

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Soil Metagenomics: Concepts and Applications

*Shrinivas N. Sabale, Padmaja P. Suryawanshi
and Krishnaraj P.U.*

“We know more about the movement of celestial bodies than about the soil underfoot.”

—Leonardo Da Vinci, circa 1500s.

Abstract

Soil is a living entity of the Earth, and considered as one of the main reservoir of microbial diversity. Studying the soil microbial diversity is very much necessary, as they play an important role in maintaining the health of soil by recycling the nutrients, creating soil structure and humus. However, the culture dependent approaches fail to provide clear estimates of the diversity and untapped resources. Hence, study of the microbial diversity using culture independent approaches become necessary. The field of metagenomics helps in studying the genomes of the diverse soil organisms collectively in their natural habitat which holds the promising for accessing novel genetic resources. Application of the metagenomics to the soil environment is very challenging due to several difficulties; one of which is co-extraction of humic acid with nucleic acids which hinder downstream high throughput processes. However, applying sequencing methods to soil microbial communities will help in uncovering the hidden resources like novel genes, biomolecules and other valuable products which are yet to be discovered or still unknown. Different culture independent techniques and applications of the metagenomics to study the abundant microflora of the complex and changing environment of soil discussed herein.

Keywords: soil microbiome, soil health, metagenomics, DGGE, 16S rDNA, metagenome sequencing

1. Introduction

Life on earth is not only dependent on the celestial environment around the earth but it is also completely influenced by the soil under our feet. Earth is a pool of different complex ecosystems and the proper functioning of those ecosystems are very much required for sustenance and evolution of the life. Soil is an important ecosystem of the earth and essential for life and one of the most valuable resources available to us, which acts as a water filter, supports plant and animal life, source of the minerals and medicines. Soil is a living entity and existence of which is defined by the occurrence of the organisms in it. The formation of soil is not only dependent on the physical or chemical activity on the rocks, but also by the continuous activity of diverse microbial species which add or enhance the

properties of soils in terms of formation of structure and function. Soil provides shelter to various soil harboring insects, reptiles, animals and a huge volume of microbes within the soil aggregates. It is the home for innumerable microbial species largely unnoticed and busy in maintaining the ecological balance of the earth. It is a complex and dynamic system, which directly or indirectly influences the food chain, nutrient cycles, and ecological equilibrium. Understanding the unexplored microbial world is high priority and is needed to come up with the ways of coping up with various climatic changes and for the welfare of every organism on the earth.

2. Soil health

Soil is an interconnected system with high levels of exchange of energy between organisms and physico-chemical components, which allows it to be a self-organized system. However, despite having the unique and incredible capacity to adapt to environmental changes, microbes are sensitive to land management and climate changes. Resistance is the capacity of the soil to maintain its health despite the magnitude of the change caused by any kind of perturbation. Resilience is the capacity of the system to return to its original state after a disturbance, which is also known as the self-healing capacity of the system (https://microbewiki.kenyon.edu/index.php/Soil_Health).

Soil has been neglected because it is termed as dirt. But from this dirt our ancestors learnt the skill of cultivation of crops and developed various cultural practices for the production of food. And our staple food started coming from the crops cultivated in the soil. Thus, agriculture—the science or art of farming, including cultivation of the soil for the growing of crops and the rearing of animals to provide food, wool, and other products came into existence (<https://en.oxforddictionaries.com/definition/agriculture>). Being a most diverse active environment with huge number of known and unknown microbial species, soil becomes a challenging ecological niche for study. The biochemistry of soil reflects several unknown functions which are very much important for sustenance of life. However, the recent technologies using heavy machinery and management practices intensified agriculture that resulted in extensive cultivation without enriching the soil. This overuse has initiated the deterioration of the cultivable farm lands through loss of fertility, vigor, soil structure and capacity to sustain life. In short, many cultivable areas have turned saline or uncultivable. Agricultural land is simultaneously being lost to non-agricultural uses. In addition, the intensive use of land in production through multiple cropping, reduced fallow periods, excessive use of agrochemicals, spread of monocultures are components of management is perceived as leading to land degradation and the undermining of its long-term productive potential (www.fao.org).

The formation of 1 cm of topsoil in its natural course requires 300 years but the resultant loss of soil due to erosion is high (Soil loss in India was 16.4 t/ha/year as reported in State of Indian Agriculture Report, 2015–2016). The population around the globe has been expanding rapidly and standing at around 7.3 billion in 2016 and still counting. With the increase in the population number of challenges around global sustainability also increases, including the need for more food and space. The demand and supply of food is the major concern for all the countries and it can be noted that we are falling short of cultivable land due to soil erosion, saline soils, excessive use of pesticides, herbicides, inorganic fertilizers and the shift of land use to housing sector. According to the soil scientist Dr. Elaine Ingham, “If we lose both bacteria and fungi, then the soil degrades”.

Overuse of chemical fertilizers and pesticides have effects on the soil organisms that are similar to human overuse of antibiotics [1]. Therefore, sustaining and improving the characteristics of soil will be an utmost priority for generations to come and we have to focus on keeping our soils alive by maintaining the cultivable land or rejuvenating the barren lands. Thus, soil health becomes one of the most important factors in terms of agriculture and forest ecosystems and very much necessary for survival of living beings.

2.1 Importance of soil microorganisms

Biological fertility of the soil comprises of the soil microbes which greatly influence the structure and functioning of the soil ecosystem. The maintenance of the physical fertility and chemical fertility is driven by the metabolic repertoire of the soil microorganisms. Soil microbes are part of a very complex food web that occurs in soil. Soil biota is indispensable for key soil function such as decomposition of soil organic matter, nutrient cycling and formation of soil aggregates and billions of microbes reside in a single gram of soil [2]. The most numerous microbes in soil are the bacteria, followed in decreasing numerical order the fungi, soil algae and soil protozoa. The work that is turned out by bacteria does benefit the plants and other living organisms. Plants are unable to absorb the nutrition from the soil without microbes working in the soil. Although microbes tend to receive more attention for the diseases they cause, they are likely to be more significant for their role in maintaining the health of the plants by protecting them from other microorganisms, providing vitamins, nutrients and influencing the developmental processes [3]. The microbial activity in the rhizosphere is essential for plant functioning as it assists the plant in nutrient uptake and offers protection against pathogen attack [4]. Rhizosphere, a region of soil which is influenced by the plant roots, a microenvironment where a great microbial diversity thrives in close association with plant roots where in various abiotic and biotic interactions take place. Abiotic factors influence the structure of microbial communities which have evolved mechanisms to deal with occupation of the space that allows them to obtain nutrients excreted by the plant roots. The organism diversity of the bulk-soil or the rhizospheric soil of different plants has been extensively studied; most studies have reported a wide range of organisms from those soil portions [5]. Surprisingly, a soil sample could contain up to 4×10^6 different taxa [6]. This suggests that soil is a large reservoir for the discovery of several compounds that may have applications in agriculture, human health or industry. The relevant insights of the studies related to soil are that each of them has shown different aspects to study and analyze. In some cases, they have provided innovative views of the microbial ecology on extreme environments for the development of life. At the same time, others have found novel biocatalysts, new antibiotics, personalized medicine, bioremediation and other potential applications in biotech industry [7].

Microbiological studies in the soil environment are hampered by the fact that the largest proportion of soil bacteria as yet cannot be cultured [8], and the microbes found using common culture methods are rarely abundant in any environmental niche from which they were cultured. The microbes isolated from the soil mostly belong to the four major phyla namely *Proteobacteria*, *Actinobacteria*, *Firmicutes* and *Bacteroidetes*, as those minute organisms are easily cultivable under laboratory conditions. Such information suggests that there is the need of molecular techniques which overcomes the need to isolate and study each organism for in depth characterization. However with the development in metagenomics a more complete picture of the rhizosphere microbiome can be unearthed [9] and many new microbial players in the rhizosphere are on their way to be discovered.

3. Soil metagenomics: a concept

Even though soil microbes dominate the biosphere, most microbes in nature have not been studied because the standard culturing techniques could help in the study of those organisms which can be cultured on artificial media and rest of the organisms are still unknown since lack of the universal artificial media and growth conditions. Instead of cultivating different organisms in pure form and studying the morphology and characteristics, metagenomics, studies all the organisms present in the environmental sample at once by the combining traditional microbiology and molecular biology. Microbial diversity is the major driving force of fundamental metabolic processes in dynamic environment like soil and a number of microbial species are associated with the plant rhizosphere. Therefore, a basic understanding of diversity of soil biota is required in order to preserve the integrity, function and long term sustainability of natural and managed terrestrial ecosystems. The opportunity that stands before microbiologists today is akin to a reinvention of the microscope in the expanse of research questions it opens to investigation [10].

Metagenomics provides a new way of examining the microbial world that not only will transform modern microbiology but has the potential to revolutionize understanding of the entire living world and their functions in different ecosystem. In metagenomics, the power of genomic analysis is applied to entire communities of microbes, bypassing the need to isolate and culture individual bacterial community members in order to identify the microbes in the community. The new approach and its attendant technologies will bring to light the myriad capabilities of microbial communities that drive the planet's energy and nutrient cycles, maintain the health of its inhabitants, and shape the evolution of life. Metagenomics is expected to generate knowledge of microbial interactions so that they can be harnessed to improve human health, food security, and energy production.

Metagenomics combines the power of genomics, bioinformatics, and systems biology. Operationally, it is novel in that it involves study of the genomes of many organisms simultaneously. It provides new access to the microbial world. Although community ecology is not new to microbiology, the power of genomics in the study of communities brings in an unparalleled opportunity.

