Metagenomics: An Effective Approach for Exploring Microbial Diversity and Functions

Nguyen Nhat Nam [1],[†], Hoang Dang Khoa Do [2],[†], Kieu The Loan Trinh [3],*, and Nae Yoon Lee [3]

1 Biotechnology Center, School of Agriculture and Aquaculture, Tra Vinh University, Tra Vinh City 87000, Vietnam
2 NTT Hi-Tech Institute, Nguyen Tat Thanh University, Ward 13, District 04, Ho Chi Minh City 72820, Vietnam
3 Department of BioNano Technology, Gachon University 1342 Seongnam-daero, Sujeong-gu, Seongnam-si 13120, Republic of Korea; nylee@gachon.ac.kr
* Correspondence: tktloan@gmail.com
† These authors contributed equally to this work.

Abstract: Various fields have been identified in the "omics" era, such as genomics, proteomics, transcriptomics, metabolomics, phenomics, and metagenomics. Among these, metagenomics has enabled a significant increase in discoveries related to the microbial world. Newly discovered microbiomes in different ecologies provide meaningful information on the diversity and functions of microorganisms on the Earth. Therefore, the results of metagenomic studies have enabled new microbe-based applications in human health, agriculture, and the food industry, among others. This review summarizes the fundamental procedures on recent advances in bioinformatic tools. It also explores up-to-date applications of metagenomics in human health, food study, plant research, environmental sciences, and other fields. Finally, metagenomics is a powerful tool for studying the microbial world, and it still has numerous applications that are currently hidden and awaiting discovery. Therefore, this review also discusses the future perspectives of metagenomics.

Keywords: metagenomics; next-generation sequencing; microbial diversity; food safety

1. Introduction

The microbial world was first discovered by Leeuwenhoek using his signature invention, the microscope [1]. The discovery of microbes triggered extensive research aiming to develop methods that could be used to culture different microorganisms. The first culture method was developed by Robert Koch and was called the solid culture media method [2]. Since then, various bacteria have been identified and explored in terms of classification, biological applications, and evolution. The taxonomic status of the microbiome has changed dramatically due to the utilization of 16S ribosomal RNA (rRNA) sequences, which has resulted in the recognition of the archaea group [3]. The 16S rRNA-based phylogenetic marker has been proposed as a crucial tool in taxonomic studies of microorganisms [4], and 16S rRNA continues to be an effective tool in microbial research today [5].

However, not all molecular studies investigating microorganisms have relied solely on 16S rRNA. For example, Handelsman et al. used genomic fragments in environmental samples to clone *E.coli* and explore new mechanisms as well as antibiotic features [6]. Consequently, they proposed the term "metagenome", which refers to a collection of genomes in the samples for studying cloning and functional analyses. The advent of sequencing techniques has accelerated the development of microbial studies based on the 16S rRNA and whole genomes. The term "metagenomics" has been used to describe studies examining the genomic data of microorganisms, and it can be divided into amplicon and shotgun metagenomics. Amplicon metagenomics studies typically explore microbial diversity, while shotgun metagenomics research is mainly focused on mining functional genes and metabolisms [7,8]. Specifically, targeted gene-based metagenomics is used to obtain portions of each of the microbes in environmental samples, such as soil, water,

air, and ice samples [9,10]. The results of shotgun metagenomics also include taxonomic diversity, as in amplicon metagenomics, and they reveal different metabolisms in the microbial community that can be inferred from whole genome sequences [8]. The microbial diversity that has been found in different environments has motivated further studies into the applications of microbial presence in agriculture, food, and pharmacy industries, in addition to human health [11–15]. Shotgun-metagenomics-associated genomic studies have shown a high potential for identifying new bacteria and viruses as well as predicting the major metabolisms in the surveyed environments [16–18]. The emergence of metagenomics has provided new paths for exploring the microbial world, which still contains hidden areas.

Previous reviews have summarized various aspects of metagenomics, including reviews of the methods and pipelines in metagenomics [19–21], as well as reviews of different applications of metagenomics in various fields [22–24]. In the present review, we summarized the principles of metagenomics using the latest approaches. We also reviewed recent metagenomic studies on human health, agriculture, food industry, environmental sciences, etc., along with their applications. Although previous metagenomic studies have reported significant results, there are still potential applications of metagenomics awaiting discovery. Therefore, we also discuss future perspectives on metagenomics (Figure 1).
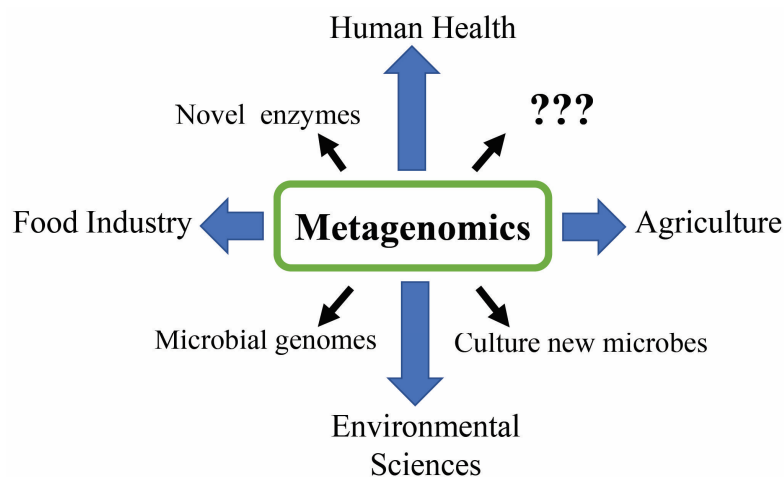

Figure 1. Contributions of metagenomics to various fields. The question marks represent the undiscovered aspects of metagenomics.

2. Fundamentals of Metagenomics

Metagenomics is the study of genomic and genetic data inferred from environmental and clinical samples. Metagenomics can be divided into two groups based on the types of data used: amplicon or targeted gene data and shotgun or untargeted gene data inferred from amplicon and shotgun sequencing, respectively (Figure 2). The data of amplicon metagenomics are amplified sequences of marker genes that include 16S/18S/26S rRNA and intergenic transcribed spacer (ITS) [20,25]. Meanwhile, the data of shotgun metagenomics include all DNA sequences in the samples. Previously, Zhang et al. attempted to classify metagenomics into functional and sequencing metagenomics [26]. In that classification, the field of studies on the discovery of new functional genes and related bioactive substances was called functional metagenomics. Meanwhile, sequencing metagenomics was used to explore the diversity of the microbial community. On the other hand, Breitwieser et al. used the term "metataxonomics" for amplicon sequencing data and "metagenomics" for shotgun sequencing data [27]. Although different terminologies have been proposed for metagenomic classification, metagenomic studies are based on two types of data: amplicon and shotgun data. Therefore, in this study, we used the terms "amplicon metagenomics" and "shotgun metagenomics" for amplicon sequencing and shotgun sequencing data, respectively.
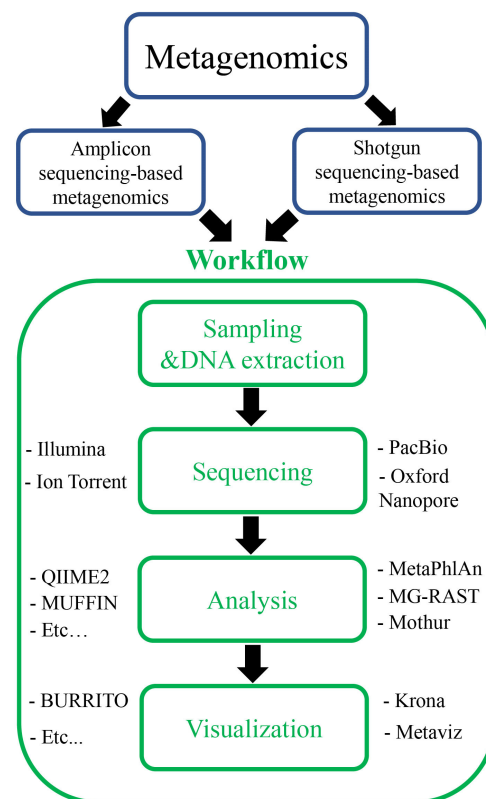
Figure 2. Classification (blue rectangles) and workflow (green rectangles) of metagenomics using different sequencing platforms and bioinformatics tools.

Although the data are different, the process of metagenomics consists of four main steps: sampling and DNA extraction, sequencing, analysis, and visualization (Figure 2). Wajid et al. compared these steps to the process of composing music, in an analogy intended to make the concept of metagenomic research more understandable to an unfamiliar readership [28]. Metagenomic studies often contain a continuous sequence of steps, in which the previous steps decide the outcomes of downstream steps. The samples for metagenomics research are collected directly from the field, so care must be taken care to prevent contamination from other sources. DNA extraction should also be conducted carefully to limit any impurities from the host DNA. In previous studies, specific protocols were developed for different samples such as human fecal samples, tropical soils, and plant tissues [29–31]. Meanwhile, commercial kits for metagenomics have also been developed such as FastDNA™Spin Kit for Soil (MP Biomedicals, Irvine, CA, USA), FavorPrep™ Soil DNA Isolation Mini Kit (Favorgen Biotech, Taiwan, China), MagAttract PowerSoil DNA KF Kit (Qiagen, Redwood, CA, USA), PureLink MicrobiomeTM Purification Kit (ThermoFisher Scientific, Waltham, MA, USA), and ZymoBIOMICS 96 Magbead DNA Kit (ZymoBIOMICS, Irvine, CA, USA). The availability of different kits and protocols triggered comparison studies examining the effectiveness of those methods for DNA extraction [32–35]. The outcomes of those studies suggested that the DNA extraction efficiency and selection of extraction protocols depended on the sample types used. Therefore, suitable protocols or kits should be considered and applied to obtain good results from metagenomic studies.

The data for metagenomic studies generally originate from DNA sequences in different environments, and these are generated using next-generation sequencing methods such as Illumina, PacBio, and Oxford Nanopore Technologies (ONT) [36]. In particular, the Illumina platform resulted in a read length of up to 300 bp, while PacBio and ONT platforms can yield long reads over 1000 bp in length. In contrast to shotgun metagenomics, in which the DNA of sufficient yield can be used for sequencing immediately, the DNA samples for amplicon metagenomics should be amplified with specific primers for targeted genes

such as 16S/18S/26S rRNA and ITS [7]. In metagenomic studies of bacteria and archaea, 16S rRNA is commonly used. For fungal and eukaryotic diversity, the small subunit (SSU) rRNA (18S), large subunit (LSU) rRNA (26S), and intergenic transcribed spacer (ITS) have been used [37–39]. In amplicon metagenomics, the lengths of the target genes (i.e., 16S, 18S, 26S rRNA, and ITS) are always greater than the sequencing outcomes of the Illumina platform; therefore, partial sequences of target genes have been amplified and used for further analyses. Previous studies have revealed conserved areas and the effectiveness of variable regions in the 16S rRNA [40,41]. For example, primer pairs have been designed based on the conserved regions and can be applied as universal primers for 16S rRNA [40]. Nine hypervariable regions, named V1 to V9, have been found to have specific properties [41]. Various regions (i.e., V1–V2, V3–V4, and V3–V5 regions) have been used to explore bacterial communities at familial, genus, and species levels [42]. Therefore, the selection of regions in the target gene should be based on the particular aims of a given study. The limitations of amplicon metagenomics can be solved using shotgun metagenomics, in which all DNA fragments are sequenced, and the entire length of the target gene can be recovered. However, the assembly process for short reads to complete the whole genes still remains a challenge because the target genes are universal and contain highly conserved regions. The long-read sequencing technique (i.e., PacBio and ONT) can overcome the limitation of short-read sequencing. However, the long-read sequencing platform approach is limited by its high cost and unstable sequencing quality. Oxford Nanopore Technologies has developed a new version for long-read sequencing techniques that produces higher-quality sequencing data (https://nanoporetech.com/q20plus-chemistry, accessed on 8 March 2023). This could reduce the cost and boost long-read-based metagenomic studies in the future.

After the sequencing step, the metagenomic data are processed through the following steps: (1) quality control of reads, (2) assembly/binning, (3) taxonomic/functional profiling, and (4) data visualization. In the first step of quality control, low-quality reads are removed from the data. Then, the remaining reads are assembled to make contigs or mapped to reference genomes. The assembly/binning results are used to predict taxonomic classification and functional mechanisms through comparison to different databases. Finally, all outcomes are visualized to provide details into the microbial composition or potential functions on the microbes. In addition to the emergence of metagenomic studies, different bioinformatic tools have been developed to effectively assist in the analysis process [21]. Wajid et al. have summarized the tools for each analysis step, including 86 tools for quality control, 48 software for assembly, 13 tools for binning, 69 tools for taxonomic classification, 27 tools for gene and functional prediction, 12 tools for metabolic profiling, 15 tools for data visualization, and 4 and 15 databases for microbial taxonomy and functional profile, respectively [28]. Different databases have previously been built for metagenomics including MetaGeneBank for human fecal specimens, Ani-malMetagenome DB, Marine Metagenomics Portal (https://mmp2.sfb.uit.no/, accessed on 8 March 2023), MGnify, TerrestrialMetagenomeDB, and MPD (a pathogen genome and metagenome database) [43–47]. In amplicon metagenomics, the main output is microbial composition, so the tools were built to clean, cluster, and quantify data, such as VSEARCH, DADA2, and Deblur [48–50]. Specifically, DADA2 can identify exact amplicon sequence variants and produce fewer false positive sequence variants than other methods [49]. Similarly, Deblur can denoise sequences and be applied to large datasets [48]. However, only single-end reads can be used as input data for Deblur. Moreover, taxonomic classification is an important output of amplicon metagenomics, which can be conducted quickly using Kraken 2 [51]. For shotgun metagenomics that are mainly focused on functional analysis, databases play an important role, and KEGG (Kyoto Encyclopedia of Genes and Genomes) and COG (Clusters of Orthologous Groups of proteins) databases have been built [52,53]. Additionally, the tools for functional profiling and prediction were developed, including PICRUSt2, MEGAN, GeneMark-HM, and Prokka [54–57]. Another database, namely Functional Annotation of Prokaryotic Taxa (FAPROTAX), contains software that can be used to convert taxonomic

profiles to putative functional profiles [58]. A previous study showed that FAPROTAX is a promising tool for predicting the function of bacteria in soil samples [59]. Similarly, FAPROTAX analysis of soil microbiota revealed a correlation between functional groups and physicochemical properties in mangrove soil [60]. Although various tools have been developed for metagenomic analysis, there are still challenges that arise during the analysis process, which were critically discussed by Breitwieser et al. [27]. Aside from the development of new bioinformatic tools, some previous tools have been upgraded with more effective results (Table 1).

Table 1. List of recently upgraded tools and platform for metagenomic analysis.

