

ACEITE PARA PUBLICAÇÃO

Título: Desenvolvimento de primers *in silico* para eDNA como ferramenta de monitoramento da comunidade fitoplanctônica em um Complexo Estuarino na Costa Amazônica

Revista: Observatório de la Economía Latinoamericana

Qualis A4 – Biotecnologia

Notificações



[olatin] Decisão do Editor

2024-07-11 03:15 PM

Jordana Adorno Furtado, Dandinaura Ferreira Silva, Nivia Sandiele de Melo Sousa, Vinicius Henrique Maciel dos Santos, Leticia Freitas Nascimento, Marcello Soares Santos Filho, Hanna Gabriely Pinto Gonçalves, Ligia Tchaicka:

Chegamos a uma decisão sobre sua submissão ao OBSERVATÓRIO DE LA ECONOMÍA LATINOAMERICANA, "Desenvolvimento de primers *in silico* para eDNA como ferramenta de monitoramento da comunidade fitoplanctônica em um Complexo Estuarino na Costa Amazônica".

Nossa decisão é: Aceitar a Submissão

Olá, boa tarde.

Após avaliação editorial, seu artigo teve parecer **POSITIVO** para publicação.

Desenvolvimento de primers *in silico* para eDNA como ferramenta de monitoramento da comunidade fitoplanctônica em um Complexo Estuarino na Costa Amazônica

Development of *in silico* primers for eDNA as a tool for monitoring the phytoplankton community in an Estuarine Complex on the Amazonian Coast

Desarrollo de primers *in silico* para eDNA como herramienta para el monitoreo de la comunidad de fitoplancton en un Complejo Estuarino de la Costa Amazónica

DOI:

Receipt of originals: / /

Acceptance for publication: / /

Jordana Adorno Furtado

Doutoranda em Biodiversidade e Biotecnologia (BIONORTE)
Instituição: Universidade Estadual do Maranhão (UEMA)
Endereço: São Luís, Maranhão, Brasil
E-mail: jordana1adorno@gmail.com

Dandinaura Ferreira Silva

Graduanda em Ciências Biológicas
Instituição: Universidade Estadual do Maranhão (UEMA)
Endereço: São Luís, Maranhão, Brasil
E-mail: dandinauraferreira@gmail.com

Nivia Sandiele de Melo Sousa

Mestranda em Ecologia e Conservação da Biodiversidade
Instituição: Universidade Estadual do Maranhão (UEMA)
Endereço: São Luís, Maranhão, Brasil
E-mail: niviasamelo@gmail.com

Vinicius Henrique Maciel dos Santos

Doutorando em Geologia e Geoquímica
Instituição: Universidade Federal do Pará (UFPA)
Endereço: Belém, Pará, Brasil
E-mail: macielviniciusoc@gmail.com

Leticia Freitas Nascimento

Graduanda em Ciências Biológicas
Instituição: Universidade Estadual do Maranhão (UEMA)

Endereço: São Luís, Maranhão, Brasil
E-mail: leticiafrenasc16@gmail.com

Marcello Soares Santos Filho

Graduando em Ciências Biológicas
Instituição: Universidade Estadual do Maranhão (UEMA)
Endereço: São Luís, Maranhão, Brasil
E-mail: marcello.s.santos.f@gmail.com

Hanna Gabriely Pinto Gonçalves

Doutoranda em Biodiversidade e Biotecnologia (BIONORTE)
Instituição: Universidade Federal do Maranhão (UFMA)
Endereço: São Luís, Maranhão e Brasil
E-mail: hanna.gabriely@discente.ufma.br

Ligia Tchaicka

Doutora em Genética e Biologia Molecular
Instituição: Universidade Estadual do Maranhão (UEMA)
Endereço: São Luís, Maranhão, Brasil
E-mail: ligia.tchaicka@uemanet.uma.br

RESUMO

Este estudo teve como objetivo selecionar iniciadores para aplicação em análises de metabarcoding (eDNA), visando o monitoramento da comunidade fitoplanctônica em um Complexo Estuarino na Costa Amazônica. Inicialmente, foi realizado um levantamento bibliográfico de trabalhos publicados entre 2005 e 2023 nas principais bases de dados (SciELO, Web of Science, ScienceDirect e Google Scholar) sobre as espécies fitoplanctônicas da região. Após isso, foi realizada a busca das sequências de referência para as espécies registradas, com foco nos genes ribossomais 16S, 18S, 23S e no gene mitocondrial Citocromo Oxidase Subunidade I (COI), contidas nos bancos genômicos (GenBank, SILVA e BOLD Systems). Para alinhar as sequências, foi usado o software BioEdit e, para avaliar a qualidade dos primers desenvolvidos, foi utilizada a ferramenta online OligoAnalyzer 3.1. Os resultados mostraram que o filo Bacillariophyta foi o mais representativo em número de espécies e sequências disponíveis. Dos marcadores escolhidos, o gene 18S ribossomal obteve a maior riqueza de sequências depositadas nos bancos de dados, especialmente no GenBank. Cinco conjuntos de primers (DFS-F1, DFS-F2, DFS-F3, DFS-F4 e DFS-F5) foram desenhados *in silico* e validados virtualmente, apresentando entre 19 e 22 pares de bases (pb). As temperaturas de fusão (Tmelting) variaram entre 55,1°C e 58°C, e a concentração de GC% variou de 40,91% a 57,89%. Portanto, esses resultados colaboram com a literatura sobre protocolos de eDNA, auxiliando na escolha do marcador molecular e dos primers mais adequados para a comunidade fitoplanctônica, aumentando a eficiência e rapidez do trabalho.

Palavras-chave: Desenho de primers, Banco de dados, eDNA, Fitoplâncton

ABSTRACT

This study aimed to select primers for application in metabarcoding (eDNA) analysis to monitor the phytoplankton community in an Estuarine Complex on the Amazon Coast. Initially, a bibliographical survey of works published between 2005 and 2023 in the principal databases (SciELO, Web of Science, ScienceDirect, and Google Scholar) on phytoplanktonic species in the region was carried out. After this, a search for reference sequences for the recorded species was carried out, focusing on the ribosomal genes 16S, 18S, 23S and the mitochondrial gene Cytochrome Oxidase Subunit I (COI), contained in genomic banks (GenBank, SILVA, and BOLD Systems). To align the sequences, we used the BioEdit software, and to assess the quality of the developed primers, we utilized the OligoAnalyzer 3.1 online tool. The results showed that the phylum Bacillariophyta was the most representative regarding the number of species and sequences available. Of the chosen markers, the 18S ribosomal gene had the highest wealth of sequences deposited in the databases, especially in GenBank. Five sets of primers (DFS-F1, DFS-F2, DFS-F3, DFS-F4, and DFS-F5) were designed *in silico* and virtually validated, presenting between 19 and 22 base pairs (bp). The melting temperatures (T_{melting}) varied between 55.1°C and 58°C, and the GC% concentration ranged from 40.91% to 57.89%. Therefore, these results collaborate with the literature on eDNA protocols, helping to choose the most suitable molecular marker and primers for the phytoplankton community, increasing the efficiency and speed of the work.

