Open access Journal International Journal of Emerging Trends in Science and Technology
DOI: https://dx.doi.org/10.18535/ijetst/v8i8.02

Molecular Techniques to Assess Microbial Community Structure, Function, and Dynamics in the Environment

Author
Abdulbasit Hussein[1]*
Department of Natural Resource Management, Haramaya University College of Agriculture and Environmental Science (HU CAES), P. O Box 138, Dire Dawa, Ethiopia

Abstract
Culture-based methods are important in studying the microbial ecology of natural and man-made environments, but they are extremely biased in assessing microbial genetic diversity by selecting a particular population of microorganisms. With the recent advances in genomics and sequencing technologies, microbial community analyzes using culture-independent molecular techniques have ushered in a new era in microbial ecology. Molecular analyzes of environmental communities have shown that the cultivable fraction represents <1% of the total number of prokaryotic species in a given sample. Molecular approaches such as genetic fingerprinting, metagenomics, metaproteomics, metatranscriptomics, and proteogenomics are critical to discovering and characterizing the enormous microbial diversity and understanding its interactions with biotic and abiotic environmental factors. This Review summarizes recent advances in molecular microbial ecology with an emphasis on novel techniques and approaches that provide new insights into the phylogenetic and functional diversity of microbial aggregates. The potential applications of the individual molecular techniques and their combination for a more comprehensive assessment of microbial diversity were illustrated using example studies.
Keywords: *DNA, RNA, microbial, function, diversity.*

Introduction

The biosphere is dominated by microorganisms and contains around 461030 prokaryotic cells (Whitman et al. 1998). This number represents at least two to three orders of magnitude more than all plant and animal cells combined. Thus, microorganisms are a very diverse group of organisms and make up about 60% of the earth's biomass (Singh et al. 2009). Microorganisms are essential components of the earth's biota and represent a large unexplored reservoir of genetic diversity.

Microorganisms are key players in important ecological processes such as soil structure formation, the breakdown of organic matter and xenobiotics as well as the recycling of essential elements (e.g. carbon, nitrogen, phosphorus and sulfur) and nutrients. Thus, microbes play a crucial role in modulating global biogeochemical cycles and influencing all life on earth (Garbeva et al. 2004). In soil ecosystems, microorganisms play a central role in suppressing soil-borne plant diseases, promoting plant growth and promoting vegetation changes (Garbeva et al. 2004). An understanding of microbial dynamics and their interactions with biotic and abiotic factors is essential in bioremediation techniques, energy generation processes and in biotechnological industries such as pharmaceuticals, food, chemistry and mining.

A variety of biochemical and molecular methods have been used to uncover the makeup of the microbial community over time and

space in response to environmental changes. These new approaches make it possible to link ecological processes in the environment with specific microbial populations. Norman Pace and his group established the concept of using molecular markers (e.g. ribosomal RNAs) for the phylogenetic characterization of microorganisms (Lane et al., 1985; Olsen et al., 1986). Karry Mullis, on the other hand, made his Nobel Prize-winning invention of the polymerase chain reaction (PCR), which enabled scientists to amplify DNA in vitro (Saiki et al., 1988). As a result of these results, along with other protocols such as the extraction of total nucleic acids from different environments (Tsai and Olson, 1991; Griffiths et al., 2000), microbial ecologists were equipped with a range of culture-independent methods of studying the diversity and dynamics of microbial communities in a much larger resolution.

2. Molecular Methods of Microbial Community Analyses

The vast majority of microbial communities in nature have not been grown in the laboratory. Hence, the primary sources of information for these uncultivated but viable organisms are their biomolecules such as nucleic acids, lipids, and proteins. Culture-independent nucleic acid approaches include analyzes of entire genomes or selected genes such as 16S and 18S rRNA (ribosomal RNA) for prokaryotes or eukaryotes. In the last few decades, microbial ecology has made enormous strides and developed a multitude of molecular techniques for describing and characterizing the phylogenetic and functional diversity of microorganisms. By and large, these techniques have been grouped into two main categories based on their ability to reveal the structure and function of microbial diversity: (1) approaches to partial community analysis and (2) approaches to whole community analysis.

2.1 Partial Community Analysis Approaches

These strategies generally include polymerase chain reaction (PCR) based methods in which total DNA / RNA extracted from an environmental sample is used as a template for the characterization of microorganisms. In principle, the PCR product generated in this way reflects a mixture of microbial gene signatures of all organisms present in a sample, including the VBNC fraction. The PCR amplification of conserved genes such as 16S rRNA from an environmental sample has been used extensively in microbial ecology primarily because these genes (1) are ubiquitous, i.e. present in all prokaryotes, (2) are structurally and functionally conserved and (3) contain variable and highly conserved regions (Hugenholtz 2002).

In addition, the appropriate gene size (~ 1,500 bp) and the growing number of 16S rRNA sequences available for comparison in sequence databases make it a gold standard in microbial ecology. By estimating the phylogenetic relationship to known microorganisms based on the homology of 16S rRNA sequences, the closest association to a new isolate or a new molecular sequence is assigned. Other conserved genes such as the beta subunit of RNA polymerase (rpoB), the beta subunit of gyrase (gyrB), recombinase A (recA) and heat shock protein (hsp60) have also been used in microbial studies and to differentiate some types of bacteria (Ghebremedhin .) used et al. 2008). The PCR products amplified from environmental DNA are analyzed primarily by (1) clone library techniques, (2) genetic fingerprints, (3) DNA microarrays, or a combination of these techniques.

