GMR The Original

FUNCTIONAL METAGENOMIC ANALYSIS OF MICROBIAL COMMUNITIES IN AGRICULTURAL ECOSYSTEMS

Vibhor Mahajan[1], Nallusamy Duraisamy[2], Dr. Kingsle Kishore Coumar M F[3], Durga B[4], Dr. Lovely[5], Nirmala A[6], Dr. Vijay Jagdish Upadhye[7], Sumitha Ramanathan[8]

[1]Centre of Research Impact and Outcome, Chitkara University, Rajpura – 140417, Punjab, India, Email: vibhor.mahajan.orp@chitkara.edu.in, ORCID: https://orcid.org/0009-0007-8272-1547
[2]Scientist, Department of Research, Meenakshi Academy of Higher Education and Research
[3]Assistant Professor cum Assistant Nursing Superintendent, Department of Medical Surgical Nursing, Kasturba Gandhi Nursing College, Sri Balaji Vidyapeeth (Deemed to be University), Puducherry, India, Email: kishore.coumar@gmail.com, ORCID: https://orcid.org/0009-0008-7959-0272
[4]Associate Professor, Meenakshi College of Allied Health Sciences, Meenakshi Medical College Hospital & Research Institute, Meenakshi Academy of Higher Education and Research
[5]Assistant Professor, Department of Biotechnology, Noida Institute of Engineering and Technology, Greater Noida, Uttar Pradesh, India
Email: lovely@niet.co.in, ORCID: https://orcid.org/0009-0006-4377-1106
[6]Associate Professor, Department of Biotechnology, Aarupadai Veedu Institute of Technology, Vinayaka Mission's Research Foundation (Deemed to be University), Tamil Nadu, India, Email: nirmalabt@avit.ac.in, ORCID: https://orcid.org/0000-0003-3992-1950
[7]Professor, Department of Microbiology, Parul Institute of Applied Sciences, Parul University, Vadodara, Gujarat, India,
Email: vijay.upadhye35296@paruluniversity.ac.in, ORCID: https://orcid.org/0000-0002-8821-1720
[8]Professor, ENT, Sree Balaji Medical College and Hospital, Bharath Institute of Higher Education and Research,
ORCID: https://orcid.org/0000-0003-2573-8784

ABSTRACT

In its current study, the authors examine the microbial communities and functional gene composition of the agricultural soils to clarify their functions on the productivity and sustainability of the ecosystem. The samples of soil were acquired at the opposite agricultural systems such as conventional and organic method of farming in order to realize variability of microbial composition as well as function. The high throughput shotgun metagenomic sequencing was performed on the total genomic DNA and then analysis using bioinformatics was conducted in which quality control, sequence assembly and functional annotation of the results with the help of the existing databases like KEGG and COG were performed. The findings showed that there was great difference in the structure of the microbial community in a variety of agricultural environments with characteristic diversity patterns depending on the management practices. The range of metabolically significant genes that were identified in functional profiling were numerous and included those of nitrogen fixation, carbon cycling, phosphorus metabolism, among other essential biogeochemical mechanisms. It is also important to note that organic farming systems had better functional diversity and greater number of genes related to nutrient recycling and soil health than conventional systems. Such findings suggest that the micro biome of the soil is essential in controlling the functions of agricultural eco systems and the functional metagenomic is important in understanding the contribution of the microbial role in soil fertility. The research gives an insightful information on the development of micro biome-based approaches to sustainable agriculture and a better way of managing the soil.

KEYWORD: Shotgun sequencing; KEGG pathway analysis; Nutrient cycling; Nitrogen fixation; Carbon metabolism; Sustainable agriculture.

1. INTRODUCTION

The communities of microbes in the soils are basic in ensuring stability, productivity, and sustainability of agriculture ecosystems. These microbial communities are known to control critical biogeochemical activity such as nutrient cycling, decomposition of organic matter, and soil structure development, and therefore, directly affect crop growth and soil fertility (Fierer, 2017; Jansson and Hofmockel, 2020). Plant-soil microbiotic relationships are becoming an important factor that defines the success of agriculture, as the microbes provide nutrient availability, pathogen control, and stress resistance to crops (Singh et al., 2020; Banerjee et al., 2018). Due to the increasing need in sustainable cultivation procedures, the structure and functioning of the soil microbial communities have become the key topic in the studies of the environment and agriculture.

The microbiological methods which are conventional in nature that almost rely on culture-dependent methods have severely restricted our knowledge about the diversity and role of soil microbiota. A huge percentage of the microorganisms found in the soil is estimated to be uncultivable in the laboratory environment and results in the incomplete representation of microbial communities and their ecological functions (Torsvik and Øvreas, 2002; Thomas et al., 2012). Even culture-neutral techniques like 16S rRNA gene sequencing can give only taxonomic information and they cannot reflect functional possibilities of microbial communities. In turn, these constraints do not provide the full picture of the process of how soil microbes may affect the ecosystem processes and productivity of the agriculture sphere.

Functional metagenomics is the new potent and groundbreaking technique that should be applied to resolve these drawbacks by conducting the direct analysis of genetic material that is obtained in the samples of the environment.

Functional metagenomics enables gene identification and characterization to be done in important metabolic pathways without microbial growth using high-throughput sequencing and tools of advanced bioinformatics (Quince et al., 2017; Knight et al., 2018). It is a technique that gives in depth information on the functioning abilities of microbial communities, their duties in nutrient conversions, stress acclimatization, and neighborhood endurance. Moreover, annotations of gene related to key biological processes are made possible through the use of databases like KEGG and COG, which allow to have a better grasp of the contribution of microbes to the agricultural ecosystem (Franzosa et al., 2018; Nayfach and Pollard, 2016).

The conversion and mobilization of nitrogen, phosphorus and carbon are considered to be among the most important processes of soil microbial communities which are involved in the nutrient cycling. Fixation and denitrification activities of microbes are important in keeping soils fertile and plant growth (Kuypers et al., 2018; Canfield et al., 2010). In the same way, microbial communities enhance carbon cycling by decomposing organic matter and sequestration of carbon which is essential to the health of soils and the regulation of climate (Kuzyakov and Blagodatskaya, 2015). Moreover, plant growth-promoting microorganisms also increase nutrient uptake, generate phytohormones and protect plants against pathogens, which are some of the factors that promote agricultural productivity (Philippot et al., 2013). These functional properties support the significance of microbial gene profile in soil of agriculture.

