Undergraduate Journal of Experimental Microbiology and Immunology (UJEMI) Vol. 26:1-13

Soil abiotic factors are not consistently associated with microbial diversity or organic matter removal intensity in regions of long-term reforestation

Mark Pitblado, Sara Kowalski, Elizabeth Vaz, Alannah Wilson

Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada

SUMMARY Healthy soil structure is necessary for the sustainable management of forest ecosystems. As a major component of the Earth's biosphere, these forest soil ecosystems play a key role in climate regulation, biogeochemical cycling, and the maintenance of biodiversity. Despite the importance of soil ecosystems, not much is known about the role of soil bacterial communities in mediating the health and productivity of the forest floor, especially in the context of deforestation. Here, we analyzed data obtained from a Long-Term Soil Productivity study to determine the effects of organic matter removal on soil abiotic factors, and the resulting impact on bacterial diversity 10 years after reforestation. By analyzing beta diversity, we found that both geographic location and soil depth were associated with differences in diversity between soil bacterial communities. Further statistical analysis also revealed significant relationships between soil depth and abiotic factors. Higher soil organic carbon, nitrogen content, and moisture content were associated with samples taken from the organic topsoil layer, and soil pH levels were more acidic in organic soil samples compared to mineral soil samples. Alpha diversity and taxonomic abundance analyses indicated that the distribution of bacterial phyla differed between geographic regions, with significantly lower diversity in British Columbia soil communities. We did not find organic matter removal to consistently impact the levels of soil organic carbon, nitrogen content, moisture content, or pH. Similarly, linear regression models for each region indicated minimal associations between soil abiotic factors and bacterial alpha diversity. Overall, these findings provide insight into the association between bacterial community composition and soil abiotic factors across a variety of geographic regions, and the impact of deforestation on these relationships.

INTRODUCTION

Forest ecosystems are composed of several interacting communities and their physical environment. Responsible for a variety of ecological processes, soil systems can be considered the backbone of the forest ecosystem, mainly functioning to provide nutrients and structure in order to sustain life (1). Although soil is known to house a myriad of microorganisms, we have just begun to appreciate the contribution of these microbial detritivores to various environmental processes (2). Research over the last two decades indicates numerous roles for microbial communities in the context of soil ecology, with soil microbes impacting processes such as organic matter decomposition (3), greenhouse gas production (3), forest floor fertility (4, 5, 6), reforestation potential after logging (4, 5), and the long-term biodiversity of forested regions (5). Together, these roles implicate soil microbial communities as key in the long-term conservation of healthy, productive forests. Thus, increased knowledge of the maintenance of soil microbial communities and the effects of deforestation on their structure has the potential to inform sustainable forest management practices, which in turn will support positive environmental change.

Given that research into the soil microbiome is still in its infancy, the function of resident bacterial species in supporting a productive and healthy community structure is still being evaluated. In general, it has been hypothesised that a more diverse bacterial community confers greater functional redundancy on soil environments. This diversity within the communities is thought to increase the stability of soil processes and systems against environmental changes and disturbances (7). However, soil communities in forested regions have been shown to be richer than those of other earth environments, such as sand and rock

Published Online: September 2021

Citation: Mark Pitblado, Sara Kowalski, Elizabeth Vaz, Alannah Wilson. 2021. Soil abiotic factors are not consistently associated with microbial diversity or organic matter removal intensity in regions of long-term reforestation. UJEMI 26:1-13

Editor: Daniela Morales, Stefanie Sternagel and Brianne Newman, University of British Columbia

Address correspondence to: https://jemi.microbiology.ubc.ca/

(8, 9). Several studies have found that Proteobacteria and Acidobacteria are generally the most abundant bacterial phyla in certain forest soil communities (8, 9, 10). In natural hardwood forests, Proteobacteria dominate, while Acidobacteria tend to dominate in coniferous forest soil populations (11). Together, these bacterial groups contribute to many essential soil processes, such as the carbon and nitrogen cycles, to help promote healthy soil structure.

Another important consideration in maintaining the structure and function of these microbial communities, is how they are influenced by soil abiotic factors. Soil conditions are defined by several physical and chemical properties that may influence the abundance and diversity of soil microbial populations. For instance, previous research has implicated Soil Moisture Content (SMC) as a primary determinant of microbial respiratory groups (12, 13). The biochemical functions of these respiratory groups further impact other abiotic factors, including soil nitrogen content (N), pH, and soil organic carbon content (SOC), which likewise affect soil microbial communities (14). Several studies have found pH to be positively correlated with microbial biomass, with lower microbial biomass associated with acidic environments (9, 15, 16, 17). Further, pH also influences the cycling of nitrogen and organic carbon within soil (18). N has been shown to be positively correlated with microbial abundance but negatively correlated with microbial diversity (19). While SOC may be associated with microbial abundance or diversity, as microbial activity contributes to the function of soil as either a carbon source or sink (20); however, its relationship with the soil microbiome remains unclear (21).

Given that abiotic factors play a substantial role in shaping the soil microbiome, conducting further comprehensive analysis is necessary to confirm previous findings and increase our understanding on the mediators of soil abiotic variables. It has been previously established that abiotic soil factors are affected by the depth at which samples are taken. Generally, it has been found that SOC and N decrease as soil depth increases (22, 23), a result of the topsoil receiving organic matter from plant and animal residues at the surface (23). As well, SOC contents tend to be more variable at all depths than N contents (23). Results surrounding the effect of soil pH, however, are less straightforward. Some reports claim that pH increases with soil depth (24, 25), while others indicate that soil depth does not cause significant acidity changes (22). Further, SMC has been associated with either decreases or increases with soil depth, depending on the environment (22). Human interaction with forested regions can also heavily affect abiotic soil factors, specifically through logging and lumber harvesting. Deforestation has been shown in various studies to decrease SOC, N, SMC, and pH (26, 27, 28, 29). Even in regions of long-term reforestation, changes in soil abiotic factors can persist. The time required to replace lost SOC and N in reforested regions may exceed 200 years depending on the extent of timber harvest (26), while decreased SMC resulting from intense logging may never recover (30). Thus, research into both the biotic and abiotic characteristics of forested regions, as well as the interactions between them, will allow for the clarification of the long-term effects of logging.

While the relationship between soil abiotic factors and bacterial communities has previously been evaluated (5, 6, 14), uncertainty still exists concerning its application to regions of reforestation. In this study, we used metadata and bacterial 16S amplicon sequences obtained from a metagenomic survey and Long-Term Soil Productivity (LTSP) study conducted by Wilhelm *et al.* in 2017 (31) to attempt to resolve the association of abiotic factors with soil bacterial diversity, within the context of several reforested ecozones across North America. We found that soil bacterial communities differed by geographic region and according to SOC, N content, pH, and SMC. In contrast, we found no correlation between soil abiotic factors and the intensity of organic matter (OM) removal. Our research gives further insights into how the abiotic factors SOC, N, SMC, and pH affect bacterial biodiversity and abundance, how soil depth affects microbial populations, and the role that organic matter removal plays in shaping a microbial community.

