Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Leaf harvesting and seasonal flooding are associated with distinct soil bacterial communities in carnauba palm groves

Elane Bezerra Silva
Universidade Federal do Ceará
Francisco Matheus Medeiros Freitas
Universidade Federal do Ceará
Murilo Sousa Almeida
Universidade Federal do Ceará
Kaio Gráculo Vieira Garcia
Universidade Federal do Ceará
Larissa dos Santos Lopes
Universidade Federal do Ceará
Vanessa Nessner Kavamura
Rothamsted International
Vânia Maria Maciel Melo
Universidade Federal do Ceará
Erika Valente de Medeiros
Universidade Federal do Agreste de Pernambuco (UFAPE)
Ademir Sérgio Ferreira Araujo
Universidade Federal do Piauí (UFPI)
Arthur Prudêncio Araujo Pereira
arthur.prudencio@ufc.br
Universidade Federal do Ceará

Article

Keywords: Caatinga, carnauba palm, extractive management, seasonal flooding, soil bacteriome, Solonetz

Posted Date: November 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11515905/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Leaf harvesting and seasonal flooding are associated with distinct soil bacterial communities in carnauba palm groves

Elane Bezerra da Silva[1], Francisco Matheus Medeiros de Freitas[2], Murilo de Sousa Almeida[3], Kaio Gráculo Vieira Garcia[4], Larissa dos Santos Lopes[5], Vanessa Nessner Kavamura[6], Vânia Maria Maciel Melo[7], Erika Valente de Medeiros[8], Ademir Sérgio Ferreira Araujo[9] and Arthur Prudêncio de Araujo Pereira[10*]

[1] Departamento de Ciências do Solo, Universidade Federal do Ceará, Campus do Pici, bloco 807 - Av. Mister Hull, s/n - Pici, Fortaleza - CE, 60455-760, Brasil. ORCID: https://orcid.org/0000-0002-3224-7791
[2] Departamento de Solos, Universidade Federal Rural do Rio de Janeiro, Rodovia BR 467, Km 07, Seropédica – RJ, 23890-000. ORCID: https://orcid.org/0009-0005-7409-0269
[3] Departamento de Ciências do Solo, Universidade Federal do Ceará, Campus do Pici, bloco 807 - Av. Mister Hull, s/n - Pici, Fortaleza - CE, 60455-760, Brasil. ORCID: https://orcid.org/0000-0002-4714-631X
[4] Departamento de Ciências do Solo, Universidade Federal do Ceará, Campus do Pici, bloco 807 - Av. Mister Hull, s/n - Pici, Fortaleza - CE, 60455-760, Brasil. ORCID: https://orcid.org/0000-0003-4980-5136
[5] Pontes Indústria Cera de Carnaúba, Rua Irineu de Souza, 10 - Álvaro Weyne, Fortaleza – CE, CEP 60337-180, Brasil.
[6] Rothamsted Research, West Common Harpenden, United Kingdon AL5 2JQ. ORCID: https://orcid.org/0000-0001-7111-7116
[7] Departamento de Biologia, Universidade Federal do Ceará, Campus do Pici - Av. Mister Hull, s/n - Pici, Fortaleza - CE, 60455-760, Brasil. ORCID: https://orcid.org/0000-0002-3302-8857
[8] Universidade Federal do Agreste de Pernambuco (UFAPE), Avenida Bom Pastor, s/n, Boa Vista, Garanhuns, Pernambuco, 55292-278, Brazil. ORCID: 0000-0001-5543-9414
[9] Universidade Federal do Piauí (UFPI), Campus Universitário Ministro Petrônio Portella, Bairro Ininga, Teresina, Piauí, 64049-550, Brazil. ORCID: 0000-0002-3212-3852
[10] Departamento de Ciências do Solo, Universidade Federal do Ceará, Campus do Pici, bloco 807 - Av. Mister Hull, s/n - Pici, Fortaleza - CE, 60455-760, Brasil. ORCID: https://orcid.org/0000-0001-9402-3243

***** Correspondence: arthur.prudencio@ufc.br

Abstract

Carnauba palm groves occupy seasonally dry floodplains where extractive management and hydrological variation may jointly shape surface-soil microbiota. We compared bacterial communities across four field conditions defined by leaf harvesting (managed/unmanaged) and water regime (seasonally flooded/non-flooded) in Caucaia,

Ceará, northeastern Brazil. Surface soil (0–10 cm) was sampled beneath six palms per area, and the 16S rRNA V4 region was sequenced on an Illumina MiSeq platform. Community composition differed significantly among conditions (PERMANOVA R^2 = 0.431, p = 0.001), with the managed, seasonally flooded area the most compositionally distinct. This area also showed the lowest estimated richness and markedly higher Na, exchangeable sodium percentage, S, Ca, Mg, clay content, and cation exchange capacity than the others. Shannon and Simpson diversity were lower under managed and flooded conditions. Bacillota and Bacillales were proportionally enriched in managed soils, while Acidobacteriota predominated in unmanaged soils. LEfSe and niche-occupancy analyses identified condition-associated taxa and a higher proportion of generalists in unmanaged areas. Exploratory co-occurrence networks varied in size and topology, though interpretation is constrained by small within-condition sample sizes. Since each management–water combination was represented by a single area, findings reflect site-condition contrasts rather than independent causal effects of harvesting or flooding. Land use, hydrology, and edaphic properties must therefore be considered jointly when assessing bacterial communities in semiarid floodplain soils.

Keywords: Caatinga; carnauba palm; extractive management; seasonal flooding; soil bacteriome; Solonetz.

Introduction

The Caatinga is a seasonally dry tropical biome restricted to Brazil. Its strong spatial and temporal variation in rainfall, soil properties, and resource availability creates heterogeneous habitats in which water availability exerts a major control on ecological processes [1,2]. Soil microorganisms mediate nutrient transformations and organic-matter turnover in these environments and can respond rapidly to natural and anthropogenic disturbance [3,4].

Carnauba (*Copernicia prunifera*) is native to northeastern Brazil and has ecological, economic, and cultural importance for regional smallholders. The palm commonly occurs on alluvial plains and other low-lying landscapes that experience pronounced seasonality, with prolonged dry periods interrupted by concentrated rainfall and seasonal flooding [5, 6]. Its leaves are harvested mainly during the dry season for wax production [7]. This

extractive practice changes the amount and timing of plant material returned to the soil and may therefore alter belowground resource availability.

Hydrological variation provides a second environmental filter. Drying, rewetting, and flooding change soil aeration, redox potential, substrate diffusion, and nutrient availability, thereby selecting taxa with different physiological strategies [8, 9, 10]. Seasonally flooded soils may favor anaerobic and facultatively anaerobic groups, whereas aerated soils support communities adapted to oxidative conditions [11]. Land use and season likewise restructure bacterial communities in Caatinga soils [12], but little is known about how these filters coincide in native carnauba groves.

Leaf harvesting and flooding may also interact indirectly through soil properties. Reduced or redistributed litter inputs can modify organic-matter supply and nutrient cycling [13], while recurrent saturation can redistribute salts and alter exchange chemistry. These effects are especially relevant in semiarid floodplains, where salinity, texture, and water availability can vary over short distances. Consequently, microbial differences attributed to management or hydrology must be interpreted together with the edaphic setting of each site.

We evaluated surface-soil bacterial communities under four field conditions combining leaf harvesting and seasonal flooding. We expected the managed, seasonally flooded condition to show the strongest shift in community composition and lower alpha diversity, accompanied by enrichment of stress-tolerant taxa. We also explored whether co-occurrence patterns and niche occupancy differ among conditions. The study was designed as a regional field comparison; therefore, its primary inference concerns the four sampled site conditions rather than isolated causal effects of management and flooding.

Results

Surface-soil properties

The four areas differed markedly in surface-soil chemistry and texture (Table 1). The +LM+SF area had the highest pH, S, Ca, Mg, Na, base saturation, exchangeable sodium percentage, clay content, and cation exchange capacity. In particular, Na and exchangeable sodium percentage were approximately one to two orders of magnitude higher in +LM+SF than in the other areas. Organic matter was highest in −LM+NF and lowest in −LM+SF. Because the values are area-specific means, they are reported descriptively rather than tested as independent treatment effects.

Soil bacterial diversity and community structure

PCoA separated the four site conditions, with the first two axes accounting for 25.6% and 11.76% of compositional variation, respectively (Figure 2). The +LM+SF samples formed the most distinct cluster, whereas +LM+NF and −LM+SF partially overlapped. The −LM+NF samples are clustered separately with comparatively low within-group spread. Overall differences among the four conditions were significant (PERMANOVA $R^2 = 0.4307$, $p = 0.001$).

