Metagenomic Analysis of Rhizospheric Microbiome of Tomato by Sequence Analysis of 16S rRNA V3–V4 Hypervariable Region

ABHISHEK KUMAR[1], SHARDENDU[2]

[1, 2] Laboratory of Environment & Biotechnology, Patna Science College, Patna University, Bihar

Abstract*— Rhizospheric microbial diversity plays an important role in maintaining the quality of soil. The diversity of these microbes depends on various factors like nutrients status, pH, Soil temperature, moisture content and the physiochemical properties of soil. In present study, we analyze the rhizospheric bacterial diversity of Solanum lycopersicum from riparian ecosystem Ganga river,Digha(25.6503• N, 85.0974•) Patna of Bihar, India using the shotgun metagenomics. Whole DNA was extracted (commercially available kits such as QIAGEN, ZYMO RESEARCH, ThermoFisher) from soil collected from rhizosphere of tomato.The sequence analysis of 16S rRNA V3–V4 (Primer-16sF:- 5' AGAGTTTGATGMTGGCTCAG3', 16sR:- 5' TTACCGCGGCMGCSGGCAC3') hypervariable region with Illumina MiSeq platform showed 280000 OTUs reads with GC content 57.5 %. The major phyla are detected as Actinobacteria (52.9%), Proteobacteria (25.32%), Firmicutes (14.62%),Bacteroidetes (1.83%) and Acidobacteria(1.47%).At genus level Kocuria (43%) is most dominant whereas at order and family level Micrococcales (34%) and Microcoleaceae (32%) respectively. The result of this study shows the diversity of rhizospheric bacteria of tomato plants in riparian ecosystem, identifying a variety of genera that could exert multiple effects on growth and health of tomato plants. It will further help in comparing the diversity with non riparian Rhizospheric microbial diversity of tomato.

Index Terms*— Metagenomic, Rhizosphere, Tomato, V3–V4, 16s rRNA

I. INTRODUCTION

Bacterial communities play a crucial role in ecosystem functioning and are involved in a range of essential ecological processes. They are responsible for the breakdown and recycling of organic matter , nutrient cycling, and the maintenance of soil structure and fertility . Bacteria are key players in the cycling of nitrogen, carbon, and other nutrients, which is essential for the growth of plants and other organisms . Found in almost every habitat on Earth, including soil, water, and the bodies of plants and animals , bacteria are particularly vital in soil ecosystems. They decompose organic matter, releasing nutrients necessary for plant growth . Additionally, bacteria are involved in plant-microbe interactions, such as the formation of nitrogen-fixing nodules on the roots of leguminous plants .

Riparian ecosystems are critical landscapes providing essential functions such as nutrient cycling, water filtration, and wildlife habitat. These ecosystems form a transition zone between aquatic and terrestrial environments, characterized by high connectivity between the two. Bacteria are crucial in maintaining the functioning of these ecosystems. Factors such as turbulence, resource availability, and edge effects influence riparian zones, making them highly diverse and productive ecosystems with numerous niches for both aquatic and soil species (Peipoch et al., 2015). The hierarchical organization of riparian ecosystems, defined by regions and physiographic identities, affects bacterial species abundance and distribution within epilithic biofilms (Peipoch et al., 2015).

Metagenomics is a powerful method for studying bacterial diversity by revealing the genomic content of entire microbial communities in a sample (Handelsman et al., 1998). This technique has been applied to various ecosystems, including the human gut, soil and aquatic environments. It enables the identification of novel bacterial species, assessment of microbial functional capabilities, and prediction of environmental impacts on microbial communities (Franzosa et al., 2015). The process involves sample collection, DNA extraction, sequencing, and bioinformatics analysis. Proper sample collection and optimized DNA extraction are crucial for obtaining high-quality sequencing data.