Tool/Platform	Functions	Year/Reference
MetaPhlAn 4	Comprehensive metagenomic taxonomic profiling	2023/[61]
HUMAnN 3	Efficient and accurate functional profiling	2021/[62]
StrainPhlAn 3/PanPhlAn 3	Nucleotide- and gene-variant-based strain profiling	2021/[62]
PhyloPhlAn 3	Phylogenetic placement and putative taxonomic assignment	2021/[62]
QIIME2	Analysis and visualization of amplicon metagenomic data	2019/[63]
DIAMOND + MEGAN	Taxonomic and functional analysis of short and long metagenomic sequence data	2021/[64]
MEGAN6 and MeganServer	Taxonomic and functional analyses and visualization of metagenomic data	2023/[65]
Kraken 2	Fast taxonomic classification	2019/[51]

For example, MetaPhlAn 4 employed a database of more than one million prokaryotic reference genomes to achieve comprehensive metagenomic taxonomic profiling [61]. HUMAnN 3, StrainPhlAn 3, PanPhlAn 3, and PhyloPhlAn 3 exhibit effectiveness in the strain-level, phylogenetic, taxonomic, and functional profiling of microbial communities [62]. Further, QIIME2 is a good bioinformatics platform for analyzing and visualizing metagenomic data [63,66,67]. Recently, the combination of DIAMOND + MEGAN and the release of MeganServer have been shown to provide a user-friendly and effective platform for exploring the taxonomic and functional analysis of short and long metagenomic sequence data [64,65]. For food-based metagenomics, Kobus et al. introduced a novel computational method, called MetaCache, which could be divided into AFS-MetaCache (based on C++) and MetaCacheSpark (based on Apache Spark), and which exhibited fast running, low false-positive rates, and high quantification accuracy features [68]. Moreover, ARG-ANNOT (Antibiotic Resistance Gene-ANNOTation) and DeepARG, which was the result of a machine-learning methodology, could detect the antibiotic resistance features of microbes in the raw reads of metagenomic data from fermented food [69]. Tools have also been created for illustrating the results of metagenomic studies, such as Metaviz, Krona, and BURRITO, and these provide interactive illustrations of metagenomic results [70–72].

The availability of various tools has led to the development of pipelines for metagenomic analyses [73]. Garfias-Gallegos et al. created an effective pipeline as the first step toward quality control in the final visualization of the metagenomic data [74]. Similarly, the Omnibus metagenome-wide association study with robustness (OMARU) pipeline was developed to explore the relationship between microbiomes and disease pathophysiology [75]. Previously, Navgire et al. summarized various pipelines for metagenomic studies of crops [19]; these pipelines were applied to different plants, such as rice, sugarcane, peanut, and wheat, and they could serve as a model for further studies in different areas. Another advancement in metagenomics is the availability of web-based tools that allow users to upload data and wait for outcomes; this reduces the time and cost for scientists who do not have enough facilities (i.e., supercomputers) for analysis in their lab. Some popular websites for metagenomics are MG-RAST, EBI MetaGenomics, IMG/M, and EDGE platform [76–79]. In these web-based platforms, the analysis pipeline commonly consists of quality control of reads, assembly, function prediction, taxonomic prediction, and the visualization of results steps using various tools and databases. Users can simply upload the metagenomic data to the platforms and select appropriate options for research purposes. The continued advent of improved technology and the availability of bioinformatic tools

have allowed for metagenomics studies to be conducted effectively and easily; however, suitable strategies inferred from the aims of the studies should be considered to obtain the best outcomes.

3. Metagenomics in the Food Industry

Among the current types of foods that are eaten often, fermented foods such as kimchi, yogurt, kefir, kombucha, tempeh, and sauerkraut rely on the presence of different microbes to achieve specific flavors [80,81]. Previously, microbial composition in foods has been checked using culture methods, which are limited in their ability to find unculturable bacteria [82]. The availability of a metagenomic approach has provided a powerful tool for exploring microbial communities in food [13,83]. For example, the bacteria in biofilms, including *Pseudomonas* spp., *Acinetobacter* spp., *Leuconostoc* spp., *Lactobacillus* spp., and *Streptococcus* spp., in food processing plants have been characterized based on 16S rRNA sequencing data [84]. In addition to amplicon metagenomics, functional metagenomic studies examining food have revealed potential applications in the food industry, such as the identification of novel enzymes for processing foods [85]. In the field of fermented foods, metagenomic analysis has resulted in new knowledge that is used for exploring, ensuring, and improving food quality [86,87]. For instance, different bacterial compositions in fermented bamboo shoots (Tuaither), soybeans (Bekang), and pork fat (Sa-um) have been surveyed using a metagenomic approach [88]. The most abundant types of bacteria in Tuaither, Bekang, and Sa-um were *Lactobacillus*, *Staphylococcus*, and *Clostridium*, respectively. These findings have revealed the correlation between the types of ingredients used and the abundance of bacteria in fermented food. Aside from indicating the diversity of microbial communities, metagenomics has enabled the differentiation of the metabolic pathways of biogenic amines caused by various microbes during the fermentation of *Brassica juncea* [89]. The connection between the predominance of microbes and metabolism was also investigated in Yucha, which is a fermented food composed of fresh fish and cooked rice that is popular in China [90]. For traditional foods such as kimchi, yogurt, kefir, kombucha, tempeh, and sauerkraut, which were developed thousands of years ago based on the experiences of our ancestors, metagenomic studies shed light on the factors (i.e., microbes and enzymes) and process involved in achieving the distinct flavors of these foods. For instance, one study found that three genera of *Leuconostoc*, *Lactobacillus*, and *Weissella* were dominant in kimchi [91]. In addition to bacteria, the existence of bacteriophages was recorded, suggesting an influence on the microbial community during fermentation. Another study on kimchi at the industrial scale revealed the presence of *Leuconostoc mesenteroides*, *Lactobacillus sakei*, *Lactobacillus plantarum*, and *Weissella koreensis* [92]. Another larger study based on 88 kimchi samples prepared at different locations, in different seasons, with various ingredients, and by several preparation methods revealed that the bacterial communities in kimchi were easily affected by many factors, while location did not have a significant effect [93]. These results have contributed to explaining the different tastes of fermented foods originating from households and industries. The outcomes of metagenomics have allowed for further research into applications related to the bacterial communities in kimchi [94,95]. Park compared the conventional conditions and CO_2-rich environments during kimchi fermentation [94]; the results showed the effectiveness of CO_2 addition on the quality, metabolisms, and alteration of microbes in kimchi products. Another study on treating kimchi with light-emitting diodes (LEDs) indicated that different wavelengths of the LED source could alter the microbial composition [95]. Consequently, the metabolomic pathways also change. Such results have exhibited the potential of using LEDs to control the quality and create new tastes in fermented foods in the future. Similarly, metagenomics has been used to sheds lights on the microbial world of other fermented foods, including yogurt, kefir, kombucha, tempeh, and sauerkraut [96–104]. In addition to exploring the role of microbes in different fermented foods, the metagenomic approach allows to identify the functional roles of bacteria during food production. Consequently, producers can alter the process to optimize product quality and reduce food waste [87].

The ability of metagenomics to provide the fundamentals of microbial communities has enabled the application of metagenomic studies to food safety in which the presence of foodborne microbial pathogens and phytopathogenic fungi was detected through high-throughput sequencing methods [105]. Furthermore, metagenomics revealed an effectiveness in detecting plant pathogens which should be monitored for global food security [106]. Previously, Tatsika et al. examined the bacterial composition of ready-to-eat vegetables using 16S rRNA and revealed different bacterial diversities in various types of salads [107]. Fortunately, there was no record of foodborne pathogenic taxa in the surveyed samples. Furthermore, their study showed that the removal of microbes on vegetables can be successfully conducted using washing methods [107]. The microbiomes on the surface of fruits such as white guava, passion fruit, and papaya harvested in Northern Argentina were also explored using shotgun metagenomics [108]. The results revealed the presence of bacteria, yeasts, and filamentous fungi on the surface as well as unidentified species, suggesting further studies examining intrinsic species in plants. The microbial composition in apple fruits before and after processing was investigated based on 16S rRNA sequences [109]. The results exhibited a reduction of microbes after processing; however, some bacteria such as *Pseudomonas* and *Ralstonia* still survive on apples, suggesting a need for monitoring these taxa on fruits. One study sequences the total RNA content in the food ingredients to test the correlation between shifts in microbiomes and contaminants [110]. Moreover, RNA sequencing allowed for detection of the viability of microbes in food. The long-read metagenomics sequencing method showed high effectiveness for the detection of Shiga toxin-producing *Escherichia coli* (STEC) contamination in water at 10^3 CFU/mL (68 reads), suggesting its potential applicability to other foodborne pathogens [111]. In addition to identifying microbes in foods, the metagenomic approach has allowed for the screening of microbial existence in factories, tools, and any state of production [112]. Although methods and bioinformatics tools have been developed for exploring microbiomes in fruits, a suitable strategy should be selected for the aims of study. Previously, Jo et al. compared DNA and mRNA libraries and three analytical methods for microbial diversity in overwintering pepper fruits and demonstrated that library types, analytical methods, and proper databases contribute to the achievement of microbiome study [113].

4. Metagenomics in Human Health

The metagenomic approach has revolutionized studies investigating the human microbiome originating from different parts of the human body, such as the skin, oral cavity, lung, and intestine [114–117]. A recent review described the advancements that have been made in technologies for exploring microbial DNA in human samples wherein host DNA depletion and microbial DNA enrichment were summarized [118]. Previously, an efficient strategy called MetaGeniE was developed to identify microbes with high specificity and sensitivity [119]. In particular, the MetaGeniE pipeline includes two parts: Read-Reduct and Patho-Detect. The former part results in high-quality reads after conducting filtration using PRINSEQ [120], BWA [121], and STAMPY [122]. In the Patho-Detect part, the filtered reads are aligned to genomic databases to identify microbial composition using BLAT [123]. Moreover, the mBodyMap containing 14,401 metagenomes related to 22 body sites and 56 human diseases was introduced as a useful database for further research into human microbiomes and related diseases [124]. Consequently, various studies have reported applications of metagenomics for surveying the pathogens in the human body, which is a necessary aspect of formulating strategies for public health [125–127]. For example, Malla et al. summarized the sequencing technologies and bioinformatic tools that were used to explore the connection between microbiomes, human health, and diseases [125]. Further, applications of metagenomics as new therapeutic approaches for diagnosis and treatment are covered and discussed. Similarly, Ko et al. outlined the use of metagenomics for pathogen surveillance (i.e., the detection of pathogens and antimicrobial resistance genes of microbes in global sewage) [126]. Further, untargeted metagenomics has revealed the dynamics of microbial communities in cystic fibrosis patients and antibiotic-resistance

genes [127]. Another study based on nanopore metagenomics showed that it could be used as a rapid and effective diagnostic tool for pneumonia pathogens [128]. Similarly, various viruses such as rhinovirus, coronavirus, parainfluenza, parechovirus, metapneumovirus, and influenza virus have been identified in clinical specimens using BLASTN [129], bowtie2 (version 2.2.5) [130], and PCR validation [131]. ONT sequencing-based metagenomics exhibited the rapid and accurate characterization of results of the bacterial communities in lower respiratory infections with high specificity and sensitivity [132]. Similarly, nanopore-sequencing-based metagenomics showed 83% sensitivity and 100% specificity for detecting the influenza virus in respiratory samples [133]. A metagenomic study on tuberculosis patients demonstrated the relationship between the pulmonary microbiome and its clinical characteristics [134]. Another review demonstrated the interactions between bacterial communities and human immunology through the developmental stages of infancy, childhood, and adulthood [135]. The results of these prior studies prove the advantages and effectiveness of next-generation sequencing-based metagenomics as a diagnostic tool for lower-respiratory-tract infections and their potential therapy for respiratory diseases [136].

In the human microbiome, the oral microbial communities have also been screened [137]. From the current oral microbiome data, about 56,213 metagenome-assembled genomes have been obtained, in which 64% of the sequences were not previously reported [115]. A survey of 47 children with dental caries and healthy dentition revealed that *Prevotella* spp., *Streptococcus mutans*, and the Epstein–Barr virus were all correlated with caries [138]. An interesting metagenomic study on the oral microbial communities in ancient humans revealed that *Anaerolineaceae* bacterium was dominant [139]. The results of that same study also indicated a significant shift in resistance to antibiotics from prehistoric humans to modern ones. Aside from providing the taxonomy and composition of the microbiome, metagenomics supplies new evidence of associations between human genetics and oral microbiomes [140]. Specifically, five loci in the human genome, including *APPL2*, *SLC2A9*, *OR11H1*, *LOC105371703*, and *MGST1*, were related to oral microbes. Moreover, the host genetics were found to be responsible for dental diseases instead of the oral microbiome, suggesting that specific therapies should be designed for each individual case. However, a recent review has emphasized the impact of oral microbiota on oral diseases [141]. The effects of oral bacteria are based on the release of pro-inflammatory cytokines. The availability of data facilitated metapangenomics of the oral microbiomes, which showed a correlation between the genomic diversity of the oral bacteria and environmental features [142]. An additional application of metagenomics for oral microbiomes is the mining of new genes. Previously, unknown acid-tolerant genes were identified in dental caries patients, suggesting the feasibility and efficiency of metagenomics for finding functional genes [143].

Most metagenomic studies on human microbiomes have focused on gut microbiota, which exhibits correlations to digestion, immunity, and diseases [144–150]. For example, a large-scale analysis of the gut microbiomes in Parkinson's disease (PD) patients revealed microbes responsible for dysbiosis and genes related to the PD mechanism [151]. These findings have provided the foundations for further studies investigating the treatment of PD using different microbial compositions. Recently, bioinformatic strategies and technological developments with various benefits and limitations for gut microbiomes have been summarized and discussed [152]. The popularity of metagenomic studies and bioinformatic tools (i.e., CheckM v1.0.11 [153], dRep v2.2.4 [154], Mash v2.1 [155], FastANI v1.1 [156], VirSorter v.1.0.5 [157], VirFinder v.1.1 [158], CONCOCT v1.1.0 [159], GTDB-Tk v1.0.2 [160], and Prodigal v2.6.3 [161]) has led to the formation of human gut microbiome database, such as the Human Reference Gut Microbiome (HRGM), gut MEtaGenome Atlas (gutMEGA), Unified Human Gastrointestinal Protein (UHGP), and Metagenomic Gut Virus catalog [162–165]. The sequence data in these databases can be easily downloaded, and they serve as useful taxonomic references for conducting metagenomics research on humans under various conditions. For example, recent findings on microbiomes have revealed significant differences between monks and control subjects [166]. Specifically, monks who practiced long-term meditation had abundant *Prevotella*, *Bacteroides*, *Mega-*

monas, and *Faecalibacterium* species. In addition, several pathways, such as those of glycan biosynthesis and lipopolysaccharide biosynthesis, were predicted to be significantly higher in person who practice meditation. By contrast, the monks had lower levels of cholesterol and apolipoprotein B, which are factors that potentially affect human health. These results indicate the benefits of meditation, which helps to reduce stress and enhance the immune system. Another review examining the health benefits of adding probiotic microorganisms showed positive impacts on human health and several diseases [167]. One of the possible outcomes of a metagenomic study is to find correlations between microbiomes and cancers. Ng et al. summarized the factors related to colorectal cancer carcinogenesis and discussed potential methods for the diagnosis and treatment of colorectal cancer [168]. Another application of metagenomics is to identify drug efficacy and toxicity through bacterial compositions. A study on 4198 individuals revealed the small and large effects of different drugs to gut microbiome [169]. A recent study on the human microbiomes of patients with gastrointestinal symptoms caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) showed a decrease in bacterial strains in these COVID-19 patients [170]. Moreover, after recovery, the compositions of the gut microbiomes in COVID-19 patients have been found to be quite different from those of the control subjects. A similar study on COVID-19 patients provided detailed information on the connection between the microbiome and disease symptoms [171]. Specifically, the opportunistic pathogen bacteria increased significantly while the protein metabolism and carbohydrate-oriented pathways were overexpressed in COVID-19 patients. Two independent research teams reported coevolution and codiversification between the gut microbiomes and human hosts [172,173]. An investigation of the gut microbiota of children suggested a new target for better growth in low- and middle-income countries: the genetic functions of microbes [174].

