CHAPTER 1

Introduction and Background to Microbiome Research

JOSEPH F. PIERRE*

Department of Nutritional Sciences, College of Agriculture and Life Science, University of Wisconsin–Madison, Wisconsin, USA
*E-mail: jpierre@wisc.edu

1.1 Introduction to the Gut Microbiota

Microbial communities colonize the gut and virtually all other body compartments, including the skin, mammary ducts, respiratory tract, and, as recent evidence supports, even the circulation. These dynamic communities are fundamentally involved in homeostasis and disease progression under environmentally and genetically shaped susceptible conditions. Through their assistance with digestion and fermentation, stimulation of nutrient absorption and endocrine regulation, production of vitamins, priming and agonism of immune education and response, and production of many small molecules, a growing number of relevant host–microbial interactions have been uncovered rapidly by microbiome researchers in recent years.[1] Despite the enormous effect microbiome research has already had on our understanding of health and many diseases, this field is still young and continues to evolve in both the amount of empirical data collected annually and through the continuous improvements in the methodological approaches and techniques used to explore the microbiome.

Scientists had historically long attempted to investigate microorganisms, but for hundreds of years the investigation into microbial life forms

Food Chemistry, Function and Analysis No. 34
Metabolism of Nutrients by Gut Microbiota
Edited by Joseph F. Pierre

Published by the Royal Society of Chemistry, www.rsc.org

was limited by what could be observed under the microscope or by indiscriminately cultivating organisms in artificial broths and solid substrates. While historical descriptions over several centuries more or less accurately described major microbial kingdoms – bacteria, fungi, and protists – the work of Louis Pasteur most notably first tied the importance of microbes to human health with his experimentation and theory of germs. Pasteur posited that certain microbes could cause specific diseases when introduced, setting into motion an unravelling of the superstition surrounding the origins of disease and improved understanding of microbial life. This new framework of empirical testing of microbes led Robert Koch to demonstrate that a single bacterium – bacilli – could cause anthrax in animals.[2] Koch subsequently developed criteria, termed Koch's postulates, for the empirical determination for the role of bacteria in disease aetiology. These postulates included the necessity to isolate suspected bacteria in all cases of a specific disease (but not from healthy individuals), successfully initiation of disease when that isolated microbe was introduced to a healthy host, and subsequent resolution of disease when that suspected microbe was eliminated.

It is now apparent that the majority of microbes that live within us are not readily grown in traditional laboratory media, instead requiring unique energetic substrates, the presence or absence of specific atmospheric gases, or specific mutualistic or even parasitic interactions with other unique members of the gut community and their metabolites to thrive. These initial challenges led to the rise of culture independent methods, including next-generation sequencing, that allowed full ecological characterization of isolated microbial DNA. Much of this early research has been around metabolism and nutrition.[3–5] In the early 2000s, these techniques first led to the rapid growth in understanding of the complex microbial communities in the gut, where the number of microbial cells is equal to the number of human cells in the human body, but with substantial genetic diversity at 100 to 150 times our own.[3,6] This diversity is encapsulated in enormous numbers of bacteria, reaching up to 10^{12} bacteria per gram in the distal intestine.[7] The enormous genetic capacity harboured by gut microbes has led to the concept that the microbiome is a virtual mammalian organ, being shaped by other host homeostatic and immune systems, and one that can subsequently be transplanted in composition and function between hosts with variable degrees of success. Microbiome research over the past two decades has done much to shift the viewpoint of our microbes towards them being helpful participants in normal development, homeostasis, and nutrition, revising an outdated concept of microbes as simply harmful pathogens and disease-causing organisms.

1.2 Approaches in Microbiome Research

In order to determine community membership, along with functional characteristics of the microbial community, a number of technologies can be used either independently or in parallel as multi-omics-based platforms. These technologies enabled culture-independent (next-generation sequencing)

or culture-dependent approaches (anaerobic chambers, fermentation cultivar systems) to identify and isolate novel microbial strains that may contribute individually or as keystone community members of broader ecology in response to various perturbations, and in particular, nutrition. The following sections outline tools and techniques used in microbiome research to explore and elucidate the effects of and interplay between nutrition and dietary intake on gut microbes.

1.2.1 Fingerprinting the Microbial Community

1.2.1.1 Microbial Gene Amplicon Sequencing Techniques

The identification and use of highly conserved regions of prokaryotic rRNA found in all bacteria and fungi enabled amplification-based sequencing approaches. Amplicons are generally 150 to 250 nucleic acid base pairs in length. Studies examining humans and animals routinely rely on the 16S rRNA marker gene amplicon sequencing platforms, which include Sanger-based sequencing, Roches454, PacBio, IonTorrent, and Illumina MiSeq/HiSeq/NexSeq platforms.[8] Despite slightly different technological approaches, these platforms can each generate millions of short read sequences (amplicons) along with unique barcodes for identifying the source of each sequence against specific samples. These sequencing platforms were complemented by advanced computation approaches allowing analysis of the millions of sequences generated, including through mothur,[9] Quantitative Insights Into Microbial Ecology (QIIME1 and 2),[10] and Minimum Entropy Decomposition (MED).[11] Each technique enables insight into microbial community composition, community diversity, and numerous methods to determine relatedness and unique signatures of microbial communities. Dependence on 16S rRNA amplicon sequencing inherently targets bacteria, which are the predominant colonizers, making up roughly 99% of microbial cell numbers, but are not the only kingdom of microorganisms, fungi and yeasts, viruses, and protists are excluded.

Advancements of rRNA amplicon sequencing approaches have recently been developed to target other kingdoms and domains, including yeasts, and are more commonly being used to study human and animal health and disease. The importance of these communities remains debated, as compared with bacterial communities that remain relatively stable within individuals,[12] the fungal membership changes more considerably between timepoints, with only 20% of species found consistently through temporal sampling.[13] These analyses are complicated by the number of environmental and dietary ingested yeast and fungal species that may be transient and limit detection of true colonizers in individuals.

