



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
DEPARTAMENTO DE BIOLOGIA
CURSO DE BACHARELADO EM CIÊNCIAS BIOLÓGICAS

BRUNA CAITANO SERPA

**Análise da diversidade viral em dados de (meta)transcriptomas de amostras de
Mangue-Vermelho (*Rhizophora mangle*)**

Recife

2026

BRUNA CAITANO SERPA

Análise da diversidade viral em dados de (meta)transcriptomas de amostras de Manguê-Vermelho (*Rhizophora mangle*)

Trabalho de Conclusão de Curso apresentado ao Curso de Bacharelado em Ciências Biológicas da Universidade Federal Rural de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Ciências Biológicas.

Orientador (a): Profa. Dra. Rosana Blawid

Recife

2026

Dados Internacionais de Catalogação na Publicação
Sistema Integrado de Bibliotecas da UFRPE
Bibliotecário(a): Suely Manzi – CRB-4 809

S481a Serpa, Bruna Caitano.
Análise da diversidade viral em dados de
(meta)transcriptomas de amostras de mangue vermelho
(*Rhizophora mangle*) / Bruna Caitano Serpa. – Recife,
2026.
50 f.; il.

Orientador(a): Rosana Blawid.

Trabalho de Conclusão de Curso (Graduação) –
Universidade Federal Rural de Pernambuco, Bacharelado
em Ciências Biológicas, Recife, BR-PE, 2026.

Inclui referências.

1. Manguezal. 2. Virus. 3. RNA. 4. Biociências - Estudo e
ensino 5. Ecossistemas costeiros. I. Blawid, Rosana, orient.
II. Título

CDD 574

BRUNA CAITANO SERPA

Análise da diversidade viral em dados de (meta)transcriptomas de amostras de Manguê-Vermelho (*Rhizophora mangle*)

Trabalho de Conclusão de Curso apresentado ao Curso de Bacharelado em Ciências Biológicas da Universidade Federal Rural de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Ciências Biológicas.

Aprovado em: 10/02/2026

BANCA EXAMINADORA

Profa. Dra. Rosana Blawid (Orientadora)
Universidade Federal Rural de Pernambuco

Ma. Dayla Bezerra (Examinador Interno)
Universidade Federal Rural de Pernambuco

Prof. Dr. Alejandro Risco (Examinador Externo)
Universidad Nacional Agraria La Molina

Dedico este trabalho a todos que direta ou indiretamente me ajudaram a chegar até aqui. Acima de tudo, ao Deus a quem sirvo, Jeová, o maior de todos os apoiadores e o que acompanhou e abençoou minha trajetória em todos os detalhes.

Vejam! Deus é a minha salvação.
Confiarei e não sentirei pavor;
Pois Jah Jeová é a minha força e o meu poder,
E ele se tornou a minha salvação. (Isaías 12:2)

RESUMO

Os manguezais são ecossistemas costeiros altamente produtivos, caracterizados por intensa ciclagem de nutrientes e elevada diversidade microbiana, incluindo comunidades virais ainda pouco compreendidas. O presente trabalho teve como objetivo caracterizar a diversidade e a ocorrência de vírus em dados (meta)transcriptômicos de diferentes tecidos de *Rhizophora mangle*. Foram analisadas 20 das 48 amostras provenientes de raízes, caule, folha e meristema. As amostras foram submetidas à montagem de genomas por meio da metodologia do *de novo* assembly, bem como à aplicação de ferramentas de alinhamento de sequências. Os resultados indicaram uma predominância de vírus pertencentes à classe *Megaviricetes*, cujo membros têm como hospedeiros principalmente algas e protistas aquáticos. Não foram detectados fitovírus, ou seja, vírus que infectam plantas, indicando que a comunidade viral identificada não está relacionada à infecção direta dos tecidos vegetais, mas provavelmente reflete a dinâmica ecológica do ecossistema de manguezal. A presença desses vírus é interpretada como uma associação indireta com *R. mangle*, possivelmente mediada pela presença de microalgas aderidas às superfícies vegetais e aos microambientes associados. Dessa forma, o estudo contribui para o conhecimento sobre a diversidade viral em manguezais e reforça a importância ecológica dos vírus eucarióticos na dinâmica de ecossistemas costeiros.

Palavras-chave: manguezal; mangue-vermelho; viroma; *Megaviricetes*; metatranscriptômica.

ABSTRACT

Mangroves are highly productive coastal ecosystems, characterized by intense nutrient cycling and high microbial diversity, including viral communities that are still poorly understood. The aim of this study was to characterize the diversity and the occurrence of viruses in (meta)transcriptomic data from different tissues of *Rhizophora mangle*. Twenty out of forty eight samples from roots, stems, leaves, and meristems were analyzed, each sample associated with different environmental conditions, such as soil hydrology and vegetation composition. The samples were subjected to *de novo* assembly and sequence alignment tools. The results showed a predominance of viruses belonging to the class *Megaviricetes*, which has algae and aquatic protists as hosts. No phytoviruses, viruses that infect plants, were found, indicating that the viral community detected is not related to direct infection of plant tissues, but likely reflects the ecological dynamics of the mangrove ecosystem. The presence of these viruses is interpreted as an indirect association with *R. mangle*, possibly mediated by the presence of microalgae attached to plant surfaces and associated microenvironments. Therefore, this study contributes to the knowledge of viral diversity in mangroves and reinforces the ecological importance of eukaryotic viruses in the dynamics of coastal ecosystems.

Keywords: mangrove; red mangrove; virome; *Megaviricetes*; metatranscriptomics.

LISTA DE ILUSTRAÇÕES

Figura 1 – Exemplo de manguezal em Semporna, Malásia	15
Figura 2 – <i>Rhizophora mangle</i> L.	16
Gráfico 1 – Identidade média das 20 amostras (BLASTn)	36
Gráfico 2 – Heatmap da média do número de hits de BLASTn para os principais vírus detectados em tecidos vegetais (raiz, folha e meristema)	37
Gráfico 3 – Frequência de ocorrência dos vírus mais abundantes identificados no conjunto das 20 amostras	38

LISTA DE TABELAS

Tabela 1 –	Dados dos SRA encontrados no NCBI	25
Tabela 2 –	Dados das amostras pós-trimagem e montagem dos contigs	28
Tabela 3 –	Resultados de BLASTn (identidade acima de 70% e tamanho acima de 500 nt)	30

SUMÁRIO (OBRIGATÓRIO)

1	INTRODUÇÃO	14
1.1	Manguezal	15
1.1.1	<i>Rhizophora mangle</i> L.	16
1.2	Montagem de Genoma	17
1.3	Fitovírus	18
2	OBJETIVOS	20
2.1	Objetivos gerais	20
2.2	Objetivos específicos	20
3	MATERIAIS E MÉTODOS	21
3.1	Análise utilizando Serratus	21
3.2	Análise de qualidade de <i>reads</i> e trimagem	21
3.3	Montagem de genomas com Megahit e BLASTx com Diamond	22
3.4	Análise de BLASTn no Geneious	22
4	RESULTADOS E DISCUSSÃO	24
4.1	Resultado do Serratus e características dos dados de SRA ...	24
4.2	Resultados da trimagem e de novo assembly	27
4.3	Resultado do BLASTx	29
4.4	Resultado do BLASTn	29
5	CONCLUSÃO	41
	REFERÊNCIAS	43

1. INTRODUÇÃO

Os manguezais são ecossistemas de transição entre rios e mares, característicos de regiões tropicais e subtropicais. Esses ambientes exercem um papel fundamental na reprodução, alimentação e proteção de diversas espécies animais e vegetais (Krug et al., 2007).

Uma das principais espécies de mangue é o *Rhizophora mangle*, ou mangue-vermelho, que desenvolveu adaptações fisiológicas e morfológicas ao ambiente inundado e salino típico desse ecossistema. Suas raízes impedem a erosão das margens, retêm sedimentos e fornecem um ambiente propício para a sobrevivência de diversas espécies marinhas (Nanni et al., 2005).

Além de sua importância ecológica, os manguezais podem abrigar diversos microrganismos, como vírus, que podem impactar tanto a flora quanto a fauna. Os vírus são seres vivos que usam a energia e maquinaria celular do hospedeiro para disseminação do seu material genético. Especificamente falando, os fitovírus, com um grande potencial de disseminação,— tendo os insetos como principais vetores, podem causar diversos danos à saúde das plantas e à agricultura (Mabola, 2021).

A identificação e caracterização dos vírus são frequentemente facilitadas por técnicas avançadas de sequenciamento, como o *Next-Generation Sequencing* (NGS, sequenciamento de alto rendimento), que permitem compreender melhor sua expressão e composição genômica e seu impacto no ecossistema. O *de novo assembly* oferece uma ferramenta poderosa para a descoberta de novas espécies e o estudo desses vírus (Liao et al., 2019).