Keywords: Primer design, Database, eDNA, Phytoplankton

RESUMEN

El objetivo de este estudio fue seleccionar primers para su aplicación en análisis de metabarcodificación (eDNA) para monitorear la comunidad fitoplanctónica en un Complejo Estuarino de la Costa Amazónica. Inicialmente, se realizó un relevamiento bibliográfico de trabajos publicados entre 2005 y 2023 en las principales bases de datos (SciELO, Web of Science, ScienceDirect y Google Scholar) sobre especies fitoplanctónicas de la región. A continuación, se realizó una búsqueda de secuencias de referencia para las especies registradas, centrándose en los genes ribosómicos 16S, 18S, 23S y el gen mitocondrial Citocromo Oxidasa Subunidad I (COI), contenidos en bases de datos genómicas (GenBank, SILVA y BOLD Systems). Se utilizó el programa BioEdit para alinear las secuencias y la herramienta en línea OligoAnalyzer 3.1 para evaluar la calidad de los primers desarrollados. Los resultados mostraron que el filo Bacillariophyta era el más representativo en cuanto al número de especies y secuencias disponibles. De los marcadores elegidos, el gen ribosómico 18S era el que contaba con mayor riqueza de secuencias depositadas en bases de datos, especialmente en GenBank. Se diseñaron *in silico* y validaron virtualmente cinco conjuntos de primers (DFS-F1, DFS-F2, DFS-F3, DFS-F4 y DFS-F5), con entre 19 y 22 pares de bases (pb). Las temperaturas de fusión oscilaron entre 55,1°C y 58°C, y la concentración de GC% entre 40,91% y 57,89%. Por lo tanto, estos resultados contribuyen a la literatura sobre protocolos de eDNA, ayudando a elegir el marcador molecular y los primers más adecuados para la comunidad fitoplanctónica, aumentando la eficiencia y la rapidez del trabajo.

Palabras clave: Diseño de primers, Base de datos, eDNA, Fitoplancton

1 INTRODUÇÃO

O fitoplâncton atua como o principal produtor primário em ecossistemas costeiros, pois realiza a ciclagem de nutrientes e contribui diretamente para a estabilização funcional e a preservação do equilíbrio biológico desses ambientes (Sun *et al.*, 2020; Yoon *et al.*, 2024). Além disso, é um excelente indicador de alterações na qualidade da água (Bilbao *et al.*, 2023; Raveh *et al.*, 2019), uma vez que é altamente sensível a perturbações quando comparado a outros organismos aquáticos (Lee *et al.*, 2023; Rocha *et al.*, 2023).

A microscopia óptica tem sido comumente a técnica mais utilizada para investigar a diversidade do fitoplâncton (Agirbas *et al.*, 2015; Lee; Won; Baek, 2020). Porém, essa técnica apresenta algumas limitações devido à falta de características diagnósticas claras, conhecimento taxonômico limitado e tempo prolongado de análise, resultando muitas vezes em informações insuficientes sobre a diversidade (Barsanti *et al.*, 2021; Gaonkar; Campbell, 2024).

O rápido crescimento das tecnologias de sequenciamento de alto rendimento possibilitou a utilização da taxonomia molecular como uma opção viável para os inventários morfotaxonômicos. A base para a utilização dessas metodologias de teste baseadas em DNA se fundamenta na ideia de que é possível identificar de maneira confiável os táxons com base em um fragmento específico de seu genoma, conhecido como "código de barras de DNA" (Pawlowski *et al.*, 2022; Taberlet *et al.*, 2012;).

O metabarcoding possibilita a identificação simultânea de táxons em amostras ambientais com base em seu DNA, utilizando o sequenciamento de genes marcadores específicos (Keck *et al.*, 2017; Zimmermann *et al.*, 2014). Esse método gera extensas informações sobre biodiversidade, pois é capaz de identificar espécies em diversas fases de vida, além de espécies crípticas e aquelas frequentemente descartadas pelos métodos tradicionais (Comtet *et al.*, 2015).

Com o avanço da bioinformática, tornou-se viável o desenvolvimento de ferramentas que oferecem suporte às pesquisas atuais, incluindo a concepção de primers

para PCR com base no *The National Center for Biotechnology Information* (NCBI) (Morais *et al.*, 2019; Muniz *et al.*, 2020). Os *primers* são sequências específicas de nucleotídeos que podem se ligar às suas sequências alvo de DNA durante a PCR (Pelt-Verkuil; Van Belkum; Hays, 2008). A finalidade da PCR *in silico* é antecipar quais produtos de PCR serão sintetizados a partir de um ou vários modelos de DNA, utilizando primers oligonucleotídicos já concebidos (Kalendar *et al.*, 2017), economizando tempo e custos adicionais (Schuler, 1997).

O Complexo Estuarino de São Marcos (CESM) caracteriza-se como um sistema de grande relevância na Costa Amazônica. Neste ambiente está inserido o Porto do Itaqui, um dos portos públicos mais importantes da América Latina, que tem sofrido fortes impactos decorrentes das atividades antrópicas (Jesus *et al.*, 2020; Oliveira *et al.*, 2019; Ribeiro *et al.*, 2023). No entanto, estudos que estabeleçam iniciadores de PCR com bom desempenho aplicados à comunidade fitoplanctônica são inexistentes.