To address these challenges, additional rRNA regions, such as the 18S rRNA subunit and the internal transcribed spacer 1 and 2 (ITS1 and ITS2) are used for phylogenetic assignment of eukaryotic microorganisms, specifically yeast and fungi.[14] Another challenge in eukaryotic microbial research is obtaining

well developed and curated databases for assigning reads generated from next generation platforms. While bacterial and archaeal databases, including SILVA and GreenGenes, are fairly well established, databases specific to yeasts and fungi are still under relatively recent but rapid curation and validation, including UNITE and Targeted Host-associated Fungi (THF). Additionally, since sequencing of complex eukaryotic communities is a relatively new technique, many of the detected organisms have not been isolated or cultured from the host, leading to challenges in accurate classification and questions about their origins from the gut, diet, and environment. As a counter example of the importance of yeast in human health and disease, the members of the well described genus of *Candida* have no known reservoir outside of the mammalian gut and are considered true mammalian gut residents.[15] The members of this genus of fungi are also well established opportunistic pathogens and are especially problematic in immune compromised individuals. However, the well-studied yeast genus *Saccharomyces*, members of which are used in fermentation of food products, remain questionable as core community residents, as their detection in the gut may be the result of dietary intake.[16] Functional understanding of fungal organisms has given some insight into the fidelity of eukaryotic colonization. For example, the yeast *Malassezia* have lost the ability to synthesize lipids and therefore require the host for their metabolic substrate, rendering them likely to be true mammalian colonizers. Other fungi, including *Debaryomyces* and *Penicillium*, although commonly found in the gut, cannot readily replicate at mammalian body temperature and their detection is, to date, considered environmental contamination.[16] Importantly, host dietary intake based on protein rich or carbohydrate rich diets has been demonstrated to shift the interkingdom dynamics between bacteria and fungi.[17] Accordingly, greater resolution of bacterial and other rare microbial members and their role in gut ecological dynamics is needed to understand the interactions between prokaryotes and eukaryotes and their synergistic function in influencing host homeostasis. These advances will probably come through synergy in deeper sequencing, further database curation, and increased sequencing coverage that will enable a more comprehensive snapshot of the complex inter-kingdom colonizing populations.

1.2.1.2 Transfer RNA (tRNA) Sequencing (seq) Techniques

Beyond 16S rRNA amplicon sequencing for microbial community fingerprinting, recent advances in microbially-derived transfer RNAs (tRNAs) that facilitate translation of messenger RNA protein are also used to distinguish microbial communities with high accuracy. This method was originally developed for isolated microbial cultures but was then coupled with enzyme treatment to analyse demethylated tRNA, allowing insights into both tRNA transcripts and tRNA post-translational modifications.[18] With similar computational approaches to those for 16S rRNA, microbial tRNA reference libraries can produce highly accurate phylogenetic community analysis. In addition to taxonomic information, data on protein expression provides an

opportunity to obtain more advanced functional insights, especially in the context of investigating the role of nutrition and dietary interactions with microbial functions.

1.2.2 Identification and Classification of Microbial Membership and Their Functions

1.2.2.1 Sequence-based Approaches

1.2.2.1.1 Shotgun Metagenomics. Large and rapid amounts of descriptive data can be obtained from amplicon sequencing-based techniques, however little functional information can be generated from taxonomic description alone. Some efforts to combine taxonomic descriptions with known metagenomes of functional genes has been made, for instance through development of the Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt), which is used to predict Kyoto Encyclopaedia of Genes and Genomes (KEGG) ortholog functions and metabolic pathways harboured by microbial communities based on 16S rRNA marker gene amplicon sequences.[19] However, as databases are incomplete and variably updated, new and rare members of the community of curated genome databases may be missing or underrepresented and these techniques are limited in capturing the true gene functions and metabolic pathways within microbial communities.[20]

The limitations of amplicon-based tools (16S, 18S, ITS rRNA) can be more readily overcome through the use of high-throughput shotgun metagenomic sequencing since this method provides untargeted collection of all genetic content isolated in a microbial sample. These approaches are also valuable because they include genetic capture of fungal, virus and bacteriophage, in addition to bacterial, genomes. To date, two general approaches are used: mapping to reference databases or *de novo* assembly of sequenced reads. For sequence mapping purposes, on-line servers are available, including Metagenomic Rapid Annotations using Subsystems Technology (MG-RAST)[21] and J. Craig Venter Institute (JCVI) Metagenomics Reports (METAREP).[22] While computer based programs, such as Human Microbiome Project Unified Metabolic Analysis Network (HUMAnN),[23] are utilized locally. *De novo* assembly is performed after functional annotations are performed, such as with platforms to reveal metagenome assembled genomes (MAGs). Further, several assembly programs are available, including khmer,[24] and visualization tools have been developed for these analyses, including analysis and visualization platform for 'omics data (Anvi'o).[25] Finally, strategies have been developed to complement metagenomic data with 16S sequencing, including ribosomal flanking region-sequencing (RiboFR-Seq), which provides 16S variable region information as well as the immediate protein-coding genes surround the 16S gene.[26] These advanced and functionally informed strategies are enabling greater insights into metabolic capacity by eliminating reliance on 16S based inferences as a stand-alone.

1.2.2.1.2 Single-cell Genomics Coupled with Next-Generation Sequencing Approaches. In addition to collective microbial community DNA, much greater functional insights have been made using single-cell sequencing. These strategies are especially useful for understanding the metabolic and functional capacity of rare or low abundance microbes within samples. This strategy is more difficult to employ as it includes initial isolation or enrichment of the microbe of interest, using flow cytometry or other antibody-based identification and enrichment strategies. Following isolation, microbe identity can be confirmed by 16S rRNA amplification and whole genome-based sequencing, with the computation analyses strategies described above. For greater review of the single-cell genomic isolation and sequencing techniques, see the technical review by Qin *et al.*[27]

