**Microbial communities in forest and wetland soil types display divergent trends in diversity and unique functional responses to varying carbon-to-nitrogen ratios**

Arsh Sharma, Ramdeep Kailay, Karmanpreet Saini, Aaron Kwai, Muhammed Atasoy

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

**ABSTRACT**

The carbon-to-nitrogen (C:N) ratio is a critical environmental factor influencing microbial growth, nutrient availability, and community dynamics. This study investigates the impact of C:N ratio variations on microbial diversity, taxonomic composition, and functional potential in forest and wetland soils. By examining metagenomic datasets from these contrasting environments, the purpose was to understand ecosystem-specific responses to changing nutrient conditions, specifically changes in carbon and nitrogen. Beta diversity analysis revealed distinct clustering of microbial communities between soil types, underscoring their compositional differences. In forest soils, the higher C:N ratio range was associated with decreased phylogenetic diversity, evenness, and functional pathway diversity, alongside upregulation of antibiotic synthesis pathways, indicating a competitive and volatile microbial environment. Conversely, wetland soils exhibited increased phylogenetic diversity with stable functional pathway diversity with increasing C:N ranges, suggesting functional redundancy despite taxonomic shifts. These findings demonstrate ecosystem-specific responses to C:N ratios, with forest soils displaying increased microbial competition and wetlands maintaining functional stability. This work highlights the importance of comparative studies to elucidate the interplay between nutrient dynamics and microbial ecology in diverse soil environments.

**INTRODUCTION**

The C:N ratio is a key environmental parameter that influences microbial growth by affecting nutrient availability [(1)](https://www.zotero.org/google-docs/?broken=4vYFHQ). It dictates the energy and nutrient dynamics within the soil, impacting microbial metabolism and consequently, community composition [(2)](https://www.zotero.org/google-docs/?broken=Fkh8YR). Nitrogen is an essential nutrient for microbial growth, and its scarcity likely imposes a strong selective pressure, favouring only those microbes that can efficiently acquire nitrogen or fix it from the atmosphere [(3)](https://www.zotero.org/google-docs/?broken=PBIM03). Previous research has shown that variations in the C:N ratio can lead to significant changes in microbial diversity and activity [(1)](https://www.zotero.org/google-docs/?broken=3MvfzV), which in turn affects soil health and ecosystem productivity [(4)](https://www.zotero.org/google-docs/?broken=8xIg17).

Soil ecosystems, such as forests and wetlands, are distinct environments that shape microbial communities and their functions in nutrient cycling and water availability [(5)](https://www.zotero.org/google-docs/?broken=z4Zxi5). While forest soils typically experience well-drained conditions, wetland soils are characterized by prolonged waterlogging, which alters nutrient dynamics [(6–9)](https://www.zotero.org/google-docs/?broken=rfSrJk). Such environmental factors, including water availability and nutrient dynamics, can also amplify the impact of disturbances on microbial communities. In forest soils, disturbances—both short-term and long-term—significantly affect microbial composition and functionality [(10)](https://www.zotero.org/google-docs/?broken=2MtRb0). Although restoration efforts demonstrate that wetlands are remarkably resilient to environmental disturbances and capable of restoring ecosystem function [(5, 6)](https://www.zotero.org/google-docs/?broken=D3QKXi), there is limited research on how these ecosystems respond at the microbial community level. In particular, little is known about wetlands responses to changes in nutrient availability, such as shifts in the C:N ratio.

We hypothesized that when comparing wetlands and forest soil types, distinct trends in diversity, taxonomy, and predicted metabolic function will be identified in terms of changing C:N ratio. By using consistent C:N ranges across both soil types, it becomes possible to investigate the C:N ratio's influence on microbial communities and their functions. Comparative studies are thus crucial to uncover the specific composition and functionality of microbes in these contrasting environments in changing C:N conditions. In this study, 2 metagenomic datasets on wetlands and forest soil were collected by Ballentine *et al.* and by Wilhelm *et al.* [(11)](https://www.zotero.org/google-docs/?broken=ATnfrY)*,* respectively. Both studies collected data on the 16s rRNA sequences and metadata of their respective soil environments.

This study revealed contrasting effects of the C:N ratio on microbial communities in forest and wetland soils. In forest soil, higher C:N ratios reduced diversity and functional pathway diversity, with upregulated antibiotic synthesis pathways, indicating a competitive and dynamic microbial environment. In wetland soil, phylogenetic diversity increased, but functional pathway diversity remained stable, suggesting a taxonomic shift while maintaining evenness and functional redundancy. Our findings highlight ecosystem-specific responses to high C:N ratios, with forest soils becoming more volatile and wetlands retaining functional stability.

**METHODS**

**Dataset Acquisition and Description.** The study utilized two primary datasets: the forest soil dataset, published by Wilhelm *et al.* [(11)](https://www.zotero.org/google-docs/?broken=pQNU5L), which analyzed soil samples from post-logging forests, and the freshwater wetlands soil dataset compiled by Balantine *et al.*, focusing on soil samples from several wetlands in Northeastern America. Although the collecting team did not publish the freshwater wetlands soil dataset, it has been previously utilized in research by Balaji *et al.* [(12)](https://www.zotero.org/google-docs/?broken=8CMEKa). Both datasets were explored for their C:N ratios and examined through 16S rRNA gene sequencing to identify microbial communities.

**Data Wrangling and Binning.** The forest soil dataset included pre-calculated C:N ratios, derived from total soil carbon and nitrogen measurements. Conversely, the freshwater wetlands soil dataset required the calculation of C:N ratios from existing total carbon and nitrogen data. C:N ratios were categorized into "Very Low" (0-10), "Low" (10-20), "Intermediate" (20-30), "High" (30-50), and "Very High" (50-127), based on established literature [(1, 13)](https://www.zotero.org/google-docs/?broken=p4lCpk). Due to their low sample sizes, the "Very Low" and "Very High" categories were excluded from further analysis.

