Aquat Ecol
https://doi.org/10.1007/s10452-025-10235-5

Beyond environmental selection: exploring the spatial structuring of a freshwater bacterial metacommunity on a broad geographical scale

Erick Mateus-Barros · Adriano Caliman · Thaís Garcia da Silva · Hugo Sarmento

Received: 27 June 2025 / Accepted: 16 September 2025

Abstract Understanding how spatial factors shape species distribution and geographic range remains a central question in community ecology, particularly for microorganisms with high dispersal capacities. Here, we applied variation partitioning to assess the impact of multiple spatial mechanisms on the β-diversity of a bacterioplankton metacommunity in 60 shallow lakes distributed across a broad landscape. Linear overland distances between sites provided the best explanation for bacterioplankton β-diversity, likely reflecting the dominance of ubiquitous, highly abundant bacteria. In contrast, rarer bacteria were associated with local neighborhoods, suggesting lower dispersal capacities. This pattern aligns with observations in other taxa and indicates a potential rescue effect. Moreover, regardless of abundance, most bacteria appear to occupy a geographic range of 600–700 km. Our results suggest that while deterministic factors play a critical role in shaping freshwater bacterial biodiversity, the spatial context also contributes to explaining bacterioplankton dissimilarity and community dynamics. Importantly, selecting the appropriate spatial metric—whether based on presence/absence or abundance data—is crucial for accurately capturing these patterns.

Keywords Variation partitioning · β-diversity partition · Spatial dynamics · 16S amplicon sequencing

Introduction

Freshwater bacteria are part of the key organisms for ecosystem functioning, playing essential roles in biogeochemical cycles and maintaining close relationships with other living organisms, including those that influence human health (Newton et al. 2011; Chiriac et al. 2022). Consequently, the relationship between microbial community structure and environmental

---

Handling Editor: Richard LaBrie.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10452-025-10235-5.

E. Mateus-Barros (✉) · H. Sarmento
Departamento de Hidrobiologia, Laboratory of Microbial Processes and Biodiversity, Universidade Federal de São Carlos (UFSCar), São Carlos, SP 13565-905, Brazil
e-mail: erickmbarros@gmail.com

E. Mateus-Barros
Grupo de Ecología Microbiana de Sistemas Acuáticos, Centro Universitario Regional del Este, Universidad de la República, Rocha, Uruguay

A. Caliman
Departamento de Ecologia, Universidade Federal do Rio Grande do Norte, Natal, RN 59078-900, Brazil

T. G. da Silva
Departamento de Ciências Biológicas, Universidade Estadual de Santa Cruz (Uesc), Ilhéus, BA 45662-900, Brazil

variation at local scales has been extensively studied over the decades. These studies consistently highlight factors such as pH, salinity, and temperature as primary drivers of microbial biodiversity across diverse contexts (Lindström et al. 2005; Zhang et al. 2021). However, the longstanding idea that their high dispersal capacity means spatial factors do not influence their biodiversity, and that this biodiversity is solely determined by local environmental conditions—i.e. everything is everywhere, but the environment selects (Baas Becking 1934)—has been challenged by recent advances in DNA sequencing and bioinformatics.

These advances have revealed that microbial biodiversity is significantly higher than previously estimated (Rappé and Giovannoni 2003; Callahan et al. 2016). This finding stems from the discovery of very low-abundance and narrow-range taxa within microbial communities, often referred to as the "rare biosphere" (Pedrós-Alió 2012). The identification of this rare biosphere has significant implications on how we interpret spatial features of biodiversity, as these organisms may respond differently to geographic distances (Mateus-Barros et al. 2021) or environmental heterogeneity (Huber et al. 2020) compared to more abundant and widely distributed microbes. Nevertheless, the relationship between microbial dynamics and spatial features remains unclear. For instance, recent studies analysing bacterial dissimilarity from distinct perspectives have reported conflicting results, likely due to the use of different methodologies. Some studies comparing the influence of organism size on spatial features of biodiversity have demonstrated a significantly higher impact of stochastic processes on microbial communities (Farjalla et al. 2012; Soininen et al. 2013). In contrast, other studies that focused solely on microorganisms have emphasized the greater relative importance of environmental filtering (Jyrkänkallio-Mikkola et al. 2017; Fillinger et al. 2019). Therefore, understanding how these organisms organize in space, and identifying mechanisms beyond environmental filtering that influence their distribution, could reveal critical insights into microbial ecology.

In this sense, study of the relative role of local and regional processes in shaping spatial community properties has received growing attention over the past few decades (e.g. Ricklefs and Jenkins 2011). Within this context, the concept of a metacommunity has emerged as a powerful ecological framework. It posits that a group of local communities are interconnected through the dispersal of multiple species, which may potentially interact with each other (Wilson 1992; Leibold et al. 2004). A well-established and straightforward way to identify the processes influencing metacommunity structure is by measuring the observed dissimilarity in community composition between pairs of local communities (Whittaker 1960, 1972). This dissimilarity is a diversity dimension known as β-diversity.

Multiple approaches have been proposed to assess and interpret patterns of β-diversity. These can be broadly categorized into qualitative and quantitative methods (Anderson et al. 2011). Qualitative approaches rely on species presence/absence data, which are more commonly available and can also be derived from other types of data, thereby facilitating access to larger datasets and insights into compositional variation. In contrast, quantitative approaches incorporate (relative) abundance, which are crucial for understanding the mechanisms that shape metacommunity structure (Anderson et al. 2011). β-diversity can be influenced by a range of intrinsic and extrinsic factors, including historical contingency, trophic level, variability in local factors and interactions, as well as the type of environment inhabited (Baselga et al. 2011). Geography also plays a critical role, both directly and indirectly, by altering environmental heterogeneity and affecting the dispersal opportunities of each species (Martiny et al. 2011; Soininen et al. 2018). This process varies across spatial scales (Vellend 2010; Baselga et al. 2011) and is not necessarily dependent on specific dispersal abilities (Gaston 2009).

In freshwater microbial communities, large-scale dispersal appears to be related to both specific lifestyles and ecological context (Huber et al. 2020; Ruiz-González et al. 2015). While some of these organisms are commonly attached to the particles they degrade to feed on, others are found moving in the water column, feeding on smaller, less complex particles available there or synthesizing nutrients from sunlight (Chiriac et al. 2022). Thus, because they are free in the water column, organisms of this second group can be more easily transported to other sites. Moreover, the composition of microbial species can differ significantly depending on the region where a community is found. For example, headwater communities are compositionally more similar to the surrounding soil communities than to those at the mouth of the same river (Ruiz-González et al. 2015). In addition, even when comparing communities in more similar contexts, such as two sites in a floodplain, the degree of connectivity between study sites directly affects the observed β-diversity (Huber et al. 2020). Finally, long-distance dispersal does not seem to be governed by a single event, but rather by small-scale dispersal events over short distances, with a linear increase in the area reached by a species (Hoetzinger et al. 2024), in a pattern likely guided by the local abundance reached by each bacterium (Mateus-Barros et al. 2021). Furthermore, research has shown that species sorting is more pronounced among the most widespread bacteria (Székely and Langenheder 2014), and that these taxa tend to have higher speciation rates (Sriswasdi et al. 2017), which can be captured at the amplicon sequence variant (ASV) level. Notably, the geographic range of freshwater bacteria remains largely unknown due to a lack of comparable studies conducted on a sufficiently large scale (but see Hoetzinger et al. 2024).

Bacteria are passive dispersers, which means that their movement across aquatic environments is driven not only by water flows (Lansac-Tôha et al. 2020, Sadeghi et al. 2024), but also by attachment to larger organisms (Grossart et al. 2010) and wind (Smith et al. 2013). In this context, some theoretical approaches can be used to determine the main way dispersal impacts bacterial β-diversity. If only 'the environment selects' as stated for many decades, the geography will only reflect the spatial structure of environmental factors (Borcard et al. 1992). On the contrary, considering the possibility of a sufficient homogeneous distribution of deterministic factors, the β-diversity will decrease as a function of increasing geographical distance (Dray et al. 2006). Other possibilities are a stronger connection by shorter distances reflecting the low dispersive capacity of the large number of rare bacteria in this data (Peres-Neto and Legendre 2010), the isolation generated by distinct drainage basins that make up this landscape which should be reflected by an increased similarity between sites located in the same watershed, and a connection between sites proportioned by river flow (Blanchet et al. 2008a, b). In light of these dispersal mechanisms, distinct theoretical approaches can help us understand how dispersal influences bacterial β-diversity.

In this study, we aimed to investigate the influence of environmental and spatial mechanisms on bacterial occupancy within a tropical bacterioplankton metacommunity. To reach this objective, we recovered the most important environmental factors and used different dispersal models to address the relative importance of these aspects to the structure of a bacterioplanktonic metacommunity across a set of 60 tropical shallow lakes scattered over a matrix of nearly 250,000 km^2. Beyond environmental factors, we sought to determine the extent of bacterial ASV occupancy by integrating spatial components from the metacommunity theoretical framework. Specifically, we aimed to elucidate the impact of dispersal mechanisms, including river flow (Fig. 1A), overland distances (Fig. 1B), neighborhood effects (Fig. 1C), and geographical barriers (Fig. 1D), on bacterial β-diversity and the size of geographic range. By adopting a comprehensive set of qualitative and quantitative methods, we sought to identify the relative importance among geographic mechanisms affecting bacterioplankton metacommunity patterns. In doing so, we aspire to address the fundamental question of how far bacteria can establish, thereby advancing our understanding of spatially related local and regional processes in microbial metacommunities.