Although, there has been major improvement in metagenomic technologies, there is a great gap of research that is critical in the characterisation of microbial communities at functional level across different agricultural systems. Most of the studies that are available pay more attention to taxonomic diversity, and less to the functional genes and metabolic pathways that have a direct impact on the soil processes and the crop performance (Delgado-Baquerizo et al., 2016). Further, the comparative studies of various agricultural management systems including conventional and organic farming have not been done yet to explain the effects of practices in details on the functionality of microbes and the sustainability of the ecosystem.

In this regard, the current research proposal will examine the microbial composition and functional gene expression of agricultural soils through the shotgun metagenomic method. This research aims at determining the fundamental pathways in the metabolism associated with the cycling of nutrients and the functioning of ecosystems by using the collected samples of soil of various agricultural systems and the functional annotation of these systems through the use of advanced bioinformatics software. The results will shed more light on how soil microbiomics may be used in the sustainability of agriculture and help to create micro biome-based solutions that would help in better soil management.

The proposed research contributes to the existing body of knowledge through the combination of microbial diversity analysis with functional metagenomic profiling of the agricultural ecosystem to reveal important gene pathways that are related to nutrient cycling and soil health. As compared to the traditional studies where the focus is on taxonomic composition, this study will include a functional approach to soil microbiomics which will provide practical information on sustainable agriculture and micro biome-based soil management methods development.

2. LITERATURE REVIEW

The microbial communities of soil constitute one of the most diverse and functionally dynamic parts in the terrestrial ecosystems and are essential in ensuring the soils have the fertility, stability of the ecosystems, and the productivity of agriculture. They include bacteria, archaea, and fungi (as well as other microorganisms) that adjust biogeochemical processes needed by the soil by reacting with physicochemical characteristics of the soil and the roots of plants (Fierer, 2017; Torsvik and Øvreås, 2002). Environmental conditions including the level of pH, moisture, and organic matter, anthropogenic actions, especially farming activities affect the composition and diversity of soil microbiota. It has been demonstrated that organic systems of farming typically maintain high levels of microbial diversity and complex ecological communities in comparison to conventional systems because of lower levels of chemical additions and higher levels of organic additions (Hartmann et al., 2015; Lupatini et al., 2017). Conversely, intensive conventional management of nutrient loss can interfere with the community structure and functional stability of microbes and hence, soil health and long-term productivity (Banerjee et al., 2019).

The development of the molecular techniques has played a great role in the comprehension of the microbial community within the soil. The culture-dependent techniques have traditionally been known to be insufficient since they only record a very small percentage of the microbial diversity within the soil (Thomas et al., 2012). With the introduction of culture-independent identification microbial taxa was made possible through the 16S rRNA gene sequencing but again this will only offer taxonomic characterization as opposed to offering functional capability information. Recently, shotgun metagenomics has become a potent method that can simultaneously examine the entire range of genetic material of samples of environmental samples and, as a result, can be used to simultaneously determine the diversity of microorganisms and their functional capacity (Quince et al., 2017; Knight et al., 2018). The possibility to use databases to label the functional genes, such as KEGG and COG, also adds to a better understanding of the role of microbes in the ecosystem processes as it helps identify the genes related to the important metabolic pathways (Franzosa et al., 2018). Nevertheless, metagenomic analyses are not yet as useful as they can be due to such limitations as the necessity to compute a lot, the lack of complete reference databases, and the inability to connect genes with a particular taxon (Nayfach & Pollard, 2016).

The use of functional metagenomic in agricultural studies has gained significance especially in the study of nutrient cycling and health of soils using microorganisms. Microorganisms in the soil are very involved in the process of transformation and mobilization of vital nutrients such as nitrogen, phosphorus, and carbon that are important in growing plants and the sustainability of the ecosystem. The metagenomic research has found that agricultural activities have a profound effect on the level of functional genes that relate to these nutrient cycles (Hu et al., 2022; Delgado-Baquerizo et al., 2016). As an illustration, organic amendments have also been found to improve genes that deal with nitrogen fixation and carbon metabolism to increase nutrient availability and soil fertility (Hu et al., 2022). Equally, a change in microbial functional profile and especially those genes associated with nitrogen cycling and decomposition of organic matter can be achieved by changes in tillage intensity and other management practices (Cloutier et al., 2023). These results indicate the relevance of functional-level analysis to the study of the role of microbial communities in ecosystem operation and the sustainability of agriculture.

Along with the nutrient cycling, the microorganisms of the soil are very important in the growth of plants and increase in crop production. Microorganisms that promote plant growth include rhizobacteria and mycorrhizal fungi that enhance nutrient uptake, synthesize phytohormone and shield plants against pathogens (Philippot et al., 2013). These positive interactions between plants and microorganisms are very fundamental in enhancing crop production and resistance, especially in adverse environmental factors. Moreover, the carbon cycling with the help of microbes is important to form organic matter in soils and sequester carbon, which is essential in preserving soil structure and countering climate change (Kuzyakov and Blagodatskaya, 2015). There are also nitrogen cycling processes that are facilitated by specialized microbial groups, which make sure that the continuous supply of bioavailable nitrogen to plants (Nitrogen fixation, nitrification, and denitrification) is ensured (Kuypers et al., 2018; Canfield et al., 2010). Altogether, these functional roles provide a high level of significance of soil microbiomes to maintaining agricultural productivity and ecosystem health.

Although these innovations have been made, there are still a lot of gaps in the existing knowledge of the soil microbial activity in the agricultural ecosystems. Most of the researches are still based on taxonomic diversity, and functional gene profiling and metabolic pathway analysis are not given much attention. In addition, there is the lack of incorporation of multi-omics methods, e.g., the metatranscriptomics and metaproteomics methods, which is required to connect the genetic potential and actual microbial activity in the soil. Also, there is limited comparative research in various agricultural systems and hence it is not possible to generalize and come up with scalable approach to sustainable agriculture. Thus, a great deal of functional metagenomic studies that are not only capable of defining the microbial diversity but also clarify the metabolic processes underlying the soil fertility and sustainability of the ecosystem is much needed to fill the gap between microbial ecology and practical applications to agriculture.