Meta in the first sense means that this new science seeks to understand biology at the aggregate level, transcending the individual organism to focus on the genes in the community and how genes might influence each other's activities in serving collective functions. In the second sense, meta also recognizes the need to develop computational methods that maximize understanding of the genetic composition and activities of communities so complex that they can only be sampled, never completely characterized.

Metagenomics, still a very new science, has already produced a wealth of knowledge about the uncultured microbial world because of its radically new ways of understanding microbial world. All metagenomics studies take the same first step: DNA is extracted directly from all the microbes living in a particular environment. The mixed sample of DNA can then be analyzed directly, or cloned into a form maintainable in laboratory bacteria, creating a library that contains the genomes of all the microbes found in that environment.

Since the invention of the simple microscope by Anton Von Leeuwenhoek we are able to identify and study only 1% of the total microbial diversity available on Earth. The rest of the 99% of the microbes are not easily accessible to us as they are busy with their life cycle and maintenance of the microbial balance of our planet. Our current knowledge about the culturable microbial techniques is unable to achieve the complete picture of the diversity and functions of the micro-organisms residing

even in a single environment. Thus the development of new techniques holds huge potential to overcome the difficulties to capture complete microbial profile.

The study of microbial diversity is very much important in a particular habitat as the microbes helps in maintenance or modification of the environment and it is also responsible for survival of surrounding organisms and also helps in the process of evolution of the other living organisms. It is therefore, time to investigate the intricate relationship between the soil microbiota and plants and how this relationship benefits both of the organisms in spite of the environmental changes.

In 1904, Lorenz Hiltner coined the term, “rhizosphere”—a word partly originated from the Greek word “rhiza”, meaning root which describes the plant-root interface [11]. Hiltner described the rhizosphere as the area around a plant root that is inhabited by a unique population of microorganisms. But in recent years based on the relative proximity and influence to root the rhizosphere is refined to include three zones. The endorhizosphere includes cortex and endodermis in which microbes occur in very close association with the roots of the plant; they are present inside the roots and referred to as endophyte at times. They are resultant of plant root exudates which attract them to colonize the roots. The rhizoplane is the medial zone directly adjacent to the root including the root epidermis and mucilage. The outermost zone is the ectorhizosphere which extends from the rhizoplane out into the bulk soil and it is the house to the organisms which are either free living or non-symbionts which are highly influenced by management or cultural practices. The rhizosphere is not a region of definable size or shape, but consists of a gradient in chemical, biological and physical properties which change both radially and longitudinally along the roots.

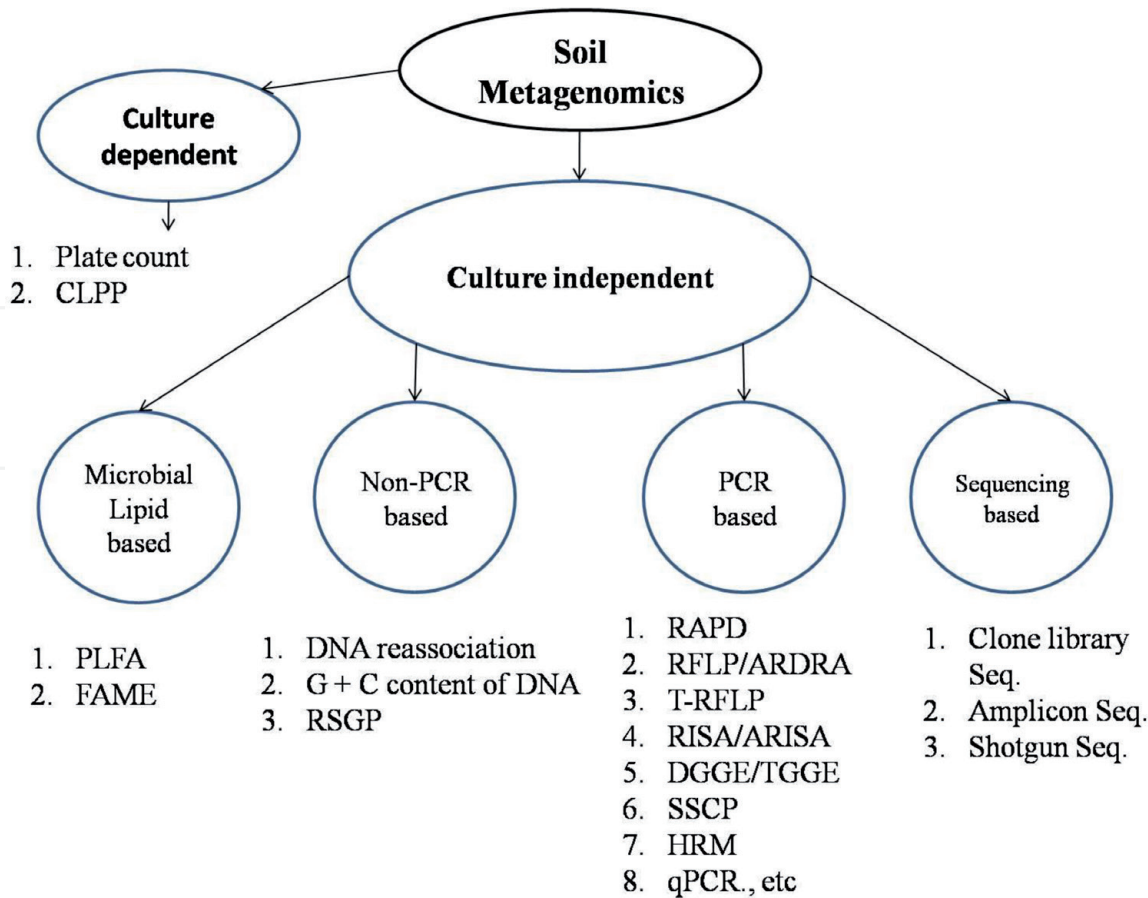


Figure 1.
Methods for studying microbial diversity harboring in the soil.

Discovering the microbiota of the soil environment requires selection of the suitable method(s) for capturing the maximum diversity at a particular time and space. It is obvious that based on a particular interest of the researcher, the approach to understand microbial diversity may be different from others or same. To determine spatial heterogeneity, considerable amount of work is still needed, for example, how representative a 0.1 mg sample of soil is with respect to the larger environment from which it was taken. Selection of proper method for extracting the meaningful information is very much important to escape the biasness of the results or to reduce the error rate in estimating the diversity by using statistical significance as well. To understand the overall microbial diversity, the researchers use one or combination of the methods given below for collecting the maximum information using their own standard sample size (**Figure 1**).

4. Culture dependent approaches

4.1 Plate count

Standard culture technique involves selective plating, direct counting viable colonies/cells and characterization of microorganisms on artificial or synthetic growth media in laboratory condition [12]. The major limitation of the culture dependent techniques is that we are unable to study the microbes which cannot be grown on artificial media. In spite of putting efforts in improving the culture media which can copy the natural conditions for the growth of the microbes, it was seen that the majority of the fraction of microbes were still unculturable. The minimal media, culture dilutions and the extended incubations helps in recovery of few slow growing microbial species which were considered as uncultivable earlier, but the abundance of the microbes in soil, which can be seen under microscope, still remains untapped or untouched. A technique has been devised for the cultivation of uncultured microorganisms from different environments that involved encapsulation of cells in gel micro-droplets for large-scale microbial cultivation under low nutrient flux conditions [13].

4.2 Community level physiological profiling (CLPP)

Garland and Mills developed a technique to assess the potential functional diversity of the bacterial population through Sole Source Carbon Utilization (SSCU) patterns [14]. Based on carbon utilization, BIOLOG (<https://biolog.com>) introduced the technique into ecological studies to estimate metabolic potential of microbial communities. The BIOLOG plate method was used for comparing metabolic activity of heterotrophic microbial communities from different habitats using redox chemistry. These plastic microtiter plates contain colorless tetrazolium dye with 95 different carbon substrates (e.g. carbohydrates, amino acids, amines, amides, carboxylic acids, etc.) in wells, and one well without any substrate used as control. Soil samples are shaken in a suitable solution; the soil solution is inoculated into the plate wells and incubated at constant temperature. Soil microorganisms oxidize the substrates in the microtiter plate wells and also reduce the tetrazolium dye to a violet formazan which is measured spectrophotometrically [15]. Currently BIOLOG has developed various microtiter plates for identification of aerobes, anaerobes, yeasts and filamentous fungi. This method is rapid, convenient and can help to identify over 2900 bacteria, fungi and yeast. But the drawback is the microbes which are cultivable and able to grow in high-nutrient conditions can only be studied. This method was used mainly to investigate microbial communities from

bulk and rhizosphere soil and to estimate the impact of stressors on soil microbial communities [15]. The BIOLOG microplates were used for identifying the bacterial population in the contaminated soils around the Explosive Institute of Riyadh community (Saudi Arabia) which shown *Proteobacteria* to be the dominant phylum in all the five regions with Burkholderiaceae, Chromatiaceae and Pseudomonadaceae common and dominant families representing 49% of the relative abundance at the family level [16]. Compared to earlier microtiter plates, BIOLOG EcoPlate are more in use for Community Level Physiological Profiling (CLPP) as it contains three replicates of 31 different carbon sources and one control well per replicate (Biolog Inc., Hayward, CA, USA) [15].

5. Culture independent approaches

The viable source of information regarding the microbial players in the soil can be discovered through the biomolecules such as lipids, DNA, RNA and proteins. The extraction procedures of the biomolecules from the soil are another challenge as the content of the soil, structure and humic acids varies from place to place, time to time. Over the years, several procedures has been devised for extracting the biomolecules such as nucleic acids but still there is the compromise in concentration or quality of the biomolecules. However, in recent years there has been huge progress in developing the new techniques for studying the microbes in soil and other environments.