Methods

Study design

This study was performed using a dataset obtained from 60 headwater shallow lakes covering a region of nearly 250,000 km^2 in São Paulo state, southeast Brazil (Fig. 1). This region has a tropical climate and is characterized by Cerrado (Brazilian savannah) and Atlantic Forest (semi-deciduous humid forest). Located in a region of intense agricultural activity, these lakes are generally small reservoirs, dammed to fulfill water needs of landowners. The study area lies within three sub-basins of the Paraná River basin (i.e. Grande, Paranapanema, and Tietê rivers) and the coastal formation, which rises rapidly from sea level to over 1000 m before descending toward the continent's interior. In line with other findings, the tropical climate prevents microbial community dynamics from being mainly driven by temperature dynamics (Fig. S1A, C). Also, possibly due to the intense

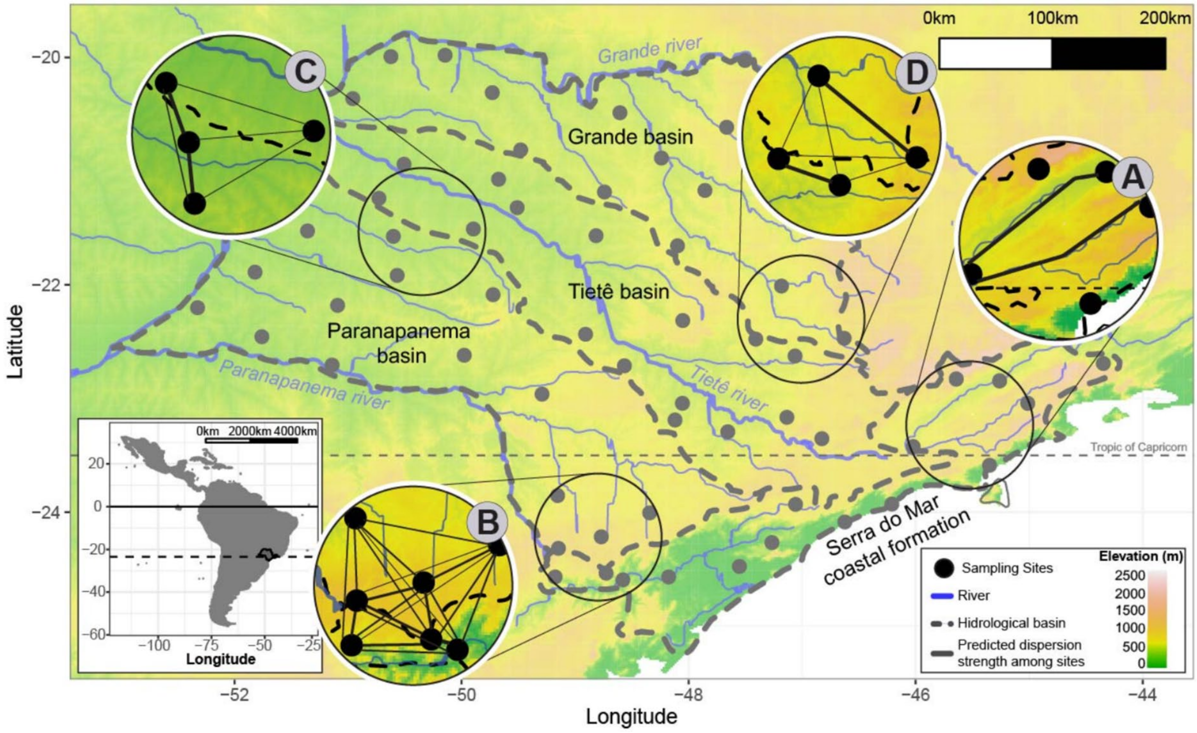

Fig. 1 Location of the sixty headwater shallow lakes that were sampled over a large tropical landscape covering four distinct hydrological basins. The main rivers (blue lines) and basin delimitations (black dashed lines) were also indicated. The large zoomed figures illustrate how each spatial aspect was considered for the variation partitioning: **A** distance between local communities through river flow, the connectivity fraction; **B** spatial overland distance between local communities, the spatial fraction; **C** isolation between local communities caused by geographical barriers, the region fraction; and **D** increased connectivity to neighboring sites, the neighborhood fraction. The line thickness in the examples represents hypothetical connectivity strength (flow of individuals) between the pairs of sites under the four dispersal possibilities described above

human activities, precipitation variation also appears not to affect these community dynamics (Fig. S1B, D), instead, they are more likely to respond to variation in pH and organic compounds (Mateus-Barros et al. 2021). Sampling was carried out between June 2012 and July 2016. To minimize bias, samples were selected at equidistant intervals (when possible), to create a homogeneous grid across the area (Fig. 1) that encompassed all geographical contexts assessable in this study.

In the field, we measured environmental variables (temperature, conductivity, pH) using a multiparameter probe (YSI, Yellow Springs, USA) and filtered sub-surface water for laboratory analyses (nutrients, carbon supply, chlorophyll-*a*, environmental DNA). Altitude and geographic coordinates were obtained using a GPS. Samples for nutrient analysis were obtained by filtering lake water through 0.45 µm polycarbonate membranes previously washed with ultrapure water to prevent carbon contamination from the filter, and stored in amber bottles in a freezer at −20 °C until analysis. Dissolved organic carbon (DOC) was measured using a TOC-V (Shimadzu®, Kyoto, Japan). Dissolved nutrient concentrations were determined using a Ion Chromatography System (Thermo Scientific®, Waltham, Massachusetts, USA), and dissolved inorganic nitrogen (DIN) was calculated by adding the values obtained from nitrite, nitrate, and ammonium. Additionally, a FS5 Spectrofluorometer (Edinburgh Instruments®, Livingston, UK) was used to estimate the Tryptophan-like fluorescent dissolved organic matter (T-FDOM), calculated as the ratio between the fluorescence of dissolved organic matter fluorescence and that of quinine sulfate (0.001 mg/L dissolved in 0.1 M H_2SO_4) at 455 nm excitation and 355 nm emission. Chlorophyll-*a* concentration (used as a proxy for trophic state) was obtained by filtering 100–500 ml of water through a glass fibre filter (Macherey–Nagel® GF-6), extracting with ethanol (90% v/v at 80 °C) in the dark (Marker 1980; Mush 1980), and quantifying by spectrophotometry (Lorenzen 1967). This dataset includes equal sample sizes across trophic state categories (oligotrophic, mesotrophic, and eutrophic) and hydrologic basins. More details on the study site and environmental variable analyses can be found in Mateus-Barros et al. (2021).

Sampling, sequencing, and bioinformatics

To analyse microbial community composition and diversity, 500 ml of surface water were pre-filtered through a glass-fiber filter of 1.2 µm mesh (BOECO® MGC) to retain eukaryotes, large particles, and attached prokaryotes. Afterwards, 200–500 ml of filtrate were passed through 0.2 µm polycarbonate membranes (Millipore® Isopore™ 0.2 µm GTBP) to retain free-living prokaryotes.

Free-living bacterial DNA was extracted using a phenol–chloroform extraction protocol. The amplification was performed with the 341F (5′-CCTACGGGNGGCWGCAG-3′) and 805R (5′-GACTACHVGGGTATCTAATCC-3′) primers (Herlemann et al. 2011). Following amplification, fragments were sequenced using the Illumina MiSeq 2×250 paired-end reads. A full description of the molecular analyses can be found in Mateus-Barros et al. (2019). Raw sequences were processed using the DADA2 pipeline (Callahan et al. 2016) implemented in the R environment (R Core Team 2019) to generate a table of Amplicon Sequence Variants (ASVs). Taxonomic identification was assigned by blasting against the SILVA database version 132 (Yilmaz et al. 2014). The initial ASV table was filtered to remove sequences assigned to the *Archaea* domain, mitochondria, and chloroplasts and was normalized to equal sampling depth to create a subsampled ASV table (14,239 reads); finally, ASVs with a total abundance of fewer than 10 reads were removed. Rarefaction curves from these samples, after rarefaction and filtering of low-abundance ASVs, can be seen in Figure S2.

Data analyses

To determine the roles of species' presence and abundance in the ecological aspects investigated here, the first step was to transform the ASVs table into two separate tables: (1) a quantitative data table with relative abundance, which was used to perform analyses based on Bray–Curtis distances, and (2) a qualitative data table using ASV presence/absence at each site, used for analyses based on Jaccard distances. The comparison between qualitative and quantitative approaches should be interpreted with caution. This is because the sequences read by NGS equipment are limited by their maximum sequencing capacity and never capture the complete number of reads present in a sample. As a result, the data have a compositional nature (Gloor et al. 2017) and should always be treated as relative rather than absolute. In any case, the variation in ASV occurrence across sites is valuable for interpreting ecological processes. All forthcoming analyses (see below) were performed using both quantitative and qualitative data matrices.

