Article ID 240204-S-09

Diversity of Soil Microorganisms: The Molecular Approach

V. Anu [1], S. Anu Rajan[2], N. Chitra[2] and V.I. Soumya[2]

[1]'Department of Agronomy, Kerala Agricultural University, College of Agriculture, Vellayani, Pin 695522'

[2] 'Department of Agricultural Microbiology, College of Agriculture, Vellayani Pin 695522'

*Corresponding Author: anu.v@kau.in

Abstract

The study of molecular ecology and diversity of soil microorganisms is crucial for understanding their role in plant health and soil quality. Several techniques have been devised to investigate the ecology, diversity, and behaviour of soil microorganisms, including PCR-based techniques, sequence-based techniques, and metagenomics. PCR-based techniques, such as fingerprinting methods and electrophoresis, allow for the identification and classification of bacteria based on their DNA sequences. Sequence-based methods that rely on the order of genetic information, such as 16S rRNA gene sequencing and multilocus sequence typing, provide more detailed information about bacterial species and their phylogenetic relationships. Metagenomics, meta transcriptomics, meta proteomics, and metabolomics are high-throughput approaches that enable the analysis of genetic material, gene expression, and protein profiles of microbial communities in their natural environments. However, these molecular methods have limitations, including the presence of enzyme-inhibiting compounds in soil and the instability of RNA molecules.

Introduction

The soil microbial population play a vital role in both plant health and soil quality. Microorganisms are a vast and varied collection of organisms that constitute approximately 60% of the Earth's total biomass. In soil, there are around 4-5 × 1030 microbial cells (Singh *et al.*, 2007). To determine an unknown microbe's functional features, link to other species in the environment, and biodiversity, taxonomic knowledge is crucial. To learn more about soil biology and the interactions between plants and microorganisms, scientists have created a variety of techniques for examining the ecology, diversity, and activity of bacteria.

The advent of PCR (polymerase chain reaction) initiated the utilization of genetics as a means of identifying microbes. Since then, a wide range of techniques—both culture-dependent and culture-independent—have been created using various guiding principles, including hybridization, 16S rRNA sequencing, and DNA sequencing. The objective of omics technologies, such as metagenomics, proteomics, lipidomics, transcriptomics, and metabolomics, is to collaboratively examine and quantify collections of biological molecules that have a role in the organization, operation, and changes of an organism or species. These methodologies possess significant practical implications and are now employed for many applications, such as transcriptional profiling, phylogeny, microbial ecology, and functional genome analysis.

Molecular based methods

The laboratory cultivation of microorganisms using specific growth media is a widely accepted method that is still used to identify microorganisms in samples from the environment. Many microorganisms in their natural environment display limited metabolic activity and/or poor rates of reproduction, rendering them viable but unable to be cultured (VBNC) species. Exclusively relying on culture-based approaches is insufficient for detecting these bacteria and, as a result, does not offer a comprehensive evaluation of microbial populations in samples. PCR enables the identification of viable, VBNC, and non-viable microorganisms by amplifying genetic sequences without the need for pre-enrichment of environmental samples.

Polyphasic taxonomy approach

The polyphasic taxonomy approach is a method that involves using many criteria to classify organisms. The term "polyphasic taxonomy," introduced by Colwell in 1970, describes the process of combining genotypic, chemotaxonomic, and phenotypic data of a microorganism to accurately classify and group the organism. The use of 16S rRNA gene sequencing and molecular fingerprinting techniques, together with other molecular markers, has become crucial in the field of microbial systematics, allowing for significant advancements in bacterial classification.

High throughput sequencing techniques

Presently, the examination of the microbial community in soil is accomplished through the utilization of contemporary molecular biology

methodologies. These approaches employ the extraction of whole DNA from soil followed by further analysis. Two separate techniques are employed for this objective. The first step involves identifying the makeup of the microorganism community by examining the sequence of barcode marker genes through the use of metabarcoding or target metagenomics. The second option involves utilizing metagenomics methodologies, such as global metagenomics and shotgun metagenomics, to determine the genetic composition of the population dwelling in a particular environment (Knight *et al.*, 2018). This process entails determining the taxonomic makeup of a community by amplifying the surrounding samples and subsequently employing high-throughput sequencing of DNA barcodes, such as 16S, ITS, or 18S. Sequencing approaches that rely on the amplification of marker genes are rapid, economical, and extensively researched.