1.2.2.1.3 Metatranscriptomics. Metagenomic shotgun sequencing and single-cell genomics used separately or in combination, can aid in identifying gene content and function of gut microbiota communities, yet the activity or abundance of microbial gene expression cannot be discerned from genomic DNA-based approaches alone. Metatranscriptomic shotgun sequencing (RNAseq) in combination with metagenomics is one strategy employed to identify the genomic potential as well as the active microbial genomes. This is carried out by isolating total RNA from the microbial community followed by enrichment for RNA [mRNA, long intergenic non-coding RNA (lincRNA), and microRNA] and fragmentation. RNA is then converted to complementary DNA *via* reverse transcriptase with oligo(dT) primers and/or random hexamers and constructed libraries can then be sequenced.[28] Despite providing insight into the activity of the microbial whole-genome, this technology can be hampered by technical issues that, to date, limit its effectiveness. For instance, integrity of gut microbial RNA can be compromised by sample collection and storage where RNA quality is compromised, leading to insufficient yields of high-quality microbial RNA, limiting purification efficiency and sequence fidelity. Furthermore, remnant RNA preservation solutions can interfere with downstream library preparation, biasing sequencing results. However, if these limitations are overcome, downstream data analysis for metatranscriptomic sequencing data then rely on similar strategies to shotgun metagenomic analysis, including Anvi'o and HUMAnN,[23,25,28] which allows for taxonomic assignment coupled with complementary identification of actively expressed gene functions and greater insight into microbial functions within a given community and environment.

1.2.2.1.4 Long Read Sequencing Technology. In contrast to amplifying restricted regions of the bacterial 16S or fungal ITS gene, other technologies have been optimized to amplify longer regions of microbial genomes. These platforms provide greater depth resolution of community membership and metabolism, but the current lack of functional annotation can be problematic. One way to overcome annotation gaps is to improve coverage and computationally perform assembly, which benefits from longer sequencing read

lengths (>10 000 bp). Long read technologies include the Oxford Nanopore MinION,[29] the Pacific BioSciences single molecule real time (SMRT) sequencer,[30] and Illumina Moleculo.[31] In depth reviews of these approaches benefits, limitations, and pitfalls are described elsewhere.[32,33]

1.2.2.2 Non Sequence-based Approaches to Identify Microbial Functions

In addition to genome and transcript analysis, greater emphasis is now being placed on the functional outputs of microbial metabolism, specifically through the quantitation of microbial modified metabolites. Microbial metabolites are abundant in the gut lumen but also throughout the systemic circulation and in host organs, where they are assessed by targeted or untargeted metabolomic, proteomic, and lipidomic approaches. These analyses are providing important insights into microbial community function and roles in mammalian biology.

Contributions of microbial metabolites are collectively part of the host's global metabonome. Untargeted and targeted methodologies are used to detect host and microbial inorganic and organic metabonomes through the use of gas chromatography–mass spectrometry (GC-MS), liquid chromatography–mass spectrometry (LC-MS), and nuclear magnetic resonance spectrometry (NMR), for assessment of organic lipids, amino acids, and simple and complex carbohydrates. Furthermore, introduction of isotope labelled dietary nutrients, especially those not digestible by host enzymes (such as C^{13} labelled inulin) enable investigators to delineate microbially produced metabolites generated from dietary components from those metabolites produced by the host.

Another approach to understanding the role of microbial generated metabolites is through the employment of germ-free animals (described in greater detail below) compared with conventional counterparts. Since germ-free animals lack any microbial colonizers, comparisons can be made of systemic and secreted metabolite profiles in the presence and absence of microbes that allow characterization of microbial generated or influenced metabolites compared with those generated by the host alone. These studies have generated evidence of unique metabolites found in the major organs, such as the kidney, heart, and brain.[34–38] In addition to lipids, amino acids, and carbohydrates, many other vastly diverse small molecules and peptides are detected, with their effects on the host remaining largely unknown.[39]

Despite the classical limitations of culturing microbes in the laboratory, another approach to understand the role of microbial metabolites is culturomics-based analysis, which complements metabolomic quantification and 16S sequencing on isolated and cultured microorganisms.[40,41] These techniques have led two databases, including the Culturable Genome Reference (CGR), which contains over 6000 indigenous isolates.[42] One strength of this type of approach is the ability to identify toxins or harmful metabolites generated by microbial isolates, especially compared with highly similar

microbes based on taxonomic assignment. For instance, functional differences have been observed in strains of *Clostridium butyricum* isolated from neonates with necrotizing enterocolitis (NEC) compared with controls,[43] indicating that the taxonomic presence of this bacteria is insufficient for disease without production of their toxins. By identifying microbes from diseases and environments, the complex metabolite signatures can be assessed by metabolomic quantification and comparison with databases, including MetaCyc,[44] Human Metabolome Database (HMDB),[45] SetupX and BinBase.[46]

1.3 *In Vivo* Models for Investigating Microbial Causality in Nutrition and Metabolism

Humans are inseparably associated with their microbiomes – often beginning at or just before birth (see Chapter 7) – and determining causal effects of microbes on the host is not directly possible. To overcome this limitation, experimental models have been developed that allow the study of microbes in microbially naive hosts, most notably through the development of germ-free animal colonies. The most common germ-free animals currently used in medical research are mice, due to their size, cost, short life cycles, similar organ and immune functions to those of humans, and the large number of genetic mutants available. Other germ-free animals include rats, guinea pigs, swine, poultry, zebrafish, fruit flies, and nematodes. The goal of germ-free animal research is to begin with an animal devoid of any detectable microbes, including bacteria, archaea, fungi, protists, bacteriophages, or viruses.

The first successful germ-free animal studies were performed by Nuttall and Thierfelder in the early 1900s, where guinea pigs were maintained under sterile conditions for two weeks, providing the first evidence that complex mammals could survive in the absence of microorganisms.[47] James Reyniers established the first sustainable germ-free mouse colony in 1931 at the University of Notre Dame. His approach was to perform caesarean section on timed pregnant mice, transfer pups into sterile isolators, and rear animals for multiple generations, an approach that has remained almost unchanged to this day. Interest in the role of microbes in health and disease has continued to grow over the past century, especially following the widespread acceptance and use of antibiotics.