To assess the role of local environmental and regional geographic features on the β-diversity observed in the bacterial metacommunity described above, we applied variation partitioning (Borcard et al. 1992) approach to dissimilarity matrices obtained from the *beta.pair* and *beta.pair.abund* functions of the *betapart* package (Baselga 2010; Baselga and Freckleton 2013). A distance-based redundancy analysis (dbRDA) was used to identify the factors employed in the variation partitioning (Legendre and Anderson 1999). Variation partitioning is a well-established approach used to determine the relative importance of different biotic and abiotic factors in shaping observed β-diversity in a given metacommunity.

To identify the dispersal features applied to variation partitioning, we used a set of eigenvector analyses. First, to test the potential connections between sites (Fig. 2A), we employed an Asymmetric Eigenvector Map (AEM) model. This model is calculated using a weight matrix that assigns "1" to river edges connecting two site pairs, while "0" is assigned to edges not connecting these sites (Blanchet et al. 2008a, b). The resulting eigenvector matrix was referred to as the connectivity fraction. Second, the linear decay of similarity was calculated using a distance-based Moran Eigenvector Map (dbMEM)

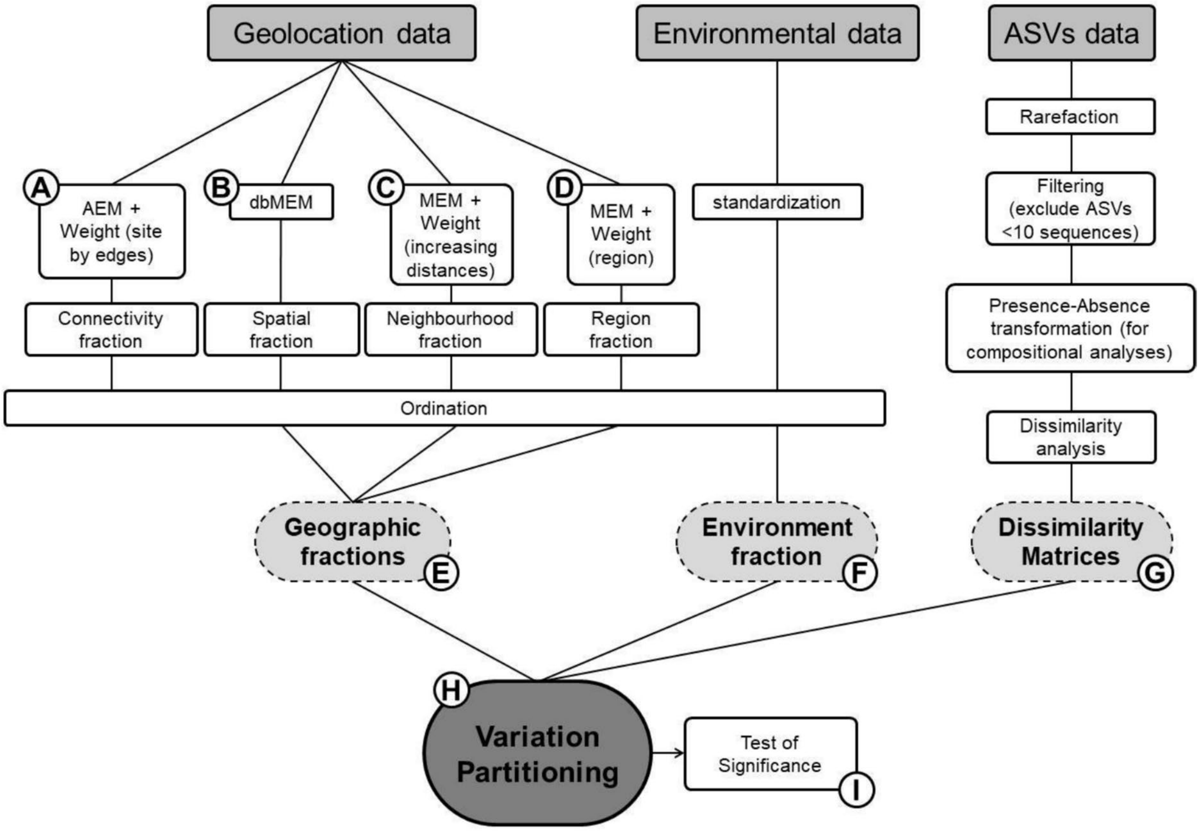

Fig. 2 Fluxogram showing all steps performed in the analysis of variation partitioning of β-diversity between spatial and environmental factors. The geolocation of each site was used to create distinct eigenvectors that model distinct geographic factors that may be impacting the metacommunity: river connectivity (**A**), overland spatial distances (**B**), neighborhood distances (**C**) and regional isolation (**D**), these matrices were filtered to the selection of relevant factors and quality check before composing the geographic fractions (**E**). The standardized environmental factors also passed through the relevance and quality checks before composing the environment fraction (**F**). By comparing these factors with qualitative and quantitative dissimilarity matrices (**G**), the variation partitioning (**H**) was performed based on a dbRDA. Finally, a CCA was used to test for significance (**I**)

approach to produce eigenvectors based on overland distances between sites (Dray et al. 2006), thereby recovering the space fraction (Fig. 2B). Third, the neighborhood fraction (Fig. 2C) was estimated using a MEM analysis (Peres-Neto and Legendre 2010) complemented by the weight formula $1-x/max(geo)$. Unlike the space fraction, this weight formula modifies the connection/distance relationship to a binomial-like shape, increasing the importance of shorter-distance connections. Finally, to assess the isolation created by drainage basins and obtain the region fraction (Fig. 2D), we used a MEM approach, determining the weight matrix by assigning “1” to site pairs located within the same basin and “0” to pairs in different regions. All eigenvectors were generated using the *adespatial* package (Dray et al. 2020).

Following these computations, all matrices underwent ordination with 999 permutations to select the most relevant eigenvectors. To mitigate issues related to spatial autocorrelation, we applied the R^2 ordination approach (Blanchet et al. 2008a, b; Bauman et al. 2018). This approach calculates an R^2 before adding each eigenvector to the ordination and terminates the analysis when the value reaches the R^2 for the global analysis, even if some significant eigenvector remain to be added (Fig. 3 shows all eigenvectors recovered). The ordination test was performed using the *ordiR2step* function of the *vegan* package (Oksanen

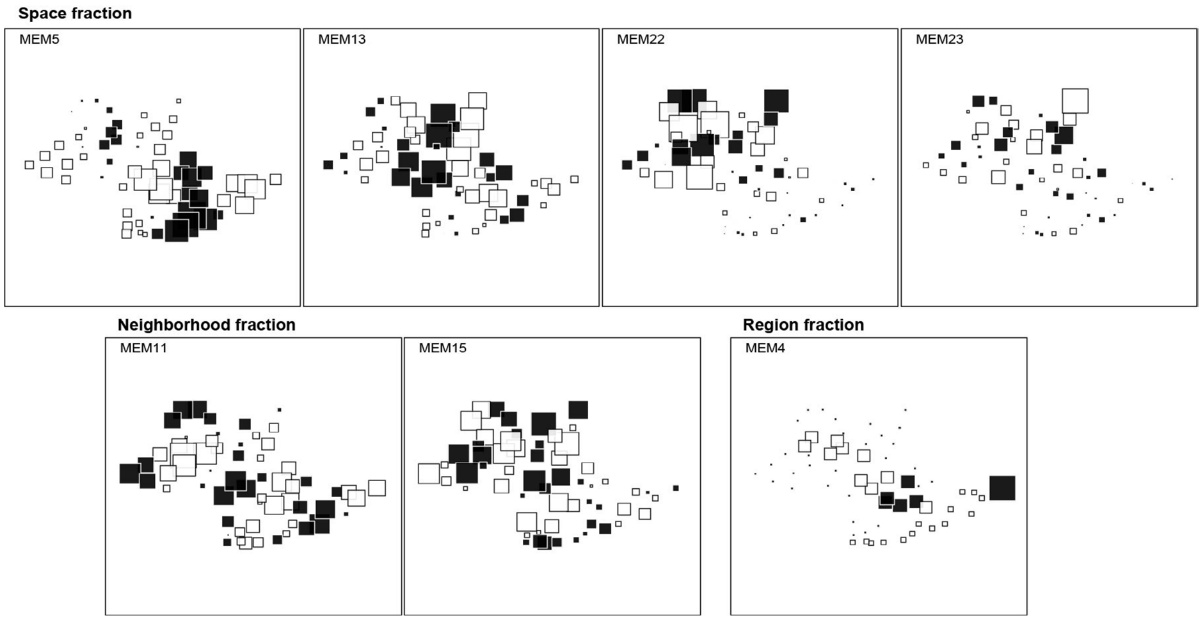

Fig. 3 All eigenvectors recovered after the ordination step. Black and white boxes indicate site groupings from each model. Only the structures that significantly match the observed microbial dissimilarity are displayed here. For the qualitative approach, the significant eigenvectors were MEM5, MEM22 (space fraction), MEM15, MEM17 (neigborhood fraction) and MEM4 (region fraction), while for the quantitative approach the significant ones were MEM13, MEM23 (space fraction), MEM11, MEM15 (neigborhood fraction) and MEM4 (region fraction)

et al. 2016). The eigenvectors derived from the connectivity were not significantly correlated with the biological data ($p > 0.05$) and were excluded from the subsequent analyses. For the environment fraction (Fig. 2G), environmental variables were standardized (except pH), and forward selection was applied to identify the relevant components.

