No connection between depth, pH and nitrogen and soil microbiome diversity at small-scale

May 8, 2018

Abstract

Those past years have seen a growing interest in microbes. Particularly, soil microbes play a crucial role in contributing to the ecosystem through many processes such as nitrogen fixation, cycling carbon, and nutrient. However, there is a lack of knowledge of what drives microbial diversity. So far, researchers have mainly focused on investigated microbial diversity at large scale. However, small scale can assist in identifying factors that influence microbial diversity in the soil. 30 samples were collected in Central London, the v4 region of 16S rRNA genes was sequenced via Illumina MiSeq and QIIMES was used to characterize soil microbial communities. Microbial diversity was unrelated to 3 metadata (pH, depth, Nitrogen) commonly found as a driver for microbial diversity at alpha and beta diversity level. This suggests that other variables, not usually considered in studies investigating on microbial diversity, could be involved in driving microbial diversity.

I. Introduction

The study of the microbiome has been growing over the past years. Microbes constitute over 90% of the planet and have been shown to play a crucial role in the human health, disease (Young *et al.,* 2017) or the ecosystem. (Finlay *et al.,*1997).

Among microbes' habitats, the soil is highly rich in diversity, with thousands of different microbial species in 1g of soil (Torsvik *et al.,*2002). Particularly, soil microbiomes interactions with the environment are important in understanding the evolution of the ecosystem. These microbes play a critical part in ecosystem functions including many processes such as cycling carbon and nutrient, nitrogen fixation and plant growth (Finlay *et al.,*1997).

Consequently, the study of microbes grew in popularity along with advances to study the microbiome. Among these, metagenomics, the study of the microbiome through sequencing and analyzing of the DNA in their environment, is the main method used to study microbes. This was allowed due to considerable technical improvements that have led to the emergence of new sequencing technologies allowing rapid and cost-effective DNA sequencing. The recent development of high-throughput sequencing techniques such as next-generation sequencing (NGS) has allowed a deeper understanding of the microbial diversity (Shendure *et al.,*2008).

Generally, NGS has greatly simplified the sequencing of amplicons of the 16S ribosomal gene which is now commonly used to explore microbial communities (Brodie *et al.,*2002, Harris *et al.,*2004). Among these, 16S rRNA gene Illumina sequencing is an effective and powerful tool to examine microbial communities and diversities in soils (Caporaso *et al.,*2011). The 16S rRNA gene is the best-known genetic marker particularly because it is specific to prokaryotes and consists of both very well conserved regions and hypervariable regions. Thus, as hypervariable regions have sequence diversity among prokaryotes species this

sequencing can generate advanced phylogenetic studies in the field of the analysis of the evolution of species.

This has further facilitated phylogenetic analysis of soil microbes, allowing a deeper understanding of relationships between soils microbiome and environment. Among these, the Earth Microbiome Project (EMP), seeks to characterize microbes taxonomic, functions and diversities in the planet. Understanding microbes' presence and function in an environment are crucial to explaining ecosystem function. Microbial communities vary according to some environmental factors. (Freedman *et al.*,2015). Especially researchers have emphasized 3 of them which have been shown to be important in driving microbial diversities in samples: depth, pH (Fierer *et al.*,2006), and nitrogen (Freedman *et al.*,2015).

However, all the factors involved in driving microbial diversities are still unknown. So far, researchers have mainly focused on investigated microbial diversity in soil samples at large scale. (Hendershot *et al.*,2017). Yet, looking for microbial diversity at small scale would allow a deeper understanding of the impact of factors driving microbial diversity in soil.

Thus, we have investigated the diversity of 30 soil samples, collected close from each other on the same day, in Central London location (Figure 1). Soil DNA was isolated and microbiome sequencing was made using Illumina-based sequencing of the v4 hypervariable region of the 16S rRNA gene. We examined (1) if our results correspond to other soil communities and (2) if pH, depth, and Nitrogen could explain microbial diversity.

