Chapter 1
The Human Genome, Microbiomes, and Disease

Jonathan H. Badger, Pauline C. Ng, and J. Craig Venter

Disease may stem from the environment, genetics, and the human microbiome. In this chapter, we discuss what is currently known about the genetic and microbial contribution to human disease. Because this is a large area, we provide a high-level review of a few diseases to reveal underlying themes of the contribution of genetics and metagenomics to human disease. As changing a microbiome is a far easier task than changing one's genome, manipulating it may prove an effective way to both treat and prevent disease where causal disease relationships can be established. Elucidating the interplay of the contributions of the genome and microbiome will be necessary to help us better understand disease.

Introduction: Variation in the Genome and Microbiome

The past decade in human genetic disease has been dominated by genome-wide association studies and, more recently, by sequencing human genomes. The goal is to discover disease genes by identifying variants that cause or increase the risk of disease. In genome-wide association studies, patients with disease are compared with unaffected controls. DNA is collected from both groups, and approximately a million genetic variants are surveyed in their genomes. Variants that are at a higher frequency in patients than in controls are associated with an increased risk of disease. More than a thousand disease-associated variants have been identified from over 400 genome-wide association studies, but much is still unknown about the genetic heritability of disease. For example, genome-wide association studies have identified 40 markers associated with height, but these markers only account for ~5% of height's heritability (Maher, 2008). Despite studying thousands of individuals, <10% of heritability is explained for most diseases (Manolio et al., 2009).

J.H. Badger (✉)
J. Craig Venter Institute, 10355 Science Center Drive, San Diego, CA, USA
e-mail: jbadger@jcvi.org

Jonathan H. Badger and Pauline C. Ng contributed equally.

K.E. Nelson (ed.), *Metagenomics of the Human Body*,
DOI 10.1007/978-1-4419-7089-3_1,

Individual human genomes have been sequenced, and there are approximately 3 million to 4 million variations with respect to the reference genome (Frazer et al., 2009). It is thought that some of these variants will cause phenotypic differences that can lead to disease or apparent physical traits. It is estimated that 3–8% of the human genome is functional (Siepel et al., 2005), so it is unlikely that all the variation in the 3 Gb human genome will lead to phenotypic differences. Rather, functional variants may be localized to the 90–240 Mb of human genome that contains transcribed coding genes, regulatory elements, RNA genes, and other functional elements.

By contrast, the human microbiome has extensive diversity. Each location (skin, mouth, intestine, etc) has its own metagenome. Recent studies have suggested that healthy individuals have up to 15,000 species-level phylotypes in their gastrointestinal tracts as determined by 16S rRNA sequencing (>97% identity) (Peterson et al., 2008) (Fig. 1.1 of this chapter), and that the two major phylogenetic groups present are the Firmicutes and Bacteriodetes. The average genome size of sequenced organisms from these groups is 3.4 Mb (Liolios et al., 2008), and the percentage of these genomes that codes for protein-coding genes is approximately 92%. Therefore, the functional part of the gastrointestinal microbiome can be estimated to be approximately 47,000 Mb (15,000 × 3.4 × 0.92), which is more than two orders of magnitude greater than the above-mentioned estimate of the functional part of the human genome.

Because human metagenomics is a nascent field, fewer individuals have been studied relative to the human genetics field. In a similar manner, the microbiome compositions observed in patients differ from those of the controls for some diseases. Like inherited genetic variation, this is not a complete predictor: species that

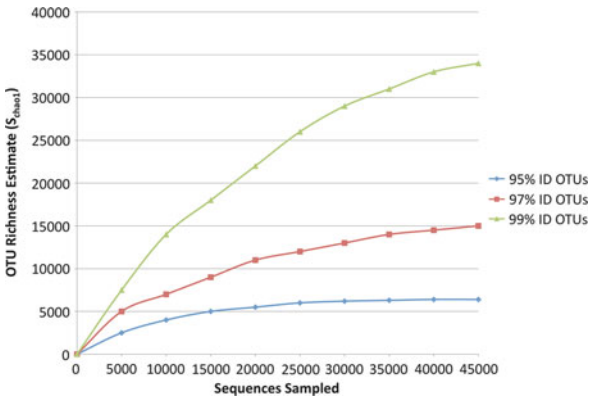

Fig. 1.1 OTU richness vs. sequences sampled (adapted from Peterson et al. (2008)). 16S rRNA collector's curves showing the number of observed operational taxonomic units (OTUs) given the number of sequences collected and a pairwise sequence identity (ID%) cutoff. A 99% ID level correlates with accepted strain classifications, likewise 97% with species and 95% with genus. When a collector's curve begins to plateau, more sequencing is unlikely to return new OTUs at that percent identity level

are predominant in the disease will also be observed in controls, and vice versa. In most cases, the microbiomes' constitution is associated rather than causative, meaning that while a correlation between the disease state and the microbiome has been established, there is no evidence that the microbial constitution caused the disease.