The remaining variables (Fig. 2E, F) were compared to both quantitative and qualitative dissimilarity matrices (Fig. 2G) through variation partitioning (Fig. 2H) (Borcard et al. 1992). Finally, a canonical correspondence analysis (CCA) (Fig. 2I) was conducted to determine the significance of each fraction. All analyses were conducted at the R environment (R Core Team 2019). To complement the variation partitioning and evaluate whether β-diversity was mainly driven by stochastic or deterministic processes, we applied the β-nearest taxon index following Stegen et al. (2013). Phylogenetic dissimilarities among samples were compared to a null model based on this metacommunity species pool. Values equal to or greater than ± 2 were interpreted as deviating significantly from the null expectation and, further, being primarily guided by deterministic processes, whereas values not significantly deviating from the null expectation were considered to be predominantly driven by stochastic processes.

To further understand the role of increasing distances on the spatial fractions recovered, we calculated the maximum distance reached by each ASV to establish nine distinct dispersal thresholds, ranging from shorter distances (0–100 km) to the maximum distances reached in this landscape (800–900 km). These thresholds were then used to group the ASVs by their frequencies and relative abundances. Also, ASVs classified in clades previously recognized for their regional abundance and/or involvement in rapid and uncontrolled growth events (i.e. blooms) were grouped. Finally, differences in abundance for each grouped clade at each threshold of maximum distance were assessed using Kruskal-Wallis and Dunn's tests to evaluate significant differences.

The raw sequence data used in this study are deposited in the NCBI repository under accession number PRJNA411849, and are part of a continental effort led by the Collaborative Network on Microbial

Aquatic Ecology in Latin America (μSudAqua). More details can be found in Metz et al. (2022).

Results

After rarefaction and filtering, a total of 3,738 ASVs classified within 22 phyla were obtained from 815,560 reads. In terms of richness, the three most represented among the ASVs were *Proteobacteria* (30% of the total), *Actinobacteriota* (26.64%), and *Patescibacteria* (20.47%). When considering the quantitative number of affiliated reads, the top three phyla were *Actinobacteriota* (54.20%), *Proteobacteria* (24.37%), and *Verrucomicrobiota* (5.60%).

Before variation partitioning, the ordination of significant environmental variables showed that, for the qualitative dataset (0.084 adj. R^2) β-diversity was related to pH (0.021 adj. R^2; $p=0.002$), DOC (0.020 adj. R^2; $p=0.002$), altitude (0.014 adj. R^2; $p=0.016$), T-FDOM (0.014 adj. R^2; $p=0.016$), and DIN (0.008 adj. R^2; $p=0.04$), while for the quantitative dataset (0.131 adj. R^2) β-diversity was related to pH (0.042 adj. R^2; $p=0.002$), DIC (0.019 adj. R^2; $p=0.006$), and altitude (0.018 adj. R^2; $p=0.016$). These variables were then selected, and the other tested variables (temperature, conductivity, and chlorophyll-*a*) were excluded from subsequent analyses. All selected variables showed substantial variation among the sites analyzed here (Table 1).

In the variation partitioning analysis (Fig. 4), all comparisons indicated significant overall correlations. The environment was the best predictor of β-diversity. For the qualitative approach the environment contributed alone with ~ 6% of observed β-diversity, while geography reached ~ 1.5%. For the quantitative approach, environment remained to explain ~ 6%, while geography reached ~ 3.7%. In this analysis, the interaction between environment and geography also showed important values, reaching 3.3 and 1.9% for qualitative and quantitative approaches, respectively. Finally, the geographic partitioning indicated that, for both qualitative and quantitative approaches, neighborhood, space, and region, in this order, explained the geographic role on observed β-diversity. However, it is important to note that the unexplained fraction consistently accounted for a substantial portion of the variance. The β-Nearest Taxon Index showed that 60.05% of the observed dissimilarities were predominantly explained by stochastic factors, while the remaining 39.95% were explained by deterministic ones (Fig. S3).

The frequency distribution of maximum overland distances covered by each ASV (Fig. 5) followed a bimodal-like shape, with one group of organisms capable of dispersing shorter distances and another group able to disperse across nearly the entire landscape. This bimodal distribution aligns with the relative abundance of each ASV, as bacteria that are more abundant also tend to disperse over greater distances. However, this pattern does not correspond to the frequency of overland distances between sites within this landscape, which follows a normal-like distribution. Regardless of abundance, most freshwater bacterial ASVs had a distribution range of 600–700 km.

When examining these ASVs based on the maximum distances they can reach (Fig. 6), we observe a shift occurring beyond the 600–700 km threshold. Bacterial taxa that surpass this range differ from those restricted to shorter distances. Within the *Actinobacteriota* phylum, the *hgcI clade* becomes increasingly prominent at greater distances ($p<0.05$),

Table 1 Average and range of tested environmental parameters in the 60 tropical shallow lakes

	Average	Standard Deviation	Minimum	Maximum
Altitude (m)*	504.83	±230.03	7	1121
pH*	6.76	±1.02	4.79	10.1
Conductivity (μS/cm)	62.13	±79.53	0	470
Chlorophyll-*a* (mg/L)	16.62	±25.13	0.092	105.19
Temperature (°C)	23.16	±4.06	15.04	31.41
DOC (mg/L)*	8.8	±7.99	1.27	42.06
DIC (mg/L)*	4.79	±3.06	0.79	13.28
DIN (mg/L)*	1.27	±5.89	0.01	45.7
T-FDOM*	1.89	±1.27	0.38	6.32

Asterisks (*) indicate the parameters selected after ordination for the variation partitioning analysis

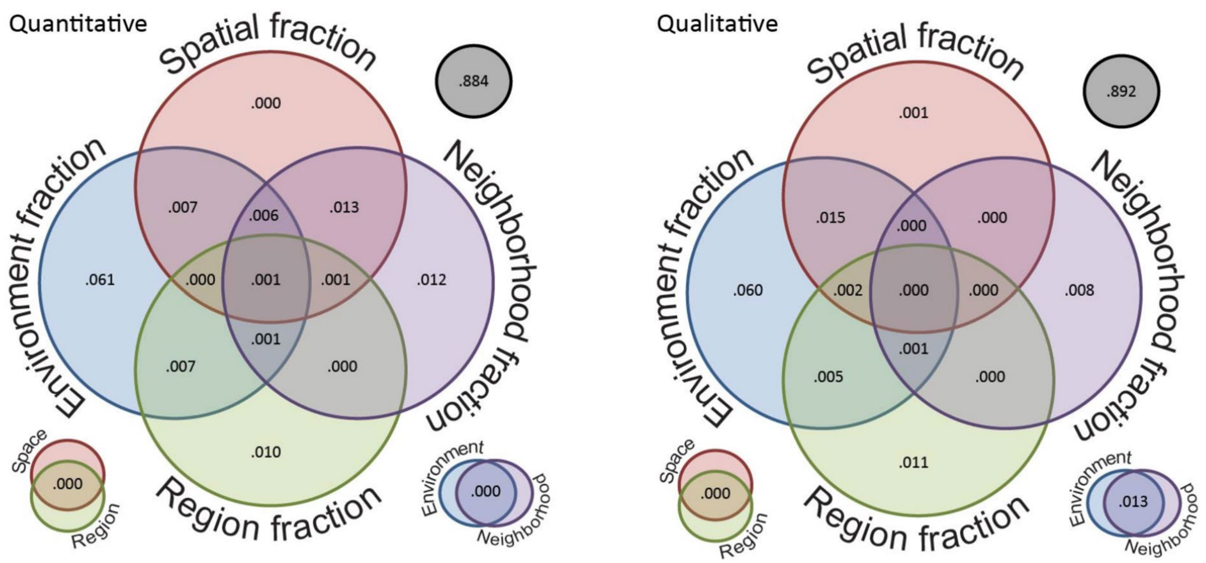

Fig. 4 Variation partitioning of bacterioplankton β-diversity for **A** qualitative and **B** quantitative data. Geographic factors are decoupled into spatial overland distance between sites (spatial fraction), increased connectivity to neighbor sites (neighborhood fraction) and the isolation promoted by geographical barriers (region fraction). The isolated circles show the non-explained fraction for each analysis