Figure 1: Geographical localization of the 29 samples visualized through google maps. Sample 515rcbc12 (purple), has been collected away from the 28 other samples (yellow). (A) All 28 other samples have been collected in Gordon Square and their coordinates are represented in different colors. (B)

II. Materials and Methods

1. Soil Samples DNA isolation

30 soil samples were collected on the 17th of October 2017 at different sites in the Central London (Figure 1). The temperature, moisture, footfall, depth of each sample

was pointed out and metadata such as pH, Nitrogen, Potassium, Phosphorus was determined using the HI3895 Soil Testing Kit. DNA was extracted from 2g of soil samples using PowerSoil® DNA Isolation Kit following manufacturers protocol. The samples were visualized, checking for quality and size of DNA, by gel electrophoresis on 1% agarose and ethidium bromide staining

2. PCR amplification of 16S rDNA v4 region

Polymerase chain reaction (PCR) was done by amplification of the V4 hypervariable region of the 16S rRNA gene for each sample using universal primers 515fB (5'-AATGATACGGCGACCACCGAGATCTACACGCT ID TATGGTAATT GT GTGYCAGCMGCCGCGGTAA-3') and 806rB (5'-CAAGCAGAAGACGGCATACGAGAT AGTCAGTCAG CC GGACTACNVGGGTWTCTAAT-3'). Each sample 515fB PCR primer was designed with 12-base ID barcoded Illumina adapter Golay barcodes allowing simultaneous sequencing and identification of multiples samples. For each sample, amplicons were generated in 3 PCRs replicate of a final volume of 25 ¦ L containing 5 pmol of each primer and using Bioline BioxMix kit following manufacturers protocol. PCR reactions were carried out in a thermocycler under following conditions: denaturation for 1 minute at 94°C, followed by 30 cycles of 15 seconds at 94°C, 45 seconds at 50°C and 30 seconds at 72°C. These steps are then followed by an extension step for 5 minutes at 72°C and are held at 4°C. Quality and size of amplicons were checked by gel electrophoresis on 1.5% agarose gel and purified using the QIAquick PCR Purification Kit following manufacturers protocol.

3. Quantification and Sequencing

Samples have been quantified using fluorescence assay using HiRange dsDNA assay kit following manufacturers protocol. The fluorescence reading of each sample was given in DNA concentration based on the machine standard curve. Then, samples were diluted and pooled to an equal concentration of 20µl aliquot of 1nM. Amplification and sequencing were realized using Illumina Miseq at UCL Genomics. The output of the sequencing is data containing DNA sequences that were further used for downstream analysis.

4. Microbiome analysis

Microbiome analysis was processed using QIIMES pipeline tool (Kuczynski *et al.,*2012) on Cartesius high-performance computer (HPC). Read 1 and Read 2 were joined for the analysis and sequences were aligned against the Greengenes database and clustered into operational taxonomy units (OTUs) at the 97% similarity level (Poretsky *et al.,*2014). The full data analysis of the 30 samples is available from GitHub repository at: *https://github.com/Ilovemicrobiome/ReportBioc3301*

III. Results

A total of 4479971 OTUs were observed and allocated between the 30 samples. Filtering of OTUs <1000 per sample was realized according to the minimum number commonly observed in 1-2 g of soil (Torsvik et al.,2002). Sample 515rcbc20 was removed from the analysis due to low (608) OTUs observed. The mean OTUs of the samples was 149332 and 515rcbc20 had the highest OTUs (1399063) observed.

Taxonomy results showed bacteria as the most abundant species representing 97.3% of samples; the only other species found was *Crenarchaeota* from Archaea. The most

abundance bacteria phyla were Proteobacteria (29.4%), Acidobacteria (18.3%), Actinobacteria (15%) and Verrucomicrobia (12.4%). The abundance of certain species was variable between samples. The most variables species were Acidobacteria and Verrucomicrobia species, varying from 4.8% and 25.6% and 6.35% to 36.57% respectively (Figure 2).