METHODS AND MATERIALS

Overview of sample collection and metadata. DNA sequencing data and associated metadata were obtained from a Long-Term Soil Productivity (LTSP) study conducted by Wilhelm *et al.* (31). Briefly, Wilhelm *et al.* collected soil samples from eighteen reforested

experimental sites located in four North American regions: British Columbia (BC), Ontario (ON), California (CA), and Texas (TX). Each collection site contained plots that experienced varying degrees of organic matter removal: OM1, the removal of tree stems (stem-only harvesting) with woody debris left behind; OM2, in which the whole tree including the branches were harvested; OM3, where the whole tree along with the forest floor (upper organic layer) was removed and; REF, reference samples were taken from neighbouring, unharvested plots. The plots were replanted with trees corresponding to the ecozone and left to reforest for 10-15 years prior to sampling. Soil was sampled in triplicate from two different layers: the organic layer (O-horizon) and the top 20 cm of the mineral layer (A-horizon), and metadata including several geographic, climatic, chemical, and physical factors were supplied for each sample. Further details on experimental design are described by Wilhelm *et al.* (31).

Preparation of amplicon sequence library. Bacterial 16S rRNA amplicon sequencing data was taken from the previously mentioned LTSP study conducted by Wilhelm *et al.* (31). In this study, the team extracted DNA from field soil samples (0.5 g) using the FastDNA Spin Kit for Soil according to the manufacturer's protocol (MPBio, Santa Ana, CA). The team prepared a bacterial 16S rRNA gene (V1 - V3 regions) library via polymerase chain reaction (PCR) according to Hartmann *et al.* (32) using the universal primers 27 F/519 R to amplify the region spanning V1-V3 (28). Further sequencing data can be found in Wilhelm *et al.* (31).

Bioinformatic analysis. Data was provided as demultiplexed sequence files and a manifest file from Wilhelm *et al*. (31). These were quality controlled using the DADA2 algorithm, and then truncated after 200 bases. We conducted phylogenetic diversity analyses in QIIME2 (33) (Supplemental Text Files: Qiime2 Code) on a remote server setup, by first generating a phylogenetic diversity tree, and then generating an alpha rarefaction plot within the QIIME2 website to select our sampling depth of 2500. These files were then exported for analysis and generation of visuals in R software environment on a local machine.

All statistical analysis and visualizations were performed in R (v 4.0.3) (R Core Team) (Supplemental Text Files: R Script). The data was cleaned to exclude ASV's with an abundance less than 10, and also to eliminate instances in which an observation was either missing (N/A) or zero, as this represented an invalid value for all variables analyzed. Following this, the phyloseq package was used to observe beta diversity between sample sites using principal coordinate analysis plots on rarefied data, both on the whole dataset and across the four regions individually (min = 5000). Calculation of observed features and Chao1 richness for alpha diversity boxplots was also conducted using rarefied data (min = 5000). A bar plot comparing phyla composition for each region was performed using the all-regions phyloseq object using base R plot functionality and then organized by region. Statistical significance for the soil depth box plots was determined via Wilcoxon signed-rank test ($p < 0.05$), and statistical significance for the organic matter removal boxplots was determined via Kruskal-Wallis H test ($p < 0.05$). In preparation for a linear regression, temporary data frames were created by binding alpha diversity metrics lists and metadata columns for each ecozone independently. Finally, the linear regression model to predict Shannon diversity from the abiotic factors of interest was performed for each of the six ecozones in the dataset using these data frames.

RESULTS

Geographic location influences soil microbiome composition. To assess the impact of geographic location on soil bacterial community composition, we compared soil bacterial communities from different regions. To this end, we analyzed data from an LTSP study performed by Wilhelm *et al.* (31). In their study, soil samples were taken from 18 LTSP study sites encompassing six different ecozones situated across North America (Fig. 1A), and a bacterial 16S amplicon sequence library was prepared. To assess beta diversity, while taking into consideration both abundance and phylogeny, we ran a weighted UniFrac analysis focused on the regional location of the samples. Distinct clustering of bacterial communities by region was observed with overlap of bacterial community clusters occurring between CA, ON, and TX, suggesting similarities in overall soil bacterial compositions between the three regions (Fig. 1B). Interestingly, BC samples clustered independently from the other three

examined regions, highlighting the largely different bacterial community composition of BC soil samples (Fig. 1B). Furthermore, when looking at the compiled data, there was no distinct community clustering for the abiotic factors pH, SOC, N, and SMC (Fig. S1). Considering the implications of these results, further analysis was performed on regionally grouped data to prevent the masking of subtle associations between microbial composition and other metadata variables.

Soil microbial communities cluster according to edaphic factors. Given the strong association between geographical location and community composition (Fig. 1B), we subset the data by region and developed PCoA plots to determine the relationship between bacterial diversity and the availability of edaphic factors. We found that soil depth had a large impact on overall bacterial beta-diversity in each of the four regions (Fig. 2A). Demonstrated by the formation of two separate clusters, bacterial communities grouped based on the depth the soil samples were taken from: the organic layer vs. the mineral layer. This result was not surprising as SOC, N, SMC, and pH level showed significant differences in relation to soil depth (Fig. S2). Across each region, high SOC, N, and SMC were associated with samples taken from the soil organic top layer, while samples taken from the mineral layer displayed significantly lower content of each abiotic factor (Fig. S2A-B). TX samples showed much lower amounts of SOC and N in the soil organic layer compared to the other four regions (Fig. S2A-B).

Further, when assessing the SOC, N, and SMC, we saw a decrease in the magnitude of difference between the organic and mineral layer samples from TX. Notably, unlike BC, ON, and TX, we found no significant difference in SMC between depths in CA (Fig. S2A-C). The opposite trend was observed when looking at soil depth and pH levels (Fig. S2D). More acidic pH was associated with samples taken from a shallow depth, whereas a relatively more basic pH was seen in the deeper, mineral layer samples. To demonstrate that this relationship between edaphic factors and soil depth played a role in the bacterial diversity observed amongst samples taken from different layers, we generated PCoA plots of weighted UniFrac distances coloured by the corresponding abiotic factor. Distinct clustering of bacterial communities associated with SOC, N, SMC, and pH were shown to be present in all four regions (Fig. 2B-E). We found that communities tended to cluster based on ‘high’ and ‘low’ SOC, N, and SMC (Fig. 2B-D), and that this pattern was stronger in samples from BC and ON. Bacterial community composition was also associated with changes in pH. Communities with an acidic pH (around 4) tended to cluster separately from those with a relatively more basic pH (around 6) (Fig. 2E). However, we found this trend to be weakly displayed across each region. Finally, we found that overall, bacterial community compositions in TX were less associated with all four abiotic factors.

