Chapter 14
Methods of Assessments of Microbial Diversity and Their Functional Role in Soil Fertility and Crop Productivity

Bhaskar Reddy, Aundy Kumar, Sahil Mehta, and Kuleshwar Prasad Sahu

Abstract In the last couple of decades, advancement in the genomic sciences coupled with computational framework has robustly accelerated the deeper understanding of the microbial diversity. The arrival of next-generation sequencing (NGS) technologies among researchers around the globe has facilitated the vast growth of public genomes as well as metagenomes. This progressive development in genome sequencing and environmental metagenomics has enabled the researcher to fully characterize the whole microbial community with detailed functional pathway mappings and enzymes discovery. Therefore, as an attempt, in the present chapter, we have described the role of NGS technologies for the assessment of microbial community coupled with bioinformatic analysis tools in soil fertility and their role in improved crop production. Furthermore, this present chapter also entails the fundamental basis and planning strategy for designing experiments as well as an analysis framework for their robust output for mankind applications.

Keywords NGS · Biochemical pathway · Metagenome · Microbial diversity · Targeted amplicon

14.1 Introduction

Since the beginning of domestication, food production for life survival is mainly performed by green plants through various agricultural process. Such productivity has been improved with the help of a biotechnological process with improved food quality and quantity (Béné et al. 2016; Jovel et al. 2016). The rapid and enlarged food feeding requirements of agricultural industries have affected the environments,

B. Reddy (✉) · A. Kumar · K. P. Sahu
Division of Plant Pathology, ICAR- Indian Agricultural Research Institute, New Delhi, India

S. Mehta
Crop Improvement Group, International Centre for Genetic Engineering and Biotechnology, New Delhi, India

S. K. Dubey, S. K. Verma (eds.), *Plant, Soil and Microbes in Tropical Ecosystems*, Rhizosphere Biology, https://doi.org/10.1007/978-981-16-3364-5_14

which is reflected by the warming condition for being environment-friendly with harmless yields along with the stability of involved resources (Gebbers and Adamchuk 2010).

Plants are surrounded by various microorganism which efficiently enriches the enhanced cell quantity rather than individually by plants. Among such microorganisms, the majority of them survive on the rhizosphere or nearby to plant root surrounding areas (Sharma et al. 2021). The useful microbiota of the rhizosphere has been found to be mediating plant growth as well as improve mineral availability (Lakshmanan et al. 2014). Although, anthropogenic actions and activities influences the soil-residing microorganisms habitat and changes their abundance, co-occurrence dynamics, biochemical pathways, and other functional contents (Mehta et al. 2021a).

The wide application of high-throughput genomics technologies enables various researchers to find the host-pathogen interaction, effect of pesticides and herbicides, and cellulolytic, xenobiotic degrading enzymes and pathways through genome mapping and their potential through molecular gene expression and correlational investigation (Keegan et al. 2016; Anamika et al. 2019; Mehta et al. 2019a; Reddy and Dubey 2021). Such progresses in molecular biology and high-throughput applications have led to the expansion of advanced automated analytic software (Lu et al. 2014). The massive advancements in genomics technologies have brought rapid developments to our understanding of cellular biology, phylogenetic relationship, microbial environments, and biochemical pathways in microbes as well as their host plants (Sahil et al. 2021; Rajput et al. 2021; Mehta et al. 2021b; Bharti et al. 2021; Mehta et al. 2019b; Reddy et al. 2019; Reddy and Dubey 2021) and are progressively unlocking new understandings and uses toward clinical care and personalized medicine (Loman and Pallen 2015; Pareek et al. 2011). Additionally, the scientific community has developed various novel tools, packages, and algorithms to process and explain the genomic data, datasets management, simple software layout, usage, and most importantly privacy of the tremendous data (Vincent et al. 2017; Anamika et al. 2019; Mehta et al. 2019a; Reddy 2019; Kumar et al. 2021).

14.2 Approaches for Soil Microbial Community Assessment

14.2.1 Overview of Microbial Diversity Methods

Since the onset of the twenty-first century, the research regarding microbial diversity was based on techniques like Denaturing Gradient Gel Electrophoresis (DGGE) and Terminal-Restriction Fragment Length Polymorphism (T-RFLP) (Mohanty et al. 2007; Ramakrishnan et al. 2001). The former technique works by using a gradient of denaturing strength on microbial DNA samples (PCR-amplified) along either the

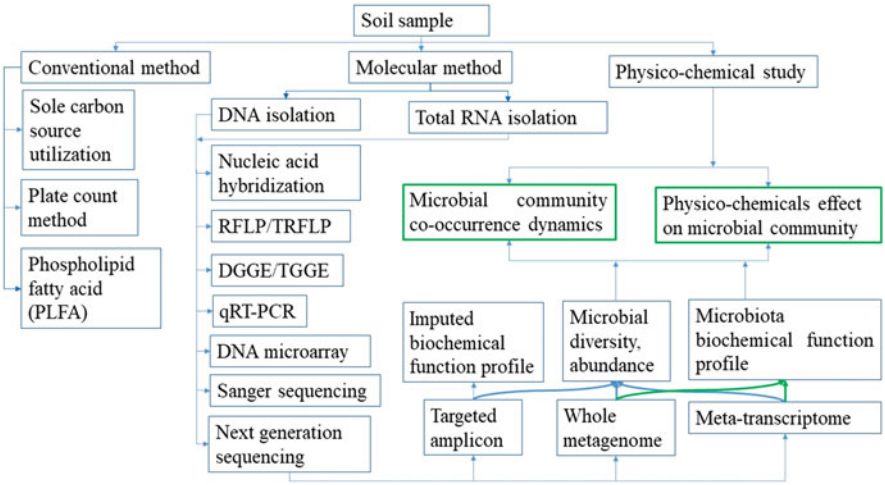

Fig. 14.1 Schematic diagram illustrating the workflow of microbial community structure assessment methodology. Note: The green color box indicates the final output through the various process

horizontal or vertical axis of a polyacrylamide gel followed by electrophoresis (Bo et al. 2020; Leite et al. 2012). While progressing through the gel, the DNA samples at different gel points get separated based on their melting domains, GC clamp, and Tm resulting in a banding pattern of single-stranded branches. In the latter technique, the difference in microbial DNA sequences is detected by a unique blotting pattern generated by using RFLP probes that hybridize specifically with restriction endonucleases-digested different lengths of fragments (Liu et al. 1997). However, with the advent of the sequencing boom, various researchers switched to sequencing-based analysis of microbial diversity. The advantages included low-cost, low rate of errors, high efficiency, high reliability, and time-to-time update. The microbial community structure assessment with various methods is schematically depicted in Fig. 14.1. In the twenty-first century, the majority of researchers are utilizing the molecular methods and depend heavily on next-generation sequencing as compared to the other approaches. NGS provides robust and detailed deep insights of community structure and underlying functional features with significantly reduced labor, time, and cost.

14.2.2 Quantitative Real-Time PCR

Through the literature survey, it has been observed that techniques like DGGE and T-RFLP are only used for qualitative analysis of microbial communities as they reveal the qualitative dynamics diagram among microbial communities (bacteria, archaea, and yeast). However, these techniques cannot be used as a quantitative method (Kanagawa 2003; Neilson et al. 2013). As the name describes, quantitative

real-time PCR (qPCR) is a simple PCR-based technique that first amplifies and then quantifies a targeted DNA enabling the users to quantify absolute as well as a relative number of gene copies from a complex DNA sample to reflect the relative abundance of the microbes (Ashajyothi et al. 2020; Bhardwaj et al. 2020). As a result, this technique is extensively applied for quantitative analysis of microbial composition in various ecological habitats such as soil (Franke-Whittle et al. 2015; Ashajyothi et al. 2020), forest soil (Bhardwaj et al. 2020), and rumen (Pitta et al. 2014; Singh et al. 2015a).