Com o aumento da poluição e do estresse ambiental nos manguezais, torna-se fundamental avaliar a presença de vírus nesse ecossistema e seus potenciais efeitos, que podem afetar inclusive a agricultura. A utilização de técnicas metagenômicas/metatranscriptômicas pode representar um ponto de partida essencial para a obtenção de dados cruciais sobre a dinâmica viral e suas interações com o ambiente (Silva, 2021).

1.1 Manguezal

O manguezal é um sistema ecológico que, embora apresente menor riqueza em espécies, é marcado por elevada abundância de populações de espécies vegetais e animais altamente adaptados a solos frequentemente inundados por maré. Essas condições ambientais tornam o manguezal um ambiente favorável à alimentação, reprodução e proteção de diversas espécies (Firme, 2003; Krug et al., 2007; Quinões, 2000).

O termo “Manguezal” é usado para definir um ecossistema formado por uma diversidade de espécies vegetais, arbóreas e arbustivas. Já o termo “Mangue” descreve as espécies vegetais que vivem no manguezal e vem do malaio "manggimanggi" e do inglês “mangrove” (Nanni et al., 2005). Por ser o elo de ligação entre o mar e a terra firme, os manguezais recebem grandes quantidades de compostos orgânicos, como restos de folhas, excrementos de animais e sais minerais provenientes da própria maré. Esse aporte confere ao ecossistema um papel fundamental no condicionamento biológico, favorecendo elevados níveis de produtividade (Schaeffer-Novelli, 1991 apud Nanni et al., 2005). Além disso, os manguezais (**Figura 1**) fornecem subsídios para outras fontes de energia, uma vez que apresentam excedente de matéria orgânica, que pode ser exportada para ecossistemas adjacentes ou armazenada localmente (Odum, 1988).

Figura 1 – Exemplo de manguezal em Pernambuco, Brasil



Fonte: https://commons.wikimedia.org/wiki/File:Manguezal_Pernambuco_2.jpg

1.1.1 *Rhizophora mangle* L. (Mangue-Vermelho)

O mangue-vermelho (*Rhizophora mangle* L., **Figura 2**) é uma das principais espécies de mangue, apresentando destacada importância econômica e ecológica. Do ponto de vista econômico, destaca-se seu elevado potencial farmacológico, com propriedades antissépticas, antimicrobianas, antiulcerogênicas, antiinflamatórias, antioxidantes, antidiarreicas, inseticidas e larvicidas. Dessa forma, a espécie pode, futuramente, constituir uma importante fonte de produtos naturais, tornando-se um grande aliado para a resolução de problemas de saúde de grande relevância (Regalado et al., 2016).

Assim como outras espécies de mangues, O mangue-vermelho desenvolveu adaptações fisiológicas e morfológicas que tornou possível a sobrevivência em ambientes encharcados e salinos, tais como mecanismos fisiológicos de filtragem de água salobra (Almeida et al., 2000), desenvolvimento de rizóforos, raízes que permitem a fixação ao substrato lamoso (Menezes, 2006 apud Prado et al., 2013), a presença de lenticelas, que possibilitam a troca gasosa entre a planta e o meio externo, e a viviparidade (Nanni et al., 2005). Assim como ocorre com outras espécies de mangue, a importância ecológica de *R. mangle* evidencia-se por atuar como um importante habitat para peixes, moluscos e crustáceos, os quais encontram nesse ambiente condições favoráveis à sobrevivência (Schaeffer- Novelli, 1995). Além disso, suas raízes funcionam como filtros na retenção dos sedimentos, além de impedir erosão. A espécie também desempenha um papel relevante como repositório de diversidade genética, sendo fundamental para a recuperação de áreas degradadas (Nanni et al., 2005).

Figura 2 – *Rhizophora mangle* L.



Fonte: <https://mangueslimposufsb.blogspot.com>

1.2 Montagem de Genoma

A montagem de genoma é o processo de reconstrução de sequências de DNA a partir de fragmentos obtidos por sequenciamento (Liao et al., 2019; Miller et al., 2010). Segundo Sohn e Nam (2018), a montagem do *de novo* de um genoma completo é frequentemente comparada a um quebra-cabeça gigante com milhões de peças. Como a sequência de um genoma inteiro não pode ser lida de uma só vez, mesmo com tecnologias de ponta, é necessário montar inúmeras sequências fragmentadas, por meio da identificação de regiões sobrepostas, até alcançar a reconstrução em nível cromossômico. A montagem de genomas possui aplicações em diversas áreas da pesquisa biológica, incluindo estudos sobre evolução de espécies, a identificação de genes específicos e a detecção de doenças (Sohn; Nam, 2018).

Os dados utilizados para a montagem de genomas são gerados por tecnologias de sequenciamento, dentre as quais destaca-se o NGS. Essa abordagem integra etapas como a preparação de amostras, sequenciamento, análises de imagem, alinhamento de genomas e métodos de montagem, permitindo a geração, a baixo custo, de grandes volumes de dados - em alguns casos, superiores a um bilhão de leituras curtas por execução do instrumento (Metzker, 2009).

Para o descobrimento de novos vírus, o processo principal é o *de novo assembly*, ou seja, a montagem de fragmentos de DNA (ou *reads*) em conjuntos sobrepostos (*contigs*) sem um genoma prévio como referência (Blawid et al., 2017). Os

algoritmos de montagem do *de novo assembly* são divididos em três estruturas básicas: OLC (*Overlap-layout-consensus*) graph, de *Bruijn graph* e *String graph*, sendo atualmente o de *Bruijn graph* o mais eficiente em resolver problemas de montagem (Blawid et al., 2017; Liao et al., 2019).

A estratégia *de novo assembly* apresenta alguns desafios e problemas. Segundo Liao (2019), existem quatro principais desafios. Primeiramente, erros de sequenciamento podem levar à construção de grafos de *Bruijn* excessivamente complexos, resultando em montagens insatisfatórias. Em segundo lugar, vieses de sequenciamento — como os observados em plataformas como a Illumina, que favorecem regiões com alto conteúdo de GC — podem gerar cobertura desigual ao longo do genoma. Em terceiro, a complexidade de regiões repetitivas pode ocasionar lacunas, rearranjos na montagem e resultados inconsistentes ou discrepantes. Por fim, muitos programas de montagem exigem elevados recursos computacionais, incluindo grande quantidade de memória RAM e, dependendo do tipo e do tamanho do genoma, longos tempos de processamento.

1.3 Fitovírus

Vírus são seres microscópicos, parasitas intracelulares obrigatórios, ou seja, extremamente dependentes do interior da célula para se replicarem. Isso significa que utilizam a energia e maquinaria molecular celular da célula hospedeira para multiplicarem novas partículas virais infectivas, ou vírions (Louten, 2016).

Vírus de plantas são um grande problema na agricultura por causarem doenças devastadoras. Por exemplo, o vírus do mosaico do tabaco (TMV) infecta plantas como o tabaco (*Nicotiana tabacum*) e a tomate (*Solanum lycopersicum*), causando mosaico foliar, deformação, nanismo e necrose (Lomonossoff, 2018). Os fitovírus desenvolveram várias estratégias de dispersão, principalmente contando com auxílio de vetores como fungos, nematóides, bactérias, ácaros e principalmente, insetos (Mabola, 2021).

Ao se alimentarem das plantas, os insetos vetores liberam partículas virais, que podem se movimentar pelos plasmodesmos de forma intercelular, e a longa distância, através do sistema vascular. Da mesma forma, ao se alimentarem de

plantas infectadas, esses insetos podem adquirir os vírions e transmiti-los a novas plantas (DURÁN, 2024).

Os fitovírus possuem uma ampla gama de hospedeiros. Uma só espécie de vírus pode infectar várias espécies de plantas. Por exemplo, o vírus do mosaico do tomateiro (ToMV) infecta tomate, tabaco, rosa, cereja, girassol, pepino e lúpulo (Bréchignac, 2005 p.58). A identificação de vírus já conhecidos em mangues pode provocar mudanças significativas na forma como se compreende a propagação de fitovírus em plantas agrônômicas, especialmente se houver a possibilidade de essas plantas atuarem como hospedeiros naturais. A análise do comportamento desses vírus nesses hospedeiros pode levar a inovações nas estratégias de manejo e controle dessas espécies virais. Além disso, a descoberta de novas espécies virais pode contribuir para o aprimoramento da classificação dos vírus de plantas, um campo em constante atualização (King et Al., 2018).

Além disso, a interação dos manguezais com o microbioma desperta grande interesse entre pesquisadores, especialmente sob a perspectiva biotecnológica (Thatoi et al., 2013). No entanto, pouco se conhece sobre a ocorrência e comportamento de vírus nesse ecossistema, de modo que este trabalho pode abrir novas perspectivas e estimular pesquisas com maior enfoque nesta área.