2.1.1 Clone Library Method

The most widely used method for analyzing PCR products amplified from an environmental sample is cloning and subsequent sequencing of the individual gene fragments (DeSantis et al. 2007). The sequences obtained are compared

with known sequences in a database such as GenBank, Ribosomal Database Project (RDP) and Greengenes. The RDP currently comprises more than 2.7 million sequences (as of May 2013) and is updated regularly. The sequencing of clone libraries from environmental samples has resulted in a wealth of information on prokaryotic diversity (Singleton et al., 2001). While clone libraries of 16S rRNA markers allow an initial overview of the diversity and identification of new taxa, a large number of clones must be sequenced in order to detect rare organisms against the background of less dominant species (Tiedje et al., 1999).

2.1.2 Genetic Fingerprinting Techniques

The genetic fingerprint creates a profile of microbial communities based on the direct analysis of PCR products amplified from environmental DNA (Muyzer 1999). These techniques include DGGE / TTGE, SSCP, RAPD, ARDRA, T-RFLP, LH-PCR, RISA and RAPD and generate a community fingerprint based on either sequence polymorphism or length polymorphism. In general, genetic fingerprinting techniques are fast and allow multiple samples to be analyzed simultaneously. Fingerprinting approaches were developed to show an effect on microbial communities or differences between microbial communities and do not offer direct taxonomic identities. The fingerprints of different samples are compared using computer-aided cluster analysis using software packages such as GelCompar, and community relationships are inferred. Community fingerprints are scored as present or absent and the similarities between samples are determined using the Jaccard coefficient.

2.1.2.1. Denaturing- or Temperature-Gradient Gel Electrophoresis

In denaturing gradient gel electrophoresis (DGGE), the PCR products are obtained from environmental DNA using primers for a specific molecular marker (e.g. 16S rRNA gene) and placed on a polyacrylamide gel that has a linear gradient of the DNA Contains denaturant, like a mixture of urea, electrophoresed and formamide (Muyzer et al. 1993). Temperature gradient gel electrophoresis (TTGE) is based on the same principle of DGGE, except that a temperature gradient is used instead of a chemical denaturant. To determine the phylogenetic identities from DGGE / TGGE fingerprints, the bands can be cut out of the gel, amplified again and sequenced or blotted on nylon membranes and hybridized with molecular probes that are specific for different taxonomic groups. DGGE analysis was used to screen the unique clones in clone libraries based on different patterns and to determine the number of functional taxonomic units.

2.1.2.2. Single-Strand Conformation Polymorphism

In single-strand conformational polymorphism (SSCP), the environmental PCR products are denatured, followed by electrophoretic separation of single-stranded DNA fragments on a non-denaturing polyacrylamide gel (Schwieger and Tebbe 1998). The separation is based on subtle differences in the sequences (often a single base pair), resulting in a differently folded secondary structure, resulting in a measurable difference in mobility in the gel. Similar to DGGE, the DNA bands can be cut out of the gel, amplified again and sequenced. However, SSCP is only suitable for small fragments (between 150 and 400 bp) (Muyzer 1999). SSCP was successfully used to differentiate pure cultures of Bacillus subtilis, Pseudomonas fluorescens and Sinorhizobium meliloti isolated from soil samples (Schwieger and Tebbe 1998).

2.1.2.3. Random Amplified Polymorphic DNA and DNA Amplification Fingerprinting

Randomly amplified polymorphic DNA (RAPD) and DNA amplification fingerprinting (DAF) techniques use PCR amplification with a short (usually ten nucleotide) primer that is randomly applied to multiple locations on the genomic DNA at a low annealing temperature, typically 35 ° C. (Franklin et al. 1999). These

methods generate PCR amplicons of different lengths in a single reaction, which are separated on agarose or polyacrylamide gel depending on the genetic complexity of the microbial communities. Because of its high speed and ease of use, RAPD / DAF has been used extensively in fingerprinting the entire microbial community structure and closely related bacterial species and strains (Franklin et al. 1999). Therefore, multiple primers and reaction conditions need to be evaluated in order to compare the relationships between microbial communities and to obtain the most diverse patterns between species or strains.

2.1.2.4. Amplified Ribosomal DNA Restriction Analysis

The amplified ribosomal DNA restriction analysis (ARDRA) is based on DNA sequence variations that are present in PCR-amplified 16S rRNA genes (Smit et al. 1997). The PCR product amplified from environmental DNA is generally digested with tetracutter restriction endonucleases and restricted fragments are separated on agarose or polyacrylamide gels. Although ARDRA provides little or no information about the types of microorganisms present in the sample, it is still useful for quickly monitoring microbial communities over time or for comparing microbial diversity in response to changing environmental conditions. ARDRA is also used to identify the unique clones and estimate OTUs in environmental clone libraries based on restriction profiles of clones (Smit et al. 1997).

2.1.2.5. Ribosomal Intergenic Spacer Analysis

Ribosomal intergenic spacer analysis (RISA) involves the PCR amplification of part of the intergenic spacer region (ISR) that is present between the small (16S) and large (23S) ribosomal subunits (Fisher and Triplett 1999). The ISR contains significant heterogeneity in both length and nucleotide sequence. By using primers that anneal to conserved regions in the 16S and 23S rRNA genes, RISA profiles can be generated from most of the dominant bacteria present in an environmental sample.

RISA provides a community-specific profile, with each group corresponding to at least one organism in the original community. The automated version of RISA is known as ARISA and involves the use of a fluorescently labeled forward primer and ISR fragments are automatically detected by a laser detector. ARISA enables the simultaneous analysis of many samples; However, technology has been shown to overestimate microbial richness and diversity.