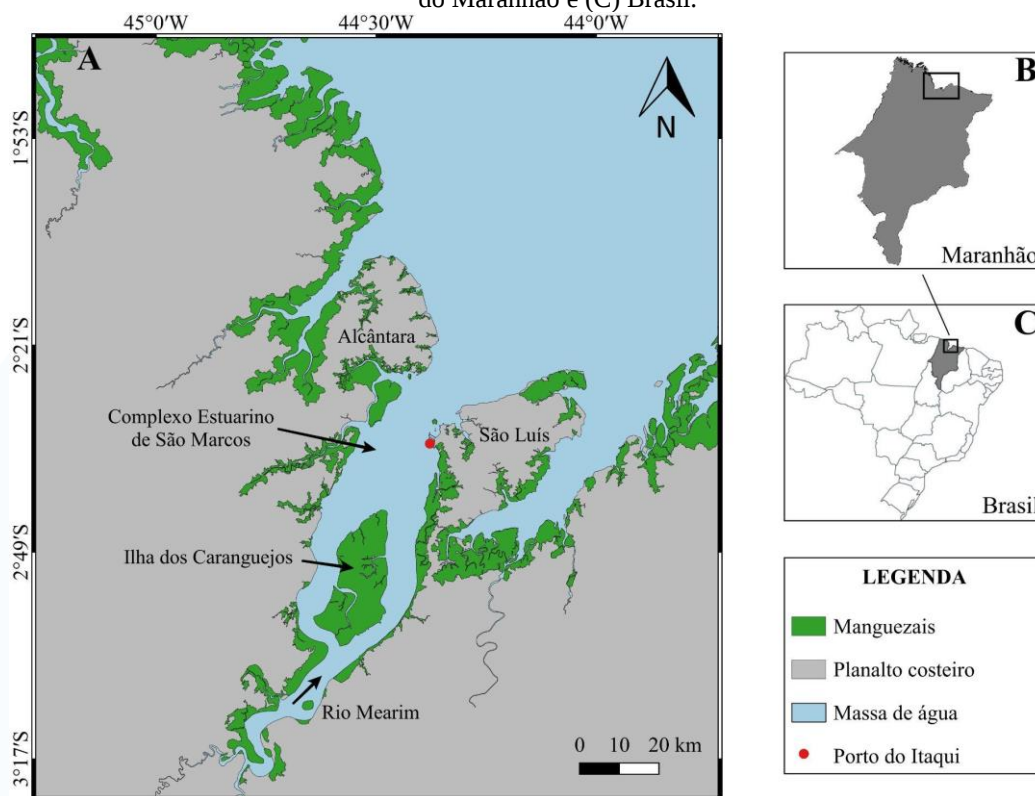
Nesse sentido, o presente estudo tem como objetivo principal o desenvolvimento *in silico* de primers para aplicação em análises de metabarcoding (eDNA), que auxiliem no monitoramento da comunidade fitoplanctônica em um Complexo Estuarino na Costa Amazônica.

2 METODOLOGIA

2.1 ÁREA DE ESTUDO

O CESM é caracterizado como uma extensa área estuarina localizada no extremo norte do Estado do Maranhão, Brasil (Figura 1). Esse ambiente está localizado em uma região conhecida como Costa dos Manguezais das Macromarés Amazônicas (CMMA), inserida entre os estados brasileiros do Amapá e Maranhão (Diniz *et al.*, 2019). Possui orientação principal NE-SW, com uma largura de 50 km na porção norte, 15 km na região central (entre Alcântara e a Ponta de São Marcos), 25 km na altura da ilha dos Caranguejos e 4 km na foz do Rio Mearim (El-Robrini *et al.*, 2006).

Figura 1. Localização da área de estudo: (A) Complexo Estuarino de São Marcos - CESM; (B) estado do Maranhão e (C) Brasil.



Fonte: Elaborado pelos autores (2024).

2.2 LISTA DE ESPÉCIES

Para este estudo, foi realizado um levantamento de trabalhos publicados entre 2005 e 2023 nas bases de dados SciELO (<https://search.scielo.org/>), Web of Science (<https://webofknowledge.com>), ScienceDirect (<https://www.sciencedirect.com/>) e Google Scholar (<https://scholar.google.com.br/>). Os descritores utilizados foram “fitoplâncton”, “Baía de São Marcos” e “Complexo Estuarino de São Marcos”, e suas respectivas traduções para o inglês. Os critérios de inclusão consideraram não só artigos e notas científicas, mas também monografias, dissertações, teses, livros, capítulos e estudos publicados em anais de eventos. A classificação taxonômica das espécies de fitoplâncton foi atualizada seguindo a base de dados internacional AlgaeBase (Guiry; Guiry, 2023).

2.3 LEVANTAMENTO DE SEQUÊNCIAS

Com base na lista de espécies obtidas, foi realizada a busca das sequências de referência focada nos genes ribossomais 16S, 18S, 23S e no gene mitocondrial Citocromo Oxidase Subunidade I (COI), nos principais bancos de dados genômicos: GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>), SILVA (<https://www.arb-silva.de/>) e BOLD Systems (<https://www.boldsystems.org/>).

2.4 DESENHO DE INICIADORES

As sequências obtidas foram convertidas para o formato FASTA no próprio sistema do banco de dados e, em seguida, foram baixadas para um software de alinhamento. Para isso, utilizou-se o software BioEdit Sequence Alignment Editor, versão 7.2.6t (Hall, 1999). Inicialmente, as sequências foram agrupadas de acordo com a ordem a que as espécies pertenciam, cada ordem foi alinhada e editada para melhor padronização das sequências. Posteriormente, esses alinhamentos foram agrupados em um único database no próprio software. Por fim, os gaps das extremidades foram removidos pois a ferramenta “Primer Blast” não permite o desenho dos primers com a presença de gaps.

Para o desenho dos primers, foram estabelecidas espécies-alvo para cada ordem com base na maior disponibilidade de dados, onde os primers foram desenvolvidos com base na região 18S. Os desenhos dos primers foram realizados na ferramenta "Primer Blast" (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>) do NCBI com os números de acessos de cada sequências alvo usado como PCR Template. Para isso, alguns parâmetros foram alterados na plataforma, são eles - PCR product size: Min: 70 e Max: 300; o Database: custom (database produzido pelos autores); Max target amplicon size: 400. Em Advanced parameters foram alterados - Primer size: Min:18, Opt: 20 e Max: 24 e em Primer GC content (%): o Min: 40.0 e Max: 60.0. Nos demais parâmetros, foram usados o default da plataforma.