The microbiome is an ecosystem in which the various members maintain equilibrium. So far, it appears that many diseases are associated with an abnormal proportion of the same taxonomic groups that are present in healthy individuals. For example, patients with Crohn's disease show a lower than normal frequency of bacteria from the phylum Bacteriodetes in their gastrointestinal tract (e.g., Gophna et al., 2006), whereas patients suffering from active celiac disease have a higher than normal frequency of Bacteriodetes (e.g., Nadal et al., 2007). Biodiversity also plays a role in the microbiome. Patients with bacterial vaginosis have a less diverse vaginal microbiome than healthy individuals, suggesting that the biodiversity normally residing in a healthy vagina may serve as a resilient buffering against disease.

Another common theme is that diseases with a microbial component are often disorders of the immune system. Psoriasis of the skin and Crohn's disease in the gut are some examples. Seeing as the immune system modulates the interaction between the microbiome and the native human body, it is not surprising that the immune system plays a huge role, but this connection may also hint at possible therapies.

Genome-wide association studies typically examine 10^6 variants and focus mainly on common variation. But recent technological advances have enabled sequencing whole genomes of human individuals, which can identify rare mutations that cause disease (Altshuler et al., 2008). Similarly, high-throughput sequencing will allow us to sequence metagenomes at greater depth to obtain finer resolution of the various species in a microbial sample.

Steps to Treatment

A human's DNA is not easily amenable to change. The genome in one cell is static and identical (for the most part) to the genomes in the individual's other cells. Thus, knowing disease-associated genetic variation does not mean you can directly treat the disease. The genome-wide association discoveries are meant to be followed up by functional studies to investigate biological pathways involved in disease, and then see which drugs target these pathways. This can take many years to translate genetics research into a clinical treatment. There are exceptions; age-related macular degeneration is one disease where treatment was rapidly implemented after genome-wide discoveries (Gehrs et al., 2006).

Microbiomes may be easier to treat. Antibiotics erase the current microbiome state. The use of probiotics, that is, ingesting bacteria to improve the intestinal microbial balance, is gaining popularity for multiple gastrointestinal diseases (Preidis and Versalovic, 2009). Probiotics (from the Greek "pro bios," meaning "for, or promoting, life") are foods or supplements containing live bacteria thought to promote health. While yogurt and other fermented products have traditionally been

considered "health food," studies of the human microbiome promise to understand what effect they have on the composition of the microbiome and ultimately health. Studies have shown that probiotics can manipulate the microbiome, as well as affect the intestinal barrier (Zyrek et al., 2007) and the immune system (Fitzpatrick et al., 2007).

Prebiotics are oligosaccharides or complex saccharides that stimulate the growth or activity of the beneficial commensal bacteria that already are present in the host. Although traditionally pro- and prebiotics have been defined in the context of the gastrointestinal tract, recent work has explored the use of probiotics (Lee et al., 2008) and prebiotics (Bockmuhl et al., 2007) on skin, and probiotics in the vagina (Mastromarino et al., 2009).

Factors in Experimental Design

Human studies can be complicated by population differences, because different ethnicities can have different patterns of linkage disequilibrium, and different allele frequencies can bias genetics studies. Furthermore, genome-wide association studies test 10^6 markers. As a consequence, a large number of individuals must be studied to get a significant *p*-value after correction for multiple testing.

The study of the microbiome has numerous complications as well. One of the most important ones is the choice of percent sequence identity to define an operational taxonomic unit (OTU). This cutoff can greatly affect the estimated sample diversity (see Fig. 1.1). A low value will group highly diverse organisms together, underestimating biodiversity, but a high value may overestimate biodiversity, particularly in the case of next-generation sequencing methods, as they often have a higher rate of sequencing error than the traditional Sanger method (Gomez-Alvarez et al., 2009).

Another choice that needs to be made is what to sequence in order to measure biodiversity. The cheapest option is to sequence one or more variable regions of the 16S rRNA gene (Huber et al., 2007). This has the advantage of not requiring assembly as such regions can fit on a single Roche 454^{TM} read. The next cheapest option is to sequence full-length (or nearly full-length) 16S rRNA sequences, typically by sequencing a forward and reverse read from a PCR-amplified product, followed by assembly. A complication in both forms of 16S rRNA diversity studies is the selection of PCR primers and annealing temperature because poor selections can greatly bias the results (Sipos et al., 2007). The most expensive, and yet potentially most informative method of studying microbial biodiversity is random environmental sequencing, known as metagenomics. More diversity is seen in the protein-coding parts of genomes from environmental sequencing than what is seen from 16S rRNA sequencing.