Metagenomics is the comprehensive examination of genetic material obtained directly from samples taken from the environment. Analyze the genetic content of various microbial communities without the requirement for individual cultivation by utilizing high-throughput sequencing techniques. Metagenomics allows for the detection and characterization of microbial species, evaluation of their functional capabilities, and investigation of genetic variation within intricate ecosystems. Metagenomics gives details about the composition and diversity of a community, as well as its capacity for various metabolic activities. Metagenomics is applicable to the study of viruses, eukaryotes, and prokaryotes.

Next-generation sequencing (NGS), a subset of 'high-throughput sequencing (HTS)' methodologies, is now the cutting-edge approach for rapid and effective genome sequencing. By utilizing this technology, we are able to analyze the intricate composition of microbial communities, particularly those found in soil, through the sequencing of several genetic markers. Next-generation sequencing refers to a high-throughput sequencing technology that is not based on the Sanger method. Mass parallel sequencing enables the concurrent sequencing of millions of DNA fragments, significantly enhancing sequencing efficiency. This method offers several advantages over first-generation sequencing techniques. It is characterized by its high-throughput capability, cost reduction, shorter sequencing length, and high precision.

PCR-based approaches for bacterial identification and analysis

The Polymerase Chain Reaction (PCR) is a technique used to detect microorganisms present in soil. PCR can improve the ability to find and recognize certain microorganisms at different levels, such as species, strain, and serovar/pathovar, by amplifying specific segments of DNA. This technique is commonly employed to amplify the 16S rRNA or its gene before conducting fingerprinting studies. This method can also be used to analyze and describe entire populations of microorganisms in samples. The molecular markers that are most frequently amplified are ribosomal RNA (rRNA) genes. These genes mainly include 16S rRNA (bacteria and archaea), 18S rRNA, and internal transcribed spacer (ITS) sections (fungi) (Nkongolo *et al.*, 2020).

DGGE/TGGE

Denaturing gradient gel electrophoresis (DGGE) or Temperature gradient gel electrophoresis (TGGE) is a method that allows for the quick analysis of bacterial communities without the need for culturing. It enables the direct assessment of microbial populations in particular habitats. The organism-specific fingerprint is obtained and can be utilized to assess the composition and abundance of the community, by analyzing the banding pattern.

Stable isotope probing (SIP)

SIP is a scientific technique used to identify and track specific microorganisms in a complex mixture by labeling their DNA or RNA with stable isotopes. It helps in studying microbial populations that are actively engaged in specific metabolic activities in the environment. The goal is to establish a connection between the phylogeny (genetic relatedness) of these microbes and their functional roles

Fluorescent in situ hybridization (FISH)

'FISH' is a technique employed to measure the occurrence and proportional representation of microbial populations within a sample from a community. Microbial cells undergo fixation and are then subjected to hybridization with specific probes, typically 15-25 base pair oligonucleotide-fluorescently labelled probes. The visualization of these cells is achieved using either epifluorescence or confocal laser microscopy. Utilizing rRNA-targeted probes for hybridization improves the analysis of uncultured bacteria and simplifies the process.

Microarray technology

DNA microarray technology is a highly effective method used extensively to investigate biological processes, such as mixed microbial communities, with regards to their taxonomy and functionality. This method is comparable to FISH, however it allows for the simultaneous examination of several genes. A DNA microarray is a small array of DNA probes that are complementary to the target DNA. These probes can be either long (about 500-5000 nucleotides) or short (15-70 base pairs) and are directly linked to a solid substrate. This setup allows for the simultaneous hybridization of many DNA samples.

Sequence based techniques for bacterial identification

Bacterial species can be identified by analysing changes in sequences of house-keeping genes. The 16S rRNA, gyrB, rpoA, rpoB, rpoC, and rpoD genes are frequently employed for bacterial identification.

