



UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECOLOGIA E EVOLUÇÃO

JOCILAINE SANTOS DE JESUS

**Diversidade Filogenética da
comunidade de microeucariotos
planctônicos da planície de inundação
do rio Araguaia por DNA metabarcoding**

Goiânia, GO
SETEMBRO/2024



UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a [Lei 9.610/98](#), o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico

Dissertação Tese Outro*: _____

*No caso de mestrado/doutorado profissional, indique o formato do Trabalho de Conclusão de Curso, permitido no documento de área, correspondente ao programa de pós-graduação, orientado pela legislação vigente da CAPES.

Exemplos: Estudo de caso ou Revisão sistemática ou outros formatos.

2. Nome completo do autor

Jocilaine Santos de Jesus

3. Título do trabalho

Diversidade Filogenética da comunidade de microeucariotos planctônicos da planície de inundação do rio Araguaia por DNA metabarcodes

4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento SIM NÃO¹

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

a) consulta ao(à) autor(a) e ao(à) orientador(a);

b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.

O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.



Documento assinado eletronicamente por **Jocilaine Santos De Jesus, Discente**, em 11/10/2024, às 12:06, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Thannya Nascimento Soares, Professora do Magistério Superior**, em 25/10/2024, às 14:46, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4893620** e o código CRC **50C096D1**.

JOCILAINE SANTOS DE JESUS



Diversidade Filogenética da comunidade de microeucariotos planctônicos da planície de inundação do rio Araguaia por DNA metabarcoding

Dissertação apresentada ao Programa de Pós-Graduação em Ecologia e Evolução do Instituto de Ciências Biológicas (ICB) da Universidade Federal de Goiás (UFG), como pré-requisito para obtenção do título de mestre em Ecologia e Evolução.

Área de Concentração: Ecologia e Evolução
Linha de Pesquisa: Ecologia Molecular e Genética da Conservação

Orientadora: Prof^a. Dra. Thannya Nascimento Soares

Coorientador: Prof. Dr. João Carlos Nabout

Goiânia, GO
SETEMBRO/2024

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Jesus, Jocilaine Santos de

Diversidade Filogenética da comunidade de microeucariotos planctônicos da planície de inundação do rio Araguaia por DNA metabarcodado [manuscrito] / Jocilaine Santos de Jesus. - 2024. 96 f.

Orientador: Profa. Dra. Thannya Nascimento Soares; co orientador Dr. João Carlos Nabout.

Dissertação (Mestrado) - Universidade Federal de Goiás, Instituto de Ciências Biológicas (ICB), Programa de Pós-Graduação em Ecologia e Evolução, Goiânia, 2024.

Bibliografia. Apêndice.

Inclui siglas, mapas, fotografias, abreviaturas, símbolos, gráfico, tabelas.

1. DNA ambiental. 2. plâncton. 3. 18S rRNA. 4. estrutura filogenética. 5. OTUs. I. Soares, Thannya Nascimento, orient. II. Título.

CDU 574



UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
ATA DE DEFESA DE DISSERTAÇÃO

Ata nº 219 da sessão de Defesa de Dissertação de **Jocilaine Santos de Jesus** que confere o título de **Mestra em Ecologia e Evolução**, na área de concentração em **Ecologia e Evolução**.

Ao/s **vinte e seis dias do mês de setembro do ano de dois mil e vinte e quatro (26/09/2024)**, a partir da(s) **14:00 horas**, no **Auditório do ICB 5**, seguindo Resolução CONSUNI/UFG Nº 141 de 13 de maio de 2022 e orientações do Ofício Circular no. 34/2022/PRPG/UFG (SEI 23070.030951/2022-07), realizou-se a sessão pública de Defesa de Dissertação intitulada “**Diversidade Filogenética da comunidade de microeucariotos planctônicos da planície de inundação do rio Araguaia por DNA metabarcoding**”. Os trabalhos foram instalados pela Orientadora **Profa. Dra. Thannya Nascimento Soares (DGen/ICB/UFG)** e pelo Co-Orientador **Prof. Dr. João Carlos Nabout (UEG)** com a participação dos demais membros da Banca Examinadora: **Prof. Dr. José Alexandre Felizola Diniz Filho (DECOL/ICB/UFG)**, membro titular interno; **Profa. Dra. Karine Borges Machado (UEG)**, membro titular externo, **cuja participação ocorreu através de videoconferência**. Durante a argüição os membros da banca **não fizeram** sugestão de alteração do título do trabalho. A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Dissertação tendo sido a candidata **aprovada** pelos seus membros. Proclamados os resultados pela **Profa. Dra. Thannya Nascimento Soares**, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, ao(s) **vinte e seis dias do mês de setembro do ano de dois mil e vinte e quatro (26/09/2024)**.

TÍTULO SUGERIDO PELA BANCA



Documento assinado eletronicamente por **Thannya Nascimento Soares, Professora do Magistério Superior**, em 26/09/2024, às 16:35, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Karine Borges Machado, Usuário Externo**, em 26/09/2024, às 16:38, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **João Carlos Nabout, Usuário Externo**, em 26/09/2024, às 16:38, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Jose Alexandre Felizola Diniz Filho, Professor do Magistério Superior**, em 26/09/2024, às 16:38, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4807568** e o código CRC **65C24F1F**.

Referência: Processo nº 23070.044576/2024-36

SEI nº 4807568

Com amor e gratidão, dedico este trabalho à minha querida avó Rosa e à minha mãe Angelina. Mulheres que, sem os privilégios da educação formal, acumularam um vasto conhecimento empírico ao longo de suas vidas, garantindo sua sobrevivência e prosperidade. Sua força e resiliência, forjadas diante das adversidades, me inspiram a cada dia.

Agradecimentos

Agradeço profundamente a todos que contribuíram direto e indiretamente para a realização deste trabalho.

Em especial:

Aos meus pais, que sempre me incentivaram a buscar conhecimento, ao meu irmão Josenilton, por sua amizade, incentivo e apoio incondicional, e por sempre acreditar em meu potencial. Ao meu companheiro David, pelo apoio, ajuda diária e pelo aconchego emocional nos dias difíceis. Aos meus amigos, pelo encorajamento, companheirismo e por tornarem meus dias mais alegres.

À minha orientadora, Dra. Thannya Nascimento Soares, pela oportunidade de ser sua orientada, por todo conhecimento compartilhado, pela paciência, compreensão e apoio nestes dois anos. Agradeço também meu coorientador, Dr. João Carlos Nabout, pelo apoio e contribuições valiosas para o desenvolvimento deste trabalho.

Aos colegas do laboratório de Genética e Biodiversidade (LGBio) pela amizade e conhecimento compartilhado, em especial ao Diego Ortiz, pela parceria na coleta e análise das amostras. À Dra. Cíntia Pelegrineti Targueta de Azevedo Brito, pela ajuda com a preparação da biblioteca e sequenciamento. À toda equipe do LGBio e à equipe do Araguaia Vivo 2023 pela contribuição na coleta das amostras e dos dados ambientais e limnológicos, sem a contribuição de cada um a execução deste trabalho não seria possível.

Ao Instituto Nacional de Ciência Tecnologia em Ecologia, Evolução e Conservação da Biodiversidade (INCT – EECBio) apoiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), à Fundação de Amparo à Pesquisa do Estado de Goiás (FAPEG), à Aliança Tropical de Pesquisa da Água (TWRA), ao Programa de Pesquisa em Biodiversidade do Araguaia (PPBio Araguaia) apoiado pelo CNPq e ao Programa Araguaia Vivo 2030 apoiado pela TWRA e FAPEG, agradeço pelo apoio financeiro e logístico para execução deste trabalho. Agradeço também ao CNPq pela concessão da bolsa, que foi meu suporte financeiro nesses dois anos. Por fim, agradeço à Universidade Federal de Goiás e ao programa de Pós-Graduação em Ecologia e Evolução, pela infraestrutura, por sua excelente qualidade de docentes e pesquisadores que foram essenciais para minha formação.

À todos, muito obrigada!

Sumário

| | |
|--|-----------|
| Resumo | 5 |
| Abstract | 6 |
| INTRODUÇÃO | 7 |
| MÉTODOS..... | 14 |
| Área de estudo..... | 14 |
| Amostragem e obtenção dos dados limnológicos e ambientais..... | 15 |
| Obtenção dos dados moleculares | 17 |
| Análises bioinformáticas e obtenção da filogenia | 21 |
| Análises de riqueza e diversidade de OTUs | 23 |
| Análises de diversidade e estrutura filogenéticas (diversidade alfa filogenética) | 23 |
| Análises ambientais e espaciais para a diversidade alfa filogenética | 24 |
| Diversidade beta filogenética..... | 25 |
| RESULTADOS | 26 |
| Análises de bioinformáticas e classificação das OTUs..... | 26 |
| Árvores filogenéticas | 33 |
| Diversidade filogenética (PD)..... | 34 |
| Estrutura filogenética da comunidade..... | 38 |
| Influência das variáveis ambientais e espaciais na diversidade alfa filogenética da comunidade | 42 |
| Diversidade beta filogenética..... | 47 |
| DISCUSSÃO | 48 |
| CONCLUSÃO..... | 59 |
| REFERÊNCIAS..... | 61 |
| APÊNDICES | 83 |

Resumo

A integração de estudos filogenéticos e de comunidades tem se mostrado bastante eficiente para desvendar os processos ecológicos e evolutivos que moldam a biodiversidade. Nesse contexto, a combinação do *metabarcoding* de eDNA com análises filogenéticas de comunidade tem se mostrado particularmente promissora para descrever os padrões de organização e montagem das comunidades, proporcionando estimativas de diversidade mais robustas e informativas. Este estudo teve como objetivo investigar a diversidade filogenética da comunidade de eucariotos capturados pelo marcador molecular 18S na planície de inundação do rio Araguaia, utilizando a técnica de *metabarcoding*. Com foco nos grupos de microeucariotos planctônicos fotossintetizantes (MEPF) buscamos compreender como fatores ambientais e espaciais influenciam a diversidade e estrutura filogenética dessas comunidades. Para isso, foram coletadas amostras de água em 140 lagos da planície de inundação do rio Araguaia. O DNA total foi extraído e a região V4 do gene 18S rRNA foi amplificada e sequenciada na plataforma Illumina MiSeq. As sequências foram tratadas utilizando o pipeline Pimba para obtenção das unidades taxonômicas operacionais (OTUs). Uma filogenia total foi construída com o totas as OTUs obtidas, e a partir dessa filogenia, foram selecionados os clados pertencentes aos grupos de MEPF para construir uma segunda filogenia, centrada apenas nos MEPF. A diversidade e estrutura filogenética foi avaliada para ambas as filogenias utilizando os índices de diversidade filogenética de Faith (PD) e índice de parentesco líquido (NRI). A influência de variáveis ambientais e espaciais na diversidade e estrutura filogenética foi analisada por meio de análises de redundância parcial (pRDA) e partição da variância. E a diversidade beta filogenética foi estimada utilizando o índice PhyloSor e particionada em aninhamento e turnover, e sua similaridade com o espaço e ambiente foi verificada com um teste parcial de Mantel. Os resultados mostram que a diversidade filogenética da comunidade de eucariotos e microeucariotos planctônicos fotossintetizantes foi significativamente menor do que o esperado ao acaso, indicando uma estruturação não aleatória da comunidade. A maioria dos lagos apresentou um padrão filogenético agrupado, sugerindo que táxons mais relacionados filogeneticamente co-ocorrem com maior frequência. A maior parte da variação na diversidade e estrutura filogenética foi explicada pelo espaço, porém o ambiente espacialmente estruturado também apresentou uma porcentagem relevante de explicação. Os filtros espaciais que descrevem a conectividade direcional entre os lagos apresentaram um maior poder de explicação do que filtros que usam distâncias lineares. A diversidade beta filogenética para os microeucariotos planctônicos fotossintetizantes foi elevada, indicando uma alta substituição de espécies entre os lagos. E a sua dissimilaridade estava positivamente correlacionada com as distâncias ambientais e espaciais. Estes resultados sugerem uma importância significativa da filtragem ambiental e dos processos de dispersão na organização da comunidade filogenética. E contribui para o preenchimento da lacuna de conhecimento sobre a diversidade filogenética de microeucariotos planctônicos da bacia hidrográfica do rio Tocantins-Araguaia, destacando a importância de considerar tanto os aspectos filogenéticos quanto os ecológicos para compreender a biodiversidade aquática.

Palavras-chave: DNA ambiental; plâncton; 18S rRNA; estrutura filogenética; OTUs.

Abstract

Integrating phylogenetic and community ecology studies has proven highly effective in unraveling the ecological and evolutionary processes that shape biodiversity. In this context, combining eDNA metabarcoding with community phylogenetic analyses holds promise for describing patterns of community organization and assembly, providing more robust and informative estimates of diversity. This study aimed to investigate the phylogenetic diversity of the eukaryotic community captured by the 18S molecular marker in the Araguaia River floodplain, using the metabarcoding technique. Focusing on groups of photosynthetic planktonic microeukaryotes (PPME), we sought to understand how environmental and spatial factors influence the diversity and phylogenetic structure of these communities. To this end, water samples were collected from 140 lakes in the Araguaia River floodplain. Total DNA was extracted, and the V4 region of the 18S rRNA gene was amplified and sequenced on the Illumina MiSeq platform. Sequences were processed using the Pimba pipeline to obtain operational taxonomic units (OTUs). A total phylogeny was constructed with all OTUs obtained, and from this phylogeny, clades belonging to the PPME groups were selected to construct a second phylogeny focused solely on PPME. Phylogenetic diversity and structure were assessed for both phylogenies using Faith's phylogenetic diversity (PD) index and the net relatedness index (NRI). The influence of environmental and spatial variables on phylogenetic diversity and structure was analyzed through partial redundancy analysis (pRDA) and variance partitioning. Phylogenetic beta diversity was estimated using the PhyloSor index and partitioned into nestedness and turnover, while its correlation with spatial and environmental factors was assessed with a partial Mantel test. The results show that the phylogenetic diversity of the eukaryotes and photosynthetic planktonic microeukaryotes communities was significantly lower than expected by chance, indicating a non-random community structure. Most lakes showed a clustered phylogenetic pattern, suggesting that more phylogenetically related taxa co-occur more frequently. Most of the variation in phylogenetic diversity and structure was explained by space, but the spatially structured environment also had a relevant percentage of explanation. Spatial filters describing the directional connectivity between lakes had a greater explanatory power than filters using linear distances. Phylogenetic beta diversity for photosynthetic planktonic microeukaryotes was high, indicating substantial species turnover among lakes, and dissimilarity was positively correlated with environmental and spatial distances. These results suggest a significant role for environmental filtering and dispersal processes in the phylogenetic community organization and contribute to filling the knowledge gap regarding the phylogenetic diversity of planktonic microeukaryotes in the Tocantins-Araguaia River basin, underscoring the importance of considering both phylogenetic and ecological aspects to understand aquatic biodiversity.

Keywords: environmental DNA; plankton; 18S rRNA; phylogenetic structure; OTUs.

INTRODUÇÃO

A integração de estudos filogenéticos e de comunidade, oferece uma poderosa ferramenta para compreender os processos ecológicos e evolutivos que moldam a biodiversidade, como apontado por Webb *et al.* (2002). Uma das formas que podemos integrar estes dois campos de estudo, é examinando a estrutura filogenética das comunidades. Este exame nos permite identificar como as espécies estão distribuídas filogeneticamente entre os habitats e observar os padrões filogenéticos de organização da comunidade (Webb *et al.*, 2002). Essa abordagem propicia uma nova perspectiva sobre os fatores que contribuem para a manutenção da diversidade e sobre a importância relativa dos diferentes processos ecológicos e evolutivos que impulsionam a montagem da comunidade (Webb *et al.*, 2002; Cavender-Bares *et al.*, 2009).

A estrutura e a diversidade filogenética de uma comunidade refletem, respectivamente, a quantidade de história evolutiva e as relações evolutivas entre as espécies que coabitam em um local. A diversidade filogenética é uma medida da quantidade de história evolutiva ou divergência evolutiva de um conjunto de espécies (Tucker *et al.*, 2017) e é dada pela soma de comprimento de ramos de uma árvore filogenética, que assume um modelo evolutivo em que características compartilhadas são explicadas pela ancestralidade comum (Faith, 1992, 2016). Assim, comunidades contendo táxons mais distantemente relacionados, capturam mais divergência evolutiva, e apresentam maiores valores de diversidade filogenética, do que comunidades que tem táxons mais intimamente relacionados (Tucker *et al.*, 2019). A estrutura filogenética é uma medida do padrão das relações evolutivas/filogenéticas das espécies de uma comunidade, medida pela distância média entre os ramos de uma filogenia (Webb *et al.*, 2002). Esses padrões podem ser classificados como agrupados, superdispersos ou aleatórios (Webb *et al.*, 2002). O padrão agrupado indica que espécies filogeneticamente próximas tendem a coexistir, frequentemente devido à conservação de características adaptativas que as tornam bem-sucedidas em ambientes similares. Já o padrão superdisperso sugere que espécies próximas evolutivamente podem evitar a competição direta, ocupando nichos ecológicos distintos. Por outro lado, o padrão aleatório pode ser resultado de processos estocásticos ou de convergência evolutiva para características semelhantes. Esses padrões são frequentemente moldados por uma combinação de fatores, incluindo a conservação de nicho ecológico, a história evolutiva e os processos ecológicos como a competição e a

dispersão (Lovette e Hochachka, 2006; Li, Ives e Waller, 2017). Tanto a estrutura como a diversidade filogenética são medidas de diversidade alfa filogenética, dadas pela riqueza evolutiva observada na comunidade (Tucker *et al.*, 2017).

A compreensão da diversidade e dos padrões filogenéticos de uma comunidade é fundamental para entender os processos ecológicos e evolutivos que moldam a biodiversidade. Esses processos são influenciados por uma complexa interação de fatores, incluindo interações ecológicas, como competição e predação (Ives e Helmus, 2011), processos neutros, como deriva genética e colonização aleatória (Hubbell, 2001). Flutuações nas condições abióticas locais e fatores históricos específicos, além de processos globais como dispersão e especiação, que juntos desempenham papéis significativos na estruturação da diversidade das comunidades (Ricklefs, 1987). Os processos relacionados a dispersão e as condições abióticas locais são amplamente estudados para comunidades aquáticas.

Na metacomunidade, definida por comunidades locais conectadas pela dispersão de múltiplas espécies que potencialmente interagem entre si (Wilson, 1992; Leibold *et al.*, 2004), a composição das comunidades locais é regulada por processos relacionados a capacidade de dispersão, fatores ambientais, interações bióticas e processos estocásticos (Leibold *et al.*, 2004; Thompson *et al.*, 2020). Os fatores ambientais, atuarão em uma escala local promovendo a filtragem ambiental que atuará no sucesso do estabelecimento e persistência dos organismos de acordo com as condições ambientais, e assim incluindo ou excluindo espécies nas comunidades de acordo com as suas tolerâncias às características do ambiente (Poff, 1997; Kraft *et al.*, 2015). A filtragem ambiental é frequentemente associada a uma estrutura filogenética agregada (Webb *et al.*, 2002; Kembel e Hubbell, 2006) e a formação de uma estrutura filogenética mais forte quando os táxons derivados são favorecidos na comunidade (Kraft *et al.*, 2007).

A capacidade de dispersão das espécies e a limitação da dispersão está associada a uma mudança na composição filogenética das comunidades (Daru *et al.*, 2017). A capacidade de dispersão afetará a distribuição espacial dos organismos na metacomunidade em uma escala regional. Essa capacidade de dispersão pode ser limitada pelo aumento do isolamento ou impulsionadas pelo efeito de massa, sendo afetados pela distância espacial entre os locais e extensão espacial pesquisada (Heino *et al.*, 2015). Em rios e riachos, a direcionalidade do fluxo de água pode desempenhar um papel crucial na dispersão, sendo fortemente afetada por efeitos de massa, o que pode levar à homogeneização das

comunidades (Heino *et al.*, 2015). Em lagos, a classificação de espécies e a limitação da dispersão são fatores importantes que estruturam a metacomunidade (Heino *et al.*, 2015).

Para separar os papéis dos processos ambientais e espaciais na organização da comunidade em ambientes aquáticos, o uso de diferentes métricas de distância espacial que especifique diferentes possibilidades de relação espacial entre os locais, como distância terrestre (Overland), distância ao longo do curso de água (Waterhouse) e processos direcionais podem ser fundamentais (Heino *et al.*, 2015). A importância da rede fluvial na explicação de padrões da metacomunidade depende de fatores como o modo de dispersão, tipo de paisagem e conectividade entre locais (Tonkin *et al.*, 2018). Podemos visualizar melhor a importância dessa rede para dispersão utilizando múltiplas métricas de distância espacial simultaneamente (Tonkin *et al.*, 2018). Devido a ampla gama de modo de dispersão de alguns organismos o uso de diferentes métricas de distância simultaneamente podem capturar melhor o papel da dispersão na metacomunidade.

A importância dos processos regionais e locais, como a dispersão e filtragem ambiental, na estrutura da comunidade variam de acordo com as características das espécies, a extensão espacial estudada e a heterogeneidade ambiental das regiões analisadas (Grönroos *et al.*, 2013; Heino *et al.* 2015). Em escalas espaciais mais finas e intermediárias, predominantemente os fatores ambientais parecem influenciar mais a estrutura das metacomunidades aquáticas (Lansac-Tôha *et al.*, 2021). Já em escalas maiores, os fatores espaciais têm um papel importante na estruturação da comunidade (Li, Tonkin e Haase, 2020). As escalas espaciais também impactam os padrões de estrutura filogenética e a força dos processos que determinam a composição filogenética das comunidades (Webb *et al.* 2002; Cavender-Bares *et al.*, 2009; Cavender-Bares, Keen e Miles, 2006).

Para os microrganismos aquáticos, os fatores ambientais são os principais reguladores de suas comunidades (Lansac-Tôha *et al.*, 2021), com as características do habitat sendo importantes fontes de variação para estes organismos (Heino *et al.*, 2014). No entanto, fatores ambientais e espaciais podem influenciar os padrões de diversidade comunitária de maneiras distintas em diferentes escalas, em comunidade de microeucariotos planctônicos (Martin *et al.*, 2021). Em escalas maiores, ambos os fatores podem ter uma influência importante, enquanto em escalas menores, o ambiente pode ter um papel predominante (Martin *et al.*, 2021; Jin *et al.*, 2022).

Os microeucariotos planctônicos desempenham um papel fundamental na manutenção da vida na Terra e nos ecossistemas aquáticos, regulando as concentrações de CO₂ na atmosfera, constituindo a base das cadeias alimentares aquáticas (Smetacek e Cloern, 2008), atuando como um importante agente na ciclagem de nutrientes e na determinação da produtividade dos ecossistemas aquáticos (Henson *et al.*, 2021; Field *et al.*, 1998). Além de serem frequentemente usados como bioindicadores da qualidade da água e estado ecológico de lagos e rios (Hardikar, Haridevi e Deshbhratar, 2022; Hanžek *et al.*, 2021; Lacdan *et al.* 2014), devido a sua sensibilidade às mudanças ambientais. Apesar de sua importância ecológica, esses organismos foram relatados como menos estudados em termos de diversidade filogenética quando comparados a comunidades de plantas terrestres e microrganismos procariontes (Vamosi *et al.*, 2009; Vamosi, 2013). Embora alguns avanços tenham sido feitos nos últimos anos (Ex. Thompson, Davies e Gonzalez, 2015; Gianuca, Pantel e De Meester, 2016; Skouroliakou *et al.*, 2022; Han *et al.*, 2022; Wang *et al.*, 2024), uma lacuna de conhecimento ainda parece persistir, dada a escassez de estudos encontrados na literatura.

Os microeucariotos planctônicos são sensíveis as condições ambientais, alterações no pH da água, por exemplo, pode levar a diminuição da riqueza e alteração da composição e diversidade dos grupos taxonômicos, impactando a montagem da comunidade e sua estrutura filogenética (Ren *et al.*, 2015). Para as comunidades de microeucariotos planctônicos fotossintetizantes, a disponibilidade de nutrientes e as características físico-químicas do ambiente, como temperatura, salinidade e oxigênio dissolvido, influenciam o crescimento, diversidade e abundância de suas comunidades (Vajravelu *et al.*, 2018). Esses microrganismos exibem uma sensibilidade às mudanças no ambiente, que podem provocar alterações na diversidade, reorganização e um aumento da instabilidade na estrutura de suas comunidades (Henson *et al.*, 2021), com possíveis consequências para os ciclos biogeoquímicos locais e até globais (Litchman *et al.*, 2015). A compreensão da diversidade e estrutura filogenética dessas comunidades é um campo de grande interesse para compreender a montagem da comunidade planctônica e os processos ecológicos envolvidos nela (Vamosi, 2013).

Os avanços recentes na área da genética molecular, tem permitido um crescente uso de técnicas moleculares para estudar comunidades microbianas aquáticas, incluindo a exploração da diversidade filogenética dessas comunidades (Grossart *et al.*, 2020). Nos últimos anos a técnica de DNA *metabarconding* tem se mostrado muito útil e promissora,

permitindo análises de muitas amostras com sensibilidade e resolução taxonômica (Santoferrara, 2019). O *metabarcoding* é uma técnica que engloba ecologia de campo, métodos moleculares e ferramentas computacionais avançadas (Ruppert, Kline e Rahman, 2019) e permite a identificação de múltiplas espécies a partir de uma amostra ambiental. Ela é capaz de detectar espécies raras e invasoras, monitorar a diversidade de comunidades, e tem se tornado uma poderosa ferramenta para a conservação (Sahu *et al.*, 2023). As sequências curtas de DNA (*amplicons*) de regiões específicas (Ex. gene 18S, 16S e ITS) gerada pela técnica, podem ser agrupadas por similaridade formando as unidades taxonômicas operacionais (OTUs) ou mantidas como variantes de sequências de *amplicons* (ASVs), e permite uma identificação eficaz da composição taxonômica microbiana (Liu *et al.*, 2021). Mas ainda existe um desafio na identificação de todos os táxons devido a incompletude dos bancos de dados (Hestetun *et al.*, 2020) e sequências mal classificadas depositadas em banco de dados públicos (Bagheri, Severin e Rajan, 2020). Apesar disso, sequências curtas vêm sendo comumente utilizadas na construção de árvores filogenéticas moleculares (Choi *et al.*, 2020; Yang *et al.*, 2021; Bock *et al.*, 2022; Wang *et al.*, 2024), principalmente devido a sua facilidade e rapidez na montagem das árvores filogenéticas.

Os estudos de DNA *metabarcoding* normalmente utilizam o DNA obtido de amostras ambientais, o eDNA (environmental DNA ou DNA ambiental), que é uma mistura complexa de DNA de muitos organismos diferentes que pode ser encontrada no solo, na água, no ar, sedimentos entre outros ambientes (Taberlet *et al.*, 2012). O *metabarcoding* de eDNA tem se mostrado eficiente e confiável para abordagens em estudos ecológicos e da biodiversidade aquática, podendo também ser utilizado como complementar a abordagens tradicionais, sendo eficiente em detecção de espécies raras ou crípticas, e garantido um monitoramento eficiente e preciso de ecossistemas aquáticos (Valentini *et al.*, 2015). O *metabarcoding* de eDNA tem sido muito usado para investigar a montagem de comunidades planctônicas (Banerji *et al.*, 2018; David *et al.*, 2021; Mo *et al.*, 2021; Liu *et al.*, 2022; Li *et al.*, 2023; Shen *et al.*, 2024). E seu uso aliado as análises de diversidade filogenética da comunidade, pode ser muito vantajoso para descrever os padrões ambientais das comunidades, fornecendo estimativas de diversidade mais consistentes e informativas (Diniz-Filho *et al.*, 2024).

A planície de inundação do rio Araguaia, com sua grande extensão e conectividade, oferece um cenário ideal para investigar a estrutura e a dinâmica das comunidades filogenéticas de microeucariotos planctônicos em ampla escala espacial. O rio Araguaia é

um importante rio da região centro-oeste do Brasil (Valente, Latrubesse e Ferreira, 2013), que possui uma importante planície de inundação, que forma um complexo mosaico de unidades inter-relacionadas, formando um sistema que regula uma importante área de ecossistemas aquáticos do Cerrado (Latrubesse e Stevaux, 2006). As planícies de inundação são áreas que são periodicamente inundadas pelo transbordamento lateral de rios ou lagos, e possui uma biota adaptada a dinâmica do pulso de inundação com estruturas comunitárias características (Junk *et al.*, 1989). O rio Araguaia faz parte da bacia do rio Tocantins-Araguaia uma das mais importantes bacias hidrográficas do Brasil que abrange duas regiões fitogeográficas distintas, o Cerrado e a Amazônia, que concentram grande biodiversidade (Latrubesse e Stevaux, 2006). Esta bacia vem sofrendo com as intensas mudanças e ameaças ambientais que tem levado a sua degradação em grande escala, colocando em risco a biodiversidade e serviços ecossistêmicos locais (Pelicice *et al.*, 2021), tornando este ecossistema uma área prioritária para o monitoramento e estudos da biodiversidade local.

O complexo mosaico de unidades inter-relacionadas da planície de inundação do rio Araguaia, cria um sistema de lagos interconectados ao curso do rio, estes ambientes são dominados por comunidades locais conectadas pela dispersão de múltiplas espécies, formando uma metacomunidade (Wilson, 1992; Leibold *et al.*, 2004) e assim integrando processos locais e regionais, como interações entre espécies, espécie-ambiente e dispersão, que atuam na montagem e estrutura da comunidade local (Logue *et al.*, 2011).

A região da bacia do Rio Araguaia, tem sido muito estudada nos últimos anos. Vários estudos tem sido feitos para avaliar a diversidade e qualidade ambiental da região, que inclui, estudos a respeito do regime hidrológico e ecossistemas da região (Aquino, Latrubesse e Souza-Filho, 2008), suas características físico-bióticas (Latrubesse e Stevaux, 2006), estudos sobre as comunidades de microrganismos aquáticos (Nabout, Nogueira e Oliveira, 2006; Machado *et al.*, 2016) e outros organismos (Tejerina-Garro *et al.*, 1998; Pereira e Colli, 2023; Paschoalini *et al.*, 2020), sobre as características limnológicas das lagoas (Alves *et al.*, 2019; Marcionílio *et al.*, 2016), uso e cobertura do solo (Martins *et al.*, 2021), estruturação das metacomunidades aquáticas (Lansac-Tôha *et al.*, 2021), estudos de paisagem (Valente, Latrubesse e Ferreira, 2013), dinâmica do mercúrio (Fernandes *et al.*, 2024; Monteiro *et al.*, 2024), entre outros. Dentre esses, foi realizado um estudo feito com técnicas moleculares sobre os padrões de diversidade de comunidades de microeucariotos planctônicos, que analisou 27 lagos na região do médio Araguaia ao longo de 500 km na

planície de inundação do rio e seus afluentes (Machado *et al.*, 2019). Neste estudo, os autores observaram uma composição de microeucariotos planctônicos espacialmente estruturada ao longo da planície, sem efeitos direto do ambiente ou espaço, mas com uma fração de explicação considerável do componente compartilhado na composição da comunidade, indicando que variáveis ambientais espacialmente estruturadas podem estar atuando para determinar a composição dessas comunidades.

O presente estudo propõe uma expansão do trabalho de Machado *et al.* (2019), em termos da amostragem de lagos, que engloba quase toda a região do médio Araguaia, perfazendo cerca de 1.500 km entre o rio Araguaia e seus afluentes, cobrindo assim, uma grande parte desse ecossistema aquático. Além disso, no presente estudo nós utilizamos uma abordagem analítica baseada na diversidade filogenética molecular, com o intuito de explorar melhor os dados obtidos e incluir os aspectos evolutivos na determinação da estrutura da comunidade e sua relação com características espaciais e ambientais. Tendo em vista estes aspectos, o objetivo deste trabalho é com o uso da ferramenta molecular de DNA *metabarcoding*, investigar a diversidade filogenética da comunidade de eucariotos capturados pelo marcador molecular 18S na planície de inundação do rio Araguaia, focando nos microeucariotos planctônicos fotossintetizantes (MEPF). Os microeucariotos planctônicos fotossintetizantes são a base da cadeia alimentar aquática e são organismos essenciais para a manutenção dos ecossistemas aquáticos (Smetacek e Cloern, 2008; Field *et al.*, 1998). Estes organismos são identificados com frequência utilizando a região 18S do RNA ribossomal (Penna *et al.*, 2017; Kezlya, Tseplik e Kulikovskiy, 2023; Akcaalan *et al.*, 2023) e já foi demonstrado que esta região tem uma maior riqueza de sequências depositadas em banco de dados públicos para estes organismos (Furtado *et al.*, 2024) o que facilita a identificação das espécies por meio do *metabarcoding*. Baseado nisso, nossos objetivos específicos são: i) Determinar a diversidade e estrutura filogenética da comunidade; ii) Avaliar a diversidade beta filogenética da comunidade; iii) Testar a influência das variáveis ambientais e espaciais para explicar a variação na diversidade filogenética da comunidade; iv) Verificar se espécies filogeneticamente próximas ocupam ambientes semelhantes; v) Verificar se os padrões de diversidade e estrutura filogenética observados são semelhantes entre a comunidade de eucariotos (comunidade total) e a comunidade de microeucariotos planctônicos fotossintetizantes (comunidade focal).