Another application of metagenomics is in the exploration of the presence of viromes and their therapeutic utilization for human health [175]. Consequently, the presence of the viromes and interaction with other viruses in the human body resulted in the development of therapeutic applications such as fecal microbiota transplantation, phage-based therapy, and oncolytic therapy [175]. A previous review summarized the detection of various viruses and bacteriophages from the human microbiome [176]. The presence of viruses was found to be related to inflammatory bowel disease, diarrhea, obesity, and diabetes [177,178]. These findings indicate the potential for developing diagnostic and therapeutic measures for viruses-related diseases. For example, bacteriophages have been shown to be effective in the treatment of patients infected with *Clostridium difficile* [179,180]. Although the applications of human virome were suggested and tested, some cautions should be addressed before wide applications to human beings, such as clinical protocols, validation of the outcomes, and details of the mechanistic interactions [175,176].

5. Metagenomics in the Environmental Sciences

Soil is an important part of the Earth that contains organic and inorganic substrates, as well as provides a living environment for many species. Metagenomics has enabled research into microbial diversity in different soil types [181]. For example, one study on the soil microbial communities in different ecosystems, such as deserts, forests, grasslands, and tundra, revealed a relationship among bacterial composition, functional genes, and the environment [182]. For example, deserts were found to have low taxonomic and functional diversity compared to other surveyed locations. Further, genes related to osmoregulation and dormancy were abundant in the desert samples. By contrast, a low number of antibiotic-resistant genes was noted from the desert bacterial community, suggesting minor competition in the desert biome. Similarly, the results of Arctic soil microbiomes from different depths have indicated a significant decrease in microbial biodiversity and notable change in the functional genes with depth [183]. Another finding is that the microbiomes in tropical rainforest soils were less diverse than those in grasslands and agricultural soils [184]. Metagenomic research has also provided an overview of the alterations in microbiomes in soils over time. Observation of the microbial compositions before and

after building an oil pipeline indicated a dramatic increase in Microcoleaceae, Pinaceae, and Williamsiaceae, along with a notable decrease in Psychromonadaceae and Leuconostocaceae in samples from 2016 compared to those from 2013 [185]. Similarly, an increase in *Nitrosospira* and *Sulfuricella* genera, which is related to cadmium tolerance, was detected in soil samples with cadmium contamination in comparison to non-contaminated soil [186]. Changes in the microbial communities in total petroleum hydrocarbon (TPH)-contaminated soil were explored, in which the portion of TPH-degrading bacteria (i.e., beta-, gamma-, and delta-proteobacteria) increased by 9% compared to the control sample [187]. Another study on mercury-contaminated soils revealed differences in microbes in the rhizospheric and bulk soils [188]. The rhizospheric soils had a high abundance of Proteobacteria, whereas Actinobacteria and Alphaproteobacteria were the majority in the bulk soil. This result revealed selective effects in contaminated soils which are useful for bioremediation. A recent observation of the microbial composition of the rhizospheres of coastal plants revealed a correlation between the soil minerals and bacterial communities, suggesting that microbes may potentially play roles in stress resistance in coastal plants [189]. The availability of soil metagenomic data resulted in the formation of the largest publicly available sequencing dataset that was managed and updated annually by the National Ecological Observatory Network (NEON) [190]. The results of soil metagenomics have indicated not only the diversity of microbes but also potential ways to clean polluted soils.

In addition to the soil environment, metagenomics has been used to assess the microbial communities in water [191]. Various water bodies, such as oceans, lakes, mangrove ecosystems, rivers, and canals, have been used for metagenomic research [192–200]. Further, results obtained using metagenomics have uncovered potentially novel deep-ocean microorganisms and diverse metabolic strategies [201]. The available marine biome data were combined to establish the MarineMetagenomeDB, which provides an effective resource for further metagenomic studies [202]. In another work, the presence of microorganisms in wastewater was studied to formulate a suitable recycling method [203]. A recent study on poly-contaminated groundwater showed that the presence of Burkholderiales can degrade various contaminants, thus indicating a suitable method for bioremediation [204]. Metagenomic tools have allowed for the assessment of the microbial communities in drinking water, which has become a useful solution for quality monitoring [205,206]. In addition to common bacteria in polluted water, the outcome of their study revealed that *Arcobacter* and *Aeromonas* could be used as pollution indicators for fecal pollution source tracking [205]. In addition to bacterial detection, metagenomics could also be used to explore the presence of viruses in drinking water [207].

Similar to the soil and water environments, the air also contains microbes. However, metagenomic studies of air samples face some challenges regarding microbial density, standardized methodologies, and bioinformatics tools [208]. Fortunately, the advent of new technologies for sample collections (i.e., the TOP filter system), DNA sequencing (i.e., Illumina, PacBio, and Oxford Nanopore Technologies platforms), and analysis pipelines (i.e., MEGAN and QIIME2) have allowed for the application of air metagenomics to explore metabolic pathways, bioremediation methods, and biogeochemical cycles [20,209]. Although the low biomass feature of the air environment has resulted in a lower number of microbes than in soil and water media, metagenomic research has revealed differing abundances of microbial communities [210]. Metagenomic studies examining the air environment at different locations revealed airborne microbiome alterations. A metagenomic analysis of 3226 air samples revealed both positive and negative correlations between anthropogenic activities and airborne communities [211]. Moreover, *Acinetobacter*, *Corynebacterium*, *Mycobacterium*, and *Staphylococcus* genera were found to be positively related to the mortality rates of patients with respiratory diseases. Further, airborne pathogens were more likely to emerge from the surfaces of the human body. Another study examining 370 air samples around the world indicated that the features of the surface environment determined the abundance of airborne bacteria [212]. Additionally, the aerial environments and the microbes in the nearby ecosystems could impact the variation of global airborne

bacteria. A notable study on 789 metagenomes collected from a single site during a one-year period has advanced our understanding of the correlations between the dynamics of bacterial composition and the Diel cycle [213]. The diversity of the microbes showed daily variations rather than fluctuations of bacterial communities on a daily or monthly basis. Further, among the environmental parameters such as temperature, humidity, and CO_2 concentration, temperature was found to be the main factor affecting microbial community dynamics. Metagenomic studies on indoor air have also exhibited a diversity of DNA and RNA viruses [214]. Hospital air directly affects patient health and is an important target for metagenomic research. Previous results based on the shotgun metagenomic approach have revealed an abundance of opportunistic pathogens such as *Aspergillus*, *Penicillium*, and *Stenotrophomonas* [215]. Moreover, a multi-drug resistant bacterial strain (i.e., *Stenotrophomonas maltophilia*) has been observed. Another metagenomic study found resistomes in *Staphylococcus*, *Micrococcus*, *Streptococcus*, and *Enterococcus* species in hospital air [216]. Moreover, the higher antimicrobial resistance associated with hospital air compared to urban ambient has been discovered [217]. Collectively, metagenomic research has shown a large diversity of microbes in the soil, water, and air environments. The presence of various microorganisms has triggered the primary application of metagenomics as a tool for the identification of specific microbes. Further, the existence of bacteria can provide necessary initial information for implementing bioremediation in polluted environments.

6. Metagenomics in Agriculture

The advantage of metagenomics in defining microbial compositions has made it a useful method for monitoring agriculturally important pathogens and diseases [218]. Functional metagenomics also provides a powerful tool for elucidating the interaction between crops and microbes to increase crop yield and to identify new genes for stress resistance in crops. To benefit from the advantages of metagenomic features, different techniques have been summarized and discussed [15,219]. For example, a review of bioinformatic tools for metagenomic studies of anaerobic digesters has shown that various approaches can be used, such as those involving artificial intelligence and neural network software [220]. Recently, aerial environmental DNA data have been used to monitor pathogens in crop fields [221]. Obviously, sustainable agricultural development requires the results of metagenomic studies [222]. Previously, microbial communities in saline environments have been identified together with three different planting systems, including a conventional system, an aerobic system, and a system of rice intensification (SRI) [223]. As a result, the SRI soil samples exhibited higher species diversity than the other methods. Moreover, varied functional properties were found in all soil samples collected from the three systems, suggesting the effectiveness of different planting systems. A screening of the microbial communities in desert farming systems revealed unexpectedly large diversity [224]. Aside from the diverse composition, unknown bacterial groups were detected, suggesting novel plant microorganism interactions. Another shotgun metagenomic research on the maize rhizosphere showed an abundance of *nif*H, *nif*A, *gro*ES, and *csp*A genes, which can be potentially employed to reduce environmental stress and enhance plant development [225]. Metagenomic studies have also enabled the detection of novel viruses belonging to Betaflexiviridae, Tombusviridae, and Geminiviridae families in maize [226]. Aside from maize, similar metagenomic research has also been conducted for other plants such as wheat, maize, sugarcane, rubber tree, and vegetables [227–232].

In addition to crop-based studies, metagenomic research has been applied to farm animals [233]. Glendinning et al. conducted a metagenomic study examining four ruminants, including cow (*Bos taurus*), sheep (*Ovis aries*), reindeer (*Rangifer tarandus*), and red deer (*Cervus elaphus*), and they constructed 391 microbial genomes, out of which 279 records were new taxa [234]. Similarly, Sato et al. assembled 146 genomes from the cattle rumen [235]. The outcomes of such metagenomic research were not only bacterial compositions but also viral information [236]. A recent study on the virus communities in chicken farms has indicated the utility of metagenomics in tracking viral pathogens [237]. Another applica-

tion of metagenomics is for the determination of the relationship between the microbial community and host nutrition and metabolisms [238]. Further, differences in microbial compositions between domestic and wild animals were addressed, and exhibited a high abundance of antimicrobial resistance genes in the microbiomes of farm animals [239]. A survey on common resistomes showed 201 antibiotic-resistant genes (ARGs) in different animal manures [240]. However, the number of ARGs was reduced significantly in commercial compost. Therefore, it is necessary to upgrade the production progress to minimize the presence of ARGs in agricultural ecosystems. Structures of the microbial communities have also shed lights on ruminal fermentation, which can be employed for further research into plant biomass degradation [241].

For aquatic animals, metagenomics is a powerful tool for exploring the microbial communities necessary to improve water quality, treat wastewater, and prevent diseases. Strategies based on different sequencing methods for exploring microbes in recirculating aquaculture systems have been established, and these have provided fundamentals for further metagenomics related to aquaculture [242]. A previous metagenomic study characterized the microbial diversity in shrimp ponds [243]. Similarly, the core bacterial genera in the gut of shrimps were identified, wherein nine taxa had strong relations to the fast growth of shrimp [244]. This observation has advanced our understanding of how to control the development of shrimp. Metagenomics has also helped assess the efficiency of newly developed technologies for aquaculture ecosystems. For example, a test of the biofloc technology for a shrimp farm revealed that biofloc-based aquaculture had more opportunistic pathogens [245]. Overall, metagenomics has exhibited significant effectiveness in agriculture for pathogen monitoring, antibiotic resistance detection, and quality improvement.

7. The Future of Metagenomics: Novel Fields and Future Perspectives

For the fields of biochemistry and biotechnology, metagenomics is a molecular tool that can be used to find new enzymes from microbial communities [246]. The workflows and strategies for mining novel enzymes through functional screening and sequence-based metagenomic approaches have been reviewed in some earlier works [247–249]. Specifically, Sung et al. developed an approach for focused identification of the NGS-based definitive enzyme research (FINDER) strategy for the rapid large-scale screening of environmental microbiota and enzymes [249]. A previous review summarized 332 industrially relevant enzymes from unculturable microorganisms [250]. Different enzymes, such as lipases, cellulases, and proteases, have been identified based on metagenomic data [251]. At present, the rapidly increasing volume of metagenomic data is expected to result in more reports on novel enzymes in the future.

For microbial diversity, metagenomic databases are useful sources for identifying new microbes in different environments. Unlike amplicon metagenomics, shotgun metagenomics provides all the DNA sequences in a sample. Therefore, various workflows for reconstructing complete genomes from metagenomic data have been developed [252–255]. A recent study reported 4142 microbial metagenome-assembled genomes in the horse gut microbiome, of which 4015 records potentially belong to new species [256]. In addition to bacterial genomes, viral genomes were successfully assembled from the metagenomic data [257]. The diatom community was also characterized using the metagenomic approach [258]. Similarly, 24 lichenized-fungal genomes were completed, and these indicated high diversity and dissimilarity in the secondary metabolite biosynthetic gene cluster of lichens [259].

Metagenomics has offered an effective method for observing uncultured microbes at the genetic level in the field of microbiology [260]. Complete genomes of unknown microorganisms have also been completed from metagenomic data. These precise genomic data have enabled the use of different strategies to convert uncultured microbes to cultured ones [261]. The availability of complete genome sequences is necessary to predict the various metabolic pathways that are needed to prepare suitable culture media; then,

uncultured microbes could be cultured using different approaches, such as ARG-based isolation, stable-isotope-probing-guided Raman-activated microbial cell sorting, gene-targeted isolation, and reverse-genomics-guided isolation. Although successful isolations of uncultured microbes have been reported, certain challenges remain, including optimization of the culture media, various sizes for the cell sorter, unknown gene expression, and DNA extraction methods.

Lastly, the recent ESM Metagenomic Atlas constructed using artificial intelligence contains more than 617 million protein structures, out of which millions were new compared to the available protein database [262]. This finding has opened new paths to explore metagenomic data using artificial intelligence, which provides powerful assistance for exploring uncovered parts of the scientific world. Obviously, metagenomics has opened new doors for exploring the microbial world on Earth. However, further discovery to achieve a better understanding requires more advancements in terms of sampling protocols, analysis pipelines, databases, and interpretation of the results.

