Turkish Computational and Theoretical Chemistry

Turkish Comp Theo Chem (TC&TC)

Volume(Issue): 9(4) – Year: 2025 – Pages: 124-131

e-ISSN: 2602-3237

https://doi.org/10.33435/tcandtc.1714834

Received: 05.05.2025 ***Accepted:*** 11.06.2025 *Research Article*

Metagenomic analysis' rhizomicrobiome' Prunus mahaleb based on 16S RNA gene sequencing on the Illumina MiSeq

Serap Çetinkaya[a], Aysun Aksu[a], Muhammed Safa Çelik[a], Hüseyin Fatih Çetinkaya[b], Burak Tüzün[c,1]

[a]Department of Molecular Biology and Genetics, Science Faculty, Sivas Cumhuriyet University, Turkey.
[b]Department of Environmental Engineering, Faculty of Engineering, Sivas Cumhuriyet University, Sivas, Turkey
[c]Plant and Animal Production Department, Technical Sciences Vocational School of Sivas, Sivas Cumhuriyet University, Sivas, Turkey.

Abstract: The study covered for the first time the investigation of the microbiome of Prunus mahaleb rhizosphere. The methodology included both metagenomic and bioinformatic means, and used the bidirectional sequencing of V3 region of 16S rRNA gene with oligo-primers universal to both Bacteria and Archaea. Results of Illumina MiSeq sequencing indicated the presence of 49 phyla, 104 classes, 242 orders, 353 families and 761 genera. The commonest genus was, not surprisingly, Pseudomonas. This genus was closely followed by uncultured genera and this finding was considered to have very meaningful implications as to the functional aspects of a biome and the importance of it was discussed in the text. A thorough insight into the rhizomicrobiome is therefore envisaged to facilitate the development of microbial fertilizers to improve plant performance and productivity impacting eventually food security.

Keywords: Bioinformatics; metagenomics; Illumina MiSeq sequencing; rhizosphere of Prunus mahaleb; V3 region of 16S rRNA

1. Introduction

Members of Bacteria live in almost every environment on earth and more than ninety percent of them await to be discovered [1] by metagenomics [2,3]. Metagenomics aims to investigate genomes or functional properties of uncultured microorganisms [4]. Functional metagenomics screens metagenomic libraries for a specific phenotype, such as halotolerance, antimicrobial- or enzymatic activity [5]. Identification approaches, on the other hand, use clones of highly conserved 16S rRNA- or other genes, or genomic DNA libraries to build whole genomes [6-8]. Hence, metagenomic technology serves as a useful in vitro tool to prepare environmental DNA samples and manipulate them for the objectives explained above [9].

Continents most probably harbor the largest microbial diversity. Their extreme spatial and chemical heterogeneity have been accounted for this diversity. One of the compelling challenges along the soil metagenomics has been the true representation of the heterogeneity of its microbial communities in evolutionary terms because on average one gram of soil sample homes around more than 4 x 10 8 bacteria [10], which amount to approximately 10,000 diverse genomes [11-14]. These estimates exceed the present annotated bacteria (16,177 species, National Centre for Biotechnology Information, 2005).

Up to a certain point random sequencing helped characterize uncultured bacteria, however, further developments are needed to achieve meaningful scientific and industrial gains [15].

The fruits and seeds of mahaleb (Prunus mahaleb) have been used as medicine throughout history. Today, it is used as a raw material and additive in the paint, food and cosmetic industry as well as

[1] Corresponding Authors

e-mail: theburaktuzun@yahoo.com

medicine. Mahalebs usually bear fruit in two colors, white and black. White mahaleb with white flowers is a deciduous tree of the Rosaceae family [16]. The fatty acid content of its seeds varies according to the color of the mahaleb fruit.

Rhizosphere bacterial flora of Prunus species has not yet been studied by uncultured methods [17]. This study found two genera, Sphingomonas and Methylobacterium in the root soil samples of Prunus that contained twenty-three genera using the sequencing of V3-hypervariable region of 16SrRNA. A similar study has recently been performed on the rhizosphere soil bacteria [18] and it has been concluded that a thorough insight into the rhizosphere microbiome would give rise to the development of microbial fertilizers to improve plant performance.