3. MATERIALS AND METHODS

Representative sites were sampled to ensure agricultural soil samples had contrasting farming systems such as conventional and organic agricultural systems to achieve variability on the microbial diversity and functional potential. These locations were chosen to vary in the type of soil, patterns of crop growth and also in the environment (temperature, moisture and content of organic matter). The samples of the soil consisted of the top layer (0-15cm) of the soil, and in this case, it is known that the top layer is where the microbial activity is more active. Multiple subsamples were taken in every site through randomized sampling method and mixed to create composite samples whereby spatial representation was achieved and heterogeneity was minimized. All the samples were taken to the laboratory in sterile and cool conditions and kept in the refrigerator at -20 °C before further examination.

A standardized protocol of soil DNA extraction was used to obtain total genomic DNA in soil samples to ensure the efficient recovery of microbial DNA with a minimum level of contamination and inhibitors. Spectrophotometry and agarose gel electrophoresis were used to determine the quality and quantity of extracted DNA. The quality of the DNA samples was then analysed by subjecting them to library preparation according to the high-throughput sequencing standard protocols. The metagenomic sequencing was done by use of an Illumina sequencing platform, which produced paired-end reads that had high coverage to give a comprehensive coverage of the microbial genomes in proportion to the samples. Fig 1 shows a general flow of the experiment that incorporated the sampling, DNA extraction, sequencing, and bioinformatics analysis of the downstream sequencing.

Data of raw sequencing were handled through reliable bioinformatics pipelines to facilitate quality and reliability of data. Primary quality control was implemented with the help of FastQC that measured quality of the sequence and the next step was the trimming and filtering of low-quality reads and adapter sequences with the help of Trimmomatic. The reads of high quality were assembled into contiguous sequences (contigs) with the help of de novo assembly programs (MEGAHIT or SPAdes). Calculation of gene prediction Assigned contigs were predicted using programs like Prokka or MetaGeneMark to find coding regions and produce gene catalogs to be used in subsequent analysis. To provide a thorough characterization of microbial metabolic potential, functional annotation of the predicted genes was conducted to use various databases. The metabolic pathways and functional modules related to nutrient cycling and other ecological processes were determined by using KEGG (Kyoto Encyclopedia of Genes and Genomes). Besides, COG (Clusters of Orthologious Groups) classification was used to classify genes into functional groups and MetaCyc and eggNOG database were used to narrow down pathway and functional annotations. This multi-database strategy facilitated the detection of strong genes that were involved in major biogeochemical activities such as fixation of nitrogen, carbon metabolism as well as phosphorous cycling.

The analyses of microbial diversity were done with an aim of determining both taxonomic and functional diversity of various agricultural systems. The alpha diversity indices such as Shannon and Simpson indices were computed to measure the evenness and the richness of species in individual samples. To compare the variability in the microbial community structure in samples and agricultural practices, multivariate analysis was carried out with Principal Component Analysis (PCA) and Principal Coordinates Analysis (PCoA) to assess the differences in beta diversity between the samples and agricultural practices. To test the necessity of noticed changes in diversity and abundance of microbial functions, statistical analyses were conducted. The differences between the groups were analyzed with the help of analysis of variance (ANOVA), whereas multivariate statistical measures were applied to detect the patterns and relationships inside the dataset. To investigate how the environmental factors impact the nature of microbial community and the ecosystem functioning, the correlation analyses were conducted to study how microbial functional profiles are associated with the soil physicochemical properties.

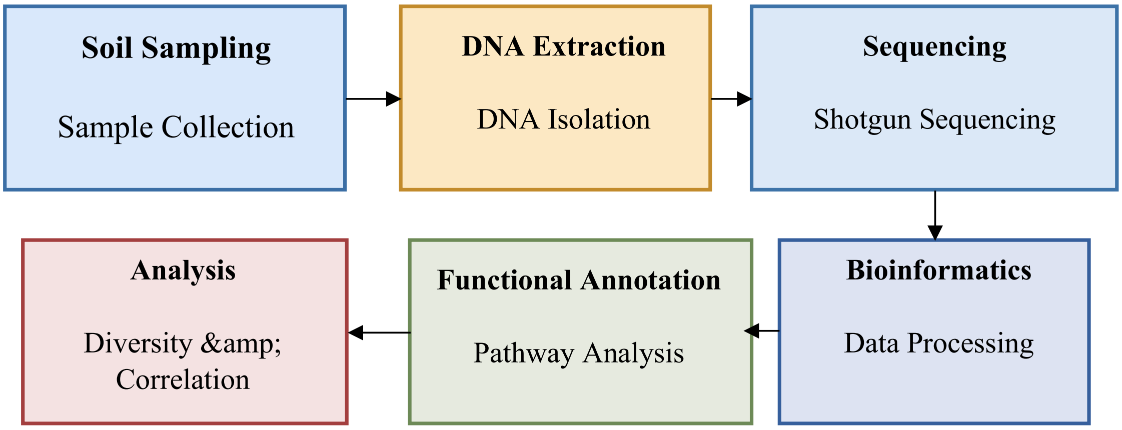


Fig 1: Workflow of Functional Metagenomic Analysis.

4. RESULTS

The meta-genomic analysis of soil samples in agriculture gave a detailed information on the diversity of microbes and distribution of functional genes and their interaction with the environment in various agricultural systems.