2.5 ANÁLISE DE QUALIDADE E VALIDAÇÃO *IN SILICO*

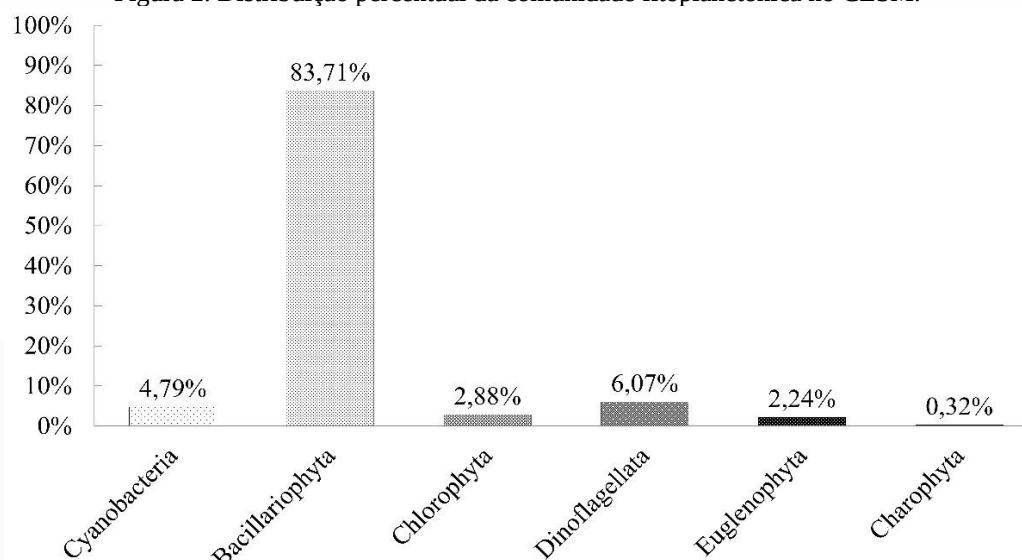
Para avaliar a qualidade dos primers desenvolvidos antes da PCR *in silico*, foi utilizada a ferramenta online OligoAnalyzer 3.1. Nessa etapa, foram avaliadas a possibilidade de formação de hairpins, self-dímeros e a temperatura de melting (°C). Posteriormente, os primers foram delimitados considerando a quantidade de guanina e citosina em suas extremidades, visando aumentar a estabilidade durante as etapas de anelamento e evitar sua dissociação (Marinho, 2021).

Para a validação dos primers foi realizado uma PCR *in silico* na ferramenta “BLAST”(https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome) do NCB. Vale salientar que a busca pelos *primers* para análise da cobertura de espécies da PCR *in silico* foi realizada em dois processos distintos. Primeiramente, foi realizada uma PCR específica para as espécies-alvo. Para isso, foi utilizado a Database “Nucleotide collection (nt)” e foram adicionadas na seção “Organism” a espécie-alvo, a ordem à qual ela pertence e as outras espécies da mesma ordem. Posteriormente, foi realizada uma PCR geral, na qual foi feita uma única busca para cada primer, incluindo todas as ordens e todas as espécies com sequências disponíveis.

3 RESULTADOS E DISCUSSÃO

O levantamento realizado contabilizou 313 táxons referentes à comunidade fitoplanctônica. Esses táxons estiveram distribuídos entre os filos Bacillariophyta (83,71%), Dinoflagellata (6,07%), Cyanobacteria (4,79%), Chlorophyta (2,88%), Euglenophyta (2,24%) e Charophyta (0,32%) (Figura 2). As diatomáceas (Bacillariophyta) foram as mais representativas com 262 táxons, sendo os gêneros *Chaetoceros* (23 táxons), *Nitzschia* (18 táxons) e *Coscinodiscus* (16 táxons) os mais abundantes.

Figura 2. Distribuição percentual da comunidade fitoplanctônica no CSM.



Fonte: Elaborado pelos autores (2024).

O filo Dinoflagellata apresentou 19 espécies, sendo o gênero *Triplos* o mais abundante com 7 espécies. Euglenophyta registrou 7 espécies, com os gêneros mais expressivos sendo *Euglena*, *Phacus* e *Trachelomonas*, cada um com 2 espécies. O filo Cyanobacteria apresentou 15 espécies, das quais o gênero *Oscillatoria* se destacou com 3 espécies. Chlorophyta apresentou 9 espécies, sem nenhum gênero expressivo, e Charophyta apresentou apenas 1 espécie.

As diatomáceas constituíram o grupo dominante em número de táxons, uma vez que são comumente encontradas em ambientes costeiros e contribuem com cerca de 20% da produção primária líquida anual em escala global (Field *et al.*, 1998; Singh *et al.*, 2024). Tais organismos ocorrem não só pela diversidade do grupo, mas por serem eurialinas e fornecerem uma rápida resposta mediante alterações ambientais (Aquino *et al.*, 2015). Esta dominância também foi observada em outros estudos realizados em estuários amazônicos (Carvalho Neto *et al.*, 2023; Cavalcanti *et al.*, 2020; Cavalcanti-Lima *et al.*, 2023; Queiroz *et al.*, 2022). Dentre os gêneros de diatomáceas identificados, os mais abundantes foram *Chaetoceros*, *Nitzschia* e *Coscinodiscus*, resultados similares foram registrados por Cavalcanti *et al.* (2018), Gama *et al.* (2011) e Matos *et al.* (2012) na mesma região.

Conforme Carvalho Neto *et al.* (2023), a dominância desses gêneros demonstra a similaridade na comunidade fitoplanctônica encontrada em estuários amazônicos, ainda que apresentem oscilações no número de espécies. Esse fator é influenciado pelas características particulares de cada ambiente, como a morfologia, aporte marinho e/ou limnético que ele possui.

Quanto às sequências obtidas para as espécies identificadas, o filo Bacillariophyta foi o mais representativo, com 119 táxons registrando a presença de sequências. Para o filo Dinoflagellata, 15 táxons apresentaram sequências disponíveis, seguido por Euglenophyta e Cyanobacteria, cada um com 5 espécies. Chlorophyta obteve sequências disponíveis para 4 táxons, enquanto Charophyta contou com apenas 1 espécie com sequência depositada no GenBank.

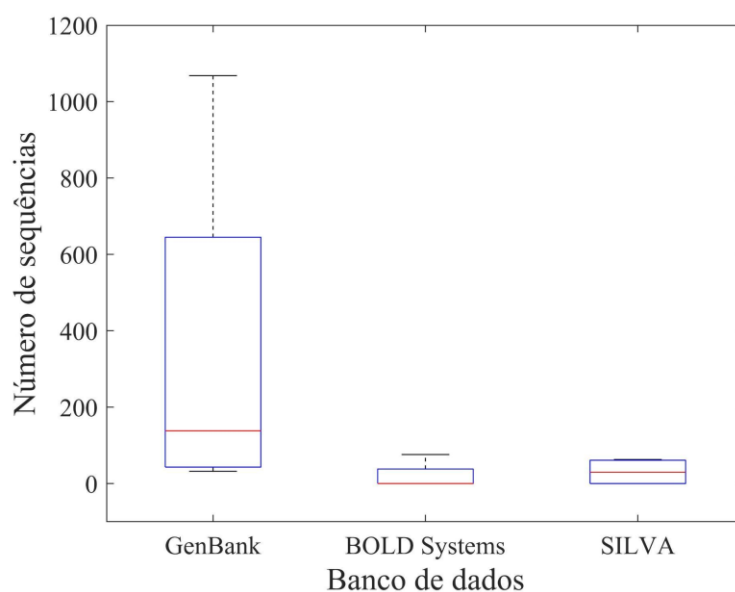
O filo Bacillariophyta apresenta uma das maiores diversidade de espécies dentre os protistas, possuem mais de 200 gêneros e 100.000 espécies (Miettinen, 2018). Essa grande riqueza de espécies e estudos favorece a disponibilidade de sequências nos bancos de dados. Conforme Gaonkar e Campbell (2024), a disponibilidade de bancos de dados de referência completos e validados taxonomicamente é essencial para que os estudos com metabarcoding possam ser eficazes.