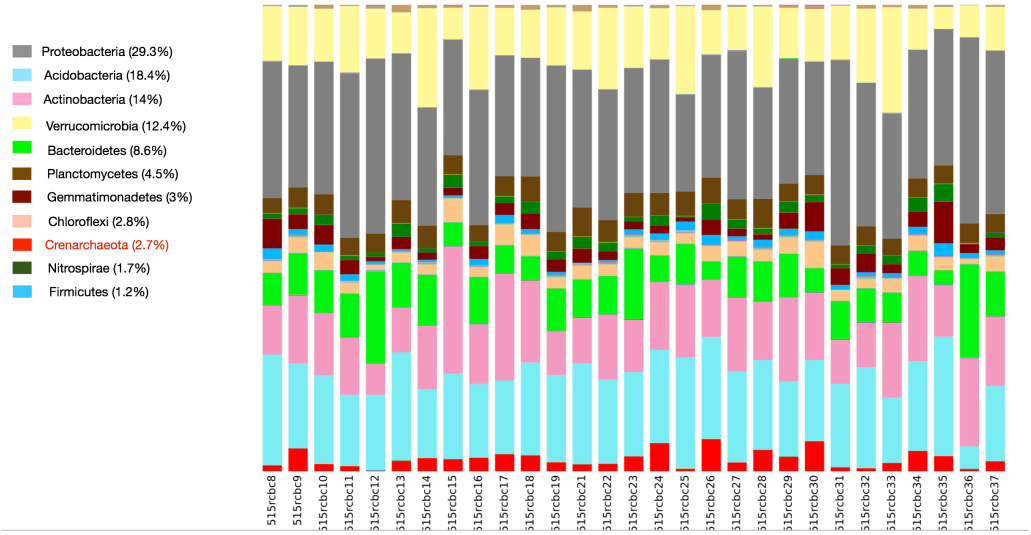

Figure 2: Taxonomy summarized at the phylum level. Total abundances (%) of dominant phyla from the 29 samples are represented based on v4 amplicon sequencing data of the 16S rRNA gene. The phylum level distribution is based against Greengenes database on 97% similarity cluster OTUs. The 10 major dominant phyla included *Proteobacteria, Acidobacteria, Verrucomicrobia, Bacteoidetes, Planctomyecetes, Gemmatimonadetes, Chloroflexi, Crenarchaeotoa, Nitrospirae and Firmicutes*. The only non-bacteria species found is represented in red and was from Archaea species.

We examined whether microbial alpha diversity across samples was related to each other. Alpha diversity was conducted assessing the number of unique OTU for each sample. Chao1 metric was used, which is based on richness and evenness among samples (Beck et al., 2010). Chao1 numbers of each sample were close to each other with a mean of 911 so there were no significant differences between samples at the alpha diversity level. The outlier is 515rcbc36 with the smallest chao1 (Table 1).

To further investigate the consistency of the samples we processed to UniFrac beta diversity which evaluate relationships among samples based on phylogenetic distances between organisms (Lozupone and Knight 2005) and applied in more than 150 publications (Lozupone et al., 2011). UniFrac distances were visualized using PCA plot (Figure 3). Samples are clustered although there is some dissimilarities especially for low abundance species. 515rcbc12 and 515rcbc36 were the most distant from other samples collected (Figure 3) which was expected for 515rcbc12, as it is the only one that was not collected in Gordon Square.

Sample	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**21**	**22**	**23**
Chao1	917	971	892	937	953	824	911	882	1007	944	909	1051	837	905	996
Sample	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	
Chao1	991	842	810	919	899	1027	912	940	818	945	858	826	747	948	

Table 1: Chao1 values for samples at 97% similarities. The outlier is 515rcbc36 is highlighted in orange for 690 sequences/sample