4.1 Sequencing Output and Data Quality

As the statistics of the sequencing and assembly of Table 1 show, the metagenomic data proves to be strong and reliable in all the cases of agricultural soils sample. Each of the samples produced a high count of raw reads, with a read count of about 45 million up to more than 52 million raw reads, signifying a deep sequencing that was able to capture soil microbial communities. After running through rigorous quality control analyses, a great percentage of reads (more than 90%) was retained as clean reads indicating the quality of the sequencing output and low contamination or technical noise. The high accuracy of base calling and sequencing reliability are also supported by the values of Q 30, which exceeds 90% in all the samples. The difference in Q30 percentages recorded in organic and conventional systems (OS1 and CS1 respectively) is small, but indicates that in the former case, the integrity of the sequence is greater, this may be attributed in part to increased microbial diversity and alleviated environmental pressure in soils managed with organic methods. On the same note, the GC content data was between 56.5 and 58.2 with a very stable and consistent GC content in the samples, with a slight variation of higher GC content in organic systems which could also be as a result of the different structure of the microbial communities.

The assembly statistics showed that there was a significant difference in farming systems. Organic soil samples had more assembled contigs (around 180,000-186,000) than conventional soils (around 155,000-162,000), and had more genomic complexity and diversity. The N50 of 1,010 bp to 1,245 bp was also significant in organic systems indicating higher quality and longer assembled sequence. Moreover, the size of the largest contigs was larger in the organic samples, which also helped to confirm the existence of more complicated and diversified microbial genomes. Gene prediction analysis revealed that there were high coding sequences in all samples with organic systems (more than 300,000 in OS1) having more predicted genes than conventional systems (around 260,000-274,000). Such a trend suggests that the functional gene pool of organic soils is more enriched, probably it is connected with high metabolic activities and the interactions with the ecological environment. The mean size of the genes was quite stable between samples and indicated a similarity in the accuracy of gene predictions.

Moreover, the findings of functional annotation showed that a larger percentage of genes in organic systems could be successfully annotated on the bases of KEGG and COG databases with the level of annotated genes being more than 250,000 in OS1 and 210,000 in CS2. This implies that besides having more genes, organic soils have a higher diversity of functionally characterised pathways. In general, the evidence shown in Table 1 demonstrates clearly that organic agricultural systems favour high sequencing depth and quality assembly as well as high functional gene diversity in contrast to conventional ones. The impact of agricultural management practices on the complexity of the microbial genomes and these findings are a good basis of future analysis of microbial composition and their functional potential.

Table 1: Sequencing Output and Assembly Statistics of Agricultural Soil Metagenomes

Parameter	**Organic System 1 (OS1)**	**Organic System 2 (OS2)**	**Conventional System 1 (CS1)**	**Conventional System 2 (CS2)**
Total Raw Reads	52,846,320	49,732,115	47,985,642	45,673,298
Clean Reads (After QC)	49,215,876	46,982,443	44,765,210	42,981,567
Q30 (%)	92.4	91.8	90.7	90.2
GC Content (%)	58.2	57.9	56.8	56.5
Total Contigs	186,245	179,832	162,514	155,276
N50 (bp)	1,245	1,198	1,042	1,010
Largest Contig (bp)	48,762	45,318	39,875	37,642
Predicted Genes	312,456	298,732	274,118	261,904
Average Gene Length (bp)	912	905	887	879
Annotated Genes (KEGG/COG)	258,764	245,118	221,903	210,547

4.2 Microbial Community Composition

The analysis of the taxonomic composition showed that bacterial communities prevailed in all samples of soil, and archaea and fungi had relatively minor contributions. Proteobacteria always appeared to have the highest relative abundance of all bacterial phyla followed by Actinobacteria, Firmicutes and Bacteroidetes. Thaumarchaiotata were the most common archaeal and there were also Ascomycota and Basidiomycota whose proportions were lower.

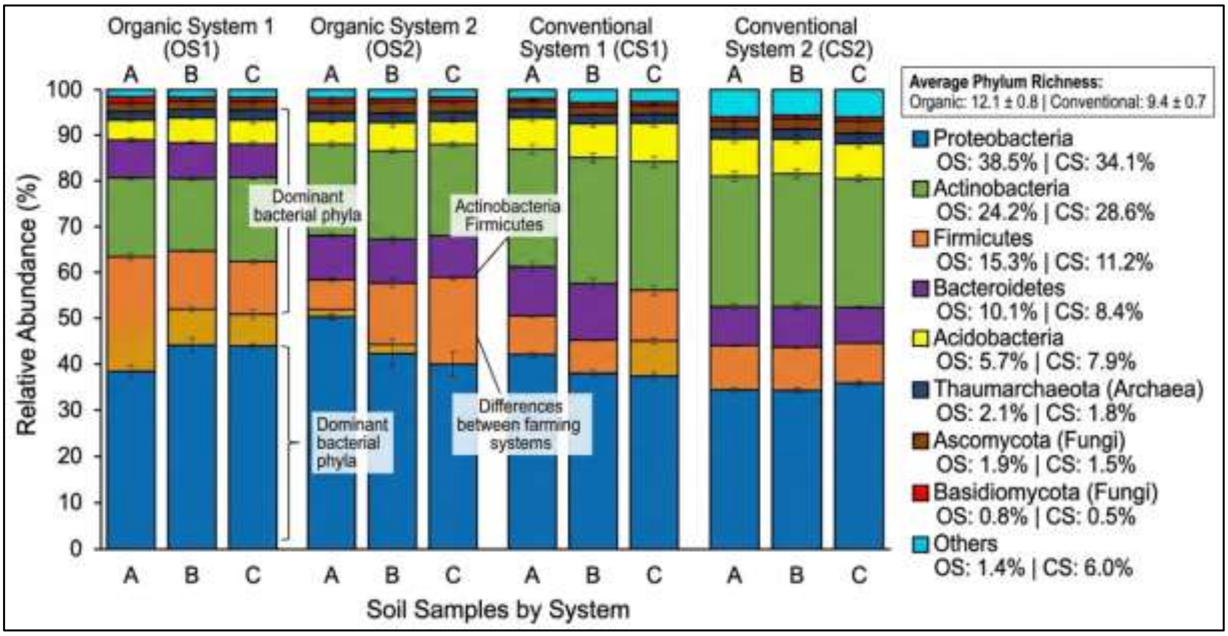


Fig 2: Relative Abundance of Soil Microbial Communities

Fig 2 showed that there were specific differences in microbial composition in the organic (OS1, OS2) and conventional (CS1, CS2) systems. Organic soils showed higher relative prevalence of Actinobacteria (about 24.2%), and Firmicutes (15.3%), which shows an increase in the activity of decomposition and the turnover of nutrients. Conversely, traditional soils also exhibited a relatively higher Acidobacteria and less proportion of beneficial microbial groups. Also, the mean phylum richness of organic system (12.1 ± 0.8) was higher than that of conventional system (9.4 ± 0.7) indicating that there were better microbial diversity in organic system with sustainable management practices. These outcomes suggest that crop cultivation has a great impact on the organization of the microbial community and ecology.