**Data Processing via QIIME2.** Sequence data from both datasets were imported and demultiplexed using QIIME2 2023.7 [(14)](https://www.zotero.org/google-docs/?broken=1uofka), followed by denoising and quality control steps conducted in DADA2 [(15)](https://www.zotero.org/google-docs/?broken=JQSTGH). For the forest soil dataset, sequences were truncated to 407 nucleotides, corresponding to the maximum sequence length that maintained a mean Phred quality score of 30. These amplicon sequence variants (ASVs) were then classified using the Silva 138-99 database [(16)](https://www.zotero.org/google-docs/?broken=yrvLBd) with primers 27F (AGAGTTTGATYMTGGCTCAG) and 534R (ATTACCGCGGCTGCTGG), targeting the V1-V3 regions of the 16S rRNA gene [(17)](https://www.zotero.org/google-docs/?broken=WyoLdt). For the freshwater wetlands dataset, sequences were truncated to 125 nucleotides, which corresponded to the maximum length of samples in the dataset and also contained a mean Phred quality score of 37. These ASVs were then classified using the Silva 138-99 database [(16)](https://www.zotero.org/google-docs/?broken=RgGq1N) with primers 515F (CACGGTCGKCGGCGCCATT) and 806R (GGACTACHVGGGTWTCTAAT), targeting the V4 region [(18)](https://www.zotero.org/google-docs/?broken=PztQJ6). To ensure data integrity, samples containing eukaryotic and mitochondrial sequences, as well as irrelevant C:N ratios (NA, “Very Low”, and “Very High”), were filtered out from both datasets. The QIIME2 [(14)](https://www.zotero.org/google-docs/?broken=ATrQMS) processing resulted in the creation of feature tables, rooted trees, taxonomy, and sample metadata for both datasets. These outputs were then merged to combine the information from both the forest soil and freshwater wetlands soil datasets. The merged data, along with the individual dataset outputs, were imported into R 4.4.1 [(19)](https://www.zotero.org/google-docs/?broken=Z43pyR) for further downstream analysis.

**Data Processing in R.** Using packages such as Phyloseq [(20)](https://www.zotero.org/google-docs/?broken=C5cC11), Ape [(21)](https://www.zotero.org/google-docs/?broken=zV995M), and Tidyverse [(22)](https://www.zotero.org/google-docs/?broken=tS0KnS) in R [(19)](https://www.zotero.org/google-docs/?broken=uA7ACu), the feature tables, rooted trees, taxonomy, and sample metadata generated from QIIME2 [(14)](https://www.zotero.org/google-docs/?broken=9XuuIw) were organized into phyloseq objects for the forest soil, freshwater wetlands soil, and combined merged datasets. Additional filtering steps included the removal of non-bacterial sequences, ASVs with less than five total counts, and samples with less than 100 reads. Samples were rarefied using rngseed = 8 to sampling depths at 2500 for the forest soil phyloseq object, 20000 for the freshwater wetlands soil phyloseq object, and 5000 for the Merged phyloseq object.

**Alpha Diversity Analysis.** Alpha diversity metrics for the forest soil and freshwater wetlands soil objects were analyzed in R [(19)](https://www.zotero.org/google-docs/?broken=kiocPc) using the following R packages: Picante [(23)](https://www.zotero.org/google-docs/?broken=DpY6XG), Phyloseq [(20)](https://www.zotero.org/google-docs/?broken=k7wYjR), Tidyverse [(22)](https://www.zotero.org/google-docs/?broken=gF73hQ), and Vegan [(24)](https://www.zotero.org/google-docs/?broken=DnVDoU). Metrics such as Shannon’s diversity index [(25)](https://www.zotero.org/google-docs/?broken=dfywAS) and Faith’s Phylogenetic Diversity (PD) [(26)](https://www.zotero.org/google-docs/?broken=sPXKo9) were then visualized using a boxplot with ggplot2 [(27)](https://www.zotero.org/google-docs/?broken=KrzAHC). Statistical significance for these metrics was computed using the Kruskal-Wallis test [(28)](https://www.zotero.org/google-docs/?broken=W1Kdgp) and visualized on the boxplot using ggsignif [(29)](https://www.zotero.org/google-docs/?broken=KMCGPj).

**Beta Diversity Analysis.** Beta diversity metrics, including Weighted UniFrac [(30)](https://www.zotero.org/google-docs/?broken=auxMg8), for the Merged phyloseq object was analyzed in R [(19)](https://www.zotero.org/google-docs/?broken=lk2vfz) using the following R packages: Phyloseq [(20)](https://www.zotero.org/google-docs/?broken=EtmLnv), Tidyverse [(22)](https://www.zotero.org/google-docs/?broken=MZSpif), and Vegan [(24)](https://www.zotero.org/google-docs/?broken=6dbtXP). Permutational multivariate analysis of variance (PERMANOVA) tests [(31)](https://www.zotero.org/google-docs/?broken=l0OaRY) were used to compute statistical significance for these metrics. A distance matrix was created for the weighted unifrac metrics using reshape2 [(32)](https://www.zotero.org/google-docs/?broken=W5D7Wi) and dplyr [(33)](https://www.zotero.org/google-docs/?broken=R9xUi4) packages, and a box plot of this matrix was made using ggplot2 [(27)](https://www.zotero.org/google-docs/?broken=UOlnka). Kruskal-Wallis [(28)](https://www.zotero.org/google-docs/?broken=cq2beX) and Dunn’s [(34)](https://www.zotero.org/google-docs/?broken=rvIGQu) tests were used to compute statistical significance using the dunn.test package [(35)](https://www.zotero.org/google-docs/?broken=okSdC6), which were visualized on the Weighted UniFrac boxplots using ggsignif [(29)](https://www.zotero.org/google-docs/?broken=kcwM3u).