2. Computational Method

The soil sample was powdered in liquid nitrogen and suspended in PBS (potassium buffer saline). Soil DNA was prepared using HotSHOT method [19]. DNA concentration was measured in a Qubit 2.0 fluorimeter using the Qubit double-chain DNA HS Assay kit and then stored at -20 °C. After quality control of the soil DNA, V3 hypervariable region of 16S rRNA gene was sequenced using universal oligonucleotides specific to Bacteria and Archaea. The primers included bar-code sequences: Forward, 5'-TCGTCGGCAGCGTCAGATGTGTATAAGAG ACAGCCTACGGGNGGCWGCAG; Reverse, 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGA GACAGGACTACHVGGGTATCTA ATCC-3'. Bidirectional library sequencing (2 x 251 b) was performed on the Illumina Miseq platform (CUTAM, Sivas Cumhuriyet University) Forward and reverse reads were matched to raw DNA sequence data in QIIME (Quantitative Insights in Microbial Ecology, join_paired_ends.py script). A table of operational taxonomic units (OTUs) in "biome" format was then created with OTUs aligned against the Greengenes database for 16S rDNA.

3. Results and discussion

Within the scope of the project, it was aimed to reveal the bacterial and archaeal profile of the *Prunus mahaleb* rhizosphere with metagenomic approaches. For this purpose, 16S rDNA V3-V4 gene regions were sequenced. The 1500 base long gene, which is usually called 16S rDNA and provides ribosomal RNA coding, is largely found in all bacteria. Although there are highly conserved regions in the gene, it also contains regions that are variable enough to allow phylogenetic classification of different bacterial species. Sequence differences in 16S rDNA regions between different species increase in proportion to evolutionary distance. For this reason, the 16S rDNA sequence is used as a marker for the identification of different bacterial species in Figure 1.

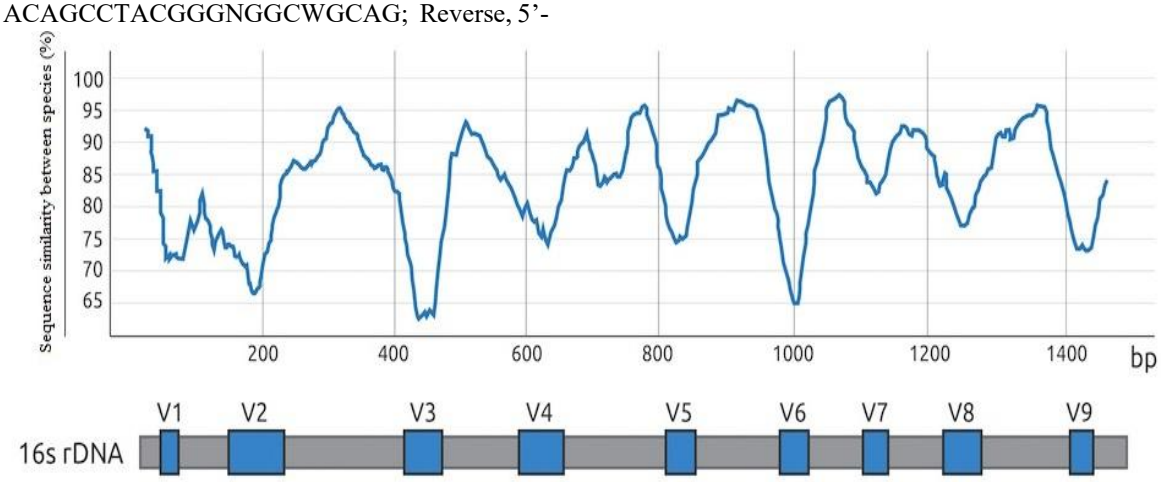

Figure 1. Diversity of the 16S rRNA gene sequence in Bacteria

To date the 16S rDNA sequences of numerous bacterial species obtained from cultured and environmental samples have been extracted by numerous studies. Sequence information has been

stored in various databases including Greengenes, Ribosomal Database Project and SILVA. High-volume sequence data obtained by methods such as next-generation sequencing can be compared with these databases, and the type of DNA fragments detected can be defined proportionally [18, 20-23].

3.1. Bioinformatics Analysis

After obtaining nucleotide information, the FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) program was used in the quality control steps. Quality control work assessed the amount of obtained data, reading quality, GC -and kmer distributions, and checked adapter contaminations. Consequently, reads of poor quality (Phred Score < Q20 and 30 bp window range) were excluded and adapter contaminants and chimeric sequences at the read tips were trimmed using the Genomes OnLine Database (GOLD) and the Trimmomatic (http://www.usadellab.org/cms/?page=trimmomatic) tool. For taxonomic profiling, reads were aligned to target organisms in SILVA (2022) database using the Ribosomal Database Project (RDP) Classifier [24]. OTU groups were then determined. R scripts were used in data reporting, statistical analysis, and data visualization studies. After these procedures, remaining sequence data were assigned to the taxonomic units with the highest similarity at different taxonomic levels. The graph below showed the total number of reads detected in each sample and the (light blue) reads defined to a taxonomic unit up to the genus level. Dark blue reads indicate readings assigned to the taxonomic unit higher than genus level in Figure 2.