2.1.3 DNA Microarrays

The PCR products amplified from the entire environmental DNA are hybridized directly to known molecular probes that are attached to the microarrays (Gentry et al. 2006). After the fluorescently labeled PCR amplicons hybridize to the probes, positive signals are assessed using confocal laser scanning microscopy. The microarray technique enables a rapid evaluation of samples with replication, which is a significant advantage in the analysis of microbial communities. In general, the hybridization signal intensity on microarrays is directly proportional to the abundance of the target organism. Cross-hybridization is a major limitation of microarray technology, especially when dealing with environmental samples. In addition, the microarray is not useful for identifying and detecting new prokaryotic taxa. The ecological importance of a genus could be completely neglected if the genus does not have a corresponding probe on the microarray. DNA microarrays used in microbial ecology can be divided into two main categories depending on the probes: (1) 16S rRNA gene microarrays and (2) functional gene arrays (FGA).

2.1.3.1. 16S rRNA gene Microarrays (PhyloChip)

The high-density, universal 16S microarray contains approximately 30,000 probes of the 16S rRNA gene targeting multiple cultured microbial species and candidate divisions

(DeSantis et al. 2007). These probes target all 121 delimited prokaryotic orders and enable the simultaneous detection of 8,741 bacterial and archaeal taxa. The PhyloChip analysis showed greater diversity than corresponding clone libraries at every taxonomic level and showed the existence of 1,3001,700 bacterial species in soil samples from uranium mines. Most of these species were members of the Proteobacteria strain and contained lines capable of uranium immobilization and metal reduction.

2.1.3.2. Functional Gene Arrays

FGA were primarily developed to detect certain metabolic groups of bacteria. Thus, FGA not only reveal the community structure, but also shed light on the in situ metabolic potential of the community. FGA contain probes from genes with known biological functions; Hence, they are also useful for linking the composition of the microbial community to ecosystem functions.

2.1.4 Quantitative PCR

Quantitative PCR (Q-PCR) or real-time PCR has been used in microbial studies to measure the frequency and expression of taxonomic and functional gene markers (Bustin et al. 2005; Smith and Osborn 2009). Q-PCR uses either intercalating fluorescent dyes such as SYBR Green or fluorescent probes (TaqMan) to measure the accumulation of amplicons in real time during each PCR cycle. The software records the increase in amplicon concentration during the early exponential phase of amplification, which enables genes (or transcripts) to be quantified when they are proportional to the starting template concentration. If Q-PCR is coupled with a previous reverse transcription reaction (RT), it can be used to quantify gene expression (RT-Q-PCR). Q-PCR is highly sensitive to the initial template concentration and measures the template frequency over a large dynamic range of about six orders of magnitude. Several sets of 16S and 5.8S rRNA gene primers have been developed for the rapid Q-PCR-based quantification of bacterial and fungal microbial communities in the soil (Fierer et al. 2005). Q-PCR has also been used successfully in environmental samples for the quantitative detection of important physiological bacterial groups such as ammonia oxidizers, methane oxidizers and sulfate reducers by targeted targeting of the amoA, pmoA and dsrA genes (Foti et al. 2007).

2.1.5 Fluorescence in Situ Hybridization

Fluorescence in situ hybridization (FISH) enables phylogenetic in situ identification and enumeration of individual microbial cells by whole cell hybridization with oligonucleotide probes (Amann et al. 1995). A large number of molecular probes targeting 16S rRNA genes have been reported at various taxonomic levels (Amann et al. 1995). The FISH probes are generally 1830 nucleotides long and contain a fluorescent dye at the 5 end that enables the detection of the probe bound to cellular rRNA by epifluorescence microscopy. In addition, the intensity of the fluorescence signals correlates with cellular rRNA levels and growth rates, which provide information about the metabolic status of the cells. FISH can be combined with flow cytometry for high resolution automated analysis of mixed microbial populations. The FISH method has been used to follow the dynamics of bacterial populations in agricultural soils that have been treated with S-triazine herbicides (Caracciolo et al. 2010).

Low signal intensity, background fluorescence, and inaccessibility of the target are common problems in FISH analysis. Liet al. (2008) developed an advanced imaging technique by combining FISH with secondary ion mass spectrometry (SIMS). In principle, the technique uses 16S rRNA probes for in situ hybridization; however, the probes are labeled with a stable isotope or element (e.g. fluorine or bromine atoms) that is rarely found in biomass. Once the probe has hybridized, the microbial identities of stable isotope-labeled cells are

simultaneously determined in situ by Nano SIMS imaging.

2.1.6 Microbial Lipid Analysis

The characterization of the microbial community by biomolecules other than nucleic acids such as lipids has been used without relying on cultivation (Banowetz et al. 2006). Fatty acids are present in a relatively constant proportion of cellular biomass, and distinctive fatty acids exist in microbial cells that can differentiate important taxonomic groups within a community. The fatty acids are extracted to the respective FAMEs by saponification and subsequent derivatization, which are then analyzed by gas chromatography. The resulting pattern is then compared to a reference FAME database to identify the fatty acids and their corresponding microbial signatures through multivariate statistical analysis.

2.2 Whole Community Analysis Approaches

Sequence analysis of 16S rRNA genes is commonly used in most microbial ecological studies. However, since it is a highly conserved molecule, the 16S rRNA gene does not provide sufficient resolution at the species and strain level (Rastogi and R. K. Sani et al. 2006). Whole genome molecular techniques provide a broader view of genetic diversity compared to PCR-based molecular approaches that target a single gene or a few. These techniques attempt to analyze all of the genetic information present in total DNA extracted from an environmental sample or pure culture.