Grande parte das sequências de referência disponíveis são para os gêneros mais frequentemente estudados, enquanto as pequenas algas (<5 microns) são pouco caracterizadas e, deste modo, não estão presentes nos bancos de dados de referência (Gaonkar; Campbell, 2024). Atualmente, o desenvolvimento de técnicas de sequenciamento de alto rendimento, integrado aos bancos de dados de sequências de referência aprimorados e bem descritos taxonomicamente, permitem uma avaliação rápida e econômica da composição fitoplanctônica em nível global (de Vargas *et al.*, 2015).

Através do levantamento de sequências feito nos bancos de dados públicos (GenBank, SILVA e BOLD Systems), foi observado que o GenBank (1376) apresentou a maior riqueza no número de sequências depositadas (Figura 3). Embora o SILVA (122) e o BOLD Systems (76) também contenham sequências, a quantidade obtida é inferior ao GenBank. Além disso, a maior parte das sequências observadas nesses bancos também

consta no GenBank. Nesse contexto, as sequências trabalhadas neste estudo foram extraídas majoritariamente dessa biblioteca.

Figura 3. Boxplots de valores de riqueza calculados para os bancos de dados genômicos.



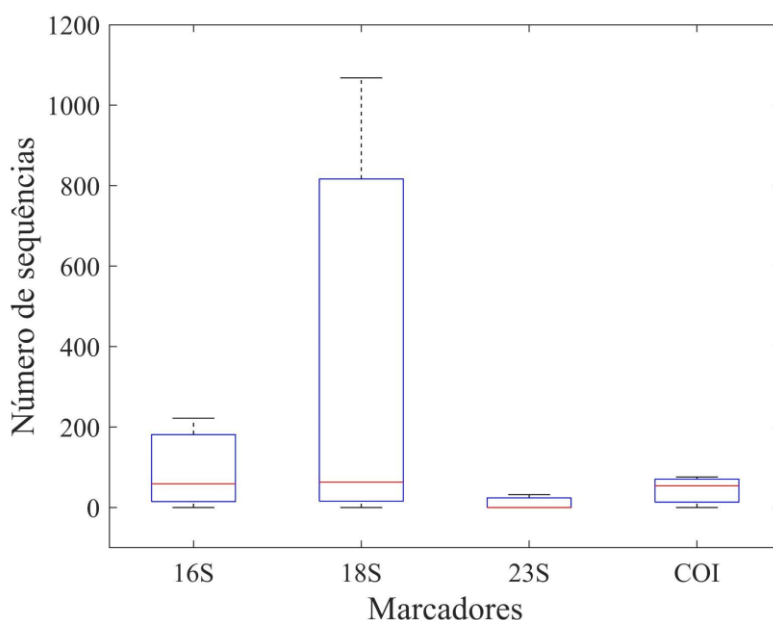
Fonte: Elaborado pelos autores (2024).

Segundo Patel; Scheible; Meiklejohn (2023), o GenBank é o maior banco de dados de sequências públicas do mundo, atualmente possuindo mais de 89 milhões de sequências na coleção Nucleotide. Os recursos disponíveis nessa base de dados a tornam uma ferramenta promissora na preparação do banco de dados a ser utilizado no metabarcoding (Ardura, 2019), uma vez que essa é a base mais completa em número de marcadores e espécies (NCBI, 2016).

Comparando a quantidade de espécies com sequências disponíveis ao total de espécies obtidas através do levantamento, foi observado que menos da metade das espécies tem sequências registradas. Assim, ainda que o protocolo de DNA Barcoding seja eficaz, nota-se que a disponibilidade das sequências de interesse nos bancos de dados é essencial para a identificação de uma espécie. A ausência dessas sequências pode tornar-se, portanto, um fator limitante em estudos que utilizam essa técnica (Cerutti-Pereyra *et al.*, 2012; Kolmann *et al.*, 2017; Martins, 2021).

O gene 18S ribossomal obteve a maior riqueza de sequências depositadas nos bancos de dados, especialmente no GenBank (Figura 4). Tais resultados corroboram os estudos realizados por Countway *et al.* (2005), Gong e Marchetti (2019) e de Vargas *et al.* (2015), que ressaltam que o gene 18S rRNA tem se destacado como o marcador mais comumente utilizado em estudos relacionados à comunidade de protistas eucariontes, tanto em ecossistemas aquáticos quanto terrestres, devido à sua alta especificidade e preservação da sequência.

Figura 4. Boxplots de valores de riqueza calculados para os marcadores analisados.



Fonte: Elaborado pelos autores (2024).

No que diz respeito aos marcadores, o gene 18S é altamente conservador em plantas e animais superiores e sua taxa evolutiva é considerada adequada em diatomáceas (Liu *et al.*, 2020). Segundo Kooistra *et al.* (2010), o gene 18S rDNA é amplamente utilizado como marcador para identificação e delineamento de espécies devido às suas regiões conservadas e hipervariáveis (V1-V9). No entanto, estudos realizados por Clarke *et al.* (2017), Zhang *et al.* (2018) e Wangenstein *et al.* (2018) indicam que o 18S possui uma menor capacidade de recuperação taxonômica quando comparado a outro marcador universal. Assim, Giebner *et al.* (2020) e Tang *et al.* (2012) recomendam seu uso para

discriminação de níveis taxonômicos mais elevados. Contudo, de acordo com Leite *et al.* (2021), essas comparações são limitadas devido ao propósito para o qual os marcadores são propostos, à biblioteca de referência de onde são extraídos e aos limiares utilizados.

Vale ressaltar que, durante a PCR específica para cada ordem, apenas a espécie-alvo *Plagiogramma* sp. não revelou similaridade com nenhum outro organismo fitoplanctônico. Na PCR geral, *Plagiogramma* sp., *Euglena* sp. e *Ankistrodesmus* sp. não apresentaram boa cobertura de espécies dentro da mesma ordem ou de ordens diferentes.