4.3 Functional Gene Profiling

Functional annotation showed that microbial communities contain a comprehensive repertoire of genes linked to relevant metabolic pathways, especially the nutrient cycling ones. Fig 3A shows that the top major KEGG functional categories were predominantly made up of metabolism related genes, then genetic information processing, and environmental information processing. Organic systems were characterized by a little higher percentage of metabolic and genetic processing pathways which are a sign of an increased functional activity. A more stringent examination of the genes on the gene level (Fig 3B) showed the predominance of nitrogen cycling genes (nifH, nifD, and nifK), which are crucial in fixing nitrogen and nitrification-related genes (amoA and hao). The genes like rbcL and rbcS represented the carbon fixation pathways whereas the phosphorus cycling was linked with gcd and pqqC. It is also important to note that Gene abundance was notably greater in organic system over conventional systems, especially the fixation pathways to nitrogen fixation and metabolism of carbon. Significant statistics in the figure show that such differences are not accidental but significantly affected by agricultural management practices. All in all, these findings prove the fact that the active participation of soil microbiomes in the nutrient transformation processes occurs, and organic systems maintain more functionally diversified microbial communities.

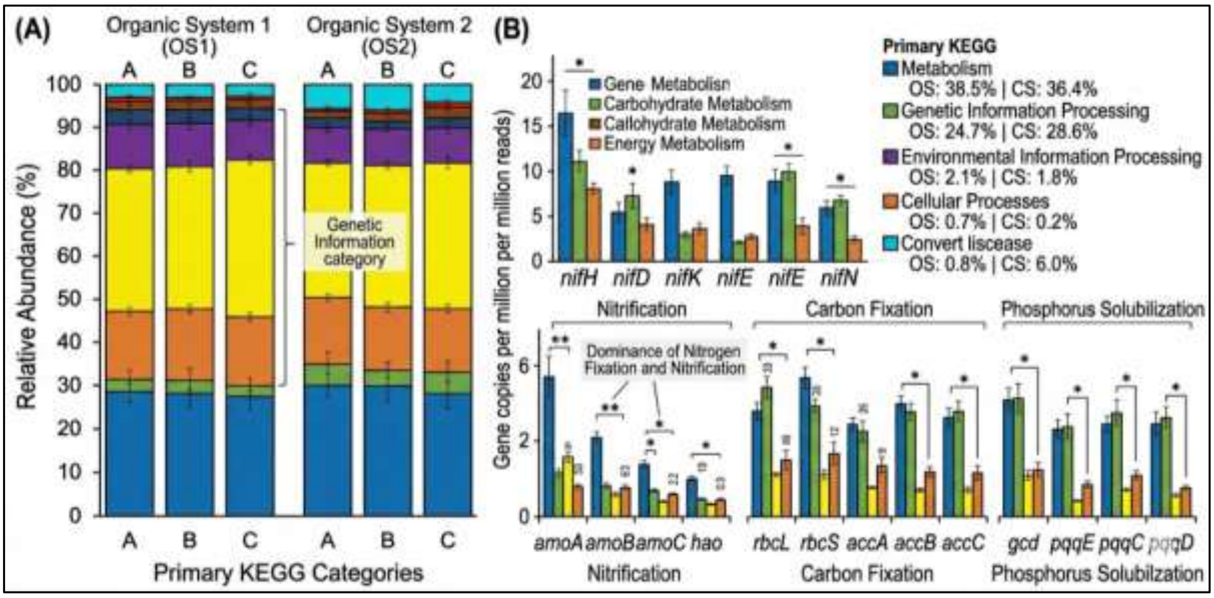


Fig 3: Functional Profiling of Microbial Communities

4.4 Comparative Analysis across Ecosystems

Multivariate analysis also helped to bring out the variance in the microbial communities in agricultural systems. According to the Principal Component Analysis (PCA) plot in Fig 4 (left panel), it is evident that the organic and conventional samples are separated on the first two principal parts with PC1 and PC2 contained 32.1 and 21.4% of the variance respectively. The concentration of samples shows that the composition of microbes is highly dependent on the farming practice. Likewise, it is clear that the Principal Coordinates Analysis (PCoA) plot in Fig 4 (right panel) indicates discrete clustering according to the profiles of functional genes with PCoA axis 1 describing 45.3% of the variance. The distinct division that is made between the organic and conventional systems indicates that the taxonomic composition and the functional gene distribution are greatly changed by the agricultural management. These differences are statistically verified with the help of the PERMANOVA ($p < 0.001$), and the results are considered extremely significant. The data provided show that organic farming enhances a more specific and functionally diverse micro-biological ecosystem than standard farming.