5.1 Microbial lipid based techniques

As we know fatty acids are the components of the cellular membrane of all living cells, and their composition can reveal the types of organisms present without actual culturing the micro-organisms. The constant proportion of fatty acids microbial cell biomass as well as the presence of signature fatty acid and their profile helps to differentiate major taxonomic group within the community.

5.1.1 Phospholipid fatty acid analysis (PLFA)

Phospholipid fatty acid analysis (PLFA) quantifies a set of biomarkers that track primarily viable biomass, avoids culturing of micro-organisms, and represents in-situ conditions. PLFA provides a community measurement that is phenotypic rather than genotypic in nature. It does not give information on species composition but rather is analogous to the ecological concept of functional groups [17]. In a comparison study regarding the PLFA analysis and 16S rRNA gene metabarcoding of bacterial communities across the biomes, the PLFA profiling has been found better in distinguishing bacterial community [18]. It was also noticed that the PLFA profiling was better at detecting community responses to heavy metal pollution [19]. The other method which is also based on fatty acid extraction and profiling is fatty acid methyl ester and was used for estimating microbial biomass and characterizing microbial community composition in soil [20].

5.1.2 Fatty acid methyl ester (FAME)

Fatty acid methyl ester (FAME) provides information on the microbial community composition based on groupings of fatty acids. The fatty acids are extracted by saponification followed by derivatization to give the respective FAMEs, which are then analyzed by gas chromatography. The pattern thus obtained is compared to a

reference FAME database to identify the fatty acids and their corresponding microbial signatures by multivariate statistical analyses [21]. Bacterial fatty acids are highly conserved due to their role in cell structure and function and are the major constituents of the lipid bilayer of bacterial membranes and lipopolysaccharides. They have been used extensively for taxonomic and identification purposes. Whole cellular FAME content is a bacterial profile and is a direct and stable expression of the cellular genome. The cellular fatty acid pattern is a phenotypic character that is not affected by mutations, acquisition or loss of plasmids. The use of fatty acid analysis by gas chromatography for the identification of bacteria is rapid, efficient, reproducible and used for the identification of both clinical and environmental isolates. It has been used to study microbial community composition and population changes due to agricultural practices. Miura and her associates reported that EL- FAME method was simple and would produce similar results to PLFA method for bacteria in both quantitative and qualitative assessments when comparing different soils across ecosystems [20].

The importance of FAME analysis for the identification of bacteria is based on the large structural differences within these molecules viz., (i) variation in length (8–20 C-atoms), (ii) presence of saturated and monounsaturated fatty acids, (iii) occurrence of branched fatty acids (iso and anteiso fatty acids or methylated within the molecule), (iv) occurrence of cyclopropane fatty acids (17:0c, 19:0c), (v) occurrence of hydroxy-fatty acids with an OH-group at position two or three of the molecule. For classification or identification of bacteria the presence of distinct fatty acids and their relative amount is analyzed and compared with the fatty acid profiles of reference strains [22].

The microbial community characterization using nucleic acids has been further discussed as follows:

5.2 Non-PCR based techniques

5.2.1 DNA re-association

DNA re-association kinetics measures the genetic complexity of the microbial community and has been used to estimate microbial diversity. Total DNA is extracted from environmental samples, purified, denatured and allowed to re-anneal. The rate of hybridization or re-association will depend on the similarity of sequences present. As the complexity or diversity of DNA sequences increases, the rate of re-association of DNA will decrease. Under specific conditions, the time needed for half of the DNA to re-associate (the half association value $C_0t_{1/2}$) can be used as a diversity index, as it takes into account both the amount and distribution of DNA re-association. The parameter controlling the re-association reaction is concentration of DNA product (C_0) and time of incubation (t), usually described as the half association value, $C_0t_{1/2}$ (the time needed for half of the DNA to re-associate). Under specific conditions, $C_0t_{1/2}$ can be used as a diversity index, as it takes into account both the amount and distribution of DNA re-association [23]. Alternatively, the similarity between communities of two different samples can be studied by measuring the degree of similarity of DNA through hybridization kinetics [24].

5.2.2 Guanine plus cytosine (G+C) content of DNA

Differences in the guanine plus cytosine (G+C) content of DNA can be used to study the bacterial diversity of soil communities [25]. It is based on the knowledge that microorganisms differ in their G+C content and that taxonomically related

groups only differ between 3 and 5%. Even though GC fractionation provides coarse level of resolution as different taxonomic groups may share the same G+C range, it is probably the only technique that is completely independent of any previous knowledge regarding which bacterial populations comprise the community or their genomic content [26]. However, this method can be used with other PCR based methods like DGGE or TGGE for better accessibility of the microbial picture.

5.2.3 Reverse sample genome probing (RSGP)

This method utilizes genome microarrays to analyze microbial community composition. RSGP has four steps: (1) isolation of genomic DNA from pure cultures; (2) cross-hybridization testing to define species with less cross hybridization (<70%). DNA fragments with greater than 70% cross-hybridization are considered to be the same species; (3) preparation of genome arrays by spotting known amounts of denatured genomic DNAs from all identified standards onto a solid support; and (4) random labeling of a defined mixture of total community DNA and internal standard, hybridization of the labeled probe with the genome array and detection and analysis of the individual dot hybridization data [27]. Due to low level of hybridization, low levels of gene expression cannot be quantitated which limits the use of this technique. However, this technique allows the coverage of the uncultured component of environmental microbial communities. Although possible in principle, the problem of linking a specific DNA fragment to a particular strain is formidable and requires extensive characterization of any metagenome through cloning and sequencing.

5.3 PCR based methods

Targeting the 16S rDNA is used extensively to study prokaryote diversity and allows identification of prokaryotes. The prediction of phylogenetic 18S rDNA and internal transcribed spacer (ITS) regions are increasingly used to study fungal communities. Soil community DNA is extracted, purified and amplified using either specific or universal primers, the resultant products are then separated by various ways and analyzed accordingly.

5.3.1 Highly repeated sequence characterization or microsatellite regions

During the process of evolution both in prokaryotic and eukaryotic organisms, there is the accumulation of highly repetitive short DNA sequences (1–10 bp) throughout their genomes, which can be used in differentiation between the organisms at species or strain level. Highly repeated sequences are also termed as microsatellite regions and have been used for identification of mycorrhiza. Fingerprints of the PCR-amplified microsatellites can be compared using similarity indices to investigate differences between or among the species. The design of primers is solely dependent on the sequence information of microsatellite regions. The use of this method to study microbial diversity may be limited depending on the complexity of the community [12].

5.3.2 Random amplified polymorphic DNA

In 1990, William and his team developed a method which includes amplification of DNA fragments by using short arbitrary primer targeting multiple loci in genomic DNA, generating unique profile (amplicons of various lengths [28]). Both genomic variations between bacterial species and genetic polymorphism between

bacterial strains could be identified based on the differences in the molecular size and the number of DNA fragments obtained (**Figure 2**). RAPD analysis was used to study metagenome diversity in soil microbial community of arid zone plants [29] and in soil affected by industrial pollutants [30].

5.3.3 Restriction fragment length polymorphism (RFLP)/amplified ribosomal DNA restriction analysis (ARDRA)

It is another tool used to study microbial diversity that relies on DNA polymorphisms. PCR amplified rDNA is digested with a 4-base pair cutting restriction enzyme. Different fragment lengths are detected using agarose or non-denaturing polyacrylamide gel electrophoresis in the case of community analysis. RFLP fingerprint can be used to measure bacterial community structure. ARDRA is a DNA fingerprinting technique based on PCR amplification of 16S ribosomal DNA using primers for conserved regions, followed by restriction enzyme digestions and agarose gel electrophoresis. ARDRA was used successfully to study and compare the microbial diversity in copper contaminated soils [31]. Sklarz and his associates evaluated the use of amplified rDNA restriction analysis assay for identification of bacterial communities and concluded that ARDRA based dendrograms may not mirror 16S rDNA sequence based phylogenetic trees [32].

5.3.4 Terminal restriction fragment length polymorphism (T-RFLP)

T-RFLP follows the same principle as RFLP except that one PCR primer is labeled with a fluorescent dye, such as TET (4,7,2V,7V-tetrachloro-6-carboxyfluorescein) or 6-FAM (phosphoramidite fluorochrome 5-carboxyfluorescein), which allows detection of only the labeled terminal restriction fragment [33]. This simplifies the banding pattern, allowing the analysis of complex communities as well as providing information on diversity as each visible band represents a single operational taxonomic unit or ribotype. The banding pattern can be used to measure species richness and evenness as well as similarities between samples. Terminal restriction fragment length polymorphism (T-RFLP) has higher resolution and is more comprehensive than cultivation-based methods [13]. This procedure can be automated to allow sampling and analysis of a large number of soil samples with recent developments in bioinformatics, several Web-based T-RFLP analysis programs have been developed, which enable researchers to rapidly assign putative identities based on a database of fragments produced by known 16S rDNA

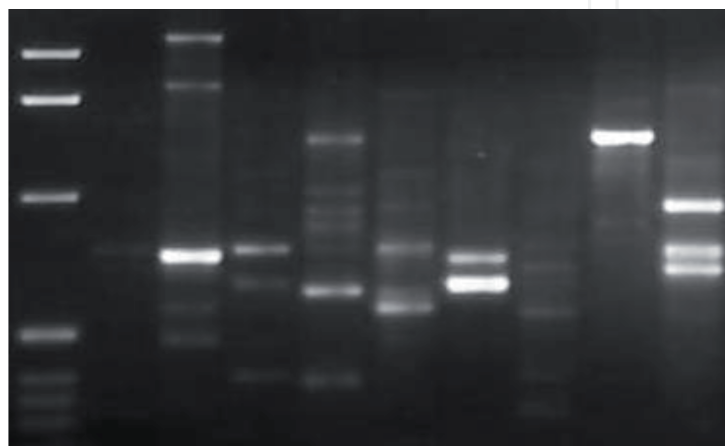


Figure 2.
RAPD profile of bulk soil DNA.

sequences. This technique has been successfully applied to the composition and diversity analysis of soil microbial communities under different environmental conditions [34–36]. T-RFLP analysis was used to reveal the composition and diversity of soil bacterial communities along an altitude gradient in Wuyi Mountains addressing relationship between the composition and structure of the soil microbial communities and the vegetation types and cause of the difference in soil microbial communities under different vegetation types [13].