References

1.   Robertson, L.A. van Leeuwenhoek microscopes—Where are they now? *FEMS Microbiol. Lett.* **2015**, *362*, fnv056. [CrossRef] [PubMed]
2.   Blevins, S.M.; Bronze, M.S. Robert Koch and the 'golden age' of bacteriology. *Int. J. Infect. Dis.* **2010**, *14*, e744–e751. [CrossRef]
3.   Woese, C.R.; Fox, G.E. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 5088–5090. [CrossRef] [PubMed]
4.   Woese, C.R. Bacterial evolution. *Microbiol. Rev.* **1987**, *51*, 221–271. [CrossRef] [PubMed]
5.   Johnson, J.S.; Spakowicz, D.J.; Hong, B.-Y.; Petersen, L.M.; Demkowicz, P.; Chen, L.; Leopold, S.R.; Hanson, B.M.; Agresta, H.O.; Gerstein, M.; et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat. Commun.* **2019**, *10*, 5029. [CrossRef]
6.   Handelsman, J.; Rondon, M.R.; Brady, S.F.; Clardy, J.; Goodman, R.M. Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products. *Chem. Biol.* **1998**, *5*, R245–R249. [CrossRef]
7.   Kamble, A.; Sawant, S.; Singh, H. 16S ribosomal RNA gene-based metagenomics: A review. *Biomed. Res. J.* **2020**, *7*, 5. [CrossRef]
8.   Quince, C.; Walker, A.W.; Simpson, J.T.; Loman, N.J.; Segata, N. Shotgun metagenomics, from sampling to analysis. *Nat. Biotechnol.* **2017**, *35*, 833–844. [CrossRef]
9.   Vecherskii, M.V.; Semenov, M.V.; Lisenkova, A.A.; Stepankov, A.A. Metagenomics: A New Direction in Ecology. *Biol. Bull.* **2021**, *48*, S107–S117. [CrossRef]
10. Garber, A.I.; Zehnpfennig, J.R.; Sheik, C.S.; Henson, M.W.; Ramírez, G.A.; Mahon, A.R.; Halanych, K.M.; Learman, D.R. Metagenomics of Antarctic Marine Sediment Reveals Potential for Diverse Chemolithoautotrophy. *mSphere* **2021**, *6*, e00770-21. [CrossRef]
11. Chevrette, M.G.; Handelsman, J. From Metagenomes to Molecules: Innovations in Functional Metagenomics Unlock Hidden Chemistry in the Human Microbiome. *Biochemistry* **2020**, *59*, 729–730. [CrossRef] [PubMed]
12. Spanogiannopoulos, P.; Bess, E.N.; Carmody, R.N.; Turnbaugh, P.J. The microbial pharmacists within us: A metagenomic view of xenobiotic metabolism. *Nat. Rev. Microbiol.* **2016**, *14*, 273–287. [CrossRef] [PubMed]
13. Yap, M.; Ercolini, D.; Álvarez-Ordóñez, A.; O'Toole, P.W.; O'Sullivan, O.; Cotter, P.D. Next-Generation Food Research: Use of Meta-Omic Approaches for Characterizing Microbial Communities along the Food Chain. *Annu. Rev. Food Sci. Technol.* **2022**, *13*, 361–384. [CrossRef] [PubMed]
14. Sabater, C.; Cobo-Díaz, J.F.; Álvarez-Ordóñez, A.; Ruas-Madiedo, P.; Ruiz, L.; Margolles, A. Novel methods of microbiome analysis in the food industry. *Int. Microbiol.* **2021**, *24*, 593–605. [CrossRef]
15. Bakshi, A.; Moin, M.; Madhav, M.S. Metagenomics in Agriculture: State-of-the-Art. In *Metagenomics: Techniques, Applications, Challenges and Opportunities*; Springer: Singapore, 2020; pp. 167–187.

16. Gounot, J.-S.; Chia, M.; Bertrand, D.; Saw, W.-Y.; Ravikrishnan, A.; Low, A.; Ding, Y.; Ng, A.H.Q.; Tan, L.W.L.; Teo, Y.-Y.; et al. Genome-centric analysis of short and long read metagenomes reveals uncharacterized microbiome diversity in Southeast Asians. *Nat. Commun.* **2022**, *13*, 6044. [CrossRef]

17. Johansen, J.; Plichta, D.R.; Nissen, J.N.; Jespersen, M.L.; Shah, S.A.; Deng, L.; Stokholm, J.; Bisgaard, H.; Nielsen, D.S.; Sørensen, S.J.; et al. Genome binning of viral entities from bulk metagenomics data. *Nat. Commun.* **2022**, *13*, 965. [CrossRef]

18. Santiago-Rodriguez, T.M.; Hollister, E.B. Unraveling the viral dark matter through viral metagenomics. *Front. Immunol.* **2022**, *13*, 1005107. [CrossRef]

19. Navgire, G.S.; Goel, N.; Sawhney, G.; Sharma, M.; Kaushik, P.; Mohanta, Y.K.; Mohanta, T.K.; Al-Harrasi, A. Analysis and Interpretation of metagenomics data: An approach. *Biol. Proced. Online* **2022**, *24*, 18. [CrossRef]

20. Liu, Y.-X.; Qin, Y.; Chen, T.; Lu, M.; Qian, X.; Guo, X.; Bai, Y. A practical guide to amplicon and metagenomic analysis of microbiome data. *Protein Cell* **2021**, *12*, 315–330. [CrossRef]

21. Roumpeka, D.D.; Wallace, R.J.; Escalettes, F.; Fotheringham, I.; Watson, M. A Review of Bioinformatics Tools for Bio-Prospecting from Metagenomic Sequence Data. *Front. Genet.* **2017**, *8*, 23. [CrossRef]

22. Tripathi, L.K.; Nailwal, T.K. Metagenomics: Applications of functional and structural approaches and meta-omics. In *Recent Advancements in Microbial Diversity*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 471–505.

23. Datta, S.; Rajnish, K.N.; Samuel, M.S.; Pugazlendhi, A.; Selvarajan, E. Metagenomic applications in microbial diversity, bioremediation, pollution monitoring, enzyme and drug discovery. A review. *Environ. Chem. Lett.* **2020**, *18*, 1229–1241. [CrossRef]

24. Prayogo, F.A.; Budiharjo, A.; Kusumaningrum, H.P.; Wijanarka, W.; Suprihadi, A.; Nurhayati, N. Metagenomic applications in exploration and development of novel enzymes from nature: A review. *J. Genet. Eng. Biotechnol.* **2020**, *18*, 39. [CrossRef] [PubMed]

25. Kinoshita, Y.; Niwa, H.; Uchida-Fujii, E.; Nukada, T. Establishment and assessment of an amplicon sequencing method targeting the 16S-ITS-23S rRNA operon for analysis of the equine gut microbiome. *Sci. Rep.* **2021**, *11*, 11884. [CrossRef]

26. Zhang, L.; Chen, F.; Zeng, Z.; Xu, M.; Sun, F.; Yang, L.; Bi, X.; Lin, Y.; Gao, Y.; Hao, H.; et al. Advances in Metagenomics and Its Application in Environmental Microorganisms. *Front. Microbiol.* **2021**, *12*, 3847. [CrossRef] [PubMed]

27. Breitwieser, F.P.; Lu, J.; Salzberg, S.L. A review of methods and databases for metagenomic classification and assembly. *Brief. Bioinform.* **2019**, *20*, 1125–1136. [CrossRef] [PubMed]

28. Wajid, B.; Anwar, F.; Wajid, I.; Nisar, H.; Meraj, S.; Zafar, A.; Al-Shawaqfeh, M.K.; Ekti, A.R.; Khatoon, A.; Suchodolski, J.S. Music of metagenomics—A review of its applications, analysis pipeline, and associated tools. *Funct. Integr. Genom.* **2022**, *22*, 3–26. [CrossRef] [PubMed]

29. Tourlousse, D.M.; Narita, K.; Miura, T.; Sakamoto, M.; Ohashi, A.; Shiina, K.; Matsuda, M.; Miura, D.; Shimamura, M.; Ohyama, Y.; et al. Validation and standardization of DNA extraction and library construction methods for metagenomics-based human fecal microbiome measurements. *Microbiome* **2021**, *9*, 95. [CrossRef]

30. Ahmadi, E.; Kowsari, M.; Azadfar, D.; Jouzani, G.S. Rapid and economical protocols for genomic and metagenomic DNA extraction from oak (*Quercus brantii Lindl.*). *Ann. For. Sci.* **2018**, *75*, 43. [CrossRef]

31. Echeverría-Beirute, F.; Varela-Benavides, I.; Jiménez-Madrigal, J.P.; Carvajal-Chacon, M.; Guzmán-Hernández, T. eDNA extraction protocol for metagenomic studies in tropical soils. *Biotechniques* **2021**, *71*, 580–586. [CrossRef]

32. Sui, H.; Weil, A.A.; Nuwagira, E.; Qadri, F.; Ryan, E.T.; Mezzari, M.P.; Phipatanakul, W.; Lai, P.S. Impact of DNA Extraction Method on Variation in Human and Built Environment Microbial Community and Functional Profiles Assessed by Shotgun Metagenomics Sequencing. *Front. Microbiol.* **2020**, *11*, 953. [CrossRef]

33. Shaffer, J.P.; Carpenter, C.S.; Martino, C.; Salido, R.A.; Minich, J.J.; Bryant, M.; Sanders, K.; Schwartz, T.; Humphrey, G.; Swafford, A.D.; et al. A comparison of six DNA extraction protocols for 16S, ITS and shotgun metagenomic sequencing of microbial communities. *Biotechniques* **2022**, *73*, 34–46. [CrossRef] [PubMed]

34. Yang, F.; Sun, J.; Luo, H.; Ren, H.; Zhou, H.; Lin, Y.; Han, M.; Chen, B.; Liao, H.; Brix, S.; et al. Assessment of fecal DNA extraction protocols for metagenomic studies. *Gigascience* **2020**, *9*, giaa071. [CrossRef] [PubMed]

35. Rehner, J.; Schmartz, G.P.; Groeger, L.; Dastbaz, J.; Ludwig, N.; Hannig, M.; Rupf, S.; Seitz, B.; Flockerzi, E.; Berger, T.; et al. Systematic Cross-biospecimen Evaluation of DNA Extraction Kits for Long- and Short-read Multi-metagenomic Sequencing Studies. *Genomics. Proteom. Bioinform.* **2022**, *20*, 405–417. [CrossRef] [PubMed]

36. Shuikan, A.; Alharbi, S.A.; Alkhalifah, D.H.M.; Hozzein, W.N. High-Throughput Sequencing and Metagenomic Data Analysis. In *Metagenomics-Basics, Methods and Applications*; IntechOpen: London, UK, 2020.

37. Lindahl, B.D.; Kuske, C.R. Metagenomics for Study of Fungal Ecology. In *The Ecological Genomics of Fungi*; John Wiley & Sons Inc: Hoboken, NJ, USA, 2013; pp. 279–303.

38. McNichol, J.; Berube, P.M.; Biller, S.J.; Fuhrman, J.A. Evaluating and Improving Small Subunit rRNA PCR Primer Coverage for Bacteria, Archaea, and Eukaryotes Using Metagenomes from Global Ocean Surveys. *mSystems* **2021**, *6*, e0056521. [CrossRef] [PubMed]

39. Martin, J.L.; Santi, I.; Pitta, P.; John, U.; Gypens, N. Towards quantitative metabarcoding of eukaryotic plankton: An approach to improve 18S rRNA gene copy number bias. *Metabarcoding Metagenom.* **2022**, *6*, e85794. [CrossRef]

40. Lan, Y.; Rosen, G.; Hershberg, R. Marker genes that are less conserved in their sequences are useful for predicting genome-wide similarity levels between closely related prokaryotic strains. *Microbiome* **2016**, *4*, 18. [CrossRef]

41. Yang, B.; Wang, Y.; Qian, P.-Y. Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. *BMC Bioinform.* **2016**, *17*, 135. [CrossRef]

42. Kim, M.; Morrison, M.; Yu, Z. Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. *J. Microbiol. Methods* **2011**, *84*, 81–87. [CrossRef]

43. Shao, L.; Liao, J.; Qian, J.; Chen, W.; Fan, X. MetaGeneBank: A standardized database to study deep sequenced metagenomic data from human fecal specimen. *BMC Microbiol.* **2021**, *21*, 263. [CrossRef]

44. Hu, R.; Yao, R.; Li, L.; Xu, Y.; Lei, B.; Tang, G.; Liang, H.; Lei, Y.; Li, C.; Li, X.; et al. A database of animal metagenomes. *Sci. Data* **2022**, *9*, 312. [CrossRef]

45. Mitchell, A.L.; Almeida, A.; Beracochea, M.; Boland, M.; Burgin, J.; Cochrane, G.; Crusoe, M.R.; Kale, V.; Potter, S.C.; Richardson, L.J.; et al. MGnify: The microbiome analysis resource in 2020. *Nucleic Acids Res.* **2019**, *48*, D570–D578. [CrossRef] [PubMed]

46. Corrêa, F.B.; Saraiva, J.P.; Stadler, P.F.; da Rocha, U.N. TerrestrialMetagenomeDB: A public repository of curated and standardized metadata for terrestrial metagenomes. *Nucleic Acids Res.* **2019**, *48*, D626–D632. [CrossRef] [PubMed]

47. Zhang, T.; Miao, J.; Han, N.; Qiang, Y.; Zhang, W. MPD: A pathogen genome and metagenome database. *Database* **2018**, *2018*, bay055. [CrossRef] [PubMed]

48. Amir, A.; McDonald, D.; Navas-Molina, J.A.; Kopylova, E.; Morton, J.T.; Zech Xu, Z.; Kightley, E.P.; Thompson, L.R.; Hyde, E.R.; Gonzalez, A.; et al. Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. *mSystems* **2017**, *2*, e00191-16. [CrossRef]

49. Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.A.; Holmes, S.P. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **2016**, *13*, 581–583. [CrossRef]

50. Rognes, T.; Flouri, T.; Nichols, B.; Quince, C.; Mahé, F. VSEARCH: A versatile open source tool for metagenomics. *PeerJ* **2016**, *4*, e2584. [CrossRef]

51. Wood, D.E.; Lu, J.; Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **2019**, *20*, 257. [CrossRef]

52. Tatusov, R.L. The COG database: A tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **2000**, *28*, 33–36. [CrossRef]

53. Kanehisa, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [CrossRef]

54. Douglas, G.M.; Maffei, V.J.; Zaneveld, J.R.; Yurgel, S.N.; Brown, J.R.; Taylor, C.M.; Huttenhower, C.; Langille, M.G.I. PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* **2020**, *38*, 685–688. [CrossRef]

55. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [CrossRef] [PubMed]

56. Lomsadze, A.; Bonny, C.; Strozzi, F.; Borodovsky, M. GeneMark-HM: Improving gene prediction in DNA sequences of human microbiome. *NAR Genom. Bioinform.* **2021**, *3*, lqab047. [CrossRef] [PubMed]

57. Huson, D.H.; Auch, A.F.; Qi, J.; Schuster, S.C. MEGAN analysis of metagenomic data. *Genome Res.* **2007**, *17*, 377–386. [CrossRef] [PubMed]

58. Louca, S.; Parfrey, L.W.; Doebeli, M. Decoupling function and taxonomy in the global ocean microbiome. *Science* **2016**, *353*, 1272–1277. [CrossRef] [PubMed]

59. Sansupa, C.; Wahdan, S.F.M.; Hossen, S.; Disayathanoowat, T.; Wubet, T.; Purahong, W. Can We Use Functional Annotation of Prokaryotic Taxa (FAPROTAX) to Assign the Ecological Functions of Soil Bacteria? *Appl. Sci.* **2021**, *11*, 688. [CrossRef]

60. Yang, Z.; Peng, C.; Cao, H.; Song, J.; Gong, B.; Li, L.; Wang, L.; He, Y.; Liang, M.; Lin, J.; et al. Microbial functional assemblages predicted by the FAPROTAX analysis are impacted by physicochemical properties, but C, N and S cycling genes are not in mangrove soil in the Beibu Gulf, China. *Ecol. Indic.* **2022**, *139*, 108887. [CrossRef]

61. Blanco-Míguez, A.; Beghini, F.; Cumbo, F.; McIver, L.J.; Thompson, K.N.; Zolfo, M.; Manghi, P.; Dubois, L.; Huang, K.D.; Thomas, A.M.; et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. *Nat. Biotechnol.* **2023**. [CrossRef]

62. Beghini, F.; McIver, L.J.; Blanco-Míguez, A.; Dubois, L.; Asnicar, F.; Maharjan, S.; Mailyan, A.; Manghi, P.; Scholz, M.; Thomas, A.M.; et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *eLife* **2021**, *10*, e65088. [CrossRef]

63. Bolyen, E.; Rideout, J.R.; Dillon, M.R.; Bokulich, N.A.; Abnet, C.C.; Al-Ghalith, G.A.; Alexander, H.; Alm, E.J.; Arumugam, M.; Asnicar, F.; et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **2019**, *37*, 852–857. [CrossRef]