MÉTODOS

Área de estudo

O rio Araguaia faz parte da bacia hidrográfica Tocantins-Araguaia, considerada a quarta maior bacia da América do Sul, com uma área de drenagem que se estende por 777.308 km² (Aquino, Latrubesse e Souza-Filho, 2009). A nascente do rio fica no sudoeste do estado de Goiás (próximo ao Parque Nacional das Emas - GO, Brasil) e deságua no Rio Tocantins (próximo a São João do Araguaia – PA, Brasil). O rio Araguaia faz a divisa dos estados de Goiás e Mato Grosso, Tocantins e Mato Grosso e Tocantins e Pará e percorre cerca de 2.100 Km até a confluência com o rio Tocantins (Aquino, Latrubesse e Souza-Filho, 2009). O rio é dividido em alto, médio e baixo Araguaia e possui uma área de escoamento, de cerca de 377.000 km² (Latrubesse e Stevaux, 2006; Aquino, Stevaux e Latrubesse, 2005). O alto Araguaia tem cerca de 450 km de extensão e uma área de drenagem de cerca de 36.400 km², vai da nascente até a região de Registro do Araguaia - GO, o médio Araguaia se estende por 1.600 km e possui uma área de drenagem maior que 300.000 km², indo de Registro do Araguaia - GO até a cidade de Conceição do Araguaia - PA (Latrubesse e Stevaux, 2006; Aquino, Latrubesse e Souza-Filho, 2009). Neste trecho se encontra os importantes rios tributários, como o rio das Mortes, rio Vermelho, rio Cristalino e rio Crixás que contribuem para o aumento da área de drenagem do médio Araguaia (Latrubesse e Stevaux, 2006). O baixo rio Araguaia tem cerca de 500 km de extensão com uma área de drenagem de cerca de 40.000 km², indo de Conceição do Araguaia – PA, até a sua confluência com o rio Tocantins (Latrubesse e Stevaux, 2006).

O rio Araguaia tem um regime hidrológico dependente do clima quente a semiúmido, característico do Cerrado (Aquino, Stevaux e Latrubesse, 2005). Apresenta uma temporada de cheia e outra de seca, de acordo com o clima do Cerrado, com picos de cheias ocorrendo entre os meses de janeiro a maio e o período de vazantes ocorrendo entre junho e setembro, variando ao longo das regiões do rio (Aquino, Latrubesse e Souza-Filho, 2008). Em razão dessa sazonalidade do regime de chuvas da região, o rio Araguaia possui uma importante planície de inundação com uma área total que pode chegar a 88.119 km² de terra coberta por água durante o período de cheia, e cerca de 2.930 km² de área úmida coberta por água durante o período de seca (Irion *et al.*, 2016). Esse regime hidrológico da região é muito importante para a manutenção dos ecossistemas aquáticos (Aquino, Latrubesse e Souza-Filho, 2008).

Amostragem e obtenção dos dados limnológicos e ambientais

Foram coletadas amostras de água de 140 lagos da planície de inundação do médio Araguaia e seus afluentes (rio Cristalino, rio das Mortes, rio Vermelho, rio do Peixe, córrego Água Limpa e rio Crixás) (Fig. 1) ao longo do mês de fevereiro de 2023, período de cheia do rio. Em cada lago foram amostrados 1,5 litros de água, coletados logo abaixo da superfície (cerca de 30 cm de profundidade). Para a coleta dos 1,5 litros de água foram utilizados três frascos de 500 ml branco leitoso. Cada 500 ml de água foi coletado de um ponto diferente na zona limnética do lago, com o intuito de aumentar a captura da diversidade biológica. Posteriormente, as amostras foram armazenadas em caixas térmicas com gelo até o transporte ao barco-hotel onde foram filtradas em até 12 horas após a coleta. Cada amostra de 500 ml foi filtrada com o uso de uma membrana de acetato de celulose com porosidade de 0,45 μm e 47 milímetros de diâmetro (Millipore), com o auxílio de uma bomba a vácuo. A membrana de porosidade 0,45 μm é capaz de capturar organismos inteiros de microeucariotos planctônicos, alguns organismos do picoplâncton (0,2-2,0 μm) e todos maiores que estes. Os equipamentos de filtração foram esterilizados com hipoclorito de sódio comercial a 10% e depois enxaguados com água destilada por três vezes. Além disso, a bancada de uso foi higienizada continuamente com álcool 70%. Após a filtração, cada membrana foi armazenada separadamente em um tubo eppendorf de 2 ml com solução tampão de Longmire (Longmire *et al.*, 1997) e mantidas refrigeradas a cerca de -4°C , até o transporte para o laboratório de Genética & Biodiversidade, da Universidade Federal de Goiás, onde foram armazenadas a -80°C .

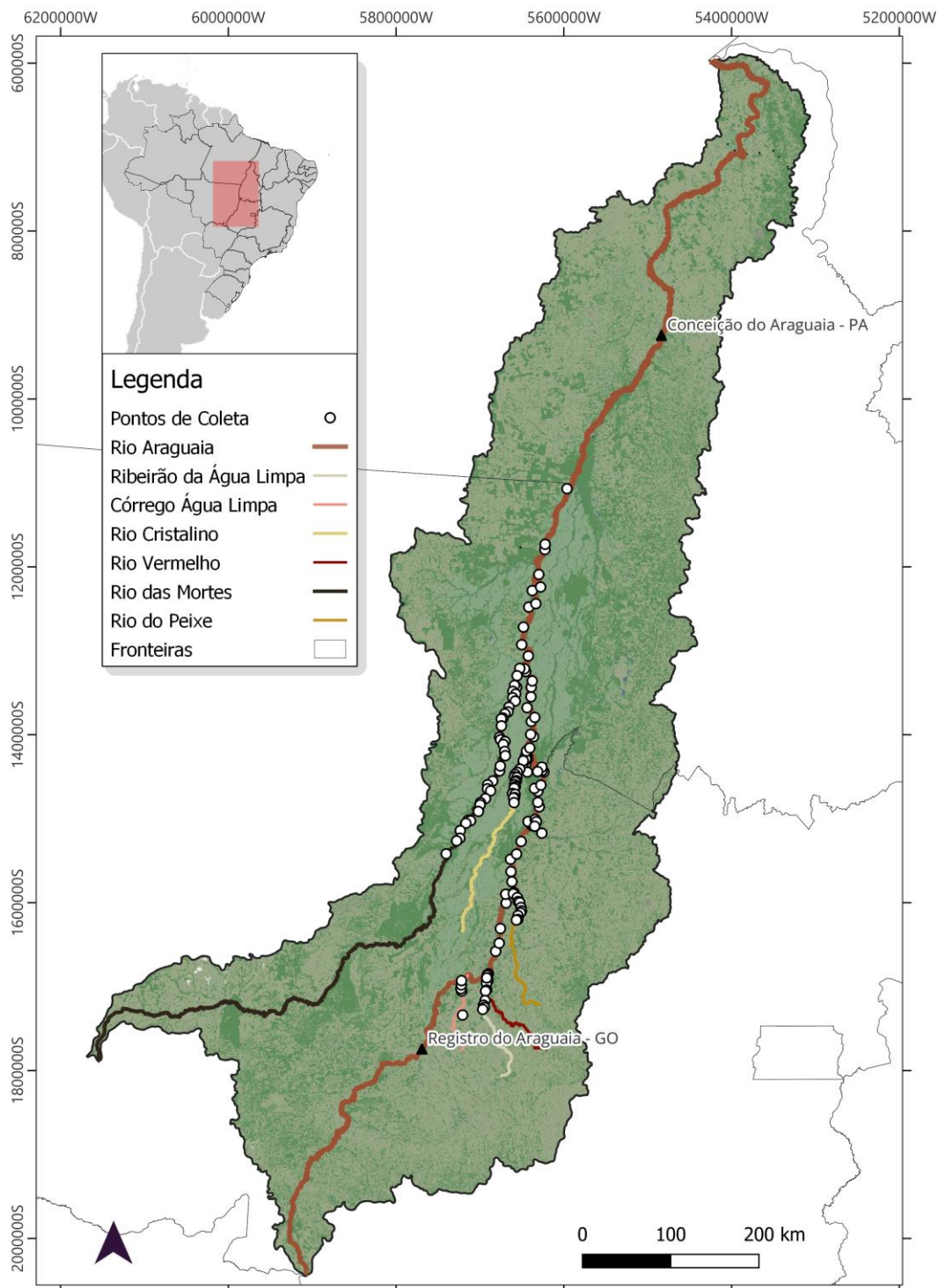


Figura 1 - Localização dos 140 pontos de coleta, ao longo do médio Araguaia e seus afluentes (rio Cristalino, rio das Mortes, rio Vermelho, rio do Peixe, córrego Água Limpa e rio Crixás).

Em cada lago amostrado foram coletadas as seguintes variáveis ambientais e limnológicas: condutividade elétrica da água (mS/cm), oxigênio dissolvido (OD – mg/L), porcentagem de oxigênio dissolvido (%OD), pH, temperatura da água (°C), turbidez (NTU- Unidade Nefelométrica de Turbidez), sólidos totais dissolvidos (TDS – g/L), potencial de

óxido-redução (ORPmV), que foram medidos em campo com o auxílio de um analisador multiparâmetro (Horiba U-50). A temperatura (°C) e a umidade do ar (%RH) foram medidas com o auxílio de um multianalisador ambiental. A profundidade (m) foi determinada por uma sonda de profundidade e a transparência (m) foi analisada com o auxílio de um disco de Secchi.

Para obtenção da concentração de pigmentos totais, em cada lago foram coletados 1 litro de água em frascos branco leitoso, a água foi filtrada com uma membrana de acetato de celulose (0,45 µm e 47 mm) com o auxílio de uma bomba à vácuo até entupimento dos poros. Após a filtração, as membranas foram armazenadas em envelopes de papel alumínio e mantidas congeladas a -20 °C até o transporte para o laboratório de Ecologia Aquática na Universidade Federal de Goiás (UFG). Para obtenção dos valores de nutrientes foram coletadas em cada lago 500 ml de água em frascos transparentes que foram mantidas congeladas a -20 °C até transporte para o laboratório de Ecologia Aquática na UFG. Em laboratório, as concentrações de pigmentos totais (mg/l) foram medidas através da extração com acetona (90%) e leitura em espectrofotômetro (Golterman, Clymo e Ohmstad, 1978). O ortofosfato (µg/L) e fósforo total (µg/L) foram determinadas usando o método de ácido ascórbico e posterior leitura em espectrofotômetro (Golterman, Clymo e Ohmstad, 1978). Os valores para nitrato (mg/L) e o nitrogênio amoniacal (mg/L) foram obtidos através do método de redução de cádmio e método salicilato, respectivamente, e posterior leitura em espectrofotômetro, como descritos em APHA (2023).

Obtenção dos dados moleculares

O DNA das amostras dos 140 lagos foi extraído a partir dos filtrados das membranas de acetato de celulose utilizando o kit comercial DNeasy *Blood and Tissue* (Qiagen). O kit Dneasy *Blood and Tissue* (Qiagen) é comumente utilizado em extrações de DNA ambiental usando filtros de membranas (Stat *et al.*, 2017; Sawaya *et al.*, 2019; Sard *et al.*, 2019; Djurhuus *et al.*, 2020). O protocolo original de tecido do kit foi modificado para aumentar a eficiência e o volume da extração (Apêndice 1). O DNA foi extraído separadamente de metade de cada membrana, totalizando 420 extrações (três membranas por ponto de coleta). A eficiência do protocolo de extração foi avaliada pela intensidade e integridade da banda de DNA genômico em gel de agarose 1% corado com brometo de etídeo. Controles negativos de extração, utilizando filtrados de água destilada, foram

realizados para monitorar possível contaminação durante o processo de extração. O DNA extraído foi quantificado por eletroforese horizontal em gel de agarose 1% corado com brometo de etídeo. Após a quantificação das 420 amostras de DNA, elas foram combinadas (juntando as três amostras de DNA de cada ponto), formando um pool de DNA único para cada um dos 140 lagos, e em seguida quantificadas novamente (Fig. 2).

A amplificação via PCR (Reação em Cadeia da Polimerase) foi realizada em triplicatas para todas as 140 amostras de DNA extraídas dos lagos, visando minimizar o viés da técnica. O par de *primers* TAREuk454FWD1 e TAREukREV3 (Stoeck *et al.*, 2010) foi utilizado para amplificar a região 18S-V4 rRNA, que possui aproximadamente 380 pares de bases (pb) e é comumente empregada em estudos de DNA *metabarcoding* para microeucariotos planctônicos. Os primers foram sintetizados com adição de sequências do adaptador Illumina (TAREuk454FWD1: 5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCAGCASCYGC GGTAATTC C-3' - TAREukREV3: 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACTTTCGTTCTTGATYRA-3'). As reações de PCR foram realizadas em um volume final de 25 µL, contendo 5,0 µL de DNA (5 ng/µL), 2,5 µL de cada primer (5 µM), 12,5 µL de Multiplex PCR (Qiagen, Alemanha) e 2,5 µL de Qsolution. As amplificações ocorreram no termociclador Applied Biosystems® Veriti, seguindo as etapas: 95°C por 15 minutos, seguido de 40 ciclos de 94°C por 30 segundos, 53°C por 30 segundos e 72°C por 30 segundos, com uma extensão final a 72°C por 10 minutos. Controles negativos de PCR foram incluídos em cada placa para monitorar a possível contaminação durante o processo. Utilizamos eletroforese horizontal em gel de agarose 1,5% corada com brometo de etídeo para avaliar a qualidade dos fragmentos amplificados, verificar visualmente a presença de contaminação via brancos/negativos, a presença da banda alvo e a consistência dos resultados entre as réplicas de PCR.

O preparo da biblioteca de DNA para sequenciamento seguiu o protocolo padrão 16S da Illumina (16S *Metagenomic Sequencing Library Preparation*, disponível em: https://support.illumina.com/downloads/16s_metagenomic_sequencing_library_preparation.html). Nesta etapa, as triplicatas de PCR de cada ponto foram combinadas, formando uma amostra por ponto, e os controles negativos que amplificaram na PCR foram incluídos no preparo da biblioteca. As amostras foram purificadas com 60 µL de beads magnéticas IPB (Illumina *Purification Beads*) para cada volume de PCR. Após a purificação foi

realizada a PCR de ciclo limitado (8 ciclos) para adição dos indexes Nextera XT (N7XX e S5XX) aos fragmentos de DNA de cada ponto, utilizando o kit KAPA Hotstart Hifi ReadyMix PCR. Em seguida, foi realizada a etapa de purificação da biblioteca com as beads magnéticas IPB para limpeza dos produtos de PCR de ciclo limitado. A integridade e o tamanho médio dos fragmentos foram avaliados a partir de 8 amostras aleatórias no Bioanalyzer com o kit DNA1000 Agilent. A quantificação das bibliotecas (ng/μL) foi realizada no Qubit Fluorometer (Thermo Scientific) com o kit High Sensitivity, utilizando diluição 1:3. As bibliotecas foram diluídas a 4 nM em Tris-HCl 10 mM. O pool das bibliotecas a 4 nM foi montado com 5 μL de cada biblioteca em um tubo de 1,5 mL e quantificado no Qubit Fluorometer para confirmar a concentração total de 4 nM. A biblioteca foi desnaturada a 4 nM e diluída a 9 pM conforme o protocolo padrão. O pool de biblioteca foi codificado com 40% de PhiX para aumentar a diversidade de fragmentos nas corridas de sequenciamento. O sequenciamento pareado (2x300) foi realizado na plataforma Illumina MiSeq utilizando o kit MiSeq Reagent v3 (600 ciclos).

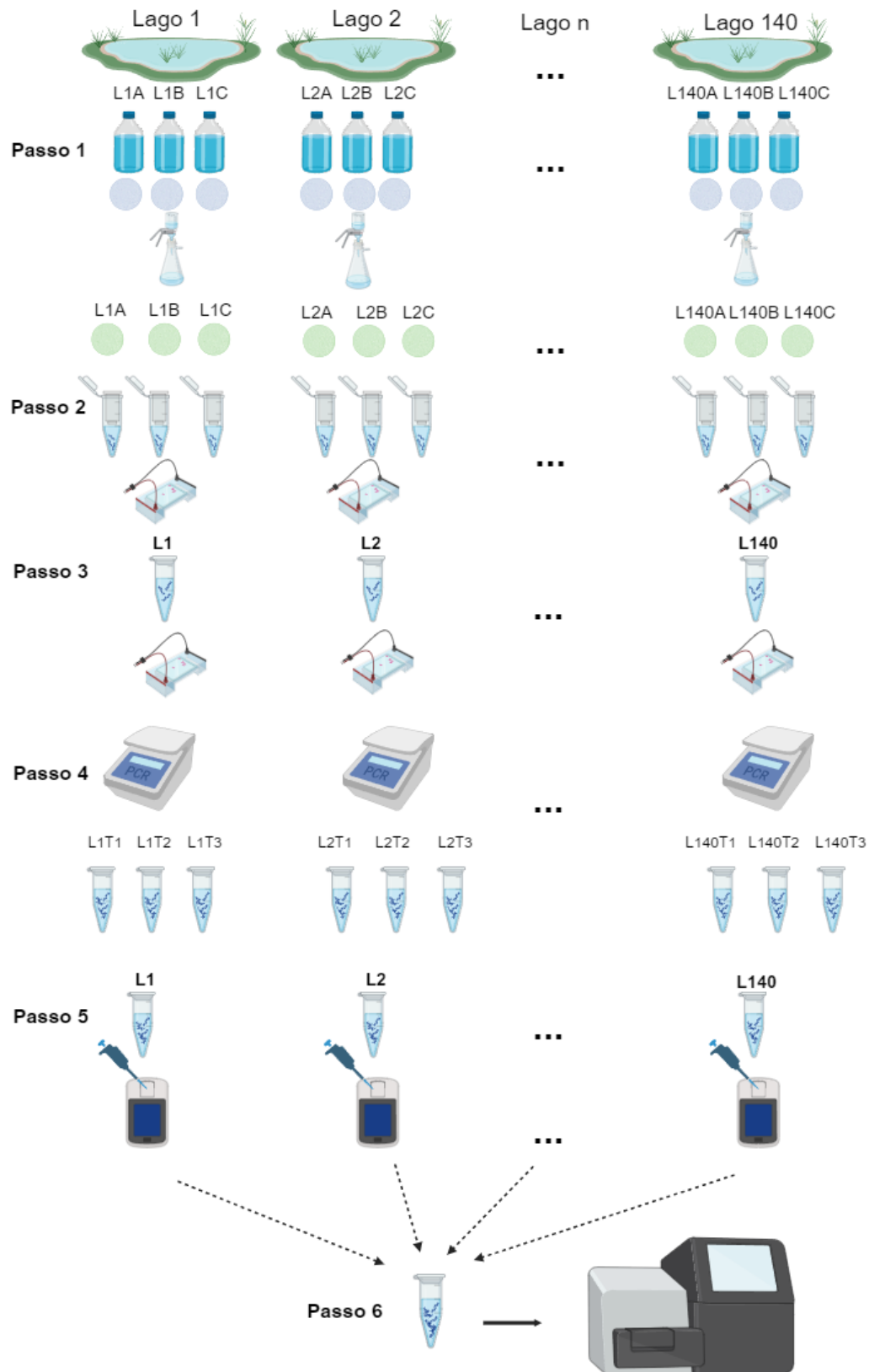


Figura 2 – Resumo esquemático das análises moleculares da coleta ao sequenciamento. Passo 1 – Coleta das amostras de água e filtragem; passo 2 – Extração do DNA e quantificação; passo 3 – Junção das amostras A, B e C e quantificação; passo 4 –Preparação e amplificação das triplicatas de PCR; passo 5 – Junção das triplicatas de PCR e quantificação; passo 6 – Formação do pool da biblioteca e sequenciamento.

Análises bioinformáticas e obtenção da filogenia

A qualidade das sequências foi avaliada utilizando o software FastQC (Andrews, 2010) e o tratamento das sequências foi realizado através do pipeline Pimba (*Pipeline for MetaBarcoding Analysis*) (Oliveira *et al.*, 2021). O pipeline é dividido em duas etapas principais: a primeira (pimba_prepare.sh) preparou os dados para as análises subsequentes, com a remoção de primers e adaptadores, tratamento de qualidade e mesclagem das sequências. Nesta etapa, também foram estabelecidos os parâmetros de qualidade para as sequências brutas (tamanho mínimo >100 pb e Phred médio > 20) e as sequências que não atenderam aos critérios foram descartadas; na segunda etapa (pimba_run.sh) foi realizada a desreplicação, classificação de abundância, descarte de sequências únicas, truncamento das sequências, remoção de sequências com 'N' > 0, definição das OTUs (Unidades Taxonômicas Operacionais) utilizando 97% de similaridade com o método UPARSE, remoção de quimeras e, por fim, a atribuição taxonômica. Para a etapa do truncamento estabelecemos um tamanho final de 390 pb para as sequências, baseado no tamanho médio dos fragmentos obtidos com truncamento 0 e no tamanho médio das sequências após a mesclagem. Para os demais parâmetros utilizamos o padrão da linha comando. A atribuição taxonômica para as sequências 18S foi feita utilizando o banco de dados SILVA (Quast *et al.*, 2013) usando uma identidade percentual de 97% e uma cobertura mínima de 90% .

As OTUs presentes no controle negativo foram removidas do conjunto total de OTUs obtidas, para garantir a confiabilidade dos dados. As OTUs restantes com atribuição taxonômica foram classificadas em grupos tróficos (heterotróficos, autotróficos e mixotróficos) com base na literatura especializada. Considerando a ampla gama de eucariotos capturados pelo marcador genético 18S-V4, foi construída uma filogenia inicial utilizando todas as OTUs obtidas, tanto as classificadas taxonomicamente quanto as sem classificação (árvore total). Posteriormente, os clados que correspondiam a microeucariotos planctônicos fotossintetizantes (MEPF) foram selecionados e foi gerada uma nova filogenia composta exclusivamente por grupos pertencentes a este grupo focal (árvore focal) (Fig. 3).

Para a construção da árvore filogenética com todas as OTUs (árvore total) as sequências foram alinhadas usando o software MAFFT v7.503 (Kato e Standley, 2013). A árvore foi construída com o método de máxima verossimilhança usando o software IQ-TREE v1.6.12 (Minh *et al.*, 2020) seguindo os parâmetros padrões do software. O modelo evolutivo mais adequado foi selecionado automaticamente pelo software ModelFinder

(Kalyaanamoorthy *et al.*, 2017), o modelo selecionado foi o GTR+F+R10, com base nos critérios BIC (Critério de Informação Bayesiano) e AIC (Critério de Informação de Akaike). A confiabilidade das ramificações da árvore foi avaliada por meio de 1000 réplicas de *bootstrap* ultrarrápido (UFBoot) (Hoang *et al.*, 2018). A visualização da árvore foi realizada nos softwares iTOL (Letunic e Bork, 2024) e FigTree (Rambaut, 2009). Após a visualização da árvore fizemos a seleção dos clados com os grupos MEPF manualmente, e removemos as OTUs que não pertencia a este grupo do conjunto total de OTUs. Selecionamos as sequências não classificadas e classificadas de OTUs que foram agrupadas aos clados do grupo focal (MEPF) para construir uma nova filogenia. Para esta segunda árvore (árvore focal) as sequências foram alinhadas usando MAFFT (v7.503) e a árvore filogenética foi construída com o IQ-TREE (v1.6.12) (Minh *et al.*, 2020) utilizando os mesmos parâmetros da árvore total.

As sequências obtidas neste trabalho serão depositadas no banco de dados do *National Center for Biotechnology Information* (NCBI).

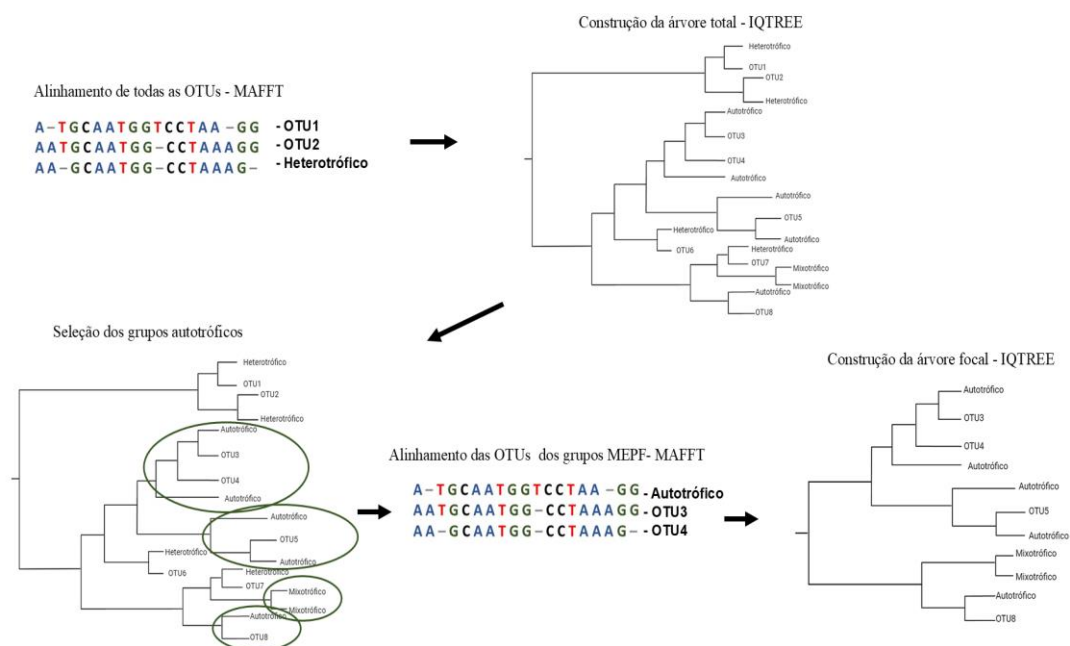


Figura 3 - Representação da construção das filogenias. Grupos MEPF = Grupos de Microeucariotos Planctônicos Fotossintetizantes. MAFFT, software de alinhamento de sequências e IQTREE software de construção de árvore filogenética.

Análises de riqueza e diversidade de OTUs

Todas as análises foram realizadas no software R (R Core Team, 2023). Para avaliar a abundância relativa dos grupos taxonômicos identificados e a abundância dos grupos tróficos em cada lago, utilizamos o pacote "phyloseq" (McMurdie e Holmes, 2013). Construímos um objeto "phyloseq" combinando a tabela de abundância de OTUs, a tabela de taxonomia e os metadados em um único objeto. Em seguida, agrupamos os dados no nível de grupos taxonômicos e grupos tróficos e transformamos as contagens absolutas em cada amostra em abundâncias relativas. A riqueza de OTUs foi avaliada a partir da tabela de abundância de OTUs com a função "diversity" do pacote "vegan" (Oksanen *et al.*, 2022). As árvores filogenéticas foram editadas e preparadas para as análises posteriores com o pacote "ape" (Paradis e Schliep, 2019) e "phytools" (Revell, 2024).

Análises de diversidade e estrutura filogenéticas (diversidade alfa filogenética)

A diversidade filogenética para ambas as árvores (árvore total e focal) foi avaliada medindo os comprimentos totais dos ramos da árvore que conectam todas as espécies presente na comunidade, para isso usamos o índice PD (Diversidade Filogenética de Faith; Faith, 1992). Para calcular o índice PD usamos a função "pd", e para testar a significância dos valores de PD usamos um modelo nulo gerado com 999 permutações usando a função "ses.pd" do pacote "picante" (Kembel *et al.*, 2010).

Para medir a estrutura filogenética da comunidade total e focal (árvore total e focal) usamos o índice NRI (Índice de Parentesco Líquido). Esse índice mostra se os táxons observados em uma comunidade estão próximos ou distantes filogeneticamente uns dos outros, o que é um padrão de estruturação filogenética, e permite ver o quanto da diversidade filogenética encontrada pode ser observada ao acaso (Han *et al.*, 2021). Valores de NRI maiores que zero indicam agrupamento filogenético e valores menores que zero indicam superdispersão filogenética. Como o NRI é uma medida padronizada da MPD (Distância Média entre Pares) (Webb *et al.*, 2002), para obter o índice NRI calculamos um modelo nulo para MPD gerado com 999 permutações com a função "ses.mpd" do pacote "picante" (Kembel *et al.*, 2010). O MPD mede a distância filogenética média entre os pares de espécies presentes na comunidade, e o NRI é dado pela diferença entre a MPD observada e a MPD esperada pelo modelo nulo, dividido pelo desvio padrão dos valores esperados (Diniz-Filho *et al.*, 2024).

A correlação entre a riqueza de OTUs e o índice PD, e a correlação de PD padronizado com o índice NRI foram verificadas com a função “cor” do pacote “stats” usando o método de Pearson (R Core Team, 2023). Para todas as análises filogenéticas, foram usados dados de incidência de OTUs (presença-ausência), transformados com a função “decostand” do pacote “vegan” (Oksanen *et al.*, 2022) e ordenados de acordo com a filogenia com a função “match.phylo.comm” do pacote “picante” (Kembel *et al.*, 2010).

Análises ambientais e espaciais para a diversidade alfa filogenética

Para avaliar a influência das variáveis ambientais e espaciais na diversidade e estrutura filogenética das comunidades, foi realizada uma análise de redundância parcial (pRDA) utilizando o pacote "vegan" (Oksanen *et al.*, 2022). Construimos as variáveis espaciais a partir das coordenadas geográficas (latitude e longitude) de cada local de coleta, utilizando dois tipos de filtros: PCNMs (Coordenadas Principais de Matrizes Vizinhas) e AEMs (Mapas de Autovetores Assimétricos). Diferentes filtros estabelecem diferentes hipóteses sobre a relação espacial entre os lagos analisados e isso é importante para separar melhor os processos ambientais e espaciais na organização da comunidade (Heino *et al.*, 2015). As PCNMs consideram as distâncias lineares e cria preditores espaciais que podem ser usados em regressão multivariada ou análise canônica. Essa técnica detecta e quantifica padrões espaciais em diferentes escalas (Borcard e Legendre, 2002). As PCNMs foram criadas com a função "pcnm" do pacote "vegan" e selecionadas com a função "ordistep" do pacote “vegan” com o modo de busca “forward” (Oksanen *et al.*, 2022) com 999 permutações, considerando a sua significância para o modelo ($p < 0,05$) (Blanchet, Legendre e Borcard, 2008a). As AEMs calculam a conectividade direcional entre os pontos de coleta, criando variáveis espaciais correspondentes a um processo direcional assimétrico que influencia a distribuição de espécies, como o fluxo de um rio (Blanchet, Legendre e Borcard, 2008b). As variáveis AEMs foram construídas com a função "aem" do pacote "AEM" (Blanchet, Legendre e Gauthier, 2015) a partir de uma matriz de conectividade direcional entre os pontos, construída de acordo com o fluxo do rio (de montante a jusante). As variáveis foram selecionadas com a função "ordistep", da mesma forma que as PCNMs.

As variáveis ambientais foram selecionadas calculando o fator de inflação da variância (VIF) para verificar a multicolinearidade entre as variáveis independentes, variáveis com $VIF > 5$ foram excluídas (Akinwande, Dikko e Samson, 2015). As variáveis

ambientais que não tinham relevância para organismos aquáticos, como temperatura e umidade do ar, foram removidas. Todas as variáveis foram padronizadas, utilizando a função "decostand" do pacote "vegan" com o método "standardize" (Oksanen *et al.*, 2022).

A partição da variância foi avaliada com a função "varpart" utilizando as variáveis ambientais e variáveis espaciais selecionadas (PCNMs e AEMs), e as matrizes com os valores de PD e NRI para cada ponto, estimados para o grupo total e focal. A significância de cada fração da variância foi verificada por testes de permutação usando a função "anova.cca" do pacote "vegan" (Oksanen *et al.*, 2022) com 999 permutações.

Para avaliar a heterogeneidade ambiental entre os lagos realizamos uma análise de componentes principais (PCA) usando a função "princomp" do pacote "stats" (R Core Team, 2023).

Diversidade beta filogenética

Para calcular a diversidade beta filogenética (diversidade filobeta) utilizamos a função "phylosor" do pacote "picante" (Kembel *et al.*, 2010). O "phylosor" mede a similaridade filogenética entre comunidades, que é definida pela fração do comprimento total dos ramos que é compartilhado entre duas comunidades (Bryant *et al.* 2008). O índice PhyloSor varia de próximo de 0 (duas comunidades compartilham apenas uma raiz muito pequena) a 1 (ambas as comunidades são compostas dos mesmos táxons) (Bryant *et al.* 2008). Calculamos a diversidade beta filogenética apenas para a árvore focal, pois a função "phylosor" exige que a árvore esteja enraizada, e a árvore total não formou raiz. Para obter os valores da partição dos componentes da diversidade beta filogenética, utilizamos a função "phylo.beta.multi" do pacote "betapart" usando o índice de Sorensen (Baselga *et al.*, 2023).

Utilizando a matriz de similaridade filobeta gerada pelo phylosor, uma matriz de distância ambiental e espacial euclidiana, avaliamos se a diversidade filobeta estava correlacionada com a distância espacial ou ambiental, a partir de um teste de mantel parcial com 999 permutações, usando a função "mantel.partial" do pacote "vegan" (Oksanen *et al.*, 2022). Para a construção da matriz de distância ambiental utilizamos apenas as variáveis ambientais que foram selecionados para o modelo da RDA relacionadas os índices de diversidade alfa filogenética da árvore focal.

RESULTADOS

Análises de bioinformáticas e classificação das OTUs

A concentração de DNA extraído variou entre 1,67 e 66,67 ng/ μ l em 110 das 140 amostras dos lagos analisados. Em 30 lagos, a concentração de DNA foi inferior ao limite de detecção visual, impossibilitando a quantificação (Fig. 4), mesmo assim todas as 140 amostras seguiram para etapa de amplificação via PCR, destas, 120 apresentaram concentração de produto amplificado suficiente para o preparo de biblioteca e sequenciamento (Tabela 1).

As sequências de produtos amplificados apresentaram uma boa qualidade, com um score Phred médio superior a 20 em praticamente toda a extensão das leituras. O comprimento médio dos fragmentos ficou entre 290 a 300 pb, indicando um bom desempenho do processo de amplificação e sequenciamento. A qualidade das leituras R1 (*forward*) e R2 (*reverse*) começou a declinar após aproximadamente 260 pb e 210 pb, respectivamente, o que é comum em sequenciamentos de pares de extremidades na plataforma Illumina MiSeq (resultados do FastQC, em Apêndice 4). O sequenciamento gerou ao todo 19.804.212 sequências brutas da região V4 do gene 18S, das quais 18.138.662 (91,59%) passaram pela filtragem de qualidade, 8.013.747 de sequências foram mescladas, gerando 8.277 OTUs. Destas, 338 (4,1%) foram removidas do conjunto de dados, por estarem presente nos controles negativos e serem consideradas contaminantes. Restando 7.939 OTUs, das quais, 1.676 (21,1%) foi possível fazer alguma atribuição taxonômica.