Finally, there are many factors that need to be considered in the study of the microbiome. Although many of them may be the same factors that influence genome-wide association studies such as age, sex, and ethnicity, there are also unique factors that need to be considered. As the microbiome is not static, time

of sampling is likely to be a factor. Time and composition of the last meal may play a major role in the gastrointestinal tract microbiome. Sanitary habits and use of cosmetics may influence the skin microbiome. Sexual activity and the current state of the menstrual cycle may influence the vaginal microbiome.

Given the diversity of the human microbiome relative to the human genome, there is a huge effort to understand the human microbiome. The Human Microbiome Project (HMP) aims to study the microbiomes of the nasal, oral, skin, gastrointestinal, and urogenital environments. By establishing reference microbiomes from healthy normal individuals, these can be compared to disease states to find potential causes and associations. Much like sequencing the human genome led to a better understanding of human genetic variation (Lander et al., 2001; Venter et al., 2001), the HMP is set up to do the same.

The Gastrointestinal Microbiome

Perhaps the body site that comes first to mind when the "human microbiome" is mentioned is the gastrointestinal system. The importance of the intestinal microbiota to health was first suggested a hundred years ago by the Russian Nobel-laureate microbiologist Ilya Ilyich Mechnikov, who hypothesized that the supposed health benefits of *kefir* (a traditional fermented milk beverage in Eastern Europe), derived from the ability of the live bacteria in the drink to colonize the intestine, displacing "unhealthy" bacteria and improving digestion (Metchnikov, 1908). Only recently have molecular methods allowed detailed studies of the relationships between disease and microbiota, particularly in relation to three disorders: Crohn's disease, celiac disease, and obesity.

Crohn's Disease

Crohn's disease is a common inflammatory bowel disease with high heritability. In Crohn's disease, it appears that the immune system misidentifies benign members of the commensal microbiota as pathogens and mounts an immune response against them (Strober et al., 2007). An analysis of three complete genome-wide scans looking at over 3,000 cases and 4,000 controls identified 31 loci involved in Crohn's disease and many were in or near immune-related genes (*IL23R*, *CCR6*, *IL12B*, *STAT3*, *ICOSLG*, *PTPN2*, *PTPN22*, *ITLN1*). The *ITLN1* gene is notable because it encodes a lectin that recognizes galactofuranosyl residues found in the cell walls of microorganisms but not present in mammals. Its physical contact with microbes may demonstrate a direct interaction between the microbiome and human genome. These 31 loci in humans explain about 20% of the heritability of Crohn's disease (Barrett et al., 2008). Either as a cause or a result of this immune response, Crohn's sufferers tend to have lower than normal levels of Firmicutes, Bacteriodetes (Frank et al., 2007; Gophna et al., 2006), and *Lachnospiraceae* (group IV and XIVa Clostridia; Peterson et al., 2008) in their gastrointestinal microbiota. As Peterson et al. mention (Peterson et al., 2008), this suggests a testable hypothesis; as the difference between

the diseased and normal microbiomes may either be causal or be reactive, restoring a normal microbiome to Crohn's sufferers would have no therapeutic effect if the microbiome is reactive to the disease, but would work as a treatment if it is causal. Peterson and colleagues propose an extensive long-term research program (Peterson et al., 2008) for understanding the nature of Crohn's and similar bowel diseases. This proposal comprises monozygotic twins studies, microbiota transplantation between healthy and diseased individuals, and characterization of the B- and T-cell repertoires of patients with differing microbiomes.

Celiac Disease

Celiac disease is an autoimmune disease in which the lining of the small intestine is damaged from eating protein gluten (found in bread, pasta, and other foods that contain wheat, barley, or rye). If an immune response is triggered due to inappropriate diet, damage can occur in the small intestine such that there is decreased nutrient absorption, and consequently vitamin deficiencies that can lead to stunted growth. Much of the known heritability (∼35%) is attributable to HLA haplotypes, which was discovered from a linkage study on 60 families (Liu et al., 2002). Genome-wide association studies on 2,000 individuals have subsequently identified an additional eight regions that account for ∼3–4% additional heritability (Hunt et al., 2008; van Heel et al., 2007). Seven of the eight celiac disease regions identify immune genes, involved in T or B cell function, indicating the importance of the immune system.