Germ-free mice are especially useful as they can be moved into dedicated sterile experimental isolators or positive pressure cages where individual microbes (monocolonization) or complex communities of microbes are introduced or transplanted (termed conventionalization or transfaunation), leading to an animal with a controlled microbiome membership (gnotobiotics; known life).[48] All cages, bedding, food and water, and other supplies are irradiated or more commonly autoclaved before introduction to the isolators or cage. Animals can be administered microbes before or at birth (by colonizing the pregnant mother) or at any point in the life cycle. Controlling the timing and composition of microbes as well as the hosts genetics,

diet, and environment, is the most robust way to causally determine the role of microbes in mammalian health, immune function, metabolism, and disease risk.

Following the advent of next generation sequencing, the utility of germ-free mice and piglets has again become popular, largely because of organ, immune, and metabolic similarities between mice, pigs, and humans depending on the focus of the study.[49,50] Comparison of host metatranscriptomes of germ-free *vs.* conventional piglets revealed that almost 70% of the transcriptome in gut and systemic organs, especially immune specific genes, are influenced by gut and mucosal associated microorganisms.[51] It should be noted that certain limitations exist in germ-free animals, as microbes and their metabolites are paramount to normal neurological, immune and physiological development, so microbial colonization, especially at later timepoints in the animals lifecycle are not fully translational to the effects of these microbes in conventionally housed and microbially competent animals. Despite these limitations, insights into the nutritional requirements of animals, contribution of microbial colonizers to host metabolism, and effects on growth rates have been gleaned from germ-free animals when compared with conventionally reared control animals.[48]

In addition to mice and pigs, other vertebrate germ-free animals have been developed, including zebrafish.[52] Fish are useful due to their extremely short life cycle, low cost of housing per animal, modifiable genetics, and transparent organs through development. In regard to nutrition, lipid digestion and absorption has been investigated in great detail using zebrafish.[53] Like all animal models, fish also have limitations in the translational relevance to humans, specifically differences in body temperature, organ arrangement and function, and vastly different microbial members that thrive in marine environments. Invertebrate germ-free animal models include *Drosophila melanogaster*[54] and *Caenorhabditis elegans*.[55] Collectively, the use of germ-free animals has contributed valuable insights into host–microbial interactions and has much to offer in advancing our understanding of host–microbial interactions in the context of nutrition in the years to come.

1.4 *In Vitro* Models to Study the Microbiome

1.4.1 Microbial Culture-based Technologies to Study Host–Microbe Interactions

In vitro model systems have been developed to further our understanding of host–microbial interactions. These simplified experimental models allow careful control of the microenvironment including nutrient composition, microbial membership, metabolite formation and release, partial gas pressures, and pH changes. One example is chemostats, which range from simple to complex equipment meant to mimic the various regions the gastrointestinal tract and which are used to explore the role of microbes and their community dynamics in response to available nutrients.[56] Perhaps the most

technologically advanced chemostat, the twin simulator of the human intestinal microbial ecosystem (TWINSHIME®) enables investigators to study both the mucosal and luminal microbial populations as available nutrients move throughout the modeled gastrointestinal tract, resulting in a final stool output.[57] Chemostats have enabled detailed understanding of microbial dietary utilization and metabolism.

1.4.2 Cell Culture Model Systems to Study Host–Microbe Interactions

Over the last decade, advances in mammalian cell culture have been made with the development of organoids, which can be generated from intestinal, hepatic, and neuronal precursor cells. Within the gut, pluripotent intestinal stem cells (piSCs) are isolated from animals or humans *via* biopsy or whole tissue sections and grown in culture to form intestinal 'miniguts'. piSCs obtained from each region of the intestine recapitulate the epithelial biology of their source, including the differentiation patterns, abundances, and functions of epithelial cells for each region of the gut, including the stomach (gastroids), small intestine (enteroids) and large intestine (colonoids).[58–60] Coculturing intestinal organoids with immune or isolated neurons (spheroids) and microbes has enabled the modeling of complex cell to cell interactions for focused investigation.

While many organoids are grown as three-dimensional structures within a collagen structure, this approach limits the study of the luminal microbes. This limitation has been overcome in-part through the use of microinjectors, which transplant small volumes of microorganisms or solutions into the basolateral organoid compartment. However, an alternative approach is to grow organoids on a flat collagen basement membrane, where they form two-dimensional monolayers. Several laboratories have used this approach to generate 'gut-on-a-chip' tools, where careful control of media through a microfluidic chamber allows the study of epithelial cell growth, dynamics, and responses to luminal stimulus.[61] The development of organoids has allowed unparalleled advancement into the study of intestinal disease, the regulation of epithelial signals that maintain gut homeostasis, microbial–host interactions, and nutrition.

1.5 Heterogeneity of Mammalian Gut Microbes – Implications for Nutritional Science

As described above, characterization of microbial community composition alone has limited functional value. This limitation is further exacerbated by the fact that humans (and other complex animals) display a large heterogeneity in their microbiome communities. For instance, two individuals with similar health status share only 10 to 30% of gut bacteria with one another. This variability it is driven by diet and environmental exposure, but also

genetics as genetically identical twins do harbor more microbial similarity than dizygotic twins or other family members when microbial membership is compared.[62–65] A deeper assessment of microbial community metatranscriptomes has revealed that even fairly dissimilar metagenome communities often contain similar functional core features, such as carbohydrate fermentation, or other pathways for surviving in the mammalian gut. Therefore, a certain degree of microbial community heterogeneity is explained by the functional redundancy in the community that is preserved rather than specific taxonomic membership of a given community. Despite the noted variability, the composition of the microbiome has been linked to diseases in human populations, such as obesity, inflammatory bowel disease (IBD), and numerous other metabolic, immunological, and neurological disorders.[64–66] The most advanced survey of the human microbiome was recently published with the second phase of the human microbiome project (HMP2), which investigated the microbiome in IBD, premature birth, and type-2 diabetes using multi-omics strategies (metagenomics, metabolomics, metatranscriptomics).[67] The results of HMP2 demonstrate that despite variability in the taxonomic community, functional differences exist under each of these common human disease states.

