Universidade Federal de São Carlos

Centro de Ciências Biológicas e da Saúde

Programa de Pós-Graduação em Ecologia e Recursos Naturais

Erick Mateus Barros

Padrões e processos que estruturam (meta)comunidades de bacterioplâncton, da escala local à intercontinental

São Carlos
2023

Universidade Federal de São Carlos

Centro de Ciências Biológicas e da Saúde

Programa de Pós-Graduação em Ecologia e Recursos Naturais

Erick Mateus Barros

Padrões e processos que estruturam (meta)comunidades de bacterioplâncton, da escala local à intercontinental

Tese apresentada ao Programa de Pós-Graduação em Ecologia e Recursos Naturais da Universidade Federal de São Carlos, como parte dos requisitos para obtenção do título de DOUTOR em CIÊNCIAS, área de concentração: Ecologia e Recursos Naturais.

Orientador: Prof. Dr. Hugo Sarmento

São Carlos
2023

UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Biológicas e da Saúde
Programa de Pós-Graduação em Ecologia e Recursos Naturais

Folha de Aprovação

Defesa de Tese de Doutorado do candidato Erick Mateus Barros, realizada em 03/03/2023.

Comissão Julgadora:

Prof. Dr. Hugo Miguel Preto de Morais Sarmento (UFSCar)

Prof. Dr. Gilmar Perbiche Neves (UFSCar)

Prof. Dr. Fernando Rodrigues da Silva (UFSCar)

Profa. Dra. Irina Izaguirre (UBA)

Prof. Dr. Juan Pablo Niño-Garcia (UA)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Ecologia e Recursos Naturais.

Agradecimentos

Aos meus pais, os primeiros e maiores patrocinadores dessa minha vontade em me tornar pesquisador.

À Gabrielle e à Sophia, minhas companheiras inseparáveis (e, pelos últimos 3 anos, de forma compulsória), nos momentos de alegria e frustração, excitação e tristeza. Aqui, agora e sempre.

Ao Prof. Dr. Hugo Sarmento, orientador, conselheiro, mestre. Por me acolher quando soubemos da Sophia e por sua paciência quando precisei estar ausente. E também por ter me guiado por todo esse caminho; entrei no laboratório sem saber segurar uma pipeta com segurança, saio doutor em ecologia microbiana, foi um longo caminho. Satisfatório, espero.

Ao Vinicius, dentre todos os meus amigos, a pessoa mais distante fisicamente, mas graças a internet, a mais próxima emocionalmente

Ao João e à Roseli, cuja disponibilidade e muita vontade foram imprescindíveis para que conseguisse me manter na pós-graduação com o iminente nascimento da Sophia e em incontáveis momentos após isso.

Aos membros das bancas de qualificação e defesa, por sua disponibilidade em participar do presente trabalho, contribuindo com dicas e sugestões para a sua melhoria.

Ao Prof. Dr. Armando Vieira e à Profa. Dra. Inessa Bagatini, responsáveis pelas coletas realizadas por todo o Estado de São Paulo, bem como à Prof. Dra. Maria Llammes e ao Prof. Dr. Juan Pablo Niño-Garcia que disponibilizaram suas bases de dados microbianas coletadas na Argentina e no Canadá, respectivamente. Gostaria de agradecer também à todos os estudantes de graduação, mestrado e doutorado que durante muitos anos realizaram e auxiliaram nas coletas de todas as amostras que acabaram sendo reunidas neste presente trabalho. Agradeço também a diversas pessoas que contribuíram para a concepção deste trabalho participando como coautores além das pessoas já citadas neste parágrafo: Dra. Clara M.

Resumo

Bactérias são organismos-chave para o funcionamento dos ecossistemas, sendo essenciais em ciclos biogeoquímicos e intimamente relacionadas com os outros seres vivos. Apesar de esses organismos terem sido estudados por décadas, as dinâmicas de sua estruturação espaço-temporal puderam ser mais bem compreendidas após o advento de abordagens moleculares independentes de cultivo. Estas novas técnicas têm evidenciado que nem todas as bactérias são ubíquas e a existência de padrões biogeográficos de microrganismos. Além disso, o estudo integrado deste grupo e macrorganismos pode impulsionar as descobertas no campo da ecologia e, em contrapartida a aplicação de modelos matemáticos já estabelecidos fornece novas perspectivas sobre a macroecologia de microrganismos. A presente tese foi dividida em duas partes. Na parte I eu trago uma introdução teórica a respeito do desenvolvimento da disciplina na última década e um sumário dos últimos avanços na área da ecologia molecular a respeito dos grupos taxonômicos mais comumente encontrados em ambientes aquáticos continentais. Na parte II, apresento os resultados de minhas pesquisas divididas em três capítulos. O primeiro capítulo desta parte trata dos processos locais que potencialmente impactam na capacidade de bactérias se manterem consistentemente abundantes ao longo do tempo, o que pode influenciar no alcance da distribuição espacial em uma metacomunidade, já que organismos mais abundantes geralmente têm maiores chances de estar presentes em todos os locais de uma determinada região. No segundo capítulo, eu analiso os padrões de dispersão mais relevantes para bactérias em uma matriz espacial com cerca de 250 mil km². Por fim, no capítulo 3 busquei entender qual o impacto do aumento da escala sobre os processos determinísticos e estocásticos que estruturam uma metacomunidade de bactérioplâncton. Este trabalho contribui para o entendimento de como microrganismos podem estar distribuídos em uma paisagem em águas continentais e quais fatores potencialmente determinam a sua distribuição e abundância.

Palavras-chave: Ecologia microbiana; estruturação espaço-temporal; paisagem; partição da variância; dinâmicas de dispersão; escala; fatores determinísticos e estocásticos; bactérias de lagos rasos; sequenciamento amplicons

Abstract

Bacteria are key organisms for ecosystem functioning, being essential in biogeochemical cycles and closely related to other living organisms. Although these organisms have been studied for decades, their dynamics of space-time structuring could be better understood after the advent of culture-independent molecular approaches. These new techniques have revealed that not all bacteria are ubiquitous and the existence of biogeographic patterns of microorganisms. In addition, the integrated study between this group and macroorganisms can boost discoveries in the field of ecology and, on the other hand, the application of already established mathematical models provides new insights into the macroecological structure of these microorganisms. This thesis was divided into two parts. In part I I bring a theoretical introduction about the development of the discipline in the last decade and a summary of the latest ecological and genomic findings regarding the taxonomic groups most commonly found in continental aquatic environments. In part II, I present the results of my research divided into three chapters. The first chapter of this part deals with the local processes that potentially impact bacteria's ability to remain consistently abundant over time, which may affect the range of spatial distribution in a metacommunity, as more abundant organisms are generally more abundant. those that are more likely to be present in all locations in a given region. In the second chapter, I analyze the most relevant dispersion patterns for bacteria in a spatial matrix of about 250 thousand km². Finally, in Chapter 3, I sought to understand the impact of scale-up on the deterministic and stochastic processes that structure a bacterioplankton metacommunity. This work contributes to the understanding of how microrganisms may be distributed over a landscape in inland waters and what factors potentially determine their distribution and abundance.

Keywords: Microbial ecology; spatiotemporal structuring; landscape; β-diversity; variation partitioning, dispersal dynamics, scale; deterministic and stochastic factors; shallow lake bacteria; amplicon sequencing

Lista de Figuras

Lista de Tabelas

Lista de Siglas e Abreviações

Ambiente

Chl *a* (Chlorophyll *a*) – Clorofila *a*;

DOC (Dissolved Organic Carbon) – Carbono Orgânico Dissolvido;

TOC (Total Organic Carbon) – Carbono Orgânico Total;

DIC (Dissolved Inorganic Carbon) – Carbono Inorgânico Dissolvido;

DIN (Dissolved inorganic nitrogen) – Nitrogênio Inorgânico Dissolvido;

DOM (Dissolved Organic Matter) – Matéria orgânica dissolvida;

    cDOM (coloured Dissolved Organic Matter) – a parcela colorida da Matéria orgânica dissolvida;

    fDOM (fluorescent Dissolved Organic Matter) – a parcela fluorescente da Matéria orgânica dissolvida;

    T-fDOM (Tryptophan-like fluorescent Dissolved Organic Matter) – A razão entre a Matéria orgânica dissolvida fluorescente e o sulfato de quinina a 455 nm excitação e 355 nm emissão;

    SR (Slope Ratio) – A razão entre os slopes recuperados nos intervalos 275-295nm e 350-400nm. Mensuração indireta do peso molecular;

    C:A peaks ratio – A razão entre os picos C (275 excitação, 304 emission) e A (260 excitation, 450 emissão). Indicativo da concentração de ácidos húmicos e fúlvicos;

    FI (Fluorescence Index) – A razão entre comprimentos de onda recuperados à 470 e 520 nm emissão, e 370 nm excitação. Indica a principal origem (água doce ou solo) da matéria orgânica;

HIX (Humification Index) – A razão entre a área da matriz de 435 to 480 nm e os picos somados à de 300 a 345 nm e de 435 a 480 nm, e 254 nm excitação. Valores altos indicam uma grande concentração de compostos húmicos;

FR (Freshness Index) – A razão entre o valor recuperado a 380 nm e o valor máximo recuperado na área à 420 a 435 emissão, e 310 nm excitação. Valores elevados indicam uma matéria orgânica produzida recentemente;

Esforços coletivos

BroaMO (Broa' Microbial Observatory) – Observatório Microbiano da Represa do Broa;

µSudAqua (Collaborative Network in Microbial Aquatic Ecology in Latin America) – Rede colaborativa em ecologia microbiana aquática da América Latina;

Análise molecular

DNA (Deoxyribonucleic Acid) – Ácido Desoxirribonucleico;

eDNA (environmental Desoxirribonucleic Acid) – DNA recuperado de amostras ambientais;

rRNA (Ribosomal Ribonucleic Acid) – Ácido Ribonucleico Ribossomal;

Amplicon – Segmento de DNA amplificado;

Conteúdo GC – estrutura gênica em que se repetem sequências Guanina-Citosina

ASV (Amplicon Sequence Variant) – variante de sequência amplicon. Uma unidade taxonômica recuperada de amostras de eDNA;

DGGE (Denaturing Gradient Gel Electrophoresis) - Eletroforese em Gel de Gradiente Desnaturante;

NGS (Next-Generation Sequencing) – Sequenciamento da Próxima Geração;

Bases de dados

NCBI (National Center for Biotechnology Information) – Centro Nacional para Informação Biotecnológica;

ENA (European Nucleotide Archive) – Arquivo Europeu de Nucleotídeos;

SILVA – Base de dados de sequências genéticas bacterianas;

SINA (SILVA Incremental Aligner) – Alinhador incremental de SILVA;

DADA2 (Divisive Amplicon Denoising Algorithm) – Algorítimo de redução de ruído Divisivo de Amplicon;

Estatística

βdiv (beta-diversity) – Beta Diversidade;

βMNTD (β-mean-nearest taxon distance) – Distância beta para o Táxon médio mais próximo. Índice de dissimilaridade filogenética;

βNTI (β-nearest taxon index) – A razão entre βMNTD da amostra e βMNTD do pool regional;

dbRDA (distance based Redudance Analysis) – Análise de redundância baseada em distâncias;

dbMEM (distance based Moran's Eigenvector Maps) – Mapas de autovetores de Moran baseados em distância;

   MEM (Moran's Eigenvector Maps) – Mapas de autovetores de Moran;

   AEM (Asymetric Eigenvector Maps) – Mapas de autovetores assimétricos;

IndVal (Indicator Value) – Valor de Espécie Indicadora;

MOStest (Mitchell-Olds & Shaw test of quadratic extremes) – Teste de Mitchell-Olds & Shaw de extremos quadráticos;

nMDS (non-metric Multidimensional scaling) – Escalamento multidimensional não-métrico;

permanova – Teste de permutações baseada em ANOVA para testar as diferenças entre amostras; difference test;

permidisp – Teste de permutações para testar a homogeneidade intra amostras;

RC$_{bray}$ (Raup-Crick dissimilarity index) – Índice de dissimilaridade de Raup-Crick [Neste caso, baseado em índice de Bray-Curtis];

SpADs (Spatial Abundance Distributions) – Distribuições espaciais de abundância;

Sumário

Apresentação

Esta tese foi construída (e sofrida) a seis mãos. Cada letra que compõe as próximas páginas desse documento foi transpirada por mim, minha esposa Gabrielle e nossa filha Sophia. Sophia veio ao mundo apenas duas semanas após meu doutorado ser iniciado e esses últimos cinco anos foram passados em atenção dividida entre coletas, modelos estatísticos, leituras e acompanhar o desenvolvimento de uma criança (com o pequeno detalhe de uma pandemia acontecendo no meio disso). Posso dizer, portanto, que esta tese não seria viável sem o enorme acolhimento que recebemos de meu orientador, o orientador de Gabrielle à época (Dr. Rhainer G. Ferreira), e de todo o corpo técnico-administrativo do Programa de Pós-Graduação em Ecologia e Recursos Naturais (PPG-ERN), pelos quais serei eternamente grato.

Nestes últimos anos de trabalho, eu busquei decompor e responder todas as dúvidas que permaneceram em minha mente após a conclusão de minha dissertação. Naquela oportunidade, eu pude ver que as bactérias encontradas em lagoas rasas amplamente distribuídas apresentam um padrão bimodal de frequências de ocupação, o que significa que alguns organismos dominam a região, estão amplamente distribuídos, enquanto outros estão isolados. Daí surgiram muitas outras questões: O que garante que uma bactéria esteja espacialmente tão bem distribuída enquanto as outras não o são? Apenas as distâncias entre sítios explicam esses padrões ou outros fatores geográficos podem impactar essa dinâmica? Será que existe uma escala onde os fatores geográficos se tornam claramente mais impactantes que a seleção local? O objetivo desta tese é tentar responder a estas perguntas, (enquanto outras foram deixadas para o futuro).

Esta tese foi basicamente dividida em duas partes:

Na 1° parte apresento uma introdução teórica dividida em dois capítulos. No capítulo I, eu busquei introduzir o leitor a uma perspectiva mais histórica sobre como a Ecologia de forma geral e a Ecologia Microbiana vêm se desenvolvendo nas últimas décadas para chegar às descobertas mais recentes e novas abordagens

metodológicas que permitiram entender melhor a dimensão da diversidade microbiana nos diversos tipos de ambientes do mundo. No capítulo II e sumarizei os últimos achados ecológicos e genômicos a respeito dos grupos taxonômicos mais comumente encontrados nos sítios em que dispus de amostras para realizar os estudos ecológicos que apresento aqui. No capítulo III eu apresento alguns achados passados que me guiaram nos estudos que conduzi durante esse mestrado e guiaram os objetivos apresentados nesta tese

A 2° parte trás os estudos realizados por mim em parceria com diversos pesquisadores do Brasil e do mundo, e que estão sendo preparados para serem submetidos logo a revistas científicas. O capítulo IV foi nomeado "Depicting spatiotemporal variables that drive dominance dynamics of lake bacteria". Nele buscamos levantar todos os processos locais que potencialmente impactam na capacidade de uma bactéria em se manter consistentemente abundante ou não ao longo do tempo (1 ano e meio), uma dinâmica relevante para os padrões espaciais em uma metacomunidade, já que os organismos mais abundantes muitas vezes são também os que possuem maior dominância regional. Fatores ambientais (pH, Nutrientes e Carbono dissolvido, Oxigênio dissolvido, entre outros), sazonalidade (caracterizadas entre períodos de seca e chuva) e interações com fitoplâncton (p. ex.: Dinophyceae; Cryptophyceae) e zooplâncton (p. ex.: *Brachionus sp.; Floscularia sp.*) foram testados. E então eu busquei associar estas variações com a capacidade de cada organismo nessa comunidade em se manter presente espacialmente e como o seu estilo de vida (revelado pela literatura) pode explicar os padrões encontrados. Este trabalho foi idealizado em parceria com Dra. Clara M. Arboleda-Baena, Dra. Mariana R. A. Costa, Dra. Karime A. Paina, MSc Gabrielle C. Pestana, Prof. Dr. Juan Pablo Niño-Garcia e Prof. Dr. Gilmar Perbiche-Neves.

O capítulo V foi chamado "Beyond environmental selection: Spatial structuring of tropical lake bacterioplankton metacommunity". Neste busquei observar quais os padrões de dispersão mais relevantes para bactérias em uma matriz espacial com cerca de 250mil km². Este trabalho foi produzido em parceria com o prof. Dr. Adriano Caliman e a Dra. Thaís Garcia da Silva.

Finalmente, o capítulo IV foi intitulado "Scale matters? The effect of spatial scale on ecological processes that drive the aquatic bacterial communities", onde eu comparei a importância relativa de fatores determinísticos e estocásticos guiando metacomunidades microbianas em distintas escalas. Para que este trabalho se tornasse possível, foram concatenados dados de 135 lagoas rasas de Argentina, Brasil e Canadá em uma única base de dados. Este trabalho, portanto, não seria possível sem a Profa. Dra. Maria Llames e o Prof. Dr. Juan Pablo Niño-Garcia que me disponibilizaram dados microbianos coletados por eles nestes países. Além deles, participaram da concepção deste capítulo Dra. Paula Huber, Dr. Pedro Junger, Dr. Sebastian Metz, Dr. Emiliano Pereira, Prof. Dr. Victor Saito e Prof. Dra. Inessa L. Bagatini.

As próximas páginas trazem o resultado de meus esforços nestes últimos anos. Obrigado a você que pretende virar esta página e continuar com a leitura, desejo-lhe uma agradável leitura.

Erick Mateus-Barros

Capítulo I -
Ecologia & Escala

"Theoretical ecology, and theoretical science more generally, relates processes that occur on different scales of space, time, and organizational complexity. Understanding patterns in terms of the processes that produce them is the essence of science, and is the key to the development of principles for management." (Levin 1992 in *The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture*)

A ecologia encontra a biogeografia

Tradicionalmente, ecologia e biogeografia se desenvolveram como disciplinas distintas. Enquanto a primeira focava em entender os padrões locais da composição de espécies, a segunda tentava entender como processos de grande escala espaço-temporal contribuíam para a distribuição atual de espécies no planeta (Jenkins e Ricklefs, 2011). O avanço tecnológico e mudanças recentes de perspectiva vêm mudando a visão de pesquisadores em ambas as áreas de atuação e aproximaram as duas disciplinas em um ponto intermediário desses dois extremos. Para a ecologia em especial, a aplicação de métodos moleculares e o acúmulo de dados e conhecimento causados pelo avanço tecnológico das últimas décadas tornou cada vez mais evidente que, para acessar algumas questões fundamentais, faz-se necessário adentrar outras escalas visando entender o que causa determinados padrões (Levin, 1992).

Esta estrutura, que tem sua semente em uma insatisfação com a abordagem vigente (Levin, 1992) fertilizada pela necessidade de entender impactos de mudanças ambientais de importância social (Chave, 2013) e regada por um avanço tecnológico vertiginoso, por si só possui ares de revolução científica (Kuhn, 2013) e vem revitalizando o estudo de comunidades biológicas trazendo novos e intrigantes paradigmas à baila.

Avanços teóricos, tais como a teoria unificada neutra (Hubbell, 2001) e as perspectivas de metapopulações (Levins, 1969) e de metacomunidades (Leibold *et al.*, 2004) consideram mudanças na composição de espécies em uma escala maior, e que vários sítios devem estar conectados e compartilhando indivíduos. Estas novas teorias, como era de se esperar, enfrentaram resistência por parte da comunidade científica em sua época, mas em especial a teoria neutra passou por extenso escrutínio até ficar claro que esta guardava em si méritos que não podiam ser deixados de lado (Leigh Jr, 2007), e então sínteses teóricas considerando o efeito conjunto de fatores locais e regionais, determinísticos e estocásticos passaram a ser consideradas (Vellend, 2010).

A estrutura da revolução ecológica

Em seu trabalho publicado em 2013, Jérome Chave lista quatro fatores que segundo ele contribuíram para mudar os paradigmas dominantes na ecologia ao longo das últimas décadas. O autor argumenta que estes fatores foram preponderantes para que essa área da ciência deixasse de focar exclusivamente nos padrões locais de composição de espécies, em direção a uma perspectiva espaço-temporal mais ampla (Chave, 2013).

Primeiro, a criação e desenvolvimento de computadores iniciaram uma chamada "era numérica" que permitiu a aplicação de modelos matemáticos antes impraticáveis por meio de caneta e papel (Chave, 2013). Exemplo interessante dessa perspectiva é o uso de redes neurais (Lecun *et al.*, 2015) para entender padrões complexos e realizar predições a respeito de como determinada comunidade ou população deverá variar no espaço e no tempo. Outra mudança notável é a criação de bancos de dados que permitem reunir em um só lugar informação que preencheria facilmente uma vida de trabalho e permitindo visualizar padrões observáveis em escalas globais (Thompson *et al.*, 2017).

Em segundo, avanços moleculares, em especial a descoberta da reação em cadeia da polimerase (polimerase chain reaction - PCR), permitiram entender relações evolutivas (Newton *et al.*, 2011), identificar espécies (Callahan *et al.*, 2016), inferir biodiversidade (Pedrós-Alió, 2012) e padrões de distribuição (Mateus-Barros *et al.*, 2021) e, a partir de técnicas mais recentes (*i.e.* High-throughput sequencing) prospectar genes, proteínas e rotas metabólicas (Chiriac, Haber*, et al.*, 2022).

Terceiro, o sensoriamento ambiental, permitiu agregar mapas e análises espaciais à perspectiva ecológica (Chave, 2013).

Quarto, a disseminação de informação e conhecimento se tornou mais eficiente e globalizada. De fato, artigos nascem em profusão, detalhes são transmitidos com facilidade, novos métodos e modelos podem ser aplicados por qualquer cientista e, quando alguma dificuldade aparece, dicas, tutoriais e fóruns estão a alguns cliques de distância.

Um exemplo recente de como todos esses avanços vem moldando estudos científicos foi o fenômeno que seguiu a pandemia do COVID-19. Desde o reconhecimento pelas autoridades globais de que o vírus estava de fato contaminando pessoas, e se espalhando pelo mundo com espantosa velocidade (também facilitada pelos avanços tecnológicos), boletins diários permitiram acompanhar com precisão de apenas algumas horas o avanço inicial do vírus e sua chegada em cada país do mundo (p. ex. Andrews *et al.*, 2020; Edrada *et al.*, 2020) em uma espécie de estudo coletivo a respeito da biogeografia e padrões de dispersão do vírus. Basta uma busca rápida para encontrarmos dados sobre o número de pessoas infectadas e falecidas em cidades, países ou em quase todo o planeta, informações estas atualizadas com apenas alguns dias de atraso (COVID-19 Dashboard); passamos a nos acostumar a ler sobre o surgimento de novas cepas assim que novas ondas de infecção surgem; também nos familiarizamos com termos técnicos, conhecidos em sua maioria apenas por profissionais e cientistas da área. Vacinas foram produzidas e testadas em tempo recorde e utilizando novas técnicas que se aproveitam de informações moleculares para criar doses quase personalizadas contra cada cepa (Ciotti *et al.*, 2020; Huang *et al.*, 2020); e finalmente, novas informações a respeito das formas que o vírus usa para infectar abundam (Banerjee *et al.*, 2020). Cientistas da área criaram um esforço coletivo mundial para divulgar informações, facilitando que tecnologias fossem produzidas e a população se mantivesse informada. E tudo isso só pôde ocorrer por causa da internet e da capacidade de processamento que computadores e smartphones alcançaram nos últimos anos.

Algo semelhante ocorreu mais recentemente dentro da própria perspectiva da Ecologia. Estudos a respeito de processos ocorrendo com organismos que podem ser vistos a olho nu se desenvolveram paralelamente aos estudos referentes a microrganismos, um padrão que vem também se revertendo nos últimos anos.

Importância de estudar bactérias

O reino Prokaryota compreende muitos dos organismos mais relevantes para a manutenção da vida como conhecemos em todos os cantos do planeta Terra.

Estes organismos simples, de vida unicelular (Chiriac, Haber*, et al.*, 2022) e em muitos casos agregada (Pascual-García *et al.*, 2014), podem ser detritívoros (Salcher *et al.*, 2015) ou autotróficos (Stanier e Cohen-Bazire, 1977), muitas vezes ambos (Ghai *et al.*, 2014) e estão, de alguma maneira, intimamente relacionados com todas as outras formas de vida terrestres.

Sempre muito abundantes e presentes em qualquer tipo de ambiente, as bactérias participam de processos considerados chave na manutenção de ecossistemas (Newton *et al.*, 2011; Chiriac, Haber*, et al.*, 2022). Estes organismos participam da ciclagem de nutrientes e matéria orgânica de diversas origens e complexidades moleculares (Kawasaki e Benner, 2006; Sarmento e Gasol, 2012) e em ciclos biogeoquímicos (Myklestad, 2000), tornando elementos como Carbono e Nitrogênio disponíveis a outros organismos (Sarmento, 2012), liberados na atmosfera por meio de respiração (Cole *et al.*, 2007), ou mineralizados (Maier, 2015).

Durante muitas décadas, a ideia vigente era que estes organismos fascinantes não poderiam nunca ser parados por barreiras naturais, ao invés disso seriam as diferentes condições ambientais locais que influenciariam a composição de espécies bacterianas (Baas Becking, 1934). A frase "Everything is everywhere, but the environment selects", uma das mais célebres já proferidas dentro das ciências microbiológicas, traduzida do original alemão, resume essa linha de raciocínio e nasce do fato que apenas um reduzido número de espécies era recuperado dos diferentes ambientes analisados à época. Somente os avanços tecnológicos alcançados nas últimas décadas permitiram identificar a existência de uma biodiversidade bacteriana maior do que se pensava e promoveu o desenvolvimento de mais uma área biológica do conhecimento: a Ecologia microbiana.

O método mais comumente utilizado para identificar microrganismos era a criação por meio de culturas microbianas em placas de Petri. Este método revolucionário promove o crescimento de organismos capazes de utilizar certos nutrientes disponibilizados para posterior identificação em microscópios (Kim *et al.*, 2019), e permanece relevante para a identificação de espécies (p. ex. Hahn, 2003;

Kim *et al.*, 2019). Porém, os avanços da biologia molecular com a descoberta da PCR (Mullis *et al.*, 1992) trouxeram em seu encalço diversas técnicas moleculares independentes de cultura capazes de recuperar informações biológicas diretamente de amostras ambientais (Taberlet *et al.*, 2012), o que permitiu identificar uma biodiversidade bacteriana muito maior que se conhecia anteriormente. Muitas bactérias incapazes de crescer nos meios de cultura tradicionais, e, portanto completamente desconhecidos anteriormente, puderam ser descobertos. Alguns organismos antes negligenciados, hoje são reconhecidos como muito abundantes e relevantes para processos e padrões ecológicos (Chiriac, Haber*, et al.*, 2022). Um ótimo exemplo disso está no clado hgcl (também conhecido como linhagem acl). Pertencente ao filo das Actinobactérias, este é um grupo de organismos capaz de alcançar até 50% da abundância relativa em diversos ambientes lacustres ao redor do mundo (Salcher *et al.*, 2010; Camara Dos Reis *et al.*, 2019; Mateus-Barros *et al.*, 2021), mas era subestimado até pouco tempo atrás, e passou a ser identificado em nível de espécie apenas a alguns anos quando novas técnicas moleculares permitiram montar o genoma completo deste grupo e à partir disso foi possível reconhecer algumas proteínas das quais eles necessitam para subsistir apesar de não produzirem, e então a adição dessas moléculas permitiu sua manutenção em laboratório (Kang, I. *et al.*, 2017; Kim *et al.*, 2019).

Essas novas tecnologias permitiram também entender que de fato existe uma microbiota rara em qualquer ambiente (Pedrós-Alió, 2006; 2012), o que implica na existência de um grupo de organismos incapaz de suplantar algumas barreiras geográficas e que o processo que torna cada ambiente único em composição de espécies não é causado apenas por fatores ambientais locais (Green e Bohannan, 2006; Martiny *et al.*, 2006); assim, uma mudança de perspectiva se fez necessária, e o estudo da estrutura espacial microbiana ganhou relevância (Acinas *et al.*, 2004).

A contribuição da ecologia para a microbiologia

Talvez a mais importante consequência do advento das técnicas moleculares foi permitir estudar a evolução bacteriana de um ponto de vista de adaptações ecológicas em contraponto a uma visão mais fisiológica de suas diferenças (Rappé e

Giovannoni, 2003). Nesse contexto, estudos microbianos voltaram-se para conceitos e teorias ecológicas, o que permitiu uma troca de conhecimento entre as duas áreas (Horner-Devine e Bohannan, 2006) e o nascimento de uma visão mais holística (Soininen *et al.*, 2018) com a oportunidade de aproveitar teorias e modelos com embasamento mais sólido a respeito de processos naturais (Horner-Devine *et al.*, 2003).

Diversos trabalhos conduzidos na esteira dessa mudança de visão puderam demonstrar que as bactérias são de fato guiados por diversos fatores ambientais como salinidade, nutrientes e temperatura (Horner-Devine *et al.*, 2003), mas também foi possível notar outras dinâmicas importantes como a existência de um decaimento da similaridade com o aumento de distâncias (Green e Bohannan, 2006), a relação entre o aumento da diversidade e o aumento da área estudada (i.e. relação taxa-área Horner-Devine *et al.*, 2004; Green e Bohannan, 2006) e que as maiores taxas de biodiversidade são observadas em um nível de heterogeneidade ambiental intermediário (i.e. Hipótese do distúrbio intermediário Horner-Devine *et al.*, 2003). Estes estudos revelam a existência de padrões biogeográficos em bactérias a aproximam estes dados a uma visão mais aproximada ao que se consensua serem os fatores que guiam animais e plantas (Horner-Devine *et al.*, 2003).

Ao mesmo tempo, estudar as bactérias espacialmente mais persistentes e abundantes ganha força atualmente, no sentido que estes constituem organismos chave para a manutenção de processos ecológicos em pequena para média escala, como em redes de interação (Sarmento, 2012), e também são importantes para processos em escala global, como os já citados ciclos biogeoquímicos (Chiriac, Haber*, et al.*, 2022), além daqueles organismos já exaustivamente conhecidos e de importância para a saúde humana (Gutierrez, 2019), mas que não parecem ser de fato incrivelmente abundantes em ambientes naturais saudáveis.

A contribuição da microbiologia para a ecologia

O desenvolvimento paralelo entre Ecologia e Microbiologia (Rappé e Giovannoni, 2003) também teve consequências práticas para os estudos ecológicos

clássicos. A mais importante delas seja talvez a falta de oportunidade em realizar estudos integrando informações entre micro- e macrorganismos (Horner-Devine e Bohannan, 2006), questão hoje já superada. Estudos holísticos estão presentes e muitas questões interessantes já foram abordadas. Atualmente sabemos, por exemplo, que a extensão da diversidade beta apresenta certa consistência através de todos os tipos de seres vivos do planeta e está relacionado com o tamanho do organismo, características ambientais, geografia e a escala de estudo (Soininen *et al.*, 2018).

Estudos microbianos também podem se provar relevantes ao fornecerem um número de indivíduos gigantesco em apenas uma gota de água, o que tem o potencial de possibilitar testes a respeito de qualquer teoria ecológica ou modelo matemático até seus limites (Barberán *et al.*, 2014; Mateus-Barros *et al.*, 2021). Outra característica interessante é a facilidade em se obter dados microbianos para qualquer tipo de estudo, já que a submissão destes em bancos de dados dedicados, abertos e gratuitos é largamente disseminada (Quast *et al.*, 2013; Metz *et al.*, 2022).

O uso de técnicas baseadas no sequenciamento do gene que codifica a subunidade 16S do RNA ribossomal microbiano tornou-se relevante e foi parte imprescindível no desenvolvimento da Ecologia Microbiana durante as últimas décadas (Rappé e Giovannoni, 2003), permitindo entender a estrutura de comunidade desse grupo (Newton *et al.*, 2011) e inferir padrões ecológicos de grande escala (Soininen *et al.*, 2018). Estas técnicas moleculares independentes de cultura, adotadas e desenvolvidas no rastro das necessidades dos ecólogos microbianos para serem capazes visualizar a biodiversidade microbiana (Rappé e Giovannoni, 2003) vem sendo cada vez mais adaptadas também a estudos com macroorganismos (Yoccoz, 2012; Barnes e Turner, 2015).

As abordagens moleculares

As novas técnicas de identificação de espécies passam pelo chamado DNA ambiental (eDNA) que consiste na coleta de amostras de solo ou água e utilização desse material para extrair e sequenciar o DNA contido nessas amostras. Assim,

tem-se uma prospecção a respeito de todos os organismos presentes no local onde for realizada a análise (Barnes e Turner, 2015). Essa técnica, aliada ao método molecular adequado permite identificar qualquer tipo de organismo e padrões muito interessantes de biodiversidade podem ser observados (Martínez *et al.*, 2015; Komura *et al.*, 2018). As duas abordagens NGS ("*New Generation Sequencing*", Sequenciamento de Nova Geração) mais utilizadas hoje em dia são chamadas Amplicon e metagenomas e utilizam esse tipo de coleta de amostras como ponto de partida.

O método de sequenciamento de amplicons (Fig. 1A) consiste no sequenciamento de algum gene de interesse. A escolha do gene a ser utilizado é talvez a etapa mais crítica, já que este deve ser evolutivamente preservado o suficiente para que seja encontrado em todos os membros do grupo a ser analisado, mas também deve apresentar variações suficientes para que seja possível identifica-las e relacioná-las com diferentes espécies ou gêneros (Quast *et al.*, 2013; Yilmaz *et al.*, 2014). Para bactérias, a tradição é amplificar o gene que transcreve o RNA da porção 16S do Ribossomo (Liu *et al.*, 2021). Este método já clássico passou por diversas melhorias ao longo dos anos e costuma ser utilizado em situações que se deseja entender de forma mais rápida qual é a composição de organismos em diferentes comunidades (Stegen *et al.*, 2013; Dini-Andreote *et al.*, 2015; Llames *et al.*, 2017; Logares *et al.*, 2018; De Melo, Bertilsson*, et al.*, 2019; Huber *et al.*, 2020; Logares *et al.*, 2020; Mateus-Barros *et al.*, 2021). Por ser um método mais barato, ainda é bastante disseminado em países em desenvolvimento, mas também vem crescendo seu uso para a identificação de organismos maiores através da coleta de amostras ambientais (Taberlet *et al.*, 2012; Yoccoz, 2012; Barnes e Turner, 2015; Thomsen e Willerslev, 2015).

Por outro lado, o método de metagenomas (Fig. 1B) consiste na obtenção de genomas completos a partir do sequenciamento de fragmentos de DNA (Liu *et al.*, 2021). Neste caso, as amostras não são filtradas, ao invés disso, o DNA ambiental é fragmentado e sequenciado diretamente. Os fragmentos de DNA obtidos passam então por sucessivas etapas de identificação e montagem para que os genomas sejam recuperados (Liu *et al.*, 2021). Em países mais desenvolvidos, a tendência dos últimos anos é de mudança em direção ao crescente uso de metagenomas.

Essa técnica, apesar de muito mais cara que a anterior, permite visualizar não somente a composição de organismos em determinado local, como também abrir espaço para o estudo de rotas metabólicas, e um acesso mais refinado aos papéis ecológicos e necessidades de cada grupo.
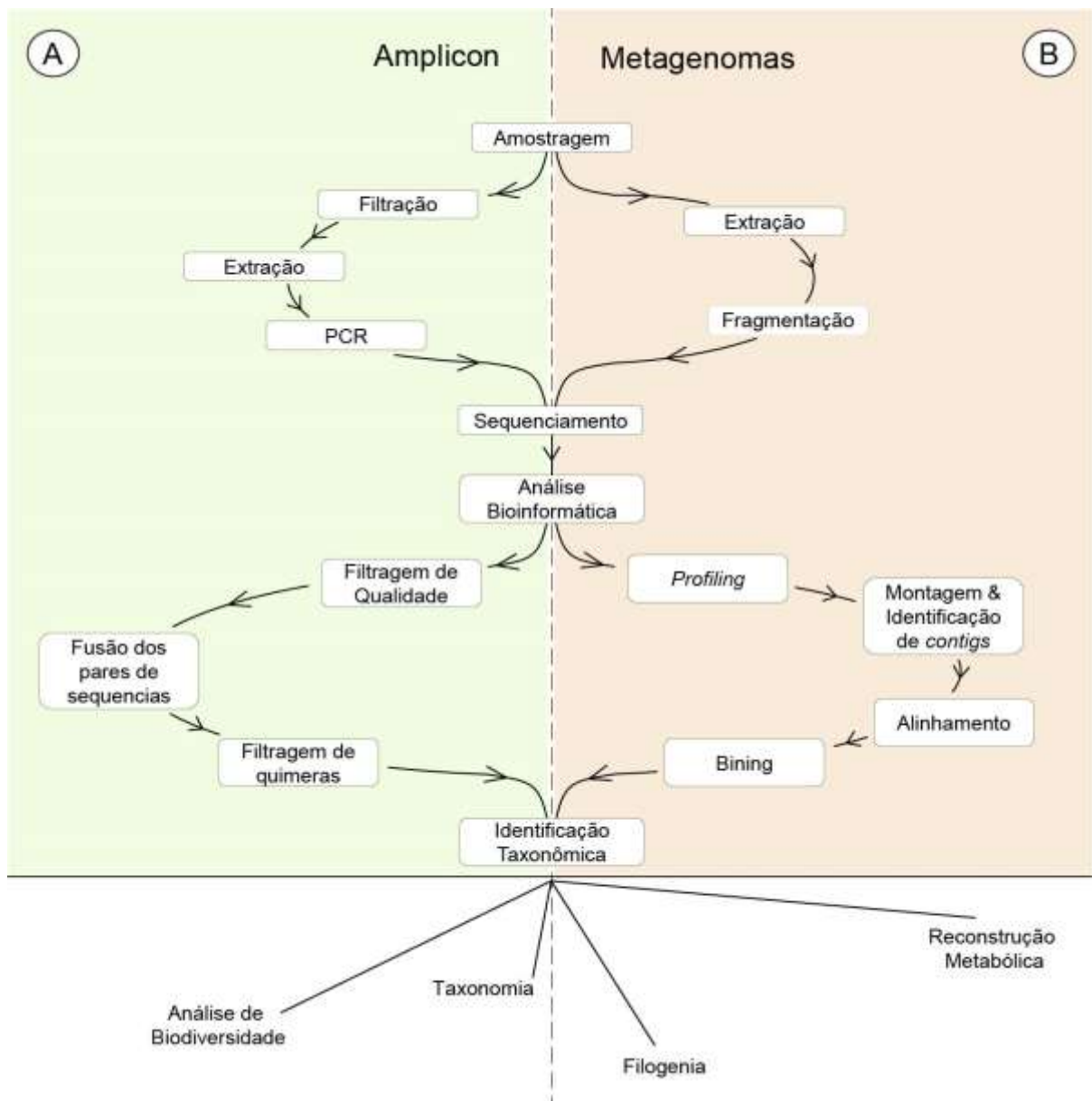
Figura 1 – Caminho (indicado pelas setas →) para obtenção de dados desde a amostragem até a identificação taxonômica para as duas abordagens mais relevantes da atualidade. O amplicon (A) consiste no sequenciamento de um gene de interesse para identificação de unidades taxonômicas. Este tipo de abordagem é melhor aproveitado para estudos de análise de biodiversidade e é uma ferramenta complementar na identificação taxonômica de uma espécies O metagenoma (B) consiste na montagem de genomas completos a partir de diversos fragmentos de DNA de distintos tamanhos.

Essa última abordagem é mais adequada para reconstruções filogenéticas e é também muito útil para reconstruções filogenéticas

Novas perspectivas, macro- e micro- estudados em conjunto

O entendimento que a estrutura de comunidades microbianas é mais similar ao observado para outros organismos do que se supunha (Horner-Devine *et al.*, 2003), aliado ao uso destes métodos moleculares e a construção de modelos matemáticos mais robustos (Horner-Devine *et al.*, 2003; Barberán *et al.*, 2014) vem ajudando a criar um *framework* (estrutura conceitual) repleto de abordagens excitantes, novas e que podem fazer com que o estudo destes organismos seja colocado na vanguarda, entre os mais relevantes estudos da ecologia. Esses avanços provavelmente serão cada vez mais aliados a outros conceitos teóricos ecológicos nos próximos anos e contribuirão para um melhor entendimento a respeito da estrutura de comunidades e biogeografia de forma geral. A seguir, apresento algumas perspectivas interessantes recentes que ilustram essa expectativa.

Em 2012, Stegen e colaboradores criaram um modelo matemático que visa prever quais fatores entre ambientais determinísticos, espaciais estocásticos e *Drift* (a variação no número de unidades taxonômicas encontrada em um sítio causada por flutuações não mensuráveis na abundância de organismos com populações pequenas) poderiam ser os principais responsáveis pela dissimilaridade observada para os organismos encontrados entre um par de sítios. O método se baseia no uso de dois modelos em sequência para calcular o efeito dessas forças ao comparar a dissimilaridade observada para a comunidade de interesse. Primeiro, aplica-se o βMNTD (β-mean-nearest taxon distance) para comparar as distâncias filogenéticas apresentadas em cada par de sítios. Para os valores significativamente distintos do modelo considera-se que as diferenças são causadas pelo efeito de fatores locais determinísticos. Os pares com valores não significativos são submetidos ao RCbray (análise de dissimilaridade de Raup-Crick baseada em distâncias de Bray-Curtis), que compara dessa vez a dissimilaridade de composição e abundância de espécies

com um modelo nulo para todas as comunidades. Valores significativos são considerados como causados por fatores espaciais estocásticos enquanto valores não significativos são considerados como causados por *Drift* (Stegen *et al.*, 2012; Stegen *et al.*, 2013). O modelo utilizado para os cálculos do βMNTD (Webb *et al.*, 2002) já recebeu críticas pertinentes (Hillerislambers *et al.*, 2012), porém alguns cálculos podem ser realizados no intuito de mensurar se de fato as distancias filogenéticas observadas estão relacionadas com a heterogeneidade ambiental do *pool* (conjunto de unidades taxonômicas encontradas na região) de sítios utilizado (i.e. teste de Mantel Filogenético ou Correlograma de Mantel Stegen *et al.*, 2012). Este método vem sendo utilizado no intuito de melhor compreender diferenças de estrutura de comunidades microbianas em distintos tipos de ambientes, tais como solo (Zhang *et al.,* 2020), sedimento (Stegen *et al.,* 2012; Stegen *et al.,* 2013), ambientes de água doce (Llames *et al.,* 2017; Huber *et al.,* 2020) e marinho (Logares *et al.,* 2020). Por meio desta abordagem é possível também distinguir comunidades microbianas da flora intestinal de habitantes de diferentes países e relacionar esses dados com o tipo de alimento majoritariamente consumido em cada nação (Martínez, Stegen et al. 2015).