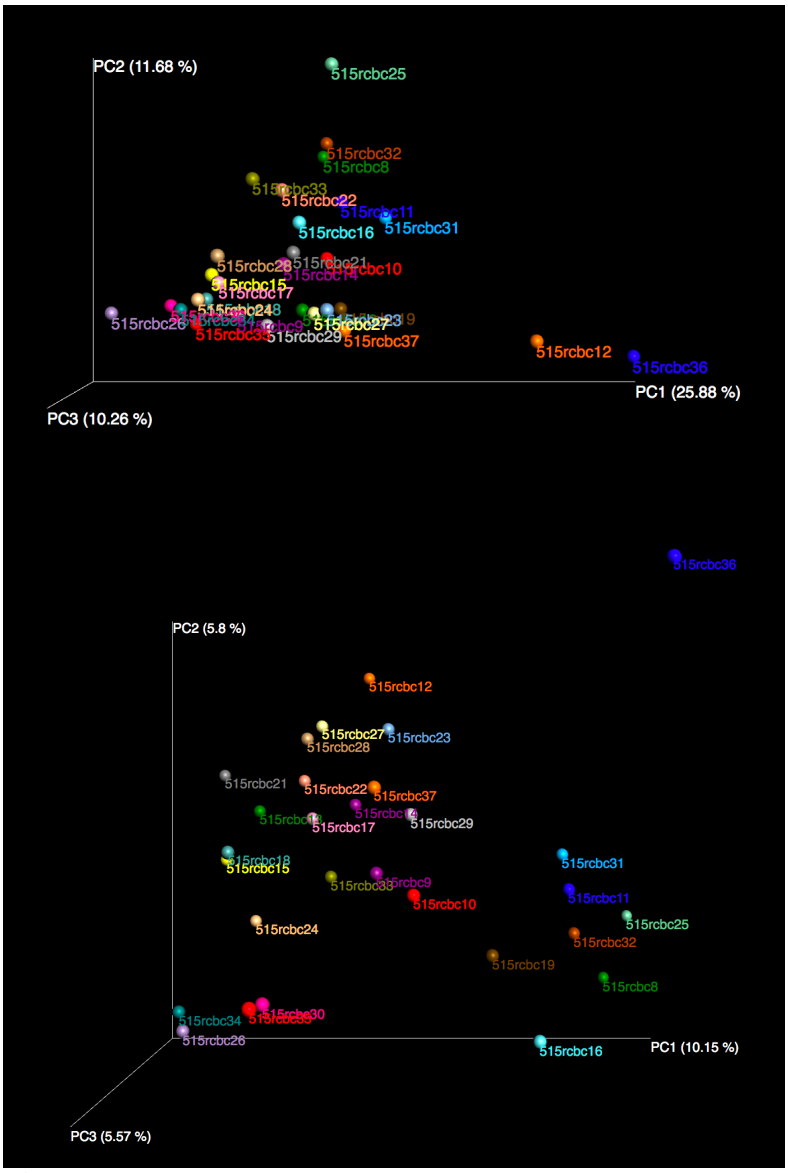

Figure 3: PCA representation of beta diversity of both Weighted (top) and Unweighted (bottom), the first one being useful in examining differences in community structure and the latter more sensitive to low-abundance differences. There is a cluster of microbial community. The outlier are 515rcbc12 and 515rcbc36.

To further examine and interpret the diversity found in samples we have applied alpha and beta analysis to 3 different metadata commonly found as a driver for diversity: depth, pH, and nitrogen. Pearson bootstrapped correlative analysis (r) (Abdi et al., 2009) was performed, relating the influence of metadata to the relative abundance of the 4 most abundant species in samples (Table 2). We did not observe strong correlations (R was close to zero) between the metadata and the species abundance. Only *Proteobacteria* abundance seemed to be lightly correlated to nitrogen (R = 0.43).

From the alpha diversity, the shape of the samples curves for different conditions of the metadata showed similar pattern and clustering (Figure 4). Their metadata (Nitrogen, pH, and depth) do not drive the significant difference in richness and evenness of samples.

To further examine this observation between samples, Mantel test (r) analysis of UniFrac was realized, which has been showed to be an efficient method in analyzing variation in beta diversity composition to metadata or distances factor (Legendre et al., 2005). For all 4 metadata, the R values are close to zero showing no correlation between metadata and samples diversity (Table 3). So, these 4 metadata are not a driver for community diversity between samples.