Chao1 richness varied among conditions and showed an exploratory management × water-regime interaction ($F = 3.92$, $p = 0.023$). Richness was highest in −LM+SF and lowest in +LM+SF; +LM+NF and −LM+NF were intermediate and did not differ from the other conditions (Figure 3A).

Shannon diversity was lower in managed than unmanaged samples ($F = 16.71$, $p = 0.00057$) and in seasonally flooded than non-flooded samples ($F = 8.26$, $p = 0.009$; Figures 3B and 3C). Simpson diversity showed the same pattern for management ($F = 15.93$, $p = 0.00072$) and water regime ($F = 6.24$, $p = 0.021$; Figures 3D and 3E). The interaction was not significant for Shannon or Simpson diversity.

Taxonomic composition of soil bacterial communities

The dominant phyla were Pseudomonadota, Actinomycetota, Bacillota, and Acidobacteriota across all four conditions (Figure 4A). Bacillota contributed a larger proportion of reads in managed areas, whereas Acidobacteriota were proportionally more abundant in unmanaged areas.

Alphaproteobacteria dominated at the class level. Bacilli, Actinobacteria, Acidobacteriae, Thermoleophilia, and Planctomycetes were also abundant (Figure 4B). Bacilli and Actinobacteria contributed more reads in managed samples, whereas Acidobacteriae and Thermoleophilia contributed more to unmanaged samples. Clostridia, Anaerolineae, and Myxococcia were proportionally more abundant under seasonally flooded conditions.

Bacillales, Rhizobiales, Gemmatales, and Acidobacteriales were the most represented orders (Figure 4C). Bacillales and Rhizobiales had higher relative abundance in managed areas, particularly under non-flooded conditions, whereas Acidobacteriales and Gemmatales contributed a larger proportion in unmanaged areas.

Condition-associated taxa

LEfSe identified taxa that discriminated among the four conditions at family and genus levels (Figure 5). Xanthobacteraceae were associated with −LM+SF. The +LM+SF condition was characterized by Bacillaceae, Anaerolineaceae, Alicyclobacillaceae, *Bacillus*, *Pullulanibacillus*, *Haloparvum*, and *Syntrophobacter*. Ktedonobacteraceae, Solirubrobacteraceae, *Bryobacter*aceae, *Conexibacter*, *Bryobacter*, and HSB OF53-F07 were associated with +LM+NF. Gemmataceae, Chthoniobacteraceae, *Candidatus Udaeobacter*, and *Bradyrhizobium* were associated with −LM+NF. These results denote relative-abundance contrasts within this dataset and do not by themselves demonstrate ecological function.

Exploratory co-occurrence networks and niche occupancy

Network size and topology varied among conditions (Figure 6). The +LM+SF, +LM+NF, −LM+SF, and −LM+NF networks contained 311, 422, 426, and 476 nodes and 1,469, 1,333, 1,640, and 2,240 edges, respectively. Corresponding mean degrees were 9.45, 6.32, 7.70, and 9.41, and edge densities were 0.0305, 0.0150, 0.0181, and 0.0198. Thus, +LM+SF had the smallest network but the highest density, whereas −LM+NF had the largest network. These descriptive metrics do not support a simple ranking of ecological stability or complexity. In CLAM comparisons, unmanaged areas contained 60.0% generalists, 20.5% flooded specialists, and 19.0% non-flooded specialists (Figure 7A). Managed areas contained 23.9% generalists, 19.9% flooded specialists, and 55.9% non-flooded specialists (Figure 7B).

Discussion

This regional field study identified pronounced differences in surface-soil bacterial communities among four carnauba-grove conditions. The managed, seasonally flooded area was the most compositionally distinct and had the lowest richness, while unmanaged areas contained more generalist ASVs. These patterns were consistent with the working expectation that the combination of harvesting and seasonal saturation would be associated with stronger environmental filtering. The design, however, does not separate those two attributes from other differences among the four areas.

The soil data provides essential context for the microbial patterns. The +LM+SF area had far higher Na and exchangeable sodium percentage, together with higher S, Ca,

Mg, clay content, base saturation, and cation exchange capacity, than the other areas (Table 1). Its distinctive PCoA position therefore likely reflects a combined edaphic template that includes salinity and texture as well as harvesting history and hydrology. Flooding can alter redox conditions, solute movement, and nutrient availability [8], while land use and seasonal variation can restructure dry-forest microbiomes [12]. The strong soil contrasts observed here preclude attributing the microbial separation to leaf removal or flooding alone.

Within this site system, the partial overlap of +LM+NF and −LM+SF suggests that different combinations of land use and hydrology can produce partly similar community configurations. The comparatively tight −LM+NF cluster indicates lower among-sample compositional variation in that area, but a single sampling campaign cannot establish temporal stability. Repeated seasonal sampling would be required to test resistance or resilience directly [14].

Alpha-diversity patterns were also condition dependent. The lowest Chao1 richness occurred in +LM+SF, whereas −LM+SF had the highest richness. This contrast shows that flooding was not uniformly associated with lower richness; its relationship with bacterial diversity depended on the management–site context. Lower Shannon and Simpson diversity in managed and flooded samples is consistent with stronger filtering of relative abundance and evenness under resource alteration, oxygen limitation, or osmotic stress [10, 15, 16]. Because the factorial terms rely on within-area tree samples, they should be interpreted as patterns in the sampled system rather than general effects across carnauba landscapes.

Broad taxonomic composition was conserved across conditions, but relative contributions shifted. Pseudomonadota and Actinomycetota are widespread in soil and encompass diverse decomposers and nutrient-transforming taxa [17, 18]. The larger contribution of Bacillota and Bacillales in managed samples is consistent with the capacity of many members of these groups to form resistant endospores and tolerate environmental fluctuation [19, 20]. The greater relative abundance of Acidobacteriota in unmanaged soils is compatible with their frequent occurrence in acidic, resource-limited soils [21], although phylum-level patterns do not resolve specific functions.

Finer taxonomic contrasts further distinguish the four conditions. The relative enrichment of Anaerolineaceae and *Syntrophobacter* in +LM+SF is compatible with the low-oxygen conditions expected during saturation, whereas *Bacillus* and *Pullulanibacillus* include taxa tolerant of environmental stress [22, 23]. *Candidatus*

Udaeobacter in −LM+NF has been associated with resource-efficient strategies in soil [24, 25]. These ecological interpretations remain hypotheses because 16S amplicons provide taxonomic profiles rather than direct measurements of metabolic activity or biogeochemical rates.

LEfSe identified condition-associated taxa, but its results require restraint. The analysis used relative abundances, nominal p values, and six samples per condition; therefore, it is best viewed as a screening step for taxa that merit validation. In particular, the presence of a taxon associated with N cycling or stress tolerance does not demonstrate that the corresponding process differed among areas. Metagenomic, transcriptomic, or process-rate measurements would be needed to test those functional interpretations.

The exploratory networks also differed in size and topology, but the available metrics do not support the original interpretation that managed systems were uniformly less connected. The +LM+SF network had fewer nodes and edges but the highest edge density, while −LM+NF had the most nodes and edges and a similar mean degree. Network size is sensitive to the number and prevalence of retained ASVs, and correlation edges represent statistical associations rather than ecological interactions [26]. With only six samples per condition and no correction for multiple testing, these networks should be used to generate hypotheses, not to infer stability, resilience, or keystone taxa.

Niche-occupancy analysis provided a clearer condition-level pattern: unmanaged areas contained a greater proportion of generalists, whereas the managed comparison was dominated by ASVs classified as specialists of the non-flooded condition. This distribution is consistent with stronger habitat filtering in the managed site pair [27, 28]. Still, specialist and generalist categories depend on the sampled environmental range and should not be extrapolated beyond these four areas without regional replication.

Conclusions

The four carnauba-grove conditions supported distinct surface-soil bacterial communities. The managed, seasonally flooded area was the most compositionally distinct and had the lowest richness; it also had exceptionally high Na, exchangeable sodium percentage, and several exchange-property values. Unmanaged areas contained more generalist ASVs, and the relative abundance of several bacterial groups varied among conditions. These results show that harvesting history, hydrological setting, and soil properties must be considered together when assessing the carnauba soil bacteriome.

Because each condition was represented by one area, the study does not isolate causal effects of leaf harvesting or flooding. Replicated landscapes, repeated seasonal sampling, and direct functional measurements are needed to test those mechanisms. Within these limits, the study establishes a regional baseline for microbial assessment and identifies salinity and surface-soil heterogeneity as priorities for the management of carnauba floodplain soils.