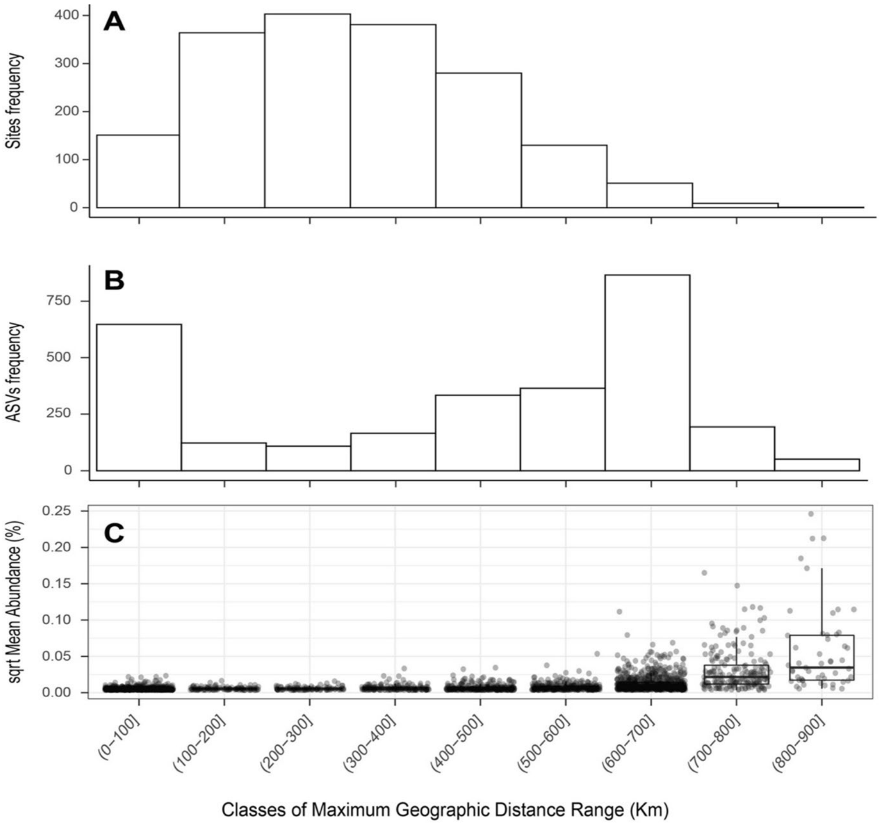

Fig. 5 Threshold of maximum overland distances achieved in this landscape. **A** The distances between sites showed a normal-like distribution frequency, while **B** the bacterioplankton had a bimodal distribution on their maximum distances reached. **C** The organisms capable of reaching the maximum distances in this landscape were also those with a greater mean relative abundance. This data was normalized by the *sqrt* function to improve visualization

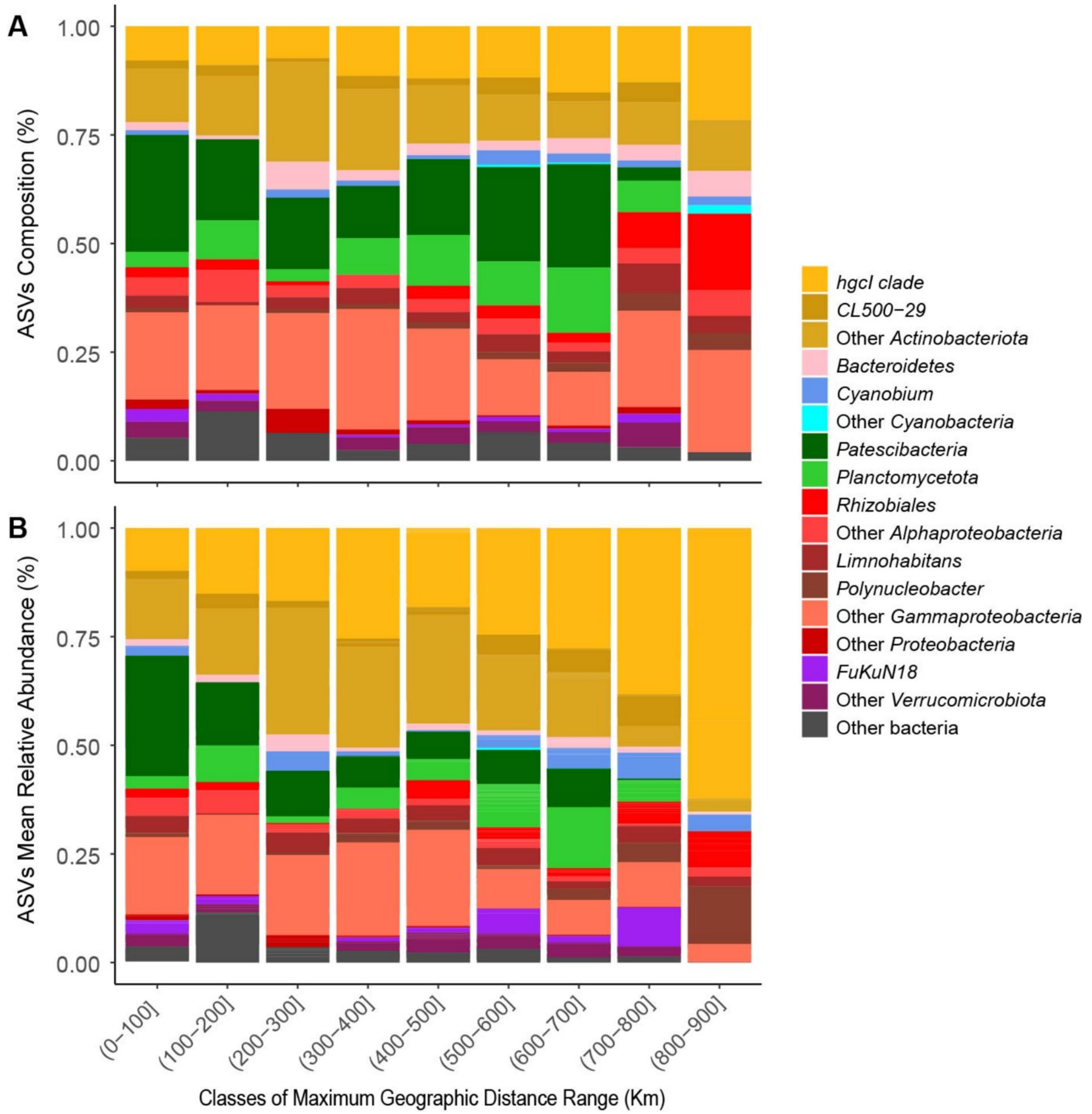

Fig. 6 The distribution of ASV representatives (**A**) and their mean relative abundance (**B**) by the thresholds of maximum distances reached. Clades previously recognized for their regional abundance and/or involvement in rapid and uncontrolled growth events (i.e. blooms) were highlighted

not only showing the highest number of representatives but also displaying a rise in relative abundance as dispersal distances increase. In contrast, bacteria belonging to *Patescibacteria* tend to decrease in representation as the maximum distance between sites increases ($p < 0.05$). The *Proteobacteria*, another group with a substantial number of representatives, maintains a relatively constant presence across this spatial spectrum. However, the relative abundance of *Proteobacteria* decreases with increasing spatial scale (*Rhizobiales*, $P < 0.05$; other_*Alphaproteobacteria*, $P < 0.05$; *Limnohabitans*, $P < 0.05$; *Polynucleobacter*,

$P < 0.05$; other G*ammaproteobacteria*, $P < 0.05$; other *Proteobacteria*, $P < 0.05$), except for the Rhizobiales family, which exhibits an increase in both the number of representatives and relative abundance at larger maximum distances. Other relevant taxonomic groups, such as Cyanobacteria (*Cyanobium*, $P < 0.05$; other *Cyanobacteria*, $P = 0.117$), Verrucomicrobiota (*FuKuN18*, $P = 0.1036$; other *Verrucomicrobia*, $P < 0.05$), and *Planctomycetes* ($p < 0.05$), show a more consistent representation across this spatial gradient. For a detailed description of the global significance and pairwise comparisons across each threshold of distances, for each clade highlighted here, see Table S1.

Discussion

In recent decades, the development of tools such as powerful computers, molecular and sensing techniques, and the facilitation of data dissemination via the internet, has contributed to a shift in dominant paradigms in ecology (Chave 2013). It has become increasingly evident that addressing certain fundamental questions requires exploring other scales to understand the drivers of observed patterns (Levin 1992). In this study, we used a large dataset of freshwater bacterial biodiversity aimed at applying some conceptual frameworks to disentangle how different geographical factors can affect bacterioplankton β-diversity in a lake metacommunity.