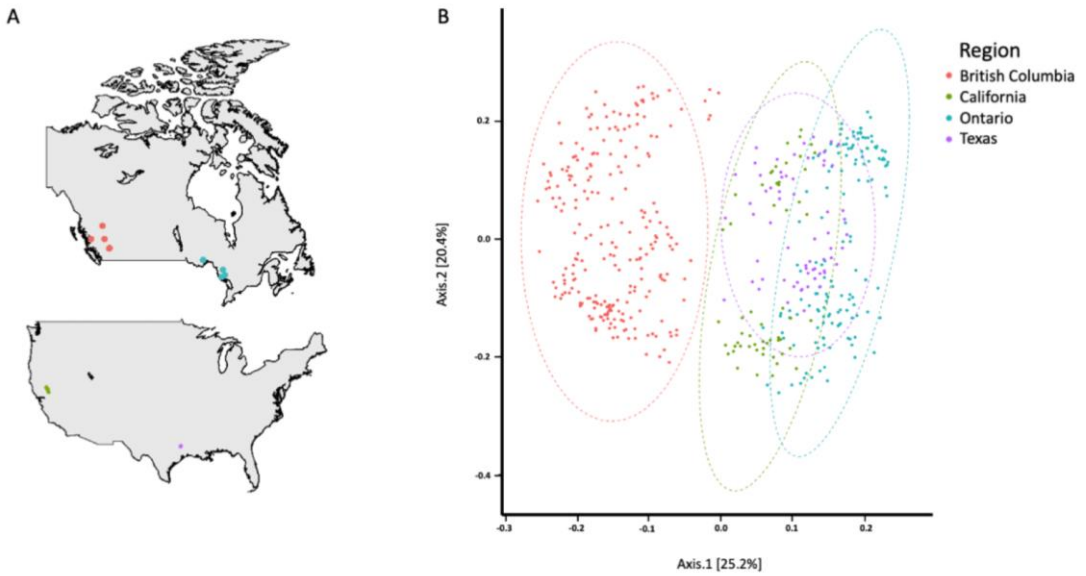

FIG. 1 Soil bacterial composition varies by region. (A) Map indicating sampled LTSP experimental sites from four North American regions: British Columbia (red), Ontario (blue), California (green), and Texas (purple). (B) PCoA plot of weighted UniFrac distances for soil samples by region. Axes represent variance explained.

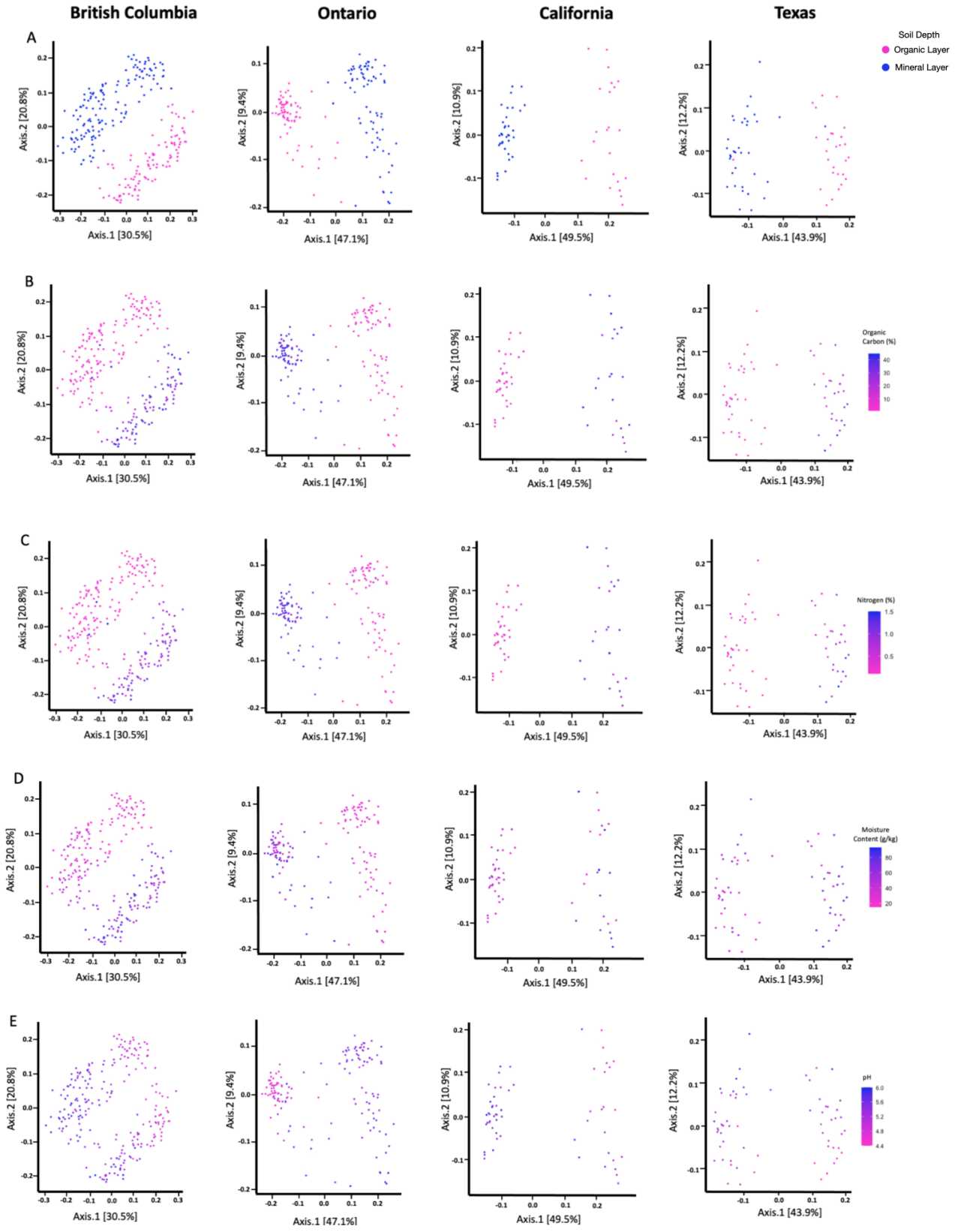

FIG. 2 Soil abiotic factors are associated with bacterial composition. Bacterial community similarities illustrated by PCoA plots of weighted UniFrac distances for each region. Plots are coloured by soil depth (A), soil organic carbon content (B), nitrogen content (C), soil moisture content (D), and pH (E). Soil organic carbon content and nitrogen content are given as a percentage of total soil mass. Axes represent variance explained.

