

**UNIVERSIDADE FEDERAL DE MINAS GERAIS**  
**Instituto de Ciências Biológicas**  
**Programa Interunidades de Pós-graduação em Bioinformática**

Carlos Willian Dias Dantas

**BIOPROSPECÇÃO DE BACTERIOFAGOS PARA O COMBATE A BACTÉRIAS  
RESISTENTES A ANTIBIÓTICOS A PARTIR DE DADOS METAGENÔMICOS**

Belo Horizonte

2025

Carlos Willian Dias Dantas

**BIOPROSPECÇÃO DE BACTERIOFAGOS PARA O COMBATE A BACTÉRIAS  
RESISTENTES A ANTIBIÓTICOS A PARTIR DE DADOS METAGENÔMICOS**

Tese apresentada ao Programa Interunidades de Pós-Graduação em Bioinformática da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Doutor em Bioinformática.

Discente: Carlos Willian Dias Dantas

Orientador: Prof. Rommel Thiago Jucá Ramos

Belo Horizonte

2025

043

Dantas, Carlos Willian Dias.

Bioprospecção de bacteriófagos para o combate a bactérias resistentes a antibióticos a partir de dados metagenômicos [manuscrito] / Carlos Willian Dias Dantas. – 2025.

91 f. : il. ; 29,5 cm.

Orientador: Prof. Rommel Thiago Jucá Ramos.

Tese (doutorado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa Interunidades de Pós-Graduação em Bioinformática.

1. Bioinformática. 2. Metagenômica. 3. Bacteriófagos. 4. Resistência Microbiana a Medicamentos. I. Ramos, Rommel Thiago Jucá. II. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. III. Título.

CDU: 573:004



UNIVERSIDADE FEDERAL DE MINAS GERAIS

**ATA**

INSTITUTO DE CIÊNCIAS BIOLÓGICAS  
PROGRAMA INTERUNIDADES DE PÓS-GRADUAÇÃO EM BIOINFORMÁTICA

**DEFESA DE TESE**

*Carlos Willian Dias Dantas*

Às quatorze horas do dia **23 de outubro de 2025**, reuniu-se, por videoconferência, a Comissão Examinadora de Tese, indicada pelo Colegiado do Programa, para julgar, em exame final, o trabalho intitulado: **""Bioprospecção de Bacteriófagos para o Combate a Bactérias Resistentes a Antibióticos a Partir de Dados Metagenômicos""**, requisito para obtenção do grau de Doutor em **Bioinformática**. Abrindo a sessão, o Presidente da Comissão, **Dr. Rommel Thiago Juca Ramos**, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato, para apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa do candidato. Logo após, a Comissão se reuniu, sem a presença do candidato e do público, para julgamento e expedição de resultado final. Foram atribuídas as seguintes indicações:

<b>Professor(a)/Pesquisador(a)</b>	<b>Instituição</b>	<b>Indicação</b>
Dr. Rommel Thiago Juca Ramos - Orientador	Universidade Federal do Pará	Aprovado
Dra. Ana Cecília Ribeiro Cruz	Instituto Evandro Chagas	Aprovado
Dr. Luis Gustavo Carvalho Pacheco	Universidade Federal da Bahia	Aprovado
Dr. Siomar de Castro Soares	Universidade Federal do Triângulo Mineiro	Aprovado
Dr. Vasco Ariston de Carvalho Azevedo	Universidade Federal de Minas Gerais	Aprovado

Pelas indicações, o candidato foi considerado: **Aprovado**

O resultado final foi comunicado publicamente ao candidato pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora.

**Belo Horizonte, 23 de outubro de 2025.**



Documento assinado eletronicamente por **Siomar de Castro Soares, Usuário Externo**, em 24/10/2025, às 14:21, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Luis Gustavo Carvalho Pacheco, Usuário Externo**, em 24/10/2025, às 16:54, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

## AGRADECIMENTOS

Gostaria de agradecer primeiramente meus pais (Leonice Dias Dantas e Antonio Carlos de Souza Dantas) e minha irmã (Lenna Karlla Dias Dantas) pelo apoio integral, amor incondicional e acompanhamento durante toda esta jornada conhecida como vida. Sem eles jamais teria chegado até aqui e com certeza esta trajetória seria muito mais difícil.

Agradeço a minha avó (Luiza Pereira da Silva) *in memoriam* por sempre acreditar em mim e por ter oferecido todo o apoio do mundo para que um dia eu pudesse terminar o doutorado.

Agradeço a minha noiva Andressa Pereira Ferreira por todos estes anos de amor, paciência e preocupação. Muito obrigado por tudo, meu amor, e que venham muito mais anos de amor e companheirismo. Especialmente por termos o mais belo presente que é o nosso filho “Benjamin Ferreira Dantas”.

A todos os integrantes Laboratório de Bioinformática e Genômica de Microrganismos (LBGM) meu muito obrigado por todo o carinho e risadas que puderam ocorrer nos prelúdios de escritas de artigos e trabalhos. Em especial a meu amigo Oscar Cardenas Alegria.

Agradeço em especial a meu orientador Rommel Thiago jucá Ramos por todo ensinamento, parceria, conselhos e orientação oferecida desde o meu mestrado. O seu apoio e ensinamentos foram essenciais para o meu desenvolvimento acadêmico. Muito obrigado por tudo!

Agradeço a todos os amigos que tive o prazer de conhecer ao longo desta caminhada. Aos passageiros e aos que continuam comigo nesta jornada saibam que guardarei para sempre os momentos que passamos juntos.

Agradeço ao Programa Interunidades de Pós-Graduação em Bioinformática da Universidade Federal de Minas Gerais, especialmente ao corpo docente e aos secretários, por sempre fazerem o possível para ensinar e inspirar os seus alunos a se tornarem cientistas neste país. Muito obrigado a todos!!

**"Você deve aproveitar os pequenos desvios.  
Ao máximo. Porque é onde você encontrará  
coisas mais importantes do que aquilo que  
você realmente deseja"**

**Ging Freecs**

## RESUMO

O crescente problema da resistência bacteriana a antibióticos, agravado pelo uso indiscriminado destes fármacos na medicina humana e veterinária, tem demandado o desenvolvimento de novas estratégias terapêuticas. Dentre essas estratégias, os bacteriófagos (fagos) destacam-se como uma alternativa promissora devido à sua especificidade e eficácia na infecção e eliminação de bactérias patogênicas, incluindo cepas multirresistentes. O uso do sequenciamento metagenômico para estudar comunidades microbianas também pode auxiliar na detecção de sequências de bacteriófagos. Nesse sentido, aproveitando a alta biodiversidade presente no bioma Amazônia, um estudo metagenômico *shotgun* foi conduzido na cidade de Bragança, Pará, com coletas de água em cinco pontos diferentes ao longo do Rio Cereja (BP1 – BP5), em triplicata, desde o trecho da nascente até uma área mais contaminada. Após a extração de DNA metagenômico, as amostras foram sequenciadas pela plataforma Illumina NovaSeq 6000 e os dados foram pré-processados com fastp seguido de montagem com metaSpades. Kraken2, KrakenTools e Microeco foram usados para caracterização taxonômica e análise de diversidade de vírus e bactérias. As correlações com os perfis de resistência e os dados taxonômicos foram normalizadas e correlacionadas por meio de análise de redundância (RDA) para identificar o ponto de maior impacto ambiental, o que mostrou BP4 como o mais contaminado. Os contigs dos 5 pontos foram analisados por VIBRANT, VirSorter2 e DeepVirFinder, e as sequências virais identificadas foram selecionadas com base na completude (> 90%), contaminação (< 10%) e critérios de qualidade viral, com remoção de provírus com base nos resultados de CheckV e ViralComplete, resultando em 19, 14, 11, 11 e 6 sequências virais selecionadas para os pontos BP1 a BP5 respectivamente. Paralelamente, "metagenomic assembled genomes" MAGs metagenômicos foram gerados e filtrados, e a identificação do táxon foi realizada usando o programa kMetaShot. Uma comparação por inferência foi feita para detectar MAGs que seriam prováveis hospedeiros das sequências virais selecionadas, o ponto BP4 por ser o mais contaminado e ter obedecido a estes critérios foi selecionado e analisado a sua relação fago – hospedeiro e eles foram avaliados quanto à presença de genes de resistência antimicrobiana (RAM) e sistemas de defesa. Foram encontrados fagos com prováveis hospedeiros de MAGs identificadas de *Citrobacter*, *Acinetobacter* e *Enterobacter*. Nenhum gene de RAM foi identificado nas sequências virais, mas RAMs foram identificados em 9 de 10 MAGs selecionados como prováveis hospedeiros bacterianos, indicando fagos potencialmente eficazes contra bactérias patogênicas resistentes a antibióticos. Análises funcionais também mostraram que MAGs selecionadas do gênero *Acinetobacter* obtiveram uma maior diversidade de sistemas de defesa contra fagos e identificação de genes Cas, sendo identificado fagos líticos para este gênero; e as MAGs que tiveram maior presença de genes de AMR tiveram identificação do sistema toxina-antitoxina MazEF, e maior identificação do sistema restrição-modificação (RM) em *Enterobacter*, sendo identificados fagos temperados para este gênero. Embora mais estudos sejam necessários para validar e robustecer as descobertas, os resultados apontam para uma abordagem promissora para orientar a bioprospecção de fagos em ambientes contaminados e prever possíveis relações fago-hospedeiro para que análises experimentais mais precisas possam ser realizadas.

Palavras-chave: Metagenômica; fagos; resistência a antibióticos; bactérias hospedeiras

## ABSTRACT

The growing problem of bacterial resistance to antibiotics, aggravated by the indiscriminate use of these drugs in human and veterinary medicine, has demanded the development of new therapeutic strategies. Among these strategies, bacteriophages (phages) stand out as a promising alternative due to their specificity and effectiveness in infecting and eliminating pathogenic bacteria, including multidrug-resistant strains. The use of metagenomic sequencing to study microbial communities can also aid in the detection of bacteriophage sequences. In this sense, taking advantage of the high biodiversity present in the Amazon biome, a shotgun metagenomic study was conducted in the city of Bragança, Pará, with water samples collected at five different points along the Cereja River (BP1 – BP5), in triplicate, from the source to a more contaminated area. After metagenomic DNA extraction, the samples were sequenced using the Illumina NovaSeq 6000 platform, and the data were preprocessed with fastp followed by assembly with metaSpades. Kraken2, KrakenTools, and Microeco were used for taxonomic characterization and diversity analysis of viruses and bacteria. Correlations with resistance profiles and taxonomic data were normalized and correlated using redundancy analysis (RDA) to identify the point of greatest environmental impact, which showed BP4 as the most contaminated. Contigs from the 5 points were analyzed by VIBRANT, VirSorter2, and DeepVirFinder, and the identified viral sequences were selected based on completeness (> 90%), contamination (< 10%), and viral quality criteria, with provirus removal based on CheckV and ViralComplete results, resulting in 19, 14, 11, 11, and 6 viral sequences selected for points BP1 to BP5, respectively. In parallel, metagenomic "metagenomic assembled genomes" (MAGs) were generated and filtered, and taxon identification was performed using the kMetaShot program. An inferential comparison was made to detect MAGs that would be likely hosts of the selected viral sequences. Point BP4, being the most contaminated and meeting these criteria, was selected and its phage-host relationship was analyzed. They were evaluated for the presence of antimicrobial resistance genes (AMRs) and defense systems. Phages with likely hosts of identified MAGs from *Citrobacter*, *Acinetobacter*, and *Enterobacter* were found. No AMR genes were identified in the viral sequences, but AMRs were identified in 9 of 10 MAGs selected as likely bacterial hosts, indicating potentially effective phages against antibiotic-resistant pathogenic bacteria. Functional analyses also showed that selected MAGs from the genus *Acinetobacter* exhibited a greater diversity of phage defense systems and identification of Cas genes, with lytic phages identified for this genus; and MAGs with a higher presence of AMR genes showed identification of the MazEF toxin-antitoxin system, and greater identification of the restriction-modification (RM) system in *Enterobacter*, with temperate phages identified for this genus. Although further studies are needed to validate and strengthen the findings, the results point to a promising approach to guide phage bioprospecting in contaminated environments and predict possible phage-host relationships so that more precise experimental analyses can be performed.

Keywords: Metagenomics; phages; antibiotics resistance; host bacteria

## LISTAS DE ABREVIATURAS E SIGLAS

<b>RAM</b>	-	Resistência Anti-Microbiana
<b>AMR</b>	-	Anti Microbial Resistance
<b>NGS</b>	-	New Generation Sequencing
<b>DNA</b>	-	Ácido DesoxirriboNucleico
<b>RNA</b>	-	Ácido RiboNucleico
<b>IDHM</b>	-	Índice de Desenvolvimento Humano Municipal
<b>HAB</b>	-	Habitantes
<b>IBGE</b>	-	Instituto Brasileiro de Geografia e Estatística
<b>MOI</b>	-	Multiplidade de Infecção
<b>DQO</b>	-	Demanda Química de Oxigênio
<b>DBO</b>	-	Demanda Biológica de Oxigênio
<b>CARD</b>	-	Comprehensive Antibiotic Resistance Database
<b>EDTA</b>	-	EthyleneDiamineTetraacetic Acid
<b>pH</b>	-	potencial Hidrogeniônico
<b>CDS</b>	-	CoDing Sequence
<b>GRAs</b>	-	Genes de Resistência à Antibióticos
<b>TMM</b>	-	Trimmed Mean of M-values
<b>SISBIO</b>	-	Sistema de Autorização e Informação em Biodiversidade
<b>SISGEN</b>	-	Sistema Nacional de Gestão do Patrimônio Genético
<b>MAGS</b>	-	Metagenomic Assembled Genomes
<b>CRISPR</b>	-	Clustered Regularly Interspaced Short Palindromic Repeat

## LISTA DE FIGURAS

- Figura 1** - 1-A Esboço da estrutura geral de um bacteriófago. 1-B descrição do ciclo lítico e lisogênico. (Adaptado de Servier Medical Art).
- Figura 2** - PONTO BRAGANÇA- Fonte: [www.googleearth.com](http://www.googleearth.com).
- Figura 3** - Fluxograma das análises *in silico* feitas para a identificação do ponto de maior contaminação e identificação de sequências virais a partir deste ponto.
- Figura 4** - Técnica de eletroforese em gel dos DNAs extraídos para cada ponto do município de Bragança-PA.
- Figura 5** - Box plots da diversidade de Shannon calculada para cada um destes dois grupos: Bactérias e Vírus, respectivamente da esquerda para a direita.
- Figura 6** - Distância de Bray-Curtis entre as regiões para bactérias e vírus, de cima para baixo, comparando as dissimilaridades entre e dentro das 5 regiões do rio.
- Figura 7** - Gráfico de barras empilhadas com os valores médios dos filos do grupo Bactéria e Virus, da esquerda para a direita respectivamente.
- Figura 8** - Mapa de calor de espécies identificadas para o grupo de bactérias (esquerda) e vírus (direita) presentes nas regiões coletadas em Bragança-PA
- Figura 9** - Gráfico de RDA que ilustra a correlação entre os perfis de resistência e as regiões de coleta.
- Figura 10** - Mapa de calor do mecanismo de resistência de MAGs pesquisado com RGI mostrando o perfil de multirresistência dos prováveis hospedeiros dos bacteriófagos selecionados.
- Figura 11** - Diagrama de Sankey ilustrando MAGs em que regiões de CRISPR e genes associados a bacteriófagos foram identificados.

## LISTA DE TABELAS

- Tabela 1** - Coordenadas geográficas dos pontos da cidade Bragança-PA em que as coletas de água foram realizadas.
- Tabela 2** - Resultados dos dados físico químicos obtidos por cada ponto na cidade de Bragança-PA. Nele estão os resultados de 10 parâmetros físico químicos coletados para cada replica e ponto.
- Tabela 3** - Valores de concentração de DNA avaliada por espectrofotometria à 260 nm e à 280 nm, usando um espectrofotômetro NanoDrop™ 1000. Da esquerda para a direita estão os resultados de concentração, índices de pureza e volume obtido. Os valores de cada ponto (Ponto 1 ao 5) estão organizados de cima para baixo.
- Tabela 4** - Número de sequências virais identificadas por cada software por ponto seguido do número de sequências totais obtidas após a concatenação e remoção de sequências redundantes e seleção final baseado nos resultados de CheckV e ViralComplete.
- Tabela 5** - Sequências virais selecionadas no ponto BP1 seguindo os critérios de seleção descritos acima e no tópico 3.4.7.
- Tabela 6** - Sequências virais em que o seu hospedeiro foi identificado usando o software iPHoP para o ponto BP1.
- Tabela 7** - Sequências virais selecionadas no ponto BP2 seguindo os critérios de seleção descritos acima e no tópico 3.4.7.
- Tabela 8** - Sequências virais em que o seu hospedeiro foi identificado usando o software iPHoP para o ponto BP2.
- Tabela 9** - Sequências virais selecionadas no ponto BP3 seguindo os critérios de seleção descritos acima e no tópico 3.4.7.
- Tabela 10** - Sequências virais em que o seu hospedeiro foi identificado usando o software iPHoP para o ponto BP3.
- Tabela 11** - Sequências virais selecionadas no ponto BP4 seguindo os critérios de seleção descritos acima e no tópico 3.4.7.
- Tabela 12** - Sequências virais em que o seu hospedeiro foi identificado usando o software iPHoP para o ponto BP4.
- Tabela 13** - Sequências virais selecionadas no ponto BP5 seguindo os critérios de seleção descritos acima e no tópico 3.4.7.

**Tabela 14** - Sequências virais em que o seu hospedeiro foi identificado usando o software iPHoP para o ponto BP5.

**Tabela 15** - MAG selecionada a partir dos critérios de contaminação, completude e classificação taxonômica usando os dados da região BP1

**Tabela 16** - MAGs selecionadas a partir dos critérios de contaminação, completude e classificação taxonômica usando os dados da região BP3

**Tabela 17** - MAG selecionada a partir dos critérios de contaminação, completude e classificação taxonômica usando os dados da região BP4

**Tabela 18** - Comparação das sequências virais identificadas que provavelmente tem como hospedeiros os MAGs selecionados a partir da tabela 16.

**Tabela 19** - MAG selecionada a partir dos critérios de contaminação, completude e classificação taxonômica usando os dados da região BP5.

**Tabela 20** - Predição do ciclo viral dos bacteriófagos selecionados da região BP4

## SUMÁRIO

1. INTRODUÇÃO.....	14
1.1 O problema da resistência aos antimicrobianos .....	14
1.2 Abordagens usadas para o combate às RAMs .....	15
1.3 Classificação, história e uso de fagos na fagoterapia.....	16
1.4 NGS, metagenômica e bioprospecção de fagos.....	19
1.5 Bioprospecção de Bacteriófagos na Amazônia .....	20
2. OBJETIVOS.....	22
2.1 Objetivo geral .....	22
2.2 Objetivos específicos .....	22
3. MATERIAL E MÉTODOS.....	23
3.1 Autorizações e dados de patrimônio genético .....	23
3.2 Características da cidade de Bragança e pontos de coletas.....	23
3.3 Filtragem e extração do DNA metagenômico .....	25
3.4 Sequenciamento e análises por bioinformática dos dados metagenômicos.....	26
3.4.1 Sequenciamento.....	26
3.4.2 Etapas e fluxograma do trabalho in silico .....	26
3.4.3 Análise de Controle e Qualidade.....	26
3.4.4 Análise do microbioma.....	27
3.4.5 Etapa de Montagem.....	27
3.4.6 Análise de Resistência .....	28
3.4.7 Análise de Sequências Virais .....	28
3.4.8 Montagem de MAGs e correlação bacteriófago-hospedeiro.....	29
4. RESULTADOS .....	31
4.1 Dados físico-químicos .....	31
4.2 Extração do DNA metagenômico .....	32
4.3 Diversidade bacteriana e viral nos dados metagenômicos.....	34
4.4 Abundância bacteriana e viral entre regiões .....	37
4.5 Associação de perfil taxonômico e resistência por análise de redundância (RDA) .....	41
4.6 Identificação de sequências de bacteriófagos de interesse a partir dos dados metagenômicos .....	43

4.6.1	Sequências virais selecionadas em BP1 .....	44
4.6.2	Sequências virais selecionadas em BP2 .....	46
4.6.3	Sequências virais selecionadas em BP3 .....	47
4.6.4	Sequências virais selecionadas em BP4 .....	49
4.6.5	Sequências virais selecionadas em BP5 .....	51
4.7	Geração de MAGs de interesse a partir dos dados metagenômicos .....	52
4.7.1	Geração de MAGs e classificação taxonômica .....	52
4.7.2	Geração de MAGs no ponto BP1 .....	53
4.7.3	Geração de MAGs no ponto BP2 .....	53
4.7.4	Geração de MAGs no ponto BP3 .....	53
4.7.5	Geração de MAGs no ponto BP4 .....	54
4.7.6	Geração de MAGs no ponto BP5 .....	56
4.8	Análises de perfil de resistência entre a relação bacteriófago-hospedeiro da região BP4 57	
4.9	Mecanismos de defesa e suscetibilidade a bacteriófagos na região BP4.....	59
4.10	Predição de presença de sistemas anti-CRISPR e do ciclo de vida viral dos bacteriófagos na região BP4 .....	62
5.	CONCLUSÕES .....	65
6.	PERSPECTIVAS .....	67
7.	REFERÊNCIAS BIBLIOGRÁFICAS .....	68
8.	ANEXOS .....	80

## 1. INTRODUÇÃO

### 1.1 O problema da resistência aos antimicrobianos

A resistência bacteriana tem tomado proporções perigosas pelo uso indevido de antibióticos em humanos e animais. Segundo a Organização Mundial de Saúde, a resistência antimicrobiana tem sido reconhecida como uma das principais ameaças à saúde global no século 21. Somente em 2019 foram estimados que cerca de 4.95 (3.62 - 6.57) milhões de mortes foram associadas a problemas relacionados à resistência bacteriana a antimicrobianos, com cerca de 1.27 (0.911 - 1.71) milhões de mortes diretamente atribuídas a resistência antimicrobiana por bactérias (RAM-Resistência Antimicrobiana) (MURRAY *et al.*, 2022).

A utilização indiscriminada de antimicrobianos na prática clínica, descarte inadequado no meio ambiente, uso indevido pela indústria agropecuária, além da prescrição indiscriminada de antibióticos devido aos protocolos de tratamento de COVID-19, são fatores que podem resultar no maior surgimento de patógenos resistentes a múltiplas drogas. Nesse cenário, observa-se um aumento crescente no número de infecções comuns que estão tornando-se cada vez mais difíceis de tratar devido ao surgimento da resistência a múltiplas drogas, o que resulta em internações mais longas com impactos nos custos médico-hospitalares e principalmente no aumento da mortalidade (CENTERS FOR DISEASE CONTROL AND PREVENTION, 2022).

O uso irrestrito de antibióticos pode influenciar na prevalência de genes de resistência a antibióticos em diversos organismos, incluindo bactérias patogênicas que possam ser uma ameaça por carrear tais genes, tornando cada vez mais difícil tratar infecções que tenham este aparato de genes de resistência (MACLEAN; MILLAN, 2019). Tais genes de resistência são encontrados em praticamente todos os tipos de ambientes, como no ar (BECSEI *et al.*, 2021; MINOR *et al.*, 2023), solo (CARDENAS ALEGRIA *et al.*, 2022; ZHENG *et al.*, 2022), e água (CARDENAS-ALEGRIA *et al.*, 2024; MACEDO *et al.*, 2024).

Neste contexto, o estudo do resistoma, conjunto de genes de resistência presente em uma amostra, tem sido utilizado para avaliar o ambiente quanto à existência de bactérias multirresistentes, para se compreender a ecologia e evolução da resistência aos antibióticos em ambientes não-clínicos, considerando que o desenvolvimento de novos antimicrobianos pela indústria farmacêutica é complexo e extremamente demorado (WALSH, 2013).

A maioria das classes de antibióticos utilizadas foi descoberta durante a chamada "era de ouro" da pesquisa de antibióticos, em que muitos foram isolados do solo e estão restritos a grupos taxonômicos derivados principalmente do grupo *Actinomyces*. No entanto, apesar da descoberta de novos antibióticos, independentemente da sua eficácia, eventualmente irá aparecer algum microrganismo com fenótipo de resistência (AMINOV, 2010).

## 1.2 Abordagens usadas para o combate às RAMs

Visto este problema, diversas abordagens alternativas têm sido adotadas para se combater esta ameaça constante de resistência à antimicrobianos (RAM) (KIRIENKO; RAHME; CHO, 2019). Dentre elas destacam-se algumas a seguir:

- a) Estratégias de terapia antivirulência (DEHBANIPOUR; GHALAVAND, 2022);
- b) Reaproveitamento de medicamentos já usados em outras finalidades que possam atuar contra organismos multirresistentes a antibióticos ou a combinação de medicamentos para aumentar a eficácia dos tratamentos (KIRIENKO; RAHME; CHO, 2019; PUSHPAKOM *et al.*, 2018);
- c) Bioprospecção de novos antibióticos a partir de outros ambientes alternativos ao encontrado no solo como os antibióticos provenientes do mar (HUGHES; FENICAL, 2010; TORTORELLA *et al.*, 2018);
- d) Bioprospecção de peptídeos antimicrobianos e outros compostos de animais e plantas que possam atuar como antimicrobianos (HANCOCK; SAHL, 2006);
- e) Bioprospecção de fagos que possam atuar contra bactérias multirresistentes (BRIVES; POURRAZ, 2020; HAMPTON; WATSON; FINERAN, 2020; LIN; KOSKELLA; LIN, 2017).