To achieve this objective, we applied various models designed to assess the influence of different potential dispersal routes using variation partitioning (e.g. Dray et al. 2006; Blanchet et al. 2008a, b; Peres-Neto and Legendre 2010). We found that the environment is the best predictor of bacterioplankton β-diversity in shallow lake metacommunities, although spatial factors also play a role, mainly for bacteria that are not capable of overcoming long distances. The importance of each factor followed a hierarchy in which β-diversity was primarily related to the environmental fraction, and afterwards by dispersion, as evidenced by distinct geographic features. This pattern was consistent across several tested factors and approaches, a result that was not previously expected, and reflects the high dispersal capacity of dominant bacteria that are able to overcome large distances and geographic barriers (Lansac-Tôha et al. 2020). However, we also identified a group of rare and less dispersed organisms that are mainly influenced by spatial features, as evidenced by Figs. 5 and 6. The role of the environment as the main driver of β-diversity is a common pattern described in the literature (e.g. Beisner et al. 2006; Winter et al. 2013; Jyrkänkallio-Mikkola et al. 2017; Fillinger et al. 2019; Porcel et al. 2025). This is especially attributed to pH, a well-known selective factor for aquatic bacteria (Lindström et al. 2005; Niño-García et al. 2016; Griffero et al. 2024; Porcel et al. 2025), which has previously been shown to be a key factor in these sites (Mateus-Barros et al. 2021). Here, we focused on refining the spatial analyses as much as possible, aiming to encompass all relevant dispersal routes that bacterioplankton can take. To achieve this objective, we applied different models that have been proposed to assess the role of geography through variation partitioning (e.g. Dray et al. 2006; Blanchet et al. 2008a, b; Peres-Neto and Legendre 2010). Despite some concerns regarding the accuracy of this approach for determining the role of environmental and spatial factors on metacommunity structuring (Gilbert and Bennett 2010), it has been demonstrated that spatial autocorrelation can be addressed through an appropriate correction (Bauman et al. 2018).

Our findings indicate that the most effective geographic predictor of bacterioplankton β-diversity in shallow lake metacommunities is the overland distance between sites (spatial fraction) in combination with the other measured fractions, with most bacterial ASVs having a distribution range of 600–700 km, and the most abundant ones presenting an even higher dispersion, reaching the maximum possible distance for this dataset (800–900 km). This pattern likely reflects the high dispersal capacity of dominant bacteria, allowing them to overcome large distances and geographic barriers (Lansac-Tôha et al. 2020). However, we also identified a set of rare and less dispersed organisms that are more consistently associated with the Neighborhood feature, suggesting a lower dispersal capacity of these organisms compared to the more abundant ones. This feature demonstrates that these organisms should not be solely guided by local environmental factors as long has been stated (Baas Becking 1934), instead, it seems to present spatial dynamics similar to what is largely observed in groups that disperse actively, like birds and mammals (Brown 1984; Matthysen 2005). This pattern indicates that bacteria can present a density-dependent dispersal (Matthysen 2005), which is consistent with the rescue effect hypothesis (Gotelli 1991), where a given area with a larger population of a species serves as a dispersal source, and their distribution range tends to correlate with the maximum local abundance they achieve (Brown 1984). This implies that organisms with greater adaptive capabilities tend to become more abundant and eventually dominate the landscape (Gaston et al. 2000). In contrast, rare organisms typically exhibit a clustered distribution and have limited dispersal capacities (Ruiz-González et al. 2015; Niño-García et al. 2016).

In this landscape, the dominant organisms were classified within the *hgcI clade*. It is a group of *Actinobacteriota* known for having a streamlined genome (Chiriac et al. 2022) and an auxotrophic lifestyle (Kim et al. 2019). This means that this group lacks the full molecular apparatus to degrade certain organic compounds required for survival. Instead, they exploit available nutrients in the environment, thereby reducing energy expenditure and increasing metabolic efficiency and reproductive rates (Chiriac et al. 2022). Despite being still poorly understood and just recently being successfully maintained in laboratory cultures (Kim et al. 2019), these organisms are largely found in freshwater environments (Glöckner et al. 2000) and generally represent more than half of 16S rRNA gene amplicon sequences (Mateus-Barros et al. 2019). In contrast, the rarer organisms belonged primarily to the *Patescibacteria*, a group largely characterized by an endosymbiont lifestyle (Castelle et al. 2018) and a consequent dependence on their hosts' dispersal capacity to colonize new sites.

Additionally, the variation partitioning analysis revealed a large portion of unexplained variation, and the complementary analysis of β-nearest taxon index has confirmed that, for the majority of site pairs, the observed dissimilarity may be explained by stochastic processes. This phenomenon could be attributed to two main factors, acting independently or in combination. First, it is possible that some crucial variables influencing the metacommunity structure were not captured. Some temporally structured environmental factors (Langenheder et al. 2012) may impact a region in ways that are more difficult to measure. Additionally, factors operating at larger scales, such as disturbances (Vellend et al. 2014) and priority effects (Siqueira et al. 2015), may be playing a significant role. Predation pressure (Livingston et al. 2017; Segovia et al. 2018) and viral lysis (Maurice et al. 2011) can also affect community structure and were not measured in this study. Furthermore, the mass flow of individuals from surrounding soils, which occurs when water flows towards the riverbed, is more common in headwater environments (Ruiz-González et al. 2015; Griffero et al. 2024) and could contribute to unexplained variance in the absence of disturbances over time. Second, this unexplained variation may be an inherent characteristic of the region, with neutral processes potentially driving community dynamics (Melo et al. 2012). This phenomenon has been previously observed in the same region for other organisms, suggesting that it may lead to increased local variability and reduced regional synchrony (Lopes et al. 2017; Zanon et al. 2018).

In our study, we successfully identified various dispersal mechanisms that may impact bacterial community composition. The dbMEM approach, commonly used in studies of this nature (e.g. Sommaruga and Casamayor 2009; Fillinger et al. 2019), effectively captured spatial variables that explained variations in dominant taxa across the landscape. Although concerns have been raised about the accuracy of this approach in determining the roles of spatial factors in metacommunity structuring (Gilbert and Bennett 2010), it has been demonstrated that spatial autocorrelation can be effectively addressed through appropriate correction methods (Bauman et al. 2018). This approach highlighted the unexpected potential role of spatially related selective pressures on these organisms. Conversely, the MEM approach provided insights into neighborhood and region aspects, which better explained variations in bacteria that were rarer and more susceptible to a certain degree of geographic isolation. This underscores the importance of employing multiple approaches to capture all geographic factors influencing metacommunity structure. Evidence suggesting that geography plays a role in bacterial distribution is growing (e.g. Stegen et al. 2013, Lindh 2017, Lansac-Tôha et al. 2020, Logares et al. 2020, Mateus-Barros et al. 2021). However, further studies covering large geographical scales, using comparable methods, and encompassing at least 700 km of overland distance between sampling sites are necessary. This emphasizes the need for continental-scale sampling programs that apply standardized protocols to comprehensively unravel the complex geographic mechanisms shaping bacterial metacommunities.

Acknowledgements We are grateful to Prof. Armando Vieira and Prof. Inessa Lacativa Bagatini for providing the samples from the Biota-FAPESP project and to his lab team Cilene C. Mori, Fabrício S. Meccheri, Guilherme P. de Moraes, Helena H. Vieira, Ingritt C. Moreira, Letícia P. Tessarolli, Lucas S. Tiberti, Luiz A. Joaquim, Naiara C. P. Archanjo and Rodrigo V. de Mello, who participated in the field sampling and performed physical and chemical analyses. This study was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—CAPES [Grant 001], Conselho Nacional de Desenvolvimento Científico e Tecnológico—CNPq [Grant Number: 304655/2025-2], and the Fundação de Amparo à Pesquisa do Estado de São Paulo—FAPESP [Grant Numbers: 2014/14139-3 and 2011/50054-4]

Author's contribution EMB, TGS and HS participated in field sampling and laboratory analyses; EMB extracted DNA, prepared samples for Illumina sequencing, run the bioinformatics and statistical analyses; EMB, AC, TGS and HS idealized this research, analysed data and wrote the manuscript; and HS obtained funding for the project.

Data availability The raw sequence data used in this study are deposited in the NCBI repository under accession number PRJNA411849.

Declarations

Conflict of interest The authors declare no competing interests.

References

Anderson MJ, Crist TO, Chase JM, Vellend M, Inouye BD, Freestone AL, Sanders NJ, Cornell HV, Comita LS, Davies KF (2011) Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. Ecol Lett 14(1):19–28. https://doi.org/10.1111/j.1461-0248.2010.01552.x

Baas Becking LGM (1934) Geobiologie of inleiding tot de milieukunde. Den Haag, Van Stockum

Baselga A (2010) Partitioning the turnover and nestedness components of beta diversity. Glob Ecol Biogeogr 19(1):134–143. https://doi.org/10.1111/j.1466-8238.2009.00490.x

Baselga A, Freckleton R (2013) Separating the two components of abundance-based dissimilarity: balanced changes in abundance vs. abundance gradients. Methods Ecol Evol 4(6):552–557. https://doi.org/10.1111/2041-210x.12029

Baselga A, Lobo JM, Svenning JC, Araújo MB (2011) Global patterns in the shape of species geographical ranges reveal range determinants. J Biogeogr 39(4):760–771. https://doi.org/10.1111/j.1365-2699.2011.02612.x

Bauman D, Drouet T, Dray S, Vleminckx J (2018) Disentangling good from bad practices in the selection of spatial or phylogenetic eigenvectors. Ecography 41(10):1638–1649. https://doi.org/10.1111/ecog.03380

Beisner BE, Peres-Neto PR, Lindström ES, Barnett A, Longhi ML (2006) The role of environmental and spatial processes in structuring lake communities from bacteria to fish. Ecology 87(12):2985–2991. https://doi.org/10.1890/0012-9658(2006)87[2985:TROEAS]2.0.CO;2

