

Review Article

Metagenomic Approaches for Understanding New Concepts in Microbial Science

Luana de Fátima Alves,¹ Cauã Antunes Westmann,² Gabriel Lencioni Lovate,¹ Guilherme Marcelino Viana de Siqueira,¹ Tiago Cabral Borelli,³ and Maria-Eugenia Guazzaroni³

¹Department of Biochemistry, Faculdade de Medicina de Ribeirão Preto, University of São Paulo, Ribeirão Preto, SP, Brazil

²Department of Cell Biology, Faculdade de Medicina de Ribeirão Preto, University of São Paulo, Ribeirão Preto, SP, Brazil

³Department of Biology, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, University of São Paulo, Ribeirão Preto, SP, Brazil

Correspondence should be addressed to Maria-Eugenia Guazzaroni; mequazzaroni@gmail.com

Received 23 February 2018; Revised 21 June 2018; Accepted 29 July 2018; Published 23 August 2018

Academic Editor: Henry Heng

Copyright © 2018 Luana de Fátima Alves et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Over the past thirty years, since the dawn of metagenomic studies, a completely new (micro) universe was revealed, with the potential to have profound impacts on many aspects of the society. Remarkably, the study of human microbiome provided a new perspective on a myriad of human traits previously regarded as solely (epi-) genetically encoded, such as disease susceptibility, immunological response, and social and nutritional behaviors. In this context, metagenomics has established a powerful framework for understanding the intricate connections between human societies and microbial communities, ultimately allowing for the optimization of both human health and productivity. Thus, we have shifted from the old concept of microbes as harmful organisms to a broader panorama, in which the signal of the relationship between humans and microbes is flexible and directly dependent on our own decisions and practices. In parallel, metagenomics has also been playing a major role in the prospection of “hidden” genetic features and the development of biotechnological applications, through the discovery of novel genes, enzymes, pathways, and bioactive molecules with completely new or improved biochemical functions. Therefore, this review highlights the major milestones over the last three decades of metagenomics, providing insights into both its potentialities and current challenges.

1. Introduction

About thirty years ago, in 1986, Pace and collaborators [1] proposed, for the first time, the revolutionary idea of cloning DNA directly from environmental samples to analyze the complexity of natural microbial populations (Figure 1, indicated as M2). The adopted strategy was based on shotgun cloning of 16S rRNA genes using purified DNA from natural samples. At that time, authors stressed that although the DNA was originated from a mixed population of microorganisms, the methodology allowed the recovery and subsequent sequencing of individual rRNA genes. Thus, by evaluating complete or partial rRNA sequences, the composition of the original microbial populations could be retrieved.

Around ten years later, in 1998, the term “metagenome” appeared, when Handelsman and collaborators [2] described the importance of soil microorganisms as sources for new natural compounds (Figure 1, indicated as M6). According to them, a new frontier in science was emerging—the mining for novel chemical compounds from uncultured microorganisms, which comprises more than 99% of the microbial diversity [3]. This new concept in microbial science opened the mind of the scientific community with respect to the astonishingly large catalogue of biochemical functions available in nature remaining to be discovered.

Currently, metagenomics is subdivided into two major approaches, which target different aspects of the local microbial community associated with a determined environment.

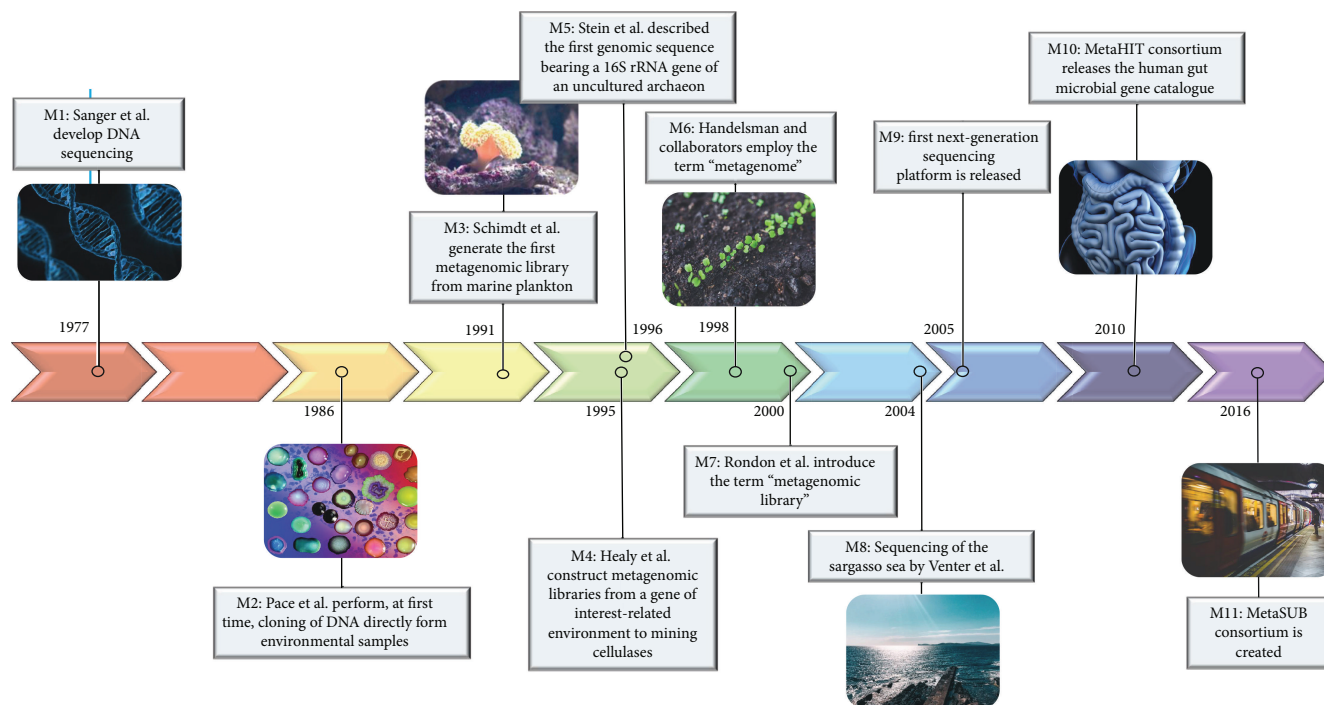


FIGURE 1: Timeline of the major advancements in metagenomics. Timeline highlighting important developments in the field over the last 40 years, since Sanger sequencing (M1), and over the last 30 years, since the first published metagenomic experiment (M2). The main metagenomic milestones are shown as M1–M11 (all of them are highlighted in the text where they were mentioned).

In the first one, the so-called structural metagenomic approach, the main focus is to study the structure of the uncultivated microbial population, which can be expanded to other properties, such as the reconstruction of the complex metabolic network established between community members (Figure 2) [4, 5]. In this sense, the microbial community structure can be defined as the population composition and its dynamics in a specific ecosystem, in response to selective pressures and spatiotemporal parameters. The study of the community structure allows a deeper understanding about the relationships between the individual components that build a community and is essential for deciphering ecological or biological functions among its members [5, 6]. In a different manner, the functional metagenomic approach aims to identify genes that code for a function of interest, which involves the generation of expression libraries with thousands of metagenomic clones followed by activity-based screenings (Figure 2) [7, 8].

It is important to highlight that 16S rRNA gene surveys are often referred to as metagenomic studies, although they are not. In the 16S rRNA gene analysis, the study is focused on a single gene used as a taxonomic marker (Figure 2). On the other hand, structural metagenomics aims to investigate the genomes of the microbial community members. In this sense, the later approach allows the overall reconstruction of the community structure, potentially revealing metabolic pathways of the whole microbiome and assigning minor or major geoecological roles to community members [4, 6, 9, 10]. We highlight that 16S rRNA surveys and metagenomics are not mutually exclusive; on the contrary,

approaches that establish a link between 16S rRNA analyses to genes or metabolic pathways have been shown to be useful in determining the functional potential of a microbiome [11–13]. Thus, the combination of these complementary strategies allows for a deeper exploration of relevant biological questions in microbial ecology such as “who are the members of the community?” and “what are their functional roles?”

At the beginning of the metagenomic studies, the use of Sanger sequencing technology [14] provided important progresses in the field [15–17] (Figure 1, indicated as M1). However, the advent of next-generation sequencing (NGS) technologies capable of sequencing millions of DNA fragments simultaneously, at a low cost, greatly bolstered the field [18–21] (Figure 1, indicated as M9). Comparatively, NGS platforms can recover up to 5000 Mb of DNA sequence per day with costs at about 0.50\$/Mb, while Sanger sequencing methodology allows about 6 Mb of DNA sequence to be created per day with costs at about 1000 higher [22].

This review focuses on how metagenomics has contributed to gain scientific comprehension in many different areas of knowledge. In this manner, milestones in metagenomics have ranged from findings with significant biotechnological impact to unexpected outcomes with high biomedical relevance, shining a light on hidden molecular components and on the connections between microbial communities and complex diseases [23–26]. We also discuss the current boundaries of the field that should be overcome for the achievement of conceptual advances in microbial science.