Nesse contexto, os bacteriófagos ou 'fagos', surgem como uma ferramenta alternativa promissora no combate a essas bactérias, uma vez que naturalmente representam uma constante ameaça a elas, coevoluindo junto com seu hospedeiro com estratégias de infecção e replicação durante sua existência. Os fagos são responsáveis por aproximadamente 20% a 40% da mortalidade diária de bactérias no mundo e exibem uma diversidade maior do que qualquer outra espécie,

estimando-se uma quantidade global de fagos em torno de  $10^{31}$  partículas virais (HAMPTON; WATSON; FINERAN, 2020).

### 1.3 *Classificação, história e uso de fagos na fagoterapia*

Os bacteriófagos (ou fagos) são vírus pequenos que contêm DNA ou RNA e têm a capacidade de infectar seletivamente bactérias. Eles são encontrados em todos os ambientes onde bactérias hospedeiras estão presentes. Os fagos manipulam o maquinário celular da bactéria para produzir partículas virais, que se multiplicam e formam novos fagos dentro do hospedeiro. Posteriormente, esses fagos recém-formados são liberados para infectar novas bactérias. De acordo com o comportamento dos fagos e seu ciclo de vida viral, eles podem ser classificados em dois tipos: temperados e virulentos (KAKASIS; PANITSA, 2019) (Figura 1).

**Temperados:** são fagos que ativam o ciclo lisogênico de um vírus, em que o material genético viral permanece dentro da bactéria hospedeira, que continua a funcionar e se reproduzir normalmente junto ao material genético viral. Essa situação persiste indefinidamente até que ocorra alguma mudança, que faça a bactéria alternar para o ciclo lítico (GARC, 2020; KAKASIS; PANITSA, 2019) (Figura 1a).

**Virulentos:** são fagos que ativam o ciclo lítico assim que introduzem seu material genético no seu hospedeiro. Neste ciclo, o material genético viral controla a maquinaria celular da bactéria hospedeira e a força a replicar o material genético viral, bem como a produzir proteínas virais que resultam na formação de novos fagos dentro do hospedeiro. Após a produção de uma grande quantidade de fagos, a membrana da bactéria rompe (lise) e libera os novos fagos, que estão prontos para infectar novas bactérias e reiniciar o ciclo (GARC, 2020) (Figura 1b).

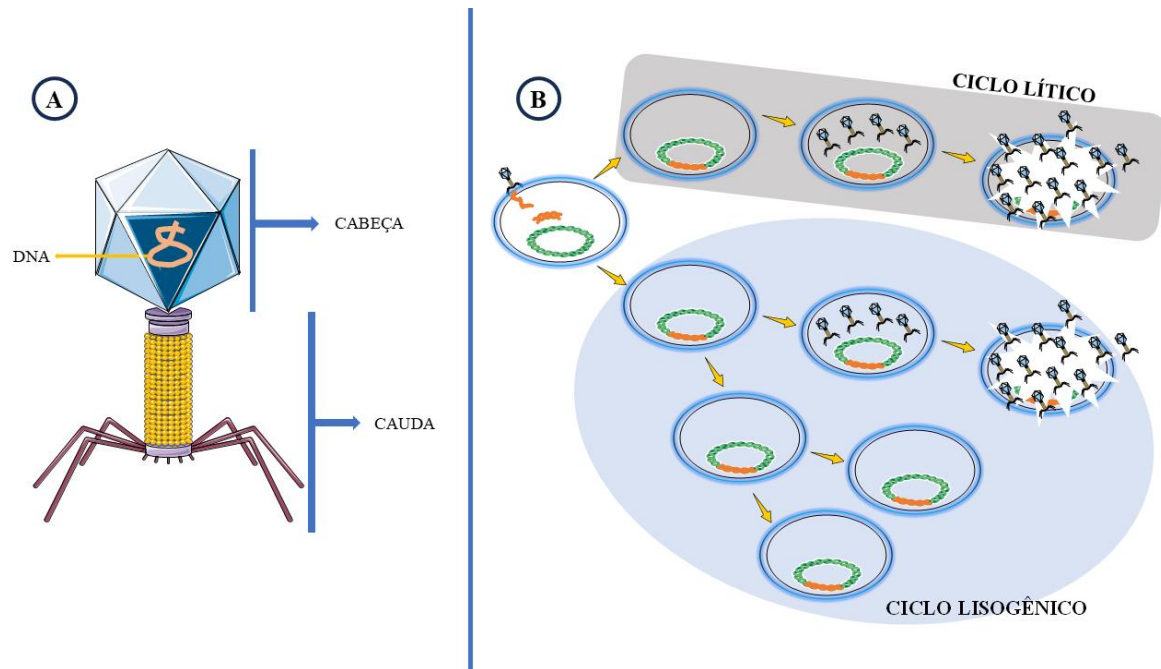


Figura 1: 1A Esboço da estrutura geral de um bacteriófago, sendo geralmente dividido entre a região da cauda e a cabeça, onde se localiza o seu material genético viral. 1B descrição do ciclo lítico e lisogênico que os bacteriófagos podem apresentar após infectar seu hospedeiro, sendo uma das características para classificar os fagos entre virulentos e temperados. (Adaptado de Servier Medical Art )

Para a obtenção de bacteriófagos para o seu uso na fagoterapia (uso de fagos para combater infecções bacterianas), é necessário seguir alguns critérios como: 1- estabelecimento de bibliotecas de fagos; 2- isolamento e identificação de bactérias patogênicas; 3- triagem de fagos eficazes contra bactérias patogênicas; 4- preparação da formulação de fagos; 5- estratégia e via de administração da preparação de fagos; 6- monitoramento da eficácia da fagoterapia; e 7- detecção do surgimento de cepas resistentes a fagos (CUI *et al.*, 2019). Para o estabelecimento de uma biblioteca de fagos, é necessário possuir uma ampla variedade dos mesmo devido a sua alta e restrita especificidade, além de eles não podem apresentar genes “não desejados” como genes de resistência à antibióticos (GRAs), genes de toxinas e eles precisam ser fagos virulentos (CUI *et al.*, 2017, 2019)

A partir destes conceitos, é necessário salientar algumas características que devem ser tomadas durante a seleção de fagos para fagoterapia . Fagos do tipo temperados não são uma boa

escolha como antimicrobianos em função de sua característica de não destruir a bactéria imediatamente e do risco de o conteúdo viral conter alguma sequência genética que promova a adição de algum fator de virulência na cepa bacteriana e/ou a adição de algum gene de resistência a antibióticos, o que acaba sendo uma característica benéfica às bactérias que são infectadas por estes tipos de fagos no ambiente por conferir esta variabilidade genética e a chance de adquirir genes de resistência (KAKASIS; PANITSA, 2019). Assim, somente fagos virulentos (que ativam o ciclo lítico) podem ser usados na fagoterapia. No entanto, fagos temperados também podem ser isolados do ambiente e posteriormente serem modificados geneticamente para tornarem-se líticos para serem usados para fagoterapia (STRATHDEE *et al.*, 2023).

O uso de bacteriófagos como forma de tratamento contra bactérias patogênicas é uma técnica antiga descoberta por volta de 1919, cerca de dois anos depois da descoberta de um “micróbio invisível” por Felix d’Herelle na qual ele batizou de bacteriófago ou “comedores de bactéria” (1917). Felix d’Herelle usou fagos isolados de culturas de fezes sobreviventes de disenteria bacilar como forma de tratamento e aplicou em 3 irmãos que tiveram a mesma comorbidade, obtendo melhoras dos pacientes dentro de 24 horas, dando assim início ao uso da fagoterapia (GARC, 2020; KAKASIS; PANITSA, 2019; STRATHDEE *et al.*, 2023).

Logo depois, devido à descoberta dos antibióticos (como a penicilina) que conseguiam eliminar uma variedade de infecções microbianas, diferente dos fagos que são bastante específicos a um grupo de bactérias, o uso de fagos se tornou restrito a regiões do leste europeu em que a disponibilidade de antibióticos era baixa. Isso levou à consolidação do uso de antibióticos na maioria das demais regiões do mundo e o consequente aumento global de microrganismos resistentes no decorrer dos anos seguintes (GARC, 2020; VENTOLA, 2015).

No entanto, terapias alternativas como esta do uso de fagos ressurgiram e com ela maior retomada dos trabalhos relacionados ao isolamento, seleção e melhoramento de fagos para serem usados contra organismos multirresistentes à antibióticos (KIRIENKO; RAHME; CHO, 2019).

Para isolar bacteriófagos, é possível coletá-los de diversos ambientes com presença de bactérias hospedeiras. No entanto, devido à grande diversidade e pequeno tamanho dos organismos, torna-se difícil e oneroso encontrar uma metodologia que permita identificar a maioria ou todos os fagos potenciais em um determinado ambiente.

#### 1.4 NGS, metagenômica e bioprospecção de fagos

Com o passar dos anos e com o desenvolvimento de novas tecnologias, surgiram formas de visualização e caracterização de fagos por meio de microscopia eletrônica, e pela análise de código genético dos fagos (DNA ou RNA) por meio de tecnologias de sequenciamento. A partir dos dados de sequenciamento obtidos, é possível caracterizar e identificar as proteínas e regiões codificantes presentes em diversos tipos de bacteriófagos (ACKERMANN, 2012; KANZI *et al.*, 2020).

Uma abordagem eficaz para o rastreamento de fagos é o uso de tecnologias avançadas de sequenciamento de próxima geração (NGS - *Next Generation Sequencing*). Esse termo refere-se a novos sequenciadores que usam abordagens de sequenciamento em paralelo de forma massiva, produzindo milhões de sequências curtas por preço e tempo menores em relação a outros métodos convencionais como o sequenciamento por Sanger (KANZI *et al.*, 2020). Com o NGS, é possível "ler" o material genético isolado de um ambiente específico de forma rápida e analisar as sequências obtidas usando ferramentas de bioinformática.

A partir dessa abordagem é possível sequenciar com agilidade fagos e outros organismos isolados em laboratório e conhecer mais sobre suas características funcionais e taxonômicas a partir de seu material genético, no entanto muitos organismos existentes não são cultiváveis ou mesmo facilmente isoláveis. Portanto, surge a necessidade de utilizar análises por metagenoma, em que todo o material genético de uma comunidade é analisado (KANZI *et al.*, 2020).

O termo metagenômica é derivado do conceito estatístico de metanálise (processo de combinar estatisticamente análises separadas) e genômica (análise compreensiva do material genético de um organismo). A partir dessa metodologia é possível obter um panorama mais abrangente do DNA total de um ambiente, incluindo sequências de organismos desconhecidos e frequentemente não cultiváveis. Essa metodologia é extremamente útil para explorar a diversidade viral e bacteriana em diferentes ambientes e pode auxiliar na descoberta de novos fagos e na compreensão de suas interações com as bactérias hospedeiras (SCHLOSS; HANDELSMAN, 2003; SIMMONDS *et al.*, 2017).

A partir de dados metagenômicos, é possível identificar partes de sequências de DNA virais presentes no material genético analisado, fornecendo uma melhor compreensão das características e quantidades de bacteriófagos de DNA e bactérias hospedeiras em um determinado ambiente (SIMMONDS *et al.*, 2017). Sendo assim, uma boa abordagem é realizar um levantamento sobre

potenciais bactérias e bacteriófagos com maior potencial de ser estudado, tanto por questões de monitoramento ambiental, quanto de seu potencial biotecnológico que tais organismos possam apresentar. Um estudo metagenômico do ambiente antes de etapas experimentais mais focadas em organismos específicos também pode preparar o pesquisador para focar em análises *in vitro* que tenham maior potencial de gerar um produto biotecnológico, partindo da identificação de organismos e genes funcionais de interesse por análises *in situ* e ,posteriormente, realizar um planejamento experimental focado nos organismos identificados por dados metagenômicos, aumentando assim as chances de obter sucesso de isolamentos *in situ*.

### 1.5 Bioprospecção de Bacteriófagos na Amazônia

Além da utilidade do uso da metagenômica para o *screening* do material genético de DNA presente no ambiente estudado, a seleção do ambiente a ser estudado para realizar a bioprospecção de fagos é uma etapa tão importante quanto para a obtenção de alta abundância de fagos prospectáveis e conseqüentemente maior rendimento de isolamento de fagos que atuem contra um determinado patógeno.

Estudos preliminares mostraram que os melhores ambientes para bioprospecção de fagos de interesse são ambientes mais contaminados, em que a bactéria patogênica esteja mais presente, como em esgotos e regiões com água contaminadas (KAKASIS; PANITSA, 2019; WEBER-DĄBROWSKA *et al.*, 2016). No entanto, levando-se em conta a possibilidade de que um ambiente com alta biodiversidade provavelmente apresenta variedade maior de organismos, é possível que a prospecção de fagos em tais regiões revele maior diversidade de fagos com potencial de combater bactérias patogênicas.

Além disso, uma alta diversidade microbiana está negativamente correlacionada com a presença de genes de resistência à antibióticos (ARGs), servindo como uma barreira contra a proliferação de tais genes (KLÜMPER *et al.*, 2024). O que torna interessante se avaliar a presença de bacteriófagos em regiões de maior impacto antropogênico em regiões de alta biodiversidade, como em cidades inseridas dentro do bioma Amazonia.

Quanto à biodiversidade, o bioma Amazônico é o bioma mais diverso do mundo; onde estima-se que ele ocupe cerca de 1/4 de toda a biodiversidade presente no planeta. Assim, buscar formas de conhecer e preservar a sua biodiversidade, mantendo a preservação deste bioma de tal

forma que valor econômico possa ser agregado à região de forma sustentável, é uma das ações necessárias para a sua preservação (BETTS; MALHI; ROBERTS, 2008; BROUWER *et al.*, 2022).

Assim, é possível identificar bacteriófagos em regiões da floresta amazônica, em regiões com as maiores alterações antropogênicas mediante a alta biodiversidade (ALVES *et al.*, 2020; BROUWER *et al.*, 2022).

A região amazônica engloba o município de Bragança, próximo ao litoral, conhecido como polo pesqueiro do estado do Pará, e responsável pela exportação de grande variedade de espécies de peixes para outras regiões. Além disso, o crescimento contínuo do município e a ocupação de áreas protegidas podem causar diversos impactos ambientais (SANTOS; QUARESMA, 2022).

Assim, considerando-se a alta biodiversidade presente na região amazônica e o risco do surgimento de novas cepas patogênicas multirresistentes ao redor do mundo, torna-se importante identificar bactérias patogênicas que apresentem elementos de resistência a antibióticos e buscar, nos ambientes em que estas bactérias estão presentes, formas de combatê-las. Neste contexto, a bioprospecção de bacteriófagos, naturalmente presente nestes ambientes, surge como uma alternativa para combater bactérias patogênicas resistentes. Além do mais, é crescente o uso de bacteriófagos nos mais variados campos, como na indústria alimentícia (MOYE; WOOLSTON; SULAKVELIDZE, 2018), em tratamentos alternativos de saúde humana (BRIVES; POURRAZ, 2020; KAKASIS; PANITSA, 2019; STRATHDEE *et al.*, 2023) e de animais (LIU *et al.*, 2020). Tornando-se assim, oportuno o uso de bacteriófagos como produtos comerciais como alternativa ao uso indiscriminado de antibióticos (BRIVES; POURRAZ, 2020; STRATHDEE *et al.*, 2023).

Além disso, ainda é possível fazer o biomonitoramento acessando os dados do sequenciamento do metagenomas, possibilitando a identificação de novos fatores de resistência e atuando no biomonitoramento de uso de antibióticos a partir dos seus resultados (CHEN *et al.*, 2007; DONATO *et al.*, 2010; SHOKRALLA *et al.*, 2012; WALSH, 2013).

Este trabalho buscou identificar bacteriófagos de interesse biotecnológico a partir de dados metagenômicos e associá-los com *Metagenomic Assembled Genomes* (MAGs) de bactérias alvos destas sequências analisadas. E buscar, nas sequências virais e MAGs selecionadas, presença de genes de resistência a antibióticos, genes de resistência a antibióticos e realizar análise funcional destas sequências selecionadas.

## 2. OBJETIVOS

### 2.1 *Objetivo geral*

Fazer a bioprospecção de fagos que possam atuar em bactérias com genótipo de resistência a partir de dados metagenômicos obtidos da cidade de Bragança-PA,

### 2.2 *Objetivos específicos*

- Analisar a diversidade e composição de bactérias e vírus a partir dos dados do sequenciamento;
- Investigar o perfil de genes de resistência entre as regiões coletadas;
- Correlacionar a composição de bactérias presentes nos dados metagenômicos com os metadados obtidos;
- Selecionar o melhor ponto a se realizar a bioprospecção de fagos;
- Identificar as sequências virais que possam atuar contra cepas resistentes a antibióticos como hospedeiras;
- Analisar as sequências virais dos dados metagenômicos para presença de genes de resistência e abundância de vírus e bactérias.

### 3. MATERIAL E MÉTODOS

#### 3.1 *Autorizações e dados de patrimônio genético*

Todos os pontos de coleta coletados (Figura 2 e 3) foram previamente autorizados pelo SISBIO (Sistema de Autorização e Informação em Biodiversidade) como “atividades com finalidade científica”, com número de protocolo de “81367-2”, e também informados ao SISGEN (Sistema Nacional de Gestão do Patrimônio Genético), tendo número de cadastro gerado de “AD70806”.

#### 3.2 *Características da cidade de Bragança e pontos de coletas*

Este trabalho foi realizado no município de Bragança, pertencente ao estado do Pará. Amostras de águas foram coletadas e filtradas no rio Cereja, que é um afluente do rio Caeté, por se tratar de um rio que corta a cidade e, assim, ter melhor ideia do impacto antropogênico causado ao longo do fluxo do corpo hídrico.

As amostras foram coletadas em 5 pontos, onde em cada ponto foram realizadas triplicatas na direção jusante. A escolha dos pontos foi baseada de acordo com os prováveis níveis de contaminação da água, onde o primeiro ponto corresponde a uma área provavelmente menos contaminada (ponto obtido em área de nascente) e os seguintes pontos seguem o caminho a jusante da água, geralmente associados a áreas de maior impacto ambiental, sendo o ponto 5 provavelmente o ponto de maior contaminação.

A cidade de Bragança foi selecionada por ser um grande polo pesqueiro do estado do Pará, responsável por exportar grande quantidade de pescado para todo o país. Além disso, no rio Cereja são relatados impactos ambientais decorrentes do crescimento urbano e ocupações em áreas antes protegidas que causaram uma série de impactos ambientais (SANTOS; QUARESMA, 2022). Sendo uma cidade com boas chances de identificação de fagos para bactérias patogênicas devido aos impactos ao longo dos anos.

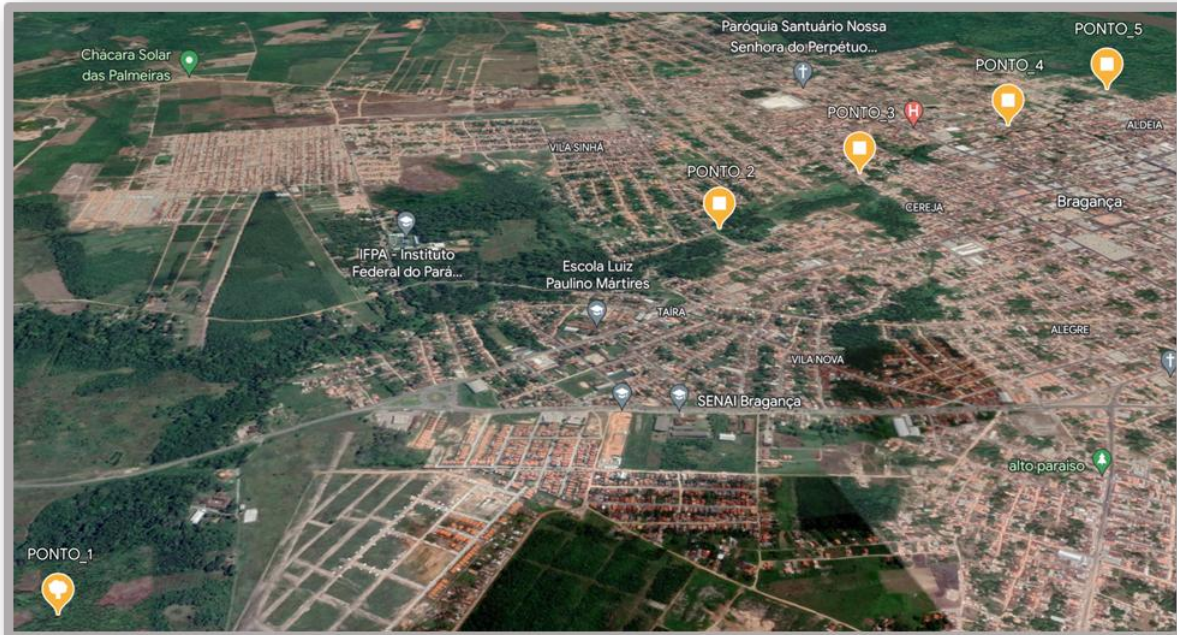


Figura 2: PONTO BRAGANÇA- Na figura acima, marcados em amarelo, estão os pontos geográficos dos locais de coletas realizados para a filtragem da água coletada, extração de material genético (DNA) e posterior sequenciamento das amostras biológicas. Foram realizadas triplicatas para cada região de coleta. Os pontos selecionados foram ordenados do ponto 1 ao 5 de acordo com as prováveis áreas de contaminação. Fonte: [www.googleearth.com](http://www.googleearth.com).

Os pontos previamente selecionados (Figura 2) foram coletados nas coordenadas geográficas mostradas a seguir (Tabela 1), em que mostram as localizações geográficas dos pontos em que as coletas foram realizadas para a cidade de Bragança-PA seguindo o protocolo descrito na metodologia (Tópico 3.2).

Durante as coletas dados físico-químicos também foram obtidos usando um sensor multiparâmetro modelo HI 9828 Hanna, em que foram obtidos cerca de 10 parâmetros físico-químicos por réplica obtida.

<b><i>Pontos de Coleta - Bragança</i></b>				
	<i>Latitude</i>	<i>Longitude</i>	<i>Elevação</i>	<i>Elevação (msl)</i>
<i>PI</i>	-1,068	-46,791	12,3	37,69

P2	-1,057	-46,777	0,4	25,79
P3	-1,054	-46,772	-151,6	9,79
P4	-1,051	-46,765	-19	6,39
P5	-1,048	-46,760	-20,9	4,49

Tabela 1: Coordenadas geográficas dos pontos da cidade Bragança-PA em que as coletas de água foram realizadas.

### 3.3 Filtragem e extração do DNA metagenômico

Durante a coleta de amostras de água foi usada uma garrafa de van-Dorn, desinfetada e lavada com água destilada que então teve seu material transferido para garrafas estéreis de polipropileno com capacidade de 5L. Antes e depois de cada coleta, a garrafa de Van Dorn foi desinfetada com uma solução de 10% de hipoclorito de sódio seguida de três lavagens com água destilada. As garrafas com amostras foram alocadas em gelo e transportadas ao laboratório. A água coletada foi primeiramente pré-filtrada (para a remoção de partículas de impurezas grandes que possam afetar a etapa de extração de DNA) e filtrada por pressão negativa através de uma membrana de nitrocelulose com poros de 14  $\mu\text{m}$  (Qually) e 0,22  $\mu\text{m}$  (Millipore) respectivamente. As membranas pré-filtradas foram descartadas e as filtradas de 0,22  $\mu\text{m}$  foram transferidas para tubos de polipropileno (tubos Falcon) de 15 mL contendo solução de preservação de DNA (50 mM Tris-HCl, 500 mM NaCl, 125 mM EDTA pH 8,0) e armazenadas em  $-80\text{ }^{\circ}\text{C}$  para posterior extração do DNA metagenômico.

O DNA metagenômico foi extraído com o kit DNeasy PowerSoil Pro (QIAGEN), seguindo o protocolo do fabricante. A concentração de DNA foi avaliada por espectrofotometria à 260 nm e à 280 nm, usando um espectrofotômetro NanoDrop™ 1000, e a integridade do DNA foi avaliada por eletroforese em gel de agarose com brometo de etídio 0,5  $\mu\text{g}/\text{mL}$ . As concentrações de DNA precisaram atingir uma concentração mínima de 50 ng/uL, com índice de pureza entre 1,4 até 2,0, para serem consideradas aptas para o sequenciamento. Para avaliar a pureza do material genético obtidos foi determinado os índices de pureza (IP) e, de acordo com a relação das absorbâncias: A260/A280 para as amostras estarem no intervalo entre 1,8 - 2,0. As amostras de DNA com valores de IP no esperado foram armazenadas a  $-20\text{ }^{\circ}\text{C}$  até o seu uso.

### 3.4 Sequenciamento e análises por bioinformática dos dados metagenômicos

#### 3.4.1 Sequenciamento

O material genético extraído foi encaminhado para sequenciamento na plataforma Illumina (NovaSeq 6000) por uma biblioteca pareada de 150 pares de bases.