Blanchet FG, Legendre P, Borcard D (2008a) Forward selection of explanatory variables. Ecology 89(9):2623–2632. https://doi.org/10.1890/07-0986.1

Blanchet FG, Legendre P, Borcard D (2008b) Modelling directional spatial processes in ecological data. Ecol Modell 215(4):325–336. https://doi.org/10.1016/j.ecolmodel.2008.04.001

Borcard D, Legendre P, Drapeau P (1992) Partialling out the spatial component of ecological variation. Ecology 73(3):1045–1055. https://doi.org/10.2307/1940179

Brown JH (1984) On the relationship between abundance and distribution of species. Am Nat 124(2):255–279. https://doi.org/10.1086/284267

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP (2016) DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods 13(7):581–583. https://doi.org/10.1038/nmeth.3869

Castelle CJ, Brown CT, Anantharaman K, Probst AJ, Huang RH, Banfield JF (2018) Biosynthetic capacity, metabolic variety and unusual biology in the CPR and DPANN radiations. Nat Rev Microbiol 16(10):629–645. https://doi.org/10.1038/s41579-018-0076-2

Chave J (2013) The problem of pattern and scale in ecology: What have we learned in 20 years? Ecol Lett 16(Suppl 1):4–16. https://doi.org/10.1111/ele.12048

Chiriac MC, Haber M, Salcher MM (2022) Adaptive genetic traits in pelagic freshwater microbes. Environ Microbiol. https://doi.org/10.1111/1462-2920.16313

Dray S, Legendre P, Peres-Neto PR (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecol Modell 196(3–4):483–493. https://doi.org/10.1016/j.ecolmodel.2006.02.015

Dray S, Bauman D, Guillaume B, Borcard D, Clappe S, Guenard G (2020) Package 'adespatial'

Farjalla VF, Srivastava DS, Marino NAC, Azevedo FD, Dib V, Lopes PM, Rosado AS, Bozelli RL, Esteves FA (2012) Ecological determinism increases with organism size. Ecology 93(7):1752–1759. https://doi.org/10.1890/11-1144.1

Fillinger L, Hug K, Griebler C (2019) Selection imposed by local environmental conditions drives differences in microbial community composition across geographically distinct groundwater aquifers. FEMS Microbiol Ecol. https://doi.org/10.1093/femsec/fiz160

Gaston KJ (2009) Geographic range limits: achieving synthesis. Proc Biol Sci 276(1661):1395–1406. https://doi.org/10.1098/rspb.2008.1480

Gaston KJ, Blackburn TM, Greenwood JJ, Gregory RD, Quinn RM, Lawton JH (2000) Abundance–occupancy relationships. J Appl Ecol 37(s1):39–59. https://doi.org/10.1046/j.1365-2664.2000.00485.x

Gilbert B, Bennett JR (2010) Partitioning variation in ecological communities: Do the numbers add up? J Appl Ecol

47(5):1071–1082. https://doi.org/10.1111/j.1365-2664.2010.01861.x

Glöckner FO, Zaichikov E, Belkova N, Denissova L, Pernthaler J, Pernthaler A, Amann R (2000) Comparative 16S rRNA analysis of lake bacterioplankton reveals globally distributed phylogenetic clusters including an abundant group of actinobacteria. Appl Environ Microbiol 66(11):5053–5065. https://doi.org/10.1128/AEM.66.11.5053-5065.2000

Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ (2017) Microbiome datasets are compositional: and this is not optional. Front Microbiol 8:2224. https://doi.org/10.3389/fmicb.2017.02224

Gotelli NJ (1991) Metapopulation models: the rescue effect, the propagule rain, and the core-satellite hypothesis. Am Nat 138(3):768–776. https://doi.org/10.1086/285249

Griffero L, Pereira-Flores E, Lescano C, Rodríguez-Gallego L, Pérez-Parada A, Alonso C (2024) Microbial communities indicate fine differences in pollution levels by emerging contaminants. Ecol Indic 169:112875. https://doi.org/10.1016/j.ecolind.2024.112875

Grossart H-P, Dziallas C, Leunert F, Tang KW (2010) Bacteria dispersal by hitchhiking on zooplankton. Proc Natl Acad Sci U S A 107(26):11959–11964. https://doi.org/10.1073/pnas.1000668107

Herlemann DP, Labrenz M, Jürgens K, Bertilsson S, Waniek JJ, Andersson AF (2011) Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. ISME J 5(10):1571–1579. https://doi.org/10.1038/ismej.2011.41

Hoetzinger M, Hahn MW, Andersson LY, Buckley N, Ramsin C, Buck M, Nuy JK, Garcia SL, Puente-Sánchez F, Bertilsson S (2024) Geographic population structure and distinct intra-population dynamics of globally abundant freshwater bacteria. ISME J. https://doi.org/10.1093/ismejo/wrae113

Huber P, Metz S, Unrein F, Mayora G, Sarmento H, Devercelli M (2020) Environmental heterogeneity determines the ecological processes that govern bacterial metacommunity assembly in a floodplain river system. ISME J. https://doi.org/10.1038/s41396-020-0723-2

Jyrkänkallio-Mikkola J, Meier S, Heino J, Laamanen T, Pajunen V, Tolonen KT, Tolkkinen M, Soininen J (2017) Disentangling multi-scale environmental effects on stream microbial communities. J Biogeogr. https://doi.org/10.1111/jbi.13002

Kim S, Kang I, Seo J-H, Cho J-C (2019) Culturing the ubiquitous freshwater actinobacterial acI lineage by supplying a biochemical ‘helper’ catalase. ISME J 13(9):2252–2263. https://doi.org/10.1038/s41396-019-0432-x

Langenheder S, Berga M, Ostman O, Szekely AJ (2012) Temporal variation of beta-diversity and assembly mechanisms in a bacterial metacommunity. ISME J 6(6):1107–1114. https://doi.org/10.1038/ismej.2011.177

Lansac-Tôha FM, Bini LM, Heino J, Meira BR, Segovia BT, Pavanelli CS, Bonecker CC, Deus CP, Benedito E, Alves GM, Manetta GI, Dias JD, Vieira LCG, Rodrigues LC, Carmo Roberto M, Brugler MR, Lemke MJ, Tessler M, DeSalle R, Mormul RP, Amadio S, Lolis SF, Jati S, Siqueira T, Silva WM, Higuti J, Lansac-Tôha FA, Martens K, Velho LFM (2020) Scale-dependent patterns of metacommunity structuring in aquatic organisms across floodplain systems. J Biogeogr 48(4):872–885. https://doi.org/10.1111/jbi.14044

Legendre P, Anderson MJ (1999) Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. Ecol Monogr 69(1):1–24. https://doi.org/10.1890/0012-9615(1999)069[0001:DBRATM]2.0.CO;2

Leibold MA, Holyoak M, Mouquet N, Amarasekare P, Chase JM, Hoopes MF, Holt RD, Shurin JB, Law R, Tilman D, Loreau M, Gonzalez A (2004) The metacommunity concept: a framework for multi-scale community ecology. Ecol Lett 7(7):601–613. https://doi.org/10.1111/j.1461-0248.2004.00608.x

Levin SA (1992) The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. Ecology 73(6):1943–1967. https://doi.org/10.2307/1941447

Lindh MV (2017) There and back again—unraveling mechanisms of bacterial biogeography in the north pacific subtropical gyre to and from station ALOHA. bioRxiv. https://doi.org/10.1101/141085

Lindström ES, Kamst-Van Agterveld MP, Zwart G (2005) Distribution of typical freshwater bacterial groups is associated with pH, temperature, and lake water retention time. Appl Environ Microbiol 71(12):8201–8206. https://doi.org/10.1128/AEM.71.12.8201-8206.2005

Livingston G, Fukumori K, Provete DB, Kawachi M, Takamura N, Leibold MA (2017) Predators regulate prey species sorting and spatial distribution in microbial landscapes. J Anim Ecol. https://doi.org/10.1111/1365-2656.12639

Logares R, Deutschmann IM, Junger PC, Giner CR, Krabberod AK, Schmidt TSB, Rubinat-Ripoll L, Mestre M, Salazar G, Ruiz-Gonzalez C, Sebastian M, de Vargas C, Acinas SG, Duarte CM, Gasol JM, Massana R (2020) Disentangling the mechanisms shaping the surface ocean microbiota. Microbiome 8(1):55. https://doi.org/10.1186/s40168-020-00827-8

Lopes VG, Castelo Branco CW, Kozlowsky-Suzuki B, Sousa-Filho IF, Souza LCE, Bini LM (2017) Predicting temporal variation in zooplankton beta diversity is challenging. PLoS ONE 12(11):e0187499. https://doi.org/10.1371/journal.pone.0187499

Lorenzen CJ (1967) Determination of chlorophyll and pheopigments: spectrophotometric equations. Limnol Oceanogr 12(2):343–346. https://doi.org/10.4319/lo.1967.12.2.0343