5.3.5 Ribosomal intergenic spacer analysis (RISA)/automated ribosomal intergenic spacer analysis (ARISA)

Ribosomal intergenic spacer analysis (RISA) is a simple, single-step PCR-based method for profiling microbial diversity that detects the variation in size of the intergenic transcribed spacer (ITS) region between the bacterial 16S and 23S rRNA genes [37]. RISA has been extensively used to profile microbial diversity in a range of environmental niche. Intergenic spacer (IGS) region between the 16S and 23S ribosomal subunits is amplified by PCR, denatured and separated on a polyacrylamide gel under denaturing conditions. This region encodes tRNAs and is useful for differentiating between bacterial strains and closely related species because of heterogeneity of the IGS length and sequence. In RISA, the sequence polymorphisms are detected using silver stain while in ARISA the forward primer is fluorescently labeled and is automatically detected using laser. ARISA is a rapid and effective community analysis technique which can be used in conjunction with other accurate labor-intensive methods (e.g., 16S rRNA gene cloning and sequencing) for fine-scale spatial and temporal resolution [38]. Delmont and co-workers evaluated the different DNA extraction protocols by using RISA in study of soil microbial communities and reported that the total community DNA extracted from different DNA extraction procedures generated the different RISA profiles [39].

5.3.6 Temperature/denaturing gradient gel electrophoresis (TGGE/DGGE)

Muyzer with his teammates expanded the use of DGGE to study microbial genetic diversity [40]. In DGGE, DNA is extracted from soil samples and amplified using PCR with universal primers targeting part of the 16S or 18S rRNA sequences. The 5'-end of the forward primer contains a 40 base pair (16S rRNA) or 50 base pair (18S rRNA) GC clamp to ensure that at least part of the DNA remains intact or to avoid the complete dissociation of the amplified products into single strands which might flow away from the gel. This is necessary as the separation of amplified DNA on a polyacrylamide gel with a gradient of increasing concentration of denaturants (formamide and urea) will occur based on melting behavior of the double-stranded DNA (**Figure 3**). TGGE uses the same principle as DGGE except the gradient is temperature rather than chemical denaturants. Polymorphism based on the separation of partially melted 16S rDNA a linear temperature gradient. It represents the sequence variations other than the restriction sites also. Sequence variation among different PCR amplicons determines the melting behavior, and therefore amplicons with different sequences stop migrating at different positions in the gel. However, it covers only less than less than 400 bp of 16S gene. Conservative fragments of available 16S rDNA sequences were mined and searched for candidate primers within the fragments by measuring the coverage rate defined as the percentage of bacterial sequences containing the target. Thirty predicted primers with a high coverage rate (>90%) were identified [41] and can be successfully used for generating DGGE fingerprints. Abundance of denitrifying genes and microbial community structure in volcanic soils [42], assessment of silver nanoparticles on soil bacterial diversity [43],

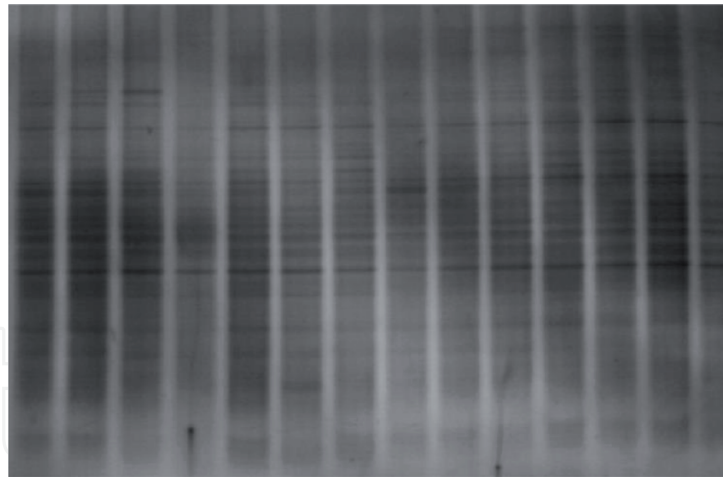


Figure 3.
DGGE profile of the 16S rDNA (V5 region) of soil rhizosphere samples.

effect of long term fertilization on bacterial and fungal diversity in brown soils [44], Changes of Soil Bacterial Diversity in a Semi-Arid Ecosystem [45] has been successfully studied by using DGGE profiles.

How to get more out of molecular fingerprints remains the question, because estimating the species diversity is the most important factor towards better understanding of the microbial load in the given soil environment. Various investigations indicate that the species richness is the simplest measure of diversity. Any characterized soil environment shows few common species but in greater abundance as compared to more uncommon species harboring in the same environment but in less number therefore, one should consider the species evenness also. A common point of agreement on the diversity is that, species richness and evenness aggregately estimates the diversity and these components should be defined so that they are independent of each other [46]. One of the most commonly used evenness measure is Pielou's evenness index, evenness expresses how evenly the individuals in the community are distributed over the different species [47].

Claude Shannon originally proposed this measure which has been useful in comparing diversity between the different habitats [48]. Shannon index is easy for calculation and interpretation, the Shannon index generally ranges between 1.5 and 3.5 for many ecological habitats. Simpson diversity is less sensitive to richness and more sensitive to evenness; whereas Shannon diversity is more sensitive to evenness. For comparing the similarity between the samples, one can apply Jaccard index or Sorensen index. They are most widely used and are based on the presence/absence of species in paired assemblages and are simple to compute [49]. A modified version of Jaccards has been proposed by Chao and his colleagues to include the effect of unseen shared species, based on either (replicated) incidence or abundance based sample data [50].

As the soils have a dynamic nature, there is always a shift in microbial population during the different time or during different seasons. At a given time, particular number of species significantly dominates over other and vice versa. Therefore, such shift in microbial load can be estimated using different comparison tools. Some of them are cluster analysis [51], moving window analysis [52] or by visual inspection [53] or the Dice index [54] which can be applied for estimation of microbial shift during a defined period of time. The percent change in microbial composition between the two sampling interval can be calculated by subtracting percent similarity (calculated by any of the similarity indices) from 100. This can be done for consecutive sampling points over a period of experimentation. Using that % change values, moving window analysis is plotted between consecutive sampling

points. The rate of change (Δt) value is calculated as the average of the respective moving window curve data points [43]. Higher Δt values represent higher shifts between two successive sampling points.

The above diversity analysis parameters are based on the relative abundance of the species in the given sample and extensively used to analyze denaturing gradient gel electrophoresis profiles.

5.3.7 DNA microarray

Nucleic acid (DNA/RNA) hybridization using specific probes is another qualitative and quantitative tool in molecular bacterial ecology. Oligonucleotide or polynucleotide probes designed from known sequences ranging in specificity from domain to species can be tagged with markers (fluorescent) at the 5'-end [55]. Based on nucleic hybridization, DNA microarrays are developed and are used to detect and identify bacterial species in soil or any other environmental samples. In this method, a single chip contains thousands of probes with high specificity can be used in identifying microbial species in soil or environmental samples. The amplified products from the soil DNA is hybridized against the known molecular probes, which are attached on the microarrays. The hybridized spots are detected and scored using microscopy. Its advantage is the rapid and replicated evaluation of the samples [56] but while analyzing the soil or environmental sample there may be risk of cross hybridization. Further, there are two different microarray chips namely 16S rRNA gene microarray or PhyloChip and Functional gene microarray or GeoChip (**Figure 4**).

PhyloChip is the most widely used phylogenetic array. In here, 16S rRNA genes are extracted followed by PCR amplification of metagenome DNA and then biotin labeled for PhyloChip hybridization, the signals are detected using digital image detection. It is Affymetrix-based technology consisting 25-mer oligonucleotide probes which differentiate between the 16S rRNA gene sequences in microbial communities. The recent version of the PhyloChip (G3) has probes targeting ~60,000 operational taxonomic units (OTUs), representing two domains (Archaea and Bacteria), 147 phyla, 1123 classes, 1219 orders, 1464



Figure 4.
 DNA microarray chip (PhyloChip or GeoChip).

families, and 10,993 subfamilies [57]. Functional gene arrays are used in targeting genes involved in various biogeochemical cycling processes which help in determining the functional composition microbial communities. GeoChip, a functional gene array is widely used which targets hundreds of functional genes involved in biogeochemical, ecological, and environmental analyses. Arrays for detecting specific functional processes, such as nitrogen cycling, methanotrophy, stress responses, hydrogen activity etc. are available. The version 5.0 GeoChip contains about 167,000 50-mer oligonucleotide probes covering ~395,000 coding sequences from >1590 functional genes related to various microbes, mineral cycling, energy metabolism, antibiotic resistance, metal homeostasis and resistance, secondary metabolism, organic remediation, stress responses, bacteriophages, and virulence [58].