O uso de metagenomas vem se tornando cada vez mais disseminado. Através de suas rotas metabólicas, é possível saber quais tipos de moléculas são preferencialmente utilizadas por cada bactéria e, portanto, se esses organismos são relevantes para algum tipo de ciclo biogeoquímico, como do enxofre (p. ex. Alfaproteobacteria, SAR11 Giovannoni, 2017), Carbono (p. ex. Gammaproteobactéria, Metylophilaceae Salcher *et al.*, 2015) e Nitrogênio (p. ex. Actinobacteriota, hgcI Ghai *et al.*, 2014). É possível também entender e identificar preferências por determinados estilos de vida (Chiriac, Haber*, et al.*, 2022). Por exemplo*,* os dois grupos mais abundantes na superfície de lagos e oceanos pertencem a filos completamente distintos, mas guardam características em comum. Os clados SAR11 (uma alfaproteobactéria) e hgcI (uma Actinobactéria) apresentam ambos um reduzido conteúdo genômico e tamanho celular, uma baixa porcentagem de conteúdo GC (estrutura gênica em que se repetem sequências Guanina-Citosina) em seu genoma e uma elevada taxa de codificação gênica. Todas essas características indicam um estilo de vida chamado "streamlined", ou seja, são

especializados em produzir apenas as proteínas essenciais à sobrevivência, relacionadas à homeostase celular e rodopsinas para uma autotrofia anaeróbica facultativa (Chiriac, Haber*, et al.*, 2022). Graças a isso, esses organismos são extremamente versáteis, o que explica sua prevalência e ubiquidade, mas também são auxotróficos (Logares *et al.*, 2010; Kim *et al.*, 2019), o que significa não serem capazes de produzir todas as proteínas que utilizam em suas rotas metabólicas e dependerem da produção realizada e exsudada por outros organismos presentes na comunidade, apesar de não parecer apresentarem qualquer relação estreita com outros grupos microbianos (Chiriac, Haber*, et al.*, 2022). Através de metagenomas também é possível identificar as proteínas necessárias ao isolamento e manutenção em laboratório de determinadas cepas bacterianas (Kim *et al.*, 2017) e também propor a existência de novas linhagens filogenéticas, por meio de uma análise mais refinada das diferenças genéticas encontradas em cada grupo (Brown *et al.*, 2015).

Outro tipo de pesquisa bastante interessante que vem ganhando espaço, apesar de insipiente, é o uso de dados moleculares para inferir a filogeografia destes organismos (Urban *et al.*, 2008). Por meio desta abordagem é possível mapear a distribuição e prováveis caminhos de dispersão de filotipos relevantes. Por meio dessa abordagem, é possível mapear espécies invasoras produtoras de toxinas do gênero *Cerratium*, por exemplo (Accattatis *et al.*, 2020); ou ainda, entender como se deu a irradiação de espécies aparentadas à partir de um ancestral marinho, como em Bathycoccaceae, Chlorophyta (Fernández *et al.*, 2017) e Methylophilaceae, Proteobacteria (Salcher *et al.*, 2019).

Capítulo II -
Bactérias-chave em ambientes lacustres:
uma breve descrição

"With the advent of cloning and sequencing ribosomal RNA genes directly from the environment to access and explore the phylogenetic diversity of natural microbial communities, it has become clear that we have barely scratched the surface at obtaining representative cultures that span the phylogenetic breadth of most of the major phyla of Bacteria." (Rappé 2003 in The uncultured microbial majority)

Durante as pesquisas que geraram os capítulos que serão apresentados nesta tese, foi possível perceber que alguns grupos bacterianos apresentam grande abundância espaço-temporalmente. Representantes dos filos Actinobacteriota, Bacteroidota, Cyanobacteria, Planctomycecota, Proteobacteria, Verrucromibiota e Patescibacteria apresentaram relevância espacial (Mateus-Barros *et al.*, 2021) e/ou temporal (Parte II desta tese, Fig. 2). Esses grupos são de fato relevantes para a composição microbiana em ambientes lacustres em todo o planeta (Newton *et al.*, 2011), tendo sua relevância reconhecida precocemente durante o desenvolvimento de técnicas moleculares baseadas em coletas de amostras ambientais (Rappé e Giovannoni, 2003), e estão entre os grupos mais estudados (Chiriac, Haber*, et al.*, 2022).

Esses grupos são sempre muito abundantes e participam de processos considerados chave na manutenção de ecossistemas (Newton *et al.*, 2011; Chiriac, Haber*, et al.*, 2022), tendo importante papel em processos de ciclagem de matéria orgânica (Kawasaki e Benner, 2006; Sarmento e Gasol, 2012) e em ciclos biogeoquímicos (Myklestad, 2000) de elementos necessários à vida como nós conhecemos; tais como Carbono, Nitrogênio, Fósforo e Enxofre (Chiriac, Haber*, et al.*, 2022). Estes organismos também participam de uma intrincada rede de interações em que participam também as algas unicelulares (picofitoplâncton). Conhecida como microbial loop (Azam *et al.*, 1983), essa rede existe através de uma dinâmica de degradação e incorporação de compostos liberados no ambiente por meio de exsudação, excreção ou lise celular (Kawasaki e Benner, 2006). Esse processo vai tornando a matéria orgânica cada vez mais lábil até que eventualmente acaba sendo incorporada à rede trófica clássica por meio de predação/pastoreio realizada por protistas heterótrofos ou zooplâncton filtradores (Sarmento, 2012), liberada na atmosfera por meio de respiração (Cole *et al.*, 2007), ou mineralizada e retida no sedimento (Maier, 2015).

Aqui, apresento uma breve descrição dos grupos bacterianos encontrados nos mais diversos ambientes. Todos os grupos descritos abaixo apresentaram grande abundância e persistência espaço-temporal e foram considerados organismos-chave para os padrões e processos que serão descritos nos capítulos subsequentes. É importante ressaltar que aqui, não tenho a intensão de criar uma

revisão detalhada a respeito destes grupos, mas levantar algumas informações a respeito da biologia destes organismos que podem ser relevantes para o entendimento dos padrões ecológicos descritos mais adiante. Para uma revisão detalhada a respeito destes e outros organismos relevantes, sugiro a leitura do artigo recentemente publicado por Chiriac e colaboradores (2022) que traz uma perspectiva metagenômica a respeito do estilo de vida que cada grupo pode apresentar. Outra revisão relevante foi apresentada por Newton e colaboradores (2011) e, apesar de já estar publicado a mais de uma década e bastante defasado a respeito das informações moleculares, ainda apresenta muita informação relevante a respeito da ecologia em cada grupo.

Para auxiliar a visualização, também construí uma árvore filogenética de referência (Fig. 2) contendo representantes de todos os grupos de destaque. Para as referências, selecionei algumas sequências (n = 16) que foram obtidas através do sequenciamento do gene 16S que codifica a porção menor do RNA ribossomal bacteriano (16S rRNA), e mantidas em bancos de dados locais construídos para responder as questões apresentadas nos próximos capítulos. Esses genes já haviam sido identificados como sendo de representantes de cada um dos distintos organismos a ser descritos a seguir. Com estas sequências de consulta em mãos, utilizei os serviços fornecidos no website "*arb-silva*" (https://www.arb-silva.de/aligner/) para construir uma árvore filogenética curada e simplificada. O site utiliza o método SINA (Pruesse *et al.*, 2012) para alinhar e classificar as sequências fornecidas ao comparar com 10 sequências de máxima similaridade possível para cada um dos genes, contidas na base de dados SILVA (Quast *et al.*, 2013; Yilmaz *et al.*, 2014). Para a classificação utilizei um índice de similaridade mínima de 90% entre as sequencias de consulta e as fornecidas pelo banco de dados. Esse valor foi considerado porque este é o mínimo para que as Patescibacterias pudessem ser classificadas. Os outros grupos não divergiram do resultado de uma consulta utilizando similaridade de 99%. A árvore filogenética foi montada utilizando o software FastTree (Price *et al.*, 2010) com o método denovo incluindo as sequências "vizinhas" (maior similaridade possível com os dados fornecidos), o modelo "GTR" e verossimilhança do tipo "Gamma". A árvore foi então curada utilizando o software Wasabi (Veidenberg *et al.*, 2015) fornecido no mesmo site e exportada para ser

editada no software TreeGraph (Stöver e Müller, 2010) onde as sequencias "vizinhas" que geravam ruído na imagem puderam ser removidas para facilitar a visualização.

Juntamente com a descrição dos clados mais relevantes, também adicionei caixas de texto que apresentam um pequeno resumo contendo algumas informações relevantes a respeito dos organismos citados. As informações contidas nessas caixas-resumo foram anteriormente revisadas por Chiriac, Haber, *et al.* (2022) e Newton *et al.* (2011). A classificação taxonômica foi obtida através da base de dados SILVA (Quast *et al.*, 2013; Yilmaz *et al.*, 2014).
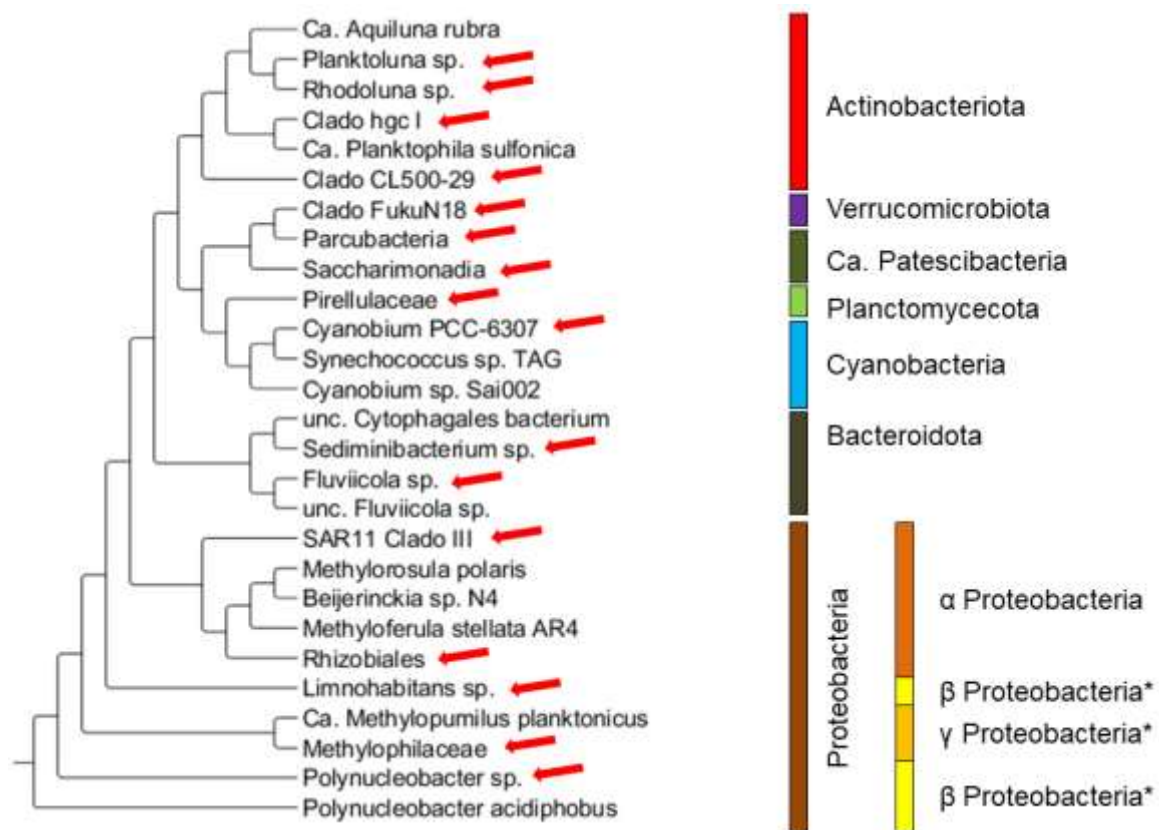
Figura 2 – Árvore filogenética indicando o grau de parentesco entre os grupos de interesse abordados neste capítulo, representados em sua maioria por espécies e gêneros. As setas vermelhas indicam nós obtidos através de sequências fornecidas, os outros nós se referem a algumas das sequências presentes no banco de dados SILVA e utilizadas como sequências aparentadas para tornar a construção da árvore mais robusta. As barras à direita indicam os filos aos quais esses genes estão afiliados e, no caso de proteobactérias, as classes. As classes Beta- e Gammaproteobacteria [indicados com * no cladograma] estão intimamente ligadas; Betaproteobacteria não constitui um clado monofilético na exclusão de Gammaproteobacteria

Actinobacteriota

Anteriormente reconhecidas como primariamente encontradas no solo, as bactérias que pertencem ao filo *Actinobacteriota*, são os organismos mais abundantes e ubíquos dentre os encontrados em ambientes aquáticos continentais (Newton *et al.*, 2011). Estes estão presentes em rios e lagos, ambientes salobros ou marinhos (Ghai *et al.*, 2014; Mizuno *et al.*, 2015), glaciais (Kang *et al.*, 2012) e de altitude (Warnecke *et al.*, 2005), mas são reinam mesmo na superfície de lagos, chegando a representar mais da metade da abundância relativa bacteriana nestes ambientes (Camara Dos Reis *et al.*, 2019; Mateus-Barros *et al.*, 2021).

Muito bem adaptadas à vida na coluna d'água estas bactérias são ubíquas e muito abundantes em lagos e reservatórios (Newton *et al.*, 2011). Estes organismos unicelulares podem apresentar genes relacionados à resistência contra o stress causado pela luz UV (Warnecke *et al.*, 2005) e produzir rodopsinas, um tipo de proteína capaz de agir na captação de energia luminosa para a produção de energia para a célula em momentos de pouca disponibilidade de compostos nutritivos (Sharma *et al.*, 2008). Os grupos mais abundantes também apresentam um reduzido tamanho celular, o que também demonstra uma adaptação para escapar do pastoreio realizado por organismos zooplanctônicos filtradores (Pernthaler *et al.*, 2001) e da predação por nanoflagelados (Chiriac, Haber*, et al.*, 2022). O tamanho dos genomas neste filo varia, porém, as linhagens mais dominantes possuem sempre um genoma formalmente nomeado como "streamlined". Organismos com este genoma simplificado são caracterizados por um reduzido tamanho molecular, acompanhado de um reduzido tamanho celular, baixa porcentagem de sequências GC e alta taxa de codificação de genes (Kim *et al.*, 2019). Esta característica simplificada lhes permite economizar energia na codificação de genes desnecessários, porém torna-os dependentes da disponibilidade de compostos exsudados por outros organismos, uma condição chamada auxotrofia (Chiriac, Haber*, et al.*, 2022).

Actinobactérias ainda possuem diversos genes relacionados à degradação e fixação de compostos ricos em Carbono (Chiriac, Haber*, et al.*, 2022) e Nitrogênio

(Ghai *et al.*, 2014; Kang, Ilnam *et al.*, 2017) inclusive os de alta complexidade originados de plantas vasculares terrestres (Ghai *et al.*, 2014), característica que evidencia sua relevância para ciclo biogeoquímico destas moléculas.

Apesar de muito abundantes e persistentes, sua fisiologia e particularidades ecológicas são ainda incertas. Isso ocorre porque o isolamento e manutenção de culturas viáveis é muito difícil (Kim *et al.*, 2019), em especial dos clados acI (hgcI) e acIV (CL500), os grupos mais abundantes em ambientes de água doce (Newton *et al.*, 2011). Porém o recente avanço tecnológico para sequenciamento metagenômico vem revolucionando esta realidade. Isso porque, o exame cauteloso de genomas completos permite avaliar todas as rotas metabólicas desses organismos e identificar compostos, metabólitos e/ou enzimas necessárias, mas não produzidas pelo próprio organismo. Por exemplo, diversas cepas do clado hgcI puderam ser recentemente isoladas após a análise mais detalhada de seus metagenomas identificar a necessidade de adicionar a enzima catalase ao meio de cultura (Kim *et al.*, 2019), enzima essa relacionada à degradação de peróxido de hidrogênio (água oxigenada) em água e oxigênio gasoso.

Clado CL500-29 (Acidimicrobiia, linhagem acIV)

O clado CL500-29, é o mais abundante da ordem Acidimicrobiia. Os organismos desse clado, também chamado de linhagem acIV, são bastante comuns em lagos, geralmente encontrados em vida livre e apresentam abundâncias pontuais, sendo mais dominantes em determinadas condição ao longo do ano (De Melo, Bertilsson*, et al.*, 2019), mas nunca espacialmente (Mateus-Barros *et al.*, 2021 mat. sup. ESM3).

O tamanho de seu conteúdo genômico pode variar bastante (Chiriac, Haber*, et al.*, 2022),

Classificação

Filo: Actinobacteriota
Classe: Actinobacteria
Ordem: Acidimicrobiia
Família: Ilumatobacteraceae
Gênero/Clado: CL500-29

Habitat & Estilo de vida

Ambientes de água doce, salobros e marinhos
Vida livre
Podem ser "Streamlined"

Tamanho estimado do genoma

1.2 ~ 3.4 Mbp

porém possuir um genoma reduzido parece estar relacionado com a elevada abundância de algumas estirpes (Ghai *et al.*, 2012).

Acidimicrobiia apresenta elevada diversidade genômica, o que dificulta entender os detalhes a respeito de seu estilo de vida e preferências metabólicas (Chiriac, Haber*, et al.*, 2022). A presença constante de rodopsinas indica uma produção energética facultativa, por outro lado, a capacidade de fixação de CO_2 ainda é incerta (Chiriac, Haber*, et al.*, 2022).

Clado hgcl (linhagem acl)

Os organismos deste grupo são normalmente os mais abundantes em ambientes lacustres (De Melo, Bertilsson*, et al.*, 2019; Mateus-Barros *et al.*, 2021). Essas bactérias podem chegar a representar mais de 50% da abundância bacteriana total (Camara Dos Reis *et al.*, 2019; Mateus-Barros *et al.*, 2021) e mais de 90% da abundância de Actinobactérias (Warnecke *et al.*, 2005) na região do eplimínio de um lago, e uma persistência média de 30% ao logo do tempo (Parte II dessa tese, Fig. 2).

Classificação taxonômica
Filo: Actinobacteriota   **Classe:** Actinobacteria   **Ordem:** Frankiales   **Família:** Sporichthyaceae   **Gênero/Clado:** hgcl
Habitat & Estilo de vida
Água doce e salobras   Vida livre   "Streamlined"   Participa na degradação de C e N
Tamanho estimado do genoma
1.16 ~ 1.55 Mbp

Podem ser encontrados também em ambientes salobros, mas nunca dominam nestas condições (Hugerth *et al.*, 2015; Mehrshad *et al.*, 2016).

Sua incrível dominância parece sempre persistir e variar muito pouco ao longo do tempo (Newton *et al.*, 2011) e, apesar de auxotróficos (Kim *et al.*, 2019), não foi observada nenhuma relação específica com qualquer outro organismo até hoje (Chiriac, Haber*, et al.*, 2022). São capazes de produzir rodopsinas e degradar diversos compostos ricos em Nitrogênio e potencialmente biodisponibilizados por diversos outros grupos microbianos (Findlay *et al.*, 2003; Ghai *et al.*, 2014).

Gêneros Planktoluna e Rhodoluna (linhagem Luna-1)

Apesar de pouco representarem a abundância de actinobactérias, esse grupo pôde ser mais bem estudado que os apresentados anteriormente, já que muitas espécies puderam ser cultivadas em laboratório e descritas (Hahn, 2009).

Esses microrganismos apresentam um reduzido tamanho celular, característica útil na fuga da predação por ciliados heterotróficos e zooplâncton filtradores (Pernthaler *et al.*, 2001; Chiriac, Haber*, et al.*, 2022). Assim como os outros grupos acima descritos, este possui genoma e conteúdo GC reduzido (Hahn *et al.*, 2014; Kang, Ilnam *et al.*, 2017; Pitt *et al.*, 2021). Também podem produzir rodopsinas (Ghai *et al.*, 2014) e carotenóides (Pitt *et al.*, 2021), compostos uteis à sua produção de energia facultativa e à preferência por viver na coluna d'água próximo à superfície (Martinez-Garcia *et al.*, 2012; Pitt *et al.*, 2021) Podem ser encontradas em estuários, porém estão mais bem adaptadas a uma vida completamente adaptada à água doce (Chiriac, Haber*, et al.*, 2022).

Tribos filogeneticamente identificadas no grupo Luna são muito dispersas por ambientes de água doce (Hahn e Pöckl, 2005; Hahn *et al.*, 2014; Lipko e Belykh, 2021), e bem adaptadas a qualquer tipo de temperatura (Hahn *et al.*, 2014; Pitt *et al.*, 2021), uma característica que parece ser relacionada com diferentes tipos de linhagens adaptados a diferentes nichos de temperatura (Hahn e Pöckl, 2005).

Classificação taxonômica
Filo: Actinobacteriota
Classe: Actinobacteria
Ordem: Micrococcales
Família: Microbacteriaceae
Gênero/Clado: Luna-1
Habitat & Estilo de vida
Água doce
Vida livre
"Streamlined"
Tamanho estimado do genoma
1.3 ~ 1.4 Mbp

Bacteroidetes

Apresentando elevada plasticidade fenotípica e metabólica, bactérias classificadas como Bacteroidota podem ser encontradas em ambientes de água

doce e marinhos, ou associados a plantas, animais e até humanos (Newton *et al.*, 2011). Esses organismos podem chegar a dominar ambientes lacustres (Pernthaler *et al.*, 2004), em especial nos ambientes com elevada entrada de matéria orgânica proveniente do solo ou após florações de microalgas e cianobactérias (Eiler e Bertilsson, 2007; Zeder *et al.*, 2009; Buchan *et al.*, 2014).

Possuindo genomas grandes e reduzido conteúdo GC (Buchan *et al.*, 2014), esse grupo pode ser encontrado aderido a partículas (Lemarchand *et al.*, 2006) e está associado à degradação de matéria orgânica de alta complexidade (Salcher, 2014). Apesar disso tudo, estão entre os grupos bacterianos de água doce menos estudados (Newton *et al.*, 2011).

Cianobactérias

Cianobactérias são (de longe) os microrganismos de água doce mais bem estudados. Isso se dá principalmente por sua importância no ciclo biogeoquímico do Nitrogênio e na produção de oxigênio gasoso (Stanier e Cohen-Bazire, 1977), mas também por sua relevância para a qualidade de vida humana, já que alguns membros deste filo podem produzir toxinas com efeitos hepáticos e neurais (Suffet *et al.*, 1995) e potencialmente letais (Huisman e Hulot, 2005; Havens, 2008). O quIssoe encorajou centenas de pesquisadores ao longo de décadas a decifrar os fatores ambientais relacionados à florações e produção dessas toxinas (Huisman e Hulot, 2005).

Estes organismos são conspícuos, podendo ser encontrados em basicamente qualquer lugar (Drakare e Liess, 2010; Ahlgren *et al.*, 2019; Camara Dos Reis *et al.*, 2019; De Melo, Bertilsson*, et al.*, 2019; Bižić *et al.*, 2020; Mateus-Barros *et al.*, 2021). São ainda capazes de realizar fotossíntese oxigênica usando principalmente o pigmento clorofila *a* (Stanier e Cohen-Bazire, 1977). Em ambientes aquáticos, algumas adaptações indicam um estilo de vida completamente voltado à vida próxima a superfície, podendo ser incrivelmente abundantes e estar aderidas a partículas (Camara Dos Reis *et al.*, 2019). É possível encontrar vesículas de ar, utilizadas na regulação da posição da célula na coluna d'água, o que maximiza o

uso da luz solar na fixação de nutrientes (Huisman *et al.*, 2018). Outra interessante adaptação está na variação de cor como uma estratégia para maximizar a utilização da onda de luz dominante dependendo do local onde o indivíduo vive (Kehoe e Gutu, 2006).

Seu genoma é relativamente grande, comparado com outros organismos anteriormente citados aqui (> 2.4 Mbp), mas seu tamanho está relacionado à complexidade do ambiente ao qual cada espécie está adaptada. Genomas de organismos tipicamente terrestres são maiores que os de água doce e marinhos (Cabello-Yeves *et al.*, 2022). Adaptações à limitada presença de enxofre em ambientes de água doce também já foram identificadas (Cabello-Yeves *et al.*, 2022).

Cyanobium Gracile (PCC-6307)

O clado *Cyanobium* compreende um dos grupos mais abundantes e conspícuos dentre as cianobactérias (Cabello-Yeves, Pedro J *et al.*, 2017; Mateus-Barros *et al.*, 2021; Callieri *et al.*, 2022). Sua presença em ambiente marinho é marcante, mas análises genéticas indicam irradiação a partir de ambientes de água doce (Cabello-Yeves, Pedro J *et al.*, 2017).

Classificação taxonômica
Filo: Cyanobacteria **Classe:** Cyanobacteriia **Ordem:** Synechococcales **Família:** Cyanobiaceae **Gênero/Clado:** Cyanobium
Habitat & Estilo de vida
Conspícuo Versátil Relevante produtor primário Forma blooms
Tamanho estimado do genoma
≥ 2.5 Mbp

O pequeno tamanho celular destes organismos (Komárek *et al.*, 1999; Kwon *et al.*, 2021), aliado a uma área de superfície (Callieri *et al.*, 2022) e tamanho do genoma (Cabello-Yeves *et al.*, 2022) relativamente grandes lhes confere versatilidade e resiliência (Callieri *et al.*, 2022) e ajudam a explicar sua ubiquidade (Callieri, 2017). Além disso, esses organismos podem ser importantes produtores primários em ambientes lacustres (Jezberová e Komárková, 2007) e também podem controlar a abundância de algas nesses ambientes (Kovács *et al.*, 2018).

Planctomycecota

O filo Planctomycecota, reúne organismos com características morfológicas únicas. Formam espaços periplasmáticos (Rappé e Giovannoni, 2003) cuja função ainda não parece ser totalmente compreendida, mas podem servir para a digestão de macromoléculas (Chiriac, Haber, *et al.*, 2022).

Possuem um grande tamanho genômico, alcançando até 12.4 Mbp (Wiegand *et al.*, 2018; Andrei *et al.*, 2019), sendo capazes de produzir enzimas para a degradação de compostos de grande peso molecular e recalcitrante proveniente de diversas fontes (Wiegand *et al.*, 2018; Andrei *et al.*, 2019). São também capazes de realizar degradação de amônio (Strous *et al.*, 1999). Apesar disso, são auxotróficos para alguns compostos (Andrei *et al.*, 2019).

Estes organismos podem ser encontrados em água doce, salobra e marinha, de vida livre ou associados a partículas e outros microrganismos (Wiegand *et al.*, 2018; Andrei *et al.*, 2019), podendo chegar a abundâncias relevantes em ambientes lacustres (Mateus-Barros *et al.*, 2021).

Proteobacteria

Um dos mais bem conhecidos grupos dentre as bactérias de água doce, o filo das Proteobactérias tem na bactéria Escherichia coli sua espécie mais famosa, e uma muito bem conhecida importância para a qualidade de vida do ser humano (Newton *et al.*, 2011). Este grupo, bastante abundante em ambientes lacustres (Mateus-Barros *et al.*, 2021), tem como principais características sua elevada plasticidade genética e fenotípica e associação com diversos relevantes ciclos biogeoquímicos, tais como Nitrogênio, Carbono e Enxofre (Newton *et al.*, 2011; Chiriac, Haber*, et al.*, 2022). Adaptações a um estilo de vida simplificado ('streamlined') (Salcher *et al.*, 2011), associado a partículas (Camara Dos Reis *et al.*, 2019) ou a um intenso afluxo de nutrientes (Gutierrez, 2019) estão presentes. Ubiquidade é comum aos gêneros mais abundantes (Morris *et al.*, 2002; Hahn *et al.*,

2022), bem como a presença em ambientes marinhos e de água doce (Chiriac, Haber*, et al.*, 2022), com proeminentes casos indicativos de adaptação e colonização de um ambiente a partir do outro (Henson *et al.*, 2018; Salcher *et al.*, 2019) apesar de sua dominância em cada tipo de ambiente depender do grupo abordado.

O filo Proteobacteria possui três grandes classes, Alphaproteobacteria, Betaproteobacteria e Gammaproteobacteria, onde estão abrigados seus mais proeminentes representantes, além de algumas outras de menor relevância.

A classe Alphaproteobacteria abriga o grupo bacteriano mais bem sucedido do ambiente marinho, o clado SAR11(Morris *et al.*, 2002). Representantes deste grupo podem também ser endossimbiontes (Batut *et al.*, 2004) e estão associados a plantas vasculares, tendo um relevante papel na fixação de Nitrogênio [i.e. Rhizobiales (Stacey, 2007)]. Outros ciclos biogeoquímicos cujo grupo parece ter relevante função são os do Carbono e Enxofre, compostos fixados por membros do clado SAR11 no oceano global (Giovannoni, 2017). Foi ainda relatado que alguns organismos podem estar associados à regulação da abundância de Cianobactérias causadoras de blooms (Berg *et al.*, 2009). O tamanho do genoma é variável a depender do estilo de vida adotado (Hahn *et al.*, 2022).

Os maiores representantes da classe Betaproteobacteria parecem estar mais intimamente relacionados com ambientes lacustres de água doce (Chiriac, Haber*, et al.*, 2022), sendo ubíquos e abundantes neste tipo de ambiente (Šlmek *et al.*, 2005; Salcher *et al.*, 2010; Mateus-Barros *et al.*, 2021). Uma demarcada partição de nicho e intensa recombinação de genes intra- clados parece ser a regra dentro deste grupo (Chiriac, Haber*, et al.*, 2022) e ajuda a explicar a extensa distribuição geográfica observada para esta classe.

Já Gammaproteobactérias podem ser abundantes em ambientes salobros (Wu *et al.*, 2006) ou com intensa eutrofização, em especial causada por atividade humana (Gutierrez, 2019) e crescem rapidamente na presença de elevada concentração de nitrogênio (Gasol *et al.*, 2002) e fósforo (Šimek *et al.*, 2006).

SAR11, Clado IIIb (linhagem LD12; Ca. Fonsibacter)

O clado IIIb da ordem SAR11 contém as bactérias que melhor puderem se adaptar à vida em água doce (Salcher *et al.*, 2011). Esse grupo é amplamente distribuído por ambientes lacustres (Lindström *et al.*, 2005; Wu *et al.*, 2006; Mateus-Barros *et al.*, 2021), porém nenhuma espécie parece ser capaz de colonizar águas salobras (Henson *et al.*, 2018; Tsementzi *et al.*, 2019), isso aliado à reduzida diversidade filigenética do clado (Logares *et al.*, 2010), evidencia um evento de isolamento

Classificação taxonômica
Filo: Proteobacteria
Classe: Alphaproteobacteria
Ordem: SAR11
Família: -
Gênero/Clado: Clade III
Habitat & Estilo de vida
Água doce
"Streamlined"
Tamanho estimado do genoma
~ 1.6 Mbp

bastante radical entre esse e outros grupos aparentados de SAR11 que impede a sua diversificação através de recombinações (Zaremba-Niedzwiedzka *et al.*, 2013). Esses organismos estão relacionados à redução de compostos sulfurosos (Henson *et al.*, 2018) e parecem preferir ambientes com temperatura mais elevada (Salcher *et al.*, 2011; Henson *et al.*, 2018).

Sua origem provavelmente ocorreu da diversificação a partir de clados de SAR11 marinhos com genoma simplificados (Logares *et al.*, 2010; Henson *et al.*, 2018), o que significa que seus genomas também são reduzidos (Salcher *et al.*, 2011), com apenas a aquisição de genes relacionados à produção de rodopsinas (Giovannoni *et al.*, 2005), e explica sua adaptação irreversível aos ambientes de água doce.

Polynucleobacter

O genero Polynucleobacter está entre os grupos bacterianos mais bem estudados vivendo em ambientes de água doce (Chiriac, Haber*, et al.*, 2022), seu genoma e tamanho celular e porcentagem de conteúdo GC são reduzidos (Chiriac,

Haber*, et al.*, 2022). Adaptações para um estilo de vida não móvel com possibilidade de quimioheterotrofia estão presentes (Hahn *et al.*, 2022), além da possibilidade de produção de rodopsinas (Hahn *et al.*, 2016), e indicam uma vida majoritariamente passada na coluna d'água, próxima à superfície. De fato, sua abundância pode diminuir gradualmente com a profundidade (Wu e Hahn, 2006).

Classificação taxonômica
Filo: Proteobacteria **Classe:** Gammaproteobacteria **Ordem:** Burkholderiales **Família:** Burkholderiaceae **Gênero/Clado:** Polynucleobacter
Habitat & Estilo de vida
Conspícuo; Partição de nicho e recombinação gênica intensas Vida livre ou endossimbiontes Streamlined
Tamanho estimado do genoma
1.5 ~ 2.5 Mbp

Este grupo possui representantes endossimbiontes (Heckmann e Schmidt, 1987) e de vida livre (Salcher *et al.*, 2008; Hahn *et al.*, 2022), chegando a alcançar grandes abundâncias em lagos (Salcher *et al.*, 2008; Hahn *et al.*, 2022).

Intensa recombinação de genes (Hoetzinger *et al.*, 2017) e relevante partição de nicho (Hahn *et al.*, 2021) parecem ser regra neste gênero, o que deve ser uma das principais causas para a ampla distribuição espacial e altas taxas de diversificação do grupo (Hoetzinger *et al.*, 2017; Hoetzinger *et al.*, 2019; Hoetzinger *et al.*, 2021). Algumas espécies com conteúdo genético muito similar apresentam diferenças ecológicas evidentes (Hahn *et al.*, 2016). Alterações na concentração de pH e Carbono Orgânico Dissolvido parecem ser fatores determinantes para essa partição, com sua variação estimulando ou inibindo o crescimento populacional de uma ou outra variante (Lindström *et al.*, 2005; Wu *et al.*, 2006).

Limnohabitans

Assim como o clado *Polynucleobacter*, o gênero *Limnohabitans* compreende um grupo bacteriano de água doce bem estudado. Estes organismos são copiotróficos e conspíquos (Kasalický *et al.*, 2013; Shabarova *et al.*, 2017), podendo chegar a representar uma grande parte da abundância bacteriana em um lago (Šlmek *et al.*, 2010).

O tamanho do seu genoma é relativamente grande em comparação com outras bactérias dominantes na coluna d'água doce (Zeng *et al.*, 2012; Kasalický *et al.*, 2018), o que lhes permite tolerar grandes variações ambientais (Kasalický *et al.*, 2013). Algumas variantes possuem inclusive a capacidade de produzir proteínas relacionadas à fotossíntese (Kasalický *et al.*, 2018), e fixação autotrófica de carbono (Zeng *et al.*, 2012).

Classificação taxonômica
Filo: Proteobacteria **Classe:** Gammaproteobacteria **Ordem:** Burkholderiales **Família:** Comamonadaceae **Gênero/Clado:** Limnohabitans
Habitat & Estilo de vida
Conspícuo Partição de nicho e recombinação gênica intensas Versátil
Tamanho estimado do genoma
3 ~ 4.7 Mbp

Esse clado parece estar mais bem adaptado a crescer rapidamente quando nutrientes entram no sistema (Šimek *et al.*, 2005; Šimek *et al.*, 2006), inclusive quando produzido e exsudado por algas (Kasalický *et al.*, 2013), apresentando notável versatilidade em relação ao tipo de composto utilizado para nutrição (Kasalický *et al.*, 2013).

A partição de nicho parece também ter um papel importante para a diversidade desse grupo, com a presença de alguns clados especializados em viver em lagos (Props e Denef, 2020) enquanto outras linhagens estão mais adaptadas à vida em rios de pequenas proporções e próximos a nascentes (Shabarova *et al.*, 2017). Além disso, um estudo recente pôde demonstrar a possibilidade de mudanças de dominância dentro deste grupo ao longo do tempo (Jezberová *et al.*, 2017).

Methylophilaceae

A Família Methylophilaceae reúne clados relacionados à degradação de compostos com 1 (um) carbono, um processo importante para o ciclo do carbono (Salcher *et al.*, 2015). Podem apresentar estilo de vida pelágica de vida livre ou associados ao sedimento, com genomas reduzidos e genes relacionados à sobrevivência com pouca disponibilidade de nutrientes (Chiriac, Haber*, et al.*, 2022).

Um aspecto interessante da história evolutiva deste grupo é que existem evidencias que indicam a passagem por dois eventos de adaptação. Os dois primeiros estão relacionados à passagem de uma vida associada à degradação de matéria orgânica em sedimentos para a vida livre na coluna d'água (Salcher *et al.*, 2019). Esta adaptação aconteceu independentemente e de forma similar em ambos grupos de vida marinha e de água doce, com uma tendência de perda de genes relacionados à motilidade e quimiotaxia

Classificação taxonômica
Filo: Proteobacteria **Classe:** Gammaproteobacteria **Ordem:** Burkholderiales **Família:** Methylophilaceae **Gênero/Clado:** -
Habitat & Estilo de vida
Conspícuo Vida livre ou associado ao sedimento Fixadoras de compostos C1
Tamanho estimado do genoma
1.3 Mbp (pelágicos); 2.3 ~ 4.6 (associados ao sedimento) Mbp

(Salcher *et al.*, 2019) e aquisição de genes para produção de rodopsinas (Chiriac, Haber*, et al.*, 2022). Há ainda um terceiro evento de adaptação de um grupo adaptado à vida pelágica em água doce para a vida pelágica marinha, com aquisição de genes relacionados à osmorregulação (Salcher *et al.*, 2019).

Verrucomicrobiotas

Verrucomicrobiota podem apresentar uma vasta gama de estilos de vida, podendo ser encontradas em ambientes terrestres, de água doce e salobros (Wagner e Horn, 2006), além de poderem estar associados a eucariotos (Wagner e Horn, 2006), partículas ou vida livre (Parveen *et al.*, 2013). Podem ser mais abundantes durante florações de algas (Eiler e Bertilsson, 2004). Como é de se esperar para um grupo com tantos possíveis estilos de vida, o tamanho de seu genoma pode variar grandemente, entre 1.8 e 6.2 Mbp (Cabello-Yeves, P. J. *et al.*, 2017). Além disso, a sua presença parece estar relacionada com um ambiente mais ácido e temperaturas mais elevadas (Lindström *et al.*, 2005), o que pode explicar sua abundância sempre relevante encontrada nas lagoas rasas tropicais apresentadas aqui (Mateus-Barros *et al.*, 2021).

Clado FukuN18

Este é outro grupo cuja abundância pode chegar a ser relevante em ambiente lacustre (Mateus-Barros *et al.*, 2021), sendo o grupo dominante deste filo (Parveen *et al.*, 2013). Estes organismos típicos de água doce (Lindström *et al.*, 2005), são majoritariamente associados a partículas (Parveen *et al.*, 2013).

Estas bactérias constituem tamanhos celulares maiores (Piwosz *et al.*, 2018) e já tiveram sua presença relacionada à predação por nanoflagelados (Piwosz *et al.*, 2018) e

Classificação taxonômica

Filo: Verrucomicrobiota
Classe: Verrucomicrobiae
Ordem: Chthoniobacterales
Família: Terrimicrobiaceae
Gênero/Clado: FukuN18

Habitat & Estilo de vida

Conspícuo
Aderidos à partícula
Chave na transferência de C à cadeia trófica clássica

Tamanho estimado do genoma

1.8 ~ 6.2 Mbp

rotíferos (Parveen *et al.*, 2013), o que pode indicar uma função ecossistêmica chave ao servir de conexão entre o loop microbiano e a cadeia trófica clássica (Piwosz *et al.*, 2018). De qualquer maneira, este constitui um grupo muito pouco estudado e seus genomas parecem não ter sido desvendados até agora.

Ca. Patescibacteria

O filo Patescibacteria foi adicionado a esse pequeno resumo por motivos opostos aos citados acima. Patescibacteria não é um filo particularmente abundante ou formador/causador de florações, porém este possui um grande número de representantes em diversos sítios (Mateus-Barros *et al.*, 2021). Ainda há pouca informação disponível a respeito da biologia deste grupo (Castelle *et al.*, 2018), porém já se sabe que seus genomas e tamanho celular são reduzidos (Castelle *et al.*, 2018; Chiriac, Bulzu*, et al.*, 2022). Seu estilo de vida principal parece ser parasítico (Castelle *et al.*, 2018), mas organismos de vida livre também estão presentes (Chiriac, Bulzu*, et al.*, 2022).

Capítulo III -
Distribuição de frequências de ocupação e abundância

"Ao contrário do que pretendem os livros didáticos, a melhor parte da ciência não está nos modelos matemáticos nem nos experimentos. Isso vem depois. O melhor da ciência emerge de um modo mais primitivo de pensar através do qual a mente do caçador vai tecendo ideias a partir de fatos velhos, metáforas novas e imagens confusas e semiensandecidas de coisas vistas recentemente. Avançar na ciência é elaborar novos padrões de pensar, que definirão por sua vez os modelos e os experimentos. Fácil de dizer, difícil de fazer" (Edward O. Wilson, 1992 em A diversidade da vida)

Durante minha dissertação, foquei em principalmente entender o padrão de distribuição de bactérias de água doce vivendo em diversas lagoas rasas próximas a nascentes e distribuídas por todo o Estado de São Paulo, em uma paisagem que compreende cerca de 250 mil km². Aquele trabalho gerou dois capítulos, o primeiro analisava duas metodologias de extração de DNA e buscava comparar os resultados de um ponto de vista de quais dados ecológicos poderiam ser interpretados de diferente forma a depender do método analisado. No segundo capítulo, buscamos aplicar a hipótese da distribuição bimodal de espécies para identificar organismos satélite e core, nessa metacomunidade. Este trabalho ainda era bastante insipiente no momento da defesa da minha dissertação em 2018, e precisou ser bastante melhorado nos anos subsequentes, já dentro do meu doutorado. Entre o documento da minha dissertação (Mateus-Barros, 2018) e o resultado final (Mateus-Barros *et al.*, 2021) publicado na revista Microbial Ecology restou pouca semelhança. De igual mesmo, ficou apenas a ideia de testar se existe bimodalidade na distribuição de bactérias de água doce.