Initial studies of the microbiota of patients with active celiac disease (Nadal et al., 2007) have revealed that their duodenal microbiota differ from healthy individuals as well as from controlled celiac patients (those who have been on a strict gluten-free diet for over a year). In particular, the microbiota of the patients with active celiac disease had significantly higher percentage of bacteria from the *Bacteroides–Prevotella* group as well as a higher percentage of *Escherichia coli* as compared to the other two groups. Although members of these groups constitute a large fraction of the commensal microbiota in healthy individuals, some *Bacteroides* species (such as *Bacteroides vulgatus*) have been shown to cause immune-system-induced intestinal inflammation in animal models (Setoyama et al., 2003). Both the patients with active and controlled celiac disease displayed reduced levels of *Lactobacillus* and *Bifidobacterium* as compared to healthy individuals.

Obesity

Family and twin studies have shown that genetic factors account for 40–70% of the population variation in BMI (Frazer et al., 2009; Siepel et al., 2005). A meta-analysis of genome-wide association studies on more than 32,000 individuals by the GIANT identified 20 loci associated with obesity. Most of the obesity markers discovered so far have minor effects on obesity (odds ratios <2). The gene FTO, the marker with the strongest association with obesity, only accounts for ∼1% of the heritability (Walley et al., 2009).

Studies comparing the microbiomes of 31 identical twin pairs (Turnbaugh et al., 2009) revealed that obese twins tended to have a higher proportion of Firmicutes and a lower proportion of Bacteriodetes in their gastrointestinal microbiomes than their normal or underweight twin. In addition, the microbiomes of obese individuals were less diverse phylogenetically than in the normal case. Interestingly, similar trends were seen in 23 pairs of fraternal twins (Turnbaugh et al., 2009), who are not closer genetically to each other than other siblings. A related study (Turnbaugh et al., 2008) involving mice also saw an increase in Firmicutes and a decrease in Bacteriodetes when the mice were fed a diet high in fat and sugar (resembling a typical "Western" diet) but not when fed a low-fat diet. This suggests a link (perhaps causal) between a high-fat, high-sugar diet and obesity that is mediated by the microbiome.

Colorectal Cancer

Colorectal cancer is the third most common cancer and approximately 650,000 people die from colorectal cancer annually. Mutations in the FAP and HNPCC genes account for approximately 5% of colorectal cancer cases (http://www.genome.gov/10000466). Whole exome sequencing of colorectal cancer tissues have identified an additional 140 genes that were enriched with somatic cancer mutations in the cancer samples (Wood et al., 2007). These genes were shown to be mutated in >10% of colorectal cancers, and in other cancers.

A diet high in red meat and fat is associated with high risk of colorectal cancer. This can possibly be traced back to the colon microbiome. Many strains of Bacteroides species can convert bile to fecapenentaenes, which may cause cancer (see Fig. 1.3) (Moore and Moore, 1995). A comparison of colorectal cancer patients, high-risk individuals with Western diets, and low-risk individuals with non-Western diets showed that the low-risk individuals had a lower percentage of *B. vulgatus* (Fig. 1.4).

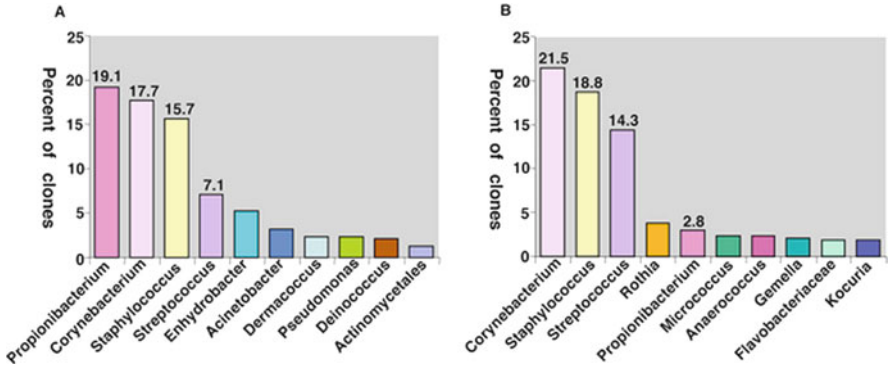

Fig. 1.2 Bacterial genera from healthy skin (**a**) and psoriatic lesions (**b**). Taken from Gao et al. (2008) under the Creative Commons Attribution License

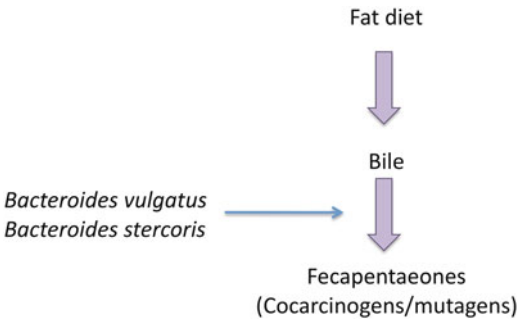

Fig. 1.3 A fatty diet can increase the risk of colorectal cancer

Vaginal Bacterial Biota

The vagina has different types of environments depending on age, hormonal fluctuations, sexual activity, sanitary habits such as douching, etc. In healthy women, the bacterial biota is dominated by lactobacilli with more diversity observed in African-American woman than in Caucasian women (Zhou et al., 2007).