#### 3.4.2 Etapas e fluxograma do trabalho *in silico*

Para obtenção de escopo geral das etapas realizadas, um fluxograma foi feito para detalhar as etapas durante cada análise e os programas usados para obtenção de resultados. A figura abaixo (Figura 3) mostra um fluxograma com os caminhos e planejamentos de etapas para a realização deste trabalho.

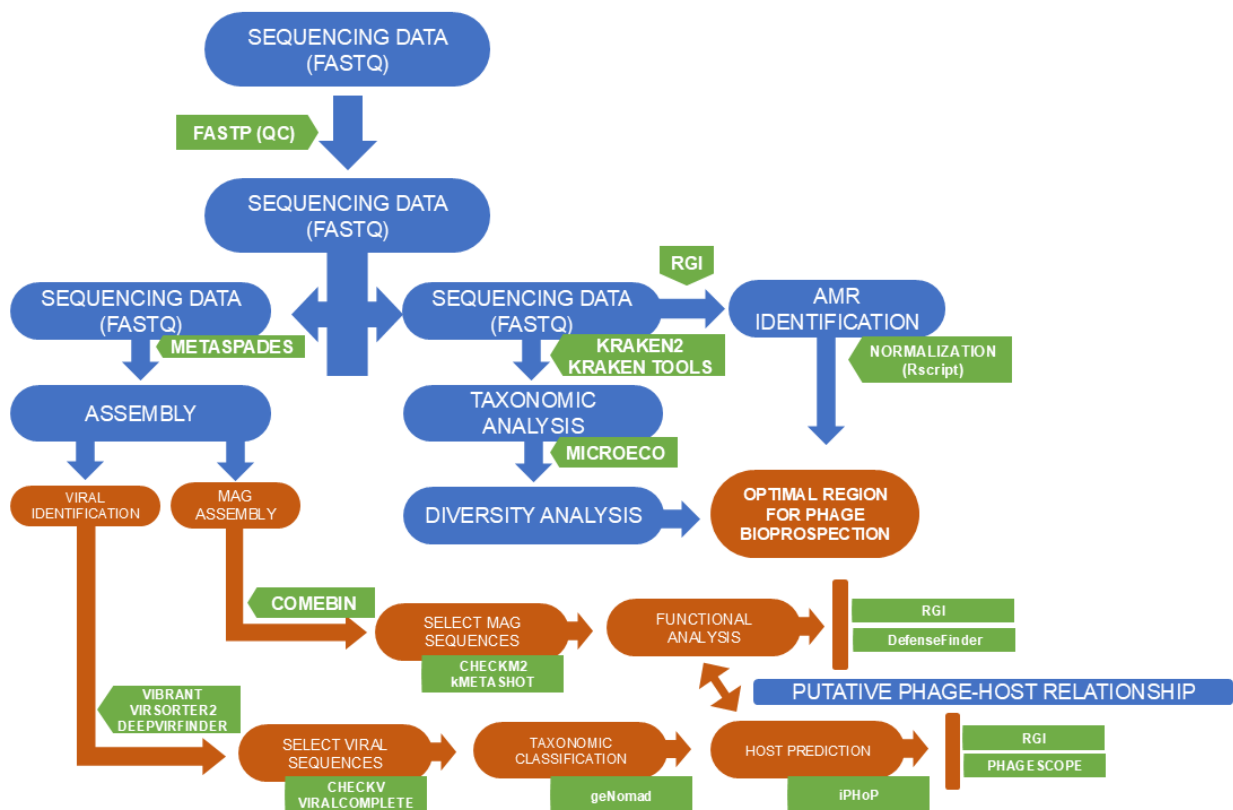


Figura 3: Fluxograma das análises *in silico* feitas para a identificação do ponto de maior contaminação e identificação de sequências virais a partir deste ponto.

#### 3.4.3 Análise de Controle e Qualidade

Os dados brutos de sequenciamento em formato FASTQ foram primeiramente rodados pelo programa FASTQC (ANDREWS; OTHERS, 2010) para avaliação das características dos dados, tamanhos e qualidades das leituras obtidas após o sequenciamento. Então, os dados foram trimados por tamanho e qualidade *Phred* usando o programa fastp (CHEN *et al.*, 2018), onde leituras de tamanho menor do que 75 pares de bases e regiões com valores de qualidade *Phred* menores que 20 foram removidas. Os dados resultantes foram usados para as etapas seguintes de análise de diversidade e composição do microbioma e montagem das leituras em *contigs*.

#### 3.4.4 *Análise do microbioma*

Para análise da diversidade e identificação de organismos presentes nos dados metagenômicos, as leituras resultantes da etapa de controle de qualidade foram usadas como entradas ao software Kraken2 (WOOD; LU; LANGMEAD, 2019), tendo como referências o banco de dados PFP (*Standard plus Refeq protozoa, fungi & plant*) gerado em 2024 obtido pela “AWS-INDEXES” com RefSeq de dados de archaea, bactéria, vírus, plasmídeos, protozoários, fungos e plantas.

Os resultados *report* obtidos do Kraken2 foram mesclados usando *scripts* presentes no Krakentools (LU *et al.*, 2022; WOOD; LU; LANGMEAD, 2019) e adaptados junto aos metadados para serem analisados no R (v. 4.2) pelo pacote *microeco* (LIU *et al.*, 2021). Nele os clados de vírus e bactérias foram filtrados, calculados e plotados análises de diversidade microbiana e posteriormente associá-los aos metadados obtidos. Usando este pacote, os dados foram pré-processados, e plotados valores como abundância taxonômica, diversidade alfa e beta, bem como análise de redundância (RDA) entre os dados de diversidade dos dados ambientais em associação aos metadados obtidos.

#### 3.4.5 *Etapa de Montagem*

Durante a etapa de montagem, o programa MetaSpades (NURK *et al.*, 2017) foi usado para a montagem das leituras em *contigs*, devido ao mesmo ter mostrado um melhor rendimento para a obtenção e caracterização de sequências virais em dados metagenômicos quando comparados com outras 16 abordagens de montagens metagenômicas (SUTTON *et al.*, 2019).

### 3.4.6 *Análise de Resistência*

A análise de resistência das leituras foi feita pela base de dados CARD (*Comprehensive Antibiotic Resistance Database- Versão 3.2.6*) pelo programa RGI (ALCOCK *et al.*, 2020), que é capaz de identificar genes de resistência, variantes, e genes de prevalência para se detectar GRAs (Genes de Resistência à Antibióticos). Estes resultados foram usados para se buscar associar os gêneros e regiões na qual estes genes de resistência estão presentes e correlacionar com a região de maior impacto ambiental.

Para se comparar o perfil do microbioma com a presença de genes de resistência a antimicrobianos (GRAs) entre as regiões foram realizados os seguintes critérios. Os resultados foram filtrados usando os critérios de seleção de leituras que tenham ao menos 80% de cobertura, e foram normalizadas seguindo a fórmula  $C = 1000 * \sqrt[2]{R}/T$ , onde C é a abundância relativa, R o número de leituras alinhadas aos genes de resistência, e T o número total de leituras (INDA-DÍAZ *et al.*, 2023). Baseado nestes parâmetros, alguns scripts foram desenvolvidos em R (v4.4.1) para gerar uma tabela de abundância com todas as regiões para verificar o perfil de resistência entre os genes funcionais. Esta tabela foi posteriormente usada como metadado para comparar o perfil do microbioma de bactérias usando Análise de redundância (RDA) para identificar as áreas mais contaminadas baseados nos GRAs identificados para cada região.

Após a identificação, os dados metagenômicos foram usados para se buscar sequências virais a partir dos dados montados, e, em paralelo, MAGs foram gerados e selecionados baseados em dados de qualidade e classificação taxonômica para a posterior identificação por inferência de possíveis bactérias hospedeiras das sequências virais.

### 3.4.7 *Análise de Sequências Virais*

Para a identificação de sequências virais, os programas utilizados foram: VIBRANT (KIEFT; ZHOU; ANANTHARAMAN, 2020), VirSorter2 (GUO *et al.*, 2021) e DeepVirFinder (REN *et al.*, 2020), por terem obtidos os maiores valores de score F1 usando dados microbianos simulados e outros dados reais para a identificação de sequências virais (DANTAS *et al.*, 2024; HO *et al.*, 2023). As sequências resultantes foram concatenadas e filtradas baseado nos seus resultados de qualidade obtidos pelos programas ViralComplete (ANTIPOV *et al.*, 2020) e checkV

(NAYFACH *et al.*, 2021), onde sequências com valores de completude maiores do que 90% e contaminação menores que 10% foram selecionados.

Para a predição de prováveis hospedeiros das sequências virais foi utilizado o pipeline iPHoP (ROUX *et al.*, 2023) usando o seu banco de dados “iPHoP\_db\_Jun25\_rw”, onde os prováveis gêneros hospedeiros foram comparados por inferência com as MAGs obtidas e classificadas taxonomicamente de acordo com o seu provável grupo hospedeiro. As sequências virais que tiveram seus hospedeiros identificados que corresponderam ao táxon obtido das MAGs do mesmo ponto foram selecionadas para análises posteriores.

As sequências virais selecionadas foram analisadas quanto a características funcionais que possam apresentar, como a presença de genes de resistência a antibióticos usando o programa RGI (ALCOCK *et al.*, 2020), presença de elementos móveis usando geNomad (CAMARGO *et al.*, 2024), presença de sequências anti-CRISPR-Cas, presença de genes de virulência e análise de ciclo viral usando a ferramenta web PhageScope (WANG *et al.*, 2024a).

#### 3.4.8 *Montagem de MAGs e correlação bacteriófago-hospedeiro*

Para realizar a etapa de montagem de genomas a partir de dados metagenômicos ou MAGs, conhecido também como “binning” foi utilizado o programa COMEBin (WANG *et al.*, 2024b) por ter sido o programa que apresentou as melhores pontuações para dados de “short-reads” quando comparado com outras ferramentas de binning (HAN; WANG; ZHU, 2025). As MAGs obtidas foram avaliadas e selecionadas de acordo com os valores de qualidade obtidos pela ferramenta CheckM2 (CHKLOVSKI *et al.*, 2023), onde foram selecionadas MAGs com valores de completude maiores do que 50% e contaminação com valores menores do que 10%.

As MAGs selecionadas foram classificadas taxonomicamente posteriormente usando a ferramenta kMetaShot (DEFAZIO *et al.*, 2024), e a partir destes resultados as MAGs que corresponderam taxonomicamente ao grupo hospedeiros identificados pelo iPHoP (ROUX *et al.*, 2023) foram selecionadas para outras análises funcionais que possam estar associadas a relação bacteriófago-hospedeiro, como identificação de genes de resistência a antibióticos usando o software RGI (ALCOCK *et al.*, 2020), presença de sequências de CRISPR-Cas, e presença de elementos associados a defesa contra bacteriófagos nas MAGs selecionadas usando o software DefenseFinder (TESSON *et al.*, 2022).



## 4. RESULTADOS

### 4.1 Dados físico-químicos

Os dados físico-químicos de cada ponto também foram obtidos paralelamente enquanto as coletas de água foram realizadas (Tabela 2).

Foram realizadas medições de 10 parâmetros físico químicos. Os parâmetros medidos foram: Potencial de pH ou potencial hidrogeniônico (pH e pHmV); temperatura em graus celcius; Pressão em mbar (milibar); Resistividade em milliohm versus cm ( $m\Omega \cdot cm$ ); condutividade elétrica em EC, quanto maior estes valores, maior a capacidade da água conduzir eletricidade devido a maior disponibilidade de íons na mesma; Disponibilidade de Sólidos Totais Dissolvidos TDS (Sólidos Totais Dissolvidos); Salinidade; Propriedades Óxido redundantes da água ORP (Potencial de Oxidação-Redução); e oxigênio dissolvido da água DO (Oxigênio Dissolvido) em ppm (partes por milhão).

Nas tabela 2 a seguir estão os resultados dos 10 parâmetros físico químicos coletados para cada replica e ponto das cidades de Bragança. Os resultados dos valores estão marcados por cor e intensidade entre vermelho, branco e verde, onde vermelho indica os maiores valores dentre os pontos, branco indica valores intermediários e verde indica os menores valores dentre os pontos (colunas).

		Potencial de H		Temperatura	Pressão	Resistência	EC	TDS	Salinidade	ORP	DO
		pH	pHmV	C	mbar	M $\Omega$ cm	uS/cm	ppm		mV	ppm
P1	P1A	3,98	186,2	27,03	1010,8	0,082	12	8	0	-147,5	7,73
	P1B	4,21	173,1	27,01	1010,7	0,077	13	9	0	-188,1	2,8
	P1C	3,95	187,9	26,94	1010,8	0,081	12	9	0	-185,9	2,96
P2	P2A	4,97	129,3	27,04	1013,2	0,0261	38	27	0,02	-175,1	4,22
	P2B	4,77	141,2	27,04	1013,4	0,0246	41	29	0,02	-171	4
	P2C	4,73	143,4	27,06	1013,5	0,0244	41	29	0,02	-175	4,21
P3	P3A	5,25	113,6	27,96	1015,1	0	0	0	0	-150,6	6,92
	P3B	5,39	105,4	27,29	1015,2	0,0103	97	68	0,04	-167,9	5,13
	P3C	5,41	104,5	27,27	1015,2	0,0102	98	69	0,04	-189,6	3,97
P4	P4A	5,36	107,1	27,03	1013,6	0,0084	119	83	0,05	-170,4	1,82
	P4B	5,27	112,5	27,02	1013,7	0,0084	120	84	0,06	-166,4	1,75

	<b>P4C</b>	5,28	111,5	26,07	1013,9	0,0166	60	42	0,03	-180	2
<b>P5</b>	<b>P5A</b>	6	70,5	26,63	1010,4	0,008	126	88	0,06	-221,2	1,32
	<b>P5B</b>	5,98	71,3	26,63	1010,5	0,008	126	88	0,06	-219,2	1,54
	<b>P5C</b>	6,41	47	26,63	1011,2	0,008	125	87	0,06	-222,8	1,51

Tabela 2: Resultados dos dados físico químicos obtidos por cada ponto na cidade de Bragança-PA. Nele estão os resultados de 10 parâmetros físico químicos coletados para cada replica e ponto. Os resultados dos valores estão destacados por cor e intensidade entre vermelho, branco e verde, onde vermelho indica os maiores valores dentre os pontos, branco indica valores intermediários e verde indica os menores valores dentre os pontos (Valores maior e menor dentro das colunas).

#### 4.2 Extração do DNA metagenômico

A partir dos materiais armazenados em TSE, foi feita a extração do DNA genômico como caracterizado na metodologia (Tópico 3.3). Os materiais genéticos obtidos foram analisados em nível de concentração e integridade.

O DNA metagenômico foi extraído com o kit DNeasy PowerSoil Pro (QIAGEN), seguindo o protocolo do fabricante. A concentração de DNA foi avaliada por espectrofotometria à 260 nm e à 280 nm, usando um espectrofotômetro NanoDrop™ 1000, e os resultados obtidos pela cidade de Bragança foram mostrados na tabela 3 a seguir.

Foram usados apenas concentrações de DNA que tiveram uma concentração mínima de 50 ng/uL, com índice de pureza entre 1,4 até 2,0, para serem consideradas aptas para o sequenciamento. Para avaliar a pureza do material genético obtidos foi determinado os índices de pureza (IP) e, de acordo com a relação das absorbâncias: A260/A280, que estejam no intervalo entre 1,8 - 2,0, assim extrações que obedeceram a esses critérios foram usadas (Tabela 3).

<b>Bragança</b>			
	<b>Concentração</b>	<b>Pureza</b>	<b>Volume</b>
	<b>ng/μL</b>	<b>IP</b>	<b>μL</b>
BP1A	63,8	1,52	50
BP1B	63,8	1,52	50
BP1C	63,8	1,52	50
BP2A	63	1,57	50

BP2B	94,3	1,52	50
BP2C	63	1,85	50
BP3A	54,7	1,65	50
BP3B	81,4	1,7	50
BP3C	150	1,7	50
BP4A	71,7	2,01	50
BP4B	66	1,76	50
BP4C	64	1,74	50
BP5A	60,1	1,69	50
BP5B	81,6	1,67	50
BP5C	52,5	1,77	50

Tabela 3: Valores de concentração de DNA avaliada por espectrofotometria à 260 nm e à 280 nm, usando um espectrofotômetro NanoDrop™ 1000. Da esquerda para a direita estão os resultados de concentração, índices de pureza e volume obtido. Os valores de cada ponto (Ponto 1 ao 5) estão organizados de cima para baixo.

A integridade do DNA foi avaliada por eletroforese em gel de agarose com brometo de etídio 0,5 ug/mL, e as amostras de DNA com valores de IP no intervalo esperado foram armazenadas a -20 °C até o seu uso. Os resultados mostraram poucos rastros pela imagem de eletroforese em gel, indicando que o DNA ambiental extraído aparenta estar íntegro (Figura 7).

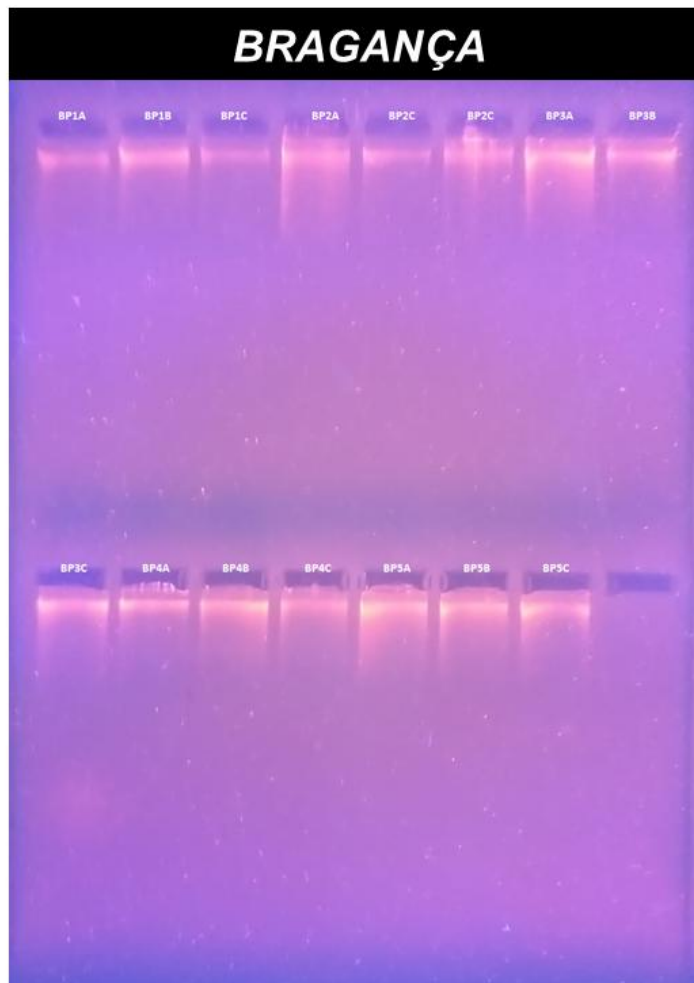


Figura 4: Técnica de eletroforese em gel dos DNAs extraídos para cada ponto do município de Bragança-PA. A eletroforese foi realizada para se comparar a integridade do material genético obtido, e ele aparenta estar íntegro pelos resultados acima.

#### 4.3 *Diversidade bacteriana e viral nos dados metagenômicos*

A diversidade alfa e beta foi calculada usando o pacote microeco, a diversidade alfa de Shannon foi usada para os grupos de bactérias e vírus analisados. Os resultados da diversidade alfa para cada amostra foram comparados com o teste não paramétrico de Wilcox, e as diferenças estatisticamente significativas entre os pontos foram rotuladas com "\*\*".

Para o grupo bacteriano, a região BP1 foi a que apresentou maiores valores de riqueza para a diversidade alfa quando comparado com as outras regiões, e a diversidade alfa BP1 mostrou-se significativamente diferente das regiões BP3, BP4 e BP5. A região BP2 mostrou-se

significativamente diferente contra BP4, e a região BP3 mostrou diferenças significativas entre os pontos BP1 e BP4. A região BP4 foi a mais significativamente diferente para todos os outros pontos neste grupo, e mostrou o menor valor de riqueza quando comparado com as outras regiões. As comparações de diversidade alfa que apresentaram um valor de p menor que 0,05 para o teste não paramétrico de Wilcoxon foram destacadas com um '\*', valores de  $p < 0,01$  foram destacados com "\*\*\*" e valores de  $p < 0,001$  com "\*\*\*\*".

Para o grupo vírus, presente no gráfico à direita, BP1 foi significativamente diferente de todos os outros 4 pontos, e teve um baixo valor de riqueza em relação aos outros pontos analisados. A região BP2 mostrou-se significativamente diferente contra BP1 e BP4, sendo BP2, BP3 e BP5 as regiões que mostraram maiores valores de riqueza para o grupo vírus. A região BP3 mostrou diferença significativa com os pontos BP4 e BP1 (Figura 5).

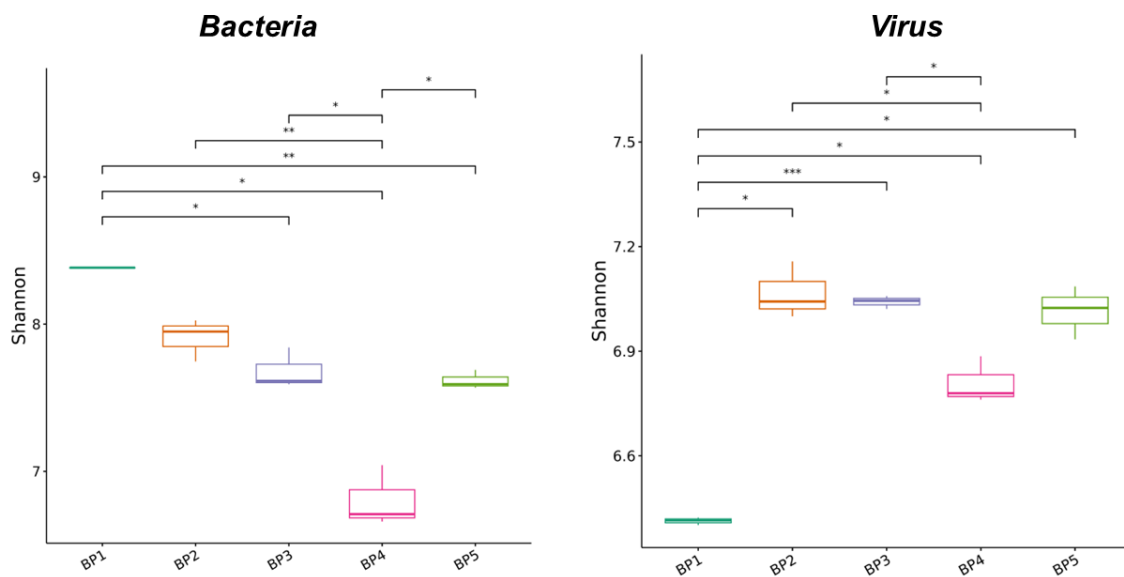


Figura 5: Box plots da diversidade de Shannon calculada para cada um destes dois grupos: Bactérias e Vírus, respectivamente da esquerda para a direita. As comparações de diversidade alfa que apresentaram um valor de p menor que 0,05 para o teste não paramétrico de Wilcoxon foram destacadas com um '\*', valores de  $p < 0,01$  foram destacados com "\*\*\*" e valores de  $p < 0,001$  com "\*\*\*\*".

Para a beta-diversidade, as análises foram feitas usando a distância de Bray-Curtis dentro e entre as regiões de cada grupo, conforme a Figura 6. Os resultados obtidos para análises de dissimilaridades entre amostras mostraram um alto valor de dissimilaridade entre a região BP1 com todas as outras regiões comparadas tanto para os grupos de bactéria quanto no grupo vírus, mostrado nos quatro primeiros box-plots da figura 6-A de cima para baixo respectivamente. Para as outras regiões, as comparações com a região BP4 também mostraram valores altos de dissimilaridade tanto para vírus quanto para bactérias, e os menores valores de dissimilaridades foram observados entre as regiões BP2 e BP3 para os dois grupos.