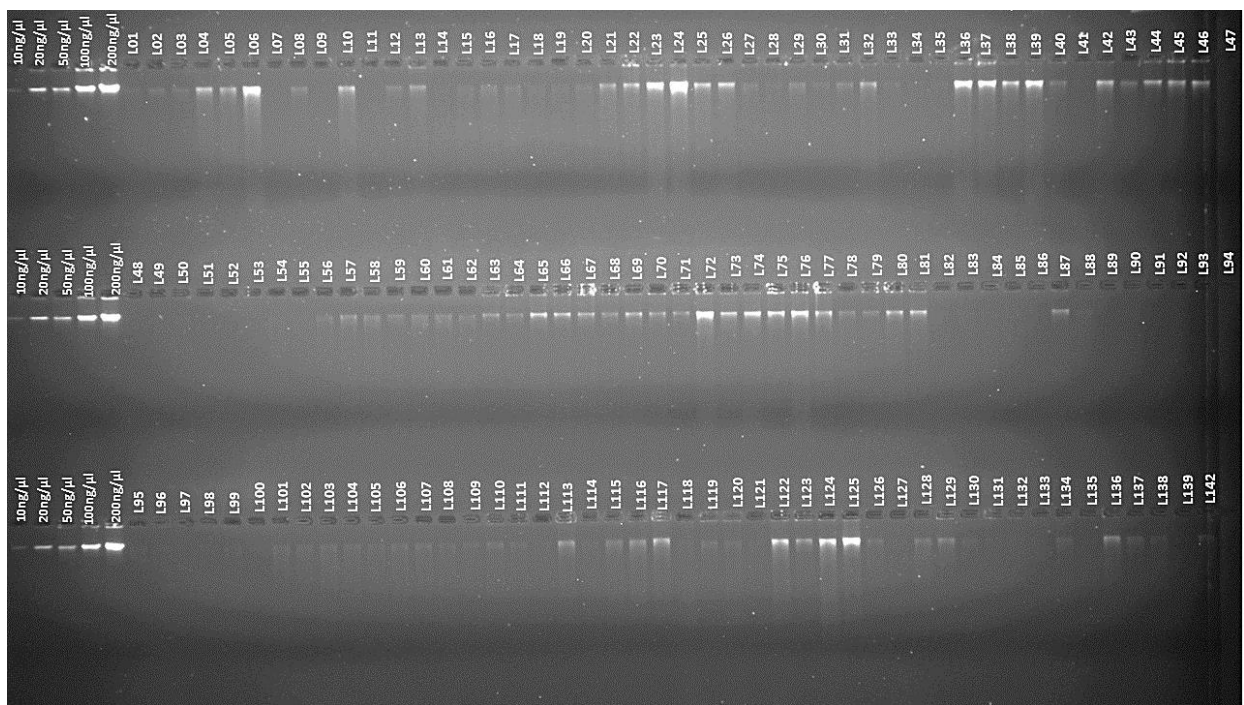


Figura 4 - Gel de agarose 1% corado com brometo de etídeo com as 140 amostras de DNA extraídas dos 140 lagos analisados. No início de cada pente temos os marcadores moleculares de 10, 20, 50, 100 e 200 ng/μl.

Tabela 1- Lista de pontos com amostras de DNA amplificadas por PCR e sequenciadas na plataforma Illumina MiSeq, com o número total de OTUs obtidas por ponto e indicação dos respectivos rios de origem.

| N° | Ponto | N° OTUs | Rio | N° | Ponto | N° OTUs | Rio | N° | Ponto | N° OTUs | Rio |
|----|---------|---------|------------|----|---------|---------|------------|-----|----------|---------|------------|
| 1 | Lago 1 | 1904 | Agua Limpa | 41 | Lago 48 | 1577 | Araguaia | 81 | Lago 94 | 1520 | das Mortes |
| 2 | Lago 2 | 887 | Agua Limpa | 42 | Lago 49 | 944 | Araguaia | 82 | Lago 95 | 1856 | das Mortes |
| 3 | Lago 3 | 730 | Agua Limpa | 43 | Lago 50 | 1541 | Araguaia | 83 | Lago 96 | 1579 | das Mortes |
| 4 | Lago 4 | 1054 | Agua Limpa | 44 | Lago 51 | 1161 | Araguaia | 84 | Lago 97 | 1338 | das Mortes |
| 5 | Lago 5 | 1075 | Vermelho | 45 | Lago 54 | 1101 | Araguaia | 85 | Lago 98 | 1590 | das Mortes |
| 6 | Lago 6 | 1311 | Vermelho | 46 | Lago 55 | 801 | Araguaia | 86 | Lago 99 | 1720 | das Mortes |
| 7 | Lago 8 | 754 | Agua Limpa | 47 | Lago 56 | 943 | Araguaia | 87 | Lago 100 | 1920 | das Mortes |
| 8 | Lago 10 | 1136 | Vermelho | 48 | Lago 57 | 1606 | Cristalino | 88 | Lago 104 | 1472 | das Mortes |
| 9 | Lago 11 | 878 | Vermelho | 49 | Lago 58 | 996 | Cristalino | 89 | Lago 105 | 1650 | das Mortes |
| 10 | Lago 12 | 931 | Vermelho | 50 | Lago 59 | 1211 | Cristalino | 90 | Lago 107 | 1501 | das Mortes |
| 11 | Lago 13 | 928 | Vermelho | 51 | Lago 60 | 951 | Cristalino | 91 | Lago 109 | 2036 | das Mortes |
| 12 | Lago 16 | 1100 | Vermelho | 52 | Lago 61 | 988 | Cristalino | 92 | Lago 110 | 1639 | das Mortes |
| 13 | Lago 17 | 736 | Vermelho | 53 | Lago 62 | 1415 | Cristalino | 93 | Lago 112 | 1525 | das Mortes |
| 14 | Lago 18 | 1027 | Vermelho | 54 | Lago 63 | 1664 | Cristalino | 94 | Lago 113 | 1911 | das Mortes |
| 15 | Lago 19 | 1179 | Vermelho | 55 | Lago 64 | 2070 | Cristalino | 95 | Lago 114 | 1505 | das Mortes |
| 16 | Lago 21 | 990 | Araguaia | 56 | Lago 65 | 1867 | Cristalino | 96 | Lago 115 | 1546 | das Mortes |
| 17 | Lago 22 | 1037 | Araguaia | 57 | Lago 66 | 1923 | Cristalino | 97 | Lago 116 | 1070 | das Mortes |
| 18 | Lago 23 | 1247 | Araguaia | 58 | Lago 67 | 1603 | Cristalino | 98 | Lago 117 | 1837 | das Mortes |
| 19 | Lago 24 | 1178 | Araguaia | 59 | Lago 68 | 1677 | Cristalino | 99 | Lago 118 | 1340 | das Mortes |
| 20 | Lago 25 | 1057 | do Peixe | 60 | Lago 69 | 1350 | Cristalino | 100 | Lago 119 | 1473 | das Mortes |
| 21 | Lago 26 | 504 | do Peixe | 61 | Lago 70 | 1441 | Cristalino | 101 | Lago 120 | 1439 | das Mortes |
| 22 | Lago 27 | 1385 | do Peixe | 62 | Lago 71 | 1790 | Cristalino | 102 | Lago 121 | 1423 | das Mortes |
| 23 | Lago 28 | 1887 | do Peixe | 63 | Lago 72 | 1824 | Cristalino | 103 | Lago 122 | 1233 | das Mortes |
| 24 | Lago 29 | 812 | do Peixe | 64 | Lago 73 | 1879 | Cristalino | 104 | Lago 123 | 851 | das Mortes |
| 25 | Lago 30 | 1201 | do Peixe | 65 | Lago 74 | 1952 | Cristalino | 105 | Lago 124 | 1197 | das Mortes |
| 26 | Lago 31 | 708 | do Peixe | 66 | Lago 75 | 1802 | Cristalino | 106 | Lago 125 | 751 | das Mortes |
| 27 | Lago 32 | 1153 | do Peixe | 67 | Lago 76 | 1417 | Cristalino | 107 | Lago 126 | 1042 | das Mortes |
| 28 | Lago 33 | 512 | do Peixe | 68 | Lago 77 | 1514 | Cristalino | 108 | Lago 128 | 1048 | das Mortes |
| 29 | Lago 34 | 1339 | do Peixe | 69 | Lago 78 | 1491 | Cristalino | 109 | Lago 129 | 655 | das Mortes |
| 30 | Lago 35 | 1020 | do Peixe | 70 | Lago 79 | 1926 | Cristalino | 110 | Lago 130 | 1120 | Araguaia |
| 31 | Lago 36 | 1036 | Araguaia | 71 | Lago 80 | 2181 | Cristalino | 111 | Lago 131 | 1062 | Araguaia |
| 32 | Lago 37 | 1157 | Araguaia | 72 | Lago 81 | 1721 | Araguaia | 112 | Lago 132 | 1234 | Araguaia |
| 33 | Lago 38 | 834 | Araguaia | 73 | Lago 85 | 1311 | Araguaia | 113 | Lago 133 | 1159 | Araguaia |
| 34 | Lago 39 | 941 | Araguaia | 74 | Lago 86 | 1686 | Araguaia | 114 | Lago 134 | 1265 | Araguaia |
| 35 | Lago 40 | 1440 | Araguaia | 75 | Lago 87 | 854 | Araguaia | 115 | Lago 135 | 1262 | Araguaia |
| 36 | Lago 42 | 1106 | Araguaia | 76 | Lago 88 | 1134 | Araguaia | 116 | Lago 136 | 1014 | Araguaia |
| 37 | Lago 43 | 1128 | Araguaia | 77 | Lago 89 | 1335 | Araguaia | 117 | Lago 137 | 1486 | Araguaia |
| 38 | Lago 44 | 915 | Crixas | 78 | Lago 91 | 1266 | Araguaia | 118 | Lago 138 | 1086 | Araguaia |
| 39 | Lago 45 | 285 | Crixas | 79 | Lago 92 | 1139 | das Mortes | 119 | Lago 139 | 1532 | Araguaia |
| 40 | Lago 46 | 1372 | Crixas | 80 | Lago 93 | 1282 | das Mortes | 120 | Lago 142 | 1561 | Araguaia |

A análise da composição taxonômica, revelou uma diversidade de grupos eucarióticos (Fig. 5), dominada pelo supergrupo SAR (Stramenopiles, Alveolata e Rhizaria), representando 42% do conjunto de dados classificados taxonomicamente, seguido por Archaeplastida (30%) e Opisthokonta (22%) e em menor proporção (6%) pelos grupos Cryptophyta, Amoebozoa e Heliozoa (Fig. 6). Entre os microeucariotos planctônicos fotossintetizantes, destacaram-se Chlorophyta como grupo mais abundante, seguido de Ochrophyta, Dinoflagellata, Cryptophyta e Charophyta (Fig. 7). Em relação ao modo de nutrição entre os grupos classificados (Tabela 2), os organismos heterotróficos

representaram 48,8% do conjunto de dados, seguidos por autotróficos (40,1%) e, em menor proporção, por mixotróficos (11,1%). No entanto, os grupos autotróficos foram numericamente mais abundantes na comunidade (Fig. 8). Por motivos não identificados, a tabela de abundância de OTUs gerada com o pipeline PIMBA possuía 20 OTUs a menos que o arquivo fasta final gerado na mesma análise. Portanto a riqueza total de OTUs para a comunidade total foi de 7.919, variando de 285 a 2.181 entre os lagos analisados (Fig. 9, A / Tabela 1). Para a comunidade focal, dos microeucariotos planctônicos fotossintetizantes, a riqueza total de OTUs foi de 2.497, variando de 114 a 880 entre os lagos analisados (Fig. 9, B).

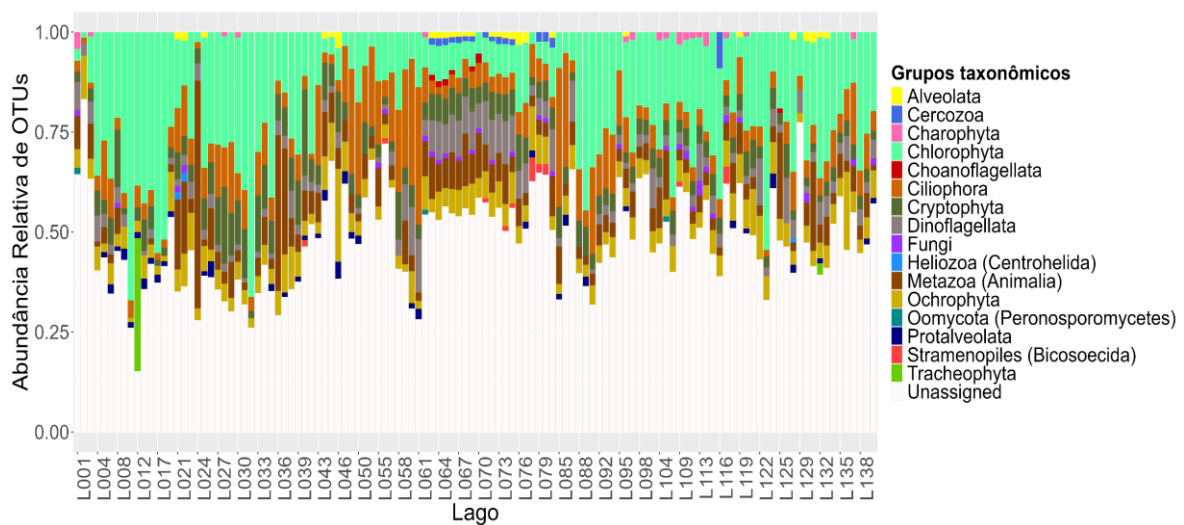


Figura 5 – Abundância relativa (>1%) dos grupos taxonômicos identificados com o gene 18S do conjunto de dados total (todas as OTUs obtidas) por lago analisado na planície de inundação do rio Araguaia. Unassigned = OTUs sem classificação taxonômica.

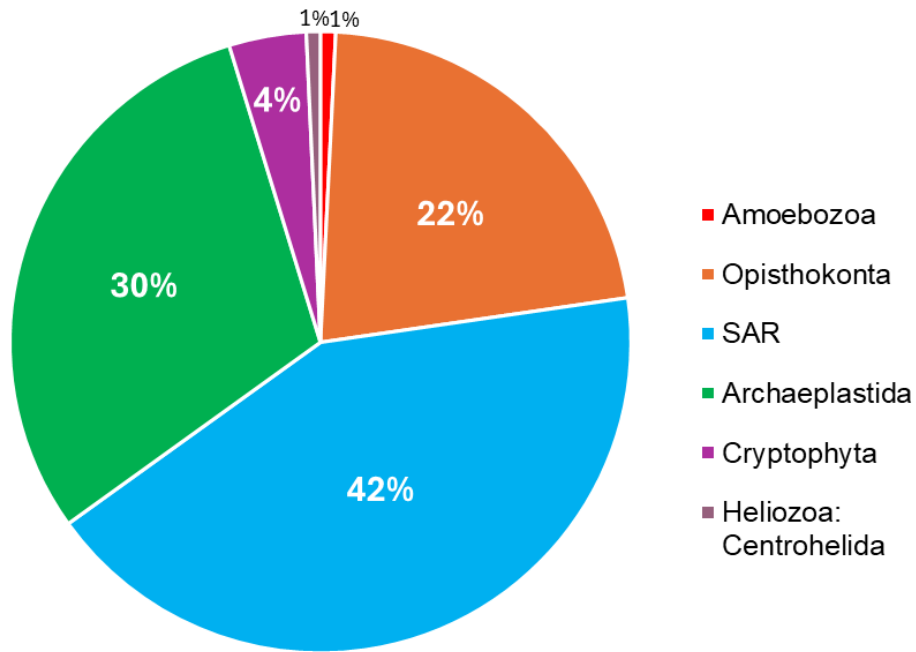


Figura 6 - Principais grupos taxonômicos identificados nos lagos analisados na planície de inundação do rio Araguaia com o gene 18S classificados no banco de dados Silva.

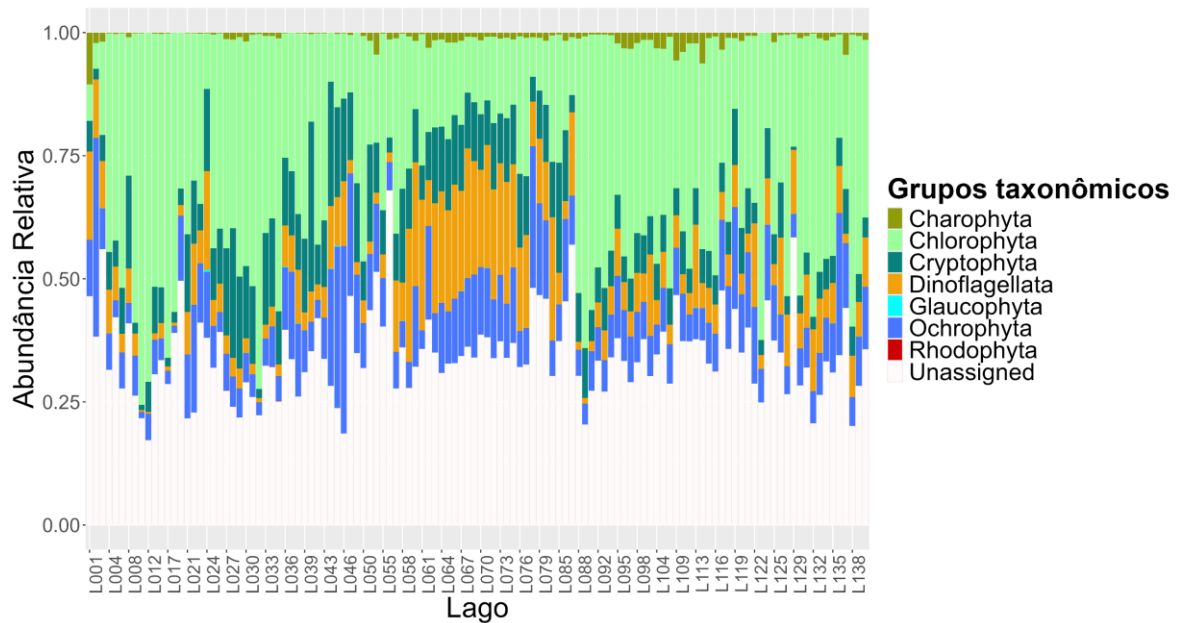


Figura 7 - Abundância relativa (>1%) dos grupos taxonômicos identificados com o gene 18S para o conjunto de dados focal (microeucariotos planctônicos fotossintetizantes) por lago analisado na planície de inundação do rio Araguaia.

Tabela 2 - Grupos taxonômicos identificados a partir das sequências de OTUs 18S identificadas com o banco de dados Silva e seus respectivos grupos tróficos de acordo com a literatura.

| Grupo taxonômico | Grupo trófico | Referências |
|--|----------------------|--|
| Chlorophyta | Autotrófico | Cormaci, Furnari e Alongi, 2014 |
| Charophyta | Autotrófico | García, 1994 |
| Stramenopiles: Ochrophyta | Autotrófico | Brown e Sorhannus, 2010 |
| Stramenopiles: Bacillariophyta | Autotrófico | Solak <i>et al.</i> , 2012 |
| Glaucochyta | Autotrófico | Price <i>et al.</i> , 2017 |
| Rhodophyta | Autotrófico | Gabrielson, Garbary e Hommersand, 1986 |
| Alveolata: Dinoflagellata | Mixotrófico | Stoecker, 1999 |
| Cryptophyta | Mixotrófico | Simon <i>et al.</i> , 2015 |
| Amoebozoa | Heterotrófico | Schilde, 2019 |
| Opisthokonta: Fungi | Heterotrófico | Wraight, Inglis e Goettel, 2007 |
| Opisthokonta: Metazoa | Heterotrófico | Ramulu, Raoult e Pontarotti, 2012 |
| Heliozoa: Centrohelida | Heterotrófico | Burki <i>et al.</i> , 2016 |
| Rhizaria: Cercozoa | Heterotrófico | Cavalier-Smith e Chao, 2003 |
| Alveolata: Ciliophora | Heterotrófico | Lynn, 2012 |
| Alveolata: Protalveolata: Colpodellida | Heterotrófico | Mylnikov, 2009 |
| Alveolata: Protalveolata: Perkinsidae | Heterotrófico | Jeon e Park, 2021 |
| Alveolata: Protalveolata: Syndiniales | Heterotrófico | Clarke <i>et al.</i> , 2019 |
| Stramenopiles: Peronosporomycetes | Heterotrófico | Islam e Hossain, 2013 |
| Opisthokonta: Ichthyosporea | Heterotrófico | Glockling, Marshall e Gleason, 2013 |
| Stramenopiles: Hyphochytriomycetes | Heterotrófico | Van der Auwera <i>et al.</i> , 1995 |
| Alveolata: Colponema | Heterotrófico | Mylnikov e Tikhonenkov, 2007 |
| Opisthokonta: Choanoflagellida | Heterotrófico | Carr <i>et al.</i> , 2008 |
| Sulcozoa: Collodictyonidae | Heterotrófico | Zhao <i>et al.</i> , 2012 |
| Alveolata: Apicomplexa | Heterotrófico | Morrison, 2009 |

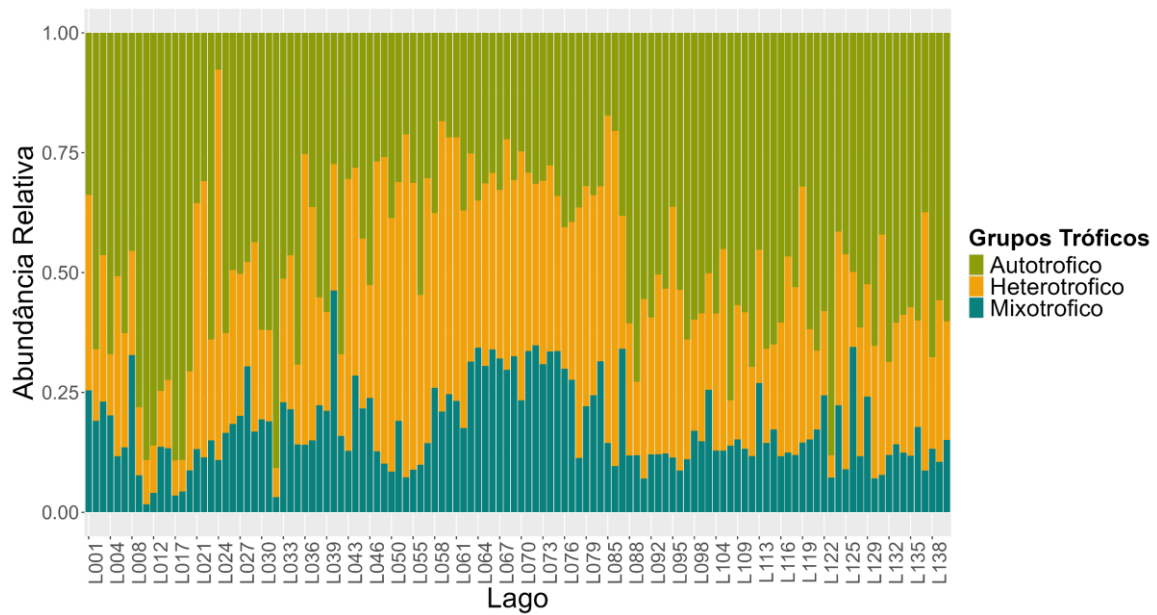


Figura 8 - Abundância relativa (> 0%) dos grupos tróficos de eucariotos obtidos a partir do gene 18S e classificados a partir da literatura por lago analisado na planície de inundação do rio Araguaia.

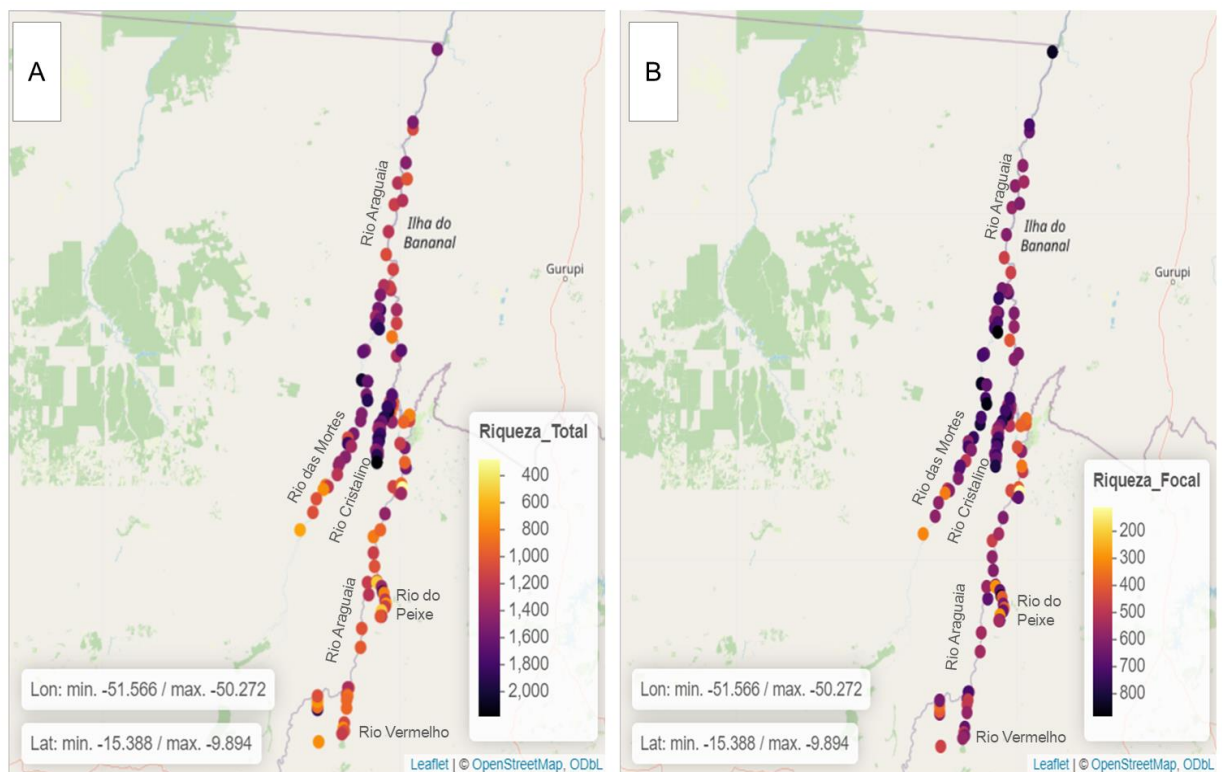


Figura 9 - Mapa de calor da riqueza de OTUs para a comunidade total (A), com todas as OTUs identificadas com o gene 18S, e para a comunidade focal (B), com as OTUs dos organismos de microeucariotos planctônicos fotossintetizantes (para ver a identificação dos lagos no mapa, ver mapa interativo no apêndice 6).

Árvores filogenéticas

A árvore filogenética total, abrangendo todas as OTUs do gene 18S, inicialmente possuía 7.939 pontas de ramo. Após a remoção de 20 OTUs não encontradas nos dados de abundância da comunidade, a árvore finalizou com 7.919 pontas de ramo (Fig. 10). A árvore filogenética focal, dos microeucariotos planctônicos fotossintetizantes, por sua vez, iniciou com 2.515 pontas de ramo. Após a exclusão de 18 OTUs não encontradas nos dados de abundância da comunidade, a árvore resultante apresentou 2.497 pontas de ramo (Fig. 11). A árvore focal apresentou uma maior quantidade de nós com alto valor de suporte, 74,1% dos nós apresentaram valores de *bootstrap* acima de 70% e 10,9% apresentaram valores de *bootstrap* menores que 50%. A árvore total apresentou 68,3% dos nós com valores de *bootstrap* acima de 70% e 14,3% dos nós com valores de *bootstrap* menores que 50%. A filogenia total também apresentou mais nós basais com baixos valores de suporte (< 20% de *bootstrap*) (ver Apêndice 5).

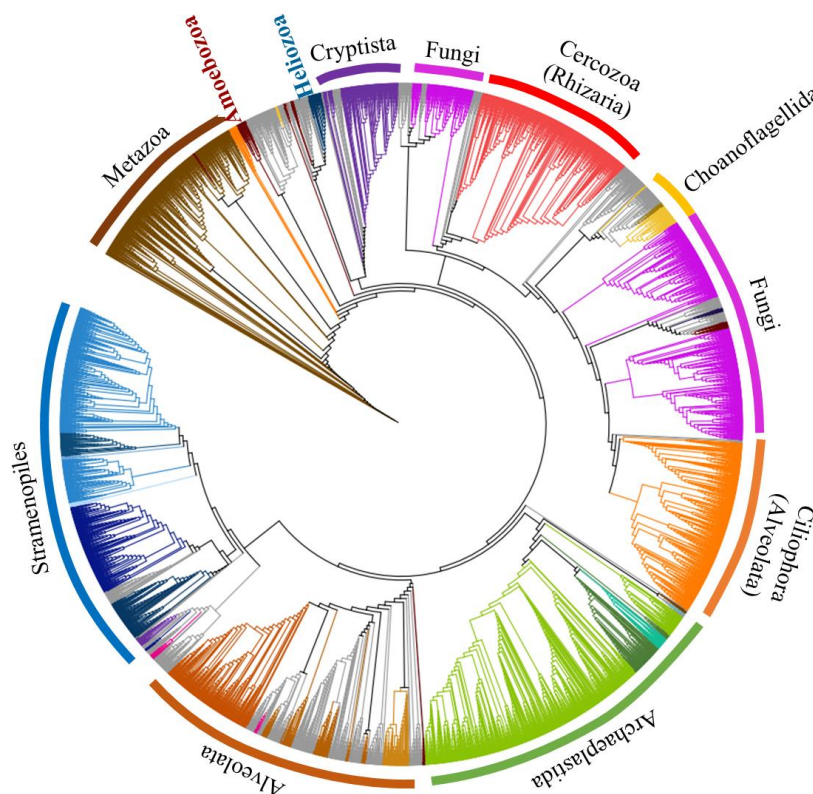


Figura 10 - Árvore filogenética total, com todo o conjunto de OTUs 18S, rodado no IQTREE com 1000 réplicas de *bootstrap* ultrarrápido. Clados não identificados estão em cinza.

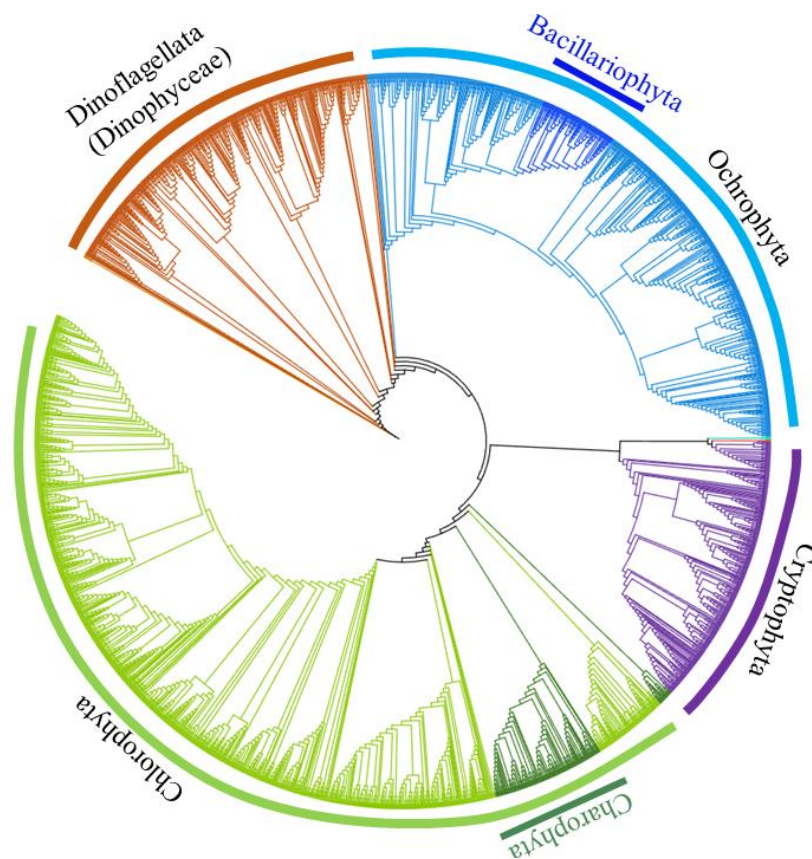


Figura 11 - Árvore focal, rodada com o conjunto de OTUs 18S selecionadas, pertencente a microeucariotos planctônicos fotossintetizantes, rodada no IQTREE com 1000 réplicas de *bootstrap* ultrarrápido. Glaucophyta (ciano) e Rhodophyta (vermelho) são grupos presentes na filogenia, mas que não foram nomeados na figura.

Diversidade filogenética (PD)

A diversidade filogenética de Faith (PD) da árvore total, variou de 111,9 (lago 45) a 472,3 (lago 80) ao longo da planície de inundação do rio Araguaia (Fig. 12, A). O lago com o menor PD (45), também foi o lago com a menor riqueza de OTUs (285) e o lago com maior PD (80) foi o lago com maior riqueza de OTUs (2.181).

A árvore focal, com apenas os organismos microeucariotos planctônicos fotossintetizantes, apresentou uma PD que variou de 14,9 (lago 45) a 63,9 (lago 80) ao longo da planície de inundação do rio Araguaia (Fig. 12, B). O lago com menor PD (45), também apresentou a menor riqueza de OTUs observada na comunidade (114), já o lago com maior PD (80) foi o nono lago com maior riqueza de OTUs (790), ficando atrás dos lagos 100 (880), 113 e 28 (858), 109 (840), 142 (836), 64 (833), 107 (806) e lago 74 (803).

A correlação entre PD e riqueza de OTUs foi alta em ambas as análises ($r_{total} = 0,97$, $r_{focal} = 0,94$) (Figs. 13 e 14). Os lagos dos rios Cristalino e das Mortes, em geral, apresentaram maiores valores de PD e riqueza de OTUs (Figs. 9 e 12).