**Differential Abundance Analysis.** Taxa that differed in abundance between C:N categories in forest soil and freshwater wetlands soil were identified using the DESeq2 [(36)](https://www.zotero.org/google-docs/?broken=D60UTo), Tidyverse [(22)](https://www.zotero.org/google-docs/?broken=kHeKbw), and Phyloseq [(20)](https://www.zotero.org/google-docs/?broken=TLo2k8) packages in R [(19)](https://www.zotero.org/google-docs/?broken=wOcpsB). The smaller of the two C:N categories was defined as the reference groups for comparisons with larger C:N categories and significantly differentially abundant taxa were filtered according to p\_adj < 0.01 and log2FoldChange ± 2. The number of hits for positive and negative differentially abundant genera were represented in a table.

**Indicator Species Analysis.** Indicator species analysis identified taxa significantly associated with varying C:N ratios in forest and wetlands soils using the following R packages: Indicspecies [(37)](https://www.zotero.org/google-docs/?broken=nSbWBH), Tidyverse [(22)](https://www.zotero.org/google-docs/?broken=lCzPtP), and Phyloseq [(20)](https://www.zotero.org/google-docs/?broken=cuaMaT). Data from phyloseq objects were transformed into compositional relative abundances and grouped to the genus levels. C:N categories served as predictors in the multipatt function, targeting taxa significantly associated with specific C:N conditions using a threshold p < 0.05.

**Predictive Functional Analysis.** Functional abundances between C:N categories in forest soil and freshwater wetlands soil were predicted using the Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt2) software [(38)](https://www.zotero.org/google-docs/?broken=1kQ6jI) after filtering out features with five or lower counts. PICRUSt2 analysis [(38)](https://www.zotero.org/google-docs/?broken=RRzo4x) aligns ASVs to reference sequences from the MetaCyc database [(39)](https://www.zotero.org/google-docs/?broken=nEN8Ch), outputting an abundance table with functional pathway annotations. The results of the differential abundance analysis of functional pathways were summarized using the DESeq2 [(36)](https://www.zotero.org/google-docs/?broken=B0UlKY), readr [(40)](https://www.zotero.org/google-docs/?broken=53YVLY), tibble [(41)](https://www.zotero.org/google-docs/?broken=IuSOLA), tidyverse [(22)](https://www.zotero.org/google-docs/?broken=X09iNc), ggprism [(42)](https://www.zotero.org/google-docs/?broken=gN7iDs), patchwork [(43)](https://www.zotero.org/google-docs/?broken=YQfOfb), ggh4x [(44)](https://www.zotero.org/google-docs/?broken=0Ypd5d), and dplyr [(33)](https://www.zotero.org/google-docs/?broken=3nX36f) packages. The smaller of the two C:N categories was defined as the reference groups for comparisons with larger C:N categories and significantly differentially abundant pathways were filtered according to p < 0.05 and log2FoldChange ± 1. The number of hits for positive and negative differentially abundant pathways were represented in a table and visualized in a log2 fold change plot with the ggpicrust2 package [(45)](https://www.zotero.org/google-docs/?broken=sUiX89)**.**

**RESULTS**

**Forest and wetlands soil are compositionally distinct and show opposing trends in microbial diversity.** Beta diversity analysis using Weighted UniFrac revealed that forest and wetland soil samples are compositionally distinct, with samples clustering separately in ordination space (Figure 1A). This result underscores significant differences in the microbial community composition between the two ecosystems. Furthermore, wetland soil exhibited higher microbial diversity than forest soil, irrespective of the C:N ratio range (Figure 1B).

To examine microbial diversity trends within each ecosystem, Faith’s Phylogenetic Diversity (PD) and Shannon diversity metrics were quantified using R. In forest soil, both Faith’s PD (Figure 2A), and Shannon’s Evenness (Figure 2B), decreased as the C:N ratio increased, indicating a decline in phylogenetic diversity and evenness with higher C:N ratios. Conversely, in wetland soil, Faith’s PD increased with rising C:N ratios (Figure 2C), suggesting an enrichment of phylogenetic diversity under these conditions. However, Shannon’s Evenness showed no significant variation across the C:N ratio range in wetland soil (Figure 2D).

**Taxonomic abundance patterns align with microbial diversity in wetland soil but diverge in forest soil.** Differential abundance analysis using DESeq2 identified unique amplicon sequence variants (ASVs) across C:N ratio ranges for both ecosystems. For both forest and wetland soil, the high C:N ratio range consistently exhibited a greater abundance of unique taxa compared to the low C:N range (Table 1). In wetland soil, these taxonomic trends correlated with increasing microbial diversity at higher C:N ratios, as indicated by Faith’s Phylogenetic Diversity (PD) (Figure 2C). However, in forest soil, despite lower microbial diversity at higher C:N ratios (Figure 2A-B), the high C:N range supported a greater abundance of unique taxa.

**Indicator Species Analysis revealed a significant presence of common indicator species across select C:N ranges in both forest and wetlands soil samples.** Analysis of indicator species provided further insight into taxonomic patterns across C:N ranges. When comparing the differential abundance of taxa across the C:N ranges of intermediate vs. low for forest samples and high vs. intermediate for wetland samples, relatively few taxa showed significant changes between these categories in each respective ecosystem (Table 1). This was supported by the Indicator Species Analysis, which revealed distinct ecological patterns (Figure 3). In forest soil, 43% of the indicator species were common between intermediate and low C:N ranges (Figure 3A), suggesting similar ecological niches in these ranges. Conversely, in wetland soil, 89% of the indicator species were common between high and intermediate C:N ranges (Figure 3B), reflecting comparable niches in these ranges. For wetlands soil samples specifically, very few, if any, unique indicator species were identified within any of the C:N ranges (Figure 3B).