Material and methods

Study area and soil sampling

The study was conducted in naturally occurring carnauba groves at Domingos Pontes Farm, Caucaia, Ceará, northeastern Brazil (3°07′32″ S, 36°24′37″ W; Figure 1). Mean annual rainfall is approximately 1,243 mm and is concentrated mainly from February to May [29, 30].

Caucaia has a predominantly tropical climate with a dry winter, although local conditions range from humid and subhumid tropical to mild tropical semiarid because of the municipality's geographic extent. The study area is classified as mild tropical semiarid, with typical temperatures of 26–33 °C [31, 30].

The soil is classified as a Solonetz according to the World Reference Base for Soil Resources [32] and as a Natric Planosol in the Brazilian Soil Classification System [33, 34]. Surface-soil chemical and particle-size properties are summarized in Table 1.

Table 1. Surface-soil chemical and particle-size properties (0–10 cm) in four carnauba-grove areas differing in leaf-harvesting status and seasonal flooding.

Soil properties	Unit	Not managed		Managed	
		Flooded	Not flooded	Flooded	Not flooded
pH	-	4.15	4.06	4.76	4.6
P	mg dm^{-3}	7.85	8.17	10.12	7.58
S	mg dm^{-3}	15.12	13.53	335.48	9.47
Ca	$mmol_c$ dm^{-3}	15.51	14.5	105.88	23.47
Mg	$mmol_c$ dm^{-3}	11.43	10.55	61.1	17.63
O.M.	$mmol_c$ dm^{-3}	27.60	32.88	28.32	34.05
K	$mmol_c$ dm^{-3}	2.68	4.77	2.71	4.72
Na	$mmol_c$ dm^{-3}	3.25	2.3	119.8	2.55

Al	$mmol_c\ dm^{-3}$	4.05	8.67	2.97	1.68
H+Al	$mmol_c\ dm^{-3}$	48.68	67.53	45.68	39.83
SB	$mmol_c\ dm^{-3}$	32.88	32.12	289.48	48.35
CEC	$mmol_c\ dm^{-3}$	79.57	99.65	335.17	88.18
V	%	41	31.83	80.17	55
m	%	11.5	22.67	2.17	3.5
ESP	%	4.17	2.33	32.67	2.83
Sand	%	89.96	81.00	72.32	88.63
Silt	%	2.23	9.49	8.21	1.19
Clay	%	7.81	9.51	19.46	10.19

Values are means of six tree-centered samples per area. Analyses followed [35]. OM, organic matter; H+Al, potential acidity; SB, sum of bases; CEC, cation exchange capacity; V, base saturation; m, aluminum saturation; ESP, exchangeable sodium percentage. Methods: pH in $CaCl_2$; OM by sodium dichromate extraction and colorimetry; P and K by ion-exchange resin; H+Al by SMP buffer; Al, Ca, and Mg by 1 mol L^{-1} KCl; S by 0.01 mol L^{-1} calcium phosphate; and Na by Mehlich 1.

The carnauba palms were identified in the field as *Copernicia prunifera* (Mill.) H. E. Moore by Dr Arthur Prudêncio de Araujo Pereira (Federal University of Ceará, Brazil), based on their solitary, erect stems, fan-shaped leaves with a waxy coating, and spiny petiole margins. The species is widely distributed in Ceará, northeastern Brazil. As this study involved soil sampling only and no plant material was collected, no voucher specimen was deposited in a herbarium.

Four areas were selected to represent combinations of two site attributes: leaf harvesting (managed or unmanaged) and water regime (seasonally flooded or non-flooded). This yielded four conditions: managed and seasonally flooded (+LM+SF), managed and non-flooded (+LM+NF), unmanaged and seasonally flooded (−LM+SF), and unmanaged and non-flooded (−LM+NF). Flooded areas accumulate water during the rainy season, whereas non-flooded areas do not. Leaf harvesting is the only reported anthropogenic activity in the managed areas and occurs from August to November; unmanaged areas had not been harvested for at least five years. Six individual palms were georeferenced within each area using Gaia GPS (version 2023.7), giving 24 tree-centered sampling points (Figure 1). Each condition was represented by one area, so the design provides spatial replication among palms within areas but no independent replication of the four area-level conditions.

Soil was collected approximately 15 cm from the stem base of each palm at 0–10 cm depth. Samples were passed through a 2-mm sieve and homogenized after visible roots and debris were removed. The processed soil was stored at 4 °C until analysis.

DNA extraction and DNA sequencing

Total DNA was extracted from 0.5 g of soil with the PowerLyzer PowerSoil DNA Isolation Kit (MoBio Laboratories, Carlsbad, CA, USA) according to the manufacturer's

protocol. DNA purity and concentration were assessed with a NanoDrop 2000 spectrophotometer (Thermo Scientific, Waltham, MA, USA).

The V4 region of the bacterial 16S rRNA gene was amplified with primers 515F (5′-GTGCCAGCMGCCGCGGTAA-3′) and 806R (5′-GGACTACHVGGGTWTCTAAT-3′) [36]. Each 25-µL reaction contained 5.0 µL of 5× buffer with 2 mM $MgCl_2$, 0.75 µL of 10 mM dNTPs, 0.75 µL of each primer (515F at 40 µM and 806R at 10 µM), 0.5 µL (1 U) of Platinum Taq DNA Polymerase High Fidelity (Invitrogen, Carlsbad, CA, USA), 2.0 µL of template DNA, and nuclease-free water to 25 µL.

Amplification comprised 95 °C for 3 min; 35 cycles of 98 °C for 20 s, 55 °C for 20 s, and 72 °C for 30 s; and a final extension at 72 °C for 3 min. Amplicons were purified with Agencourt AMPure XP beads (Beckman Coulter, Brea, CA, USA) and quantified with a Qubit 2.0 fluorometer and the dsDNA BR Assay Kit (Invitrogen). Equimolar amplicons were pooled, adjusted to 2 nM, denatured, diluted to 8 pM, supplemented with 20% PhiX, and sequenced on an Illumina MiSeq platform.

Data analysis

Sequences were processed in QIIME 2 (version 2024.10) [37]. Amplicon sequence variants (ASVs) were inferred with DADA2 after quality filtering, dereplication, error-model estimation, paired-read processing, and chimera removal [38]. Taxonomy was assigned against SILVA release 138 [39] with the classify-consensus-vsearch method, and sequences not assigned to Bacteria were removed. Representative sequences were aligned with MAFFT [40] and a phylogenetic tree was generated with FastTree [41]. The representative sequences, ASV table, taxonomy, and sample metadata were imported into R 4.4.0 [42] and organized with phyloseq [43]. Ecological analyses used vegan [44]. The raw sequencing data generated in this study are available in the NCBI Sequence Read Archive (SRA) under BioProject accession number PRJNA1542438 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1542438/).

Beta diversity was evaluated from an ASV-level dissimilarity matrix. Principal-coordinate analysis (PCoA) summarized compositional variation, and PERMANOVA tested differences among the four site conditions after homogeneity of multivariate dispersion was assessed. Given the absence of independent area-level replication, the PERMANOVA is interpreted as a comparison among the sampled conditions rather than as a causal factorial test of management and water regime.

Alpha diversity was described as Chao1 richness and the Shannon and Simpson indices. Residual normality and variance homogeneity were evaluated with Shapiro–Wilk and Levene tests. When assumptions were met, exploration two-way ANOVA partitioned variation between leaf-harvesting status, water regime, and their interaction at the tree-sample level. Tukey's honestly significant difference test was used for post hoc comparisons when appropriate, with $\alpha = 0.05$. Because management and water regime were not replicated at the area level, these tests quantify within-study contrasts and do not establish independent area-level effects.

Community composition was summarized at phylum, class, and order levels after unassigned taxa were excluded. Relative abundance was calculated within each sample

and summarized by condition. LEfSe, implemented in MicrobiomeAnalyst [45], was used as an exploratory screening method for taxa that discriminated among conditions (Kruskal–Wallis $p < 0.05$; LDA score > 2.0) [46]. Because the test used nominal p values and relative-abundance data, identified taxa are described as condition-associated rather than as confirmed biomarkers or functional indicators.

Exploratory bacterial co-occurrence networks were inferred separately for +LM+SF, +LM+NF, −LM+SF, and −LM+NF with the microeco package in R [47]. The ASV table, taxonomy, and sample metadata were subset consistently for each condition, and low-abundance ASVs (relative abundance < 0.0005) were excluded.