64. Bağcı, C.; Patz, S.; Huson, D.H. DIAMOND+MEGAN: Fast and Easy Taxonomic and Functional Analysis of Short and Long Microbiome Sequences. *Curr. Protoc.* **2021**, *1*, e59. [CrossRef]

65. Gautam, A.; Zeng, W.; Huson, D.H. MeganServer: Facilitating interactive access to metagenomic data on a server. *Bioinformatics* **2023**, *39*, btad105. [CrossRef] [PubMed]

66. Fung, C.; Rusling, M.; Lampeter, T.; Love, C.; Karim, A.; Bongiorno, C.; Yuan, L. Automation of QIIME2 Metagenomic Analysis Platform. *Curr. Protoc.* **2021**, *1*, e254. [CrossRef] [PubMed]

67. Estaki, M.; Jiang, L.; Bokulich, N.A.; McDonald, D.; González, A.; Kosciolek, T.; Martino, C.; Zhu, Q.; Birmingham, A.; Vázquez-Baeza, Y.; et al. QIIME 2 Enables Comprehensive End-to-End Analysis of Diverse Microbiome Data and Comparative Studies with Publicly Available Data. *Curr. Protoc. Bioinform.* **2020**, *70*, e100. [CrossRef] [PubMed]

68. Kobus, R.; Abuín, J.M.; Müller, A.; Hellmann, S.L.; Pichel, J.C.; Pena, T.F.; Hildebrandt, A.; Hankeln, T.; Schmidt, B. A big data approach to metagenomics for all-food-sequencing. *BMC Bioinform.* **2020**, *21*, 102. [CrossRef]

69. Walsh, L.H.; Coakley, M.; Walsh, A.M.; O'Toole, P.W.; Cotter, P.D. Bioinformatic approaches for studying the microbiome of fermented food. *Crit. Rev. Microbiol.* **2022**, 1–33. [CrossRef]

70. McNally, C.P.; Eng, A.; Noecker, C.; Gagne-Maynard, W.C.; Borenstein, E. BURRITO: An Interactive Multi-Omic Tool for Visualizing Taxa–Function Relationships in Microbiome Data. *Front. Microbiol.* **2018**, *9*, 365. [CrossRef]

71. Ondov, B.D.; Bergman, N.H.; Phillippy, A.M. Interactive metagenomic visualization in a Web browser. *BMC Bioinform.* **2011**, *12*, 385. [CrossRef]

72. Wagner, J.; Chelaru, F.; Kancherla, J.; Paulson, J.N.; Zhang, A.; Felix, V.; Mahurkar, A.; Elmqvist, N.; Bravo, H.C. Metaviz: Interactive statistical and visual analysis of metagenomic data. *Nucleic Acids Res.* **2018**, *46*, 2777–2787. [CrossRef]

73. Ye, D.-D.; Fan, M.-M.; Guan, Q.; Chen, H.-J.; Ma, Z.-S. A review on the bioinformatics pipelines for metagenomic research. *Zool. Res.* **2012**, *33*, 574–585. [CrossRef]

74. Garfias-Gallegos, D.; Zirión-Martínez, C.; Bustos-Díaz, E.D.; Arellano-Fernández, T.V.; Lovaco-Flores, J.A.; Espinosa-Jaime, A.; Avelar-Rivas, J.A.; Sélem-Mójica, N. Metagenomics Bioinformatic Pipeline. In *Plant Comparative Genomics*; Springer: New York, NY, USA, 2022; pp. 153–179.

75. Kishikawa, T.; Tomofuji, Y.; Inohara, H.; Okada, Y. OMARU: A robust and multifaceted pipeline for metagenome-wide association study. *NAR Genom. Bioinform.* **2022**, *4*, lqac019. [CrossRef]

76. Glass, E.M.; Wilkening, J.; Wilke, A.; Antonopoulos, D.; Meyer, F. Using the Metagenomics RAST Server (MG-RAST) for Analyzing Shotgun Metagenomes. *Cold Spring Harb. Protoc.* **2010**, *2010*, pdb.prot5368. [CrossRef]

77. Hunter, S.; Corbett, M.; Denise, H.; Fraser, M.; Gonzalez-Beltran, A.; Hunter, C.; Jones, P.; Leinonen, R.; McAnulla, C.; Maguire, E.; et al. EBI metagenomics—A new resource for the analysis and archiving of metagenomic data. *Nucleic Acids Res.* **2014**, *42*, D600–D606. [CrossRef] [PubMed]

78. Markowitz, V.M.; Chen, I.-M.A.; Chu, K.; Szeto, E.; Palaniappan, K.; Pillay, M.; Ratner, A.; Huang, J.; Pagani, I.; Tringe, S.; et al. IMG/M 4 version of the integrated metagenome comparative analysis system. *Nucleic Acids Res.* **2014**, *42*, D568–D573. [CrossRef] [PubMed]

79. Li, P.-E.; Lo, C.-C.; Anderson, J.J.; Davenport, K.W.; Bishop-Lilly, K.A.; Xu, Y.; Ahmed, S.; Feng, S.; Mokashi, V.P.; Chain, P.S.G. Enabling the democratization of the genomics revolution with a fully integrated web-based bioinformatics platform. *Nucleic Acids Res.* **2017**, *45*, 67–80. [CrossRef] [PubMed]

80. Varzakas, T. Microbiology of Fermented Foods and Beverages. *Foods* **2020**, *9*, 1660. [CrossRef] [PubMed]

81. Voidarou, C.; Antoniadou, M.; Rozos, G.; Tzora, A.; Skoufos, I.; Varzakas, T.; Lagiou, A.; Bezirtzoglou, E. Fermentative Foods: Microbiology, Biochemistry, Potential Human Health Benefits and Public Health Issues. *Foods* **2020**, *10*, 69. [CrossRef] [PubMed]

82. Giraffa, G.; Neviani, E. DNA-based, culture-independent strategies for evaluating microbial communities in food-associated ecosystems. *Int. J. Food Microbiol.* **2001**, *67*, 19–34. [CrossRef]

83. Kergourlay, G.; Taminiau, B.; Daube, G.; Vergès, M.-C.C. Metagenomic insights into the dynamics of microbial communities in food. *Int. J. Food Microbiol.* **2015**, *213*, 31–39. [CrossRef]

84. Guzmán, A.C.; Hurtado, M.I.G.; Cuesta-Astroz, Y.; Torres, G. Metagenomic characterization of bacterial biofilm in four food processing plants in Colombia. *Braz. J. Microbiol.* **2020**, *51*, 1259–1267. [CrossRef]

85. Coughlan, L.M.; Cotter, P.D.; Hill, C.; Alvarez-Ordóñez, A. Biotechnological applications of functional metagenomics in the food and pharmaceutical industries. *Front. Microbiol.* **2015**, *6*, 672. [CrossRef]

86. De Filippis, F.; Parente, E.; Ercolini, D. Metagenomics insights into food fermentations. *Microb. Biotechnol.* **2017**, *10*, 91–102. [CrossRef] [PubMed]

87. Srinivas, M.; O'Sullivan, O.; Cotter, P.D.; Sinderen, D.V.; Kenny, J.G. The Application of Metagenomics to Study Microbial Communities and Develop Desirable Traits in Fermented Foods. *Foods* **2022**, *11*, 3297. [CrossRef]

88. Deka, P.; Mehetre, G.T.; Lalnunmawii, E.; Upadhyaya, K.; Singh, G.; Hashem, A.; Al-Arjani, A.-B.F.; Abd_Allah, E.F.; Singh, B.P. Metagenomic Analysis of Bacterial Diversity in Traditional Fermented Foods Reveals Food-Specific Dominance of Specific Bacterial Taxa. *Fermentation* **2021**, *7*, 167. [CrossRef]

89. Yu, Y.; Li, L.; Xu, Y.; Li, H.; Yu, Y.; Xu, Z. Metagenomics Reveals the Microbial Community Responsible for Producing Biogenic Amines during Mustard [*Brassica juncea* (L.)] Fermentation. *Front. Microbiol.* **2022**, *13*, 824644. [CrossRef]

90. Zhang, J.; Wang, X.; Huo, D.; Li, W.; Hu, Q.; Xu, C.; Liu, S.; Li, C. Metagenomic approach reveals microbial diversity and predictive microbial metabolic pathways in Yucha, a traditional Li fermented food. *Sci. Rep.* **2016**, *6*, 32524. [CrossRef]

91. Jung, J.Y.; Lee, S.H.; Kim, J.M.; Park, M.S.; Bae, J.-W.; Hahn, Y.; Madsen, E.L.; Jeon, C.O. Metagenomic Analysis of Kimchi, a Traditional Korean Fermented Food. *Appl. Environ. Microbiol.* **2011**, *77*, 2264–2274. [CrossRef]

92. Ahn, J.-Y.; Min, J.; Lee, S.-H.; Jang, A.; Park, C.-K.; Kwon, S.-D.; Park, S.-K.; Lee, K.; Kim, Y.-H. Metagenomic analysis for identifying *Kimchi* sp. during the industrial-scale batch fermentation. *Toxicol. Environ. Health Sci.* **2014**, *6*, 8–15. [CrossRef]

93. Lee, M.; Song, J.H.; Jung, M.Y.; Lee, S.H.; Chang, J.Y. Large-scale targeted metagenomics analysis of bacterial ecological changes in 88 kimchi samples during fermentation. *Food Microbiol.* **2017**, *66*, 173–183. [CrossRef]

94. Park, D.H. Effects of carbon dioxide on metabolite production and bacterial communities during kimchi fermentation. *Biosci. Biotechnol. Biochem.* **2018**, *82*, 1234–1242. [CrossRef]

95. Oh, Y.-J.; Park, Y.-R.; Hong, J.; Lee, D.-Y. Metagenomic, Metabolomic, and Functional Evaluation of Kimchi Broth Treated with Light-Emitting Diodes (LEDs). *Metabolites* **2021**, *11*, 472. [CrossRef]

96. Demirci, T.; Akın, N.; Öztürk, H.İ.; Oğul, A. A metagenomic approach to homemade back-slopped yogurts produced in mountainous villages of Turkey with the potential next-generation probiotics. *LWT* **2022**, *154*, 112860. [CrossRef]

97. Kumar, M.R.; Yeap, S.K.; Mohamad, N.E.; Abdullah, J.O.; Masarudin, M.J.; Khalid, M.; Leow, A.T.C.; Alitheen, N.B. Metagenomic and phytochemical analyses of kefir water and its subchronic toxicity study in BALB/c mice. *BMC Complement. Med. Ther.* **2021**, *21*, 183. [CrossRef] [PubMed]

98. Yerlikaya, O.; Akan, E.; Kinik, Ö. The metagenomic composition of water kefir microbiota. *Int. J. Gastron. Food Sci.* **2022**, *30*, 100621. [CrossRef]

99. Patel, S.H.; Tan, J.P.; Börner, R.A.; Zhang, S.J.; Priour, S.; Lima, A.; Ngom-Bru, C.; Cotter, P.D.; Duboux, S. A temporal view of the water kefir microbiota and flavour attributes. *Innov. Food Sci. Emerg. Technol.* **2022**, *80*, 103084. [CrossRef]

100. Kaashyap, M.; Cohen, M.; Mantri, N. Microbial Diversity and Characteristics of Kombucha as Revealed by Metagenomic and Physicochemical Analysis. *Nutrients* **2021**, *13*, 4446. [CrossRef]

101. Pradhan, S.; Prabhakar, M.R.; Parvathy, K.R.K.; Dey, B.; Jayaraman, S.; Behera, B.; Paramasivan, B. Metagenomic and physicochemical analysis of Kombucha beverage produced from tea waste. *J. Food Sci. Technol.* **2022**, *60*, 1088–1096. [CrossRef] [PubMed]

102. Yarlina, V.P.; Andoyo, R.; Djali, M.; Lani, M.N. Metagenomic Analysis for Indigenous Microbial Diversity in Soaking Process of Making Tempeh Jack Beans (*Canavalia ensiformis*). *Curr. Res. Nutr. Food Sci.* **2022**, *10*, 620–632. [CrossRef]

103. Yulandi, A.; Suwanto, A.; Waturangi, D.E.; Wahyudi, A.T. Shotgun metagenomic analysis reveals new insights into bacterial community profiles in tempeh. *BMC Res. Notes* **2020**, *13*, 562. [CrossRef]

104. Leech, J.; Cabrera-Rubio, R.; Walsh, A.M.; Macori, G.; Walsh, C.J.; Barton, W.; Finnegan, L.; Crispie, F.; O'Sullivan, O.; Claesson, M.J.; et al. Fermented-Food Metagenomics Reveals Substrate-Associated Differences in Taxonomy and Health-Associated and Antibiotic Resistance Determinants. *mSystems* **2020**, *5*, e00522-20. [CrossRef]

105. Shen, Y.; Nie, J.; Kuang, L.; Zhang, J.; Li, H. DNA sequencing, genomes and genetic markers of microbes on fruits and vegetables. *Microb. Biotechnol.* **2021**, *14*, 323–362. [CrossRef]

106. Piombo, E.; Abdelfattah, A.; Droby, S.; Wisniewski, M.; Spadaro, D.; Schena, L. Metagenomics Approaches for the Detection and Surveillance of Emerging and Recurrent Plant Pathogens. *Microorganisms* **2021**, *9*, 188. [CrossRef] [PubMed]

107. Tatsika, S.; Karamanoli, K.; Karayanni, H.; Genitsaris, S. Metagenomic Characterization of Bacterial Communities on Ready-to-Eat Vegetables and Effects of Household Washing on their Diversity and Composition. *Pathogens* **2019**, *8*, 37. [CrossRef] [PubMed]

108. Vermote, L.; Verce, M.; Mozzi, F.; De Vuyst, L.; Weckx, S. Microbiomes Associated with the Surfaces of Northern Argentinian Fruits Show a Wide Species Diversity. *Front. Microbiol.* **2022**, *13*, 2620. [CrossRef]

109. Wicaksono, W.A.; Buko, A.; Kusstatscher, P.; Sinkkonen, A.; Laitinen, O.H.; Virtanen, S.M.; Hyöty, H.; Cernava, T.; Berg, G. Modulation of the food microbiome by apple fruit processing. *Food Microbiol.* **2022**, *108*, 104103. [CrossRef] [PubMed]

110. Beck, K.L.; Haiminen, N.; Chambliss, D.; Edlund, S.; Kunitomi, M.; Huang, B.C.; Kong, N.; Ganesan, B.; Baker, R.; Markwell, P.; et al. Monitoring the microbiome for food safety and quality using deep shotgun sequencing. *npj Sci. Food* **2021**, *5*, 3. [CrossRef] [PubMed]

111. Maguire, M.; Kase, J.A.; Roberson, D.; Muruvanda, T.; Brown, E.W.; Allard, M.; Musser, S.M.; González-Escalona, N. Precision long-read metagenomics sequencing for food safety by detection and assembly of Shiga toxin-producing *Escherichia coli* in irrigation water. *PLoS ONE* **2021**, *16*, e0245172. [CrossRef]

112. De Filippis, F.; Valentino, V.; Alvarez-Ordóñez, A.; Cotter, P.D.; Ercolini, D. Environmental microbiome mapping as a strategy to improve quality and safety in the food industry. *Curr. Opin. Food Sci.* **2021**, *38*, 168–176. [CrossRef]

113. Jo, Y.; Back, C.-G.; Kim, K.-H.; Chu, H.; Lee, J.H.; Moh, S.H.; Cho, W.K. Comparative Study of Metagenomics and Metatranscriptomics to Reveal Microbiomes in Overwintering Pepper Fruits. *Int. J. Mol. Sci.* **2021**, *22*, 6202. [CrossRef]