**Predictive functional analysis reveals contrasting metabolic pathway abundance changes across C:N ratios in forest and wetland soils.** A predictive functional analysis using PICRUSt2 was conducted to evaluate whether changes in the abundance of unique taxonomic groups correspond with shifts in unique metabolic pathways across C:N ranges. In forest soil, the high C:N range led to the downregulation of unique metabolic pathways, with a greater number of pathways downregulated at higher C:N ratios compared to the lower and intermediate C:N ranges (Table 2). In contrast, wetland soil exhibited minor changes in the regulation of unique metabolic pathways across C:N ranges (Table 2), suggesting functional redundancy within the microbial communities across all ranges.

**DISCUSSION**

This study reveals distinct patterns in microbial diversity, taxonomic composition, and functional potential between forest and wetland soils, influenced by varying C:N ratios. Beta diversity analysis highlighted the compositional divergence of microbial communities between these ecosystems, with forest and wetland soils clustering separately in ordination space (Figure 1A). Furthermore, wetland soil samples displayed higher diversity compared to forest soil samples at all C:N ranges (Figure 1B). This finding does not validate previous findings, which highlighted inconsistent diversity differences across forest and wetland samples [(46)](https://www.zotero.org/google-docs/?broken=piiCl3). This difference in finding could be due to the unique wetland environments analyzed.

Beyond the compositional divergence observed between forest and wetland ecosystems, microbial diversity within each ecosystem exhibited contrasting trends: wetland soils supported higher overall diversity and phylogenetic enrichment at elevated C:N ratios (Figure 2C), whereas forest soils displayed a decline in diversity and evenness under the same conditions (Figure 2A-B). These findings align with previous studies, where low C:N ratios in forest soil have been shown to promote bacterial diversity [(47, 48)](https://www.zotero.org/google-docs/?broken=bclKPQ). In contrast, artificially constructed wetlands supplemented with litter containing high C:N ratios demonstrated higher microbial diversities [(49)](https://www.zotero.org/google-docs/?broken=uwFqrT), supporting our observations of phylogenetic enrichment in wetland soils at elevated C:N ratios (Figure 2C). These patterns suggest that the microbial community responses to C:N ratio variations are ecosystem-specific, with forest soils favoring lower C:N ratios for higher diversity and wetlands benefiting from higher C:N ratios.

A notable point regarding the diversity results is that wetlands soil samples did not show significant differences across C:N ranges when using Shannon’s Evenness (Figure 2D), despite showing a trend when using Faith’s Phylogenetic Diversity (Figure 2C). This stability in evenness suggests that while the phylogenetic diversity of species may vary across C:N ranges, the overall balance in how they are distributed within the community does not shift dramatically. The relative proportions of different species are maintained, indicating that no single species or group of species becomes significantly more or less dominant. This is further emphasized by the result garnered from Indicator Species Analysis, which showed that none of the 3 C:N ranges have high proportions of unique indicator species (Figure 3B). Current literature remains limited regarding the connection between C:N ratio and species indicative of specific C:N ranges. Our result may help elucidate this connection for wetland soil systems, showing that C:N ranges do not drive differentiation of indicator species.

Taxonomic analyses further emphasized the divergent patterns in diversity for both soil types. In wetland soils, the enrichment of microbial diversity at higher C:N ratios (Figure 2C) was accompanied by an increased abundance of unique taxa (Table 1), suggesting a synergistic relationship between taxonomic and phylogenetic diversity in this environment. In contrast, forest soils exhibited a paradox: despite a decline in microbial diversity and evenness at higher C:N ratios (Figure 2A-B), these conditions still supported a greater abundance of unique taxonomic groups compared to other C:N ranges (Table 1).

Predictive functional analyses offered further insight into the divergent characteristics across the two soil types. In forest soil, the high C:N range was marked by the downregulation of most unique metabolic pathways (Table 2). However, of the few upregulated metabolic synthesis pathways at the high C:N range, all corresponded to antibiotic synthesis (Figure S1A). Previous studies have demonstrated that high C:N ratios in forest soils increase carbon availability relative to nitrogen, resulting in greater nitrogen limitation [(3, 50)](https://www.zotero.org/google-docs/?broken=5trH5b). This limitation reduces the diversity of microorganisms capable of surviving under such conditions [(3)](https://www.zotero.org/google-docs/?broken=I9ABa6). Furthermore, it is well-established that soil microbes upregulate antibiotic production pathways when competing for scarce resources, such as nitrogen [(51)](https://www.zotero.org/google-docs/?broken=8nDMpE). This resource-constrained environment may favor a microbial community dominated by highly competitive taxa, which outcompete less specialized organisms. This may explain the observed increase in unique taxonomic groups at the high C:N range (Table 1) in forest soil, suggesting that specific taxa thrive through competitive mechanisms like antibiotic production.This phenomenon may also explain why diversity (Figure 2A-B) decreases at the high C:N range, as large proportions of closely related taxa could be dominating the microbial community.

Conversely, wetland soil samples displayed functional stability across C:N ranges, with minimal changes in the abundance of unique pathways (Table 2), indicative of functional redundancy within the microbial community. This stability is further supported by the absence of significant differences in indicator species across the C:N ranges (Figure 3B), suggesting that no particular taxa are uniquely associated with any of the three C:N categories. Such an absence of distinct indicator species, in conjunction with no significant differences in Shannon's Evenness (Figure 2D), points to a balanced, consistent microbial community, where species are evenly distributed and functional roles are maintained across C:N ranges. Such functional redundancy highlights the adaptability of wetland microbes, and may serve to ensure the maintenance of critical ecological functions despite changing C:N conditions. Previous studies have similarly emphasized the finely balanced nutrient cycling and energy flow dynamics characteristic of wetland ecosystems [(52)](https://www.zotero.org/google-docs/?broken=yNsLKs). Together, these findings may work synergistically to underscore the buffering capacity of wetland microbial communities, enabling them to uphold ecosystem stability and resilience despite fluctuations in nutrient availability.