Bacterial vaginosis is a common disease in women whose symptoms include abnormal vaginal discharge and increased susceptibility to HIV. The vaginal microbiome of women with bacterial vaginosis is more diverse than that of healthy women. Healthy women's vaginas were initially thought to be dominated by *Lactobacillus* species. *Lactobacillus* may provide a protective environment against pathogens, and disturbance of the ecology may eventually lead to bacterial vaginosis. It is thought that *Lactobacillus* metabolizes the glycogen to lactic acid in the vaginal epithelium to produce a low pH environment that prevents the growth of pathogenic organisms (Fig. 1.4) (Zhou et al., 2007).

However, based on 16S rRNA sequencing, *Lactobacillus* is not always found on healthy human vaginal epitheliums (Hyman et al., 2005; Zhou et al., 2007). *Gardnerella vaginalis* and *Atopobium vaginae* can also be detected in healthy women (Hyman et al., 2005), even though they are associated with bacterial vaginosis (Srinivasan and Fredricks, 2008). It has been suggested that the co-occurrence of *G. vaginalis* and *A. vaginae* may play a role in bacterial vaginosis (Bradshaw et al., 2006).

Human genetic variation also plays a role in HIV susceptibility. The CCR5 is the major HIV co-receptor expressed in the female genital tract (Patterson et al., 1998). In Caucasians, an allelic variant of the CCR5 with a 32-bp deletion that results in a frameshift and a nonfunctional receptor protects against HIV infection (Samson et al., 1996). CCR5 mRNA is significantly increased in a woman with STDs and inflammatory conditions, so the bacterial environment in conjunction with a woman's genetics could determine her susceptibility to HIV.

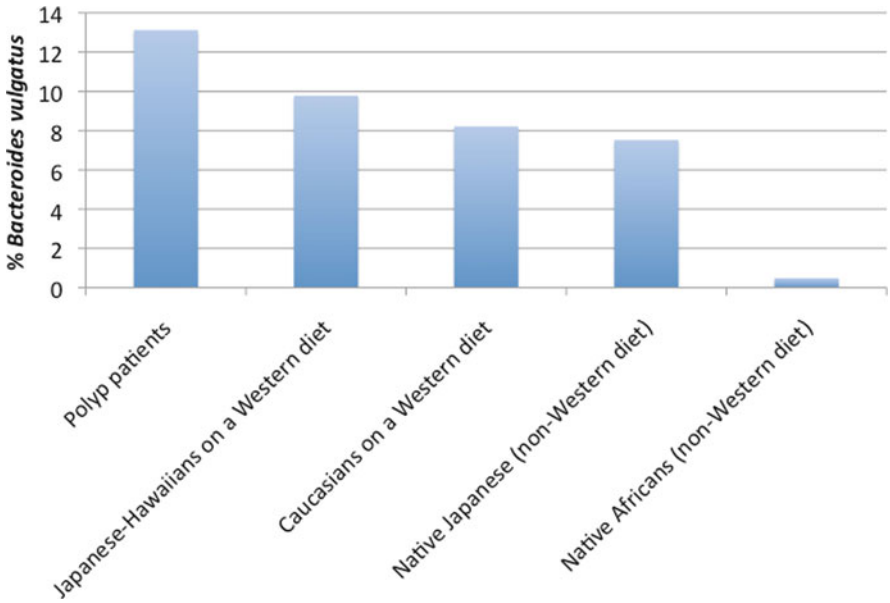

Fig. 1.4 Intestinal microflorae differ between individuals with low and high risk for colorectal cancer. From percentages in Moore and Moore, 1995

Oral

The oral environment is challenging because it undergoes constant changes in pH, redox potential, atmospheric conditions, salinity, and water activity from saliva (Avila et al., 2009). People who are diabetic, HIV positive, pregnant, lactating, or who have recently taken antibiotics can also have altered oral bacterial compositions (Lepp et al., 2004). Within the oral cavity, different environments (e.g., soft tissues and plaques) are dominated by different species (Aas et al., 2005). More than 700 species reside in the oral cavity, and some of these can cause disease. *Porphyromonas gingivalis*, *Tannerella forsythia*, and *Treponema denticola* have been found to be associated with periodontal disease; *Streptococcus mutans*, *Lactobacillus* spp., *Bifidobacterium* spp., and *Atopobium* spp. have been found to be associated with dental caries.