Group	Correlation		
	Depth	pH	Nitrogen
Proteobacteria	-0.034	0.12	0.43
Acidobacteria	0.20	0.087	-0.19
Actinobacteria	-0.18	-0.14	-0.06
Verrucomicrobia	$6.93\ 10^{-4}$	-0.18	-0.29

Table 2: Correlations test (Pearson) results of the 4 most abundant species (*Proteobacteria, Acidobacrteria, Actinobacteria and Verrucomicrobia*) with metada (depth, pH, nitrogen and potassium) estimated using Pearson bootstrapped. Pearson correlation (r) The only small correlation found is in orange, but for all metadata R were close to zero showing that metadata are not a driver on community diversities found between samples.

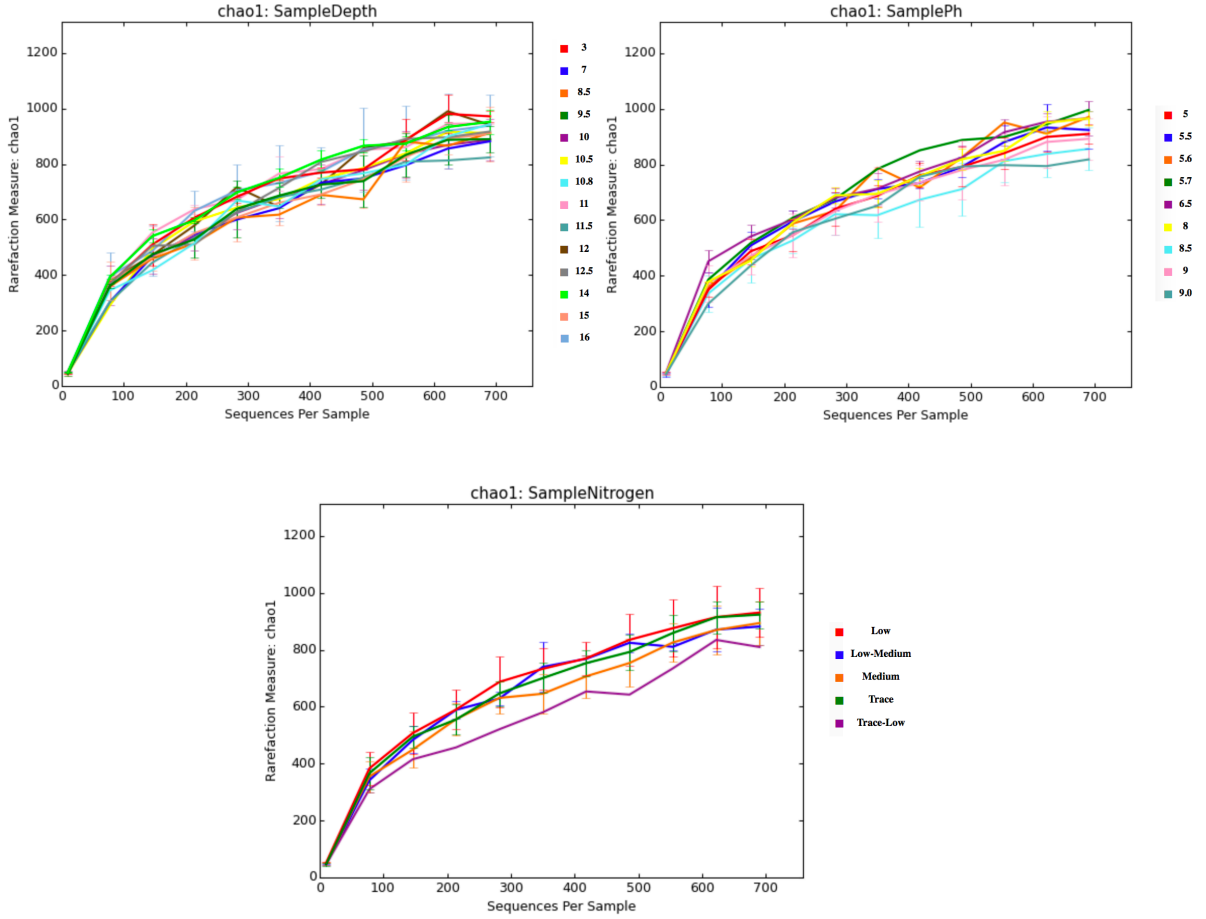

Figure 4: Influence of 3 environmental factors: depth, pH and nitrogen on richness and evenness of samples using chao1 representation.