**Study limitations.** The main limitation of this study was that the datasets were acquired from two separate sources [(11)](https://www.zotero.org/google-docs/?broken=6eOvyf). Analyzing two different datasets introduces the possibility of confounding environmental factors, differences in sample size, time of collection, and sample collection methodology. Specifically in terms of the methodological differences for the two datasets, it remains unknown. This is because the wetlands dataset by Balantine *et al.* was never published in a primary research paper. This introduces ambiguities in sample collection methodology of the wetlands dataset and makes it difficult to draw conclusions relating the two unique environments, if sampling methods were significantly different. Comparing the different datasets, there were also inconsistencies in sample sizes. The forest soil had a much larger sample size (574 total samples) in comparison to wetlands soil (105 total samples) after filtering out “NA” and removing samples with C:N ranges categorized as “Very Low” and “Very High”. A different sampling depth was also used for each specific dataset. For the forest soil dataset, a sampling depth of 2,500 was used. For the wetlands dataset, a sampling depth of 20,000 was chosen. While merging the datasets with a standardized sampling depth (5,000) helped ensure consistency, it could not fully address the imbalance in statistical representation between the two environments.

Another limitation of our study was being unable to account for confounding variables. Previous literature has shown that other factors included in our metadata, such as pH and soil depth, have also been linked to impacting microbial diversity in soil [(48, 53–55)](https://www.zotero.org/google-docs/?broken=o5DSSu). Being unable to account for these variables restricts the ability to make strong conclusions about the findings and connections seen. Furthermore, the taxonomic results gathered show many uncultured groups or taxa that are poorly resolved in their taxonomic assignment. This limited the study’s ability to investigate the roles of organisms found to be related to specific C:N ranges, particularly for forest soil. These organisms cannot be identified through the use of our tools until classification and documentation of them has occurred.

**Conclusion.** This study confirms that C:N ratio variations exert ecosystem-specific effects on microbial diversity, taxonomic composition, and functional potential in forest and wetland soils. The results show that in forest soils, the high C:N range decreases phylogenetic diversity and functional pathway diversity, while promoting competitive dynamics through upregulation of antibiotic synthesis pathways. Conversely, wetland soils exhibit increased phylogenetic diversity at higher C:N ratios and static functional pathway regulation across C:N ranges. This highlights functional redundancy despite taxonomic shifts. These findings demonstrate that nutrient availability, as reflected by C:N ratio, drives distinct microbial community structures and functions in these contrasting soil ecosystems. This work underscores the importance of nutrient dynamics in shaping microbial ecology and provides a framework for future comparative studies.

**Future directions.** To address a major limitation of this study—the potential influence of confounding variables on the results—future research that examines the role of C:N ratios on microbial composition while controlling for these variables would strengthen the link between C:N ratios and the observed findings. If the abundance of samples permits this, both datasets used in this study contain several factors that could be used as controls, such as pH, elevation, humidity, and soil depth.

Having observed minimal changes in the regulation of metabolic pathways across C:N ranges in wetland soil samples, this warrants further investigation into the common metabolic activities of the microbial communities regardless of C:N ranges. This study focused on elucidating the patterns of metabolic regulation in correlation to C:N ranges, however, investigating the specific roles of these pathways in the maintenance of the wetland ecosystem could strengthen and further explain our results.

Although the upregulation of antibiotic synthesis pathways was observed in forest soil samples at high C:N ranges, this finding is based on predictive functional analyses and requires experimental validation to confirm whether significant antibiotic production is indeed occurring in the soil environment under these conditions. Experimental validation could involve direct measurements of antibiotic compounds in the soil, such as through metabolite profiling [(56)](https://www.zotero.org/google-docs/?broken=0W5tg0), or assessing microbial activity related to antibiotic production using culture-based techniques or gene expression assays. This validation is essential to establish a stronger causal link between high C:N ratios and the competitive shifts observed within the microbial community. Such shifts, as inferred from predictive data, suggest heightened microbial competition. However, without experimental confirmation of active antibiotic production, the extent to which these pathways contribute to the observed community dynamics remains speculative. By confirming this connection, future research could validate the impact of antibiotic synthesis on microbial communities in forest soil. Experimental approaches should include gene knockout studies targeting antibiotic synthesis in high C:N environments. Monitoring changes in microbial composition over time following these genetic interventions could offer valuable insights into the ecological roles and impacts of these pathways.

**ACKNOWLEDGEMENTS**

This research was conducted at the University of British Columbia, located on the traditional, ancestral, and unceded territory of the Musqueam People. We sincerely thank Dr. Evelyn Sun, Rituparna Banerjee, and the MICB 475 teaching team for their mentorship and guidance throughout this project. We are also grateful to the UBC Department of Microbiology and Immunology for providing funding and resources that made this work possible. Additionally, we extend our gratitude to Balantine *et al.* and Wilhelm *et al.* for sharing the datasets utilized in our study. There are no conflicts of interest.

**CONTRIBUTIONS**

**RK:** Carried out QIIME2 pipeline, R code and helped make figure legends for Figures 1-3, Table 1-2, and Supplementary Figure 1, wrote the methods, helped AS write the abstract, introduction, results, discussion, limitations, conclusion, future directions, and added citations.

**AS:** Involved in generation of figures, including figure and table titles for Figures 1-3, Table 1-2, and Supplementary Figure 1. Carried out the writing of the descriptions of the figures and tables. The writing of the manuscript components: Abstract, Introduction, Results, Discussion, Conclusion, Limitations, Future Studies, and Conclusion was done in equal collaboration with RK.

**KS:** Helped modify Figure 3 coding and appearance, and helped write Introduction, results, discussion and edit the final manuscript.