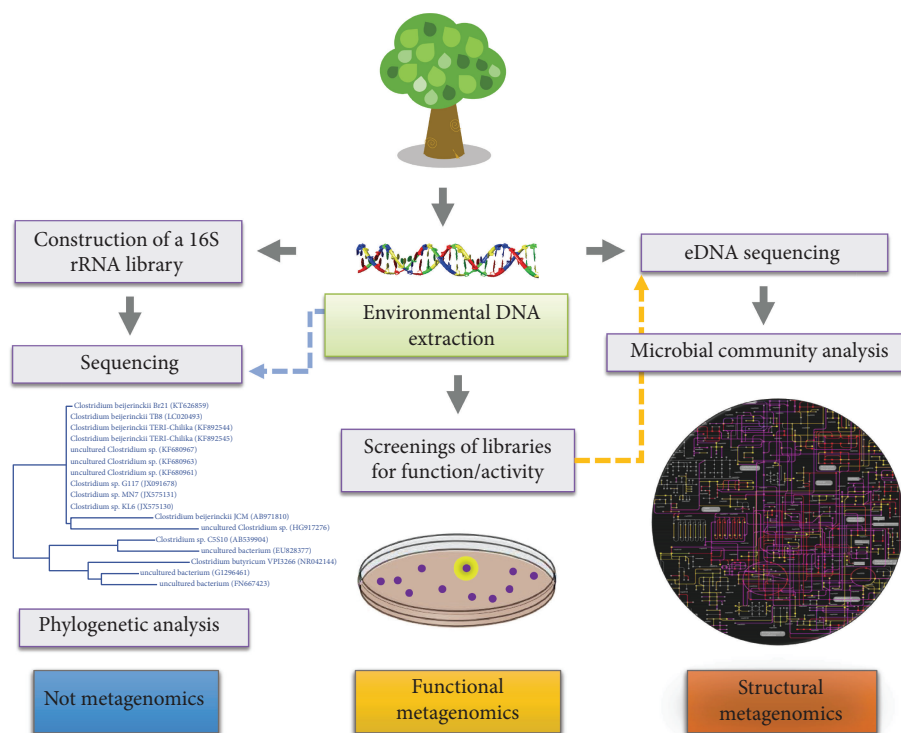


FIGURE 2: The metagenomics framework and its two main approaches. Both structural and functional metagenomic approaches are the main strategies for exploring key ecological and biotechnological features in environmental samples, respectively. Additionally, 16S rRNA gene surveys can work in synergy with metagenomics for further understanding of microbial ecology.

2. Milestones in Metagenomics

In 1991, Schmidt and collaborators generated the first metagenomic library using DNA from marine picoplankton [27], and, some years later, Healy et al. constructed metagenomic libraries from an enriched consortia sampled from cellulose digesters to mine genes encoding cellulases [28] (Figure 1, indicated as M3 and M4, resp.). In this context, the idea of screening metagenomic libraries from specific environments was introduced, allowing that the number and properties of the retrieved genes (e.g., enzymes) could be correlated to the conditions of the source environments. However, only in 2000, Rondon and collaborators [29] coined the term “metagenomic libraries,” by generating libraries in BACs (bacterial artificial chromosome) using DNA from soil samples (Figure 1, indicated as M7). Furthermore, the authors also performed phylogenetic analysis of 16S rRNA sequences and identified a number of clones expressing heterologous genes in functional screenings using *Escherichia coli* as host.

Since then, a large amount of data has been generated using metagenomic approaches and impacting different areas of high applicability in our society (Figure 2). Here, we highlighted the main milestones in metagenomic studies that defined the field in distinct contexts. Although many of the advances in the field can be credited to novel sequencing approaches and the development of new computational methodologies to analyze the generated data, in this review, we have focused on highlighting biological discoveries through metagenomics rather than describing these more

technical strategies, which have been addressed in recent reviews [30–32].

2.1. Genomic and Taxonomic Novelties in Environments. Ever since the proposal of using molecular markers such as 16S (or 18S) rRNA for comparing species similarity, in the late 1970s, followed by the outlining of a third “urkingdom” (the Archaeobacteria), our knowledge of organism relatedness has taken a great leap forward. Taxonomic classification started to rely on a “comparative approach that can measure the degree of difference in comparable structures,” which not only allowed a more resolved phylogeny and a less biased organization based on the *Prokaryotae* versus *Eukaryotae* dichotomy but also made it possible to better understand how life on earth has come to be [33]. In this context, metagenomic approaches have been used to generate data of novel genomes from otherwise uncultivated organisms, deepening the framework of genomic tools available for comparison and study.

The shotgun sequencing of the Sargasso Sea waters, by Venter and colleagues in 2004 [34], is one of the most illustrative examples of how metagenomics are a feasible way to accumulate genomic knowledge (Figure 1, indicated as M8). In this study, Venter and collaborators have recovered almost 1.5 Gbp of microbial DNA sequences from microbial populations of three marine sites using en masse whole genome shotgun sequencing from filtered sea water. This leads to the finding of almost 70 thousand novel genes among the roughly 1.2 million genes by ORF (open reading frame)

identification and alignment of the putative protein products. Among the main findings, the researchers described a novel ammonia oxidation pathway uninhibited by UV light, putative genes for trace metal resistance such as arsenate and copper, and, additionally, up to 782 new proteorhodopsin-like receptors genes. The latter added an insight to the exotic marine photoheterotrophic lifestyle first described by Béjà and colleagues only a few years before [35]. Further, the metagenomic study of the Sargasso Sea allowed the identification of 148 previously unknown rRNA bacterial phylotypes. However, this probably was an underestimation of the environment's total genomic pool, given that the difference in the number of rRNA coding genes within the rRNA operon between species may result in biases in PCR studies with under representation of some of the community constituents, as also discussed by Klappenbach and colleagues [36].

In this context, a recent study based on metagenomic data has shown that about 10% of environmental bacterial or archaeal sequences might not be recovered when using a targeted PCR survey with the most common primers for SSU rRNA [37]. For instance, a study led by Brown and colleagues in 2015 [38] first described the Candidate Phyla Radiation (CPR) bacterial lineage comprising at least 25 new bacterial phyla, which make up to at least 15% to the Bacteria domain. These novel organisms lack many biosynthetic pathways and possess unusual features in their small genomes, such as self-splicing introns within the rRNA genes and novel ribosome structure, providing insight to the organelle's evolution trajectory. It is worth noting that all complete CPR genomes curated in the work have only one copy of the 16S rRNA gene and many of the organisms would evade detection by 16S rRNA gene amplicon surveys, even though corresponding to such a high percentage of bacterial diversity in environments. In their work, Venter and collaborators have attempted to address this issue and to better elucidate the phylogenetic relationships in the recovered genetic material by employing six other phylogenetic markers, such as heat shock protein 70 (HSP70) and elongation factor Tu (EF-Tu), as well as several methods to estimate species diversity, which have resulted in the estimation of over 1000 species [34].

Metagenomic studies also shed a light on challenges in the field of evolutionary biology, such as in the understanding of sexual reproduction as a constraint on genomic variation [39]. Previous to the metagenomic era, it was believed that many microbial species should be genomic clones—asexual reproduction was assumed to produce identical clones. Yet, metagenomic analyses unexpectedly revealed that most microbial species were not clonal [34]. Thus, asexual reproduction present in bacteria should increase genetic variation providing evolutionary diversity for future environmental challenges [40]. This finding was essential to support the conclusion that sexual reproduction acts as a constraint on genomic and epigenetic variation, thereby limiting adaptive evolution [39].

In recent years, new sampling and sequencing techniques allowed researchers to further explore life diversity improving phylogenetic, genomic, and ecological notions previously

established. Rinke and colleagues, for instance, revealed the “microbial dark matter” in a single-cell genomics study that comprised over 20 major uncultivated archaeal and bacterial lineages [41]. They first described archaeal sigma factors and the first reported lateral gene transfer from a eukaryote to an archaeon. Another insight of how the archaeal/eukaryotic relationship came to be and how the modern eukaryotic cell was formed was found by Spang and collaborators [42], who have identified a candidate archaeal phylum which forms a monophyletic group with eukaryotes. It also possesses genes encoding proteins similar to those related to cell shape formation processes in eukaryotic cells, which might suggest sophisticated membrane remodeling and vesicular trafficking processes in eukaryotic cells even before the acquisition of mitochondrion.

2.2. Innovative Functions in Uncultivable Microbes.

Sequenced-based analysis in metagenomics can be accomplished, overall, by following one of the following paths: (i) sequencing all clones with a phylogenetic marker indicating the potential taxonomic source of the DNA fragment or (ii) sequencing random fragments until a gene of interest is found followed by sequencing of the adjacent regions to find taxonomic markers. The former method was developed by Stein's group and described the first genomic sequence bearing a 16S rRNA gene of an uncultured archaeon (Figure 1, indicated as M5) [43]. This provided the first highlights of the metagenomics capabilities for unravelling novel genes, functions, and taxonomic groups. In this study, with a marine picoplankton assemblage collected at eastern North Pacific, a 38.5 kbp fragment containing an archaeal 16S rRNA gene was isolated for the first time. Among other features, the discovery of an RNA helicase and a glutamate semialdehyde aminotransferase, which were still unknown in archaeal organisms, was reported. The metagenomic library in this study had an average fragment size of 40 kb and was based on fosmid backbones. The screening was very labor intensive, relying on Southern blotting with phylogenetic probes for 16S rRNA gene, followed by sequencing of selected clones through automated Sanger or shotgun methods.

It did not take long for the same group to push the boundaries and further advance the incipient field of metagenomics [35, 44]. From the surface waters of the Californian coast, a marine planktonic bacterial assemblage was collected for metagenomics analysis. Hitherto, bacteriorhodopsins were considered unique features of halophilic archaea; however, the small ribosome subunit gene revealed its source as a gammaproteobacterium (uncultivable bacterial SAR86 lineage). It was also shown that the protein was functional when cloned into *E. coli*, presenting similar kinetics to the archaeal-related cognates [35]. In this study, both metagenomic library and fragment average sizes were much larger than in the previous study and the screening process was accelerated (6240 screened clones with an average size of 80 kb) [44]. These improvements were only possible due to the establishment of BACs and PCR-based screening as new tools for metagenomic studies. Thus, it did not take long for novel proteorhodopsins to be detected in other populations of planktonic marine bacteria [45]. This was considered the first

great “metagenomic success,” allowing the adoption of these new techniques among laboratories around the world.

2.3. Deciphering and Rebuilding Microbial Communities. By starting with an extremely simple microbial community—an acid mine drainage (AMD) microbial community—evidenced by an initial group-specific fluoresce in situ hybridization (FISH) assay, Tyson and collaborators [46] were able to perform the first assembly of genomes directly from environmental samples. In this study, they have obtained two nearly complete genomes of *Leptospirillum* group II and *Ferroplasma* type II and partially recovered another three genomes. The initial characterization of this microenvironment, a biofilm with rather extreme conditions like very acidic pH (approximately 0.83), revealed the presence of archaea (*Thermoplasmatales*) and bacteria (*Leptospirillum*, *Sulfobacillus*, and *Acidimicrobium*), with the domination of *Leptospirillum* group II. The low diversity of the system was reported as possibly related to the extreme environmental conditions. Afterwards, the DNA sequencing of a biofilm sample suggested adaptive molecular traits of the community to survive in this environment—such as homologous recombination forming mosaic genome types—and metabolic adaptations, such as abundance of genes related to ferrous iron oxidation.