Marker A (1980) The measurement of photosynthetic pigments in freshwaters and standardization of methods: conclusions and recommendations. Arch Hydrobiol Ergebn Limnol 14:91–106

Martiny JBH, Eisen JA, Penn K, Allison SD, Horner-Devine MC (2011) Drivers of bacterial β-diversity depend on spatial scale. Proc Natl Acad Sci U S A 108(19):7850–7854. https://doi.org/10.1073/pnas.1016308108

Mateus-Barros E, Meneghine AK, Bagatini IL, Fernandes CC, Kishi LT, Vieira AAH, Sarmento H (2019) Comparison of two DNA extraction methods widely used in aquatic microbial ecology. J Microbiol Methods 159:12–17. https://doi.org/10.1016/j.mimet.2019.02.005

Mateus-Barros E, de Melo ML, Bagatini IL, Caliman A, Sarmento H (2021) Local and geographic factors shape the occupancy-frequency distribution of freshwater bacteria. Microb Ecol. https://doi.org/10.1007/s00248-020-01560-3

Matthysen E (2005) Density-dependent dispersal in birds and mammals. Ecography 28(3):403–416. https://doi.org/10.1111/j.0906-7590.2005.04073.x

Maurice CF, Mouillot D, Bettarel Y, De Wit R, Sarmento H, Bouvier T (2011) Disentangling the relative influence of bacterioplankton phylogeny and metabolism on lysogeny in reservoirs and lagoons. ISME J 5(5):831–842. https://doi.org/10.1038/ismej.2010.181

Melo AS, Schneck F, Hepp LU, Simões NR, Siqueira T, Bini LM (2012) Focusing on variation: methods and applications of the concept of beta diversity in aquatic ecosystems. Acta Limnol Bras 23(3):318–331. https://doi.org/10.1590/s2179-975x2012005000010

Metz S, Huber P, Mateus-Barros E, Junger PC, de Melo M, Bagatini IL, Izaguirre I, Câmara dos Reis M, Llames ME, Accattatis V, Quiroga MV, Devercelli M, Schiaffino MR, Niño-García JP, Bastidas Navarro M, Modenutti B, Vieira H, Saraceno M, Sabio CA, y, García, Pereira E, González-Revello A, Piccini C, Unrein F, Alonso C, Sarmento H (2022) A georeferenced rRNA amplicon database of aquatic microbiomes from South America. Sci Data 9(1):565. https://doi.org/10.1038/s41597-022-01665-z

Mush E (1980) Comparison of different methods for chlorophyll and phaeopigment determination. Arch Hydrobiol Beih 14:14–36

Newton RJ, Jones SE, Eiler A, McMahon KD, Bertilsson S (2011) A guide to the natural history of freshwater lake bacteria. Microbiol Mol Biol Rev 75(1):14–49. https://doi.org/10.1128/MMBR.00028-10

Niño-García JP, Ruiz-González C, del Giorgio PA (2016) Interactions between hydrology and water chemistry shape bacterioplankton biogeography across boreal freshwater networks. ISME J. https://doi.org/10.1038/ismej.2015.226

Oksanen J, Blanchet FG,Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E and Wagner H (2016). Vegan: community ecology package

Pedrós-Alió C (2012) The rare bacterial biosphere. Ann Rev Mar Sci 4:449–466. https://doi.org/10.1146/annurev-marine-120710-100948

Peres-Neto PR, Legendre P (2010) Estimating and controlling for spatial structure in the study of ecological communities. Glob Ecol Biogeogr 19(2):174–184. https://doi.org/10.1111/j.1466-8238.2009.00506.x

Porcel S, Sabio y García CA, Schiaffino MR, Chaparro G, Diéguez MdC, Lancelotti J, Roesler I, Izaguirre I (2025) How do environmental and spatial factors influence beta diversity and composition of abundant and rare bacterioplankton communities in lakes from the Patagonian plateaus? Freshw Biol. https://doi.org/10.1111/fwb.70047

R Core Team (2019) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Rappé MS, Giovannoni SJ (2003) The uncultured microbial majority. Annu Rev Microbiol 57(1):369–394. https://doi.org/10.1146/annurev.micro.57.030502.090759

Ricklefs RE, Jenkins DG (2011) Biogeography and ecology: towards the integration of two disciplines. Philos Trans R Soc Lond B Biol Sci 366(1576):2438–2448. https://doi.org/10.1098/rstb.2011.0066

Ruiz-González C, Niño-García JP, del Giorgio PA (2015) Terrestrial origin of bacterial communities in complex boreal freshwater networks. Ecol Lett 18(11):1198–1206. https://doi.org/10.1111/ele.12499

Sadeghi J, Venney CJ, Wright S, Watkins J, Manning D, Bai E, Frank C, Heath DD (2024) Aquatic bacterial community connectivity: the effect of hydrological flow on community diversity and composition. Environments 11(5):90. https://doi.org/10.3390/environments11050090

Segovia BT, Meira BR, Lansac-Toha FM, Amadeo FE, Unrein F, Velho LFM, Sarmento H (2018) Growth and cytometric diversity of bacterial assemblages under different top–down control regimes by using a size-fractionation approach. J Plankton Res 40(2):129–141. https://doi.org/10.1093/plankt/fbx071

Siqueira T, Lacerda CGLT, Saito VS (2015) How does landscape modification induce biological homogenization in tropical stream metacommunities? Biotropica 47(4):509–516. https://doi.org/10.1111/btp.12224

Smith DJ, Timonen HJ, Jaffe DA, Griffin DW, Birmele MN, Perry KD, Ward PD, Roberts MS (2013) Intercontinental dispersal of bacteria and archaea by transpacific winds. Appl Environ Microbiol 79(4):1134–1139. https://doi.org/10.1128/AEM.03029-12

Soininen J, Korhonen JJ, Luoto M (2013) Stochastic species distributions are driven by organism size. Ecology 94(3):660–670. https://doi.org/10.1890/12-0777.1

Soininen J, Heino J, Wang J (2018) A meta-analysis of nestedness and turnover components of beta diversity across organisms and ecosystems. Glob Ecol Biogeogr 27(1):96–109. https://doi.org/10.1111/geb.12660

Sommaruga R, Casamayor EO (2009) Bacterial 'cosmopolitanism' and importance of local environmental factors for community composition in remote high-altitude lakes. Freshw Biol 55(5):994–1005. https://doi.org/10.1111/j.1365-2427.2008.02146.x

Sriswasdi S, Yang C-C, Iwasaki W (2017) Generalist species drive microbial dispersion and evolution. Nat Commun 8(1):1162. https://doi.org/10.1038/s41467-017-01265-1

Stegen JC, Lin X, Fredrickson JK, Chen X, Kennedy DW, Murray CJ, Rockhold ML, Konopka A (2013) Quantifying community assembly processes and identifying features that impose them. ISME J 7(11):2069–2079. https://doi.org/10.1038/ismej.2013.93

Székely AJ, Langenheder S (2014) The importance of species sorting in bacterial communities differs between habitat generalists and specialists. FEMS Microbiol Ecol 87(1):102–112. https://doi.org/10.1111/1574-6941.12195

Vellend M (2010) Conceptual synthesis in community ecology. Q Rev Biol 85(2):183–206. https://doi.org/10.1086/652373

Vellend M, Srivastava DS, Anderson KM, Brown CD, Jankowski JE, Kleynhans EJ, Kraft NJB, Letaw AD, Macdonald AAM, Maclean JE, Myers-Smith IH, Norris AR, Xue X (2014) Assessing the relative importance of neutral stochasticity in ecological communities. Oikos 123(12):1420–1430. https://doi.org/10.1111/oik.01493

Whittaker RH (1960) Vegetation of the Siskiyou Mountains, Oregon and California. Ecol Monogr 30(3):279–338. https://doi.org/10.2307/1943563

Whittaker RH (1972) Evolution and measurement of species diversity. Taxon 21(2–3):213–251. https://doi.org/10.2307/1218190

Wilson DS (1992) Complex interactions in metacommunities, with implications for biodiversity and higher levels of selection. Ecology 73(6):1984–2000. https://doi.org/10.2307/1941449

Winter C, Matthews B, Suttle CA (2013) Effects of environmental variation and spatial distance on bacteria, archaea and viruses in sub-polar and arctic waters. ISME J 7(8):1507–1518. https://doi.org/10.1038/ismej.2013.56

Yilmaz P, Parfrey L, Yarza P, Gerken J, Pruesse E, Quast C, Schweer T, Peplies J, Ludwig W, Glockner F (2014) The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. Nucleic Acids Res 42:D643–D648

Zanon JE, Rodrigues L, Bini LM (2018) Hard to predict: synchrony in epiphytic biomass in a floodplain is independent of spatial proximity, environmental distance, and environmental synchrony. Ecol Indic 93:379–386. https://doi.org/10.1016/j.ecolind.2018.05.023

Zhang G, Bai J, Tebbe CC, Zhao Q, Jia J, Wang W, Wang X, Yu L (2021) Salinity controls soil microbial community structure and function in coastal estuarine wetlands. Environ Microbiol 23(2):1020–1037. https://doi.org/10.1111/1462-2920.15281