Given the large degree of variability in the gut microbiome, investigators need to be aware of limitations when planning, designing, and carrying out nutritional science research in humans and animal models. Specific to experimental animals, it is known that the sources of commercially available research animals contain disparate microbiome compositions that can alter the physiological outcomes of research studies.[68,69] For example, a classical example was that mice from Taconic Farms contain segmented filamentous bacteria (SFB), while mice from the Jackson (JAX) laboratory do not. Comparison of mice between these vendors demonstrated that SFB were strong drivers of T helper (Th) 17 cluster of differentiation 4 positive (CD4+) T cells[70] and disease progression was initiated by the presence of these bacteria. In addition, each individual animal facility, different rooms within the same facilities, and even individual cages within a given room can contain variable microbiomes.[71,72] Subtle differences in the gut microbiota of research animals can alter the immunological and metabolic outcomes of animals, or lead to different microbiome communities and functions on different diets. One approach to these challenges has been to routinely mix cage bedding during experiments to normalize a core microbiome, and bank that bedding for future experimental replicates.[73] Furthermore, the source of diets, ventilation systems, and even the investigators handling the animals can have environmental effects on gut microbiomes. These limitations can be overcome through thoughtful experimental design, frequent sterilization of surfaces and gloves used to handle animals, and the use of biological safety cabinets to prevent novel microbial introduction. Finally, appropriately powered and careful data analysis can be employed to successfully carry out studies of host–microbial interactions in nutritional science research.

1.6 Summary

The microbiome is now appreciated to influence host homeostasis and metabolism. While many techniques, approaches, and experimental models are available to examine microbial interactions in the setting of nutritional science, tools and methodologies used to explore microbial functional effects on the host continue to evolve. In the setting of nutrition, the timing and composition of diet is one of the most profound influencers of microbial community structure and function.[74] Ingested nutrients form the energetic substrates that are utilized by the host and microbes alike, directly or indirectly. In addition, microbes ferment substrates and generate *de novo* nutrients otherwise unavailable to the host. In return, the host generates secreted metabolites that help stabilize the microbial population within the gut. The role of dietary composition therefore fundamentally orchestrates the host–microbial ecology in meaningful ways. The chapters that follow here will focus on the role of major dietary macronutrient intake, interactions between microbes and drug metabolism, the state of prebiotics, probiotics, and synbiotics, the role of microbes in bariatric surgically induced weight loss, and how initial microbial colonizers following birth shape lifelong chronic disease risk and metabolism.

References

1. I. Montoliu, U. Genick, M. Ledda, S. Collino, F.-P. Martin, J. le Coutre and S. Rezzi, *Genes Nutr.*, 2013, **8**, 19–27.
2. L. S. King, *J. Hist. Med. Allied Sci.*, 1952, **VII**, 350–361.
3. R. E. Ley, D. A. Peterson and J. I. Gordon, *Cell*, 2006, **124**, 837–848.
4. B. L. Cantarel, V. Lombard and B. Henrissat, *PLoS One*, 2012, **7**, e28742.
5. P. J. Turnbaugh, R. E. Ley, M. A. Mahowald, V. Magrini, E. R. Mardis and J. I. Gordon, *Nature*, 2006, **444**, 1027–1131.
6. R. Sender, S. Fuchs and R. Milo, *PLoS Biol.*, 2016, **14**, e1002533.
7. E. T. Hillman, H. Lu, T. Yao and C. H. Nakatsu, *Microbes Environ.*, 2017, **32**, 300–313.
8. J. G. Caporaso, C. L. Lauber, W. A. Walters, D. Berg-Lyons, C. A. Lozupone, P. J. Turnbaugh, N. Fierer and R. Knight, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108** Suppl 1, 4516–4522.
9. P. D. Schloss, S. L. Westcott, T. Ryabin, J. R. Hall, M. Hartmann, E. B. Hollister, R. A. Lesniewski, B. B. Oakley, D. H. Parks, C. J. Robinson, J. W. Sahl, B. Stres, G. G. Thallinger, D. J. Van Horn and C. F. Weber, *Appl. Environ. Microbiol.*, 2009, **75**, 7537–7541.
10. J. G. Caporaso, J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Peña, J. K. Goodrich, J. I. Gordon, G. A. Huttley, S. T. Kelley, D. Knights, J. E. Koenig, R. E. Ley, C. A. Lozupone, D. McDonald, B. D. Muegge, M. Pirrung, J. Reeder, J. R. Sevinsky, P. J. Turnbaugh, W. A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld and R. Knight, *Nat. Methods*, 2010, **7**, 335–336.