2.2.1 DNA–DNA Hybridization Kinetics

Whole Genome DNADNA Hybridization (DDH) provides a true genome-wide comparison between organisms. A value of 70% DDH has been suggested as the recommended standard for the differentiation of bacterial species (Goris et al. 2007). Typically bacterial species with 70% or more genomic DNA similarities. Although DDH techniques were originally developed for pure culture comparisons, they have been modified for use in analyzing entire microbial communities. In the DDH technique, the total community DNA extracted from an environmental sample is denatured and then incubated under conditions that allow them to hybridize or reassociate. The rate of DNA reassociation correlates with the genomic complexity (diversity) present in the sample. If the sample has high sequence diversity, the rate of DNA reassociation will decrease. Under defined conditions, the time required for half of the DNA to reassociate is proportional to the genomic diversity and can be used as a diversity index.

2.2.2 Guanine plus Cytosine Content Fractionation

Different prokaryotic groups differ in their guanine plus cytosine (G + C) content of the DNA, and phylogenetically related bacterial groups differ only by 35% in their G + C content (Nsslein and Tiedje 1999). Thus, fractionation of all community DNA can be achieved by density gradient centrifugation based on the G + C content. The technique creates a fractional profile of all community DNA and displays the relative abundance of DNA (hence taxa) as a function of G + C content. All of the community DNA is physically separated into highly purified fractions, each of which represents a different G + C content that can be analyzed by additional molecular techniques such as DGGE / ARDRA to better assess overall community diversity. G + C fractionation has been widely used in studying microbial communities in soil to assess the effect of various treatments or management practices (e.g., changing vegetation, grazing, applying pesticides, and applying compost).

2.2.3 Whole Microbial Genome Sequencing

Exploring microbial systems through whole genome analysis is a comprehensive and integrated approach to understanding microbial ecology and function. Whole microbial genomes are sequenced using a shotgun cloning method that involves (1) extraction of DNA from pure cultures, (2) random fragmentation of

the obtained genomic DNA into small fragments, (3) ligation and cloning of DNA fragments into plasmid vectors, and (4) bidirectional sequencing of DNA fragments. As soon as the sequences have been obtained, they are aligned and assembled into finished sequences using specialized computer programs such as MEGAN (MEtaGenome ANalyzer) (Huson et al. 2007). Sequencing the entire genome has provided unprecedented insights into microbial processes at the molecular level and has potential applications in individual and community ecology, bioenergy production, biological remediation, human and plant health and in various industries (Ikeda et al. 2003).

2.3. Application of "-omics" Technologies

The -omics-based technologies are equipped with novel sequencing techniques, which are often referred to as next generation sequencing (NGS) or deep sequencing. Compared to traditional Sanger sequencing, these newer sequencing formats allow a much faster production of gibases from sequence information in a few days. A number of NGS platforms, e.g. B. 454 Pyrosequencing (Roche), MiSeq, Solexa (Illumina) and Ion Torrent are widely used by biologists (Mardis, 2013).

2.3.1. Metagenomics

Metagenomics is the study of collective microbial genomes that are obtained directly from environmental samples and do not rely on cultivation or prior knowledge of the microbial communities (Riesenfeld et al. 2004). Metagenomics is also known by other names such as environmental genomics or community genomics or microbial ecogenomics. Essentially, metagenomics does not include methods that only interrogate selected genes amplified by PCR (e.g. genetic fingerprinting techniques), as they do not provide any information about the genetic diversity beyond the genes to be amplified. Metagenomic studies have been carried out in different environments such as soil, phyllosphere, ocean and acid mine runoff and have given access to the phylogenetic and functional diversity of uncultivated microorganisms (Handelsman 2004). Therefore, metagenomics is critical to understanding the biochemical role of uncultivated microorganisms and their interaction with other biotic and abiotic factors.

The construction of a metagenomic library comprises the following steps: (1) isolation of total DNA from an environmental sample, (2) shotgun cloning of random DNA fragments into a suitable vector and (3) transformation of the clones into a host bacterium and screening for positive ones Clones. Metagenomic libraries with small DNA fragments in the range of 23 kb provide better coverage of the metagenome of an environment than those with larger fragments. Small insert DNA libraries are also useful to screen for phenotypes encoded by individual genes and to reconstruct the metagenomes for genotypic analysis. Large fragment metagenomic libraries (100,200 kb) are desirable when exploring multi-gene biochemical pathways. Metagenomic libraries could be screened either by sequential metagenomic analysis, which involves massive, high throughput sequencing, or by functional screening of expressed phenotypes. Sequence-driven massive metagenomic sequencing of the entire genome sheds light on many important genomic features such as redundancy of functions in a community, genomic organizations, and traits acquired from uniquely related taxa through horizontal gene transfer (Handelsman 2004).

Function-driven metagenomic analysis (functional metagenomics) screens libraries based on the expression of a selected phenotype on a specific medium. A wide variety of biochemical activities have been discovered in environmental metagenomic libraries. Function-driven metagenomic approaches require the successful expression of a gene of interest in a heterologous host such as E. coli. Thus, a major limitation is that most of the environmental genes in E. coli are expressed very little or not

at all. In some cases, improved gene expression can be achieved by transforming metagenomic DNA into several additional surrogate hosts such as Streptomyces, Bacillus, Pseudomonas, and Agrobacterium. Strategies that can enhance the heterologous expression of unknown genes in host cells are highly desirable. In a metagenomic library, the frequency of active gene clones expressing a phenotype is typically very low.