114. Wei, Q.; Li, Z.; Gu, Z.; Liu, X.; Krutmann, J.; Wang, J.; Xia, J. Shotgun metagenomic sequencing reveals skin microbial variability from different facial sites. *Front. Microbiol.* **2022**, *13*, 2779. [CrossRef]

115. Zhu, J.; Tian, L.; Chen, P.; Han, M.; Song, L.; Tong, X.; Sun, X.; Yang, F.; Lin, Z.; Liu, X.; et al. Over 50,000 Metagenomically Assembled Draft Genomes for the Human Oral Microbiome Reveal New Taxa. *Genomics. Proteom. Bioinform.* **2022**, *20*, 246–259. [CrossRef]

116. Hu, Y.; Cheng, M.; Liu, B.; Dong, J.; Sun, L.; Yang, J.; Yang, F.; Chen, X.; Jin, Q. Metagenomic analysis of the lung microbiome in pulmonary tuberculosis—A pilot study. *Emerg. Microbes Infect.* **2020**, *9*, 1444–1452. [CrossRef] [PubMed]

117. Hiseni, P.; Rudi, K.; Wilson, R.C.; Hegge, F.T.; Snipen, L. HumGut: A comprehensive human gut prokaryotic genomes collection filtered by metagenome data. *Microbiome* **2021**, *9*, 165. [CrossRef] [PubMed]

118. Shi, Y.; Wang, G.; Lau, H.C.-H.; Yu, J. Metagenomic Sequencing for Microbial DNA in Human Samples: Emerging Technological Advances. *Int. J. Mol. Sci.* **2022**, *23*, 2181. [CrossRef]

119. Rawat, A.; Engelthaler, D.M.; Driebe, E.M.; Keim, P.; Foster, J.T. MetaGeniE: Characterizing Human Clinical Samples Using Deep Metagenomic Sequencing. *PLoS ONE* **2014**, *9*, e110915. [CrossRef] [PubMed]

120. Schmieder, R.; Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **2011**, *27*, 863–864. [CrossRef]

121. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [CrossRef] [PubMed]

122. Lunter, G.; Goodson, M. Stampy: A statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Res.* **2011**, *21*, 936–939. [CrossRef]

123. Kent, W.J. BLAT—The BLAST-Like Alignment Tool. *Genome Res.* **2002**, *12*, 656–664. [CrossRef]

124. Jin, H.; Hu, G.; Sun, C.; Duan, Y.; Zhang, Z.; Liu, Z.; Zhao, X.-M.; Chen, W.-H. mBodyMap: A curated database for microbes across human body and their associations with health and diseases. *Nucleic Acids Res.* **2022**, *50*, D808–D816. [CrossRef]

125. Malla, M.A.; Dubey, A.; Kumar, A.; Yadav, S.; Hashem, A.; Abd_Allah, E.F. Exploring the Human Microbiome: The Potential Future Role of Next-Generation Sequencing in Disease Diagnosis and Treatment. *Front. Immunol.* **2019**, *9*, 2868. [CrossRef]

126. Ko, K.K.K.; Chng, K.R.; Nagarajan, N. Metagenomics-enabled microbial surveillance. *Nat. Microbiol.* **2022**, *7*, 486–496. [CrossRef] [PubMed]

127. Suttner, B.; Johnston, E.R.; Orellana, L.H.; Rodriguez-R, L.M.; Hatt, J.K.; Carychao, D.; Carter, M.Q.; Cooley, M.B.; Konstantinidis, K.T. Metagenomics as a Public Health Risk Assessment Tool in a Study of Natural Creek Sediments Influenced by Agricultural and Livestock Runoff: Potential and Limitations. *Appl. Environ. Microbiol.* **2020**, *86*, e02525-19. [CrossRef] [PubMed]

128. Yang, L.; Haidar, G.; Zia, H.; Nettles, R.; Qin, S.; Wang, X.; Shah, F.; Rapport, S.F.; Charalampous, T.; Methé, B.; et al. Metagenomic identification of severe pneumonia pathogens in mechanically-ventilated patients: A feasibility and clinical validity study. *Respir. Res.* **2019**, *20*, 265. [CrossRef] [PubMed]

129. Cameron, M.; Williams, H.E. Comparing Compressed Sequences for Faster Nucleotide BLAST Searches. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2007**, *4*, 349–364. [CrossRef]

130. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [CrossRef]

131. Gong, Y.-N.; Yang, S.-L.; Chen, G.-W.; Chen, Y.-W.; Huang, Y.-C.; Ning, H.-C.; Tsao, K.-C. A metagenomics study for the identification of respiratory viruses in mixed clinical specimens: An application of the iterative mapping approach. *Arch. Virol.* **2017**, *162*, 2003–2012. [CrossRef]

132. Charalampous, T.; Kay, G.L.; Richardson, H.; Aydin, A.; Baldan, R.; Jeanes, C.; Rae, D.; Grundy, S.; Turner, D.J.; Wain, J.; et al. Nanopore metagenomics enables rapid clinical diagnosis of bacterial lower respiratory infection. *Nat. Biotechnol.* **2019**, *37*, 783–792. [CrossRef]

133. Lewandowski, K.; Xu, Y.; Pullan, S.T.; Lumley, S.F.; Foster, D.; Sanderson, N.; Vaughan, A.; Morgan, M.; Bright, N.; Kavanagh, J.; et al. Metagenomic Nanopore Sequencing of Influenza Virus Direct from Clinical Respiratory Samples. *J. Clin. Microbiol.* **2019**, *58*, e00963-19. [CrossRef]

134. Ding, L.; Liu, Y.; Wu, X.; Wu, M.; Luo, X.; Ouyang, H.; Xia, J.; Liu, X.; Ding, T. Pathogen Metagenomics Reveals Distinct Lung Microbiota Signatures between Bacteriologically Confirmed and Negative Tuberculosis Patients. *Front. Cell. Infect. Microbiol.* **2021**, *11*, 708827. [CrossRef]

135. Di Simone, S.K.; Rudloff, I.; Nold-Petry, C.A.; Forster, S.C.; Nold, M.F. Understanding respiratory microbiome–immune system interactions in health and disease. *Sci. Transl. Med.* **2023**, *15*, abq5126. [CrossRef]

136. Diao, Z.; Han, D.; Zhang, R.; Li, J. Metagenomics next-generation sequencing tests take the stage in the diagnosis of lower respiratory tract infections. *J. Adv. Res.* **2022**, *38*, 201–212. [CrossRef] [PubMed]

137. Gao, C.; Li, X.; Zhao, X.; Yang, P.; Wang, X.; Chen, X.; Chen, N.; Chen, F. Standardized studies of the oral microbiome: From technology-driven to hypothesis-driven. *iMeta* **2022**, *1*, e19. [CrossRef]

138. Baker, J.L.; Morton, J.T.; Dinis, M.; Alvarez, R.; Tran, N.C.; Knight, R.; Edlund, A. Deep metagenomics examines the oral microbiome during dental caries, revealing novel taxa and co-occurrences with host molecules. *Genome Res.* **2021**, *31*, 64–74. [CrossRef] [PubMed]

139. Ottoni, C.; Borić, D.; Cheronet, O.; Sparacello, V.; Dori, I.; Coppa, A.; Antonović, D.; Vujević, D.; Price, T.D.; Pinhasi, R.; et al. Tracking the transition to agriculture in Southern Europe through ancient DNA analysis of dental calculus. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2102116118. [CrossRef]

140. Liu, X.; Tong, X.; Zhu, J.; Tian, L.; Jie, Z.; Zou, Y.; Lin, X.; Liang, H.; Li, W.; Ju, Y.; et al. Metagenome-genome-wide association studies reveal human genetic impact on the oral microbiome. *Cell Discov.* **2021**, *7*, 117. [CrossRef]

141. Chattopadhyay, I.; Lu, W.; Manikam, R.; Malarvili, M.B.; Ambati, R.R.; Gundamaraju, R. Can metagenomics unravel the impact of oral bacteriome in human diseases? *Biotechnol. Genet. Eng. Rev.* **2022**, *39*, 85–117. [CrossRef]

142. Utter, D.R.; Borisy, G.G.; Eren, A.M.; Cavanaugh, C.M.; Welch, J.L.M. Metapangenomics of the oral microbiome provides insights into habitat adaptation and cultivar diversity. *Genome Biol.* **2020**, *21*, 293. [CrossRef]

143. Cheng, X.; He, F.; Sun, P.; Chen, Q. Identification of unknown acid-resistant genes of oral microbiotas in patients with dental caries using metagenomics analysis. *AMB Express* **2021**, *11*, 39. [CrossRef]

144. Cui, H.; Li, Y.; Zhang, X. An overview of major metagenomic studies on human microbiomes in health and disease. *Quant. Biol.* **2016**, *4*, 192–206. [CrossRef]

145. Cani, P.D. Human gut microbiome: Hopes, threats and promises. *Gut* **2018**, *67*, 1716–1725. [CrossRef]

146. Nielsen, J.; Ji, B. New insight into the gut microbiome through metagenomics. *Adv. Genom. Genet.* **2015**, *5*, 77–91. [CrossRef]

147. The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* **2012**, *486*, 207–214. [CrossRef] [PubMed]

148. Le Chatelier, E.; Nielsen, T.; Qin, J.; Prifti, E.; Hildebrand, F.; Falony, G.; Almeida, M.; Arumugam, M.; Batto, J.-M.; Kennedy, S.; et al. Richness of human gut microbiome correlates with metabolic markers. *Nature* **2013**, *500*, 541–546. [CrossRef] [PubMed]

149. Claesson, M.J.; Jeffery, I.B.; Conde, S.; Power, S.E.; O'Connor, E.M.; Cusack, S.; Harris, H.M.B.; Coakley, M.; Lakshminarayanan, B.; O'Sullivan, O.; et al. Gut microbiota composition correlates with diet and health in the elderly. *Nature* **2012**, *488*, 178–184. [CrossRef]

150. Wang, W.-L.; Xu, S.-Y.; Ren, Z.-G.; Tao, L.; Jiang, J.-W.; Zheng, S.-S. Application of metagenomics in the human gut microbiome. *World J. Gastroenterol.* **2015**, *21*, 803. [CrossRef]

151. Wallen, Z.D.; Demirkan, A.; Twa, G.; Cohen, G.; Dean, M.N.; Standaert, D.G.; Sampson, T.R.; Payami, H. Metagenomics of Parkinson's disease implicates the gut microbiome in multiple disease mechanisms. *Nat. Commun.* **2022**, *13*, 6958. [CrossRef]

152. Yen, S.; Johnson, J.S. Metagenomics: A path to understanding the gut microbiome. *Mamm. Genome* **2021**, *32*, 282–296. [CrossRef]

153. Parks, D.H.; Imelfort, M.; Skennerton, C.T.; Hugenholtz, P.; Tyson, G.W. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **2015**, *25*, 1043–1055. [CrossRef]

154. Olm, M.R.; Brown, C.T.; Brooks, B.; Banfield, J.F. dRep: A tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* **2017**, *11*, 2864–2868. [CrossRef]

155. Ondov, B.D.; Treangen, T.J.; Melsted, P.; Mallonee, A.B.; Bergman, N.H.; Koren, S.; Phillippy, A.M. Mash: Fast genome and metagenome distance estimation using MinHash. *Genome Biol.* **2016**, *17*, 132. [CrossRef]

156. Jain, C.; Rodriguez-R, L.M.; Phillippy, A.M.; Konstantinidis, K.T.; Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **2018**, *9*, 5114. [CrossRef] [PubMed]

157. Roux, S.; Enault, F.; Hurwitz, B.L.; Sullivan, M.B. VirSorter: Mining viral signal from microbial genomic data. *PeerJ* **2015**, *3*, e985. [CrossRef] [PubMed]

158. Ren, J.; Ahlgren, N.A.; Lu, Y.Y.; Fuhrman, J.A.; Sun, F. VirFinder: A novel k-mer based tool for identifying viral sequences from assembled metagenomic data. *Microbiome* **2017**, *5*, 69. [CrossRef] [PubMed]

159. Alneberg, J.; Bjarnason, B.S.; de Bruijn, I.; Schirmer, M.; Quick, J.; Ijaz, U.Z.; Lahti, L.; Loman, N.J.; Andersson, A.F.; Quince, C. Binning metagenomic contigs by coverage and composition. *Nat. Methods* **2014**, *11*, 1144–1146. [CrossRef] [PubMed]

160. Chaumeil, P.-A.; Mussig, A.J.; Hugenholtz, P.; Parks, D.H. GTDB-Tk: A toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **2019**, *36*, 1925–1927. [CrossRef]

161. Hyatt, D.; Chen, G.-L.; LoCascio, P.F.; Land, M.L.; Larimer, F.W.; Hauser, L.J. Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform.* **2010**, *11*, 119. [CrossRef]

162. Kim, C.Y.; Lee, M.; Yang, S.; Kim, K.; Yong, D.; Kim, H.R.; Lee, I. Human reference gut microbiome catalog including newly assembled genomes from under-represented Asian metagenomes. *Genome Med.* **2021**, *13*, 134. [CrossRef]

163. Zhang, Q.; Yu, K.; Li, S.; Zhang, X.; Zhao, Q.; Zhao, X.; Liu, Z.; Cheng, H.; Liu, Z.-X.; Li, X. gutMEGA: A database of the human gut MEtaGenome Atlas. *Brief. Bioinform.* **2021**, *22*, bbaa082. [CrossRef]

164. Almeida, A.; Nayfach, S.; Boland, M.; Strozzi, F.; Beracochea, M.; Shi, Z.J.; Pollard, K.S.; Sakharova, E.; Parks, D.H.; Hugenholtz, P.; et al. A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat. Biotechnol.* **2021**, *39*, 105–114. [CrossRef]

165. Nayfach, S.; Páez-Espino, D.; Call, L.; Low, S.J.; Sberro, H.; Ivanova, N.N.; Proal, A.D.; Fischbach, M.A.; Bhatt, A.S.; Hugenholtz, P.; et al. Metagenomic compendium of 189,680 DNA viruses from the human gut microbiome. *Nat. Microbiol.* **2021**, *6*, 960–970. [CrossRef]

166. Sun, Y.; Ju, P.; Xue, T.; Ali, U.; Cui, D.; Chen, J. Alteration of faecal microbiota balance related to long-term deep meditation. *Gen. Psychiatry* **2023**, *36*, e100893. [CrossRef] [PubMed]

167. Dahiya, D.; Nigam, P.S. The Gut Microbiota Influenced by the Intake of Probiotics and Functional Foods with Prebiotics Can Sustain Wellness and Alleviate Certain Ailments like Gut-Inflammation and Colon-Cancer. *Microorganisms* **2022**, *10*, 665. [CrossRef] [PubMed]

168. Ng, C.; Li, H.; Wu, W.K.K.; Wong, S.H.; Yu, J. Genomics and metagenomics of colorectal cancer. *J. Gastrointest. Oncol.* **2019**, *10*, 1164–1170. [CrossRef]

169. Nagata, N.; Nishijima, S.; Miyoshi-Akiyama, T.; Kojima, Y.; Kimura, M.; Aoki, R.; Ohsugi, M.; Ueki, K.; Miki, K.; Iwata, E.; et al. Population-level Metagenomics Uncovers Distinct Effects of Multiple Medications on the Human Gut Microbiome. *Gastroenterology* **2022**, *163*, 1038–1052. [CrossRef] [PubMed]