SparCC correlations were calculated within the microeco workflow [48]. Edges were retained when $|r| \geq 0.9$ and the nominal p value was < 0.05; p values were not adjusted for multiple testing. Nodes represent ASVs and edges represent non-directional associations, not direct ecological interactions. Because each network was based on six samples, network topology was used only for exploratory description and was not subjected to inferential comparison among conditions.

Modules were identified with the fast-greedy community-detection algorithm. Networks were exported in GEXF format and visualized in Gephi [49]. Node color denotes phylum, node size is proportional to degree, and edge color distinguishes positive from negative associations.

Niche occupancy was evaluated with the multinomial species-classification method implemented by the clamtest function in vegan. The supermajority rule was applied with K = 2/3 and $p < 0.005$ [50]. Pairwise analyses compared flooded and non-flooded conditions separately within unmanaged and managed areas.

Data availability

The raw sequencing data generated in this study are available in the NCBI Sequence Read Archive (SRA) under BioProject accession number PRJNA1542438 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1542438/).

References

1. Leal, I. R., Silva, J. M. C., Tabarelli, M. & Lacher, T. E. Jr. Changing the course of biodiversity conservation in the Caatinga of northeastern Brazil. *Conserv. Biol.* **19**, 701–706 (2005). https://doi.org/10.1111/j.1523-1739.2005.00703.x

2. Silva, J. M. C., Barbosa, L. C. F., Leal, I. R. & Tabarelli, M. The Caatinga: Understanding the challenges. *Biodivers. Conserv.* **26**, 2317–2326 (2017). https://doi.org/10.1007/978-3-319-68339-3_1

3. Mendes, L. W., Tsai, S. M., Navarrete, A. A., Holland, T. C. & Kuramae, E. E. Soil-borne microbiome: Linking diversity to function. *Microb. Ecol.* **70**, 255–265 (2015).https://doi.org/10.1007/s00248-014-0559-2

4. Araujo, A. S. F., de Medeiros, E. V., da Costa, D. P., de Araujo Pereira, A. P. & Mendes, L. W. From desertification to restoration in the Brazilian semiarid region: Unveiling the potential of land restoration on soil microbial properties. *J. Environ. Manage.* **351**, 119746 (2024). https://doi.org/10.1016/j.jenvman.2023.119746

5. Prado, D. E. As caatingas da América do Sul: Histórico, conceitos e vegetação atual. In *Ecologia e conservação da Caatinga* (eds Leal, I. R., Tabarelli, M. & Silva, J. M. C.) 3–73 (Editora Universitária da UFPE, 2003).

6. Queiroz, L. P., Cardoso, D., Fernandes, M. F. & Moro, M. F. Diversity and evolution of flowering plants of the Caatinga domain. *J. Syst. Evol.* **55**, 229–246 (2017). https://doi.org/10.1007/978-3-319-68339-3_2

7. Gomes, R. L. F. & Nascimento, I. B. A cadeia produtiva da carnaúba no Nordeste brasileiro. *Rev. Econ. Nordeste* **37**, 353–370 (2006).

8. Ponnamperuma, F. N. The chemistry of submerged soils. Adv. *Agron.* **24**, 29–96 (1972). https://doi.org/10.1016/S0065-2113(08)60633-1

9. Placella, S. A., Brodie, E. L. & Firestone, M. K. Rainfall-induced carbon dioxide pulses result from sequential resuscitation of phylogenetically clustered microbial groups. *Proc. Natl. Acad. Sci. USA* **109**, 10931–10936 (2012). https://doi.org/10.1073/pnas.1204306109

10. Manzoni, S., Schimel, J. P. & Porporato, A. Responses of soil microbial communities to water stress: Results from a meta-analysis. *Ecology* **93**, 930–938 (2012). https://doi.org/10.1890/11-0026.1

11. Conrad, R. Microbial ecology of methanogens and methanotrophs. *Adv. Agron.* **96**, 1–63 (2007). https://doi.org/10.1016/S0065-2113(07)96005-8

12. Lacerda-Júnior, G. V., Noronha, M. F., Cabral, L., Delforno, T. P., de Sousa, S. T. P., Fernandes-Júnior, P. I. & Oliveira, V. M. Land use and seasonal effects on the soil microbiome of a Brazilian dry forest. *Front. Microbiol.* **10**, 648 (2019). https://doi.org/10.3389/fmicb.2019.00648

13. Berg, B. Decomposition patterns for foliar litter—A theory for influencing factors. *Soil Biol. Biochem.* **78**, 222–232 (2014). https://doi.org/10.1016/j.soilbio.2014.08.005

14. Allison, S. D. & Martiny, J. B. Resistance, resilience, and redundancy in microbial communities. *Proc. Natl. Acad. Sci. USA* **105**, 11512–11519 (2008). https://doi.org/10.1073/pnas.080192510

15. Schimel, J. P. Life in dry soils: Effects of drought on soil microbial communities and processes. Annu. *Rev. Ecol. Evol. Syst.* **49**, 409–432 (2018). https://doi.org/10.1146/annurev-ecolsys-110617-062614

16. Ansari, J., Bardhan, S., Eivazi, F., Anderson, S. H. & Mendis, S. S. Bacterial community diversity for three selected land use systems as affected by soil moisture regime. *Appl. Soil Ecol.* **192**, 105100 (2023). https://doi.org/10.1016/j.apsoil.2023.105100

17. Delgado-Baquerizo, M. et al. A global atlas of the dominant bacteria found in soil. *Science* **359**, 320–325 (2018). https://10.1126/science.aap9516

18. Viruel, E., Fontana, C. A., Puglisi, E., Nasca, J. A., Banegas, N. R. & Cocconcelli, P. S. Land-use change affects the diversity and functionality of soil bacterial communities in semi-arid Chaco region, Argentina. *Appl. Soil Ecol.* **172**, 104362 (2022). https://doi.org/10.1016/j.apsoil.2021.104362

19. Galperin, M. Y. Genome diversity of spore-forming Firmicutes. In *The bacterial spore: From molecules to systems* (eds Driks, A. & Eichenberger, P.) 1–18 (ASM Press, 2016). https://doi.org/10.1128/9781555819323.ch1

20. Mandic-Mulec, I., Stefanic, P. & van Elsas, J. D. Ecology of Bacillaceae. *Microbiol. Spectr.* **3**, TBS-0017-2013 (2015). https://doi.org/10.1128/9781555819323.ch3

21. Kielak, A. M., Barreto, C. C., Kowalchuk, G. A., van Veen, J. A. & Kuramae, E. E. The ecology of Acidobacteria: Moving beyond genes and genomes. *Front. Microbiol.* **7**, 744 (2016). https://doi.org/10.3389/fmicb.2016.00744

22. Yamada, T., Sekiguchi, Y., Hanada, S., Imachi, H., Ohashi, A., Harada, H. & Kamagata, Y. Anaerolinea thermolimosa sp. nov., *Levilinea saccharolytica* gen. nov., sp. nov. and *Leptolinea tardivitalis* gen. nov., sp. nov., and description of the new classes Anaerolineae and Caldilineae in the bacterial phylum Chloroflexi. Int. J. *Syst. Evol. Microbiol.* **56**, 1331–1340 (2006). https://doi.org/10.1099/ijs.0.64169-0

23. Fierer, N. Embracing the unknown: Disentangling the complexities of the soil microbiome. *Nat. Rev. Microbiol.* **15**, 579–590 (2017). https://doi.org/10.1038/nrmicro.2017.87

24. Brewer, T. E., Handley, K. M., Carini, P., Gilbert, J. A. & Fierer, N. Genome reduction in an abundant and ubiquitous soil bacterium '*Candidatus Udaeobacter copiosus*'. *Nat. Microbiol.* **2**, 1–7 (2017). https://doi.org/10.1038/nmicrobiol.2016.198

25. Willms, I. M., Rudolph, A. Y., Göschel, I., Bolz, S. H., Schneider, D., Penone, C. et al. Globally abundant "*Candidatus Udaeobacter*" benefits from release of antibiotics in soil and potentially performs trace gas scavenging. *mSphere* **5**, e00128-20 (2020). https://doi.org/10.1128/msphere.00186-20

26. Barberán, A., Bates, S. T., Casamayor, E. O. & Fierer, N. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J.* **6**, 343–351 (2012). https://doi.org/10.1038/ismej.2011.119

27. Pandit, S. N., Kolasa, J. & Cottenie, K. Contrasts between habitat generalists and specialists: An empirical extension to the basic metacommunity framework. *Ecology* **90**, 2253–2262 (2009). https://doi.org/10.1890/08-0851.1

28. Zhang, Z., Wu, X., Zhang, J., Liu, Y., Luo, W. & Mou, G. Community assembly of bacterial generalists and specialists and their network characteristics in different altitudinal soils on Fanjing Mountain in Southwest China. *Catena* **238**, 107863 (2024). https://doi.org/10.1016/j.catena.2024.107863

29. Hiera, M. D., Lima Júnior, A. F. & Zanella, M. E. Tendência da precipitação no estado do Ceará. *Rev. Bras. Climatol.* **24**, 1–22 (2019). https://doi.org/10.5380/abclima.v24i0.54207

30. IPECE. Perfil municipal (Instituto de Pesquisa e Estratégia Econômica do Ceará, 2017).

31. Muniz, L. F. et al. Classificação climática para o estado do Ceará (2017).

32. IUSS Working Group WRB. World Reference Base for Soil Resources: International soil classification system for naming soils and creating legends for soil maps, 4th edn (International Union of Soil Sciences, Vienna, 2022).