In other metagenomic studies, Tringe et al. [5] performed comparisons of the composition and functionality of microbial communities from two nutrient-poor and two enriched-nutrient environments. This approach was mainly concerned with gene function rather than genome composition, overcoming limitations to genome assembly from complex environments. Authors showed that gene function and structure differ in nutrient-limited (Sargasso Sea and AMD) versus nutrient-abundant (Minnesota farm soil and deep-sea “whale fall” carcasses) environments. Some gene functions were exclusive to specific environments, for instance, (i) cellobiose phosphorylases were only found at the agricultural soil and (ii) light-driven proton pumps are only found at the Sargasso Sea samples, whereas no photoreceptors were found at the deep-sea samples.

The development of functional approaches in large scales also provided novel insights into communities’ key metabolic process. A comparative functional profiling of 9 biomes was performed by Dinsdale and collaborators [47], describing how different biological traits play essential roles in each environment. For instance, authors showed an abundance of virulence genes in fish- and terrestrial-animal-associated metagenomes in comparison to the other biomes. In contrast, virome analysis of the total number of biomes showed a more uniform gene functional composition due to phages playing similar roles in different environments [47].

2.4. City-Scale Molecular Profile of DNA. The last decades of metagenomics have shed a light on the human microbiome and its profound influence on a wide range of aspects which were previously regarded as solely (epi-) genetically encoded, such as diseases susceptibility, immunological response, and social and nutritional behaviors [48–52]. Although there is still much to learn in this subject, recent studies are targeting

not only the inside but also the external part of the human world, also known as microbiomes from human-built environments [53–55]. In this context, a myriad of anthropocentric utensils and physical spaces has been assessed, such as kitchen sponges [56], dollar bills [57], ATMs (Automated Teller Machines) [58], homes [59], hospitals [60], subways [61, 62], food production sites [63], and even the International Space Station [64]. The main goal of this new research branch is to provide a framework for understanding the relationship between human societies and microbial communities, ultimately optimizing both human health and productivity.

A seminal research in the context of urban areas was conducted by Afshinnekoo et al. [61], which sampled different features of New York subways and reported that nearly 1700 microbial taxa were dominated mostly by human skin bacteria and to a lesser extent by microbes from the human gastrointestinal and urogenital tracts. Almost half the DNA present on the subway surfaces matched no known organism. Although results showed that the bacteria found in the subways were mostly harmless, several pathogenic agents, including fragments of the plague and anthrax genomes were detected. This was the beginning of an international consortium called The Metagenomics and Metadesign of Subways and Urban Biomes (MetaSUB) that has been sampling urban microbiomes throughout the world [62] (Figure 1, indicated as M11). An important scientific agenda was also launched in 2017, the microbiomes of the built environment [65]. Its main objective is to assess the current state of knowledge on indoor microbiomes and also to map out research agendas and advise government agencies on how living spaces can be designed “to support occupant health and well-being.” Other studies have assessed complementary aspects of this matter, such as the influence of landscape connectivity in microbial diversity [66], the influence of green areas in urban spaces [67, 68], and the consequences of excessive time expenditure indoors in the context of both human health and environmental microbiomes [54, 69].

Altogether, those studies allowed the depiction of a few important conclusions regarding human-built spaces and microbiomes, further explained by Stephens et al. [70]: (i) culture-independent methods are essential for those surveys, (ii) indoor spaces often harbor unique microbial communities, tightly related to the indoor sources—mostly humans and pets, (iii) building design and operation can directly modulate indoor microbial communities, and (iv) it is possible to optimize human health by exposure to certain microbial groups. Consequently, society moved from the old concept of microbes as harmful organisms to a new view in which the interaction between humans and microorganisms can be flexible and directly dependent on our own decisions and practices. Further studies shall reveal novel rules on “good living standards” for both humans and microbes in built environments.

2.5. The Human Microbiome. The concept of the human microbiome, embracing the idea that human beings are highly susceptible to the microbial communities that live in and on our bodies, was an indubitable milestone in

metagenomics with large repercussion in many areas. In this sense, scientific contributions involving metagenomic approaches rapidly highlighted the evidences showing that these microbiomes play key roles in human health and disease.

The human gut microbiota—the collection of microorganisms that compose the human gut—is composed by up to 10^{14} microorganisms [71, 72] including bacteria, viruses, fungi, and protozoa. Deciphering the function and composition of our microbiome—the collective genomes of the microbial community resident—is a challenge that has been explored by researchers in a series of initiatives like the Human Microbiome Project (HMP), the Integrative Human Microbiome Project (iHMP), and the MetaHIT (METAge-nomics of the Human Intestinal Tract) [73–75] (Figure 1, indicated as M10). The findings of these projects have provided valuable data about the function of the human microbiome in different tracts (e.g., nasal, oral, skin, gastro-intestinal, and urogenital). Particularly, advances in molecular biology procedures, next-generation DNA sequencing, and *omics* techniques have allowed to access not only to the microbial genetic diversity but also to the understanding of the physiology and the lifestyle of our microbiome. In this sense, it was demonstrated that gut microorganisms perform several elemental functions like synthesis of essential amino acids and vitamins and processing of cellulosic material [76], playing an important role in a number of human health aspects [77].

In a series of studies coming out of the Gordan lab at Washington University School of Medicine in St. Louis, Ley and collaborators [78] showed that obesity has a microbial component. To explore the relation between gut microbial ecology and body fat in humans, authors studied 12 obese people, who were randomly assigned to two types of low-calorie diet. The composition of their gut microbiota was monitored over the course of one year by sequencing 16S ribosomal RNA genes from stool samples [78]. Obtained data showed that two groups of beneficial bacteria were dominant in the human gut, the Bacteroidetes and the Firmicutes. In addition, they showed that the percentage of Bacteroidetes correlated with the percentage of loss weight. In other studies [79], they found that the transplantation of the microbiota from obese mice to lean mice could lead to an increase of body fat in transplanted mice when compared with transplantation from lean mice microbiota.

The other important outcome from the human gut microbiome studies was regarding antibiotic resistance. In order to determine an “antibiotic resistance potential,” Forslund et al. [80] performed a quantitative gut metagenomic analysis of known resistance genes from people of three countries. In this study, authors showed that antibiotic resistance gene abundance in the general human population is correlated with the length of antibiotic usage in livestock [80–82]. In another study, Raymond and coworkers [83] showed that the initial gut microbiome affects its recomposition after antibiotics treatment. They administrated two second-generation antibiotics (cephalosporin and cefprozil) in healthy individuals and showed that antibiotics altered the microbiome of healthy volunteers in an interindividual

manner, allowing the emergence of potentially pathogenic *Enterobacteriaceae*—in a subgroup of patients—probably related to a decreased initial gut microbiome diversity in those individuals.

Besides obesity and antibiotic resistance, the human gut microbiome has been associated with several diseases as type 2 diabetes, cardiovascular and inflammatory bowel diseases, and even cancer [84]. Some studies have also associated gut microbiome with intestinal immunity. It has been shown that a healthy microbiota improves local expression of a Toll-like receptor (TLR) [85] which recognizes the PAMPs (pathogen-associated molecular patterns) expressed by a broad range of infective agents and improves the percentage of antigen-presenting cells, differentiated T cells, and lymphoid follicles [86, 87]. Besides the local immunity, the gut microbiota affects the systemic immunity by increasing splenic $CD4^+$ T cells and systematic antibody expression [88]. Consequently, the global role of the gut microbiota in intestinal immunity has increased the interest of the scientific community in developing techniques that improve human health by manipulating the gut microbiota.

Nowadays, researchers are exploiting these important results for medical applications; for instance, fecal microbiota transplantation (FMT) has been used to eliminate *Clostridium difficile* recurrent infection by transplantation of healthy microbiota in human patients [89]. Besides that, FMT has been successfully used in treatment of inflammatory bowel disease, functional gastrointestinal disorders, hepatic encephalopathy, obesity, and metabolic syndrome [24, 90].

2.6. Biomedical Significance. Findings from studies of the gut microbiome shed a light over a number of diseases directly impacted by it, becoming a promising scope for advances in understanding and treating of complex diseases. Among them, Crohn’s disease [91], rheumatoid arthritis [92], obesity [93, 94], type 1 and type 2 diabetes [95–97], breast cancer [98, 99], and atherosclerosis [100] can be cited as associated to the gut microbiome. Thus, researchers are interested in finding biomarkers and microbiome-based signatures for use in diagnostics, prognostics, and treatment of patients with diseases related to the human microbiome, describing important targets with biomedical significance that could be useful for public health.

In this context, Yu and collaborators [101] performed a metagenome-wide association study (MGWAS) using stool samples from 74 Chinese individuals with colorectal cancer and 54 controls, aiming to identify noninvasive biomarkers for colorectal cancer. Authors found that, besides known colorectal cancer-associated species such as *Fusobacterium nucleatum* and *Peptostreptococcus stomatis*, the other 20 microbial gene markers could differentiate colorectal cancer from control patients. In order to define a “worldwide” signature for colorectal cancer identification, they validated four of these gene markers in Danish-, French-, and Austrian-published cohorts. This result indicates that the four biomarkers validated in individuals from different countries might be used to early diagnosis of colorectal cancer even in different populations with different gut microbiome structures and is a promise for early noninvasive diagnosis of the

disease. Later, Yu's group [102] developed a new diagnostic tool for colorectal cancer using the four biomarkers validated in 2015. For this, they applied a probe-based duplex qPCR assay for quantifying these bacterial genes and showed that one of these genes can discriminate colorectal cancer from control individuals with 77.7% of sensitivity and 79.5% of specificity. Moreover, combining these four bacterial genes with a fecal immunochemical test improved the diagnostic, displaying a sensitivity of 92.8% and a specificity of 81.5%.