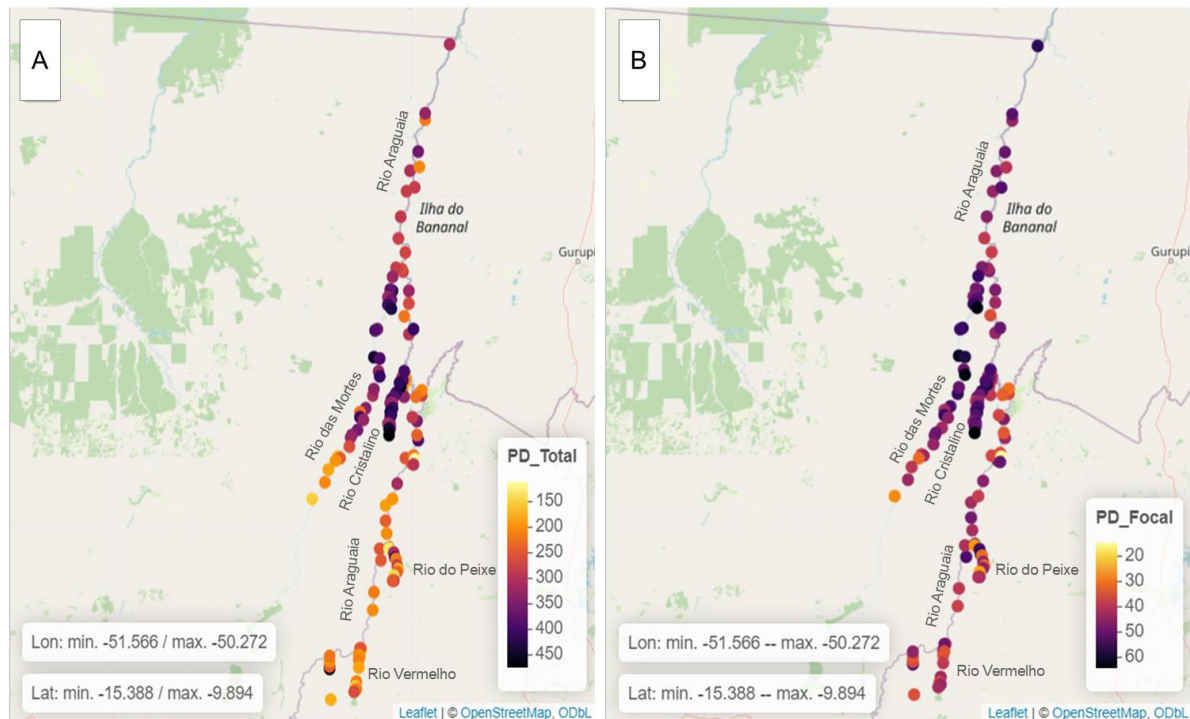


Figura 12 - Mapa de calor dos valores de PD (Diversidade Filogenética de Faith) por ponto para a árvore total (A), com todas as OTUs obtidas para o gene 18S e para a árvore focal (B), com as OTUs relacionadas aos grupos de microeucariotos planctônicos fotossintetizantes, ao longo do rio Araguaia e seus afluentes (para ver a numeração dos lagos no mapa, ver mapa interativo no apêndice 6).

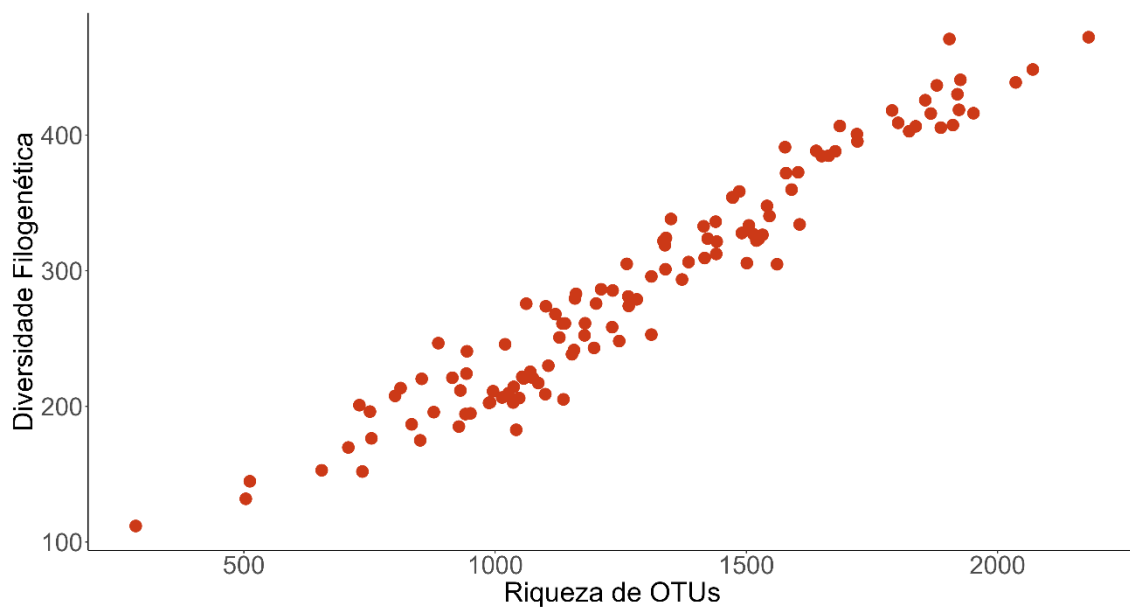


Figura 13 - Relação entre diversidade filogenética (PD) e riqueza de OTUs ($r = 0,97$) para comunidade total dos organismos identificados com o gene 18S na planície de inundação do rio Araguaia.

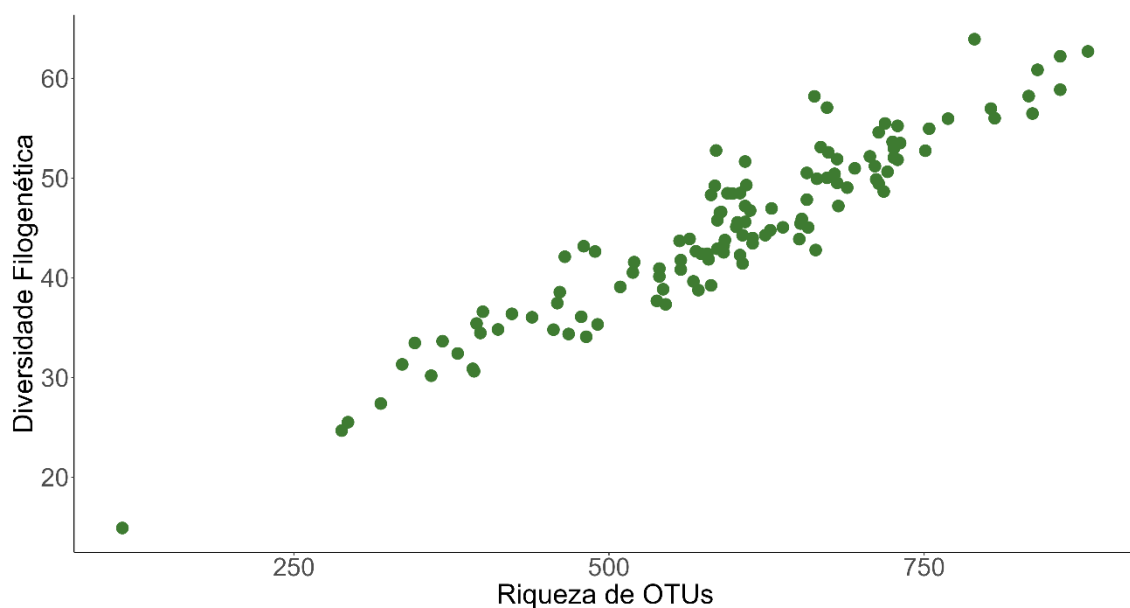


Figura 14 - Relação entre diversidade filogenética (PD) e riqueza de OTUs ($r = 0,94$) para comunidade focal de microeucariotos planctônicos fotossintetizantes identificados com o gene 18S na planície de inundação do rio Araguaia.

Todos os resultados de PD observado para ambas as árvores (total e focal) em cada lago foi significativamente diferente daquele esperado sob o modelo nulo, que é baseado na distribuição aleatória das espécies na árvore filogenética (Fig. 15 e 16). A diversidade filogenética de Faith observada para ambas as árvores foi menor do que a esperada sob o modelo nulo, com uma diferença maior para a árvore total (Fig. 15).

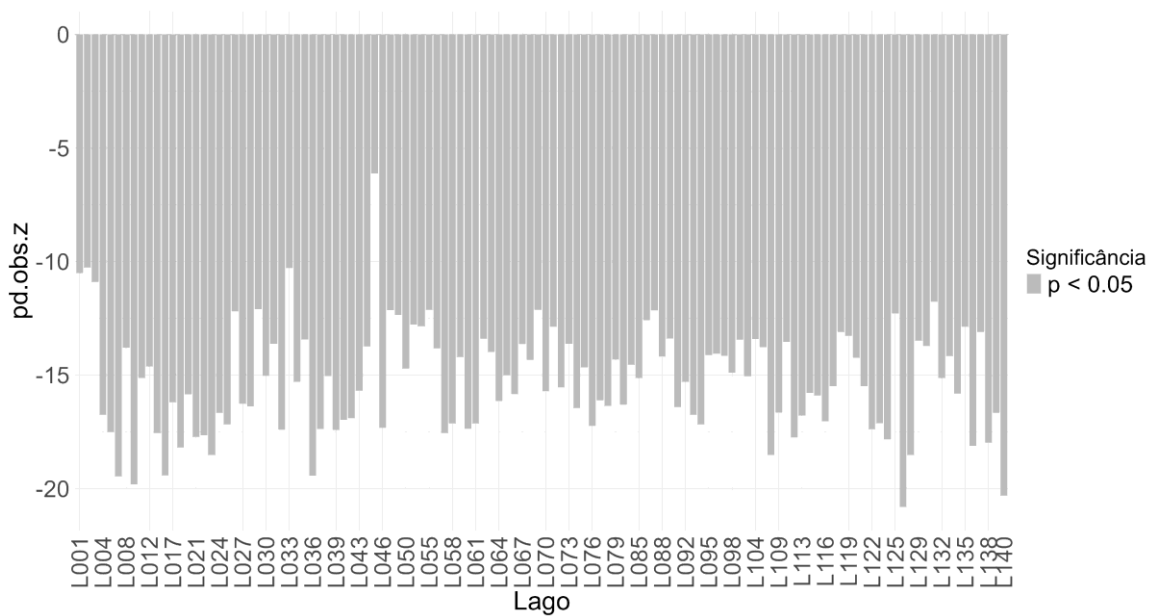


Figura 15 - Modelo nulo da diversidade filogenética (PD) da árvore total, com todas as OTUs obtidas para o gene 18S, gerado com 999 permutações usando a função “ses.pd”. Valores significativos estatisticamente ($p < 0,05$) estão em cinza.

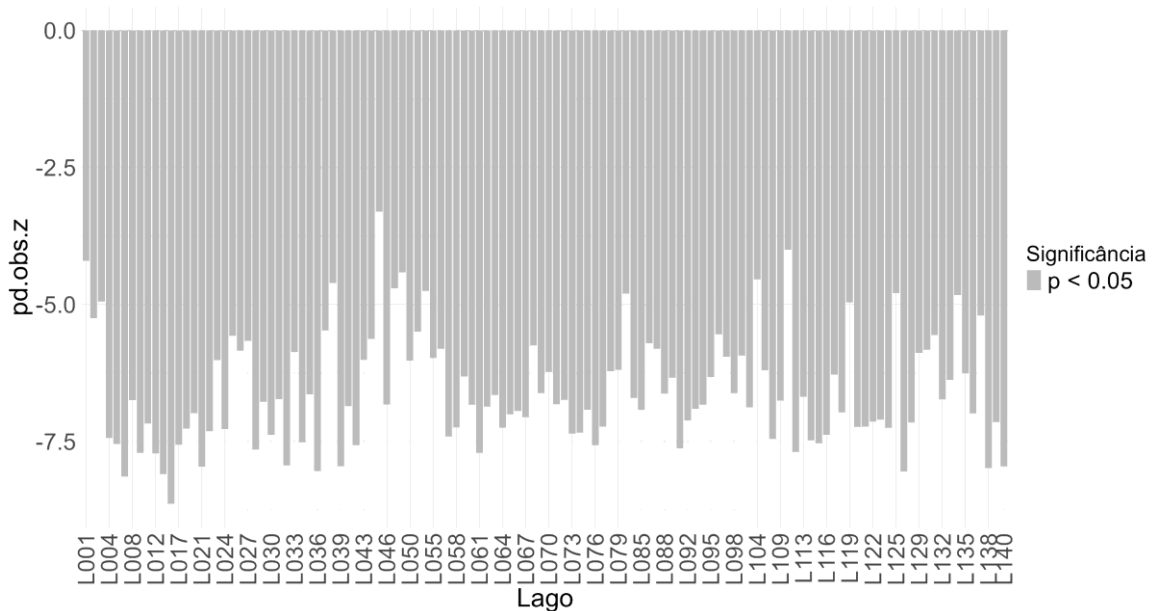


Figura 16 - Modelo nulo da diversidade filogenética (PD) da árvore focal, com as OTUs relacionadas aos grupos de microeucariotos planctônicos fotossintetizantes, gerado com 999 permutações usando a função “ses.pd”. Valores significativos estatisticamente ($p < 0,05$) estão em cinza.

Estrutura filogenética da comunidade

Os valores de NRI para ambas as árvores foram maiores que 0, indicando um agrupamento filogenético. A árvore total, apresentou valores de NRI variando de 0,95 (lago 45) a 18 (lago 17) (Fig. 17 e 19, A) e a árvore focal apresentou valores de NRI variando de 0,27 (lago 48) a 12,43 (lago 17) (Fig. 18 e 19, B). Lagos pertencentes ao rio Cristalino e próximos a ele apresentaram os menores valores de NRI tanto para a árvore total como para a árvore focal (Fig. 19). O NRI apresentou uma correlação negativa com o PD padronizado (pd.obs.z) tanto para a árvore total ($r = -0,71$) (Fig. 20) como para a árvore focal ($r = -0,74$) (fig. 21).

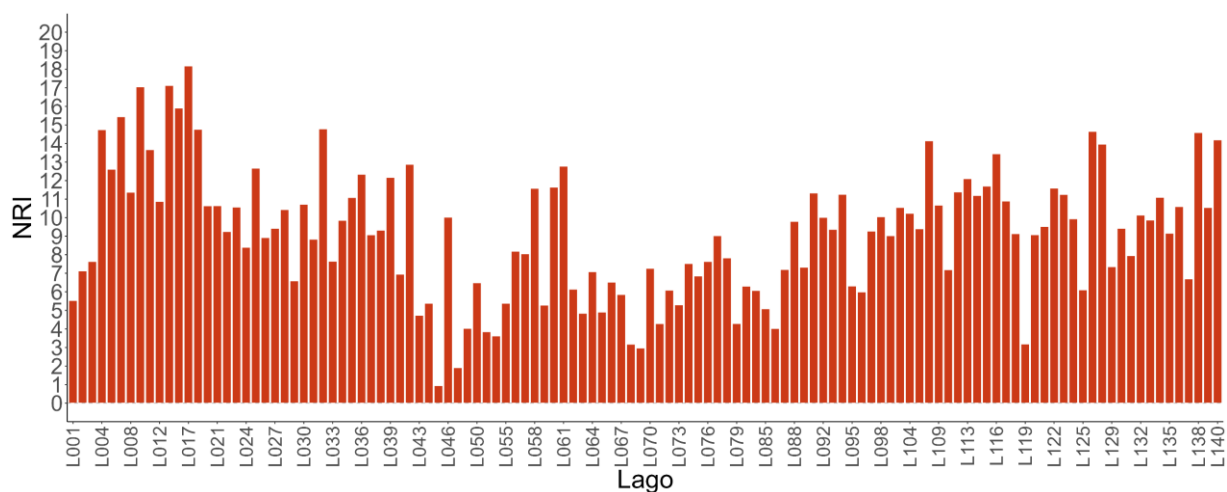


Figura 17 - Valores de NRI (Índice de Parentesco Líquido) para a árvore total, com todas as OTUs obtidas para o gene 18S, por lago analisado na planície de inundação do rio Araguaia.

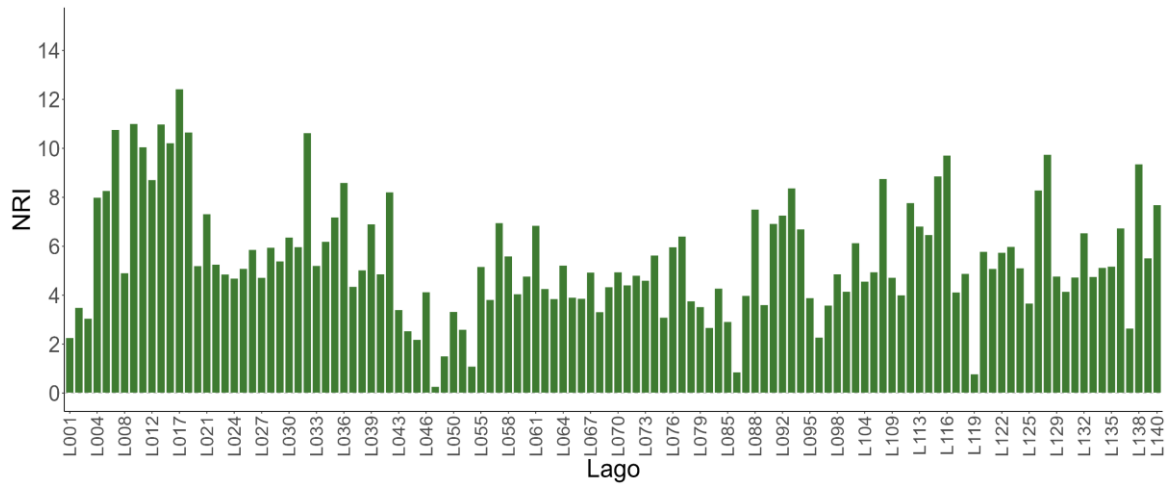


Figura 18 - Valores de NRI (Índice de Parentesco Líquido) para a árvore focal, com os microeucariotos plancônicos fotossintetizantes, por lago analisado na planície de inundação do rio Araguaia.

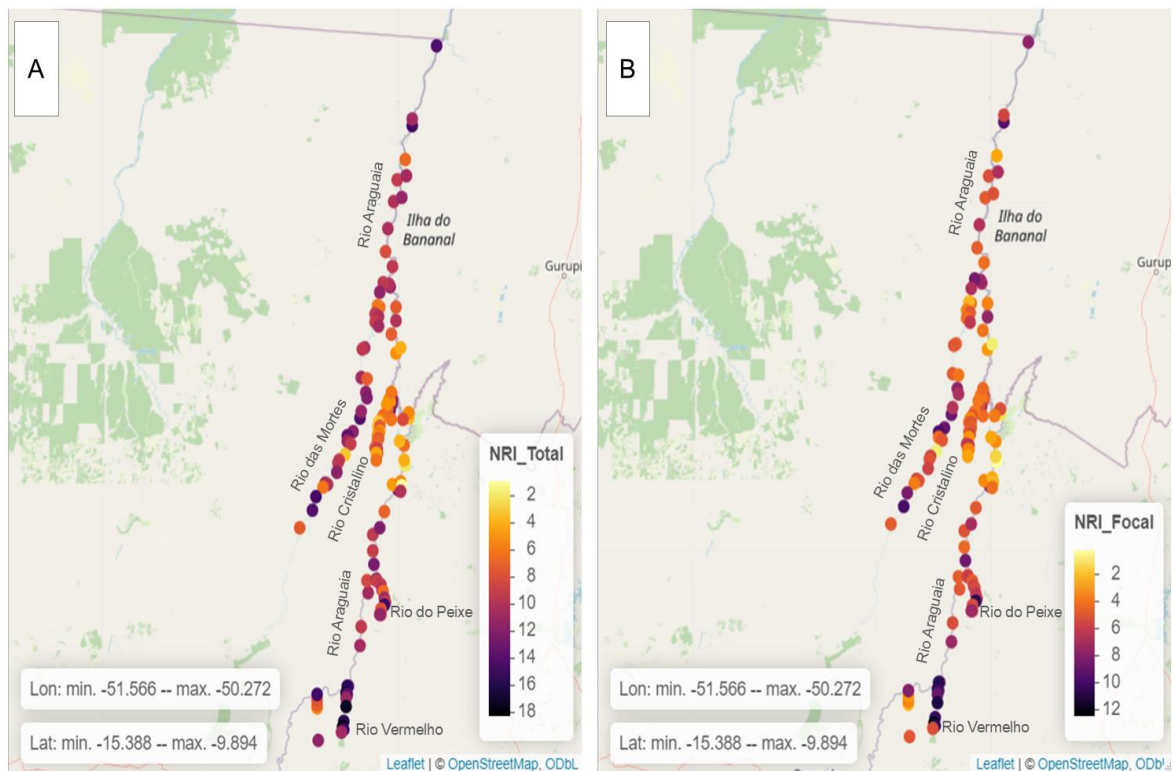


Figura 19 - Mapa de calor dos valores de NRI (Índice de Parentesco Líquido) da árvore total (A), com todas as OTUs obtidas para o gene 18S, e árvore focal (B), com as OTUs relacionadas aos grupos de microeucariotos plancônicos fotossintetizantes, ao longo da planície de inundação do rio Araguaia (para ver a identificação dos lagos no mapa, ver mapas interativos no apêndice 6).

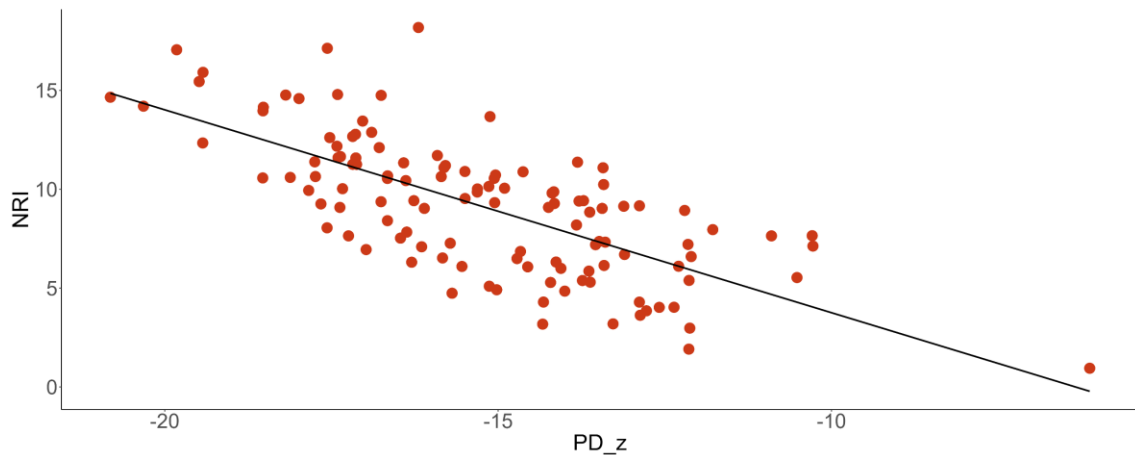


Figura 20 – Relação entre o índice de parentesco líquido (NRI) e a diversidade filogenética de Faith padronizada (PD_z) ($r = -0,71$) para a árvore total com todos as OTUs obtidas com o gene 18S na planície de inundação do rio Araguaia.

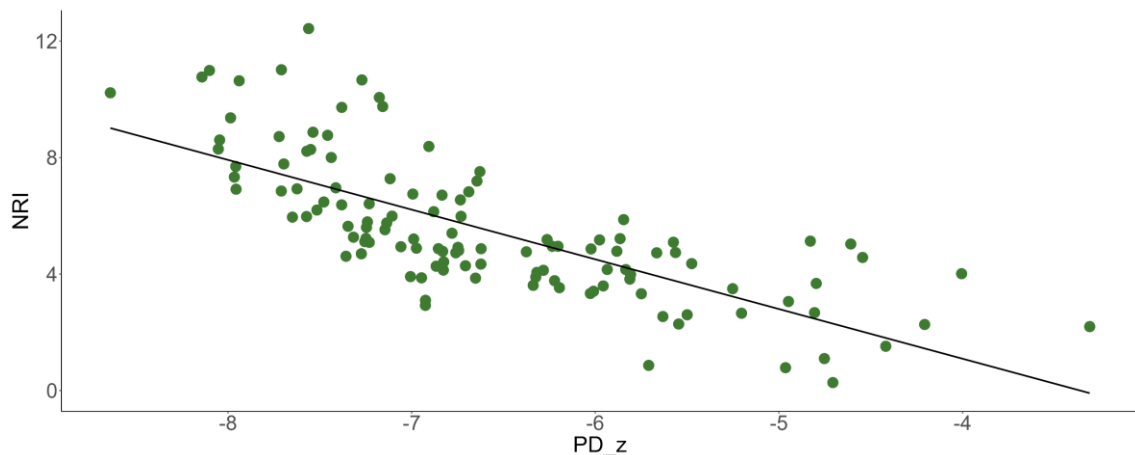


Figura 21 - Relação entre o índice de parentesco líquido (NRI) e a diversidade filogenética de Faith padronizada (PD_z) ($r = -0,74$) para a árvore focal, com os microeucariotos planctônicos fotossintetizantes identificados com o gene 18S na planície de inundação do rio Araguaia.

O modelo nulo para MPD (Dispersão Média Ponderada) que é o NRI não padronizado, revelou que o agrupamento filogenético, dado por $mpd.obs$ (observado) < $mpd.esp$ (esperado), foi significativo para a maior parte dos lagos analisados, exceto para o lago 45 para a árvore total (Fig. 22) e para os lagos 48, 49, 54, 86 e 119 para a árvore focal (Fig. 23).

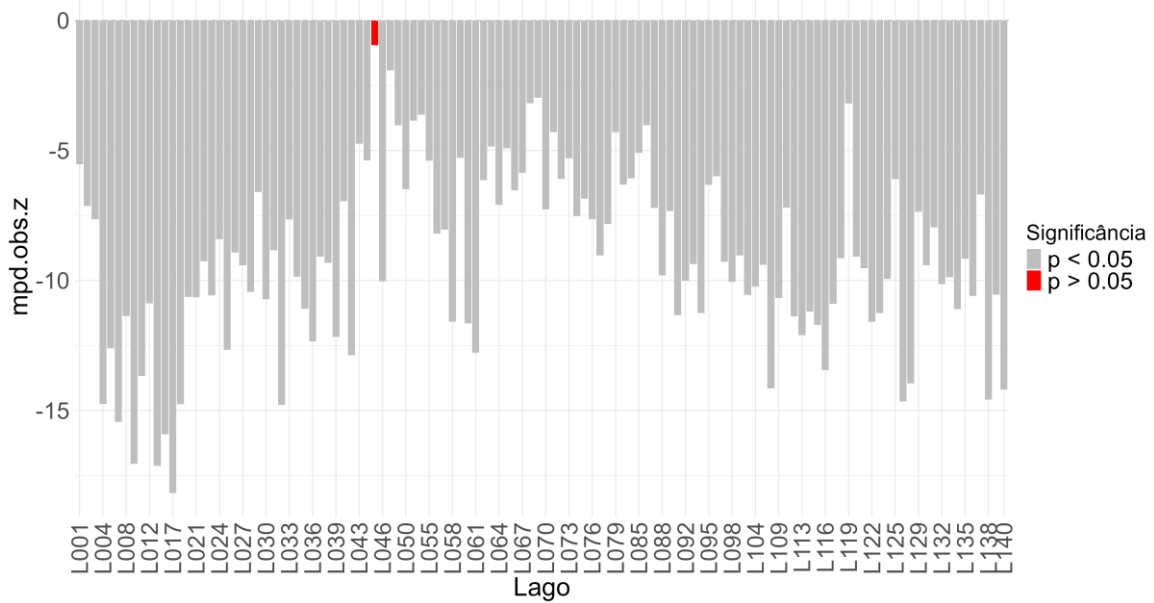


Figura 22 - Modelo nulo de MPD (Dispersão Média Ponderada), o NRI não padronizado, da árvore total, com todas as OTUs obtidas com o gene 18S, gerado com 999 permutações usando a função “ses.mpd”. Valores significativos estatisticamente ($p < 0,05$) estão em cinza, valores não significativos estatisticamente ($p > 0,05$) estão em vermelho.

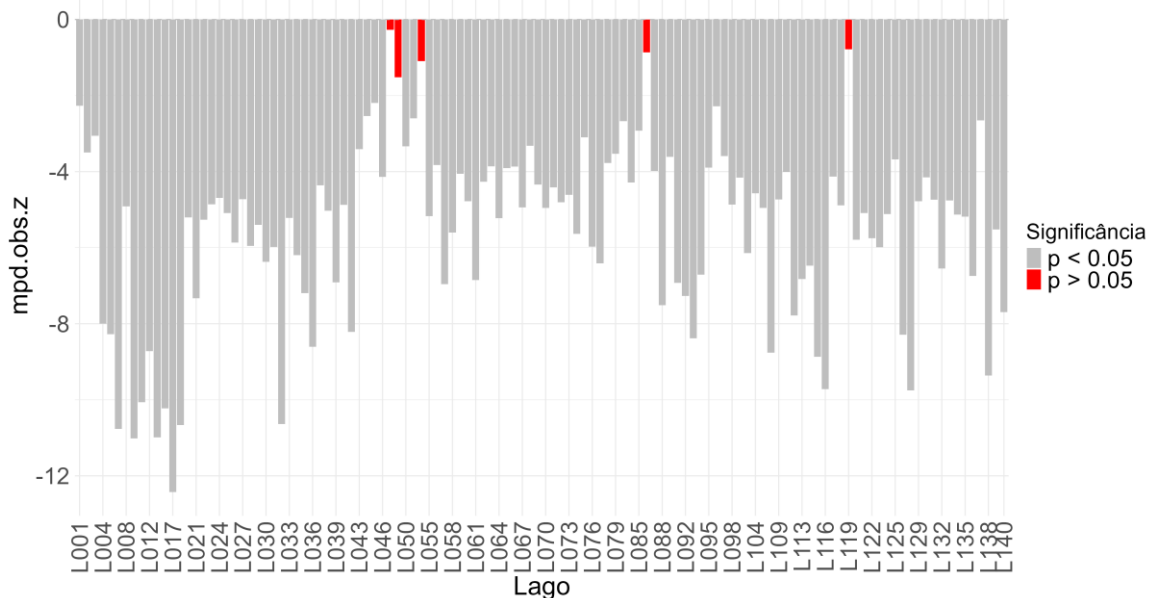


Figura 23 - Modelo nulo de MPD (Dispersão Média Ponderada), o NRI não padronizado, da árvore focal, com as OTUs relacionadas aos grupos de microeucariotos plancctônicos fotossintetizantes, gerado com 999 permutações usando a função “ses.mpd”. Valores significativos estatisticamente de NRI ($p < 0,05$) estão em cinza, valores não significativos estatisticamente ($p > 0,05$) estão em vermelho.

Influência das variáveis ambientais e espaciais na diversidade alfa filogenética da comunidade

Para investigar a influência das variáveis ambientais e variáveis espaciais obtidas com os filtros espaciais PCNMs na variação da diversidade filogenética da árvore total, foram selecionadas treze variáveis ambientais com baixo índice de multicolinearidade ($VIF < 5$) e cinco variáveis espaciais (autovetores de PCNMs de ampla e fina escala) que foram estatisticamente significativas ($p < 0,05$) para PD da árvore total. As variáveis selecionadas estão detalhadas na tabela do Apêndice 7. Juntas, às variáveis ambientais e espaciais explicaram 55% da variação na diversidade filogenética (PD) da árvore total. A maior parte desta variação foi explicada pelo componente compartilhado do ambiente e espaço (ambiente espacialmente estruturado) (24%), enquanto a menor proporção foi atribuída ao efeito puramente ambiental (15%) e 16% da variação foi atribuída ao efeito espacial puro. Os 45% restante da variação não pôde ser explicada pelo modelo utilizado (Fig. 24, A). Para a diversidade filogenética da árvore focal foram selecionadas as mesmas treze variáveis ambientais, e sete variáveis espaciais (autovetores de PCNMs) de fina e ampla escala (ver apêndice 7). Juntas estas variáveis explicaram 41% da variação de PD da árvore focal. O efeito espacial puro e o efeito das variáveis ambientais espacialmente estruturadas explicaram a maior parte da variação (17% cada) enquanto o efeito ambiental puro explicou apenas 7% da variação. No entanto a maior parte da variação dos dados (59%) não pôde ser explicada pelo modelo (Fig. 24, B).

Para a estrutura filogenética (NRI) da árvore total foram selecionadas também as mesmas treze variáveis ambientais, e seis variáveis espaciais (autovetores de PCNMs) de ampla e fina escala (ver apêndice 7). Juntas as variáveis ambientais e espaciais explicaram 48% da variação do NRI para a árvore total. A maior parte desta variação foi explicada pelo efeito puramente espacial (27%), seguido pelo efeito do ambiente espacialmente estruturado (15%), e a menor proporção (6%) foi explicada pelo efeito puramente ambiental. O restante (52%) da variação dos dados não pôde ser explicada pelo modelo (Fig. 24, C). Para a estrutura filogenética da árvore focal foram selecionadas as treze variáveis ambientais e sete variáveis espaciais (autovetores de PCNMs) de ampla e fina escala (ver apêndice 7). Juntas estas variáveis explicaram 48% da variação nos valores de NRI para a árvore focal. A maior parte desta variação foi atribuída ao efeito puramente

espacial (24%), seguido pelo efeito do ambiente espacialmente estruturado (16%) e em menor proporção pelo efeito puramente ambiental, que só explicou 8% da variação. Os outros 52% da variação dos dados não pôde ser explicada pelo modelo (Fig. 24, D).

Todos as frações da partição da variância tanto para a diversidade filogenética, como para a estrutura filogenética para ambas as árvores foram estatisticamente significativas ($p < 0,05$).