Metadata	Mantel r statistic	p- value
pH	0.037	0.273
Depth	0.090	0.239
Nitrogen	-0.016	0.446
Potassium	0.13	0.106

Metadata	Mantel r statistic	p- value
pH	0.022	0.320
Depth	0.042	0.346
Nitrogen	0.038	0.299
Potassium	0.054	0.260

Table 3: Correlations test results for weighted (*top*) and unweighted (*bottom*) of different metadata (pH, depth, Nitrogen and Potassium) on community diversity estimated using Mantel tests. Mantel r statistic for all meta were close to zero showing that metadata are not a driver on community diversities found between samples.

IV. Discussion

In this study, we examined the microbial community structure at small-scale in Central London location. The most abundant species in soil samples were *Proteobacteria* which has also been observed in other studies with similar abundance, around 25-40%. In concordance with our observations, both *Acidobacteria* (Jones *et al.,*2009) and *Verrucomicrobia* are commonly found in soil (Chiang et al., 2018; Janssen,2006). We found an average of 12.3% of Verrucomicrobia consistent with other researchers (Janssen,2006; Buckley and Schmidt, 2001). Among Archaea species, we found that *Crenarchaeota* was the only group present (2.7%) in accord with researchers that found Crenarchaeota as the predominant Archaea species in soil (Kemnitz et al., 2007).

Samples were clustered for both alpha and beta diversity highlighting their similarities in both species richness and microbial community. From our observations, samples 515rcbc12 and 515rcbc36 had the most distant community diversity from other samples. 515rcbc12 has been collected away Gordon Square where other samples were collected and therefore was expected to be more diverse from other samples. Sample 515rcbc12 has the highest community diversity (OTUs = 1399063) observed and is distant in community diversity from other samples. Thus, these findings support that the spatial heterogeneity in an important driver for microbial diversity in soils (Ellner *et al.,*2001). 515rcbc36 has also been found to be distant in term of diversity from other samples, especially in low-abundance species, although it has been collected in close proximity with others 27 samples. Such divergence observed especially from samples that are close, such as 515rcbc28 or 515rcbc35, is unlikely (Figure 1). Therefore, we suggested that sample 515rcbc36 was contaminated although further analysis is required to confirmed that hypothesis.

We found dissimilarities in species abundance between different soil samples. These dissimilarities were also observed at community levels, mainly for low-abundance communities (Figure) that have also been observed in soil microbiome researchers (Lauber *et al.,*2009). There was a lack of correlation between diversity and these

metadata pointed out by both alpha and beta diversity (R lied between -0.016 and 0.090) and at the species level. However, many pieces of research related microbial diversity, at both community and species level, to environmental factors such as pH, Nitrogen, and depth (Tripathi *et al.*,2012). Especially, pH was found in many papers to regulate the most soil microbial diversity being in different environments including grasslands, with the highest diversity found at neutral pH even at small scale (Kaiser et al.,2016). A negative correlation between pH and Acidobacteria (Jones *et al.*,2009) and Proteobacteria (Spain *et al.*,2009) are commonly observed but the correlations we observed were low; (R = 0.087) and (R = 0.12) respectively. Some researchers revealed the negative correlation of microbial diversity with Nitrogen concentration (Li et al., 2016). Although, we were able to find a small correlation (R= 0.43) between nitrogen and Proteobacteria abundance which has also been observed in other papers, usually stronger (Dai Z, *et al.*,2018), however, this was not observed at community diversity level (R = -0.016).