2.3.2. Metaproteomics

Metaproteomics, also known as environmental proteomics, deals with the large-scale study of proteins that are expressed by microbial communities in the environment at a given point in time (Wilmes and Bond 2006; Keller and Hettich 2009). Protein biomarkers are more reliable than other cell molecules such as lipids and nucleic acids and provide a clearer picture of metabolic functions than functional genes or even the corresponding mRNA transcripts of microbial communities (Wilmes and Bond 2006). Although methods such as SIP / MARFISH were developed for structural function analysis of microbial communities, these methods only provide information about microbial communities that are associated with a particular biogeochemical process (e.g., compared to these methods, proteomics offers a comprehensive approach to qualitatively and For example, proteomic profiling of microbial communities provides critical information about protein abundance and protein interactions that could not be achieved with molecular DNA / RNA techniques such as metatranscriptomics and metagenomics (Keller and Hettich 2009). The physiological reactions of microbial communities due to a stressful state could not be achieved can be identified using an altered proteofingerprint that reflects changes in the functional status of the communities Such metagenomic sequences can be linked in order to link metabolic functions with individual microbial species.

The metaproteome analysis involves the extraction of total proteins from an environmental sample. Although in-situ protein analysis methods allow exhaustive extraction, a significant part of the protein comes from other organisms such as protozoa, fungi and multicellular organisms, which makes the taxonomic characterization of proteins even more difficult (Keller and Hettich 2009). Therefore, in some cases, microbial cells are first separated from the surrounding matrix by ultracentrifugation and then analyzed, whereby a much higher quality and quantity of bacterial proteins can be obtained. Once the total protein is obtained, it is separated by one-dimensional and two-dimensional electrophoresis to create a common proteofingerprint. After separation, protein spots are digested and identified using a variety of powerful analytical methods. Due to the development of chromatographic and mass spectroscopic techniques (MS-based proteomics), proteomic profiling of microbial communities with high throughput is currently possible. The highly efficient mass spectrometry with integrated liquid chromatography enables a highly sensitive and fast identification of proteins. A metaproteomic approach was used to identify proteins involved in the biodegradation of chlorophenoxy acid in soil samples (Benndorf et al. 2007). Soil samples were first enriched for chlorophenoxy acid degrading bacteria by incubation with 2, 4-D for a period of 22 days. After incubation, protein extracts were isolated from the soil and separated by SDS-PAGE and protein bands were identified by liquid chromatography in conjunction with mass spectroscopy.

2.3.3. Metatranscriptomics

Metatranscriptomics (or environmental transcriptomics) enables the monitoring of microbial gene expression profiles in natural environments by studying the global transcription of genes by randomly sequencing mRNA transcripts pooled from microbial communities at a specific time and location

(Moran 2009). Metatranscriptomics is particularly suitable for measuring changes in gene expression and its regulation in relation to changing environmental conditions. A method for selective enrichment of mRNA by subtractive hybridization of rRNA was developed and evaluated for gene transcript analysis of marine and freshwater bacterioplankton communities, which showed the presence of many transcripts associated with biogeochemical processes such as sulfur oxidation (soxA), assimilation of C1 compounds (fdh1B) and nitrogen uptake through polyamine degradation (aphA) (Poretsky et al. 2005). More recently, a double RNA method has been developed to analyze the entire RNA pool of a community, as it naturally contains not only functionally but also taxonomically relevant molecules, i.e. mRNA or rRNA (Urich et al. 2008). . This offers the possibility to study both the structural and the biochemical activity of microbes in a single experiment. The rRNA tags provided data on the phylogenetic composition of microbial communities in the soil and showed that actinobacteria and proteobacteria were the most common, while Crenarchaeota were less common in soil samples.

2.4. linking of Microbial Community Structure and Function

Understanding how microbial communities function in natural environments is a central goal of microbial ecology. RNA extracted from environmental samples provides more valuable information than DNA to reveal active microbial communities versus dormant microbial communities (Torsvik, 2002). This is because rRNA and mRNA are considered indicators of functionally active microbial populations. The amount of rRNA in a cell is roughly correlated with the growth activity of bacteria, and mRNA from functional genes enables the detection and identification of bacteria that, under certain conditions, actually express key enzyme activities (Wellington et al. 2003). Microbial catabolic diversity could also be studied through enzyme-encoding genes that are involved in the utilization of specific carbon substrates such as chitin, cellulose and lipids (Torsvik and vres 2002). The diversity of lipase-producing microorganisms in the glacier floor was investigated by PCR amplification of lipase genes, and sequence analyzes showed the existence of several new lipase-producing organisms in the soil (Yuhong et al. 2009). More advanced methods that use stable isotopes, such as Stable Isotope Sonding (SIP), Microautoradiography FISH (MARFISH), and Raman FISH, provide more detailed insights into the metabolic activities of microbial communities and are discussed in the following sections.

2.5.1 Stable Isotope Probing

SIP includes the offer of a stable isotope-labeled substrate for microbial communities, the use of which is of interest to decipher a key biogeochemical process (Wellington et al. 2003). Active microbial communities that use the labeled substrate during growth incorporate the isotopes into their biomass. The labeled biomolecules (e.g. DNA, RNA, phospholipid fatty acids [PLFA]) are then separated from the biomass using various biochemical methods and the phylogenetic identity of the microorganisms that metabolize the substrate is determined using molecular techniques. SIP, based on DNA biomarkers, involves labeling DNA with 13C, which could be separated from 12C by CsCl equilibrium density gradient centrifugation. The 13C-labeled DNA could be analyzed by genetic fingerprinting or clone library techniques, resulting in the identification of microorganisms.