Os manguezais sofrem cada vez mais com a poluição e o estresse ambiental, e há uma escassez nas pesquisas envolvendo este tema (Silva, 2021, p.76). Sabe-se que o estresse ecológico pode levar a mudanças nas propriedades de vírus patogênicos, facilitando sua dispersão (Bréchignac, 2005 p.58). Assim, os vírus podem ser ótimos indicadores ambientais (Zagorodini, 1999, p.86). Portanto, é importante avaliar quais são os vírus presentes nos manguezais e como eles podem afetá-lo. O sequenciamento de alto rendimento pode ser um começo de uma solução para este problema, tendo a capacidade de identificar vírus integrados no genoma da planta.

2. OBJETIVOS

2.1 Objetivos Gerais

Identificar a diversidade e quantificar a ocorrência viral em dados de metatranscriptômica (RNA-seq) depositados no NCBI (*National Center for Biotechnology Information*) de mangue vermelho (*Rhizophora mangle* L.), visando à compreensão da composição viral da planta e do ambiente de manguezal.

2.1 Objetivos específicos

1. Detectar, e analisar a diversidade viral em dados de metatranscriptômica (RNA-seq) de *R. mangle*;
2. Identificar as espécies virais mais frequentes nas amostras;
3. Relacionar a composição viral encontrada com o ecossistema de manguezal nos diferentes tecidos.

3. MATERIAIS E MÉTODOS

3.1 Análise utilizando Serratus

O projeto utilizou dados de *RNA-Seq* provenientes de materiais disponíveis no NCBI, especialmente no *Sequence Read Archive* (SRA) , um banco de dados que armazena sequências de DNA e RNA obtidos de sequenciamentos de nova geração. Houve pré análise usando a ferramenta Serratus (<https://serratus.io/explorer>), com o objetivo de auxiliar na escolha dos dados para serem utilizados. Foi escolhida a opção NT search, ou opção de busca de alinhamentos por nucleotídeos, e a ferramenta foi explorada por cada um dos 20 SRA, com identidade de alinhamento de no mínimo 75% e um score mínimo de 50.

3.2 Análise de qualidade de *reads* e trimagem

Para identificar os genomas virais presentes nessas amostras, seguimos o método descrito por Blawid et al., 2017, com modificações. A análise de qualidade dos *reads* foi realizado com o auxílio do programa FastQC v0.11.9 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), além de trimagem de adaptadores e por qualidade pelo TRIMMOMATIC v.0.36 (BOLGER et al., 2014). Esse processo foi importante para a remoção de sequências e regiões com baixa qualidade e que pudessem causar erros na análise dos dados. O parâmetro ILLUMINACLIP foi utilizado para remoção de adaptadores Illumina, e para isso foi utilizado o arquivo de sequência TruSeq3-PE.fa, com tolerância de 2 mismatches permitidos, *match* entre o adaptador e *reads* com pontuação 30 e com comprimento mínimo de 10 para corte (ILLUMINACLIP:TruSeq-PE.fa:2:30:10). O parâmetro LEADING é utilizado para remover bases com pouca qualidade. E o parâmetro TRAILING é usado para remover bases de baixa qualidade no fim da sequência. Para ambos, foi especificado uma qualidade mínima 3 (LEADING:3 TRAILING:3). O parâmetro SLIDINGWINDOW é uma ferramenta para remover bases com qualidade abaixo de uma média limite. Para o cálculo da média foi especificada uma 'janela' de 4 bases, com qualidade média 15 exigida. (SLIDINGWINDOW:4:15). *Reads* com comprimento menor que 36 nucleotídeos (nt) foram removidos (MINLEN:36). A

ferramenta BBTools (<https://sourceforge.net/projects/bbmap/>) foi usada para remoção de duplicatas. Para isso, foi usado a função `./clumpy.sh`.

3.3 Montagem de genomas com Megahit e BLASTx com Diamond

Após isso, foi realizado o '*de novo assembly*', ou seja, a construção de *contigs* com vários fragmentos de DNA, pelo MEGAHIT v.1.2.9 (Dinghua et al., 2015) usando valores *K* de 21 a 99. Foram utilizados os *contigs* obtidos para performar o BLASTx, ou seja, para identificar sequências de nucleotídeos similares com base no potencial de codificação, no Diamond v.2.1.11 (Buchfink, 2015), na qual o banco de dados foi construído manualmente a partir de sequências de aminoácidos do NCBI Vírus em agosto de 2025. Obtiveram-se os arquivos em *output* no formato DAA (DIAMOND alignment archive: `otfmt 100`), onde permitiu que o arquivo fosse utilizado no programa MEGAN v.6.25.10 (Huson et al., 2007), com os parâmetros de 70% de similaridade com *contigs* de no mínimo 300 aminoácidos (aa) para calcular e explorar o conteúdo taxonômico dos contigs, a fim de resumir e ordenar os resultados.

3.4 Análise de BLASTn no Geneious

O programa Geneious v.11.0.5 (<https://www.geneious.com>) foi usado para a realização de um BLASTn. Para a análise dos resultados de identidade de sequências, foram avaliados os parâmetros *E-value*, *bit score*, *pairwise identity* e o comprimento das sequências (*sequence length*), quantidades de *hits* para cada espécie identificada. Em cada amostra, foi feita uma tabela com os vírus mais encontrados (levando em consideração a identidade acima de 70% e tamanho acima de 500 nucleotídeos) e as medidas estatísticas adequadas para cada parâmetro. O *E-value* indica a probabilidade de um alinhamento acontecer ao acaso, quanto menor a pontuação, mais significativo é o alinhamento. Foi aplicado o cálculo da mediana pois é menos sensível às variações extremas do parâmetro. O *Bit Score* expressa a qualidade e confiabilidade do alinhamento, onde valores mais altos indicam alinhamentos mais confiáveis. Também foi representado pela mediana para evitar a influência de valores muito altos ou baixos. O *Pairwise Identity* é um

indicador do grau de similaridade entre as sequências, pois expressa a porcentagem de posições idênticas entre elas. Por apresentar menor dispersão e não possuir valores extremos, foi utilizada a média como cálculo para esse parâmetro. O comprimento de sequência foi levado em consideração para evitar interpretações enviesadas baseadas em alinhamentos muito curtos ou fragmentados, e foi expresso pela média.

4. RESULTADOS E DISCUSSÃO

4.1 Serratus e características dos dados de SRA

Foram encontrados 48 RNA-Seq de *Rhizophora mangle*, provenientes de quatro projetos diferentes. Destes, foram escolhidas 20 amostras (**Tabela 1**), provenientes de dois projetos, provenientes da Universidad del Valle e da UNICAMP. Os dados, com tamanho em média de 324,01 Mb, foram separados por tecidos (raíz, folha, meristema e tronco). Não foram encontrados resultados na ferramenta Serratus. Ao verificarmos a análise feita pelo NCBI, encontramos uma predominância de vírus da família *Siphoviridae* (**Tabela 1**). Embora muitas amostras não indicassem o vírus pertencente a tal família, ou o gênero ou a espécie, em algumas amostras houve o aparecimento da espécie de fago *Propionibacterium phage* PHL092M00. Ao realizar uma pesquisa no site do ICTV (*International Committee on Taxonomy of Viruses*), porém, foi descoberto que a família *Siphoviridae* foi desfeita, e que a espécie *Propionibacterium phage* PHL092M00 pertence à classe *Caudoviricetes* e ao gênero *Pahexavirus*. Com a mudança de classificação pelo ICTV, o gênero *Pahexavirus* não foi atribuído a nenhuma família. A mudança foi realizada com o objetivo de organizar a classificação com base em características genômicas, não apenas morfológicas (WEI; YOU, 2026).

O gênero *Pahexavirus* reúne um grupo de bacteriófagos de DNA de fita dupla (dsDNA) classificados na classe *Caudoviricetes*, vírus com estrutura do tipo *head-tail* (cabeça icosaédrica e cauda) que infectam principalmente bactérias do gênero *Cutibacterium*, antes chamado *Propionibacterium* (Li et Al. 2024). Este gênero vem chamando atenção pela sua estabilidade físico-química, atividade lítica e anti-biofilme, o que indica um forte potencial biotecnológico e terapêutico no controle de bactérias patogênicas (Li et Al. 2024).

A detecção de fagos em amostras de *R. mangle* pode ser explicada pelo fato de tais vírus estarem associados ao microbioma ambiental, não à planta em si. Membros da classe *Caudoviricetes* estão amplamente distribuídos em ambientes aquáticos e sedimentares, e a alta densidade microbiana, formação de biofilmes e elevada carga de matéria orgânica favorece a diversidade de fagos (Zhu et al., 2022). Além disso, a estabilidade físico-química e a sua atuação em biofilmes são características

compatíveis com condições estuarinas e variações de salinidade (Firme, 2003; Kim et al.), podendo tanto os vírus da classe *Caudoviricetes* como seus hospedeiros ter um potencial de refletir as características físico-químicas do solo. O achado de vírus do gênero *Pahexavirus* se relaciona mais com a presença de seus hospedeiros, seja na planta ou no sedimento, do que qualquer interação direta com os tecidos vegetais. O fato de bactérias do gênero *Propionibacterium* serem bactérias do microbioma cutâneo pode indicar uma influência antrópica vinda de efluentes e material de origem orgânica.