Agora, nós pudemos observar que de fato, a distribuição de ocupação para as bactérias presentes nestes ambientes é bimodal. Neste sentido, dois distintos grupos se destacam: (i) as bactérias "Satélite" são aquelas muito pouco persistentes espacialmente e totalizam centenas de unidades taxonômicas, e (ii) as bactérias "Core" que, em oposição, são organismos ubíquos (podem ser encontrados em todos os lugares, onipresentes). Nós aproveitamos a oportunidade para analisar como era a distribuição de abundância desses organismos core e pudemos notar uma predominância de padrões bimodais e normais nessa distribuição. Isso indica que esses organismos são na verdade muito abundantes em alguns sítios e mais raros em outros, e que esses sítios onde são mais abundantes possivelmente servem como uma fonte de indivíduos de onde saem e se dispersam para o restante dos sítios nessa paisagem. Finalmente, comparamos estes dados biológicos com dados ambientais e de distância espacial e notamos que esses padrões são guiados principalmente em função da variação do pH, mas também movidos em menor força pelas distâncias entre sítios, apesar de não termos tido a chance de entender melhor de que forma esses fatores ambientais estavam influenciando a biodiversidade nessa metacomunidade.

Após esse estudo, muitas questões permaneceram e eu tive a oportunidade de guiar esta tese no sentido de tentar sanar algumas dessas questões.

O objetivo geral é buscar entender como a variação espaço-temporal de fatores determinísticos e estocásticos impacta no padrão de estruturação de comunidades microbianas em lagoas rasas de água doce em diferentes escalas de observação.

No primeiro trabalho que apresento a seguir (Capítulo IV) busquei responder às perguntas: Quais são os fatores locais que estão relacionados com a persistência de determinadas bactérias? O que potencialmente previne que uma bactéria abundante seja também parte do grupo de organismos core em uma metacomunidade lacustre? Alguns organismos raros podem persistir ao longo do ano, mas é possível que façam parte do core espacial? Para isso, busquei descobrir quais organismos representam um grupo Core no que diz respeito à persistência temporal em uma lagoa tropical (Fig. 3A). Estes organismos Core foram então comparados com diversos fatores ambientais e interações com outras bactérias também muito abundantes ou persistentes e também picoeucariotos para tentar encontrar alguma relação que indique o que pode potencializar a dominância de certos organismos enquanto previne a dominância de outros. O intuito era de discutir

como esse padrão pode estar relacionado com o padrão bimodal de distribuição espacial de unidades taxonômicas nesta mesma região.

Após, no capítulo V, busquei responder: Quais fatores espaciais estão relacionados com a distribuição espacial dessas bactérias? Para tal, apliquei o método da Partição da Variância e comparei as dissimilaridades biológicas entre diversos sítios (Fig. 3B) contra eigenvetores recuperados de modelos matemáticos que buscam representar o efeito de tipos distintos de distâncias espaciais sobre as dissimilaridades biológicas observadas (Legendre e Legendre, 2012). Quatro modelos foram testados. O primeiro (dbMEM – "distance based Moran's Eigenvector Maps) representa as distâncias lineares entre sítios e a capacidade dispersiva sobre a terra que bactérias podem apresentar. Também apliquei o método MEM ("Moran's Eigenvector Maps") utilizando duas tabelas de distâncias como referência, uma para representar o efeito de uma relação maior de troca de indivíduos entre sítios vizinhos que a maiores distâncias, outro para evidenciar o efeito do isolamento geográfico proporcionado pelas bacias hidrográficas nessa região. E finalmente, utilizei o AEM ("Asymetric Eigenvector Maps") para representar a conexão destes sítios via fluxo dos rios. Dessa forma abarquei as formas de distribuição mais comuns disponíveis para bactérias e busquei entender qual a importância relativa destes para este reino.

E finalmente, no capítulo VI busquei a resposta para: O aumento da escala de observação também muda a relação entre fatores que são vistos como primariamente guiando essas comunidades? Para isso, apliquei o método de Stegen e colaboradores (2012) e comparei os dados coletados no Brasil com outros dois sets de dados semelhantes coletados na Argentina e no Canada (Fig. 3C) buscando entender quais entre ambientais determinísticos e espaciais estocásticos são relevantes em três escalas: *Regional*, comparando sítios em cada set de dados, *Continental*, comparando dados de Brasil e Argentina e *Intercontinental*, adicionando a estes os padrões observados no Canadá.
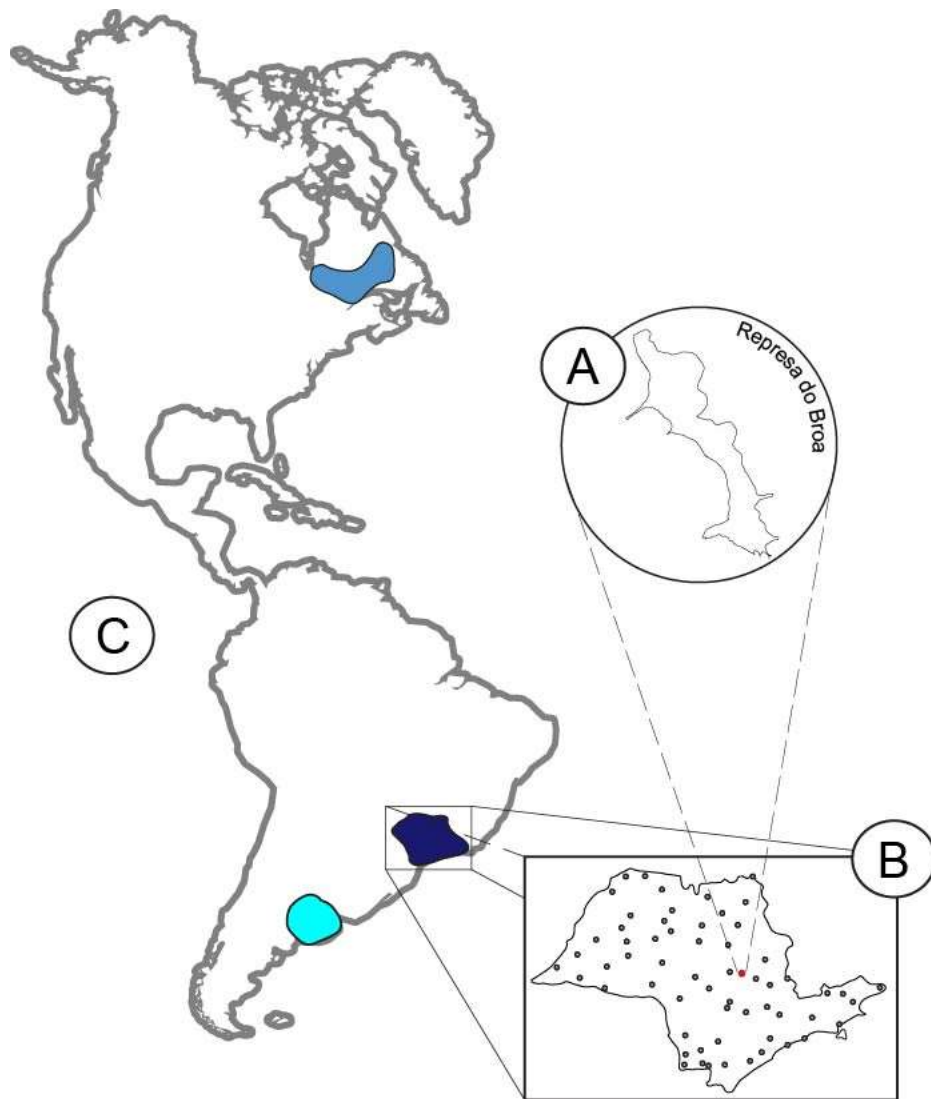
Figura 3 – Escalas geográficas compreendidas nesta tese. Para o estudo que visou abordar os aspectos locais que potencialmente guiam a dominância de certas bactérias no espaço, foram coletadas amostras mensais na Represa do Broa (A), um reservatório raso e pequeno. Para o capítulo que aborda os aspectos espaciais relacionados à dispersão destes microrganismos, um set de dados contendo 60 amostras coletadas em lagoas rasas distribuídas por uma larga área e alcançam distâncias máximas de 822 km (B). Estes dados foram agregados a uma base de dados maior contendo amostras semelhantes de Argentina (distância máxima = ~547 km) e Canada (distância máxima = ~1137 km) (C) e utilizados no estudo a respeito de como fatores determinísticos e estocásticos impactam as dinâmicas bacterianas em distintas escalas.

Capítulo IV -
Depicting spatiotemporal variables
that drive the dynamics of dominance in lake bacteria

As bactérias são organismos-chave para funções ecológicas e estruturação de comunidades. Apesar disso, características espaço-temporais só puderam ser mais bem compreendidas após o advento de abordagens moleculares independentes de cultura. Aqui, nós investigamos quais processos temporais locais potencialmente contribuem para a dominância local de organismos regionalmente bem distribuídos. Para isso, nós agrupamos os microrganismos em distintas categorias considerando sua abundância média e persistência. Em geral, os organismos classificados como "Core" (elevada abundância e persistência) são tipicamente conhecidos por viver na coluna d'água em lagos continentais, apresentando um estilo de vida simplificado e dependem de nutrientes exsudados por outras bactérias e fitoplâncton. Apesar disso, não parecem estar relacionados com grupos biológicos específicos. As bactérias "Persistentes" (baixa abundância, elevada persistência) parecem estar mais intimamente relacionadas com o aporte de nutrientes ricos em ácido húmido vindos do solo ao redor ou são conhecidos por se relacionarem com fitoplâncton formadores de florações. Os dois grupos são capazes de dispersar e colonizar sítios semelhantes ao estudado aqui, porém o segundo grupo provavelmente seria filtrado em comunidades aquáticas menos impactadas pelo ambiente terrestre. Os organismos "Abundantes" (elevada abundância, menor persistência) são fortemente selecionados por características locais, e não são eficientes colonizadores para a vizinhança. Finalmente, zooplâncton parece estar mais relacionado com organismos "Transientes" (abundantes pelo menos uma vez) ou "Satélites" (reduzida abundância e persistência). Os achados apresentados aqui indicam a importância de processos bottom-up guiando a seleção de bactérias localmente que dominam a região, em contraste a processos top-down que parecem apresentar uma relevância mais reduzida. Apesar disso, há algum indicativo que a predação desempenha um papel em manter os organismos mais raros.

Introduction

Bacteria are critical for ecological functions and community structure, being essential in biogeochemical cycles (Myklestad, 2000), nutrient cycling (Kawasaki and Benner, 2006; Sarmento and Gasol, 2012) and closely related to other living creatures (Sarmento, 2012; Martínez *et al.*, 2015). Although these organisms have been studied for decades, spatiotemporal features could be better understood only after the advent of culture-independent molecular approaches (Taberlet *et al.*, 2012). These new techniques have been highlighting that, contrary to what was thought (Baas Becking, 1934), not all bacteria are well spread, and the study of their biogeography can unveil new patterns (Martiny *et al.*, 2006). Also, the integrated analysis of both micro- and macroorganisms became possible (Horner-Devine e Bohannan, 2006), which can boost the discoveries in the ecology field (Acinas *et al.*, 2004; Barberán *et al.*, 2014). Notably, studying bacteria, the possibility of applying well-established mathematical models provides new insights into their macroecological structure (e.g. Lindh *et al.*, 2017; Soininen *et al.*, 2018).

The core-satellite hypothesis (Hanski, 1982) states that the frequency of occupancy in a metapopulation should be bimodal. This spatial pattern of species distribution indicates that, when the sampling effort is ideal, most of the organisms found should be present in one or few sites, while others will be distributed throughout the entire landscape (Gleason, 1929). This distribution dichotomy evidences the existence of two distinct groups of organisms regarding occupancy: on one hand, there are the satellites which are those restricted to one or few sites, while on the other, the most persistent organisms are called core (Hanski, 1982). This phenomenon seems to be mainly related to abundance (Papp and Izsák, 1997) since these organisms are generally those that manage to (re)colonize new sites in a dispersal wave that gradually increases the area covered by a species (Brown, 1984).

This model emerges as a simple way to explore the macroecological structure of microorganisms (Lindh *et al.*, 2017; Jeong *et al.*, 2020; Mateus-Barros *et al.*, 2021; Escalas *et al.*, 2022). From this perspective, it is possible to recover some interesting information about biodiversity, since its pattern depends on the dispersion capacity

inherent to each organism that should impact this structural feature. For example, some aquatic bacteria can be regionally persistent even when they are not abundant (Lindh *et al.*, 2017), but zooplanktonic metacommunities do not seem to have a bimodal distribution (Soininen and Heino, 2005). Additionally, while the satellite encompasses the major part of regional diversity (Mateus-Barros *et al.*, 2021), the core is composed of organisms that are potentially key to ecosystem functions (Fonte *et al.*, 2021). The core-satellite approach is also a valuable approach to retrieving information concerning the metacommunity structure (*i.e.* a set of local communities linked by a flux of individuals). By using this approach it is possible to recover details on topics such as colonization and extinction dynamics, adaptiveness and dispersal capacity, niche partition, and competition (Gaston *et al.*, 2000; Mehranvar and Jackson, 2001; Lindh *et al.*, 2017).

The bimodal distribution of species is probably directly related to the abundance variation through space and time. Locally, environmental factors like pH and temperature (Lindström *et al.*, 2005; Niño-García *et al.*, 2016; Mateus-Barros *et al.*, 2021) can determine which groups will likely be more abundant and, consequently, have increased chances to disperse a greater number of individuals, and dominate spatially (Gaston *et al.*, 2000). A long-term study on the dynamics of a phytoplanktonic community in a tropical lake showed that the change in environmental conditions changed dominance features, and allowed a non-seasonal change of ecological state, from filamentous cyanobacteria to green algae (Rugema *et al.*, 2019).

Biological interactions can also be a relevant feature to be considered. A microbial community is connected in an intricate interaction web (i.e. microbial loop Azam *et al.*, 1983; Sarmento, 2012) by compounds incorporation made available by exudation, excretion, and cell lysis by other organisms (Kawasaki e Benner, 2006; Sarmento and Gasol, 2012). The most abundant bacterial clades living in the freshwater and marine epilimnion are specialized to optimize their genomes. They maintain genes necessary to key functions (Escalas *et al.*, 2022) and use the resources released by other bacteria and picophytoplankton (Sarmento, 2012; Sarmento and Gasol, 2012; Chiriac, Haber*, et al.*, 2022). Predation and grazing are other factors that can reduce bacterial abundance (Livingston *et al.*, 2017) and

maintaining the community evenness (Segovia *et al.*, 2018). A reduced cell size and aggregation are adaptations that can be observed in scenarios of high top-down control (Pernthaler *et al.*, 2001; Chiriac, Haber*, et al.*, 2022), also, the community controlling done by *Daphinia sp.* in temperate environments is a well-documented pattern (Zöllner *et al.*, 2003); and a strategy of maintaining regional rarity to avoid predation is also possible for bacteria communities (Lindh *et al.*, 2017). Besides that, the relationship between these variables and bimodality remains an interesting question to be depicted.

In this sense, our aim was to investigate the local temporal processes that may contribute to the abundance and persistence of bacteria, and their roles in guarantying or preventing certain organisms from being included in the core of a metacommunity. For this purpose, we used a dataset collected monthly in a tropical shallow reservoir to test the impact of a series of local environmental and biological variables on the persistence and abundance variation of bacteria that are dominant aiming to understand what may make them temporally core or not. We hypothesize that the organisms present in the core were those who have a more generalist lifestyle and with a capacity of withstanding the annual environmental heterogeneity, in opposition to the transient bacteria, which should be recognized by a bloom-forming lifestyle. Complementarily, the bacteria which are persistent and not abundant should be those most sensitive to bottom-up and/or top-down effects in the food chain. On the other hand, those bacteria abundant and not persistent should be mainly represented by more sensitive organisms to environmental heterogeneity through time.

Methods

The Broa's Microbial Observatory (BroaMO)

The Broa reservoir (47°53'44.2"W, 22°10'58.2"S), also called Lobo reservoir is a small-sized and shallow lake, dammed to fulfill local human necessities for water and energy supplies, and recreation. This site is a well-studied place with environmental factors and distinct biological compartments being studied since the

1970s (p. ex. Rocha and Matsumura-Tundisi, 1976), but that never had a microbial long-term observatory settled before. Also, this reservoir is close to the lab's university, which allows for bringing water to the laboratory and easily conducting some sampling steps. All these characteristics made it the perfect place to initiate a long-term sampling effort to study the microbial community structure and functions. The Broa' Microbial Observatory (called hereafter BroaMO) had its activities initiated on 15th March 2018 and has been collected monthly until now. This Microbial Observatory was idealized as part of the long-term microbial observatories' network which was settled in many countries in Latin America. This project is part of the "Collaborative Network in Microbial Aquatic Ecology in Latin America" (µSudAqua) (for more details, see the network's website https://microsudaqua.netlify.app/en/).

This paper is the first report about microbial structuring found at BroaMO and encompasses data from the first 15 sampling efforts (March of 2018 to May 2019), and a complete cycle of dry and rainy seasons.

Sampling procedures and molecular analyses

The sampling has been taken each month, always during the second fortnight in the morning. In the field, we take some of the environmental data such as water temperature and Dissolved Oxygen using a multi-parameter probe, and the water column transparency with a Secchi disk, for further calculations of the Euphotic Zone(Cunha *et al.*, 2013). We also collected subsurface water samples to take further analyses at the laboratory. At the laboratory, we filtered water through membranes of distinct pore sizes aiming to recover a set of environmental and biological data. We filtered ~500 mL of water through a 0.45 µm mash glass fiber filter to retain material to further chlorophyll-a (chl-a) measurements, while the filtrate was read in a TOC-V (Shimadzu®, Kyoto, Japan) to estimate Dissolved Organic Carbon (DOC) concentration. Also, 150-200 mL passed through 0.22 µm polycarbonate membranes to retain the environmental DNA (eDNA), while the filtrate was read in an FS5 Spectrofluorometer (Edinburgh Instruments®, Livingston, UK) to estimate features of Dissolved Organic Matter (DOM). We also measured the pH using a pHmeter at the laboratory.

The filtered material for chl-*a* concentration was extracted with ethanol (90% v/v at 80 °C) in the dark (Marker, 1980; Mush, 1980), quantified by spectrophotometry (Lorenzen, 1967), and used as a proxy of the trophic state as calculated by Cunha and collaborators (2013), in an approach adapted from Carlson (1977) to tropical environments.

The filtrated material for DOM was analyzed to recover information about their Coloured (cDOM) and Fluorescent (fDOM) fractions as described in detail by de Melo and collaborators (2019). These measurements were done in the dark, at room temperature, and using a 0.01 m quartz cuvette (1 nm intervals and dwell time of 0.2 s). For the cDOM analysis, absorbance spectra from 200 to 800 nm were recovered and the resulting values were compared against an ultrapure water blank spectrum to remove the interference of water. The logaritimized resultant values can be interpreted as an exponential equation and the extraction of slope values from distinct fragments along this line allows to recover DOM data. The Slope Ratio (SR) is the ratio between slopes recovered from the 275-295nm and 350-400nm intervals and, the resultant value can be interpreted as inversely related to molecular weights (Helms *et al.*, 2008). Also, the C:A peaks ratio was calculated by dividing the values at 275 excitation, 304 emission (C peak) and 260 excitation, 450 emission (A peak) nm; in this case, smaller values indicate a DOM with high humic content while greater values indicate a DOM with high fulvic content (Hansen *et al.*, 2016). For the fDOM analysis, the excitation-emission matrices were performed. The wavelengths ranged from 240 to 450 nm in 10 nm increment excitation and from 300 to 560 nm in 2 nm increments emission. The dwell time was 0.25 s and a bandwidth of 5 nm. The matrices were then corrected in successive steps to guarantee the quality of DOM measurements. First, the matrices were corrected against an ultrapure water matrix to remove background peaks; also, the reference lamp signal was used to clear it from the sample, and the fluorescence spectra were corrected for inner filter effects with the absorbance-based approach (Lakowicz, 2006; Kothawala *et al.*, 2013). Finally, the matrices were calibrated to Raman units (Lawaetz and Stedmon, 2009). The corrections were conducted using the software Matlab and the DOMcorr toolbox (Murphy *et al.*, 2010). After the corrections, some other measurements were recovered from the excitation-emission matrices. The Fluorescence Index (FI) is

calculated as the ratio between wavelengths at 470 and 520 nm emission and 370 nm excitation (Cory *et al.*, 2010), and indicates the source in which the DOM was mainly produced. DOM mainly derived from the surrounding soils are indicated by lower FI values, on the other hand, higher FI values indicate an autochthonous DOM derived from microbial metabolism (Mcknight *et al.*, 2001). To calculate the Humification Index (HIX), it is necessary to measure the ratio between the matrix area from 435 to 480 nm and the summed peaks area from 300 to 345 nm and from 435 to 480 nm, at 254 nm excitation. In this case, greater values indicate a DOM content with higher concentrations of humic compounds (Ohno, 2002). Finally, the Freshness Index (FR) was calculated as the ratio between the value at 380 nm and the maximum value found in the area from 420 to 435 emission, at 310 nm excitation. Higher values of FR imply recently produced DOM (Parlanti *et al.*, 2000).

For the molecular procedures, eDNA was extracted using a Qiagen® DNeasy PowerWater kit. The extracted material was amplified using the 16S rRNA gene for bacteria with 341F (5′-CCTACGGGNGGCWGCAG-3′) and 805R (5′-GACTACHVGGGTATCTAATCC-3′) primers (Herlemann *et al.*, 2011; Apprill *et al.*, 2015) and the 18S rRNA gene for eukaryotes with TAReukFWD1 (5'-CCAGCASCYGCGGTAATTCC-3') and TAReukREV3 (TTTCGTTCTTGATYRA 5'-CA-3') primers (Stoeck *et al.*, 2010). The amplified material was sequenced in an Illumina MiSeq plataform and processed using the DADA2 pipeline (Callahan *et al.*, 2016) to check quality, remove chimeras and assign taxonomy using the SILVA 132 database (Quast *et al.*, 2013; Yilmaz *et al.*, 2014). The DADA2 was conducted in the R software (R Core Team, 2019). The rRNA amplicon results have been deposited in the NCBI repository and are summarized in the μSudAqua datapaper (Metz *et al.*, 2022).

To guarantee the quality of final analyses, the ASVs table passed through some filtering steps. First, genes belonging to mitochondria, chloroplasts, archaea, and vascular plants and animals were removed from the tables, also, the dataset was rarefied by the lowest sample richness (24,860 reads for 16S rRNA, 88,399 reads for 18S rRNA), and after, the ASVs that did not reach 10 reads summing all samples were removed. Also, because of some issues due to the sequencing, which were previously reported (Gloor *et al.*, 2017), we will always treat these datasets in terms

of relative abundances. And last, but not least we used the Bray Curtis (dis)similarity when applicable because this index is more appropriate to our datasets, which have excesses of zeros (Anderson *et al.*, 2011).

Statistical analyses

In order to determine if the data obtained in BroaMO is relevant to understand which local processes potentially guide the bacterial dynamics of metacommunity dominance, we tested whether these data met an assumption regarding the core-satellite hypothesis. In this sense, an organism that is regionally core should dominate locally, which means that it should be persistent over time and preferentially also very abundant, so this organism would be capable to easily recolonize other sites in the region.

To test this assumption, we applied the Mitchell-Olds & Shaw test of quadratic extremes (Mitchell-Olds and Shaw, 1987) to determine if there is a bimodal distribution on the bacteria persistence combined with an abundance distribution chart. The bimodality test could determine if there are organisms capable to persist all over the recovered timespan here analyzed and the distribution chart complement this analysis showing if the persistent bacteria are also the most abundant ones. After this first assumption check, we could identify a set of organisms that fall over six categories concerning their persistence and abundance over time (Tab. 1). The persistence was determined by the number of times each ASV was found, while their abundance was determined by the rare-abundant biosphere proposal (Pedrós-Alió, 2006; 2012). In this case, the organisms that persisted over all samples and had a mean relative abundance greater than 1% were defined as "Core". The "Persistent" category encompasses the organisms that persisted over time, but not reached a great mean relative abundance, on the other hand, the not persistent organisms that showed a great mean relative abundance were categorized as "Abundant". Complementarily, the organisms that do not reach a great mean abundance, but were abundant in each least one month were categorized under the "Transient", and the organisms categorized as "Satellite" were those who do not reach 0.1% of mean

relative abundances in this dataset. All the others which do not fall under the previously described categories were called "Non-Abundant".

Table 1 – Persistence-abundance categories

		Persistent over time	
		Yes	No
Abundance	>0.1% (mean)	Core	Abundant
	>0.1% (at least once)	-	Transient
	<0.1% (mean)	Persistent	Non-abundant
	<0.01% (mean)	-	Satellite

To determine the relationship between the temporal variations of bacteria abundances and the environmental factors (Tab. S1) we applied the distance-based Redundancy Analysis (dbRDA) and the Mantel test. For both analyses, the biological data was transformed into distance matrices based on Bray-Curtis and the environmental data was standardized (unless for pH).

All statistical analyses described until now were performed using the R software (R Core Team, 2019) and functions provided by the vegan package (Oksanen *et al.*, 2016).

To further analyze the co-occurrence network of microbial communities across the dry and rainy seasons, we constructed a network using the maximal information coefficient analysis with the software MICtools (Albanese *et al.*, 2018). The ASVs were classified according to their abundance and persistence throughout the sampling campaigns as Core, Satellite, Abundant, Persistent, Non-abundant or Transient (Tab. 1). To identify ASVs associated with a given season based on specificity and fidelity, we used the Indicator Value (IndVal) (Dufrêne and Legendre, 1997) with the labdsv R package (Roberts e Roberts, 2016). Network visualization was performed by the software Cytoscape: https://cytoscape.org (Franz *et al.*, 2016).

Results

After rarefaction and filtering, we had 369,946 reads and 1,085 ASVs from 25 phyla in the 16S rRNA data, and 1'323,379 reads and 1,479 ASVs from 8 phyla in the 18S rRNA data. About the 16S rRNA genes, the phyla that showed greater mean relative abundances were Actinobacteria (40.16%), Proteobacteria (23.81%), Verrucomicrobiota (11.24%), Cyanobacteria (10.03%) and Bacteroidota (9.58%). Some clades worldwide found in freshwater lacustrine environments also stand out in this environment (Fig. 1A), they are the hgcI (Actinobacteria, 29.76%), Polynucleobacter (Gammaproteobacteria, 4.41%) and Limnohabitans (Gammaproteobacteria, 2.94%), FuKuN18 (Verrucomicrobiota, 5.34%) and Cyanobium (Cyanobacteria, 9.85%). Concerning the Eukaryotes (Fig. 1B), the Alveolata (37.76%) and Opisthokonta (26.41%) emerged as the most abundant phyla, which Ciliophora (17.65%) and Dinoflagellata (15.55%), Metazoa (21.95%) and Fungi (3.89%) being their most prominent Classes respectively. Alveolata, Metazoa and other picoeukaryotes were considered in further analyses to recover information about co-occurrence dynamics because they can explain bacteria abundance patterns. Despite their relevance in freshwater environments, we do not use Fungi data in further analyses because the 18S rRNA gene is not meaningful for this group, and their relative abundance is probably underestimated here.
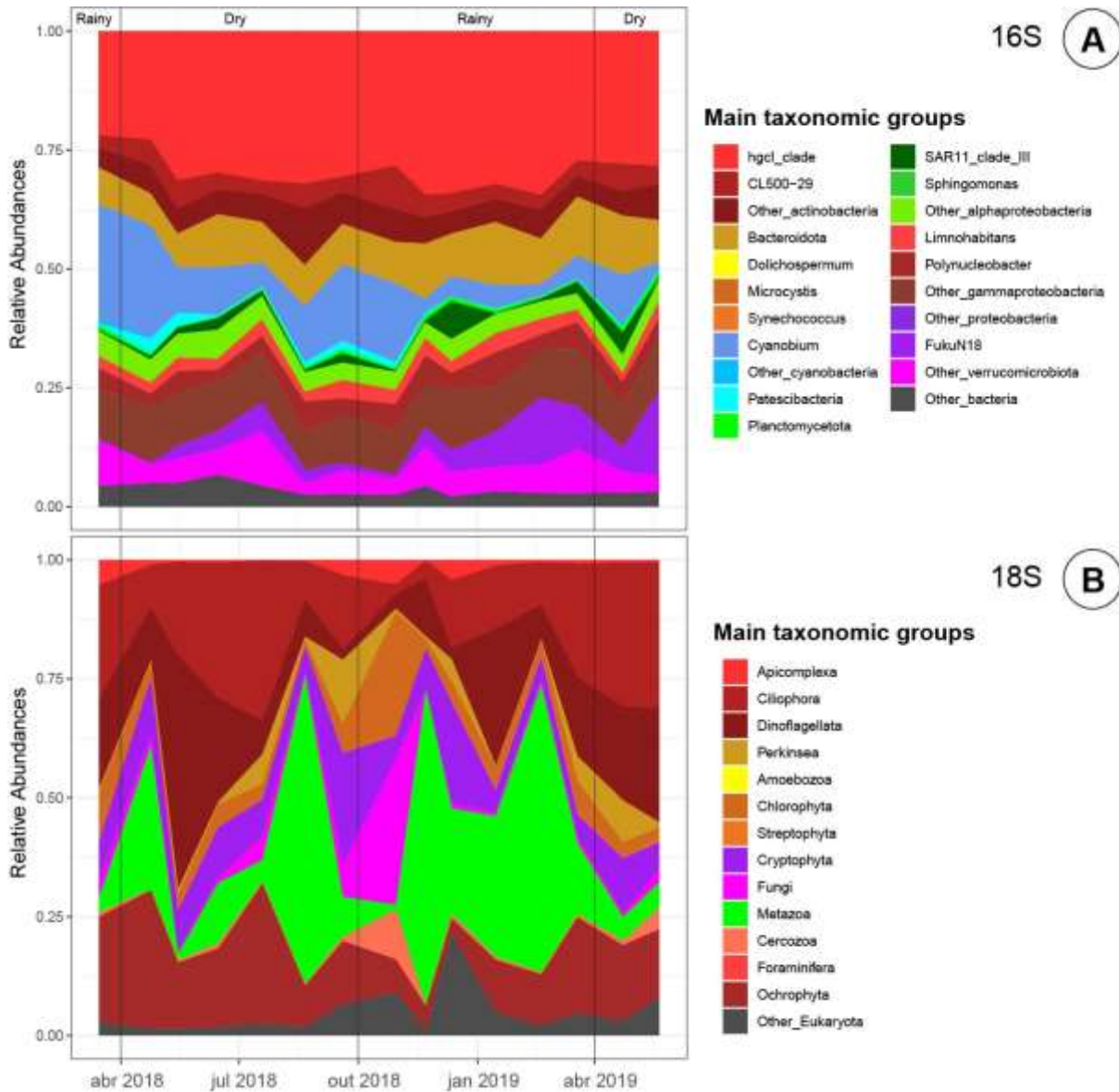
Figure 1 – Abundance thought time for Bacteria (A) and Micro-eukaryotes (B). Bacteria known as relevant in literature were also showed in this plot, it is the case of hgcI and CL500-29 clades (Actinobacteriota), Microcystis and Synechococcus (Cyanobacteria), SAR 11, Sphingomonas, Limnohabitans and Polynucleobacter (Proteobacteria), FuKuN18 clade (Verrucomicrobiota). Other relevant phyla like Bacteroidetes and Planctomycetota were highlighted. For the eukaryotes, Ciliophora, Dinoflagellata, Chlorophyta, , Metazoa and Ochrophyta stand out.

Concerning the core-satellite hypothesis, we recovered a significant bimodal distribution ($p < 0.05$) for bacterial persistence (Fig. 2A) in our dataset. The most abundant organisms were also those who normally persisted during all timespan encompassed in this study (Fig. 2B). These results reinforce the idea that the core bacteria are those who easily recolonize the neighborhood, in a process called

rescue effect, and reinforce our assumptions to consider this site as a good model to understand potential local factors shaping the spatial bimodal feature.
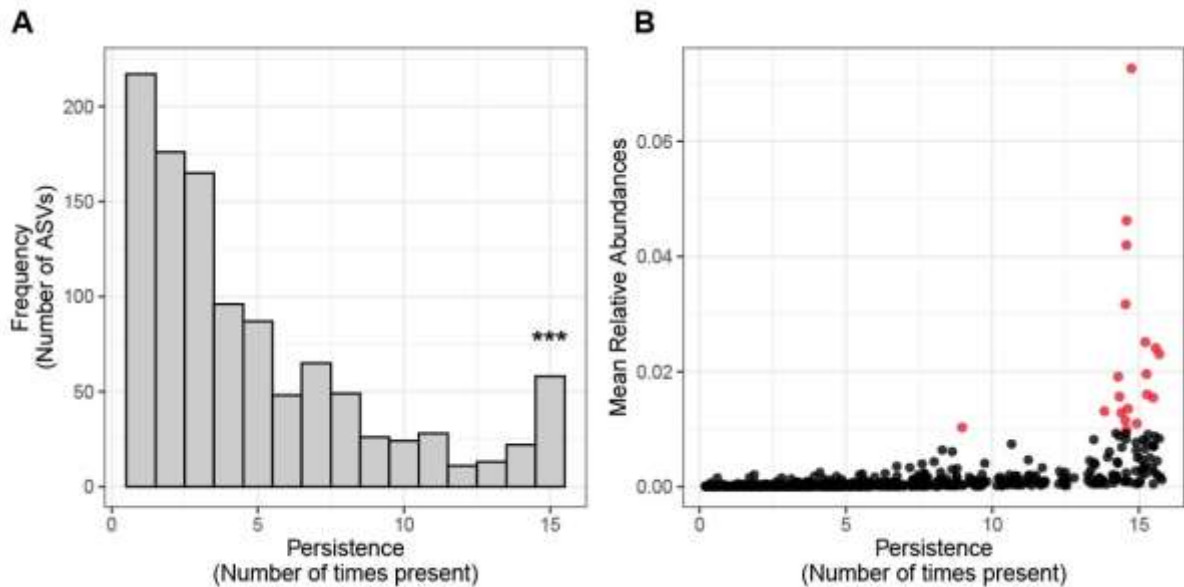
Figure 2 – A) Persistence-frequency distribution of bacteria thought time in one site sampled monthly 15 times. The asterisks (***) represent a significant result in the MOStest ($p < 0.05$), which means that there is a consistent hump in each extreme in this graph. B) Persistence-Abundance distribution plot, showing that the most abundant organisms are normally those who persist in this site; the point colored in red are those representing ASVs that showed e mean relative abundance greater than 1%

The dbRDA (Fig. 4) and Mantel test (Tab. 2) analyses showed that SR, FI, C:A ratio, Euphotic Zone, and the Trophic State are related to bacteria diversity. Temperature and Dissolved Oxygen also presented relevance in the dbRDA, despite being less closely related to the samples than the others previously cited. Concerning the DOM measurements, FI and C:A ratio presented values lower than 1.32 and 0.6 respectively, which indicates that the DOM originated from the surrounding soils. Also, FR was never greater than 0.7, which indicates an old DOM content, while HIX was ever greater than 1, which indicates that the DOM is mainly composed of humic acids. Finally, SR showed values around 1, which represent compounds of moderate molecular height. About the other environmental samples that showed a significant relationship with microbial diversity, the euphotic zone had a mean value of 1.129 m, while the trophic state presented a mean value of 52.6 (Oligotrophic as stated by Cunha *et al.*, 2013). The major part of samples in the dry seasons sound to be

related to the higher values of C:A ratio, FR, dissolved oxygen, and euphotic zone, while the rainy season samples were mainly related to HIX, temperature, and trophic state. Despite being classified as a rainy sample, BroaMO #1 was more closely related to other samples from the dryer season, on the other hand, samples #14 and #15 were closely related to rainy samples despite being from a dryer season. The dbRDA also showed that the variation in environmental samples was not significantly related to the Core nor the Abundant bacteria, but showed a relationship with the Persistent and Transient ones, in a pattern like the one shown by the entire metacommunity.

The Mantel results indicated that the Persistent bacteria presented a greater number of environmental factors in which it was closely related (i.e. SR, C:A ratio, and euphotic zone; Tab. 2), on the other hand, the Abundant bacteria presented fewer number of relationships (i.e. trophic state, Tab. 2). The temperature range was related only with the Non-abundant bacteria, on the other hand, the transient bacteria do not show the greatest relationship which any environmental factor. The Eukaryotes showed closer relationships with SR, FI, and the euphotic zone.

Figure 3 – dbRDA results showing the relationship between each sample and measured environmental factors (A). Yellow circles indicate samples collected in the dry season while blue circles indicate samples collected in the rainy season. The decomposition between bacterial groups showed a non-significant relationship for "Core" (B) and "Abundant" (C) (p > 0.05), otherwise, the relationship was significant (p < 0.05) for the "Persistent" (D) and "Transient" (E) groups. Dry season samples are directly related to the C:A ratio and Dissolved Oxygen, while the rainy samples were directly related to the Trophic State (TSI), SR, euphotic zone, and temperature. This pattern remains similar for the Persistent and Transient organisms

Table 2 – Mantel test results between bacteria relative abundances and environmental variables. Only interaction strength values for significant results (p < 0.05) were shown. Temperature, pH and Dissolved Oxygen were obtained *in situ* using a multiparameter probe; DOC and DOM were measured by different laboratory equipment. The Slope Ratio (SR), Fluorescence Index (FI), Freshness Index (FR), Humification Index (HIX) and C:A Ratio were derived from fluorescent measurements of DOM at distinct excitation/emission values (nm). The Euphotic Zone was calculated from Sechhi disk measurement obtained *in situ* and the Trophic State values were calculated from chlorophyll *a* concentrations (mg/L).

	Tempera-ture (°C)	pH	Dissolved O$_2$ (mg/L)	DOC (mg/L)	SR	FI	FR	HIX	C:A ratio	Euphotic Zone (m)	Trophic state
Bacteria (16S)	n.s.	n.s.	n.s.	n.s.	0.228	0.262	n.s.	n.s.	0.381	0.386	0.353
Core	n.s.	n.s.	n.s.	n.s.	0.183	0.319	n.s.	n.s.	0.172	0.229	0.179
Abundant	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	0.240
Persistent	n.s.	n.s.	n.s.	n.s.	0.250	n.s.	n.s.	n.s.	0.509	0.445	0.262
Transient	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	0.382	0.294	0.382
Non-Abundant	0.172	n.s.	n.s.	n.s.	0.170	0.157	n.s.	n.s.	0.509	0.239	0.401
Eukaryotes (18S)	n.s.	n.s.	n.s.	n.s.	0.444	0.417	n.s.	n.s.	n.s.	0.352	0.207

To examine the patterns of co-occurrence between bacteria, phytoplankton, and zooplankton, we applied co-occurrence networks with the temporal variation of relative abundances for these groups (Fig. 5). This analysis of co-occurrence only included bacteria identified as Core, Abundant, Persistent and Transient, while the entire biodiversity of phytoplankton and zooplankton was used. From 1,405 filtered ASVs (including bacteria, phytoplankton, and zooplankton), we could recover 463 nodes and 2,528 links between them.

