R. *et al.* (2016) Gut microbiota regulate motor deficits and neuroinflammation in a model of Parkinson's disease. *Cell* 167, 1469–1480
86. Zhou, Y. *et al.* (2012) Promiscuous cross-seeding between bacterial amyloids promotes interspecies biofilms. *J. Biol. Chem.* 287, 35092–35103
87. Chen, S.G. *et al.* (2016) Exposure to the functional bacterial amyloid protein curli enhances alpha-synuclein aggregation in aged Fischer 344 rats and *Caenorhabditis elegans*. *Sci. Rep.* 6, 34477
88. Miller, L.H. *et al.* (1976) The resistance factor to *Plasmodium vivax* in blacks: the Duffy-blood-group genotype, FyFy. *N. Engl. J. Med.* 295, 302–304
89. Culleton, R.L. *et al.* (2008) Failure to detect *Plasmodium vivax* in West and Central Africa by PCR species typing. *Malar. J.* 7, 174–182
90. Jarolim, P. *et al.* (1991) Deletion in erythrocyte band 3 gene in malaria-resistant Southeast Asian ovalocytosis. *Proc. Natl. Acad. Sci. U. S. A.* 88, 11022–11026
91. Cattani, J.A. *et al.* (1987) Hereditary ovalocytosis and reduced susceptibility to malaria in Papua New Guinea. *Trans. R. Soc. Trop. Med. Hyg.* 81, 705–709
92. Genton, B. *et al.* (1995) Ovalocytosis and cerebral malaria. *Nature* 378, 564–568
93. Hill, A.V. *et al.* (1991) Common West African HLA antigens are associated with protection from severe malaria. *Nature* 352, 595–600
94. Ackerman, H. *et al.* (2005) A comparison of case-control and family-based association methods: the

example of sickle-cell and malaria. *Ann. Hum. Genet.* 69, 559–565

95. Andersen, K.G. *et al.* (2012) Genome-wide scans provide evidence for positive selection of genes implicated in Lassa fever. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 367, 868–877
96. Rowland-Jones, S. *et al.* (1995) HIV-specific cytotoxic T-cells in HIV-exposed but uninfected Gambian women. *Nat. Med.* 1, 59–64
97. Samson, M. *et al.* (1996) Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* 382, 722–725
98. Huang, Y. *et al.* (1996) The role of a mutant CCR5 allele in HIV-1 transmission and disease progression. *Nat. Med.* 2, 1240–1243
99. Glass, R.I. *et al.* (1985) Predisposition for cholera of individuals with O blood group possible evolutionary significance. *Am. J. Epidemiol.* 121, 791–796
100. Saleh, M. *et al.* (2004) Differential modulation of endotoxin responsiveness by human caspase-12 polymorphisms. *Nature* 429, 75–79
101. Xue, Y. *et al.* (2006) Spread of an inactive form of caspase-12 in humans is due to recent positive selection. *Am. J. Hum. Genet.* 78, 659–670