The graph below that included minimum, maximum, median and first/third quarterly values of each genus, listed the 20 bacterial genera with the highest incidence in Figure 3.

3.2. Phylum

Metagenomic analysis indicated a total of 49 phyla and the commonest 20 were shown below in Figure 4.

3.3. Class

A total of 104 classes were identified. The top 20 classes with the highest prevalence were shown in Figure 5.

The number of orders assigned was 242 and again the most prominent 20 were presented in Figure 6.

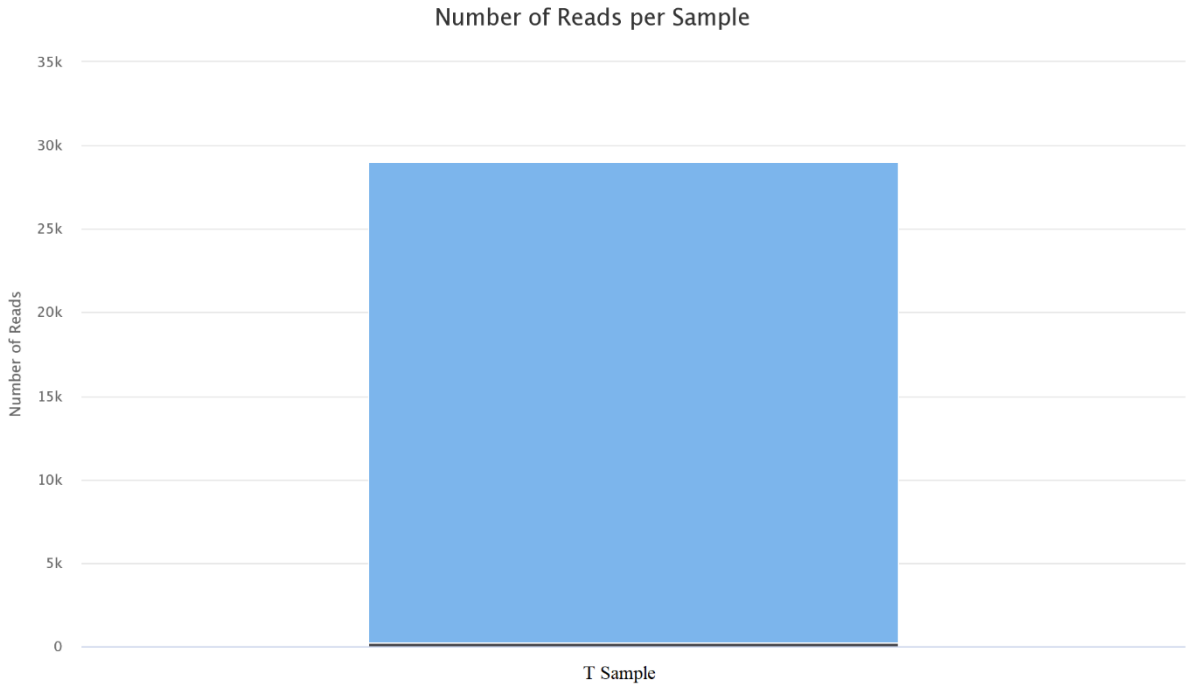

Figure 2. Number of reads