In the same way, Pascal et al. [103] defined a microbial signature for Crohn's disease. They performed a MGWAS using samples from 2045 individuals from four countries (Spain, Belgium, UK, and Germany) and found eight groups of microorganisms that could be used to discriminate between Crohn's disease and ulcerative colitis (the two main inflammatory bowel diseases that share many immunologic, epidemiologic, and clinical features). Then, they developed an algorithm that showed specificity of approximately 90% of Crohn's disease detection when compared with healthy control, anorexia, ulcerative colitis, and inflammatory bowel syndrome patients. Similarly, Looma et al. 2017 [104] defined a gut microbiome signature for the diagnosis of advanced fibrosis in individuals with nonalcoholic fatty liver disease using a metagenomic analysis. Analogously, a similar approach was used to distinguish between type 2 diabetes individuals and nondiabetic controls [97]. By means of a sequence-based profiling metagenomic approach, authors showed that type 2 diabetes individuals were characterized by an increase of opportunist pathogens, an enrichment of sulphate reduction genes and oxidative stress resistance genes, and a decrease in butyrate-producing species. Taken together, these gut metagenomic markers might become a powerful tool for the diagnosis of patients with the disease.

2.7. Mining of Microbial Genes In Vivo. New outcomes from the studies of the human microbiome inspired novel biological questions related to the microbial fitness at diverse human tracts. Thus, elucidating the set of genes involved in the colonization and maintenance of the intestinal microbiota would provide valuable tools to further engineering of probiotics or to enhance the survival of certain microorganisms directly related to healthy conditions.

In this way, the *in vivo* temporal functional metagenomics approach, developed by Yaung and collaborators [105], allowed them to mine microbial genes associated to microbial fitness in the mammalian gastrointestinal tract. The temporal sequencing platform employed in this study allowed the detection of genes that confer microbial fitness using a mouse as a model host. Thus, authors inoculated germ-free mice with *E. coli* transformed with a library composed of 2–5 kb fragments of the *Bacteroides thetaiotaomicron* (Bt) complete genome. Through a 28-day experiment of kinetic monitoring of the enriched clones inside the mice's gastrointestinal tract, they were able to identify genes related to colonization and maintenance of the bacteria inside the tract. The study revealed that different sets of genes were enriched in the pool of bacteria at contrasting times. For instance, during the early phase of colonization, genes responsible for the synthesis of polysaccharides and

lipopolysaccharides (LPS) were significantly present in the pool, expanding *E. coli*'s LPS biosynthesis repertoire with the acquired Bt biosynthetic genes. Those changes in the LPS synthesis could provide the *E. coli* with different antigenicity and enhance resistance to barriers for colonization in the gut. Nevertheless, a distinct set of genes was present in the long-term experiment, mostly related to sugar metabolism or transport. Furthermore, except for a mutation in the *galK* chromosomal gene, the recipient strain maintained its genetic stability [105].

The work is the first one to use temporal-functional metagenomics to describe how temporal data can contribute to the discovery of genes with functions of interest, once most of the genes involved in the GI tract community fitness would not be found if the data was from only an endpoint [105]. Temporal approaches like these could also bring interesting insights about interaction dynamics and fitness of microbes in other environments (such as bacteria associated to plant growth or parasitic interactions), unraveling the genes involved in a microbial community structure and metabolism. Moreover, identified genes could be used as novel drug target genes in pathogenic bacteria.

2.8. Biotechnological Impact. Biotechnology is one of the most favored fields by the metagenomic era. As microorganisms are the major source of biocatalysts for industrial purposes, increasing the repertoire of biochemical transformations available for biotechnological solutions is of high relevance [106]. Since Healy and collaborators [28] introduced the idea of constructing metagenomic libraries from a gene of interest-related environment, functional and sequence-based metagenomics have been shown to be effective in the identification of novel genes that confer resistance to extreme conditions, enzymes, antibiotics, and other bioactive molecules derived from a variety of environments (Table 1) [107–111]. Furthermore, functional metagenomic approaches made it possible to identify novel biological parts with specific activities without the need for isolation and cultivation of microorganisms.

In this context, a report from Ferrer and collaborators [112] should be highlighted. Authors constructed a metagenomic library of DNA from a cow rumen in a phage lambda vector and performed functional screenings for different carbohydrate-active enzymes. Considering that cow rumen microorganisms are specialized in degradation of cellulosic plant material, the sample used for library construction should be enriched in biomass-degrading genes. The success of the approach was exposed by the high rate of recovering of clones with different hydrolytic activities (22 clones), being among them are 9 esterases, 12 endo- β -1,4-glucanase, and 1 cyclodextrinase. Moreover, after DNA sequencing analysis of the enzymes, 8 could not be found in any sequenced genomic data, revealing that 36% of the recovered enzymes were completely new.

Other interesting works related to exploring functional genes in leaf-cutter ant fungus gardens were carried out to determine enzymes and pathways involved in symbiosis between leaf-cutter ants and their cultivar. By using metagenomic approaches, authors have determined the microbial

TABLE 1: Genes discovered through metagenomic approaches with high biotechnological potential.

Function/gene target	DNA source	Library size	Screening method*	Number of hits found	Biotechnological relevance	Reference
<i>Enzymes</i>						
Esterases, endo- β -1,4-glucanases, and cyclodextrinase	Cow rumen	1.1 Gb	Function based	22	Eight enzymes (36%) were entirely new	[112]
Laccase	Water from South China Sea	1.4 Gb	Sequencing based	1	High chloride resistance and ability to decolorize industrial dyes	[136]
Naphthalene dioxygenase	Oil-contaminated soil	294 Mb	Function based	2	Applicable in oil-contaminated soil/water	[137]
Oxygenases	Artificially polluted soil	5.2 Gb	Function based	29	Applicable in oil-contaminated soil/water	[126]
Cutinases	Leaf-branch compost	735 Mb	Function based	19	Potential application in polyethylene terephthalate (PET) degradation	[138]
Phenol hydroxylases and catechol 2,3-dioxygenases	Wastewater treatment plant	495 Mb	Function based	413	Potential use in aromatic compound degradation	[139]
Carboxylesterase	Marine water	~1.3 Gb	Function based	95	Cold-active and salt-resistant enzyme	[140]
Cellulase/esterase	Water lakes	1.86 Gb	Function based	3	New cellulase	[141]
Cellulase	Soil	Not found	Function based	1	Halo- and thermotolerant enzyme	[142]
β -Glucosidase	Hydrothermal spring water	Not found	Function based	1	Thermotolerant and heat-active enzyme	[143]
Lipase/protease/hemolysins/biosurfactants	Slaughterhouse drain	~884 Mb	Function based	22	Antimicrobial activity	[144]
<i>Genes that confer resistance to extreme conditions</i>						
Acid resistance genes	Plankton and rhizosphere from Tinto River	~2.3 Gb	Function based	15	Genes involved in acid resistance	[118]
Nickel resistance genes	Rhizosphere of <i>E. andevalensis</i> from Tinto River	2.15 Gb	Function based	13	Genes related to nickel resistance	[117]
Salt resistance genes	Brine and rhizosphere from Es Trenc saltern	2.15 Gb	function-based	11	Genes conferring salt resistance	[116]
Arsenic resistance genes	Headwater from Tinto River	151 Mb	Function based	18	Genes involved in arsenic resistance	[145]
<i>Regulatory sequences</i>						
Constitutive promoters	Soil from secondary Atlantic Forest	~500 Mb	Function based	33	Use as “biobricks”	[135]
<i>Pathways/systems/operons</i>						
Naphthalene-degrading system	Naphthalene-contaminated groundwater	~283 Mb	Sequencing based	3	Pollutant-degrading enzyme systems	[146]
Dioxygenase-degrading cluster	Forest soil	260–815 bp	Sequencing based	11	Degrading phenoxyalkanoic acid (PAA) herbicides avoiding groundwater contamination	[147]
NRPS biosynthetic pathway	Tunicate consortium in Florida Keys	~280 Mb	Sequencing based	1	ET-743 biosynthetic pathway; anticancer molecule	[148]
<i>Bioactive molecules</i>						
Pigmentation producing and antibacterial activity	Soil	Not found	Function based	45	Potential new molecules to be used as antibiotics	[125]

TABLE 1: Continued.

Function/gene target	DNA source	Library size	Screening method*	Number of hits found	Biotechnological relevance	Reference
Turbomycin A and B	Soil	~1 Gb	Function based	3	Antibiotic activity	[16]
Antimicrobial small molecules	Soil	~720 Mb	Function based	4	Antibiotic activity	[149]

*All genes discovered through sequencing-based methodologies were experimentally tested for their related functions.

composition of the fungus gardens [113, 114] and how the plant biomass degrading process in this microbial community occurs, showing a number of novel cellulases involved in it [113–115].

Alternative relevant metagenomic studies take advantage of the genetic potential of microbe inhabitants of extreme environments, such as high or low temperatures, salinity, acidity, pressure, radiation, or high concentrations of heavy metals (Table 1) [116–118]. Deciphering microbial diversity and metabolic activities of microorganisms in extreme conditions reveals the biochemical strategies used by them to survive under extreme conditions. This, in turn, can be used to expand the capability of survival of bacteria used in industrial processes. In addition, those organisms are interesting sources for enzymatic activities with specific and unusual features. Table 1 summarises some of the most noticeable genes that have been discovered until now using the strategies described above.

3. Conclusions and Perspectives

Over the last three decades, since the first studies using the concept of metagenomics, extraordinary advances in the field have been achieved. Collective intelligence from a plethora of experts in diverse fields (such as biologists, biochemists, geneticists, physicists, and computer scientists) was imperative for answering central biological questions and for bringing biotechnological solutions in a myriad of different fields.

Understanding properties such as structure, diversity, richness, and dynamics of microbial communities is essential for unravelling the underlying processes that govern the organization of those systems. However, for a more comprehensive analysis, it is crucial to integrate information from both the microbial community and the environment it is embedded in. The macrodynamics of physical spaces and the interactions between their components directly modulate the microbiological universe (and vice versa). Thus, understanding the physical, chemical, and relational aspects of an environment can provide insightful predictions about its microinhabitants, whereas the reverse process, depicting an environment from a collection of “microbial footprints,” although more challenging, is also attainable [119].