British Columbia samples differ in Alpha diversity and taxonomic abundance. In order to assess the overall diversity of the dataset, we calculated alpha diversity by two metrics. Given that our previous analysis had shown that different regions were highly variable in terms of community, we subset by region (Fig. 3A). BC samples differed significantly in their taxonomic composition, and thus also in their alpha diversity. Both alpha diversity metrics, observed features and Chao1, were lower in BC samples than in other regions (Fig. 3A-B). While ON, CA, and TX had mean observed feature counts just under 400, the BC samples showed under 200 uniquely observed features. Additionally, ON, CA, and TX samples had higher, very similar alpha diversities. The notches overlap for CA, ON, and TX populations in all three plots, indicating similar populations (Fig. 3A-B). Investigating bacterial community composition further, we found that bacterial populations differed significantly in BC samples compared to other regions, which likely contributed to the difference in alpha diversity. There were a higher proportion of Proteobacteria in the BC samples than in ON, CA, and TX samples (Fig. 3C), although Proteobacteria were the most abundant phyla in all regions. Further, several bacterial phyla were reduced in quantity in the BC samples, including WPS2 (Eremiobacterota), WS2, and WS3 (Latescibacteria), which are represented in pink. Notably, Acidobacteria were abundant in all regions (Fig. 3C).

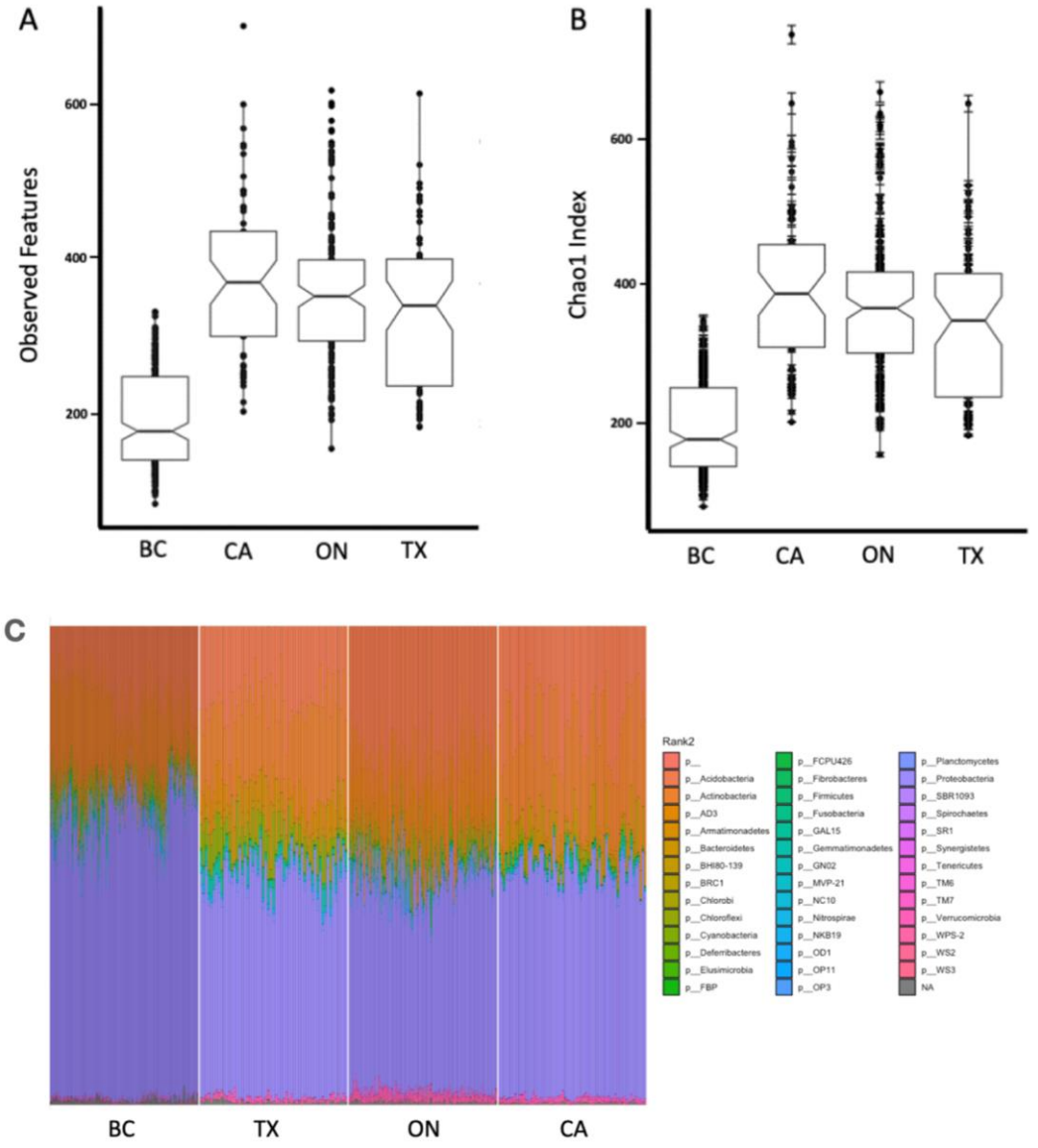

FIG. 3 Microbial diversity and taxonomic abundance differ in BC compared to TX, ON and CA. Comparing alpha diversity between regions using the observed features metric (A) and Chao1, which measures richness (B). Medians which fall outside the notches of other groups are considered significantly different. (C) Relative taxonomic abundance of microbial phyla by region. Classifications are resolved to the phylum level, with NA (gray) representing the relative abundance of sequences that do not have a phylum-level annotation, and blank (salmon) representing sequences that do have annotations but were not resolved to the phylum level.

Organic matter removal intensity does not consistently influence soil organic carbon and nitrogen content. After we identified the association between soil microbiome composition and SOC and N content, we next investigated whether these abiotic factors were influenced by the intensity of OM removal. Wilhelm *et al.* (31) obtained soil samples from plots that had been subjected to varying intensities of OM removal. Reforestation had occurred naturally for 11-17 years before soil samples were collected. We found that SOC and N both varied between different OM removal intensities, but our data did not indicate any consistent association between the values of either abiotic factor or OM removal intensity (Fig. 4). OM3 samples tended to have lower SOC and N than other OM removal intensities. Interestingly, this trend was not observed in samples from TX. Taking this together with the overall SOC and N content compared between soil depth, TX soil samples seem to have a different SOC and N profile compared to other regions. While we also investigated the potential influence of OM removal intensity on pH and SMC, we did not identify any significant correlations or trends between these factors (Fig. S3).