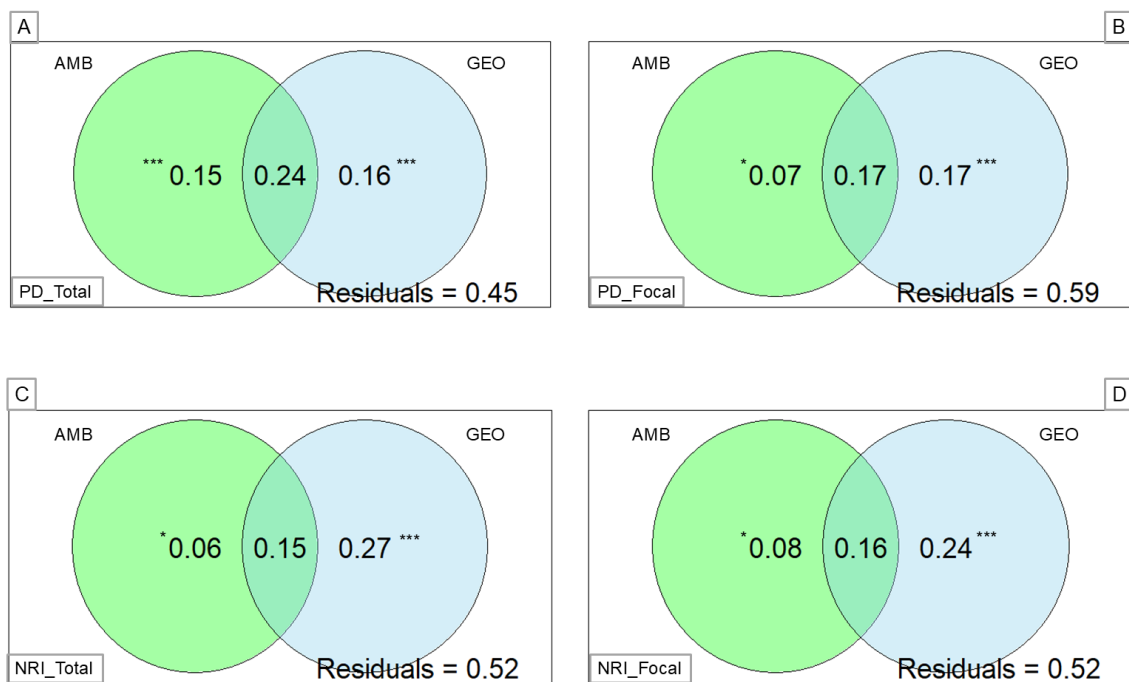


Figura 24 - Diagrama de Venn da partição de variância da diversidade (PD) e estrutura filogenética (NRI) de OTUs para a árvore total (PD_Total e NRI_Total), e a árvore focal (PD_Focal e NRI_Focal), em relação ao ambiente (AMB) e espaço (GEO), com os filtros espaciais PCNMs (Coordenadas Principais de Matrizes Vizinhas). * = $p < 0,05$ e *** = $p < 0,001$.

Para avaliar a influência das variáveis ambientais e variáveis espaciais, obtidas com os filtros espaciais AEMs, na diversidade filogenética (PD) da árvore total, foram selecionadas treze variáveis ambientais e dezesseis variáveis espaciais (autovetores de AEMs, com $p < 0,05$) de ampla e fina escala (tabela de variáveis em apêndice 7). Juntas

estas variáveis explicaram 75,5% da variação da PD da árvore total. A maior proporção dessa explicação é atribuída ao efeito espacial puro (36%), enquanto o ambiente espacialmente estruturado explicou a segunda maior proporção (32,5%) e o ambiente puro apresentou o menor percentual de explicação (7%). Os 24,5% da variação restante não pôde ser explicada pelo modelo (Fig. 25, A). Para a PD da árvore focal, foram selecionadas treze variáveis ambientais e vinte e uma variáveis espaciais (autovetores de AEMs) de ampla e fina escala espacial (ver apêndice 7). Juntas, estas variáveis explicaram 72% da variação da diversidade filogenética da árvore focal. A maior parte (48%) é atribuída ao efeito espacial puro, a segunda maior proporção (23%) é atribuída ao efeito do ambiente espacialmente estruturado e a menor (1%) é atribuída ao efeito ambiental puro. E 28% da variação não pôde ser explicada pelo modelo (Fig. 25, B).

Para avaliar a influência das variáveis ambientais e variáveis espaciais, obtidas com os filtros espaciais AEMs, na estrutura filogenética (NRI) da árvore total, foram selecionadas treze variáveis ambientais e quatorze variáveis espaciais de ampla e fina escala (ver variáveis na tabela do apêndice 7). Juntas estas variáveis explicaram 64,8% da variação de NRI da árvore total. A maior proporção (43,3%) é atribuída ao efeito espacial puro, seguida pelo efeito do ambiente espacialmente estruturado (17,1%), e a menor proporção (4,4%) da variação dos dados foi atribuída ao efeito ambiental puro. Os 35,2% restante da variação dos dados não pôde ser explicada pelo modelo (Fig. 25, C). Para NRI da árvore focal, foram selecionadas treze variáveis ambientais e treze variáveis espaciais de ampla e fina escala (ver apêndice 7). Junto essas variáveis explicaram 58% da variação da estrutura filogenética da árvore focal. O efeito espacial puro explicou a maior proporção (34%), seguido pelo efeito do ambiente espacialmente estruturado (20%) e por último pelo efeito ambiental puro, que explicou a menor proporção da variação (4%). Os 42% restante da variação dos dados não pôde ser explicada pelo modelo (Fig. 25, D).

Todos as frações da partição da variância tanto para a diversidade filogenética, como para a estrutura filogenética para ambas as árvores foram estatisticamente significativas ($p < 0,05$), exceto para a fração 'a' (variância única explicada pelo ambiente controlada pelo espaço) da PD da árvore focal ($p = 0,2$).

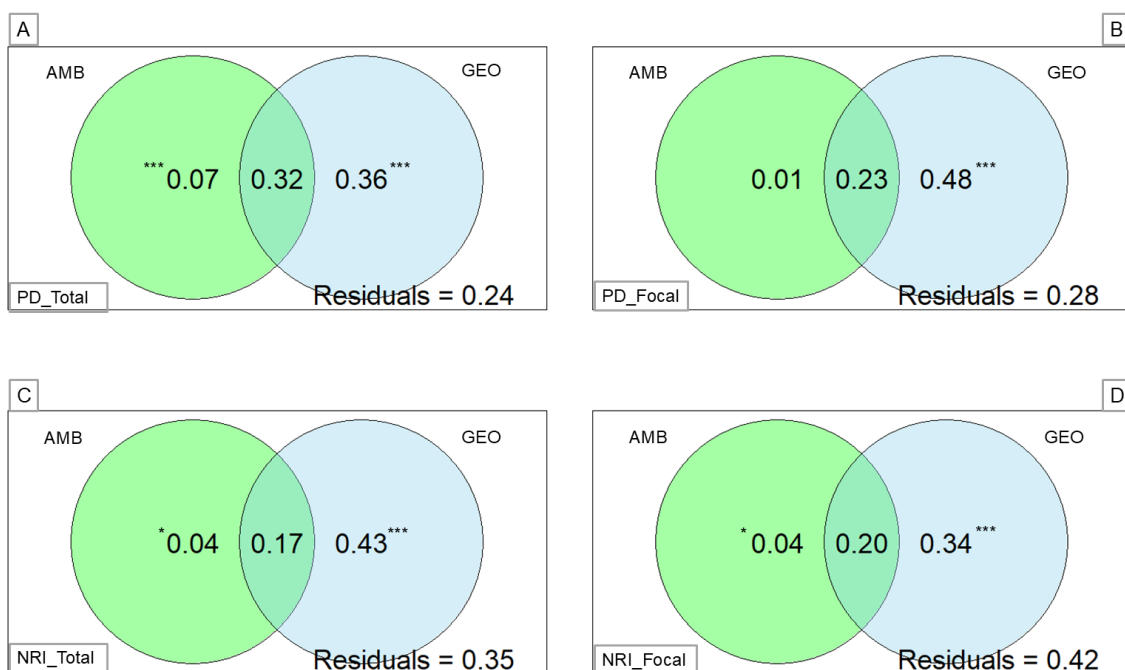


Figura 25 - Diagrama de Venn da partição de variância da diversidade (PD) e estrutura filogenética (NRI) de OTUs para a árvore total (PD_Total e NRI_Total), e a árvore focal (PD_Focal e NRI_Focal), em relação ao ambiente (AMB) e espaço (GEO), com os filtros espaciais AEMs (Mapas de Autovetores Assimétricos). * = $p < 0,05$ e *** = $p < 0,001$.

A análise de componentes principais (PCA) das variáveis ambientais e limnológicas mostrou que os dois primeiros componentes principais explicaram 35,3% da variação nos dados, sendo 23,3% atribuído à primeira dimensão e 12% à segunda dimensão da PCA. As variáveis que mais contribuíram para a variação explicada pelas dimensões 1 e 2 foram: fósforo (P), sólidos totais dissolvidos (TDS), potencial de óxido redução (ORP), pigmentos totais (CHL_T), turbidez (NTU), temperatura da água (T_agua), profundidade (PROF), condutividade (COND) e pH. O fósforo (P), sólidos totais dissolvidos (TDS), potencial de óxido redução (ORP), clorofila total (CHL_T), turbidez (NTU) e pH foram as variáveis mais importantes no componente principal 1, e temperatura da água (T_agua), profundidade (PROF), condutividade (COND) e pH foram as variáveis mais importantes no componente principal 2 (Fig. 26).

No aspecto geral os lagos apresentaram uma heterogeneidade considerável entre eles. Embora não tenha sido observada uma clara associação entre as características ambientais dos lagos e os rios aos quais pertencem, exceto pelos lagos pertencentes ao rio do Peixe e rio das Mortes que apresentaram características ambientais mais homogêneas entre eles. Os lagos do rio Araguaia mostraram a maior amplitude de variação nas características ambientais, enquanto os lagos do rio do Peixe e do rio das Mortes apresentaram menor variabilidade. Os lagos 52, 53, 54, 83, 87 e 11 se destacaram por suas características mais distintas (Fig. 27).

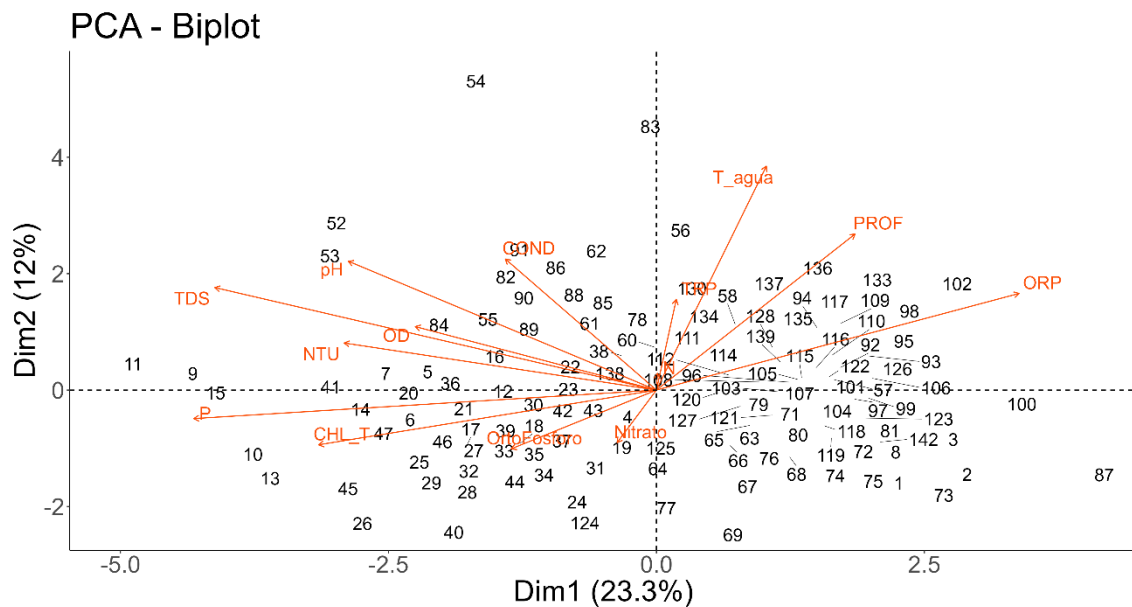


Figura 26 - Análise de componentes principais (PCA) de 14 variáveis ambientais e limnológicas coletadas nos 140 lagos do rio Araguaia e seus afluentes com a ordenação dos pontos de coleta. As variáveis ambientais e limnológicas usadas foram: Sólidos totais dissolvidos (TDS), potencial hidrogeniônico (pH), oxigênio dissolvido (OD), potencial de óxido redução (ORP), unidade nefelométrica de turbidez (NTU), temperatura da água (T_agua), condutividade (COND), transparência da água (TRP), profundidade (PROF), pigmentos totais (CHL_T), nitrogênio (N), nitrato, fósforo (P) e ortofosfato.

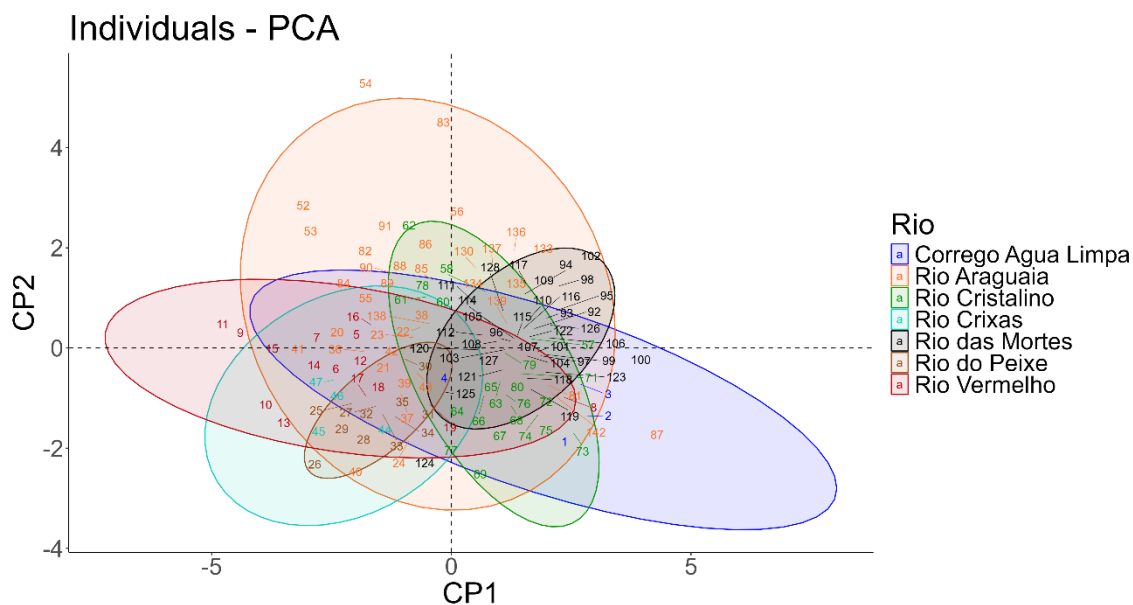


Figura 27 - Ordenação dos pontos da análise de componentes principais (PCA) de 14 variáveis ambientais e limnológicas coletadas nos 140 lagos do rio Araguaia e seus afluentes com o destaque para os rios ao qual os lagos pertencem. Azul = Córrego Água Limpa, laranja = Rio Araguaia, verde = Rio Cristalino, ciano = Rio Crixás, preto = Rio das Mortes, marrom = Rio do Peixe e vermelho = Rio Vermelho.

Diversidade beta filogenética

A diversidade beta filogenética para os microeucariotos planctônicos fotossintetizantes (árvore focal) ficou entre 0,34 e 0,85, muitos pontos apresentaram valores de diversidade filobeta similares (Fig. 28). Os lagos 21 e 22 apresentaram maior similaridade filogenética (0,85) seguido pelos lagos 72 e 73 (0,80). Os lagos 1 e 45 e 80 e 45 (0,34) apresentaram a menor similaridade filogenética. Os resultados do teste de mantel mostraram que existe uma relação positiva e significativa entre a dissimilaridade ambiental e a dissimilaridade da diversidade beta filogenética ($r = 0,29$, $p = 0,001$). Para a dissimilaridade espacial em linha reta também foi observado uma relação positiva e significativa com a dissimilaridade da diversidade beta filogenética ($r = 0,14$, $p = 0,007$). A diversidade beta total obtida através do índice se Sorensen (β SOR) foi de 0,96, 0,94 deste valor foi atribuído ao turnover (β tur) e 0,02 ao aninhamento (β nes) de OTUs na comunidade focal.

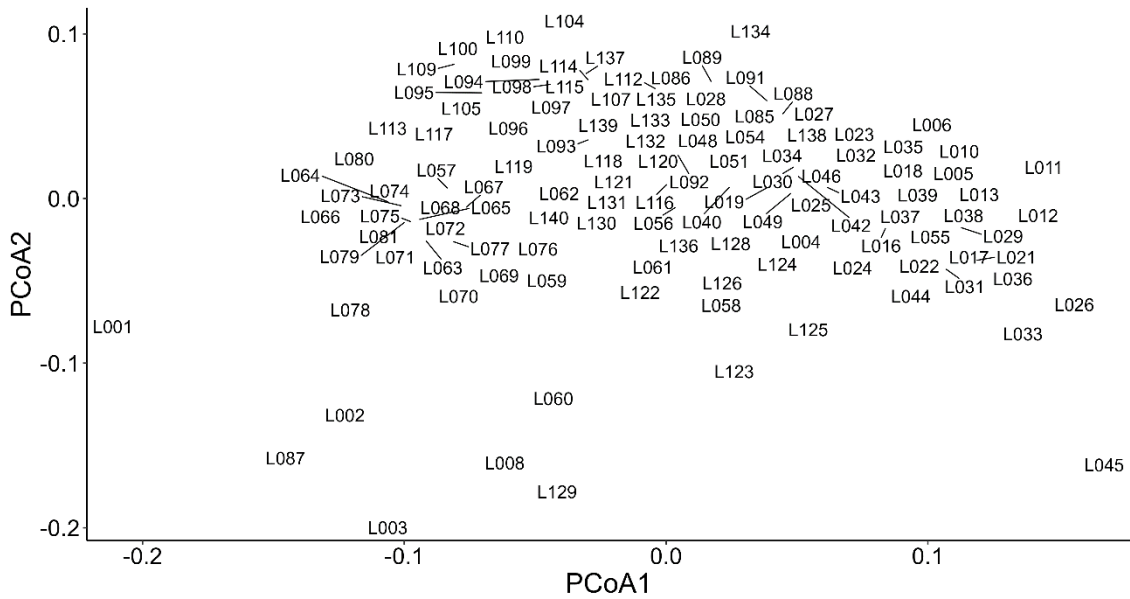


Figura 28 - PCoA da dissimilaridade da diversidade beta filogenética (diversidade filobeta) entre os lagos analisados na planície de inundação do rio Araguaia para a árvore focal, com os microeucariotos planctônicos fotossintetizantes.

DISCUSSÃO

O presente estudo teve o objetivo de investigar a diversidade filogenética da comunidade de eucariotos, capturados pelo marcador molecular 18S, da planície de inundação do rio Araguaia usando a técnica de DNA *metabarcoding*, com foco nos microeucariotos planctônicos fotossintetizantes. Especificamente, determinamos a diversidade filogenética e a estrutura filogenética da comunidade por meio das métricas PD e NRI. Avaliamos a diversidade beta filogenética da comunidade, por meio da métrica phylosor e, por fim, testamos a influência das variáveis ambientais e espaciais na explicação da variação da diversidade e estrutura filogenética da comunidade ao longo da planície de inundação. Nossos resultados mostraram que a comunidade total de eucariotos e microeucariotos planctônicos fotossintetizantes acessados pela região v4 do gene 18S apresentaram diversidades filogenéticas significativamente menores do que as esperadas sob o modelo nulo e essa diversidade foi altamente correlacionada com a riqueza de OTUs. Ambos os conjuntos de dados (total e focal) também apresentaram um padrão filogenético agrupado significativo para a maioria dos lagos. A maior parte da variação da diversidade e estrutura filogenética explicada, foi atribuída ao efeito espacial puro, no entanto o ambiente espacialmente estruturado também representou uma porcentagem significativa

de explicação dos dados. Os resultados foram coincidentes para ambos os filtros espaciais, no entanto os filtros espaciais que calculam a conectividade direcional entre os lagos (AEMs) tiveram uma maior porcentagem de explicação do que os filtros espaciais de distâncias lineares (PCNMs). A maior parte da diversidade beta filogenética (diversidade filobeta) observada para a comunidade focal foi atribuída ao turnover, indicando uma alta substituição de OTUs entre os lagos analisados, e a sua dissimilaridade foi positivamente correlacionada com a dissimilaridade ambiental e espacial.

As comunidades foram dominadas pelo supergrupo SAR (Stramenopiles, Alveolata e Rhizaria), que contém organismos heterotróficos, autotróficos e mixotróficos e por Archaeplastida, um grupo autotrófico, destacando a importância e abundância do grupo autotrófico na composição e diversidade da comunidade de microeucariotos planctônicos. Isso também é apoiado pela maior abundância relativa de OTUs autotróficas na comunidade, apesar de entre as OTUs classificadas os heterotróficos serem melhor representados em presença. Dentre os microeucariotos planctônicos fotossintetizantes o grupo das Chlorophyta foi o mais abundante, este parece ser o grupo de microeucariotos fotossintetizantes predominante na bacia do rio Tocantins-Araguaia, como foi observado também em outros estudos na região (Nabout, Nogueira e Oliveira, 2006; Nabout *et al.*, 2007; Nabout *et al.*, 2009; Machado *et al.*, 2016; Machado *et al.*, 2019; Silva *et al.*, 2024).

A riqueza de OTUs para ambos os conjuntos de dados analisados, assim como a diversidade filogenética, parece apresentar um gradiente latitudinal, aumentando em direção a latitudes menores, à jusante do rio, com as regiões dos rios das Mortes e Cristalino apresentando alta riqueza de OTUs e diversidade filogenética. Resultados semelhantes de um gradiente latitudinal para riqueza de espécies de microeucariotos planctônicos fotossintetizantes, na estação de cheia do rio Araguaia, foi observado por Nabout *et al.* (2007). Isso pode estar associado as características ambientais dos lagos ou processos de dispersão favorecidos pela maior conectividade dos lagos e fluxo do rio. Fatores relacionados ao tamanho do lago, produtividade, qualidade da água e níveis de predação controlam a riqueza de espécies de microeucariotos planctônicos em lagos, no entanto eventos de dispersão também podem contribuir para o aumento da riqueza (Hobæk, Manca e Andersen, 2002). Um padrão latitudinal de riqueza também foi observado para espécies de microeucariotos planctônicos fotossintetizantes em lagos da Patagônia e da Antártida, e sua explicação foi atribuída a fatores ambientais e espaciais da região (Izaguirre *et al.*, 2016).

A árvore filogenética total, com todas as OTUs identificadas com o gene 18S-V4, mostrou um baixo suporte em nós basais, isso provavelmente foi causado pela presença de grupos muito distantes filogeneticamente e pela ausência de táxons que representassem todos os clados presente na filogenia. Nós não resolvidos na filogenia podem afetar os resultados da análise filogenética de comunidade (Okuno *et al.*, 2022). Foi relatado que sequências curtas de DNA (p. ex.: DNA barcode) fornece mais informações sobre diferenças nas distâncias filogenéticas entre espécies e gêneros intimamente relacionados (Okuno *et al.*, 2022), resolvendo melhor os nós terminais das filogenias, mas fornecendo uma baixa resolução em nós basais. Em nossas filogenias, construídas com sequências de DNA metabarcodes, também observamos uma melhor resolução em nós terminais do que nos nós basais. Apesar de não oferecer uma boa resolução basal, filogenias com sequências curtas de DNA tem mostrado vantagens em relações às filogenias tradicionais, permitindo a incorporação na árvore filogenética de organismos não identificados taxonomicamente ou mesmo identificados, mas em níveis taxonômicos altos (Yi *et al.*, 2010; Heckenhauer *et al.*, 2017). Um estudo de Yi *et al.* (2010) demonstrou que a inclusão de sequências não identificadas ao conjunto de sequências identificadas para inferência filogenética de um grupo de interesse, melhorou a inferência filogenética de eucariotos microbianos, mudando a topologia da árvore e melhorando o suporte de nós.

A falta de resolução filogenética afeta as métricas de diversidade filogenética e estrutura filogenética, principalmente quando é usado árvores grandes sem boa resolução basal (Swenson, 2009). Swenson (2009) demonstrou que a diversidade filogenética é cada vez mais subestimada à medida que as filogenias se tornam menos resolvidas e quando árvores maiores são usadas. Park, Worthington e Xi (2018) argumentam que filogenias grandes, como as filogenias comunitárias podem enviesar estimativas da diversidade filogenética e levar a uma classificação incorreta da diversidade da comunidade, devido a subamostragem dos táxons.

Diferenças na amostragem de táxons, afeta a resolução e topologia das árvores filogenéticas, o que pode afetar as métricas de diversidade (Park, Worthington e Xi, 2018). Uma alternativa para melhorar a resolução e precisão de filogenias com muitos grupos distintos, seria realizar uma amostragem mais ampla de táxons (Hillis *et al.*, 2003) e incluir táxons não necessariamente de interesse do estudo como marcadores de posição na filogenia (Park, Worthington e Xi, 2018). A inclusão de sequências multilocus (sequências de diferentes marcadores genéticos) e o uso conjunto de árvores de restrição como um

modelo de como as espécies estão relacionadas, ao utilizar sequências curtas de DNA para construção de filogenias, pode contribuir para aumentar a resolução topológica e o suporte filogenético para árvores de comunidades (Kress *et al.*, 2010; Liu *et al.*, 2019; Macías-Hernández *et al.*, 2020). Isso pode ajudar a garantir que a análise filogenética resultante respeite os padrões filogenéticos previamente conhecidos para o grupo estudado, aumentando a robustez da inferência filogenética.

Ao analisar a diversidade filogenética da comunidade total, com todos as OTUs, e da comunidade focal, com os microeucariotos planctônicos fotossintetizantes, observamos um padrão similar de diversidade filogenética entre ambos os grupos. Isso sugere que os microeucariotos fotossintetizantes desempenham um papel crucial na manutenção da diversidade filogenética da comunidade total ou/e que este grupo foi mais bem capturado com os primers usados para o gene 18S-V4. Estes fatores podem ter contribuído com a melhor resolução da árvore filogenética para o grupo focal de fotossintetizantes.

Para ambas os conjuntos de dados (total e focal), a diversidade filogenética de Faith observada foi menor do que a esperada ao acaso para todos os lagos analisados. PD é uma medida de diversidade alfa que mede a riqueza dentro de cada local e essa riqueza é comparada com uma riqueza predita por um modelo nulo (Tucker *et al.*, 2017). Menores valores de PD do que o esperado por um modelo nulo indica menor quantidade de história evolutiva acumulada no local e menor divergência evolutiva observada para o conjunto de OTUs, indicando a presença de mais OTUs filogeneticamente relacionadas na comunidade. Valores mais elevados de PD indicam concentração de linhagens profundas e ramos longos na filogenia, revelando uma maior quantidade de história evolutiva acumulada no local (Diniz-Filho *et al.*, 2024; Tucker *et al.*, 2019). As unidades dos comprimentos de ramos, determinam a interpretação precisa de PD, estes podem estar em unidades de mudanças observáveis em características, distâncias genéticas ou tempo evolutivo, mais no geral eles são descritos como estimativas de história evolutiva, informação ou divergência evolutiva (Tucker *et al.*, 2019). Em nosso caso, o tamanho dos ramos representa distâncias/divergência genética entre as OTUs. Os valores de PD são sensíveis a qualidade da estimativa de comprimento dos ramos da filogenia, como destacado por Faith (1992) ao formular a métrica, e recentemente lembrado por Cardillo (2023), que destaca a importância de saber qual o grau de inconsistência que a incerteza na inferência filogenética causa nas estimativas de PD, algo ainda pouco explorado.

A diversidade filogenética apresentou uma forte correlação positiva com a riqueza de OTUs, um padrão frequentemente relatado para a métrica PD (Helmus *et al.*, 2007; Tucker *et al.*, 2019). Essa correlação é tão forte, que alguns autores argumentam que a riqueza de espécies pode ser um bom substituto para medidas de diversidade filogenética (Polasky *et al.*, 2001; Rodrigues, Brooks e Gaston, 2005), no entanto outros autores se opõem a isso, ao dizer que o uso da riqueza como um proxy para a diversidade filogenética não é adequado devido a padrões conflitantes entre a riqueza de táxons e PD (Forest *et al.*, 2007; Pio *et al.*, 2011). Tucker e Cadotte (2013) argumentam que apesar de existir uma correlação entre a riqueza de espécies (SR) e PD, a força dessa correlação pode ser afetada por fatores, como: a estrutura da árvore filogenética, tamanho do pool de espécies e pela estrutura espacial das áreas de distribuição de espécies. Por exemplo, em árvores com mais informações nos ramos terminais é esperada uma alta correlação para SR e PD, assim como se ocorrer uma alta autocorrelação espacial em presença de espécies e quando temos um pool de espécies pequenos. Já em pools regionais relativamente grandes é mais comum aparecer locais onde a correlação entre SR e PD é baixa (Tucker e Cadotte, 2013).

Ambas os conjuntos de dados (total e focal) apresentaram um padrão filogenético agrupado (NRI >0), e esse padrão foi significativo para a maioria dos lagos. O agrupamento filogenético é gerado pela maior proximidade filogenética das espécies presente na comunidade (Webb *et al.*, 2002). Em comunidades de microrganismos aquáticos tem sido comum encontrar um padrão filogenético agrupado (Horner-Devine e Bohannan, 2006; Barberán e Casamayor, 2010; Pontarp *et al.*, 2012; Keck e Kahlert, 2019; Chen *et al.*, 2023), isso pode indicar uma maior importância da filtragem ambiental na montagem da comunidade de microrganismos, já que estes exibem maior sensibilidade as condições ambientais (Ren *et al.*, 2015). Um padrão filogenético agrupado é uma evidência de que filtros ambientais estão atuando na comunidade, fazendo com que táxons mais filogeneticamente relacionados que tem tolerâncias ambientais semelhantes sejam selecionados na comunidade (Webb *et al.*, 2002; Cavender-Bares *et al.*, 2004). Do contrário observaríamos um padrão superdisperso, com uma abundância de táxons distantemente relacionados, o que poderia indicar exclusão competitiva entre espécies filogeneticamente próximas (Webb *et al.*, 2002; Cavender-Bares *et al.*, 2004). No entanto, estes podem não ser os únicos mecanismos envolvidos na formação da estrutura filogenética da comunidade, podendo até causarem um efeito inverso aquele esperado, com a competição levando ao agrupamento se uma característica filogeneticamente conservada determinar se as espécies

são boas competidoras (Vamosi *et al.*, 2009). Ou superdisperso quando traços ecológicamente importantes para especialização de habitats tem uma história evolutiva convergente (Cavender-Bares *et al.*, 2004; Cavender-Bares, Keen e Miles, 2006). Uma estrutura filogenética aleatória como observada para os lagos 45 da árvore total e para os lagos 48, 49, 54, 86 e 119 da árvore focal, pode ser um indicativo de interações neutras, ou que a força dos processos ambientais é equilibrada ou fracas, e características filogenéticas aleatórias estruturam a comunidade (Kembel e Hubbell, 2006).

É esperado que o NRI seja estreitamente correlacionado ao PD padronizado (Vamosi *et al.*, 2009). Observamos que essa correlação é negativa, onde é observado altos valores de diversidade filogenética é observado baixos valores de NRI, os lagos pertencentes ao rio Cristalino, por exemplo, apresentaram alta riqueza e diversidade filogenética, e baixa estrutura filogenética, apresentando os menores valores de NRI tanto para a árvore total como para a árvore focal. Isso indica que onde temos uma maior diversidade filogenética temos táxons que são mais distantemente relacionados reduzindo o agrupamento filogenético no local.

Assim como o PD, o NRI também pode ser influenciado pela topologia e resolução da árvore filogenética (Kembel e Hubbell, 2006; Swenson, 2009; Muscarella *et al.*, 2014; Park, Worthington e Xi, 2018). É mais provável encontrar valores de NRI que estão mais próximos da expectativa nula com filogenias com baixa resolução, principalmente ao usar árvores filogenéticas grandes (Swenson, 2009). O uso de grandes arvores com nós basais mal resolvidos, afeta a detecção da estrutura filogenética da comunidade (Swenson, 2009). O NRI é um método que possui um excelente poder de detecção de padrões resultantes da filtragem ambiental, no entanto é mais sensível às questões dos nós mal resolvidos na filogenia (Kraft *et al.* 2007; Pausas e Verdú, 2010), o que pode enviesar os resultados para a métrica. Apesar da baixa resolução da filogenia afetar as métricas de diversidade, Swenson (2009) demonstrou que os padrões de diversidade entre filogenias bem resolvidas e filogenias com baixa resolução permaneceu parecidos, ele observou que os valores para as métricas de diversidade foram altamente correlacionados entre as filogenias com baixa e alta resolução filogenética. Portanto, podemos supor que nossos resultados estão no caminho certo, os padrões de diversidade e estrutura filogenética observados neste estudo são consistentes e indicam uma estruturação não aleatória da comunidade. Embora a resolução das árvores filogenéticas, especialmente da árvore total, possa influenciar as estimativas quantitativas, os padrões gerais observados são robustos, segundo a literatura,

e sugerem a importância de processos ecológicos e evolutivos na organização da comunidade.

Além da resolução da filogenia, a riqueza de espécies e a escala do estudo também podem afetar o padrão de estrutura filogenética identificado na comunidade. A variação na riqueza de espécies provocada por filtros ambientais pode influenciar a interpretação do NRI, principalmente quando a filtragem ambiental forte favorece um clado de tamanho moderado, fazendo com que o NRI se torne cada vez mais negativo à medida que a riqueza de espécies aumenta (Sandel, 2018). A escala do estudo pode influenciar não só quais padrões são identificados, mas também a importância dos processos que atuam na comunidade (Vamosi *et al.*, 2009). Em pequenas escalas espaciais frequentemente encontramos uma estrutura filogenética superdispersa, mas em grandes escalas espaciais é comum encontrar um padrão agrupado (Cavender-Bares, Keen e Miles, 2006; Swenson *et al.*, 2007). Em escala regional ou em grandes escalas espaciais, a biogeografia da especiação pode impulsionar o agrupamento filogenético em comunidades locais, enquanto o tipo de especiação e a limitação da dispersão local pode levar a comunidades superdispersas (Vamosi *et al.*, 2009).