II. LITRATURE REVIEW

Tomato (Solanum lycopersicum) is a widely cultivated fruit-bearing plant native to western South America, grown globally on approximately 4.8 million hectares (FAO, 2018). As a member of the solanaceous vegetable crops, tomatoes hold significant global importance. Typically grown as annuals due to their frost sensitivity, tomato plants have pinnate leaves with serrated edges, fine hairs, and a distinct odor when crushed. The yellow, star-shaped flowers are borne in clusters. Technically a fruit, tomatoes are commonly treated as vegetables in culinary contexts, with varieties exhibiting a range of sizes, shapes, colors, and flavors, including red, yellow, orange, green, and purple. Their adaptability to various climates and high nutritional value have driven increased cultivation. Extensively researched for genetics, genomics, and breeding, tomatoes are a model for studying fruit quality, stress tolerance, and physiological traits.

Metagenomics studies the genetic and physiological traits of microbial populations without traditional culturing. By isolating DNA directly from environmental samples, it accesses uncultured organisms. Early advancements include cloning DNA into phage vectors and creating metagenomic libraries from mixed samples and seawater, revealing uncultured archaeons (De Long et al., 1989). Soil DNA posed challenges due to extraction difficulties but was eventually successful. Traditional microbiology focuses on culturable microorganisms, though only about 1% can be cultured in vitro (Amann et al., 1995). Molecular methods now identify both cultured and previously uncultured microbes.

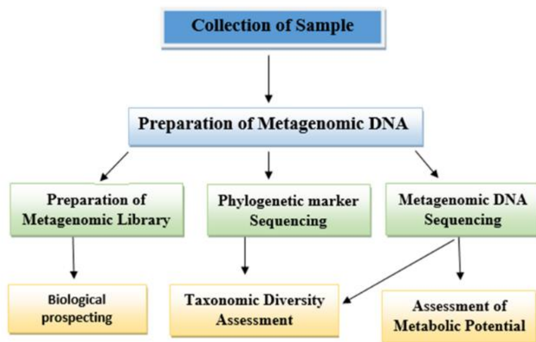


Figure 1: - Flow chart of Metagenomic process

These upgraded tools and platforms reflect the ongoing advancements in Metagenomic analysis, empowering researchers to delve deeper into microbial ecosystems and their roles in various environments.

III. METHODOLOGY

The extraction of DNA from soil samples for metagenomic analysis starts with the careful collection of soil using sterile tools and containers to prevent contamination. Samples are typically taken from the top 5-10 cm of soil and stored at appropriate temperatures to preserve microbial DNA integrity. The collected soil is homogenized using methods such as a mortar and pestle with liquid nitrogen or a bead beater to ensure an even distribution of microorganisms. Microbial cells are lysed to release their DNA by adding a lysis buffer with detergents and sometimes Proteinase K, followed by mechanical disruption. The resulting lysate is purified through a series of steps, starting with centrifugation to remove soil debris. The supernatant, containing the DNA, undergoes phenol:chloroform:isoamyl alcohol extraction for further purification and is precipitated using isopropanol or ethanol. The DNA pellet is washed with cold ethanol, air-dried, and resuspended in a buffer. The quality and quantity of the DNA are assessed using spectrophotometry and agarose gel electrophoresis, aiming for 260/280 readings of ~1.8 to 2 (Robe et al., 2003; Taberlet et al., 2012; Zhou et al., 1996).

For PCR amplification of the V3-V4 region of the 16S rRNA gene, a high-fidelity TAQ Master Mix and specific primers (16sF: 5’-AGAGTTTGATGMTGGCTCAG-3’ and 16sR: 5’-TTACCGCGGCMGCSGGCAC-3’) are used. The PCR mix includes high-fidelity DNA polymerase, 0.5 mM dNTPs, 3.2 mM $MgCl_2$, and a PCR enzyme buffer. The reaction utilizes 40 ng of extracted DNA and 10 pM of each primer, with thermal cycling conditions involving initial denaturation at 95°C, followed by 25 cycles of 95°C for 15 seconds, 60°C for 15 seconds, and 72°C for 2 minutes, concluding with a final extension at 72°C for 10 minutes.