Tabela 1 – Dados dos SRA encontrados no NCBI e analisados neste estudo

Bioprojetos (número de acesso)	Estudo SRA (número de acesso)	Descrição (folha, raiz, meristema e caule)	Size (Mb)	Taxonomia NCBI (vírus, %)
PRJNA476540	SRR7366900	Folha	448.16 Mb	Siphoviridae: <0.01%
PRJNA476540	SRR7366901	Folha	437.43 Mb	Human betaherpesvirus 7: <0.01% Siphoviridae: <0.01%
PRJNA476540	SRR7366902	Folha	376.16 Mb	Propionibacterium phage QueenBey: <0.01% Streptococcus phage 5093: <0.01%
PRJNA476540	SRR7366903	Meristema	252.38 Mb	Siphoviridae: <0.01%
PRJNA476540	SRR7366904	Meristema	441.24 Mb	Enterobacteria phage S13: <0.01% Siphoviridae: <0.01%
PRJNA476540	SRR7366905	Meristema	144.29 Mb	Enterobacteria

				phage S13: <0.01%
PRJNA476540	SRR7366906	Meristema	114.71 Mb	Propionibacterium phage PHL113M01: <0.01%
PRJNA476540	SRR7366911	Meristema	289.98 Mb	Propionibacterium phage PHL092M00: <0.01%
PRJNA476540	SRR7366912	Meristema	291.90 Mb	Propionibacterium phage PHL092M00: <0.01%
PRJNA476540	SRR7366891	Raíz	395.55 Mb	Deep-sea thermophilic phage D6E, 0.01%
PRJNA476540	SRR7366892	Raíz	356.28 Mb	Siphoviridae, 0.01%
PRJNA476540	SRR7366893	Raíz	496.89 Mb	Propionibacterium phage Enoki: <0.01% Streptococcus virus 7201: <0.01% Streptococcus virus Sfi21: <0.01%
PRJNA476540	SRR7366895	Folha	767.19 Mb	não encontrado
PRJNA476540	SRR7366897	Raíz	551.62 Mb	Propionibacterium phage P100D: <0.01% Streptococcus phage 5093: <0.01%
PRJNA476540	SRR7366898	Raíz	456.75 Mb	Geobacillus vírus E3: <0.01%
PRJNA476540	SRR7366899	Folha	464.88 Mb	Propionibacterium phage

				PHL301M00: <0.01%
PRJNA1215819	SRR32130051	Folha	21.26 Mb	uncultured marine virus: <0.01%
PRJNA1215819	SRR32130052	Folha	39.13 Mb	Propionibacteriu m phage P100_A: <0.01%
PRJNA1215819	SRR32130053	Raiz	29.21 Mb	Beihai mollusks virus 1: <0.01%
PRJNA1215819	SRR32130054	Raiz	57.00 Mb	uncultured cyanophage: <0.01%

Fonte: Serpa, 2026.

4.2 Trimagem e de novo assembly

Ao fazer a análise das amostras no FastQC v0.11.9, pudemos verificar que as mesmas já tinham passado por algum tipo de processamento para a retirada dos adaptadores. O relatório do FastQC mostrou algumas classificações “atenção” e “reprovado” já esperados para dados de *RNA-Seq*, como por exemplo, o conteúdo por sequência-base, onde é esperado uma distribuição não uniforme de bases para os 10 a 15 primeiros nucleotídeos (Hansen et al., 2010). Por isso, ao fazermos uma trimagem no TRIMMOMATIC v.0.36, optou-se por não usar a opção ‘*HEADCROP*’, responsável por uniformizar as sequências-base. Caso isso fosse feito, correria o risco de perder quantidades significativas de dados para a pesquisa.

Foi usada a ferramenta BBTools para remover duplicatas e fazer com que os dois sequenciamentos pair-end (*forward* e *reverse*) estivessem no mesmo tamanho. Houve uma trimagem pelo TRIMMOMATIC, onde bases no início e no fim dos *reads* (comandos LEADING e TRAILING, respectivamente) que estavam abaixo do limite de qualidade especificado, bem como *reads* com tamanho abaixo do determinado (comando MINLEN) foram eliminados. Os SRA possuem uma média de 4.312.337 *reads*. Após a trimagem, passaram a ter uma média de 3.090.508 *reads*. Utilizando a ferramenta MEGAHIT, foram formados em média, para cada amostra, 27.371 *contigs*, com tamanhos entre 200 e 11.500 pares de base (**Tabela 2**).

Tabela 2 – Dados das amostras pós-trimagem e montagem dos *contigs*

SRA	Quantidade e total de <i>reads</i>	Quantidade de <i>reads</i> pós-trimagem	Número de <i>contigs</i> (total)	Comprimento do menor <i>contig</i> (bp)	comprimento do maior <i>contig</i> (bp)
SRR7366901	5729255	5697982	24177	205	11326
SRR7366902	4865475	4832714	29861	208	7525
SRR7366903	3592036	1890774	27317	201	7408
SRR7366904	6428111	2903775	29163	208	8064
SRR7366905	2079318	1262669	22378	207	7510
SRR7366906	1685154	973413	18867	207	6114
SRR7366911	4101993	26144	29685	207	9886
SRR7366912	4135830	1922124	31149	207	7388
SRR7366891	5081516	5040537	26144	209	11349
SRR7366892	4648284	4611505	27425	207	9854
SRR7366893	6503242	6447739	6447739	206	6133
SRR7366895	9485921	9433864	23644	208	7544
SRR7366897	7165275	7095671	31033	208	11203
SRR7366898	5950731	2777975	29484	203	8717
SRR7366899	5918290	5878581	27499	203	7479
SRR32130051	348909	341379	4387	250	1650
SRR32130052	640178	606978	3365	226	1776
SRR32130053	478677	467580	4110	227	5024
SRR32130054	937828	337834	4172	227	3578

4.3 Análise do BLASTx

Após o BLASTx ser realizado no DIAMOND, foram analisados os *contigs* no programa MEGAN com os parâmetros de 70% de similaridade com *contigs* de no mínimo 300 aa. Em cada amostra obtivemos no DIAMOND uma média de 2019 sequências alinhadas ao banco de dados do NCBI Vírus. Após as amostras serem inseridas no MEGAN, dessa média de 2019 sequências, apenas uma média de 41 *hits* foram considerados virais. Durante o processo de alinhamento com o banco de dados viral, observou-se que o banco promoveu alinhamentos incorretos, associando sequências que, na realidade, correspondiam a material genético de origem vegetal. Essa constatação foi possível a partir da extração das sequências inicialmente classificadas como virais e da realização de análises complementares utilizando as ferramentas BLASTx do NCBI, as quais confirmaram a verdadeira origem vegetal desses fragmentos.

Ao ajustar os parâmetros no MEGAN (30% de identidade e 100 aa), a espécie *Coccolithovirus huxleyi* e o vírus *Micromonas pusilla* vírus 12t aparecem com poucos hits, de forma esporádica com 35% de identidade, apenas se restringido o alinhamento para busca de megaviricetes. O MEGAN utiliza um banco de dados próprio (fev/2022), incompatível com o Geneious, o que impede verificar a presença de *Biavirus raunefjordenense* e *Tethysvirus ontarioense* no banco de dados.

4.4 Análise do BLASTn

A **Tabela 3** apresenta os vírus mais encontrados considerando o conjunto das 20 amostras analisadas. Observou-se que *Prymnesium kappa* virus (*Biavirus raunefjordenense*) foi detectado em todas as amostras analisadas (100%). O *Chrysochromulina parva* virus BQ2 (*Tethysvirus ontarioense*) foi identificado em 17 amostras, correspondendo a uma frequência de 85%. O *Micromonas pusilla* vírus 12T ocorreu em 15 amostras (75%), enquanto o *Emiliania huxleyi* vírus 86 (*Coccolithovirus huxleyi*) apresentou a menor frequência entre os vírus analisados, estando presente em 13 amostras (65%). (**Gráfico 3**). A porcentagem de identidade das 20 amostras está no **Gráfico 1**. Todas as espécies são da classe *Megaviricetes*, sendo *Biavirus raunefjordenense* e *Tethysvirus ontarioense* pertencentes a ordem

Imitervirales, e *Coccolithovirus huxleyi* e *Micromonas pusilla* virus 12T pertencentes a ordem *Algavirales*. Não foram achados espécies virais de plantas. *Biavirus raunefjordense* apresentou as maiores médias de hits em todos os tecidos, especialmente na raiz, *Micromonas pusilla* virus 12T apresentou uma distribuição homogênea, enquanto *Tethysvirus ontarioense* e *Coccolithovirus huxleyi* mostraram redução consistente no meristema (**Gráfico 2**).