In a general overview, we were able to identify eight ASVs that can be key in this community (Fig. 5). Four key ASVs were present in both seasons, These organisms were identified as hgcI clade (Bacteria, Actinobacteria), FuKuN18 clade (Bacteria, Verrucomicrobiota), Dinophyceae (phytoplankton) and Bacillariophyta (Phytoplankton). Concerning the organisms related to the dry season, the other two were prominent: Polynucleobacter (Bacteria, gammaproteobacteria) and a non-identified clade of Actinobacteria. For the rainy season, the other two key bacteria were Polynucleobacter (Gammaproteobacteria) and hgcI clade (Actinobacteria). All

these organisms were well linked to other ASVs, including several connections between each other. The clade hgcI is who presented a greater number of co-occurrences and was linked to other Actinobacteriota, Bacteroidota, Proteobacteria, Verrucomicrobiota, Cyanobacteria, and also Phytoplankton and Zooplankton. Cyanobium (Cyanobacteria) sounds to blooming in the early-dry season and FukuN18 (Verrucomicrobiota) in the late-rainy season. All other bacteria were well-distributed throughout the year (Fig. 5, Fig. 6).
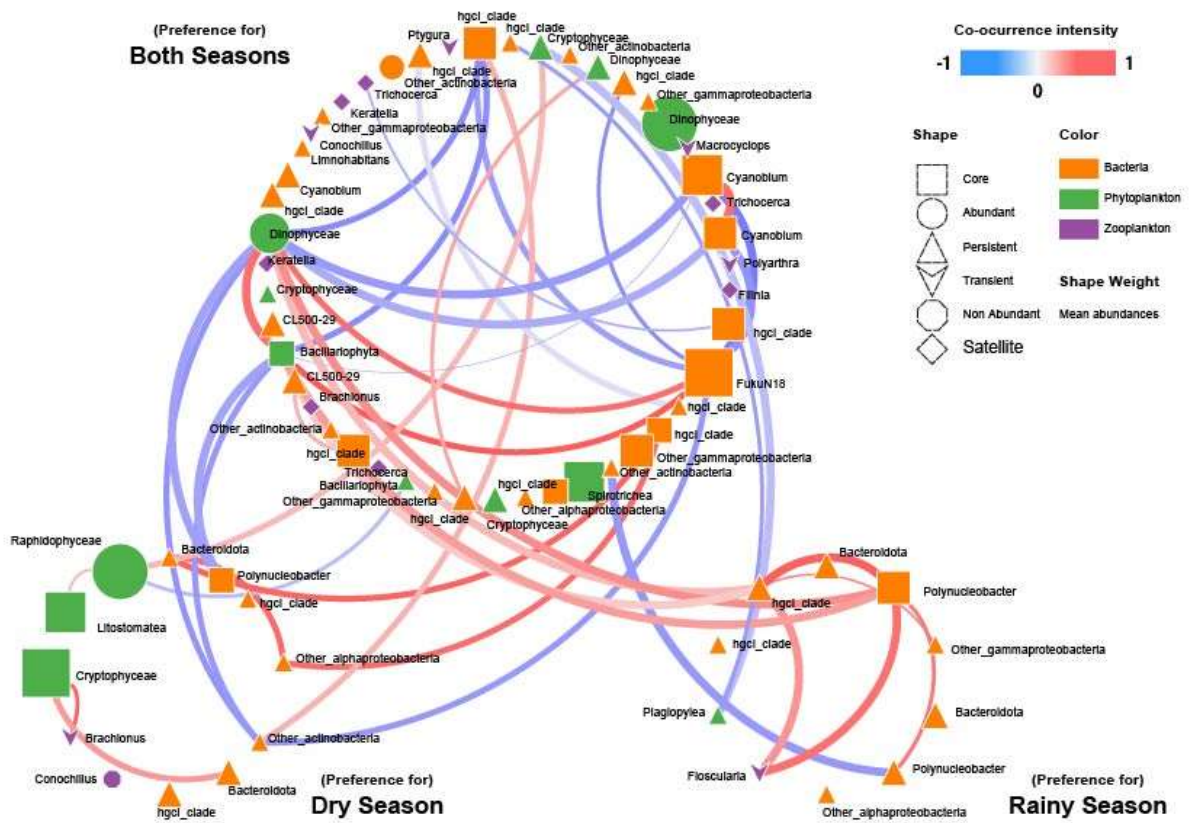
Figure 4 – Co-occurrence networks considering bacteria (orange), phytoplankton (green), and zooplankton (purple) taxonomic units as recovered by 16S and 18S rDNA sampling during a year. Blue lines indicate negative co-occurrences, while red lines indicate positive co-occurrences. Squared shapes indicate Core organisms, round shapes are the Abundant ones and triangles are the Persistent. Inverted triangles are the Transient, octagons represent Non-abundant ASVs and diamond shapes are for the Satellite. The shape weight is the mean abundance of each ASV
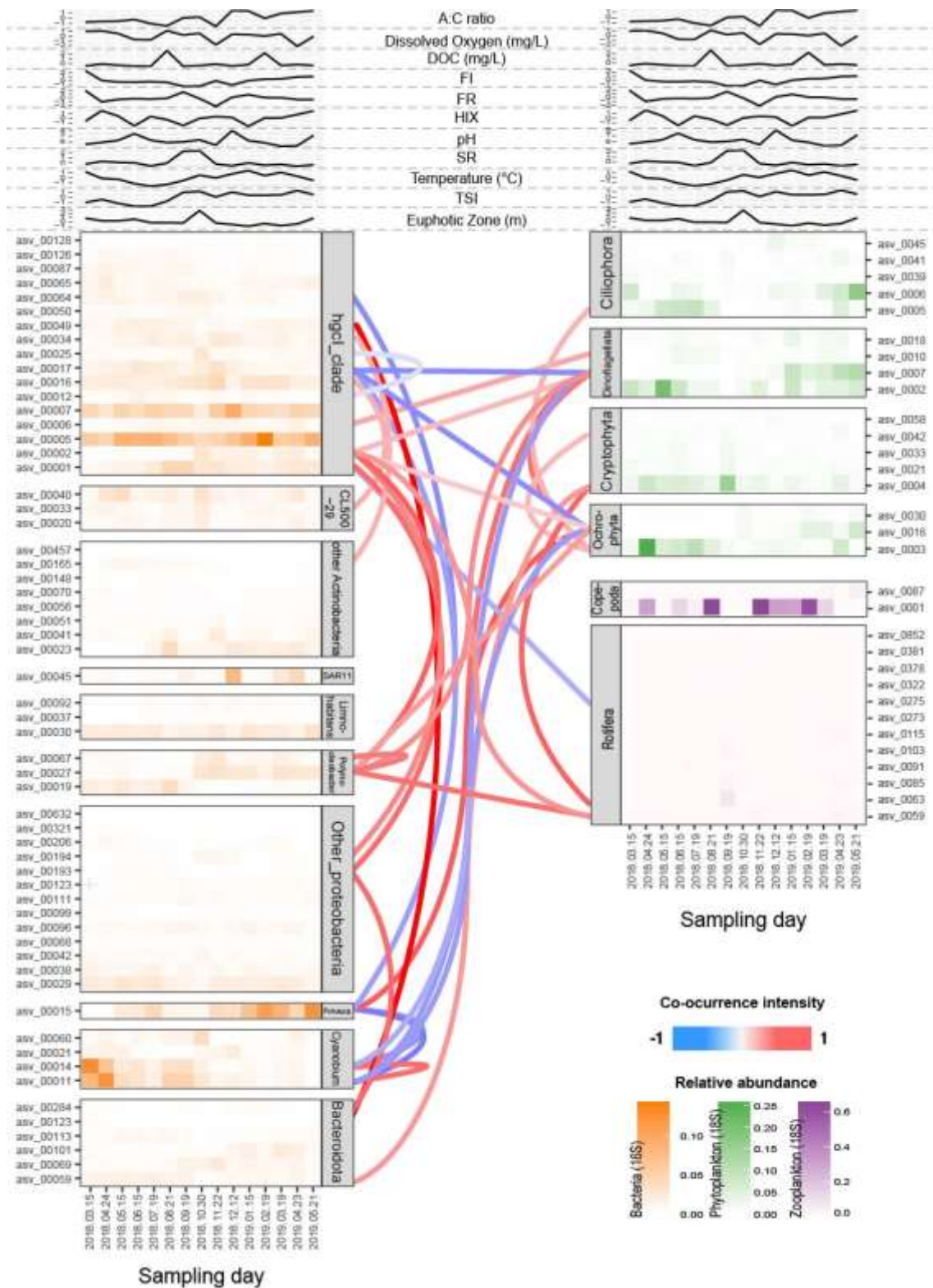
Figure 5 – Summary board showing environmental (upper board) and ASVs (bottom board) abundance variances through time for bacteria (orange), phytoplankton (green), and zooplankton (purple). Blue lines indicate negative co-occurrences and red lines indicate positive co-occurrences. ASVs assigned as hgcI clade showed a greater number of connections with other groups and showed a consistent abundance variation over time. Cyanobium and Dinoflagelatta ASVs showed an increment in abundance in the early dry season, while FuKuN18 and Ciliophora ASVs showed an increased abundance in the late rainy season

The negative co-occurrences were found for organisms presenting distinct preferences for dry and rainy seasons ($p > 0.05$), these connections were expected and removed from figure 5. For the remaining negative relationships, a deeper observation (shown in Fig. S1) indicates that it occurs as a result of less pronounced preferences for one or other seasons.

Zooplankton co-occurs with rarer organisms, mainly phytoplankton (Fig. S2). An outstanding exception is ASV #0059, classified as Floscularia, it is a small filterer organism that co-occurred negatively with ASV #0002, a Core hgcl clade, and some other transient bacteria that belonged to distinct groups like cyanobacteria, bacteroidota, and proteobacteria.

Discussion

The capacity to maintain large local abundances is one of the most important aspects of spatial dominance for distinct species, as it allows an organism to recolonize neighbor sites (Brown, 1984; Papp and Izsák, 1997; Gaston *et al.*, 2000). Here, we investigated the local temporal factors that have the potential to guide freshwater bacteria abundances, and in this way allow these organisms to be spatially Core or not. We discriminated against the bacteria by their mean local abundances and annual persistence as Core, Abundant, Persistent, Transient, Non-abundant and Satellite and, based on these different classifications, we focused on finding clues that would lead us to recognize the different factors capable of forcing these organisms into each box.

Concerning the characteristics of these different classes of organisms, we could identify that the Core bacteria can adopt two strategies, some maintain a similar relative abundance all along the year, while others maintain a minor abundance and form blooms when the local conditions allow it. The first strategy is observed in the hgcl clade (also known as acl lineage), this clade groups Actinobacteria that adopted a streamlined lifestyle. The streamlined organisms are those that have a reduced genome (Ghai *et al.*, 2014) and present auxotrophy (Kim

et al., 2019). This means that they do not produce all proteins they need, and instead take advantage of molecules exuded by other bacteria and picoeukaryotes (Newton *et al.*, 2011; Chiriac, Haber*, et al.*, 2022). Despite being dependent on other organisms, a capacity to absorb Carbon and Nitrogen from distinct sources, including a facultative quimiotrophy, is notable (Findlay *et al.*, 2003; Ghai *et al.*, 2014), and a closer relationship to a group was never observed (Chiriac, Haber*, et al.*, 2022). All these characteristics help to explain their persistence and great abundance observed. The other strategy showed by Core organisms has representatives of distinct phyla. ASVs concerning clades such as *Polynucleobacter* (phylum Proteobacteria), FuKuN18 (phylum Verrucomicrobiota), and *Cyanobium gracile* (phylum Cyanobacteria) showed a great abundance increment in distinct moments and a clear preference for the dry or rainy season. All these organisms have in common the relationship with specific environmental conditions (Lindström *et al.*, 2005; Wu *et al.*, 2006) and can be associated to organic particles (Parveen *et al.*, 2013) or are related to algal blooms that allow them to grow (Eiler and Bertilsson, 2004; Chiriac, Haber*, et al.*, 2022). An interesting characteristic of representatives of Polynucleobacter that also help to explain this pattern is that the group presents intense ecological niche partitions despite being hardly discernible into distinct variants (Hahn *et al.*, 2016; Hahn *et al.*, 2021).

In contrast to the Core, the Persistent bacteria showed a clear relationship with environmental factors, in special those related to a relevant input of nutrients from soil into this reservoir. This group also presented ASVs classified at the clades previously described for the Core, which are known as typically inhabiting the water column in lakes (Chiriac, Haber*, et al.*, 2022)*,* but this group also has other representatives. Some of them are associated with water column habitats with presenting lesser dominance compared to the Core (Salcher *et al.*, 2011; Camara Dos Reis *et al.*, 2019; De Melo, Bertilsson*, et al.*, 2019; Mateus-Barros *et al.*, 2021). Also, there are bacteria adapted to degrade more recalcitrant organic matter of terrestrial origin or methyl compounds (Chiriac, Haber*, et al.*, 2022). These characteristics explain the relationship of this group with DOM measurements that indicated a great input of soil humic matter in this community and pinpoint a huge distinction to the Core. These organisms spread and maintain a notable presence at

sites with similar characteristics in this landscape (Mateus-Barros *et al.*, 2021), but they would probably be filtered if a sample was taken in a site with a lesser connection to the surrounding soil, as a greater lake or river mouth, in contrast to the Core, which is formed by organisms also found in these types of sites (Ruiz-González *et al.*, 2015). Another hypothesis that can help to explain this group is a strategy to be rarer. This pattern was already found in marine environments (Lindh *et al.*, 2017), and may be interpreted by both, a strategy to avoid or a susceptibility to predation/grazing by zooplankton and heterotrophic nanoflagelates. Also, this highlight that in some contexts, the greater dispersal capacity of bacteria may overcome the necessity of being very abundant to colonize an entire landscape (Soininen *et al.*, 2018).

Similarly to the Core, the Abundant bacterium (ASV #23, Ca. Planktophila), do not showed a relationship with the environmental factors, but also not seems to co-occur with the other relevant organisms. As we do not found a relation of this group with others, it is hard to identify a potential role for it, but this pattern may be explained by a good adaptation to some environmental characteristic(s) of this site that overpasses the seasons, and it is capable to disperse but is quickly filtered where this characteristic is not present. This may be explained by other variables that we did not measured, however, in this case, other organisms, in special gammaproteobacteria (Newton *et al.*, 2011; Chiriac, Haber*, et al.*, 2022) should also be in this group.

The variation in bacteria biodiversity was also evidenced in this study. The changing of the season from dry to rainy implied an increment in DOM from the soil, DOC concentrations, temperature, and trophic state, and a reduction of the euphotic zone, with a consequent increment in the nutrient availability for bacteria. The co-occurrences network showed that in general, negative co-occurrences can be related to specific preferences for dry or rainy seasons. This was expected when comparing the organisms that showed a significant correlation for one or other season (and because of that, these links were removed from figure 5), but a deeper analysis, which can be found in Fig. S1, showed that this relationship remains true for the ASVs that did not show a significant correlation with the season. Some of the most abundant organisms showed a prominent variation in their abundances in a timespan

that ever belongs to 2-3 months, as a bloom-forming dynamic. The blooms occurred during the early dry and late rainy seasons and were accompanied by the increment in the phytoplanktonic abundance concerning the phyla Dinoflagellata and Ochrophyta. The first involved Cyanobium ASVs, while the second had a representative of FuKuN18. An interesting pattern is that the early dry bloom ended before the phytoplankton becomes more abundant, but the late rainy bloom occurred concomitantly with the algal blooms. This can be relevant because these two bacteria are known by being related to phytoplankton blooms (Chiriac, Haber, *et al.*, 2022). While the Cyanobacteria are also known as bloom-forming organisms, and can perform a chemical controlling of phytoplanktonic blooms (Kovács *et al.*, 2018), the Verrucomicrobiota normally are associated to phytoplankton by using compounds exuded during bloom events (Eiler and Bertilsson, 2004).

Zooplankton co-occurs more with Transient bacteria and Satellite phytoplankton. Several bacteria have a reduced cell size associated with a streamlined lifestyle and an adaptation that is useful to avoid predation (Chiriac, Haber, *et al.*, 2022). It is, for example, the case of hgcI clade (Neuenschwander *et al.*, 2018). On the other hand, several other groups have a relatively longer genome size and, as a consequence, bigger cell size. These organisms are more sensitive to predation and grazing made by zooplankton and heterotrophic nanoflagellates (Parveen *et al.*, 2013; Piwosz *et al.*, 2018), and are hypothesized to be the link between microbial interaction networks and classical food chains (Piwosz *et al.*, 2018). Here, we could not identify a great number of co-occurrences between zooplankton and bacteria, but this process can potentially be very relevant in natural communities, in special for tropical environments, in which biological interactions should play a greater role in guiding microbial communities (Sarmento, 2012). Experimental procedures demonstrated that it can increase the degree of unexplained variations over time (Livingston *et al.*, 2017) and can be related to the maintaining evenness in the community (Segovia *et al.*, 2018). In this sense, this was an aspect less consistent than expected and may occur due to a lower accuracy of these approaches applied here to detect these patterns.

Conclusions

The core-satellite hypothesis is an interesting and simple alternative to recovering information about the spatial distribution of bacteria. This approach has been used to recover details on microbes in terrestrial, freshwater, and marine realms (Lindh *et al.*, 2017; Mateus-Barros *et al.*, 2021; Escalas *et al.*, 2022). Here, we investigated the temporal variation in local environmental and biological aspects that can potentially contribute to guiding bacteria to be part of the Core pool of organisms. Our findings highlight the importance of organic compounds to determine dominance patterns in shallow lake environments. On one hand, the Core is basically composed of organisms typical of freshwater environments and described as capable to grow using compounds exuded by other bacteria or phytoplankton, on the other hand, Persistent bacteria are closely related to the input of recalcitrant humic DOM from the surrounding soils.

These patterns indicate the relevance of bottom-up processes guiding the selection of bacteria that will be capable to persist in these locations highly impacted by matter carried from the soil. As these organisms are incredible dispersers (Soininen *et al.*, 2018), it implies that they will also persist spatially (Mateus-Barros *et al.*, 2021). In contrast, the top-down processes seem to have a reduced selective relevance, but there is some indication that predation plays a role in maintaining the rarer organisms rarer (Segovia *et al.*, 2018).

Capítulo V -
Beyond environmental selection: Spatial structuring of
tropical lake bacterioplankton metacommunity

Determinar a importância relativa dos fatores ambientais e de dispersão que estruturam a abundância e a composição das espécies é uma questão central na ecologia de comunidades. Isso é especialmente verdadeiro para grandes dispersores como microrganismos, que ainda carecem de uma avaliação mais abrangente (ou seja, táxons abundantes e raros). Aqui, avaliamos a diversidade bacteriana com sequenciamento de amplicon 16S rRNA e aplicamos uma abordagem estatística mais completa por meio de particionamento de variação para recuperar a influência de fatores ambientais e de dispersão na diversidade beta (βdiv) em uma metacomunidade bacterioplanctônica composta por 60 lagos rasos em regiões de cabeceira distribuídos por uma ampla paisagem (~250.000 km2). Em geral, βdiv foi melhor explicado pela filtragem ambiental do que pela dispersão. No entanto, diferentes fatores espaciais foram importantes para determinar a metacomunidade bacterioplanctônica, dependendo se as métricas βdiv deram mais importância a espécies raras (dados qualitativos) ou abundantes (dados quantitativos). Para dados quantitativos, a maior parte da variação espacial βdiv foi explicada pelo isolamento geográfico (diferentes bacias hidrográficas), enquanto a vizinhança (aumento da conexão entre locais mais próximos) e as distâncias terrestres explicaram melhor os dados qualitativos. Sugerimos que, embora os fatores determinísticos permaneçam primordiais para direcionar a biodiversidade do bacterioplâncton de água doce, os contextos espaciais também podem ajudar a explicar a variação do bacterioplâncton resultando em alta dissimilaridade e elucidar como alguns organismos podem acabar dominando ou sendo dominados nessas comunidades

Introduction

Understanding the relative role of local and regional processes in space-related community properties has received growing attention over the last decades. In this context, the metacommunity concept has emerged as an ecological synthesis that states that a set of local communities are connected by dispersal of multiple potentially interacting species (Wilson, 1992; Leibold *et al.*, 2004). In this case, local determinants act together with stochastic factors at major or minor intensity to determine the regional biodiversity in a process also related to regional dispersion (Lowe e Mcpeek, 2014; Vellend *et al.*, 2014). Together, these factors shape the spatial patterns of species' compositions in different ways depending on the spatial scale (Vellend, 2010) and organisms' intrinsic features (De Bie *et al.*, 2012; Soininen *et al.*, 2018), modulating the observed beta-diversity (βdiv), which is the extent of dissimilarity in community composition between a pair of local communities (Whittaker, 1960; Whittaker, 1972).

The βdiv depends on intrinsic and extrinsic factors such as specific dispersal capacity, trophic level, variation of local factors and interactions, and the realm inhabited (Leibold *et al.*, 2004; Soininen *et al.*, 2018). For example, βdiv is lower in marine systems compared to freshwater and terrestrial systems (Drakare *et al.*, 2006) and higher in tropical regions (Soininen *et al.*, 2018). Geographic factors are also an important aspect to be considered as they impact a community both directly and indirectly since environmental heterogeneity and dispersal probability may change with distance (Martiny *et al.*, 2011; Soininen *et al.*, 2018). In microorganisms, βdiv is generally high as the dissimilarity observed in distinct realms all around the world is about 60–70%, mainly due to the replacement of species (i.e., spatial turnover) (Soininen *et al.*, 2018). Recent studies that assessed the joint structure of macro- and microorganism communities have shown the importance of body size in mediating the relative importance of local and spatial processes on species distribution and abundance, and showed that the spatial community structure of microorganisms should be mainly governed by stochasticity (Farjalla *et al.*, 2012; Soininen *et al.*, 2013). On the other hand, some previous studies have challenged this general view and highlighted a higher relative importance of environmental filters shaping bacterial community composition (Beisner *et al.*, 2006; Winter *et al.*, 2013;

Jyrkänkallio-Mikkola *et al.*, 2017; Fillinger *et al.*, 2019). However, these studies had some limitations such as using fingerprinting (e.g. DGGE), a molecular method to assess prokaryotic diversity that recovers only the most abundant taxa, or using a simplistic mathematical approach that covered only one or a few geographic aspects acting on these communities (e.g. overland linear distance).

In this context, as bacteria have high diversity and respond differently to geographic distances (Mateus-Barros *et al.*, 2021) or environmental heterogeneity (Huber *et al.*, 2020) and can be found in larger numbers than any other organism, this group has emerged as key models to study metacommunity structure (Barberán *et al.*, 2014) and to test βdiv patterns in distinct ecosystems and environmental contexts (e.g. Sommaruga and Casamayor, 2009; Caruso *et al.*, 2011; Jyrkänkallio-Mikkola *et al.*, 2017; Fillinger *et al.*, 2019). Despite this, little is still known about the processes that shape microbial metacommunities compared to macroorganisms (Soininen, 2012).

Distinct methods have already been proposed to assess and interpret βdiv. Some consider presence/absence data (qualitative approach) and others also incorporate the importance of species relative abundance (quantitative approach) (Anderson *et al.*, 2011). The first provides insights into compositional variation and gives greater weight to the importance of rare species, while the second considers the species abundance, an important aspect of metacommunity structure (Anderson *et al.*, 2011). Abundance may be a key characteristic to be considered because there is a relationship to the dispersal capacity of an organism. The site where a species is more abundant serves as a dispersal source and their range of distribution seems to be related to the maximum abundance it reaches locally (Brown, 1984), which means that the organisms with greater adaptive capacities tends to become abundant and end up dominating the landscape (Gaston *et al.*, 2000), while rare organisms should present a clustered distribution and a limited dispersal capacity.

Recent studies applying and exploring newer genomic sequencing methods to determine the taxonomic identification of microorganisms (e.g. Callahan *et al.*, 2016) reported that microbial biodiversity is at least two orders of magnitude higher than previously believed. This is caused by the identification of very uncommon taxonomic

units in the communities where they are found, which is the so-called rare biosphere (Pedrós-Alió, 2012). The identification of this rare biosphere certainly affects the way we interpret βdiv in the sense that the qualitative approach gives equal weight to organisms with distinct abundances, which tends to increase the importance of the rarer ones, while the quantitative approach gives more importance to the variation in abundance of those more common species (Anderson *et al.*, 2011), and this dichotomy may be an important aspect for determining the observed βdiv (Melo *et al.*, 2012).

Bacteria are passive dispersers, which means that they move by attaching to larger organisms (Grossart *et al.*, 2010) or by being carried by water flows (Lansac-Tôha *et al.*, 2020) and wind (Smith *et al.*, 2013). In this context, some theoretical approaches can be used to determine the main way dispersal impacts bacterial βdiv. 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 βdiv 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.*, 2008b).

In this study, we aimed to evaluate how environment and dispersal can contribute to explain βdiv. 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 tropical shallow lakes scattered over a matrix of nearly 250,000 km^2. We addressed this objective using new generation sequencing (NGS) to recover highly detailed taxonomic units and applying two different approaches (qualitative and quantitative) to observe the impacts of those on the observed βdiv. We also recovered the impact of local environmental factors and time on this metacommunity. We predicted that the relative impact of the different dispersal models would depend

on the approach used. For the quantitative approach, we expected to see a βdiv mainly associated with environmental factors and a linear decay on connectivity between sites; on the other hand, for the qualitative approach we expected to see a βdiv mainly associated with an increased decay on connectivity between sites associated with a greater connection between neighbour sites and isolation promoted by geographic barriers in distinct drainage basins.

Methods

Study Design

This study was performed using a dataset obtained from 60 headwater shallow lakes that cover a region of nearly 250,000 km^2 in São Paulo state, southeast Brazil (Fig. 7). This region has a tropical climate and is comprised of Cerrado vegetation (Brazilian savannah) and Atlantic Forest (semi-deciduous humid forest). These lakes are generally small reservoirs, dammed to fulfil water necessities of landowners. The samplings were carried out between June 2012 and July 2016. In the field, we measured environmental variables (temperature, conductivity, pH) using a multi-parameter probe (YSI, Yellow Springs, USA) and filtered sub-surface water for laboratory analyses (nutrients, carbon supply, chlorophyll-*a*, e-DNA). The altitude and location were obtained using a GPS. Samples for nutrient analyses were obtained by filtering lake water through 0.45 µm polycarbonate membranes previously washed with ultrapure water to prevent carbon contamination caused by the filter and stored in amber bottles in a freezer at -20 °C until analysis. Then, a TOC-V (Shimadzu®, Kyoto, Japan) was used to obtain dissolved organic carbon (DOC) and Dissolved Inorganic Carbon (DIC). Dissolved inorganic nitrogen (DIN) was calculated by summing the values obtained from nitrite, nitrate, and ammonium recovered by an Ion Chromatography System (Thermo Scientific®, Waltham, Massachusetts, USA). Moreover, a FS5 Spectrofluorometer (Edinburgh Instruments®, Livingston, UK) was used to estimate the Tryptophan-like fluorescent dissolved organic matter (T-FDOM), calculated by the ratio between dissolved organic matter fluorescence and quinine sulfate (0.001 mg/L dissolved in 0.1 M

H2SO4) at 455 nm excitation and 355 nm emission. Chlorophyll-*a* concentration (a proxy of trophic state) was obtained by filtering 100–500 ml of water through a glass fibre filter (Macherey-Nagel® GF-6), extracted with ethanol (90% v/v at 80 °C) in the dark (Marker, 1980; Mush, 1980), and quantified by spectrophotometry (Lorenzen, 1967). In this dataset, we have equal sample sizes across trophic state categories (oligotrophic, mesotrophic, and eutrophic) and hydrologic basins. More details about the study site description and environmental variable analyses can be found in Mateus-Barros *et al.* (2021).
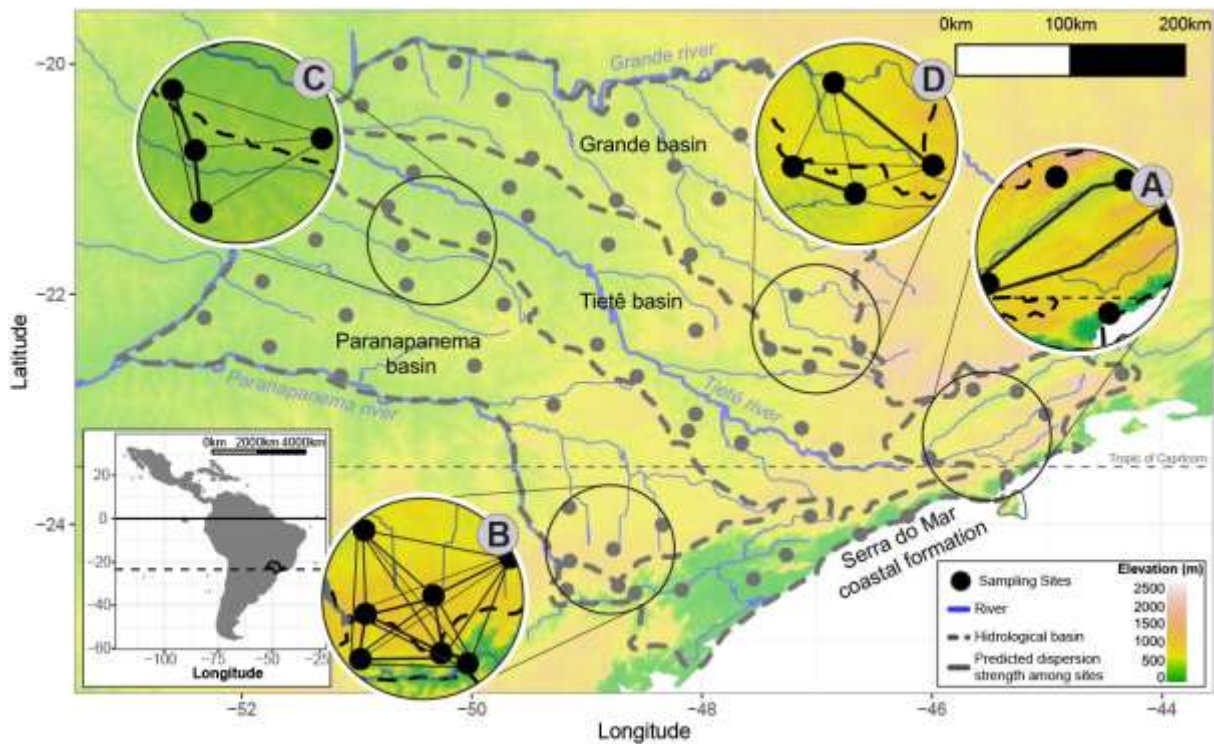
Figure 6 – Location of the sixty headwater shallow lakes sampled scattered over a large tropical landscape that covered four distinct hydrological basins. The main rivers (blue lines) and basins delimitation (black dashed lines) were also indicated. The large zoomed figures illustrate how each spatial aspect was considered to the variation partitioning: if the main form of dispersion is through the river flow to other lake, then the connected lakes should present similar composition (A); if any geographical barrier is capable of stopping the bacterial dispersal, then only the overland distances should be considered and they have equal chances to reach each site (B); if the increasing distances make this dispersal increasingly hard, then the chance of one individual reach neighbour sites is greater than reaching distant ones (C); finally, if a geographical barrier can stop the bacterial dispersal, then it is easier to reach the sites at the same hydrological region even when a pair of sites is spatially more distant than a neighbour site located at the adjacent hydrological region (D). The line thickness in the examples represents a hypothetical connectivity strength between the pairs of sites under the four dispersal possibilities described above

Molecular *Analyses and Bioinformatics*

A full description of the molecular analyses is provided in Mateus-Barros *et al.* (2019; 2021). In summary, 500 ml of water was pre-filtered through a glass-fibre filter with 1.2 µm mesh (BOECO® MGC) to retain eukaryotes, large particles, and attached prokaryotes. Afterwards, 200–500 ml of filtrate was passed through 0.2 µm polycarbonate membranes (Millipore® Isopore™ 0.2 µm GTBP) to retain free-living prokaryotes. Free-living bacterial DNA was extracted from half of the filter using a phenol-chloroform extraction. Using a KAPA HiFi HotStart ReadyMix PCR Kit (Kapa Biosystems®), the V3–V4 regions of 16S rRNA were amplified with the 341F (5′-CCTACGGGNGGCWGCAG-3′) and 805R (5′-GACTACHVGGGTATCTAATCC-3′) primers (Herlemann *et al.*, 2011). After amplification, fragments were sequenced in an Illumina MiSeq platform. Then, the software R (R Core Team, 2019) was used to process the data with the DADA2 pipeline (Callahan *et al.*, 2016) and taxonomy assignment. The taxonomy was assigned by blasting against the SILVA version 132 database (Quast *et al.*, 2013; Yilmaz *et al.*, 2014) to obtain an Amplicon Sequence Variants (ASVs) table. The 16S rRNA amplicon results have been deposited in the NCBI repository under accession number PRJNA411849.

Prior to the statistical analyses, the ASVs table passed through some filtering steps to guarantee the quality of final analyses. First, sequences identified as mitochondria, chloroplasts, and archaea were removed, then the table was rarefied by the lowest sample richness (14,239 reads), and finally, the ASVs that did not sum 10 reads in all sampling sites were removed.

Data analyses

To determine the roles of species' presence and abundance on the ecological aspects investigated here, the first step was to transform the ASVs table to recover two different tables: (1) quantitative data table with relative-abundance data, which was used to perform analyses based on Bray-Curtis distances and (2) qualitative data by using ASV presence/absence at each site to perform 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 lines and never recover the complete number of reads contained in a sample, and because of that the results have a compositional nature (Gloor *et al.*, 2017) and should be always treated as relative instead of absolute. In any case, the variation that each ASV presents per site is valuable to the interpretation of ecological processes. All forthcoming analyses (see below) were performed on both quantitative and qualitative data matrices.

To determine the importance of environmental and dispersal factors in shaping the observed βdiv in this bacterioplanktonic metacommunity, we applied variation partitioning (Borcard *et al.*, 1992). In this case, we used dissimilarity matrices obtained from the *beta.pair* and *beta.pair.abund* functions for qualitative and quantitative data, respectively. These functions are part of the *betapart* package (Baselga, 2010; Baselga and Freckleton, 2013) and applied a distance-based redundancy analysis (dbRDA) to determine the factors used in the variation partitioning (Legendre e Anderson, 1999). For the dispersal features, we used eigenvector analyses to try to recover the ways bacteria can disperse through the landscape. Here, we considered four main dispersal possibilities. First, the possibility of a connection between sites promoted by a river flow (Fig. 7A); in this case, we used an AEM (Asymmetric Eingenvector Map) model. In this model, we provided a weight matrix considering the river segments connecting two sites; the edges that reflect a water flux to the observed site were assigned a value of "1" while "0" was assigned to the other edges not related to the site (Blanchet *et al.*, 2008b). The resultant eigenvector matrix was called connectivity fraction (Fig. S3A). Second, to observe the linear decay of similarity and obtain what we called the space fraction (Fig. 7B, Fig. S3B), we used a dbMEM (distance-based Moran Eigenvector Map) approach to produce eigenvectors based on the overland distances between sites (Dray *et al.*, 2006). Here, the connection between sites is linearly related to the increment in distance. Third, complementary to the previous model, the neighborhood fraction (Fig. 7C, Fig. S3C) was estimated with a MEM analysis (Peres-Neto e Legendre, 2010) and a weight table was calculated by the formula *1 - x/max(geo)*. In contrast to the space fraction, this weight formula modifies the connection/distance relationship to a binomial-like shape that increments the

importance of shorter distance connections. The last was the region fraction (Fig. 7D, Fig. S3D) that also used a MEM approach, but in this case the weight table was based on drainage basins. We assigned a value of "1" to the site-pairs located in the same basin and "0" to the pairs located in distinct regions. In this case, the eigenvectors reflected the isolation promoted by these geological formations. All these eigenvectors were recovered by the *adespatial* package (Dray *et al.*, 2020). Afterwards, all matrices were submitted to an ordination based on 999 permutations to select the relevant eigenvectors. In the ordination step, we did not recover any relevant eigenvector constructed by the connectivity fraction, which was expected as we have only headwater shallow lakes. In this stage the forward selection approach was used to avoid the spatial autocorrelation artifact that arises from the classical ordination and overestimates the impact of spatial factors on the metacommunity (Blanchet *et al.*, 2008a; Bauman *et al.*, 2018). This approach adds an r-squared calculation to the classical ordination; in this sense, an r-square is calculated before the ordination and after the selection of each eigenvector. When the r-square calculated for the eigenvectors overcomes the r-square calculated for the global analysis the ordination is stopped and, even if other eigenvectors present significant correlation with the biological data, they are excluded from the downstream analyses (Fig. S4 shows all eigenvectors recovered). The ordination test was conducted with the *ordiR2step* function of the *vegan* package (Oksanen *et al.*, 2016). For the environment fraction (Fig. S3F), the environmental factors were standardized (except pH) and forward selection was applied to select the relevant components. Complementarily, we obtained the time fraction by transforming the collection dates to Julian days; in this way, the value "1" was attributed to day 01 of January of the first collection year (2012) and the number of days to each sampling was added to it (Fig. S3G).

Each geographic, environmental, and time fraction were compared considering biological dissimilarities through variation partitioning (Fig. S3H). This analysis was done as a distance-based Redundancy Analysis (dbRDA) to compare their contribution in the quantitative and qualitative βdiv (Fig. S3I). After the partition, we performed a canonical correspondence analysis (CCA) to check fraction significance (Fig. S3J). As a complement, we also compared the geographic fractions

between them to see which is more relevant to βdiv and sub-grouped the environment fraction between the pH and the other relevant components. This was done because pH is consistently reported as the most relevant factor shaping freshwater bacterioplankton (Lindström *et al.*, 2005; Mateus-Barros *et al.*, 2021). The connectivity fraction was not used in this analysis as no relevant eigenvector was recovered. These analyses were conducted with vegan package (Oksanen *et al.*, 2016).

Complementarily, to measure the degree of βdiv and determine the impact of abundance on this metacommunity, we performed a Raup-Crick dissimilarity analysis, which is different from classic analyses because it controls the local richness of βdiv. The analysis uses all the species found to construct a regional pool and calculate a random distribution of species; then, each paired dissimilarity is compared with the random result to measure whether the site-pairs are more dissimilar or similar than expected by chance (Anderson *et al.*, 2011). The resultant values fall between 1 and -1, which means that βdiv values that are close to the random expectation receive values closer to 0, while positive values represent a greater βdiv than expected by chance and negative values a greater similarity. We used two different approaches, one to lead with the qualitative data (Chase *et al.*, 2011) and the other for the quantitative data (Stegen *et al.*, 2013).

The R scripts used in these analyses can be found at https://github.com/LMPB/Variation-Partitioning.

Results

After rarefaction and filtering, 3,738 ASVs were obtained from 815,560 reads and classified within 22 phyla. The phylum Proteobacteria was the most represented (30% of the total number of ASVs); other well-represented phyla were Actinobacteria (26.64%) and Patescibacteria (20.47%). Concerning the sum of reads, we found that 54.20% were Actinobacteria, 24.37% Proteobacteria, and 5.60% Verrucomicrobia. Before variation partitioning, the ordination of significant environmental variables showed that for the qualitative dataset (0.084 adj. R^2) βdiv was related to pH (0.021

adj. R²; p = 0.002), DOC (0.020 adj. R²; p = 0.002), altitude (0.014 adj. R²; p = 0.016), T-FDOM (0.014 adj. R²; p = 0.016), and DIN (0.008 adj. R²; p = 0.04), while for the quantitative (0.131 adj. R²) there was relationship with pH (0.042 adj. R²; p = 0.002), DIC (0.019 adj. R²; p = 0.006), and altitude (0.018 adj. R²; p = 0.016). These variables were then selected, and the other tested variables (temperature, conductivity, and chl-*a*) were not used in any subsequent analysis. All the selected variables showed a great variation among the sites analyzed here (Tab. 1).