**AK:** Helped write results for Table 1 and Table 2, conclusion, study limitations, and helped edit the final manuscript.

**MA:** Helped with the Figure 2 and Table 1 legends, and helped write the Introduction and limitation sections.

**DATA AVAILABILITY**

The forest soil dataset can be accessed from the Short Read Archive under accession PRJEB8599 [(11)](https://www.zotero.org/google-docs/?broken=pbuy6P). Scripts used to generate data via QIIME2 and RStudio are available at <https://github.com/arshsharma14/Team_8.git>**.**

**REFERENCES**

1. [**Brust GE.** 2019. Chapter 9 - Management Strategies for Organic Vegetable Fertility, p. 193–212. *In* **Biswas, D, Micallef, SA** (eds.), Safety and Practice for Organic Food*. Academic Press*.](https://www.zotero.org/google-docs/?broken=zNpw1z)
2. [**Wan X, Huang Z, He Z, Yu Z, Wang M, Davis MR, Yang Y.** 2015. Soil C:N ratio is the major determinant of soil microbial community structure in subtropical coniferous and broadleaf forest plantations. *Plant Soil* **387**:103–116.](https://www.zotero.org/google-docs/?broken=j9G6Rl)
3. [**Tg I, Haq I, Kalamdhad AS**. 2022. 14 - Factors affecting anaerobic digestion for biogas production: a review, p. 223–233. *In* **Hussain, C, Hait, S** (eds.), Advanced Organic Waste Management. *Elsevier*.](https://www.zotero.org/google-docs/?broken=NcuX5U)
4. [**Ostrowska A, Porębska G.** 2015. Assessment of the C/N ratio as an indicator of the decomposability of organic matter in forest soils. *Ecol Indic* **49**:104–109.](https://www.zotero.org/google-docs/?broken=omM5qG)
5. [**Schoenholtz SH, Van Miegroet H, Burger J. 2000.** A review of chemical and physical properties as indicators of forest soil quality: challenges and opportunities. *For Ecol Manag* **138**:335–356.](https://www.zotero.org/google-docs/?broken=KSCql1)
6. [**Truu M, Juhanson J, Truu J.** 2009. Microbial biomass, activity and community composition in constructed wetlands. *Sci Total Environ* **407**:3958–3971.](https://www.zotero.org/google-docs/?broken=Ynkv9x)
7. [**Faulkner SP, Richardson CJ.** 1989. Physical and Chemical Characteristics of Freshwater Wetland SoilsConstructed Wetlands for Wastewater Treatment. CRC Press.](https://www.zotero.org/google-docs/?broken=JmJsYD)
8. [**Bodelier PLE, Dedysh SN.** 2013. Microbiology of wetlands. *Front Microbiol* **4**:79.](https://www.zotero.org/google-docs/?broken=VygTp0)
9. [**Chen Z, Zhang C, Liu Z, Song C, Xin S.** 2023. Effects of Long-Term (17 Years) Nitrogen Input on Soil Bacterial Community in Sanjiang Plain: The Largest Marsh Wetland in China. *Microorganisms* **11**:1552.](https://www.zotero.org/google-docs/?broken=Ue5AuI)
10. [**Baldrian P. 2017.** Forest microbiome: diversity, complexity and dynamics. *FEMS Microbiol Rev* **41**:109–130.](https://www.zotero.org/google-docs/?broken=O8qjXx)
11. [**Wilhelm RC, Cardenas E, Maas KR, Leung H, McNeil L, Berch S, Chapman W, Hope G, Kranabetter JM, Dubé S, Busse M, Fleming R, Hazlett P, Webster KL, Morris D, Scott DA, Mohn WW.** 2017. Biogeography and organic matter removal shape long-term effects of timber harvesting on forest soil microbial communities. *ISME J* **11**:2552–2568.](https://www.zotero.org/google-docs/?broken=a9B8WP)
12. [**Balaji T, Banka A, Kuder D, Sidhu R.** 2023. Microbial diversity and population density is positively correlated in New York State freshwater wetlands. *Undergrad J Exp Microbiol Immunol* **28**.](https://www.zotero.org/google-docs/?broken=VUIeMd)
13. [Carbon:Nitrogen. Plant Divers Co. https://www.coversandco.ca/carbon-to-nitrogen. Retrieved 24 October 2024.](https://www.zotero.org/google-docs/?broken=B8Bv9D)
14. [**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 Y-X, 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, Caporaso JG.** 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**:852–857.](https://www.zotero.org/google-docs/?broken=tax2rW)
15. [**Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP.** 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**:581–583.](https://www.zotero.org/google-docs/?broken=GW7FOK)
16. [**Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO.** 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**:D590–D596.](https://www.zotero.org/google-docs/?broken=61rgPl)
17. [**Abellan-Schneyder I, Matchado MS, Reitmeier S, Sommer A, Sewald Z, Baumbach J, List M, Neuhaus K.** Primer, Pipelines, Parameters: Issues in 16S rRNA Gene Sequencing. *mSphere* **6**:e01202-20.](https://www.zotero.org/google-docs/?broken=QUbl15)
18. [**Bokulich NA, Joseph CML, Allen G, Benson AK, Mills DA.** 2012. Next-Generation Sequencing Reveals Significant Bacterial Diversity of Botrytized Wine. *PLoS ONE* **7**:e36357.](https://www.zotero.org/google-docs/?broken=VTjGhg)
19. **R Core Team.** 2023. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
20. **McMurdie PJ, Holmes S.** 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* **8**:e61217.
21. **Paradis E, Claude J, Strimmer K**. 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* **20:**289–290.
22. [**Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H.** 2019. Welcome to the Tidyverse. J Open Source Softw **4**:1686.](https://www.zotero.org/google-docs/?broken=zw7Fjc)
23. **Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO.** 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**:1463–1464.
24. [**Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, O’Hara RB, Solymos P, Stevens MHH, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Carvalho G, Chirico M, De Caceres M, Durand S, Evangelista HBA, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill MO, Lahti L, McGlinn D, Ouellette M-H, Ribeiro Cunha E, Smith T, Stier A, Ter Braak CJF, Weedon J.** 2001. vegan: Community Ecology Package.](https://www.zotero.org/google-docs/?broken=UzmYou)
25. [**Shannon CE.** 1948. A mathematical theory of communication. *Bell Syst Tech* J **27:**379–423.](https://www.zotero.org/google-docs/?broken=Cb7hQZ)