33. Santos, H. G. et al. *Sistema Brasileiro de Classificação de Solos*, 5th edn (Embrapa, 2018).

34. Amaral, R. V. do. O manejo extrativista influencia a dinâmica do carbono e nutrientes em carnaubais nativos? Dissertação de mestrado, Universidade Federal do Ceará (2023).

35. Donagema, G. K., Campos, D. V. B., Calderano, S. B., Teixeira, W. G. & Viana, J. H. M. (eds) *Manual de métodos de análise de solo*, 2nd edn rev. (Embrapa Solos, Rio de Janeiro, 2011).

36. Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., Fierer, N. & Knight, R. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. USA* **108**, 4516–4522 (2011). https://doi.org/10.1073/pnas.1000080107

37. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. *Biotechnol.* **37**, 852–857 (2019). https://doi.org/10.1038/s41587-019-0209-9

38. Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A. & Holmes, S. P. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016). https://doi.org/10.1038/nmeth.3869

39. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P. et al. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013). https://doi.org/10.1093/nar/gks1219

40. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013). https://doi.org/10.1093/molbev/mst010

41. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490 (2010). https://doi.org/10.1371/journal.pone.0009490

42. R Core Team. R: A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria, 2024). https://www.R-project.org/

43. McMurdie, P. J. & Holmes, S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**, e61217 (2013). https://doi.org/10.1371/journal.pone.0061217

44. Oksanen, J. et al. vegan: Community Ecology Package. R package version 2.6-4 (2022). https://CRAN.R-project.org/package=vegan

45. Chong, J., Liu, P., Zhou, G. & Xia, J. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. *Nat. Protoc.* **15**, 799–821 (2020). https://doi.org/10.1038/s41596-019-0264-1

46. Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S. & Huttenhower, C. Metagenomic biomarker discovery and explanation. *Genome Biol.* **12**, R60 (2011). http://genomebiology.com/2011/11/6/R60

47. Liu, C., Cui, Y., Li, X. & Yao, M. microeco: An R package for data mining in microbial community ecology. *FEMS Microbiol. Ecol.* **97**, fiaa255 (2021). https://doi.org/10.1093/femsec/fiaa255

48. Friedman, J. & Alm, E. J. Inferring correlation networks from genomic survey data. *PLoS Comput. Biol.* **8**, e1002687 (2012). https://doi.org/10.1371/journal.pcbi.1002687

49. Bastian, M., Heymann, S. & Jacomy, M. Gephi: An open source software for exploring and manipulating networks. In *Proceedings of the International AAAI Conference on Web and Social Media* **3**, 361–362 (2009). https://doi.org/10.1609/icwsm.v3i1.13937

50. Chazdon, R. L. et al. A novel statistical method for classifying habitat generalists and specialists. *Ecology* **92**, 1332–1343 (2011). https://doi.org/10.1890/10-1345.1

Acknowledgements

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil (CAPES), Finance Code 001. Additional support was provided by the Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP) through the Universal Call 06/2023 (Process UNI-0210-00092.01.00/23) for the project "The microbiome associated with carnauba plants (*Copernicia prunifera*): ecology and biotechnological applications". The authors acknowledge Rothamsted Research for providing the infrastructure and support necessary for the data analysis.

Author Contributions

Elane Bezerra da Silva: Data curation, Investigation, Writing – original. **Francisco Matheus Medeiros de Freitas:** Data curation, Investigation, Writing – original draft. **Murilo de Sousa Almeida:** Data curation. **Kaio Gráculo Vieira Garcia:** Writing – review and editing. **Larissa dos Santos Lopes:** Data curation, Investigation. **Vanessa Nessner Kavamura Noguchi:** Investigation, Writing – original draft, Writing – review & editing. **Vânia Maria Maciel Melo:** Writing – review and editing**. Erika Valente de Medeiros:** Writing – review and editing**. Ademir Sérgio Ferreira Araujo:** Writing – review and editing**. Arthur Prudêncio de Araujo Pereira:** Conceptualization, Supervision, Writing – review & editing.

Declarations

Competing interests

The authors declare no competing interests

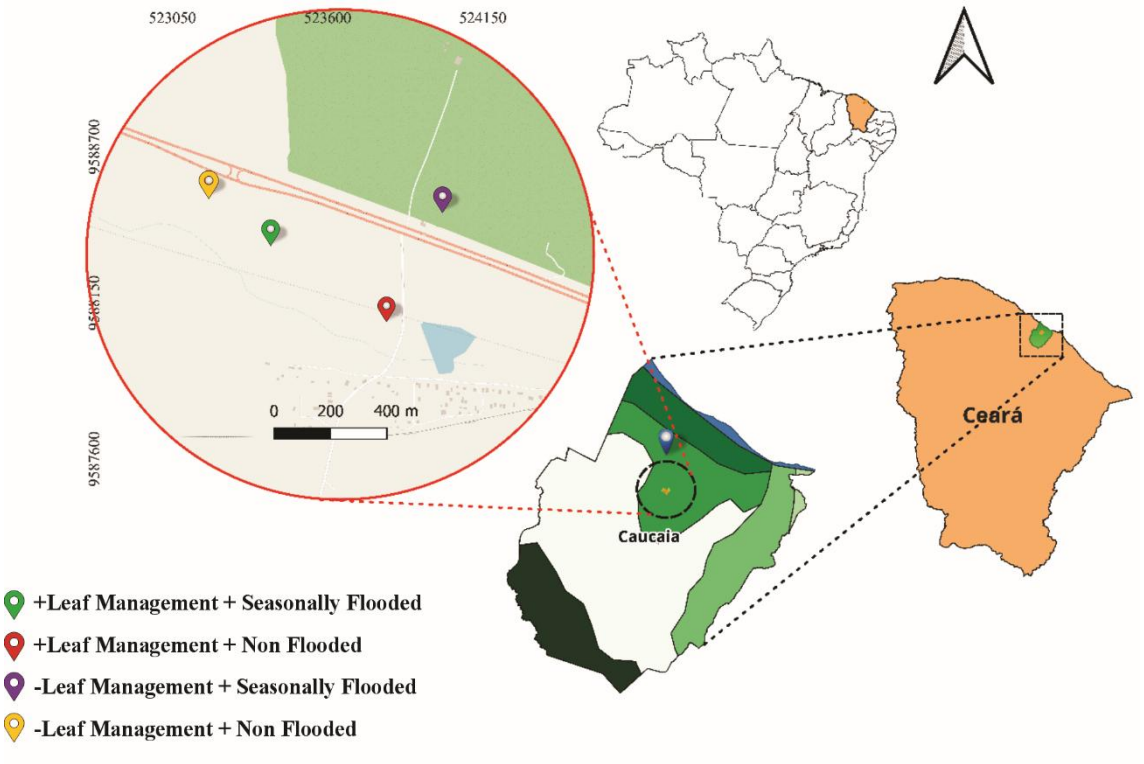


Figure 1. Location of the study area in a native *Copernicia prunifera* (carnauba) farm in the municipality of Caucaia, Ceará State, northeastern Brazil. The map shows the geographic position of the study area within Brazil and Ceará State, with a zoomed inset illustrating the spatial distribution of the evaluated palms. Points represent four conditions: +Leaf management + Seasonally flooded (green), +Leaf management + Non-flooded (red), −Leaf management + Seasonally flooded (purple), and −Leaf management + Non-flooded (yellow).