14.2.3 Isolation, Library Preparation, and Sequencing

In case of eukaryotic microbes, for example, the isolated fungus is characterized by the internal transcriber region using ITS-4 and ITS-5 markers to confirm the fungus genus, species, and purity of the isolate. Once the fungus is confirmed, genomic DNA (gDNA) is isolated from the pure fungus and further used for library preparation. Generally, each NGS sample processing known as library preparation starts with the shearing/fragmentation/tagmentation of gDNA into desired fragments and followed by end repair. After the end repair, each sample is usually subjected to multiplexing through adapter and barcode/index ligation reaction, referred as sequencing libraries. The prepared library is subjected to the quality and quantity check to make sure prepared libraries are suitable for sequencing. As samples are barcoded, the various samples cab be pooled together through normalization, and then equimolar pooling is carried out. Next, a pooled library is placed for clonal amplification through emulsion PCR (emPCR, in 454 GS FLX and Ion Torrent) and bridge amplification (cluster generation, in Illumina). In 454 GS FLX and Ion Torrent, after emPCR, sample is processed for recovery and enrichment (Fig. 14.2). The finally enriched sample was loaded in a chip and then placed in a machine for sequencing. Whereas in the Illumina platform, the sequencing is followed immediately after the step of cluster generation/cluster amplification. Each machine-generated sequenced sample is stored in the form of nucleotides fastq files, which is a standard output format (Barriuso et al. 2011; Endrullat et al. 2016).

14.2.4 Brief Summary of Sequencing by Reversible Termination

In the year 2006, the instrument Illumina Genome Analyzer (SOLEXA) was launched based on sequencing by reversible termination technology. In this technology, the subjected study material was prepared through random fragmentation, which was followed by the ligation of oligonucleotide adaptors and indexes, referred as prepared libraries which would be subjected to sequencing in the machine. The

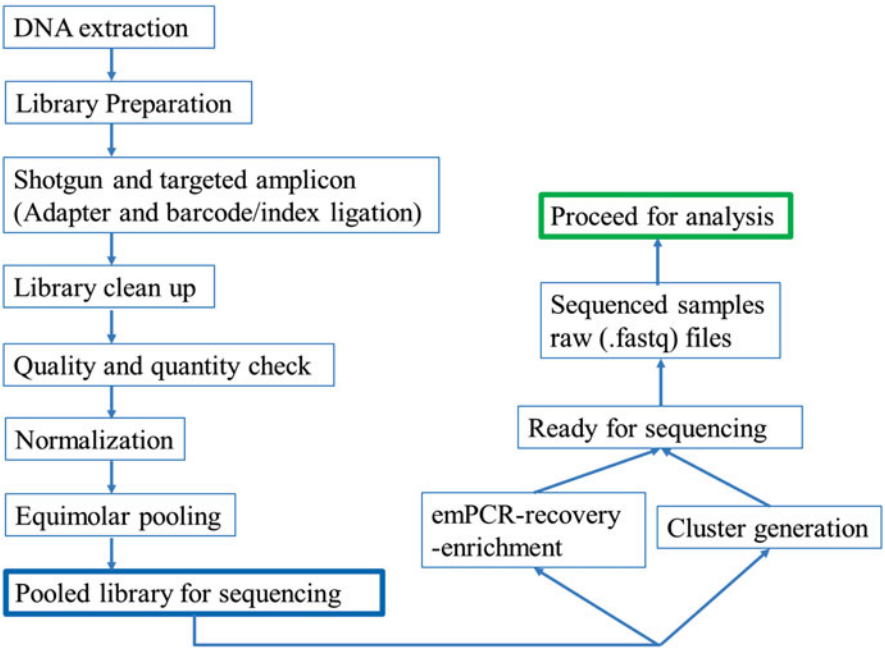

Fig. 14.2 Simple workflow of next-generation sequencing library preparation and sequencing protocol

extensive details of library amplification and sequencing are described further (Adessi et al. 2000; Fedurco et al. 2006; Ju et al. 2006). This technology offers the following two distinct kinds of library preparation, while both kind of libraries are sequenced on compatible Illumina sequencing machine in a default sequencing chemistry.

- Paired-end library preparation:
  - The paired-end (PE) sequencing libraries are prepared using instrument compatible library preparation kit.
  - The insert size of the PE sequencing libraries is usually in range of ~300–550 bp.
  - Each sequencing library will be individually indexed/barcoded for sequencing.
  - Mostly used by researchers for sequencing genomes and metagenomes.
- Mate Pair library preparation:
  - Preparation of Mate Pair library with a jumping distance of 3 and 8 KB average insert size.
  - Each sequencing library will be individually indexed/barcoded for sequencing.
  - Usually used for genome gap finish and polishing.

14.2.5 Third-Generation Sequencing Technology

The third-generation sequencer comprises of DNA sequencing without applying the PCR extension, as extension introduces a bias in sequenced base, and the existence of high GC content influences both depth and coverage. The key advantage of this technology is the longer reads with an average length of 5000–10,000 bases. In this sequencing, single-molecule real-time (SMRT) technology-based first commercial instrument was PacBio Sequel released by Pacific Biosciences and mechanism described here (Eid et al. 2009). The sequenced data (base) output of the PacBio RS II instrument is 0.5–1 billion bases in a single SMRT cell with a higher error rate (10–15%). Another third-generation instrument is the MinIon instrument marketed by Oxford Nanopore Technology in the year of 2014. Specifically, in this sequencing technology, the sample is subjected to a nano-sized pore through electrophoresis, using electrolytic solutions with a fixed electric field. As the template passes through the nanopore, a change in current occurs, and the resultant magnitude is recorded. Compared to PacBio, MinIon instrument is smaller in size and less cost-effective. However, the obtained bases (sequences) display a correctness of near about 88% (Laszlo et al. 2014).

14.3 NGS Reads Processing

Initially in all kind of NGS-based studies, the quality screening and filtration of generated poor bases and reads is a prerequisite. The schematic workflow of NGS reads processing illustrated in Fig. 14.3. The quality passed reads subjected to various kinds of analyses such as whole-genome assembly, metagenome, meta-transcriptome, variant calling, and gene expression. In general, for the targeted amplicon sequencing driven taxonomic classification involves the quality passed reads clustering, operational taxonomic unit (OTU) picking, and then OTUs taxonomic classification. On the other hand, targeted amplicon, whole metagenome, and meta-transcriptome approach utilize the reads alignment against the reference database, followed to taxonomic and functional annotation. Whole metagenome and meta-transcriptome classification using de novo assembly provide much more detailed insights of studied samples with significantly increased cost and computation time. Taxonomic classification provides the insights of phylogenetic classification, alpha diversity (number of OTUs, Species richness, Chao1, Shannon index, and Simpson index), beta diversity such as principal coordinate analyses (PCoA), and taxa abundance (number of specific phyla or genera count or percentage). The reads functional classification provides the descriptive insights of underlying metabolic machinery categories obtained against a specific database. A database such as KEGG pathways depicts the classified reads into various biological pathways such as starch and sucrose metabolism, sulfur, propionate, butyrate, and methanogenesis (Anamika et al. 2019). Database CAZymes provide the reads with a property of