5.3.8 Single-strand conformation polymorphism (SSCP)

SSCP was developed to detect gene polymorphism in human DNA and mutations by comparing PCR products [59]. DNA fragments are amplified, followed by denaturation and separation in non-denaturing polyacrylamide gel. Single-stranded DNA is separated on a polyacrylamide gel based on differences in mobility caused by their folded secondary structure. DNA fragments of equal size with no denaturant results into folding and movement, depending on the DNA sequences. Single stranded DNA forms secondary structures which are unique to its nucleotide sequence. This secondary structure hinders the movement of DNA in polyacrylamide gel and hence, the banding pattern is obtained at different position from the fragment of equal size but amplified from genetically different species. The advantage of SSCP over DGGE is that here GC clamp is not required during PCR but has a limitation with that there is high rate of re-annealing of DNA strands after an initial denaturation during electrophoresis which can however be overcome using a phosphorylated primer during PCR [21]. Smalla and her team studied the bacterial diversity of the soils by assessing the DGGE, T-RFLP, and SSCP fingerprints of 16S rRNA gene fragments. Although, the fragments amplified comprised different variable regions and lengths, DGGE, T-RFLP and SSCP analyses led to similar findings: (a) a clustering of fingerprints which correlated with soil physico-chemical properties, (b) little variability between the replicates of the same soil, (c) the patterns of the two brown soils were more similar to each other than to those of the other two soils, and (d) the fingerprints of the different soil types revealed significant differences in a permutation test [60]. SSCP fingerprints were also used in study of microbial diversity in landslide soils; the more detailed profile of fungal diversity was obtained using this method [59].

5.3.9 High resolution melt curve (HRM)

High resolution melt curve typically involve qPCR amplification followed by a melting curve collected using a fluorescent dye. It follows extraction of nucleic acids and qPCR with fluorescent labeled dye. Following amplification and quantification of the 16S rRNA gene, a high-resolution melting curve analysis is obtained. The procedure melts the amplified 16S rRNA gene products between the temperatures ranging from 72°C and ending at 95°C, with fluorescence measurements taken at every 0.1°C increment. Melting curves are normalized to relative fluorescence units (RFU) in a specified “melt region” (e.g., 83.5–89.5°C), thereby negating the effect of absolute RFU values. The melt region is autocalled by the melt analysis software (Precision Melt Analysis) [61]. The effects of herbicide on soil bacterial diversity were efficiently studied using HRM analysis [62].

5.3.10 Real time PCR (qPCR)

Real time PCR (qPCR) helps in real time quantification of the active population present in the environmental sample by targeting ribosomal genes specific to different species of the microbial community. It gives an idea of phylogenetic community composition by assessing the range of phyla or classes [63]. Quantitative PCR (qPCR), or real-time PCR, has been used in microbial investigations to measure the abundance and expression of taxonomic and functional gene markers. Unlike traditional PCR, which relies on end-point detection of amplified genes, qPCR uses either intercalating fluorescent dyes such as SYBR Green or fluorescent probes (TaqMan) to measure the accumulation of amplicons in real time during each cycle of the PCR. Software records the increase in amplicon concentration during the early exponential phase of amplification which enables the quantification of genes (or transcripts) when they are proportional to the starting template concentration. When qPCR is coupled with a preceding reverse transcription (RT) reaction, it can be used to quantify gene expression (RT-qPCR). qPCR is highly sensitive to starting template concentration and measures template abundance in a large dynamic range of around six orders of magnitude. Several sets of 16S and 5.8S rRNA gene primers have been designed for rapid qPCR based quantification of soil bacterial and fungal microbial communities [64]. Aparna and his co-workers investigated the quantification of 16S rDNA, *amoA* and *nifH* genes in organic and inorganic cropping soil using qPCR, which shown a clear 1.8 fold increase in both organic cropping and organic orchard soils whereas the abundance of *amoA* gene decreased by 22- and 11-folds in organic cropping and orchards [65].

5.4 Sequencing based methods

5.4.1 Clone library sequencing

Clone library sequencing involves extraction of environmental DNA followed by amplification of partial or full length of 16S rRNA (27f (5'-AGRGTTCGAT CMTGGCTCAG)) and 1492R (5'-GGTTACCTTGTTACGACTT). The amplified sequences are then ligated and cloned in a suitable vector. Further the clones containing organism-specific 16S rRNA gene fragments are purified and sequenced from each terminus. Sequences are then assembled and checked for QC. Phylum, class, order, family, subfamily, or OTU placement are determined when a clone surpasses similarity thresholds of 80, 85, 90, 92, 94, or 97%, respectively. When similarity to nearest database sequence falls below 94%, the clone is considered to represent a novel subfamily and a novel class was denoted when similarity is less than 85% [66]. A comprehensive study was done to characterize the relative abundance, diversity and composition of acidobacterial communities across a range of soil types using clone library analyses [67]. Vázquez and co-workers collected the soil and sediment samples from the coastal region in response to diesel spill, based on DGGE analysis they selected six soils and two sediment samples for identification of bacterial community structure using clone library analyses [68].

5.4.2 Amplicon sequencing

Soil DNA is extracted and then 16S/18S rDNA genes are amplified by using specific set of specific primers targeting variable regions of 16S/18S rDNA, followed by purification of fragments using magnetic beads, then adapters are ligated and the library of fragments (clones) is amplified and the samples are sequenced using NGS platform. The dataset obtained after sequencing can be

compared with Ribosomal Database Project (RDP) for identification of microbial community harboring in environmental sites [69]. Using NGS, it is possible to resolve highly complex microbiota compositions with greater accuracy, as well as to link microbial community diversity with niche function. The soil DNA was extracted and the two step PCR amplification was done using domain-specific primer, 515F/806R (for prokaryotes), F1427/R1616 (for eukaryotes) or ITS1/ITS2 (for fungi) followed by purification and amplicon sequencing using illumina platform for identification of soil bacterial and fungal community. [70]. Schöler and colleagues has stated brief highlights regarding the crucial steps that should be considered for accurate analysis and data interpretation while opting for amplicon sequencing using marker genes [71].

5.4.3 Shotgun metagenome sequencing

Exploration of microbial universe in environmental systems through shotgun metagenome sequencing allows us to investigate deeper strata of the microbial communities and provides an unbiased view on the phylogentic and functional composition of the environmental microbial communities [21]. For sequencing of the soil or environmental DNA the steps involved are, extraction of high quality total community DNA followed by fragmentation to obtain desired length of fragments, which are further purified followed by amplification and sequencing using desired or available sequencing platform. The data set thus obtained is analyzed with the offline (MEGAN) or web- based software (MG-RAST) for visualization or comparison of the pictures of microbial world. Although shotgun metagenomic sequencing does not involve the biased amplification of 16S rRNA genes, the relative organism abundances inferred from metagenomic sequences vary significantly depending on the DNA extraction and sequencing protocol utilized [72]. Utilizing the Illumina sequencing platform, the impact of N fertilization on soil microbial communities was studied, where the field was cultivated under soybean and corn [73]. Luo and is team reported whole-genome shotgun metagenomic analysis of microbial communities of temperate grassland soils that experienced 2°C infrared heating for 10 years and observed that the heated communities showed significant shifts in composition and predicted metabolism, and these shifts were community wide as opposed to being attributable to a few taxa. Key metabolic pathways such as cellulose degradation (~13%), CO₂ production (~10%), and to nitrogen cycling (~12%), were enriched under heating, which was consistent with independent physicochemical measurements. [74].

6. Applications of metagenomics in soil environment

Soil habitats contain the greatest microbial diversity of all the other environments on earth. So far, metagenomic approaches able to scratch the surface of the genomic, metabolic and phylogenetic diversity stored in the soil metagenome. Being a most diverse and challenging environment, soil holds an unlimited resource for the discovery of novel genes, enzymes, natural products, bioactive compounds, and bioprocesses. Soil metagenomic methods, specifically isolation of soil DNA followed by construction and screening of clone libraries, enable to look at more complete picture of soil microbial communities and to better understand their interactions. This methodology is of great potential for use in the studies of soil microbial communities and their functional genes, and in the discovery of new biocatalysts for industry. The sustainable economic future of modern industrialized

societies requires the development of novel molecules, enzymes, processes, products, and applications. Application of metagenomics in soil had helped many microbiologists for uncovering its huge potential by overcoming the need for culturing the microbes in pure form as well as capturing the unculturable ones. High-throughput and sensitive screening methods are employed to overcome the complexity of soil metagenome. In general, screening of soil metagenome libraries relies on metabolic activity (function-driven approach) or on nucleotide sequence (sequence driven approach) whichever is suitable and feasible. By employing one or combination of these methods, researchers can discover vast number of novel products of industrial or agricultural use.

6.1 Soil health

Soil microbes are essential for the proper functioning (i.e., health) of a soil. Yet, there is little information about how to assess the life present in soil to determine if a given soil is healthy or not. Till date, biological measures of soil health have been centered on biological functions, such as respiration or nitrogen mineralization. Soil metagenomics is a promising approach in describing the functional potential of the soil microbial community, which might yield greater insight into the health of a soil than taxonomy-based metrics. Soil health cannot be measured directly, so quality/health of soil is evaluated by indicators. Structural and functional diversity of microbes, presence/absence of important microbial players, microbial activity (respiration, DNA replication and cell division), nutrient cycling, production of cofactors and secondary metabolites, and response to biotic and abiotic stress can be used as biological indicators of soil health. Biological activity of the soil can be known by estimating dehydrogenase enzyme in the soils [75]. β -glucosidase involves in hydrolysis and biodegradation of various β -glucoside present in decomposing plant debris in soil. β -glucosidase is characteristically useful as a soil quality indicator, and may give a reflection of past biological activity, the capacity of soil to stabilize the soil organic matter, and can be used to detect management effect on soils [76]. The amount of these enzymes activity indicates the biological capacity of soil, for the enzymatic conversion of the substrate and also has an important role in the ecology of microorganisms in the ecosystem. The degradation and hydrolysis of chitin is achieved by Chitinase a key enzyme responsible for the same. Its presence in different forms in the ecosystem has demonstrated its effectiveness in the control of soil-borne diseases. Arylsulphatases are responsible for the hydrolysis of sulphate esters in the soil [77] and are secreted by bacteria into the external environment as a response to sulphur limitation. Phosphatases are good indicator of soil fertility and are believed to play critical roles in P cycles [78]. When there is a signal indicating P deficiency in the soil, acid phosphatase secretion from plant roots is increased to enhance the solubilization and remobilization of phosphate, thus influencing the ability of the plant to cope with P-stressed conditions [79].