Tabela 3 – Resultados de BLASTn (identidade acima de 70% e tamanho acima de 500 nt)

SRA	Blastn/espécie viral	contigs (hits)	tamanho médio dos contigs (nt)	Identidade média (%)	Bit-score	E-value
SRR7366900	<i>Biavirus raunefjordense</i>	12	833	70%	216 789	2.80e-61
	<i>Micromonas pusilla</i> virus 12T	5	1525	69.1%	655 007	0
	<i>Tethysvirus ontarioense</i>	4	1748	63.9%	275 398	2.91e-65
SRR7366901	<i>Biavirus raunefjordense</i>	11	860	67.5%	209 575	1.74e-62
	<i>Coccolithovirus huxleyi</i>	5	611	67.1%	176 213	1.45e-41
	<i>Micromonas pusilla</i> virus 12T	5	1361	69.2%	472 867	4.95e-131
	<i>Tethysvirus ontarioense</i>	3	1781	64.1%	297 039	5.32e-78

SRR7366 902	<i>Biavirus raunefjordense</i>	10	822	67.3%	205 969	6.2e-52
	Micromonas pusilla virus 12T	4	1399	69.1%	552 666	1.32e-81
	<i>Tethysvirus ontarioense</i>	4	1895	65.8%	301 547	2.53e-78
	<i>Coccolithovirus huxleyi</i>	3	594	66%	101 373	6.00e-19
SRR7366 903	<i>Biavirus raunefjordense</i>	6	654	68,12%	240 232	2,12e-60
	Micromonas pusilla virus 12T	6	1259	70.0%	424 176	2.53e-178
	<i>Tethysvirus ontarioense</i>	3	1526	64%	258 266	3.01e-66
	<i>Coccolithovirus huxleyi</i>	3	588	66.8%	176 213	1.39e-41
SRR7366 904	<i>Biavirus raunefjordense</i>	11	833	66.6%	209 575	9.54e-52
	Micromonas pusilla virus 12T	6	1094	69.4%	416 061	1.16e-133
	<i>Tethysvirus ontarioense</i>	2	1311	68%	259 618	1.59e-63
SRR7366 905	<i>Biavirus raunefjordense</i>	5	846	67%	205 067	1.41e-49
	Micromonas pusilla virus 12T	4	1085	70.3%	451 677	2.09e-78

	<i>Tethysvirus ontarioense</i>	3	1580	63.9%	258 266	3.05e-66
	Prochlorococcus phage P-TIM68	3	549	69.4%	229 412	6.54e-58
SRR7366 906	Micromonas pusilla virus 12T	4	1351	70.5%	669 434	0
	<i>Tethysvirus ontarioense</i>	3	1616	65.7%	688 37	2.94e-66
	<i>Coccolithovirus huxleyi</i>	2	625	65.9%	138 793	3.30e-19
	<i>Biavirus raunefjordense</i>	2	960	68.3%	352 943	3.72e-52
SRR7366 911	<i>Biavirus raunefjordense</i>	9	737	66.9%	203 264	2.54e-48
	Micromonas pusilla virus 12T	4	1330	70.5%	669 434	0
	<i>Tethysvirus ontarioense</i>	3	1553	64%	229 412	1.18e-57
SRR7366 912	<i>Biavirus raunefjordense</i>	10	791	67.1%	218 141	7.15e-52
	Micromonas pusilla virus 12T	7	866	69.9%	263 676	1.20e-78
	<i>Coccolithovirus huxleyi</i>	4	604	67.2%	138 793	7.20e-42
	<i>Tethysvirus ontarioense</i>	3	1616	64%	258 266	2.77e-66
SRR7366 891	<i>Biavirus raunefjordense</i>	13	826	66.9%	208 674	6.80e-54

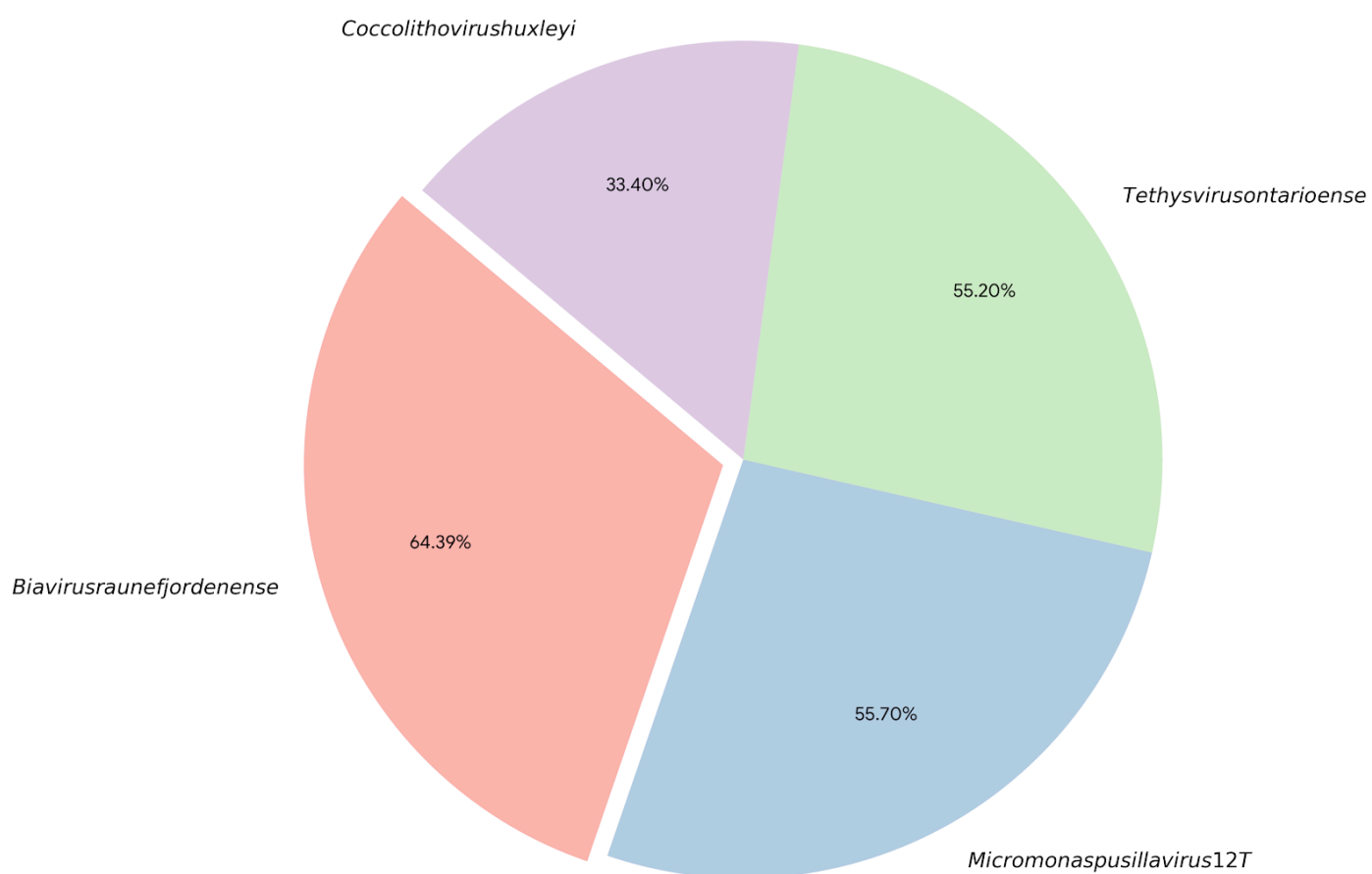
	<i>nse</i>					
	<i>Tethysvirus ontarioense</i>	5	1673	63.9%	258 717	2.88e-65
	<i>Coccolithovirus huxleyi</i>	5	656	67.3%	176 213	1.41e-41
	Micromonas pusilla virus 12T	4	1615	70%	676 647	0
SRR7366 893	<i>Biavirus raunefjordense</i>	11	852	67.5%	216 789	7.14e-54
	Micromonas pusilla virus 12T	5	1385	69%	655 007	0
	<i>Coccolithovirus huxleyi</i>	4	702	67%	156 827	2.66e-30
	<i>Tethysvirus ontarioense</i>	3	1781	67%	297 039	5.13e-78
SRR7366 892	<i>Biavirus raunefjordense</i>	11	831	67.3%	208 674	1.22e-51
	Micromonas pusilla virus 12T	5	1605	69%	669 434	0
	<i>Coccolithovirus huxleyi</i>	5	656	67%	176 213	1.52e-41
	<i>Tethysvirus ontarioense</i>	4	1748	64%	275 398	2.89e-65