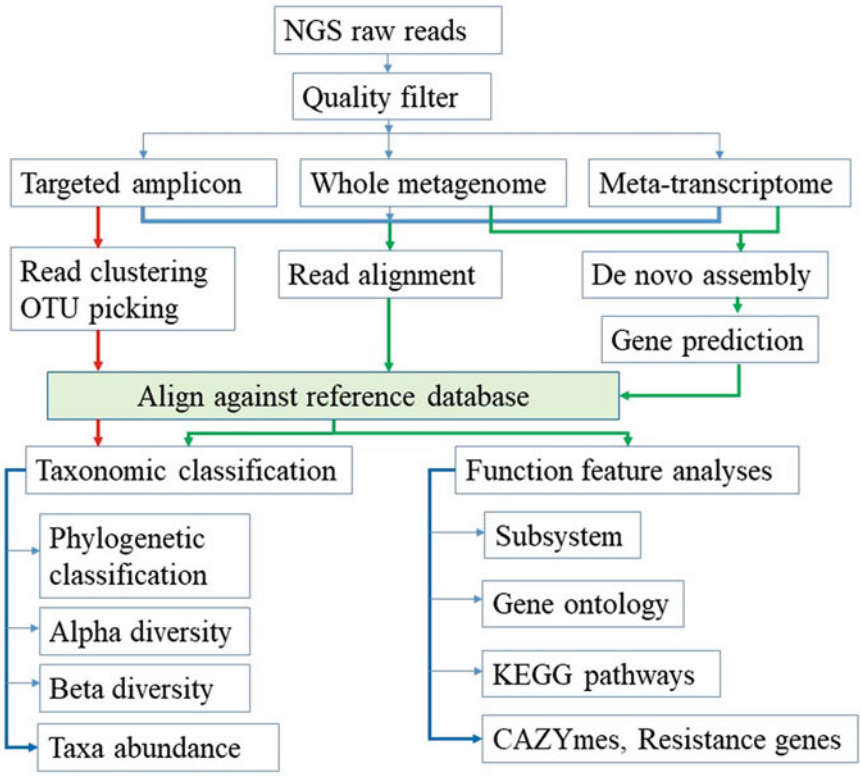

Fig. 14.3 Schematic workflow of NGS raw reads for taxonomic classification and function annotation

cellulose, pectin, and hemicellulose degradation, which offers the metagenomic study could provide the genomic details with great importance for industrial applications.

14.3.1 Generated Read Quality Filtration

Initially, the raw NGS files are processed for the filtering criteria, that is, any read with base quality score $Q < 20$ is filtered, then following to read trimming from 5′ end and 3′ end, if required. The machine-generated raw reads are filtered for the removal of poor bases and reads to obtain high-quality cleaned data. Few quality filtration tools are Trimmomatic, Cutadapt, Trim Galore, PRINSEQ, etc. (Del Fabbro et al. 2013; Pfeifer 2017).

14.3.2 De Novo Assembly of Sequenced Reads Microbial Communities

The quality passed reads are utilized for metagenome assembly, which describes the various steps together as input of fragmented large number of short DNA reads, and placing them back in overlapping fashion generates the original DNA sequence. The word de novo means starting from the beginning. Assemblies can be produced which have fewer gaps, less or no misassemblies, and fewer errors by tweaking the input parameters. The usually used tools for sequenced genome assembly are based on the command-line interface (CLI). Among that, meta-Velvet, Meta-IDBA, MetaSPAdes, and MEGAHIt are widely used. Such assembler algorithm, input data format, and requirements are presented in Table 14.1. This step is performed to optimize the generated assemblies by combining overlapping contigs and introducing appropriate gaps. Some of the scaffolding tools are SSPACE, PBJelly, gapCLoser, etc. More descriptive comparisons are provided here (Vollmers et al. 2017; Ayling et al. 2019).

14.3.3 Analysis of Microbial Diversity

Determination of microbial community in studied ecosystem samples provides the composition of the microbial diversity and composition under the influence of environmental factors and their co-occurrence. To find the community composition, there are various tools available among the scientific community to achieve their objectives. Among that, majority were read alignment against reference database-based annotation such as MG-RAST, MEGAN, EBI-Metagenome, QIIME, and RDP. The further advancement in annotation methodology, approaches such as k-mer, composition, and alignment-free tools, becomes available (Table 14.2). These tools enabled the scientific community to analyze the microbiota associated

Table 14.1 List of some tools available for metagenome assembly currently used by researchers

Assembler	Algorithm	Assembly method	Standard input	Read length	Output format	Availability
MetaMOSS	de Bruijn multiple Kmer	Denovo	fastq, fasta	Arbitrary	fasta	Open source
MetaSPAdes	De Bruijn graphs	Denovo	fastq, fasta	Arbitrary	fasta	Open source
MEGAHIT	de Bruijn graph	Denovo	fastq, fasta	Arbitrary	fasta	Open source
Meta-Velvet	de Bruijn graph	Denovo	fastq	Arbitrary	fasta	Open source
Meta-IBDA	de Bruijn graph	Denovo	fastq	Arbitrary	fasta	Open source
Ray Meta	de Bruijn graph	Denovo	fastq	Arbitrary	fasta	Open source
PRICE	Hybrid	Denovo	fastq	Arbitrary	fasta	Open source

Table 14.2 List of tools employed for 16S rRNA and whole shotgun metagenome data analysis

Available tools	Input	Output	Availability
For 16S rRNA, 18S rRNA, and fungal ITS			
QIIME 1, 2	sff, fasta, fastq	biom, txt	CLI
MOTHUR	sff, fasta, fastq	biom, txt	CLI
RDP	fasta, fastq	txt	CLI, web server
MG-RAST	fasta, fastq	txt	Web server
MEGAN 5, 6	txt, xml, sam	txt	GUI
EBI metagenome	fastq	biom, txt	Web server
MGX	fasta, fastq	txt	GUI
Hybrid_Denovo	fastq	biom, txt	CLI
For shotgun/whole metagenome			
KAAS	fasta	.txt, html	CLI, web server
MG-RAST	fasta, fastq	txt	Web server
MEGAN 5, 6	txt, xml, sam	txt	GUI
InterProScan	fasta	txt	Web server, CLI
dbCAN	fasta	txt	Web server, CLI
RapSearch	fastq	txt	CLI
Diamond	fastq	txt, sam	CLI
BLAST$^+$	fasta	txt, sam	Web server, CLI
EBI metagenome	fasta, fastq	biom, txt	Web server
Kaiju	fasta, fastq	txt	CLI, web server
Kraken	fasta, fastq	txt	CLI
k-Salm	fastq	txt, sam	CLI
CLARK	fasta, fastq	txt	CLI

with dormancy and sporulation, stress response genes, acetogenesis, methanogenesis, carbohydrate, protein metabolism, antibiotic, metal ion resistance genes, and aromatic compound metabolism (Roumpeka et al. 2017; Tamames et al. 2019).

Additionally, tools are also available that automates the matched reads were post-processed to find the community structure such as MG-RAST and EBI-Metagenomics including simple statistical graphical plots (Table 14.2). However, alignment/sequence matching against the reference sequence requires high computation power as the number of reads and length increases, which makes quite challenging and time-consuming tasks such as BLAST$^+$. Meanwhile, methodological advancements, such as *k*-mer and composition-based binning, facilitated the robust way analysis in limited time. In the *k*-mer approach, reads are converted into a small subset of 6 bases, 11 bases, and/or 22 bases called *k*-mers of similar sequences. The generated read *k*-mer composition is then compared to a reference database, and hits are counted to a known organism. For such a task, there are numerous tools available like Kraken, k-SALM, Kaiju, Klark, and the Ray Meta (Table 14.2). Additionally, various web servers are now available for automated whole genome

such as RAST, GenSAS, and metagenome annotation for taxonomic and functional annotations such as MG-RAST, EBI-Metagenome, and GALAXY (Roumpeka et al. 2017; Tamames et al. 2019).

14.3.4 Classification of Microbial Diversity with Bioinformatic Tools

14.3.4.1 MG-RAST

Out of all tools, Metagenomics Rapid Annotation using Subsystem Technology (MG-RAST) is the most popular, structured-web server for the analysis of microbial communities abundance at a taxonomic and functional level with graphical result visualization (Keegan et al. 2016). MG-RAST consists of various integrated tools and databases to determine the taxonomic and functional classification of NGS raw datasets. It takes the NGS raw input in form of single-end or pair-end sequencing reads and followed by quality processing. Quality passed reads were then automatically submitted for taxonomic and function analysis. After analysis, the user/researcher can visualize and download the entire result against the various databases. For the functional classification of metagenomes, MG-RAST offers various databases such as subsystem, COG, NOG, and KEGG. These all are hierarchical (up to level 4) type databases that enable the researcher to comprehensively determine the functional roles of sequences obtained from metagenomes. Further KEGG databases extensively provide the mapping of metagenomic sequences to the biochemical pathways such as sulfur metabolism, acetogenesis, methanogenesis, propionic acid metabolism, and starch and sucrose metabolism. The various kinds of visualization are bar, stacked, rarefaction, principal component analysis (PCoA), network, and pathways map.