Outros fatores que afetam o padrão de estrutura filogenética identificado na comunidade são a escala taxonômica e o tamanho da comunidade. Estudos demonstraram que à medida que a escala taxonômica vai se tornando mais fina é comum identificar um padrão superdisperso na comunidade (Swenson *et al.*, 2006; Cavender-Bares, Keen e Miles, 2006), e à medida que se aumenta a escala taxonômica o inverso acontece. O tamanho da comunidade local, em relação ao pool regional também influenciam o poder estatístico das análises filogenéticas, a presença de comunidades locais muito pequenas ou muito grande em relação ao pool regional podem aumentar as taxas de erros do tipo 2 (rejeitar uma hipótese alternativa quando ela é verdadeira) (Kraft *et al.*, 2007). Já para comunidades de tamanho intermediário entre 30 e 60% do pool regional há uma melhora no poder estatístico, dependendo da métrica utilizada (Kraft *et al.*, 2007). Em nosso estudo observamos um padrão filogenético agrupado como é esperado, segundo a literatura, para estudos em grandes escalas espaciais com uma escala taxonômica grosseira. O tamanho da comunidade local representou menos de 30% do pool regional para o conjunto de dados total e menos de 40% do pool regional no conjunto de dados focal, no entanto isso não afetou a detecção de um agrupamento filogenético estatisticamente significativo para ambos os conjuntos de dados.

A maior parte da variação da diversidade e estrutura filogenética foi explicada pelo efeito espacial puro, para ambos os filtros espaciais, no entanto os filtros espaciais que calcula a conectividade direcional dos lagos ao longo do rio (AEMs) tiveram uma maior porcentagem de explicação do que os filtros espaciais com distâncias lineares (PCNMs). Um efeito significativo do componente espacial puro sugere a atuação de processos relacionados a dispersão ocorrendo na metacomunidade. O efeito combinado do ambiente e espaço também representou uma porção significativa de explicação para ambos os filtros, indicando que algumas variáveis ambientais estão estruturadas no espaço. A influência das variáveis ambientais e espaciais na diversidade (PD) e agrupamento filogenético (NRI), avaliada com os filtros espaciais PCNMs mostraram que o espaço (puro e compartilhado) apresentou uma maior influência na diversidade e estrutura filogenética para ambas as árvores, e na maioria dos casos o espaço puro apresentou uma maior porcentagem de explicação, exceto para a estrutura filogenética da árvore total, onde o efeito do ambiente estruturado no espaço apresentou uma porcentagem de explicação maior. Com os filtros espaciais AEMs, foi observado que o espaço (puro e compartilhado) também apresentou uma maior influência na variação de PD e NRI para ambas as árvores, no entanto, o espaço puro apresentou uma maior porcentagem de explicação para a variação da diversidade e estrutura filogenética. A influência ambiental pura apesar de significativa para maioria das métricas, no geral foi muito baixa para ambas as árvores com ambos os filtros espaciais, sendo menor quando os filtros AEMs foram usados. Esses resultados indicam que tanto a classificação de espécies como a dispersão são fatores importantes na estruturação da metacomunidade, no entanto os processos relacionados a dispersão parecem ter um papel predominante.

Preditores espaciais também foram o fator principal na estruturação da metacomunidade de microeucariotos planctônicos fotossintetizantes analisadas sob um aspecto taxonômico e funcional na planície de inundação do rio Araguaia (Silva *et al.*, 2024). Os autores argumentam que a homogeneização das condições bióticas causada pelo pulso de inundação, pode ter diminuído a força dos filtros ambientais sobre as metacomunidades da região. A homogeneização dos habitats após as inundações derivadas do aumento da conectividade durante o período de águas altas é considerada um padrão geral para os ambientes de planície de inundação (Thomaz, Bini e Bozelli, 2007). No entanto, Nabout *et al.* (2009) ao investigar como a composição das comunidades de microeucariotos planctônicos fotossintetizantes foram afetadas por fatores ambientais e

espaciais na estação seca e chuvosa na bacia do rio Araguaia, não encontraram nenhuma evidência de processos espaciais, nem ambientais atuando na estruturação das comunidades em nenhuma das estações. Durante a fase de cheia de planícies de inundação os processos espaciais podem ser mais influentes, devido as altas taxas de dispersão das espécies (Heino *et al.*, 2015). No entanto, essa dispersão pode ser limitada com o aumento da extensão espacial e distância espacial entre os locais (Heino *et al.*, 2015) principalmente para organismos que se dispersam passivamente (De Bie *et al.* 2012). Mesmo em sistemas altamente interconectado, as condições ambientais podem ser fortes o suficiente para estruturar as comunidades locais (Cottenie *et al.*, 2003) e a dispersão pode ser resultado da rastreabilidade da heterogeneidade ambiental, mesmo em grandes escalas espaciais (Cottenie, 2005) ao invés de ser provocada por efeito de massa, por exemplo.

Embora o padrão frequentemente encontrado para microeucariotos planctônicos, é o ambiente atuando como fator principal na estruturação da comunidade, indicando que a classificação de espécies é o processo mais importante nas metacomunidades de microrganismos aquáticos (Cottenie *et al.*, 2003; Cottenie, 2005; Barberán e Casamayor, 2010; De Bie *et al.* 2012; Huszar *et al.*, 2015), alguns estudos mostram que o espaço também pode atuar como um fator significativo (Jin *et al.*, 2022; Xiao *et al.*, 2022; Huang *et al.*, 2023). O ambiente também pode estar estruturado no espaço, e nesse caso fatores ambientais e espaciais atuam conjuntamente na estruturação da comunidade (Bortolini *et al.*, 2017; Jin *et al.*, 2022). Para planícies de inundação, o período de águas altas parece aumentar a importância do fator espacial em comunidades de microeucariotos planctônicos (Chaparro, O'Farrell, eHein, 2023; Huang *et al.*, 2023). No entanto, a força desses fatores pode mudar dependendo da escala espacial. Em escalas menores podemos ver uma forte atuação dos efeitos ambientais (Declerck *et al.*, 2011; Machado *et al.*, 2016), mas em grandes escalas fatores espaciais apresentam uma significância maior na estruturação da comunidade (Heino *et al.*, 2010; Declerck *et al.*, 2011). De Bie *et al.* (2012) argumentam que com o aumento da escala espacial, as forças que regulam as metacomunidades tendem a mudar, passando de controladas pelo ambiente para controladas pela dispersão. No entanto, isso pode não ser verdade para todos os microrganismos planctônicos, e mesmo em grandes escalas espaciais o ambiente pode permanecer como forte controlador da estrutura da comunidade (Astorga *et al.*, 2012), ou o espaço pode apresentar uma influência significativa mesmo em pequenas escalas espaciais (Soininen *et al.*, 2007). Apesar disso tudo, temos que considerar que o componente espacial também pode capturar parte da

variação ambiental não medida pelas variáveis ambientais incluídas no modelo (Cottenie, 2005) o que pode aumentar o poder de explicação do componente espacial.

O uso de diferentes filtros espaciais pode gerar diferentes resultados para a influência do espaço na diversidade e estrutura filogenética da comunidade aquática. Em nosso estudo observamos que os filtros espaciais que consideram a conectividade e direcionalidade do fluxo da água foram melhores preditores da variação da estrutura e diversidade filogenética da comunidade. Ao trabalhar com ambientes aquáticos, as distâncias baseadas em curso d'água pode capturar padrões espaciais melhor e explicar mais variância do que as distâncias lineares (Landeiro *et al.*, 2011). Quando variáveis espaciais são os principais preditores da estrutura da metacomunidade, distâncias do curso de água podem ser mais importantes, resultando em melhores estimativas dos condutores espaciais do que distâncias terrestres (Padiál *et al.*, 2014). Liu *et al.* (2013) observaram que filtros que consideram apenas distâncias terrestres eram preditores fracos para padrões espaciais de metacomunidades aquáticas, e que filtros que consideravam a conectividade e direcionalidade do fluxo de água eram melhores preditores para compreensão da estrutura espacial de microrganismos de dispersão passiva. Por outro lado, Grönroos *et al.* (2013) observaram que para macroinvertebrados aquáticos os filtros espaciais usando diferentes distâncias foram ambos descritores pobres para descrever a estrutura da comunidade. Em contraste aos nossos resultados, Machado *et al.* (2019) não observaram efeitos significativos do espaço na composição de OTUs de comunidades de microeucariotos planctônicos na planície de inundação do rio Araguaia ao usar diferentes filtros espaciais (AEMs e PCNMs). Isso pode ter ocorrido devido aos tipos de autovetores de AEMs e PCNMs utilizados, enquanto eles utilizaram principalmente autovetores que captam a variação espacial em uma escala mais ampla, nós utilizamos autovetores que capturam a variação em fina e ampla escala. O uso de autovetores de ampla (autovetores com números mais baixos) e fina escala (autovetores com números mais altos) indicam processos diferentes, com os autovetores de ampla escala sendo mais propensos a indicar efeito de massa, enquanto autovetores maiores podem indicar processos relacionados a limitação da dispersão (Heino *et al.*, 2015). Bortolini *et al.* (2017) ao utilizar filtros espaciais AEMs de fina e ampla escala espacial, observaram uma influência do espaço na biomassa de microeucariotos planctônicos fotossintetizantes, sugerindo que os processos de dispersão influenciavam a estruturação da comunidade e que a configuração espacial é fundamental em modelos de metacomunidade.

A diversidade beta filogenética total observada para a comunidade focal foi alta, e sua maior parte foi atribuída ao componente turnover, indicando uma alta substituição de OTUs entre os lagos analisados. Um alto valor de diversidade beta para microeucariotos planctônicos, com uma grande proporção dessa diversidade associada a substituição de OTUs ao longo da planície de inundação do rio Araguaia, também foi observado por Machado *et al.* (2019). No entanto, eles estavam analisando a diversidade beta, e nós aqui analisamos a diversidade beta filogenética, mas como a diversidade beta taxonômica e filogenética são positivamente correlacionadas (Corbelli *et al.*, 2015) acreditamos que a diversidade beta pode ser usada como um proxy para a diversidade beta filogenética. Nabout *et al.* (2007) também observaram uma diversidade beta maior no período de cheia do rio Araguaia, e atribuíram este fato às chuvas isoladas ao longo do rio levarem a um aumento da heterogeneidade ambiental na região. A alta heterogeneidade ambiental é coincidente com um aumento na diversidade beta de microeucariotos planctônicos (Declerck *et al.*, 2011; Wojciechowski *et al.*, 2017). Em outros trabalhos na região do rio Araguaia na época de cheia foi observado que os lagos apresentavam características ambientais heterogêneas (Machado *et al.*, 2019; Nabout *et al.*, 2009). Neste trabalho observamos uma heterogeneidade considerável entre os lagos, mas não ao ponto de criar agrupamentos claros de características ambientais entre os tributários, o que pode ter ocorrido devido a homogeneização das características dos lagos devido a conectividade maior exibida em períodos de cheias do rio (Thomaz, Bini e Bozelli, 2007).

Observamos que a dissimilaridade da diversidade beta filogenética, aumentou com o aumento das distâncias ambientais, revelando a importância da heterogeneidade ambiental para a diversidade beta filogenética. O componente espacial também teve um efeito significativo, porém menor, sugerindo a atuação de algum mecanismo de dispersão. Padrões contrastantes são encontrados para microrganismos, onde a diversidade beta filogenética apresenta maior correlação com as distâncias espaciais, destacando a importância da dispersão para esses organismos (Wang *et al.*, 2013). Para microeucariotos planctônicos fotossintetizantes, o ambiente parece ser o principal fator influenciando a dissimilaridade da diversidade beta (Qu *et al.*, 2018; Huszar *et al.*, 2015), no entanto o espaço também pode ter um efeito significativo para a dissimilaridade beta dessas comunidades (Jin *et al.*, 2022), ou ambos os fatores podem atuarem juntos (Jamoneau *et al.*, 2018). Huszar *et al.* (2015) argumentam que a identificação de uma correlação entre dissimilaridade espacial e de diversidade beta em locais mais próximos pode ser devido ao

compartilhamento de características ambientais exibida por estes locais, indicando uma dissimilaridade espacial estruturada pelo ambiente. Para a região do rio Araguaia, Nabout *et al.* (2007) observaram que a similaridade da comunidade de microeucariotos planctônicos fotossintetizantes foi influenciada principalmente pelo espaço, em contraste aos resultados que observamos para a diversidade beta filogenética. O aumento da dissimilaridade da diversidade beta filogenética com o aumento das distâncias ambientais, pode indicar uma similaridade ecológica para as comunidades de microeucariotos planctônicos fotossintetizantes. Comunidades com maior proporção de espécies intimamente relacionadas, podem estar em ambientes ecologicamente mais semelhantes (Burns e Strauss, 2011). No entanto, Swenson *et al.* (2012) argumentam que a relação filogenética nem sempre é um bom preditor para similaridade ecológica, porque a relação filogenética muitas vezes não é um bom indicador de características (Mazel *et al.*, 2018). Pois as espécies podem ser filogeneticamente diferentes, mas fenotipicamente semelhantes (Tucker *et al.*, 2019).

CONCLUSÃO

A comunidade total de eucariotos e a comunidade focal de microeucariotos planctônicos fotossintetizantes, da planície de inundação do rio Araguaia mostrou padrões semelhantes de riqueza, diversidade e estrutura filogenética. Para ambas, foi observada uma diversidade filogenética menor do que a esperada ao acaso e uma estrutura filogenética agrupada, sugerindo que a montagem da comunidade é influenciada por processos não aleatórios, como limitação da dispersão e filtragem ambiental. A variação da diversidade e estrutura filogenética foi explicada principalmente pelos fatores espaciais, com filtros espaciais que modelam um processo de conectividade direcional entre os lagos (AEMs) apresentando um maior poder de explicação, revelando a importância dos processos de dispersão para a estruturação filogenética da metacomunidade. A diversidade beta filogenética da comunidade de microeucariotos planctônicos fotossintetizantes foi alta e predominantemente influenciada pela substituição de espécies. A dissimilaridade foi correlacionada positivamente com as distâncias ambientais e espaciais, porém as distâncias ambientais apresentaram uma correlação maior com a dissimilaridade da diversidade filobeta.

Este estudo demonstra a eficiência da técnica de *metabarcoding de eDNA* para o monitoramento das comunidades de microeucariotos planctônicos e avaliação da diversidade e estrutura filogenética destas comunidades, ajudando a elucidar questões ecológicas e evolutivas a respeito da montagem das metacomunidades aquáticas. Embora reconheçamos que existem algumas incertezas associadas aos nossos resultados, eles são consistentes e apresenta um padrão geral robusto sobre a diversidade e estrutura filogenética da comunidade de microeucariotos planctônicos fotossintetizantes da planície de inundação do rio Araguaia. Sendo assim, este trabalho vem contribuir para o preenchimento da lacuna de conhecimento sobre a diversidade filogenética da comunidade de microeucariotos planctônicos da bacia do rio Tocantins-Araguaia. E contribui com o aumento do conhecimento da diversidade filogenética para microeucariotos planctônicos fotossintetizantes no geral. E por fim, destaca a importância de considerar tanto os aspectos filogenéticos quanto os ecológicos para compreender as comunidades de microeucariotos planctônicos.

REFERÊNCIAS

AKCAALAN, R. *et al.* Does environmental DNA reflect the actual phytoplankton diversity in the aquatic environment? Case study of marine mucilage in the Sea of Marmara. *Environmental science and pollution research international*, v. 30, n. 28, p. 72821–72831, 2023.

AKINWANDE, M. O.; DIKKO, H. G.; SAMSON, A. Variance inflation factor: As a condition for the inclusion of suppressor variable(s) in regression analysis. *Open journal of statistics*, v. 05, n. 07, p. 754–767, 2015.

ALVES, M. T. R. *et al.* A snapshot of the limnological features in tropical floodplain lakes: the relative influence of climate and land use. *Acta Limnologica Brasiliensia*, v. 31, p. e10, 2019.

ANDREWS, S. FastQC: a quality control tool for high throughput sequence data. *Babraham Bioinformatics*, 2010. Disponível em: <<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>>. Acesso em: 23 de março de 2023.

APHA (AMERICAN PUBLIC HEALTH ASSOCIATION). *Standard Methods for the examination of water and wastewater*. 24^a Ed. Washington, DC: APHA/WEF/AWWA, 2023. p.1516.

AQUINO, S.; LATRUBESSE, E. M.; SOUZA-FILHO, E. E. Caracterização hidrológica e geomorfológica dos afluentes da Bacia do Rio Araguaia. *Revista Brasileira de Geomorfologia*, v. 10, n. 1, 2009.

AQUINO, S.; LATRUBESSE, E. M.; SOUZA-FILHO, E. E. Relações entre o regime hidrológico e os ecossistemas aquáticos da planície aluvial do rio Araguaia. *Acta Scientiarum. Biological Sciences*, v. 30, n. 4, p. 361-369, 2008.

AQUINO, S.; STEVAUX, J. C.; LATRUBESSE, Edgardo Manuel. Regime hidrológico e aspectos do comportamento morfohidráulico do rio Araguaia. *Revista brasileira de geomorfologia*, v. 6, n. 2, 2005.

ASTORGA, A. *et al.* Distance decay of similarity in freshwater communities: do macro- and microorganisms follow the same rules? *Global ecology and biogeography: a journal of macroecology*, v. 21, n. 3, p. 365–375, 2012.

BAGHERI, H.; SEVERIN, A. J.; RAJAN, H. Detecting and correcting misclassified sequences in the large-scale public databases. *Bioinformatics (Oxford, England)*, v. 36, n. 18, p. 4699–4705, 2020.

BANERJI, A. *et al.* Spatial and temporal dynamics of a freshwater eukaryotic plankton community revealed via 18S rRNA gene metabarcoding. *Hydrobiologia*, v. 818, n. 1, p. 71–86, 2018.

BARBERÁN, A.; CASAMAYOR, E. O. Global phylogenetic community structure and β -diversity patterns in surface bacterioplankton metacommunities. *Aquatic microbial ecology: international journal*, v. 59, p. 1–10, 2010.

BASELGA, A. *et al.* betapart: Partitioning Beta Diversity into Turnover and Nestedness Components, 2023. R package version 1.6. Disponível em: <<https://CRAN.R-project.org/package=betapart>>. Acesso em: 05 de maio de 2024.

BLANCHET, F. G.; LEGENDRE, P.; BORCARD, D. Forward selection of explanatory variables. *Ecology*, v. 89, n. 9, p. 2623–2632, 2008a.

BLANCHET, F. G.; LEGENDRE, P.; BORCARD, D. Modelling directional spatial processes in ecological data. *Ecological modelling*, v. 215, n. 4, p. 325–336, 2008b.

BLANCHET, F.G.; LEGENDRE, P.; GAUTHIER, O. AEM: Tools to construct Asymmetric e igenvector maps (AEM) spatial variables, 2015. R package version 0.6/r127. Disponível em: <<https://R-Forge.R-project.org/projects/sedar/>>. Acesso em: 10 de junho de 2024.

- BOCK, C. *et al.* Phylogenetic and functional diversity of Chrysophyceae in inland waters. *Organisms, diversity & evolution*, v. 22, n. 2, p. 327–341, 2022.
- BORCARD, D.; LEGENDRE, P. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological modelling*, v. 153, n. 1-2, p. 51-68, 2002.
- BORTOLINI, J. C. *et al.* Environmental and spatial processes influencing phytoplankton biomass along a reservoirs-river-floodplain lakes gradient: A metacommunity approach. *Freshwater biology*, v. 62, n. 10, p. 1756–1767, 2017.
- BROWN, J. W.; SORHANNUS, U. A molecular genetic timescale for the diversification of autotrophic Stramenopiles (Ochrophyta): Substantive underestimation of putative fossil ages. *PloS one*, v. 5, n. 9, p. e12759, 2010.
- BRYANT, J. A. *et al.* Microbes on mountainsides: contrasting elevational patterns of bacterial and plant diversity. *Proceedings of the National Academy of Sciences*, v. 105, n. supplement_1, p. 11505-11511, 2008.
- BURKI, F. *et al.* Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. *Proceedings. Biological sciences*, v. 283, n. 1823, p. 20152802, 2016.
- BURNS, J. H.; STRAUSS, S. Y. More closely related species are more ecologically similar in an experimental test. *Proceedings of the National Academy of Sciences of the United States of America*, v. 108, n. 13, p. 5302–5307, 2011.
- CARDILLO, M. Phylogenetic diversity in conservation: A brief history, critical overview, and challenges to progress. *Cambridge Prisms: Extinction*, v. 1, n. e11, 2023.

CARR, M. *et al.* Molecular phylogeny of choanoflagellates, the sister group to Metazoa. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, v. 105, n. 43, p. 16641–16646, 2008.

CAVALIER-SMITH, T.; CHAO, E. E.-Y. Phylogeny and classification of phylum Cercozoa (protozoa). *Protist*, v. 154, n. 3–4, p. 341–358, 2003.

CAVENDER-BARES, J. *et al.* Phylogenetic overdispersion in Floridian oak communities. *The American naturalist*, v. 163, n. 6, p. 823–843, 2004.

CAVENDER-BARES, J. *et al.* The merging of community ecology and phylogenetic biology. *Ecology letters*, v. 12, n. 7, p. 693–715, 2009.

CAVENDER-BARES, J.; KEEN, A.; MILES, B. Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. *Ecology*, v. 87, n. sp7, p. S109–S122, 2006.

CHEN, Y. *et al.* Bacterial communities exhibit apparent phosphate concentration-related patterns of community composition, alpha diversity, and phylogenetic structure in the subtropical Daya Bay. *Frontiers in marine science*, v. 9, 2023.

CHOI, C. J. *et al.* Seasonal and geographical transitions in eukaryotic phytoplankton community structure in the Atlantic and Pacific oceans. *Frontiers in microbiology*, v. 11, p. 11:542372, 2020.

CLARKE, L. J. *et al.* A globally distributed Syndiniales parasite dominates the Southern Ocean micro-eukaryote community near the sea-ice edge. *The ISME journal*, v. 13, n. 3, p. 734–737, 2019.

CORBELLI, J. M. *et al.* Integrating taxonomic, functional and phylogenetic beta diversities: Interactive effects with the biome and land use across taxa. *PloS one*, v. 10, n. 5, p. e0126854, 2015.

CORMACI, M.; FURNARI, G.; ALONGI, G. Marine benthic flora of the Mediterranean Sea: Chlorophyta. Bulletin of the Gioenia Academy of Natural Sciences of Catania, v. 47, n. 377, p. FP11-FP436, 2014.

COTTENIE, K. *et al.* Zooplankton metacommunity structure: Regional vs. Local processes in highly interconnected ponds. Ecology, v. 84, n. 4, p. 991–1000, 2003.

COTTENIE, K. Integrating environmental and spatial processes in ecological community dynamics. Ecology letters, v. 8, n. 11, p. 1175–1182, 2005.

DARU, B. H. *et al.* Understanding the processes underpinning patterns of phylogenetic regionalization. Trends in ecology & evolution, v. 32, n. 11, p. 845–860, 2017.

DAVID, G. M. *et al.* Environmental drivers of plankton protist communities along latitudinal and vertical gradients in the oldest and deepest freshwater lake. Environmental microbiology, v. 23, n. 3, p. 1436–1451, 2021.

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.

DECLERCK, S. A. J. *et al.* Scale dependency of processes structuring metacommunities of cladocerans in temporary pools of High-Andes wetlands. Ecography, v. 34, n. 2, p. 296–305, 2011.

DINIZ-FILHO, J. A. F. *et al.* Environmental DNA and biodiversity patterns: a call for a community phylogenetics approach. Perspectives in ecology and conservation, v. 22, n. 1, p. 15–23, 2024.

DJURHUUS, A. *et al.* Environmental DNA reveals seasonal shifts and potential interactions in a marine community. Nature Communications, v. 11, p. 254, 2020.

FAITH, D. P. Conservation evaluation and phylogenetic diversity. Biological conservation, v. 61, n. 1, p. 1-10, 1992.

FAITH, D. P. The PD phylogenetic diversity framework: Linking evolutionary history to feature diversity for biodiversity conservation. In: PELLENS, R.; GRANDCOLAS, P. (eds) Biodiversity Conservation and Phylogenetic Systematics. Topics in Biodiversity and Conservation, v. 14. Cham: Springer International Publishing, 2016. p. 39–56.

FERNANDES, I. O. *et al.* Mercury distribution in organisms, litter, and soils of the Middle Araguaia floodplain in Brazil. Environmental science and pollution research international, v. 31, n. 14, p. 20925–20940, 2024.

FIELD, C. B. *et al.* Primary production of the biosphere: integrating terrestrial and oceanic components. science, v. 281, n. 5374, p. 237-240, 1998.

FOREST, F. *et al.* Preserving the evolutionary potential of floras in biodiversity hotspots. Nature, v. 445, n. 7129, p. 757–760, 2007.

FURTADO, J. A. *et al.* Desenvolvimento de primers in silico para eDNA como ferramenta de monitoramento da comunidade fitoplanctônica em um Complexo Estuarino na Costa Amazônica. Observatório de la Economía Latinoamericana, v. 22, n. 7, p. e5905, 2024.

GABRIELSON, P. W.; GARBARY, D.; HOMMERSAND, M. H. Systematics of red algae (Rhodophyta). Critical reviews in plant sciences, v. 3, n. 4, p. 325–366, 1986.

GARCÍA, A. Charophyta: their use in paleolimnology. Journal of paleolimnology, v. 10, n. 1, p. 43–52, 1994.

GIANUCA, A. T.; PANTEL, J. H.; DE MEESTER, L. Disentangling the effect of body size and phylogenetic distances on zooplankton top-down control of algae. Proceedings. Biological sciences, v. 283, n. 1828, p. 20160487, 2016.

GLOCKLING, S. L.; MARSHALL, W. L.; GLEASON, F. H. Phylogenetic interpretations and ecological potentials of the Mesomycetozoea (Ichthyosporae). Fungal ecology, v. 6, n. 4, p. 237–247, 2013.

GOLTERMAN, H. L.; CLYMO, R. S.; OHMSTAD, M. A. M. Methods for physical and chemical analysis of fresh waters. IBP Handbook, v. 8, p. 215, 1978.

GRÖNROOS, M. *et al.* Metacommunity structuring in stream networks: roles of dispersal mode, distance type, and regional environmental context. *Ecology and evolution*, v. 3, n. 13, p. 4473–4487, 2013.

GROSSART, H. P. *et al.* Linking metagenomics to aquatic microbial ecology and biogeochemical cycles. *Limnology and oceanography*, v. 65, n. S1, p. S2-S20, 2020.

HAN, D. *et al.* Phylogenetic diversity and spatiotemporal dynamics of bacterial and microeukaryotic plankton communities in Gwangyang Bay of the Korean Peninsula. *Scientific reports*, v. 12, n. 1, p. 2980, 2022.

HAN, D. *et al.* Survey of Bacterial Phylogenetic Diversity During the Glacier Melting Season in an Arctic Fjord. *Microbial ecology*, v. 81, p. 579–591, 2021.

HANŽEK, N. *et al.* Assessing ecological status in karstic lakes through the integration of phytoplankton functional groups, morphological approach and environmental DNA metabarcoding. *Ecological Indicators*, v. 131, p. 108166, 2021.

HARDIKAR, R.; HARIDEVI, C. K.; DESHBHRATAR, S. Phytoplankton community indices in assessing the ecological quality status (EcoQS) of a coastal ecosystem (Thane Creek, India): Mono-metric vs multi-metric approach. *Estuarine, Coastal and Shelf Science*, v. 274, p. 107922, 2022.

HECKENHAUER, J. *et al.* Plant DNA barcodes and assessment of phylogenetic community structure of a tropical mixed dipterocarp forest in Brunei Darussalam (Borneo). *PloS one*, v. 12, n. 10, p. e0185861, 2017.

HEINO, J. *et al.* Geographical patterns of micro-organismal community structure: are diatoms ubiquitously distributed across boreal streams? *Oikos* (Copenhagen, Denmark), v. 119, n. 1, p. 129–137, 2010.

HEINO, J. *et al.* Metacommunity organisation, spatial extent and dispersal in aquatic systems: patterns, processes and prospects. *Freshwater biology*, v. 60, n. 5, p. 845–869, 2015.

HEINO, J. *et al.* Microbial diversity and community–environment relationships in boreal streams. *Journal of Biogeography*, v. 41, n. 12, p. 2234–2244, 2014.

HELMUS, M. R. *et al.* Phylogenetic measures of biodiversity. *The American naturalist*, v. 169, n. 3, p. E68–83, 2007.

HENSON, S. A. *et al.* Future phytoplankton diversity in a changing climate. *Nature communications*, v. 12, n. 1, p. 1–8, 2021.

HESTETUN, J. T. *et al.* Significant taxon sampling gaps in DNA databases limit the operational use of marine macrofauna metabarcoding. *Marine biodiversity*, v. 50, n. 5, p. 70, 2020.

HILLIS, D. M. *et al.* Is sparse taxon sampling a problem for phylogenetic inference? *Systematic biology*, v. 52, n. 1, p. 124–126, 2003.

HOANG, D. T. *et al.* UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular biology and evolution*, v. 35, n. 2, p. 518–522, 2018.

HOBÆK, A.; MANCA, M.; ANDERSEN, T. Factors influencing species richness in lacustrine zooplankton. *Acta oecologica* (Montrouge, France), v. 23, n. 3, p. 155–163, 2002.

HORNER-DEVINE, M. C.; BOHANNAN, B. J. M. Phylogenetic clustering and overdispersion in bacterial communities. *Ecology*, v. 87, n. sp7, p. S100–S108, 2006.

HUANG, Z. *et al.* Seasonal variation of phytoplankton community assembly processes in Tibetan Plateau floodplain. *Frontiers in microbiology*, v. 14, p. 14:1122838, 2023.

HUBBELL, S. P. The unified neutral theory of biodiversity and biogeography. *Monographs in Population Biology*, v. 32, 2001.

HUSZAR, V. L. M. *et al.* Environmental and not spatial processes (directional and non-directional) shape the phytoplankton composition and functional groups in a large subtropical river basin. *Journal of plankton research*, p. fbv084, 2015.

IRION, G. *et al.* Araguaia River floodplain: Size, age, and mineral composition of a large tropical Savanna wetland. *Wetlands (Wilmington, N.C.)*, v. 36, n. 5, p. 945–956, 2016.

ISLAM, M. T.; HOSSAIN, M. M. Biological control of peronosporomycete phytopathogen by bacterial antagonist. Em: *Bacteria in Agrobiolgy: Disease Management*. Berlin, Heidelberg: Springer Berlin Heidelberg, p. 167–218, 2013.

IVES, A. R.; HELMUS, M. R. Generalized linear mixed models for phylogenetic analyses of community structure. *Ecological Monographs*, v. 81, n. 3, p. 511-525, 2011.

IZAGUIRRE, I. *et al.* Drivers of phytoplankton diversity in Patagonian and Antarctic lakes across a latitudinal gradient (2150 km): the importance of spatial and environmental factors. *Hydrobiologia*, v. 764, n. 1, p. 157–170, 2016.

JAMONEAU, A. *et al.* Beta diversity of diatom species and ecological guilds: Response to environmental and spatial mechanisms along the stream watercourse. *Freshwater biology*, v. 63, n. 1, p. 62–73, 2018.

JEON, B. S.; PARK, M. G. A Novel Parasitoid of Marine Dinoflagellates, *Pararosarium dinoexitiosum* gen. et sp. nov. (Perkinsozoa, Alveolata), Showing Characteristic Beaded Sporocytes. *Frontiers in microbiology*, v. 12, p.748092, 2021.

JIN, L. *et al.* The scale-dependence of spatial distribution of reservoir plankton communities in subtropical and tropical China. *The Science of the total environment*, v. 845, n. 157179, p. 157179, 2022.

KALYAANAMOORTHY, S. *et al.* ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature methods*, v. 14, n. 6, p. 587–589, 2017.

KATOH, K.; STANDLEY, D. M. MAFFT: multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, v. 30, n. 4, p. 772-780, 2013.

KECK, F.; KAHLERT, M. Community phylogenetic structure reveals the imprint of dispersal-related dynamics and environmental filtering by nutrient availability in freshwater diatoms. *Scientific reports*, v. 9, n. 1, p. 11590, 2019.

KEMBEL, S. W.; HUBBELL, S. P. The phylogenetic structure of a neotropical forest tree community. *Ecology*, v. 87, n. sp7, p. S86–S99, 2006.

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

KEZLYA, E.; TSEPLIK, N.; KULIKOVSKIY, M. Genetic markers for metabarcoding of freshwater microalgae: Review. *Biology*, v. 12, n. 7, 2023.

KRAFT, N. J. B. *et al.* Community assembly, coexistence and the environmental filtering metaphor. *Functional ecology*, v. 29, n. 5, p. 592–599, 2015.

KRAFT, N. J. B. *et al.* Trait evolution, community assembly, and the phylogenetic structure of ecological communities. *The American naturalist*, v. 170, n. 2, p. 271–283, 2007.

KRESS, W. J. *et al.* Advances in the use of DNA barcodes to build a community phylogeny for tropical trees in a Puerto rican forest dynamics plot. *PloS one*, v. 5, n. 11, p. e15409, 2010.

LACDAN, N. F. *et al.* Assessing water quality of Dao River, Batangas using phytoplankton biomonitoring. *International Journal of Current Science*, n. 12, p. 98-102, 2014.

LANDEIRO, V. L. *et al.* Spatial eigenfunction analyses in stream networks: do watercourse and overland distances produce different results?: Watercourse or overland distances? *Freshwater biology*, v. 56, n. 6, p. 1184–1192, 2011.

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, 2021.

LATRUBESSE, E. M.; STEVAUX, J. C. Características físico-bióticas e problemas ambientais associados à planície aluvial do Rio Araguaia, Brasil Central. *Revista Geociências-UNG-Ser*, v. 5, n. 1, p. 65-73, 2006.

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.

LETUNIC, I.; BORK, P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic acids research*, v. 52, n. W1, p. W78–W82, 2024.

LI, D.; IVES, A. R.; WALLER, D. M. Can functional traits account for phylogenetic signal in community composition?. *New Phytologist*, v. 214, n. 2, p. 607-618, 2017.