6.2 Industrial use

Activity-based screening has the potential to detect entirely novel genes encoding new types and classes of enzymes or to identify the new bioactive compounds. In addition, it is selective for full-length genes and functional gene products. Cellulases has been isolated from various natural environments like soil, rumen, compost soil and many more using metagenomic technique by constructing the metagenomic libraries followed by screening of the biologically active clones.

Cellulases are used in animal feeds for improving the nutritional quality and digestibility, in processing of fruit juices, and in baking, while de-inking of paper is yet another emerging application [80]. Alvarez with his team isolated and characterized a novel cellulase by functional screening of a metagenomic library derived from sugarcane field land soil [65]. Lipases have been found in many species of animals and plants. The enzymes from microbial sources (such as bacteria, yeast and fungi) are currently receiving particular attention because of their actual and potential applications in industry mainly in the detergents, oils and fats, dairy and pharmaceutical industries [80]. A novel metagenomic xylanase has been isolated from compost-soil metagenome that shows alkali stability and thermostability, thus bearing a potential application in paper and pulp industry in pulp bleaching [81]. Vidya with her co-associates isolated a thermostable and calcium-dependent amylase from a soil by constructing and screening soil metagenomic library and suggested its applications in baking and de-starching [82].

6.3 Antibiotics discovery

Since the discovery of first antibiotic there has been tremendous improvement in human health. However by traditional techniques we can tap only known antibiotics identified from known organisms. However, soil metagenomics presents a greater opportunity for discovery of vast number of antibiotics yet to be discovered. Researchers have isolated new antibiotics, one of such example is Turbomycin A and B from a metagenomic library of soil microbial DNA [83]. Some of them, isolated the genes encoding enzymes required for synthesis of other antibiotics e.g., Voget and co-workers found one amidase-positive clone while general screening of a soil metagenomic library for biocatalysts [84]. Amidases are used in the biosynthesis of β -lactam antibiotics [85]. Chang and Brady reported a gene cluster *bor* from soil that encodes indolotryptoline based compounds, a small and relatively rare family of natural products that exhibit potent activity against certain tumor cell lines [86]. Through activity-based screening of *E. coli* plasmid library from four agricultural soil samples 45 clones were isolated which were showing resistant to tetracycline, chloramphenicol, kanamycin, minocycline, gentamicin, amikacin, aminoglycosides, streptomycin and rifampicin [87].

6.4 Bioremediation

Due to industrialization, dumping or burying of harmful wastes in the soil or water stream resulted in degradation of surrounding cultivable or arable lands. By tapping these affected soil environments one can discover the microbes degrading the harmful hydrocarbons. As microbes are directly involved in carbon cycles, they may play a role in breaking down the carbon present in the harmful hydrocarbons. Using a function-driven metagenomic approach, new metabolic pathways involved in the biodegradation of aromatic compounds can be discovered. Many rhizospheric microbes produce biosurfactant; these biomolecules play vital role in motility, signaling, and biofilm formation, indicating that biosurfactant governs plant-microbe interaction. In agriculture, biosurfactants can be used for plant pathogen elimination and for increasing the bioavailability of nutrient for beneficial plant associated microbes. Biosurfactants can widely be applied for improving the agricultural soil quality by soil remediation. These biomolecules can replace the harsh surfactant presently being used in million dollar pesticide industries [88]. Studies related to adaptation of microbes in toxic environments may give rise to trace new metagenomic communities useful for efficient bioremediation.

6.5 Agriculture

Soil microbes play an important role in triggering the growth, stress responses and defense in plants. Understanding the relationship between the soil microbiota and plants using soil metagenomics will be very helpful in designing the crop systems. Metagenomics study of the soils supplemented with organic manures obtained from various farm animals would be very much helpful in formulating the fertilization strategies and reducing the dependence on inorganic fertilizers. For sustainable agricultural production, beneficial microbes of agricultural importance can serve as a crucial alternative. Metagenomics help in the prediction of microbial community structure and, therefore, can tackle and address fundamental scientific questions related to agriculturally important microorganisms. This approach has been successfully explored for the assessment of the diazotrophs belonging to the rhizosphere of native red kidney beans (RKB) of the Western Indian Himalaya by targeting *nifH* [89].

7. Conclusion

Soil being a complex environment, understanding of the interactions within microbial communities and soil environments is very much necessary for designing the strategies for sustainable agriculture, bioremediation, and human welfare. Soil metagenomic data will reveal potential activities present in microbial communities which can be harnessed for future good use. The huge diversity of soil microorganisms, together with the heterogeneity of the soil environment, hinders analyses of microbial diversity, structure and the linking functional processes. One of the major challenges for soil metagenomics is to develop methods to capture the heterogeneity and dynamics of complex soil microbial communities, both over time and spatially. However, new methods will greatly increase the number of samples that can be analyzed in the future. All the methods used for the investigating microbial diversity and activity contain inherent biases and it is necessary to understand the underlying mechanisms in order to be aware of the drawbacks and limitations, and to appreciate the strengths and weaknesses of each approach [9]. Nevertheless, these methods are starting to dissect the soil microbial biomass and the soil metagenome, and will, in the future, enable a greatly improved understanding of microbial community dynamics and interactions relevant to soil functions.

The TerraGenome International Soil Metagenome Sequencing Consortium has been dedicated towards soil metagenomics and helps in co-coordinating the global researchers for the discovery of the soil hidden treasures (<http://www.terragenome.org/>).

IntechOpen

IntechOpen

Author details

Shrinivas N. Sabale, Padmaja P. Suryawanshi and Krishnaraj P.U.*
University of Agricultural Sciences, Dharwad, Karnataka, India

*Address all correspondence to: krishnarajpu@gmail.com

IntechOpen

© 2019 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Aktar W, Sengupta D, Chowdhury A. Impact of pesticides use in agriculture: Their benefits and hazards. *Interdisciplinary Toxicology*. 2009;**2**(1):1-2. DOI: 10.2478/v10102-009-0001-7
- [2] Christopher J. *Living Soils: The Role of Microorganisms in Soil Health*. Strategic Analysis Paper. Nedlands: Future Directions International; 2017
- [3] Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM. Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products. *Chemistry & Biology*. 1998;**5**(10):R245-R249. DOI: 10.1016/S1074-5521(98)90108-9
- [4] Berendsen RL, Pieterse CM, Bakker PA. The rhizosphere microbiome and plant health. *Trends in Plant Science*. 2012;**17**(8):478-486. DOI: 10.1016/j.tplants.2012.04.001
- [5] Daniel R. The soil metagenome—A rich resource for the discovery of novel natural products. *Current Opinion in Biotechnology*. 2004;**15**(3):199-204. DOI: 10.1016/j.copbio.2004.04.005
- [6] Hernandez-Leon R, Martínez-Trujillo M, Valencia-Cantero E, Santoyo G. Construction and characterization of a metagenomic DNA library from the rhizosphere of wheat (*Triticum aestivum*). *Phyton (Buenos Aires)*. 2012;**81**:133-137
- [7] Bashir Y, Pradeep Singh S, Kumar Konwar B. Metagenomics: An application based perspective. *Chinese Journal of Biology*. 2014:1-7. DOI: 10.1155/2014/146030
- [8] Doornbos RF, van Loon LC, Bakker PA. Impact of root exudates and plant defense signaling on bacterial communities in the rhizosphere. A review. *Agronomy for Sustainable Development*. 2012;**32**(1):227-243. DOI: 10.1007/s13593-011-0028-y
- [9] Hirsch PR, Mauchline TH, Clark IM. Culture-independent molecular techniques for soil microbial ecology. *Soil Biology and Biochemistry*. 2010;**42**(6):878-887. DOI: 10.1016/j.soilbio.2010.02.019
- [10] National Research Council. *The New Science of Metagenomics: Revealing the Secrets of our Microbial Planet*. Washington, DC: National Academies Press; 2007
- [11] Hartmann A, Rothballer M, Schmid M. Lorenz Hiltner, a pioneer in rhizosphere microbial ecology and soil bacteriology research. *Plant and Soil*. 2008;**312**(1-2):7-14. DOI: 10.1007/s11104-007-9514-z
- [12] Kirk JL, Beaudette LA, Hart M, Moutoglis P, Klironomos JN, Lee H, et al. Methods of studying soil microbial diversity. *Journal of Microbiological Methods*. 2004;**58**(2):169-188. DOI: 10.1016/j.mimet.2004.04.006
- [13] Wu Z, Lin W, Li B, Wu L, Fang C, Zhang Z. Terminal restriction fragment length polymorphism analysis of soil bacterial communities under different vegetation types in subtropical area. *PLoS One*. 2015;**10**(6):e0129397. DOI: 10.1371/journal.pone.0129397
- [14] Garland JL, Mills AL. Classification and characterization of heterotrophic microbial communities on the basis of patterns of community-level sole-carbon-source utilization. *Applied and Environmental Microbiology*. 1991;**57**(8):2351-2359
- [15] Stefanowicz A. The biolog plates technique as a tool in ecological studies of microbial communities. *Polish Journal of Environmental Studies*. 2006;**15**(5):669-676