14.3.4.2 MEGAN

MEtaGenome Analyzer (MEGAN) is a comprehensive locally installation-based stand-alone tool for microbial communities' abundance taxonomic and function analysis. The MEGAN primary requirement is that the sequences should be homology aligned against the database. The aligned sequences are imported/subjected as input to MEGAN and then parsed to taxonomic and functional profiles. In MEGAN, similar to MG-RAST, a researcher can map sequences against subsystem, COG, NOG, and KEGG databases; MEGAN also provides various kinds of visualization and biochemical pathway mappings. MEGAN taxonomic and functional classification can be visualized at various hierarchical levels along with significant statistical values (Huson et al. 2007).

14.3.4.3 QIIME

Quantitative Insights Into Microbial Ecology (QIIME) version 2.0 is a comprehensive tool for the targeted amplicon taxonomic classification and abundance estimation. It is a stand-alone, pipeline nature which consists of various integrated tools such as OTU picking, OTU classification, OTU rarefication, alpha and beta diversity estimation, statistical analysis, and OTU network-based co-occurrence determination. QIIME accepts barcoded, non-barcoded, and single, pair-end raw and quality passed reads. Using this tool, the researcher can classify efficiently amplicon reads such as 16S rRNA and 18S rRNA, fungal ITS, and functional marker-based community classification such as *pmoA*. QIIME also provides integrated rarefaction and statistical graph visualization (Bolyen et al. 2019).

14.3.4.4 MGnify

MGnify is a part of ENA (European Nucleotide Archive) infrastructure and a web server for the analysis of microbial communities' abundance at a taxonomic and functional level with graphical result visualization. For analyzing reads using this tool, the user is required to first deposit the raw read to the ENA database as per the standard of Genome Standard Consortium (GSC). EBI-metagenome enables the researcher to determine the targeted amplicon and whole metagenome taxonomic profile against 16S rRNA and 18S rRNA database, whereas functional classification is performed using gene ontology (GO) approach in a three main broad category, e.g., biological process, molecular function, and cellular component. The EBI graphical visualization includes a bar plot, pie chart, and PCoA plot (Mitchell et al. 2020).

14.3.5 Analysis of Microbial Community Metabolic Potential

The standard metagenome functional annotation pipeline is illustrated in Fig. 14.3, which consists of gene scanning (gene prediction), aligning against the reference sequence, taxonomy, function, and metabolic pathway assignment. The progressive advancement in genomes and metagenomes sequencing has led the development of numerous bioinformatics software for the prediction of genes and gene models. Further, as bioinformatic knowledgebase advanced, it offered to the development of various automated whole genome and metagenome data-based microbial genome binning and functional annotation, while requiring high computation resources (Roumpeka et al. 2017; Vincent et al. 2017). Lately, these developments have even opened up the possibility of "microbiome gene modifications" using CRISPR/Cas technology that will boom the genome editing of higher eukaryotes, especially host plants (Mehta et al. 2020; Dilawari et al. 2021).

Generally, the NGS machines from shotgun metagenome generate the read length from 50 to 600 base. Among that, majority were ranged from 300 to 600 bases, depending on the sequencing platform and chemistry. These short reads are assembled into longer sequences called contigs in a process called assembly. The assembly of short sequences becomes more important when the objective is to find the functional gene and metabolic pathways (Vollmers et al. 2017; Mitchell et al. 2020). Because, earlier, the input DNA is randomly fragmented into short fragments and then sequenced which used to result in a very poor quality of reads which contains a very high number of poor base quality scores. However, using the third generation, the read length is increased to more than 10 K bases, as well as poor base calling. Hence, a combination of both generation sequencers is more reliable for full-length functional gene discovery in genomes and metagenomes. At the current time, numerous tools are available for genomics and metagenomic data analysis. These tools mainly vary from algorithms and code language. Other variations include hardware requirements, user interface, installations, and user-interface (Roumpeka et al. 2017; Vollmers et al. 2017).

For the genome and metagenome functional annotation tools details, algorithm, input data type, and dependencies are given in Table 14.3. In the alignment

Table 14.3 List of software used for gene identification and prediction in genomes and metagenomes

Tools	Input	Single/paired-end	Output format	Availability	Suitability
Reference based					
BLAST+	fasta,fastq	Both	txt, sam, xml	Open source	Genome, metagenome
InterProScan	fasta	Single	txt, xml	Open source	Genome, metagenome
DIAMOND	fasta,fastq	Both	txt, sam, xml	Open source	Genome, metagenome
Usearch	fasta,fastq	Both	standard	Open source	Metagenome
RAPSearch	fasta,fastq	Both	standard	Open source	Genome metagenome
PALADIN	fasta,fastq	Both	standard	Open source	Metagenome
GhostX	fasta	Single	txt, html	Open source	Genome metagenome
Blast2GO	fasta,fastq	Single	txt, xml	License	Genome
Ab-initio gene prediction					
Meta-GeneMark	Fasta,fatsq	Single	txt	Open source	Metagenome
GLIMMER	fasta	Single	txt	Open source	Genome
GLIMMER -MG	Fasta,fatsq	Single	txt	Open source	Metagenome
AUGUSTUS	fasta	Single	txt, gff	Open source	Genome
FragGeneScan	fasta,fastq	Single, paired	txt	Open source	Metagenome
GeneMark	fasta	Single	txt, gff	Open source	Genome
ORF finder	fasta	Single	txt	Open source	Genome
Prodigal	fasta	Single	txt, gff	Open source	Genome

approach, the quality passed reads are matched against reference databases such as NCBI nr and NCBI RefSeq databases using sequence similarity search tools such as DIAMOND, PALADIN, RAPSearch, VSEARCH, and BLAST^{+}. The blast search utilizes the alignment of query sequences against the previously known reference sequence and classifies the sequence to their affiliation to taxonomy and function. InterProScan performs the identification of protein family, conserved domains, and superfamilies in the query sequence (Yadav et al. 2020).

14.4 Application of NGS Technology to Assess Microbial Diversity with Soil Fertility

Earth planet soil is the fundamental site for maintaining the ecological process and equilibrium maintenance. Soil provides the primary site for crop production, vegetation, life survival, biological, various hydrological, and economical processes. Among the biological process, microorganism plays various essential role such as mineralization, nutrient recycle, and maintenance of soil health. Hence, the protection of soil health for prolonged fertility in the agricultural system is highly important. Doran and Zeiss (2000) described health as the potential of soil functionality within an ecosystem and land use borders for sustainable biological productivity, improvement of environmental quality, and enhancement of animal and plant health. In agricultural practice, the microorganism ecosystem is generally balance-altering and dynamics of the microbial community.

It is generally achieved in the agricultural ecosystem through microbes-plant interaction and forms the important phenomenon of soil ecosystems (Bélanger and Avis 2002). In the landscape system, microbes are abundantly distributed in soil, which consists of useful and harmful communities. The plant root-adhered soil bacteria significantly contribute to the enhancement of soil property and release of phosphatase, dehydrogenase, mineralization, and various self-defense molecules such as secondary metabolites (Haas and Keel 2003) and stabilization of soil characteristics (Miller and Jastrow 2000). The microorganism-mediated soil fertility improvement involved (1) nitrogen fixation, (2) phosphate solubilization, (3) siderophore production, and (4) phytohormone production.