Os melhores primers foram selecionados com base na capacidade de identificar um maior número de ordens distintas na PCR *in silico* (Tabela 1). Desse modo, eles conseguiram identificar entre 40 e 81 espécies de ordens diferentes e amplificaram fragmentos de 100 a 300 pb. Além disso, os primers selecionados possuem um tamanho entre 19 e 22 pb, temperaturas de fusão (*Melting*) entre 55,1°C e 58°C e uma concentração de GC% variando de 40,91% a 57,89%.

Tabela 1- Descrição dos primers escolhidos.

Primers	Sequência (5'→ 3')	Pares de bases	GC	Melt temp (°C)	Self dimer	Hairpin (°C/kcla.mol ⁻¹)
DFS- F1	TGATTACGTCCCTGCCCTTTG (F)	21	52.38	57.3	-6.3	17.9/0.47
	TGACACCAATAAAGGCAACAGT (R)	22	40.91	54.9	-3.14	25.7/-0.04
DFS-F2	TTTGTACACACCGCCCGTC (F)	19	57.89	58.0	-7.55	14.8/0.65
	GTCCCAGTAAAGGAACCAGACC (R)	22	54.55	57.2	-4.64	29.9/-0.28
DFS- F3	GCCAGTAGTCATACGCTCGTC (F)	21	57.14	57.3	-4.95	23.5/0.1
	TCACCACATCATCCCGAAGG (R)	21	55.00	56.9	-3.61	44.7/-0.89
DFS-F4	ATTGGAGGGCAAGTCTGGTG (F)	20	55.00	57.5	-3.9	32.2/-0.63
	CACACAATTAGTGCGACCGAC (R)	21	55.00	56.3	-5.36	46.2/-1.74
DFS-F5	GGGGAGTATGGTCGCAAGATT (F)	21	52.38	56.9	-3.61	43.4/-0.74
	GTCCCTCTAAGAAGTTGCCC (R)	20	55.00	55.1	-3.17	16/0.59

Fonte: Elaborado pelos autores (2024).

Tais resultados se assemelham aos registrados por Sharma (2021), que em seu estudo indicou que os melhores resultados são obtidos em temperaturas de melting entre 52 e 60 °C. Muniz *et al.* (2020) destacam que a concentração de GC dos primers projetados está dentro das diretrizes recomendadas, entre 40% e 60%, o que é semelhante aos parâmetros utilizados nos estudos de Kalendar; Lee; Schulman (2011), que investigaram ferramentas para o design de primers *in silico*.

Quanto aos valores dos self-dímeros, a variação ocorreu entre -3,9 e -7,55. O hairpin apresentou, para todas as espécies de fitoplâncton-alvo, valores de temperatura de melting abaixo de 50 °C. Ao observar o Delta G, houve uma variação de -1,74 a 0,65. Esses parâmetros são importantes no desenho de primers, pois evitam desperdício durante a execução de uma PCR *in vitro* (Nascimento, 2021).

4 CONCLUSÃO

Os resultados obtidos neste estudo revelaram que a comunidade fitoplanctônica do CESM é similar à observada em outros estuários amazônicos, especialmente pela dominância de espécies de diatomáceas em relação aos demais grupos. O filo Bacillariophyta obteve o maior número de sequências disponíveis nos bancos de dados públicos, o que auxilia diretamente na realização de estudos ligados ao metabarcoding.

Os primers desenvolvidos neste estudo (DFS-F1, DFS-F2, DFS-F3, DFS-F4 e DFS-F5) foram eficazes na identificação de fitoplâncton a nível de espécie na região. Apesar de algumas restrições apontadas por autores, o marcador 18S mostrou-se eficiente para a comunidade fitoplanctônica estudada. Esses resultados enriquecem a literatura sobre protocolos eDNA, auxiliando na escolha do marcador molecular e dos primers, tornando o trabalho mais eficiente e prático.

AGRADECIMENTOS

Gostaríamos de agradecer ao Laboratório de Biodiversidade Molecular (LaBimol) pelo suporte, à Fundação de Amparo à Pesquisa do Estado do Maranhão (FAPEMA) e ao Programa de Pós-Graduação em Biodiversidade e Biotecnologia (BIONORTE) pelo apoio financeiro, e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio acadêmico.

REFERÊNCIAS

- AGIRBAS, E. *et al.* Phytoplankton community composition in the south-eastern Black Sea determined with pigments measured by HPLC-CHEMTAX analyses and microscopy cell counts. **Journal of the Marine Biological Association of the United Kingdom**, v. 95, n. 1, p. 35–52, 2015. <https://doi.org/10.1017/S0025315414001040>
- ARDURA, A. Species-specific markers for early detection of marine invertebrate invaders through eDNA methods: gaps and priorities in GenBank as database example. **Journal for Nature Conservation**, v. 47, p. 51-57, 2019. <https://doi.org/10.1016/j.jnc.2018.11.005>
- AQUINO, E. P. *et al.* Phytoplankton in a tropical estuary, Northeast Brazil: composition and life forms. **Check List**, v. 11, p. 1633, 2015. <https://doi.org/10.15560/11.3.1633>
- BARSANTI, L.; BIRINDELLI, L.; GUALTIERI, P. Water monitoring by means of digital microscopy identification and classification of microalgae. **Environmental Science: Processes & Impacts**, v. 23, p. 1443–1457, 2021. <https://doi.org/10.1039/D1EM00258A>
- BILBAO, J. *et al.* Phytoplankton community composition in relation to environmental variability in the Urdaibai estuary (SE Bay of Biscay): Microscopy and eDNA metabarcoding. **Marine Environmental Research**, v. 191, p. 106175, 2023. <https://doi.org/10.1016/j.marenvres.2023.106175>
- CARVALHO NETO, F. C. M. *et al.* Fitoplâncton do estuário do Emboraí Velho (Nova Olinda - PA). **Observatorio de la Economía Latinoamericana**, v. 21, n. 6, p. 5492-5519, 2023. <https://doi.org/10.55905/oelv21n6-122>
- CAVALCANTI, L.F. *et al.* Structure of microphytoplankton community and environmental variables in a macrotidal estuarine complex, São Marcos Bay, Maranhão – Brazil. **Brazilian Journal of Oceanography**, v. 66, p. 283–300, 2018. <https://doi.org/10.1590/s1679-87592018021906603>
- CAVALCANTI, L.F. *et al.* Patterns of phytoplankton structure in response to environmental gradients in a macrotidal estuary of the Equatorial Margin (Atlantic coast, Brazil). **Estuarine, Coastal and Shelf Science**, v. 245, p. 106969, 2020. <https://doi.org/10.1016/j.ecss.2020.106969>
- CAVALCANTI-LIMA, L.F. *et al.* Effects of climate, spatial and hydrological processes on shaping phytoplankton community structure and β -diversity in an estuary-ocean continuum (Amazon continental shelf, Brazil). **Journal of Sea Research**, v. 193, p. 102384, 2023. <https://doi.org/10.1016/j.seares.2023.102384>