SRR7366 895	<i>Biavirus raunefjordense</i>	10	867	67.2%	227 158	1.14e-62
	Micromonas pusilla virus 12T	5	1557	69%	655 007	0
	<i>Tethysvirus ontarioense</i>	4	1748	64%	275 398	2.83e-65
	<i>Coccolithovirus huxleyi</i>	4	617	67%	138 793	8.77e-40
SRR7366 897	<i>Biavirus raunefjordense</i>	11	859	67.4%	221 297	7.1e-56
	Micromonas pusilla virus 12T	5	1605	69%	669 434	0
	<i>Tethysvirus ontarioense</i>	4	1748	64%	275 398	2.81e-65
	<i>Coccolithovirus huxleyi</i>	3	723	67%	176 213	1.38e-41
SRR7366 898	<i>Biavirus raunefjordense</i>	12	850	67%	213 182	4.66e-54
	-	6	635	67%	156 827	2.85e-30
	Micromonas pusilla virus 12T	5	1605	69%	669 434	0
	<i>Tethysvirus ontarioense</i>	4	1748	64%	275 398	2.87e-65
SRR7366 899	<i>Biavirus raunefjordense</i>	10	806	67%	243.380	3.4e-61

	Micromonas pusilla virus 12T	5	1548	69%	616.685	0
	<i>Tethysvirus ontarioense</i>	4	1748	64%	275.398	2.8e-65
SRR3213 0051	<i>Biavirus raunefjordense</i>	3	364	69.9%	206 870	4.30e-42
	Dishui lake phycodnavirus 1	2	340	70%	147 359	1.55e-20
	Synechococcus phage ACG-2014f	2	400	77.6%	309 662	1.44e-26
SRR3213 0052	Bacillus phage G	3	390	66.7%	977 665	4.05e-19
	<i>Biavirus raunefjordense</i>	2	359	68.8%	581 447	3.91e-19
SRR3213 0053	Synechococcus phage S-SCSM1	4	669	77%	505 327	4.59e-139
	<i>Biavirus raunefjordense</i>	4	415	70%	191 992	1.6e-42
	Prochlorococcus phage P-TIM68	3	390	68%	147 359	1.25e-33
	<i>Tethysvirus ontarioense</i>	2	401	69%	168 098	2.36e-24
SRR3213 0054	Prochlorococcus phage P-TIM68	11	761	68%	221 297	7.0e-51

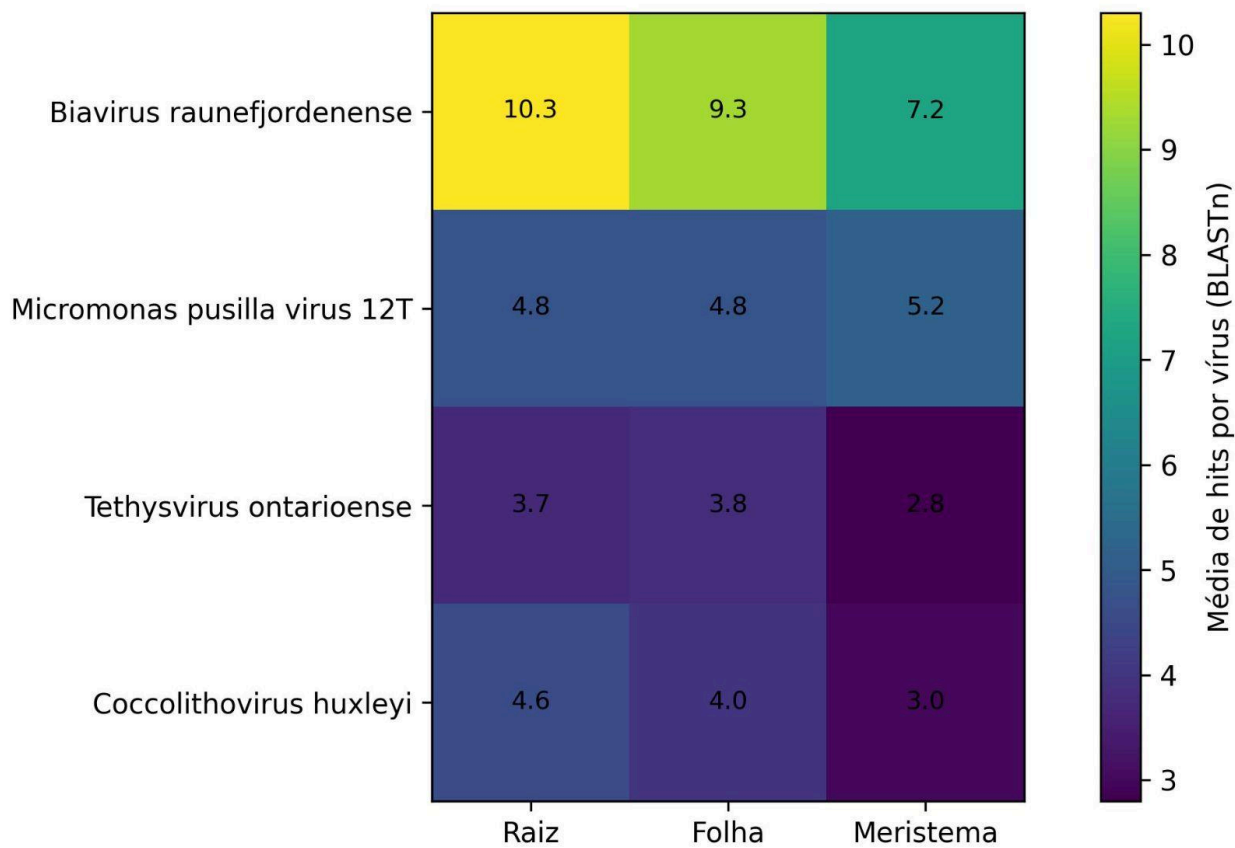
Synechococcus phage S-SCSM1	7	837	77%	639 679	0
Synechococcus phage S-CAM9 isolate 1109NB16	3	720	77%	552 215	2.47e-93

Gráfico 1 – Porcentagem de Identidade nucleotídica média das 20 amostras analisadas por BLASTn



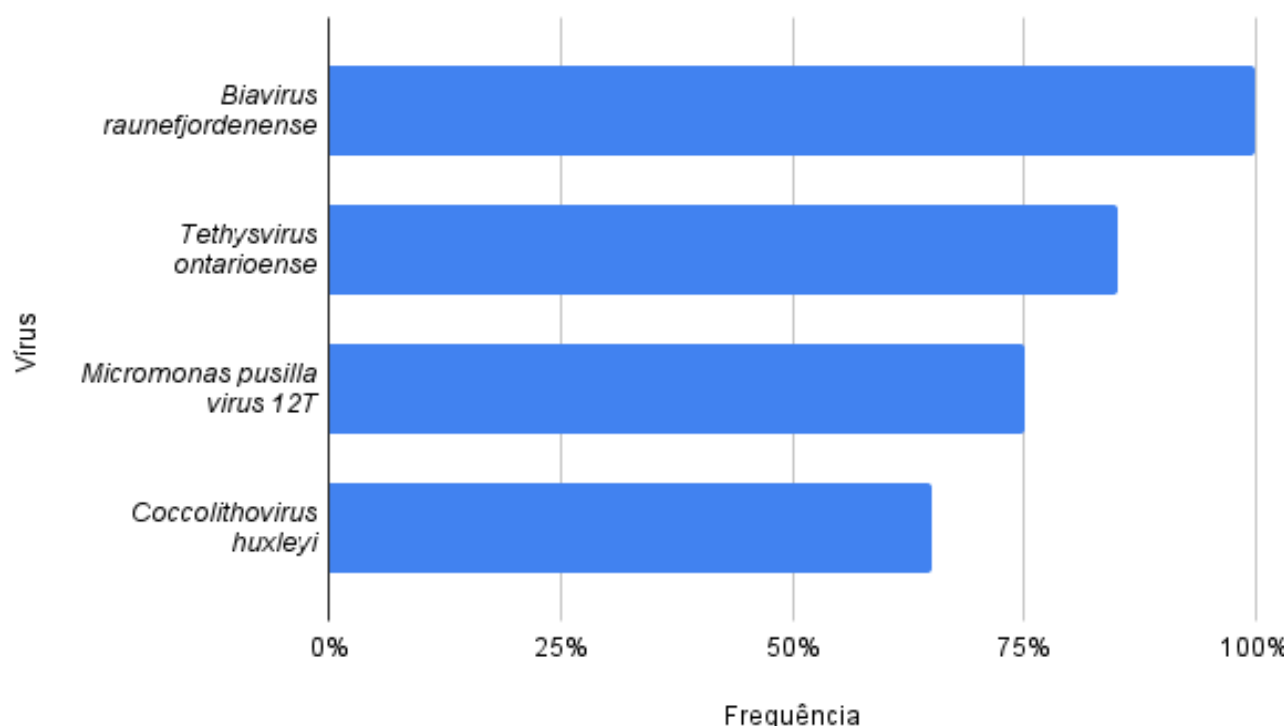
Fonte: Serpa, 2026

Gráfico 2 – Heatmap da média do número de hits de BLASTn para os principais vírus detectados em tecidos vegetais (raiz, folha e meristema)