Periodontal disease (gingivitis and periodontitis) is a chronic bacterial infection. Besides the pathogenic microflora that causes periodontitis, genetics can also contribute to disease risk. Genetic studies have shown that approximately half of the population variance in chronic periodontitis is attributable to genetic factors. A genome-wide association study on aggressive periodontitis identified an intronic variant in the gene *GLT6D1*, which is expressed in gingiva (Schaefer et al., 2010). The G allele of the variant is found in 12% more of the periodontitis cases than in controls, and the variant appears to lie in a GATA-3 transcription factor binding site. Furthermore, the authors show that the disease-associated G allele has reduced

binding affinity for GATA-3. GATA-3 regulates Th2 cells and is involved in immunity against extracellular parasites. Further work is needed to elucidate possible interactions between GLT6D1 and the oral microbiome.

It is interesting to observe that the oral microbiome from children with severe cavities is much less diverse than their healthy counterparts (Li et al., 2007). Similar to Crohn's disease in the gastrointestinal tract and bacterial vaginosis in the vagina, the affected individuals have less diverse microflora than healthy individuals.

Skin Microbiome

The skin is the largest organ in the body, and within the skin, there are microenvironments such as sebaceous (e.g., upper chest and back), moist (e.g., armpit and inside of nostril), and dry (e.g., forearm and buttock) (Grice et al., 2009). Based on 16S rRNA sequencing, the majority of microbes are Proteobacteria (Grice et al., 2008). Skin also shows a low level of interpersonal variation in contrast to the gastrointestinal tract (Grice et al., 2008).

Psoriasis

Psoriasis is a common skin condition that causes skin redness and irritation. It is a chronic autoimmune disease that afflicts ~2% of the European population. The number of species per skin sample is not significantly different between healthy individuals and those with psoriasis, but the composition differs (Fig. 1.5). *Streptococcus* spp. is found more frequently in psoriatic lesions than in normal skin from the same patients. Meanwhile, *Propionibacterium* is observed more frequently on normal skin (21% from healthy, 12% from normal skin on psoriatic

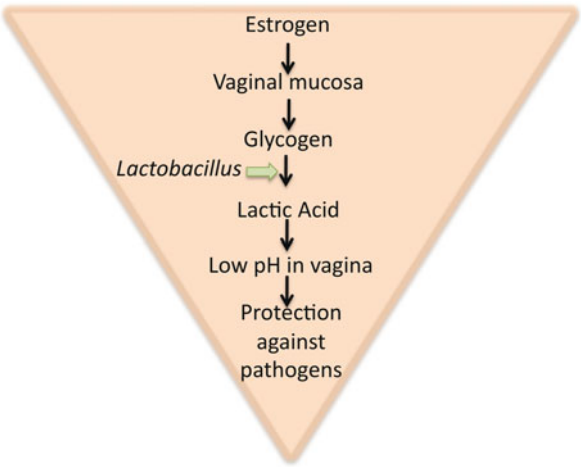

Fig. 1.5 The role of microbiota in a healthy vagina

Table 1.1 Genes with nearby variants that are strongly associated with psoriasis (Nair et al., 2009)

Notable nearby genes	Location of associated variant	Role in immune response
MHC		
HLA-C	13 kb upstream	Major histocompatibility complex class I heavy chain receptor which presents small peptides to the immune system, so that the immune system can recognize foreign antigens
IL-23 signaling		
IL12B	24 kb downstream	IL12B encodes subunit of IL-23
IL23A	3.7 kb downstream	encodes a subunit of IL-23
IL23R	Intron	subunit of IL-23 receptor
NF-κB Pathway		
TNFAIP3	Intron	TNFα-induced protein 3 inhibits NF-κB activation and terminates NF-κB-mediated responses
TNIP1	1 kb upstream	TNFAIP3-interacting protein 1
Other		
IL13	Nonsynonymous Coding	Modulate humoral immune response mediated by Th2 cells
STAT2	Intron	Responds to interferon. STAT2 interacts with P300/CBP, which is thought to be involved in the process of blocking IFN-α response by adenovirus

patients) than on psoriatic lesions (2.9%). It is not known if *Propionibacterium* protects the host, and removal of *Propionibacterium* therefore plays a causative role. Alternatively, *Propionibacterium* could simply be correlated to healthy skin, and a decreased number simply indicates its displacement from more aggressive microbes (Gao et al., 2008).

A genome-wide association study revealed that many of the variants that increase the risk of psoriasis are near, or in genes, that play a role in the human immune system (Table 1.1) (Nair et al., 2009). For example, variants in the IL-23 pathway (the IL12B, IL23A, IL23R genes) are associated with psoriasis. IL-23 signaling promotes cellular immune responses by promoting the survival and expansion of a subset of T cells that protects epithelia against microbial pathogens (Altshuler et al., 2008).