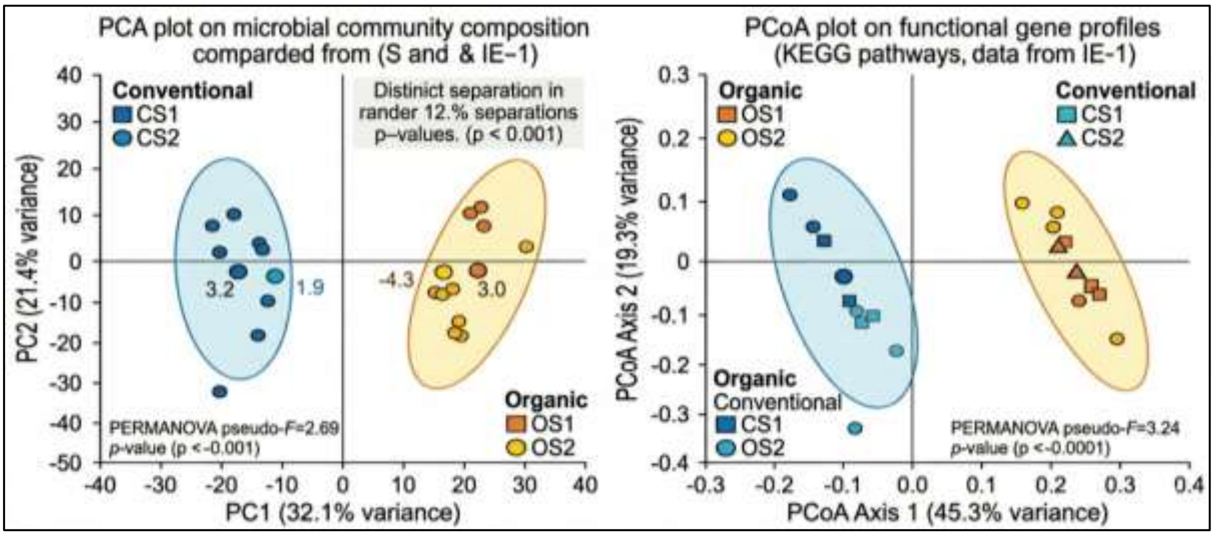
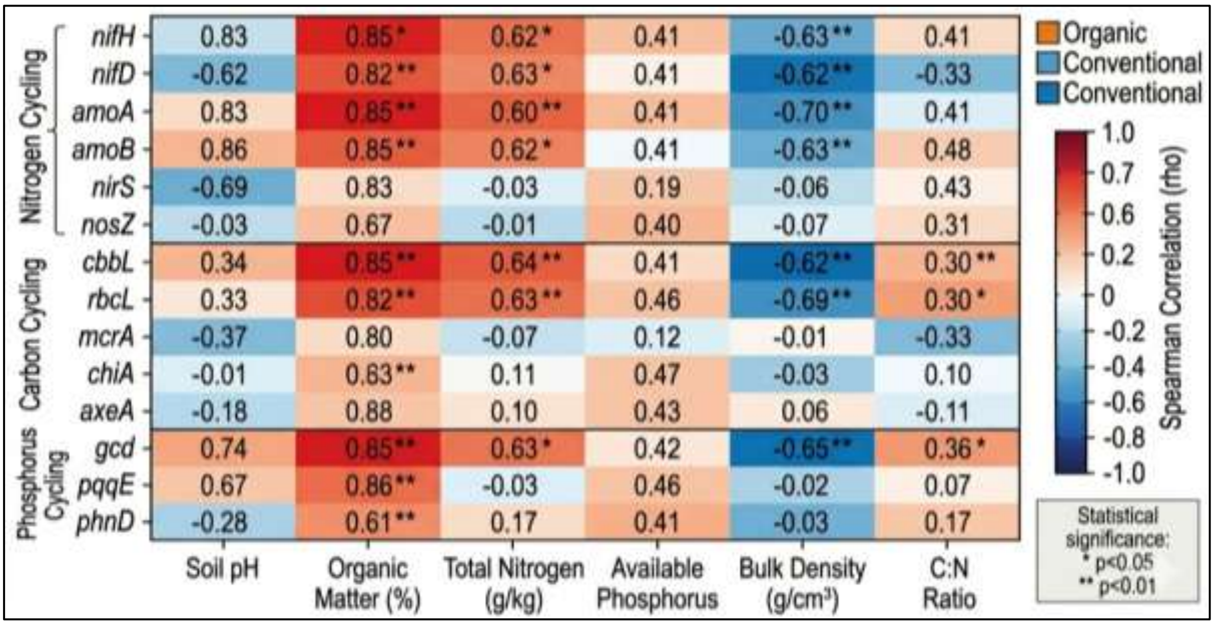


Fig 4: Multivariate Analysis of Microbial Community Structure.

4.5 Correlation with Soil Parameters

Correlation analysis was used to investigate the relationship of microbial functional genes with soil physicochemical properties as demonstrated in Fig 5. The heatmap demonstrates that there are positive correlations between the content of organic matter and the genes of nitrogen cycling (e.g., nifH, amoA) and carbon metabolism (e.g., rbcL), which are strong and positive. On the other hand, bulk density reported negative relationships with a number of functional genes implying that compacted soils can restrict microbial activities and diminish functional diversity. The change of soil pH was also variable, with some genes having positive correlation and others having negative relations, which demonstrates the complicated impact of pH on the ecological state of microbes. Moreover, there were positive associations between total nitrogen and available phosphorus and key functional genes associated with nutrient cycling that support the hypothesis that microbial communities play a role in the regulation of soil fertility. The fact that the correlations are statistically significant ($p < 0.05$ and $p < 0.01$) was another indication of the strength of these interactions. On the whole, the correlation study indicates that soil properties and microbial functional potential are highly interdependent, which is why environmental factors are essential in the context of forming soil microbiomes.



		Soil pH	Organic Matter (%)	Total Nitrogen (g/kg)	Available Phosphorus	Bulk Density (g/cm³)	C:N Ratio
Nitrogen Cycling	*nifH*	0.83	0.85*	0.62*	0.41	-0.63**	0.41
	nifD	-0.62	0.82**	0.63*	0.41	-0.62**	-0.33
	amoA	0.83	0.85**	0.60**	0.41	-0.70**	0.41
	amoB	0.86	0.85**	0.62*	0.41	-0.63**	0.48
	nirS	-0.69	0.83	-0.03	0.19	-0.06	0.43
	nosZ	-0.03	0.67	-0.01	0.40	-0.07	0.31
Carbon Cycling	*cbbL*	0.34	0.85**	0.64**	0.41	-0.62**	0.30**
	rbcL	0.33	0.82**	0.63**	0.46	-0.69**	0.30*
	mcrA	-0.37	0.80	-0.07	0.12	-0.01	-0.33
	chiA	-0.01	0.83**	0.11	0.47	-0.03	0.10
	axeA	-0.18	0.88	0.10	0.43	0.06	-0.11
Phosphorus Cycling	*gcd*	0.74	0.85**	0.63*	0.42	-0.65**	0.36*
	pqqE	0.67	0.86**	-0.03	0.46	-0.02	0.07
	phnD	-0.28	0.61**	0.17	0.41	-0.03	0.17

Fig 5: Correlation between Microbial Functions and Soil Properties.