26. [**Faith DP.** 1992. Conservation evaluation and phylogenetic diversity. *Biol Conserv* **61:**1–10.](https://www.zotero.org/google-docs/?broken=nYtRx9)
27. **Wickham H.** 2009. ggplot2: Elegant Graphics for Data Analysis. Springer, New York, NY. https://link.springer.com/10.1007/978-0-387-98141-3. Retrieved 6 December 2023.
28. [**Kruskal WH, Wallis WA.** 1952. Use of Ranks in One-Criterion Variance Analysis. J Am Stat Assoc.](https://www.zotero.org/google-docs/?broken=lFDCqM)
29. [**Ahlmann-Eltze C, Patil I.** 2021. ggsignif: R Package for Displaying Significance Brackets for “ggplot2.” OSF https://doi.org/10.31234/osf.io/7awm6.](https://www.zotero.org/google-docs/?broken=l46fRL)
30. [**Lozupone C, Knight R.** 2005. UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. *Appl Environ Microbiol* **71:**8228–8235.](https://www.zotero.org/google-docs/?broken=XM218z)
31. **Anderson.** 2001.[A new method for non‐parametric multivariate analysis of variance.](https://www.zotero.org/google-docs/?broken=y5cQ4M) *Austral Ecol* **26**:32–46.
32. **Wickham H.** 2007.Reshaping Data with the reshape Package. *J Stat Softw* **21**:1–20.
33. **Wickham H, François R, Henry L, Müller K.** 2022. dplyr: A Grammar of Data Manipulation.
34. [**Dunn OJ.** 1964. Multiple Comparisons Using Rank Sums. Technometrics 6:241–252.](https://www.zotero.org/google-docs/?broken=G6K04G)
35. [**Dinno A**. 2024. dunn.test: Dunn’s Test of Multiple Comparisons Using Rank Sums (1.3.6).](https://www.zotero.org/google-docs/?broken=DQEbQl)
36. [**Love MI, Huber W, Anders S.** 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol **15**:550.](https://www.zotero.org/google-docs/?broken=nUQA8l)
37. **Cáceres MD, Legendre P.** 2009. Associations between species and groups of sites: indices and statistical inference. *Ecology* **90**:3566–3574.
38. [**Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MGI.** 2020. PICRUSt2 for prediction of metagenome functions. *Nat Biotechnol* **38:**685–688.](https://www.zotero.org/google-docs/?broken=HKuRZ7)
39. [MetaCyc database of metabolic pathways and enzymes - a 2019 update | Nucleic Acids Research | Oxford Academic. https://academic.oup.com/nar/article/48/D1/D445/5581728. Retrieved 9 December 2024.](https://www.zotero.org/google-docs/?broken=KyxZjH)
40. [**Wickham H, Hester J, Bryan J.** 2015. readr: Read Rectangular Text Data.](https://www.zotero.org/google-docs/?broken=8hZFmZ)
41. **Müller K, Wickham H.** 2023. tibble: Simple Data Frames.
42. [**Dawson C.** 2021. ggprism: A “ggplot2” Extension Inspired by “GraphPad Prism.”](https://www.zotero.org/google-docs/?broken=s3epSb)
43. [**Pedersen TL**. 2024. patchwork: The Composer of Plots (1.3.0).](https://www.zotero.org/google-docs/?broken=rLY0Gd)
44. [**Brand T van den.** 2024. ggh4x: Hacks for “ggplot2” (0.2.8).](https://www.zotero.org/google-docs/?broken=p8v31V)
45. **Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MGI.** 2020. PICRUSt2 for prediction of metagenome functions. 6. *Nat Biotechnol* **38**:685–688.
46. [**Wang X, Zhang Z, Yu Z, Shen G, Cheng H, Tao S.** 2020. Composition and diversity of soil microbial communities in the alpine wetland and alpine forest ecosystems on the Tibetan Plateau. *Sci Total Environ* **747**:141358.](https://www.zotero.org/google-docs/?broken=lbwO4k)
47. [**Shen F, Wu J, Fan H, Liu W, Guo X, Duan H, Hu L, Lei X, Wei X.** 2019. Soil N/P and C/P ratio regulate the responses of soil microbial community composition and enzyme activities in a long-term nitrogen loaded Chinese fir forest. *Plant Soil* **436:**91–107.](https://www.zotero.org/google-docs/?broken=4yeudv)
48. [**Delgado-Baquerizo M, Reich PB, Khachane AN, Campbell CD, Thomas N, Freitag TE, Abu Al-Soud W, Sørensen S, Bardgett RD, Singh BK**. 2017. It is elemental: soil nutrient stoichiometry drives bacterial diversity. *Environ Microbiol* **19**:1176–1188.](https://www.zotero.org/google-docs/?broken=rkhwbL)
49. [**Ping Y, Pan X, Li W, Wang J, Cui L.** 2019. The soil bacterial and fungal diversity were determined by the stoichiometric ratios of litter inputs: evidence from a constructed wetland. *Sci Rep* **9**:13813.](https://www.zotero.org/google-docs/?broken=6xBohQ)
50. [**Tian Q, Wang X, Wang D, Wang M, Liao C, Yang X, Liu F**. 2017. Decoupled linkage between soil carbon and nitrogen mineralization among soil depths in a subtropical mixed forest. *Soil Biol Biochem* **109**:135–144.](https://www.zotero.org/google-docs/?broken=vP1ATx)
51. [**Newman DJ, Cragg GM, Snader KM**. 2003. Natural products as sources of new drugs over the period 1981-2002. *J Nat Prod* **66**:1022–1037.](https://www.zotero.org/google-docs/?broken=2Org0S)
52. [**Wang C, Yu J, Zhang J, Zhu B, Zhao W, Wang Z, Yang T, Yu C**. 2024. A review of factors affecting the soil microbial community structure in wetlands. *Environ Sci Pollut Res* **31**:46760–46768.](https://www.zotero.org/google-docs/?broken=QOO5wP)
53. [**Kaiser K, Wemheuer B, Korolkow V, Wemheuer F, Nacke H, Schöning I, Schrumpf M, Daniel R**. 2016. Driving forces of soil bacterial community structure, diversity, and function in temperate grasslands and forests. *Sci Rep* **6**:33696.](https://www.zotero.org/google-docs/?broken=BIOZQG)
54. [**Nacke H, Thürmer A, Wollherr A, Will C, Hodac L, Herold N, Schöning I, Schrumpf M, Daniel R.** 2011. Pyrosequencing-Based Assessment of Bacterial Community Structure Along Different Management Types in German Forest and Grassland Soils. *PLOS ONE* **6**:e17000.](https://www.zotero.org/google-docs/?broken=5LTJNf)
55. [**Fierer N, Jackson RB**. 2006. The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci* **103**:626–631.](https://www.zotero.org/google-docs/?broken=EuEOmh)
56. [**Clarke CJ, Haselden JN**. 2008. Metabolic profiling as a tool for understanding mechanisms of toxicity. *Toxicol Pathol* **36**:140–147.](https://www.zotero.org/google-docs/?broken=YdGok7)