Fonte: Serpa, 2026

Gráfico 3 –Frequência de ocorrência dos vírus mais abundantes identificados no conjunto das 20 amostras analisadas



Fonte: Serpa, 2026

Os vírus da classe *Megaviricetes* são chamados de ‘vírus gigantes’ e fazem parte do filo *Nucleocytoviricota*, onde pertencem os vírus nucleocitoplasmáticos de DNA gigantes (NCLDVs). Representam os maiores vírus em termos de tamanho de vírions e comprimento de genoma, chegando a ultrapassar os de algumas bactérias (Alyward, 2023; Monttinen, 2021). Tal classe possui três ordens: *Algavirales*, *Imitervirales* e *Primascovirales*, e diversas famílias (Alyward, 2023). São vírus de dsDNA, com capacidade de replicação tanto no núcleo como no citoplasma do hospedeiro, e tem como genes em comum a DNA polymerase B, genes que codificam a proteína do capsídeo, entre outros (Nash, 2025 apud Monttinen, 2021). Tem como hospedeiros uma gama de algas e eucariotos, e preferência por ecossistemas aquáticos (Nash, 2025 apud Moniruzzaman, 2020).

A elevada frequência de vírus associados a algas e eucariotos observadas nas amostras está em consonância com a literatura referente ao viroma de ambientes costeiros e estuarinos. Por exemplo, ambientes com solo alagado ou intensamente

encharcados tendem a apresentar um maior acúmulo de matéria orgânica e elevada produtividade microbiana, favorecendo a ocorrência de microalgas e protistas, sendo eles os principais hospedeiros dos vírus da classe *Megaviricetes* (Thatoi et al., 2013; Ghizelini et al., 2012). Ambientes de manguezal caracterizam-se por elevada produtividade, intensa ciclagem de carbono e nutrientes e forte acoplamento entre os compartimentos terrestre e aquático, condições que favorecem o desenvolvimento de comunidades microbianas diversas, bem como ativos ciclos biogeoquímicos (Alongi, 2014). Estudos de diversidade microbiana em manguezais apontam que esse ecossistema abriga uma ampla gama de microrganismos, como bactérias, fungos e microalgas, participantes na decomposição, fixação de Nitrogênio e outros processos (Thatoi et al., 2013; Ghizelini et al., 2012). No contexto do presente estudo, a presença consistente de vírus/espécies virais como *Biavirus raunefjordenense*, *Tethysvirus ontarioense*, *Micromonas pusilla virus 12T* e *Coccolithovirus huxleyi* sugere que interações vírus–fitoplâncton desempenham papel relevante na estruturação das comunidades microbianas locais. Vírus da classe *Megaviricetes* possuem um papel ecológico importante na alteração da dinâmica dos seus hospedeiros, como por exemplo, no controle da proliferação de algas (Nash, 2025). Além disso, contribuem para a exportação de carbono e em ambientes marinhos através da morte dos hospedeiros (Kim et al, 2023), assim como afetam o metabolismo de nitrogênio em tais (Nash, 2025 apud Liang, 2024). Trabalhos anteriores indicam que a lise viral de microalgas realizadas por vírus da classe *Megaviricetes* contribui para a liberação de matéria orgânica dissolvida, intensificando o chamado *viral shunt* e influenciando diretamente os fluxos de carbono e nutrientes em ecossistemas costeiros (Wilhelm & Suttle, 1999; Danovaro et al., 2011; Nash, 2025). Assim, os resultados obtidos reforçam a importância dos vírus associados a microalgas como componentes-chave da dinâmica ecológica de manguezais, atuando de forma indireta nos processos biogeoquímicos relacionados à produtividade e à ciclagem de matéria orgânica em sistemas dominados por *R. mangle*. Uma pesquisa realizada por Rignonato et al. em 2012 sobre diversidade de cianobactérias na filosfera de manguezal mostrou que diferentes tecidos podem hospedar distintas assembleias microbianas, indicando que a estrutura da planta pode influenciar na distribuição dos microrganismos.

Esses padrões ecológicos apoiam a interpretação de que a presença de megaviricetes nas amostras de raiz, caule, folha e meristema de *Rhizophora mangle* provavelmente está relacionada à heterogeneidade microambiental e à dinâmica de algas e protistas que coexistem com a planta, e não a uma infecção direta dos tecidos vegetais.

5. CONCLUSÃO

O presente estudo teve como objetivo caracterizar a diversidade e viral em dados (meta)transcriptômicos de *Rhizophora mangle*. A partir da análise de 20 amostras selecionadas, foi identificado um conjunto consistente de vírus eucarióticos pertencentes à classe *Megaviricetes*.

A predominância de vírus da classe *Megaviricetes*, todos associados a microalgas e outros eucariotos aquáticos, aliada à ausência de vírus tipicamente fitopatogênicos, sugere que a comunidade viral detectada reflete majoritariamente a dinâmica ecológica do ambiente de manguezal, e não processos de infecção direta dos tecidos vegetais de *R. mangle*. Esses resultados reforçam a hipótese de que os vírus identificados estejam associados de forma indireta à planta, possivelmente por meio da presença de microalgas, protistas e biofilmes aderidos às superfícies vegetais ou presentes nos microambientes adjacentes às raízes, folhas, caules e meristemas.

Nesse contexto, a elevada frequência de vírus associados a microalgas observada neste estudo concorda com a literatura sobre viromas costeiros e estuarinos, que aponta os *Megaviricetes* como componentes da dinâmica ecológica desses ambientes, atuando no controle populacional de hospedeiros, na liberação de matéria orgânica dissolvida e na modulação dos fluxos de carbono e nutrientes.

Do ponto de vista metodológico, os resultados obtidos demonstram o potencial da abordagem (meta)transcriptômica aliada ao *de novo assembly* como ferramenta eficaz para a caracterização de comunidades virais em ecossistemas complexos, como os manguezais.

O presente trabalho não incluiu anotação genômica e funcional das sequências virais detectadas. Assim, estudos futuros que abordem a montagem dos genomas virais, bem como anotação de genes e vias funcionais, poderão fornecer informações mais precisas sobre mecanismos de infecção, estratégias de replicação e possíveis interações entre vírus e hospedeiro, bem como de ambiente de manguezal.

De modo geral, o presente trabalho apresenta um conhecimento inicial da diversidade viral em *Rhizophora mangle*, destacando a relevância ecológica dos vírus da classe *Megaviricetes* nesses ambientes, abrindo o caminho para pesquisas

futuras com abordagens genômicas, funcionais e ecológicas. A ampliação dessas análises poderá oferecer uma visão aprofundada no papel dos vírus em ecossistemas de manguezal.

REFERÊNCIAS

- ABERGEL, C.; CLAVERIE, J.-M. **Mimivirus and its virophage**. *Annual Review of Genetics*, v. 52, p. 89–111, 2018.
- ALONGI, D. M. **Carbon cycling and storage in mangrove forests**. *Annual Review of Marine Science*, v. 6, p. 195–219, 2014.
- 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: 10 set. 2024.
- AYLWARD, F. O. et al. **Taxonomic update for giant viruses in the order Imitervirales (phylum Nucleocytoviricota)**. *Archives of Virology*, v. 168, n. 11, p. 283, 2023.
- BABRAHAM BIOINFORMATICS. **FastQ Screen**. Disponível em: https://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/. Acesso em: 10 set. 2024.
- BANKEVICH, A. et al. **SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing**. *Journal of Computational Biology*, v. 19, n. 5, p. 455–477, 2012.
- BBMAP. **BBMap short read aligner**. Disponível em: <https://sourceforge.net/projects/bbmap/>. Acesso em: 10 set. 2024.
- BECKAGE, N. E.; GELMAN, D. B. **Wasp parasitoid disruption of host development: implications for new biologically based strategies for insect control**. *Annual Review of Entomology*, Palo Alto, v. 49, p. 299–330, 2004. DOI: 10.1146/annurev.ento.49.061802.123324.
- BRÉCHIGNAC, F.; DESMET, G. **Equidosimetry: ecological standardization and equidosimetry for radioecology and environmental ecology**. Dordrecht: Springer, [s.d.].
- CAMPILLO-BALDERAS, J. A.; LAZCANO, A.; BECERRA, A. **Viral genome size distribution does not correlate with the antiquity of the host lineages**. *Frontiers in Ecology and Evolution*, v. 3, p. 1–8, 2015.
- CHAPMAN, B. S.; CHRISTMANN, J. L.; THATCHER, E. F. **Bioinformatics for undergraduates: steps toward a quantitative bioscience curriculum**. *Biochemistry and Molecular Biology Education*, v. 34, n. 3, p. 180–186, 2006.
- CHEN, Hao; ZHANG, Weijia; LI, Xiefei; PAN, Yingjie; YAN, Shuling; WANG, Yongjie. **The genome of a prasinoviruses-related freshwater virus reveals unusual diversity of phycodnaviruses**. *BMC Genomics*, Londres, v. 19, art. 49, 2018. DOI: 10.1186/s12864-018-4432-4. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/29334892/>. Acesso em: 21 jan. 2026.