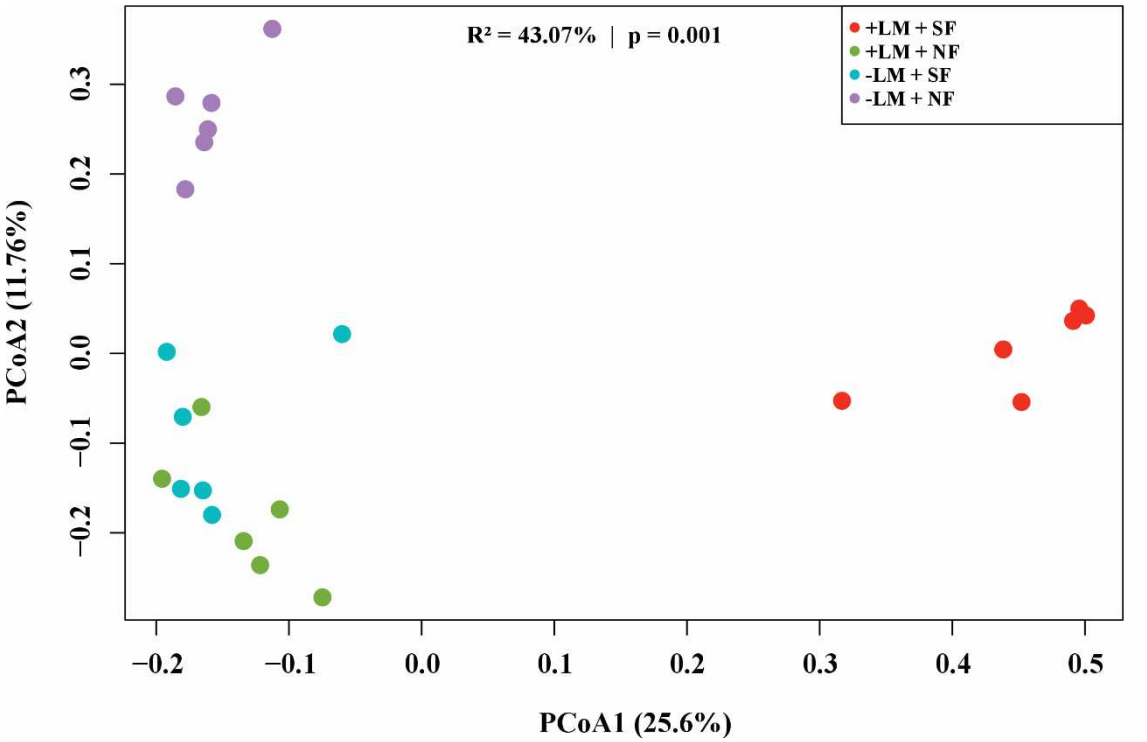


Figure 2. Principal-coordinate analysis of ASV-level bacterial community dissimilarity among four carnauba-grove conditions. PCoA1 and PCoA2 explain 25.6% and 11.76% of the variation, respectively. Differences among conditions were tested by PERMANOVA (R^2 = 0.4307, p = 0.001; n = 6 tree-centered samples per condition). +LM+SF, managed and seasonally flooded; +LM+NF, managed and non-flooded; −LM+SF, unmanaged and seasonally flooded; −LM+NF, unmanaged and non-flooded.

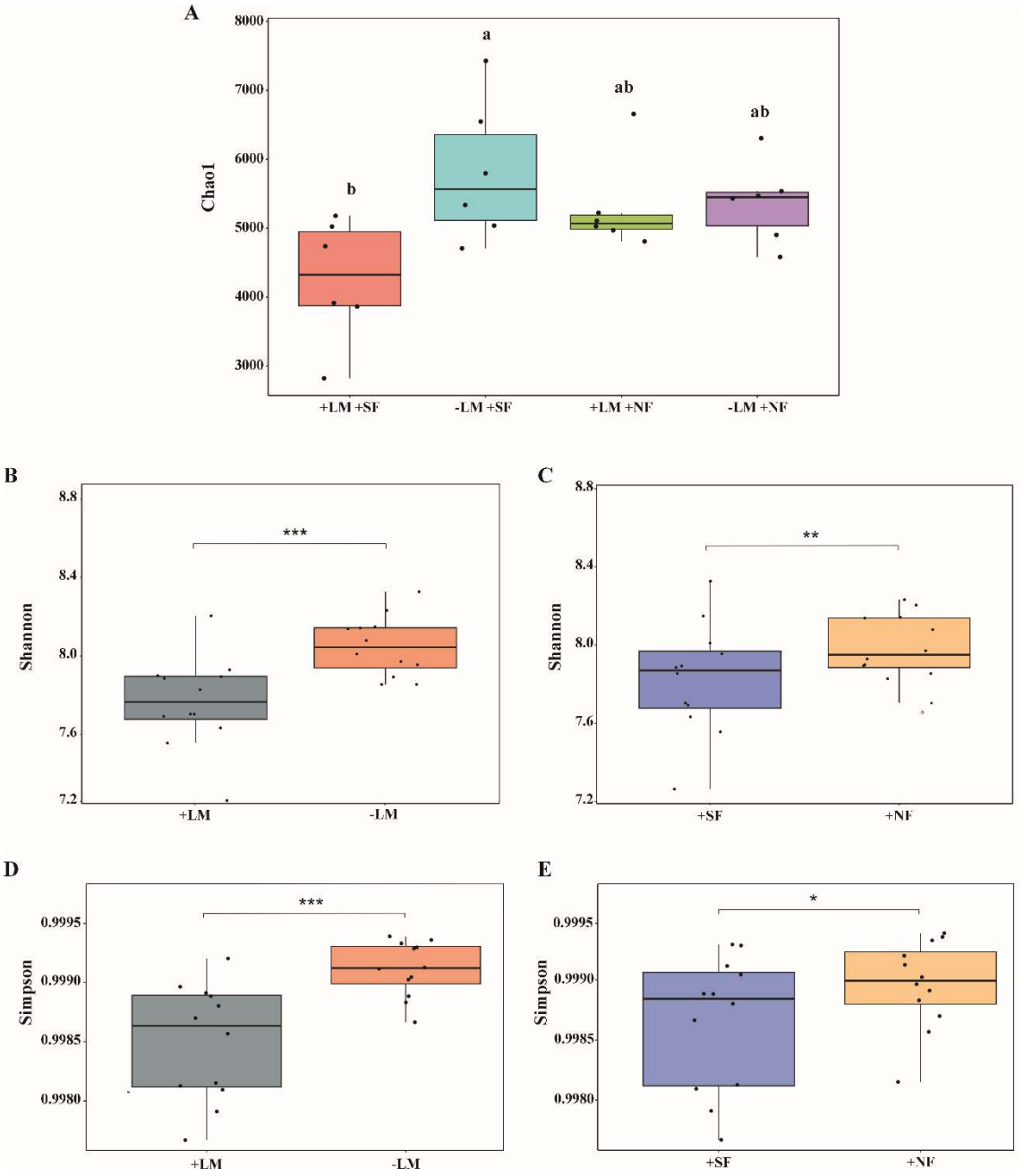


Figure 3. Alpha diversity of soil bacterial communities in four carnauba-grove conditions (n = 6 tree-centered samples per condition). (A) Chao1 richness; different lowercase letters indicate significant Tukey contrasts among conditions ($p < 0.05$). (B, C) Shannon diversity by leaf-harvesting status and water regime. (D, E) Simpson diversity by leaf-harvesting status and water regime. Asterisks denote $p < 0.05$ (*), $p < 0.01$ (**), and $p < 0.001$ (***). Abbreviations are defined in Figure 2.