- [16] Al-Dhabaan FA, Bakhali AH. Analysis of the bacterial strains using Biolog plates in the contaminated soil from Riyadh community. *Saudi Journal of Biological Sciences*. 2017;**24**(4):901-906. DOI: 10.1016/j.sjbs.2016.01.043
- [17] Joergensen RG, Wichern F. Quantitative assessment of the fungal contribution to microbial tissue in soil. *Soil Biology and Biochemistry*. 2008;**40**(12):2977-2991. DOI: 10.1016/j.soilbio.2008.08.017
- [18] Docherty KM, Borton HM, Espinosa N, Gebhardt M, Gil-Loaiza J, Gutknecht JL, et al. Key edaphic properties largely explain temporal and geographic variation in soil microbial communities across four biomes. *PLoS One*. 2015;**10**(11):e0135352. DOI: 10.1371/journal.pone.013535
- [19] Chodak M, Gołębiewski M, Morawska-Płoskonka J, Kuduk K, Niklińska M. Diversity of microorganisms from forest soils differently polluted with heavy metals. *Applied Soil Ecology*. 2013;**64**:7-14. DOI: 10.1016/j.apsoil.2012.11.004
- [20] Miura T, Makoto K, Niwa S, Kaneko N, Sakamoto K. Comparison of fatty acid methyl ester methods for characterization of microbial communities in forest and arable soil: Phospholipid fraction (PLFA) versus total ester linked fatty acids (EL-FAME). *Pedobiologia*. 2017;**63**:14-18. DOI: 10.1016/j.pedobi.2017.04.002
- [21] Rastogi G, Sani RK. Molecular techniques to assess microbial community structure, function, and dynamics in the environment. In: Ahmad et al., editors. *Microbes and Microbial Technology: Agricultural and Environmental Applications*. New York: Springer Science+Business Media, LLC; 2011. DOI: 10.1007/978-1-4419-7931-5_2
- [22] Sharmili AS, Ramasamy P. Fatty acid methyl ester (FAME) analysis of moderately thermophilic bacteria isolated from the Coramandel Coast, Chennai, Tamilnadu. *European Journal of Experimental Biology*. 2016;**6**(5):1-7
- [23] Torsvik V, Daae FL, Sandaa RA, Øvreås L. Novel techniques for analysing microbial diversity in natural and perturbed environments. *Journal of Biotechnology*. 1998;**64**(1):53-62. DOI: 10.1016/S0168-1656(98)00103-5
- [24] Fakruddin M, Mannan K. Methods for analyzing diversity of microbial communities in natural environments. *Ceylon Journal of Science (Biological Sciences)*. 2013;**42**(1):19-33. DOI: 10.4038/cjsbs.v42i1.5896
- [25] Nüsslein K, Tiedje JM. Soil bacterial community shift correlated with change from forest to pasture vegetation in a tropical soil. *Applied and Environmental Microbiology*. 1999;**65**(8):3622-3626
- [26] Holben WE. GC Fractionation allows comparative total microbial community analysis, enhances diversity assessment, and facilitates detection of minority populations of bacteria. In: *Handbook of Molecular Microbial Ecology. I: Metagenomics and Complementary Approaches*. New York: John Wiley & Sons, Inc.; 2011. pp. 183-196
- [27] Greene EA, Voordouw G. Analysis of environmental microbial communities by reverse sample genome probing. *Journal of Microbiological Methods*. 2003;**53**(2):211-219. DOI: 10.1016/S0167-7012(03)00024-1
- [28] Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*. 1990;**18**(22):6531-6535
- [29] Sharma G, Verma HN, Sharma R. RAPD analysis to study metagenome diversity in soil microbial community

of arid zone plants. Proceedings of the National Academy of Sciences, India Section B: Biological Sciences. 2013;**83**(2):135-139. DOI: 10.1007/s40011-012-0041-1

[30] Tilwari A, Chouhan D, Sharma R. Random amplified polymorphic DNA (RAPD) analysis of microbial community diversity in soil affected by industrial pollutants: Reference to Mandideep Industrial Area. African Journal of Microbiology Research. 2013;**7**(30):3933-3942. DOI: 10.5897/AJMR12.1720

[31] Smit E, Leeftang P, Wernars K. Detection of shifts in microbial community structure and diversity in soil caused by copper contamination using amplified ribosomal DNA restriction analysis. FEMS Microbiology Ecology. 1997;**23**(3):249-261. DOI: 10.1111/j.1574-6941.1997.tb00407.x

[32] Sklarz MY, Angel R, Gillor O, Soares MI. Evaluating amplified rDNA restriction analysis assay for identification of bacterial communities. Antonie Van Leeuwenhoek. 2009;**96**(4):659. DOI: 10.1007/s10482-009-9380-1

[33] Liu WT, Marsh TL, Cheng H, Forney LJ. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Applied and Environmental Microbiology. 1997;**63**(11):4516-4522

[34] Leckie SE. Methods of microbial community profiling and their application to forest soils. Forest Ecology and Management. 2005;**220**(1-3):88-106. DOI: 10.1016/j.foreco.2005.08.007

[35] Singh BK, Nazaries L, Munro S, Anderson IC, Campbell CD. Use of multiplex terminal restriction fragment length polymorphism for rapid and simultaneous analysis of different

components of the soil microbial community. Applied and Environmental Microbiology. 2006;**72**(11):7278-7285. DOI: 10.1128/AEM.00510-06

[36] Joo S, Lee SR, Park S. Monitoring of phytoplankton community structure using terminal restriction fragment length polymorphism (T-RFLP). Journal of Microbiological Methods. 2010;**81**(1):61-68. DOI: 10.1016/j.mimet.2010.01.025

[37] Borneman J, Triplett EW. Molecular microbial diversity in soils from eastern Amazonia: Evidence for unusual microorganisms and microbial population shifts associated with deforestation. Applied and Environmental Microbiology. 1997;**63**:2647-2653

[38] Fisher MM, Triplett EW. Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. Applied and Environmental Microbiology. 1999;**65**(10):4630-4636

[39] Delmont TO, Robe P, Cecillon S, Clark IM, Constancias F, Simonet P, et al. Accessing the soil metagenome for studies of microbial diversity. Applied and Environmental Microbiology. 2011;**77**(4):1315-1324. DOI: 10.1128/AEM.01526-10

[40] Muyzer G, De Waal EC, Uitterlinden AG. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Applied and Environmental Microbiology. 1993;**59**(3):695-700

[41] Wang Y, Qian PY. Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. PLoS One. 2009;**4**(10):e7401. DOI: 10.1371/journal.pone.0007401

- [42] Carvajal AM, Vargas RA, Alfaro M. Abundance of denitrifying genes and microbial community structure in volcanic soils. *Journal of Soil Science and Plant Nutrition*. 2016;**16**(3):677-688. DOI: 10.4067/S0718-95162016005000049
- [43] Marzorati M, Wittebolle L, Boon N, Daffonchio D, Verstraete W. How to get more out of molecular fingerprints: Practical tools for microbial ecology. *Environmental Microbiology*. 2008;**10**(6):1571-1581. DOI: 10.1111/j.1462-2920.2008.01572.x
- [44] Luo P, Han X, Wang Y, Han M, Shi H, Liu N, et al. Influence of long-term fertilization on soil microbial biomass, dehydrogenase activity, and bacterial and fungal community structure in a brown soil of Northeast China. *Annals of Microbiology*. 2015;**65**(1):533-542. DOI: 10.1007/s13213-014-0889-9
- [45] Ding GC, Piceno YM, Heuer H, Weinert N, Dohrmann AB, Carrillo A, et al. Changes of soil bacterial diversity as a consequence of agricultural land use in a semi-arid ecosystem. *PLoS One*. 2013;**8**(3):e59497. DOI: 10.1371/journal.pone.0059497
- [46] Jost L. The relation between evenness and diversity. *Diversity*. 2010;**2**(2):207-232. DOI: 10.3390/d2020207
- [47] Peet RK. The measurement of species diversity. In: *Section of Ecology and Systematic*. Vol. 5. Ithaca, New York: Langmuir Laboratory, Cornell University; 1974. pp. 285-307
- [48] Shannon CE. A mathematical theory of communication. *Bell System Technical Journal*. 1948;**27**(3):379-423
- [49] Magurran AE. *Ecological Diversity and its Measurement*. Princeton, New Jersey: Princeton University Press; 1988
- [50] Chao A, Chazdon RL, Colwell RK, Shen TJ. A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecology Letters*. 2005;**8**(2):148-159. DOI: 10.1111/j.1461-0248.2004.00707.x
- [51] Seghers D, Top EM, Reheul D, Bulcke R, Boeckx P, Verstraete W, et al. Long-term effects of mineral versus organic fertilizers on activity and structure of the methanotrophic community in agricultural soils. *Environmental Microbiology*. 2003;**5**(10):867-877. DOI: 10.1046/j.1462-2920.2003.00477.x
- [52] Possemiers S, Verthé K, Uyttendaele S, Verstraete W. PCR-DGGE-based quantification of stability of the microbial community in a simulator of the human intestinal microbial ecosystem. *FEMS Microbiology Ecology*. 2004;**49**(3):495-507. DOI: 10.1016/j.femsec.2004.05.002
- [53] Lozada M, Figuerola EL, Itria RF, Erijman L. Replicability of dominant bacterial populations after long-term surfactant-enrichment in lab-scale activated sludge. *Environmental Microbiology*. 2006;**8**(4):625-638. DOI: 10.1111/j.1462-2920.2005.00939.x
- [54] LaPara TM, Nakatsu CH, Pantea LM, Alleman JE. Stability of the bacterial communities supported by a seven-stage biological process treating pharmaceutical wastewater as revealed by PCR-DGGE. *Water Research*. 2002;**36**(3):638-646. DOI: 10.1016/S0043-1354(01)00277-9
- [55] Theron J, Cloete TE. Molecular techniques for determining microbial diversity and community structure in natural environments. *Critical Reviews in Microbiology*. 2000;**26**(1):37-57. DOI: 10.1080/10408410091154174
- [56] Gentry TJ, Wickham GS, Schadt CW, He Z, Zhou J. Microarray

applications in microbial ecology research. *Microbial Ecology*. 2006;**52**(2):159-175. DOI: 10.1007/s00248-006-9072-6