Microbial diversity is not consistently associated with soil abiotic factors. While we did not find a consistent relationship between OM removal intensity and any of our abiotic factors of interest, we still wanted to assess whether these abiotic factors were associated with soil microbial diversity. In order to reduce the variation introduced by geographic location, we subset samples by ecozone rather than by region. Each of the four regions was divided into ecozones, and we performed a linear regression analysis for samples within each ecozone (Table 1). While each abiotic factor was a significant predictor of Shannon diversity in at least one ecozone, none of the factors were able to predict alpha diversity across all ecozones. SOC, N, and pH were significantly correlated with alpha diversity in three of the six ecozones, and SMC was correlated with alpha diversity in a single ecozone, however none of these correlations were of a meaningful magnitude.

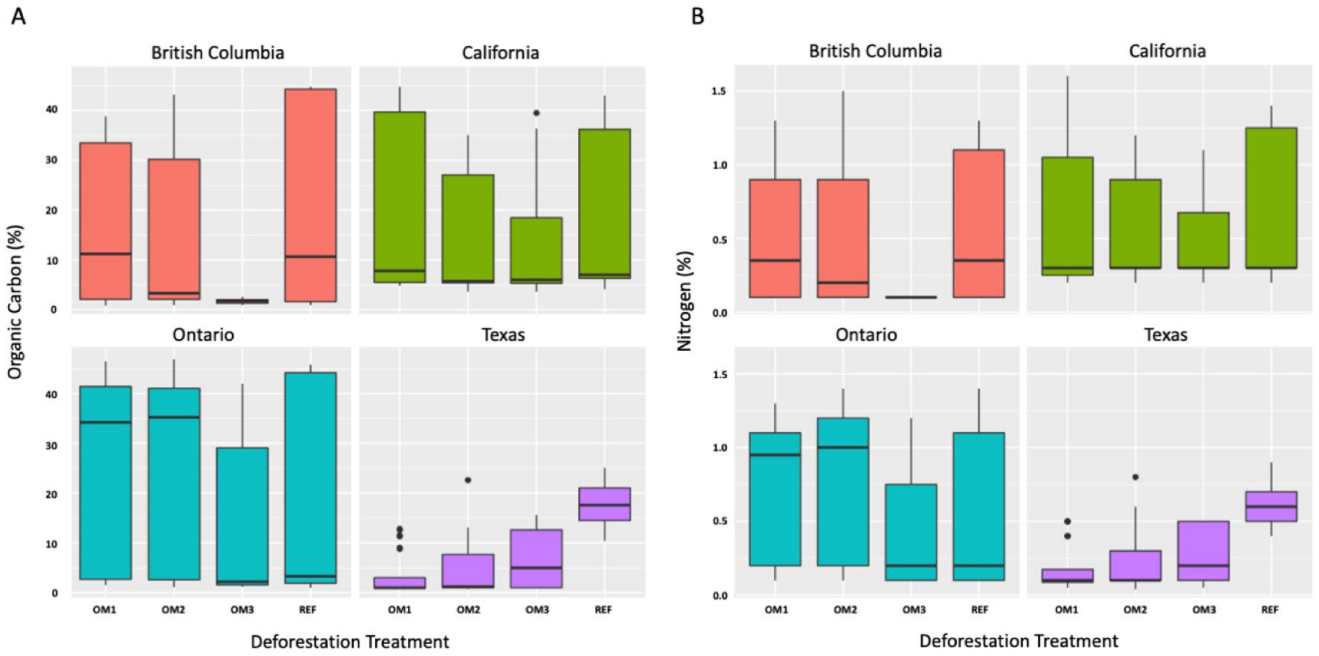

FIG. 4 Overall organic carbon and nitrogen content do not vary by intensity of organic matter removal. Organic carbon content (A) and nitrogen content (B) as a percentage of total soil mass. Deforestation treatments included removal of tree branches (OM1), removal of tree trunks and branches (OM2), or removal of whole trees and the entire topsoil layer (OM3). Reference samples were taken from nearby unharvested plots. Boxes indicate the 1st quartile, median, and 3rd quartile; whiskers indicate minimum and maximum values; dots indicate outliers. OM = organic matter; REF = reference.

TABLE 1. Linear regression of Shannon diversity vs abiotic factors by ecozone.

Ecozone		Std Estimate	Standard Error	Pr(>|t|)
IDFBC	*Total Nitrogen*	-0.07006	0.03348	0.03890
	Total Carbon	-0.00200	0.00096	0.03980
	pH	0.05909	0.06711	0.38100
	Moisture Content	-0.00079	0.00081	0.33800
SBSBC	*Total Nitrogen*	-0.15176	0.05346	0.00516
	Total Carbon	-0.00368	0.00140	0.00952
	pH	0.21961	0.04928	0.00002
	Moisture Content	0.00006	0.00106	0.95200
BSON	*Total Nitrogen*	-0.12777	0.08756	0.15000
	Total Carbon	-0.00308	0.00218	0.16200
	pH	0.20531	0.08473	0.01830
	Moisture Content	-0.00230	0.00225	0.31000
JPON	*Total Nitrogen*	-0.18190	0.06320	0.00503
	Total Carbon	-0.00477	0.00168	0.00558
	pH	0.12090	0.04837	0.01430
	Moisture Content	-0.00468	0.00197	0.01970
LPTX	*Total Nitrogen*	-0.16397	0.17279	0.34600
	Total Carbon	-0.00658	0.00566	0.24900
	pH	-0.10760	0.10470	0.30800
	Moisture Content	-0.00243	0.00447	0.58900
PPCA	*Total Nitrogen*	0.08803	0.09527	0.35900
	Total Carbon	0.00144	0.00279	0.60800
	pH	0.01923	0.06099	0.75400
	Moisture Content	0.00179	0.00297	0.54900

DISCUSSION

In this study, we investigated the relationship between bacterial diversity and the soil abiotic factors SOC, N, SMC, and pH and assessed the influence of OM removal on these abiotic factors. We found that the composition of soil bacterial communities differed by geographic region, and that BC communities were less diverse than those of other regions. Our analyses also indicated that abiotic factors varied by soil depth, while OM removal did not consistently influence the levels of soil abiotic factors.

Our results indicate that soil depth is significantly correlated with the edaphic factors SOC, N, pH, and SMC. Specifically, we found that SOC, N, and SMC decreased as soil depth increased, which is consistent with findings from published literature (34, 35, 36). Previous studies have also shown that in forest soil, pH is positively correlated with soil depth in upper soil layers (from 0 to 30 cm deep), but negatively correlated with soil depth in lower soil layers (below 30 cm) (24). Our analyses found that pH was significantly higher with increased soil depth, which aligns with published findings, as our samples were taken within the upper soil layer, at depths of 10 cm and 30 cm. Taken together, our results support previous research that suggests that SOC and N content are negatively correlated with soil pH, as low pH supports the accumulation of organic matter (24).