The Raup-Crick measurement (Fig. 8) showed that, considering the quantitative approach (Fig. 8A) showed that most sites are more dissimilar than expected while the qualitative approach (Fig. 8B), the tendency was to consider the majority of site-pairs as more similar than expected by chance.
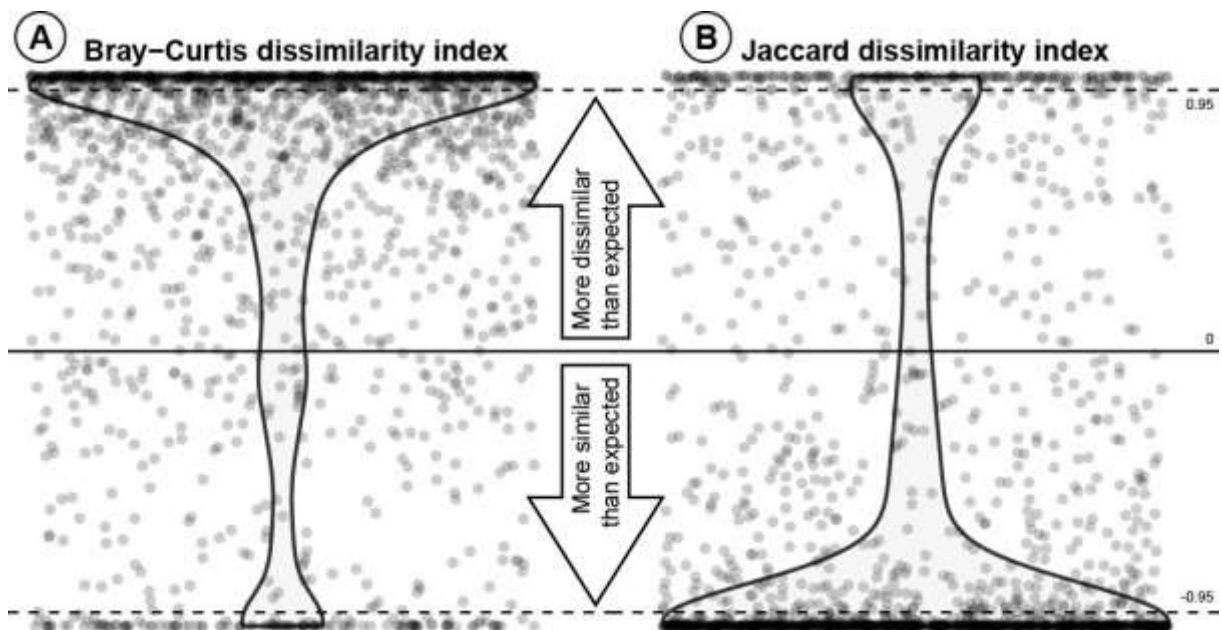
Figure 7 – Raup-Crick dissimilarity results considering the bacterioplankton composition (A) and abundances (B) in each site pairs. The values recovered from the analysis vary between -1 (more similar than expected by chance) and 1 (more dissimilar than expected by chance). Near to zero results can be interpreted as showing differences governed by drift

In the variation partitioning (Fig. 9), all comparisons presented a significant overall correlation. The fractions also showed significant results in the analyses, but the test for pure fractions showed that geography (for both qualitative and

93

quantitative approaches) and time (for qualitative approach) consistently presented non-significant results when compared with the environment. Also, in the comparison of all geographic factors between them, pure fractions of space and neighborhood were not significant in the quantitative approach (see Tab. S2 for significance values). Variation partitioning also showed that the environment was the best predictor of βdiv. For the qualitative approach the environment contributed alone with ~6% of observed βdiv, while time and geography did not reach 1%. For the quantitative approach, environment explained ~7.5%, time ~2.2%, and geography less than 1%. In this analysis, the interaction between environment, time, and geography also showed important values, and in some cases, it was greater than pure time and geography (Fig. 9). The partition of the environment between pH and the other factors showed that pH is a better predictor of βdiv, accounting alone for almost the same values as the other factors when considering the qualitative data and doubling their comparative importance for the quantitative data. Finally, the partition of geography indicated that, for the qualitative approach, neighborhood, space, and region explained the variation in this order, while for the quantitative approach, region and the interaction between neighborhood and space explained the geographic role on observed βdiv.
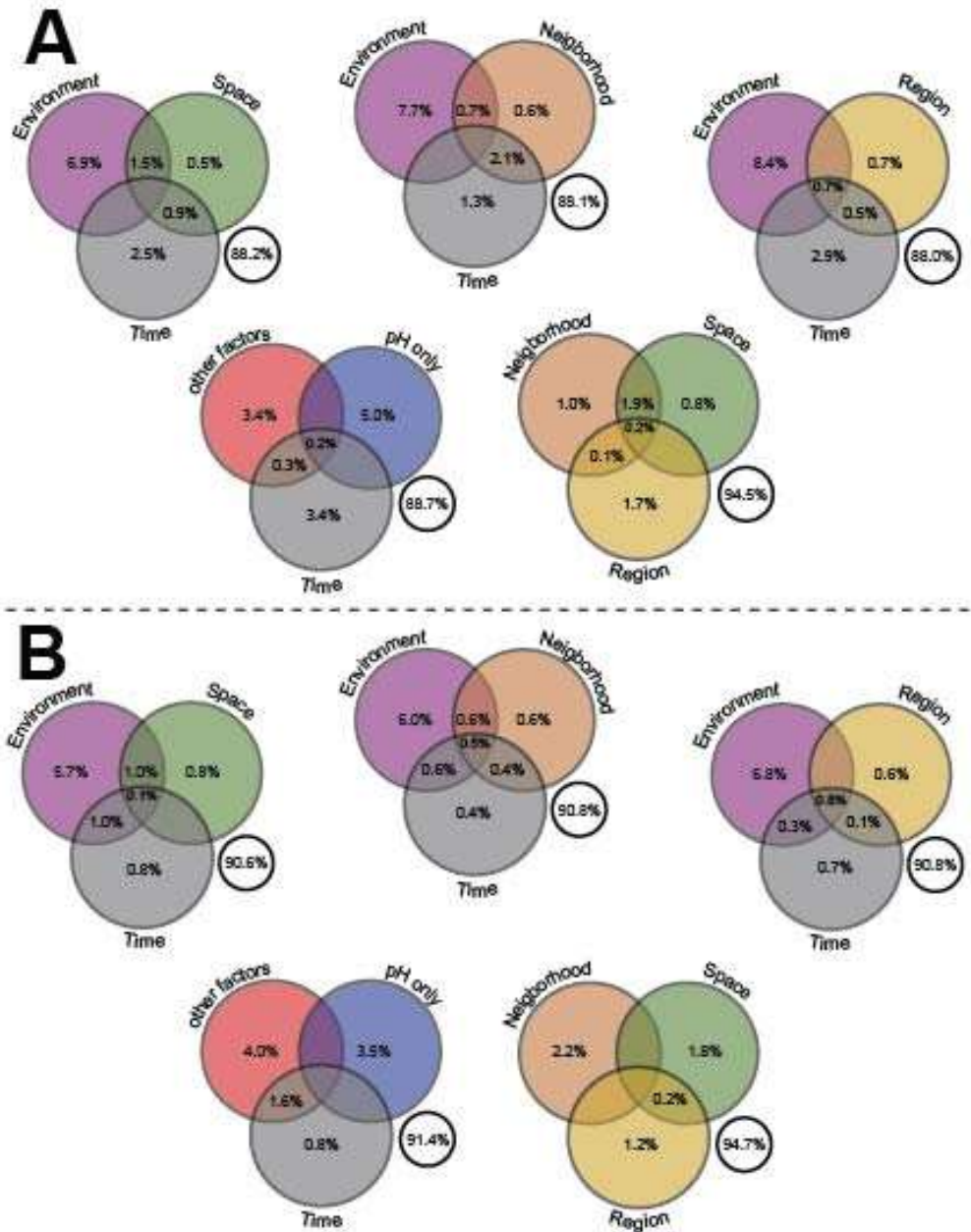
Figure 8 – Variation partitioning of bacterioplankton βdiv for qualitative (A) and quantitative (B) data for environmental (purple) and geographic (green) factors decoupled into their relevant components. The environment, was partitioned into pH (blue) an all other significant factor (orange); also, the geography was represented by the spatial overland distance between sites (space fraction; green), increased connectivity to neighbor sites (neighborhood fraction ; brown) and the isolation promoted by geographical barrier (region fraction; yellow); the time fraction (grey) was also represented. The isolated white circles show the non-explained fraction for each analysis

As evidenced by the occupancy range of each ASV (Fig.10), the abundant bacteria are those also capable of reaching greater distances in this landscape. We used the maximum distance to determine nine thresholds of maximum range and to

further apply the variation partitioning and verify which fraction (i.e. environment, geography, and time) better explain the βdiv according to the dispersion capacity of each bacterium. The environment is the best predictor of the better dispersers' βdiv, but its explicability decreases with the maximum range until it is overtaken by the spatial fractions at ~400–500 km distances and less. Unfortunately, we could not recover information for bacteria that travels less than 300 km as the analysis returned errors caused by the excess of zeros. The number of eigenvectors recovered for the space fraction was similar to those recovered for the global analysis but varied for the neighborhood and region fractions, increasing in number as the smallest ranging bacteria were being addressed. The unique observable difference between the qualitative and quantitative approaches is that the second shows a great increment on the spatial fraction explicability at the maximum range, which may indicate that it was caused by the change in abundance of these dominant organisms and a potential selective pressure spatially correlated. Table S2 and Figure S4 show more information about all fractions and significance tests.
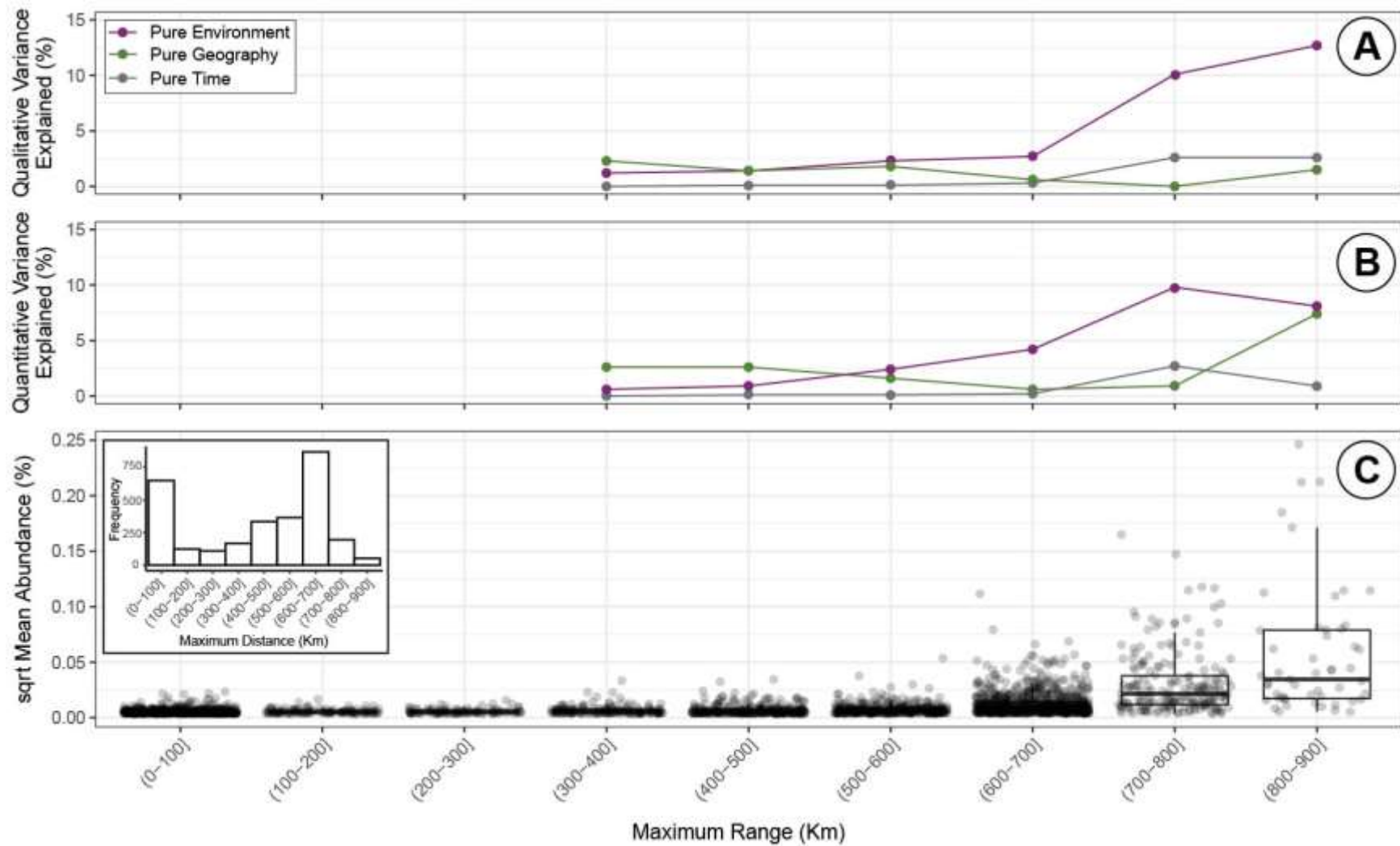
Figure 9 – Maximum distance bacteria can overcome compared with the relative importance of pure environment (purple), pure geography (green), and pure time (grey) for qualitative (A) and quantitative (B) approaches. The bottom panel (C) shows the comparison between squared mean abundances and maximum distance thresholds and the frequency distribution of maximum distances in which bacteria were capable to overcome.

For all analyses of variation partitioning, the unexplained fraction was always high, accounting for at least 76% of variance.

Discussion

In this study we found that the environment is the best predictor of the bacterioplankton βdiv in shallow lake metacommunities, but spatial factors also play a role. The importance of each factor followed a hierarchy in which the βdiv was mainly related to the environmental fraction followed by the variation through time and finally dispersion, as evidenced by distinct geographic features. This pattern is consistent through several tested factors and approaches, a pattern that was not previously expected, and probably 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 could also identify a set of rare and less dispersed organisms that are mainly guided by spatial features, as evidenced by Figure 10.

The role of the environment as the main driver of βdiv 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). This is especially due to pH, a well-known selective factor for aquatic bacteria (Lindström *et al.*, 2005; Niño-García *et al.*, 2016), which is a previously demonstrated pattern for these sites (Mateus-Barros *et al.*, 2021). Here, we focused on refining the spatial analyses as much as we could, seeking to cover all relevant dispersal routes bacterioplankton can take. To reach 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.*, 2008b; Peres-Neto and Legendre, 2010). Despite some concern 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 the spatial autocorrelation can be dealt with by an appropriate correction (Bauman *et al.*, 2018).

This high βdiv degree can also be linked to the high number of ASVs found, as it may lead to a higher number of negative interactions, which should increase the

effect of local environmental filtering, and consequently boost βdiv (Whittaker, 1972). On the contrary, it can promote the selection of a similar set of organisms better adapted to the factors found in this region, and consequently increase regional dominance, which can boost the similarity when the variation in local richness is controlled (Chase *et al.*, 2011). The Raup-Crick analysis, which was meant to overcome the effects of variation on local richness (Chase *et al.*, 2011), showed that these local communities were in fact more similar than expected when the qualitative data was considered, but the quantitative data showed that these communities were still dissimilar in many cases. The variation partitioning applied to different thresholds of maximum range showed a similar pattern when comparing qualitative and quantitative approaches at the maximum ranges. These two combined results seem to indicate that much of the spatial features observed here are in fact reflecting spatially-related variation in local selective pressures which in turn affect local abundances, and therefore the variation in dominance through this landscape could be detected and measured (Anderson *et al.*, 2011).

Variation partitioning recovered a large portion of the unexplained variation. We believe that this probably occurred due to two main factors acting alone or together: we may not have sampled some important variables from the processes that shape this metacommunity or this can in fact be a characteristic of this region and we are observing the action of neutral processes driving the communities (Melo *et al.*, 2012). For the first assumption, it was previously demonstrated that predation can influence a bacterial community in no disturbance conditions and it can increase the fraction of unexplained variations over time (Livingston *et al.*, 2017). Moreover, some temporally structured environmental factor may impact the metacommunity (Langenheder *et al.*, 2012) in a way we were not able to observe. Also, some environmental stochastic factor, for example, a disturbance (Vellend *et al.*, 2014), or a spatial deterministic factor, like a priority effect (Siqueira *et al.*, 2015), may be playing a strong role in this metacommunity; however, those factors were not measured in this study. Finally, recent studies showed that the flow of bacteria from the soil impacts headwater microbial diversity (Ruiz-González *et al.*, 2015; Caillon *et al.*, 2021). For the second assumption, the Raup-Crick analysis recovered high percentages of variation nearly to the null-model, which can be interpreted as the

metacommunity being mainly governed by drift (Stegen *et al.*, 2013). This inexplicability was previously observed in this region for other organisms and it was suggested that the characteristics of this region can lead the metacommunity to an increased relative importance of local variability and decreased regional synchrony (Lopes *et al.*, 2017; Zanon *et al.*, 2018). In a recent study, Huber *et al.* (2020) hypothesized that environments that have intermediate degrees of environmental heterogeneity should be mainly governed by stochastic processes. Complementarily, tropical environments have higher metabolic rates (Amado *et al.*, 2013; Freitas *et al.*, 2017), which may increase the importance of neutral processes to the community assembly (Saito *et al.*, 2021). In any case, an explicability of about 10% of total variation is common for this type of analysis (Melo *et al.*, 2012).

We also hypothesized that the quantitative approach would reflect the impact of the abundant bacteria on this metacommunity and be related to the environment and the linear decay in space (i.e. space fraction), while the qualitative approach would be more related to neighbourhood and region fractions and rare bacteria. These predictions were partially corroborated in the sense that the approaches do not precisely reflect the relationship between abundance and geography/environment fractions, which were in fact observed. By scrutinizing the variation partitioning results we can determine how each fraction is impacting this metacommunity. So, the space fraction (i.e. linear decay on between-site relationship) is consistently related to bacterial βdiv through the maximum distance thresholds analyzed, but the spatially-related local selection evidenced by the unique inconsistency between qualitative and quantitative approaches in this analysis indicates that the geographic fraction is more closely related to the regional variation in abundance of dominant taxa. On the other hand, the increment in importance of geographic fractions at smaller ranges occurs because of the addition of eigenvectors related to neighborhood (i.e. abrupt decay in the between-sites relationship) and region (i.e. isolation by geographic barriers) fractions; it also demonstrates a relationship between these two fractions and rare bacteria. It is important to point out that despite these results showing a relationship between different geographic aspects for rare and abundant bacteria, the recovered values were small. It makes the delimitation of another study encompassing a wider area to better understand the influence of these aspects on

bacteria necessary (Lansac-Tôha *et al.*, 2020) and maybe also figure out the impact of scale for this group.

Conclusions

The βdiv is a key tool for infering biological differences between sites. This method has been refined in recent years (Tuomisto, 2010; Anderson *et al.*, 2011; Chase *et al.*, 2011), as well as the variation partitioning approach (Bauman *et al.*, 2018), which allows inferences into the patterns that impact organisms in natural communities (Soininen *et al.*, 2018). In the meantime, much effort has been made to better understand which processes guide bacterial communities (e.g. Stegen *et al.*, 2013) and to integrate bacterial studies into classical ecology (Barberán *et al.*, 2014). In this study, we were able to recover different dispersal aspects that may influence bacterial community composition. The dbMEM approach, which is the method most commonly used in studies of this type, recovered the spatial variables that best explained the spatial variation of abundance on dominant taxa and highlighted a potential space-related selective pressure on these organisms. On the other hand, the MEM approach allied to distinct strength data to recover neighborhood and region aspects were those who better explained the variation in bacteria that were rare and susceptible to geographic isolation. This result illustrates the importance of exploring different approaches aiming to recover all geographic aspects influencing the metacommunity. Although this factor normally shows a weaker influence on biological variation, it plays a role in structuring the metacommunity in a direct way, as demonstrated here, but also indirectly by opposing or reinforcing the effects of stochastic and deterministic forces (Vellend *et al.*, 2014). The evidence that, to some degree, geography influences bacterial distribution is increasing (e.g. Stegen *et al.*, 2013; Lindh *et al.*, 2017; Lansac-Tôha *et al.*, 2020; Logares *et al.*, 2020; Mateus-Barros *et al.*, 2021), but studies that unveil the exact scale it overcomes the environmental factors, and how this change can potentially impacts bacterial metacommunites are still necessary (Lansac-Tôha *et al.*, 2020).

Capítulo VI - Scale matters?
The effect of spatial scale on ecological processes
that drive the aquatic bacterial communities

Embora as bactérias desempenhem um papel fundamental em todos os ambientes, a verdadeira extensão de sua diversidade só foi descoberta com o avanço das ferramentas moleculares. Esses avanços permitiram entender e quantificar os processos ecológicos que impulsionam a composição das comunidades bacterianas. No entanto, ainda faltam informações sobre como os processos ecológicos mudam em escalas espaciais. Reunimos dados de sequenciamento de amplicon 16S rRNA de 135 lagos rasos da Argentina, Brasil e Canadá, e comparamos as características determinísticas e estocásticas com modelos nulos para determinar quais processos impulsionam a metacomunidade bacterioplanctônica em três escalas espaciais (regional, subcontinental e intercontinental). Observamos que os processos ecológicos determinantes das comunidades bacterianas foram diferentes em cada região, dependendo dos contextos locais: seleção homogênea (principalmente no Brasil) e limitação da dispersão combinada com deriva (principalmente na Argentina) foram os processos mais proeminentes, enquanto no Canadá houve uma contribuição igual de ambos os processos. Um padrão geral emergiu de nossa análise de que o processo ecológico dominante mudou de seleção para deriva e, finalmente, para limitação de dispersão com distâncias crescentes, não importando a escala considerada. Descobrimos que os processos ecológicos que moldam uma metacomunidade bacteriana mudam com as distâncias geográficas, em um padrão que se repete em escalas continentais e intercontinentais. Portanto, pelo menos na escala dos continentes norte e sul-americanos, os processos ecológicos que moldam uma metacomunidade bacteriana não eram dependentes de escala.

Introduction

Understanding and measuring the ecological processes that drive communities' composition remains a fundamental question in ecology. Usually, the processes are classified as deterministic or stochastic and act jointly in major or minor proportion depending on local conditions to shape the community assembly (Melo *et al.,* 2012). Deterministic processes are measurable and can be directly related to the capacity of a population or community to lead with their changes over space and time. This perspective has been studied in community ecology over decades and is commonly related to the multidimensional niche concept (Hutchinson, 1959; Vandermeer, 1972), like biological interactions and environmental filters (Hillerislambers *et al.,* 2012). On the other hand, stochastic processes, which were more recently proposed under the umbrella of neutral theory (Hubbell, 2001), are those that unpredictably impact a community, and have been related to spatial factors and drift (Vellend *et al.,* 2014). The recognition that these two high-level processes are not acting alone allowed the conceptualization of new perspectives in community ecology. A great example is the metacommunity concept, which considers that a set of local communities are connected by migration (Leibold *et al.,* 2004), and provides a powerful lens that has been extensively used to see the assembly of communities simultaneously in both local and regional scales.

Bacteria are closely related to any other Earth life form and are fundamental to most known biogeochemical cycles. These organisms are widely found (Green and Bohannan, 2006), and take part in ecological webs by using and cycling nutrients (Sarmento and Gasol, 2012; Sarmento *et al.,* 2013), as nutritive preys (Khan e Siddiqui, 2014; Kavagutti *et al.,* 2019), or even as effective population growth controllers (Paliwal, 2017). Despite its remarkable importance, we could figure out the extent of their biodiversity only when newer molecular tools applied to microbial taxonomic identification became widespread and, since that, new theoretical perspectives on how the structuring community were proposed (e.g. Pedrós-Alió, 2012). Also, these tools made it possible to interpret the value of stochastic processes allied to the already known deterministic roles (Barberán *et al.,* 2014).

Studying how and at which intensity these assembly processes guide bacterial community dissimilarity remains a fundamental question to microbial ecology, and has been measured extensively. In this context, Stegen and collaborators (2012), proposed a quantitative tool to determine the relative importance of deterministic and stochastic forces guiding a metacommunity. Since its proposal, this method was applied to multiple microbial metacommunities and has been providing valuable information on the processes that guide microbial biodiversity in different contexts, such as gut microbiomes (Martínez *et al.*, 2015), sediments (Stegen *et al.*, 2012; Stegen *et al.*, 2013), soils (Zhang *et al.*, 2020), lakes (Llames *et al.*, 2017; Logares *et al.*, 2018), wetlands (Huber *et al.*, 2020) and ocean (Logares *et al.*, 2020), ; and have been demonstrating a variable prevalence of deterministic selection and stochastic dispersal factors. The predominant process in a region may also vary in space and time, depending on environmental heterogeneity (Huber *et al.*, 2020). For example, in soil bacteria, the degree of explicability for all processes fluctuates from the tropics to temperate latitudes (Zhang *et al.*, 2020). Also, in a scenario of ecological succession, it is also possible to observe a succession of processes' prevalence from an initial predominance of drift, through a rise of deterministic pressures and reaching a climax with more prominent spatial stochasticity (Dini-Andreote *et al.*, 2015). Despite this extensive scrutiny of microbial metacommunities using this approach, a robust analysis of how ecological processes change across spatial scales is still lacking.

In this study, our objective is to determine the effects of spatial scale on the contribution of different ecological processes that shape bacterioplanktonic communities. For this purpose, we concatenated data from shallow lakes' bacterial diversity in three regions: The pampean region in Argentina, the southeastern Brazilian coast and highlands, and a boreal biome in Canada. We organized this data to observe the main processes at three distinct scales: (1) the regional scale encompasses the interactions within each one of the regions and comprehends distances between sites of one thousand kilometers, (2) the continental scale considers a connection between Argentinian and Brazilian sites and distances of about three thousand kilometers, and finally (3) the intercontinental scale, in which the Canadian data is lined up with the other two and considers distances greater than fifteen thousand kilometers (Fig. 11). To build our hypotheses, we set up some

statements as a framework (resumed in Tab. 3). First, headwater freshwaters receive a greater number of terrestrial organisms (Ruiz-González *et al.*, 2015); in this case, the aquatic biodiversity may vary depending on terrestrial processes, like human activities for example (Banerjee *et al.*, 2019). Second, the degree of environmental heterogeneity may change the process that is primarily guiding the metacommunity assembly (Huber *et al.*, 2020). Also, the temperature may influence microbial biodiversity, as this environmental factor may raise metabolic rates and lead to processes that are neutral and independent of species identities, for example, by accelerating life cycles with a consequent growth and losses of individuals and reduction of competition opportunities (Saito *et al.*, 2021). And finally, the dispersal capacity of an organism is preponderant to whether spatial processes impact a community or not (Wang *et al.*, 2013), and therefore should be decreasingly noted as the scale at which processes are observed is increased. In this context, we hypothesized that (I) at a lower scale, we will observe distinct processes governing the metacommunity assembly depending on the context in each region inserted: for (Ia) the Argentinian data, we expect to see a prevalence of stochastic processes due to an increased possibility of environmental heterogeneity; for (Ib) the Brazilian dataset, we expect a prevalence of homogeneous selection due to increased contact with soil environment caused by voluminous tropical rainy events and human land use (agriculture and livestock) that should homogenize de microbial biodiversity, also the elevated metabolic rates should contribute to increasing drift guiding these communities; and for (Ic) the Canadian dataset, we expect the predominance of heterogeneous selection and dispersal limitation due to moderate environmental heterogeneity and a certain degree of isolation by distance greater than the other two. Also, we expect that (II) the increasing scales will show a substitution in the main assembly processes from deterministic to stochastic factors, with (IIa) an increment of dispersal limitation on the median scale, and (IIb) drift at a larger scale.

Table 3 – Expected main processes for each environmental context and scale

| | | Framework | | | Expected main processes |
	Mass Effect (Ruiz-González et al., 2015)	Environmental heterogeneity (Huber et al., 2020)	Temperature (Saito et al., 2021)	Dispersal rate (Wang et al., 2013)	
Argentina	+	++	+	+++	Stochastic processes
Brazil	+++	+	+++	+++	Homogeneous Selection + Stochastic processes (mainly Drift)
Canada	++	++	+	++	Heterogeneous Selection + Stochastic processes (mainly dispersal limitation)
Increasing Scale	++	increase	++	decrease	Heterogeneous Selection + Dispersal limitation -> Drift

(left margin label: **Scale**)

Methods

Data sources

For this study we compiled a dataset that brought together 135 samples of shallow lakes near headwaters in three different regions of the American continent (Fig. 11). The Argentinian samples were collected at the central plain area, in the called Pampa ecoregion (Iriondo, 1989). The climate is temperate and humid, with around 700 mm of annual precipitation that can vary enormously depending on the location (Northeast to Southwest) and climate event (la niña, el niño) (Diovisalvi *et al.*, 2015). In this region, shallow lakes are numerous and produced by rivers and wind, and the vegetation is dominated by grasses (Iriondo, 1989). The Brazilian data comprises a transition area between the Cerrado and Atlantic forest biomes in the coast and highlands southwestern region. In this region, there are two marked climatic periods, dry winter and rainy summer (Ratter *et al.*, 1997), and the annual precipitation is about 1500 mm. The Cerrado soil is typically acid and nutrient-poor (Haridasan, 2008), with periodic fire episodes during the dry season (Coutinho, 1990). The Atlantic forest grows over richer soils and presents higher humidity throughout the year (Joly *et al.*, 2014). In this region, lakes are uncommon, but rivers of any order are dammed to fulfill human necessities.
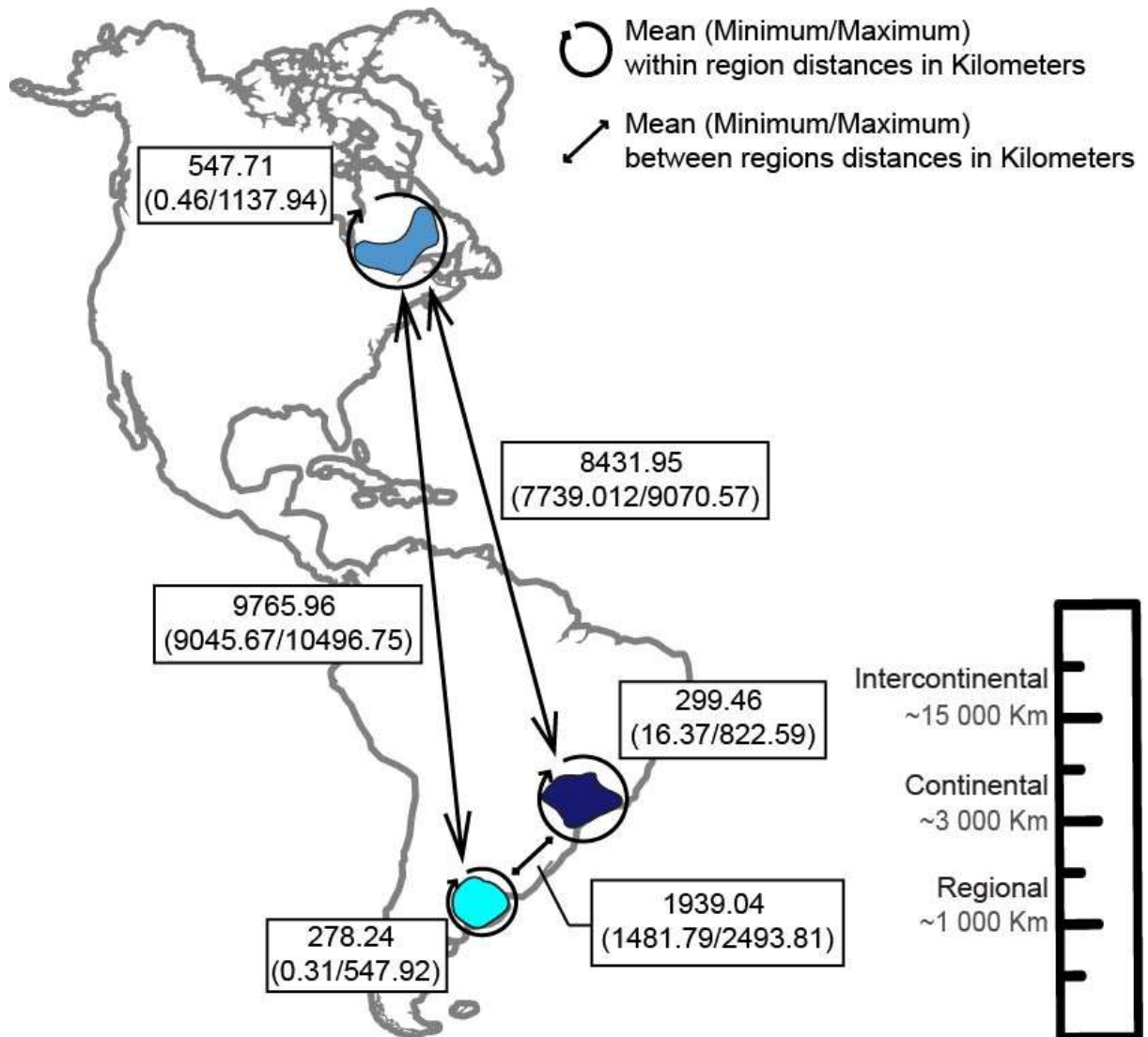
Figure 10 – The regions in which data was gathered. The ruler illustrates the scales considered for this study; the boxes show mean (main value), minimum and maximum (between parentheses) distances between sites are in the boxes within and between regions in kilometers. At the regional scale, the distances between sites are about one thousand kilometers and encompass sites from the Argentinian Pampas (cyan), the Brazilian coast/highlands (dark blue), and the Canadian boreal ecozone (blue). The continental scale considers the distances between Argentina and Brazil and about three thousand kilometers. The intercontinental scale brings Canada together whit the others and considers distances of about fifteen thousand kilometers

The data from South America is part of a consortium that made a recent effort to gather data on bacterioplankton across the continent (Metz *et al.*, 2022). The Canadian data is part of a sampling effort that sampled rivers and lakes of Québec, Canada (Niño-García *et al.*, 2016). It is a boreal ecozone that comprises several ecoregions slightly isolated from each other (Niño-García *et al.*, 2016). The

vegetation varies from deciduous to conifer forests, as well as the soils, vary from glacial deposits over sedimentary rocks to a maritime climate (Rasilo *et al.*, 2015). In this region, the temperatures can be negative and annual precipitation varies by about 900 mm (Ruiz-González *et al.*, 2015). Because of the harsh climatic conditions most of the year, the samples were sampled during the summer period (Niño-García *et al.*, 2016). This data can be found at the European Nucleotide Archive (ENA) under accession number PRJEB11530.

From all environmental samples, temperature, pH, chlorophyll *a* and Dissolved Organic Carbon (DOC) could be recovered from all samples and, complementary, altitude, lake area and perimeter were obtained from georeferencing software using each geolocation (Tab. 4).

Table 4 – Environmental data for each environmental context and scale. For each environmental variable the main values are means, while the values between parentheses are the minimum and maximum respectively

		Altitude (m)	Lake area (km²)	Lake perimeter (km)	Temperature (°C)	pH	Chlorophyll *a* (µg/L)	DOC (mg/L)
Argentina	mean	70.44	22.077	30.562	18.540	8.754	89.251	94.294
	(max/min)	(246/-16)	(157.812/0.197)	(141.166/1.846)	(30.0/7.0)	(9.4/8.0)	(981.063/1.580)	(539/1.680)
Brazil	mean	478.52	0.016	121.541	23.102	6.730	15.992	9.337
	(max/min)	(1121/7)	(0.096/0.0001)	(1195/0.021)	(30.6/16.1)	(10.1/5.03)	(105.193/0.092)	(42.060/2.04)
Canada	mean	435.72	2.894	2.894	18.363	6.656	1.506	6645.542
	(max/min)	(846/89)	(9.415/0.002)	(20.778/0.167)	(23.74/13)	(8.07/5.05)	(5.5/0.19)	(17113/2.06)
Continental Scale	mean	274.48	11.047	76.052	20.821	7.742	52.621	51.815
	(max/min)	(1121/-16)	(157.812/0.0001)	(1195/0.021)	(30.6/7)	(10.1/5.03)	(981.063/0.092)	(539/1.68)
Inter-continental Scale	mean	328.227	7.508	51.666	20.002	7.380	35.583	2249.72
	(max/min)	(1121/-16)	(153.513/0.0001)	(1195/0.021)	(30.6/7)	(10.1/5.03)	(981.063/0.092)	(17113/1.68)

Standardization of datasets and bioinformatics

Concerning the amplification and sequencing of 16S rRNA for the taxonomic identification of bacterioplankton, not all samples used the same primer. For the samples from Argentina and Brazil, the V3-V4 region were sequenced using the 341F (5′-CCTACGGGNGGCWGCAG-3′) and 805R (5′-GACT ACHVGGGTATCTAATCC-3′) (Herlemann *et al.*, 2011), while the samples from Canada were sequenced with 515F (5′-GTGCCAGCMGCCGCGGTAA-3′) and 806R (5′-GACTACHVGGGTATCTAAT-3') (Caporaso *et al.*, 2012), that also recovers the

V3-V4 region, but resulting in shorter sequences. Aiming to standardize the sequences length, all sequences were cut using cutadapt (Martin, 2011), using as reference the adapters that recovered shorter sequences. After this, we used the DADA2 pipeline (Callahan *et al.*, 2016) in the software R (R Core Team, 2019) to guaranty sequences quality, eliminate chimeras and assign taxonomy using the SILVA version 128 database (Quast *et al.*, 2013; Yilmaz *et al.*, 2014) as reference to recover an unified Amplicon Sequence Variants (ASVs) table. Then, the ASVs table was rarefied by the lowest sample richness (12,682 reads), and the ASVs that do not have at least 10 reads in the sum of the reads in all sampling sites were removed.

Statistical procedures

To determine if the main factors shaping the observed turnover between pairs of sites are due to deterministic or stochastic processes, we applied a framework that uses two statistical analyses sequentially, comparing the observed features to null-models and defining the main process assembling microbial communities (Stegen *et al.*, 2013). The first step uses phylogenetic trees through the β-mean-nearest taxon distance (βMNTD). In this analysis, the phylogenetic distance of each taxon in one site is compared with its closest relative in the other, and the generated values are compared with the value found for the phylogenetic pool of ASVs constructed using all samples to determine if the calculated difference are due to selective pressure or not. For each region and scale, a distinct phylogenetic pool was constructed considering only the ASVs found in each context. The resulting values are the β-nearest taxon index (βNTI) and if it is greater than 2 or smaller than -2, there is a correlation greater than expected by chance. In the case of values >2 the communities are experiencing a heterogeneous selection, on the other hand, values <-2 indicate communities experiencing a homogeneous selection. This analysis assumes that closely related taxa present closely related habitat preferences (Losos, 2008), and, because of that, it is necessary to test this assumption through the phylogenetic signal approach that should show a significant relationship between environmental factors and taxa at short phylogenetic distances (Webb *et al.*, 2002) (Fig. S5). Pairs that do not show significant values of βNTI are considered as being

mainly assembled by stochastic factors and analyzed afterward. The second step uses the Raup-Crick dissimilarity, in this case, a regional pool of ASVs' composition is used to calculate a Bray-Curtis based β-diversity and this is compared to the β-diversity for each pair of sites (Chase *et al.*, 2011). For each region and scale, a distinct ASVs' pool was constructed considering only the ASVs found in each context. This analysis return values that fall between 1 and -1 and are considered significantly different from the null expectation when are greater than 0.95 and smaller than -0.95. Values >0.95 indicate that the assembly in these communities is mainly governed by Dispersal limitation combined with Drift, while values <-0.95 indicate the action of a homogenizing dispersal. Results that fall between these values represent pairs of sites equality structured to the random expectation, and therefore, assembled by Drift acting alone (Stegen *et al.*, 2013). This analysis uses ASVs abundances and a Bray-Curtis dissimilarity approach to calculate the β-diversity (Stegen *et al.*, 2013), because of that (and henceforward in this study) it is called RC_{bray}. All the statistical analyses above described were conducted using the software R (R Core Team, 2019). The βMNTD/βNTI analysis used the packages *picante* (Kembel *et al.*, 2010) and *vegan* (Oksanen *et al.*, 2016), and the RC_{bray} uses a function developed by Chase *et al.* (2011) and adapted to abundances by Stegen *et al.* (2012).

To examine the accuracy and prediction power of our theoretical framework, we applied some other statistics. About the assumption that these headwater shallow lakes should be intimately related to the soil environment, we used georeferencing software to find the highest point in the surrounding landscape and the distance of this point to the lake margin. With this information, we could calculate a slope value for each lake and, allied to annual precipitation information, we could infer a potential of bacterial flow from the soil to freshwater and compare it with the community assembly results. The relationship between bacterioplankton and the environmental factors showed in Tab. 4 was accessed with a non-metric Multidimensional scaling (nMDS).

Complementarily, we also compared the environmental heterogeneity and dispersal limitation (showed as the increasing distance between a pair of sites) with biological dissimilarity to see the direct influence of these variables on biological

dissimilarities. To the biological dissimilarity, we applied the beta diversity partitioning analysis using Bray-Curtis index (Baselga *et al.*, 2017). To the environmental dissimilarity, we applied a distance-based ordination (Oksanen *et al.*, 2016), to obtain a unified environmental heterogeneity scale that considers all environmental factors together. Finally, we calculated the overland spatial dissimilarity using a distance-based Moran Eigenvector Map (dbMEM) approach (Dray *et al.*, 2006).

Results

After rarefaction and filtering, we obtained 1'690'098 reads scattered over 5'162 ASVs and 29 Phyla. Actinobacteriota (33.23%), Proteobacteria (31.32%), Bacteroidota (12.72%), Verrucomicrobiota (8.06%), and Cyanobacteria (4.52%) were the phyla that most contributed to the abundance in these communities, while Proteobacteria (1'506), Bacteroidota (1'024), Actinobacteriota (770), Verrucomicrobiota (414) and Patescibacteria (384) were the phyla that had the major number of ASVs identified within then. Of all these ASVs, 88 (1.7%) were shared between all regions, while 4579 (88.7%) were unique for one of them; in which 2005 (38.8%) were found only in Argentina, 994 (19.3%) from Brazil, and 1'580 (30.6%) uniquely found in Canada. Argentina and Brazil shared 132 (2.6%) ASVs, Argentina and Canada had 125 (2.4%) ASVs, while 238(4.6%) were in both Brazil and Canada.

Complementarily, we compared the relationship between the occupancy capacity of each ASV with their (log) mean relative abundances in all scales (Fig. 12). This distribution could be decomposed into linear equations which can highlight some interesting biodiversity differences for the ecological contexts. First, the linear coefficients found in the Argentinian and Canadian regions (-9.95 and -9.83, respectively) greater than in Brazil (linear coefficient = - 10.28) indicates that there is a bigger number of rarer ASVs in these temperate regions, on the other hand, there are slightly larger values in their slopes (0.24 and 0.2, respectively. Brazil = 0.13), which indicate that there is a bigger number of ubiquitous ASVs in Brazil, despite these ASVs sound to be less abundant in these samples.