11. A. M. Eren, H. G. Morrison, P. J. Lescault, J. Reveillaud, J. H. Vineis and M. L. Sogin, *ISME J.*, 2014, **9**, 968–979.
12. C. A. Lozupone, J. I. Stombaugh, J. I. Gordon, J. K. Jansson and R. Knight, *Nature*, 2012, **489**, 220–230.
13. H. E. Hallen-Adams, S. D. Kachman, J. Kim, R. M. Legge and I. Martínez, *Fungal Ecol.*, 2015, **15**, 9–17.
14. E. A. Witherden, S. Shoaie, R. A. Hall and D. L. Moyes, *J. Fungi*, 2017, 56.
15. T. A. Auchtung, T. Y. Fofanova, C. J. Stewart, A. K. Nash, M. C. Wong, J. R. Gesell, J. M. Auchtung, N. J. Ajami and J. F. Petrosino, *mSphere*, 2018, e00092-18.
16. R. A. Hall and M. C. Noverr, *Curr. Opin. Microbiol.*, 2017, **40**, 58–64.
17. C. Hoffmann, S. Dollive, S. Grunberg, J. Chen, H. Li, G. D. Wu, J. D. Lewis and F. D. Bushman, *PLoS One*, 2013, **8**, e66019.
18. M. H. Schwartz, H. Wang, J. N. Pan, W. C. Clark, S. Cui, M. J. Eckwahl, D. W. Pan, M. Parisien, S. M. Owens, B. L. Cheng, K. Martinez, J. Xu, E. B. Chang, T. Pan and A. M. Eren, *Nat. Commun.*, 2018, **9**, 5353.
19. M. G. I. Langille, J. Zaneveld, J. G. Caporaso, D. McDonald, D. Knights, J. A. Reyes, J. C. Clemente, D. E. Burkepile, R. L. Vega Thurber, R. Knight, R. G. Beiko and C. Huttenhower, *Nat. Biotechnol.*, 2013, **31**, 814–821.
20. T. J. Sharpton, *Front. Plant Sci.*, 2014, **5**, 209.
21. F. Meyer, D. Paarmann, M. D'Souza, R. Olson, E. Glass, M. Kubal, T. Paczian, A. Rodriguez, R. Stevens, A. Wilke, J. Wilkening and R. Edwards, *BMC Bioinf.*, 2008, **9**, 386.
22. J. Goll, D. B. Rusch, D. M. Tanenbaum, M. Thiagarajan, K. Li, B. A. Methé and S. Yooseph, *Bioinformatics*, 2010, **26**, 2631–2632.
23. S. Abubucker, N. Segata, J. Goll, A. M. Schubert, J. Izard, B. L. Cantarel, B. Rodriguez-Mueller, J. Zucker, M. Thiagarajan, B. Henrissat, O. White, S. T. Kelley, B. Methé, P. D. Schloss, D. Gevers, M. Mitreva and C. Huttenhower, *PLoS Comput. Biol.*, 2012, **8**, e1002358.
24. M. R. Crusoe, H. F. Alameldin, S. Awad, E. Boucher, A. Caldwell, R. Cartwright, A. Charbonneau, B. Constantinides, G. Edvenson, S. Fay, J. Fenton, T. Fenzl, J. Fish, L. Garcia-Gutierrez, P. Garland, J. Gluck, I. González, S. Guermond, J. Guo, A. Gupta, J. R. Herr, A. Howe, A. Hyer, A. Härpfer, L. Irber, R. Kidd, D. Lin, J. Lippi, T. Mansour, P. McA'Nulty, E. McDonald, J. Mizzi, K. D. Murray, J. R. Nahum, K. Nanlohy, A. J. Nederbragt, H. Ortiz-Zuazaga, J. Ory, J. Pell, C. Pepe-Ranney, Z. N. Russ, E. Schwarz, C. Scott, J. Seaman, S. Sievert, J. Simpson, C. T. Skennerton, J. Spencer, R. Srinivasan, D. Standage, J. A. Stapleton, S. R. Steinman, J. Stein, B. Taylor, W. Trimble, H. L. Wiencko, M. Wright, B. Wyss, Q. Zhang, E. Zyme and C. T. Brown, *F1000Research*, 2015, 900.
25. A. M. Eren, Ö. C. Esen, C. Quince, J. H. Vineis, H. G. Morrison, M. L. Sogin and T. O. Delmont, *PeerJ*, 2015, **3**, e1319.
26. Y. Zhang, P. Ji, J. Wang and F. Zhao, *Nucleic Acids Res.*, 2016, **44**, e99.
27. J. Qin, R. Li, J. Raes, M. Arumugam, K. S. Burgdorf, C. Manichanh, T. Nielsen, N. Pons, F. Levenez, T. Yamada, D. R. Mende, J. Li, J. Xu, S. Li, D. Li, J. Cao, B. Wang, H. Liang, H. Zheng, Y. Xie, J. Tap, P. Lepage,

M. Bertalan, J.-M. Batto, T. Hansen, D. Le Paslier, A. Linneberg, H. B. Nielsen, E. Pelletier, P. Renault, T. Sicheritz-Ponten, K. Turner, H. Zhu, C. Yu, S. Li, M. Jian, Y. Zhou, Y. Li, X. Zhang, S. Li, N. Qin, H. Yang, J. Wang, S. Brunak, J. Doré, F. Guarner, K. Kristiansen, O. Pedersen, J. Parkhill, J. Weissenbach, B. MetaHIT Consortium, P. Bork, S. D. Ehrlich and J. Wang, *Nature*, 2010, **464**, 59–65.

28. S. Bashiardes, G. Zilberman-Schapira and E. Elinav, *Bioinf. Biol. Insights*, 2016, **2016**, 19–25.
29. A. S. Mikheyev and M. M. Y. Tin, *Mol. Ecol. Resour.*, 2014, **14**, 1097–1102.
30. A. McCarthy, *Chem. Biol.*, 2010, **17**, 675–676.
31. R. C. McCoy, R. W. Taylor, T. A. Blauwkamp, J. L. Kelley, M. Kertesz, D. Pushkarev, D. A. Petrov and A.-S. Fiston-Lavier, *PLoS One*, 2014, **9**, e106689.
32. S. Koren and A. M. Phillippy, *Curr. Opin. Microbiol.*, 2015, **23**, 110–120.
33. M. O. Pollard, D. Gurdasani, A. J. Mentzer, T. Porter and M. S. Sandhu, *Hum. Mol. Genet.*, 2018, **27**, R234–R241.
34. F.-P. J. Martin, M.-E. Dumas, Y. Wang, C. Legido-Quigley, I. K. S. Yap, H. Tang, S. Zirah, G. M. Murphy, O. Cloarec, J. C. Lindon, N. Sprenger, L. B. Fay, S. Kochhar, P. van Bladeren, E. Holmes and J. K. Nicholson, *Mol. Syst. Biol.*, 2007, 112.
35. S. P. Claus, T. M. Tsang, Y. Wang, O. Cloarec, E. Skordi, F.-P. Martin, S. Rezzi, A. Ross, S. Kochhar, E. Holmes and J. K. Nicholson, *Mol. Syst. Biol.*, 2008, **4**, 219.
36. W. R. Wikoff, A. T. Anfora, J. Liu, P. G. Schultz, S. A. Lesley, E. C. Peters and G. Siuzdak, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 3698–3703.
37. M. Matsumoto, R. Kibe, T. Ooga, Y. Aiba, S. Kurihara, E. Sawaki, Y. Koga, Y. Benno, J. Qin, G. W. Tannock, P. V. Kirjavainen, T. Arvola, S. J. Salminen, E. Isolauri, J. Penders, E. E. Stobberingh, P. A. van den Brandt, C. Thijs, I. Sobhani, P. J. Turnbaugh, Y. Song, C. Liu, S. M. Finegold, H. A. Gordon, E. Bruckner-Kardoss, B. S. Wostmann, A. Suau, S. R. Gill, K. Kurokawa, S. P. Claus, S. P. Claus, P. B. Mortensen, M. R. Clausen, M. Matsumoto, K. Kakizoe, Y. Benno, M. Matsumoto, Y. Benno, L. M. Raamsdonk, Y. Ohashi, W. R. Wikoff, M. Li, F. P. Martin, F. P. Martin, B. Vitali, T. Soga, M. R. Monton, T. Soga, W. G. Bergen, G. Wu, C. C. Metges, T. Soga, T. Mitsuoka, K. Hayakawa, J. Neu, M. Matsumoto, S. Kurihara, R. Kibe, H. Ashida, Y. Benno, A. E. Pegg, P. P. McCann, G. M. Murphy, C. Loser, A. Eisel, D. Harms, U. R. Folsch, G. D. Lux, L. J. Marton, S. B. Baylin, M. Zhang, C. H. Clarke, D. M. Shankel, T. Eisenberg, M. Matsumoto, Y. Benno, M. Matsumoto, M. Sakamoto, Y. Benno, M. Burrell, C. C. Hanfrey, E. J. Murray, N. R. Stanley-Wall, A. J. Michael, T. Chinen, M. A. Wozny, M. P. Bryant, L. V. Holdeman, W. E. Moore, M. H. Floch, T. Pedron, P. Sansonetti, P. J. Sansonetti, R. Medzhitov, Y. Guigoz, J. Dore, E. J. Schiffrin, T. Ooga, T. Matsuki, M. Matsumoto, M. Sakamoto and Y. Benno, *Sci. Rep.*, 2012, **2**, 59–65.
38. M. Matsumoto, R. Kibe, T. Ooga, Y. Aiba, E. Sawaki, Y. Koga and Y. Benno, *Front. Syst. Neurosci.*, 2013, **7**, 9.