CHINA NATIONAL GENBANK. **Pahexavirus genomic records**. CNGB Database. Disponível em: https://db.cngb.org/data_resources/organism/1982282. Acesso em: 17 jan. 2026.

CLAVERIE, J.-M.; ABERGEL, C. **Mimivirus and its virophage**. *Annual Review of Genetics*, v. 52, p. 89–111, 2018.

COMPEAU, P. E. C.; PEVZNER, P. A.; TESLER, G. **How to apply de Bruijn graphs to genome assembly**. *Nature Biotechnology*, v. 29, n. 11, p. 987–991, 2011.

DE CASTRO, Edouard et al. **ViralZone 2024 provides higher-resolution images and advanced virus-specific resources**. *Nucleic Acids Research*, v. 52, n. D1, p. D817-D821, 5 jan. 2024.

DELANEY NASH, D. et al. **Hybrid sequencing reveals the genome of a *Chrysochromulina parva* virus and highlights its distinct replication strategy**. *bioRxiv*, 2025. DOI: 10.1101/2025.01.22.633263.

FINGU-MABOLA, J. C.; FRANCIS, F. **Aphid–plant–phytovirus pathosystems: influencing factors from vector behaviour to virus spread**. *Agriculture*, v. 11, n. 6, p. 502, 2021.

GENEIOUS. **Geneious bioinformatics software**. Disponível em: <https://www.geneious.com/>. Acesso em: 10 set. 2024.

GHIZELINI, A. M.; MENDONÇA-HAGLER, L. C. S.; MACRAE, A. **Microbial diversity in Brazilian mangrove sediments: a mini review**. *Brazilian Journal of Microbiology*, v. 43, n. 4, p. 1242–1254, 2012. DOI: 10.1590/S1517-83822012000400002.

HANSEN, K. D.; BRENNER, S. E.; DUDOIT, S. **Biases in Illumina transcriptome sequencing caused by random hexamer priming**. *Nucleic Acids Research*, v. 38, n. 12, e131, 2010. DOI: 10.1093/nar/gkq224.

HUSON, D. H. et al. **MEGAN analysis of metagenomic data**. *Genome Research*, v. 17, n. 3, p. 377–386, 2007. DOI: 10.1101/gr.5969107.

KIM, J.; LEE, H.; PARK, S. **vB_CacS-HV1 as a Novel Pahexavirus Bacteriophage with Lytic and Anti-Biofilm Potential against *Cutibacterium acnes***. *Microorganisms*, v. 12, n. 8, 2024. DOI: <https://doi.org/10.3390/microorganisms12081566>. Acesso em: 17 jan. 2026.

KIM, K. E. et al. **Covariance of marine nucleocytoplasmic large DNA viruses with eukaryotic plankton communities in the sub-Arctic Kongsfjorden ecosystem**. *Microorganisms*, v. 11, n. 1, 2023.

KRUG, L.; LEÃO, C.; AMARAL, S. **Dinâmica espaço-temporal de manguezais no Complexo Estuarino de Paranaguá e relação com dados socioeconômicos**. 2006. Disponível em: <http://marte.dpi.inpe.br/col/dpi.inpe.br/sbsr@80/2006/11.14.20.50/doc/2753-2760.pdf>; .. Acesso em: 10 set. 2024.

- LIAO, X. et al. **Current challenges and solutions of de novo assembly.** *Quantitative Biology*, v. 7, n. 2, p. 90–109, 2019.
- LIANG, J. L. et al. **Unraveling the habitat preferences, ecological drivers, potential hosts, and auxiliary metabolism of soil giant viruses across China.** *Microbiome*, v. 12, n. 1, 2024.
- LOUTEN, J. **Virus structure and classification.** In: *Essential Human Virology*. London: Academic Press, 2016. p. 19–29.
- LOMONOSSOFF, George P.; WEGE, Christina. Chapter six – TMV particles: the journey from fundamental studies to bionanotechnology applications. In: **ADVANCES IN VIRUS RESEARCH**. San Diego: Academic Press, 2018. v. 102, p. 149–176.
- METZKER, M. L. **Sequencing technologies — the next generation.** *Nature Reviews Genetics*, v. 11, n. 1, p. 31–46, 2009.
- MILLER, J. R.; KOREN, S.; SUTTON, G. **Assembly algorithms for next-generation sequencing data.** *Genomics*, v. 95, n. 6, p. 315–327, 2010.
- MONIRUZZAMAN, M. et al. **Dynamic genome evolution and complex virocell metabolism of globally distributed giant viruses.** *Nature Communications*, v. 11, n. 1, 2020.
- MÖNTTINEN, H. A. M. et al. **The genomes of nucleocytoplasmic large DNA viruses: viral evolution writ large.** *Microbial Genomics*, v. 7, n. 9, 2021.
- NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION. **Taxonomy: Pahexavirus.** NCBI Taxonomy Browser. Disponível em: <https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=1982251>. Acesso em: 17 jan. 2026.
- NAGELKERKEN, I.; BLABER, S. J. M.; BOOTH, D.; GREEN, P.; HAYWOOD, M.; KIRKMAN, H.; MEYNECKE, J. O.; PENDER, P.; SHEAVES, M.; BRODIE, J.; LONERAGAN, N. R. **The habitat function of mangroves for terrestrial and marine fauna: a review.** *Aquatic Botany*, Amsterdam, v. 89, n. 2, p. 155–185, 2008. DOI: 10.1016/j.aquabot.2007.12.007.
- ODUM, E. P. **Ecologia.** Rio de Janeiro: Guanabara Koogan, 1988.
- PEVSNER, J. **Bioinformatics and functional genomics.** Chichester: Wiley-Blackwell, 2015.
- RIGONATO, J. et al. **Cyanobacterial diversity in the phyllosphere of a mangrove forest.** *FEMS Microbiology Ecology*, v. 80, n. 2, p. 312–323, 2012.
- SCHAEFFER-NOVELLI, Y. **Manguezal: ecossistema entre a terra e o mar.** São Paulo: Caribbean Ecological Research, 1995.
- SILVA, A. M. et al. **Tailed bacteriophages (Caudoviricetes) dominate the microbiome of a diseased stingless bee.** *Genetics and Molecular Biology*, v. 47,

2023. DOI: <https://doi.org/10.1590/1678-4685-GMB-2023-0120>. Acesso em: 17 jan. 2026

STRAND, M. R.; BURKE, G. R. **Polydnaviruses: nature's genetic engineers.** *Annual Review of Virology*, Palo Alto, v. 1, p. 333–354, 2014. DOI: 10.1146/annurev-virology-031413-085451.

SUTTLE, C. A. **Marine viruses: major players in the global ecosystem.** *Nature Reviews Microbiology*, v. 5, p. 801–812, 2007.

THATOI, H. et al. **Biodiversity and biotechnological potential of microorganisms from mangrove ecosystems: a review.** *Annals of Microbiology*, v. 63, p. 1–19, 2013. DOI: 10.1007/s13213-012-0442-7.

UNIPROT CONSORTIUM. **UniProt: a worldwide hub of protein knowledge.** Disponível em: <https://www.uniprot.org/>. Acesso em: 10 set. 2024.

XU, Shengzhong; ZHOU, Liang; LIANG, Xiaosha; ZHOU, Yifan; CHEN, Hao; YAN, Shuling; WANG, Yongjie. **Novel cell-virus-virophage tripartite infection systems discovered in the freshwater Lake Dishui Lake in Shanghai, China.** *Journal of Virology*, Washington, v. 94, n. 11, e00149-20, 2020. DOI: 10.1128/JVI.00149-20. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/32188734/>. Acesso em: 21 jan. 2026.

ZHU, Yilin; SHANG, Jiayu; PENG, Cheng; SUN, Yanni. **Phage family classification under Caudoviricetes: A review of current tools using the latest ICTV classification framework.** *arXiv preprint*, 2022. Disponível em: <https://arxiv.org/abs/2209.01942>. Acesso em: 17 jan. 2026.