14.4.1 Microbial Community Diversity and Composition

The profile and function of soil microbes are connected with variable plants via litter quality, biomass production, root exudates, and root-shoot carbon allocation (Porazinska et al. 2003; Potthoff et al. 2006). Plant-derived alteration in litter inputs affects the microbial diversity and functionality (Habekost et al. 2008; Strecker et al. 2016). Lange et al. (2014) reported that species richness is the fundamental basis of

soil microbial community biomass, whereas the ratio of fungi to bacteria was positively affected by active group richness of plants and the existence of legumes. Also, the richness of plant species effect on soil microbial biomass was facilitated through nitrogen inputs and its concentration (Eisenhauer et al. 2010; Bessler et al. 2012).

The descriptive determination of plant microbiota interaction provides not only the remarkable supports for plant biology but additionally the identification and characterization of biochemical machinery for their potential application in biotechnological uses. For example, it can be utilized for improving plant health and growth, development of disease resistance, and various other resistance such as salt, biotic, and abiotic resistance variety development. Further development in the genomic studies facilitated the identification of various biological and biochemical function like virulence (Reddy et al. 2014), resistance against antibiotics and metals (Reddy and Dubey 2019), and energy production through detritus material (Yadav et al. 2020), core microbiome (Kumar et al. 2021) which play a significant role in the agriculture sector (Rialch et al. 2019, Sahu et al. 2020). Thus, detailed information on microbial community and functional ability of soil and rhizospheric microbiota facilitates the manipulation of environmental situations (Alisoltani et al. 2019).

The robust development in high-throughput sequencing technology and the release of vast organism species, strain genomes, and metagenomic studies extensively facilitated the deeper understanding of biochemical pathways (Loman and Pallen 2015; Singh et al. 2015b; Reddy et al. 2019). The technologies available in the twenty-first century have tremendous potential for the illustration/depiction of the taxonomic profile of microbial communities along with the determination of function metabolic pathways. However, the determination of such a taxonomic and functional profile is a tedious process for the microbiologist and hence requires strong computational skills as it consists of pipelines of distinct integrated tools. Although function and metabolic potential determination of microbial communities through metagenome and metatranscriptome are highly suitable for researchers as it provides vast information about the specific function-associated microbial communities (Singh et al. 2015b; Reddy et al. 2019; Reddy 2019).

14.4.2 Application of High-Throughput Sequencing on Soil Fertility

As per the glossary of Soil Science Society of America (SSSA), the soil can be formally defined as complex unconsolidated mixtures of minerals, organic matter, air, water, and countless (non) decayed organisms on the immediate earth's surface (Soil Science Society of America 2020). It forms the "vital skin of the earth" as it supports the earth's life web that consists of plants, animals, humans, and microbes. Since the beginning of civilization, soil fertility seems to sustain the plant's growth and agricultural yield (Sharma et al. 2021). It has been reported to be affected by

both genetic (parent material and related-characteristics) and environmental factors (climate, time, landscape, amendments, and macro-, and microorganisms) (Davies et al. 2019; Lisuma et al. 2020). In the present times of modern agriculture, the maintenance of soil fertility is typically required which is achieved by following soil evaluation and conservation practices. One such method is to use the metagenome sequencing for analyzing the soil fertility for various geographical areas. This has been already done significantly by various researchers as sequencing integrated soil fertility management around the globe over the last decade.

One of the very conclusive observations on establishing the role of sequencing in understanding microbial diversity in soil and correlating it with soil fertility was reported by Xue et al. (2011). In their study, they summarized the effect of consecutive years of mono-cropping on microbial populations and diversity. Furthermore, they introduced the advantages of 454 GS-FLX pyrosequencing high-sequencing method for the analysis of microbial populations and diversity. By using pyrosequencing in 146 different soil samples across the globe, Bates et al. (2011) observed consistent correlation among the soil C:N ratio with an abundance of two archaeal members. In the very next year, Hiiesalu et al. (2012) directly compared the multi-time point grassland plant richness below the soil surface by using accurate 454 sequencing of the chloroplast *trn*L(UAA) and related the variations in microbial composition to the fertility of the soil. Gigliotti and group observed the effect of organic addition amendments to the soils results in enhancement of nutrients as well as organic matter, C sequestration, and changes in microbial activity and biodiversity structure (Gigliotti et al. 2013). Furthermore, the use of pyrosequencing revealed that bacterial phyla and fungi species are related to the organic matter turnover in soil. In another study report, the effect of biochar use on re-wiring composition and function of microbes residing in fertile agricultural soils using 16S rRNA tag sequences showed significant differences in the composition of microbial community and the correlation patterns (Nielsen et al. 2014).

By using 454 pyrosequencing, Franke-Whittle and colleagues revealed the significant differences in microbial communities (fungi and bacteria) between replant and fallow soils. Furthermore, they urged to reveal the functional role of associated genera with soil fertility (Franke-Whittle et al. 2015). By employing the pyrosequencing of ITS2 amplicons, Sterkenburg and group observed significant changes in the composition of fungal communities related to plant nutrition and decomposition along a soil fertility gradient in a boreal forest. Through their experiment, they revealed the composition significantly varies at the levels of species, genera, as well as orders. Further, they revealed that ascomycetes fungi were dominant in less fertile forests, while the fungi related to basidiomycetes were highly abundant in more fertile forests, hummus, and litter (Sterkenburg et al. 2015). In a similar manner, the direct impact of fertilization on the composition of belowground arbuscular mycorrhizal (AM) fungi along the gradient of soil fertility was studied by Liu et al. (2015). They revealed the fertilizer application caused remarkable changes in the genus richness of AM fungi and over-dispersion statistically when fertilizers were applied at higher treatments (Liu et al. 2015).

As per the various experiments, it has been an established fact that soil pH apart from climatic conditions and management practices also regulated the soil fertility as well as impacted the diversity of below-ground communities. This was further supported by the findings by Jeanbille et al. (2016), who characterized the significant differences in bacterial communities enriched with acidic (nutrient-poor) and alkaline soils. Li et al. (2017) highlighted the role of C/N- and C/P-based shifts occurring in succession, composition, and diversity (alpha and beta) of microbial communities along a soil fertility gradient in paddy cultivation (Li et al. 2017). Tu et al. (2018) evaluated the significant effect of fertilizer application on the soil bacteria richness and role related to fertility assessed through 16S rRNA sequencing in dragon tree plantations (Tu et al. 2018). Recently, Burke and group characterized the responses and quantified a high degree of fungal communities in the beech-maple forest. Furthermore, they inferred the fungal taxa strongly associated with P-availability (Burke et al. 2019). More recently, Guo et al. (2020) evidenced the complexity of fungal assemblage in the soil directly correlates with soil fertility gradient by collecting various soil samples from tea plantations and sequencing them further with the Illumina MiSeq platform. In another study, Lisuma et al. (2020) reported work on tobacco plants grown in different Tanzanian landscape soils and cropping patterns linked the changes in rhizospheric bacterial composition with the soil fertility using 16S rRNA sequencing. Furthermore, they inferred the tobacco's rhizospheric bacterial diversity influences the solubilities of various macronutrients such as phosphorous, potassium, sulfur, as well as fix total N in the soil.

14.4.3 Role of High-Throughput Sequencing on Microbial Diversity and Crop Productivity

Ever since their origin millions of years ago, plants have existed in contact with microbes. Among the multitude of host functions that microbes control are nutrient uptake, protection, and phenology (Friesen et al. 2011). The identified microbial composition associated with plant root and their manipulation can be utilized for significantly boosting the quality of crop production by using beneficial microbiomes in agricultural systems (Bakker et al. 2012; Mueller and Sachs 2015). After studies to demonstrate that rhizobium nodules are colonized and the nitrogen fixed for their plant hosts, the Department of Agriculture (USA) advised inoculation of legume crops (Schneider 1892).