This correlation between soil depth and SOC and N is particularly notable given the association of these factors with distinct clusters of bacterial communities. When clustered

based on weighted UniFrac distances, samples within each region formed two well-defined clusters. These clusters corresponded with soil depth, with one cluster containing samples from the organic layer and the other containing samples from the mineral layer. Levels of SOC and N were also tightly associated with this clustering; SOC and N tended to be high in the organic layer cluster and low in the mineral layer cluster. These results suggest that soil microbiome composition is associated with SOC and N content, as the bacterial communities analyzed are most similar to one another when levels of SOC and N are comparable.

In addition to edaphic factors, we found that soil microbiome composition was also influenced by geographic region, and that BC differed from other regions with regard to bacterial diversity. Alpha diversity comparisons of both observed features and Shannon diversity indicated that BC samples exhibited lower abundance and were less evenly distributed than soil bacterial communities in other regions. While not conclusive, this may provide evidence as to why the BC samples were separated from the other regions on the PCoA plots, or alternatively that there are unique characteristics of the BC ecosystems that affect diversity differently than the other regions. We also assessed the taxonomic composition of the soil microbiome in each region. The results of this analysis agree with previous literature which suggests that Proteobacteria and Acidobacteria account for large portions of forest soil bacteria, both generally (8, 9, 10), and in BC specifically (37). The Proteobacteria phylum consists of the classes Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria, and Deltaproteobacteria (37). The increased proportion of Proteobacteria and Acidobacteria in BC samples likely contributes to the reduced alpha diversity of these samples. However, given the physiological diversity of Proteobacteria (37), it is difficult to predict the metabolic and biochemical consequences of their increased proportion on bacterial soil communities in BC.

One of the aims of our analyses was to investigate the impact of OM removal on edaphic factors, and consequently on soil bacterial diversity. Various studies have shown that deforestation results in decreased SOC, N, SMC, and pH (26, 27, 28, 29), but these conclusions are not supported by our results. We did not find a correlation between the intensity of OM removal and any of our edaphic factors of interest. To begin to explain the inconsistency between our results and those of previous studies, it should be noted that the samples considered in our analyses were collected from sites at which reforestation had occurred for over a decade after OM removal. In contrast, some of the previous studies (26, 28) collected samples from plots at which reforestation had not occurred, either because the sites had only recently been harvested, or because urbanization prevented reforestation. Thus, any changes in edaphic factors that may occur during the process of reforestation may contribute to differences between our results and those of previous studies.

Additionally, the LTSP study conducted by Wilhelm *et al.* (31) included samples from sites subjected to three different intensities of OM removal, while some of the published studies cited above (26, 27, 28) considered only a single OM removal treatment. The increased resolution of the data used in our analyses may help to explain the contradiction with previous results. For our results to show a significant correlation between OM removal intensity and an abiotic factor, the level of the abiotic factor would have to consistently increase or decrease between samples from each of the three OM removal intensities, as well as the reference samples. That is, the REF–OM1, OM1–OM2, and OM2–OM3 differences must all trend in the same direction. This more stringent significance criteria used in our analyses may contribute to the discrepancy between our findings and those of previous studies.

This discrepancy may also be due to variability introduced by geographic location. In the present study we could not address this potential source of variability by subsetting samples by region before assessing the relationship between OM removal and abiotic factors, as this would have considerably reduced our sample size, and consequently prevented us from drawing any statistically significant conclusions. While we did not find an overall correlation between OM removal intensity and any of our abiotic factors of interest, our analyses did indicate that SOC and N were significantly decreased in ON and BC samples from OM3-treated plots. Additionally, we found that pH was significantly increased in OM3-treated samples from ON and BC. These results agree with our finding that SOC and N content are negatively correlated with pH in upper soil layers.

To gain a more comprehensive understanding of the associations between edaphic factors and bacterial diversity, we performed a linear regression analysis. We had hoped that this analysis would serve as the first step in formulating a model to describe changes in bacterial diversity in relation to soil abiotic factors. However, our regression results did not indicate consistent relationships between soil bacterial diversity and any of our abiotic factors of interest. Despite this, our results do provide insight into which abiotic factors may be better predictors of bacterial diversity. SOC, N, and pH were each correlated with Shannon diversity in at least three ecozones, while SMC was only correlated with diversity in one of the six ecozones. Considering this, future attempts to define a descriptive model of soil bacterial diversity should consider SOC, N, and pH, but not necessarily SMC.

Conclusions This study utilized bioinformatic methods to assess associations between organic matter removal, soil abiotic factors, and the soil microbiome. Soil bacterial communities were found to be most similar between samples from the same geographic region. Additionally, soil abiotic factors were found to vary by soil depth, with soil depth being negatively correlated with SOC, N, and SMC, and positively correlated with pH. No correlation was found between OM removal and changes in edaphic factors, and edaphic factors did not consistently correlate with bacterial diversity. This study serves to confirm previously published results regarding the impact of soil depth on abiotic factors, and to broaden our understanding of the potential effects of OM removal on the soil microbiome.

Future Directions While this investigation has provided valuable insight into the determinants of bacterial biodiversity in forest soil communities, the Wilhelm *et al.* (31) dataset contains additional relevant metadata that may be investigated further. Subsetting the data by another metric, such as OM removal, as opposed to region, may provide valuable insight. Additionally, the remaining metadata categories not included in this analysis, mainly precipitation, soil bulk density, and compaction treatment, could be introduced as covariates.

The data from the Wilhelm *et al.* LTSP study was composed of samples from a variety of diverse geographic regions (31). While this large range of sampling sites supports the generalizability of our analyses, it also introduces a high degree of variability to the data. Herein lies the dilemma of our analysis: sub-setting the data by region may reveal a well-defined correlation between OM removal and a given variable, but it will not allow for the generalization of that correlation. The opposite is true for a large dataset encompassing several regions. While it may be more reasonable to generally apply any conclusions drawn from such a dataset, the correlations would be difficult to identify given the large variability in the dataset. Given that we did subset the data by region, however, if a similar study were to be conducted, it would be beneficial to have more sample replicates for a given region, in order to increase the amount of analyzed data.

Additionally, our findings indicated that SOC, N, and pH are important factors associated with soil bacterial diversity in reforested regions, but that SMC may not play as critical a role. This may suggest that SMC should not be a focus of future soil bacterial diversity analyses, however, more research may be needed, preferably with a larger dataset, in order to confirm this suggestion. Given that our investigation has demonstrated SOC, N, and pH to be seemingly better predictors of diversity, future studies should aim efforts towards these abiotic factors.