170. Ke, S.; Weiss, S.T.; Liu, Y.-Y. Dissecting the role of the human microbiome in COVID-19 via metagenome-assembled genomes. *Nat. Commun.* **2022**, *13*, 5235. [CrossRef] [PubMed]

171. Mancabelli, L.; Milani, C.; Fontana, F.; Lugli, G.A.; Tarracchini, C.; Viappiani, A.; Ciociola, T.; Ticinesi, A.; Nouvenne, A.; Meschi, T.; et al. Untangling the link between the human gut microbiota composition and the severity of the symptoms of the COVID-19 infection. *Environ. Microbiol.* **2022**, *24*, 6453–6462. [CrossRef]

172. Suzuki, T.A.; Fitzstevens, J.L.; Schmidt, V.T.; Enav, H.; Huus, K.E.; Ngwese, M.M.; Grießhammer, A.; Pfleiderer, A.; Adegbite, B.R.; Zinsou, J.F.; et al. Codiversification of gut microbiota with humans. *Science* **2022**, *377*, 1328–1332. [CrossRef]

173. Shahab, M.; Shahab, N. Coevolution of the Human Host and Gut Microbiome: Metagenomics of Microbiota. *Cureus* **2022**, *14*, e26310. [CrossRef]

174. Robertson, R.C.; Edens, T.J.; Carr, L.; Mutasa, K.; Gough, E.K.; Evans, C.; Geum, H.M.; Baharmand, I.; Gill, S.K.; Ntozini, R.; et al. The gut microbiome and early-life growth in a population with high prevalence of stunting. *Nat. Commun.* **2023**, *14*, 654. [CrossRef]

175. Bai, G.-H.; Lin, S.-C.; Hsu, Y.-H.; Chen, S.-Y. The Human Virome: Viral Metagenomics, Relations with Human Diseases, and Therapeutic Applications. *Viruses* **2022**, *14*, 278. [CrossRef]

176. Santiago-Rodriguez, T.M.; Hollister, E.B. Human Virome and Disease: High-Throughput Sequencing for Virus Discovery, Identification of Phage-Bacteria Dysbiosis and Development of Therapeutic Approaches with Emphasis on the Human Gut. *Viruses* **2019**, *11*, 656. [CrossRef] [PubMed]

177. Clooney, A.G.; Sutton, T.D.S.; Shkoporov, A.N.; Holohan, R.K.; Daly, K.M.; O'Regan, O.; Ryan, F.J.; Draper, L.A.; Plevy, S.E.; Ross, R.P.; et al. Whole-Virome Analysis Sheds Light on Viral Dark Matter in Inflammatory Bowel Disease. *Cell Host Microbe* **2019**, *26*, 764–778.e5. [CrossRef]

178. Cao, Z.; Sugimura, N.; Burgermeister, E.; Ebert, M.P.; Zuo, T.; Lan, P. The gut virome: A new microbiome component in health and disease. *eBioMedicine* **2022**, *81*, 104113. [CrossRef] [PubMed]

179. Zuo, T.; Wong, S.H.; Lam, K.; Lui, R.; Cheung, K.; Tang, W.; Ching, J.Y.L.; Chan, P.K.S.; Chan, M.C.W.; Wu, J.C.Y.; et al. Bacteriophage transfer during faecal microbiota transplantation in Clostridium difficile infection is associated with treatment outcome. *Gut* **2017**, *67*, 634–643. [CrossRef] [PubMed]

180. Ott, S.J.; Waetzig, G.H.; Rehman, A.; Moltzau-Anderson, J.; Bharti, R.; Grasis, J.A.; Cassidy, L.; Tholey, A.; Fickenscher, H.; Seegert, D.; et al. Efficacy of Sterile Fecal Filtrate Transfer for Treating Patients with Clostridium difficile Infection. *Gastroenterology* **2017**, *152*, 799–811.e7. [CrossRef] [PubMed]

181. Sabale, S.N.; Suryawanshi, P.P.; Krishnaraj, P.U. Soil Metagenomics: Concepts and Applications. In *Metagenomics-Basics, Methods and Applications*; IntechOpen: London, UK, 2020.

182. Fierer, N.; Leff, J.W.; Adams, B.J.; Nielsen, U.N.; Bates, S.T.; Lauber, C.L.; Owens, S.; Gilbert, J.A.; Wall, D.H.; Caporaso, J.G. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 21390–21395. [CrossRef]

183. Wu, X.; Chauhan, A.; Layton, A.C.; Vetter, M.C.Y.L.; Stackhouse, B.T.; Williams, D.E.; Whyte, L.; Pfiffner, S.M.; Onstott, T.C.; Vishnivetskaya, T.A. Comparative Metagenomics of the Active Layer and Permafrost from Low-Carbon Soil in the Canadian High Arctic. *Environ. Sci. Technol.* **2021**, *55*, 12683–12693. [CrossRef]

184. Karthikeyan, S.; Orellana, L.H.; Johnston, E.R.; Hatt, J.K.; Löffler, F.E.; Ayala-del-Río, H.L.; González, G.; Konstantinidis, K.T. Metagenomic Characterization of Soil Microbial Communities in the Luquillo Experimental Forest (Puerto Rico) and Implications for Nitrogen Cycling. *Appl. Environ. Microbiol.* **2021**, *87*, e00546-21. [CrossRef]

185. Bonomo, M.G.; Calabrone, L.; Scrano, L.; Bufo, S.A.; Di Tomaso, K.; Buongarzone, E.; Salzano, G. Metagenomic monitoring of soil bacterial community after the construction of a crude oil flowline. *Environ. Monit. Assess.* **2022**, *194*, 48. [CrossRef]

186. Feng, G.; Xie, T.; Wang, X.; Bai, J.; Tang, L.; Zhao, H.; Wei, W.; Wang, M.; Zhao, Y. Metagenomic analysis of microbial community and function involved in cd-contaminated soil. *BMC Microbiol.* **2018**, *18*, 11. [CrossRef]

187. Kim, J.-W.; Hong, Y.-K.; Kim, H.-S.; Oh, E.-J.; Park, Y.-H.; Kim, S.-C. Metagenomic Analysis for Evaluating Change in Bacterial Diversity in TPH-Contaminated Soil after Soil Remediation. *Toxics* **2021**, *9*, 319. [CrossRef] [PubMed]

188. González, D.; Robas, M.; Fernández, V.; Bárcena, M.; Probanza, A.; Jiménez, P.A. Comparative Metagenomic Study of Rhizospheric and Bulk Mercury-Contaminated Soils in the Mining District of Almadén. *Front. Microbiol.* **2022**, *13*, 797444. [CrossRef] [PubMed]

189. Vu, M.T.; Geraldi, A.; Do, H.D.K.; Luqman, A.; Nguyen, H.D.; Fauzia, F.N.; Amalludin, F.I.; Sadila, A.Y.; Wijaya, N.H.; Santoso, H.; et al. Soil Mineral Composition and Salinity Are the Main Factors Regulating the Bacterial Community Associated with the Roots of Coastal Sand Dune Halophytes. *Biology* **2022**, *11*, 695. [CrossRef] [PubMed]

190. Werbin, Z.R.; Hackos, B.; Lopez-Nava, J.; Dietze, M.C.; Bhatnagar, J.M. The National Ecological Observatory Network's soil metagenomes: Assembly and basic analysis. *F1000Research* **2022**, *10*, 299. [CrossRef] [PubMed]

191. Staley, C.; Sadowsky, M.J. Application of metagenomics to assess microbial communities in water and other environmental matrices. *J. Mar. Biol. Assoc. UK* **2016**, *96*, 121–129. [CrossRef]

192. Cabello-Yeves, P.J.; Callieri, C.; Picazo, A.; Mehrshad, M.; Haro-Moreno, J.M.; Roda-Garcia, J.J.; Dzhembekova, N.; Slabakova, V.; Slabakova, N.; Moncheva, S.; et al. The microbiome of the Black Sea water column analyzed by shotgun and genome centric metagenomics. *Environ. Microbiome* **2021**, *16*, 5. [CrossRef]

193. Wang, Y.; Liao, S.; Gai, Y.; Liu, G.; Jin, T.; Liu, H.; Gram, L.; Strube, M.L.; Fan, G.; Sahu, S.K.; et al. Metagenomic Analysis Reveals Microbial Community Structure and Metabolic Potential for Nitrogen Acquisition in the Oligotrophic Surface Water of the Indian Ocean. *Front. Microbiol.* **2021**, *12*, 518865. [CrossRef]

194. Liem, M.; Regensburg-Tuïnk, T.; Henkel, C.; Jansen, H.; Spaink, H. Microbial diversity characterization of seawater in a pilot study using Oxford Nanopore Technologies long-read sequencing. *BMC Res. Notes* **2021**, *14*, 42. [CrossRef]

195. Loza, A.; García-Guevara, F.; Segovia, L.; Escobar-Zepeda, A.; Sanchez-Olmos, M.D.C.; Merino, E.; Sanchez-Flores, A.; Pardo-Lopez, L.; Juarez, K.; Gutierrez-Rios, R.-M. Definition of the Metagenomic Profile of Ocean Water Samples from the Gulf of Mexico Based on Comparison with Reference Samples from Sites Worldwide. *Front. Microbiol.* **2022**, *12*, 781497. [CrossRef]

196. Man, T.D.; Thoa, N.K.; Viet, N.Q.; Minh, P.T.T.; Ha, P.T.; Thuy, T.T.; Tu, H.M.; Xuan, L.T.T.; Mai, B.T. Metagenomics analysis of marine eukaryotic community in water and sediments at Lang Co-Da Nang sea by throughput 18S rRNA gene sequencing. *Tạp Chí Khoa Học Và Công Nghệ Biển* **2021**, *21*, 85–94. [CrossRef]

197. Ngugi, D.K.; Salcher, M.M.; Andrei, A.-S.; Ghai, R.; Klotz, F.; Chiriac, M.-C.; Ionescu, D.; Büsing, P.; Grossart, H.-P.; Xing, P.; et al. Postglacial adaptations enabled colonization and quasi-clonal dispersal of ammonia-oxidizing archaea in modern European large lakes. *Sci. Adv.* **2023**, *9*, adc9392. [CrossRef] [PubMed]

198. Ghosh, A.; Saha, R.; Bhadury, P. Metagenomic insights into surface water microbial communities of a South Asian mangrove ecosystem. *PeerJ* **2022**, *10*, e13169. [CrossRef] [PubMed]

199. Nguyen, S.G.; Raza, S.; Ta, L.T.; Le, L.-A.T.; Ho, C.T.; Unno, T. Metagenomic investigation of the seasonal distribution of bacterial community and antibiotic-resistant genes in Day River Downstream, Ninh Binh, Vietnam. *Appl. Biol. Chem.* **2022**, *65*, 26. [CrossRef]

200. Curran, J.F.; Zaggia, L.; Quero, G.M. Metagenomic Characterization of Microbial Pollutants and Antibiotic- and Metal-Resistance Genes in Sediments from the Canals of Venice. *Water* **2022**, *14*, 1161. [CrossRef]

201. Acinas, S.G.; Sánchez, P.; Salazar, G.; Cornejo-Castillo, F.M.; Sebastián, M.; Logares, R.; Royo-Llonch, M.; Paoli, L.; Sunagawa, S.; Hingamp, P.; et al. Deep ocean metagenomes provide insight into the metabolic architecture of bathypelagic microbial communities. *Commun. Biol.* **2021**, *4*, 604. [CrossRef]

202. Nata'ala, M.K.; Santos, A.P.A.; Kasmanas, J.C.; Bartholomäus, A.; Saraiva, J.P.; Silva, S.G.; Keller-Costa, T.; Costa, R.; Gomes, N.C.M.; Ponce de Leon Ferreira de Carvalho, A.C.; et al. MarineMetagenomeDB: A public repository for curated and standardized metadata for marine metagenomes. *Environ. Microbiome* **2022**, *17*, 57. [CrossRef]

203. Weerakoon, W.M.T.D.N.; Seneviratne, K.N.; Jayathilaka, N. Metagenomic analysis of wastewater for water quality assessment. In *Metagenomics to Bioremediation*; Elsevier: Amsterdam, The Netherlands, 2023; pp. 285–309.

204. Hauptfeld, E.; Pelkmans, J.; Huisman, T.T.; Anocic, A.; Snoek, B.L.; von Meijenfeldt, F.A.B.; Gerritse, J.; van Leeuwen, J.; Leurink, G.; van Lit, A.; et al. A metagenomic portrait of the microbial community responsible for two decades of bioremediation of poly-contaminated groundwater. *Water Res.* **2022**, *221*, 118767. [CrossRef]

205. Acharya, K.; Blackburn, A.; Mohammed, J.; Haile, A.T.; Hiruy, A.M.; Werner, D. Metagenomic water quality monitoring with a portable laboratory. *Water Res.* **2020**, *184*, 116112. [CrossRef]

206. Gomez-Alvarez, V.; Liu, H.; Pressman, J.G.; Wahman, D.G. Metagenomic Profile of Microbial Communities in a Drinking Water Storage Tank Sediment after Sequential Exposure to Monochloramine, Free Chlorine, and Monochloramine. *ACS EST Water* **2021**, *1*, 1–1283. [CrossRef]

207. Bibby, K.; Crank, K.; Greaves, J.; Li, X.; Wu, Z.; Hamza, I.A.; Stachler, E. Metagenomics and the development of viral water quality tools. *npj Clean Water* **2019**, *2*, 9. [CrossRef]

208. Behzad, H.; Gojobori, T.; Mineta, K. Challenges and Opportunities of Airborne Metagenomics. *Genome Biol. Evol.* **2015**, *7*, 1216–1226. [CrossRef] [PubMed]

209. Unterwurzacher, V.; Bruck, S.; Biedermann, M.; Pogner, C.; Konlechner, A.; Tondl, G.; Berger, H.; Pfeifer, C.; Strauss, J.; Gorfer, M.; et al. Development and Validation of a Simple Bioaerosol Collection Filter System Using a Conventional Vacuum Cleaner for Sampling. *Aerosol Sci. Eng.* **2021**, *5*, 404–418. [CrossRef]

210. Yooseph, S.; Andrews-Pfannkoch, C.; Tenney, A.; McQuaid, J.; Williamson, S.; Thiagarajan, M.; Brami, D.; Zeigler-Allen, L.; Hoffman, J.; Goll, J.B.; et al. A Metagenomic Framework for the Study of Airborne Microbial Communities. *PLoS ONE* **2013**, *8*, e81862. [CrossRef] [PubMed]

211. Jiang, X.; Wang, C.; Guo, J.; Hou, J.; Guo, X.; Zhang, H.; Tan, J.; Li, M.; Li, X.; Zhu, H. Global Meta-analysis of Airborne Bacterial Communities and Associations with Anthropogenic Activities. *Environ. Sci. Technol.* **2022**, *56*, 9891–9902. [CrossRef]

212. Zhao, J.; Jin, L.; Wu, D.; Xie, J.; Li, J.; Fu, X.; Cong, Z.; Fu, P.; Zhang, Y.; Luo, X.; et al. Global airborne bacterial community—Interactions with Earth's microbiomes and anthropogenic activities. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2204465119. [CrossRef]

213. Gusareva, E.S.; Acerbi, E.; Lau, K.J.X.; Luhung, I.; Premkrishnan, B.N.V.; Kolundžija, S.; Purbojati, R.W.; Wong, A.; Houghton, J.N.I.; Miller, D.; et al. Microbial communities in the tropical air ecosystem follow a precise diel cycle. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 23299–23308. [CrossRef]