The plant microbiome's normal ecological roles leading to plant development, growth, and survival against biotic and abiotic stresses are well recorded (Turner et al. 2013; Müller and Ruppel 2014; Mehta et al. 2021a). Because of their close plant associations, the endophytic microbiome is believed to affect plant growth and production more specifically than epiphytic microbiomes. NGS-based metagenomic analysis is currently widely used to analyze plant endophytic microbiomes, contributing to an increased understanding of the profiles and roles of microbiomes. The

endophytic microbiome co-operating plant is now considered a new source of bio-inoculants to improve agricultural productivity. In recent decades, the plants are being inoculated with individual microbes to facilitate growth, nitrogen and phosphorus absorption (Afzal and Bano 2008), drought tolerance (Eke et al. 2019), and resistance to disease (Ashajyothi et al. 2020). However, this initiative was mostly centered almost on an individual strains of microbial species with occurrence of variable performance, which is usually due to the difficulty and habitat settings of experimental site or inoculation place. It is in general requirement for the understanding and administration of the diversified beneficial microbial consortia in cultivation sites to improve soil fertility and enhance support for plant growth. Several initiatives have been taken in the recent past for the above purpose (Reid and Greene 2013; Gilbert et al. 2014; Alivisatos et al. 2015; Stulberg et al. 2016; APS 2016). Because, identification of the "core microbiome" will help to identify plant-associated microbes that should be prioritized for further research and deceptive experiments (Bulgarelli et al. 2012; Lundberg et al. 2012; Sahu et al. 2020). Plant microbiota is highly diverse, yet not all of these microbes play functionally important roles in their host's biology. Defining the core microbiome enables researchers to filter out transient associations and refine the focus on stable taxa with a greater likelihood of influencing host phenotype. In comparison to the very profound sequence of a few plant microbiomes, NGS-based surveys of large numbers of microbiomes of the same plant species from different environments will help in higher progress against that target and follow-up selective cultivation of the candidate core microbiome.

14.5 Conclusion

In the present chapter, we have summarized the various approaches for the characterization of soil microbial community and their function. Furthermore, the involvement of various NGS technology and computational tools for the classification of raw reads has been also covered. The functional classification approach potentially offers the determination of various biochemical pathways and mining of enzymes for uses in industrial applications. The detailed information of the soil community offers the design of policy for manipulating soil microbes, enhancing fertility sustainably, and increasing chances of providing better crop productivity with increased economical values to the farmers, society, and whole mankind. Keeping this point in a long way to the future, the NGS-based assessment will facilitate the development of a sustainable management system for soil fertility and disease prevention.

Acknowledgments All authors are highly thankful to the authority of the departments and institutions for their support in doing this research. All the authors together duly acknowledge the infrastructural support from the Director of ICAR-IARI and Head, Division of Plant Pathology, and also thank-full to World Bank sponsored ICAR-NAHEP-CAAST, New Delhi, for financial support. Furthermore, Director of ICGEB, as well as group leader of Crop Improvement group, ICGEB New Delhi, is also gratefully acknowledged.

References

Adessi C, Matton G, Ayala G et al (2000) Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms. Nucleic Acids Res 28(20):e87–e87

Afzal A, Bano A (2008) Rhizobium and phosphate solubilizing bacteria improve the yield and phosphorus uptake in wheat (Triticum aestivum). Int J Agric Biol 10(1):85–88

Alisoltani A, Abia ALK, Bester L (2019) Shared microbiome in different ecosystems: a meta-omics perspective. In: Tripathi V, Kumar P, Tripathi P, Kishore A, Kamle M (eds) Microbial genomics in sustainable agroecosystems, vol 2. Springer, Singapore, pp 1–20. https://doi.org/10.1007/978-981-32-9860-6_1

Alivisatos AP, Blaser MJ, Brodie EL et al (2015) A unified initiative to harness earth's microbiomes. Science 350(6260):507–508

American Phytopathological Society (APS) (2016) Phytobiomes: a roadmap for research and translation. Available at http://www.phytobiomes.org/Roadmap/Documents/PhytobiomesRoadmap.pdf

Anamika MS, Singh B et al (2019) Databases: a weapon from the arsenal of bioinformatics for plant abiotic stress research. In: Wani S (ed) Recent approaches in omics for plant resilience to climate change. Springer, Cham, pp 135–169. https://doi.org/10.1007/978-3-030-21687-0_7

Ashajyothi M, Kumar A, Sheoran N et al (2020) Black pepper (Piper nigrum L.) associated endophytic Pseudomonas putida BP25 alters root phenotype and induces defense in rice (Oryza sativa L.) against blast disease incited by Magnaporthe oryzae. Biol Control 143:104181

Ayling M, Clark MD, Leggett RM (2019) New approaches for metagenome assembly with short reads. Brief Bioinform 21(2):584–594. https://doi.org/10.1093/bib/bbz020

Bakker MG, Manter DK, Sheflin AM et al (2012) Harnessing the rhizosphere microbiome through plant breeding and agricultural management. Plant Soil 360(1–2):1–13

Barriuso J, Valverde JR, Mellado RP (2011) Estimation of bacterial diversity using next generation sequencing of 16S rDNA: a comparison of different workflows. BMC Bioinform 12(1):473. https://doi.org/10.1186/1471-2105-12-473

Bates ST, Berg-Lyons D, Caporaso JG et al (2011) Examining the global distribution of dominant archaeal populations in soil. ISME J 5(5):908–917

Bélanger RR, Avis TJ (2002) Ecological processes and interactions occurring in leaf surface fung. In: Lindow SE, Hecht-Poinar EI, Elliott VJ (eds) Phyllosphere microbiology. APS Press, St. Paul, MN, pp 193–207

Béné C, Arthur R, Norbury H et al (2016) Contribution of fisheries and aquaculture to food security and poverty reduction: assessing the current evidence. World Dev 79:177–196. https://doi.org/10.1016/j.worlddev.2015.11.007

Bessler H, Oelmann Y, Roscher C et al (2012) Nitrogen uptake by grassland communities: contribution of N 2 fixation, facilitation, complementarity, and species dominance. Plant Soil 358(1–2):301–322

Bhardwaj Y, Reddy B, Dubey SK (2020) Temporal shift in methanotrophic community and methane oxidation potential in forest soils of dry tropics: high-throughput metagenomics approach. Biol Fertil Soils 56(6):859–867

Bharti J, Sahil MS et al (2021) Mitogen-activated protein kinase, plants, and heat stress. In: Husen A (ed) Harsh environment and plant resilience: molecular and functional aspects. Springer Nature, Cham, pp 323–354. https://doi.org/10.1007/978-3-030-65912-7_13

Bo B, Kim S-A, Han NS (2020) Bacterial and fungal diversity in Laphet, traditional fermented tea leaves in Myanmar, analyzed by culturing, DNA amplicon-based sequencing, and PCR-DGGE methods. Int J Food Microbiol 320:108508

Bolyen E, Rideout JR, Dillon MR et al (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 37(8):852–857. https://doi.org/10.1038/s41587-019-0209-9

Bulgarelli D, Rott M, Schlaeppi K et al (2012) Revealing structure and assembly cues for arabidopsis root-inhabiting bacterial microbiota. Nature 488(7409):91–95