Acne

Acne vulgaris is the most common skin disorder and drives a multi-billion dollar industry. There are two stages of the disorder: (1) blackheads and whiteheads are formed in a noninflammatory phase and (2) the blackheads and whiteheads develop into inflamed lesions in an inflammatory phase (Leeming et al., 1985).

In the first stage, *Pityrosporum* spp., Stapylococci, and Propionibacteria were found more frequently at blackheads and whiteheads than at normal follicles, but were not found in all of these sites (Leeming et al., 1985). Mature blackheads are more frequently colonized, and it may be that microbial colonization is not necessary for whitehead/blackhead formation, but that colonization occurs as the whitehead/blackhead enlarges.

Propionibacterium acnes, which is present in both normal and acne-afflicted follicles, could be responsible for instigating inflammation in the second stage of acne (Burkhart et al., 1999). *Propionibacterium acnes* directly secretes factors that attract immune cells, and acne patients have an immune response to *P. acnes*. Thus, controlling the skin microbiome by decreasing the *P. acnes* population and/or inhibiting its production of inflammatory factors could be effective in treating acne.

Conclusion

In this chapter, we review some diseases known to involve both human genetics and the human microbiome. Undoubtedly, more relationships between microbiome and human genetics will be elucidated with further research. However, from this chapter we can summarize some general trends across diseases that are evident today. First, the disease state is often associated with a reduced level of diversity in the microbiome (Turnbaugh et al., 2009; Zhou et al., 2007; Li et al., 2007) as compared to the healthy state. This may mean that the disease state represents a breakdown in the complex ecosystem that is the microbiome. Second, many diseases associated with the microbiome are associated with problems in the immune system, particularly disorders that involve misidentifying commensals as pathogens (e.g., Crohn's disease, celiac disease, acne, and psoriasis).

The future of microbiome studies will likely include attempts at curing diseases through the modification of the microbiomes of diseased individuals. This should not be done in a vacuum but should be combined with genome-wide association studies and studies of the immune system in order to make the best use of the results. The synergy of both metagenomics and human genetics holds the promise to improve human health for many common diseases.

References

Aas JA, Paster BJ, Stokes LN, Olsen I, Dewhirst FE (2005) Defining the normal bacterial flora of the oral cavity. J Clin Microbiol 43:5721–5732

Altshuler D, Daly MJ, Lander ES (2008) Genetic mapping in human disease. Science 322:881–888

Avila M, Ojcius DM, Yilmaz O (2009) The oral microbiota: living with a permanent guest. DNA Cell Biol 28:405–411

Barrett JC et al (2008) Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat Genet 40:955–962

Bockmuhl D et al (2007) Prebiotic cosmetics: an alternative to antibacterial products. Int J Cosmet Sci 29:63–64