In this context, one of the most proximal models of the study is our own and the built spaces we create and live in. From our bodies to our cities and far away, the Earth is heavily packed with microbes [120–122]. A “reference man” (one who is 70 kilograms, 20–30 years old, and 1.7 meters tall) contains on average about 30 trillion human cells

and 39 trillion bacteria [123] and emits bacteria at rates of over a million biological particles per hour [124]. Then, it is daunting to ask can we understand the “maketh men” through its microbiome? Nowadays, we know that our personal ecosystem of microbes is shed on everything we touch and everywhere we go as “molecular echoes.” Thus, can we trace back an individual lifetime through metagenomics? What about the lifetime of a whole society? In reverse, can we use metagenomics to guide the rational design of novel-built environments (indoors and outdoors) for artificially selecting microbial communities, which will ultimately contribute to human health? The answer to all those exciting questions is yes, we can—at least, partially—as it was described in diverse examples along this revision, which are enabling us to move towards the right answers.

In contrast, despite the notable milestones reached in the field, there are still crucial challenges that need to be faced in order to delineate new conceptual advances in microbial science. Development of novel bioinformatic tools specific for metagenomic analysis is still necessary, once the next-generation sequencing platforms are generating an increasing amount of data that is not directly proportional to its biological significance. That is, there is an enormous quantity of information in sequenced data that need to be transformed into biological understating. In the near future, huge data processing and analysis in an integrated way with data already known will be the main challenge of the field [30].

On the other hand, the success of the function-based screenings depends on factors like the size of the gene in metagenomic DNA, its abundance in the samples, the efficiency of the screening method, the host vector system used, and the heterologous expression of the gene. Nowadays, after overcoming of some critical limitations related to host-biased screenings, researchers have used alternative hosts besides *E. coli* to perform the screenings [125–130]. For this, the use of broad host range vectors, able to replicate in different hosts, is required. Although a large collection of broad host range vectors is available [131], we still need to create *a la carte* vectors specific for some microorganisms that are essential to certain industrial uses. In this sense, strategies involving synthetic biological approaches [132] have been crucial to develop new smart screening methods. Engineering biological circuits—the so-called biosensors—have accelerated the identification of positive activities in metagenomic screenings including millions of clones. Interesting examples are the substrate-induced gene expression (SIGEX) and the product-induced gene expression (PIGEX) approaches [17, 133] or a riboswitch-based selection system initially constructed

for mining thiamine uptake functions [134], but generalizable to other compounds. Furthermore, developing new methods for expanding the search space of functional metagenomics from enzymes to novel genetic elements such as regulators, promoters, and cis-regulatory sequences is very important for both mining biological parts and understanding their natural diversity [17, 130, 133, 135].

In summary, current challenges in metagenomics that need to be addressed can be divided into two main groups: (i) the development of novel bioinformatic tools and (ii) the generation of novel molecular tools. The first group comprises the necessity of dealing with the colossal amount of information delivered from novel sequencing approaches, as previously described. Thus, it is imperative to transform the metagenomic information overload into biological understanding. The second challenge is related to the generation of novel molecular approaches such as merging metagenomic and synthetic biology to delineate novel strategies for activity-driven screening. Existing functional screening methods usually have low rates of gene target identification. Therefore, the construction of novel synthetic circuits able to detect enzymatic activities—or other target gene output—present in the cloned metagenomic fragments is essential to improve the screening efficiency of metagenomic libraries. In this manner, by combining the collective efforts of specialists for overcoming the previously described challenges, it will be possible to integrate emerging concepts and dive deeper into the universe of metagenomics expanding the current knowledge in a myriad of areas. Shedding a light on the “hidden” world of uncultured microorganisms—and its inherent biochemical treasures—shall tell us stories not only about the multitude of Microverses that surround us but also about ourselves.

Conflicts of Interest

The authors have no conflict of interest to declare.

Acknowledgments

The authors are grateful to the anonymous reviewer for her/his careful reading of our manuscript and her/his many insightful comments and suggestions. This work was supported by the National Counsel of Technological and Scientific Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico 472893/2013–0) and by Young Research Wards by the Sao Paulo State Foundation (Fundação de Amparo à Pesquisa do Estado de São Paulo, Award no. 2015/04309–1). Luana de Fátima Alves, Cauã Antunes Westmann, and Tiago Cabral Borelli are beneficiaries of Fundação de Amparo à Pesquisa do Estado de São Paulo fellowships (Award nos. 2016/06323–4, 2016/05472–6, and 2017/20818–9, resp.).

References

- [1] N. R. Pace, D. A. Stahl, D. J. Lane, and G. J. Olsen, “The analysis of natural microbial populations by ribosomal RNA sequences,” in *Advances in Microbial Ecology*, M. K. Cou, Ed., pp. 1–55, Springer, Boston, MA, USA, 1986.
- [2] J. Handelsman, M. R. Rondon, S. F. Brady, J. Clardy, and R. M. Goodman, “Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products,” *Chemistry & Biology*, vol. 5, no. 10, pp. R245–R249, 1998.
- [3] R. D. Sleator, C. Shortall, and C. Hill, “Metagenomics,” *Letters in Applied Microbiology*, vol. 47, no. 5, pp. 361–366, 2008.
- [4] J. Handelsman, “Metagenomics: application of genomics to uncultured microorganisms,” *Microbiology and Molecular Biology Reviews*, vol. 69, no. 1, pp. 195–195, 2005.
- [5] S. G. Tringe, C. von Mering, A. Kobayashi et al., “Comparative metagenomics of microbial communities,” *Science*, vol. 308, no. 5721, pp. 554–557, 2005.
- [6] J. M. Vieites, M. E. Guazzaroni, A. Beloqui, P. N. Golyshin, and M. Ferrer, “Metagenomics approaches in systems microbiology,” *FEMS Microbiology Reviews*, vol. 33, no. 1, pp. 236–255, 2009.
- [7] C. Schmeisser, H. Steele, and W. R. Streit, “Metagenomics, biotechnology with non-culturable microbes,” *Applied Microbiology and Biotechnology*, vol. 75, no. 5, pp. 955–962, 2007.
- [8] M. E. Guazzaroni, R. Silva-Rocha, and R. J. Ward, “Synthetic biology approaches to improve biocatalyst identification in metagenomic library screening,” *Microbial Biotechnology*, vol. 8, no. 1, pp. 52–64, 2015.
- [9] M. E. Guazzaroni, A. Beloqui, P. N. Golyshin, and M. Ferrer, “Metagenomics as a new technological tool to gain scientific knowledge,” *World Journal of Microbiology and Biotechnology*, vol. 25, no. 6, pp. 945–954, 2009.
- [10] S. Louca, M. F. Polz, F. Mazel et al., “Function and functional redundancy in microbial systems,” *Nature Ecology & Evolution*, vol. 2, no. 6, pp. 936–943, 2018.
- [11] M. Li, B. Wang, M. Zhang et al., “Symbiotic gut microbes modulate human metabolic phenotypes,” *Proceedings of the National Academy of Sciences*, vol. 105, no. 6, pp. 2117–2122, 2008.
- [12] M. G. I. Langille, J. Zaneveld, J. G. Caporaso et al., “Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences,” *Nature Biotechnology*, vol. 31, no. 9, pp. 814–821, 2013.
- [13] D. E. Hunt, Y. Lin, M. J. Church et al., “Relationship between abundance and specific activity of bacterioplankton in open ocean surface waters,” *Applied and Environmental Microbiology*, vol. 79, no. 1, pp. 177–184, 2013.
- [14] F. Sanger, G. M. Air, B. G. Barrell et al., “Nucleotide sequence of bacteriophage phi X174 DNA,” *Nature*, vol. 265, no. 5596, pp. 687–695, 1977.
- [15] M. Breitbart, I. Hewson, B. Felts et al., “Metagenomic analyses of an uncultured viral community from human feces,” *Journal of Bacteriology*, vol. 185, no. 20, pp. 6220–6223, 2003.
- [16] D. E. Gillespie, S. F. Brady, A. D. Bettermann et al., “Isolation of antibiotics turbomycin a and B from a metagenomic library of soil microbial DNA,” *Applied and Environmental Microbiology*, vol. 68, no. 9, pp. 4301–4306, 2002.
- [17] T. Uchiyama, T. Abe, T. Ikemura, and K. Watanabe, “Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes,” *Nature Biotechnology*, vol. 23, no. 1, pp. 88–93, 2005.
- [18] P. J. Turnbaugh, V. K. Ridaura, J. J. Faith, F. E. Rey, R. Knight, and J. I. Gordon, “The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic

- mice,” *Science Translational Medicine*, vol. 1, no. 6, p. 6ra14, 2009.
- [19] S. Sunagawa, L. P. Coelho, S. Chaffron et al., “Structure and function of the global ocean microbiome,” *Science*, vol. 348, no. 6237, article 1261359, 2015.
- [20] A. Klindworth, E. Pruesse, T. Schweer et al., “Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies,” *Nucleic Acids Research*, vol. 41, no. 1, pp. e1–11, 2013.
- [21] A. Oulas, C. Pavlouidi, P. Polymenakou et al., “Metagenomics: tools and insights for analyzing next-generation sequencing data derived from biodiversity studies,” *Bioinformatics and Biology Insights*, vol. 9, pp. BBI.S12462–BBI.S12488, 2015.
- [22] M. Kircher and J. Kelso, “High-throughput DNA sequencing – concepts and limitations,” *BioEssays*, vol. 32, no. 6, pp. 524–536, 2010.
- [23] E. Le Chatelier, MetaHIT consortium, T. Nielsen et al., “Richness of human gut microbiome correlates with metabolic markers,” *Nature*, vol. 500, no. 7464, pp. 541–546, 2013.
- [24] D. Kao, B. Roach, H. Park et al., “Fecal microbiota transplantation in the management of hepatic encephalopathy,” *Hepatology*, vol. 63, no. 1, pp. 339–340, 2016.
- [25] N. Qin, F. Yang, A. Li et al., “Alterations of the human gut microbiome in liver cirrhosis,” *Nature*, vol. 513, no. 7516, pp. 59–64, 2014.
- [26] H. K. Pedersen, V. Gudmundsdottir, H. B. Nielsen et al., “Human gut microbes impact host serum metabolome and insulin sensitivity,” *Nature*, vol. 535, no. 7612, pp. 376–381, 2016.
- [27] T. M. Schmidt, E. F. DeLong, and N. R. Pace, “Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing,” *Journal of Bacteriology*, vol. 173, no. 14, pp. 4371–4378, 1991.
- [28] F. G. Healy, R. M. Ray, H. C. Aldrich, A. C. Wilkie, L. O. Ingram, and K. T. Shanmugam, “Direct isolation of functional genes encoding cellulases from the microbial consortia in a thermophilic, anaerobic digester maintained on lignocellulose,” *Applied Microbiology and Biotechnology*, vol. 43, no. 4, pp. 667–674, 1995.
- [29] M. R. Rondon, P. R. August, A. D. Bettermann et al., “Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms,” *Applied and Environmental Microbiology*, vol. 66, no. 6, pp. 2541–2547, 2000.
- [30] A. Escobar-Zepeda, A. V.-P. de León, and A. Sanchez-Flores, “The road to metagenomics: from microbiology to DNA sequencing technologies and bioinformatics,” *Frontiers in Genetics*, vol. 6, 2015.
- [31] D. D. Roumpeka, R. J. Wallace, F. Escalettes, I. Fotheringham, and M. Watson, “A review of bioinformatics tools for bio-prospecting from metagenomic sequence data,” *Frontiers in Genetics*, vol. 8, 2017.
- [32] F. P. Breitwieser, J. Lu, and S. L. Salzberg, “A review of methods and databases for metagenomic classification and assembly,” *Briefings in Bioinformatics*, pp. 1–15, 2017.
- [33] C. R. Woese and G. E. Fox, “Phylogenetic structure of the prokaryotic domain: the primary kingdoms,” *Proceedings of the National Academy of Sciences*, vol. 74, no. 11, pp. 5088–5090, 1977.
- [34] J. C. Venter, K. Remington, J. F. Heidelberg et al., “Environmental genome shotgun sequencing of the Sargasso Sea,” *Science*, vol. 304, no. 5667, pp. 66–74, 2004.
- [35] O. Béjà, L. Aravind, E. V. Koonin et al., “Bacterial rhodopsin: evidence for a new type of phototrophy in the sea,” *Science*, vol. 289, no. 5486, pp. 1902–1906, 2000.
- [36] J. A. Klappenbach, J. M. Dunbar, and T. M. Schmidt, “rRNA operon copy number reflects ecological strategies of bacteria,” *Applied and Environmental Microbiology*, vol. 66, no. 4, pp. 1328–1333, 2000.
- [37] E. A. Elloe-Fadrosh, N. N. Ivanova, T. Woyke, and N. C. Kyrpides, “Metagenomics uncovers gaps in amplicon-based detection of microbial diversity,” *Nature Microbiology*, vol. 1, no. 4, 2016.
- [38] C. T. Brown, L. A. Hug, B. C. Thomas et al., “Unusual biology across a group comprising more than 15% of domain Bacteria,” *Nature*, vol. 523, no. 7559, pp. 208–211, 2015.
- [39] R. Gorelick and H. H. Q. Heng, “Sex reduces genetic variation: a multidisciplinary review,” *Evolution*, vol. 65, no. 4, pp. 1088–1098, 2011.
- [40] H. H. Q. Heng, “Elimination of altered karyotypes by sexual reproduction preserves species identity,” *Genome*, vol. 50, no. 5, pp. 517–524, 2007.
- [41] C. Rinke, P. Schwientek, A. Sczyrba et al., “Insights into the phylogeny and coding potential of microbial dark matter,” *Nature*, vol. 499, no. 7459, pp. 431–437, 2013.
- [42] A. Spang, J. H. Saw, S. L. Jørgensen et al., “Complex archaea that bridge the gap between prokaryotes and eukaryotes,” *Nature*, vol. 521, no. 7551, pp. 173–179, 2015.
- [43] J. L. Stein, T. L. Marsh, K. Y. Wu, H. Shizuya, and E. F. DeLong, “Characterization of uncultivated prokaryotes: isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon,” *Journal of Bacteriology*, vol. 178, no. 3, pp. 591–599, 1996.
- [44] O. Béjà, M. T. Suzuki, E. V. Koonin et al., “Construction and analysis of bacterial artificial chromosome libraries from a marine microbial assemblage,” *Environmental Microbiology*, vol. 2, no. 5, pp. 516–529, 2000.
- [45] O. Béjà, E. N. Spudich, J. L. Spudich, M. Leclerc, and E. F. DeLong, “Proteorhodopsin phototrophy in the ocean,” *Nature*, vol. 411, no. 6839, pp. 786–789, 2001.
- [46] G. W. Tyson, J. Chapman, P. Hugenholtz et al., “Community structure and metabolism through reconstruction of microbial genomes from the environment,” *Nature*, vol. 428, no. 6978, pp. 37–43, 2004.
- [47] E. A. Dinsdale, R. A. Edwards, D. Hall et al., “Functional metagenomic profiling of nine biomes,” *Nature*, vol. 452, no. 7187, pp. 629–632, 2008.
- [48] J. Lloyd-Price, A. Mahurkar, G. Rahnavard et al., “Strains, functions and dynamics in the expanded human microbiome project,” *Nature*, vol. 550, pp. 61–66, 2017.
- [49] J. A. Gilbert, R. A. Quinn, J. Debelius et al., “Microbiome-wide association studies link dynamic microbial consortia to disease,” *Nature*, vol. 535, no. 7610, pp. 94–103, 2016.
- [50] A. L. Kau, P. P. Ahern, N. W. Griffin, A. L. Goodman, and J. I. Gordon, “Human nutrition, the gut microbiome and the immune system,” *Nature*, vol. 474, no. 7351, pp. 327–336, 2011.
- [51] E. A. Archie and J. Tung, “Social behavior and the microbiome,” *Current Opinion in Behavioral Sciences*, vol. 6, pp. 28–34, 2015.

- [52] T. Kuntz and J. Gilbert, "Does the brain listen to the gut?," *eLife*, vol. 5, 2016.
- [53] A. J. Prussin and L. C. Marr, "Sources of airborne microorganisms in the built environment," *Microbiome*, vol. 3, no. 1, p. 78, 2015.
- [54] S. W. Kembel, E. Jones, J. Kline et al., "Architectural design influences the diversity and structure of the built environment microbiome," *The ISME Journal*, vol. 6, no. 8, pp. 1469–1479, 2012.
- [55] M. H. Y. Leung and P. K. H. Lee, "The roles of the outdoors and occupants in contributing to a potential pan-microbiome of the built environment: a review," *Microbiome*, vol. 4, no. 1, pp. 21–15, 2016.
- [56] T. Thomas, L. Moitinho-Silva, M. Lurgi et al., "Diversity, structure and convergent evolution of the global sponge microbiome," *Nature Communications*, vol. 7, 2016.
- [57] J. M. Maritz, S. A. Sullivan, R. J. Prill, E. Aksoy, P. Scheid, and J. M. Carlton, "Filthy lucre: a metagenomic pilot study of microbes found on circulating currency in New York City," *PLoS One*, vol. 12, no. 4, pp. e0175527–e0175516, 2017.
- [58] H. M. Bik, J. M. Maritz, A. Luong, H. Shin, M. G. Dominguez-Bello, and J. M. Carlton, "Microbial community patterns associated with automated teller machine keypads in New York City," *mSphere*, vol. 1, no. 6, pp. e00226–e00216, 2016.
- [59] R. I. Adams, A. C. Bateman, H. M. Bik, and J. F. Meadow, "Microbiota of the indoor environment: a meta-analysis," *Microbiome*, vol. 3, no. 1, p. 49, 2015.
- [60] S. G. Tringe, T. Zhang, X. Liu et al., "The airborne metagenome in an indoor urban environment," *PLoS One*, vol. 3, no. 4, article e1862, 2008.
- [61] E. Afshinnekoo, C. Meydan, S. Chowdhury et al., "Geospatial resolution of human and bacterial diversity with city-scale metagenomics," *Cell Systems*, vol. 1, no. 1, pp. 72–87, 2015.
- [62] The MetaSUB International Consortium, "The Metagenomics and Metadesign of the Subways and Urban Biomes (MetaSUB) International Consortium inaugural meeting report," *Microbiome*, vol. 4, no. 1, p. 24, 2016.
- [63] N. A. Bokulich, Z. T. Lewis, K. Boundy-Mills, and D. A. Mills, "A new perspective on microbial landscapes within food production," *Current Opinion in Biotechnology*, vol. 37, pp. 182–189, 2016.
- [64] J. M. Lang, D. A. Coil, R. Y. Neches et al., "A microbial survey of the International Space Station (ISS)," *Peer J*, vol. 5, article e4029, 2017.
- [65] National Academies of Sciences, *Engineering and MNA of ED on E and PSH and MDD on E and LSB on I and the CEB on ES and T. Microbiomes of the Built Environment*, National Academies Press, Washington, D.C. USA, 2017.
- [66] J. S. Griffin, N. Lu, N. Sangwan et al., "Microbial diversity in an intensively managed landscape is structured by landscape connectivity," *FEMS Microbiology Ecology*, vol. 93, no. 10, pp. 1–12, 2017.
- [67] G. Mhuireach, B. R. Johnson, A. E. Altrichter et al., "Urban greenness influences airborne bacterial community composition," *Science of The Total Environment*, vol. 571, pp. 680–687, 2016.
- [68] A. T. Reese, A. Savage, E. Youngsteadt et al., "Urban stress is associated with variation in microbial species composition - but not richness - in Manhattan," *The ISME Journal*, vol. 10, no. 3, pp. 751–760, 2016.
- [69] J. A. Gilbert, "How do we make indoor environments and healthcare settings healthier?," *Microbial Biotechnology*, vol. 10, no. 1, pp. 11–13, 2017.
- [70] B. Stephens, "What have we learned about the microbiomes of indoor environments?," *mSystems*, vol. 1, no. 4, 2016.
- [71] S. R. Gill, M. Pop, R. T. DeBoy et al., "Metagenomic analysis of the human distal gut microbiome," *Science*, vol. 312, no. 5778, pp. 1355–1359, 2006.
- [72] J. W. Arnold, J. Roach, and M. A. Azcarate-Peril, "Emerging technologies for gut microbiome research," *Trends in Microbiology*, vol. 24, no. 11, pp. 887–901, 2016.
- [73] The Integrative HMP, (iHMP) Research Network consortium, "The integrative human microbiome project: dynamic analysis of microbiome-host omics profiles during periods of human health and disease," *Cell Host & Microbe*, vol. 16, no. 3, pp. 276–289, 2014.
- [74] The Human Microbiome Project Consortium, "Structure, function and diversity of the healthy human microbiome," *Nature*, vol. 486, no. 7402, pp. 207–214, 2012.
- [75] MetaHIT Consortium, J. Qin, R. Li et al., "ARTICLES A human gut microbial gene catalogue established by metagenomic sequencing," *Nature*, vol. 464, no. 7285, pp. 59–65, 2010.
- [76] F. Bäckhed, R. E. Ley, J. L. Sonnenburg, D. A. Peterson, and J. I. Gordon, "Host-bacterial mutualism in the human intestine," *Science*, vol. 307, no. 5717, pp. 1915–1920, 2005.
- [77] P. C. Watts, K. R. Buley, S. Sanderson, W. Boardman, C. Ciofi, and R. Gibson, "Parthenogenesis in Komodo dragons," *Nature*, vol. 444, no. 7122, pp. 1021–1022, 2006.
- [78] R. E. Ley, P. J. Turnbaugh, S. Klein, and J. I. Gordon, "Microbial ecology: human gut microbes associated with obesity," *Nature*, vol. 444, no. 7122, pp. 1022–1023, 2006.
- [79] P. J. Turnbaugh, R. E. Ley, M. A. Mahowald, V. Magrini, E. R. Mardis, and J. I. Gordon, "An obesity-associated gut microbiome with increased capacity for energy harvest," *Nature*, vol. 444, no. 7122, pp. 1027–1031, 2006.
- [80] K. Forslund, S. Sunagawa, J. R. Kultima et al., "Country-specific antibiotic use practices impact the human gut resistome," *Genome*, vol. 23, no. 7, pp. 1163–1169, 2013.
- [81] K. Forslund, S. Sunagawa, L. P. Coelho, and P. Bork, "Metagenomic insights into the human gut resistome and the forces that shape it," *BioEssays*, vol. 36, no. 3, pp. 316–329, 2014.
- [82] T. S. Ghosh, S. S. Gupta, G. B. Nair, and S. S. Mande, "In silico analysis of antibiotic resistance genes in the gut microflora of individuals from diverse geographies and age-groups," *PLoS One*, vol. 8, no. 12, article e83823, 2013.
- [83] F. Raymond, A. A. Ouameur, M. Déraspe et al., "The initial state of the human gut microbiome determines its reshaping by antibiotics," *The ISME Journal*, vol. 10, no. 3, pp. 707–720, 2016.
- [84] R. K. Singh, H.-W. Chang, D. Yan et al., "Influence of diet on the gut microbiome and implications for human health," *Journal of Translational Medicine*, vol. 15, no. 1, pp. 73–17, 2017.
- [85] A. Lundin, C. M. Bok, L. Aronsson et al., "Gut flora, Toll-like receptors and nuclear receptors: a tripartite communication that tunes innate immunity in large intestine," *Cellular Microbiology*, vol. 10, no. 5, pp. 1093–1103, 2008.
- [86] Y. K. Lee and S. K. Mazmanian, "Has the microbiota played a critical role in the evolution of the adaptive immune system?," *Science*, vol. 330, no. 6012, pp. 1768–1773, 2010.