Burke DJ, Carrino-Kyker SR, Burns JH (2019) Is it climate or chemistry? Soil fungal communities respond to soil nutrients in a multi-year high-resolution analysis. Ecosphere 10(10):e02896
Davies G, Kirkpatrick J, Cameron E et al (2019) Ecosystem engineering by digging mammals: effects on soil fertility and condition in Tasmanian temperate woodland. R Soc Open Sci 6 (1):180621
Del Fabbro C, Scalabrin S, Morgante M et al (2013) An extensive evaluation of read trimming effects on illumina ngs data analysis. PLoS One 8(12):e85024. https://doi.org/10.1371/journal.pone.0085024
Dilawari R, Kaur N, Priyadarshi N et al (2021) Genome editing: a tool from the vault of science for engineering climate-resilient cereals. In: Husen A (ed) Harsh environment and plant resilience: molecular and functional aspects. Springer Nature, Cham, pp 45–72. https://doi.org/10.1007/978-3-030-65912-7_3
Doran JW, Zeiss MR (2000) Soil health and sustainability: managing the biotic component of soil quality. Appl Soil Ecol 15(1):3–11. https://doi.org/10.1016/S0929-1393(00)00067-6
Eid J, Fehr A, Gray J et al (2009) Real-time DNA sequencing from single polymerase molecules. Science 323(5910):133–138
Eisenhauer N, Beßler H, Engels C et al (2010) Plant diversity effects on soil microorganisms support the singular hypothesis. Ecology 91(2):485–496
Eke P, Kumar A, Sahu KP et al (2019) Endophytic bacteria of desert cactus (Euphorbia trigonas Mill) confer drought tolerance and induce growth promotion in tomato (Solanum lycopersicum L.). Microbiol Res 228(2019):126302
Endrullat C, Glokler J, Franke P et al (2016) Standardization and quality management in next-generation sequencing. Appl Transl Genom 10:2–9. https://doi.org/10.1016/j.atg.2016.06.001
Fedurco M, Romieu A, Williams S et al (2006) BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. Nucleic Acids Res 34(3):e22–e22
Franke-Whittle IH, Manici LM, Insam H et al (2015) Rhizosphere bacteria and fungi associated with plant growth in soils of three replanted apple orchards. Plant Soil 395(1–2):317–333
Friesen ML, Porter SS, Stark SC et al (2011) Microbially mediated plant functional traits. Annu Rev Ecol Evol Syst 42(1):23–46
Gebbers R, Adamchuk VI (2010) Precision agriculture and food security. Science 327 (5967):828–831. https://doi.org/10.1126/science.1183899
Gigliotti G, Pezzolla D, Zadra C et al (2013) Dynamics of organic matter and microbial populations in amended soil: a multidisciplinary approach. EGUGA:EGU2013-4405
Gilbert JA, Jansson JK, Knight R (2014) The earth microbiome project: successes and aspirations. BMC Biol 12(1):69
Guo J, Ling N, Chen Z et al (2020) Soil fungal assemblage complexity is dependent on soil fertility and dominated by deterministic processes. New Phytol 226(1):232–243
Haas D, Keel C (2003) Regulation of antibiotic production in root-colonizing Peudomonas spp. and relevance for biological control of plant disease. Annu Rev Phytopathol 41:117–153. https://doi.org/10.1146/annurev.phyto.41.052002.095656
Habekost M, Eisenhauer N, Scheu S et al (2008) Seasonal changes in the soil microbial community in a grassland plant diversity gradient four years after establishment. Soil Biol Biochem 40 (10):2588–2595
Hiiesalu I, Oepik M, Metsis M et al (2012) Plant species richness belowground: higher richness and new patterns revealed by next-generation sequencing. Mol Ecol 21(8):2004–2016
Huson DH, Auch AF, Qi J et al (2007) MEGAN analysis of metagenomic data. Genome Res 17 (3):377–386. https://doi.org/10.1101/gr.5969107
Jeanbille M, Buée M, Bach C et al (2016) Soil parameters drive the structure, diversity and metabolic potentials of the bacterial communities across temperate beech forest soil sequences. Microb Ecol 71(2):482–493
Jovel J, Patterson J, Wang W et al (2016) Characterization of the gut microbiome using 16S or shotgun metagenomics. Front Microbiol 7:459. https://doi.org/10.3389/fmicb.2016.00459

Ju J, Kim DH, Bi L et al (2006) Four-color DNA sequencing by synthesis using cleavable fluorescent nucleotide reversible terminators. Proc Natl Acad Sci U S A 103 (52):19635–19640. https://doi.org/10.1073/pnas.0609513103

Kanagawa T (2003) Bias and artifacts in multitemplate polymerase chain reactions (PCR). J Biosci Bioeng 96(4):317–323

Keegan KP, Glass EM, Meyer F (2016) MG-RAST, a metagenomics service for analysis of microbial community structure and function. Methods Mol Biol 1399:207–233. https://doi.org/10.1007/978-1-4939-3369-3_13

Kumar M, Kumar A, Sahu KP et al (2021) Deciphering core-microbiome of rice leaf endosphere: revelation by metagenomic and microbiological analysis of aromatic and non-aromatic genotypes grown in three geographical zones. Microbiol Res 246:126704. https://doi.org/10.1016/j.micres.2021.126704

Lakshmanan V, Selvaraj G, Bais HP (2014) Functional soil microbiome: belowground solutions to an aboveground problem. Plant Physiol 166(2):689–700

Lange M, Habekost M, Eisenhauer N et al (2014) Biotic and abiotic properties mediating plant diversity effects on soil microbial communities in an experimental grassland. PLoS One 9(5): e96182

Laszlo AH, Derrington IM, Ross BC et al (2014) Decoding long nanopore sequencing reads of natural DNA. Nat Biotechnol 32(8):829

Leite AM, Mayo B, Rachid CT et al (2012) Assessment of the microbial diversity of Brazilian kefir grains by PCR-DGGE and pyrosequencing analysis. Food Microbiol 31(2):215–221

Li W, Chen X, Liu M et al (2017) Shifts in microbial communities with increasing soil fertility across a chronosequence of paddy cultivation in subtropical China. Appl Soil Ecol 120:153–159

Lisuma JB, Zuberi Z, Ndakidemi PA et al (2020) Linking rhizosphere bacterial diversity and soil fertility in tobacco plants under different soil types and cropping pattern in Tanzania: a pilot study. Heliyon 6(7):e04278

Liu W-T, Marsh TL, Cheng H et al (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl Environ Microbiol 63(11):4516–4522

Liu Y, Johnson NC, Mao L et al (2015) Phylogenetic structure of arbuscular mycorrhizal community shifts in response to increasing soil fertility. Soil Biol Biochem 89:196–205

Loman NJ, Pallen MJ (2015) Twenty years of bacterial genome sequencing. Nat Rev Microbiol 13 (12):787–794. https://doi.org/10.1038/nrmicro3565

Lu YF, Goldstein DB, Angrist M et al (2014) Personalized medicine and human genetic diversity. Cold Spring Harb Perspect Med 4(9):a008581. https://doi.org/10.1101/cshperspect.a008581

Lundberg DS, Lebeis SL, Paredes SH et al (2012) Defining the core Arabidopsis thaliana root microbiome. Nature 488(7409):86–90

Mehta S, James D, Reddy MK (2019a) Omics technologies for abiotic stress tolerance in plants: current status and prospects. In: Wani S (ed) Recent approaches in omics for plant resilience to climate change. Springer, Cham, pp 1–34. https://doi.org/10.1007/978-3-030-21687-0_1

Mehta S, Singh B, Dhakate P et al (2019b) Rice, marker-assisted breeding, and disease resistance. In: Wani S (ed) Disease resistance in crop plants. Springer, Cham, pp 83–111. https://doi.org/10.1007/978-3-030-20728-1_5

Mehta S, Lal SK, Sahu KP et al (2020) CRISPR/Cas9-edited rice: a new frontier for sustainable agriculture. In: Rakshit A, Singh H, Singh A, Singh U, Fraceto L (eds) New frontiers in stress management for durable agriculture. Springer, Singapore, pp 427–458. https://doi.org/10.1007/978-981-15-1322-0_23

Mehta S, Singh B, Patra A et al (2021a) Maize microbiome: current insights for the sustainable agriculture. In: Solanki MK, Kashyap PL, Ansari RA (eds) Microbiomes and plant health. Academic Press, Massachusetts, pp 267–297. https://doi.org/10.1016/B978-0-12-819715-8.00009-4

Mehta S, Chakraborty A, Roy A et al (2021b) Fight hard or die trying: current status of lipid signaling during plant–pathogen interaction. Plants 10(6):1098

Miller RM, Jastrow JD (2000) Mycorrhizal fungi influence soil structure. In: Kapulnik Y, Douds DD (eds) Arbuscular mycorrhizas: physiology and function. Springer, Dordrecht, pp 3–18. https://doi.org/10.1007/978-94-017-0776-3_1

Mitchell AL, Almeida A, Beracochea M et al (2020) MGnify: the microbiome analysis resource in 2020. Nucleic Acids Res 48(D1):D570–D578. https://doi.org/10.1093/nar/gkz1035