About the processes that mainly influenced the community assembly (Fig. 13), we found quite different results at the regional scale in each region: in the Argentinian sites, the community assembly seems to be more impacted by stochastic processes, in special by the dispersal limitation combined with Drift, and less for the Drift acting alone. Here, the second most important assembling process was homogeneous selection, and is possible to see that some differences were probably caused by Heterogeneous selection. On the contrary, for the Brazilian metacommunity, Homogeneous selection was the most identified process, followed by Drift acting alone and dispersal limitation combined with Drift. The Brazilian metacommunity was unique in that also showed the Homogenizing dispersal as an important assembly process. The Canadian metacommunity showed a balanced number of pairs of sites guided by Homogeneous selection, Dispersal limitation combined with Drift and Drift actin alone. At the continental scale, the trend was to an assembly mainly guided for stochastic processes; dispersal limitation combined with Drift was the main process, followed by Homogeneous selection. Drift acting alone and Heterogeneous selection was also detected for this scale. Finally, at a larger scale, we found similar importance of deterministic and stochastic processes guiding these communities at the larger scale. The homogeneous selection was the most prominent process, followed by dispersal limitation acting in combination with Drift and Drift. We also plotted the βNTI and RCBray against the geographic distances and calculated the adjusted R-squared to observe the distribution of results with increasing distances.
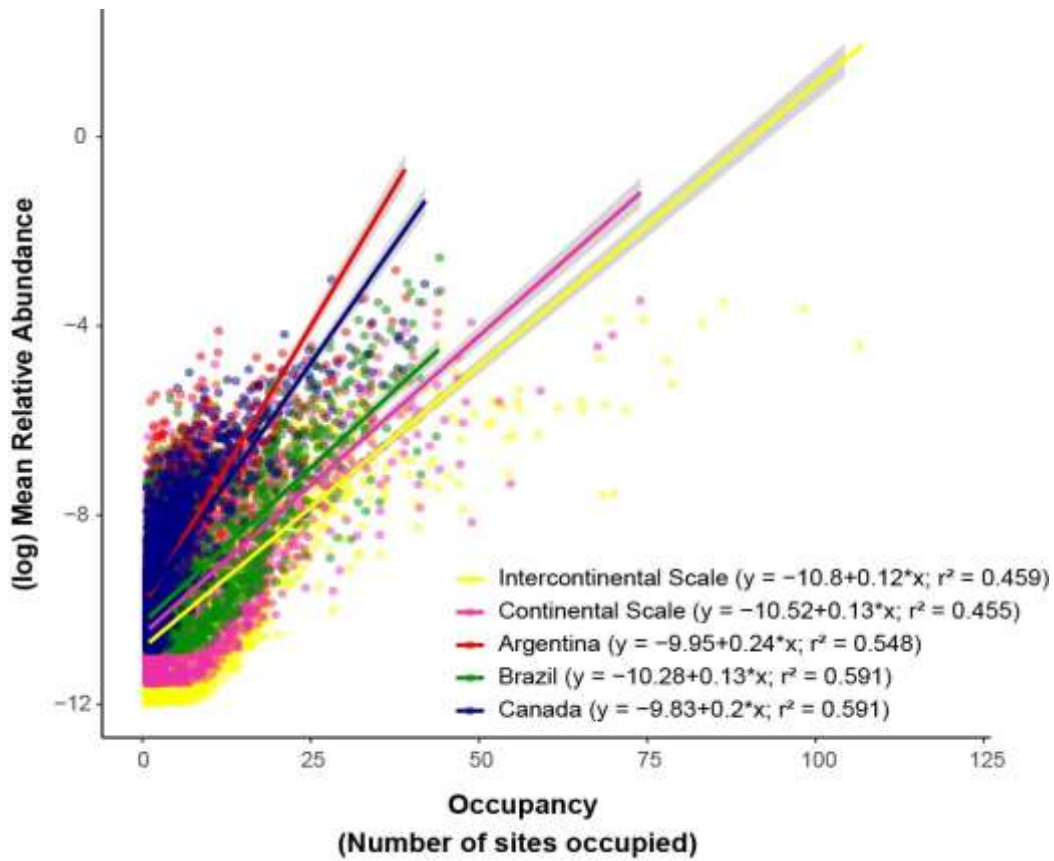
Figure 11 – Sites occupancy of each ASV against their log mean relative abundances for Argentina (red), Brazil (green) and Canada (blue) at smaller scale, and also for Continental (pink) and Intercontinental (yellow) scales
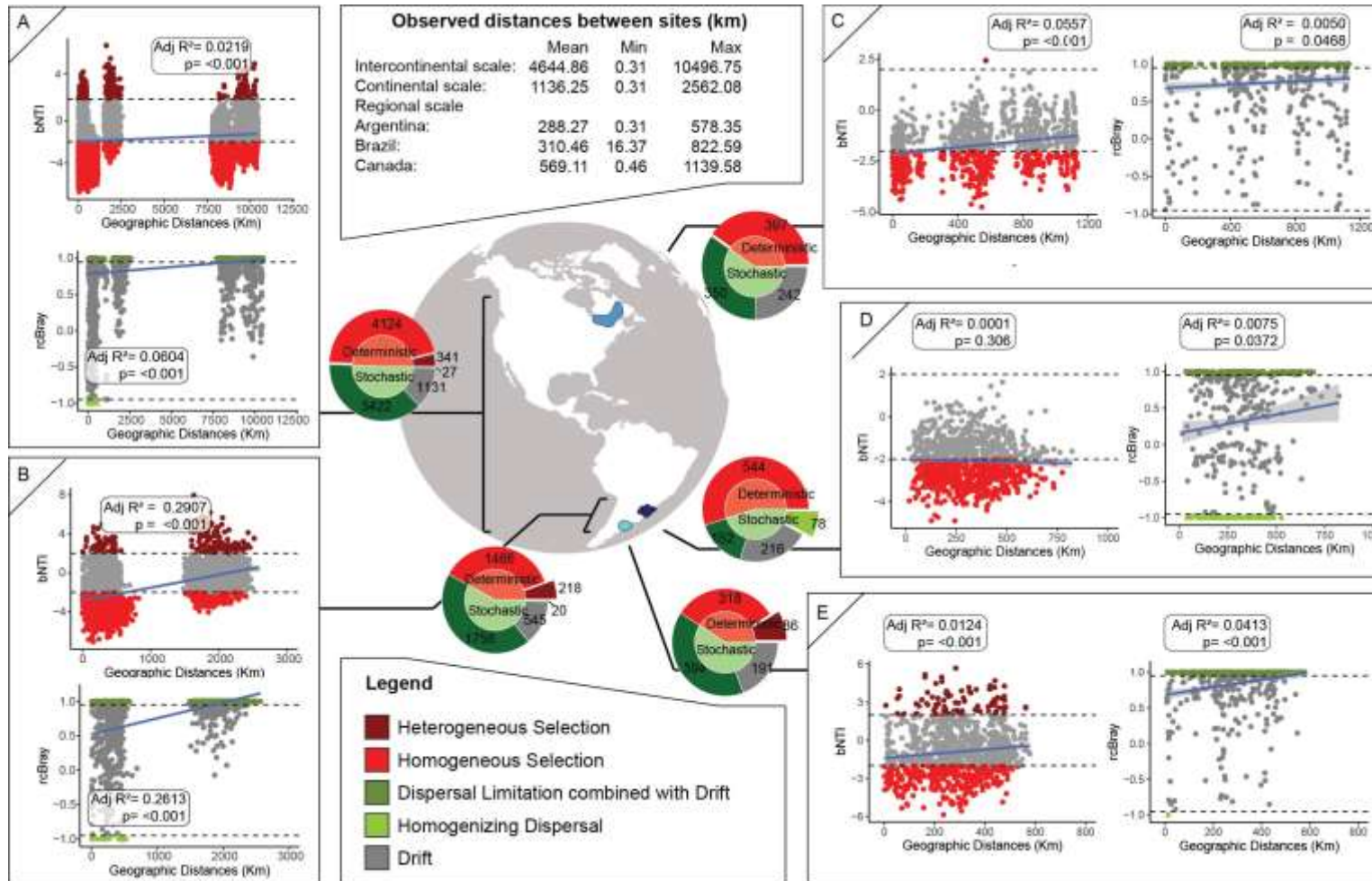
Figure 12 – High-level metacommunity assembly results for Intercontinental (A) and Continental scales (B), also, at regional scale, for Canada (C), Brazil (D) and Argentina (E). Pie plots show the relative importance of each process in each region/scale while the charts show the tendency of distribution for each process with increasing distances between pairs of sites

114

A pattern that was consistently observed through all the scales were to a major number of observed homogeneous selection in direction to not significant βNTI results, combined with a trend of communities being mostly guided by Drift acting alone, and a gradual substitution to Dispersal limitation combined with Drift. The Brazilian data was the only one that did not precisely follow this pattern, as the tendency here was to maintain a great number of sites assembled by Homogeneous selection even at larger distances. It is also important to highlight that the adjusted r-squared values remain low for all environments, which indicates a weak correlation of this pattern even when significant values were found.

The rain flow potential (Tab. 6) can be inferred by the direct observation of landscape slopes around the shallow lakes compared with annual precipitation in each region. The Argentinian landscape showed fewer slopes allied to the fewer annual precipitations, which can indicate a lesser connection between freshwater and soil bacterial communities. On the other hand, the Brazilian landscape has greater slopes and huge annual precipitation, indicating a mass effect potential from the surrounding soils to shallow lake metacommunity through the entire landscape. The Canadian landscape had great values of annual precipitation and landscape slopes to de lakes, but slightly smaller than observed in Brazil, which also indicates a smaller mass effect in this region.

Table 5 – Rainfall flow potential for each region/scale of this study. The main values represent the mean. The values between parentheses are the minimum and maximum respectively

		Max distance (m)	Altitude (m)	Slope	Annual precipitation (mm)
Argentina	mean	557.8667	15.644	0.042	~ 700
	(max/min)	(3580/28.5)	(86/1)	(0.170/0.005)	
Brazil	mean	1026.798	88.267	0.105	~ 1500
	(max/min)	(3910/30.6)	(898/1)	(1.021/0.02)	
Canada	mean	794.809	75.244	0.093	~ 900
	(max/min)	(3330/21.2)	(308/1)	(0.286/0.009)	
Continental Scale	mean	792.332	51.956	0.074	
	(max/min)	(3910/28.5)	(898/1)	(1.021/0.005)	
Inter-continental Scale	mean	793.158	59.718	0.079	
	(max/min)	(3910/21.2)	(898/1)	(1.021/0.005)	

Concerning the comparison between bacteria dissimilarity and environmental and spatial heterogeneities (Fig. 14), we could observe an increment in biological

dissimilarities closest related to the increment in environmental heterogeneity. This pattern is observed across different regions and scales. On the contrary, each region shows a particular relationship with space distances. The Argentinian data presented a limited relationship between increasing biological dissimilarity and increasing spatial distances, on the opposite, the Brazilian have two groups of small and great dissimilarity not related to the distance increase. The Canadian data also presented a different pattern, in which some clear groups were formed by the increasing distances presenting always a moderate to strong biological dissimilarity. For the continental and intercontinental scales, there is a strong shape of increased biological dissimilarity with increasing distances between sites.
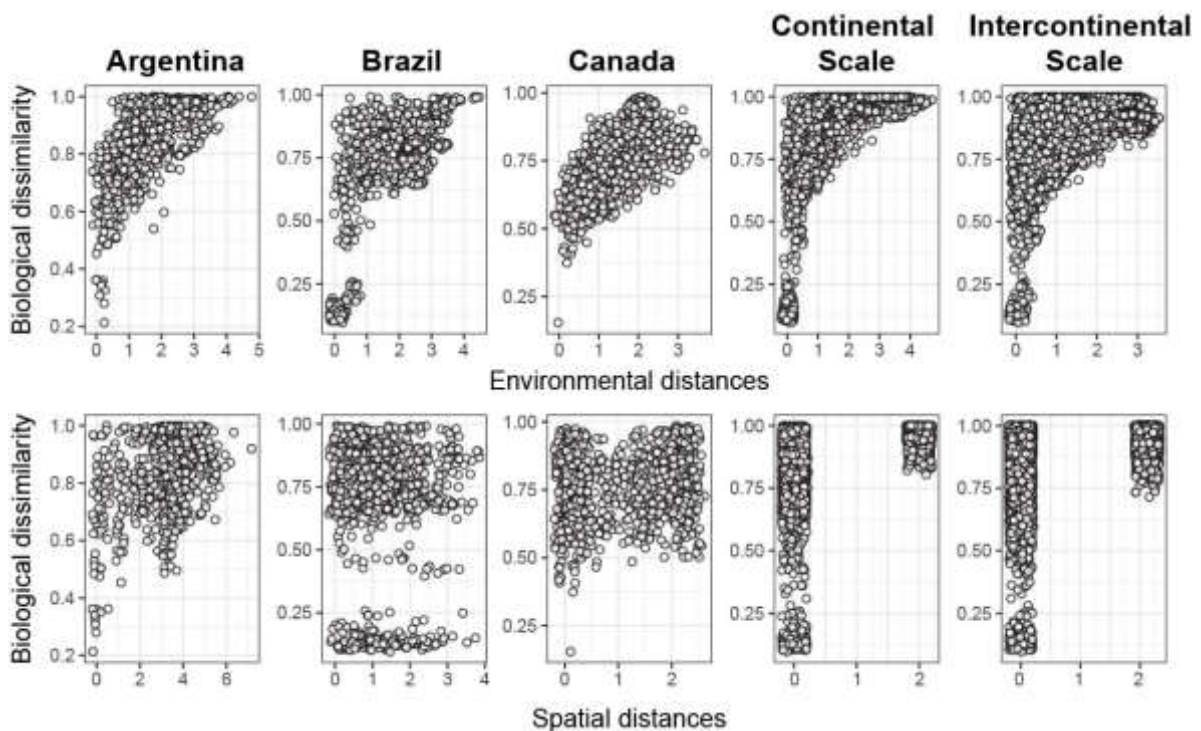
Figure 13 – Between sites dissimilarity compared with environmental (top) and spatial (bottom) dissimilarities for each region/scale here studied

The nMDS (Fig. 15) showed that, despite a similar within-region dissimilarity (permidisp < 0.05), each region seems to be related to different environmental factors, which reflects a significant dissimilarity between regions (permanova > 0.05). Argentina was positively related with chlorophyll a, pH, and lake area and negatively

with altitude, indicating a stronger correlation between bacterial biodiversity and variation in biochemical factors that may be indicative of biological interactions or large-scale deterministic processes like rain regimes, for example. Brazil showed a positive relationship between lake perimeter and temperature, which indicates a possible correlation with metabolic issues. Finally, Canada was positively related to DOC, which indicates a stronger correlation between nutritional inputs and availability.
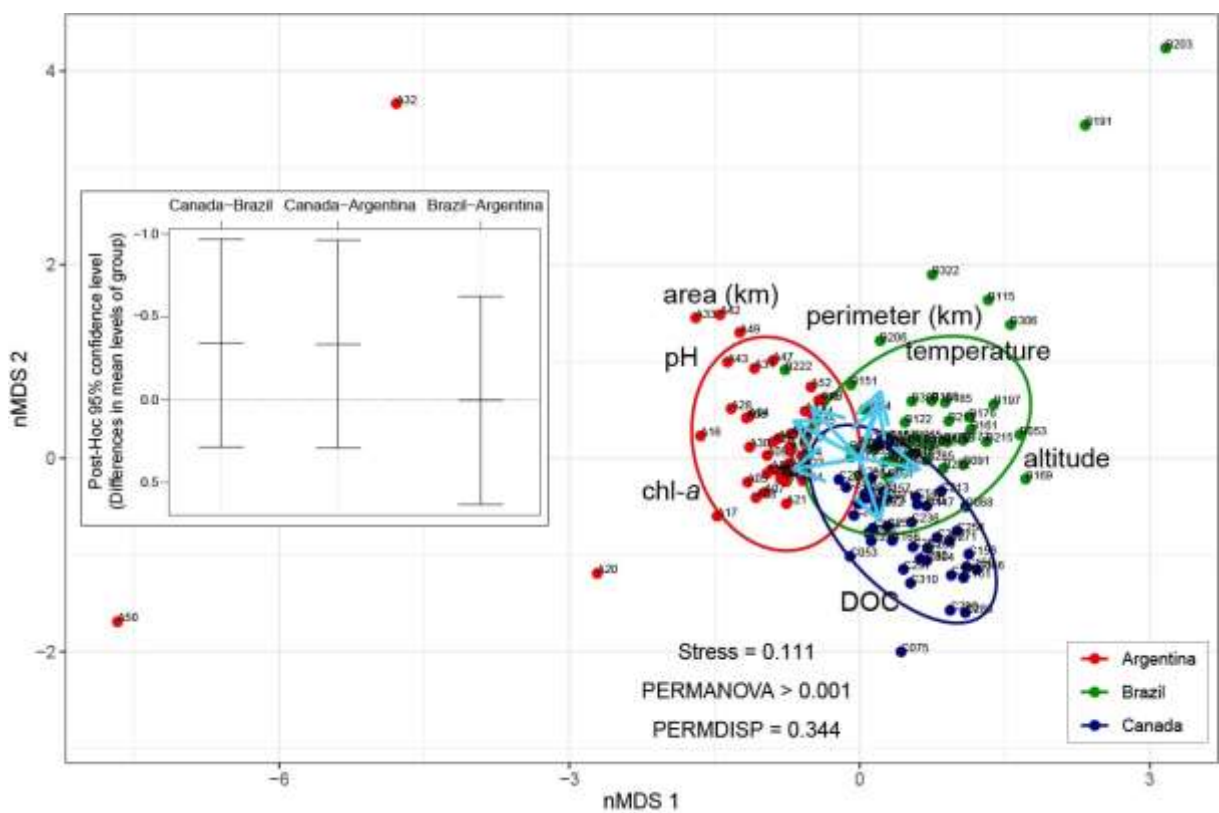
Figure 14 – nMDS showing within (Permdisp < 0.05) and between (Permanova > 0.05) regions dissimilarity values for Argentina (red), Brazil (green) and Canada (blue) shallow lakes. The blue arrows indicate the envvectors calculated by the nMDS to reach the dissimilarities observed in this plot. The smaller box shows the *post hoc* analysis for each pair of regions

Discussion

In this study, we applied a two-step null model comparison to detect the main processes assembling different shallow lake metacommunities at three spatial scales. This method assumes that closely related organisms have similar ecological niches and, therefore, are similarly selected by deterministic processes (Webb *et al.*, 2002). This approach is prone to some criticism, as not all possible selective pressures can be accessed by this method (Hillerislambers *et al.*, 2012). Despite this, the method remains a robust tool to infer the high-level deterministic and stochastic processes prevalence in metacommunity assembly. Also, to guarantee that the observed metacommunity follows this assumption, a phylogenetic signal (Losos, 2008) should be tested and show a correlation between environmental factors and ASVs at short phylogenetic distances (Stegen *et al.*, 2013). Similar to what was previously reported for other environments (Stegen *et al.*, 2013; Dini-Andreote *et al.*, 2015; Llames *et al.*, 2017; Huber *et al.*, 2020), the phylogenetic signal was seen here in distinct contexts and scales. Also, this method has the advantage of disentangling environmental data to determine the importance of deterministic factors, which prevents underestimation issues by unmeasured environmental variables.

The partition of the environmental fraction indicated that the pH alone represented more than half of all variation explained by environmental factors (Chapter V of this thesis). While we are starting to describe the impacts of space on the structure of microbial communities (e.g. Mateus-Barros *et al.*, 2021), the importance of the environment is well documented (Baas Becking, 1934). For inland environments, pH has been reported repeatedly as the principal driver of bacterial community composition (Lindström *et al.*, 2005; Niño-García *et al.*, 2016; Mateus-Barros *et al.*, 2021). Some experimental and empiric studies have already demonstrated that this factor can act favoring organisms with preferences for humic and acidic conditions (Lindström *et al.*, 2005). This factor can be especially relevant in tropical regions due to high organic matter decomposition rates (Amado *et al.*, 2013; Freitas *et al.*, 2017; De Melo, Kothawala*, et al.*, 2019), and, in special headwater environments, that frequently receive large volumes of environmental and biological matter from the soil (Ruiz-González *et al.*, 2015). This factor seems to play

a greater role in the dominant bacteria, while other factors, such as dissolved nutrients and physical components, can also impact the rarer ones (Mateus-Barros *et al.*, 2021). In any case, the identification of acidophilic tropical bacteria remains an open field.

At the regional scale, we found that homogenous selection, dispersal limitation combined with Drift and Drift acting alone are the main high-level processes assembling shallow lake metacommunities in three different regions of the world, but the main process varies, which was expected and evidenced by context-dependent metacommunity assembly. The environmental heterogeneity seems to be the principal determinant process in each region (Huber *et al.*, 2020), which, in turn, seems to be determined by the connection between the lake and its surrounding soils, through a mass flow that can homogenize the metacommunity (Ruiz-González *et al.*, 2015) and after filtered by common taxa to freshwater environments (Niño-García *et al.*, 2016). If the mass flux weakens, the metacommunity passes through a heterogeneous selection. Spatial isolation also has its role, but to a weaker degree (Mateus-Barros *et al.*, 2021) and explains part of stochastic processes. In general, Homogenizing Dispersal is not a relevant process shaping communities and sounds to reflect a closer connection between freshwater and soil environments.

In Argentina, we found increased importance of stochastic processes. In this region, shallow lake metacommunity experiences a great variation in the annual precipitation (Diovisalvi *et al.*, 2015). Also, the landscape here is relatively flat; the reduced slopes found on the ground around the lakes can contribute to homogenizing the landscape during events of a significant flood, but also decreases the potential mass flow from the soil during the dryer seasons. These patterns may evidence a greater environmental heterogeneity and contribute to explaining the emergence of assembly processes explained mainly by heterogeneous selection (Huber *et al.*, 2020). We expected to observe the greater importance of stochastic processes in Brazil, which we have not found. Our assumption was based on the fact that biomes that experience increased temperatures may end up showing more prominent unpredictable assembly processes. This should occur basically by metabolic features, which should accelerate the pace of life in these locations and, in consequence, accelerate biological interactions, reproductive events, and so on,

which reduces the time for deterministic processes and increases stochastic features like births and deaths (Saito *et al.*, 2021). An expected consequence should be a greater diversity in warmer realms, but it is not the case here. The decreased bacterial diversity in Brazil may be explained by intense agriculture in this region that reduces the number of soil key taxa (Banerjee *et al.*, 2019). These soil organisms should be carried to water during rain events (Ruiz-González *et al.*, 2015), which contributes to the highest number of detected deterministic and stochastic homogenizing processes. For the Canadian shallow lakes, we found similar numbers of assembly processes guided by homogeneous selection, dispersal limitation, and drift, which partially corroborated our initial hypothesis. In this region, the environmental pressures and mass flow from the soil are huge (Niño-García *et al.*, 2016). Also, this region encompassed six sub-basins (Niño-García *et al.*, 2016) with a higher degree of spatial isolation that can be noted when geographic dissimilarity was related to any other feature. These combined homogenizing and isolation features probably explain the main observed assembly processes.

Concerning the increasing distances, a pattern that arises and is scale-independent is a substitution from the homogeneous selection at smaller scales, through an increment of Drift acting alone, and ending with higher numbers of dispersal limitation acting in concert with Drift. When present, homogenizing dispersal is observed only at smaller distances. The analysis of spatial versus environmental factors influencing a microbial metacommunity usually found that the deterministic factors are the main drivers in these metacommunities (Mateus-Barros *et al.*, 2021). It generally is attributed to the fact that these organisms are great dispersers and, therefore, only environmental factors can determine dissimilarities(Nabout *et al.*, 2009), but more detailed scrutiny of spatial features sheds light on bacterial slight responses to space (Chapter V of this thesis). These responses lead to an alternative interpretation that this weak relationship may occur because the scale of observation is not larger enough. Here, we could observe a pattern that emerges from the increasing scale: the relationship between distances and biological dissimilarity becomes evident from the continental scale. Also, at larger scales, the number of assembly processes assigned to dispersal limitation sounds to be more concentrated in the direction of the upper right side of the graph. These results challenge this

dominant reasoning for microbes, but it is in agreement with other studies focusing on bigger freshwater organisms, where environmental factors were inversely correlated to scale, while spatial determinants were directly correlated with scale (Melo *et al.*, 2012). In this sense, an analysis at even greater scales than presented here should show only assembly processes assigned to dispersal limitation.

Conclusions

Studying the combined impact of deterministic and stochastic assembly processes on metacommunity is gaining attention (Vellend *et al.*, 2014). Here, we could demonstrate that the bacterioplankton can be impacted by deterministic and stochastic processes, simultaneously acting to structure the metacommunities. These two high-level act jointly to a major or minor degree depending on the distance between sites, a pattern also observed by other organisms (Melo *et al.*, 2012). The most important assembly processes at shorter distances are the selection, which homogenizes the metacommunity. This process changes slowly to Drift and, after, to Dispersal limitation combined whit Drift. Contrary to our expectations, the role of spatial scales seems to be to turn this substitution sequence more prominent, by making it more hardly perceivable in smaller-scale contexts.

Observações Finais

O avanço de tecnologias de sequenciamento massivo de DNA vem contribuindo para um significativo avanço das ciências biológicas e ambientais. Em especial, a Ecologia Microbiana passa por um avassalador aumento no número de descobertas sendo produzidas nos últimos anos. Técnicas de sequenciamento independentes de cultura permitem melhor entender como bactérias interagem com o ambiente ao redor (Mateus-Barros *et al.*, 2021), quais tipos de estio de vida são capazes de manter (Chiriac, Haber*, et al.*, 2022) e até quais rotas metabólicas estão a disposição em seu genomas (Ghai *et al.*, 2014; Salcher *et al.*, 2015).

A noção de que estes fantásticos organismos seriam guiados apenas por variações nos fatores locais ambientais já foi superada (Martiny *et al.*, 2006), e crescem as evidências de que questões espaciais também tem seu papel em guiar a distribuição de abundância e biodiversidade bacteriana (Stegen *et al.*, 2013; Dini-Andreote *et al.*, 2015; Martínez *et al.*, 2015; Lindh *et al.*, 2017; Logares *et al.*, 2018; Logares *et al.*, 2020). Por outro lado, estes organismos podem também se provar valiosos em estudos ecológicos mais amplos, já que milhares de sequências de DNA bacteriano podem ser encontradas em apenas uma gota d'água (Barberán *et al.*, 2014).

Na presente tese, busquei explorar como os fatores ambientais e espaciais podem conjuntamente impactar nas dinâmicas que determinam a biodiversidade de bactérias que vivem na coluna d'água de lagoas rasas. Tive a oportunidade de acessar diversas bases de dados coletadas pelo continente Americano e busquei aplicar o maior número de análises estatísticas que pude para buscar entender como funcionam alguns padrões espaciais da diversidade bacteriana.

Espero com este trabalho ter contribuído para o avanço no entendimento de como esses organismos podem estar distribuídos sobre uma paisagem e quais fatores potencialmente guiam sua distribuição.

Referências

ACCATTATIS, V. et al. Identifying invaders: the case of Ceratium furcoides (Gonyaulacales, Dinophyceae) in South America. **Journal of Phycology**, 2020. ISSN 0022-3646.

ACINAS, S. G. et al. Fine-scale phylogenetic architecture of a complex bacterial community. **Nature,** v. 430, n. 6999, p. 551-554, 2004. ISSN 1476-4687.

AHLGREN, N. A. et al. Multi-year dynamics of fine-scale marine cyanobacterial populations are more strongly explained by phage interactions than abiotic, bottom-up factors. **Environ Microbiol,** v. 21, n. 8, p. 2948-2963, Aug 2019. ISSN 1462-2920 (Electronic), 1462-2912 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/31106939 >.

ALBANESE, D. et al. A practical tool for maximal information coefficient analysis. **GigaScience,** v. 7, n. 4, p. giy032, 2018. ISSN 2047-217X.

AMADO, A. M. et al. Tropical freshwater ecosystems have lower bacterial growth efficiency than temperate ones. **Front Microbiol,** v. 4, p. 167, 2013. ISSN 1664-302X (Electronic) 1664-302X (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/23801986 >.

ANDERSON, M. J. et al. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. **Ecology letters,** v. 14, n. 1, p. 19-28, 2011. ISSN 1461-023X.

ANDREI, A.-Ş. et al. Niche-directed evolution modulates genome architecture in freshwater Planctomycetes. **The ISME journal,** v. 13, n. 4, p. 1056-1071, 2019. ISSN 1751-7370.

ANDREWS, M. et al. First confirmed case of COVID-19 infection in India: A case report. **The Indian journal of medical research,** v. 151, n. 5, p. 490, 2020.

APPRILL, A. et al. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. 2015.

AZAM, F. et al. The ecological role of water-column microbes in the sea. **Marine ecology progress series**, p. 257-263, 1983. ISSN 0171-8630.

BAAS BECKING, L. G. M. **Geobiologie of inleiding tot de milieukunde**. Den Haag: Van Stockum, 1934.

BANERJEE, A. K.; BEGUM, F.; RAY, U. Mutation hot spots in Spike protein of COVID-19. **Preprints,** v. 2020, p. 2020040281, 2020.

BANERJEE, S. et al. Agricultural intensification reduces microbial network complexity and the abundance of keystone taxa in roots. **ISME J,** v. 13, n. 7, p. 1722-1736, Jul 2019. ISSN 1751-7370 (Electronic) 1751-7362 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/30850707 >.

BARBERÁN, A.; CASAMAYOR, E. O.; FIERER, N. The microbial contribution to macroecology. **Front Microbiol,** v. 5, p. 203, 2014. ISSN 1664-302X (Print), 1664-302X (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/24829564 >.

BARNES, M. A.; TURNER, C. R. The ecology of environmental DNA and implications for conservation genetics. **Conservation Genetics,** v. 17, n. 1, p. 1-17, 2015. ISSN 1566-0621.

BASELGA, A. Partitioning the turnover and nestedness components of beta diversity. **Global Ecology and Biogeography,** v. 19, n. 1, p. 134-143, 2010. ISSN 1466-8238. Disponível em: < http://dx.doi.org/10.1111/j.1466-8238.2009.00490.x >.

BASELGA, A.; FRECKLETON, R. Separating the two components of abundance-based dissimilarity: balanced changes in abundance vs. abundance gradients. **Methods in Ecology and Evolution,** v. 4, n. 6, p. 552-557, 2013. ISSN 2041210X.

BASELGA, A. et al. betapart: Partitioning Beta Diversity into Turnover and Nestedness Components. 2017. Disponível em: < https://CRAN.R-project.org/package=betapart >.

BATUT, J.; ANDERSSON, S. G.; O'CALLAGHAN, D. The evolution of chronic infection strategies in the α-proteobacteria. **Nature Reviews Microbiology,** v. 2, n. 12, p. 933-945, 2004. ISSN 1740-1534.

BAUMAN, D. et al. Disentangling good from bad practices in the selection of spatial or phylogenetic eigenvectors. **Ecography,** v. 41, n. 10, p. 1638-1649, 2018. ISSN 09067590.

BEISNER, B. E. et al. The role of environmental and spatial processes in structuring lake communities from bacteria to fish. **Ecology,** v. 87, n. 12, p. 2985-2991, 2006. ISSN 1939-9170.

BERG, K. A. et al. High diversity of cultivable heterotrophic bacteria in association with cyanobacterial water blooms. **The ISME journal,** v. 3, n. 3, p. 314-325, 2009. ISSN 1751-7370.

BIŽIĆ, M. et al. Aquatic and terrestrial cyanobacteria produce methane. **Science advances,** v. 6, n. 3, p. eaax5343, 2020. ISSN 2375-2548.

BLANCHET, F. G.; LEGENDRE, P.; BORCARD, D. Forward selection of explanatory variables. **Ecology,** v. 89, n. 9, p. 2623-2632, 2008a. ISSN 1939-9170. Disponível em: < http://dx.doi.org/10.1890/07-0986.1 >.

_____. Modelling directional spatial processes in ecological data. **Ecological Modelling,** v. 215, n. 4, p. 325-336, 2008b. ISSN 03043800.

BORCARD, D.; LEGENDRE, P.; DRAPEAU, P. Partialling out the Spatial Component of Ecological Variation. **Ecology,** v. 73, n. 3, p. 1045-1055, 1992. ISSN 0012-9658. Disponível em: < https://esajournals.onlinelibrary.wiley.com/doi/abs/10.2307/1940179 >.

BROWN, C. T. et al. Unusual biology across a group comprising more than 15% of domain Bacteria. **Nature,** v. 523, n. 7559, p. 208-211, 2015. ISSN 1476-4687.

BROWN, J. H. On the relationship between abundance and distribution of species. **The american naturalist,** v. 124, n. 2, p. 255-279, 1984. ISSN 0003-0147.

BUCHAN, A. et al. Master recyclers: features and functions of bacteria associated with phytoplankton blooms. **Nature Reviews Microbiology,** v. 12, n. 10, p. 686-698, 2014. ISSN 1740-1534.

CABELLO-YEVES, P. J. et al. Elucidating the picocyanobacteria salinity divide through ecogenomics of new freshwater isolates. **BMC biology,** v. 20, n. 1, p. 1-24, 2022. ISSN 1741-7007.

CABELLO-YEVES, P. J. et al. Reconstruction of Diverse Verrucomicrobial Genomes from Metagenome Datasets of Freshwater Reservoirs. **Front Microbiol,** v. 8, p. 2131, 2017. ISSN 1664-302X (Print), 1664-302X (Electronic). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/29163419 >.

CABELLO-YEVES, P. J. et al. Novel Synechococcus genomes reconstructed from freshwater reservoirs. **Frontiers in microbiology,** v. 8, p. 1151, 2017. ISSN 1664-302X.

CAILLON, F. et al. Soil microbial inoculation during flood events shapes headwater stream microbial communities and diversity. **Microb Ecol**, Feb 2 2021. ISSN 1432-184X (Electronic) 0095-3628 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/33532913 >.

CALLAHAN, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. **Nat Methods,** v. 13, n. 7, p. 581-3, Jul 2016. ISSN 1548-7105 (Electronic), 1548-7091 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/27214047 >.

CALLIERI, C. Synechococcus plasticity under environmental changes. **FEMS microbiology letters,** v. 364, n. 23, p. fnx229, 2017. ISSN 1574-6968.

CALLIERI, C.; CABELLO-YEVES, P. J.; BERTONI, F. The "Dark Side" of Picocyanobacteria: Life as We Do Not Know It (Yet). **Microorganisms,** v. 10, n. 3, p. 546, 2022. ISSN 2076-2607.

CAMARA DOS REIS, M. et al. Spatial heterogeneity and hydrological fluctuations drive bacterioplankton community composition in an Amazon floodplain system. **PLoS One,** v. 14, n. 8, p. e0220695, 2019. ISSN 1932-6203 (Electronic), 1932-6203 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/31398199 >.

CAPORASO, J. G. et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. **The ISME journal,** v. 6, n. 8, p. 1621-1624, 2012. ISSN 1751-7370.

CARLSON, R. E. A trophic state index for lakes. **Limnology and Oceanography,** v. 22, n. 2, p. 361-369, 1977. ISSN 1939-5590. Disponível em: < http://dx.doi.org/10.4319/lo.1977.22.2.0361 >.

CARUSO, T. et al. Stochastic and deterministic processes interact in the assembly of desert microbial communities on a global scale. **ISME J,** v. 5, n. 9, p. 1406-13, Sep 2011. ISSN 1751-7370 (Electronic), 1751-7362 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/21368908 >.

CASTELLE, C. J. et al. Biosynthetic capacity, metabolic variety and unusual biology in the CPR and DPANN radiations. **Nature Reviews Microbiology,** v. 16, n. 10, p. 629-645, 2018. ISSN 1740-1534.

CHASE, J. M. et al. Using null models to disentangle variation in community dissimilarity from variation in α-diversity. **Ecosphere,** v. 2, n. 2, p. art24, 2011. ISSN 2150-8925.

CHAVE, J. The problem of pattern and scale in ecology: what have we learned in 20 years? **Ecology Letters,** v. 16 Suppl 1, p. 4-16, May 2013. ISSN 1461-0248 (Electronic), 1461-023X (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/23351093 >.

CHIRIAC, M. C. et al. Ecogenomics sheds light on diverse lifestyle strategies in freshwater CPR. **Microbiome,** v. 10, n. 1, p. 1-21, 2022. ISSN 2049-2618.

CHIRIAC, M. C.; HABER, M.; SALCHER, M. M. Adaptive genetic traits in pelagic freshwater microbes. **Environ Microbiol**, Dec 13 2022. ISSN 1462-2920 (Electronic), 1462-2912 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/36513610 >.

CIOTTI, M. et al. The COVID-19 pandemic. **Critical reviews in clinical laboratory sciences,** v. 57, n. 6, p. 365-388, 2020. ISSN 1040-8363.

COLE, J. J. et al. Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon budget. **Ecosystems,** v. 10, n. 1, p. 172-185, 2007. ISSN 1435-0629.

CORY, R. M. et al. Singlet Oxygen in the Coupled Photochemical and Biochemical Oxidation of Dissolved Organic Matter. **Environmental Science & Technology,** v. 44, n. 10, p. 3683-3689, 2010/05/15 2010. ISSN 0013-936X. Disponível em: < http://dx.doi.org/10.1021/es902989y >.

COUTINHO, L. M. Fire in the ecology of the Brazilian cerrado. In: (Ed.). **Fire in the tropical biota**: Springer, 1990. p.82-105.

COVID-19 Dashboard. Center for Systems Science and Engeneering, Disponível em: < https://coronavirus.jhu.edu/map.html >. Acesso em: 29/12/2022.

CUNHA, D. G. F.; CALIJURI, M. D. C.; LAMPARELLI, M. C. A trophic state index for tropical/subtropical reservoirs (TSItsr). **Ecological Engineering,** v. 60, p. 126-134, 11// 2013. ISSN 0925-8574. Disponível em: < http://www.sciencedirect.com/science/article/pii/S0925857413003091 >.

DE BIE, T. et al. Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. **Ecology letters,** v. 15, n. 7, p. 740-747, 2012. ISSN 1461-023X.

DE MELO, M. L. et al. Flood pulse regulation of bacterioplankton community composition in an Amazonian floodplain lake. **Freshwater Biology,** v. 64, n. 1, p. 108-120, 2019. ISSN 00465070.

DE MELO, M. L. et al. Linking dissolved organic matter composition and bacterioplankton communities in an Amazon floodplain system. **Limnology and Oceanography,** v. 65, p. 63-76, 2019. ISSN 0024-3590 (Electronic), 1939-5590 (Linking).

DINI-ANDREOTE, F. et al. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. **Proceedings of the National Academy of Sciences,** v. 112, n. 11, p. E1326-E1332, 2015. ISSN 0027-8424.

DIOVISALVI, N. et al. Shallow lakes from the Central Plains of Argentina: an overview and worldwide comparative analysis of their basic limnological features. **Hydrobiologia,** v. 752, p. 5-20, 2015. ISSN 0018-8158.

DRAKARE, S.; LENNON, J. J.; HILLEBRAND, H. The imprint of the geographical, evolutionary and ecological context on species-area relationships. **Ecol Lett,** v. 9, n. 2, p. 215-27, Feb 2006. ISSN 1461-0248 (Electronic) 1461-023X (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/16958886 >.

DRAKARE, S.; LIESS, A. Local factors control the community composition of cyanobacteria in lakes while heterotrophic bacteria follow a neutral model. **Freshwater Biology,** v. 55, n. 12, p. 2447-2457, 2010. ISSN 00465070.

DRAY, S. et al. **Package 'adespatial'** 2020.

DRAY, S.; LEGENDRE, P.; PERES-NETO, P. R. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). **Ecological Modelling,** v. 196, n. 3-4, p. 483-493, 2006. ISSN 03043800.

DUFRÊNE, M.; LEGENDRE, P. Species assemblages and indicator species: the need for a flexible asymmetrical approach. **Ecological monographs,** v. 67, n. 3, p. 345-366, 1997. ISSN 1557-7015.

EDRADA, E. M. et al. First COVID-19 infections in the Philippines: a case report. **Tropical medicine and health,** v. 48, p. 1-7, 2020.

EILER, A.; BERTILSSON, S. Composition of freshwater bacterial communities associated with cyanobacterial blooms in four Swedish lakes. **Environmental Microbiology,** v. 6, n. 12, p. 1228-1243, 2004. ISSN 1462-2912. Disponível em: < https://sfamjournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.1462-2920.2004.00657.x >.

_____. Flavobacteria blooms in four eutrophic lakes: linking population dynamics of freshwater bacterioplankton to resource availability. **Applied and Environmental Microbiology,** v. 73, n. 11, p. 3511-3518, 2007. ISSN 0099-2240.

ESCALAS, A. et al. Macroecological distributions of gene variants highlight the functional organization of soil microbial systems. **The ISME Journal,** v. 16, n. 3, p. 726-737, 2022/03/01 2022. ISSN 1751-7370. Disponível em: < https://doi.org/10.1038/s41396-021-01120-8 >.

FARJALLA, V. F. et al. Ecological determinism increases with organism size. **Ecology,** v. 93, n. 7, p. 1752-1759, 2012. ISSN 1939-9170. Disponível em: < http://dx.doi.org/10.1890/11-1144.1 >.

FERNÁNDEZ, L. D. et al. Geographical distance and local environmental conditions drive the genetic population structure of a freshwater microalga (Bathycoccaceae; Chlorophyta) in Patagonian lakes. **FEMS Microbiology Ecology,** v. 93, n. 10, 2017. ISSN 0168-6496. Disponível em: < https://doi.org/10.1093/femsec/fix125 >. Acesso em: 3/19/2023.

FILLINGER, L.; HUG, K.; GRIEBLER, C. Selection imposed by local environmental conditions drives differences in microbial community composition across geographically distinct groundwater aquifers. **FEMS Microbiology Ecology,** v. 95, n. 11, 2019. ISSN 0168-6496

1574-6941.

FINDLAY, S. E. et al. Metabolic and structural response of hyporheic microbial communities to variations in supply of dissolved organic matter. **Limnology and oceanography,** v. 48, n. 4, p. 1608-1617, 2003. ISSN 0024-3590.

FONTE, L. F. M. D. et al. Amphibian diversity in the Amazonian floating meadows: a Hanski core-satellite species system. **Ecography,** v. 44, n. 9, p. 1325-1340, 2021. ISSN 0906-7590. Disponível em: < https://onlinelibrary.wiley.com/doi/abs/10.1111/ecog.05610 >.

FRANZ, M. et al. Cytoscape. js: a graph theory library for visualisation and analysis. **Bioinformatics,** v. 32, n. 2, p. 309-311, 2016. ISSN 1460-2059.

FREITAS, R. et al. Productivity and rainfall drive bacterial metabolism in tropical cascading reservoirs. **Hydrobiologia,** v. 809, n. 1, p. 233-246, 2017. ISSN 0018-8158

1573-5117.

GASOL, J. M. et al. A transplant experiment to identify the factors controlling bacterial abundance, activity, production, and community composition in a eutrophic canyon-shaped reservoir. **Limnology and Oceanography,** v. 47, n. 1, p. 62-77, 2002. ISSN 0024-3590.

GASTON, K. J. et al. Abundance–occupancy relationships. **Journal of Applied Ecology,** v. 37, n. s1, p. 39-59, 2000. ISSN 1365-2664.

GHAI, R.; MCMAHON, K. D.; RODRIGUEZ-VALERA, F. Breaking a paradigm: cosmopolitan and abundant freshwater actinobacteria are low GC. **Environmental microbiology reports,** v. 4, n. 1, p. 29-35, 2012. ISSN 1758-2229.

GHAI, R. et al. Key roles for freshwater Actinobacteria revealed by deep metagenomic sequencing. **Mol Ecol,** v. 23, n. 24, p. 6073-90, Dec 2014. ISSN 1365-294X (Electronic), 0962-1083 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/25355242 >.

GILBERT, B.; BENNETT, J. R. Partitioning variation in ecological communities: do the numbers add up? **Journal of Applied Ecology,** v. 47, n. 5, p. 1071-1082, 2010. ISSN 00218901.

GIOVANNONI, S. J. SAR11 Bacteria: The Most Abundant Plankton in the Oceans. **Ann Rev Mar Sci,** v. 9, p. 231-255, Jan 3 2017. ISSN 1941-0611 (Electronic), 1941-0611 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/27687974 >.

GIOVANNONI, S. J. et al. Proteorhodopsin in the ubiquitous marine bacterium SAR11. **Nature,** v. 438, n. 7064, p. 82-85, 2005. ISSN 1476-4687.

GLEASON, H. The significance of Raunkiaer's law of frequency. **Ecology,** v. 10, n. 4, p. 406-408, 1929. ISSN 0012-9658.

GLOOR, G. B. et al. Microbiome Datasets Are Compositional: And This Is Not Optional. **Front Microbiol,** v. 8, p. 2224, 2017. ISSN 1664-302X (Print) 1664-302X (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/29187837 >.

GREEN, J.; BOHANNAN, B. J. Spatial scaling of microbial biodiversity. **Trends Ecol Evol,** v. 21, n. 9, p. 501-7, Sep 2006. ISSN 0169-5347 (Print), 0169-5347 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/16815589 >.

GROSSART, H.-P. et al. Bacteria dispersal by hitchhiking on zooplankton. **Proceedings of the National Academy of Sciences,** v. 107, n. 26, p. 11959-11964, 2010. ISSN 0027-8424.

GUTIERREZ, T. Marine, Aerobic Hydrocarbon-Degrading Gammaproteobacteria: Overview. In: MCGENITY, T. J. (Ed.). **Taxonomy, Genomics and Ecophysiology of Hydrocarbon-Degrading Microbes**. Cham: Springer International Publishing, 2019. p.143-152. ISBN 978-3-030-14796-9.

HAHN, M. W. Isolation of strains belonging to the cosmopolitan Polynucleobacter necessarius cluster from freshwater habitats located in three climatic zones. **Appl Environ Microbiol,** v. 69, n. 9, p. 5248-54, 2003. ISSN 0099-2240 (Print), 0099-2240 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/12957910 >.