Burkhart CG, Burkhart CN, Lehmann PF (1999) Acne: a review of immunologic and microbiologic factors. Postgrad Med J 75:328–331
Bradshaw CS et al (2006) The association of *Atopobium vaginae* and *Gardnerella vaginalis* with bacterial vaginosis and recurrence after oral metronidazole therapy. J Infect Dis 194:828–836
Fitzpatrick LR et al (2007) Effects of the probiotic formulation VSL#3 on colitis in weanling rats. J Pediatr Gastroenterol Nutr 44:561–570
Frank DN et al (2007) Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. Proc Natl Acad Sci USA 104:13780–13785
Frazer KA, Murray SS, Schork NJ, Topol EJ (2009) Human genetic variation and its contribution to complex traits. Nat Rev Genet 10:241–251
Gao Z, Tseng CH, Strober BE, Pei Z, Blaser MJ (2008) Substantial alterations of the cutaneous bacterial biota in psoriatic lesions. PLoS One 3:e2719
Gehrs KM, Anderson DH, Johnson LV, Hageman GS (2006) Age-related macular degeneration – emerging pathogenetic and therapeutic concepts. Ann Med 38:450–471
Gomez-Alvarez V, Teal TK, Schmidt TM (2009) Systematic artifacts in metagenomes from complex microbial communities. ISME J 3:1314–1317
Gophna U, Sommerfeld K, Gophna S, Doolittle WF, Veldhuyzen van Zanten SJ (2006) Differences between tissue-associated intestinal microfloras of patients with Crohn's disease and ulcerative colitis. J Clin Microbiol 44:4136–4141
Grice EA et al (2008) A diversity profile of the human skin microbiota. Genome Res 18:1043–1050
Grice EA et al (2009) Topographical and temporal diversity of the human skin microbiome. Science 324:1190–1192
Huber JA et al (2007) Microbial population structures in the deep marine biosphere. Science 318:97–100
Hunt KA et al (2008) Newly identified genetic risk variants for celiac disease related to the immune response. Nat Genet 40:395–402
Hyman RW et al (2005) Microbes on the human vaginal epithelium. Proc Natl Acad Sci USA 102:7952–7957
Lander ES et al (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921
Lee J, Seto D, Bielory L (2008) Meta-analysis of clinical trials of probiotics for prevention and treatment of pediatric atopic dermatitis. J Allergy Clin Immunol 121:116–121 e111
Leeming JP, Holland KT, Cunliffe WJ (1985) The pathological and ecological significance of microorganisms colonizing acne vulgaris comedones. J Med Microbiol 20:11–16
Lepp PW et al (2004) Methanogenic Archaea and human periodontal disease. Proc Natl Acad Sci USA 101:6176–6181
Liolios K, Mavromatis K, Tavernarakis N, Kyrpides NC (2008) The Genomes On Line Database (GOLD) in 2007: status of genomic and metagenomic projects and their associated metadata. Nucleic Acids Res 36:D475–479
Liu J et al (2002) Genomewide linkage analysis of celiac disease in Finnish families. Am J Hum Genet 70:51–59
Li Y, Ge Y, Saxena D, Caufield PW (2007) Genetic profiling of the oral microbiota associated with severe early-childhood caries. J Clin Microbiol 45:81–87
Maher B (2008) Personal genomes: the case of the missing heritability. Nature 456:18–21
Manolio TA et al (2009) Finding the missing heritability of complex diseases. Nature 461:747–753
Mastromarino P et al (2009) Effectiveness of lactobacillus-containing vaginal tablets in the treatment of symptomatic bacterial vaginosis. Clin Microbiol Infect 15:67–74
Metchnikov I (1908) The prolongation of life: optimistic studies. G.P. Putnam's Sons, New York & London
Moore WE, Moore LH (1995) Intestinal floras of populations that have a high risk of colon cancer. Appl Environ Microbiol 61:3202–3207
Nadal I, Donat E, Ribes-Koninckx C, Calabuig M, Sanz Y (2007) Imbalance in the composition of the duodenal microbiota of children with coeliac disease. J Med Microbiol 56: 1669–1674

Nair RP et al (2009) Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways. Nat Genet 41:199–204
Patterson BK et al (1998) Repertoire of chemokine receptor expression in the female genital tract: implications for human immunodeficiency virus transmission. Am J Pathol 153:481–490
Peterson DA, Frank DN, Pace NR, Gordon JI (2008) Metagenomic approaches for defining the pathogenesis of inflammatory bowel diseases. Cell Host Microbe 3:417–427
Preidis GA, Versalovic J (2009) Targeting the human microbiome with antibiotics, probiotics, and prebiotics: gastroenterology enters the metagenomics era. Gastroenterology 136:2015–2031
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
Schaefer AS et al (2010) A genome-wide association study identifies GLT6D1 as a susceptibility locus for periodontitis. Hum Mol Genet 19:553–562
Setoyama H, Imaoka A, Ishikawa H, Umesaki Y (2003) Prevention of gut inflammation by bifidobacterium in dextran sulfate-treated gnotobiotic mice associated with bacteroides strains isolated from ulcerative colitis patients. Microbes Infect 5:115–122
Siepel A et al (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res 15:1034–1050
Sipos R et al (2007) Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targetting bacterial community analysis. FEMS Microbiol Ecol 60:341–350
Srinivasan S, Fredricks DN (2008) The human vaginal bacterial biota and bacterial vaginosis. Interdiscip Perspect Infect Dis 2008:750479
Strober W, Fuss I, Mannon P (2007) The fundamental basis of inflammatory bowel disease. J Clin Invest 117:514–521
Turnbaugh PJ et al (2009) A core gut microbiome in obese and lean twins. Nature 457:480–484
Turnbaugh PJ, Backhed F, Fulton L, Gordon JI (2008) Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. Cell Host Microbe 3:213–223
van Heel DA et al (2007) A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. Nat Genet 39:827–829
Venter JC et al (2001) The sequence of the human genome. Science 291:1304–1351
Walley AJ, Asher JE, Froguel P (2009) The genetic contribution to non-syndromic human obesity. Nat Rev Genet 10:431–442
Wood LD et al (2007) The genomic landscapes of human breast and colorectal cancers. Science 318:1108–1113
Zhou X et al (2007) Differences in the composition of vaginal microbial communities found in healthy Caucasian and black women. ISME J 1:121–133
Zyrek AA et al (2007) Molecular mechanisms underlying the probiotic effects of *Escherichia coli* Nissle 1917 involve ZO-2 and PKCzeta redistribution resulting in tight junction and epithelial barrier repair. Cell Microbiol 9:804–816