LI, F.; TONKIN, J. D.; HAASE, P. Local contribution to beta diversity is negatively linked with community-wide dispersal capacity in stream invertebrate communities. *Ecological indicators*, v. 108, n. 105715, p. 105715, 2020.

LI, T. *et al.* Eukaryotic plankton community assembly and influencing factors between continental shelf and slope sites in the northern South China Sea. *Environmental research*, v. 216, n. 114584, p. 114584, 2023.

LITCHMAN, E. *et al.* Global biogeochemical impacts of phytoplankton: a trait-based perspective. *Journal of ecology*, v. 103, n. 6, p. 1384-1396, 2015.

LIU, J. *et al.* Effects of connectivity, dispersal directionality and functional traits on the metacommunity structure of river benthic diatoms. *Journal of biogeography*, v. 40, n. 12, p. 2238–2248, 2013.

LIU, J. *et al.* The use of DNA barcodes to estimate phylogenetic diversity in forest communities of southern China. *Ecology and evolution*, v. 9, n. 9, p. 5372–5379, 2019.

LIU, S. *et al.* Composition and spatial-temporal dynamics of phytoplankton community shaped by environmental selection and interactions in the Jiaozhou Bay. *Water research*, v. 218, n. 118488, p. 118488, 2022.

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.

LOGUE, J. B. *et al.* Empirical approaches to metacommunities: a review and comparison with theory. *Trends in ecology & evolution*, v. 26, n. 9, p. 482–491, 2011.

LONGMIRE, J. L. *et al.* Use of “Lysis Buffer” in DNA isolation and its implication for museum collections. Lubbock, TX: Museum of Texas Tech University, 1997.

LOVETTE, I. J.; HOCHACHKA, W. M. Simultaneous effects of phylogenetic niche conservatism and competition on avian community structure. *Ecology*, v. 87, n. sp7, p. S14–S28, 2006.

LYNN, D. H. Ciliophora. In: *Encyclopedia of Life Sciences*. Chichester: John Wiley & Sons, 2012. p. 1-12. Disponível em: <<http://dx.doi.org/10.1002/9780470015902.a0001966.pub3>>. Acesso em: 19 de agosto de 2024.

MACHADO, K. B. *et al.* Comparing the effects of landscape and local environmental variables on taxonomic and functional composition of phytoplankton communities. *Journal of Plankton Research*, v. 38, n. 5, p. 1334-1346, 2016.

MACHADO, K. B. *et al.* Diversity patterns of planktonic microeukaryote communities in tropical floodplain lakes based on 18S rDNA gene sequences. *Journal of Plankton Research*, v. 41, n. 3, p. 241-256, 2019.

MACÍAS-HERNÁNDEZ, N. *et al.* Building a robust, densely-sampled spider tree of life for ecosystem research. *Diversity*, v. 12, n. 8, p. 288, 2020.

MARCIONÍLIO, S. M. L. O. *et al.* Environmental factors affecting chlorophyll-a concentration in tropical floodplain lakes, Central Brazil. *Environmental Monitoring and Assessment*, v. 188, p. 1-9, 2016.

MARTIN, G. K. *et al.* Freshwater zooplankton metapopulations and metacommunities respond differently to environmental and spatial variation. *Ecology*, v. 102, n. 1, 2021.

MARTINS, P. R. *et al.* Terrain units, land use and land cover, and gross primary productivity of the largest fluvial basin in the Brazilian Amazonia/Cerrado ecotone: the Araguaia River basin. *Applied Geography*, v. 127, p. 102379, 2021.

MAZEL, F. *et al.* Prioritizing phylogenetic diversity captures functional diversity unreliably. *Nature communications*, v. 9, n. 1, p. 2888, 2018.

MCMURDIE, P. J.; HOLMES, S. *Phyloseq*: An R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one*, v. 8, n. 4, p. e61217, 2013.

MINH, B.Q. *et al.* IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.*, v. 37, p. 1530-1534, 2020.

MO, Y. *et al.* Low shifts in salinity determined assembly processes and network stability of microeukaryotic plankton communities in a subtropical urban reservoir. *Microbiome*, v. 9, n. 1, 2021.

MONTEIRO, L. C. *et al.* Local and landscape factors influencing mercury distribution in water, bottom sediment, and biota from lakes of the Araguaia River floodplain, Central Brazil. *The Science of the total environment*, v. 908, n. 168336, p. 168336, 2024.

MORRISON, D. A. Evolution of the Apicomplexa: where are we now? *Trends in parasitology*, v. 25, n. 8, p. 375–382, 2009.

MYLNIKOV, A. P. Ultrastructure and phylogeny of colpodellids (Colpodellida, Alveolata). *Biology bulletin of the Russian Academy of Sciences*, v. 36, n. 6, p. 582–590, 2009.

MYLNIKOV, A. P.; TIKHONENKOV, D. V. A new species of soil predatory flagellate, *Colponema edaphicum* sp. n., from Vorontsovskaya Cave, North Caucasus (Protista, Alveolata: Colponemidae). *Zoosystematica Rossica*, v. 16, n. 1, p. 1-4, 2007.

NABOUT, J. C. *et al.* No evidence for environmental and spatial processes in structuring phytoplankton communities. *Acta oecologica (Montrouge, France)*, v. 35, n. 5, p. 720–726, 2009.

NABOUT, J. C. *et al.* Phytoplankton diversity (alpha, beta, and gamma) from the Araguaia River tropical floodplain lakes (central Brazil). *Hydrobiologia*, v. 575, n. 1, p. 455–461, 2007.

NABOUT, J. C.; NOGUEIRA, I. de S.; OLIVEIRA, L. G. Phytoplankton community of floodplain lakes of the Araguaia River, Brazil, in the rainy and dry seasons. *Journal of Plankton Research*, v. 28, n. 2, p. 181-193, 2006.

OKSANEN, J. *et al.* Package ‘vegan’. Community ecology package. Versão do pacote R 2.6-4., v. 2, n. 9, p. 1-295, 2022.

OKUNO, S. *et al.* Community phylogeny and spatial scale affect phylogenetic diversity metrics in a species-rich rainforest in Borneo. *Ecology and evolution*, v. 12, n. 11, 2022.

OLIVEIRA, R. R. M. *et al.* PIMBA: A Pipeline for MetaBarcoding Analysis. In: STADLER, P.F., WALTER, M.E.M.T., HERNANDEZ-ROSALES, M., BRIGIDO, M.M. (eds) *Advances*

in Bioinformatics and Computational Biology. BSB 2021. Lecture Notes in Computer Science. 1^a Ed, v. 13063. Switzerland: Springer, 2021. p. 106-116.

PADIAL, A. A. *et al.* Dispersal ability determines the role of environmental, spatial and temporal drivers of metacommunity structure. *PloS one*, v. 9, n. 10, p. e111227, 2014.

PARADIS, E.; SCHLIEP, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, v. 35, p. 526–528, 2019.

PARK, D. S.; WORTHINGTON, S.; XI, Z. Taxon sampling effects on the quantification and comparison of community phylogenetic diversity. *Molecular ecology*, v. 27, n. 5, p. 1296–1308, 2018.

PASCHOALINI, M. *et al.* On the brink of isolation: Population estimates of the Araguaia River dolphin in a human-impacted region in Brazil. *Plos One*, v. 15, n. 4, p. e0231224, 2020.

PAUSAS, J. G.; VERDÚ, M. The Jungle of methods for evaluating phenotypic and phylogenetic structure of communities. *Bioscience*, v. 60, n. 8, p. 614–625, 2010.

PELICICE, F. M. *et al.* Large-scale degradation of the Tocantins-Araguaia River basin. *Environmental Management*, v. 68, p. 445-452, 2021.

PENNA, A. *et al.* Analysis of phytoplankton assemblage structure in the Mediterranean Sea based on high-throughput sequencing of partial 18S rRNA sequences. *Marine genomics*, v. 36, p. 49–55, 2017.

PEREIRA, A. C.; COLLI, G. R. Landscape features affect caiman body condition in the middle Araguaia River floodplain. *Animal Conservation*, v. 26, n. 4, p. 531-545, 2023.

PIO, D. V. *et al.* Spatial predictions of phylogenetic diversity in conservation decision making: Patterns of phylogenetic diversity. *Conservation biology: the journal of the Society for Conservation Biology*, v. 25, n. 6, p. 1229–1239, 2011.

POFF, N. L. Landscape filters and species traits: Towards mechanistic understanding and prediction in stream ecology. *Journal of the North American Benthological Society*, v. 16, n. 2, p. 391–409, 1997.

POLASKY, S. *et al.* A comparison of taxonomic distinctness versus richness as criteria for setting conservation priorities for North American birds. *Biological conservation*, v. 97, n. 1, p. 99–105, 2001.

PONTARP, M. *et al.* Phylogenetic analysis suggests that habitat filtering is structuring marine bacterial communities across the globe. *Microbial ecology*, v. 64, n. 1, p. 8–17, 2012.

PRICE, D. C. *et al.* Glaucophyta. In: Archibald, J., Simpson, A., Slamovits, C. (eds) *Handbook of the Protists*. Cham: Springer International Publishing, p. 23–87, 2017.

QU, Y. *et al.* Riverine phytoplankton shifting along a lentic-lotic continuum under hydrological, physiochemical conditions and species dispersal. *The Science of the total environment*, v. 619–620, p. 1628–1636, 2018.

QUAST, C. *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* v. 41, p. 1–7, 2013.

R CORE TEAM. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing, 2023. Disponível em: <<https://www.R-project.org/>>. Acessado em: 10 de agosto de 2023.

RAMBAUT, A. FigTree. Tree figure drawing tool. 2009. Disponível em: <<http://tree.bio.ed.ac.uk/software/figtree/>>. Acessado em: 15 de abril de 2024.

RAMULU, H. G.; RAOULT, D.; PONTAROTTI, P. The rhizome of life: what about metazoa? *Frontiers in cellular and infection microbiology*, v. 2, 2012.

REN, L. *et al.* pH influences the importance of niche-related and neutral processes in lacustrine bacterioplankton assembly. *Applied and Environmental Microbiology*, v. 81, n. 9, p. 3104-3114, 2015.

REVELL, L. J. Phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). *PeerJ*, v. 12, p. e16505, 2024.

RICKLEFS, R. E. Community Diversity: Relative Roles of Local and Regional Processes. *Science*, v. 235, n. 4785, p. 167–171, 1987.

RODRIGUES, A. S. L.; BROOKS, T. M.; GASTON, K. J. Integrating phylogenetic diversity in the selection of priority areas for conservation: does it make a difference? In: PURVIS, A.; GITTLEMAN, J. L.; BROOKS, T. (Eds.). *Phylogeny and Conservation*. Conservation Biology. Cambridge: Cambridge University Press, 2005. p. 101–119.

RUPPERT, K. M.; KLINE, R. J.; RAHMAN, M. S. Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Global Ecology and Conservation*, v. 17, p. e00547, 2019.

SAHU, A. *et al.* Environmental DNA (eDNA): Powerful Technique for Biodiversity Conservation. *Journal for Nature Conservation*, v. 71, p. 126325, 2023.

SANDEL, B. Richness-dependence of phylogenetic diversity indices. *Ecography*, v. 41, n. 5, p. 837–844, 2018.

SANTOFERRARA, L. F. Current practice in plankton metabarcoding: optimization and error management. *Journal of Plankton Research*, v. 41, n. 5, p. 571-582, 2019.

SARD, N. M. *et al.* Comparison of fish detections, community diversity, and relative abundance using environmental DNA metabarcoding and traditional gears. *Environmental DNA*, v. 1, n. 4, p. 368-384, 2019.

- SAWAYA, N. A. *et al.* Assessing eukaryotic biodiversity in the Florida Keys National Marine Sanctuary through environmental DNA metabarcoding. *Ecology and Evolution*, v. 9, n. 3, p. 1029-1040, 2019.
- SCHILDE, C. Amoebozoa. In: SCHMIDT, T. M. (Eds) *Encyclopedia of Microbiology*. 4 Ed., Academic Press, p. 99-106, 2019.
- SHEN, Y. *et al.* Environmental DNA metabarcoding revealing the distinct responses of phytoplankton and zooplankton to cascade dams along a river-way. *Ecological indicators*, v. 166, n. 112545, p. 112545, 2024.
- SILVA, L. B. *et al.* Spatial structure of phytoplankton metacommunities and their functional groups in a neotropical floodplain. *Biota neotropica*, v. 24, n. 2, p. e20231553, 2024.
- SIMON, M. *et al.* Marked seasonality and high spatial variability of protist communities in shallow freshwater systems. *The ISME journal*, v. 9, n. 9, p. 1941–1953, 2015.
- SKOUROLIAKOU, D.-I. *et al.* Stochastic and deterministic processes regulate phytoplankton assemblages in a temperate coastal ecosystem. *Microbiology spectrum*, v. 10, n. 6, 2022.
- SMETACEK, V.; CLOERN, J. E. On phytoplankton trends. *Science*, v. 319, n. 5868, p. 1346-1348, 2008.
- SOININEN, J. *et al.* Neutrality, niches, and determinants of plankton metacommunity structure across boreal wetland ponds. *Ecoscience*, v. 14, n. 2, p. 146–154, 2007.
- SOLAK, C. N. *et al.* A review of investigations on diatoms (Bacillariophyta) in Turkish inland waters. *Nova Hedwigia, Beiheft*, v. 141, p. 431-462, 2012.
- STAT, M. *et al.* Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. *Scientific Reports*, v. 7, n. a. 12240, 2017.

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.

STOECKER, D. K. Mixotrophy among Dinoflagellates¹. *The journal of eukaryotic microbiology*, v. 46, n. 4, p. 397–401, 1999.

SWENSON, N. G. *et al.* Phylogenetic and functional alpha and beta diversity in temperate and tropical tree communities. *Ecology*, v. 93, n. sp8, p. S112–S125, 2012.

SWENSON, N. G. *et al.* The influence of spatial and size scale on phylogenetic relatedness in tropical forest communities. *Ecology*, v. 88, n. 7, p. 1770–1780, 2007.

SWENSON, N. G. *et al.* The problem and promise of scale dependency in community phylogenetics. *Ecology*, v. 87, n. 10, p. 2418–2424, 2006.

SWENSON, N. G. Phylogenetic resolution and quantifying the phylogenetic diversity and dispersion of communities. *PloS one*, v. 4, n. 2, p. e4390, 2009.

TABERLET, P. *et al.* Environmental DNA. *Molecular ecology*, v. 21, n. 8, p. 1789-1793, 2012.

TEJERINA-GARRO, F. L.; FORTIN, R.; RODRÍGUEZ, M. A. Fish community structure in relation to environmental variation in floodplain lakes of the Araguaia River, Amazon Basin. *Environmental Biology of fishes*, v. 51, p. 399-410, 1998.

THOMAZ, S. M.; BINI, L. M.; BOZELLI, R. L. Floods increase similarity among aquatic habitats in river-floodplain systems. *Hydrobiologia*, v. 579, n. 1, p. 1–13, 2007.

THOMPSON, P. L. *et al.* A process-based metacommunity framework linking local and regional scale community ecology. *Ecology letters*, v. 23, n. 9, p. 1314–1329, 2020.

THOMPSON, P. L.; DAVIES, T. J.; GONZALEZ, A. Ecosystem functions across trophic levels are linked to functional and phylogenetic diversity. *PloS one*, v. 10, n. 2, p. e0117595, 2015.

TONKIN, J. D. *et al.* The role of dispersal in river network metacommunities: Patterns, processes, and pathways. *Freshwater biology*, v. 63, n. 1, p. 141–163, 2018.

TUCKER, C. M. *et al.* A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological reviews of the Cambridge Philosophical Society*, v. 92, n. 2, p. 698–715, 2017.

TUCKER, C. M. *et al.* Assessing the utility of conserving evolutionary history. *Biological reviews of the Cambridge Philosophical Society*, v. 94, n. 5, p. 1740–1760, 2019.

TUCKER, C. M.; CADOTTE, M. W. Unifying measures of biodiversity: understanding when richness and phylogenetic diversity should be congruent. *Diversity & distributions*, v. 19, n. 7, p. 845–854, 2013.

VAJRAVELU, M. *et al.* Seasonal influence of physico-chemical parameters on phytoplankton diversity, community structure and abundance at Parangipettai coastal waters, Bay of Bengal, South East Coast of India. *Oceanologia*, v. 60, n. 2, p. 114-127, 2018.

VALENTE, C. R.; LATRUBESSE, E. M.; FERREIRA, L. G. Relationships among vegetation, geomorphology and hydrology in the Bananal Island tropical wetlands, Araguaia River basin, Central Brazil. *Journal of South American Earth Sciences*, v. 46, p. 150-160, 2013.

VALENTINI, A. *et al.* Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular ecology*, v. 25, n. 4, p. 929-942, 2015.

VAMOSI, S. M. *et al.* Emerging patterns in the comparative analysis of phylogenetic community structure. *Molecular ecology*, v. 18, n. 4, p. 572–592, 2009.

VAMOSI, S. M. Phylogenetic community ecology as an approach for studying old ideas on competition in the plankton: opportunities and challenges. *Journal of limnology*, v. 73, n. s1, 2013.

VAN DER AUWERA, G. *et al.*, The phylogeny of the Hyphochytriomycota as deduced from ribosomal RNA sequences of *Hyphochytrium catenoides*., *Molecular Biology and Evolution*, Volume 12, Issue 4, p. 671–678, 1995.

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.

WANG, X. *et al.* Community organization and network complexity and stability: contrasting strategies of prokaryotic versus eukaryotic microbiomes in the Bohai Sea and Yellow Sea. *mSphere*, p. 0:e00395-24, 2024.

WEBB, C. O. *et al.* Phylogenies and community ecology. *Annual review of ecology and systematics*, v. 33, n. 1, p. 475-505, 2002.

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.

WOJCIECHOWSKI, J. *et al.* Temporal variation in phytoplankton beta diversity patterns and metacommunity structures across subtropical reservoirs. *Freshwater biology*, v. 62, n. 4, p. 751–766, 2017.

WRAIGHT, S. P.; INGLIS, G. D.; GOETTEL, M. S. Fungi. In: LACEY, L.A., KAYA, H.K. (eds) *Field Manual of Techniques in Invertebrate Pathology*. Dordrecht: Springer Netherlands, 2007. p. 223–248.

XIAO, Z. *et al.* Geographic pattern of phytoplankton community and their drivers in lakes of middle and lower reaches of Yangtze River floodplain, China. *Environmental science and pollution research international*, v. 29, n. 55, p. 83993–84005, 2022.

YANG, J. *et al.* Seasonal and spatial patterns of eukaryotic phytoplankton communities in an urban river based on marker gene. *Scientific reports*, v. 11, n. 1, 2021.

YI, Z. *et al.* Increasing taxon sampling using both unidentified environmental sequences and identified cultures improves phylogenetic inference in the Prorodontida (Ciliophora, Prostomatea). *Molecular phylogenetics and evolution*, v. 57, n. 2, p. 937–941, 2010.

ZHAO, S. *et al.* Collodictyon--an ancient lineage in the tree of eukaryotes. *Molecular biology and evolution*, v. 29, n. 6, p. 1557–1568, 2012.

APÊNDICES

Apêndice 1. Protocolo modificado do kit Dneasy® Blood & Tissue (Qiagen)

Notas antes de começar

- Efetuar todos os passos de centrifugação à temperatura ambiente (15-25°C).
- Redissolver quaisquer precipitados em Tampão AL e Tampão ATL.
- Adicionar etanol aos concentrados de Buffer AW1 e Buffer AW2.
- Equilibrar o tecido congelado ou os pellets de células à temperatura ambiente.
- Pré-aquecer uma incubadora a 65°C.*
- Consultar o manual para o pré-tratamento de tecido fixado, inseto, bactéria ou outro material.

1a. Tecido: Cortar o tecido (≤ 10 mg de baço ou ≤ 25 mg de outro tecido) em pedaços pequenos e colocar num tubo de microcentrifugação de 1,5 ml. Para caudas de roedores, utilizar 1 (rato) ou 2 (ratinho) caudas com 0,4-0,6 cm de comprimento. Adicionar 180 μ l de tampão ATL. Adicionar 20 μ l de proteinase K, misturar em vórtex e incubar a 56°C até à lise completa. Agitar ocasionalmente no vortex durante a incubação. Agitar em vórtice 15 s diretamente antes de passar ao passo 2.

***Modificação:** Adicionar beads para maceração mecânica e 380 μ l de tampão ATL (quando usar toda ou metade da membrana para a extração), adicionar 20 μ l de proteinase K, misturar em vórtex. Incubar a 65°C por 10 min.

1b. Sangue não nucleado: Pipetar 20 μ l de proteinase K para um tubo de microcentrifugação de 1,5 ml ou 2 ml. Adicionar 50-100 μ l de sangue tratado com anticoagulante. Ajustar o volume para 220 μ l com PBS. Avançar para o passo 2.

1c. Sangue nucleado: Pipetar 20 μ l de proteinase K para um tubo de microcentrifugação de 1,5 ml ou 2 ml. Adicionar 5-10 μ l de sangue tratado com anticoagulante. Ajustar o volume para 220 μ l com PBS. Avançar para o passo 2.

1d. Células cultivadas: Centrifugar um máximo de 5×10^6 células durante 5 minutos a 300 x g (190 rpm). Ressuspender em 200 μ l de PBS. Adicionar 20 μ l de proteinase K. Avançar para o passo 2.

2. Adicionar 200 μ l de tampão AL. Misturar bem, agitando em vórtice.

Para as amostras de sangue incubarem a 56°C durante 10 min.

Modificação: Adicionar 400 μ l de tampão AL, quando utilizar 380 μ l de tampão ATL.

3. Adicionar 200 μ l de etanol (96-100%). Misturar bem, agitando em vórtice.

Modificação: Adicionar 400 μ l de etanol, quando utilizar 380 μ l de tampão ATL.

4. Pipete a mistura para uma coluna de centrifugação DNeasy Mini spin colocada num tubo coletor de 2 ml. Centrifugar a $\geq 6000 \times g$ (8000 rpm) durante 1 min. Jogar fora o fluxo e o tubo coletor.

5. Coloque a coluna DNeasy Mini spin em um novo tubo coletor de 2 ml. Adicionar 500 μl de tampão AW1. Centrifugue durante 1 min a $\geq 6000 \times g$ (8000 rpm). Jogar fora o fluxo e o tubo coletor.

6. Coloque a coluna DNeasy Mini spin em um novo tubo coletor de 2 ml, adicione 500 μl de tampão AW2 e centrifugue durante 3 min a $20.000 \times g$ (14.000 rpm). Jogar fora o fluxo e o tubo coletor.

7. Transfira a coluna DNeasy Mini spin para um novo tubo de microcentrifugação de 1,5 ml ou 2 ml.

8. Eluir o DNA adicionando 200 μl de Buffer AE ao centro da membrana da coluna DNeasy Mini spin. Incubar durante 1 minuto à temperatura ambiente (15-25°C). Centrifugar durante 1 min a $\geq 6000 \times g$ (8000 rpm).

Modificação: Eluir o DNA adicionado 50 μl de água nuclease-free aquecida a 56° C. Incubar por 15 minutos a temperatura ambiente. Centrifugar durante 1 min a $\geq 6000 \times g$ (8000 rpm).

9. Opcional: Repetir o passo 8 para aumentar o rendimento de DNA.

Modificação: Fazer uma segunda eluição em um novo tubo e armazenar a primeira eluição.

OBS: Para algumas amostras 100 μl do tampão de armazenamento foram adicionados juntamente com as membranas para realização da extração. Essas amostras eram aquelas cujo tampão de Longmire usado no armazenamento das membranas continha muito material particulado suspenso, liberado das membranas durante o período de armazenamento.

Apêndice 2. Geis de agarose da extração de DNA dos filtrados das membranas de acetato de celulose utilizando o kit comercial DNeasy Blood and Tissue (Qiagen).

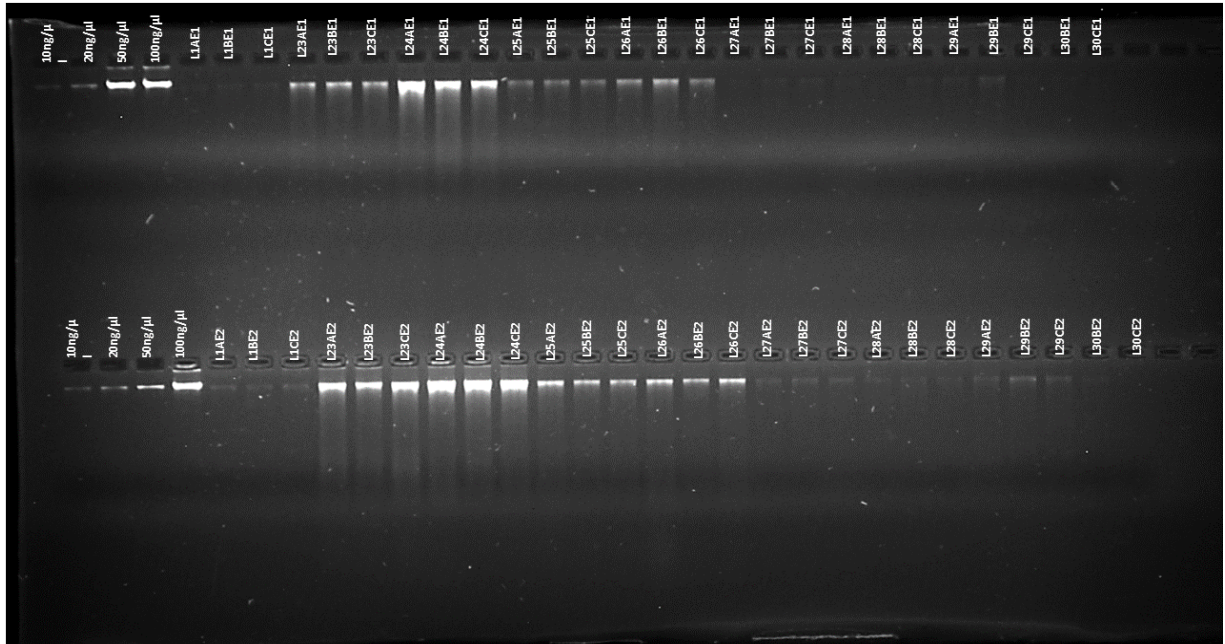


Figura 1S. Gel de agarose 1% corado com brometo de etídeo da extração de DNA das triplicatas dos lagos 01 e 23 a 30 com o protocolo de tecido modificado. No início de cada pente temos os marcadores moleculares de 10, 20, 50, 100 e 200 ng/μl.

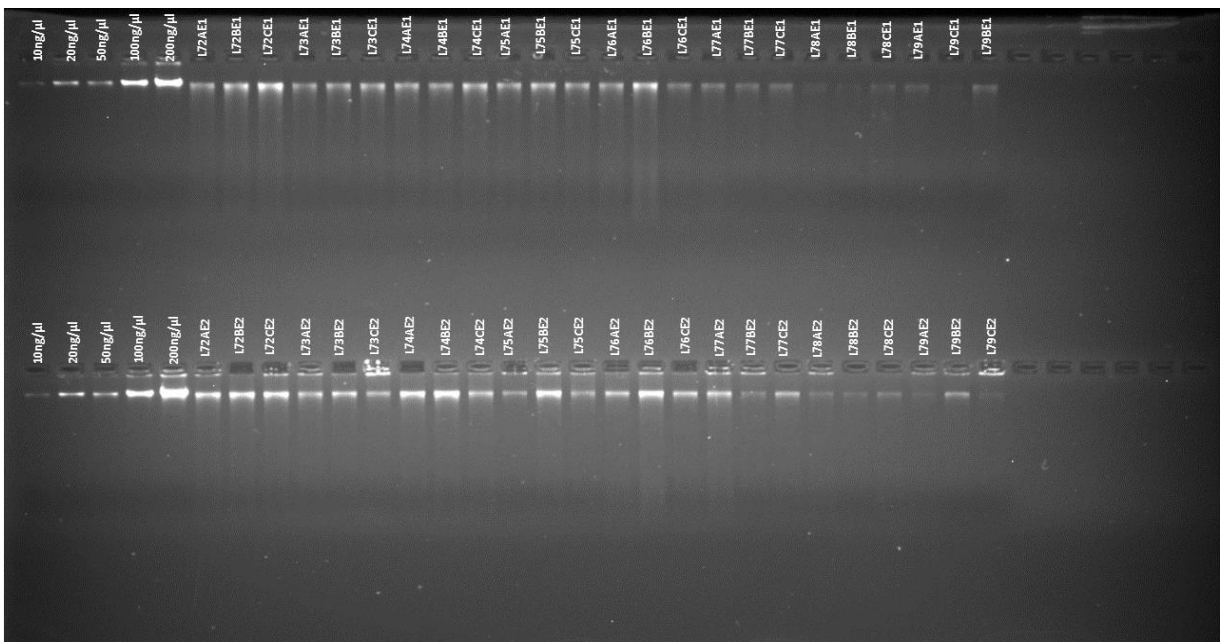


Figura 2S. Gel de agarose 1% corado com brometo de etídeo da extração de DNA das triplicatas dos lagos 72 a 79 com o protocolo de tecido modificado. No início de cada pente temos os marcadores moleculares de 10, 20, 50, 100 e 200 ng/μl.

Apêndice 3. Géis de amplificação do gene 18S-V4 através da PCR utilizando par de *primers* TAREuk454FWD1 e TAREukREV3 (Stoeck *et al.*, 2010).

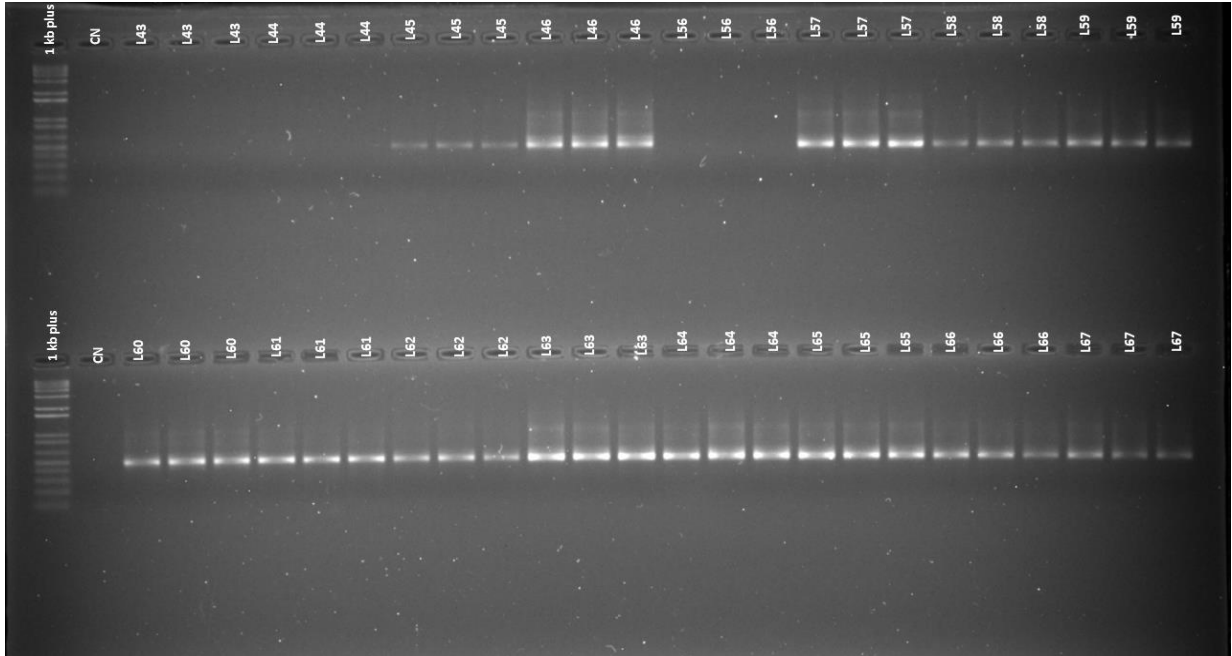


Figura 3S. Gel de agarose 1.5% corado com brometo de etídeo da amplificação do gene 18S-V4 dos pontos L43 a L67. Marcador 1kb plus.

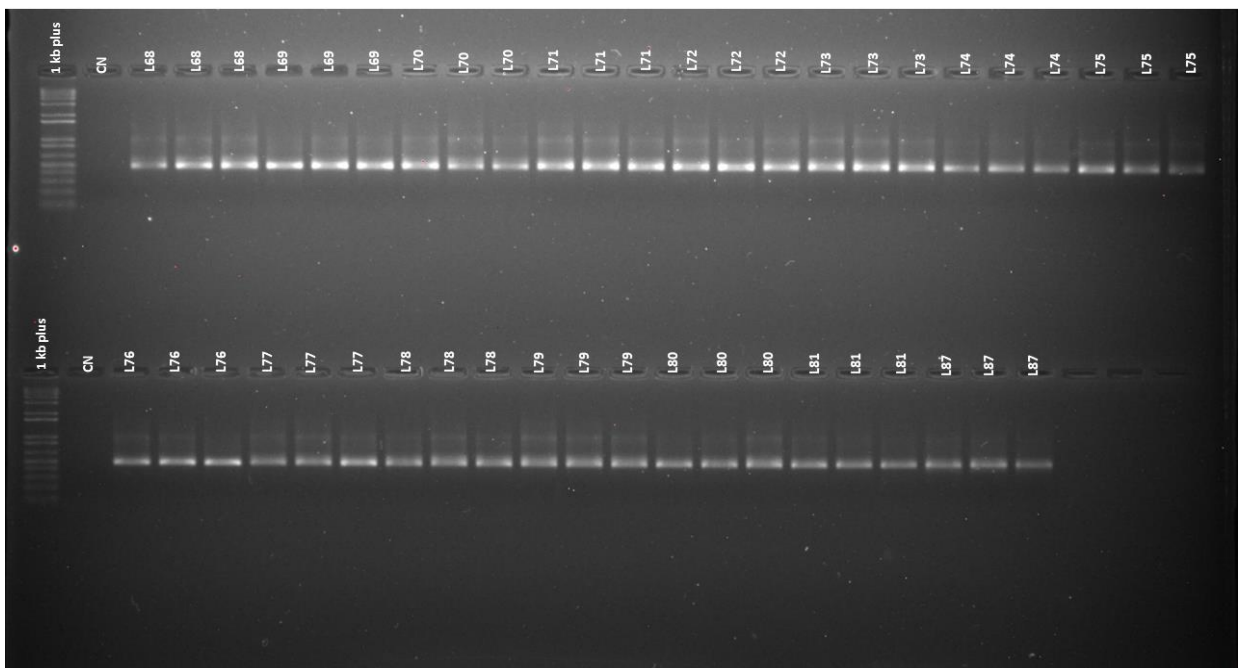


Figura 4S. Gel de agarose 1.5% corado com brometo de etídeo da amplificação do gene 18S-V4 dos pontos L68 a L87. Marcador 1kb plus.