We observed a high clustering of samples microbial community within the samples even with high variations in depth (3 to 16 cm) pH (5 to 9), and nitrogen concentration (trace to medium) suggesting that these metadata did not significantly impact microbial diversity. These findings support a recent study conducted at macroscale which found dissimilarities and no correlation between variables, such as the pH and suggested the involvement of other factors such as climatic or historical contexts (Hendershot *et al.*,2017). However, at small scale, these latter factors would not be significant variables to explain diversity. We, therefore, suggested that factors that drive soil microbial diversities differ from places.

Researchers on microbial diversity are now evolving on finding new factors that might be implicated in driving diversity. Some other factors such as potassium or vegetation have been investigated in our study, available in GitHub repository, but were not a significant driver for microbial diversity either. The lack of correlation with these metadata to explain microbial diversity at small scale further interrogates on more factors contributing to microbial diversity. We proposed that human activities could also alter soil microbiome composition which has been reported by other studies explaining diversity for low abundance species (Leff *et al.*,2015). Also, some biotic factors that were not measured in our studies such as eukaryotes microbes or viruses have crucial roles in driving microbial diversity in soil (King *et al.*,2016). Some studies showed that bacterial species presence is driven by the interaction between abiotic environment and other species (Lawrence *et al.*,2012). Thus, using techniques such as whole genome sequencing of eukaryotes and viruses in microbial studies could allow a deeper understanding of factors that drive microbial diversity (Thompson *et al.*,2017)

V. Conclusion

Our analysis is consistent with researchers on soil microbiome and could contribute to the EMP database. However, we could not find any correlation between 3 major drivers commonly observed for diversity: depth, pH, and nitrogen at small-scale. This raised the problem on the existence of more factors, not known or commonly explored on microbial studies, to explain differences in microbial community in the soil. Combining techniques might be an answer to have a deeper understanding of factors that drive microbes' diversity.

VI. References

1. Young VB. The role of the microbiome in human health and disease: An introduction for clinicians. Vol. 356, BMJ (Online). 2017.

2. Finlay BJ, Maberly SC, Cooper JI. Microbial Diversity and Ecosystem Function. Oikos [Internet]. 1997;80(2):209. Available from: http://www.jstor.org/stable/3546587?origin=crossref

3. Torsvik V, Øvreås L. Microbial diversity and function in soil: From genes to ecosystems. Curr Opin Microbiol. 2002;5(3):240–5.

4. Shendure J, Ji H. Next-generation DNA sequencing. Vol. 26, Nature Biotechnology. 2008. p. 1135–45.

5. Brodie E, Edwards S, Clipson N. Bacterial community dynamics across a floristic gradient in a temperate upland grassland ecosystem. Microb Ecol. 2002;44(3):260–70.

6. Harris JK, Kelley ST, Pace NR. New Perspective on Uncultured Bacterial Phylogenetic Division OP11. Appl Environ Microbiol. 2004;70(2):845–9.

7. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci [Internet]. 2011;108(Supplement_1):4516–22. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.1000080107

8. Freedman Z, Zak DR. Soil bacterial communities are shaped by temporal and environmental filtering: Evidence from a long-term chronosequence. Environ Microbiol. 2015;17(9):3208–18.

9. Fierer N, Jackson RB. The diversity and biogeography of soil bacterial communities. Proc Natl Acad Sci U S A. 2006;103(3):626–31.

10. Hendershot JN, Read QD, Henning JA, Sanders NJ, Classen AT. Consistently inconsistent drivers of microbial diversity and abundance at macroecological scales. In: Ecology. 2017. p. 1757–63.

11. Kuczynski J, Stombaugh J, Walters WA, González A, Caporaso JG, Knight R. Using QIIME to analyze 16s rRNA gene sequences from microbial communities. Curr Protoc Microbiol. 2012;(SUPPL.27).

12. Poretsky R, Rodriguez-R LM, Luo C, Tsementzi D, Konstantinidis KT. Strengths and limitations of 16S rRNA gene amplicon sequencing in revealing temporal microbial community dynamics. PLoS One. 2014;9(4).