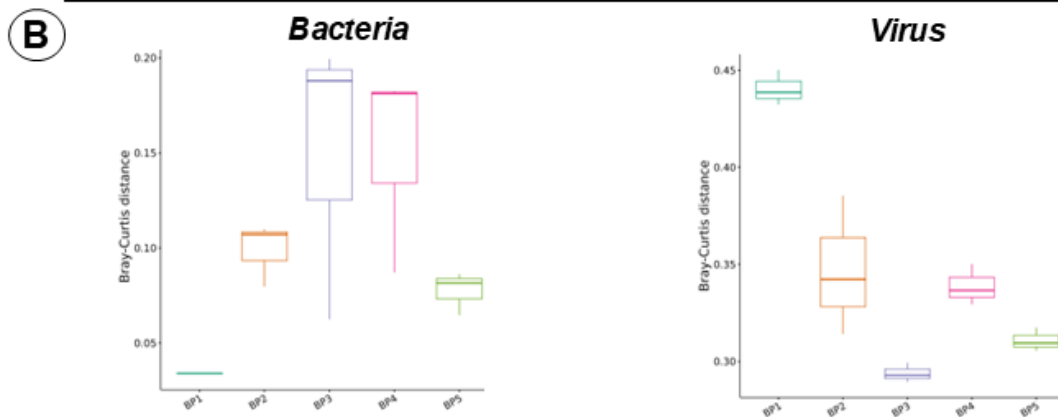
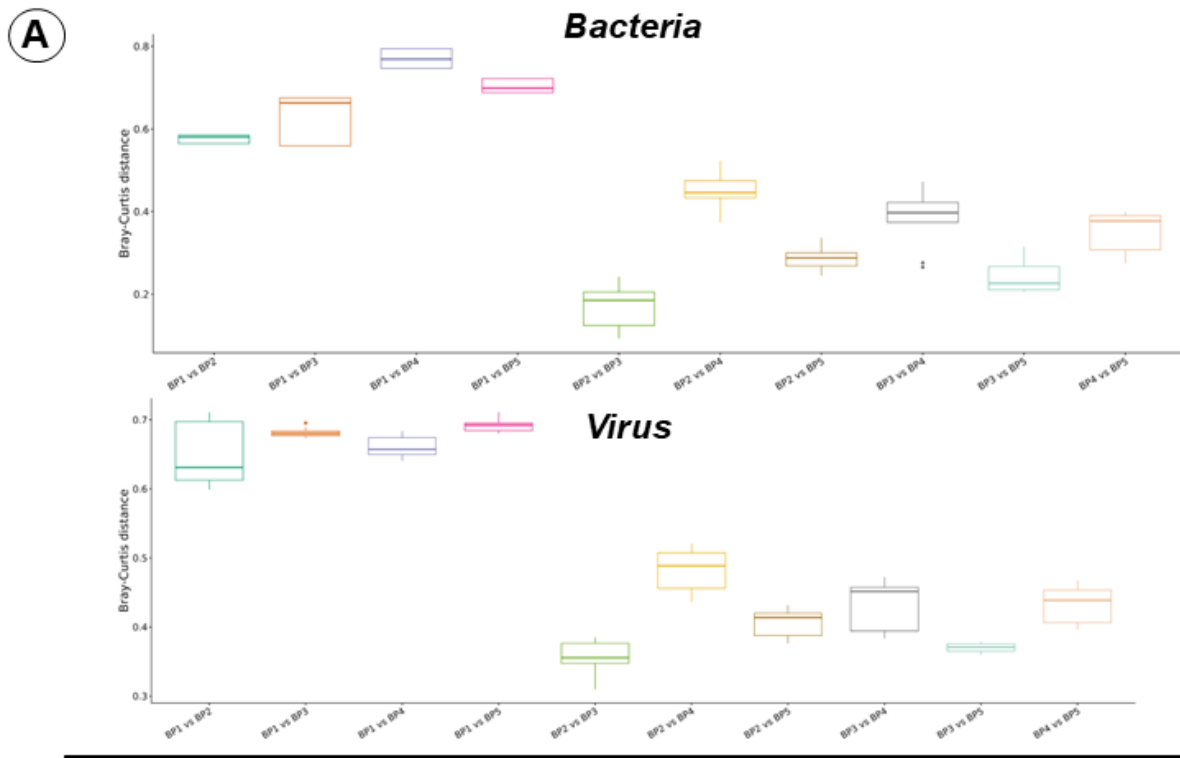


Figura 6: Distância de Bray-Curtis entre as regiões para bactérias e vírus, de cima para baixo, comparando as dissimilaridades entre e dentro das 5 regiões do rio. A - Box-plots da distância de Bray-Curtis entre pares de regiões comparando-se de cima para baixo, bactérias e vírus. B - Box-plot da distância de Bray-Curtis para mostrar as dissimilaridades dentro das 5 regiões do rio para bactérias e vírus, da esquerda para a direita, respectivamente.

Quanto às análises de beta diversidade dentro dos grupos mostrado na Figura 6-B, observou-se valores baixos de dissimilaridade no grupo Bactéria para a região BP1 seguida da região BP5, sendo as demais regiões que apresentaram maiores valores de dissimilaridade do maior para menor BP3, BP4 e BP2, em que suas medianas variaram de 0.1 a 0.2. Para o grupo Vírus, a distância de Bray-Curtis mostrou um maior valor em BP1 com mediana próximo de 0.45, enquanto nas outras regiões os maiores valores de maior para menor foram encontrados em BP2, BP4 e BP5 que tiveram suas medianas variando de 0.35 a 0.3. BP3 mostrou um menor valor de dissimilaridade em relação às outras regiões com uma mediana com valor abaixo de 0.3 (Figura 6-B).

#### 4.4 Abundância bacteriana e viral entre regiões

Para caracterizar as espécies mais comuns em cada região com base nos dados de leituras metagenômicas, um gráfico de barras empilhadas dos valores médios foi gerado de acordo com a abundância relativa de cada região em nível de filo. Os resultados mostraram que o filo *Pseudomonadota* é o mais abundante em todas as 5 regiões, com a abundância relativa aumentando ligeiramente até que este filo tenha valor mais abundante em BP4, seguido por uma diminuição nesse filo em BP5. *Actinomycetota* e *Bacillota* apresentaram uma ligeira diminuição de BP1 para BP4, com um aumento desse filo na região BP5, e *Bacteroidota* apresentou crescimento contínuo das regiões BP1 para BP5 (Figura 7).

*Pseudomonadota*, formalmente conhecido como *Proteobacteria*, é geralmente o filo mais abundante entre os ambientes naturais e pode estar associado ao impacto antrópico, devido a maioria das bactérias patogênicas estarem presentes nesse filo.

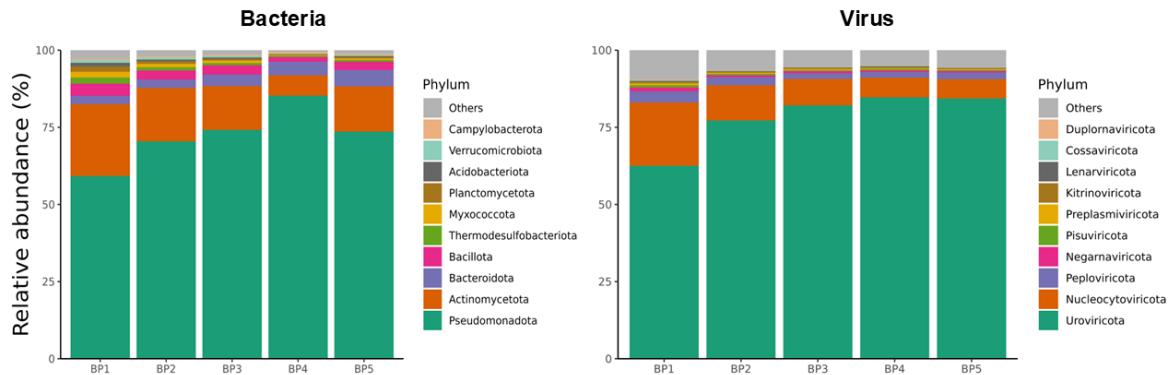


Figura 7: Gráfico de barras empilhadas com os valores médios dos filios do Grupo Bactéria e Vírus, da esquerda para a direita respectivamente. A abundância relativa foi ordenada de baixo para cima e indica a abundância relativa de cada filo em uma escala percentual em comparação com os outros vírus identificados.

Para os vírus identificados nesses dados, os *Uroviricota* foram o filo mais abundante em todas as regiões, apresentando mais de 50% da abundância relativa (%) presente em todos os vírus identificados, com sua média de abundância relativa aumentando ligeiramente do BP1 ao BP4, enquanto a média dos *Nucleocytoviricota* diminuiu ligeiramente do BP1 ao BP5 (Figura 7).

Os *Uroviricota* incluem a classe *Caudoviricetes*, conhecida por ser uma das classes de vírus mais predominantes presentes no planeta, com cerca de 96% de todos os vírus conhecidos sendo dessa classe (FOKINE; ROSSMANN, 2014), e sua abundância corresponde à alta abundância dos *Pseudomonadota*, visto que algumas de suas bactérias hospedeiras provavelmente estão presentes nesse filo (ZHOU *et al.*, 2023).

Os *Nucleocytoviricota* correspondem a vírus com genomas grandes e mais complexos, com alta diversidade de genes homólogos presentes em todos os outros grupos, o que torna sua classificação mais difícil. Eles apresentam um formato icosaédrico e são mais conhecidos pelos vírus gigantes presentes nesse filo, como alguns *Mimivírus* (XIAN; XIAO, 2020).

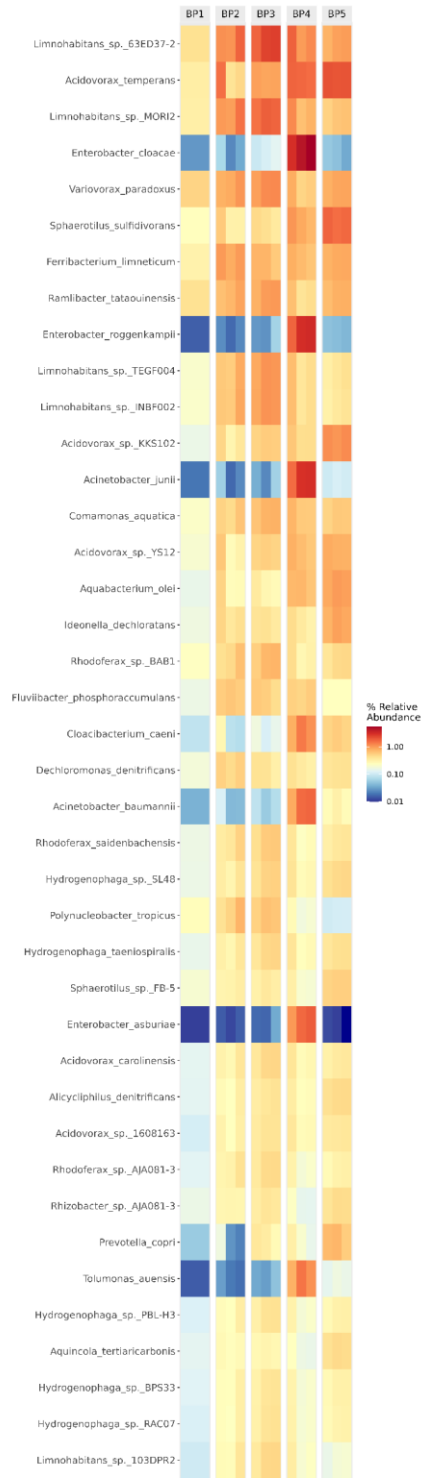
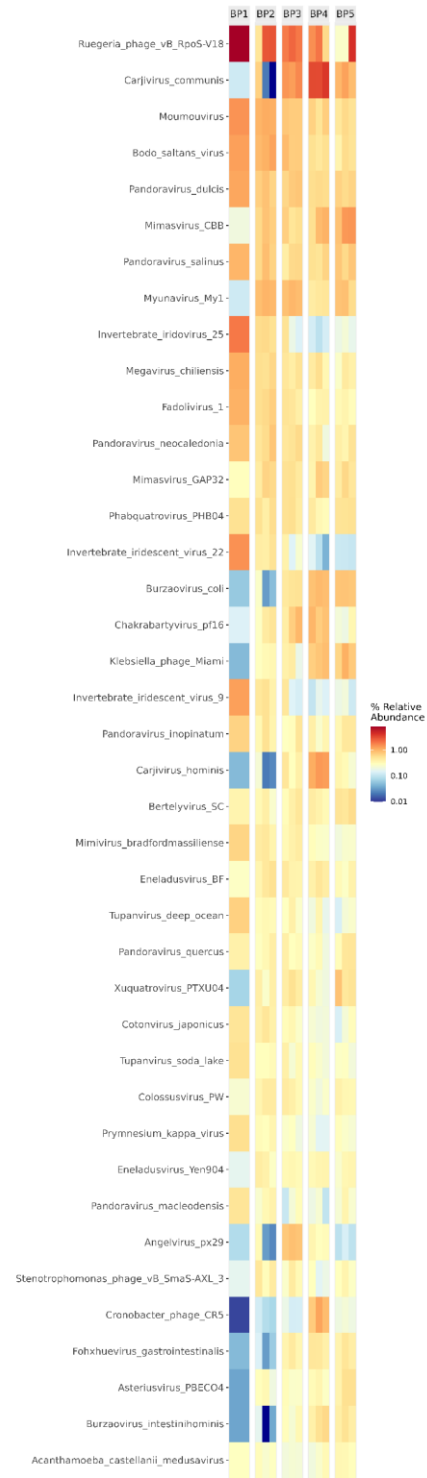
**(A)****Bacteria****(B)****Virus**

Figura 8: Mapa de calor de espécies identificadas para o grupo de bactérias (esquerda) e vírus (direita) presentes nas regiões coletadas em Bragança-PA

Para as espécies identificadas de bactérias e vírus em cada uma dessas regiões, observou-se um aumento de enterobactérias associadas a infecções patogênicas a partir do BP1 até o BP4, sendo o BP4 o mais abundante, seguido por uma ligeira diminuição de algumas destas bactérias no BP5. Algumas das espécies patogênicas identificadas com maior abundância no BP4 são *Enterobacter cloacae*, *Enterobacter roggenkampii*, *Acinetobacter junii*, *Cloacibacterium caeni* e *Enterobacter asburiae*.

A maioria dessas bactérias identificadas está associada a infecções patogênicas relacionadas ao trato urinário, pneumonia ou infecções de feridas, sendo todas consideradas bactérias oportunistas (ELBEHIRY *et al.*, 2024; GUPTA *et al.*, 2019; MOSAFFA *et al.*, 2024), com exceção do *Cloacibacterium caeni*, que é encontrado em águas residuais (CHUN *et al.*, 2017).

A partir dos vírus identificados observados na Figura 8, na BP1, foi demonstrado que os vírus conhecidos por infectar insetos foram mais abundantes na área menos impactada, a partir da nascente do rio, diminuindo ligeiramente até a BP4. Os vírus invertebrados identificados incluíram: iridovírus de invertebrados 25, vírus iridescente de invertebrados 22 e vírus iridescente de invertebrados 9.

Considerando que as bactérias patogênicas foram mais abundantes na região BP4, os bacteriófagos identificados, conhecidos por terem hospedeiros como enterobactérias e actinobactérias, incluem as bactérias mais abundantes na BP4 (Figura 8). Os bacteriófagos mais proeminentes na BP4 foram, em nível de espécie, os seguintes: *Carjivirus communis*, *Carjivirus hominis* e *Cronobacter fago CR5*. Embora não tão conhecidos quanto esses vírus, também foi observado um aumento de bacteriófagos conhecidos por terem bactérias patogênicas como hospedeiras a partir do BP2, como *Xuquatrovirus PTXU04*, *Burzaovirus coli*, *Chakrabartyvirus pf16*, *Asteriusvirus PBECO4* e *Klebsiella phage Miami*. A partir do BP3, também foi observado um aumento de bacteriófagos relacionados a bactérias patogênicas como hospedeiras, como *Angelvirus px29*, *Fohxhuevirus gastrointestinalis* e *Burzaovirus intestinhominis*. Embora não tão elevado quanto o BP4, dois bacteriófagos pareceram apresentar um ligeiro aumento de bacteriófagos de bactérias patogênicas, como *Carjivirus communis* e *Carjivirus hominis*. Curiosamente, também foi identificado em quase todas as regiões, e com uma maior abundância relativa presente em BP1, sequências que correspondiam a um bacteriófago de nome “Ruegeria

phage vB\_RpoS-V18” que tem uma bactéria heterotrófica marinha conhecida como *Ruegeria pomeroyi* como hospedeira, que é conhecida por degradar 2,3-di-hidroxiopropano-1-sulfonato (DHPS) que é produzido por um grande número de algas marinhas presentes no ambiente (COOPER; SMITH, 2015). Essa interação foi demonstrada com *Thalassiosira pseudonana* que produz o composto 2,3-di-hidroxiopropano-1-sulfonato (DHPS), alimentando-se de *Ruegeria pomeroyi*, uma vez que esta espécie possui genes relacionados com o catabolismo de DHPS, e em troca, fornece vitamina B12 para apoiar o crescimento da *Thalassiosira pseudonana* (DURHAM *et al.*, 2015).

#### 4.5 Associação de perfil taxonômico e resistência por análise de redundância (RDA)

A partir do perfil taxonômico obtido para o grupo Bactéria, e o perfil de resistência obtido pelo programa RGI para gerar uma tabela normalizada com os mecanismos de resistência como metadados foi gerado um gráfico RDA com o objetivo de identificar qual região apresentaria maior correlação com bactérias patogênicas, bem como a presença de genes de resistência identificados, e a partir disso, selecionar a região mais contaminada por ser ideal para identificar os bacteriófagos ideais para uso contra bactérias patogênicas com GRAs presentes na água (Figura 9).

Os resultados mostraram que os dados das regiões analisadas explicaram cerca de 88.89% da variação dos dados, com significância estatística  $p\text{-value} = 0.001$  para 999 permutações. A variável Grupo (que contém as regiões analisadas) foi altamente significativa ( $F = 179,03$ ;  $p = 0,001$ ) para o modelo de RDA gerado. Quanto às variáveis de mecanismos de resistência analisadas, houve uma forte correlação com valores de  $r^2 > 0.97$  e um  $p\text{-value}$  menor do que 0.001 para estes dados. Os mecanismos mais fortemente relacionados foram “efluxo de antibióticos” ( $r^2 = 0,986$ ) e “substituição de alvo antibiótico” ( $r^2 = 0,985$ ), seguidos de “alteração do alvo antibiótico”, “redução da permeabilidade dos antibióticos” e “inativação de antibióticos”, com valores de  $r^2 \geq 0,978$ . Contudo, a análise de variância revelou que apenas “alteração do alvo do antibiótico” ( $F = 36,68$ ;  $p = 0,002$ ), “efluxo de antibióticos” ( $F = 29,31$ ;  $p = 0,001$ ) e “proteção do alvo do antibiótico” ( $F = 6,21$ ;  $p = 0,043$ ) contribuíram de forma significativa e independente para a variação explicada pelo modelo (Figura 9).

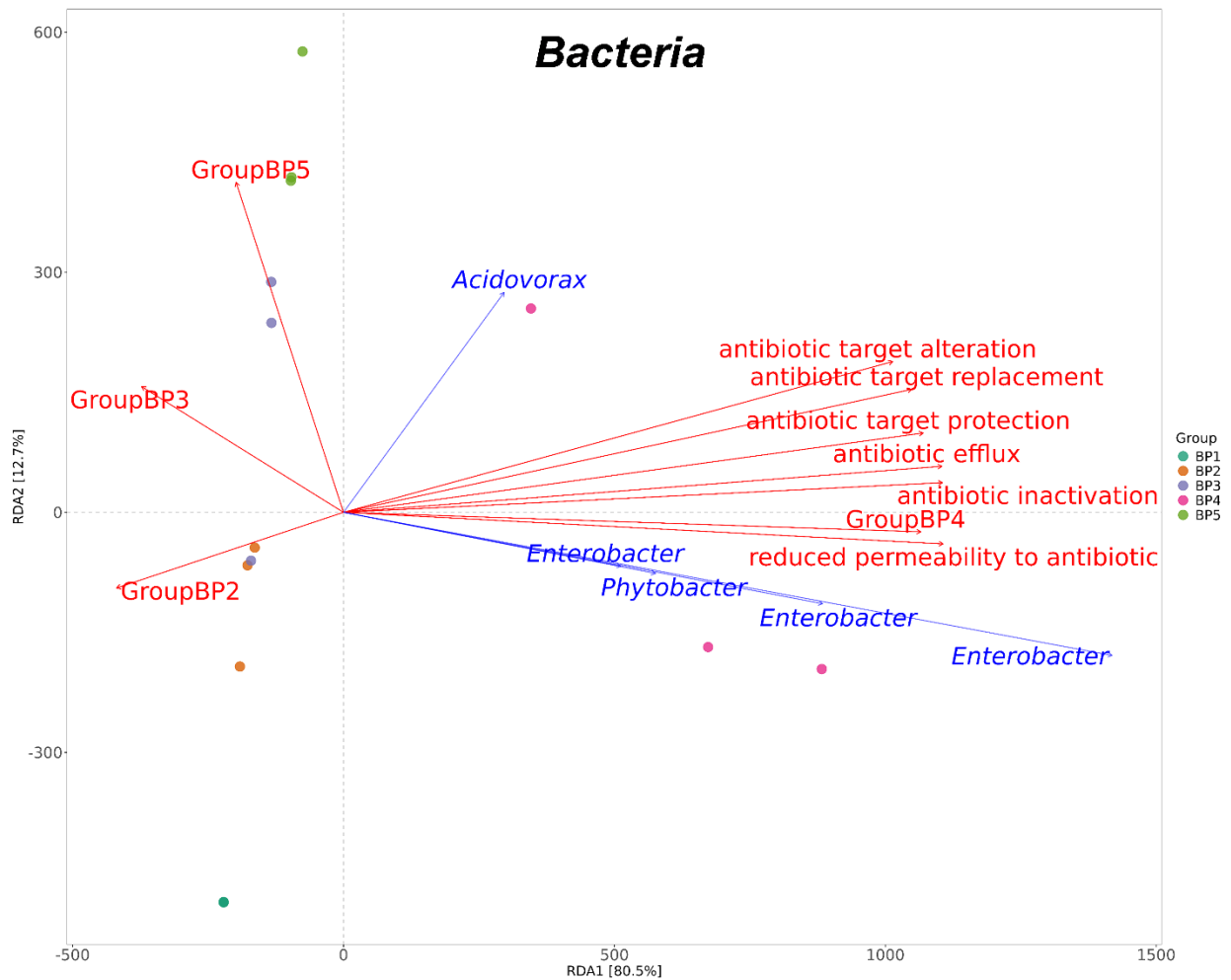


Figura 9: O gráfico de RDA ilustra a correlação entre os perfis de resistência e as regiões de coleta. Observou-se que os genes associados aos mecanismos de resistência apresentaram maior associação com a região BP4, além de correlação relevante com gêneros bacterianos como *Enterobacter*, *Phytobacter* e *Acidovorax*.

Esses resultados sugerem que a abundância dos gêneros *Enterobacter*, *Phytobacter* e *Acidovorax* apresentou correlação com a região BP4, e ela está correlacionada com perfis de mecanismos de resistência a antibióticos associados à alteração, substituição e proteção do alvo antibiótico, efluxo, inativação e redução da permeabilidade dos antibióticos. Com base nisso e no fato de que a identificação de bactérias patogênicas foi mais abundante na região BP4, essa região foi selecionada como mais apta para identificar sequências de bacteriófagos presentes nos dados metagenômicos que possam atuar no combate a bactérias patogênicas.

Análises de correlação com os dados físico-químicos obtidos também foram realizadas com os dados metagenômicos obtidos (figura não mostrada). Os dados da região explicaram cerca de 87.12% da variação dos dados, com significância estatística  $p\text{-value} = 0.001$  para 999 permutações. Houve uma certa correlação com valores de  $r^2 > 0.70$  e um  $p\text{-value}$  menor do que 0.001 para 4 dados físico-químicos que foram: pH, EC, TDS e salinidade, com valores de  $r^2$  variando de 0.8239, 0.7068, 0.7078 e 0.7199 respectivamente.

#### 4.6 Identificação de sequências de bacteriófagos de interesse a partir dos dados metagenômicos

Para a identificação das sequências virais, foram utilizados os softwares VIBRANT (KIEFT; ZHOU; ANANTHARAMAN, 2020), VirSorter2 (GUO *et al.*, 2021) e DeepVirFinder (REN *et al.*, 2020), com comprimento mínimo de sequências de 1.000 bp. As sequências identificadas foram concatenadas e as sequências redundantes removidas para então serem analisadas quanto a sua qualidade, completude e contaminação usando os softwares ViralComplete (ANTIPOV *et al.*, 2020) e CheckV (NAYFACH *et al.*, 2021), em que sequências de qualidade média ou superior, completude  $> 90\%$ , contaminação  $< 10\%$  e que não fossem provírus foram selecionadas. Após usar esses filtros foram obtidas 19, 14, 11, 11 e 6 sequências virais para os pontos BP1, BP2, BP3, BP4 e BP5 respectivamente (Tabela 4).

	Sequências Virais identificadas				
	BP1	BP2	BP3	BP4	BP5
<b>DeepVirFinder</b>	10446	17596	15475	15136	13120
<b>VIBRANT</b>	1994	2935	2090	1881	962
<b>VirSorter2</b>	7572	11986	9622	8900	6304
<b>Total</b>	14313	22920	20450	21083	17664
<b>Sequências selecionadas</b>	19	14	11	11	6

Tabela 4: Número de sequências virais identificadas por cada software por ponto seguido do número de sequências totais obtidas após a concatenação e remoção de sequências redundantes e seleção final baseada nos resultados de CheckV e ViralComplete.

Para a classificação taxonômica das sequências virais selecionadas foi utilizado a ferramenta geNomad (CAMARGO *et al.*, 2024), e nela, foi possível classificar taxonomicamente, em nível de classe, algumas sequências virais.