16S rRNA gene sequencing

The utilization of the 16S rRNA gene for identifying purposes is justified by its presence in all species that execute the same function. The sequence size is around 1500 bp, which is relatively straightforward to sequence and provides ample information for the identification and analysis of phylogeny. 16S rRNA gene sequencing is both speedy and precise and does not necessitate the use of specialized equipment or skill. Databases can also be used to retrieve sequences of similar species for comparative phylogenetic research. Subsequently, sequence-comparing software packages like BLAST and CLUSTAL X is employed to align the 16S rRNA gene sequence. The degree of genetic similarity between bacterial species can be examined by creating a phylogenetic tree or dendrogram using tree-making software like PAUP, PHYLIP, and MEGA 4.

Multilocus sequence typing (MLST)

It is a method used to analyse the genetic variation in several loci of an organism's genome. MLST focuses on sequencing internal portions of several house-keeping genes.

qPCR (quantitative PCR)

Quantitative PCR is a commonly used method in microbial ecology to measure the number of genes or transcripts in environmental materials. This technology enhances and supplements fingerprint and sequencing techniques. qPCR employs a fluorescent dye, usually SYBR Green, or a fluorescent probe such as TaqMan probes or 50-nuclease probes, to see the real-time amplification of a product. Furthermore, it facilitates understanding of the reaction's dynamics, enabling the determination of the initial amount of the product.

Limitations of molecular methods

The presence of enzyme-inhibiting organic substances, such as humic and fulvic acids, along with decreased extraction efficiency caused by the attachment of nucleic acids to soil particles, inadequate cell disintegration, and contamination by DNase and RNase enzymes. Bead-beating extraction methods are widely used and have been shown to be the most effective in addressing the issue of nucleic acid binding to soil particles. RNA-based research has greater challenges due to the heightened instability of RNA molecules in comparison to DNA. Establishing RNase-free environments poses a challenging endeavour due to the widespread occurrence and enduring functionality of RNases. Furthermore, it is common for mRNA to undergo fragmentation prior to cell lysis due to the concurrent processes of transcription and translation in archaeal and bacterial cells. When utilizing mRNA-focused techniques like as microarray and meta transcriptomics analysis, it is frequently recommended to perform a step to remove rRNA. This is because only a small portion, specifically up to 5%, of the overall RNA sample obtained consists of messenger RNA (mRNA).

Conclusion

In conclusion, employing various PCR-based techniques, such as fingerprinting methods and electrophoresis, allow for the identification and classification of bacteria in a speedy and effective way. Sequence-based techniques, such as 16S rRNA gene sequencing and multilocus sequence typing, provide more detailed information about bacterial species and their phylogenetic relationships. high-throughput approaches can be employed to enable the analysis of genetic material, gene expression, and protein profiles of microbial communities in their natural environments.

References

Knight, R., Vrbanac, A., Taylor, B.C., Aksenov, A., Callewaert, C.., Debelius, J., Gonzalez, A..,Kosciolek, T., McCall, L.-I. and McDonald, D. (2018). Best practices for analysing microbiomes. *Nature Reviews Microbiology*. 16, 410–422. https://doi.org/10.1038/s41579-018-0029-9

Singh, B.K., Campbell, C.D., Sorenson, S.J., Zhou, J. (2009). Soil genomics. *Nature Reviews*

Microbiology. 7, 756. https://doi.org/10.1038/nrmicro2119-c1

Nkongolo, K.K., Narendrula-Kotha, R. (2020). Advances in monitoring soil microbial community dynamic and function. *Journal of Applied Genetics*. 61, 249–263. https://doi.org/10.1007/s13353-020-00549-5

Rastogi, G. and Sani, R.K. (2011). Molecular Techniques to Assess Microbial Community Structure, Function, and Dynamics in the Environment. In Microbes and Microbial Technology. (Eds.); Ahmad, I., Ahmad, F. and Pichtel, J. Springer: New York, NY, USA, 2011. pp. 29–57.

* * * * * * * *