5. DISCUSSION

The current research study offers an in-depth information on the functional and taxonomic dynamics of soil microbial communities in the various agricultural systems with the main focus on the essential nature of functional metagenomics in the context of ecosystem functions. The presented changes in microbial diversity and abundance of genes vividly show that microbiomes of soil are very sensitive to agricultural management activities. The increased sequencing depth, prediction of genes and functional annotations of the higher order observed in organic systems suggest that there is a more sophisticated and active microbial community. It is an indication that organic farming methods foster an appropriate environment on microbial growth and functional differentiation, which increase the resilience of the ecosystem. Among the most interesting results of this work, the prevalence of genes related to the nutrient cycling, especially, the nitrogen metabolism, the carbon fixation and the phosphorus cycling should be singled out. The pathways are essential to the agricultural productivity and soil fertility. This is because the presence of nitrogen fixation genes is more in organic systems implying that biological nitrogen is more available and therefore the fetal reliance on synthetic fertilizers is minimized. In the same manner, the enhancement of the carbon metabolism pathways indicates augmentation of the microbial activity in the breakdown of organic matter and sequestration of carbon, which is very vital in the stability of the soil structure and long term fertility. The availability of phosphorus cycling genes is another contribution towards how microbial communities have eased the nutrient availability to the plant uptake and thus, led to sustainable management of soils.

There are also microbial communities, which play a central role in facilitating the crop productiveness after plant-microbe interactions. The greater abundance of functionally active microbial populations in organic systems indicates that organic systems are more able to provide the mechanism of enhanced plant growth promotion, such as solubilization of nutrients, phytohormone synthesis, and suppression of pathogens. These interactions improve the health and the resilience of plants, especially when the environment is subjected to a stress environment. Conversely, it was found that microbial diversity and functional redundancy in the conventional systems is relatively lower thus suggesting a possible decrease in ecosystem stability that might be as a result of intense agricultural inputs e.g. chemical fertilizers and pesticides. The result of this paper is similar to that of the earlier metagenomic research reports which have claimed considerable changes in the composition of microbial communities and functional genes profile as a response to varying agricultural activities. Previous studies have established that organic the additions and less tillage activities raise microbial diversity and increase functioning genes related with nutrient cycling. The present paper conveys these findings by giving a more in-depth functional outlook, including the individual metabolic routes that become enriched in the agricultural systems that are sustainable. Also, the discerned microbial community segregation experienced in the multivariate analyses promotes the opinion that the agricultural management practices are some of the main determinants of microbial ecological structure.

Agricultural practices have an apparent effect on the microbiome in the soil, both in terms of taxonomic distribution and in terms of functional aspects. Organic systems encourage more diverse and functionally diverse microbiome that lead to better soil health and ecosystem sustainability. Conversely, traditional methods can cause simplified microbial communities that have limited functional ability that could influence long-term productivity of soil. Such discoveries affirm the need to embrace sustainable agricultural production that promotes the diversity of microbes and the environment. Ecologically it has been noted that there exists the strong correlation between microbial functional genes and soil physicochemical properties, which indicates the interdependence of soil ecosystems. Organic matter content, pH and availability of nutrients are some of the factors that determine the structure and functioning of microbial communities. This rationale gives the positive correlation of organic matter and key functional genes, which indicates that soil amendments are capable of improving the functioning of microorganisms and the ecosystem to a great extent.

Biotechnologically, the knowledge obtained in terms of functional metagenomic analysis provides new opportunities in terms of designing microbiome-based agricultural solutions. The discovery of the important functional genes in cycling of nutrients gives the prospects of designing biofertilizers and microbial consortia to enhance soil fertility and crop production. Also, microbial functional profile monitoring can be useful in precision agriculture through the provision of soil health-specific interventions. All in all, this paper strengthens the relevance of functional metagenomics as a potent approach to the study of the ecology of soil microorganisms and its application in sustainable agricultural practices. The combination between taxonomic and functional analysis gives a complete picture of how microbes contribute to the processes of the ecosystem, which makes it essential to conduct further research in the given field, to come up with new solutions to agricultural problems in the world.

6. LIMITATIONS

Although this study has yielded useful information on the functional metagenomic properties of microbial communities in the soil of agricultural ecosystems, it has some limitations, which ought to be mentioned. To begin with, the research was carried out in a small number of very few sampling sites thus it might limit generalization of the results. The microbial communities that occur in the soil are very heterogeneous and affected by the site-specific environmental conditions, hence larger geographic sampling would give a more detailed insight on the microbial diversity and functionality in various agro ecosystems. Second, high-throughput sequencing was used; thus, the sequencing depth limitation could have prevented the identification of low-abundance or rare microbial taxa and functional genes. These rare microbial communities can be more of an ecological importance, and underrepresentation could result into an incomplete description of soil micro biome.

Third, the functional annotation of metagenomic data is based on the existing reference databases like KEGG, COG, and eggNOG that might not be quite reflective of the diversity of microorganisms of the environmental type. Consequently, there are chances that a number of the predicted genes can not be annotated or wrongfully grouped, thus restricting the understanding of functional possibilities of microbes. This dependency of genetic information databases shows the necessity of the constant development and improvement of genetic databases.

Lastly, the functional potential of the microbial communities under study is mostly analysed based on the metagenomic data, and no active or passive experimentation of gene expression or activity was done. This is hindered by the fact that some complementary methods like metatranscriptomics, metaproteomics or biochemical assays are not done to establish the fact that identified genes are actively involved in processes of an ecosystem. In general, despite the fact that the findings are very informative in terms of microbial diversity and functional pathways, these limitations in the future studies will increase the precision, depth and applicability of the metagenomic studies in agricultural systems.