Para a identificação de prováveis hospedeiros das sequências virais foi utilizada a ferramenta iPHoP (ROUX *et al.*, 2023) (versão 1.4.1) utilizando o banco de dados padrão (“Jun\_2025\_pub\_rw”), e uma lista de prováveis hospedeiros foi gerada com o objetivo de comparar com os MAGs posteriormente identificados. No entanto, algumas sequências virais não puderam ter seus hospedeiros identificados pelo iPHoP, portanto, não foram consideradas para análise posterior.

#### 4.6.1 Sequências virais selecionadas em BPI

Seguindo os critérios de seleção de sequências virais a partir dos resultados do CheckV e ViralComplete, foram identificadas 19 sequências virais mostradas na Tabela 5 abaixo.

<b>contig_id</b>	<b>contig_length</b>	<b>provirus</b>	<b>checkv_quality</b>	<b>completeness ViralComplete</b>	<b>contamination</b>
NODE_52	41432	No	Medium-quality	98.1%	0
NODE_27	54305	No	Complete	126.3%	0
NODE_42	43494	No	High-quality	100.3%	0
NODE_40	43985	No	Medium-quality	90.4%	0
NODE_21	61116	No	Medium-quality	97.5%	0
NODE_18	63509	No	High-quality	147.2%	0
NODE_26	55440	No	Medium-quality	99.4%	0
NODE_11	74789	No	Medium-quality	96.1%	0
NODE_35	46590	No	High-quality	92.4%	0
NODE_69	37775	No	Medium-quality	92.7%	0
NODE_14	67391	No	Medium-quality	117.0%	0
NODE_41	43911	No	High-quality	116.7%	0
NODE_66	38444	No	Medium-quality	110.2%	0
NODE_9	75510	No	Complete	195.2%	0
NODE_77	35758	No	High-quality	108.7%	0

NODE_23	57955	No	Medium-quality	90.4%	0
NODE_64	38865	No	Complete	98.6%	0
NODE_44	42374	No	Medium-quality	101.0%	0
NODE_36	45932	No	High-quality	123.0%	0

Tabela 5: Sequências virais selecionadas no ponto BP1 seguindo os critérios de seleção descritos acima e no tópico 3.4.7.

Com estas 19 sequências selecionadas, usou-se o software iPHoP para a identificação de prováveis bactérias hospedeiras destas sequências, no entanto, somente o “NODE\_36” foi identificado, gerando os resultados disponibilizados na Tabela 6 abaixo.

Virus	Host genus	Confidence score
NODE_36	f_Bryobacteraceae;g_JADGNR01	90.30
	f_Bryobacteraceae;g_JAIQFL01	90.30

Tabela 6: Sequências virais em que o seu hospedeiro foi identificado usando o software iPHoP para o ponto BP1.

Esta provável sequência viral mostrada na Tabela 6 aparenta ser um bacteriófago de bactérias da família *Bryobacteraceae* que é uma família relacionada a bactérias responsáveis na fixação de nitrogênio para o meio ambiente, além de ser uma família pouco estudada, mas conhecida por estar presente no ciclo biogeoquímico de ambientes poucos impactados (GONÇALVES *et al.*, 2024). Sendo assim, um indicador do pouco impacto ambiental presente na região.

A nível de classificação taxonômica, todos os contigs virais foram classificados como pertencentes à classe *Caudoviricetes*, sendo o NODE\_42 classificado como ordem *Fredfastierviridae* além de pertencer à classe *Caudoviricetes*.

Como a região BP1 foi o ponto de menor contaminação por atividades antropogênicas analisado, e observou-se um baixo valor de riqueza para o ponto BP1 (Figura 5) em comparação com outras regiões, no entanto, esta região também foi a que mais foi identificada sequências de vírus de qualidade. Este resultado pode ter ocorrido devido às seguintes possibilidades. 1- A baixa riqueza presente nos dados metagenômicos terem oferecidos menos espécies para serem analisadas

e assim ter sido possível obter mais sequências virais que aumentaram a qualidade durante a montagem para este domínio; 2 – Os resultados de diversidade podem ter identificados poucas sequências virais devido ao seu banco de dados do kraken2 ter poucas sequências que se identificaram com os dados metagenômicos, daí o baixo valor de riqueza obtido, já para a identificação de sequências virais é possível que durante a etapa de montagem tenha gerado contigs virais a partir de leituras virais de matéria escura (*Viral Dark Matter*) (SANTIAGO-RODRIGUEZ; HOLLISTER, 2022) que não foram identificados pelo banco de dados do kraken2. Estes contigs devem ter facilitado a identificação de sequências virais novas pelos programas usados, resultando nas 19 sequências selecionadas.

#### 4.6.2 Sequências virais selecionadas em BP2

Seguindo os critérios de seleção de sequências virais a partir dos resultados do CheckV e ViralComplete, foram identificadas 14 sequências virais mostradas na Tabela 7 abaixo.

contig_id	contig_length	provirus	checkv_quality	completeness ViralComplete	Contamination
NODE_206	36025	No	Medium-quality	79,85	0
NODE_176	39051	No	Complete	100	0
NODE_161	41015	No	High-quality	100	0
NODE_3298	6495	No	Complete	100	0
NODE_147	43536	No	High-quality	98,36	0
NODE_153	42391	No	Complete	100	0
NODE_165	40344	No	High-quality	100	0
NODE_2	38091	No	High-quality	99,5	0
NODE_117	50218	No	Complete	100	0
NODE_61	74846	No	Medium-quality	86,64	0
NODE_190	37654	No	Medium-quality	57,07	0
NODE_212	35156	No	Complete	100	0
NODE_173	39325	No	Medium-quality	65,54	0
NODE_85	61424	No	High-quality	100	0

Tabela 7: Sequências virais selecionadas no ponto BP2 seguindo os critérios de seleção descritos no tópico 3.4.7.

Usando o software iPHoP somente o “NODE\_161” foi identificado, gerando os resultados disponibilizados na Tabela 8 abaixo.

<b>Virus</b>	<b>Host genus</b>	<b>Confidence score</b>
NODE_161	f_Burkholderiaceae;g_JAEWVY01	95.70
	f_Burkholderiaceae;g_Limnohabitans_A	91.60

Tabela 8: Sequências virais em que o seu hospedeiro foi identificado usando o software iPHoP para o ponto BP2.

A nível de classificação taxonômica, quase todos os contigs virais foram classificados como pertencentes à classe *Caudoviricetes*, sendo o NODE\_153 e o NODE\_147 classificados como pertencentes à ordem *Fredfastierviridae* e à classe *Caudoviricetes*. Sendo o NODE\_3298 o único que foi classificado como pertencente à família *Tombusviridae*, que se trata de um vírus de RNA de plantas (LOZIER *et al.*, 2024). Como estes dados se tratam de dados de metagenoma, a origem desta classificação pode ser devido a um falso positivo ou a algum outro erro de classificação desta sequência viral.

Dentre as 14 sequências virais identificadas como possíveis bacteriófagos no ponto BP2, apenas uma pode ter seu hospedeiro identificado usando uma pontuação de confiança mínimo de 90%. Esta sequência viral presente na tabela 8 aparenta ter uma sequência viral de um bacteriófago com hospedeiro de família *Burkholderiaceae*, sendo esta família associada com supressão de crescimento de fungos e promover o crescimento de plantas (CARRIÓN *et al.*, 2018). Quanto ao gênero *Limnohabitans*, que seria um dos prováveis hospedeiros do NODE\_161, é um gênero normalmente encontrados em regiões de águas doces e seu crescimento está associado com a proliferação de algas, e em regiões com baixa presença deste gênero se correlaciona com a proliferação de cianobactérias (JEZBEROVÁ *et al.*, 2017).

#### 4.6.3 Sequências virais selecionadas em BP3

Seguindo os critérios de seleção de sequências virais a partir dos resultados do CheckV e ViralComplete, foram identificadas 11 sequências virais mostradas na Tabela 9 abaixo.

contig_id	contig_length	provirus	checkv_quality	completeness	contamination
NODE_169	44727	No	Complete	100	0
NODE_83	62770	No	High-quality	100	0
NODE_187	40945	No	Complete	100	0
NODE_30121	2277	No	Medium-quality	54,72	0
NODE_181	42392	No	Complete	100	0
NODE_205	38938	No	Medium-quality	84,45	0
NODE_192	40344	No	High-quality	100	0
NODE_1	34368	No	Medium-quality	89,52	0
NODE_288	31828	No	Medium-quality	84,93	0
NODE_116	54370	No	High-quality	100	0
NODE_199	39331	No	Medium-quality	65,55	0

Tabela 9: Sequências virais selecionadas no ponto BP3 seguindo os critérios de seleção descritos no tópico 3.4.7.

A nível de classificação taxonômica, todos os contigs virais foram classificados como pertencentes à classe *Caudoviricetes*, sendo o NODE\_181 classificado como ordem *Fredfastierviridae* além de pertencer à classe *Caudoviricetes*.

Com estas 11 sequências selecionadas, usou-se o software iPHoP para a identificação de prováveis bactérias hospedeiras destas sequências, gerando os resultados disponibilizados na Tabela 10 abaixo.

Virus	Host genus	Confidence score
NODE_187	f_Burkholderiaceae;g_JAEWVY01	95.70
	f_Burkholderiaceae;g_Limnohabitans_A	90.70
	f_Burkholderiaceae;g_Melaminivora	90.30
	f_Burkholderiaceae;g_Rhodoferax_A	90.30
NODE_83	f_Burkholderiaceae;g_Limnohabitans_A	95.00

Tabela 10: Sequências virais em que o seu hospedeiro foi identificado usando o software iPHoP para o ponto BP3.

As duas sequências virais encontradas aparentam ser bacteriófagos de bactérias da família *Burkholderiaceae*, descritas e também encontradas na região BP2, além de apresentar prováveis hospedeiros o gênero *Limnohabitans* para as duas sequências virais e que também foram descritas no tópico anterior.

#### 4.6.4 Sequências virais selecionadas em BP4

Seguindo os critérios de seleção de sequências virais a partir dos resultados do CheckV e ViralComplete, foram identificadas 11 sequências virais mostradas na Tabela 11 abaixo.

contig_id	contig_length	provirus	checkv_quality	completeness	Contamination
NODE_162	47079	No	Medium-quality	118.0%	0
NODE_294	36883	No	High-quality	98.5%	0
NODE_177	45190	No	High-quality	100.4%	0
NODE_173	45349	No	Medium-quality	91.5%	0
NODE_481	30446	No	Medium-quality	90.2%	0
NODE_9535	6431	No	High-quality	110.3%	0
NODE_4769	9526	No	High-quality	133.8%	0
NODE_190	44288	No	High-quality	108.7%	0
NODE_382	33205	No	Complete	99.1%	0
NODE_1	34368	No	Medium-quality	113.3%	0
NODE_249	39316	No	Medium-quality	94.6%	0

Tabela 11: Sequências virais selecionadas no ponto BP4 seguindo os critérios de seleção descritos no tópico 3.4.7.

A nível de classificação taxonômica, quase todos os contigs virais foram classificados como pertencentes à classe *Caudoviricetes*, sendo o NODE\_4769 e o NODE\_9535 classificados como pertencentes à família *Inoviridae*, sendo esta uma família de bacteriófagos não envelopados de fita simples de DNA (KNEZEVIC; ADRIAENSSENS; ICTV REPORT CONSORTIUM, 2021; SZÉKELY; BREITBART, 2016).

Com estas 11 sequências selecionadas, usou-se o software iPHoP para a identificação de prováveis bactérias hospedeiras destas sequências, gerando os resultados disponibilizados na Tabela 12 abaixo.

Virus	Host genus	Confidence score
NODE_162	f_Burkholderiaceae;g_Limnohabitans_A	90.10
NODE_173	f_Enterobacteriaceae;g_Enterobacter	96.40
NODE_177	f_Bacteroidaceae;g_Bacteroides	99.60
	f_Bacteroidaceae;g_Phocaeicola	98.70
NODE_190	f_Enterobacteriaceae;g_Enterobacter	93.80
NODE_294	f_Burkholderiaceae;g_JAHFQI01	94.90
	f_Burkholderiaceae;g_Malikia	90.20
NODE_382	f_Enterobacteriaceae;g_Enterobacter	94.30
	f_Enterobacteriaceae;g_Citrobacter	90.70
NODE_4769	f_Moraxellaceae;g_Acinetobacter	98.20
NODE_481	f_Moraxellaceae;g_Acinetobacter	100.00
NODE_9535	f_Moraxellaceae;g_Acinetobacter	94.50

Tabela 12: Sequências virais em que o seu hospedeiro foi identificado usando o software iPHoP para o ponto BP4.

Das sequências virais analisadas, alguns dos gêneros hospedeiros identificados foi *Limnohabitans* para a sequência NODE\_162. O NODE\_173 e o NODE\_190 parece ter o gênero *Enterobacter* como o hospedeiro mais provável, e este gênero é um grupo patogênico conhecido que pertence ao grupo de bactérias ESKAPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* e *Enterobacter spp*), conhecido por possuir um conjunto de características que lhe permitem apresentar resistência a múltiplos fármacos e ser um grupo de bactérias patogênicas para humanos e outros animais (MILLER; ARIAS, 2024).

O NODE\_177 aparenta ser um bacteriófago dos gêneros *Bacteroides* e *Phocaeicola*, que são conhecidos por estarem normalmente presentes na microbiota intestinal (DA SILVA MORAIS *et al.*, 2024; ZAFAR; SAIER, 2021).

O NODE\_382 apresentou os gêneros *Citrobacter* e *Enterobacter* como prováveis hospedeiros sendo o primeiro associado com infecções oportunistas normalmente encontrados em recém nascidos e o segundo está presente no grupo ESKAPE (JABEEN *et al.*, 2023; MILLER;

ARIAS, 2024). O NODE\_294 apresentou o gênero *Malikia* e JAHFQI01 como prováveis hospedeiros, sendo *Burkholderiaceae* a família comum destes gêneros identificados, sendo elas relacionadas a atividades antifúngicas capazes de proteger o crescimento de plantas que convivam com este grupo de bactérias (CARRIÓN *et al.*, 2018).

As outras 3 sequências virais (NODE\_4769, NODE\_481 e NODE\_9535) demonstraram ser prováveis bacteriófagos do gênero *Acinetobacter*, gênero relacionado a uma ampla gama de infecções nosocomiais e resistência a múltiplos medicamentos; além disso, uma espécie deste gênero inclui o grupo ESKAPE *A. baumannii* (CARVALHEIRA; SILVA; TEIXEIRA, 2021). Além disso, o iPHoP não identificou o hospedeiro bacteriano de 2 sequências virais (NODE\_1 e NODE\_249).

#### 4.6.5 Sequências virais selecionadas em BP5

Seguindo os critérios de seleção de sequências virais a partir dos resultados do CheckV e ViralComplete, foram identificadas 6 sequências virais mostradas na Tabela 13 abaixo.

contig_id	contig_length	provirus	checkv_quality	completeness	contamination
NODE_1	370461	No	High-quality	98.09	0
NODE_11	130998	No	Medium-quality	51.53	0
NODE_118	37038	No	High-quality	96.48	0
NODE_120	36831	No	High-quality	94.94	0
NODE_125	36315	No	Complete	100	0
NODE_69	48594	No	Medium-quality	81.35	0

Tabela 13: Sequências virais selecionadas no ponto BP5 seguindo os critérios de seleção descritos no tópico 3.4.7.

A nível de classificação taxonômica, todos os contigs virais identificados neste ponto foram classificados como pertencentes à classe *Caudoviricetes*.

Com estas 6 sequências selecionadas, usou-se o software iPHoP para a identificação de prováveis bactérias hospedeiras destas sequências, gerando os resultados disponibilizados na Tabela 14 abaixo.

Virus	Host genus	Confidence score
NODE 120	f Burkholderiaceae;g Malikia	90.20
NODE 1	f Rhodanobacteraceae;g Luteibacter	91.10
NODE 69	f Burkholderiaceae;g Limnohabitans A	93.60

Tabela 14: Sequências virais em que o seu hospedeiro foi identificado usando o software iPHoP para o ponto BP5.

Dentre as 3 sequências, duas apresentam a família *Burkholderiaceae* que está associada com atividade antifúngica e consequente proteção para o crescimento de plantas e o gênero *Limnohabitans* para o NODE\_69 que seu crescimento está associado a proliferação de algas e a sua baixa presença está associada à proliferação de cianobactérias (CARRIÓN *et al.*, 2018; JEZBEROVÁ *et al.*, 2017). Sendo esta família e gênero de hospedeiros também encontrados na região BP2, BP3 e BP4 dos pontos analisados.

O NODE\_1 apresentou a família *Rhodanobacteraceae* e o gênero *Luteibacter* como provável hospedeiro desta sequência viral, sendo este gênero associado com bactérias que crescem em raízes de plantas e produtoras de substâncias que ajudem a proteger de patógenos (AHMED *et al.*, 2025; JOHANSEN *et al.*, 2005).

#### 4.7 Geração de MAGs de interesse a partir dos dados metagenômicos

##### 4.7.1 Geração de MAGs e classificação taxonômica

Para gerar os Genomas Metagenômicos Montados (MAGs), utilizou-se o software COMEBin (WANG *et al.*, 2024b) para cada região analisada separadamente seguidas de uma análise de completude e contaminação usando o software CheckM2 (CHKLOVSKI *et al.*, 2023), com isso, os MAGs que obtiveram um valor mínimo de completude de 50% e um valor máximo de contaminação de 10% foram selecionados. Estes MAGs foram classificados taxonomicamente usando o software kMetaShot (DEFAZIO *et al.*, 2024), sendo que os MAGs classificados foram analisados quanto à possibilidade por inferência de serem prováveis hospedeiros das sequências virais selecionadas e analisadas em cada ponto. Estes hospedeiros e sequências virais que foram

selecionados por inferência foram analisados quanto à presença de genes de resistência, presença de sistemas CRISPR-Cas e quanto à presença de genes associados à infecção por fagos.

#### 4.7.2 *Geração de MAGs no ponto BP1*

Para o ponto BP1 usando o software COMEBin cerca de 106 MAGs foram geradas. Destas, foi selecionada apenas uma MAG que obteve os padrões de seleção declarados (Tabela 15).

<b>Bin</b>	<b>Completeness</b>	<b>Contamination</b>	<b>organism_name</b>
117.fa	99,91	1,2	<i>Sphingopyxis soli</i>

Tabela 15: MAG selecionada a partir dos critérios de contaminação, completude e classificação taxonômica usando os dados da região BP1

Apesar de ter tido uma MAG selecionada que obedeceram aos critérios estabelecidos, não foi possível estabelecer uma inferência bactéria-hospedeiro com os dados obtidos para o ponto BP1, pois o grupo taxonômico obtido a partir da MAG selecionada não corresponde com os prováveis hospedeiros das sequências virais selecionadas.

#### 4.7.3 *Geração de MAGs no ponto BP2*

Para o ponto BP2 usando o software COMEBin cerca de 185 MAGs foram geradas. No entanto, não foi possível identificar MAGs que obedeceram aos critérios estabelecidos. Assim, nenhuma MAG foi selecionada e o ponto BP2 também não foi selecionado para realizar análise de relação bactéria-hospedeiro por não ter sido identificada relação de bactérias e vírus para tal.

#### 4.7.4 *Geração de MAGs no ponto BP3*

Para o ponto BP3 usando o software COMEBin cerca de 286 MAGs foram geradas. Destas, apenas duas MAGs que obtiveram os padrões de seleção declarados (Tabela 16).

<b>Bin</b>	<b>Completeness</b>	<b>Contamination</b>	<b>organism_name</b>
------------	---------------------	----------------------	----------------------

59084.fa	56,07	1,65	<i>Prevotella copri</i>
60784.fa	50,12	2,29	<i>Cloacibacterium normanense</i>

Tabela 16: MAGs selecionadas a partir dos critérios de contaminação, completude e classificação taxonômica usando os dados da região BP3

Apesar de ter tido duas MAGs selecionadas que obedeceram aos critérios estabelecidos, não foi possível estabelecer uma inferência bactéria-hospedeiro com os dados obtidos para o ponto BP3, pois o grupo taxonômico obtido a partir das MAGs selecionadas não correspondeu aos prováveis hospedeiros das sequências virais selecionadas na mesma região.

#### 4.7.5 Geração de MAGs no ponto BP4

Para o ponto BP4 usando o software COMEBin cerca de 342 MAGs foram geradas. Destas foram selecionadas 26 MAGs que obtiveram os padrões de seleção de completude >50%, contaminação <10% e que foram classificadas taxonomicamente por kMetaShot (Tabela 17).

Bin	Completeness	Contamination	organism_name
56583	66,74	1,43	<i>Acidovorax caeni</i>
41845	56,37	5,01	<i>Acidovorax</i> sp. 210-6
53866	52,44	0,56	<i>Acidovorax</i> sp. IB03
70207	61,88	5,08	<i>Acinetobacter baumannii</i> 532413
13504	80,75	2,99	<i>Acinetobacter brisouii</i>
9515	79,2	1,2	<i>Acinetobacter junii</i> NIPH 182
1227	91,93	2,59	<i>Acinetobacter soli</i>
49724	56,94	3,14	<i>Acinetobacter</i> sp. ME22
26282	71,23	5,13	<i>Acinetobacter</i> sp. WCHA29
11970	77,18	9,71	<i>Acinetobacter tandoii</i>
2166	90,17	2,54	<i>Acinetobacter ursingii</i> DSM 16037
16394	68,08	6	<i>Aeromonas hydrophila</i> 187
2339	84,34	7,32	<i>Aeromonas hydrophila</i> pc104A
90932	54,23	6,9	<i>Aquitalea magnusonii</i>
21009	63,69	5,08	<i>Aquitalea</i> sp. ASV15
4843	84,63	6,6	<i>Citrobacter freundii</i> ATCC 8090
58820	50,62	2,71	<i>Cloacibacterium normanense</i>
68074	64,93	4,97	<i>Dechloromonas agitata</i> is5
6313	74,28	5,44	<i>Enterobacter</i> sp. RHBSTW-00175
80424	59,84	0,95	<i>Epilithonimonas hominis</i>

33139	68,6	5,15	Kluyvera georgiana ATCC 51603
45173	98,39	0,22	Moraxella osloensis
101627	57,95	2,69	Prevotella copri
42954	81,37	1,7	Rivicola pingtungensis
64428	68,94	2,35	Vogesella urethralis
9992	92,13	0,59	Empedobacter falsenii

Tabela 17: MAG selecionada a partir dos critérios de contaminação, completude e classificação taxonômica usando os dados da região BP4

Dos 26 MAGs selecionados após o processo de “binning”, 10 MAGs com probabilidade de serem hospedeiros bacterianos das sequências virais identificadas foram selecionados para análises posteriores.

A partir destes resultados foi possível estabelecer uma relação bactéria-hospedeiro por inferência para 10 MAGs bacterianas que se apresentaram como prováveis hospedeiras de 6 sequências virais identificadas no ponto BP4 (Tabela 18). Estas MAGs foram posteriormente analisadas quanto a presença de genes de resistência usando os softwares RGI, presença de elementos associados a defesa contra bacteriófagos usando o software DefenseFinder (TESSON *et al.*, 2022) que possam conferir resistência a fagos.

<b>Contig_id (virus)</b>	<b>Host genus</b>	<b>MAG (Bacteria)</b>	<b>organism_name</b>
NODE_382	<i>g__Citrobacter</i>	4843.fasta	<i>Citrobacter freundii</i> ATCC 8090
NODE_4769	<i>g__Acinetobacter</i>	70207.fasta	<i>Acinetobacter baumannii</i> 532413
		13504.fasta	<i>Acinetobacter brisouii</i>
		9515.fasta	<i>Acinetobacter junii</i> NIPH 182
NODE_481		1227.fasta	<i>Acinetobacter soli</i>
		49724.fasta	<i>Acinetobacter</i> sp. ME22
		11970.fasta	<i>Acinetobacter tandoii</i>
NODE_9535		26282.fasta	<i>Acinetobacter</i> sp. WCHA29
		2166.fasta	<i>Acinetobacter ursingii</i> DSM 16037
NODE_173		<i>g__Enterobacter</i>	6313.fasta
NODE_190			

NODE_382			
----------	--	--	--

Tabela 18: Comparação das sequências virais identificadas que provavelmente tem como hospedeiros os MAGs selecionados a partir da tabela 17. A identificação taxonômica desses MAGs selecionados mostra que, dentre as 11 sequências virais selecionadas, 6 tiveram seus prováveis hospedeiros também identificados durante as montagens dos MAGs.

A partir desses resultados, o bacteriófago NODE\_382 teve como provável hospedeiro o gênero *Citrobacter* e o gênero *Enterobacter* e seus MAGs com táxons correspondentes para hospedeiros foram *Citrobacter freundii* (4843.fasta) e *Enterobacter* sp. (6313.fasta) respectivamente. O grupo *Citrobacter* está associado com infecções hospitalares por ter um comportamento de patógeno oportunista, além de ter crescido o número de ocorrências de aumento de resistência a antibióticos para este gênero (JABEEN *et al.*, 2023). Já o gênero *Enterobacter* é um conhecido patógeno pertencente ao grupo ESKAPE, notadamente conhecido por apresentar genes de multirresistência a antibióticos.