[57] Zhou J, He Z, Yang Y, Deng Y, Tringe SG, Alvarez-Cohen L. High-throughput metagenomic technologies for complex microbial community analysis: open and closed formats. *MBio*;6:e02288-14. DOI: 10.1128/mBio.02288-14

[58] Zengler K, Walcher M, Clark G, Haller I, Toledo G, Holland T, et al. High-throughput cultivation of microorganisms using microcapsules. *Methods in Enzymology*. 2005;**397**:124-130. DOI: 10.1016/S0076-6879(05)97007-9

[59] Schmalenberger A, Tebbe CC. Profiling the diversity of microbial communities with single-strand conformation polymorphism (SSCP). In: *Environmental Microbiology*. Totowa, NJ: Humana Press; 2014. pp. 71-83. DOI: 10.1007/978-1-62703-712-9_6

[60] Smalla K, Oros-Sichler M, Milling A, Heuer H, Baumgarte S, Becker R, et al. Bacterial diversity of soils assessed by DGGE, T-RFLP and SSCP fingerprints of PCR-amplified 16S rRNA gene fragments: Do the different methods provide similar results? *Journal of Microbiological Methods*. 2007;**69**(3):470-479. DOI: 10.1016/j.mimet.2007.02.014

[61] Hjelmsø MH, Hansen LH, Bælum J, Feld L, Holben WE, Jacobsen CS. High-resolution melt analysis for rapid comparison of bacterial community compositions. *Applied and Environmental Microbiology*. 2014;**80**(12):3568-3575. DOI: 10.1128/AEM.03923-13

[62] Arabet D, Tempel S, Fons M, Denis Y, Jourlin-Castelli C, Armitano J,

et al. Effects of a sulfonylurea herbicide on the soil bacterial community. *Environmental Science and Pollution Research*. 2014;**21**(8):5619-5627. DOI: 10.1007/s11356-014-2512-9

[63] Yergeau E, Bezemer TM, Hedlund K, Mortimer SR, Kowalchuk GA, Van Der Putten WH. Influences of space, soil, nematodes and plants on microbial community composition of chalk grassland soils. *Environmental Microbiology*. 2010;**12**(8):2096-2106. DOI: 10.1111/j.1462-2920.2009.02053.x

[64] Fierer N, Jackson JA, Vilgalys R, Jackson RB. Assessment of soil microbial community structure by use of taxon-specific quantitative PCR assays. *Applied and Environmental Microbiology*. 2005;**71**(7):4117-4120. DOI: 10.1128/AEM.71.7.4117-4120.2005

[65] Aparna K, Pasha MA, Rao DLN, Krishnaraj PU. Organic amendments as ecosystem engineers: Microbial, biochemical and genomic evidence of soil health improvement in a tropical aridzone field site. *Ecological Engineering*. 2014;**71**:268-277. DOI: 10.1016/j.ecoleng.2014.07.016

[66] DeSantis TZ, Brodie EL, Moberg JP, Zubietta IX, Piceno YM, Andersen GL. High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. *Microbial Ecology*. 2007;**53**(3):371-383. DOI: 10.1007/s00248-006-9134-9

[67] Jones RT, Robeson MS, Lauber CL, Hamady M, Knight R, Fierer N. A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. *The ISME Journal*. 2009;**3**(4):442-453. DOI: 10.1038/ismej.2008.127

[68] Vázquez S, Monien P, Minetti RP, Jürgens J, Curtosi A, Primitz JV, et al.

Bacterial communities and chemical parameters in soils and coastal sediments in response to diesel spills at Carlini Station, Antarctica. *Science of the Total Environment*. 2017;**605**:26-37. DOI: 10.1016/j.scitotenv.2017.06.129

[69] Rosselli R, Romoli O, Vitulo N, Vezzi A, Campanaro S, De Pascale F, et al. Direct 16S rRNA-seq from bacterial communities: A PCR-independent approach to simultaneously assess microbial diversity and functional activity potential of each taxon. *Scientific Reports*. 2016;**6**:32165. DOI: 10.1038/srep32165

[70] Tkacz A, Hortala M, Poole PS. Absolute quantitation of microbiota abundance in environmental samples. *Microbiome*. 2018;**6**(1):110. DOI: 10.1186/s40168-018-0491-7

[71] Schöler A, Jacquiod S, Vestergaard G, Schulz S, Schlöter M. Analysis of soil microbial communities based on amplicon sequencing of marker genes. *Biology and Fertility of Soils*. 2017;**53**(5):485-489. DOI: 10.1007/s00374-017-1205-1

[72] Morgan JL, Darling AE, Eisen JA. Metagenomic sequencing of an in vitro-simulated microbial community. *PLoS One*. 2010;**5**(4):e10209. DOI: 10.1371/journal.pone.0010209

[73] Orellana LH, Chee-Sanford JC, Sanford RA, Löffler FE, Konstantinidis KT. Year-round shotgun metagenomes reveal stable microbial communities in agricultural soils and novel ammonia oxidizers responding to fertilization. *Applied and Environmental Microbiology*. 2018;**84**(2):e01646-e01617. DOI: 10.1128/AEM.01646-17

[74] Luo C, Rodriguez-R LM, Johnston ER, Wu L, Cheng L, Xue K, et al. Soil microbial community responses to a decade of warming as revealed by comparative metagenomics.

Applied and Environmental Microbiology. 2014;**80**(5):1777-1786. DOI: 10.1128/AEM.03712-13

[75] Burns RG. Enzyme activity in soil: Some theoretical and practical considerations. In: *Soil Enzymes*. London: Academic Press; 1978. pp. 295-340

[76] Ndiaye EL, Sandeno JM, McGrath D, Dick RP. Integrative biological indicators for detecting change in soil quality. *American Journal of Alternative Agriculture*. 2000;**15**(1):26-36. DOI: 10.1017/S0889189300008432

[77] Kertesz MA, Mirleau P. The role of soil microbes in plant sulphur nutrition. *Journal of Experimental Botany*. 2004;**55**(404):1939-1945. DOI: 10.1093/jxb/erh176

[78] Speir TW, Ross DJ. Soil phosphatase and sulphatase. In: Burns RG, editor. *Soil Enzymes*. London, UK: Academic; 1978. pp. 197-250

[79] Versaw WK, Harrison MJ. A chloroplast phosphate transporter, PHT2; 1, influences allocation of phosphate within the plant and phosphate-starvation responses. *The Plant Cell*. 2002;**14**(8):1751-1766. DOI: 10.1105/tpc.002220

[80] Nazir A. Review on metagenomics and its applications. *Imperial Journal of Interdisciplinary Research*. 2016;**2**:10

[81] Verma D, Kawarabayasi Y, Miyazaki K, Satyanarayana T. Cloning, expression and characteristics of a novel alkalistable and thermostable xylanase encoding gene (Mxyl) retrieved from compost-soil metagenome. *PLoS One*. 2013;**8**(1):e52459. DOI: 10.1371/journal.pone.0052459

[82] Vidya J, Swaroop S, Singh SK, Alex D, Sukumaran RK, Pandey A. Isolation and characterization of a novel α -amylase from a metagenomic library

of Western Ghats of Kerala, India.
Biologia. 2011;**66**(6):939. DOI: 10.2478/
s11756-011-0126-y

Western Indian Himalayan native red
kidney beans (*Phaseolus vulgaris* L.). 3
Biotech. 2015;**5**:433-441. DOI: 10.1007/
s13205-014-0238-5

[83] Gillespie DE, Brady SF,
Bettermann AD, Cianciotto NP,
Liles MR, Rondon MR, et al. Isolation
of antibiotics turbomycin a and B from
a metagenomic library of soil microbial
DNA. Applied and Environmental
Microbiology. 2002;**68**(9):4301-4306.
DOI: 10.1128/AEM.68.9.4301-4306

[84] Voget S, Leggewie C, Uesbeck A,
Raasch C, Jaeger KE, Streit WR.
Prospecting for novel biocatalysts
in a soil metagenome. Applied
and Environmental Microbiology.
2003;**69**(10):6235-6242. DOI: 10.1128/
AEM.69.10.6235-6242.2003

[85] Ghazanfar S, Azim A,
Ghazanfar MA, Anjum MI, Begum I.
Metagenomics and its application in
soil microbial community studies:
Biotechnological prospects. Journal
of Animal & Plant Sciences.
2010;**6**(2):611-622

[86] Chang F, Brady SF. Discovery of
indolotryptoline antiproliferative agents
by homology-guided metagenomic
screening. Proceedings of the National
Academy of Sciences of the United
States of America. 2013;**110**(7):2478-
2483. DOI: 10.1073/pnas.1218073110

[87] Su JQ, Wei B, Xu CY, Qiao M,
Zhu YG. Functional metagenomic
characterization of antibiotic resistance
genes in agricultural soils from China.
Environment International. 2014;**65**:9-
15. DOI: 10.1016/j.envint.2013.12.010

[88] Sachdev DP, Cameotra SS.
Biosurfactants in agriculture. Applied
Microbiology and Biotechnology.
2013;**97**(3):1005-1016. DOI: 10.1007/
s00253-012-4641-8

[89] Suyal DC, Yadav A, Shouche Y,
Goel R. Diversified diazotrophs
associated with the rhizosphere of