214. Rosario, K.; Fierer, N.; Miller, S.; Luongo, J.; Breitbart, M. Diversity of DNA and RNA Viruses in Indoor Air as Assessed via Metagenomic Sequencing. *Environ. Sci. Technol.* **2018**, *52*, 1014–1027. [CrossRef]

215. King, P.; Pham, L.K.; Waltz, S.; Sphar, D.; Yamamoto, R.T.; Conrad, D.; Taplitz, R.; Torriani, F.; Forsyth, R.A. Longitudinal Metagenomic Analysis of Hospital Air Identifies Clinically Relevant Microbes. *PLoS ONE* **2016**, *11*, e0160124. [CrossRef]

216. Li, X.; Wu, Z.; Dang, C.; Zhang, M.; Zhao, B.; Cheng, Z.; Chen, L.; Zhong, Z.; Ye, Y.; Xia, Y. A metagenomic-based method to study hospital air dust resistome. *Chem. Eng. J.* **2021**, *406*, 126854. [CrossRef]

217. Wu, D.; Jin, L.; Xie, J.; Liu, H.; Zhao, J.; Ye, D.; Li, X. Inhalable antibiotic resistomes emitted from hospitals: Metagenomic insights into bacterial hosts, clinical relevance, and environmental risks. *Microbiome* **2022**, *10*, 19. [CrossRef]

218. Iquebal, M.A.; Jagannadham, J.; Jaiswal, S.; Prabha, R.; Rai, A.; Kumar, D. Potential Use of Microbial Community Genomes in Various Dimensions of Agriculture Productivity and Its Management: A Review. *Front. Microbiol.* **2022**, *13*, 708335. [CrossRef] [PubMed]

219. Nwachukwu, B.C.; Babalola, O.O. Metagenomics: A Tool for Exploring Key Microbiome with the Potentials for Improving Sustainable Agriculture. *Front. Sustain. Food Syst.* **2022**, *6*, 886987. [CrossRef]

220. Upadhyay, A.; Kovalev, A.A.; Zhuravleva, E.A.; Kovalev, D.A.; Litti, Y.V.; Masakapalli, S.K.; Pareek, N.; Vivekanand, V. A Review of Basic Bioinformatic Techniques for Microbial Community Analysis in an Anaerobic Digester. *Fermentation* **2023**, *9*, 62. [CrossRef]

221. Giolai, M.; Verweij, W.; Pearson, N.; Nicholson, P.; Leggett, R.M.; Clark, M.D. Air-seq: Measuring air metagenomic diversity in an agricultural ecosystem. *bioRxiv* **2022**. [CrossRef]

222. Goel, R.; Suyal, D.C.; Narayan; Dash, B.; Soni, R. Soil Metagenomics: A Tool for Sustainable Agriculture. In *Mining of Microbial Wealth and MetaGenomics*; Springer: Singapore, 2017; pp. 217–225.

223. Rokins, P.D.; Gopal, N.O.; Anandham, R.; Saraswathi, R. The Impact of Different Planting Systems on the Bacterial Diversity of Rice Cultivated in Saline Soil Based on 16S rRNA Gene-Based Metagenomic Insights. *Agriculture* **2022**, *12*, 1624. [CrossRef]

224. Lian, W.-H.; Mohamad, O.A.A.; Dong, L.; Zhang, L.-Y.; Wang, D.; Liu, L.; Han, M.-X.; Li, S.; Wang, S.; Antunes, A.; et al. Culturomics- and metagenomics-based insights into the microbial community and function of rhizosphere soils in Sinai desert farming systems. *Environ. Microbiome* **2023**, *18*, 4. [CrossRef]

225. Omotayo, O.P.; Igiehon, O.N.; Babalola, O.O. Microbial Genes of Agricultural Importance in Maize Rhizosphere Unveiled through Shotgun Metagenomics. *Span. J. Soil Sci.* **2022**, *12*, 6. [CrossRef]

226. Lappe, R.R.; Elmore, M.G.; Lozier, Z.R.; Jander, G.; Miller, W.A.; Whitham, S.A. Metagenomic identification of novel viruses of maize and teosinte in North America. *BMC Genom.* **2022**, *23*, 767. [CrossRef]

227. Babalola, O.O.; Omotayo, O.P.; Igiehon, N.O. Survey of Maize Rhizosphere Microbiome Using Shotgun Metagenomics. *Microbiol. Resour. Announc.* **2021**, *10*, e01309-20. [CrossRef]

228. Graziano, S.; Caldara, M.; Gullì, M.; Bevivino, A.; Maestri, E.; Marmiroli, N. A Metagenomic and Gene Expression Analysis in Wheat (*T. durum*) and Maize (*Z. mays*) Biofertilized with PGPM and Biochar. *Int. J. Mol. Sci.* **2022**, *23*, 10376. [CrossRef]

229. Pang, Z.; Dong, F.; Liu, Q.; Lin, W.; Hu, C.; Yuan, Z. Soil Metagenomics Reveals Effects of Continuous Sugarcane Cropping on the Structure and Functional Pathway of Rhizospheric Microbial Community. *Front. Microbiol.* **2021**, *12*, 627569. [CrossRef] [PubMed]

230. Effendi, Y.; Aini, N.; Pambudi, A.; Sasaerila, H.Y. Metagenomics analysis of soil microbial communities in plant agroforestry system rubber tree (*Hevea brasiliensis*)—Ganyong (*Canna* sp.). *IOP Conf. Ser. Earth Environ. Sci.* **2020**, *468*, 012045. [CrossRef]

231. Keenum, I.; Wind, L.; Ray, P.; Guron, G.; Chen, C.; Knowlton, K.; Ponder, M.; Pruden, A. Metagenomic tracking of antibiotic resistance genes through a pre-harvest vegetable production system: An integrated lab-, microcosm- and greenhouse-scale analysis. *Environ. Microbiol.* **2022**, *24*, 3705–3721. [CrossRef] [PubMed]

232. Akinola, S.; Ayangbenro, A.; Babalola, O. Metagenomic Insight into the Community Structure of Maize-Rhizosphere Bacteria as Predicted by Different Environmental Factors and Their Functioning within Plant Proximity. *Microorganisms* **2021**, *9*, 1419. [CrossRef]

233. Wani, A.K.; Hashem, N.M.; Akhtar, N.; Singh, R.; Madkour, M.; Prakash, A. Understanding microbial networks of farm animals through genomics, metagenomics and other meta-omic approaches for livestock wellness and sustainability—A Review. *Ann. Anim. Sci.* **2022**, *22*, 839–853. [CrossRef]

234. Glendinning, L.; Genç, B.; Wallace, R.J.; Watson, M. Metagenomic analysis of the cow, sheep, reindeer and red deer rumen. *Sci. Rep.* **2021**, *11*, 1990. [CrossRef]

235. Sato, Y.; Takebe, H.; Oishi, K.; Yasuda, J.; Kumagai, H.; Hirooka, H.; Yoshida, T. Identification of 146 Metagenome-assembled Genomes from the Rumen Microbiome of Cattle in Japan. *Microbes Environ.* **2022**, *37*, ME22039. [CrossRef]

236. Kwok, K.T.T.; Nieuwenhuijse, D.F.; Phan, M.V.T.; Koopmans, M.P.G. Virus Metagenomics in Farm Animals: A Systematic Review. *Viruses* **2020**, *12*, 107. [CrossRef]

237. Kwok, K.T.T.; de Rooij, M.M.T.; Messink, A.B.; Wouters, I.M.; Smit, L.A.M.; Cotten, M.; Heederik, D.J.J.; Koopmans, M.P.G.; Phan, M.V.T. Establishing farm dust as a useful viral metagenomic surveillance matrix. *Sci. Rep.* **2022**, *12*, 16308. [CrossRef]

238. Dixit, S.; Kumar, S.; Sharma, R.; Banakar, P.S.; Singh, M.; Keshri, A.; Tyagi, A.K. Rumen multi-omics addressing diet–host–microbiome interplay in farm animals: A review. *Anim. Biotechnol.* **2022**, 1–19. [CrossRef]

239. Skarżyńska, M.; Leekitcharoenphon, P.; Hendriksen, R.S.; Aarestrup, F.M.; Wasyl, D. A metagenomic glimpse into the gut of wild and domestic animals: Quantification of antimicrobial resistance and more. *PLoS ONE* **2020**, *15*, e0242987. [CrossRef] [PubMed]

240. Qiu, T.; Huo, L.; Guo, Y.; Gao, M.; Wang, G.; Hu, D.; Li, C.; Wang, Z.; Liu, G.; Wang, X. Metagenomic assembly reveals hosts and mobility of common antibiotic resistome in animal manure and commercial compost. *Environ. Microbiome* **2022**, *17*, 42. [CrossRef] [PubMed]

241. Shi, T.; Zhang, T.; Wang, X.; Wang, X.; Shen, W.; Guo, X.; Liu, Y.; Li, Z.; Jiang, Y. Metagenomic Analysis of in Vitro Ruminal Fermentation Reveals the Role of the Copresent Microbiome in Plant Biomass Degradation. *J. Agric. Food Chem.* **2022**, *70*, 12095–12106. [CrossRef]

242. Rieder, J.; Kapopoulou, A.; Bank, C.; Adrian-Kalchhauser, I. Metagenomics and metabarcoding experimental choices and their impact on microbial community characterization in freshwater recirculating aquaculture systems. *Environ. Microbiome* **2023**, *18*, 8. [CrossRef] [PubMed]

243. Le, L.-A.; Nguyen-Hoang, T.-P.; Huynh, V.-P.; Nguyen, T.-H.; Nguyen, T.-V.; Ho-Huynh, T.-D. Microbiome dataset analysis from a shrimp pond in Ninh Thuan, Vietnam using shotgun metagenomics. *Data Brief* **2020**, *31*, 105731. [CrossRef]

244. Amin, M.; Kumala, R.R.C.; Mukti, A.T.; Lamid, M.; Nindarwi, D.D. Metagenomic profiles of core and signature bacteria in the guts of white shrimp, Litopenaeus vannamei, with different growth rates. *Aquaculture* **2022**, *550*, 737849. [CrossRef]

245. Chen, X.; He, Z.; Zhao, J.; Liao, M.; Xue, Y.; Zhou, J.; Hoare, R.; Monaghan, S.J.; Wang, N.; Pang, H.; et al. Metagenomic Analysis of Bacterial Communities and Antibiotic Resistance Genes in *Penaeus monodon* Biofloc-Based Aquaculture Environments. *Front. Mar. Sci.* **2022**, *8*, 762345. [CrossRef]

246. Wani, A.K.; Rahayu, F.; Kadarwati, F.T.; Suhara, C.; Singh, R.; Dhanjal, D.S.; Akhtar, N.; Mir, T.G.; Chopra, C. Metagenomic screening strategies for bioprospecting enzymes from environmental samples. *IOP Conf. Ser. Earth Environ. Sci.* **2022**, *974*, 012003. [CrossRef]

247. Robinson, S.L.; Piel, J.; Sunagawa, S. A roadmap for metagenomic enzyme discovery. *Nat. Prod. Rep.* **2021**, *38*, 1994–2023. [CrossRef]

248. Boddu, R.S.; Prabhakar, K.A.; Divakar, K. Metagenomic Bioprospecting of Uncultivable Microbial Flora in Soil Microbiome for Novel Enzymes. *Geomicrobiol. J.* **2022**, *39*, 97–106. [CrossRef]

249. Sung, J.; Lee, Y.; Cho, Y.; Shin, M.; Lee, S.; Lee, H.; Koh, H.; Bae, J.; Shin, J.; Kim, H.J.; et al. A large-scale metagenomic study for enzyme profiles using the focused identification of the NGS-based definitive enzyme research (FINDER) strategy. *Biotechnol. Bioeng.* **2021**, *118*, 4360–4374. [CrossRef] [PubMed]

250. Berini, F.; Casciello, C.; Marcone, G.L.; Marinelli, F. Metagenomics: Novel enzymes from non-culturable microbes. *FEMS Microbiol. Lett.* **2017**, *364*, fnx211. [CrossRef] [PubMed]

251. Patel, T.; Chaudhari, H.G.; Prajapati, V.; Patel, S.; Mehta, V.; Soni, N. A brief account on enzyme mining using metagenomic approach. *Front. Syst. Biol.* **2022**, *2*, 45. [CrossRef]

252. Yang, C.; Chowdhury, D.; Zhang, Z.; Cheung, W.K.; Lu, A.; Bian, Z.; Zhang, L. A review of computational tools for generating metagenome-assembled genomes from metagenomic sequencing data. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 6301–6314. [CrossRef] [PubMed]

253. Setubal, J.C. Metagenome-assembled genomes: Concepts, analogies, and challenges. *Biophys. Rev.* **2021**, *13*, 905–909. [CrossRef]

254. Chivian, D.; Jungbluth, S.P.; Dehal, P.S.; Wood-Charlson, E.M.; Canon, R.S.; Allen, B.H.; Clark, M.M.; Gu, T.; Land, M.L.; Price, G.A.; et al. Metagenome-assembled genome extraction and analysis from microbiomes using KBase. *Nat. Protoc.* **2023**, *18*, 208–238. [CrossRef]

255. Goussarov, G.; Mysara, M.; Vandamme, P.; Van Houdt, R. Introduction to the principles and methods underlying the recovery of metagenome-assembled genomes from metagenomic data. *Microbiologyopen* **2022**, *11*, e1298. [CrossRef]

256. Li, C.; Li, X.; Guo, R.; Ni, W.; Liu, K.; Liu, Z.; Dai, J.; Xu, Y.; Abduriyim, S.; Wu, Z.; et al. Expanded catalogue of metagenome-assembled genomes reveals resistome characteristics and athletic performance-associated microbes in horse. *Microbiome* **2023**, *11*, 7. [CrossRef]

257. Bassi, C.; Guerriero, P.; Pierantoni, M.; Callegari, E.; Sabbioni, S. Novel Virus Identification through Metagenomics: A Systematic Review. *Life* **2022**, *12*, 2048. [CrossRef]

258. Alindonosi, A.R.; Baeshen, M.N.; Elsharawy, N.T. Prospects for Diatoms Identification Using Metagenomics: A Review. *Appl. Ecol. Environ. Res.* **2021**, *19*, 4281–4298. [CrossRef]

259. Llewellyn, T.; Nowell, R.W.; Aptroot, A.; Temina, M.; Prescott, T.A.K.; Barraclough, T.G.; Gaya, E. Metagenomics Shines Light on the Evolution of "Sunscreen" Pigment Metabolism in the *Teloschistales* (Lichen-Forming Ascomycota). *Genome Biol. Evol.* **2023**, *15*, evad002. [CrossRef] [PubMed]

260. Handelsman, J. Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiol. Mol. Biol. Rev.* **2004**, *68*, 669–685. [CrossRef] [PubMed]

261. Liu, S.; Moon, C.D.; Zheng, N.; Huws, S.; Zhao, S.; Wang, J. Opportunities and challenges of using metagenomic data to bring uncultured microbes into cultivation. *Microbiome* **2022**, *10*, 76. [CrossRef] [PubMed]

262. Lin, Z.; Akin, H.; Rao, R.; Hie, B.; Zhu, Z.; Lu, W.; Smetanin, N.; Verkuil, R.; Kabeli, O.; Shmueli, Y.; et al. Evolutionary-scale prediction of atomic level protein structure with a language model. *bioRxiv* **2022**. [CrossRef]