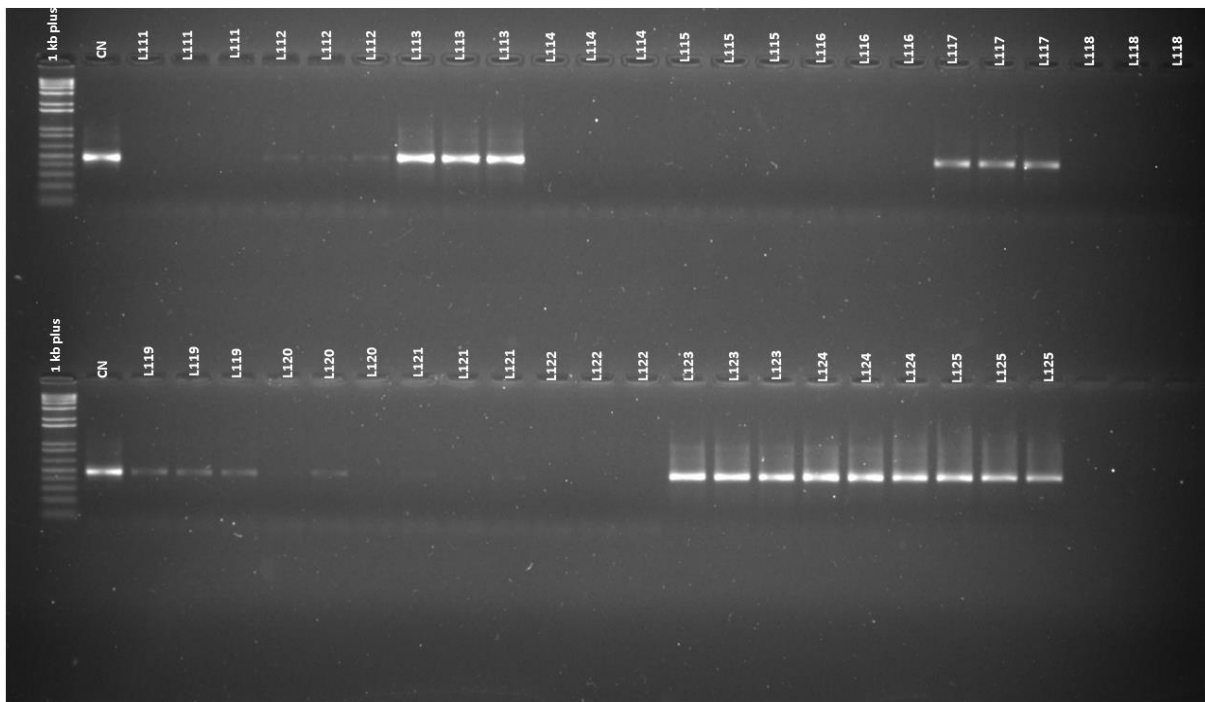


Figura 5S. Gel de agarose 1.5% corado com brometo de etídeo da amplificação do gene 18S-V4 dos pontos L111 a L125. Marcador 1kb plus.

Apêndice 4. Análise de qualidade das sequências geradas no sequenciamento na plataforma Illumina MiSeq utilizando o software FASTQC.

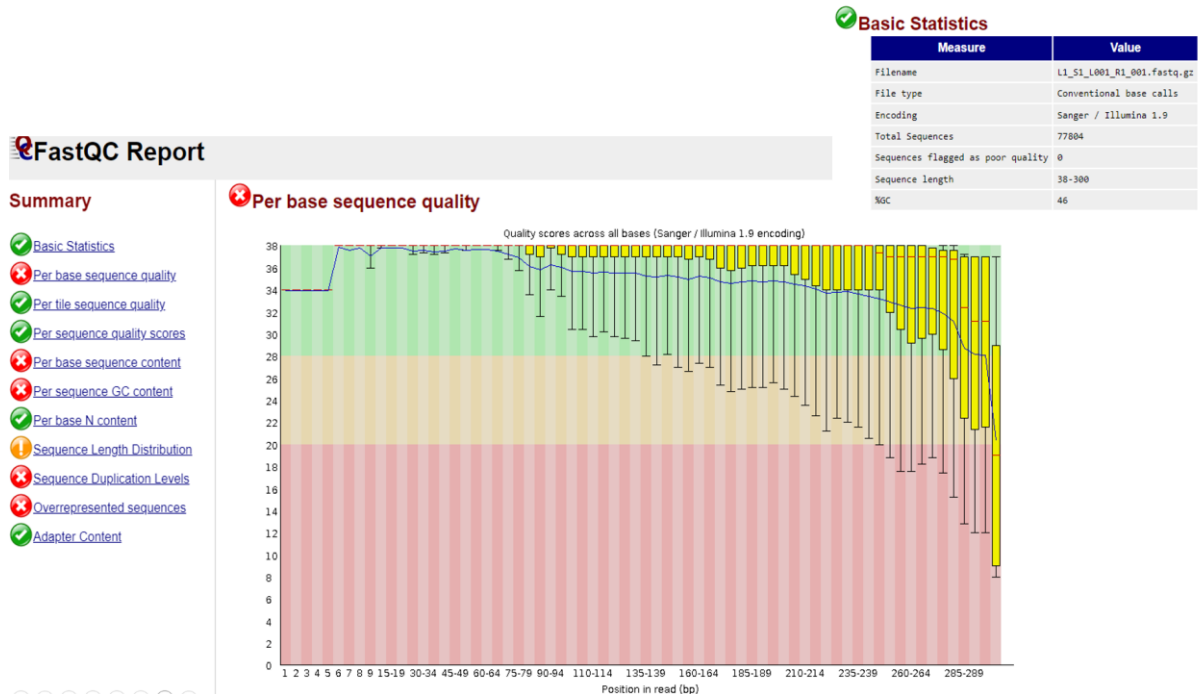


Figura 6S. Análise de qualidade da sequência L1_R1 (forward) por base.

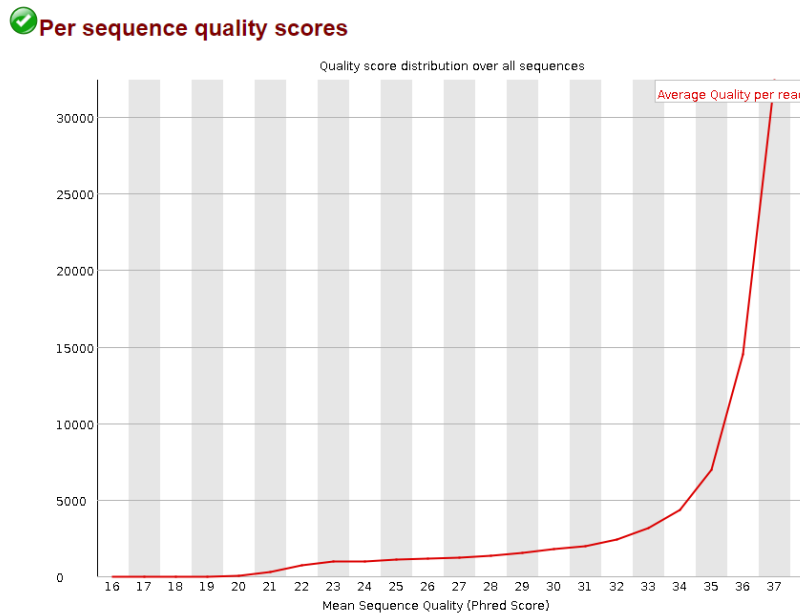


Figura 7S. Score de qualidade médio geral de todas as sequências de L1_R1.

Sequence Length Distribution

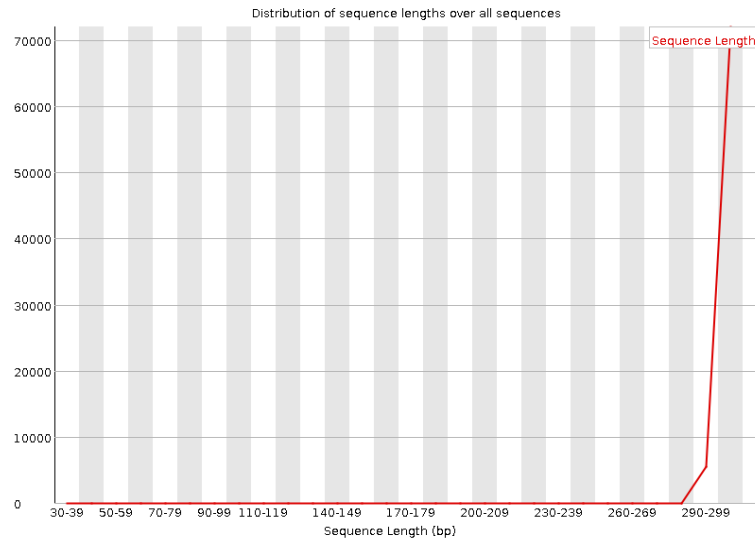


Figura 8S. Distribuição do tamanho geral de todas as sequências L1_R1.

Summary

- Basic Statistics
- Per base sequence quality
- Per tile sequence quality
- Per sequence quality scores
- Per base sequence content
- Per sequence GC content
- Per base N content
- Sequence Length Distribution
- Sequence Duplication Levels
- Overrepresented sequences
- Adapter Content

Basic Statistics

| Measure | Value |
|-----------------------------------|----------------------------|
| Filename | L1_S1_L001_R2_001.fastq.gz |
| File type | Conventional base calls |
| Encoding | Sanger / Illumina 1.9 |
| Total Sequences | 77804 |
| Sequences flagged as poor quality | 0 |
| Sequence length | 35-300 |
| %GC | 42 |

Per base sequence quality

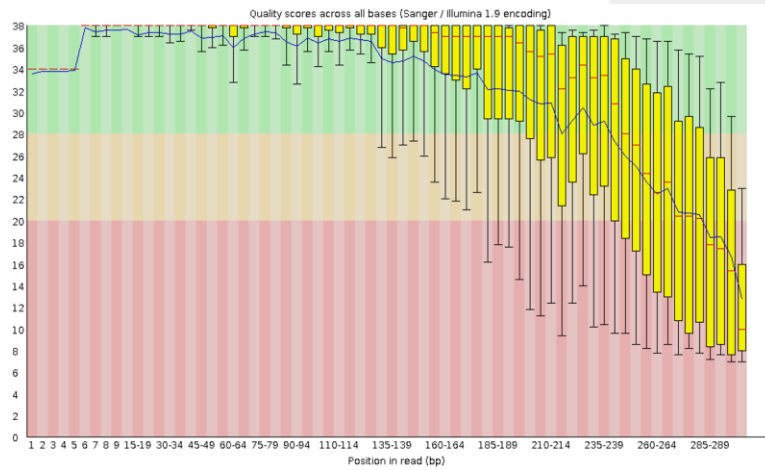


Figura 9S. Análise de qualidade das sequências L1_R2 (reverse) por base.

✔ **Per sequence quality scores**

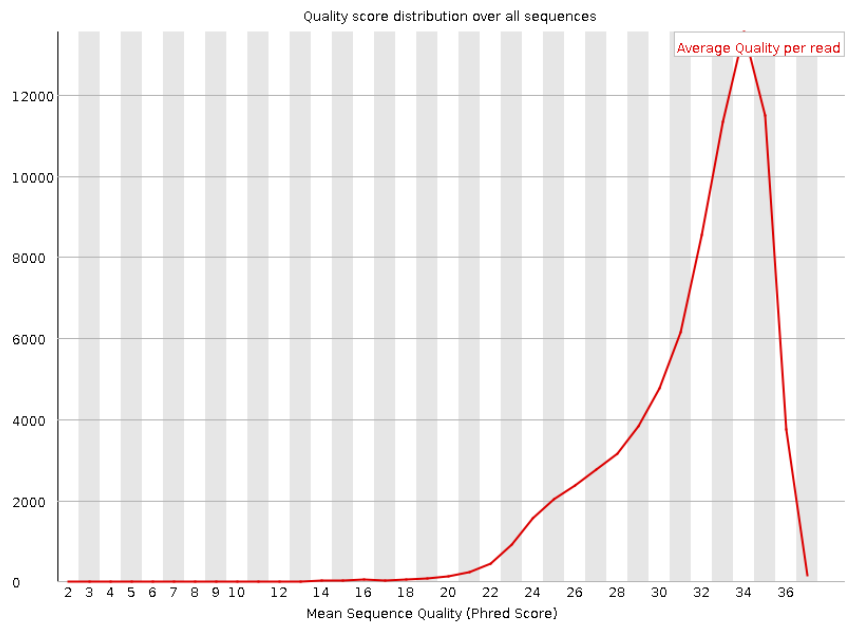


Figura 10S. Score de qualidade médio geral de todas as seqüências de L1_R2.

⚠ **Sequence Length Distribution**

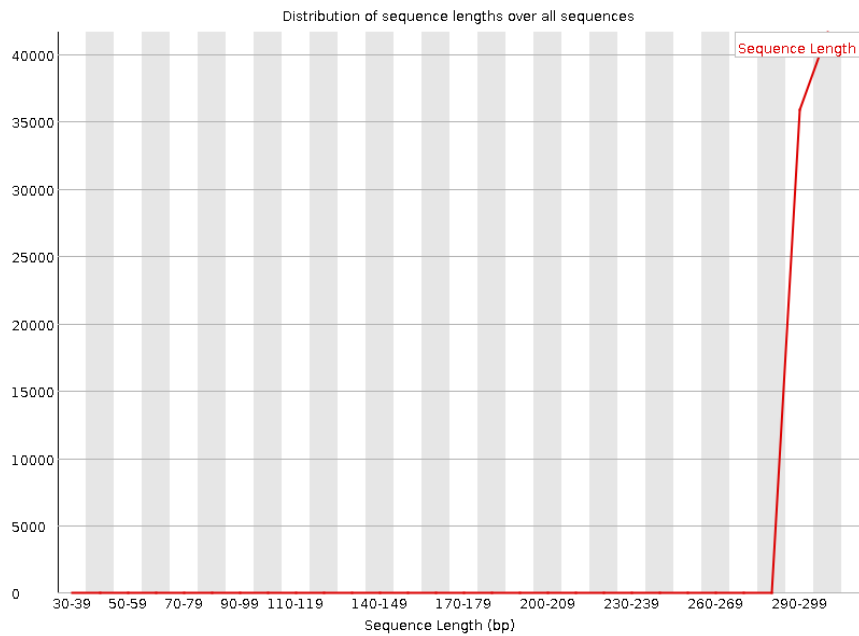


Figura 11S. Distribuição do tamanho geral de todas as seqüências L1_R2.

Summary

- Basic Statistics
- Per base sequence quality
- Per tile sequence quality
- Per sequence quality scores
- Per base sequence content
- Per sequence GC content
- Per base N content
- Sequence Length Distribution
- Sequence Duplication Levels
- Overrepresented sequences
- Adapter Content

Basic Statistics

| Measure | Value |
|-----------------------------------|------------------------------|
| Filename | L72_564_L001_R1_001.fastq.gz |
| File type | Conventional base calls |
| Encoding | Sanger / Illumina 1.9 |
| Total Sequences | 95958 |
| Sequences flagged as poor quality | 0 |
| Sequence length | 39-300 |
| %GC | 47 |

Per base sequence quality

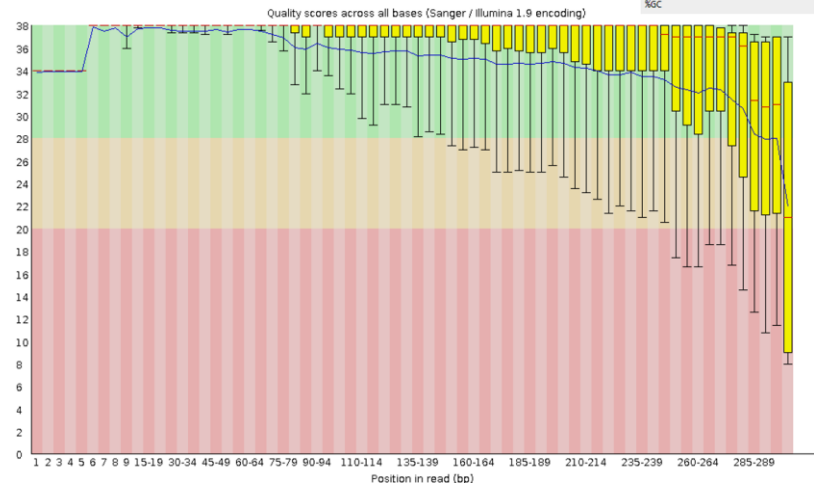


Figura 12S. Análise de qualidade das seqüências L72_R1 (forward) por base.

Per sequence quality scores

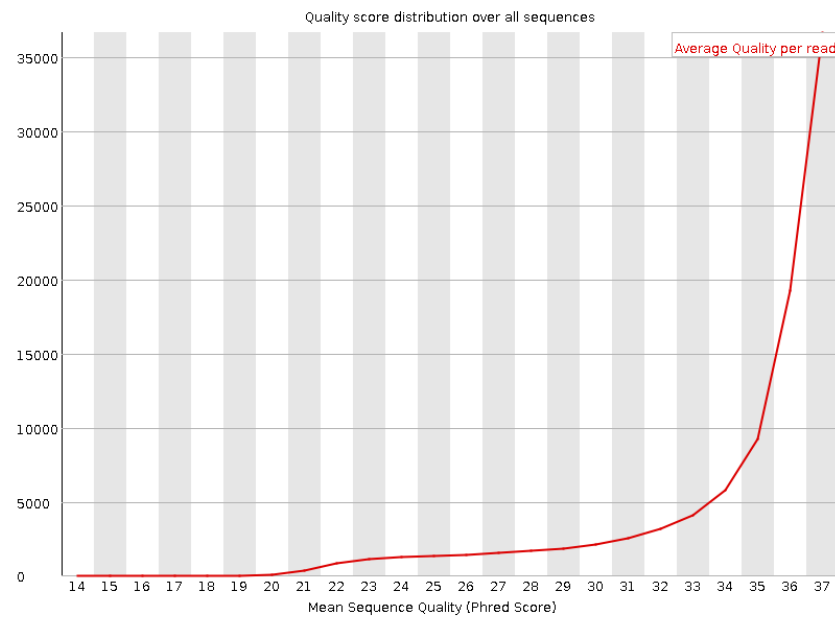


Figura 13S. Score de qualidade médio geral de todas as seqüências de L72_R1.

Sequence Length Distribution

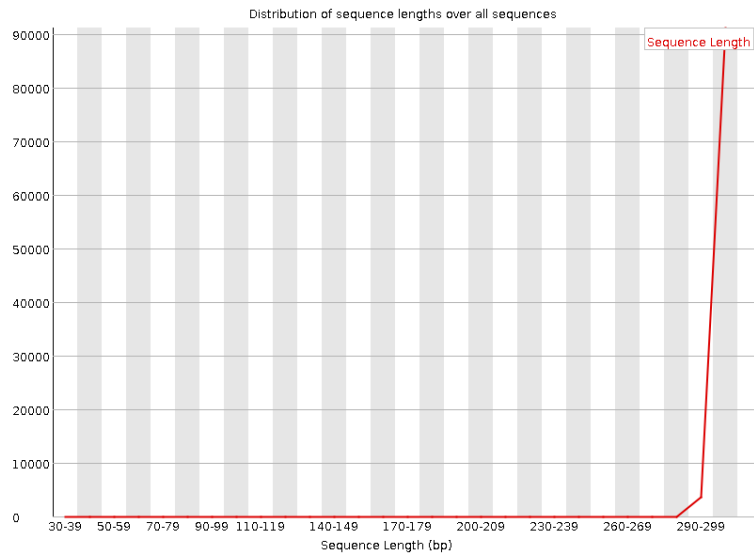


Figura 14S. Distribuição do tamanho geral de todas as sequências L72_R1.

Summary

- ✓ Basic Statistics
- ✗ Per base sequence quality
- ✓ Per tile sequence quality
- ✓ Per sequence quality scores
- ✗ Per base sequence content
- ✗ Per sequence GC content
- ✓ Per base N content
- ! Sequence Length Distribution
- ✗ Sequence Duplication Levels
- ✗ Overrepresented sequences
- ✓ Adapter Content

Basic Statistics

| Measure | Value |
|-----------------------------------|------------------------------|
| Filename | L72_564_L001_R2_001.fastq.gz |
| File type | Conventional base calls |
| Encoding | Sanger / Illumina 1.9 |
| Total Sequences | 95958 |
| Sequences flagged as poor quality | 0 |
| Sequence length | 35-300 |
| %GC | 43 |

Per base sequence quality

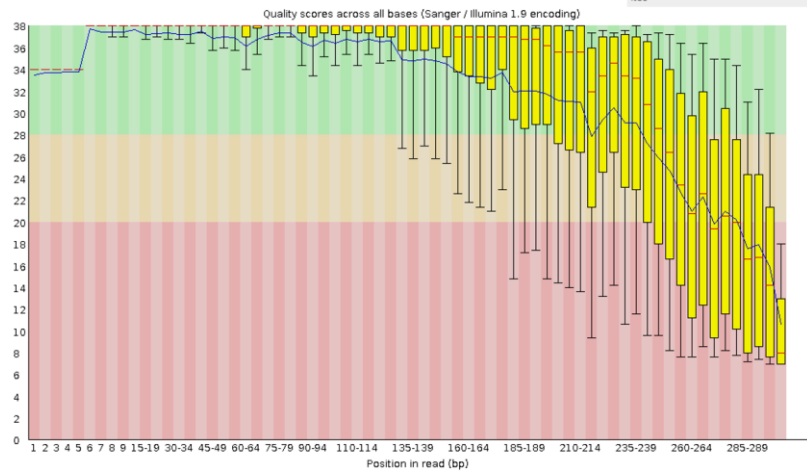


Figura 15S. Análise de qualidade das sequências L72_R2 (reverse) por base.

✔ Per sequence quality scores

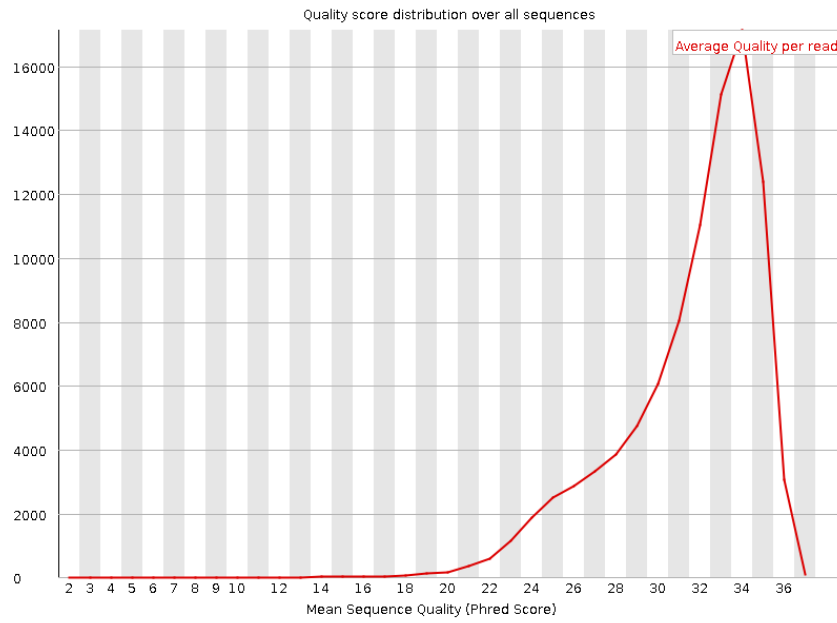


Figura 16S. Score de qualidade médio geral de todas as seqüências de L72

! Sequence Length Distribution

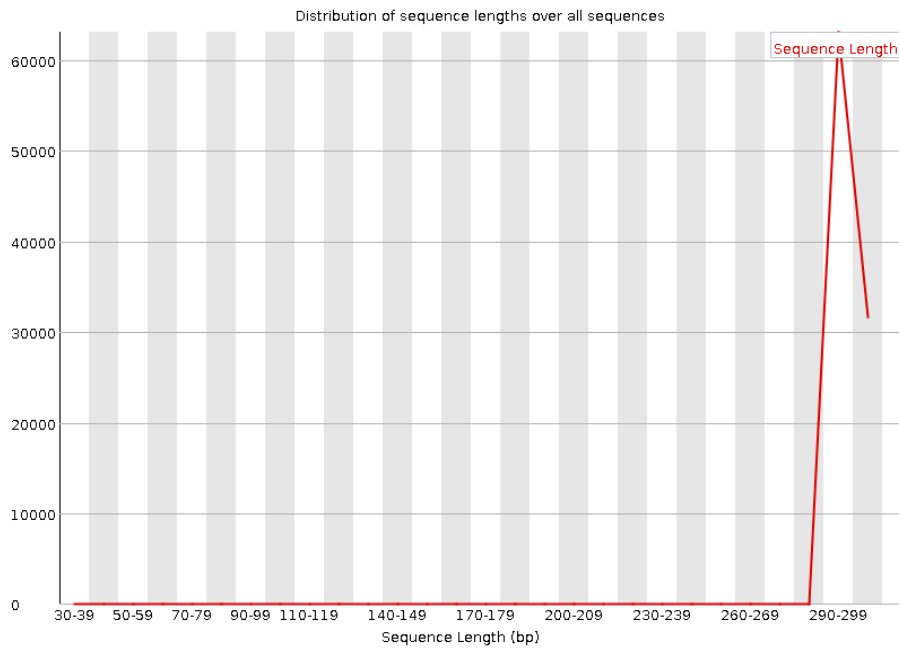


Figura 17S. Distribuição do tamanho geral de todas as seqüências L72_R2.

Apêndice 5. Valores de bootstraps para a árvore focal e árvore total gerada com o software IQTREE com 1000 bootstrap ultrarrápidos.

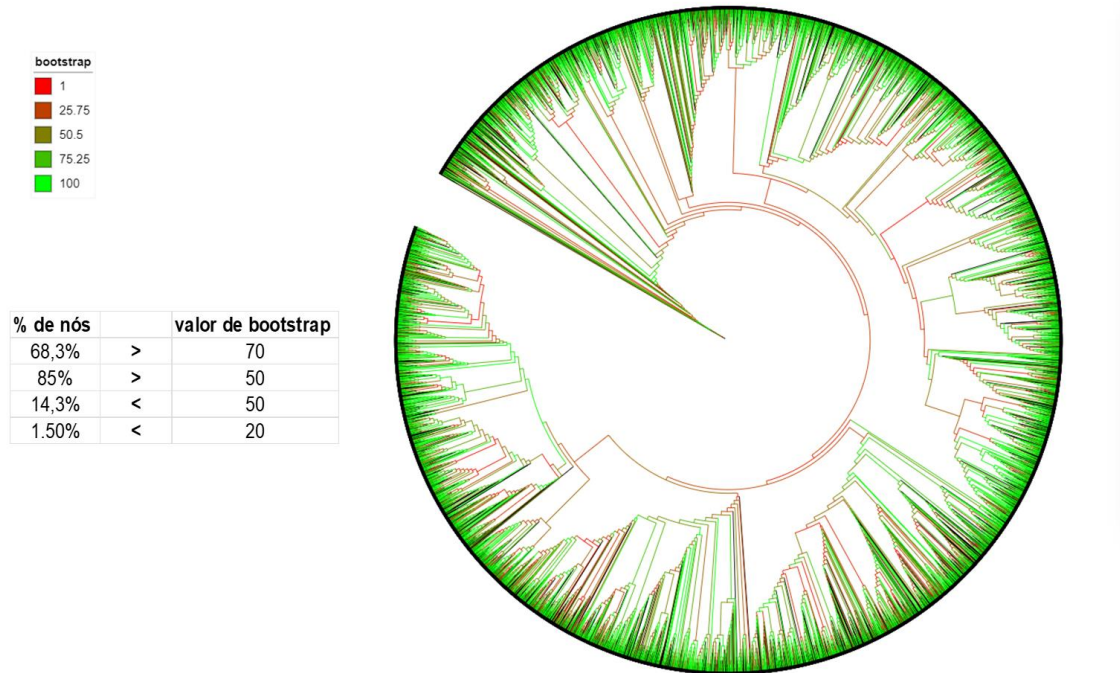


Figura 18S - Árvore filogenética total, com todas as OTUs identificadas com o gene 18S, mostrando os valores de bootstrap associados aos nós da filogenia, mostrado por padrão em cores, verde representando altos valores de bootstrap e vermelho representando baixos valores de bootstrap. E uma tabela mostrando a porcentagem de nós com valores de bootstraps acima ou abaixo do especificado.

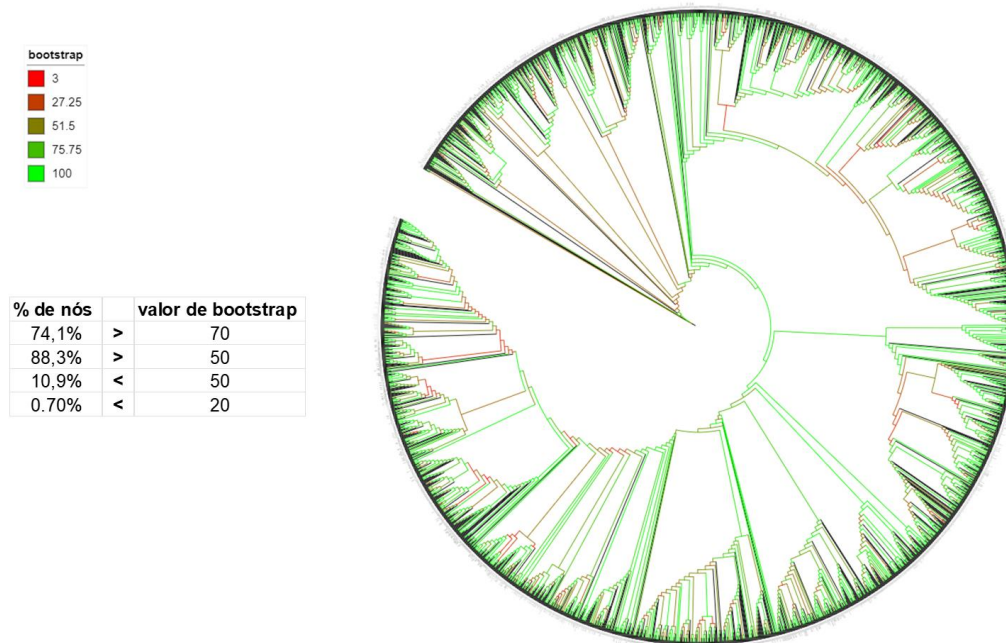


Figura 19S - Árvore filogenética focal, com as OTUs que representam os grupos de microeucariotos planctônicos fotossintetizantes identificadas com o gene 18S, mostrando os valores de bootstrap associados aos nós da filogenia, mostrado por padrão em cores, verde representando altos valores de bootstrap e vermelho representando baixos valores de bootstrap. E uma tabela mostrando a porcentagem de nós com valores de bootstraps acima ou abaixo do especificado.

Apêndice 6. A tabela com as variáveis ambientais e limnológicas dos 140 lagos do rio Araguaia e seus afluentes, juntamente com a tabela com o número de sequências obtidas no sequenciamento e análises de bioinformática e a tabela de atribuição taxonômica das OTUs do gene 18S-V4, classificadas no banco de dados SILVA (Quast *et al.*, 2013) e os mapas interativos de riqueza, diversidade e estrutura filogenética se encontram em uma pasta do drive, que pode ser acessada pelo link a seguir : https://drive.google.com/drive/folders/11R1MIEqUGQcYoBG9hw2C_bxaggCgxcu?usp=sharing

Apêndice 7. Tabela com as variáveis ambientais e espaciais selecionadas para inclusão no modelo de partição da variância para a comunidade total e focal.

Tabela 1S - Variáveis ambientais e espaciais selecionadas para o modelo de análise de redundância parcial (pRDA) e partição da variância para as diferentes métricas de diversidade (NRI e PD) para a árvore total e árvore focal.

| | Autovetores espaciais selecionados para o modelo de pRDA | | Váriaveis ambientais selecionadas para o modelo de pRDA |
|------------------|--|--|--|
| | PCNMs | AEMs | |
| PD_total | PCNM1, PCNM2, PCNM9, PCNM10, PCNM43 | V2, V6, V1, V71, V73, V31, V8, V29, V60, V14, V95, V106, V16, V94, V76, V97 | T_agua, pH, TRP, ORP, COND, PROF, NTU, OD, TDS, P, OrtoFosforo, N, Nitrato |
| PD_focal | PCNM2, PCNM9, PCNM1, PCNM10, PCNM48, PCNM20, PCNM14 | V1, V6, V2, V71, V8, V31, V106, V73, V76, V94, V57, V4, V60, V62, V12, V32, V5, V68, V29, V34, V77 | T_agua, pH, TRP, ORP, COND, PROF, NTU, OD, TDS, P, OrtoFosforo, N, Nitrato |
| NRI_total | PCNM1, PCNM3, PCNM5, PCNM40, PCNM32, PCNM41 | V2, V4, V5, V16, V12, V43, V26, V42, V82, V17, V95, V9, V84, V106 | T_agua, pH, TRP, ORP, COND, PROF, NTU, OD, TDS, P, OrtoFosforo, N, Nitrato |
| NRI_focal | PCNM1, PCNM40, PCNM3, PCNM5, PCNM41, PCNM46, PCNM19 | V2, V4, V7, V16, V9, V43, V5, V46, V15, V14, V42, V59, V26 | T_agua, pH, TRP, ORP, COND, PROF, NTU, OD, TDS, P, OrtoFosforo, N, Nitrato |