39. M. S. Donia and M. A. Fischbach, *Science*, 2015, **349**, 1254766.
40. J.-C. Lagier, S. Khelaifia, M. T. Alou, S. Ndongo, N. Dione, P. Hugon, A. Caputo, F. Cadoret, S. I. Traore, E. H. Seck, G. Dubourg, G. Durand, G. Mourembou, E. Guilhot, A. Togo, S. Bellali, D. Bachar, N. Cassir, F. Bittar, J. Delerce, M. Mailhe, D. Ricaboni, M. Bilen, N. P. M. Dangui Nieko, N. M. Dia Badiane, C. Valles, D. Mouelhi, K. Diop, M. Million, D. Musso, J. Abrahão, E. I. Azhar, F. Bibi, M. Yasir, A. Diallo, C. Sokhna, F. Djossou, V. Vitton, C. Robert, J. M. Rolain, B. La Scola, P.-E. Fournier, A. Levasseur and D. Raoult, *Nat. Microbiol.*, 2016, **1**, 16203.
41. N. Singhal, M. Kumar, P. K. Kanaujia and J. S. Virdi, *Front. Microbiol.*, 2015, **6**, 791.
42. M. Kogawa, M. Hosokawa, Y. Nishikawa, K. Mori and H. Takeyama, *Sci. Rep.*, 2018, **8**, 2059.
43. N. Cassir, S. Benamar, J. B. Khalil, O. Croce, M. Saint-Faust, A. Jacquot, M. Million, S. Azza, N. Armstrong, M. Henry, P. Jardot, C. Robert, C. Gire, J.-C. Lagier, E. Chabrière, E. Ghigo, H. Marchandin, C. Sartor, P. Boutte, G. Cambonie, U. Simeoni, D. Raoult and B. La Scola, *Clin. Infect. Dis.*, 2015, **61**, 1107–1115.
44. R. Caspi, H. Foerster, C. A. Fulcher, P. Kaipa, M. Krummenacker, M. Latendresse, S. Paley, S. Y. Rhee, A. G. Shearer, C. Tissier, T. C. Walk, P. Zhang and P. D. Karp, *Nucleic Acids Res.*, 2008, **36**, D623–D631.
45. D. S. Wishart, Y. D. Feunang, A. Marcu, A. C. Guo, K. Liang, R. Vázquez-Fresno, T. Sajed, D. Johnson, C. Li, N. Karu, Z. Sayeeda, E. Lo, N. Assempour, M. Berjanskii, S. Singhal, D. Arndt, Y. Liang, H. Badran, J. Grant, A. Serra-Cayuela, Y. Liu, R. Mandal, V. Neveu, A. Pon, C. Knox, M. Wilson, C. Manach and A. Scalbert, *Nucleic Acids Res.*, 2018, **46**, D608–D617.
46. D. S. Wishart, *Briefings Bioinf.*, 2007, **8**, 279–293.
47. G. H. F. Nuttall and H. Thierfelder, *Hoppe-Seyler's Zeitschrift für Physiol. Chemie*, 1897, **22**, 62–73.
48. T. Luckey, *Germfree Life and Gnotobiology*, Elsevier Science, 1963.
49. M. Wang and S. M. Donovan, *ILAR J.*, 2015, **56**, 63–73.
50. M. R. Charbonneau, D. O'Donnell, L. V Blanton, S. M. Totten, J. C. C. Davis, M. J. Barratt, J. Cheng, J. Guruge, M. Talcott, J. R. Bain, M. J. Muehlbauer, O. Ilkayeva, C. Wu, T. Struckmeyer, D. Barile, C. Mangani, J. Jorgensen, Y. Fan, K. Maleta, K. G. Dewey, P. Ashorn, C. B. Newgard, C. Lebrilla, D. A. Mills and J. I. Gordon, *Cell*, 2016, **164**, 859–871.
51. J. Sun, H. Zhong, L. Du, X. Li, Y. Ding, H. Cao, Z. Liu and L. Ge, *Sci. Rep.*, 2018, **8**, 10745.
52. E. Melancon, S. Gomez De La Torre Canny, S. Sichel, M. Kelly, T. J. Wiles, J. F. Rawls, J. S. Eisen and K. Guillemin, *Methods Cell Biol.*, 2017, 61–100.
53. S. Wong, W. Z. Stephens, A. R. Burns, K. Stagaman, L. A. David, B. J. M. Bohannan, K. Guillemin and J. F. Rawls, *mBio*, 2015, **6**, e00687-15.
54. C. Kietz, V. Pollari and A. Meinander, *Curr. Protoc. Toxicol.*, 2018, **77**, e52.
55. F. Zhang, M. Berg, K. Dierking, M.-A. Félix, M. Shapira, B. S. Samuel and H. Schulenburg, *Front. Microbiol.*, 2017, **8**, 485.