Para as 3 sequências virais identificadas (NODE\_4769, NODE\_481 e NODE\_9535) como tendo o gênero *Acinetobacter* como provável hospedeiro, 7 MAGs foram taxonomicamente classificadas como *Acinetobacter*, e, dentre elas, foi encontrado um MAG pertencente ao grupo ESKAPE (*Acinetobacter baumannii* 532413 – 70207.fasta), e outros 2 com potencial patogenicidade (*Acinetobacter junii* – 9515.fasta e *Acinetobacter* sp. ME22 – 49724.fasta) (AGUILAR-VERA *et al.*, 2024; CARVALHEIRA; SILVA; TEIXEIRA, 2021).

Para o bacteriófago *Enterobacter*, das 3 sequências virais identificadas que têm o *Enterobacter* como provável hospedeiro (NODE\_382, NODE\_173 e NODE\_190) correspondeu-se a um MAG taxonomicamente identificado como *Enterobacter* (*Enterobacter* sp. RHBSTW-00175 – 6313.fasta), sendo este gênero altamente relacionado com patogenicidade por estar presente no grupo ESKAPE e conhecido por apresentar multirresistência a antibióticos (MILLER; ARIAS, 2024).

Para o ponto BP5 usando o software COMEBin cerca de 358 MAGs foram geradas. No entanto, apenas duas MAGs que obtiveram os padrões de seleção definidos foram selecionadas (Tabela 19).

Bin	Completeness	Contamination	organism_name
88625.fa	52,46	2,17	Vogesella urethralis
75364.fa	52,72	4,3	Flavobacterium stagni
68021.fa	53,22	1,99	Rivicola pingtungensis
91515.fa	53,89	1,36	Chroococidiopsis sp. CCALA 051
24675.fa	59,39	2,26	Acidovorax caeni
23008.fa	64,79	6,86	Acinetobacter haemolyticus
30797.fa	71,44	3,93	Cloacibacterium normanense
17738.fa	77,49	0,88	Prevotella copri
5041.fa	83,05	1,71	Acinetobacter sp. WCHA29

Tabela 19: MAG selecionada a partir dos critérios de contaminação, completude e classificação taxonômica usando os dados da região BP5

Apesar de ter tido duas MAGs selecionadas que obedeceram aos critérios estabelecidos, não foi possível estabelecer uma inferência bactéria-hospedeiro com os dados obtidos para o ponto BP5, pois o grupo taxonômico obtido a partir das MAGs selecionadas não correspondeu aos prováveis hospedeiros das sequências virais selecionadas na mesma região.

#### 4.8 Análises de perfil de resistência entre a relação bacteriófago-hospedeiro da região BP4

Além dos MAGs que corresponderam por inferência aos hospedeiros identificados pelas sequências virais e este critério só ter sido possível ser feito na região BP4, ela foi escolhida como o ponto de maior potencial de se identificar bacteriófagos de interesse que possam atuar em bactérias patogênicas. Esta região também foi a mesma em que bactérias patogênicas foram identificadas durante as análises de taxonomia (Figura 8) e também foi a região em que mais foram identificadas bactérias com perfil de resistência a antibióticos e mecanismos de resistência através do plot de RDA (Figura 9). Dando ênfase ao gênero *Enterobacter* que correlacionou ao perfil de resistência identificado na região BP4 por este gênero também ter sido identificado durante a geração de MAGs e quanto a identificação de 3 sequências de bacteriófagos que tenham este gênero como hospedeiro.

Embora os bacteriófagos potencializem a eficácia do tratamento quando combinados com antibióticos, eles também podem oferecer resistência antimicrobiana a outras bactérias por meio da Transferência Horizontal de Genes (THG) (OLAWADE *et al.*, 2024; ROMERO-CALLE *et al.*, 2019). Para verificar se as sequências virais ou os hospedeiros de MAGs inferidos seriam portadores de genes de resistência antimicrobiana, foram pesquisadas as sequências com RGI em ambos os casos e foram plotados os resultados caso alguma resistência antimicrobiana fosse identificada.

Ao selecionar os MAGs presentes na tabela 18 como hospedeiros inferidos das sequências virais, pesquisamos os genes de resistência antimicrobiana presentes neles para verificar se esses prováveis hospedeiros também poderiam ser portadores de genes de multirresistência a antibióticos.

Quando as sequências virais foram analisadas com RGI, nenhum gene de resistência foi encontrado, mas para a maioria dos hospedeiros virais inferidos, alguns genes associados a mecanismos de resistência para o perfil de AMR foram encontrados (Figura 10).

A partir dos resultados mostrados na figura 10, a resistência a múltiplos fármacos foi encontrada em *Citrobacter freundii*, *Enterobacter sp.*, *Acinetobacter ursingi*, *Acinetobacter tandoi*, *Acinetobacter junii* e *Acinetobacter balmani*. A *Enterobacter sp.* Junto a *Citrobacter freundii* apresentou a maior presença de genes de mecanismo de resistência, em que foram identificados mecanismos relacionados à resistência à permeabilidade reduzida a antibióticos, alteração do alvo antibiótico, inativação de antibióticos e efluxo de antibióticos.

*Acinetobacter balmani*, que pertence ao grupo ESKAPE, apresentou mecanismos de resistência de alteração do alvo antibiótico, inativação de antibióticos e efluxo de antibióticos. *Acinetobacter tandoi* e *Acinetobacter junii* apresentaram alteração do alvo antibiótico e efluxo antibiótico; *Acinetobacter ursingi* apresentou inativação e efluxo antibiótico.

Alguns outros MAGs identificaram apenas um mecanismo de resistência, que foi *Acinetobacter sp. ME22* com inativação antibiótica; e *Acinetobacter soli* e *Acinetobacter brisouii* que apresentaram efluxo antibiótico como mecanismo de resistência. E em *Acinetobacter sp. WCHA29* nenhum mecanismo de resistência foi identificado neste MAG.

Curiosamente, os MAGs com maior probabilidade de apresentar multirresistência à RAM identificados que foram *Citrobacter freundii*, *Enterobacter sp.* também foram considerados pela literatura como tendo resistência a múltiplos fármacos, por serem gêneros presentes no grupo

ESKAPE ou por serem patógenos emergentes, representando desafios terapêuticos para eles (AGUILAR-VERA *et al.*, 2024; MILLER; ARIAS, 2024; WONG *et al.*, 2017).

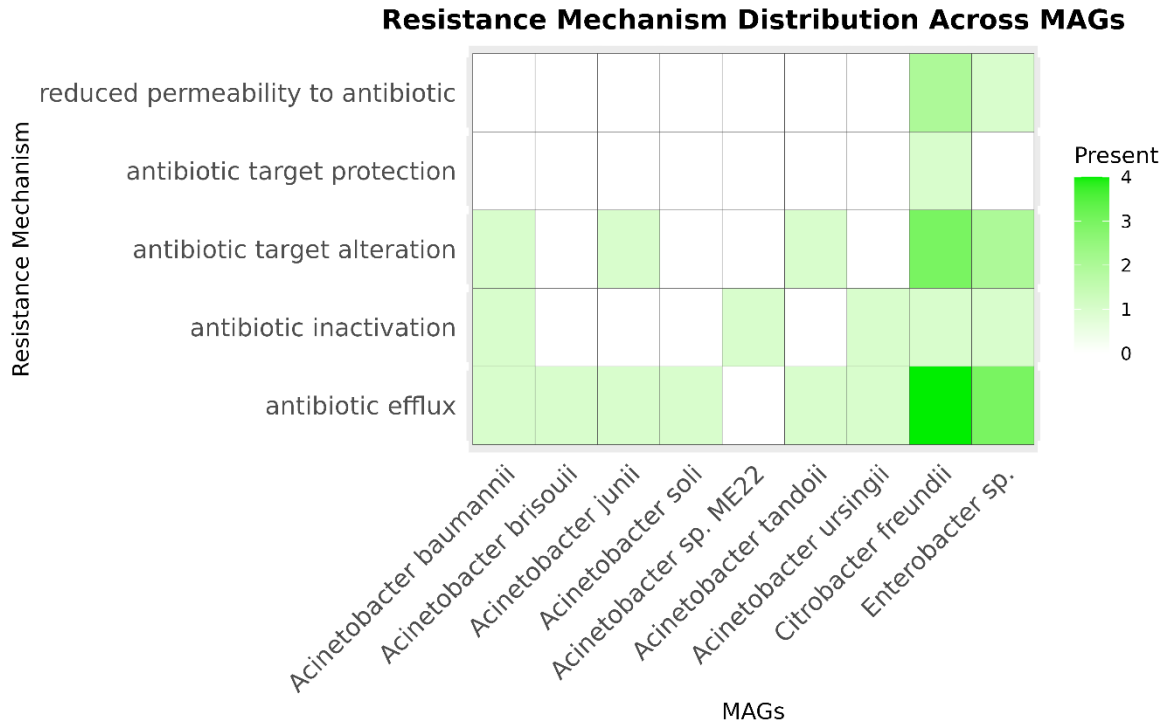


Figura 10: Mapa de calor do mecanismo de resistência de MAGs pesquisado com RGI mostrando o perfil de multirresistência dos prováveis hospedeiros dos bacteriófagos selecionados.

Estes resultados mostram que bacteriófagos identificados nesta região BP4 possuem uma maior chance de poder atuar contra esses gêneros que apresentaram um maior número de genes de resistência a antibióticos identificados.

#### 4.9 Mecanismos de defesa e suscetibilidade a bacteriófagos na região BP4

Foi utilizado o software DefenseFinder (TESSON *et al.*, 2022) para realizar uma análise funcional para detecção de genes de defesa contra vírus possivelmente presentes nas MAGs selecionadas (Tabela 18). A partir disso, usou-se um script em R para visualizar por um gráfico de

Sankey as anotações relacionadas à presença de sistemas de defesa encontrados para estes MAGs que foram: *gcu167* (KIEFFER *et al.*, 2025), dCTPdeaminase (HSUEH *et al.*, 2022), Wadjet (TESSON *et al.*, 2022), VP1817 (GETZ *et al.*, 2025), Septu (STOKAR-AVIHAIL *et al.*, 2023), SanaTA (SBERRO *et al.*, 2013), RM (SNEPPEN *et al.*, 2015), MazEF (HAZAN; ENGELBERG-KULKA, 2004), Cas (presentes no sistema CRISPR) (KOONIN; MAKAROVA, 2024; LOUREIRO; DA SILVA, 2019) e AbiD (mecanismo Abi) (CHOPIN; CHOPIN; BIDNENKO, 2005) que são sistemas de defesa que a bactéria pode usar para se proteger de ataque contra bacteriófagos (Figura 11).

Esta análise mostrou que MAGs que apresentaram uma maior presença de mecanismos de resistência a antibióticos, como a *Citrobacter freuri* e *Enterobacter sp.* foram as MAGs que foi identificado o sistema MazEF que codifica um sistema de toxina-antitoxina capaz de gerar uma morte celular programada em células infectadas por bacteriófagos, diminuindo a prevalência de fagos no ambiente (HAZAN; ENGELBERG-KULKA, 2004; KELLY *et al.*, 2023). Para a MAG de *Enterobacter sp.* também foi identificado sistemas de restrição-modificação (RM) (SNEPPEN *et al.*, 2015) em maior quantidade quando comparado com as outras MAGs. Vale ressaltar que este sistema RM é um dos mais abundantes sistemas de defesa, representando uma barreira contra a disseminação de plasmídeos e Transferência Horizontal de Genes (THG) (SHAW; ROCHA; MACLEAN, 2023); e que bacteriófagos não temperados (líticos) apresentam maior facilidade em resistir ao sistema de restrição-modificação (RM) (SHAW; ROCHA; MACLEAN, 2023), identificado na MAG de *Enterobacter sp.*.

Para a MAG de *Acinetobacter ursingii*, foi identificado sistemas de restrição-modificação (RM) e sistemas Cas, que, em conjunto, podem apresentar uma maior proteção contra fagos (DUPUIS *et al.*, 2013). A MAG de *Acinetobacter tandoi* apresentou os sistemas Septu I (STOKAR-AVIHAIL *et al.*, 2023), Wadjet (TESSON *et al.*, 2022; WU *et al.*, 2024) e Cas, enquanto a *Acinetobacter sp. WCHA29* apresentou *gcu167* (KIEFFER *et al.*, 2025) e sistemas de restrição-modificação (RM), sendo esta MAG a que nenhum mecanismo de resistência à antibióticos foi identificado (Figura 10).

A MAG de *Acinetobacter soli* foi identificado o sistema VP1817 que faz parte de integrons que agem como bibliotecas de defesa contra fagos (GETZ *et al.*, 2025) e o dCTP deaminase que usa uma abordagem de depleção de dCTPs e conseqüentemente dificultando a replicação do DNA

viral na célula (HSUEH *et al.*, 2022). A MAG de *Acinetobacter junii* identificou os sistemas Cas e SanaTA (KELLY *et al.*, 2023; SBERRO *et al.*, 2013) como mecanismo de defesa.

Em *Acinetobacter brisouii* identificou-se sistemas de restrição-modificação (RM) (SNEPPEN *et al.*, 2015) e o sistema Abi (AbiD) ou de infecção abortiva em que durante a adsorção do fago seu desenvolvimento é interrompido e a célula inicia um processo de morte celular como mecanismo de defesa contra a disseminação do fago (CHOPIN; CHOPIN; BIDNENKO, 2005). Enquanto a MAG de *Acinetobacter baumannii* identificou-se o sistema de defesa do tipo Cas, que estão presentes dentro de sistemas CRISPR (LOUREIRO; DA SILVA, 2019).

Quanto à presença de regiões de defesa por CRISPR (*Clustered Regularly Interspaced Short Palindromic Repeat*), as MAGs que mais foram identificadas com sistemas Cas do CRISPR foram de *Acinetobacter ursingii*, *Acinetobacter tandoi*, *Acinetobacter junii* e a *Acinetobacter balmani*. Das 3 MAGs com maior identificação de sistemas Cas tiveram também uma menor presença de mecanismos de resistência, com exceção da *Acinetobacter balmani* que foi identificado 3 mecanismos de defesa contra antibiótico: efluxo de antibiótico, alteração do alvo antibiótico e inativação de antibióticos (Figura 10 e 11).

A identificação deste sistema de defesa por CRISPR é interessante por ele poder oferecer proteção contra bacteriófagos e outros elementos móveis, sendo um sistema imune adaptativo contra vírus em que se usa uma sequência de RNA guia que irá se ligar a estes elementos e clivar estas regiões, evitando a infecção viral e garantindo a sobrevivência da bactéria (KOONIN; MAKAROVA, 2024).

Na MAG de *Acinetobacter sp. ME22* nenhum mecanismo de defesa foi encontrado usando o DefenseFinder (TESSON *et al.*, 2022). Este resultado pode ter sido obtido devido a 3 possibilidades: 1- vieses durante a geração das MAGs; 2- Ao fato de regiões de defesa não terem sido obtidas durante a geração desta MAG que teve uma completude de apenas 56.94% (Tabela 17), abrindo a possibilidade de que genes de defesa contra fagos possam estar presentes nestas regiões não observadas; ou 3-Pela ausência de tais mecanismos nesta MAG, ou por não apresentar mecanismos de defesas conhecidos pela ferramenta utilizada.

## Defense type Distribution Across MAGs

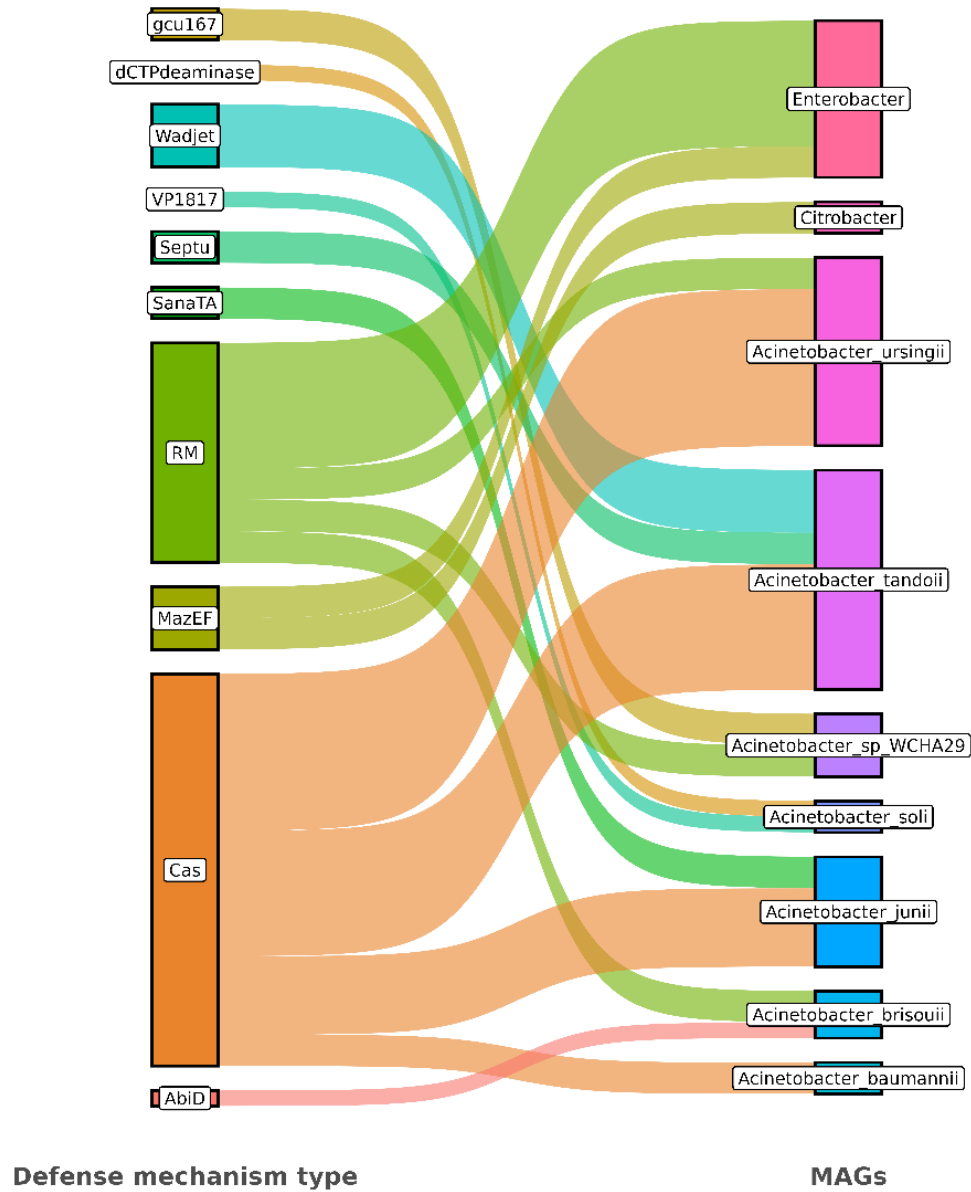


Figura 11: Diagrama de Sankey ilustrando MAGs e os tipos de sistemas de defesa contra vírus que foram identificados usando a ferramenta DefenseFinder.

### 4.10 Predição de presença de sistemas anti-CRISPR e do ciclo de vida viral dos bacteriófagos na região BP4

Para a análise das sequências virais selecionadas usou-se a ferramenta web PhageScope (WANG et al., 2024a) para verificar se estas sequências apresentariam algum sistema anti-CRISPR que pode atuar como defesa contra o sistema CRISPR presente em algumas bactérias, e buscar prever o ciclo viral destas sequências virais selecionadas.

Os resultados obtidos não mostraram nenhum sistema anti-CRISPR em nenhuma das sequências analisadas, indicando que este mecanismo de defesa contra CRISPR possa não estar presentes nestes bacteriófagos.

Quanto à predição do ciclo viral das sequências analisadas, observou-se 6 sequências virais com provável ciclo lítico, enquanto outras 5 foram identificados como fagos temperados (Tabela 20)

<i>phageid</i>	<i>lifestyle</i>	<i>likely Phage-host genus</i>
NODE_4769	virulent	Acinetobacter
NODE_481	virulent	Acinetobacter
NODE_9535	virulent	Acinetobacter
NODE_177	virulent	Bacteroides/Phocaeicola
NODE_382	temperate	Citrobacter/Enterobacter
NODE_173	temperate	Enterobacter
NODE_190	temperate	Enterobacter
NODE_162	virulent	Limnohabitans
NODE_294	temperate	Malikia
NODE_1	virulent	Not identified
NODE_249	temperate	Not identified

Tabela 20: Predição do ciclo viral dos bacteriófagos selecionados da região BP4

Entretanto, os bacteriófagos associados às MAGs identificadas no tópico 4.65 como mais suscetíveis a causar infecção foram preditos como temperados (Tabela 20). Esse resultado sugere que, em uma eventual bioprospecção desses fagos, deve-se considerar sua natureza temperada.

Nesse contexto, as sequências virais correspondentes aos bacteriófagos NODE\_173, NODE\_190 e NODE\_382 poderiam ser isoladas experimentalmente por meio de abordagens específicas para a captura de fagos temperados e, subsequentemente, modificadas geneticamente para a conversão em fagos líticos, visando sua aplicação no controle de bactérias resistentes a antibióticos como potenciais produtos biotecnológicos.

Entre os fagos identificados como líticos, todas as sequências virais associadas ao gênero *Acinetobacter* foram classificadas como bacteriófagos líticos (Tabela 20). Notavelmente, esse mesmo grupo apresentou a maior frequência de sistemas de defesa do tipo CRISPR (Cas), detectados em 4 das 8 MAGs analisadas (Tabela 18), e uma maior diversidade também por 9 de 10 sistemas de defesa em 7 de 8 MAGs analisadas terem sido identificados em *Acinetobacter* (Figura 11), o que sugere uma possível pressão seletiva como mecanismo de defesa contra esses fagos. Esses achados indicam que a bioprospecção de bacteriófagos associados a *Acinetobacter* no ponto BP4 provavelmente demandará estratégias experimentais voltadas especificamente para o isolamento de fagos líticos.

## 5. CONCLUSÕES

Os resultados deste trabalho mostraram que a partir dos dados metagenômicos obtidos para a cidade de Bragança-PA foi possível caracterizar e analisar a diversidade e composição do microbioma de vírus e bactérias presentes em diferentes pontos do rio Cereja.

Também foi possível investigar o perfil de genes de resistência entre as regiões coletadas e correlacionar estes dados para identificar que a região com o maior potencial de estar contaminada.

Os resultados obtidos permitiram correlacionar a composição de bactérias presentes nos dados metagenômicos com o perfil de resistência, e com isso, identificar a presença dos gêneros provavelmente associado ao perfil de resistência e o ponto de maior contaminação selecionado.

A partir da caracterização dos perfis de vírus e bactérias presentes e correlação com os metadados de perfil de resistência, foi possível identificar que o ponto BP4 seria o mais contaminado, e, portanto, o com maior probabilidade de se encontrar bacteriófagos com perfil de interesse biotecnológico para combater bactérias patogênicas que tenham resistência a antibióticos. Este ponto, de forma interessante, se encontra entre um hospital da região (Figura 2), indicando um possível descarte irregular que possa ter gerado este acúmulo de bactérias patogênicas e de genes de resistência a antibióticos.

Este trabalho também foi capaz de identificar e selecionar algumas sequências virais com boa qualidade de bacteriófagos de cada ponto coletado, e, dentre elas, selecionar quais teriam maior chances de serem bacteriófagos de bactérias patogênicas que tenham interesse biotecnológico. MAGs foram geradas em todas as regiões identificadas e com isso, buscou-se às que teriam maior probabilidade de serem hospedeiras das sequências virais obtidas. Observou-se também que apenas na região mais contaminada foi possível estabelecer esta relação bactéria-hospedeiro e que elas sejam de MAGs que tenham alguma patogenicidade e presença de genes de resistência a antibióticos.

Por fim, as bactérias e bacteriófagos que foram selecionados por inferência pela sua relação bactéria-hospedeiro foram selecionados e analisados separadamente quanto à presença de genes de resistência e análise de genes funcionais para as sequências virais e MAGs selecionadas; e quanto à presença de mecanismos de defesa nas MAGs e presença de sistemas anti-CRISPR-Cas para as sequências virais selecionadas. Foi observado que a presença de genes associados a fagos e regiões que tenham o sistema CRISPR presentes nas MAGs obtidas, e com isso, observou-se que as MAGs

que tiveram uma maior identificação de genes de resistência a antibióticos (*Citrobacter freuri* e *Enterobacter sp.*), foi identificado o sistema de toxina-antitoxina MazEF, e em *Enterobacter sp.* o sistema de restrição-modificação (RM), sendo identificados fagos temperados para estas MAGs nos dados metagenômicos. De forma análoga, para os bacteriófagos que têm como hospedeiro o gênero *Acinetobacter*, os resultados sugerem que estratégias experimentais direcionadas ao isolamento de fagos líticos podem aumentar a probabilidade de recuperação de vírus de interesse biotecnológico.

Assim, os dados metagenômicos apresentados permitiram delinear perspectivas sobre a dinâmica bacteriófago-hospedeiro em ambientes de maior contaminação, constituindo uma abordagem promissora para orientar experimentos *in vitro* direcionados ao isolamento e caracterização de bacteriófagos com potencial aplicação no controle de bactérias resistentes a antibióticos.