**TABLES**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Quantities of Unique ASVs** | | | |
| **Forest soil** | | **Freshwater Wetland soil** | |
|  | **Increased Abundance** | **Decreased Abundance** | **Increased Abundance** | **Decreased Abundance** |
| **High vs**  **Low** | 124 | 96 | 773 | 407 |
| **High vs Intermediate** | 81 | 30 | 115 | 69 |
| **Intermediate vs Low** | 16 | 18 | 746 | 698 |

**Table 1.** **Increased abundance of distinct ASVs is observed at higher C:N ratios in both ecosystems.** The table presents the number of amplicon sequence variants (ASVs) with increased or decreased abundance in pairwise comparisons between low, intermediate, and high carbon-to-nitrogen (C:N) ratios for both soil types. The second group listed per pairwise comparison serves as the reference group. Colour intensity corresponds to the magnitude of unique ASVs, with darker shades indicating higher quantities. Statistical thresholds for the analysis are set at log2FoldChange ± 2 and p\_adj < 0.01.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Quantities of Unique Metabolic Pathways** | | | |
| **Forest soil** | | **Freshwater Wetland soil** | |
|  | **Upregulated** | **Downregulated** | **Upregulated** | **Downregulated** |
| **High vs**  **Low** | 6 | 31 | 2 | 1 |
| **High vs Intermediate** | 2 | 23 | 0 | 4 |
| **Intermediate vs Low** | 2 | 1 | 7 | 1 |

**Table 2.** **Downregulation of unique metabolic pathways seen at high C:N range in forest soil samples. wetland soil samples show similar pathways across C:N ranges.** The table presents the number of unique metabolic pathways with increased or decreased abundance in pairwise comparisons between low, intermediate, and high carbon-to-nitrogen (C:N) ratios. The second group listed per pairwise comparison serves as the reference group. Differential abundance analysis was conducted with statistical thresholds set at log2FoldChange ± 1 and p < 0.05. Colour intensity corresponds to the magnitude of unique pathways, with darker shades indicating higher quantities.

**FIGURE CAPTIONS**

|  |  |
| --- | --- |
| **A** | **B** |

**Figure 1. Forest and wetland soil are significantly different in terms of beta diversity regardless of C:N range.** Weighted Unifrac was measured across forest soil and wetlands soil samples with their associated carbon-to-nitrogen (C:N) ranges of low, intermediate, and high and visualized as a Principal Coordinate Analysis (PCoA) plot (A) and a box plot (B). Statistical comparisons for the PCoA plot (A) were conducted using Permutational multivariate analysis of variance (PERMANOVA). Ellipses (A) represent a 95% confidence interval. Statistical comparisons for the box plot (B) were conducted using the Kruskal-Wallis and Dunn’s tests indicated by asterisks (\*\*\* = p < 0.001). Colours denote soil ecosystem and respective C:N categories

|  |  |  |
| --- | --- | --- |
| **A** | **B** |  |
| **C** | **D** |  |

**Figure 2. Alpha diversity metrics reveal opposing trends between ecosystems as C:N ratio increases.** Faith’s Phylogenetic Diversity (A, C) and Shannon’s Evenness (B, D) were measured for forest soil (A, B) and freshwater wetlands soil (C, D) samples. All diversity comparisons are conducted across carbon-to-nitrogen (C:N) ranges of low, intermediate, and high within each soil type. Statistical comparisons were conducted using the Kruskal-Wallis test, with significant pairwise differences between C:N categories (Low, Intermediate, and High) indicated by asterisks (\* = p < 0.05, \*\* = p < 0.01, \*\*\* = p < 0.001, \*\*\*\* = p < 0.0001). Colours denote C:N categories.

|  |  |
| --- | --- |
| **A** | **B** |

**Figure 3. Indicator Species Analysis displays high common indicator species across select C:N ranges in forest and wetlands soil samples.** Indicator taxa for forest soil (A) and wetlands soil (B) were identified, filtered for those that were significant (p < 0.05), and the quantities were visualized as venn diagrams. The indicator taxa were resolved to the genus level.