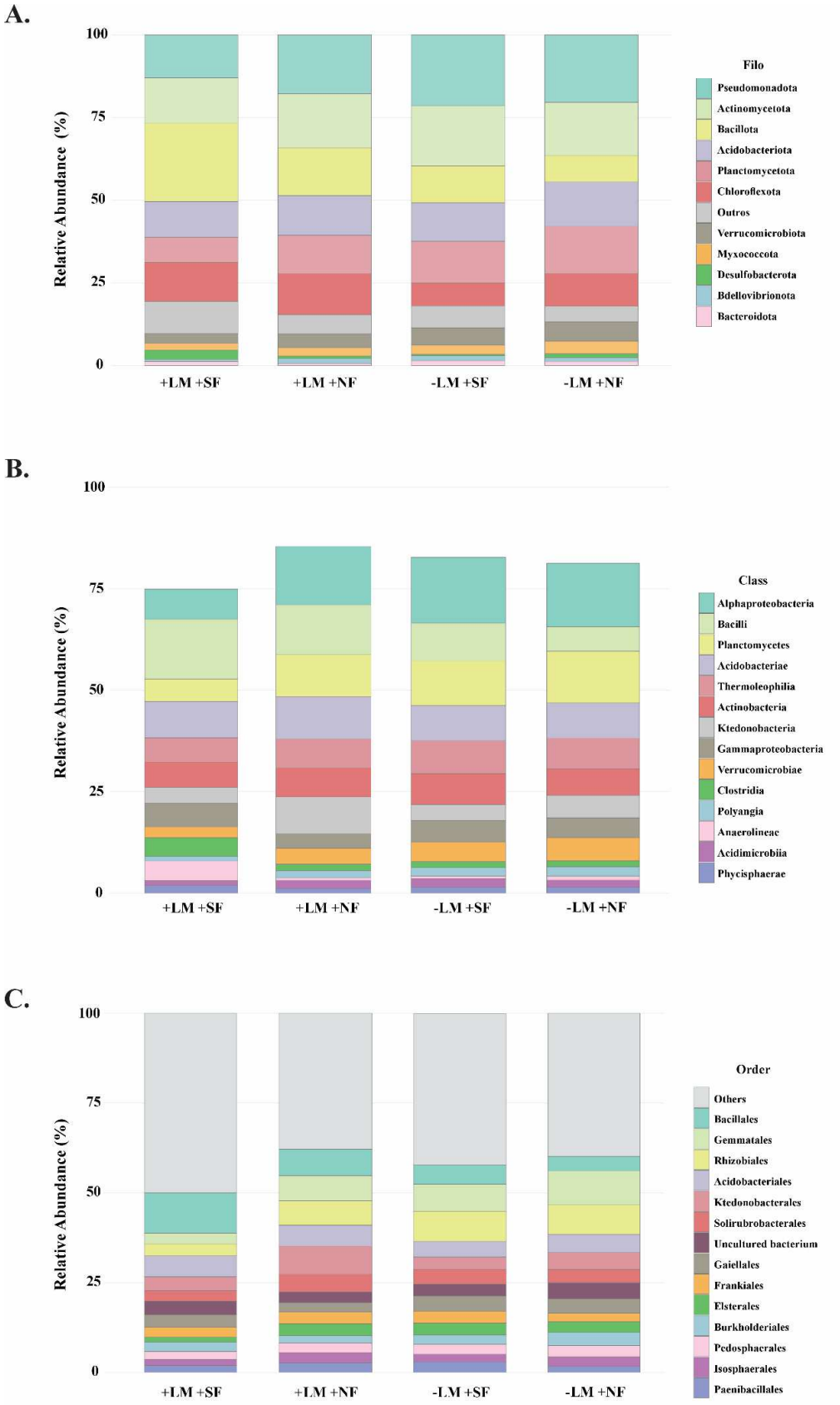

A.
Relative Abundance (%)
100
75
50
25
0
+LM +SF
+LM +NF
-LM +SF
-LM +NF
Filo
Pseudomonadota
Actinomycetota
Bacillota
Acidobacteriota
Planctomycetota
Chloroflexota
Outros
Verrucomicrobiota
Myxococcota
Desulfobacterota
Bdellovibrionota
Bacteroidota
B.
Relative Abundance (%)
100
75
50
25
0
+LM +SF
+LM +NF
-LM +SF
-LM +NF
Class
Alphaproteobacteria
Bacilli
Planctomycetes
Acidobacteriae
Thermoleophilia
Actinobacteria
Ktedonobacteria
Gammaproteobacteria
Verrucomicrobiae
Clostridia
Polyangia
Anaerolineae
Acidimicrobiia
Phycisphaerae
C.
Relative Abundance (%)
100
75
50
25
0
+LM +SF
+LM +NF
-LM +SF
-LM +NF
Order
Others
Bacillales
Gemmatales
Rhizobiales
Acidobacteriales
Ktedonobacterales
Solirubrobacterales
Uncultured bacterium
Gaiellales
Frankiales
Elsterales
Burkholderiales
Pedosphaerales
Isosphaerales
Paenibacillales

Figure 4. Relative abundance of dominant bacterial groups at (A) phylum, (B) class, and (C) order levels in four carnauba-grove conditions. Bars summarize the proportion of assigned sequences within each condition. Abbreviations are defined in Figure 2.

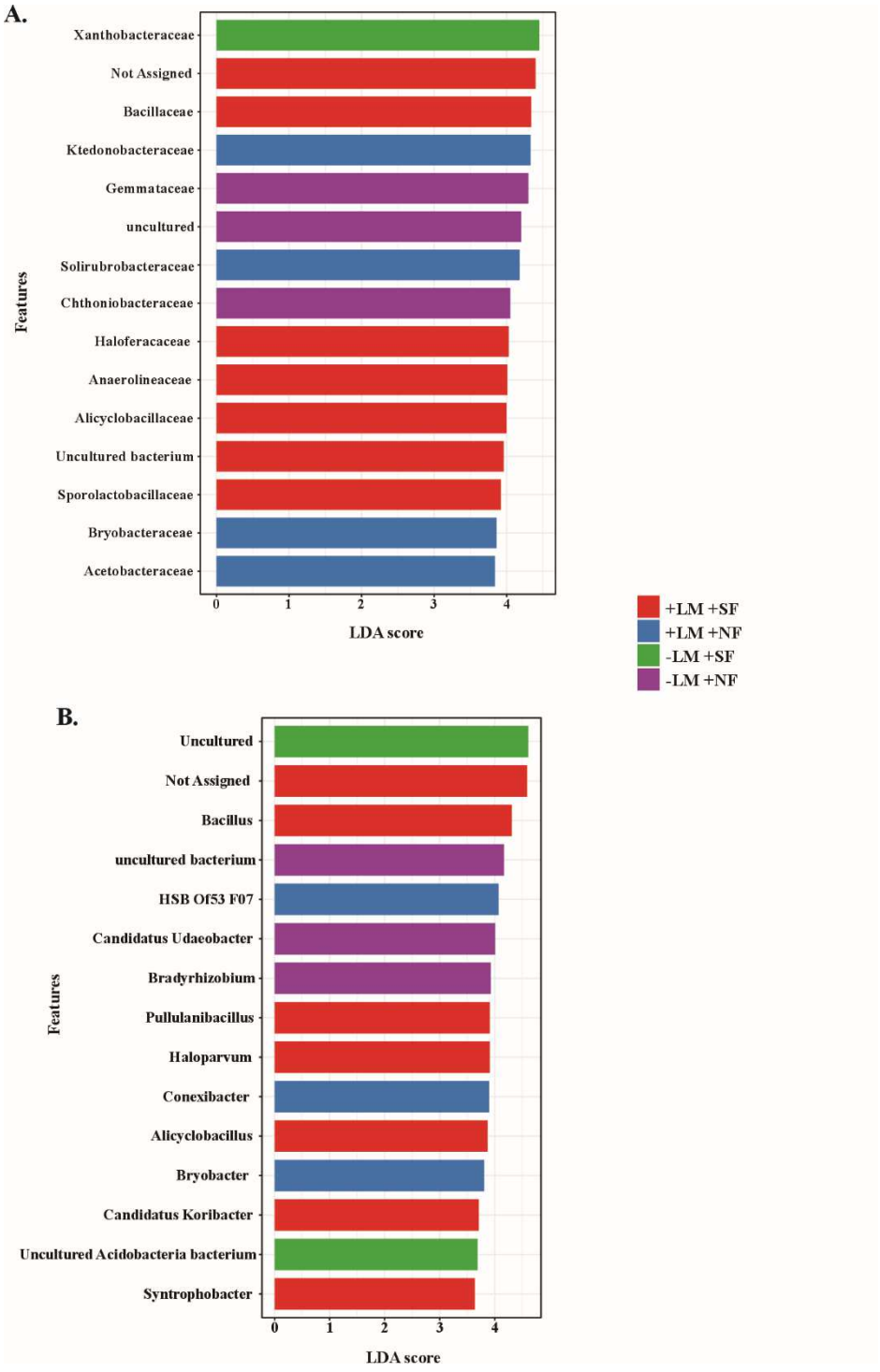


Figure 5. LEfSe-identified taxa discriminating among the four carnauba-grove conditions at (A) family and (B) genus levels. Bars show log10 LDA scores. Only taxa with a nominal Kruskal–Wallis $p < 0.05$ and LDA score > 2.0 are displayed. Colors indicate the condition with the highest relative abundance. Abbreviations are defined in Figure 2.