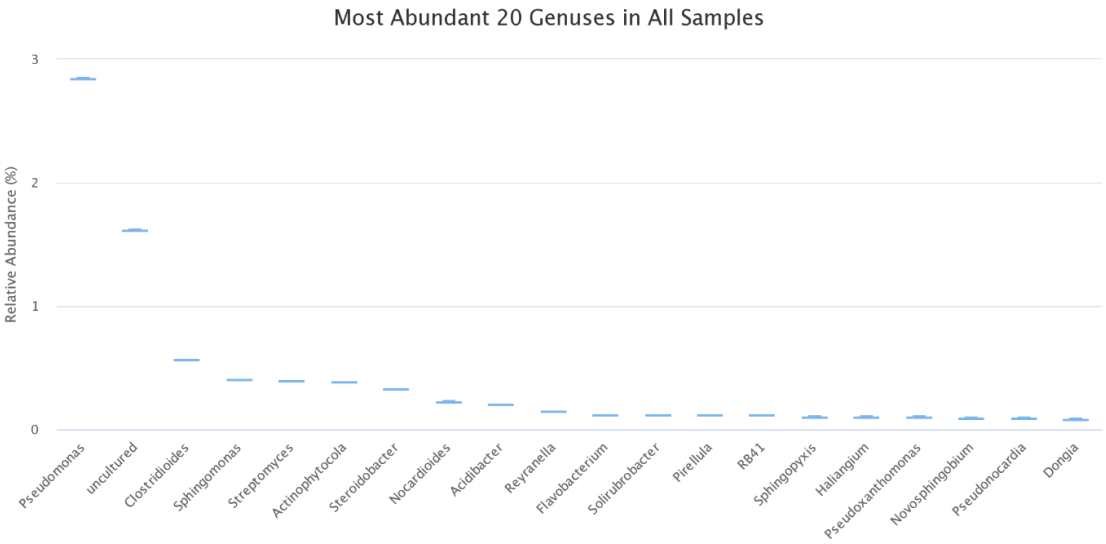

Figure 3. Distribution and abundance of 20 predominant genera

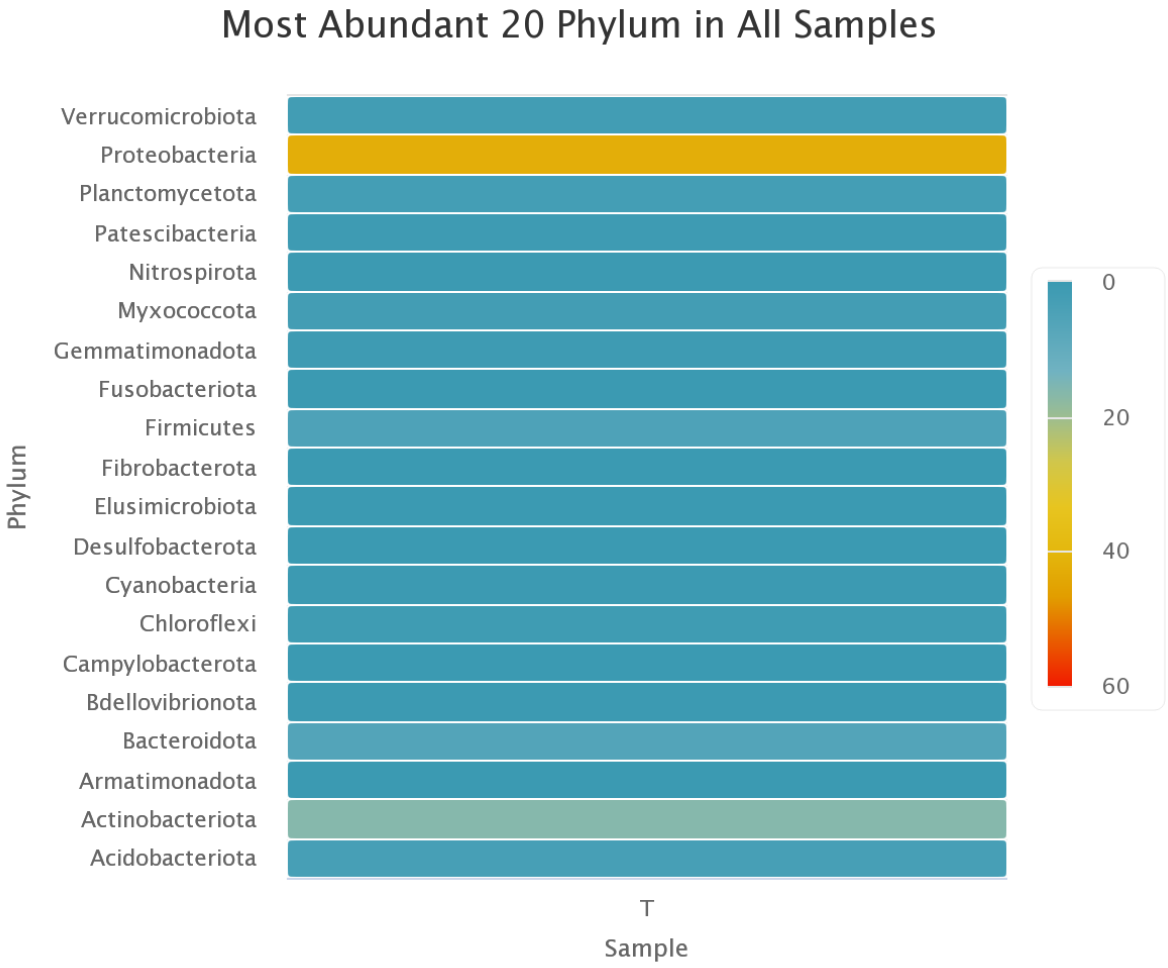

Figure 4. Phyla distribution and abundance

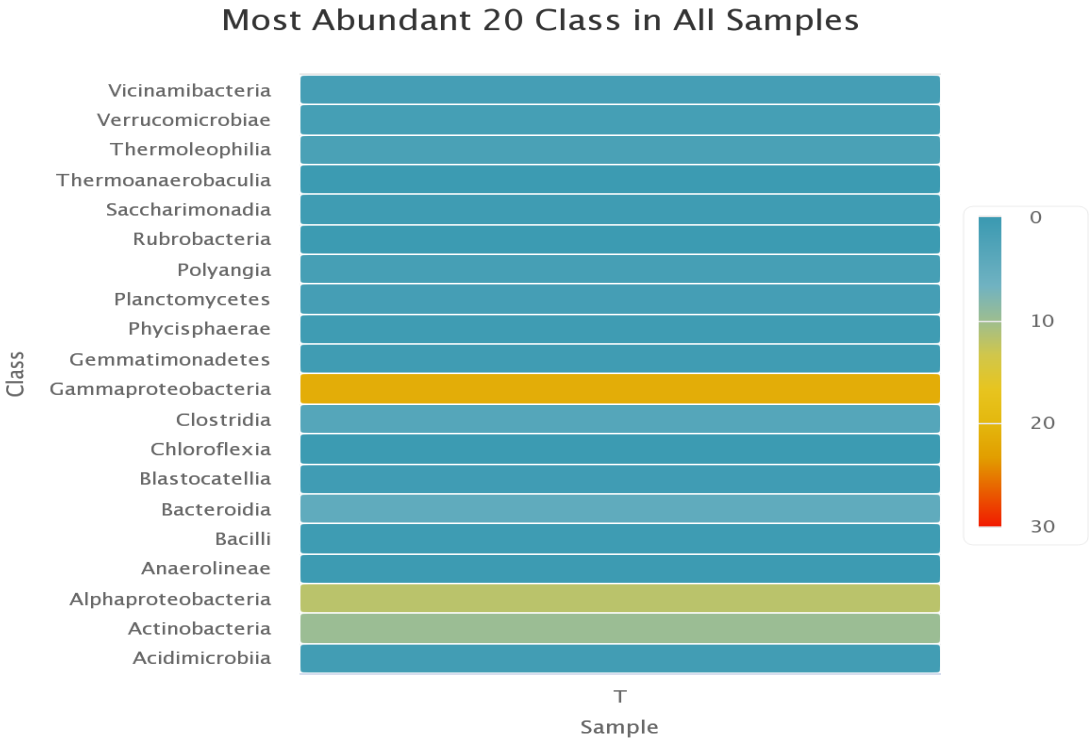

Figure 5. Class distribution and abundance

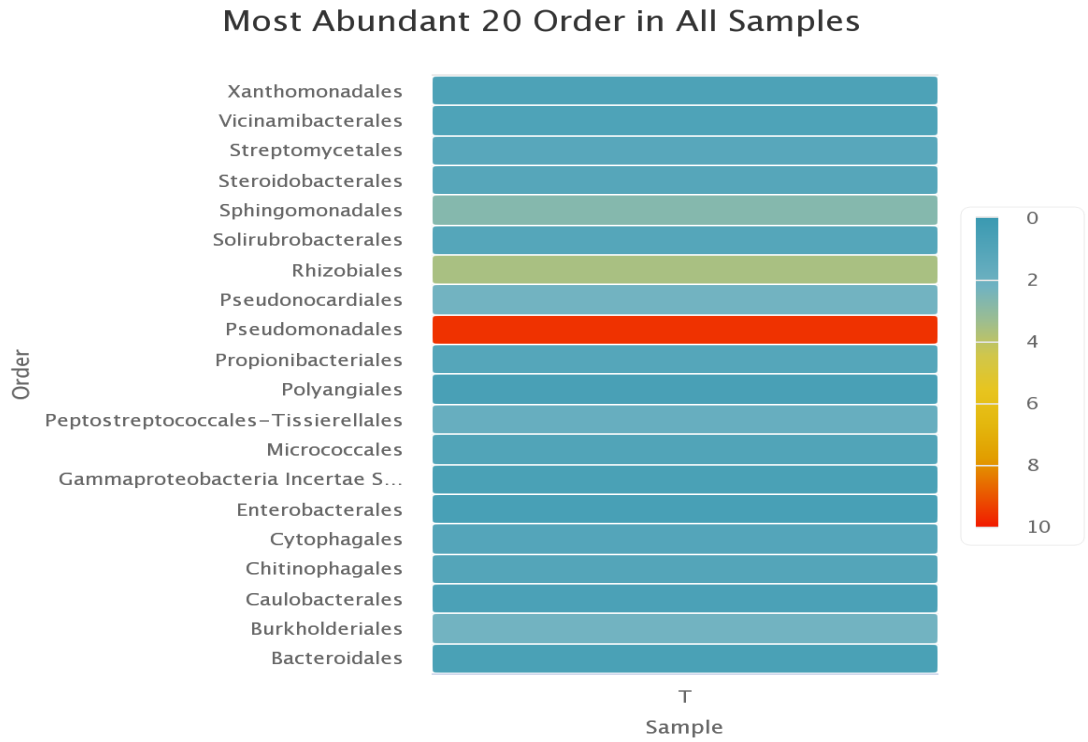

Figure 6. Order distribution and abundance

3.4. **Family**

The sequence data demonstrated the presence of 353 families, and 20 predominant members were shown in Figure 7.

3.5. Genus

The work demonstrated the existence of 761 genera and prevailing members were indicated below in Figure 8.

3.6. Diversity

The diversity index (also called phylogenetic indices or phylogenetic metrics) is a quantitative measure that reflects how many different organisms are present in a dataset (a community) as well as phylogenetic relationships (co-distribution, species affinity, species richness). Within the scope of the project, Shannon index [25] was used for diversity calculation and Shannon, Simpson and Inverse Simpson indices of each species and the number of species were specified in Table 1. The R:vegan [26] package was used to calculate the indices.