In our study, soil sampling depth was a key factor in limiting the metadata. Given that SOC and N content are so closely associated with soil depth, and that only two soil depths (10 cm and 30 cm) were surveyed, the SOC and N contents in the dataset tended to be either “high” or “low.” Due to the lack of moderate levels of SOC and N, this may have impacted any analyses that were conducted, especially the linear regression model. Thus, in the future, more varied soil sampling depths should be collected for any studies attempting a similar analysis.

In several areas, we found that geographic location had a large influence on bacterial diversity and abundance. This is likely due to the difference in soil composition, as well as other variables that may not have been recorded in the metadata of this study. Among these variables are different parent materials, particle size distribution, and mineral content (38).

Thus, our findings echo those of previous papers, which maintain that several variables affect the composition of bacterial soil communities.

ACKNOWLEDGEMENTS

We would like to acknowledge the teaching time of Microbiology 447, Dr. David Oliver and Dr. Stephan Koenig for their time and patience in helping us learn something new. We would also like to acknowledge Reynold Farrera, our designated teaching assistant, who provided feedback and mentorship. Lastly, we would like to thank the Department of Microbiology at the University of British Columbia for providing the resources, community, and opportunity to work and ponder.

CONTRIBUTIONS

M.P. wrote the QIIME2 and R code and created the tables and figures. S.K. wrote the methods section and contributed to the abstract, introduction, and results. E.V. contributed to the abstract, introduction, results, and future directions. A.W. wrote the conclusion and contributed to the abstract and introduction. All collaborated for the discussion section of the manuscript.

REFERENCES

1. **Doran JW, Zeiss MR**. 2000. Soil health and sustainability: managing the biotic component of soil quality. *Appl Soil Ecol* **15**:3-11.
2. **Maier RM, Pepper IL, Gerba CP**. 2009. Introduction to Environmental Microbiology. *Environmental Microbiology* 3-7.
3. **Kirschbaum MUF, Bin Guo L, Gifford RM**. 2008. Why does rainfall affect the trend in soil carbon after converting pastures to forests? A possible explanation based on nitrogen dynamics. *For Ecol Manage* **255**:2990-3000.
4. **Jastrzębska E, Kucharski J**. 2008. Dehydrogenases, urease and phosphatases activities of soil contaminated with fungicides. *Plant Soil Environ* **53**:51-57.
5. **Wyszkowska J, Boros E, Kucharski J**. 2007. Effect of interactions between nickel and other heavy metals on the soil microbiological properties. *Plant Soil Environ* **53**:544-552.
6. **Chen Q-L, Ding J, Zhu Y-G, He J-Z, Hu H-W**. 2020. Soil bacterial taxonomic diversity is critical to maintaining the plant productivity. *Environ Int* **140**:105766.
7. **Isobe K, Ise Y, Kato H, Oda T, Vincenot CE, Koba K, Tateno R, Senoo K, Ohte N**. 2020. Consequences of microbial diversity in forest nitrogen cycling: diverse ammonifiers and specialized ammonia oxidizers. *ISME J* **14**:12-25.
8. **Roesch LFW, Fulthorpe RR, Riva A, Casella G, Hadwin AKM, Kent AD, Daroub SH, Camargo FAO, Farmerie WG, Triplett EW**. 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J* **1**:283-290.
9. **Lauber CL, Hamady M, Knight R, Fierer N**. 2009. Pyrosequencing-Based Assessment of Soil pH as a Predictor of Soil Bacterial Community Structure at the Continental Scale. *Appl Environ Microbiol* **75**:5111-5120.
10. **Nemergut DR, Cleveland CC, Wieder WR, Washenberger CL, Townsend AR**. 2010. Plot-scale manipulations of organic matter inputs to soils correlate with shifts in microbial community composition in a lowland tropical rain forest. *Soil Biol Biochem* **42**:2153-2160.
11. **Lin Y-T, Jangid K, Whitman WB, Coleman DC, Chiu C-Y**. 2011. Change in bacterial community structure in response to disturbance of natural hardwood and secondary coniferous forest soils in central Taiwan. *Microb Ecol* **61**:429-437.
12. **Li S, Xu M, Sun B**. 2014. Long-term hydrological response to reforestation in a large watershed in southeastern China. *Hydrol Process* **28**:5573-5582.
13. **Linn, D.M. & Doran, J.W**. 1984. Effect of water-filled pore space carbon dioxide and nitrous oxide production in tilled and nontilled soils. *Soil Sci Soc Am J* **48**:1267-1272.
14. **Giacometti C, Demyan MS, Cavani L, Marzadori C, Ciavatta C, Kandeler E**. 2013. Chemical and microbiological soil quality indicators and their potential to differentiate fertilization regimes in temperate agroecosystems. *Appl Soil Ecol* **64**:32-48.
15. **Zhalnina K, Dias R, de Quadros PD, Davis-Richardson A, Camargo FAO, Clark IM, McGrath SP, Hirsch PR, Triplett EW**. 2015. Soil pH Determines Microbial Diversity and Composition in the Park Grass Experiment. *Microb Ecol* **69**:395-406.
16. **Rousk J, Bååth E, Brookes PC, Lauber CL, Lozupone C, Caporaso JG, Knight R, Fierer N**. 2010. Soil bacterial and fungal communities across a pH gradient in an arable soil. *ISME J* **4**:1340-1351.
17. **Fierer N, Jackson RB**. 2006. The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci U S A* **103**:626-631.
18. **Kemmitt SJ, Wright D, Goulding KWT, Jones DL**. 2006. pH regulation of carbon and nitrogen dynamics in two agricultural soils. *Soil Biol Biochem* **38**:898-911.