13. Beck J, Schwanghart W. Comparing measures of species diversity from incomplete inventories: an update. Methods Ecol Evol [Internet]. 2010;1(1):38–44. Available from: http://doi.wiley.com/10.1111/j.2041-210X.2009.00003.x

14. Lozupone C, Knight R. UniFrac: A new phylogenetic method for comparing microbial communities. Appl Environ Microbiol. 2005;71(12):8228–35.

15. Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. UniFrac: An effective distance metric for microbial community comparison. Vol. 5, ISME Journal. 2011. p. 169–72.

16. Abdi H, Edelman B, Valentin D, Dowling WJ. Correlation. In: Experimental Design & Analysis for Psychology [Internet]. 2009. p. 16–38. Available from: http://www.oup.com/uk/orc/bin/9780199299881/ch02.pdf%5Cnhttp://ukcatalogue.oup.com

/product/9780199299881.do?keyword=Experimental+Design+&+Analysis+for+Psychology+&sortby=bestMatches

17. Jones RT, Robeson MS, Lauber CL, Hamady M, Knight R, Fierer N. A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. ISME J. 2009;3(4):442–53.

18. Chiang E, Schmidt ML, Berry MA, Biddanda BA, Burtner A, Johengen TH, Palladino D, Denef VJ. Verrucomicrobia are prevalent in north-temperate freshwater lakes and display class-level preferences between lake habitats. 2018;3

19. Janssen PH. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. Vol. 72, Applied and Environmental Microbiology. 2006. p. 1719–28.

20. Buckley DH, Schmidt TM. Environmental factors influencing the distribution of rRNA from Verrucomicrobia in soil. FEMS Microbiol Ecol. 2001;35(1):105–12.

21. Ellner SP, McCauley E, Kendall BE, Briggs CJ, Hosseini PR, Wood SN, et al. Habitat structure and population persistence in an experimental community. Nature. 2001;412(6846):538–43.

22. Lauber CL, Hamady M, Knight R, Fierer N. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. Appl Environ Microbiol. 2009;75(15):5111–20.

23. Tripathi BM, Kim M, Singh D, Lee-Cruz L, Lai-Hoe A, Ainuddin AN, et al. Tropical Soil Bacterial Communities in Malaysia: PH Dominates in the Equatorial Tropics Too. Microb Ecol. 2012;64(2):474–84.

24. Kaiser K, Wemheuer B, Korolkow V, Wemheuer F, Nacke H, Schöning I, et al. Driving forces of soil bacterial community structure, diversity, and function in temperate grasslands and forests. Sci Rep. 2016;6.

25. Spain AM, Krumholz LR, Elshahed MS. Abundance, composition, diversity and novelty of soil Proteobacteria. ISME J. 2009;3(8):992–1000.

26. Li Q, Song X, Gu H, Gao F. Nitrogen deposition and management practices increase soil microbial biomass carbon but decrease diversity in Moso bamboo plantations. Sci Rep. 2016;6.

27. Dai Z, Su W, Chen H, Barberán A, Zhao H, Yu M, Yu L, Brookes P, Schadt CW, Chang SX, Xu J. Long-term nitrogen fertilization decreases bacterial diversity and favors the growth of Actinobacteria and Proteobacteria in agro-ecosystems across the globe. Ecol Monogr. Global Change Biology. 2018.

28. Leff JW, Jones SE, Prober SM, Barberán A, Borer ET, Firn JL, et al. Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe. Proc Natl Acad Sci [Internet]. 2015;112(35):10967–72. Available from: http://www.pnas.org/lookup/doi/10.1073/pnas.1508382112

29. King AJ, Bueno de Mesquita CP, Schmidt SK, Farrer EC, Suding KN. Incorporating biotic factors in species distribution modeling: are interactions with soil microbes important? Ecography (Cop). 2016;39(10):970–80.

30. Lawrence D, Fiegna F, Behrends V, Bundy JG, Phillimore AB, Bell T, et al. Species interactions alter evolutionary responses to a novel environment. PLoS Biol. 2012;10(5).

31. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, et al. A communal catalogue reveals Earth's multiscale microbial diversity. Nature. 2017;551(7681):457–63.