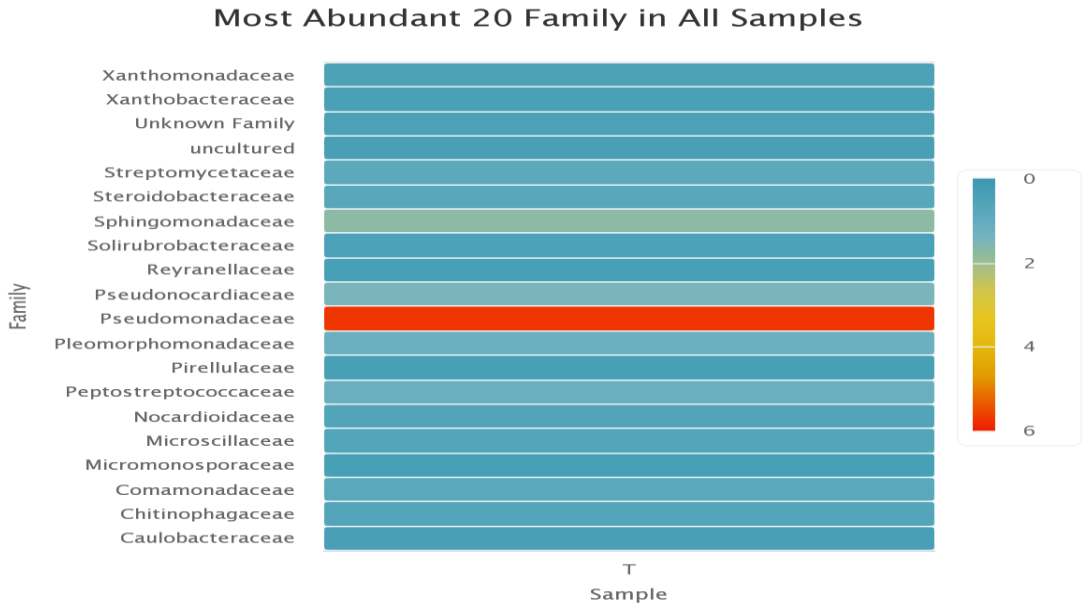

Figure 7. Family distribution and abundance

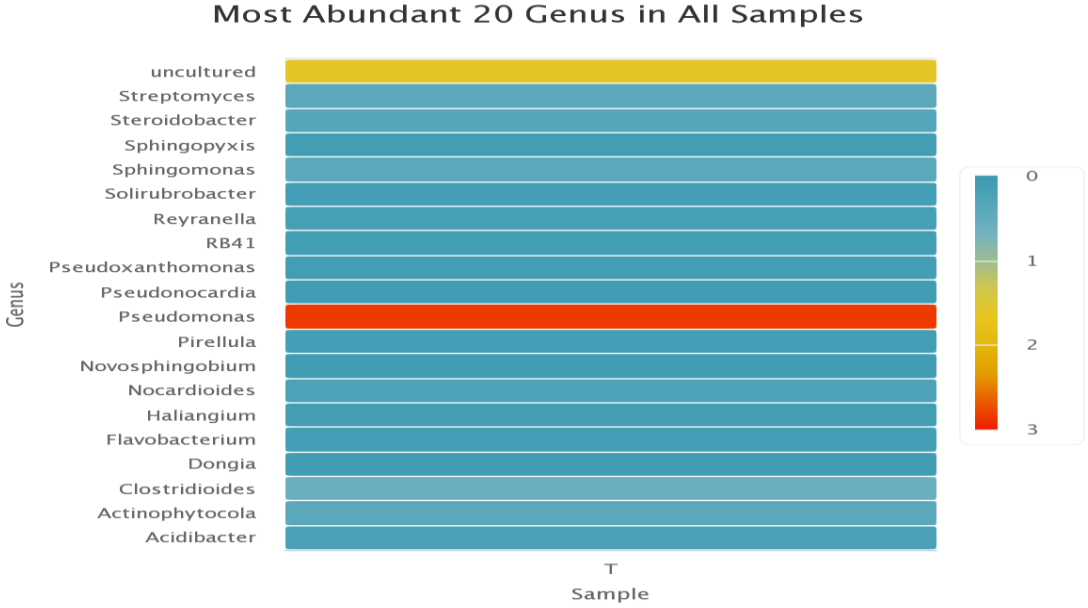

Figure 8. Predominant 20 genera

Table 1 Diversity indices

Sample	Shannon Index	Simpson Index	Inverse Simpson Index	Number of Species
T	4,2	0,93	14,3	760

4. Conclusions

It was not surprising that sequence data evaluated at different taxonomic levels produced concordant results as the study used only one identifier, the nucleotide sequence of V3 of 16S rRNA gene. It was also not surprising that the predominant bacteria included the members of Pseudomonas. The most striking point of this research, however, was that uncultured genera constituted the second most predominant genera. This finding could have very important implications as to the propagation of uncultured bacteria that could be made possible using "natural" environments by imitating the Prunus mahaleb rhizosphere conditions, in this case, and then investigate the microbiomes by metagenomics. Thus, further studies could be performed first in the "open" and then in the laboratory using both agricultural- and molecular means, respectively, with the aim to tilt the predominancy toward uncultured genera because abundance of Pseudomonas may not be directly relevant to its functional importance on the plant growth, health, and fertility and a combinatorial interplay might have taken place.