19. **Wang Q, Wang C, Yu W, Turak A, Chen D, Huang Y, Ao J, Jiang Y, Huang Z**. 2018. Effects of Nitrogen and Phosphorus Inputs on Soil Bacterial Abundance, Diversity, and Community Composition in Chinese Fir Plantations. *Front Microbiol* **9**:1543.
20. **Maron PA, Sarr A, Kaisermann A, Lévêque J, Mathieu O, Guigue J, Karimi B, Bernard L, Dequiedt S, Terrat S, Chabbi A, Ranjard L**. 2018. High Microbial Diversity Promotes Soil Ecosystem Functioning. *Appl Environ Microbiol* **84**:e02738-17.
21. **Trivedi P, Anderson IC, Singh BK**. 2013. Microbial modulators of soil carbon storage: integrating genomic and metabolic knowledge for global prediction. *Trends Microbiol* **21**:641-651.
22. **Emiru N, Gebrekidan H**. 2013. Effect of land use changes and soil depth on soil organic matter, total nitrogen and available phosphorus contents of soils in Senbat watershed, western Ethiopia. *Am J Agric Biol Sci* **8**:7.
23. **Zhou W, Han G, Liu M, Zeng J, Liang B, Liu J, Qu R**. 2020. Determining the Distribution and Interaction of Soil Organic Carbon, Nitrogen, pH and Texture in Soil Profiles: A Case Study in the Lancangjiang River Basin, Southwest China. *Forests* **11**:532.
24. **Zhou W, Han G, Liu M, Li X**. 2019. Effects of soil pH and texture on soil carbon and nitrogen in soil profiles under different land uses in Mun River Basin, Northeast Thailand. *PeerJ* **7**:e7880.
25. **Han G, Li F, Tang Y**. 2015. Variations in soil organic carbon contents and isotopic compositions under different land uses in a typical karst area in Southwest China. *Geochem J* **49**:63-71.
26. **Jurgensen MF, Harvey AE, Graham RT, Page-Dumroese DS, Tonn JR, Larsen MJ, Jain TB**. 1997. Impacts of Timber Harvesting on Soil Organic Matter, Nitrogen, Productivity, And Health of Inland Northwest Forests. *For Sci* **43**:234-251.
27. **Martínez-Mena M, Rogel JA, Castillo V, Albaladejo J**. 2002. Organic carbon and nitrogen losses influenced by vegetation removal in a semiarid Mediterranean soil. *Biogeochemistry* **61**:309-321.
28. **Sahani U, Behera N**. 2001. Impact of deforestation on soil physicochemical characteristics, microbial biomass, and microbial activity of tropical soil. *Land Degrad Dev* **12**:93-105.
29. **Brais S, Camiré C, Paré D**. 1995. Impacts of whole-tree harvesting and winter windrowing on soil pH and base status of clayey sites of northwestern Quebec. *Can J For Res* **25**:997-1007.
30. **Unger IM, Goyne KW, Kremer RJ, Kennedy AC**. 2012. Microbial community diversity in agroforestry and grass vegetative filter strips. *Agrofor Syst* **87**:395-402.
31. **Wilhelm RC, Cardenas E, Leung H, Maas K, Hartmann M, Hahn A, Hallam S, Mohn WW**. 2017. A metagenomic survey of forest soil microbial communities more than a decade after timber harvesting. *Sci Data* **4**:170092.
32. **Hartmann M, Lee S, Hallam SJ, Mohn WW**. 2009. Bacterial, archaeal and eukaryal community structures throughout soil horizons of harvested and naturally disturbed forest stands. *Environ Microbiol* **11**:3045-3062.
33. **Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu YX, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, and Caporaso JG**. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**:852-857.
34. **Singh A, Phogat VK, Dahiya R, Batra SD**. 2014. Impact of long-term zero till wheat on soil physical properties and wheat productivity under rice–wheat cropping system. *Soil Tillage Res* **140**:98-105.
35. **Zhao X, Xue J-F, Zhang X-Q, Kong F-L, Chen F, Lal R, Zhang H-L**. 2015. Stratification and Storage of Soil Organic Carbon and Nitrogen as Affected by Tillage Practices in the North China Plain. *PLoS One* **10**:e0128873.
36. **Craib IJ**. 1929. PhD thesis. Some Aspects of Soil Moisture in the Forest. Yale University, New Haven, CT.
37. **Chow ML, Radomski CC, McDermott JM, Davies J, Axelrood PE**. 2002. Molecular characterization of bacterial diversity in Lodgepole pine (*Pinus contorta*) rhizosphere soils from British Columbia forest soils differing in disturbance and geographic source. *FEMS Microbiol Ecol* **42**:347-357.
38. **Luo B, Chen X, Ding L, Huang Y, Zhou J, Yang T**. 2015. Response Characteristics of Soil Fractal Features to Different Land Uses in Typical Purple Soil Watershed. *PLoS One* **10**:e0122842.

R Packages

R. R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: https://www.R-project.org/

tidyverse. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen T, Miller E, Bache S, Müller K, Ooms J, Robinson D, Seidel D, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. 2019. Welcome to the Tidyverse. *JOSS* **4**:1686.

vegan. Oksanen J, Blanchet JF, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2020. vegan: Community Ecology Package. R package version 2.5-7. Available from: https://CRAN.R-project.org/package=vegan

phyloseq. McMurdie PJ, Holmes S. 2013. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**(4):e61217.

DESeq2. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**(12):550.

ggplot2. Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.

dplyr. Hadley Wickham, Romain François, Lionel Henry and Kirill Müller. 2020. dplyr: A Grammar of Data Manipulation. R package version 1.0.2. Available from: https://CRAN.R-project.org/package=dplyr

ggthemes. Arnold JB. 2019. ggthemes: Extra Themes, Scales and Geoms for 'ggplot2'. R package version 4.2.0. Available from: https://CRAN.R-project.org/package=ggthemes

ggpubr. Alboukadel K. 2020. ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.4.0. Available from: https://CRAN.R-project.org/package=ggpubr

viridis. Garnier S. 2018. viridis: Default Color Maps from 'matplotlib'. R package version 0.5.1. Available from: https://CRAN.R-project.org/package=viridis

ape. Paradis E, Schliep K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**:526-528.

plyr. Wickham H. 2011. The Split-Apply-Combine Strategy for Data Analysis. *Journal of Statistical Software* **40**(1):1-29. URL http://www.jstatsoft.org/v40/i01/.

maps. Original code: Becker RA, Wilks AR. R version: Brownrigg R. Enhancements: Minka TP, Deckmyn A. 2018. maps: Draw Geographical Maps. R package version 3.3.0. Available from: https://CRAN.R-project.org/package=maps

mapdata. Original code: Becker RA, Wilks AR. R version: Brownrigg R. 2018. mapdata: Extra Map Databases. R package version 2.3.0. Available from: https://CRAN.R-project.org/package=mapdata

scales. Wickham H, Seidel D. 2020. scales: Scale Functions for Visualization. R package version 1.1.1. Available from: https://CRAN.R-project.org/package=scales

ggsignif. Ahlmann-Eltze C. 2019. ggsignif: Significance Brackets for 'ggplot2'. R package version 0.6.0. Available from: https://CRAN.R-project.org/package=ggsignif

Hmisc. Original code: Harrell FE Jr. Contributions: Dupont C. 2020. Hmisc: Harrell Miscellaneous. R package version 4.4-2. Available from: https://CRAN.R-project.org/package=Hmisc

aod. Lesnoff M, Lancelot R. 2012. aod: Analysis of Overdispersed Data. R package version 1.3.1. Available from: http://cran.r-project.org/package=aod