Mohanty SR, Bodelier PL, Conrad R (2007) Effect of temperature on composition of the methanotrophic community in rice field and forest soil. FEMS Microbiol Ecol 62(1):24–31

Mueller UG, Sachs JL (2015) Engineering microbiomes to improve plant and animal health. Trends Microbiol 23(10):606–617

Müller T, Ruppel S (2014) Progress in cultivation-independent phyllosphere microbiology. FEMS Microbiol Ecol 87(1):2–17

Neilson JW, Jordan FL, Maier RM (2013) Analysis of artifacts suggests DGGE should not be used for quantitative diversity analysis. J Microbiol Methods 92(3):256–263

Nielsen S, Minchin T, Kimber S et al (2014) Comparative analysis of the microbial communities in agricultural soil amended with enhanced biochars or traditional fertilisers. Agric Ecosyst Environ 191:73–82

Pareek CS, Smoczynski R, Tretyn A (2011) Sequencing technologies and genome sequencing. J Appl Genet 52(4):413–435. https://doi.org/10.1007/s13353-011-0057-x

Pfeifer SP (2017) From next-generation resequencing reads to a high-quality variant data set. Heredity (Edinb) 118(2):111–124. https://doi.org/10.1038/hdy.2016.102

Pitta DW, Parmar N, Patel AK et al (2014) Bacterial diversity dynamics associated with different diets and different primer pairs in the rumen of Kankrej cattle. PLoS One 9(11):e111710

Porazinska DL, Bardgett RD, Blaauw MB et al (2003) Relationships at the aboveground–belowground interface: plants, soil biota, and soil processes. Ecol Monogr 73(3):377–395

Potthoff M, Steenwerth KL, Jackson LE et al (2006) Soil microbial community composition as affected by restoration practices in California grassland. Soil Biol Biochem 38(7):1851–1860. https://doi.org/10.1016/j.soilbio.2005.12.009

Rajput LS, Aggarwal SK, Mehta S et al (2021) Role of WRKY transcription factor superfamily in plant disease management. In: Giri B, Sharma MP (eds) Plant stress biology. Springer, Singapore, pp 335–361. https://doi.org/10.1007/978-981-15-9380-2_11

Ramakrishnan B, Lueders T, Dunfield PF et al (2001) Archaeal community structures in rice soils from different geographical regions before and after initiation of methane production. FEMS Microbiol Ecol 37(2):175–186

Reddy B (2019) Advancements in microbial genome sequencing and microbial community characterization. In: Tripathi V, Kumar P, Tripathi P, Kishore A, Kamle M (eds) Microbial genomics in sustainable agroecosystems. Springer, Singapore, pp 87–113

Reddy B, Dubey SK (2019) River Ganges water as reservoir of microbes with antibiotic and metal ion resistance genes: High throughput metagenomic approach. Environ Pollut 246:443–451. https://doi.org/10.1016/j.envpol.2018.12.022

Reddy B, Dubey SK (2021) Exploring the allochthonous pollution influence on bacterial community and co-occurrence dynamics of River Ganga water through 16S rRNA-tagged amplicon metagenome. Environ Sci Pollut Res 28:26990–27005

Reddy B, Singh KM, Patel AK et al (2014) Insights into resistome and stress responses genes in Bubalus bubalis rumen through metagenomic analysis. Mol Biol Rep 41(10):6405–6417

Reddy B, Pandey J, Dubey SK (2019) Assessment of environmental gene tags linked with carbohydrate metabolism and chemolithotrophy associated microbial community in River Ganga. Gene 704:31–41. https://doi.org/10.1016/j.gene.2019.04.004

Reid A, Greene SE (2013) How microbes can help feed the world. Issues 105(31):33–37

Rialch I, Singh S, Singh R et al (2019) Plant and microbial genomics in crop improvement. In: Tripathi V, Kumar P, Tripathi P, Kishore A, Kamle M (eds) Microbial genomics in sustainable agroecosystems, vol 2. Springer, Singapore, pp 215–230. https://doi.org/10.1007/978-981-32-9860-6_13

Roumpeka DD, Wallace RJ, Escalettes F et al (2017) A review of bioinformatics tools for bio-prospecting from metagenomic sequence data. Front Genet 8:23. https://doi.org/10.3389/fgene.2017.00023

Sahil KR, Patra A et al (2021) Expression and regulation of stress-responsive genes in plants under harsh environmental conditions. In: Husen A (ed) Harsh environment and plant resilience: molecular and functional aspects. Springer Nature, Cham, pp 25–44. https://doi.org/10.1007/978-3-030-65912-7_2

Sahu KP, Kumar A, Patel A et al (2020) Rice blast lesions: an unexplored phyllosphere microhabitat for novel antagonistic bacterial species against Magnaporthe oryzae. Microb Ecol. https://doi.org/10.1007/s00248-020-01617-3

Schneider A (1892) Observations on some American rhizobia. Bull Torrey Bot Club 19(7):203–218

Sharma P, Sharma MMM, Malik A et al (2021) Rhizosphere, rhizosphere biology, and rhizospheric engineering. In: Mohamed HI, El-Beltagi HEDS, Abd-Elsalam KA (eds) Plant growth-promoting microbes for sustainable biotic and abiotic stress management. Springer, Cham, pp 577–624. https://doi.org/10.1007/978-3-030-66587-6_21

Singh K, Jisha T, Reddy B et al (2015a) Microbial profiles of liquid and solid fraction associated biomaterial in buffalo rumen fed green and dry roughage diets by tagged 16S rRNA gene pyrosequencing. Mol Biol Rep 42(1):95–103

Singh KM, Patel AK, Shah RK et al (2015b) Potential functional gene diversity involved in methanogenesis and methanogenic community structure in Indian buffalo (Bubalus bubalis) rumen. J Appl Genet 56(3):411–426

Soil Science Society of America (2020) Wisconsin, United Nations. https://www.soils.org/. Accessed 6 Oct 2020

Sterkenburg E, Bahr A, Brandström Durling M et al (2015) Changes in fungal communities along a boreal forest soil fertility gradient. New Phytol 207(4):1145–1158

Strecker T, Macé OG, Scheu S et al (2016) Functional composition of plant communities determines the spatial and temporal stability of soil microbial properties in a long-term plant diversity experiment. Oikos 125(12):1743–1754

Stulberg E, Fravel D, Proctor LM et al (2016) An assessment of US microbiome research. Nat Microbiol 1(1):15015

Tamames J, Cobo-Simon M, Puente-Sanchez F (2019) Assessing the performance of different approaches for functional and taxonomic annotation of metagenomes. BMC Genomics 20 (1):960. https://doi.org/10.1186/s12864-019-6289-6

Tu J, Qiao J, Zhu Z et al (2018) Soil bacterial community responses to long-term fertilizer treatments in Paulownia plantations in subtropical China. Appl Soil Ecol 124:317–326

Turner TR, James EK, Poole PS (2013) The plant microbiome. Genome Biol 14(6):1–10

Vincent AT, Derome N, Boyle B et al (2017) Next-generation sequencing (NGS) in the microbiological world: how to make the most of your money. J Microbiol Methods 138:60–71. https://doi.org/10.1016/j.mimet.2016.02.016

Vollmers J, Wiegand S, Kaster A-K (2017) Comparing and evaluating metagenome assembly tools from a microbiologist's perspective - not only size matters. PLoS One 12(1):e0169662. https://doi.org/10.1371/journal.pone.0169662

Xue C, Huang Q, Ling N et al (2011) Analysis, regulation and high-throughput sequencing of soil microflora in mono-cropping system. Acta Pedol Sin 48(3):612–618

Yadav S, Reddy B, Dubey SK (2020) De novo genome assembly and comparative annotation reveals metabolic versatility in cellulolytic bacteria from cropland and forest soils. Funct Integr Genomics 20(1):89–101. https://doi.org/10.1007/s10142-019-00704-0