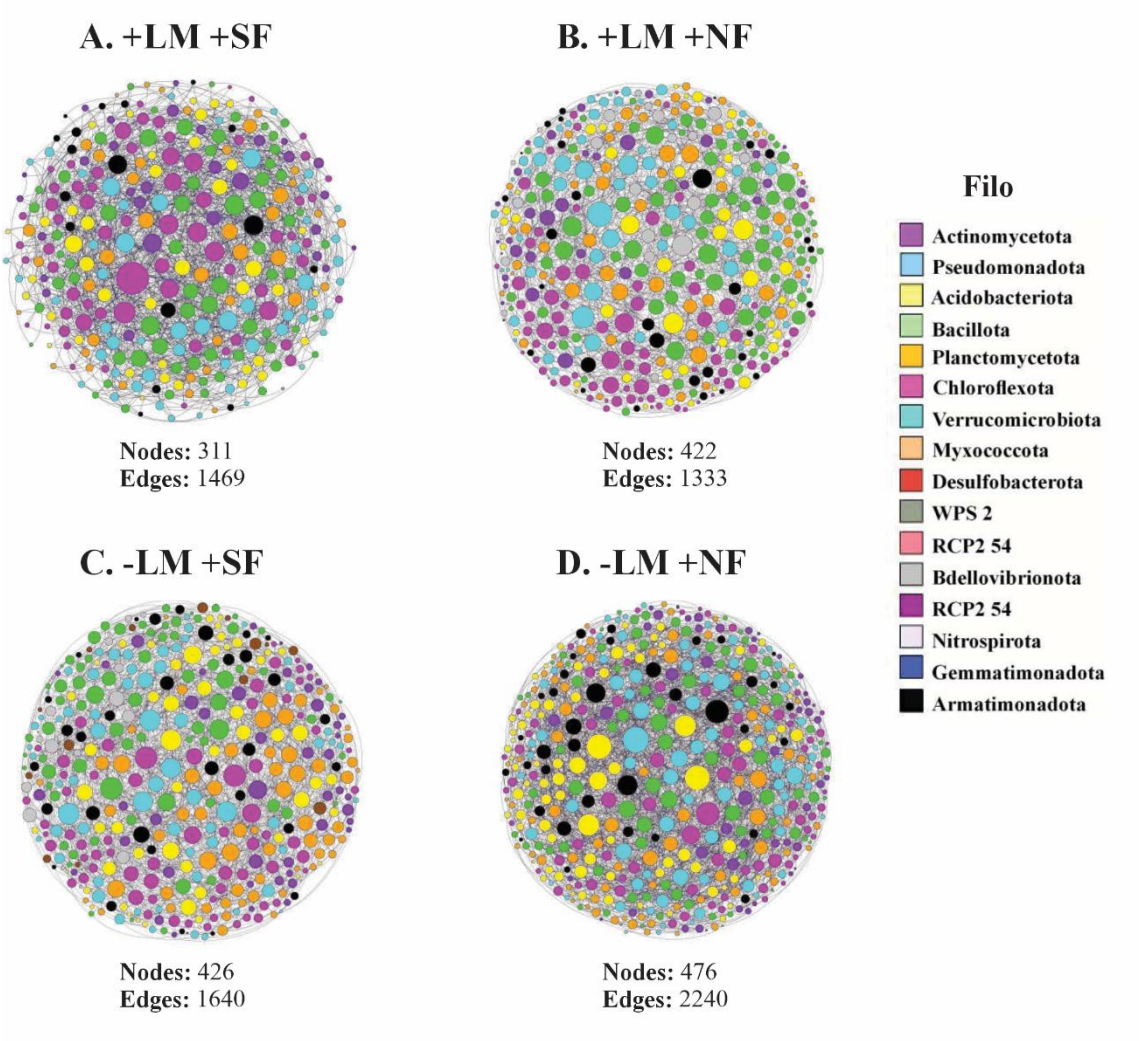


Figure 6. Exploratory bacterial co-occurrence networks for (A) +LM+SF, (B) +LM+NF, (C) −LM+SF, and (D) −LM+NF. Nodes represent ASVs, node color denotes phylum, node size is proportional to degree, and edges represent retained SparCC associations ($|r| \geq 0.9$; nominal $p < 0.05$). Positive and negative associations are shown in dark blue and red, respectively. Networks are descriptive because each condition contained six samples and p values were not adjusted for multiple testing. Abbreviations are defined in Figure 2.

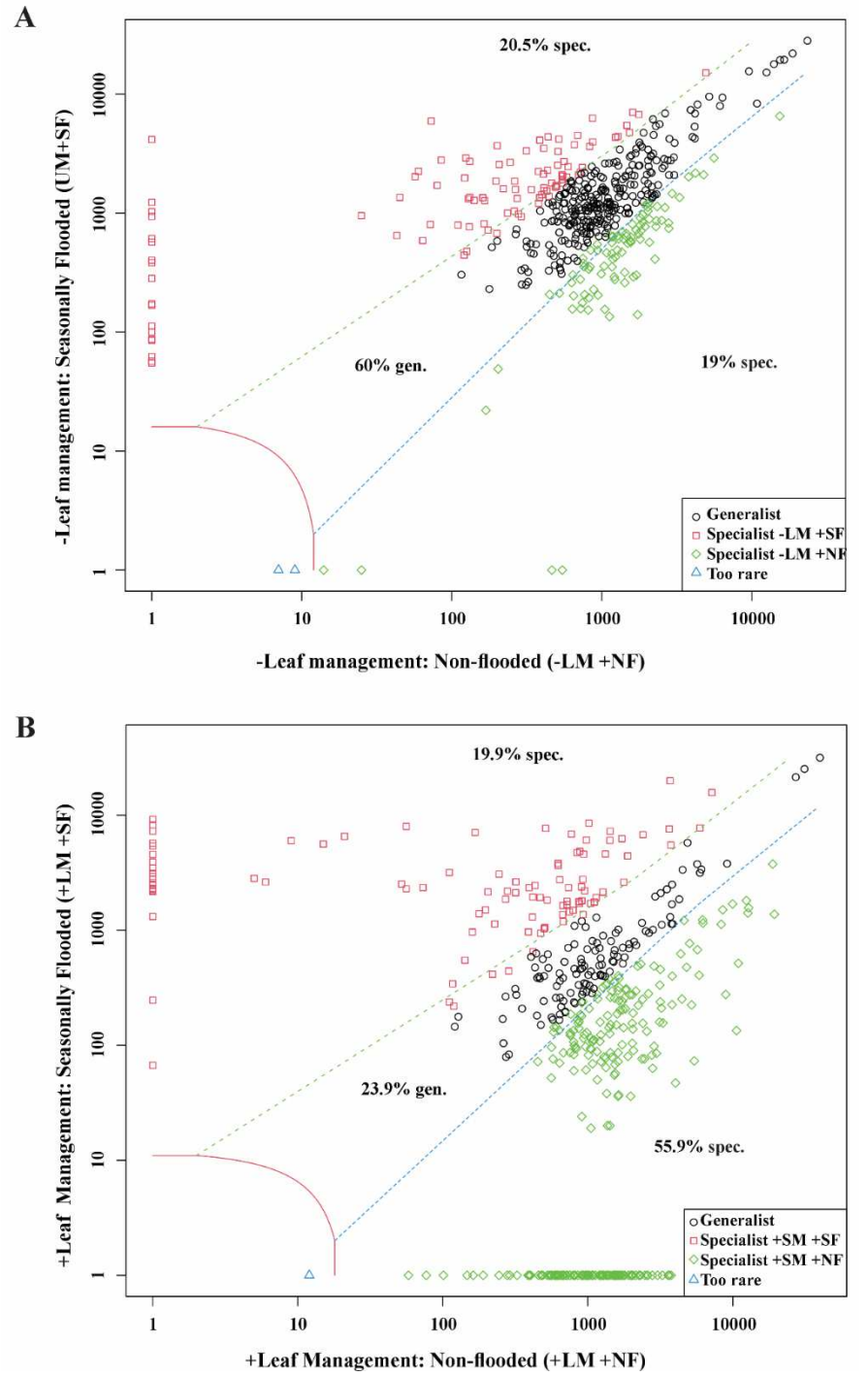


Figure 7. CLAM niche-occupancy classification of ASVs in (A) unmanaged and (B) managed areas, comparing seasonally flooded and non-flooded conditions. Each point represents an ASV positioned by its abundance in the two conditions. Categories were assigned with the supermajority rule ($K = 2/3$, $p < 0.005$) as generalists, condition specialists, or too rare to classify. Abbreviations are defined in Figure 2.