- [87] Y. Belkaid and T. W. Hand, "Role of the microbiota in immunity and inflammation," *Cell*, vol. 157, no. 1, pp. 121–141, 2014.
- [88] M. C. Noverr and G. B. Huffnagle, "Does the microbiota regulate immune responses outside the gut?," *Trends in Microbiology*, vol. 12, no. 12, pp. 562–568, 2004.
- [89] J. R. Allegretti, J. R. Korzenik, and M. J. Hamilton, "Fecal microbiota transplantation via colonoscopy for recurrent *C. difficile* infection," *Journal of Visualized Experiments*, no. 94, article e52154, 2014.
- [90] S. Gupta, E. Allen-Vercoe, and E. O. Petrof, "Fecal microbiota transplantation: in perspective," *Therapeutic Advances in Gastroenterology*, vol. 9, no. 2, pp. 229–239, 2015.
- [91] P. Seksik, L. Rigottier-Gois, G. Gramet et al., "Alterations of the dominant faecal bacterial groups in patients with Crohn's disease of the colon," *Gut*, vol. 52, no. 2, pp. 237–242, 2003.
- [92] J. U. Scher, A. Sczesnak, R. S. Longman et al., "Expansion of intestinal *Prevotella copri* correlates with enhanced susceptibility to arthritis," *eLife*, vol. 2, 2013.
- [93] L. Zhao, "The gut microbiota and obesity: from correlation to causality," *Nature Reviews Microbiology*, vol. 11, no. 9, pp. 639–647, 2013.
- [94] C. Menni, M. A. Jackson, T. Pallister, C. J. Steves, T. D. Spector, and A. M. Valdes, "Gut microbiome diversity and high-fibre intake are related to lower long-term weight gain," *International Journal of Obesity*, vol. 41, no. 7, pp. 1099–1105, 2017.
- [95] M. Knip and H. Siljander, "The role of the intestinal microbiota in type 1 diabetes mellitus," *Nature Reviews Endocrinology*, vol. 12, no. 3, pp. 154–167, 2016.
- [96] J. L. Han and H. L. Lin, "Intestinal microbiota and type 2 diabetes: from mechanism insights to therapeutic perspective," *World Journal of Gastroenterology*, vol. 20, no. 47, pp. 17737–17745, 2014.
- [97] J. Qin, Y. Li, Z. Cai et al., "A metagenome-wide association study of gut microbiota in type 2 diabetes," *Nature*, vol. 490, no. 7418, pp. 55–60, 2012.
- [98] C. M. Velicer, S. R. Heckbert, J. W. Lampe, J. D. Potter, C. A. Robertson, and S. H. Taplin, "Antibiotic use in relation to the risk of breast cancer," *JAMA*, vol. 291, no. 7, pp. 827–835, 2004.
- [99] H. T. Sørensen, M. V. Skriver, S. Friis, J. K. McLaughlin, W. J. Blot, and J. A. Baron, "Use of antibiotics and risk of breast cancer: a population-based case-control study," *British Journal of Cancer*, vol. 92, no. 3, pp. 594–596, 2005.
- [100] Z. Wang, E. Klipfell, B. J. Bennett et al., "Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease," *Nature*, vol. 472, no. 7341, pp. 57–63, 2011.
- [101] J. Yu, Q. Feng, S. H. Wong et al., "Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer," *Gut*, vol. 66, no. 1, pp. 70–78, 2016.
- [102] Q. Liang, J. Chiu, Y. Chen et al., "Fecal bacteria act as novel biomarkers for noninvasive diagnosis of colorectal cancer," *Clinical Cancer Research*, vol. 23, no. 8, pp. 2061–2070, 2017.
- [103] V. Pascal, M. Pozuelo, N. Borrueal et al., "A microbial signature for Crohn's disease," *Gut*, vol. 66, no. 5, pp. 813–822, 2017.
- [104] R. Loomba, V. Seguritan, W. Li et al., "Gut microbiome-based metagenomic signature for non-invasive detection of advanced fibrosis in human nonalcoholic fatty liver disease," *Cell Metabolism*, vol. 25, no. 5, pp. 1054–1062.e5, 2017.
- [105] S. J. Young, L. Deng, N. Li et al., "Improving microbial fitness in the mammalian gut by in vivo temporal functional metagenomics," *Molecular Systems Biology*, vol. 11, no. 3, p. 788, 2015.
- [106] L. Fernández-Arrojo, M. E. Guazzaroni, N. López-Cortés, A. Beloqui, and M. Ferrer, "Metagenomic era for biocatalyst identification," *Current Opinion in Biotechnology*, vol. 21, no. 6, pp. 725–733, 2010.
- [107] M. Ferrer, A. Ghazi, A. Beloqui et al., "Functional metagenomics unveils a multifunctional glycosyl hydrolase from the family 43 catalysing the breakdown of plant polymers in the calf rumen," *PLoS One*, vol. 7, no. 6, article e38134, 2012.
- [108] M. V. Del Pozo, L. Fernández-Arrojo, J. Gil-Martínez et al., "Microbial β -glucosidases from cow rumen metagenome enhance the saccharification of lignocellulose in combination with commercial cellulase cocktail," *Biotechnology for Biofuels*, vol. 5, no. 1, pp. 73–13, 2012.
- [109] C. Thompson, W. Beys-Da-Silva, L. Santi et al., "A potential source for cellulolytic enzyme discovery and environmental aspects revealed through metagenomics of Brazilian mangroves," *AMB Express*, vol. 3, no. 1, p. 65, 2013.
- [110] C. Yang, Y. Xia, H. Qu et al., "Discovery of new cellulases from the metagenome by a metagenomics-guided strategy," *Biotechnology for Biofuels*, vol. 9, no. 1, p. 138, 2016.
- [111] S. Courtois, C. M. Cappellano, M. Ball et al., "Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products," *Applied and Environmental Microbiology*, vol. 69, no. 1, pp. 49–55, 2003.
- [112] M. Ferrer, O. V. Golyshina, T. N. Chernikova et al., "Novel hydrolase diversity retrieved from a metagenome library of bovine rumen microflora," *Environmental Microbiology*, vol. 7, no. 12, pp. 1996–2010, 2005.
- [113] F. O. Aylward, K. E. Burnum, J. J. Scott et al., "Metagenomic and metaproteomic insights into bacterial communities in leaf-cutter ant fungus gardens," *The ISME Journal*, vol. 6, no. 9, pp. 1688–1701, 2012.
- [114] G. Suen, J. J. Scott, F. O. Aylward, and C. R. Currie, "The microbiome of leaf-cutter ant fungus gardens," in *Handbook of Molecular Microbial Ecology II: Metagenomics in Different Habitats*, F. J. de Bruijn, Ed., John Wiley & Sons, Inc., Hoboken, NJ, USA, 2011.
- [115] F. O. Aylward, K. E. Burnum-Johnson, S. G. Tringe et al., "Leucoagaricus gongylophorus produces diverse enzymes for the degradation of recalcitrant plant polymers in leaf-cutter ant fungus gardens," *Applied and Environmental Microbiology*, vol. 79, no. 12, pp. 3770–3778, 2013.
- [116] S. Mirete, M. R. Mora-Ruiz, M. Lamprecht-Grandío, C. G. de Figueras, R. Rosselló-Móra, and J. E. González-Pastor, "Salt resistance genes revealed by functional metagenomics from brines and moderate-salinity rhizosphere within a hypersaline environment," *Frontiers in Microbiology*, vol. 6, 2015.
- [117] S. Mirete, C. G. De Figueras, and J. E. González-Pastor, "Novel nickel resistance genes from the rhizosphere metagenome of plants adapted to acid mine drainage," *Applied and Environmental Microbiology*, vol. 73, no. 19, pp. 6001–6011, 2007.
- [118] M. E. Guazzaroni, V. Morgante, S. Mirete, and J. E. González-Pastor, "Novel acid resistance genes from the metagenome of the Tinto River, an extremely acidic