7. FUTURE PERSPECTIVES

Future studies on the soil microbiome should go beyond the conventional approach of metagenomic profiling at a single point in time in favor of a more multi-omics platform to determine better the functionality of the microbios in agricultural systems. Integration of metatranscriptomics and metaproteomics will provide the possibility to identify active genes and proteins, which will shed light on the activity of microbes in real-time as opposed to the genetic potential only. These methods will assist in filling the gap between predicted functions and real metabolic functions and provide a more precise picture of the role of microbes in the cycle of nutrients and health in the soil.

Moreover, the further application of artificial intelligence (AI) and models of machine learning has great potential in the development of functional metagenomic analysis. Complex patterns in large-scale sequencing datasets can be detected with the help of AI-based predictive models and this approach will allow predicting more accurately the response of the metabolic pathways, microbial interactions, and ecosystems to environmental changes. They may also be used to discover new functional genes, and annotate older uncharacterized sequences, overcoming existing limitations on database-based analyses.

The other significant avenue is that of formulation of microbial based biofertilizers and artificial microbial consortia to be used with the view of improving soil fertility and crop productivity. With the help of the information on functional metagenomics, particular microbial strains or groups, which have desirable qualities, including nitrogen fixation, phosphorus solubilization, and stress tolerance, can be enriched and utilized in an agricultural environment selectively. This can help to decrease the reliance on chemical fertilizers and in order to leverage sustainable farming. Moreover, real-time tracking of the microbiomes of the soil could be applied by new technologies on sensors and sequencing, which can offer moving data on the diversity and functional activity of the microbes in different conditions of the environment. High precision agriculture can be reinforced through such monitoring systems since interventions can be conducted in time due to indications of soil health and microbial response. Altogether, the combination of the innovative technologies of omics, AI-based analytics, and applied microbiome engineering should revolutionize the agricultural practice, allowing to introduce the approach to soil management that could be considered as data-driven, sustainable, and resilient.

CONCLUSION

This paper will show that functional metagenomic analysis can give invaluable hints into the composition and metabolic capabilities of the soil microbial communities in agricultural ecosystems. The discovery of the main functional genes related to nutrient cycling, such as the involvement of nitrogen, carbon, and phosphorus metabolism in the soil, which contributes to the meaningfulness of microbiomics in preserving soil productivity and fertility. The differences in the presence between organic and conventional agricultural systems are also observed and further support the effects of agriculture on microbial diversity and functional ability. Through combination of microbial functional profiling and environmental parameters, this study has provided useful insights in the formulation of sustainable farming strategies. In addition, the results will enable the development of micro biome-based agricultural innovations, including biofertilizer and precision soil management devices, to create agriculture that is more environmentally resilient and more productive.

REFERENCES

1. Banerjee, S., Schlaeppi, K., & Van Der Heijden, M. G. (2018). Keystone taxa as drivers of microbiome structure and functioning. *Nature Reviews Microbiology*, *16*(9), 567-576.
2. Chen, M., Ma, Y., Wu, S., Zheng, X., Kang, H., Sang, J., & Bao, Y. (2021). Genome Warehouse: a public repository housing genome-scale data. *Genomics, proteomics & bioinformatics*, *19*(4), 584-589.
3. Cloutier, M., Alcaide, T., Duiker, S., & Bruns, M. A. (2023). Tillage intensity and plant rhizosphere selection shape bacterial-archaeal assemblage diversity and nitrogen cycling genes. *Soil and Tillage Research*, *225*, 105525.
4. Delgado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-González, A., Eldridge, D. J., Bardgett, R. D., & Fierer, N. (2018). A global atlas of the dominant bacteria found in soil. *Science*, *359*(6373), 320-325.
5. Fierer, N. (2017). Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature reviews microbiology*, *15*(10), 579-590.
6. Franzosa, E. A., McIver, L. J., Rahnavard, G., Thompson, L. R., Schirmer, M., Weingart, G., & Huttenhower, C. (2018). Species-level functional profiling of metagenomes and metatranscriptomes. *Nature methods*, *15*(11), 962-968.
7. Hu, X., Gu, H., Liu, J., Wei, D., Zhu, P., Cui, X. A., & Wang, G. (2022). Metagenomics reveals divergent functional profiles of soil carbon and nitrogen cycling under long-term addition of chemical and organic fertilizers in the black soil region. *Geoderma*, *418*, 115846.
8. Jansson, J. K., & Hofmockel, K. S. (2020). Soil microbiomes and climate change. *Nature Reviews Microbiology*, *18*(1), 35-46.
9. Knight, R., Vrbanac, A., Taylor, B. C., Aksenov, A., Callewaert, C., Debelius, J., & Dorrestein, P. C. (2018). Best practices for analysing microbiomes. *Nature Reviews Microbiology*, *16*(7), 410-422.
10. Li, Y., Song, D., Dang, P., Wei, L., Qin, X., & Siddique, K. H. (2020). Combined ditch buried straw return technology in a ridge–furrow plastic film mulch system: Implications for crop yield and soil organic matter dynamics. *Soil and Tillage Research*, *199*, 104596.
11. Nayfach, S., & Pollard, K. S. (2016). Toward accurate and quantitative comparative metagenomics. *Cell*, *166*(5), 1103-1116.
12. Pittelkow, C. M., Liang, X., Linquist, B. A., Van Groenigen, K. J., Lee, J., Lundy, M. E., ... & Van Kessel, C. (2015). Productivity limits and potentials of the principles of conservation agriculture. *Nature*, *517*(7534), 365-368.
13. Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., & Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nature biotechnology*, *35*(9), 833-844.
14. Rodríguez del Río, Á. Scheu, S., & Rillig, M. C. (2025). Soil microbial responses to multiple global change factors as assessed by metagenomics. *Nature Communications*, *16*(1), 5058.
15. Singh, B. K., Trivedi, P., Egidi, E., Macdonald, C. A., & Delgado-Baquerizo, M. (2020). Crop microbiome and sustainable agriculture. *Nature Reviews Microbiology*, *18*(11), 601-602.
16. Thomas, T., Gilbert, J., & Meyer, F. (2012). Metagenomics-a guide from sampling to data analysis. *Microbial informatics and experimentation*, *2*(1), 3.