CERUTTI-PEREYRA F. *et al.* Identification of Rays through DNA Barcoding: An Application for Ecologists. **PLOS ONE**, v. 7 (6), p. e36479–e36479, 2012. <https://doi.org/10.1371/journal.pone.0036479>

CLARKE, L. J. *et al.* Effect of marker choice and thermal cycling protocol on zooplankton DNA metabarcoding studies. **Ecology and Evolution**, v. 7, p. 873 – 883, 2017. <https://doi.org/10.1002/ece3.2667>

COMTET, T. *et al.* DNA (meta)barcoding of biological invasions: a powerful tool to elucidate invasion processes and help managing aliens. **Biological invasions**, v. 17, p. 905-922. 2015. <https://doi.org/10.1007/s10530-015-0854-y>

COUNTWAY, P. D. *et al.* Protistan diversity estimates based on 18S rDNA from seawater incubations in the Western North Atlantic. **Journal of Eukaryotic Microbiology**, v. 52 (2), p. 95-106, 2005. <https://doi.org/10.1111/j.1550-7408.2005.05202006.x>

DE VARGAS, C. *et al.* Eukaryotic plankton diversity in the sunlit global ocean. **Science**, v. 348 p. 1261605, 2015. <https://doi.org/10.1126/science.1261605>

DINIZ, C. *et al.* Brazilian Mangrove Status: Three Decades of Satellite Data Analysis. **Remote Sensing**, v.11, n. 7, p. 808, 2019. <https://doi.org/10.3390/rs11070808>

EL-ROBRINI, M. *et al.* In: MUEHE, D. **Erosão e progradação no litoral brasileiro**. Brasília: MMA, p. 41-86, 2006.

FIELD, C.B. *et al.* Primary production of the biosphere: Integrating terrestrial and oceanic components. **Science**, v. 281, p. 237–240, 1998.

GAMA, L. R. *et al.* Microfitoplâncton das baías do Golfão Maranhense e Litoral Oriental do Estado do Maranhão. **Boletim do Laboratório de Hidrobiologia**, v. 24, n. 1, p. 13–26, 2011.

GAONKAR, C. C.; CAMPBELL, L. A full-length 18S ribosomal DNA metabarcoding approach for determining protist community diversity using Nanopore sequencing. **Ecology and Evolution**, v. 14, 2024. <https://doi.org/10.1002/ece3.11232>

GIEBNER, H. *et al.* Comparing diversity levels in environmental samples: DNA sequence capture and metabarcoding approaches using 18S and COI genes. **Molecular Ecology Resources**, v. 20, p. 333–1345, 2020. <https://doi.org/10.1111/1755-0998.13201>

GONG, W.; MARCHETTI, A. Estimation of 18S gene copy number in marine eukaryotic plankton using a next-generation sequencing approach. **Frontiers in Marine Science**, v. 6, p. 219, 2019. <https://doi.org/10.17615/th53-ak73>

GUIRY, M.D.; GUIRY, G.M., 2023. **AlgaeBase**. World-wide electronic publication. National University of Ireland, Galway.

HALL, T.A. BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. **Nucleic Acids Symposium Series**, v. 41, p. 95-98, 1999.

JESUS, W.B. *et al.* Biomarcadores enzimáticos e histológicos em brânquias de *Ucides cordatus* (Linnaeus, 1763) (Crustacea, Brachyura, Ucididae) indicativos de impactos ambientais em uma região portuária do nordeste do Brasil. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 72, n. 4, p. 1413–1423, 2020. <https://doi.org/10.1590/1678-4162-11666>

KALENDAR, R.; LEE, D.; SCHULMAN, A. H. Java web tools for PCR, *in silico* PCR, and oligonucleotide assembly and analysis. **Genomics**, v. 98, n. 2, p.137–144, 2011. <https://doi.org/10.1016/j.ygeno.2011.04.009>

KALENDAR, R. *et al.* FastPCR: An *in silico* tool for fast primer and probe design and advanced sequence analysis. **Genomics**, v. 109 (3-4), p. 312–319, 2017. <https://doi.org/10.1016/j.ygeno.2017.05.005>

KECK, F. *et al.* Freshwater biomonitoring in the information age. **Frontiers in Ecology and the Environment**, v. 15, p. 266–274, 2017. <https://doi.org/10.1002/fee.1490>

KOLMANN, M. A. *et al.* DNA barcoding reveals the diversity of sharks in Guyana coastal markets. **Neotropical Ichthyology**, v. 15, (4), p. e170097, 2017. <http://dx.doi.org/10.1590/1982-0224-20170097>

KOOISTRA, W. H. *et al.* Comparative molecular and morphological phylogenetic analyses of taxa in the Chaetocerotaceae (Bacillariophyta). **Phycologia**, v. 49, n. 5, p. 471–500, 2010. <https://doi.org/10.2216/09-59.1>

LEE, J. *et al.* Methodological advances and future directions of microalgal bioassays for evaluation of potential toxicity in environmental samples: A review. **Environment International**, v. 173, p. 107869, 2023. <https://doi.org/10.1016/j.envint.2023.107869>

LEE, M.; WON, N.I.; BAEK, S.H. Comparison of HPLC Pigment Analysis and Microscopy in Phytoplankton Assessment in the Seomjin River Estuary, Korea. **Sustainability**, v. 12, n. 4, p. 1675, 2020. <https://doi.org/10.3390/su12041675>

LEITE, B.R. *et al.* Comparing species detection success between molecular markers in DNA metabarcoding of coastal macroinvertebrates. **Metabarcoding and Metagenomics**, v. 5, p. e70063, 2021. <https://doi.org/10.3897/mbmg.5.70063>

LIU, M. *et al.* Diatom DNA barcodes for forensic discrimination of drowning incidents. **FEMS Microbiology Letters**, v. 367, n. 17, 2020. <https://doi.org/10.1093/femsle/fnaa145>

MARINHO, E. I. N. **Desenho de primers *in silico* para o gene 16S rRNA**. 2021. 41f. Trabalho de Conclusão de Curso (Licenciatura em Ciências Biológicas) - Instituto Federal do Piauí, Campus Pedro II, 2021.

MARTINS, T. O. M. **DNA Barcode e a identificação de espécies de Raias exploradas pela pesca**. 2021. 25 f. Trabalho de conclusão de curso (Graduação em Bacharelado Interdisciplinar em Ciência e Tecnologia do Mar) - Instituto do Mar, Universidade Federal de São Paulo, Santos, 2021.