In recent years, with advances in imaging and spectroscopic techniques, SIP has been combined with other techniques such as FISH and Raman microscopy to simultaneously study the taxonomic identity and activity of microbial communities upon single cell dissolution (Huang et al. 2007). In the RamanFISH method,

environmental samples are incubated with a substrate that is labeled with the stable 13C isotope. After the incorporation, the spectral profiles of uncultivated microbial cells are generated with single-cell resolution using Raman microscopy, which measures the laser light that is scattered by chemical bonds between various cell biomarkers. Huang et al. (2007) used the RamanFISH method to study naphthalene-degrading Pseudomonas communities in groundwater. Their results, based on differences in the 13C content of the different Pseudomonad cells, suggested that different Pseudomonas species and even members of the same species vary in their ability to degrade naphthalene.

2.5.2 Micro autoradiography

Microautoradiography (MAR) is based on the fact that metabolically active cells that use radioactively labeled substrate can be visualized by exposure to radiation-sensitive silver halide emulsions (Okabe et al. 2004). The emulsion is applied to cells mounted on a microscope slide. After exposure, excited silver ions precipitate as black grains of metallic silver inside or next to the cells, which can be observed by transmission electron microscopy. Commonly used radiolabelled substrates include glucose, acetate, and amino acids, which provide a general overview of all metabolic diversity. More specific substrates along with selective growth (incubation) conditions were used to identify important physiological processes in situ. For example, radioactively labeled iron or sulfate can be provided under controlled anaerobic conditions in order to identify the iron- or sulfate-reducing microbial communities. When MAR is used in combination with FISH (MARFISH), it enables simultaneous phylogenetic identification of active cells that are consuming the radioactive substrate (Rogers et al. 2007).

MARFISH was slightly modified, resulting in other methods such as STAR (Substrate Tracking Autoradiography) FISH. However, STAR FISH only differs from MARFISH in methodological details, and the basic principle of the technique remains the same. Nielsen et al. (2003) developed a quantitative MAR (QMAR) FISH approach that can even detect individual cells due to its improved fixation protocol and the use of an internal standard of bacteria with known specific radioactivity. The MARFISH technique was used to study the autotrophic nitrifying bacteria in biofilms (Okabe et al. 2005). The uptake of 14C-labeled products from nitrifying bacteria by heterotrophic bacteria was visualized directly by MARFISH. The results showed that members of Chloroflexi and Cytophaga Flavobacterium play an important role in capturing the dead biomass and metabolites of nitrifying bacteria and ultimately preventing the accumulation of organic waste products in the biofilms.

2.5.3 Isotope Array

Isotope arrays enable functional and phylogenetic screening of active microbial communities with high throughput. The technique uses a combination of SIP to monitor substrate uptake profiles and microarray technology to decipher the taxonomic identities of active microbial communities (Adamczyk et al. 2003). In principle, samples are incubated with a 14C-labeled substrate, which is incorporated into microbial biomass in the course of growth. The 14C-labeled rRNA is separated from unlabeled rRNA and then labeled with fluorochromes. Fluorescently labeled rRNA is hybridized to a phylogenetic microarray, followed by scanning for radioactive and fluorescent signals. The technology thus enables the parallel investigation of the composition of the microbial community and the specific substrate consumption by metabolically active microorganisms of complex microbial communities. The great strengths of the technology are that it does not contain an amplification step and is therefore free from biases associated with PCR. The limitations of

the technology include the difficulties of obtaining high quality rRNA and the detection of low abundance but active microbial populations from environmental samples (Adamczyk et al. 2003).

Conclusion

Post-genomic molecular approaches allow us to question the structural and functional diversity of microbial communities in the environment, showing that we have only scratched the surface of the genetic and metabolic diversity present in the most abundant organisms on earth, prokaryotes is. Understanding the functional roles of uncultivated organisms remains a daunting task, as most of the genes identified do not have homologous representatives in databases. Although significant advances have been made in characterizing microbial communities using metagenomic, metatranscriptomic, and proteogenomic approaches, many technical challenges remain, including DNA, RNA, and protein extraction from environmental samples, mRNA instability, and low abundance of certain gene transcripts in total RNA. Next generation sequencing techniques are still in development and many technological innovations specifically tailored to environmental samples are expected in these techniques.

The development of bioinformatic tools is also required to assess the tremendous amount of information generated by whole genome analysis and metagenomic and metatranscriptomic approaches. The quantitative assessment of microbial communities is the most challenging because of the significant biases associated with nucleic acid isolation and PCR, and requires more advanced DNA / RNA extraction techniques for environmental samples. All of the molecular approaches available for structural and functional analysis of communities have advantages and limitations, and none provide full access to the genetic and functional diversity of complex microbial communities. A combination of several techniques should be used to study the diversity, function and ecology of microorganisms. Culture-based and culture-independent molecular techniques are neither contradicting nor exclusive and should be viewed as complementary. An interdisciplinary systems approach that includes multiple omics technologies to uncover the interactions between genes, proteins and environmental factors is needed to gain new insights into environmental microbiology. The development of multi-omics approaches will be a focus of research in the coming years.