ACKNOWLEDGEMENT

The study was partly supported by Cumhuriyet University Scientific Research Projects Commission.

Declarations

Conflict of interest The authors declare that they have no conflict of interests.

Funding Not applicable

Ethical approval Not applicable.

Informed consent Not applicable.

Research involving human and animal rights This article does not contain any studies involving animals performed by any of the authors.

References

[1] R.I. Amann, B.J. Binder, R.J. Olson, S.W. Chisholm, R. Devereux, & D. Stahl, Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Applied and environmental microbiology, 56(6) (1990) 1919-1925.

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

[3] P. López-García, & D. Moreira, Tracking microbial biodiversity through molecular and genomic ecology. Research in microbiology, 159(1) (2008) 67-73.

[4] E. Gabor, K. Liebeton, F. Niehaus, J. Eck, & P. Lorenz, Updating the metagenomics toolbox. Biotechnology Journal: Healthcare Nutrition Technology, 2(2) (2007) 201-206.

[5] E.A. Dinsdale, R.A. Edwards, D. Hall, F. Angly, M. Breitbart, J.M. Brulc, ... & F. Rohwer, Functional metagenomic profiling of nine biomes. Nature, 452(7187) (2008) 629-632.

[6] K.J. Hoff, M. Tech, T. Lingner, R. Daniel, B. Morgenstern, & P. Meinicke, Gene prediction in metagenomic fragments: a large scale machine learning approach. BMC bioinformatics, 9 (2008) 1-14.