- environment,” *Environmental Microbiology*, vol. 15, no. 4, pp. 1088–1102, 2013.
- [119] D. R. Garza, M. C. van Verk, M. A. Huynen, and B. E. Dutilh, “Towards predicting the environmental metabolome from metagenomics with a mechanistic model,” *Nature Microbiology*, vol. 3, no. 4, pp. 456–460, 2018.
- [120] A. P. Alivisatos, M. J. Blaser, E. L. Brodie et al., “A unified initiative to harness Earth’s microbiomes,” *Science*, vol. 350, no. 6260, pp. 507–508, 2015.
- [121] L. R. Thompson, J. G. Sanders, D. McDonald et al., “A communal catalogue reveals Earth’s multiscale microbial diversity,” *Nature*, vol. 551, no. 7681, pp. 457–463, 2017.
- [122] C. Schmidt, “Living in a microbial world,” *Nature Biotechnology*, vol. 35, no. 5, pp. 401–403, 2017.
- [123] R. Sender, S. Fuchs, and R. Milo, “Revised estimates for the number of human and bacteria cells in the body,” *PLoS Biology*, vol. 14, no. 8, pp. e1002533–e1002514, 2016.
- [124] J. F. Meadow, A. E. Altrichter, A. C. Bateman et al., “Humans differ in their personal microbial cloud,” *Peer J*, vol. 3, article e1258, 2015.
- [125] J. W. Craig, F.-Y. Chang, J. H. Kim, S. C. Obiajulu, and S. F. Brady, “Expanding small-molecule functional metagenomics through parallel screening of broad-host-range cosmid environmental DNA libraries in diverse proteobacteria,” *Applied and Environmental Microbiology*, vol. 76, no. 5, pp. 1633–1641, 2010.
- [126] H. Nagayama, T. Sugawara, R. Endo et al., “Isolation of oxygenase genes for indigo-forming activity from an artificially polluted soil metagenome by functional screening using *Pseudomonas putida* strains as hosts,” *Applied Microbiology and Biotechnology*, vol. 99, no. 10, pp. 4453–4470, 2015.
- [127] B. Leis, A. Angelov, M. Mientus et al., “Identification of novel esterase-active enzymes from hot environments by use of the host bacterium *Thermus thermophilus*,” *Frontiers in Microbiology*, vol. 6, 2015.
- [128] A. Angelov, M. Mientus, S. Liebl, and W. Liebl, “A two-host fosmid system for functional screening of (meta) genomic libraries from extreme thermophiles,” *Systematic and Applied Microbiology*, vol. 32, no. 3, pp. 177–185, 2009.
- [129] T. Aakvik, K. F. Degnes, R. Dahlsrud et al., “A plasmid RK2-based broad-host-range cloning vector useful for transfer of metagenomic libraries to a variety of bacterial species,” *FEMS Microbiology Letters*, vol. 296, no. 2, pp. 149–158, 2009.
- [130] N. I. Johns, A. L. C. Gomes, S. S. Yim et al., “Metagenomic mining of regulatory elements enables programmable species-selective gene expression,” *Nature Methods*, vol. 15, no. 5, pp. 323–329, 2018.
- [131] R. Silva-Rocha, E. Martínez-García, B. Calles et al., “The standard European vector architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes,” *Nucleic Acids Research*, vol. 41, no. D1, pp. D666–D675, 2013.
- [132] L. D. F. Alves, R. Silva-Rocha, and M.-E. Guazzaroni, “Enhancing metagenomic approaches through synthetic biology,” in *Functional Metagenomics: Tools and Applications María-Eugenia Guazzaroni, T. C. Charles, M. R. Liles, and A. Sessitsch, Eds.*, pp. 1–14, Springer International Publishing, Berlin, 1st edition, 2017.
- [133] T. Uchiyama and K. Miyazaki, “Product-induced gene expression, a product-responsive reporter assay used to screen metagenomic libraries for enzyme-encoding genes,” *Applied and Environmental Microbiology*, vol. 76, no. 21, pp. 7029–7035, 2010.
- [134] H. J. Genee, A. P. Bali, S. D. Petersen et al., “Functional mining of transporters using synthetic selections,” *Nature Chemical Biology*, vol. 12, no. 12, pp. 1015–1022, 2016.
- [135] C. A. Westmann, L. d. F. Alves, R. Silva-Rocha, and M.-E. Guazzaroni, “Mining novel constitutive promoter elements in soil metagenomic libraries in *Escherichia coli*,” *Frontiers in Microbiology*, vol. 9, 2018.
- [136] Z. Fang, T. Li, Q. Wang et al., “A bacterial laccase from marine microbial metagenome exhibiting chloride tolerance and dye decolorization ability,” *Applied Microbiology and Biotechnology*, vol. 89, no. 4, pp. 1103–1110, 2011.
- [137] A. Ono, R. Miyazaki, M. Sota, Y. Ohtsubo, Y. Nagata, and M. Tsuda, “Isolation and characterization of naphthalene-catabolic genes and plasmids from oil-contaminated soil by using two cultivation-independent approaches,” *Applied Microbiology and Biotechnology*, vol. 74, no. 2, pp. 501–510, 2007.
- [138] S. Sulaiman, S. Yamato, E. Kanaya et al., “Isolation of a novel cutinase homolog with polyethylene terephthalate-degrading activity from leaf-branch compost by using a metagenomic approach,” *Applied and Environmental Microbiology*, vol. 78, no. 5, pp. 1556–1562, 2012.
- [139] C. C. Silva, H. Hayden, T. Sawbridge et al., “Identification of genes and pathways related to phenol degradation in metagenomic libraries from petroleum refinery wastewater,” *PLoS One*, vol. 8, no. 4, pp. e61811–e61811, 2013.
- [140] A. Tchigvintsev, H. Tran, A. Popovic et al., “The environment shapes microbial enzymes: five cold-active and salt-resistant carboxylesterases from marine metagenomes,” *Applied Microbiology and Biotechnology*, vol. 99, no. 5, pp. 2165–2178, 2015.
- [141] H. C. Rees, S. Grant, B. Jones, W. D. Grant, and S. Heaphy, “Detecting cellulase and esterase enzyme activities encoded by novel genes present in environmental DNA libraries,” *Extremophiles*, vol. 7, no. 5, pp. 415–421, 2003.
- [142] R. Garg, R. Srivastava, V. Brahma, L. Verma, S. Karthikeyan, and G. Sahni, “Biochemical and structural characterization of a novel halotolerant cellulase from soil metagenome,” *Scientific Reports*, vol. 6, no. 1, pp. 1–15, 2016.
- [143] C. Schröder, S. Elleuche, S. Blank, and G. Antranikian, “Characterization of a heat-active archaeal β -glucosidase from a hydrothermal spring metagenome,” *Enzyme and Microbial Technology*, vol. 57, pp. 48–54, 2014.
- [144] S. Thies, S. C. Rausch, F. Kovacic et al., “Metagenomic discovery of novel enzymes and biosurfactants in a slaughterhouse biofilm microbial community,” *Scientific Reports*, vol. 6, no. 1, pp. 1–12, 2016.
- [145] V. Morgante, S. Mirete, C. G. de Figueras, M. Postigo Cacho, and J. E. González-Pastor, “Exploring the diversity of arsenic resistance genes from acid mine drainage microorganisms,” *Environmental Microbiology*, vol. 17, no. 6, pp. 1910–1925, 2015.
- [146] Y. Wang, Y. Chen, Q. Zhou et al., “A culture-independent approach to unravel uncultured bacteria and functional genes in a complex microbial community,” *PLoS One*, vol. 7, no. 10, article e47530, 2012.
- [147] A. Zaprasia, Y. J. Liu, S. J. Liu, H. L. Drake, and M. A. Horn, “Abundance of novel and diverse *tfdA*-like genes, encoding putative phenoxyalkanoic acid herbicide-degrading

- dioxygenases, in soil," *Applied and Environmental Microbiology*, vol. 76, no. 1, pp. 119–128, 2009.
- [148] C. M. Rath, B. Janto, J. Earl et al., "Meta-omic characterization of the marine invertebrate microbial consortium that produces the chemotherapeutic natural product ET-743," *ACS Chemical Biology*, vol. 6, no. 11, pp. 1244–1256, 2011.
- [149] M. N. IA, C. L. Tiong, C. Minor et al., "Expression and isolation of antimicrobial small molecules from soil DNA libraries," *Journal of Molecular Microbiology and Biotechnology*, vol. 3, no. 2, pp. 301–308, 2001.