HAHN, M. W. Description of seven candidate species affiliated with the phylum Actinobacteria, representing planktonic freshwater bacteria. **International journal of systematic and evolutionary microbiology,** v. 59, n. 0 1, p. 112, 2009.

HAHN, M. W. et al. Opening a next-generation black box: Ecological trends for hundreds of species-like taxa uncovered within a single bacterial> 99% 16S rRNA operational taxonomic unit. **Molecular ecology resources,** v. 21, n. 7, p. 2471-2485, 2021. ISSN 1755-098X.

HAHN, M. W. et al. Complete ecological isolation and cryptic diversity in Polynucleobacter bacteria not resolved by 16S rRNA gene sequences. **The ISME journal,** v. 10, n. 7, p. 1642-1655, 2016. ISSN 1751-7370.

HAHN, M. W. et al. Fourteen new Polynucleobacter species: P. brandtiae sp. nov., P. kasalickyi sp. nov., P. antarcticus sp. nov., P. arcticus sp. nov., P. tropicus sp. nov., P. bastaniensis sp. nov., P. corsicus sp. nov., P. finlandensis sp. nov., P. ibericus sp. nov., P. hallstattensis sp. nov., P. alcilacus sp. nov., P. nymphae sp. nov., P. paludilacus sp. nov. and P. parvulilacunae sp. nov. **International Journal of Systematic and Evolutionary Microbiology,** v. 72, n. 6, p. 005408, 2022. ISSN 1466-5026.

HAHN, M. W.; PÖCKL, M. Ecotypes of planktonic Actinobacteria with identical 16S rRNA genes adapted to thermal niches in temperate, subtropical, and tropical freshwater habitats. **Applied and environmental microbiology,** v. 71, n. 2, p. 766-773, 2005. ISSN 0099-2240.

HAHN, M. W. et al. Rhodoluna lacicola gen. nov., sp. nov., a planktonic freshwater bacterium with stream-lined genome. **International journal of systematic and evolutionary microbiology,** v. 64, n. Pt 9, p. 3254, 2014.

HANSEN, A. M. et al. Optical properties of dissolved organic matter (DOM): Effects of biological and photolytic degradation. **Limnology and Oceanography**, p. n/a-n/a, 2016. ISSN 00243590.

HANSKI, I. Dynamics of Regional Distribution: The Core and Satellite Species Hypothesis. **Oikos,** v. 38, n. 2, p. 210-221, 1982. ISSN 00301299, 16000706. Disponível em: < http://www.jstor.org/stable/3544021 >.

HARIDASAN, M. Nutritional adaptations of native plants of the cerrado biome in acid soils. **Brazilian Journal of Plant Physiology,** v. 20, p. 183-195, 2008. ISSN 1677-9452.

HAVENS, K. E. Cyanobacteria blooms: effects on aquatic ecosystems. **Cyanobacterial harmful algal blooms: state of the science and research needs**, p. 733-747, 2008.

HECKMANN, K.; SCHMIDT, H. J. Polynucleobacter necessarius gen. nov., sp. nov., an obligately endosymbiotic bacterium living in the cytoplasm of Euplotes aediculatus. **International Journal of Systematic and Evolutionary Microbiology,** v. 37, n. 4, p. 456-457, 1987. ISSN 1466-5026.

HELMS, J. R. et al. Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. **Limnology and Oceanography,** v. 53, n. 3, p. 955-969, 2008.

HENSON, M. W. et al. Cultivation and genomics of the first freshwater SAR11 (LD12) isolate. **The ISME journal,** v. 12, n. 7, p. 1846-1860, 2018. ISSN 1751-7370.

HERLEMANN, D. P. et al. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. **The ISME journal,** v. 5, n. 10, p. 1571-1579, 2011. ISSN 1751-7362.

HILLERISLAMBERS, J. et al. Rethinking Community Assembly through the Lens of Coexistence Theory. **Annual Review of Ecology, Evolution, and Systematics,** v. 43, n. 1, p. 227-248, 2012. ISSN 1543-592X

1545-2069.

HOETZINGER, M. et al. Continental-Scale Gene Flow Prevents Allopatric Divergence of Pelagic Freshwater Bacteria. **Genome Biology and Evolution,** v. 13, n. 3, 2021. ISSN 1759-6653. Disponível em: < https://doi.org/10.1093/gbe/evab019 >. Acesso em: 12/27/2022.

HOETZINGER, M. et al. Microdiversification of a pelagic Polynucleobacter species is mainly driven by acquisition of genomic islands from a partially interspecific gene pool. **Applied and environmental microbiology,** v. 83, n. 3, p. e02266-16, 2017. ISSN 0099-2240.

HOETZINGER, M. et al. Polynucleobacter paneuropaeus sp. nov., characterized by six strains isolated from freshwater lakes located along a 3000 km north-south cross-section across Europe. **Int J Syst Evol Microbiol,** v. 69, n. 1, p. 203-213, 2019. ISSN 1466-5026 (Print).

HORNER-DEVINE, M. C.; BOHANNAN, B. J. Unifying ecology to include all creatures great and small. **Trends in ecology & evolution,** v. 21, n. 9, p. 473, 2006. ISSN 0169-5347.

HORNER-DEVINE, M. C.; CARNEY, K. M.; BOHANNAN, B. J. An ecological perspective on bacterial biodiversity. **Proc Biol Sci,** v. 271, n. 1535, p. 113-22, 2003. ISSN 0962-8452 (Print), 0962-8452 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/15058386 >.

HORNER-DEVINE, M. C. et al. A taxa-area relationship for bacteria. **Nature,** v. 432, n. 7018, p. 750-753, 12/09/print 2004. ISSN 0028-0836. Disponível em: < http://dx.doi.org/10.1038/nature03073 >.

HUANG, J. et al. Pharmacological therapeutics targeting RNA-dependent RNA polymerase, proteinase and spike protein: from mechanistic studies to clinical trials for COVID-19. **Journal of clinical medicine,** v. 9, n. 4, p. 1131, 2020. ISSN 2077-0383.

HUBBELL, S. P. **The Unified Neutral Theory of Biodiversity and Biogeography.** 2001.

HUBER, P. et al. Environmental heterogeneity determines the ecological processes that govern bacterial metacommunity assembly in a floodplain river system. **ISME J**, Jul 27 2020. ISSN 1751-7370 (Electronic), 1751-7362 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/32719401 >.

HUGERTH, L. W. et al. Metagenome-assembled genomes uncover a global brackish microbiome. **Genome biology,** v. 16, n. 1, p. 1-18, 2015. ISSN 1474-760X.

HUISMAN, J. et al. Cyanobacterial blooms. **Nature Reviews Microbiology,** v. 16, n. 8, p. 471-483, 2018. ISSN 1740-1534.

HUISMAN, J.; HULOT, F. D. Population dynamics of harmful cyanobacteria. In: (Ed.). **Harmful cyanobacteria**: Springer, 2005. p.143-176.

HUTCHINSON, G. E. Homage to Santa Rosalia or why are there so many kinds of animals? **The American Naturalist,** v. 93, n. 870, p. 145-159, 1959. ISSN 0003-0147.

IRIONDO, M. Quaternary lakes of Argentina. **Palaeogeography, Palaeoclimatology, Palaeoecology,** v. 70, n. 1-3, p. 81-88, 1989. ISSN 0031-0182.

JENKINS, D. G.; RICKLEFS, R. E. Biogeography and ecology: two views of one world. **Philosophical Transactions of the Royal Society B: Biological Sciences,** v. 366, n. 1576, p. 2331-2335, 2011.

JEONG, S. Y.; CHOI, J. Y.; KIM, T. G. Coordinated Metacommunity Assembly and Spatial Distribution of Multiple Microbial Kingdoms within a Lake. **Microb Ecol,** v. 79, p. 801-814, Nov 8 2020. ISSN 1432-184X (Electronic) 0095-3628 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/31705158 >.

JEZBEROVÁ, J. et al. The Limnohabitans Genus Harbors Generalistic and Opportunistic Subtypes: Evidence from Spatiotemporal Succession in a Canyon-Shaped Reservoir. **Applied and Environmental Microbiology,** v. 83, n. 21, p. e01530-17, 2017. Disponível em: < https://journals.asm.org/doi/abs/10.1128/AEM.01530-17 >.

JEZBEROVÁ, J.; KOMÁRKOVÁ, J. Morphological transformation in a freshwater Cyanobium sp. induced by grazers. **Environmental microbiology,** v. 9, n. 7, p. 1858-1862, 2007. ISSN 1462-2912.

JOLY, C. A.; METZGER, J. P.; TABARELLI, M. Experiences from the Brazilian Atlantic Forest: ecological findings and conservation initiatives. **New phytologist,** v. 204, n. 3, p. 459-473, 2014. ISSN 0028-646X.

JYRKÄNKALLIO-MIKKOLA, J. et al. Disentangling multi-scale environmental effects on stream microbial communities. **Journal of Biogeography**, p. n/a-n/a, 2017. ISSN 1365-2699. Disponível em: < http://dx.doi.org/10.1111/jbi.13002 >.

KANG, I. et al. The first complete genome sequences of the acI lineage, the most abundant freshwater Actinobacteria, obtained by whole-genome-amplification of dilution-to-extinction cultures. **Scientific reports,** v. 7, n. 1, p. 1-14, 2017. ISSN 2045-2322.

KANG, I. et al. The first complete genome sequences of the acI lineage, the most abundant freshwater Actinobacteria, obtained by whole-genome-amplification of dilution-to-extinction cultures. **Sci Rep,** v. 7, p. 42252, 2017. ISSN 2045-2322 (Electronic), 2045-2322 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/28186143 >.

KANG, I. et al. **Genome sequence of "Candidatus Aquiluna" sp. strain IMCC13023, a marine member of the Actinobacteria isolated from an arctic fjord**: Am Soc Microbiol 2012.

KASALICKÝ, V. et al. The diversity of the Limnohabitans genus, an important group of freshwater bacterioplankton, by characterization of 35 isolated strains. **PloS one,** v. 8, n. 3, p. e58209, 2013. ISSN 1932-6203.

KASALICKÝ, V. et al. Aerobic Anoxygenic Photosynthesis Is Commonly Present within the Genus Limnohabitans. **Applied and Environmental Microbiology,** v. 84, n. 1, p. e02116-17, 2018. Disponível em: < https://journals.asm.org/doi/abs/10.1128/AEM.02116-17 >.

KAVAGUTTI, V. S. et al. Phage-centric ecological interactions in aquatic ecosystems revealed through ultra-deep metagenomics. **Microbiome,** v. 7, n. 1, p. 1-15, 2019. ISSN 2049-2618.

KAWASAKI, N.; BENNER, R. Bacterial release of dissolved organic matter during cell growth and decline: molecular origin and composition. **Limnology and Oceanography,** v. 51, n. 5, p. 2170-2180, 2006. ISSN 0024-3590.

KEHOE, D. M.; GUTU, A. Responding to color: the regulation of complementary chromatic adaptation. **Annu. Rev. Plant Biol.,** v. 57, p. 127-150, 2006. ISSN 1543-5008.

KEMBEL, S. W. et al. Picante: R tools for integrating phylogenies and ecology. **Bioinformatics,** v. 26, n. 11, p. 1463-1464, 2010. ISSN 1460-2059.

KHAN, N. A.; SIDDIQUI, R. Predator vs aliens: bacteria interactions with Acanthamoeba. **Parasitology,** v. 141, n. 7, p. 869-874, 2014. ISSN 0031-1820.

KIM, B. R. et al. Deciphering Diversity Indices for a Better Understanding of Microbial Communities. **J Microbiol Biotechnol,** v. 27, n. 12, p. 2089-2093, 2017. ISSN 1738-8872 (Electronic), 1017-7825 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/29032640 >.

KIM, S. et al. Culturing the ubiquitous freshwater actinobacterial acI lineage by supplying a biochemical 'helper'catalase. **The ISME journal,** v. 13, n. 9, p. 2252-2263, 2019. ISSN 1751-7370.

KOMÁREK, J.; KOPECKÝ, J.; CEPÁK, V. Generic characters of the simplest cyanoprokaryotes Cyanobium, Cyanobacterium and Synechococcus. **Cryptogamie Algologie,** v. 20, n. 3, p. 209-222, 1999. ISSN 0181-1568.

KOMURA, T. et al. DNA barcoding reveals seasonal shifts in diet and consumption of deep-sea fishes in wedge-tailed shearwaters. **PLoS One,** v. 13, n. 4, p. e0195385, 2018. ISSN 1932-6203.

KOTHAWALA, D. N. et al. Inner filter correction of dissolved organic matter fluorescence. **Limnology and Oceanography: Methods,** v. 11, n. 12, p. 616-630, 2013. ISSN 1541-5856.

KOVÁCS, A. W.; TÓTH, V. R.; PÁLFFY, K. The effects of interspecific interactions between bloom forming cyanobacteria and Scenedesmus quadricauda (chlorophyta) on their photophysiology. **Acta Biologica Hungarica,** v. 69, n. 2, p. 210-223, 2018. ISSN 0236-5383.

KUHN, T. S. **A estrutura das revoluções científicas**. 12° edição. Editora Perspectiva, 2013. ISBN 9897026789.

KWON, D. R. et al. New records of the genus Cyanobium and Cyanobium gracile (Synechococcales, Cyanophyceae) in Korean freshwater. **Korean Journal of Environmental Biology,** v. 39, n. 1, p. 32-38, 2021. ISSN 1226-9999.

LAKOWICZ, J. R. **Principles of fluorescence spectroscopy**. Springer, 2006.

LANGENHEDER, S. et al. Temporal variation of beta-diversity and assembly mechanisms in a bacterial metacommunity. **ISME J,** v. 6, n. 6, p. 1107-14, Jun 2012. ISSN 1751-7370 (Electronic) 1751-7362 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/22158394 >.

LANSAC-TÔHA, F. M. et al. Scale-dependent patterns of metacommunity structuring in aquatic organisms across floodplain systems. **Journal of Biogeography,** v. 48, n. 4, p. 872-885, 2020. ISSN 0305-0270 (Electronic) 1365-2699 (Linking).

LAWAETZ, A. J.; STEDMON, C. A. Fluorescence intensity calibration using the Raman scatter peak of water. **Applied spectroscopy,** v. 63, n. 8, p. 936-940, 2009. ISSN 0003-7028.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature,** v. 521, n. 7553, p. 436-44, 2015. ISSN 1476-4687 (Electronic), 0028-0836 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/26017442 >.

LEGENDRE, P.; ANDERSON, M. J. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. **Ecological monographs,** v. 69, n. 1, p. 1-24, 1999. ISSN 1557-7015.

LEGENDRE, P.; LEGENDRE, L. **Numerical ecology**. Elsevier, 2012. ISBN 0444538690.

LEIBOLD, M. A. et al. The metacommunity concept: a framework for multi-scale community ecology. **Ecology Letters,** v. 7, n. 7, p. 601-613, 2004. ISSN 1461023X, 14610248.

LEIGH JR, E. G. Neutral theory: a historical perspective. **Journal of Evolutionary Biology,** v. 20, n. 6, p. 2075-2091, 2007. ISSN 1010-061X. Disponível em: < https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1420-9101.2007.01410.x >.

LEMARCHAND, C. et al. Community composition and activity of prokaryotes associated to detrital particles in two contrasting lake ecosystems. **FEMS microbiology ecology,** v. 57, n. 3, p. 442-451, 2006. ISSN 1574-6941.

LEVIN, S. A. The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. **Ecology,** v. 73, n. 6, p. 1943-1967, 1992. ISSN 1939-9170.

LEVINS, R. Some demographic and genetic consequences of environmental heterogeneity for biological control. **Bulletin of the Entomological society of America,** v. 15, n. 3, p. 237-240, 1969. ISSN 0013-8754.

LINDH, M. V. There And Back Again -- Unraveling Mechanisms Of Bacterial Biogeography In The North Pacific Subtropical Gyre To And From Station ALOHA. **bioRxiv**, 2017.

LINDH, M. V. et al. Metapopulation theory identifies biogeographical patterns among core and satellite marine bacteria scaling from tens to thousands of kilometers. **Environ Microbiol,** v. 19, n. 3, p. 1222-1236, Mar 2017. ISSN 1462-2920 (Electronic), 1462-2912 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/28028880 >.

LINDSTRÖM, E. S.; KAMST-VAN AGTERVELD, M. P.; ZWART, G. Distribution of typical freshwater bacterial groups is associated with pH, temperature, and lake water retention time. **Applied and environmental microbiology,** v. 71, n. 12, p. 8201-8206, 2005. ISSN 0099-2240.

LIPKO, I.; BELYKH, O. Environmental features of freshwater planktonic actinobacteria. **Contemporary Problems of Ecology,** v. 14, n. 2, p. 158-170, 2021. ISSN 1995-4263.

LIU, Y.-X. et al. A practical guide to amplicon and metagenomic analysis of microbiome data. **Protein & Cell,** v. 12, n. 5, p. 315-330, 2021/05/01 2021. ISSN 1674-8018. Disponível em: < https://doi.org/10.1007/s13238-020-00724-8 >.

LIVINGSTON, G. et al. Predators regulate prey species sorting and spatial distribution in microbial landscapes. **J Anim Ecol**, Jan 31 2017. ISSN 1365-2656 (Electronic) 0021-8790 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/28138991 >.

LLAMES, M. E. et al. Interplay between stochastic and deterministic processes in the maintenance of alternative community states in Verrucomicrobia-dominated shallow lakes. **FEMS Microbiol Ecol,** v. 93, n. 7, 2017. ISSN 1574-6941 (Electronic), 0168-6496 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/28582516 >.

LOGARES, R. et al. Infrequent transitions between saline and fresh waters in one of the most abundant microbial lineages (SAR11). **Molecular biology and evolution,** v. 27, n. 2, p. 347-357, 2010. ISSN 1537-1719.

LOGARES, R. et al. Disentangling the mechanisms shaping the surface ocean microbiota. **Microbiome,** v. 8, n. 1, p. 55, 2020. ISSN 2049-2618 (Electronic), 2049-2618 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/32312331 >.

LOGARES, R. et al. Contrasting prevalence of selection and drift in the community structuring of bacteria and microbial eukaryotes. **Environ Microbiol**, 2018. ISSN 1462-2920 (Electronic), 1462-2912 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/29727053 >.

LOPES, V. G. et al. Predicting temporal variation in zooplankton beta diversity is challenging. **PLoS One,** v. 12, n. 11, p. e0187499, 2017. ISSN 1932-6203 (Electronic)

1932-6203 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/29095892 >.

LORENZEN, C. J. Determination of chlorophyll and pheo-pigments: spectrophotometric equations. **Limnology and oceanography,** v. 12, n. 2, p. 343-346, 1967. ISSN 1939-5590.

LOSOS, J. B. Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. **Ecol Lett,** v. 11, n. 10, p. 995-1003, Oct 2008. ISSN 1461-0248 (Electronic) 1461-023X (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/18673385 >.

LOWE, W. H.; MCPEEK, M. A. Is dispersal neutral? **Trends Ecol Evol,** v. 29, n. 8, p. 444-50, Aug 2014. ISSN 1872-8383 (Electronic) 0169-5347 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/24962790 >.

MAIER, R. M. Chapter 16 - Biogeochemical Cycling. In: PEPPER, I. L.;GERBA, C. P.*, et al* (Ed.). **Environmental Microbiology (Third Edition)**. San Diego: Academic Press, 2015. p.339-373. ISBN 978-0-12-394626-3.

MARKER, A. The measurement of photosynthetic pigments in freshwaters and standardization of methods: conclusions and recommendations. **Arch. Hydrobiol. Ergebn. Limnol.,** v. 14, p. 91-106, 1980.

MARTIN, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. **EMBnet. journal,** v. 17, n. 1, p. 10-12, 2011. ISSN 2226-6089.

MARTINEZ-GARCIA, M. et al. High-throughput single-cell sequencing identifies photoheterotrophs and chemoautotrophs in freshwater bacterioplankton. **The ISME journal,** v. 6, n. 1, p. 113-123, 2012. ISSN 1751-7370.

MARTÍNEZ, I. et al. The gut microbiota of rural papua new guineans: composition, diversity patterns, and ecological processes. **Cell Rep,** v. 11, n. 4, p. 527-38, Apr 28 2015. ISSN 2211-1247 (Electronic). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/25892234 >.

MARTINY, J. B. H. et al. Microbial biogeography: putting microorganisms on the map. **Nat Rev Micro,** v. 4, n. 2, p. 102-112, 02//print 2006. ISSN 1740-1526. Disponível em: < http://dx.doi.org/10.1038/nrmicro1341 >.

MARTINY, J. B. H. et al. Drivers of bacterial β-diversity depend on spatial scale. **Proceedings of the National Academy of Sciences of the United States of America,** v. 108, n. 19, p. 7850-7854, 2011. Disponível em: < https://www.scopus.com/inward/record.uri?eid=2-s2.0-79956366900&partnerID=40&md5=7666a2fb87e8376d6fe5ac7fba32ddd5 >.

MATEUS-BARROS, E. Macroecologia Microbiana: dispersão bacteriana em lagos rasos distribuídos pelo estado de São Paulo. 2018.

MATEUS-BARROS, E. et al. Local and Geographic Factors Shape the Occupancy-Frequency Distribution of Freshwater Bacteria. **Microbial Ecology**, 2021. ISSN 1432-184X. Disponível em: < https://doi.org/10.1007/s00248-020-01560-3 >.

MATEUS-BARROS, E. et al. Comparison of two DNA extraction methods widely used in aquatic microbial ecology. **J Microbiol Methods,** v. 159, p. 12-17, Feb 6 2019. ISSN 1872-8359 (Electronic) 0167-7012 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/30738110 >.

MCKNIGHT, D. M. et al. Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. **Limnology and Oceanography,** v. 46, n. 1, p. 38-48, 2001. ISSN 0024-3590.

MEHRANVAR, L.; JACKSON, D. A. History and taxonomy: their roles in the core-satellite hypothesis. **Oecologia,** v. 127, n. 1, p. 131-142, Mar 2001. ISSN 1432-1939 (Electronic) 0029-8549 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/28547164 >.

MEHRSHAD, M. et al. Genome reconstruction from metagenomic data sets reveals novel microbes in the brackish waters of the Caspian Sea. **Applied and environmental microbiology,** v. 82, n. 5, p. 1599-1612, 2016. ISSN 0099-2240.

MELO, A. S. et al. Focusing on variation: methods and applications of the concept of beta diversity in aquatic ecosystems. **Acta Limnologica Brasiliensia,** v. 23, n. 3, p. 318-331, 2012. ISSN 2179-975X

0102-6712.

METZ, S. et al. A georeferenced rRNA amplicon database of aquatic microbiomes from South America. **Scientific Data,** v. 9, n. 1, p. 565, 2022/09/13 2022. ISSN 2052-4463. Disponível em: < https://doi.org/10.1038/s41597-022-01665-z >.

MITCHELL-OLDS, T.; SHAW, R. G. REGRESSION ANALYSIS OF NATURAL SELECTION: STATISTICAL INFERENCE AND BIOLOGICAL INTERPRETATION. **Evolution,** v. 41, n. 6, p. 1149-1161, 1987. ISSN 1558-5646. Disponível em: < http://dx.doi.org/10.1111/j.1558-5646.1987.tb02457.x >.

MIZUNO, C. M.; RODRIGUEZ-VALERA, F.; GHAI, R. Genomes of planktonic Acidimicrobiales: widening horizons for marine Actinobacteria by metagenomics. **mBio,** v. 6, n. 1, Feb 10 2015. ISSN 2150-7511 (Electronic). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/25670777 >.

MORRIS, R. M. et al. SAR11 clade dominates ocean surface bacterioplankton communities. **Nature,** v. 420, n. 6917, p. 806-810, 2002. ISSN 1476-4687.

MULLIS, K. et al. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. **Biotechnology Series**, p. 17-17, 1992. ISSN 0740-7378.

MURPHY, K. R. et al. Measurement of dissolved organic matter fluorescence in aquatic environments: an interlaboratory comparison. **Environmental science & technology,** v. 44, n. 24, p. 9405-9412, 2010. ISSN 0013-936X.

MUSH, E. Comparison of different methods for chlorophyll and phaeopigment determination. **Arch. Hydrobiol. Beih,** v. 14, p. 14-36, 1980.

MYKLESTAD, S. M. Dissolved organic carbon from phytoplankton. In: (Ed.). **Marine chemistry**: Springer, 2000. p.111-148.

NABOUT, J. C. et al. No evidence for environmental and spatial processes in structuring phytoplankton communities. **Acta Oecologica,** v. 35, n. 5, p. 720-726, 2009. ISSN 1146609X.

NEUENSCHWANDER, S. M. et al. Microdiversification in genome-streamlined ubiquitous freshwater Actinobacteria. **ISME J,** v. 12, n. 1, p. 185-198, Jan 2018. ISSN 1751-7370 (Electronic) 1751-7362 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/29027997 >.

NEWTON, R. J. et al. A guide to the natural history of freshwater lake bacteria. **Microbiol Mol Biol Rev,** v. 75, n. 1, p. 14-49, Mar 2011. ISSN 1098-5557 (Electronic), 1092-2172 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/21372319 >.

NIÑO-GARCÍA, J. P.; RUIZ-GONZÁLEZ, C.; DEL GIORGIO, P. A. Interactions between hydrology and water chemistry shape bacterioplankton biogeography across boreal freshwater networks. **ISME J**, 02/05/online 2016. ISSN 1751-7370. Disponível em: < http://dx.doi.org/10.1038/ismej.2015.226 >.

OHNO, T. Fluorescence inner-filtering correction for determining the humification index of dissolved organic matter. **Environmental science & technology,** v. 36, n. 4, p. 742-746, 2002. ISSN 0013-936X.

OKSANEN, J. et al. Vegan: community ecology package. 2016.

PALIWAL, D. **Identification and characterisation of new aphid killing bacteria for use as biological pest control agents**. 2017.   University of Reading

PAPP, L.; IZSÁK, J. Bimodality in occurrence classes: a direct consequence of lognormal or logarithmic series distribution of abundances: a numerical experimentation. **Oikos,** v. 79, n. 1, p. 191-194, 1997. ISSN 0030-1299.

PARLANTI, E. et al. Dissolved organic matter fluorescence spectroscopy as a tool to estimate biological activity in a coastal zone submitted to anthropogenic inputs. **Organic geochemistry,** v. 31, n. 12, p. 1765-1781, 2000. ISSN 0146-6380.

PARVEEN, B. et al. Temporal dynamics and phylogenetic diversity of free-living and particle-associated Verrucomicrobia communities in relation to environmental variables in a mesotrophic lake. **FEMS Microbiology Ecology,** v. 83, n. 1, p. 189-201, 2013. ISSN 0168-6496. Disponível em: < https://doi.org/10.1111/j.1574-6941.2012.01469.x >. Acesso em: 12/27/2022.

PASCUAL-GARCÍA, A.; TAMAMES, J.; BASTOLLA, U. Bacteria dialog with Santa Rosalia: Are aggregations of cosmopolitan bacteria mainly explained by habitat filtering or by ecological interactions? **BMC microbiology,** v. 14, n. 1, p. 284, 2014. ISSN 1471-2180.

PEDRÓS-ALIÓ, C. Marine microbial diversity: can it be determined? **Trends Microbiol,** v. 14, n. 6, p. 257-63, 2006. ISSN 0966-842X (Print), 0966-842X (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/16679014 >.

_____. The rare bacterial biosphere. **Ann Rev Mar Sci,** v. 4, p. 449-66, 2012. ISSN 1941-1405 (Print), 1941-0611 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/22457983 >.

PERES-NETO, P. R.; LEGENDRE, P. Estimating and controlling for spatial structure in the study of ecological communities. **Global Ecology and Biogeography,** v. 19, n. 2, p. 174-184, 2010. ISSN 1466822X.

PERNTHALER, J. et al. Predator-specific enrichment of actinobacteria from a cosmopolitan freshwater clade in mixed continuous culture. **Applied and Environmental Microbiology,** v. 67, n. 5, p. 2145-2155, 2001. ISSN 1098-5336.

PERNTHALER, J. et al. Bloom of filamentous bacteria in a mesotrophic lake: identity and potential controlling mechanism. **Applied and Environmental Microbiology,** v. 70, n. 10, p. 6272-6281, 2004. ISSN 0099-2240.

PITT, A. et al. Aquiluna borgnonia gen. nov., sp. nov., a member of a Microbacteriaceae lineage of freshwater bacteria with small genome sizes. **International Journal of Systematic and Evolutionary Microbiology,** v. 71, n. 5, p. 004825, 2021. ISSN 1466-5026.

PIWOSZ, K. et al. Determining lineage-specific bacterial growth curves with a novel approach based on amplicon reads normalization using internal standard (ARNIS). **The ISME Journal,** v. 12, n. 11, p. 2640-2654, 2018/11/01 2018. ISSN 1751-7370. Disponível em: < https://doi.org/10.1038/s41396-018-0213-y >.

PRICE, M. N.; DEHAL, P. S.; ARKIN, A. P. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. **PLOS ONE,** v. 5, n. 3, p. e9490, 2010. Disponível em: < https://doi.org/10.1371/journal.pone.0009490 >.

PROPS, R.; DENEF, V. J. Temperature and Nutrient Levels Correspond with Lineage-Specific Microdiversification in the Ubiquitous and Abundant Freshwater Genus *Limnohabitans*. **Applied and Environmental Microbiology,** v. 86, n.

10, p. e00140-20, 2020. Disponível em: < https://journals.asm.org/doi/abs/10.1128/AEM.00140-20 >.

PRUESSE, E.; PEPLIES, J.; GLÖCKNER, F. O. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. **Bioinformatics,** v. 28, n. 14, p. 1823-1829, 2012. ISSN 1367-4803. Disponível em: < https://doi.org/10.1093/bioinformatics/bts252 >. Acesso em: 12/24/2022.

QUAST, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. **Nucleic Acids Research,** v. 41, p. D590-D596, 2013.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria: R Foundation for Statistical Computing 2019.

RAPPÉ, M. S.; GIOVANNONI, S. J. The uncultured microbial majority. **Annual review of microbiology,** v. 57, n. 1, p. 369-394, 2003. ISSN 0066-4227.

RASILO, T.; PRAIRIE, Y. T.; DEL GIORGIO, P. A. Large-scale patterns in summer diffusive CH4 fluxes across boreal lakes, and contribution to diffusive C emissions. **Glob Chang Biol,** v. 21, n. 3, p. 1124-39, Mar 2015. ISSN 1365-2486 (Electronic)

1354-1013 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/25220765 >.

RATTER, J. A.; RIBEIRO, J. F.; BRIDGEWATER, S. The Brazilian cerrado vegetation and threats to its biodiversity. **Annals of botany,** v. 80, n. 3, p. 223-230, 1997. ISSN 0305-7364.

ROBERTS, D. W.; ROBERTS, M. D. W. Package 'labdsv'. **Ordination and multivariate,** v. 775, 2016.

ROCHA, O.; MATSUMURA-TUNDISI, T. Atlas do zooplâncton (Represa do Broa, São Carlos). **São Carlos: Universidade Federal de São Carlos,** v. 1, 1976.

RUGEMA, E. et al. Long-term change of phytoplankton in Lake Kivu: The rise of the greens. **Freshwater Biology**, 2019. ISSN 0046-5070

1365-2427.

RUIZ-GONZÁLEZ, C.; NIÑO-GARCÍA, J. P.; DEL GIORGIO, P. A. Terrestrial origin of bacterial communities in complex boreal freshwater networks. **Ecology Letters,** v.

18, n. 11, p. 1198-1206, 2015. ISSN 1461-0248. Disponível em: <
http://dx.doi.org/10.1111/ele.12499 >.

SAITO, V. S.; PERKINS, D. M.; KRATINA, P. A Metabolic Perspective of Stochastic
Community Assembly. **Trends in Ecology & Evolution**, 2021. ISSN 01695347.

SALCHER, M. M. Same same but different: ecological niche partitioning of planktonic
freshwater prokaryotes. **Journal of Limnology,** v. 73, n. 1s, p. 74-87, 2014. ISSN
1129-5767.

SALCHER, M. M. et al. The ecology of pelagic freshwater methylotrophs assessed
by a high-resolution monitoring and isolation campaign. **The ISME journal,** v. 9, n.
11, p. 2442-2453, 2015. ISSN 1751-7370.

SALCHER, M. M.; PERNTHALER, J.; POSCH, T. Spatiotemporal distribution and
activity patterns of bacteria from three phylogenetic groups in an oligomesotrophic
lake. **Limnology and Oceanography,** v. 55, n. 2, p. 846-856, 2010. ISSN 0024-
3590.

_____. Seasonal bloom dynamics and ecophysiology of the freshwater sister clade
of SAR11 bacteria 'that rule the waves'(LD12). **The ISME journal,** v. 5, n. 8, p. 1242-
1252, 2011. ISSN 1751-7370.

SALCHER, M. M. et al. Spatio-temporal niche separation of planktonic
Betaproteobacteria in an oligo-mesotrophic lake. **Environmental Microbiology,** v.
10, n. 8, p. 2074-2086, 2008. ISSN 1462-2912.

SALCHER, M. M. et al. Evolution in action: habitat transition from sediment to the
pelagial leads to genome streamlining in Methylophilaceae. **The ISME journal,** v. 13,
n. 11, p. 2764-2777, 2019. ISSN 1751-7370.

SARMENTO, H. New paradigms in tropical limnology: the importance of the microbial
food web. **Hydrobiologia,** v. 686, n. 1, p. 1-14, 2012. ISSN 0018-8158 (Electronic),
1573-5117 (Linking).

SARMENTO, H.; GASOL, J. M. Use of phytoplankton-derived dissolved organic
carbon by different types of bacterioplankton. **Environ Microbiol,** v. 14, n. 9, p.
2348-60, Sep 2012. ISSN 1462-2920 (Electronic), 1462-2912 (Linking). Disponível
em: < http://www.ncbi.nlm.nih.gov/pubmed/22639946 >.

SARMENTO, H. et al. Phytoplankton species-specific release of dissolved free amino acids and their selective consumption by bacteria. **Limnology and Oceanography,** v. 58, n. 3, p. 1123-1135, 2013. ISSN 00243590.

SEGOVIA, B. T. et al. Growth and cytometric diversity of bacterial assemblages under different top–down control regimes by using a size-fractionation approach. **Journal of Plankton Research,** v. 40, n. 2, p. 129-141, 2018. ISSN 0142-7873

1464-3774.

SHABAROVA, T. et al. Distribution and ecological preferences of the freshwater lineage L im A (genus L imnohabitans) revealed by a new double hybridization approach. **Environmental microbiology,** v. 19, n. 3, p. 1296-1309, 2017. ISSN 1462-2912.

SHARMA, A. K. et al. Actinorhodopsins: proteorhodopsin-like gene sequences found predominantly in non-marine environments. **Environmental microbiology,** v. 10, n. 4, p. 1039-1056, 2008. ISSN 1462-2912.

ŠIMEK, K. et al. Influence of Top-Down and Bottom-Up Manipulations on the R-BT065 Subcluster of β-Proteobacteria, an Abundant Group in Bacterioplankton of a Freshwater Reservoir. **Applied and Environmental Microbiology,** v. 71, n. 5, p. 2381-2390, 2005. Disponível em: < https://journals.asm.org/doi/abs/10.1128/AEM.71.5.2381-2390.2005 >.

ŠIMEK, K. et al. Influence of top-down and bottom-up manipulations on the R-BT065 subcluster of β-Proteobacteria, an abundant group in bacterioplankton of a freshwater reservoir. **Applied and Environmental Microbiology,** v. 71, n. 5, p. 2381-2390, 2005. ISSN 0099-2240.

ŠIMEK, K. et al. Maximum growth rates and possible life strategies of different bacterioplankton groups in relation to phosphorus availability in a freshwater reservoir. **Environmental Microbiology,** v. 8, n. 9, p. 1613-1624, 2006. ISSN 1462-2912.

ŠIMEK, K. et al. Broad habitat range of the phylogenetically narrow R-BT065 cluster, representing a core group of the betaproteobacterial genus Limnohabitans. **Applied and environmental microbiology,** v. 76, n. 3, p. 631-639, 2010. ISSN 0099-2240.

SIQUEIRA, T.; LACERDA, C. G. L. T.; SAITO, V. S. How does landscape modification induce biological homogenization in tropical stream metacommunities? **Biotropica,** v. 47, n. 4, p. 509-516, 2015. ISSN 0006-3606.

SMITH, D. J. et al. Intercontinental dispersal of bacteria and archaea by transpacific winds. **Applied and environmental microbiology,** v. 79, n. 4, p. 1134-1139, 2013. ISSN 0099-2240.

SOININEN, J. Macroecology of unicellular organisms - patterns and processes. **Environ Microbiol Rep,** v. 4, n. 1, p. 10-22, Feb 2012. ISSN 1758-2229 (Electronic) 1758-2229 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/23757224 >.

SOININEN, J.; HEINO, J. Relationships between local population persistence, local abundance and regional occupancy of species: distribution patterns of diatoms in boreal streams. **Journal of Biogeography,** v. 32, n. 11, p. 1971-1978, 2005. ISSN 0305-0270 (Electronic) 1365-2699 (Linking).

SOININEN, J.; HEINO, J.; WANG, J. A meta-analysis of nestedness and turnover components of beta diversity across organisms and ecosystems. **Global Ecology and Biogeography,** v. 27, n. 1, p. 96-109, 2018. ISSN 1466822X.

SOININEN, J.; KORHONEN, J. J.; LUOTO, M. Stochastic species distributions are driven by organism size. **Ecology,** v. 94, n. 3, p. 660-670, 2013. ISSN 00129658, 19399170. Disponível em: < http://www.jstor.org/stable/23436269 >.

SOMMARUGA, R.; CASAMAYOR, E. O. Bacterial 'cosmopolitanism' and importance of local environmental factors for community composition in remote high-altitude lakes. **Freshw Biol,** v. 55, n. 5, p. 994-1005, May 2009. ISSN 0046-5070 (Print) 0046-5070 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/20543908 >.

STACEY, G. Chapter 10 - The Rhizobium-Legume Nitrogen-Fixing Symbiosis. In: BOTHE, H.;FERGUSON, S. J*., et al* (Ed.). **Biology of the Nitrogen Cycle**. Amsterdam: Elsevier, 2007. p.147-163. ISBN 978-0-444-52857-5.

STANIER, R.; COHEN-BAZIRE, G. Phototrophic prokaryotes: the cyanobacteria. **Annual review of microbiology,** v. 31, n. 1, p. 225-274, 1977. ISSN 0066-4227.

STEGEN, J. C. et al. Quantifying community assembly processes and identifying features that impose them. **ISME J,** v. 7, n. 11, p. 2069-79, 2013. ISSN 1751-7370 (Electronic), 1751-7362 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/23739053 >.

STEGEN, J. C. et al. Stochastic and deterministic assembly processes in subsurface microbial communities. **ISME J,** v. 6, n. 9, p. 1653-64, 2012. ISSN 1751-7370

(Electronic), 1751-7362 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/22456445 >.

STOECK, T. et al. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. **Molecular ecology,** v. 19, p. 21-31, 2010. ISSN 0962-1083.

STÖVER, B. C.; MÜLLER, K. F. TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. **BMC Bioinformatics,** v. 11, n. 1, p. 7, 2010/01/05 2010. ISSN 1471-2105. Disponível em: < https://doi.org/10.1186/1471-2105-11-7 >.

STROUS, M. et al. Missing lithotroph identified as new planctomycete. **Nature,** v. 400, n. 6743, p. 446-449, 1999. ISSN 1476-4687.

SUFFET, I.; MALLEVIALLE, J.; KAWCZYNSKI, E. **Advances in taste-and-odor treatment and control**. American Water Works Association, 1995. ISBN 0898677440.

TABERLET, P. et al. Environmental DNA. **Molecular Ecology,** v. 21, n. 8, p. 1789-1793, 2012. ISSN 1365-294X. Disponível em: < http://dx.doi.org/10.1111/j.1365-294X.2012.05542.x >.

THOMPSON, L. R. et al. A communal catalogue reveals Earth's multiscale microbial diversity. **Nature**, 2017. ISSN 1476-4687 (Electronic), 0028-0836 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/29088705 >.

THOMSEN, P. F.; WILLERSLEV, E. Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. **Biological Conservation,** v. 183, p. 4-18, 2015. ISSN 00063207.

TSEMENTZI, D. et al. Ecogenomic characterization of widespread, closely-related SAR11 clades of the freshwater genus "Candidatus Fonsibacter" and proposal of Ca. Fonsibacter lacus sp. nov. **Systematic and applied microbiology,** v. 42, n. 4, p. 495-505, 2019. ISSN 0723-2020.

TUOMISTO, H. A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. **Ecography,** v. 33, n. 1, p. 2-22, 2010. ISSN 1600-0587. Disponível em: < http://dx.doi.org/10.1111/j.1600-0587.2009.05880.x >.

URBAN, M. C. et al. The evolutionary ecology of metacommunities. **Trends Ecol Evol,** v. 23, n. 6, p. 311-7, 2008. ISSN 0169-5347 (Print), 0169-5347 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/18439707 >.

VANDERMEER, J. H. Niche Theory. **Annual Review of Ecology and Systematics,** v. 3, p. 107-132, 1972. ISSN 00664162. Disponível em: < http://www.jstor.org/stable/2096844 >. Acesso em: 2022/08/01/.

VEIDENBERG, A.; MEDLAR, A.; LÖYTYNOJA, A. Wasabi: An Integrated Platform for Evolutionary Sequence Analysis and Data Visualization. **Molecular Biology and Evolution,** v. 33, n. 4, p. 1126-1130, 2015. ISSN 0737-4038. Disponível em: < https://doi.org/10.1093/molbev/msv333 >. Acesso em: 12/24/2022.

VELLEND, M. Conceptual synthesis in community ecology. **The Quarterly review of biology,** v. 85, n. 2, p. 183-206, 2010. ISSN 0033-5770.

VELLEND, M. et al. Assessing the relative importance of neutral stochasticity in ecological communities. **Oikos,** v. 123, n. 12, p. 1420-1430, 2014. ISSN 00301299.

WAGNER, M.; HORN, M. The Planctomycetes, Verrucomicrobia, Chlamydiae and sister phyla comprise a superphylum with biotechnological and medical relevance. **Curr Opin Biotechnol,** v. 17, n. 3, p. 241-9, Jun 2006. ISSN 0958-1669 (Print), 0958-1669 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/16704931 >.