The amplified products are purified with Ampure beads, and an additional PCR with Illumina barcoded

adapters is performed to prepare sequencing libraries. These libraries are purified, quantified using a Qubit dsDNA High Sensitivity assay kit, and sequenced using an Illumina MiSeq with a 2x300PE v3 sequencing kit. Bioinformatics analysis of the raw data includes quality control with FASTQC and MULTIQC, trimming with TRIMGALORE, and processing through QIIME, MOTHUR, KRAKEN, or BRACKEN for tasks such as read merging, chimera removal, OTU abundance calculation, and taxonomic classification using databases like SILVA, GREENGENES, and NCBI. This workflow ensures accurate identification and analysis of microbial communities, aiding in pathogen identification and understanding microbial diversity.

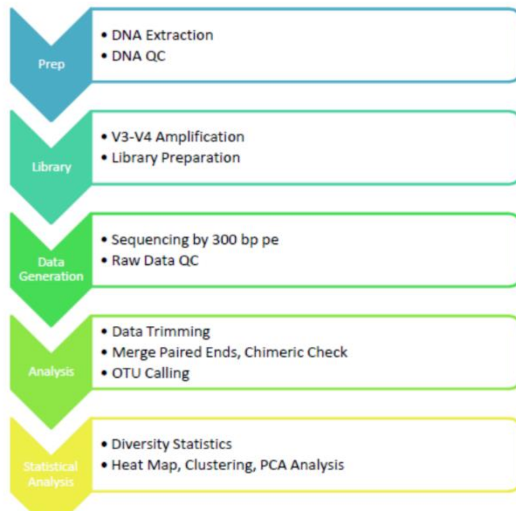


Figure 2:- Next Generation Sequencing Protocol in form of flow chart

IV. RESULT & DISCUSSION

Metagenomic studies often focus on analyzing the prokaryotic 16S ribosomal RNA gene (16S rRNA), which spans approximately 1,500 base pairs and contains nine variable regions interspersed among conserved ones. These variable regions are essential for phylogenetic classification at the genus or species level within diverse microbial populations and are thus the primary target of sequencing efforts (Woese et al., 1990; J et al., 2007). This approach will facilitate the exploration of bacterial communities in our sample Soil1T16s.

Taxonomical assignment is a crucial step in analyzing microbiome datasets. Initially, it involves filtering out reads not assigned to a phylum from the raw data to ensure analysis quality. QIIME 2's third-party plugins offer interactive, real-time, multi-taxonomical level Krona plots, enhancing visualization. These plots are saved as .html files, accessible via any internet browser. The tool provides overview images of the interactive tool for samples Soil1T16s, aiding in understanding microbial community composition. (Fig 3)

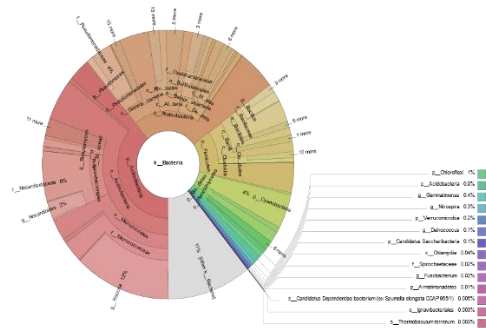
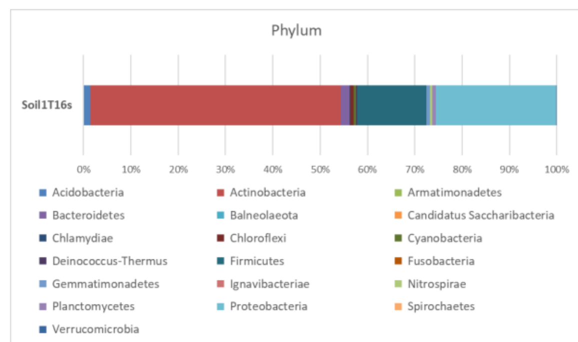


Figure 3: Krona plot of sample Soil1T16s

The analysis of 16S ribosomal RNA sequencing data from soil samples unveils the intricate microbial communities inhabiting these ecosystems. Actinobacteria emerges as the dominant phylum in Soil1T16s, with markedly higher counts in Soil1T16s (10,707 reads). Conversely, Proteobacteria prevails as the most abundant phylum in Soil1T16s exhibits 5,126 reads for the same phylum. This observation underscores the prevalence of Proteobacteria in soil ecosystems, particularly away from water bodies, indicating their crucial role in organic matter decomposition and nutrient cycling processes in agricultural land (Hilal et al., 2021; Finley et al., 2016; Cobaugh et al., 2015).