MATOS, J. B. *et al.* Caracterização quali-quantitativa do fitoplâncton da zona de arrebentação de uma praia amazônica. **Acta Botanica Brasilica**, v. 26, n. 4, p. 979–990, 2012. <https://doi.org/10.1590/s0102-33062012000400028>

MIETTINEN, A. Diatoms in Arctic regions: Potential tools to decipher environmental changes. **Polar Science**, v. 18, p. 220–226, 2018. <https://doi.org/10.1016/j.polar.2018.04.001>

MORAIS, R.C.S. **Ensaio de multiplex PCR em tempo real (TaqMan probe) para identificação de Leishmania spp. relacionadas com a etiologia da leishmaniose tegumentar Americana**. 2019. 129f. Tese (Doutorado) – Fundação Oswaldo Cruz, Centro de Pesquisas Aggeu Magalhães, CPQAM, Recife, 2019.

MUNIZ, H. V. L. *et al.* Desenvolvimento de primers *in silico* de *Mycobacterium tuberculosis* da região 16s para o diagnóstico da tuberculose. **Interfaces Científicas - Saúde e Ambiente**, v. 8, n. 2, p. 77–86, 2021. <https://doi.org/10.17564/2316-3798.2020v8n2p77-86>

NASCIMENTO, J. Í. S. **Identificação e validação *in silico* de potenciais marcadores moleculares para o gênero Leishmania**. 2021. 59f. Trabalho de Conclusão de Curso (Licenciatura em Ciências Biológicas) Instituto Federal do Piauí, Campus Pedro II, 2021.

NCBI, 2016. <https://www.ncbi.nlm.nih.gov/genbank/>

OLIVEIRA, S. R. S. D. *et al.* Enzymatic and Histological Biomarkers in *Ucides cordatus* (Crustacea, Decapoda) in an Industrial Port on the North Coast of Brazil. **Bulletin of Environmental Contamination and Toxicology**, v. 102, p. 802–810, 2019. <https://doi.org/10.1007/s00128-019-02594-1>

PATEL, K. K.; SCHEIBLE, M. K. R.; MEIKLEJOHN, K. A. Interlaboratory study to assess the practical utility of OSAC proposed standard 2021-S-0006: Standard for the use of GenBank for taxonomic assignment of wildlife. **Forensic Science International:**

Animals and Environments, v. 4, p. 100067, 2023.
<https://doi.org/10.1016/j.fsiae.2023.100067>

PAWLOWSKI, J. *et al.* Environmental DNA metabarcoding for benthic monitoring: A review of sediment sampling and DNA extraction methods. **Science of The Total Environment**, v. 818, p. 151783, 2022. <https://doi.org/10.1016/j.scitotenv.2021.151783>

PELT-VERKUIL, E.; VAN BELKUM, A.; HAYS, J. P. **Principles and technical aspects of PCR amplification**, Springer, 332 pp, 2008.

QUEIROZ, J.B.M. *et al.* Phytoplankton of the shipping sector of São Marcos Bay (Amazon Coast): A potential risk area for the establishment of non-indigenous species. **Regional Studies in Marine Science**, v. 49, p. 102121, 2022. <https://doi.org/10.1016/j.rsma.2021.102121>

RAVEH, O. *et al.* Phytoplankton response to N-rich well amelioration brines: a mesocosm study from the southeastern Mediterranean Sea. **Marine Pollution Bulletin**, v. 146, p. 355–365, 2019. <https://doi.org/10.1016/j.marpolbul.2019.06.067>

RIBEIRO, E. B. *et al.* Gill and hepatic histological alterations in *Sciades herzbergii* resulting from trace element contamination in the Port of São Luiz, Brazil. **Brazilian Journal of Biology**, v. 83, p. e274069, 2023. <https://doi.org/10.1590/1519-6984.274069>

ROCHA, B.S. *et al.* Assessing resilience and sensitivity patterns for fish and phytoplankton in French lakes. **Global Ecology and Conservation**, v. 43, p. e02458, 2023. <https://doi.org/10.1016/j.gecco.2023.e02458>

SCHULER, G. D. Sequence mapping by electronic PCR. **Genome research**, v. 7, n. 5, p. 541-550, 1997. <https://doi.org/10.1101/gr.7.5.541>

SHARMA, M. Basic Concepts of Primer Designing: A Mini Review. **International Journal of Latest Trends in Engineering and Technology**, v. 17, n. 4, p. 010-012, 2021. <http://dx.doi.org/10.21172/1.174.03>

SINGH, P. *et al.* Marine sedimentary ancient DNA from Antarctic diatoms. **Palaeogeography, Palaeoclimatology, Palaeoecology**, v. 640, p. 112090, 2024. <https://doi.org/10.1016/j.palaeo.2024.112090>

SUN, Y. *et al.* Disentangling effects of river inflow and marine diffusion in shaping the planktonic communities in a heavily polluted estuary. **Environmental Pollution**, v. 267, p. 115414, 2020. <https://doi.org/10.1016/j.envpol.2020.115414>

TABERLET, P. *et al.* Soil sampling and isolation of extracellular DNA from large amount of starting material suitable for metabarcoding studies. **Molecular Ecology**, v. 21, p. 1816–1820, 2012. <https://doi.org/10.1111/j.1365-294x.2011.05317.x>

TANG, C. Q. *et al.* The widely used small subunit 18S rDNA molecule greatly underestimates true diversity in biodiversity surveys of the meiofauna. **Proceedings of the National Academy of Sciences**, v. 109, n. 40, p. 16208 - 16212, 2012. <https://doi.org/10.1073/pnas.1209160109>

WANGENSTEEN, O. S. *et al.* DNA metabarcoding of littoral hard-bottom communities: high diversity and database gaps revealed by two molecular markers. **PeerJ**, v. 6, p. e4705, 2018. <https://doi.org/10.7717/peerj.4705>

YOON, S. J. *et al.* Phytoplankton assemblage responses to massive freshwater inputs and anthropogenic toxic substances contamination in the Geum River Estuary, South Korea. **Marine Pollution Bulletin**, v. 199, p. 116020, 2024. <https://doi.org/10.1016/j.marpolbul.2023.116020>

ZHANG, G.K. *et al.* Metabarcoding using multiplexed markers increases species detection in complex zooplankton communities. **Evolutionary Applications**, v. 11, p. 1901–1914, 2018. <https://doi.org/10.1111/eva.12694>

ZIMMERMANN, J. *et al.* Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. **Molecular Ecology Resources**, v. 15, p. 526–542, 2014. <https://doi.org/10.1111/1755-0998.12336>