References

1. Adamczyk, J., Hesselsoe, M., Iversen, N., Horn, M., Lehner, A., Nielsen, P.H., Schloter, M., Roslev, P. and Wagner, M., 2003. The isotope array, a new tool that employs substrate-mediated labeling of rRNA for determination of microbial community structure and function. *Applied and environmental microbiology*, *69*(11), pp.6875-6887.
2. Amann, R.I., Ludwig, W. and Schleifer, K.H., 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological reviews*, *59*(1), pp.143-169.
3. Banowetz, G.M., Whittaker, G.W., Dierksen, K.P., Azevedo, M.D., Kennedy, A.C., Griffith, S.M. and Steiner, J.J., 2006. Fatty acid methyl ester analysis to identify sources of soil in surface water. *Journal of environmental quality*, *35*(1), pp.133-140.
4. Benndorf, D., Balcke, G.U., Harms, H. and Von Bergen, M., 2007. Functional metaproteome analysis of protein extracts from contaminated soil and groundwater. *The ISME journal*, *1*(3), pp.224-234.
5. Brodie, E.L., DeSantis, T.Z., Parker, J.P.M., Zubietta, I.X., Piceno, Y.M. and Andersen, G.L., 2007. Urban aerosols harbor diverse and dynamic bacterial populations. *Proceedings of*

the National Academy of Sciences, *104*(1), pp.299-304.

6. Bustin, S.A., Benes, V., Nolan, T. and Pfaffl, M.W., 2005. Quantitative real-time RT-PCR–a perspective. *Journal of molecular endocrinology*, *34*(3), pp.597-601.
7. Caracciolo, A.B., Bottoni, P. and Grenni, P., 2010. Fluorescence in situ hybridization in soil and water ecosystems: a useful method for studying the effect of xenobiotics on bacterial community structure. *Toxicological & Environmental Chemistry*, *92*(3), pp.567-579.
8. Fierer, N. and Jackson, R.B., 2006. The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences*, *103*(3), pp.626-631.
9. Fierer, N., Jackson, J.A., Vilgalys, R. and Jackson, R.B., 2005. Assessment of soil microbial community structure by use of taxon-specific quantitative PCR assays. *Applied and environmental microbiology*, *71*(7), pp.4117-4120.
10. Fisher, M.M. and Triplett, E.W., 1999. Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. *Applied and environmental microbiology*, *65*(10), pp.4630-4636.
11. Foti, M., Sorokin, D.Y., Lomans, B., Mussman, M., Zacharova, E.E., Pimenov, N.V., Kuenen, J.G. and Muyzer, G., 2007. Diversity, activity, and abundance of sulfate-reducing bacteria in saline and hypersaline soda lakes. *Applied and environmental microbiology*, *73*(7), pp.2093-2100.
12. Franklin, R.B., Taylor, D.R. and Mills, A.L., 1999. Characterization of microbial communities using randomly amplified polymorphic DNA (RAPD). *Journal of Microbiological Methods*, *35*(3), pp.225-235.
13. Garbeva, P.V., Van Veen, J.A. and Van Elsas, J.D., 2004. Microbial diversity in soil: selection of microbial populations by plant and soil type and implications for disease suppressiveness. *Annu. Rev. Phytopathol.*, *42*, pp.243-270.
14. Gentry, T.J., Wickham, G.S., Schadt, C.W., He, Z. and Zhou, J., 2006. Microarray applications in microbial ecology research. *Microbial ecology*, *52*(2), pp.159-175.
15. Ghebremedhin, B., Layer, F., Konig, W. and Konig, B., 2008. Genetic classification and distinguishing of Staphylococcus species based on different partial gap, 16S rRNA, hsp60, rpoB, sodA, and tuf gene sequences. *Journal of clinical microbiology*, *46*(3), pp.1019-1025.
16. Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P. and Tiedje, J.M., 2007. DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *International journal of systematic and evolutionary microbiology*, *57*(1), pp.81-91.
17. Griffiths, R.I., Whiteley, A.S., O'Donnell, A.G. and Bailey, M.J., 2000. Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA-and rRNA-based microbial community composition. *Applied and environmental microbiology*, *66*(12), pp.5488-5491.
18. Huang, W.E., Stoecker, K., Griffiths, R., Newbold, L., Daims, H., Whiteley, A.S. and Wagner, M., 2007. Raman- FISH: combining stable- isotope Raman spectroscopy and fluorescence in situ hybridization for the single cell analysis of identity and function. *Environmental microbiology*, *9*(8), pp.1878-1889.
19. Hugenholtz, P., 2002. Exploring prokaryotic diversity in the genomic era. *Genome biology*, *3*(2), pp.1-8.
20. Huson, D.H., Auch, A.F., Qi, J. and Schuster, S.C., 2007. MEGAN analysis of metagenomic data. *Genome research*, *17*(3), pp.377-386.
21. Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M. and Ōmura, S., 2003.

Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis. *Nature biotechnology*, *21*(5), pp.526-531.