Soil1T16s: Top 5 abundant Phylum

1. Actinobacteria (10,707)
2. Proteobacteria (5,126)
3. Firmicutes (2,961)
4. Acidobacteria (299)
5. Bacteroidetes (371)



CONCLUSION

In our study, Soil1T16s was collected from the tomato rhizosphere in the Gangetic delta region,within the same tomato rhizosphere area. Analysis of soil samples near the Ganga River indicated a higher abundance of Actinobacteria, particularly evident in the Soil1T16s dataset. Furthermore, Phylum Firmicutes exhibited a notable presence, especially in Soil1T16s, with a total of 6,106 reads. Firmicutes are commonly observed as dominant phyla in cultivation fields, consistent with the regular cultivation area represented by Soil1T16s (Sharmin et al., 2013). Conversely, Proteobacteria emerged as the dominant phylum in Soil1T16s showed 5,126 reads. This underscores the thriving of Proteobacteria in soil environments distant from water bodies, highlighting their crucial role in organic matter decomposition and nutrient cycling processes in agricultural land (Hilal et al., 2021; Finley et al., 2016; Cobaugh et al., 2015).

REFERENCES

[1] Amann, R.I., Ludwig, W., and Schleifer, K.H. 1995. Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol. Rev.* 59, 143–169.

[2] Cobaugh, K. L., Schaeffer, S. M., & DeBruyn, J. M. (2015). Functional and structural succession of soil microbial communities below decomposing human cadavers. *PloS one*, *10*(6), e0130201.

[3] De Long, E. F., et al. (1989). Characterization of uncultivated prokaryotes: isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon. Journal of bacteriology, 178(3), 591-599.

[4] Finley, S. J., Pechal, J. L., Benbow, M. E., Robertson, B. K., & Javan, G. T. (2016). Microbial signatures of cadaver gravesoil during decomposition. *Microbial ecology*, *71*, 524-529.

[5] Franzosa, E. A., Hsu, T., Sirota-Madi, A., Shafquat, A., Abu-Ali, G., Morgan, X. C., ...&Huttenhower, C. (2015). Sequencing and beyond: integrating molecular'omics'for microbial community profiling. Nature Reviews Microbiology, 13(6), 360-372.

[6] Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J., & Goodman, R. M. (1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chemistry & biology*, *5*(10), R245-R249.

[7] Hilal, M. G., Yu, Q., Zhou, R., Wang, Y., Feng, T., Li, X., & Li, H. (2021). Exploring microbial communities, assessment methodologies and applications of animal's carcass decomposition: a review. *FEMS Microbiology Ecology*, *97*(8), fiab098.

[8] Peipoch, M., Foulquier, A., Bornette, G., & Céréghino, R. (2015). The hierarchical organization of riparian ecosystems: A new conceptual framework. *Freshwater Science*, 34(1), 262-272.

[9] Robe, P., Nalin, R., Capellano, C., Vogel, T. M., &Simonet, P. (2003). Extraction of DNA from soil. European Journal of Soil Biology, 39(4), 183-190. DOI: 10.1016/S1164-5563(03)00033-5.

[10] Sharmin, F., Wakelin, S., Huygens, F., & Hargreaves, M. (2013). Firmicutes dominate the bacterial taxa within sugar-cane processing plants. *Scientific reports*, *3*(1), 3107.

[11] Taberlet, P., Coissac, E., Hajibabaei, M., &Rieseberg, L. H. (2012). Environmental DNA. Molecular Ecology, 21(8), 1789-1793. DOI: 10.1111/j.1365-294X.2012.05542.x.

[12] Woese, C. R., Kandler, O., &Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences*, *87*(12), 4576-4579.

[13] Zhou, J., Bruns, M. A., &Tiedje, J. M. (1996). DNA recovery from soils of diverse composition. Applied and Environmental Microbiology, 62(2), 316-322. PMID: 8593015.