56. M. Guzman-Rodriguez, J. A. K. McDonald, R. Hyde, E. Allen-Vercoe, E. C. Claud, P. M. Sheth and E. O. Petrof, *Methods*, 2018, **149**, 31–41.
57. L. Liu, J. Firrman, C. Tanes, K. Bittinger, A. Thomas-Gahring, G. D. Wu, P. Van den Abbeele and P. M. Tomasula, *PLoS One*, 2018, **13**, e0197692.
58. S. Bartfeld and H. Clevers, *J. Mol. Med.*, 2017, 729–738.
59. N. Sachs, Y. Tsukamoto, P. Kujala, P. J. Peters and H. Clevers, *Development*, 2017, **144**, 1107–1112.
60. H. Yu, N. M. Hasan, J. G. In, M. K. Estes, O. Kovbasnjuk, N. C. Zachos and M. Donowitz, *Annu. Rev. Physiol.*, 2017, **79**, 291–312.
61. A. Bein, W. Shin, S. Jalili-Firoozinezhad, M. H. Park, A. Sontheimer-Phelps, A. Tovaglieri, A. Chalkiadaki, H. J. Kim and D. E. Ingber, *Cell. Mol. Gastroenterol. Hepatol.*, 2018, **5**, 659–668.
62. T. Yatsunenko, F. E. Rey, M. J. Manary, I. Trehan, M. G. Dominguez-Bello, M. Contreras, M. Magris, G. Hidalgo, R. N. Baldassano, A. P. Anokhin, A. C. Heath, B. Warner, J. Reeder, J. Kuczynski, J. G. Caporaso, C. A. Lozupone, C. Lauber, J. C. Clemente, D. Knights, R. Knight and J. I. Gordon, *Nature*, 2012, **486**, 222–227.
63. P. J. Turnbaugh, M. Hamady, T. Yatsunenko, B. L. Cantarel, A. Duncan, R. E. Ley, M. L. Sogin, W. J. Jones, B. A. Roe, J. P. Affourtit, M. Egholm, B. Henrissat, A. C. Heath, R. Knight and J. I. Gordon, *Nature*, 2009, **457**, 480–484.
64. P. Lepage, R. Häsler, M. E. Spehlmann, A. Rehman, A. Zvirbliene, A. Begun, S. Ott, L. Kupcinskas, J. Doré, A. Raedler and S. Schreiber, *Gastroenterology*, 2011, **141**, 227–236.
65. J. K. Goodrich, E. R. Davenport, M. Beaumont, M. A. Jackson, R. Knight, C. Ober, T. D. Spector, J. T. Bell, A. G. Clark and R. E. Ley, *Cell Host Microbe*, 2016, **19**, 731–743.
66. V. K. Ridaura, J. J. Faith, F. E. Rey, J. Cheng, A. E. Duncan, A. L. Kau, N. W. Griffin, V. Lombard, B. Henrissat, J. R. Bain, M. J. Muehlbauer, O. Ilkayeva, C. F. Semenkovich, K. Funai, D. K. Hayashi, B. J. Lyle, M. C. Martini, L. K. Ursell, J. C. Clemente, W. Van Treuren, W. A. Walters, R. Knight, C. B. Newgard, A. C. Heath and J. I. Gordon, *Science*, 2013, **341**, 1241214.
67. The Integrative HMP (iHMP) Research Network Consortium, The Integrative Human Microbiome Project, *Nature*, 2019, **569**, 641–648.
68. P. Rausch, M. Basic, A. Batra, S. C. Bischoff, M. Blaut, T. Clavel, J. Gläsner, S. Gopalakrishnan, G. A. Grassl, C. Günther, M. Hirose, S. Ibrahim, G. Loh, J. Mattner, S. Nagel, O. Pabst, F. Schmidt, B. Siegmund, T. Strowig, V. Volynets, S. Wirtz, S. Zeissig, Y. Zeissig, A. Bleich and J. F. Baines, *Int. J. Med. Microbiol.*, 2016, **306**, 343–355.
69. K. D. Parker, S. E. Albeke, J. P. Gigley, A. M. Goldstein and N. L. Ward, *Front. Microbiol.*, 2018, **9**, 1598.
70. I. I. Ivanov, K. Atarashi, N. Manel, E. L. Brodie, T. Shima, U. Karaoz, D. Wei, K. C. Goldfarb, C. A. Santee, S. V. Lynch, T. Tanoue, A. Imaoka, K. Itoh, K. Takeda, Y. Umesaki, K. Honda and D. R. Littman, *Cell*, 2009, **139**, 485–498.

71. H. E. Jakobsson, A. M. Rodríguez-Piñeiro, A. Schütte, A. Ermund, P. Boysen, M. Bemark, F. Sommer, F. Bäckhed, G. C. Hansson and M. E. V. Johansson, *EMBO Rep.*, 2015, **16**, 164–177.
72. R. N. N. Carmody, G. K. K. Gerber, J. M. M. Luevano, D. M. M. Gatti, L. Somes, K. L. L. Svenson and P. J. J. Turnbaugh, *Cell Host Microbe*, 2014, **17**, 72–84.
73. E. B. Chang, J. F. Pierre, R. Hinterleitner, R. Bouziat, N. Hubert, V. Leone, J. Miyoshi and B. Jabri, *Data Brief*, 2018, 387–393.
74. L. A. David, C. F. Maurice, R. N. Carmody, D. B. Gootenberg, J. E. Button, B. E. Wolfe, A. V. Ling, A. S. Devlin, Y. Varma, M. A. Fischbach, S. B. Biddinger, R. J. Dutton and P. J. Turnbaugh, *Nature*, 2014, **505**, 559–563.