WANG, J. et al. Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. **The ISME Journal,** v. 7, n. 7, p. 1310-1321, 2013. ISSN 1751-7362

1751-7370.

WARNECKE, F. et al. Abundances, identity, and growth state of actinobacteria in mountain lakes of different UV transparency. **Applied and environmental microbiology,** v. 71, n. 9, p. 5551-5559, 2005. ISSN 0099-2240.

WEBB, C. O. et al. Phylogenies and Community Ecology. **Annual Review of Ecology and Systematics,** v. 33, n. 1, p. 475-505, 2002. ISSN 0066-4162.

WHITTAKER, R. H. Vegetation of the Siskiyou Mountains, Oregon and California. **Ecological Monographs,** v. 30, n. 3, p. 279-338, 1960. ISSN 1557-7015. Disponível em: < http://dx.doi.org/10.2307/1943563 >.

WHITTAKER, R. H. Evolution and measurement of species diversity. **Taxon,** v. 21, n. 2-3, p. 213-251, 1972. ISSN 0040-0262.

WIEGAND, S.; JOGLER, M.; JOGLER, C. On the maverick Planctomycetes. **FEMS microbiology reviews,** v. 42, n. 6, p. 739-760, 2018. ISSN 1574-6976.

WILSON, D. S. Complex Interactions in Metacommunities, with Implications for Biodiversity and Higher Levels of Selection. **Ecology,** v. 73, n. 6, p. 1984-2000, 1992. ISSN 1939-9170. Disponível em: < http://dx.doi.org/10.2307/1941449 >.

WINTER, C.; MATTHEWS, B.; SUTTLE, C. A. Effects of environmental variation and spatial distance on bacteria, archaea and viruses in sub-polar and arctic waters. **ISME J,** v. 7, n. 8, p. 1507-18, Aug 2013. ISSN 1751-7370 (Electronic) 1751-7362 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/23552622 >.

WU, Q. L.; HAHN, M. W. High predictability of the seasonal dynamics of a species-like Polynucleobacter population in a freshwater lake. **Environmental Microbiology,** v. 8, n. 9, p. 1660-1666, 2006. ISSN 1462-2912.

WU, Q. L. et al. Bacterioplankton community composition along a salinity gradient of sixteen high-mountain lakes located on the Tibetan Plateau, China. **Applied and Environmental Microbiology,** v. 72, n. 8, p. 5478-5485, 2006. ISSN 0099-2240.

YILMAZ, P. et al. The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. **Nucleic Acids Research,** v. 42, p. D643-D648, 2014.

YOCCOZ, N. G. The future of environmental DNA in ecology. **Molecular Ecology,** v. 21, n. 8, p. 2031-2038, 2012. ISSN 1365-294X. Disponível em: < http://dx.doi.org/10.1111/j.1365-294X.2012.05505.x >.

ZANON, J. E.; RODRIGUES, L.; BINI, L. M. Hard to predict: Synchrony in epiphytic biomass in a floodplain is independent of spatial proximity, environmental distance, and environmental synchrony. **Ecological Indicators,** v. 93, p. 379-386, 2018. ISSN 1470160X.

ZAREMBA-NIEDZWIEDZKA, K. et al. Single-cell genomics reveal low recombination frequencies in freshwater bacteria of the SAR11 clade. **Genome biology,** v. 14, n. 11, p. 1-14, 2013. ISSN 1474-760X.

ZEDER, M. et al. A small population of planktonic Flavobacteria with disproportionally high growth during the spring phytoplankton bloom in a prealpine

lake. **Environmental microbiology,** v. 11, n. 10, p. 2676-2686, 2009. ISSN 1462-2912.

ZENG, Y. et al. **Genome sequences of two freshwater betaproteobacterial isolates, Limnohabitans species strains Rim28 and Rim47, indicate their capabilities as both photoautotrophs and ammonia oxidizers**: Am Soc Microbiol 2012.

ZHANG, X. et al. Local community assembly mechanisms shape soil bacterial beta diversity patterns along a latitudinal gradient. **Nat Commun,** v. 11, n. 1, p. 5428, 2020. ISSN 2041-1723 (Electronic), 2041-1723 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/33110057 >.

ZÖLLNER, E. et al. Cascading predation effects of Daphnia and copepods on microbial food web components. **Freshwater Biology,** v. 48, n. 12, p. 2174-2193, 2003. ISSN 0046-5070. Disponível em: < https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2426.2003.01158.x >.

Material Suplementar

Depicting spatiotemporal variables that drive dominance dynamics of lake bacteria

Table S1 – Environmental data for each sample collected at the Broa's Microbial Observatory and used in this study. Temperature, pH and Dissolved O_2 were obtained through a multiparameter probe, while the Euphotic Zone was calculated using a Secchi disk measurement. The Dissolved Organic Carbon (DOC) and Dissolved Organic Matter (DOM) and Chlorophyll *a* (chl *a*) concentrations were measured in laboratory; the DOM was used to obtain the as fluorescence index, freshness index, humification index and A:C ratio, whitch were used to infer the quality of this material, while the chl *a* measurements served as a proxy to calculate de Trophic State

| ID	Collection Date	Season	Temperature (°C)	pH	Dissolved O_2 (mg/L)	DOC (mg/L)	Slope Ratio	Fluorescence Index	Freshness Index	Humification Index	A:C Ratio	Euphotic Zone (m)	Trophic State
BroaMO_01	Mar-15-2018	Rainy	27.61	5.63	10.86	3.784	0.981	1.332	0.695	4.979	2.478	1.34	50.01
BroaMO_02	Apr -24-2018	Dry	23.69	5.94	11.1	7.823	1.143	1.229	0.577	7.82	2.485	1.18	51.54
BroaMO_03	May-15-2018	Dry	23.17	6.29	9.26	5.204	1.087	1.215	0.603	6.359	2.49	1.17	49.63
BroaMO_04	Jun-15-2018	Dry	20.56	7.45	4.87	3.188	1.063	1.213	0.61	3.683	2.514	1.28	48.16
BroaMO_05	Jul-19-2018	Dry	18.97	6.33	4.68	2.722	0.858	1.218	0.606	6.229	2.457	1.06	48.16
BroaMO_06	Aug-21-2018	Dry	20	5.92	11.29	37.63	1.043	1.211	0.614	5.38	2.424	1.06	50.67
BroaMO_07	Sep-19-2018	Dry	22.37	6.03	8.16	2.824	1.77	1.169	0.681	3.772	2.537	1.04	55.12
BroaMO_08	Oct-30-2018	Rainy	25.6	6.38	9.35	4.474	1.81	1.17	0.612	5.72	2.548	1.83	55.6
BroaMO_09	Nov-22-2018	Rainy	23.99	5.48	1.13	7.515	0.986	1.226	0.53	6.191	2.418	1.04	52.8
BroaMO_10	Dec-12-2018	Rainy	26.37	7.98	8.68	3.663	0.878	1.183	0.625	5.959	2.637	0.94	55.21
BroaMO_11	Jan-15-2019	Rainy	28.34	6.25	8.47	6.447	1.057	1.228	0.652	3.564	2.633	0.83	53.69
BroaMO_12	Feb-19-2019	Rainy	25.38	5.88	7.47	32.58	0.912	1.244	0.622	5.837	2.53	1	53.69
BroaMO_13	Mar-19-2019	Rainy	27.26	5.27	9.06	4.798	0.964	1.246	0.62	5.85	2.598	0.87	53.84
BroaMO_14	Apr-23-2019	Rainy	24.95	5.36	0	6.605	0.846	1.267	0.602	6.769	2.617	0.97	55.98
BroaMO_15	May-21-2019	Dry	22.48	7.16	7.22	4.709	0.935	1.271	0.601	7.738	2.635	1.34	53.99
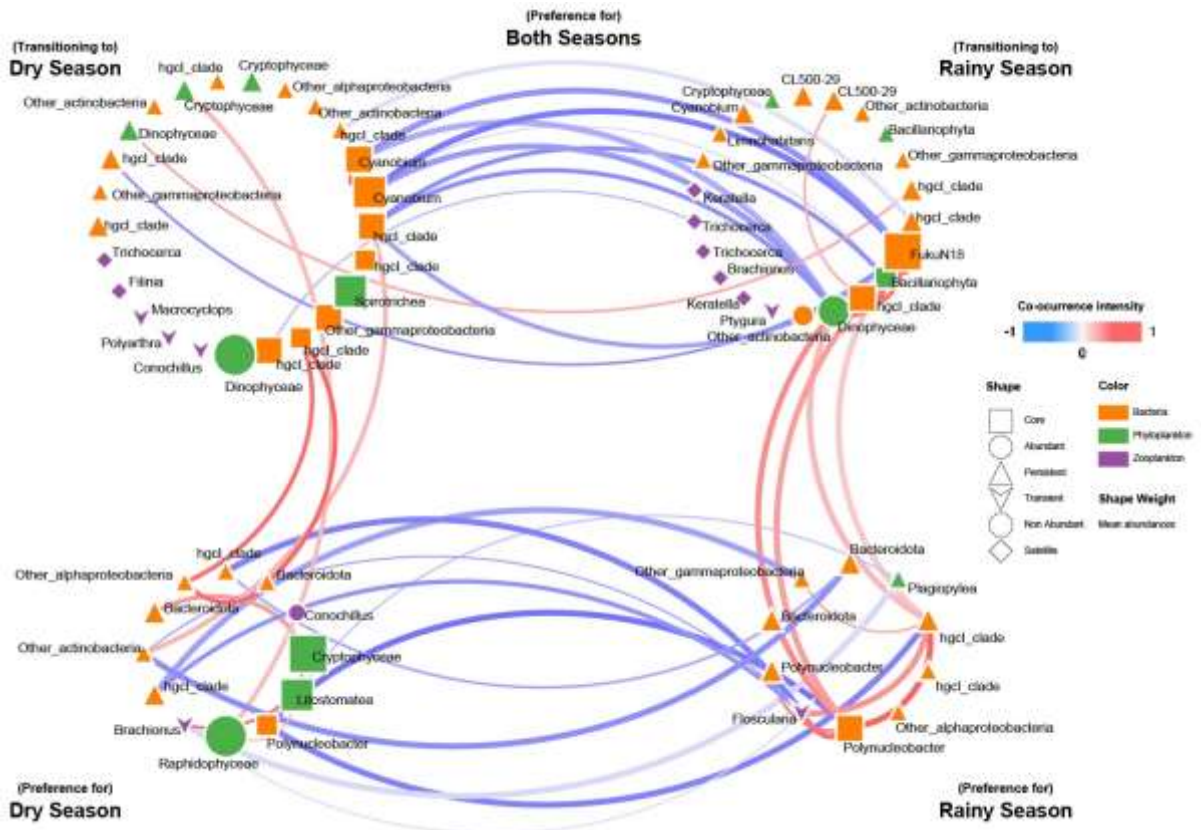
Figure S1 – Co-occurrence networks considering bacteria (orange), phytoplankton (green) and zooplankton (purple) taxonomic units as recovered by 16S and 18S rDNA sampling. Blue lines indicate negative co-occurrences, while red lines indicate positive co-occurrences. Squared shapes indicate Core organisms, round shapes are the Abundant ones and triangles are the Persistent. Inverted triangles are the Transient, octagons represent Non-abundant ASVs and diamond shapes are for the Satellite. The shape weight is the mean abundance of each ASV.
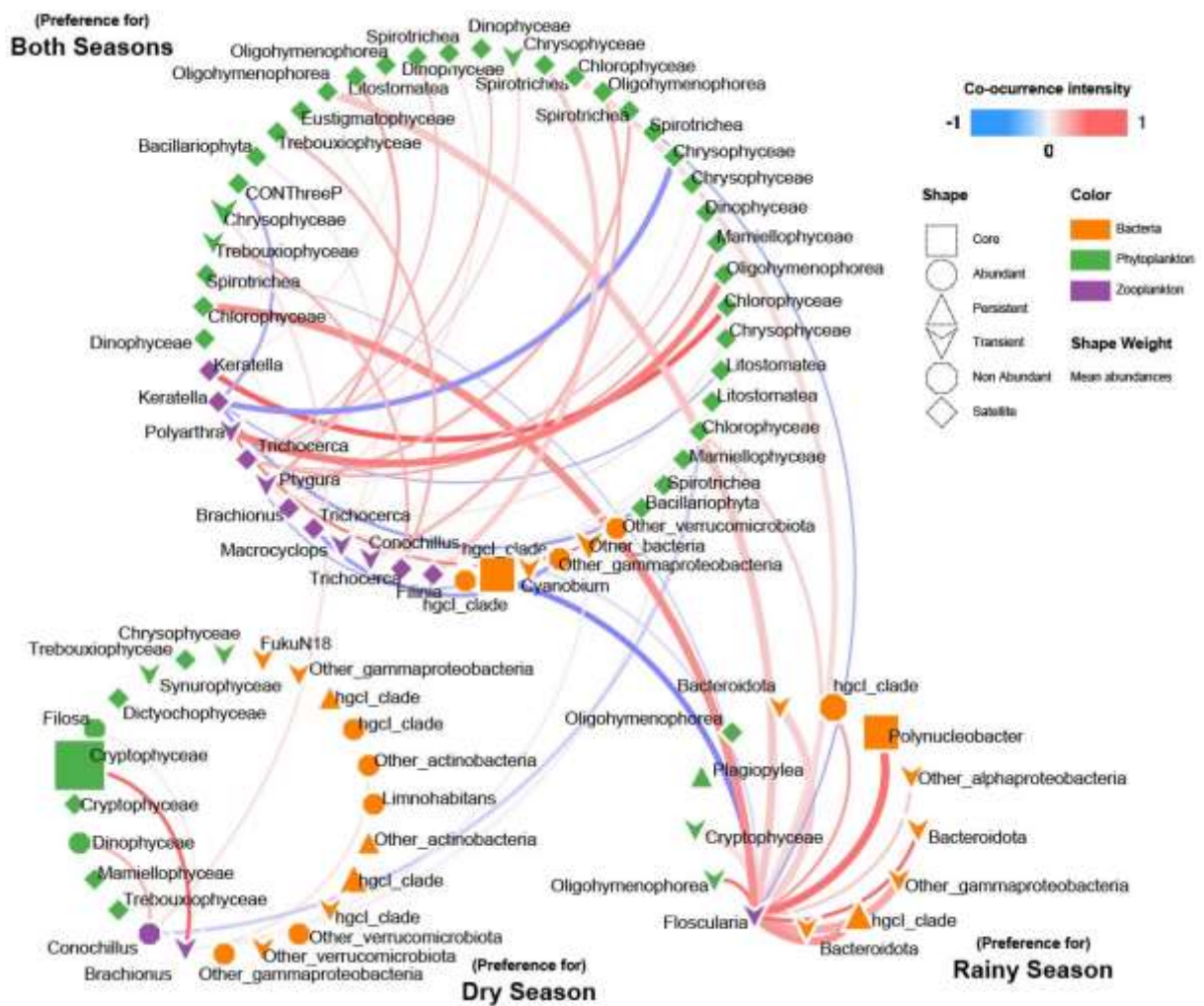
Figure S2 – Co-occurrence networks considering showing the connections of zooplankton (purple) with bacteria (orange)and phytoplankton (green) taxonomic units as recovered by 16S and 18S rDNA sampling. Blue lines indicate negative co-occurrences, while red lines indicate positive co-occurrences. Squared shapes indicate Core organisms, round shapes are the Abundant ones and triangles are the Persistent. Inverted triangles are the Transient, octagons represent Non-abundant ASVs and diamond shapes are for the Satellite. The shape weight is the mean abundance of each ASV.

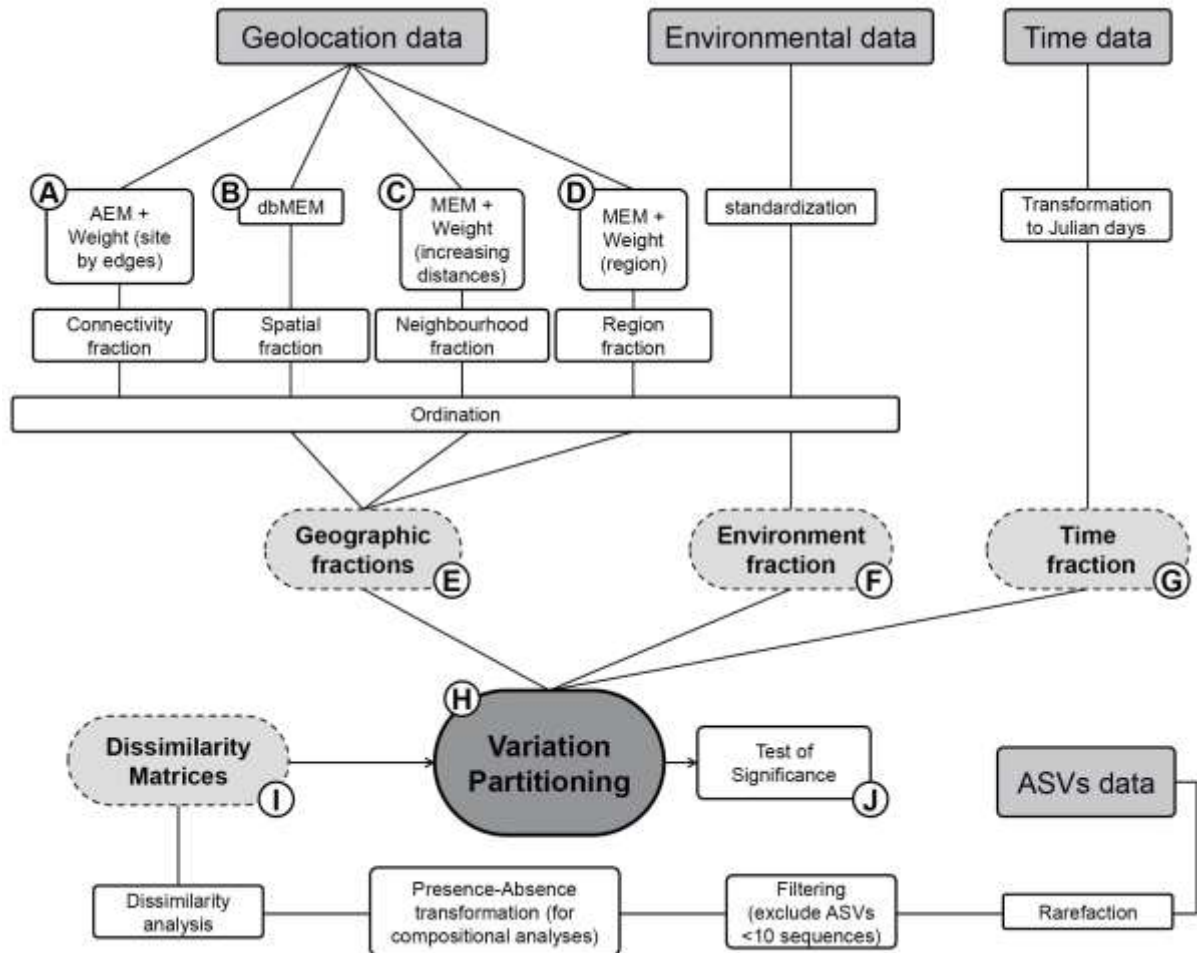Beyond environmental selection: Spatial structuring of tropical lake bacterioplankton metacommunity

Figure S3 – Fluxogram showing all steps performed in the analysis of variation partitioning of βdiv between spatial and environmental factors. The geolocation of each site was used to create distinct eigenvectors that models 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) and the time fraction (G) was obtained by the transformation for Julian date. The variation partitioning (H) based on a dbRDA was than performed to test the importance of these fractions to the dissimilarity matrices (I) and a CCA was used to test for significance (J)
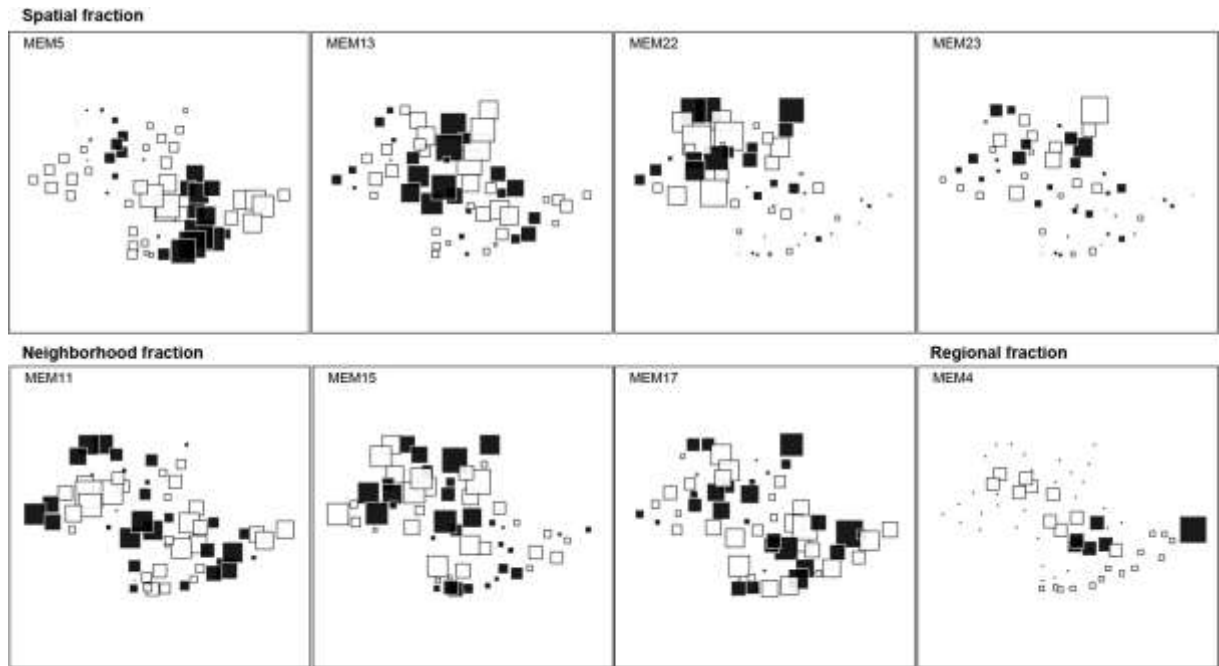
Figure S4 – All eigenvectors recovered after the ordination step. For the qualitative approach, the significant eigenvectors were MEM5, MEM22 (space fraction), MEM15, MEM17 (neigborhood fraction) and MEM4 (region fraction), while for the quantitative approach the significant ones were MEM13, MEM23 (space fraction), MEM11, MEM15 (neigborhood fraction) and MEM4 (region fraction). Here, the eingenvectors were plotted against latitude (y axis) and longitude (x axis); the square sizes indicate the value attributed to each site, which may be negative (white squares) or positive (black squares). Squares of same color and similar sizes indicate sites more closely related

157

Table S2 – Variation partitioning results for the metacommunity considering qualitative and quantitative approaches as showed in the Figure 13

Qualitative approach

Compared fractions	All fractions	Fraction 1	Fraction 2	Fraction 3	Fraction 1 vs Fraction 2	Fraction 1 vs Fraction 3	Fraction 2 vs Fraction 3	Pure 1	Pure 2	Pure 3
Environment vs Space vs Time	0.001	0.001	0.019	0.004	0.001	0.001	0.029	0.001	0.079	0.046
Environment vs Neighborhood vs Time	0.001	0.001	0.014	0.008	0.001	0.001	0.056	0.001	0.163	0.191
Environment vs Region vs Time	0.001	0.001	0.012	0.004	0.001	0.001	0.038	0.001	0.085	0.101
Other environment vs pH vs Time	0.001	0.003	0.005	0.01	0.001	0.001	0.001	0.002	0.001	0.058
Space vs Neigborhood vs Region	0.002	0.016	0.01	0.015	0.003	0.003	0.001	0.013	0.01	0.01

Quantitative approach

Compared fractions	All fractions	Fraction 1	Fraction 2	Fraction 3	Fraction 1 vs Fraction 2	Fraction 1 vs Fraction 3	Fraction 2 vs Fraction 3	Pure 1	Pure 2	Pure 3
Environment vs Space vs Time	0.001	0.001	0.01	0.004	0.001	0.001	0.004	0.002	0.271	0.001
Environment vs Neighborhood vs Time	0.001	0.001	0.004	0.004	0.001	0.001	0.004	0.001	0.178	0.031
Environment vs Region vs Time	0.001	0.001	0.012	0.003	0.001	0.001	0.001	0.001	0.155	0.002
Other environment vs pH vs Time	0.001	0.001	0.004	0.004	0.001	0.001	0.001	0.005	0.001	0.002
Space vs Neighborhood vs Region	0.002	0.006	0.003	0.013	0.023	0.034	0.016	0.288	0.138	0.011

Table S3 (Part 1/2) – Variation partitioning results for the metacommunity considering maximum distances thresholds as showed in the Figure 14. Both qualitative and quantitative approaches were considered

Qualitative Approach

Filter	env	env-spa	spa	spa-time	time	time-env	both	nothing	env fraction	space fraction	neighbour fraction	region fraction
(800-900]	12.7%	3.2%	1.5%	0.1%	2.6%	NA	NA	81.2%	FC + pH + DIN + DIC	20 + 1	23 + 2	NA
(700-800]	10.1%	0.7%	NA	0.5%	2.6%	NA	NA	86.8%	pH + FC + alt + DIN + DIC	13	NA	NA
(600-700]	2.7%	2.6%	0.6%	0.5%	0.3%	NA	0.2%	91.7%	pH + DOC + alt + FC	5 + 23	15 + 14	4
(500-600]	2.3%	1.0%	1.8%	0.5%	0.1%	NA	0.6%	93.9%	pH + DOC +alt + FC	5 + 22 + 23 + 9	15 + 17	4
(400-500]	1.4%	0.6%	1.4%	0.5%	0.1%	NA	0.7%	95.5%	pH + DOC + alt	3+ 5 + 20	1 + 15 + 7	4 + 1 + 3
(300-400]	1.2%	0.9%	2.3%	1.0%	NA	NA	0.3%	94.5%	alt + pH + DOC + DIC + temp	5 + 2 + 3 + 19	1 + 11 + 14 + 2 + 8	3 + 1
(200-300]	NA	NA	NA	NA	NA	NA	NA	NA	ERROR	ERROR	ERROR	ERROR
(100-200]	NA	NA	NA	NA	NA	NA	NA	NA	ERROR	ERROR	ERROR	ERROR
(0-100]	NA	NA	NA	NA	NA	NA	NA	NA	ERROR	ERROR	ERROR	ERROR

Quantitative Approach

Filter	env	env-spa	spa	spa-time	time	time-env	both	nothing	env fraction	space fraction	neighbour fraction	region fraction
(800-900]	8.1%	5.1%	7.4%	2.1%	0.9%	0.5%	NA	76.0%	pH + DIC + alt + temp	23 + 2 + 13 + 1	11 + 23	4
(700-800]	9.8%	NA	0.9%	0.4%	2.7%	NA	0.7%	86.8%	pH + DIC + DIN + alt	NA	NA	4
(600-700]	4.2%	1.1%	0.6%	0.6%	0.2%	NA	1.7%	91.6%	pH + alt + DOC + FC	23	15 + 7	4
(500-600]	2.4%	0.9%	1.6%	0.5%	0.1%	NA	0.9%	94.1%	pH + DOC + alt + FC	5 + 22	15	4 + 1
(400-500]	0.9%	1.0%	2.6%	0.5%	0.1%	0.2%	0.4%	94.3%	alt + pH + DOC	5 + 3 + 12 + 22	1 + 15 + 7	4 + 3
(300-400]	0.6%	1.2%	2.6%	1.3%	NA	0.1%	0.2%	94.4%	alt + pH + DOC + temp	5 + 23	1 + 2 + 11 + 8 + 14 + 15	1 + 3 + 4
(200-300]	NA	NA	NA	NA	NA	NA	NA	NA	ERROR	ERROR	ERROR	ERROR
(100-200]	NA	NA	NA	NA	NA	NA	NA	NA	ERROR	ERROR	ERROR	ERROR
(0-100]	NA	NA	NA	NA	NA	NA	NA	NA	ERROR	ERROR	ERROR	ERROR

Table S3 (Part 2/2)

[a+b+c+d+e+f+g] All	[a+d+f+g] Space fraction	[b+d+b+g] env fraction	[a+d+b+g] time fraction	[a+d+b+g] Space vs env	[b+e+c+g] env vs time	[a+f+c+g] time vs space	[a] pure env	[b] pure space	[c] pure time
0.001	0.014	0.001	0.052	0.001	0.001	0.04	0.001	0.242	0.008
0.001	0.035	0.001	0.038	0.001	0.001	0.035	0.001	0.452	0.034
0.001	0.002	0.001	0.007	0.001	0.001	0.005	0.005	0.244	0.135
0.001	0.001	0.001	0.009	0.001	0.001	0.009	0.005	0.021	0.386
0.001	0.001	0.001	0.004	0.001	0.001	0.063	0.032	0.012	0.246
0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.017	0.002	0.449
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
[a+b+c+d+e+f+g] All	[a+d+f+g] Space fraction	[b+d+b+g] env fraction	[a+d+b+g] time fraction	[a+d+b+g] Space vs env	[b+e+c+g] env vs time	[a+f+c+g] time vs space	[a] pure env	[b] pure space	[c] pure time
0.001	0.001	0.001	0.011	0.001	0.001	0.002	0.002	0.006	0.068
0.001	0.008	0.001	0.004	0.001	0.001	0.003	0.001	0.15	0.006
0.001	0.002	0.001	0.005	0.003	0.001	0.012	0.003	0.307	0.367
0.001	0.001	0.001	0.023	0.001	0.001	0.005	0.005	0.023	0.457
0.001	0.001	0.001	0.006	0.001	0.001	0.008	0.124	0.011	0.45
0.001	0.001	0.001	0.002	0.001	0.006	0.001	0.247	0.008	0.902
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Scale matters? The effect of spatial scale on ecological processes that drive aquatic bacterial communities
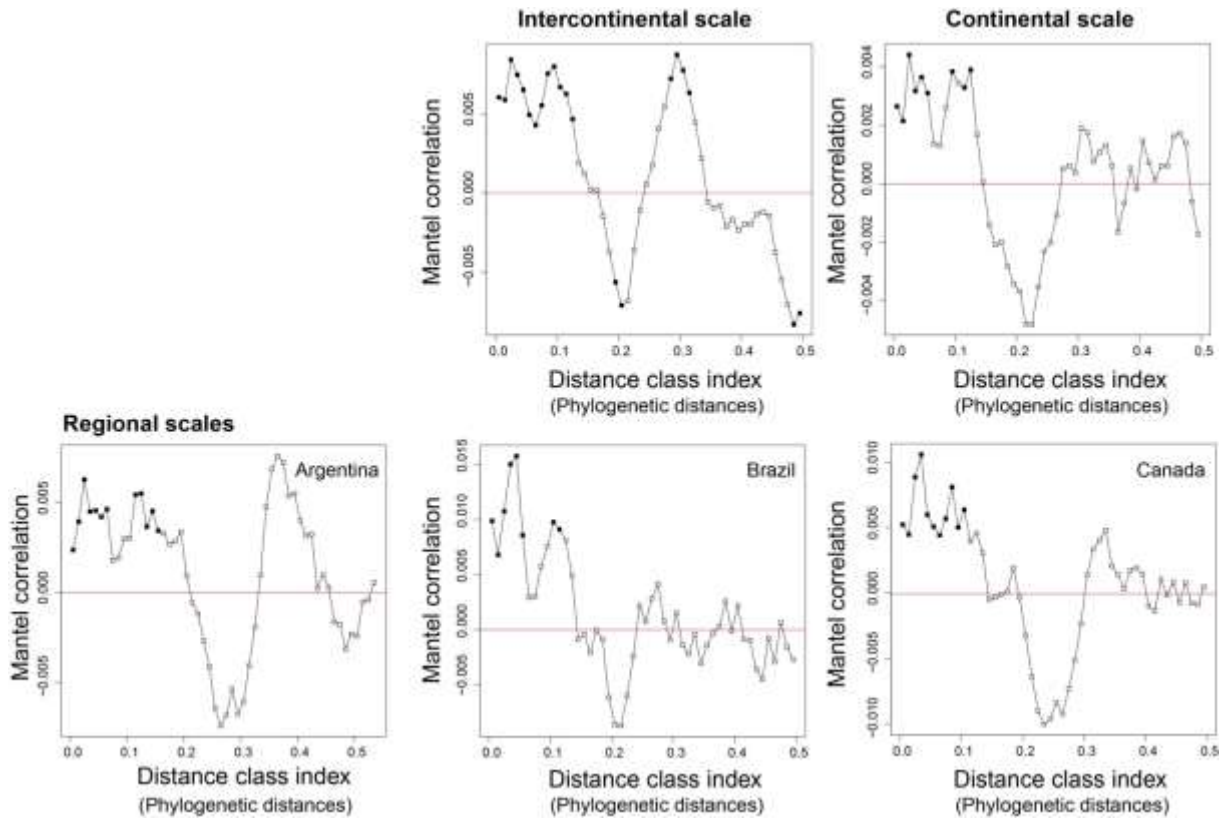
Figure S5 – Mantel correlograms for all scales here addressed. The correlogram was obtained by the analysis of between sites phylogenetic dissimilarities against environmental dissimilarities. The environmental dissimilarities were obtained by using the measurements of Altitude (m), lake area (km^2) and perimeter (km), pH, Chlorophyll a (mg/L) and DOC (mg/L). The black squares represent significative correlations (p < 0.05) between phylogenetic and environmental distances, which means that closely related microrganisms were similarly selected by the similar environmental conditions, in all addressed scales

Anexos

Artigos publicados em revistas indexadas durante o curso deste doutorado.
Organizados por ordem de publicação

Mateus-Barros *et al.* (2019) *Journal of Microbiological Methods*

Contents lists available at ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

ELSEVIER

Comparison of two DNA extraction methods widely used in aquatic microbial ecology

Erick Mateus-Barros[a,b], Aylan K. Meneghine[a], Inessa Lacativa Bagatini[c], Camila C. Fernandes[d,e], Luciano T. Kishi[d,e], Armando A.H. Vieira[c], Hugo Sarmento[a,*]

[a] *Universidade Federal de São Carlos (UFSCar), Department of Hydrobiology, Laboratory of Microbial Processes and Biodiversity, São Carlos, SP 13565-905, Brazil*
[b] *Post Graduate Program in Ecology and Natural Resources (PPGERN), UFSCar, São Carlos, SP 13565-905, Brazil*
[c] *UFSCar, Department of Botany, Laboratory of Phycology, São Carlos, SP 13565-905, Brazil*
[d] *Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias, Departamento de Tecnologia, Laboratório de Bioquímica de Microrganismos e Plantas – LBMP, Jaboticabal, SP 14884-900, Brazil*
[e] *UNESP, Faculdade de Ciências Agrárias e Veterinárias, Departamento de Tecnologia, Laboratório Multiusuário Centralizado para Sequenciamento de DNA em Larga Escala e Análise de Expressão Gênica – LMSeq, Jaboticabal, SP 14884-900, Brazil*

Pestana *et al.* (2020) *Bird Study*

BTO
Looking out for birds

Check for updates

The influence of parent body colouration and nesting habitat on bird nest predation

Gabrielle Cristina Pestana, Erick Mateus-Barros and Rhainer Guillermo-Ferreira

Departamento de Hidrobiologia, Centro de Ciências Biológicas e da Saúde, Universidade Federal de São Carlos, São Carlos, Brasil

Mateus-Barros *et al.* (2021) ***Microbial Ecology***

MICROBIOLOGY OF AQUATIC SYSTEMS

Check for updates

Local and Geographic Factors Shape the Occupancy-Frequency Distribution of Freshwater Bacteria

Erick Mateus-Barros [1,2] · Michaela L. de Melo [1,3] · Inessa L. Bagatini [4] · Adriano Caliman [5] · Hugo Sarmento [1]

Pestana *et al.* (2021) ***International Journal of Odonatology***

Gabrielle Cristina Pestana [1,2], Erick Mateus-Barros [1,2], Leandro Schlemmer Brasil [3] and Rhainer Guillermo-Ferreira [1,4]*

A scientometric analysis on pre- and post-copulatory traits in Odonata

Correa *et al.* (2022) ***Journal of Veterinary Behavior***

Is item hiding a good enrichment strategy to reduce stereotypic behaviors and increase social interactions in captive female spectacled bears?

Check for updates

Ana Julia Sant'Ana Correa [a], Erick Mateus Barros [a], Vinicius Marques Lopez [b], Rhainer Guillermo-Ferreira [c,*]

[a] *Universidade Federal de São Carlos, São Carlos, SP, Brazil*
[b] *Universidade de São Paulo, Ribeirão Preto, SP, Brazil*
[c] *Universidade Federal do Triângulo Mineiro, Uberaba, MG, Brazil*

scientific **data**

Check for updates

OPEN

DATA DESCRIPTOR

A georeferenced rRNA amplicon database of aquatic microbiomes from South America

Sebastian Metz[1,2], Paula Huber[3,4], Erick Mateus-Barros[4], Pedro C. Junger[4], Michaela de Melo[4,5], Inessa Lacativa Bagatini[6], Irina Izaguirre[7], Mariana Câmara dos Reis[4,8], Maria E. Llames[1], Victoria Accattatis[3], María Victoria Quiroga[1], Melina Devercelli[3], María Romina Schiaffino[9,10], Juan Pablo Niño-García[11], Marcela Bastidas Navarro[12], Beatriz Modenutti[12], Helena Vieira[6,13], Martín Saraceno[7], Carmen Alejandra Sabio y García[7], Emiliano Pereira[14], Alvaro González-Revello[15,16], Claudia Piccini[17], Fernando Unrein[1], Cecilia Alonso[14] & Hugo Sarmento[4] ✉

[1]Laboratorio de Ecología Acuática, Instituto Tecnológico de Chascomús (INTECH), UNSAM-CONICET, Av. Intendente Marino Km 8.200, (7130) Chascomús, Buenos Aires, Argentina. [2]Sorbonne Université, CNRS, UMR7144 Adaptation et Diversité en Milieu Marin, Ecology of Marine Plankton (ECOMAP), Station Biologique de Roscoff SBR, 29680, Roscoff, France. [3]Laboratorio de Plancton Instituto Nacional de Limnología (INALI), CONICET-UNL, Ciudad Universitaria, Paraje El Pozo, C. P, 3000, Santa Fe, Argentina. [4]Laboratory of Microbial Processes & Biodiversity, Departamento de Hydrobiologia, Universidade Federal de São Carlos (UFSCar). Rodovia Washington Luiz, São Carlos, São Paulo, 13565-905, Brazil. [5]University of Quebec at Montreal, Department of Biological Science. Centre-Ville, C.P, 8888, Montreal, (Quebec), Canada. [6]Laboratório de Ficologia, Departamento de Botânica, Universidade Federal de São Carlos (UFSCar). Rodovia Washington Luiz, São Carlos, São Paulo, 13565-905, Brazil. [7]Depto. de Ecologia, Genética y Evolución, IEGEBA (UBA-CONICET), Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires (UBA). Intendente Güiraldes 2160, C1428EHA, Buenos Aires, Argentina. [8]Sorbonne Université/Centre National de la Recherche Scientifique, UMR 7144, Adaptation et Diversité en Milieu Marin, Station Biologique de Roscoff, Place Georges Teissier, Roscoff, France. [9]Departamento de Ciencias Básicas y Experimentales, Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA), Roque Sáenz Peña 456, 6000, Junín, Argentina. [10]Centro de Investigaciones y Transferencia del Noroeste de la Provincia de Buenos Aires (CITNOBA) – UNNOBA-UNSAdA-CONICET, Monteagudo 2772, 2700, Buenos Aires, Argentina. [11]Escuela de Microbiología, Universidad de Antioquia, Cl. 67 ##53-108, Medellín, Antioquia, Colombia. [12]Instituto de Investigaciones en Biodiversidad y Medio Ambiente (INIBIOMA), CONICET-Universidad Nacional del Comahue, Quintral 1250, R8400, San Carlos de Bariloche, Río Negro, Argentina. [13]Czech Academy of Sciences, Biology Centre, Hydrobiology Institute, Na Sádkách 702/7, 370 05, Ceske Budejovice, Czechia. [14]Centro Universitario Regional del Este. Universidad de la República, Ruta nacional N° 9 intersección con ruta N°15, CP 270000, Rocha, Uruguay. [15]Departamento de Ciencia y Tecnología de los Alimentos, Facultad de Veterinaria, UDELAR. Alberto Lasplaces 1620, CP 11600, Montevideo, Uruguay. [16]Departamento de Genómica, Instituto de Investigaciones Biológicas Clemente Estable, MEC. Av. Italia 3318, 11600, Montevideo, Departamento de Montevideo, Uruguay. [17]Laboratorio de Ecología Microbiana Acuática, Departamento de Microbiología, Instituto de Investigaciones Biológicas Clemente Estable, MEC. Av. Italia 3318, 11600, Montevideo, Departamento de Montevideo, Uruguay. ✉e-mail: hsarmento@ufscar.br

Stadler *et al.* (2023) **Frontiers in Water**

Artigo submetido em revistas indexadas e sob revisão por pares neste momento

Mateus-Barros *et al.* (in prep) **PeerJ**

AUTHORS

👤 Erick Mateus-Barros
📍 Programa de Pós-Graduação em Ecologia e Recursos Naturais (PPG-ERN), Universidade Federal de São Carlos (UFSCar), São Carlos, São Paulo, Brasil
📍 Departamento de Hidrobiologia, Universidade Federal de São Carlos, São Carlos, São Paulo, Brasil

👤 Adriano Caliman
📍 Departamento de Ecologia, Universidade Federal do Rio Grande do Norte (UFRN), Natal, Rio Grande do Norte, Brasil

👤 Thaís Garcia da Silva
📍 Departamento de Hidrobiologia, Universidade Federal de São Carlos, São Carlos, São Paulo, Brasil
📍 Departamento de Botânica, Universidade Federal de São Carlos, São Carlos, São Paulo, Brazil

👤 Hugo Sarmento
📍 Departamento de Hidrobiologia, Universidade Federal de São Carlos, São Carlos, São Paulo, Brasil