22. Keller, M. and Hettich, R., 2009. Environmental proteomics: a paradigm shift in characterizing microbial activities at the molecular level. *Microbiology and molecular biology reviews*, *73*(1), pp.62-70.
23. Lane, D.J., Pace, B., Olsen, G.J., Stahl, D.A., Sogin, M.L. and Pace, N.R., 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proceedings of the National Academy of Sciences*, *82*(20), pp.6955-6959.
24. Li, T., Wu, T.D., Mazéas, L., Toffin, L., Guerquin- Kern, J.L., Leblon, G. and Bouchez, T., 2008. Simultaneous analysis of microbial identity and function using NanoSIMS. *Environmental microbiology*, *10*(3), pp.580-588.
25. Mardis, E.R., 2013. Next-generation sequencing platforms. *Annual review of analytical chemistry*, *6*, pp.287-303.
26. Markowitz, V.M., Chen, I.M.A., Palaniappan, K., Chu, K., Szeto, E., Grechkin, Y., Ratner, A., Anderson, I., Lykidis, A., Mavromatis, K. and Ivanova, N.N., 2010. The integrated microbial genomes system: an expanding comparative analysis resource. *Nucleic acids research*, *38*(suppl_1), pp.D382-D390.
27. Moran, M.A., 2009. Metatranscriptomics: Eavesdropping on Complex Microbial Communities-Large-scale sequencing of mRNAs retrieved from natural communities provides insights into microbial activities and how they are regulated. *Microbe*, *4*(7), p.329.
28. Muyzer, G., 1999. Genetic fingerprinting of microbial communities–present status and future perspectives. In *Proceedings of the 8th International Symposium on Microbial Ecology* (Vol. 1, No. 10). Atlantic Canada Society for Microbial Ecology.
29. Muyzer, G., De Waal, E.C. and Uitterlinden, A.G., 1993. 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*, *59*(3), pp.695-700.
30. Nielsen, J.L., Christensen, D., Kloppenborg, M. and Nielsen, P.H., 2003. Quantification of cell- specific substrate uptake by probe- defined bacteria under in situ conditions by microautoradiography and fluorescence in situ hybridization. *Environmental Microbiology*, *5*(3), pp.202-211.
31. Okabe, S., Kindaichi, T. and Ito, T., 2004. MAR-FISH—An ecophysiological approach to link phylogenetic affiliation and in situ metabolic activity of microorganisms at a single-cell resolution. *Microbes and environments*, *19*(2), pp.83-98.
32. Okabe, S., Kindaichi, T. and Ito, T., 2005. Fate of 14C-labeled microbial products derived from nitrifying bacteria in autotrophic nitrifying biofilms. *Applied and Environmental Microbiology*, *71*(7), pp.3987-3994.
33. Poretsky, R.S., Bano, N., Buchan, A., LeCleir, G., Kleikemper, J., Pickering, M., Pate, W.M., Moran, M.A. and Hollibaugh, J.T., 2005. Analysis of microbial gene transcripts in environmental samples. *Applied and Environmental Microbiology*, *71*(7), pp.4121-4126.
34. Rastogi, G., Stetler, L.D., Peyton, B.M. and Sani, R.K., 2009. Molecular analysis of prokaryotic diversity in the deep subsurface of the former Homestake gold mine, South Dakota, USA. *The Journal of Microbiology*, *47*(4), pp.371-384.
35. Riesenfeld, C.S., Schloss, P.D. and Handelsman, J., 2004. Metagenomics: genomic analysis of microbial communities. *Annu. Rev. Genet.*, *38*, pp.525-552.
36. Rogers, S.W., Moorman, T.B. and Ong, S.K., 2007. Fluorescent in situ hybridization and

micro-autoradiography applied to ecophysiology in soil.

37. Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A., 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, *239*(4839), pp.487-491.
38. Schwieger, F. and Tebbe, C.C., 1998. A new approach to utilize PCR–single-strand-conformation polymorphism for 16S rRNA gene-based microbial community analysis. *Applied and Environmental Microbiology*, *64*(12), pp.4870-4876.
39. Singh, B.K., 2009. Organophosphorus-degrading bacteria: ecology and industrial applications. *Nature Reviews Microbiology*, *7*(2), pp.156-164.
40. Singleton, D.R., Furlong, M.A., Rathbun, S.L. and Whitman, W.B., 2001. Quantitative comparisons of 16S rRNA gene sequence libraries from environmental samples. *Applied and environmental microbiology*, *67*(9), pp.4374-4376.
41. Smit, E., Leeflang, P. and Wernars, K., 1997. Detection of shifts in microbial community structure and diversity in soil caused by copper contamination using amplified ribosomal DNA restriction analysis. *FEMS microbiology ecology*, *23*(3), pp.249-261.
42. Smith, C.J. and Osborn, A.M., 2009. Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. *FEMS microbiology ecology*, *67*(1), pp.6-20.
43. Tiedje, J.M., Asuming-Brempong, S., Nüsslein, K., Marsh, T.L. and Flynn, S.J., 1999. Opening the black box of soil microbial diversity. *Applied soil ecology*, *13*(2), pp.109-122.
44. Torsvik, V. and Øvreås, L., 2002. Microbial diversity and function in soil: from genes to ecosystems. *Current opinion in microbiology*, *5*(3), pp.240-245.
45. Tsai, Y.L. and Olson, B.H., 1991. Rapid method for direct extraction of DNA from soil and sediments. *Applied and environmental microbiology*, *57*(4), pp.1070-1074.
46. Urich, T., Lanzén, A., Qi, J., Huson, D.H., Schleper, C. and Schuster, S.C., 2008. Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome. *PloS one*, *3*(6), p.e2527.
47. Wellington, E.M., Berry, A. and Krsek, M., 2003. Resolving functional diversity in relation to microbial community structure in soil: exploiting genomics and stable isotope probing. *Current opinion in microbiology*, *6*(3), pp.295-301.
48. Whitman, W.B., Coleman, D.C. and Wiebe, W.J., 1998. Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences*, *95*(12), pp.6578-6583.
49. Wilmes, P. and Bond, P.L., 2006. Metaproteomics: studying functional gene expression in microbial ecosystems. *Trends in microbiology*, *14*(2), pp.92-97.
50. Zhang, Y., Shi, P., Liu, W., Meng, K., Bai, Y., Wang, G., Zhan, Z. and Yao, B., 2009. Lipase diversity in glacier soil based on analysis of metagenomic DNA fragments and cell culture. *Journal of microbiology and biotechnology*, *19*(9), pp.888-897.