## 6. PERSPECTIVAS

Com o desenvolvimento e barateamento das tecnologias de sequenciamento, espera-se que este trabalho possa auxiliar em metodologias futuras de busca de pontos ótimos de coleta para bioprospecção de bacteriófagos que tenham interesse biotecnológico para atuar contra bactérias patogênicas que tenham resistência a antibióticos e direcionar experimentos para aumentar as chances de obtenção de bacteriófagos de interesse a partir dos resultados obtidos pelos dados metagenômicos.

Espera-se que se possa usar estas tecnologias de sequenciamento para realizar uma “leitura” dos locais, em que, além do biomonitoramento, se possa prospectar potenciais bacteriófagos a serem isolados experimentalmente para usá-lo no combate a bactérias resistentes a antibióticos antes mesmos de iniciar os testes experimentais para o seu isolamento, barateando custos e sendo mais específico e direcionado para a obtenção de bacteriófagos de interesse.

Considerando os resultados obtidos pelo ponto mais “contaminado” foi possível identificar que as MAGs *Citrobacter freuri* e *Enterobacter sp.* que foram as que mais apresentaram potencial de se identificar sequências virais de interesse que possam atuar contra bactérias patogênicas e que tenha resistência a antibiótico. Estes resultados podem orientar em trabalhos futuros para realizar experimentos *in vitro* mais direcionados, e, portanto, com maior potencial de isolar e identificar bacteriófagos de maior interesse biotecnológico com um menor risco de não identificação por se tratar de um experimento mais direcionado ao hospedeiro que já se tem conhecimento da sua presença a partir dos resultados obtidos.

A partir dos resultados obtidos no ponto mais contaminado, verificou-se que as MAGs classificadas como *Citrobacter freundii* e *Enterobacter sp.* apresentaram o maior potencial para a identificação de sequências virais de interesse, associadas a bactérias patogênicas resistentes a antibióticos. No entanto, como as sequências virais selecionadas foram preditas majoritariamente como fagos temperados, futuros desenhos experimentais voltados ao isolamento desse tipo de fago provavelmente seriam mais adequados para seu isolamento e bioprospecção.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

ACKERMANN, H.-W. Bacteriophage electron microscopy. **Advances in virus research**, v. 82, p. 1–32, 1 jan. 2012. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/22420849>>. Acesso em: 20 jun. 2023.

AGUILAR-VERA, A.; BELLO-LÓPEZ, E.; PANTOJA-NUÑEZ, G. I.; RODRÍGUEZ-LÓPEZ, G. M.; MORALES-ERASTO, V.; CASTILLO-RAMÍREZ, S. *Acinetobacter junii*: an emerging One Health pathogen. **mSphere**, v. 9, n. 5, p. e0016224, 29 maio 2024. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/38606973>>.

AHMED, W.; YE, W.; PAN, J.; LIU, S.; JI, W.; ZHOU, S.; WANG, F.; LI, Z.; MOHANY, M.; WANG, X. Evaluation the role of *Luteibacter pinisoli* DP2-30 in mitigating pine wilt disease caused by *Bursaphelenchus xylophilus* through modulation of host microbiome. **Frontiers in Plant Science**, v. 16, n. March, p. 1–19, 5 mar. 2025. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fpls.2025.1515506/full>>.

ALCOCK, B. P.; RAPHENYA, A. R.; LAU, T. T. Y.; TSANG, K. K.; BOUCHARD, M.; EDALATMAND, A.; HUYNH, W.; NGUYEN, A. L. V.; CHENG, A. A.; LIU, S.; MIN, S. Y.; MIROSHNICHENKO, A.; TRAN, H. K.; WERFALLI, R. E.; NASIR, J. A.; OLONI, M.; SPEICHER, D. J.; FLORESCU, A.; SINGH, B.; FALTYN, M.; HERNANDEZ-KOUTOUCHEVA, A.; SHARMA, A. N.; BORDELEAU, E.; PAWLOWSKI, A. C.; ZUBYK, H. L.; DOOLEY, D.; GRIFFITHS, E.; MAGUIRE, F.; WINSOR, G. L.; BEIKO, R. G.; BRINKMAN, F. S. L.; HSIAO, W. W. L.; DOMSELAAR, G. V.; MCARTHUR, A. G. CARD 2020: Antibiotic resistance surveillance with the comprehensive antibiotic resistance database. **Nucleic Acids Research**, v. 48, n. D1, p. D517–D525, 1 jan. 2020. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/31665441/>>. Acesso em: 19 jun. 2021.

ALVES, J.; DIAS, L.; MATEUS, J.; MARQUES, J.; GRAÇAS, D.; RAMOS, R.; SELDIN, L.; HENRIQUES, I.; SILVA, A.; FOLADOR, A. Resistome in Lake Bolonha, Brazilian Amazon: Identification of Genes Related to Resistance to Broad-Spectrum Antibiotics. **Frontiers in Microbiology**, v. 11, n. February, p. 1–13, 2020.

AMINOV, R. I. A brief history of the antibiotic era: lessons learned and challenges for the future. **Frontiers in microbiology**, v. 1, n. DEC, p. 134, 2010. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/21687759>>.

ANDREWS, S.; OTHERS. **FastQC: a quality control tool for high throughput sequence data. 2010** <https://www.Bioinformatics.Babraham.Ac.Uk/Projects/Fastqc/2010>. Disponível em: <<https://www.bibtex.org/bibtex/f230a919c34360709aa298734d63dca3>>. Acesso em: 22 out. 2019.

ANTIPOV, D.; RAIKO, M.; LAPIDUS, A.; PEVZNER, P. A. Metaviral SPAdes: assembly of viruses from metagenomic data. **Bioinformatics (Oxford, England)**, v. 36, n. 14, p. 4126–4129, 15 ago. 2020. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/32413137>>.

BECSEI, Á.; SOLYMOSI, N.; CSABAI, I.; MAGYAR, D. Detection of antimicrobial resistance genes in urban air. **MicrobiologyOpen**, v. 10, n. 6, p. e1248, nov. 2021. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/34964297>>.

BETTS, R. A.; MALHI, Y.; ROBERTS, J. T. The future of the Amazon: new perspectives from climate, ecosystem and social sciences. **Philosophical transactions of the Royal Society of London. Series B, Biological sciences**, v. 363, n. 1498, p. 1729–35, 27 maio 2008. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/18267894>>.

BRIVES, C.; POURRAZ, J. Phage therapy as a potential solution in the fight against AMR: obstacles and possible futures. **Palgrave Communications**, v. 6, n. 1, p. 100, 19 dez. 2020. Disponível em: <<http://dx.doi.org/10.1057/s41599-020-0478-4>>.

BROUWER, R.; PINTO, R.; DUGSTAD, A.; NAVRUD, S. The economic value of the Brazilian Amazon rainforest ecosystem services: A meta-analysis of the Brazilian literature. **PLoS one**, v. 17, n. 5, p. e0268425, 2022. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/35588116>>.

CAMARGO, A. P.; ROUX, S.; SCHULZ, F.; BABINSKI, M.; XU, Y.; HU, B.; CHAIN, P. S. G.; NAYFACH, S.; KYRPIDES, N. C. Identification of mobile genetic elements with geNomad. **Nature biotechnology**, v. 42, n. 8, p. 1303–1312, ago. 2024. Disponível em: <<http://dx.doi.org/10.1038/s41587-023-01953-y>>.

CARDENAS-ALEGRIA, O. V.; FERREIRA, V. B. C.; NOGUERA, W. G.; MARTINS, D. T.; MARTINS NETO, A. P.; MONTEIRO PONTES, P. R.; LOPES CAVALCANTE, R. B.; AGUIAR ALVES, S. I.; LUIZ DA COSTA DA SILVA, A.; GOMES COSTA, R.; FRANCO DE LOS SANTOS, E. F.; AZEVEDO, V. A. de C.; RAMOS, R. T. J. Microbiome analyses of the Uraim River in the Amazon and georeferencing analyses to establish correlation with anthropogenic impacts of land use. **Frontiers in Environmental Science**, v. 12, n. August, p. 1–12, 6 ago. 2024. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fenvs.2024.1404230/full>>.

CARDENAS ALEGRIA, O.; PIRES QUARESMA, M.; DIAS DANTAS, C. W.; SILVA GUEDES LOBATO, E. M.; DE OLIVEIRA ARAGÃO, A.; PATROCA DA SILVA, S.; COSTA BARROS DA SILVA, A.; RIBEIRO CRUZ, A. C.; RAMOS, R. T. J.; CARNEIRO, A. R. Impacts of soybean agriculture on the resistome of the Amazonian soil. **Frontiers in microbiology**, v. 13, p. 948188, 2022. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/36160259>>.

CARRIÓN, V. J.; CORDOVEZ, V.; TYC, O.; ETALO, D. W.; DE BRUIJN, I.; DE JAGER, V. C. L.; MEDEMA, M. H.; EBERL, L.; RAAIJMAKERS, J. M. Involvement of Burkholderiaceae and sulfurous volatiles in disease-suppressive soils. **The ISME journal**, v. 12, n. 9, p. 2307–2321, set. 2018. Disponível em: <<http://dx.doi.org/10.1038/s41396-018-0186-x>>.

CARVALHEIRA, A.; SILVA, J.; TEIXEIRA, P. Acinetobacter spp. in food and drinking water - A review. **Food microbiology**, v. 95, n. July 2020, p. 103675, maio 2021. Disponível em: <<https://doi.org/10.1016/j.fm.2020.103675>>.

CENTERS FOR DISEASE CONTROL AND PREVENTION. **COVID-19: U.S. Impact on**

**Antimicrobial Resistance, Special Report 2022** Atlanta, GA: U.S. Department of Health and Human Services, CDC; 2022. [s.l.: s.n.]. Disponível em: <<https://stacks.cdc.gov/view/cdc/117915>>.

CHEN, J.; YU, Z.; MICHEL, F. C.; WITTUM, T.; MORRISON, M. Development and application of real-time PCR assays for quantification of erm genes conferring resistance to macrolides-lincosamides-streptogramin B in livestock manure and manure management systems. **Applied and environmental microbiology**, v. 73, n. 14, p. 4407–16, jul. 2007. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/17496134>>.

CHEN, S.; ZHOU, Y.; CHEN, Y.; GU, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. **Bioinformatics (Oxford, England)**, v. 34, n. 17, p. i884–i890, 1 set. 2018. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/30423086>>.

CHKLOVSKI, A.; PARKS, D. H.; WOODCROFT, B. J.; TYSON, G. W. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. **Nature methods**, v. 20, n. 8, p. 1203–1212, ago. 2023. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/37500759>>.

CHOPIN, M.-C.; CHOPIN, A.; BIDNENKO, E. Phage abortive infection in lactococci: variations on a theme. **Current opinion in microbiology**, v. 8, n. 4, p. 473–9, ago. 2005. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/15979388>>.

CHUN, B. H.; LEE, Y.; JIN, H. M.; JEON, C. O. Cloacibacterium caeni sp. nov., isolated from activated sludge. **International journal of systematic and evolutionary microbiology**, v. 67, n. 6, p. 1688–1692, jun. 2017. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/28150572>>.

COOPER, M. B.; SMITH, A. G. Exploring mutualistic interactions between microalgae and bacteria in the omics age. **Current opinion in plant biology**, v. 26, n. Figure 1, p. 147–53, ago. 2015. Disponível em: <<http://dx.doi.org/10.1016/j.pbi.2015.07.003>>.

CUI, Z.; GUO, X.; DONG, K.; ZHANG, Y.; LI, Q.; ZHU, Y.; ZENG, L.; TANG, R.; LI, L. Safety assessment of Staphylococcus phages of the family Myoviridae based on complete genome sequences. **Scientific reports**, v. 7, n. January, p. 41259, 24 jan. 2017. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/28117392>>.

CUI, Z.; GUO, X.; FENG, T.; LI, L. Exploring the whole standard operating procedure for phage therapy in clinical practice. **Journal of translational medicine**, v. 17, n. 1, p. 373, 14 nov. 2019. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/31727099>>.

DA SILVA MORAIS, E.; GRIMAUD, G. M.; WARDA, A.; STANTON, C.; ROSS, P. Genome plasticity shapes the ecology and evolution of Phocaeicola dorei and Phocaeicola vulgatus. **Scientific reports**, v. 14, n. 1, p. 10109, 2 maio 2024. Disponível em: <<https://doi.org/10.1038/s41598-024-59148-7>>.

DANTAS, C. W. D.; MARTINS, D. T.; NOGUEIRA, W. G.; ALEGRIA, O. V. C.; RAMOS, R.

T. J. Tools and methodology to in silico phage discovery in freshwater environments. **Frontiers in microbiology**, v. 15, n. May, p. 1390726, 31 maio 2024. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fmicb.2024.1390726/full>>.

DEFAZIO, G.; TANGARO, M. A.; PESOLE, G.; FOSSO, B. kMetaShot: a fast and reliable taxonomy classifier for metagenome-assembled genomes. **Briefings in bioinformatics**, v. 26, n. 1, 22 nov. 2024. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/39749666>>.

DEHBANIPOUR, R.; GHALAVAND, Z. Anti-virulence therapeutic strategies against bacterial infections: recent advances. **Germs**, v. 12, n. 2, p. 262–275, 2022.

DONATO, J. J.; MOE, L. A.; CONVERSE, B. J.; SMART, K. D.; BERKLEIN, F. C.; MCMANUS, P. S.; HANDELSMAN, J. Metagenomic analysis of apple orchard soil reveals antibiotic resistance genes encoding predicted bifunctional proteins. **Applied and environmental microbiology**, v. 76, n. 13, p. 4396–401, jul. 2010. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/20453147>>.

DUPUIS, M.-È.; VILLION, M.; MAGADÁN, A. H.; MOINEAU, S. CRISPR-Cas and restriction-modification systems are compatible and increase phage resistance. **Nature communications**, v. 4, n. May, p. 2087, 2013. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/23820428>>.

DURHAM, B. P.; SHARMA, S.; LUO, H.; SMITH, C. B.; AMIN, S. A.; BENDER, S. J.; DEARTH, S. P.; VAN MOOY, B. A. S.; CAMPAGNA, S. R.; KUJAWINSKI, E. B.; ARMBRUST, E. V.; MORAN, M. A. Cryptic carbon and sulfur cycling between surface ocean plankton. **Proceedings of the National Academy of Sciences of the United States of America**, v. 112, n. 2, p. 453–7, 13 jan. 2015. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/25548163>>.

ELBEHIRY, A.; AL SHOAIABI, M.; ALZHRANI, H.; IBRAHEM, M.; MOUSSA, I.; ALZABEN, F.; ALSUBKI, R. A.; HEMEG, H. A.; ALMUTAIRI, D.; ALTHOBAITI, S.; ALANAZI, F.; ALOTAIBI, S. A.; ALMUTAIRI, H.; ALZHRANI, A.; ABU-OKAIL, A. Enterobacter cloacae from urinary tract infections: frequency, protein analysis, and antimicrobial resistance. **AMB Express**, v. 14, n. 1, p. 17, 8 fev. 2024. Disponível em: <<https://doi.org/10.1186/s13568-024-01675-7>>.

FOKINE, A.; ROSSMANN, M. G. Molecular architecture of tailed double-stranded DNA phages. **Bacteriophage**, v. 4, n. 1, p. e28281, 1 jan. 2014. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/24616838>>.

GARC, P. **Bacteriophages**. [s.l.] MDPI, 2020.

GETZ, L. J.; FAIRBURN, S. R.; VIVIAN LIU, Y.; QIAN, A. L.; MAXWELL, K. L. Integrons are anti-phage defence libraries in *Vibrio parahaemolyticus*. **Nature Microbiology**, v. 10, n. 3, p. 724–733, 2025. Disponível em: <<https://doi.org/10.1038/s41564-025-01927-7>>.

GUO, J.; BOLDUC, B.; ZAYED, A. A.; VARSANI, A.; DOMINGUEZ-HUERTA, G.;

DELMONT, T. O.; PRATAMA, A. A.; GAZITÚA, M. C.; VIK, D.; SULLIVAN, M. B.; ROUX, S. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. **Microbiome**, v. 9, n. 1, p. 37, 1 fev. 2021. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/33522966>>.

GUPTA, V.; YE, G.; OLESKY, M.; LAWRENCE, K.; MURRAY, J.; YU, K. Trends in resistant Enterobacteriaceae and Acinetobacter species in hospitalized patients in the United States: 2013–2017. **BMC infectious diseases**, v. 19, n. 1, p. 742, 23 ago. 2019. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/31443635>>.

HAMPTON, H. G.; WATSON, B. N. J.; FINERAN, P. C. The arms race between bacteria and their phage foes. **Nature**, v. 577, n. 7790, p. 327–336, 2020. Disponível em: <<http://dx.doi.org/10.1038/s41586-019-1894-8>>.

HAN, H.; WANG, Z.; ZHU, S. Benchmarking metagenomic binning tools on real datasets across sequencing platforms and binning modes. **Nature communications**, v. 16, n. 1, p. 2865, 24 mar. 2025. Disponível em: <<http://dx.doi.org/10.1038/s41467-025-57957-6>>.

HANCOCK, R. E. W.; SAHL, H.-G. Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. **Nature biotechnology**, v. 24, n. 12, p. 1551–7, dez. 2006. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/17160061>>.

HAZAN, R.; ENGELBERG-KULKA, H. Escherichia coli mazEF-mediated cell death as a defense mechanism that inhibits the spread of phage P1. **Molecular genetics and genomics : MGG**, v. 272, n. 2, p. 227–34, set. 2004. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/15316771>>.

HO, S. F. S.; WHEELER, N. E.; MILLARD, A. D.; VAN SCHAIK, W. Gauge your phage: benchmarking of bacteriophage identification tools in metagenomic sequencing data. **Microbiome**, v. 11, n. 1, p. 84, 21 abr. 2023. Disponível em: <<https://www.biorxiv.org/content/10.1101/2021.04.12.438782v2%0Ahttps://www.biorxiv.org/content/10.1101/2021.04.12.438782v2.abstract>>.

HSUEH, B. Y.; SEVERIN, G. B.; ELG, C. A.; WALDRON, E. J.; KANT, A.; WESSEL, A. J.; DOVER, J. A.; RHOADES, C. R.; RIDENHOUR, B. J.; PARENT, K. N.; NEIDITCH, M. B.; RAVI, J.; TOP, E. M.; WATERS, C. M. Phage defence by deaminase-mediated depletion of deoxynucleotides in bacteria. **Nature microbiology**, v. 7, n. 8, p. 1210–1220, ago. 2022. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/35817890>>.

HUGHES, C. C.; FENICAL, W. Antibacterials from the sea. **Chemistry (Weinheim an der Bergstrasse, Germany)**, v. 16, n. 42, p. 12512–25, 8 nov. 2010. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf>>.

INDA-DÍAZ, J. S.; LUND, D.; PARRAS-MOLTÓ, M.; JOHNNING, A.; BENGTSSON-PALME, J.; KRISTIANSSON, E. Latent antibiotic resistance genes are abundant, diverse, and mobile in human, animal, and environmental microbiomes. **Microbiome**, v. 11, n. 1, p. 44, 8 mar. 2023.

Disponível em: <<https://doi.org/10.1186/s40168-023-01479-0>>.

JABEEN, I.; ISLAM, S.; HASSAN, A. K. M. I.; TASNIM, Z.; SHUVO, S. R. A brief insight into *Citrobacter* species - a growing threat to public health. **Frontiers in antibiotics**, v. 2, n. December, p. 1276982, 2023. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/39816660>>.

JEZBEROVÁ, J.; JEZBERA, J.; ZNACHOR, P.; NEDOMA, J.; KASALICKÝ, V.; ŠIMEK, K. The Limnohabitans Genus Harbors Generalistic and Opportunistic Subtypes: Evidence from Spatiotemporal Succession in a Canyon-Shaped Reservoir. **Applied and environmental microbiology**, v. 83, n. 21, 1 nov. 2017. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/28842542>>.

JOHANSEN, J. E.; BINNERUP, S. J.; KROER, N.; MØLBAK, L. *Luteibacter rhizovicinus* gen. nov., sp. nov., a yellow-pigmented gammaproteobacterium isolated from the rhizosphere of barley (*Hordeum vulgare* L.). **International journal of systematic and evolutionary microbiology**, v. 55, n. Pt 6, p. 2285–2291, nov. 2005. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/16280484>>.

KAKASIS, A.; PANITSA, G. Bacteriophage therapy as an alternative treatment for human infections. A comprehensive review. **International Journal of Antimicrobial Agents**, v. 53, n. 1, p. 16–21, 2019. Disponível em: <<https://doi.org/10.1016/j.ijantimicag.2018.09.004>>.

KANZI, A. M.; SAN, J. E.; CHIMUKANGARA, B.; WILKINSON, E.; FISH, M.; RAMSURAN, V.; DE OLIVEIRA, T. Next Generation Sequencing and Bioinformatics Analysis of Family Genetic Inheritance. **Frontiers in genetics**, v. 11, n. October, p. 544162, 23 out. 2020. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fgene.2020.544162/full>>.

KELLY, A.; ARROWSMITH, T. J.; WENT, S. C.; BLOWER, T. R. Toxin-antitoxin systems as mediators of phage defence and the implications for abortive infection. **Current opinion in microbiology**, v. 73, p. 102293, jun. 2023. Disponível em: <<https://doi.org/10.1016/j.mib.2023.102293>>.

KIEFFER, N.; HIPÓLITO, A.; ORTIZ-MIRAVALLÉS, L.; BLANCO, P.; DELOBELLE, T.; VIZUETE, P.; OJEDA, F. M.; JOVÉ, T.; JURENAS, D.; GARCÍA-QUINTANILLA, M.; CARVALHO, A.; DOMINGO-CALAP, P.; ESCUDERO, J. A. Mobile integrons encode phage defense systems. **Science (New York, N.Y.)**, v. 388, n. 6747, p. eads0915, 8 maio 2025. Disponível em: <<http://biorxiv.org/lookup/doi/10.1101/2024.07.02.601719>>.

KIEFT, K.; ZHOU, Z.; ANANTHARAMAN, K. VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. **Microbiome**, v. 8, n. 1, p. 90, 10 jun. 2020. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/32522236>>.

KIRIENKO, N. V.; RAHME, L.; CHO, Y.-H. Editorial: Beyond Antimicrobials: Non-traditional Approaches to Combating Multidrug-Resistant Bacteria. **Frontiers in cellular and infection microbiology**, v. 9, n. October, p. 343, 2019. Disponível em:

<<http://www.ncbi.nlm.nih.gov/pubmed/31681623>>.

KLÜMPER, U.; GIONCHETTA, G.; CATÃO, E.; BELLANGER, X.; DIELACHER, I.; ELENA, A. X.; FANG, P.; GALAZKA, S.; GORYLUK-SALMONOWICZ, A.; KNEIS, D.; OKOROAFOR, U.; RADU, E.; SZADZIUL, M.; SZEKERES, E.; TEBAN-MAN, A.; COMAN, C.; KREUZINGER, N.; POPOWSKA, M.; VIERHEILIG, J.; WALSH, F.; WOEGERBAUER, M.; BÜRGMANN, H.; MERLIN, C.; BERENDONK, T. U. Environmental microbiome diversity and stability is a barrier to antimicrobial resistance gene accumulation. **Communications biology**, v. 7, n. 1, p. 706, 8 jun. 2024. Disponível em: <<http://dx.doi.org/10.1038/s42003-024-06338-8>>.

KNEZEVIC, P.; ADRIAENSSENS, E. M.; ICTV REPORT CONSORTIUM. ICTV Virus Taxonomy Profile: Inoviridae. **The Journal of general virology**, v. 102, n. 7, p. 10–11, jul. 2021. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/34227934>>.

KOONIN, E. V.; MAKAROVA, K. S. CRISPR in mobile genetic elements: counter-defense, inter-element competition and RNA-guided transposition. **BMC biology**, v. 22, n. 1, p. 295, 18 dez. 2024. Disponível em: <<https://doi.org/10.1186/s12915-024-02090-x>>.

LIN, D. M.; KOSKELLA, B.; LIN, H. C. Phage therapy: An alternative to antibiotics in the age of multi-drug resistance. **World journal of gastrointestinal pharmacology and therapeutics**, v. 8, n. 3, p. 162–173, 6 ago. 2017. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/28828194>>.

LIU, C.; CUI, Y.; LI, X.; YAO, M. microeco: an R package for data mining in microbial community ecology. **FEMS microbiology ecology**, v. 97, n. 2, p. 1–9, 26 jan. 2021. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/33332530>>.

LIU, J.; GAO, S.; DONG, Y.; LU, C.; LIU, Y. Isolation and characterization of bacteriophages against virulent *Aeromonas hydrophila*. **BMC Microbiology**, v. 20, n. 1, p. 1–13, 2020.

LOUREIRO, A.; DA SILVA, G. J. Crispr-cas: Converting a bacterial defence mechanism into a state-of-the-art genetic manipulation tool. **Antibiotics**, v. 8, n. 1, 2019.

LOZIER, Z.; HILL, L.; SEMMANN, E.; MILLER, W. A. A proposed new Tombusviridae genus featuring extremely long 5' untranslated regions and a luteo/polerovirus-like gene block. **Frontiers in Virology**, v. 4, n. August, p. 1–21, 2024.