[7] C.S. Riesenfeld, P.D. Schloss, & J. Handelsman, Metagenomics: genomic analysis of microbial communities. Annu. Rev. Genet., 38(1) (2004) 525-552.

[8] R.D. Sleator, C. Shortall, & C. Hill, Metagenomics. Letters in applied microbiology, 47(5) (2008) 361-366.

[9] M. Langer, E.M. Gabor, K. Liebeton, G. Meurer, F. Niehaus, R. Schulze, ... & P. Lorenz, Metagenomics: an inexhaustible access to nature's diversity. Biotechnology Journal: Healthcare Nutrition Technology, 1(7-8) (2006) 815-821.

[10] D.D. Richter, & D. Markewitz, How deep is soil?. BioScience, 45(9) (1995) 600-609.

[11] V. Torsvik, F.L. Daae, R.A. Sandaa, & L. Øvreås, L. Novel techniques for analysing microbial diversity in natural and perturbed environments. Journal of biotechnology, 64(1) (1998) 53-62.

[12] V. Torsvik, & L. Øvreås, Microbial diversity and function in soil: from genes to ecosystems. Current opinion in microbiology, 5(3) (2002) 240-245.

[13] V. Torsvik, R. Sørheim, & J. Goksøyr, Total bacterial diversity in soil and sediment communities—a review. Journal of

Industrial Microbiology and Biotechnology, 17(3-4) (1996) 170-178.

[14] W.F. Doolittle, Phylogenetic classification and the universal tree. Science, 284(5423) (1999). 2124-2128.

[15] R. Daniel, The metagenomics of soil. Nature reviews microbiology, 3(6) (2005) 470-478.

[16] A. Özbey, N. Öncül, Z. Yıldırım, & M. Yıldırım, Mahlep ve mahlep ürünleri. Journal of Agricultural Faculty of Gaziosmanpaşa University (JAFAG), 2011(2) (2011) 153-158.

[17] Y. Jo, J.K. Cho, H. Choi, H. Chu, S. Lian, & W.K. Cho, Bacterial communities in the phylloplane of Prunus species. Journal of Basic Microbiology, 55(4) (2011) (2015) 504-508.

[18] C. Oliveira, E. Shakiba, D. North, M. McGraw, E. Ballard, M. Barrett-D'Amico, ... & Y. Rahmatallah, 16S rRNA gene-based metagenomic analysis of rhizosphere soil bacteria in Arkansas rice crop fields. Agronomy, 12(1) (2022) 222.

[19] G.E. Truett, P. Heeger, R.L. Mynatt, A.A. Truett, J.A. Walker, & M.L. Warman, Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). Biotechniques, 29(1) (2000) 52-54.

[20] F. Durazzi, C. Sala, G. Castellani, G. Manfreda, D. Remondini, & A. De Cesare, Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota. Scientific reports, 11(1) (2021) 3030.

[21] E. Fadeev, M.G. Cardozo-Mino, J.Z. Rapp, C. Bienhold, I. Salter, V. Salman-Carvalho, ... & A. Boetius, Comparison of two 16S rRNA primers (V3–V4 and V4–V5) for studies of arctic microbial communities. Frontiers in microbiology, 12 (2021) 637526.

[22] E.V. Grigorova, N.L. Belkova, U.M. Nemchenko, E.S. Klimenko, A.V. Pogodina, A.I. Romanitsa, ... & L.V. Rychkova, Metasequencing of V3-V4 variable regions of 16S rRNA gene in opportunistic microbiota and gut biocenosis in obese adolescents. Bulletin of Experimental Biology and Medicine, 170(3) (2021) 321-325.

[23] D. Peterson, K.S. Bonham, S. Rowland, C.W. Pattanayak, R. Consortium, & V. Klepac-Ceraj, Comparative analysis of 16S rRNA gene and metagenome sequencing in pediatric gut microbiomes. Frontiers in microbiology, 12 (2021) 670336.

[24] Q. Wang, G.M. Garrity, J.M. Tiedje, & J.R. Cole, Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and environmental microbiology, 73(16) (2007) 5261-5267.

[25] I.F. Spellerberg, & P.J. Fedor, A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the 'Shannon–Wiener'Index. Global ecology and biogeography, 12(3) (2003) 177-179.

[26] J. Oksanen, F.G. Blanchet, R. Kindt, P. Legendre, P.R. Minchin, R.B. O'hara, ... & M.J. Oksanen, Package 'vegan'. Community ecology package, version, 2(9) (2013) 1-295.