LU, J.; RINCON, N.; WOOD, D. E.; BREITWIESER, F. P.; POCKRANDT, C.; LANGMEAD, B.; SALZBERG, S. L.; STEINEGGER, M. Metagenome analysis using the Kraken software suite. **Nature protocols**, v. 17, n. 12, p. 2815–2839, dez. 2022. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/36171387>>.

MACEDO, D. B.; DOS ANJOS, T. M. C.; DE LOS SANTOS, E. F. F.; RODRIGUES, M. D. N.; ALEGRIA, O. V. C.; RAMOS, R. T. J. New perspectives on metagenomic analysis for pathogen monitoring in sustainable freshwater aquaculture production: a systematic review. **Frontiers in Freshwater Science**, v. 2, 2 out. 2024. Disponível em: <<https://www.frontiersin.org/articles/10.3389/ffwsc.2024.1459233/full>>.

MACLEAN, R. C.; MILLAN, A. S. The evolution of antibiotic resistance. **Science**, v. 365, n. 6458, p. 1082–1083, 2019.

MILLER, W. R.; ARIAS, C. A. ESKAPE pathogens: antimicrobial resistance, epidemiology, clinical impact and therapeutics. **Nature reviews. Microbiology**, v. 22, n. 10, p. 598–616, out. 2024. Disponível em: <<http://dx.doi.org/10.1038/s41579-024-01054-w>>.

MINOR, N. R.; RAMUTA, M. D.; STAUSS, M. R.; HARWOOD, O. E.; BRAKEFIELD, S. F.; ALBERTS, A.; VUYK, W. C.; BOBHOLZ, M. J.; ROSINSKI, J. R.; WOLF, S.; LUND, M.; MUSSA, M.; BEVERSDORF, L. J.; ALIOTA, M. T.; O'CONNOR, S. L.; O'CONNOR, D. H. Metagenomic sequencing detects human respiratory and enteric viruses in air samples collected from congregate settings. **Scientific reports**, v. 13, n. 1, p. 21398, 4 dez. 2023. Disponível em: <<https://doi.org/10.1038/s41598-023-48352-6>>.

MOSAFFA, F.; SAFFARI, F.; VEISI, M.; TADJROBEHKAR, O. Some virulence genes are associated with antibiotic susceptibility in *Enterobacter cloacae* complex. **BMC infectious diseases**, v. 24, n. 1, p. 711, 19 jul. 2024. Disponível em: <<https://doi.org/10.1186/s12879-024-09608-2>>.

MOYE, Z. D.; WOOLSTON, J.; SULAKVELIDZE, A. Bacteriophage applications for food production and processing. **Viruses**, v. 10, n. 4, p. 1–22, 2018.

MURRAY, C. J.; IKUTA, K. S.; SHARARA, F.; SWETSCHINSKI, L.; ROBLES AGUILAR, G.; GRAY, A.; HAN, C.; BISIGNANO, C.; RAO, P.; WOOL, E.; JOHNSON, S. C.; BROWNE, A. J.; CHIPETA, M. G.; FELL, F.; HACKETT, S.; HAINES-WOODHOUSE, G.; KASHEF HAMADANI, B. H.; KUMARAN, E. A. P.; MCMANIGAL, B.; AGARWAL, R.; AKECH, S.; ALBERTSON, S.; AMUASI, J.; ANDREWS, J.; ARAVKIN, A.; ASHLEY, E.; BAILEY, F.; BAKER, S.; BASNYAT, B.; BEKKER, A.; BENDER, R.; BETHOU, A.; BIELICKI, J.; BOONKASIDECHA, S.; BUKOSIA, J.; CARVALHEIRO, C.; CASTAÑEDA-ORJUELA, C.; CHANSAMOUTH, V.; CHAURASIA, S.; CHIURCHIÙ, S.; CHOWDHURY, F.; COOK, A. J.; COOPER, B.; CRESSEY, T. R.; CRIOLLO-MORA, E.; CUNNINGHAM, M.; DARBOE, S.; DAY, N. P. J.; DE LUCA, M.; DOKOVA, K.; DRAMOWSKI, A.; DUNACHIE, S. J.; ECKMANN, T.; EIBACH, D.; EMAMI, A.; FEASEY, N.; FISHER-PEARSON, N.; FORREST, K.; GARRETT, D.; GASTMEIER, P.; GIREF, A. Z.; GREER, R. C.; GUPTA, V.; HALLER, S.; HASELBECK, A.; HAY, S. I.; HOLM, M.; HOPKINS, S.; IREGBU, K. C.; JACOBS, J.; JAROVSKY, D.; JAVANMARDI, F.; KHORANA, M.; KISSOON, N.; KOBEISSI, E.; KOSTYANEV, T.; KRAPP, F.; KRUMKAMP, R.; KUMAR, A.; KYU, H. H.; LIM, C.; LIMMATHUROTSAKUL, D.; LOFTUS, M. J.; LUNN, M.; MA, J.; MTURI, N.; MUNERA-HUERTAS, T.; MUSICHA, P.; MUSSI-PINHATA, M. M.; NAKAMURA, T.; NANAVATI, R.; NANGIA, S.; NEWTON, P.; NGOUN, C.; NOVOTNEY, A.; NWAKANMA, D.; OBIERO, C. W.; OLIVAS-MARTINEZ, A.; OLLIARO, P.; OOKO, E.; ORTIZ-BRIZUELA, E.; PELEG, A. Y.; PERRONE, C.; PLAKKAL, N.; PONCE-DE-LEON, A.; RAAD, M.; RAMDIN, T.; RIDDELL, A.; ROBERTS, T.; ROBOTHAM, J. V.; ROCA, A.; RUDD, K. E.; RUSSELL, N.; SCHNALL, J.; SCOTT, J. A. G.; SHIVAMALLAPPA, M.; SIFUENTES-OSORNIO, J.; STEENKESTE, N.; STEWARDSON, A. J.; STOEVA, T.; TASAK, N.; THAIPRAKONG, A.;

THWAITES, G.; TURNER, C.; TURNER, P.; VAN DOORN, H. R.; VELAPHI, S.; VONGPRADITH, A.; VU, H.; WALSH, T.; WANER, S.; WANGRANGSIMAKUL, T.; WOZNIAK, T.; ZHENG, P.; SARTORIUS, B.; LOPEZ, A. D.; STERGACHIS, A.; MOORE, C.; DOLECEK, C.; NAGHAVI, M. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. **The Lancet**, v. 399, n. 10325, p. 629–655, 12 fev. 2022. Disponível em: <<http://www.thelancet.com/article/S0140673621027240/fulltext>>. Acesso em: 4 abr. 2023.

NAYFACH, S.; CAMARGO, A. P.; SCHULZ, F.; ELOE-FADROSH, E.; ROUX, S.; KYRPIDES, N. C. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. **Nature Biotechnology**, v. 39, n. 5, p. 578–585, 2021. Disponível em: <<http://dx.doi.org/10.1038/s41587-020-00774-7>>.

NURK, S.; MELESHKO, D.; KOROBAYNIKOV, A.; PEVZNER, P. A. metaSPAdes: a new versatile metagenomic assembler. **Genome research**, v. 27, n. 5, p. 824–834, maio 2017. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/28298430>>.

OLAWADE, D. B.; FAPOHUNDA, O.; EGBON, E.; EBIESUWA, O. A.; USMAN, S. O.; FARONBI, A. O.; FIDELIS, S. C. Phage therapy: A targeted approach to overcoming antibiotic resistance. **Microbial pathogenesis**, v. 197, n. August, p. 107088, dez. 2024. Disponível em: <<https://doi.org/10.1016/j.micpath.2024.107088>>.

PUSHPAKOM, S.; IORIO, F.; EYERS, P. A.; ESCOTT, K. J.; HOPPER, S.; WELLS, A.; DOIG, A.; GUILLIAMS, T.; LATIMER, J.; MCNAMEE, C.; NORRIS, A.; SANSEAU, P.; CAVALLA, D.; PIRMOHAMED, M. Drug repurposing: Progress, challenges and recommendations. **Nature Reviews Drug Discovery**, v. 18, n. 1, p. 41–58, 2018.

REN, J.; SONG, K.; DENG, C.; AHLGREN, N. A.; FUHRMAN, J. A.; LI, Y.; XIE, X.; POPLIN, R.; SUN, F. Identifying viruses from metagenomic data using deep learning. **Quantitative biology (Beijing, China)**, v. 8, n. 1, p. 64–77, mar. 2020. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1007/s40484-019-0187-4>>.

ROMERO-CALLE, D.; GUIMARÃES BENEVIDES, R.; GÓES-NETO, A.; BILLINGTON, C. Bacteriophages as Alternatives to Antibiotics in Clinical Care. **Antibiotics (Basel, Switzerland)**, v. 8, n. 3, p. 1–5, 4 set. 2019. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/31487893>>.

ROUX, S.; CAMARGO, A. P.; COUTINHO, F. H.; DABDOUB, S. M.; DUTILH, B. E.; NAYFACH, S.; TRITT, A. iPHoP: An integrated machine learning framework to maximize host prediction for metagenome-derived viruses of archaea and bacteria. **PLoS biology**, v. 21, n. 4, p. e3002083, abr. 2023. Disponível em: <<http://dx.doi.org/10.1371/journal.pbio.3002083>>.

SANTIAGO-RODRIGUEZ, T. M.; HOLLISTER, E. B. Unraveling the viral dark matter through viral metagenomics. **Frontiers in immunology**, v. 13, n. September, p. 1005107, 2022. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/36189246>>.

SANTOS, M. R. da S.; QUARESMA, M. de N. S. **Análise socioambiental da degradação ao longo do Rio Cereja, Bragança-Pará**. [s.l.] Atena Editora, 2022. 1–23 p.

SBERRO, H.; LEAVITT, A.; KIRO, R.; KOH, E.; PELEG, Y.; QIMRON, U.; SOREK, R. Discovery of functional toxin/antitoxin systems in bacteria by shotgun cloning. **Molecular cell**, v. 50, n. 1, p. 136–48, 11 abr. 2013. Disponível em: <<http://dx.doi.org/10.1016/j.molcel.2013.02.002>>.

SCHLOSS, P. D.; HANDELSMAN, J. Biotechnological prospects from metagenomics. **Current opinion in biotechnology**, v. 14, n. 3, p. 303–10, jun. 2003. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/12849784>>.

SHAW, L. P.; ROCHA, E. P. C.; MACLEAN, R. C. Restriction-modification systems have shaped the evolution and distribution of plasmids across bacteria. **Nucleic acids research**, v. 51, n. 13, p. 6806–6818, 21 jul. 2023. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/37254807>>.

SHOKRALLA, S.; SPALL, J. L.; GIBSON, J. F.; HAJIBABAEI, M. Next-generation sequencing technologies for environmental DNA research. **Molecular ecology**, v. 21, n. 8, p. 1794–805, abr. 2012. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/22486820>>.

SIMMONDS, P.; ADAMS, M. J.; BENKŐ, M.; BREITBART, M.; BRISTER, J. R.; CARSTENS, E. B.; DAVISON, A. J.; DELWART, E.; GORBALENYA, A. E.; HARRACH, B.; HULL, R.; KING, A. M. Q.; KOONIN, E. V.; KRUPOVIC, M.; KUHN, J. H.; LEFKOWITZ, E. J.; NIBERT, M. L.; ORTON, R.; ROOSSINCK, M. J.; SABANADZOVIC, S.; SULLIVAN, M. B.; SUTTLE, C. A.; TESH, R. B.; VAN DER VLUGT, R. A.; VARSANI, A.; ZERBINI, F. M. Consensus statement: Virus taxonomy in the age of metagenomics. **Nature reviews. Microbiology**, v. 15, n. 3, p. 161–168, mar. 2017. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/28134265>>.

SNEPPEN, K.; SEMSEY, S.; SESHASAYEE, A. S. N.; KRISHNA, S. Restriction modification systems as engines of diversity. **Frontiers in Microbiology**, v. 6, n. JUN, p. 1–6, 2 jun. 2015. Disponível em: <<http://journal.frontiersin.org/Article/10.3389/fmicb.2015.00528/abstract>>.

STOKAR-AVIHAIL, A.; FEDORENKO, T.; HÖR, J.; GARB, J.; LEAVITT, A.; MILLMAN, A.; SHULMAN, G.; WOJTANIA, N.; MELAMED, S.; AMITAI, G.; SOREK, R. Discovery of phage determinants that confer sensitivity to bacterial immune systems. **Cell**, v. 186, n. 9, p. 1863–1876.e16, 27 abr. 2023. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/37030292>>.

STRATHDEE, S. A.; HATFULL, G. F.; MUTALIK, V. K.; SCHOOLEY, R. T. Phage therapy: From biological mechanisms to future directions. **Cell**, v. 186, n. 1, p. 17–31, 2023. Disponível em: <<https://doi.org/10.1016/j.cell.2022.11.017>>.

SUTTON, T. D. S.; CLOONEY, A. G.; RYAN, F. J.; ROSS, R. P.; HILL, C. Choice of assembly software has a critical impact on virome characterisation. **Microbiome**, v. 7, n. 1, p. 1–15, 2019.

SZÉKELY, A. J.; BREITBART, M. Single-stranded DNA phages: from early molecular biology tools to recent revolutions in environmental microbiology. **FEMS microbiology letters**, v. 363, n. 6, p. 1–9, mar. 2016. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/26850442>>.

TESSON, F.; HERVÉ, A.; MORDRET, E.; TOUCHON, M.; D'HUMIÈRES, C.; CURY, J.; BERNHEIM, A. Systematic and quantitative view of the antiviral arsenal of prokaryotes. **Nature communications**, v. 13, n. 1, p. 2561, 10 maio 2022. Disponível em: <<https://www.nature.com/articles/s41467-022-30269-9>>.

TORTORELLA, E.; TEDESCO, P.; ESPOSITO, F. P.; JANUARY, G. G.; FANI, R.; JASPARS, M.; DE PASCALE, D. Antibiotics from deep-sea microorganisms: Current discoveries and perspectives. **Marine Drugs**, v. 16, n. 10, p. 1–16, 2018.

VENTOLA, C. L. The antibiotic resistance crisis: part 1: causes and threats. **P & T: a peer-reviewed journal for formulary management**, v. 40, n. 4, p. 277–83, abr. 2015. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/B9781483197111500223>>.

WALSH, F. The multiple roles of antibiotics and antibiotic resistance in nature. **Frontiers in Microbiology**, v. 4, n. AUG, p. 3389, 2013. Disponível em: <<http://journal.frontiersin.org/article/10.3389/fmicb.2013.00255/abstract>>.

WANG, R. H.; YANG, S.; LIU, Z.; ZHANG, Y.; WANG, X.; XU, Z.; WANG, J.; LI, S. C. PhageScope: a well-annotated bacteriophage database with automatic analyses and visualizations. **Nucleic acids research**, v. 52, n. D1, p. D756–D761, 5 jan. 2024a. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/37904614>>.

WANG, Z.; YOU, R.; HAN, H.; LIU, W.; SUN, F.; ZHU, S. Effective binning of metagenomic contigs using contrastive multi-view representation learning. **Nature communications**, v. 15, n. 1, p. 585, 17 jan. 2024b. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/38233391>>.

WEBER-DĄBROWSKA, B.; JOŃCZYK-MATYSIAK, E.; ŻACZEK, M.; ŁOBOCKA, M.; ŁUSIAK-SZELACHOWSKA, M.; GÓRSKI, A. Bacteriophage Procurement for Therapeutic Purposes. **Frontiers in microbiology**, v. 7, n. AUG, p. 1177, 12 ago. 2016. Disponível em: <<http://journal.frontiersin.org/Article/10.3389/fmicb.2016.01177/abstract>>.

WOOD, D. E.; LU, J.; LANGMEAD, B. Improved metagenomic analysis with Kraken 2. **Genome biology**, v. 20, n. 1, p. 257, 2019. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/31779668>>.

WU, Y.; GARUSHYANTS, S. K.; VAN DEN HURK, A.; APARICIO-MALDONADO, C.; KUSHWAHA, S. K.; KING, C. M.; OU, Y.; TODESCHINI, T. C.; CLOKIE, M. R. J.; MILLARD, A. D.; GENÇAY, Y. E.; KOONIN, E. V.; NOBREGA, F. L. Bacterial defense systems exhibit synergistic anti-phage activity. **Cell host & microbe**, v. 32, n. 4, p. 557– 572.e6, 10 abr. 2024. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/38402614>>.

XIAN, Y.; XIAO, C. Current capsid assembly models of icosahedral nucleocytoviricota viruses. **Advances in virus research**, v. 108, p. 275–313, 2020. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/33837719>>.

ZAFAR, H.; SAIER, M. H. Gut Bacteroides species in health and disease. **Gut microbes**, v. 13, n.

1, p. 1–20, 2021. Disponível em: <<https://doi.org/10.1080/19490976.2020.1848158>>.

ZHENG, D.; YIN, G.; LIU, M.; HOU, L.; YANG, Y.; VAN BOECKEL, T. P.; ZHENG, Y.; LI, Y. Global biogeography and projection of soil antibiotic resistance genes. **Science advances**, v. 8, n. 46, p. eabq8015, 18 nov. 2022. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/36383677>>

ZHOU, K.; WONG, T. Y.; LONG, L.; ANANTHARAMAN, K.; ZHANG, W.; WONG, W. C.; ZHANG, R.; QIAN, P.-Y. Genomic and transcriptomic insights into complex virus-prokaryote interactions in marine biofilms. **The ISME journal**, v. 17, n. 12, p. 2303–2312, dez. 2023. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/37875603>>.

## 8. ANEXOS

### PRODUÇÕES

#### Metagenômica

##### 1. Impacts of soybean agriculture on the resistome of the Amazonian soil.

 frontiers | Frontiers in Microbiology

TYPE Original Research  
PUBLISHED 09 September 2022  
DOI 10.3389/fmicb.2022.948188

 Check for updates

#### OPEN ACCESS

EDITED BY  
Xiaolin Hou,  
Beijing University of Agriculture, China

REVIEWED BY  
Rahul Bhadouria,  
University of Delhi, India  
Li Lin,  
Anhui Agricultural University, China

\*CORRESPONDENCE  
Oscar Cardenas Alegria  
oscar.alegria@icb.ufpa.br

<sup>†</sup>These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION  
This article was submitted to Antimicrobials, Resistance and Chemotherapy, a section of the journal Frontiers in Microbiology

RECEIVED 19 May 2022  
ACCEPTED 16 August 2022  
PUBLISHED 09 September 2022

## Impacts of soybean agriculture on the resistome of the Amazonian soil

Oscar Cardenas Alegria<sup>1\*†</sup>, Marielle Pires Quaresma<sup>1†</sup>, Carlos Willian Dias Dantas<sup>2</sup>, Elaine Maria Silva Guedes Lobato<sup>3</sup>, Andressa de Oliveira Aragão<sup>1</sup>, Sandro Patroca da Silva<sup>4</sup>, Amanda Costa Barros da Silva<sup>1</sup>, Ana Cecília Ribeiro Cruz<sup>4</sup>, Rommel Thiago Jucá Ramos<sup>1,2</sup> and Adriana Ribeiro Carneiro<sup>1</sup>

<sup>1</sup>Laboratory of Genomic and Bioinformatics, Center of Genomics and System Biology, Institute of Biological Sciences, Federal University of Pará, Belém, Brazil, <sup>2</sup>Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte, Brazil, <sup>3</sup>Department of Soil Science, Federal Rural University of the Amazon, Paragominas, Brazil, <sup>4</sup>Department of Arbovirology and Hemorrhagic Fevers, Evandro Chagas Institute-IEC/SVS/MS, Ananindeua, Brazil

The soils of the Amazon are complex environments with different organisms cohabiting in continuous adaptation processes; this changes significantly when these environments are modified for the development of agricultural activities that alter the chemical, macro, and microbiological compositions.

##### 2. Influence of soil management for soybean production under microbial diversity in amazon soils


**OPEN ACCESS**

EDITED BY  
 Jehangir Bhadha,  
 University of Florida—Everglades  
 Research & Education Center,  
 United States

REVIEWED BY  
 Ademir Araujo,  
 Federal University of Piauí, Brazil  
 Nan Xu,  
 University of Florida, United States

\*CORRESPONDENCE  
 Oscar Victor Cardenas-Alegria,  
 oscar.alegria@icb.ufpa.br

SPECIALTY SECTION  
 This article was submitted to Soil  
 Processes,  
 a section of the journal  
 Frontiers in Environmental Science

RECEIVED 22 August 2022

ACCEPTED 31 October 2022

PUBLISHED 18 November 2022

# Influence of soil management for soybean production under microbial diversity in amazon soils

Oscar Victor Cardenas-Alegria<sup>1\*</sup>, Carlos Willian Dias Dantas<sup>2</sup>,  
 Elaine Maria Silva Guedes Lobato<sup>3</sup>, Sandro Patroca Da Silva<sup>4</sup>,  
 Cássia de Sousa Moraes<sup>1</sup>, Ana Cecília Ribeiro Cruz<sup>4</sup>,  
 Rommel Thiago Jucá Ramos<sup>1,2</sup> and Adriana Ribeiro Carneiro<sup>1</sup>

<sup>1</sup>Laboratory of Genomic and Bioinformatics, Center of Genomics and System Biology, Institute of Biological Sciences, Federal University of Pará, Belém, Brazil, <sup>2</sup>Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte, Brazil, <sup>3</sup>Federal Rural University of the Amazon, Paragominas, Brazil, <sup>4</sup>Evandro Chagas Institute - IEC/SVS/MS, Department of Arbovirology and Hemorrhagic Fevers, Ananindeua, Brazil

The tropical Amazon has a unique biodiversity that has been affected by the

### 3. Characterization of antibiotic resistance genes in soils from agroecosystems of the Brazilian Amazon



## OPEN ACCESS

EDITED BY  
Yaping Lin,  
Minzu University of China, ChinaREVIEWED BY  
Ajaya Kumar Rout,  
Rani Lakshmi Bai Central Agricultural  
University, India  
Mohamed A. M. Atia,  
Agricultural Research Center, Egypt  
Chenwei Zheng,  
Arizona State University, United States  
Lyudmyla Symochko,  
University of Coimbra, Portugal\*CORRESPONDENCE  
Taynara Cristina Santos Tavares  
✉ narataavares51@gmail.com<sup>†</sup>These authors have contributed equally to  
this work and share first authorshipRECEIVED 08 October 2024  
ACCEPTED 29 April 2025  
PUBLISHED 27 May 2025

# Characterization of antibiotic resistance genes in soils from agroecosystems of the Brazilian Amazon

Taynara Cristina Santos Tavares<sup>1\*†</sup>, Livia Freitas da Silva Pinto<sup>1†</sup>,  
Oscar Victor Cardenas-Alegria<sup>2</sup>, Carlos William Dias Dantas<sup>2</sup>,  
Sandro Patroca da Silva<sup>3</sup>, Ana Cecilia Ribeiro Cruz<sup>3</sup>,  
Anibal Coutinho do Rêgo<sup>4</sup>, Hervé Louis Ghislain Rogez<sup>1</sup>,  
Rommel Thiago Juca Ramos<sup>2</sup>, Cristian Faturi<sup>5</sup> and  
Adriana Ribeiro Carneiro Nunes<sup>1</sup><sup>1</sup>Center for the Valorization of Bioactive Compounds From the Amazon, Federal University of Pará, Belém, Brazil, <sup>2</sup>Simulation and Computational Biology Laboratory (SIMBIC), High Performance Computing Center (CCAD), Federal University of Pará, Belém, PA, Brazil, <sup>3</sup>Department of Arbovirology and Hemorrhagic Fevers, Evandro Chagas Institute—IEC, Ananindeua, Brazil, <sup>4</sup>Department of Animal Science, Federal University of Ceará, Fortaleza, CE, Brazil, <sup>5</sup>Institute of Animal Health and Production, Federal Rural University of Amazonia, Belém, PA, Brazil

## 4. CrAssphage distribution analysis in an Amazonian river based on metagenomic sequencing data and georeferencing

DETAILS
RELATIONS



**Applied and Environmental  
Microbiology**  
Volume 91, Issue 5  
May 2025

ARTICLE

**CrAssphage distribution analysis in an Amazonian river based on metagenomic sequencing data and georeferencing**

[View article page](#)

David Tavares Martins, Oscar Victor Cardenas Alegria, Carlos William Dias Dantas, Edian Fran... [See all authors](#)

**CITE**

Copyright © 2025 Martins et al.  
<https://doi.org/10.1128/aem.01470-24>



ATCC 100 YEARS

Advertisement End

Applied and Environmental  
Microbiology®ENVIRONMENTAL MICROBIOLOGY  
May 2025 Volume 91 Issue 5 e01470-24  
<https://doi.org/10.1128/aem.01470-24>

### CrAssphage distribution analysis in an Amazonian river based on metagenomic sequencing data and georeferencing

David Tavares Martins 1,2,3,4, Oscar Victor Cardenas Alegria 1,2,3,4, Carlos William Dias Dantas 3,4,6, Edian Franklin Franco De Los Santos 6, Paulo Rógenes Monteiro Pontes 7, Rosane Barbosa Lopes Cavalcante 7, Rommel Thiago Juca Ramos 1,2,3,4

<sup>1</sup>Laboratory of Bioinformatics and Genomics of Microorganisms, Federal University of Pará-UFGPA, Belém, Pará, Brazil<sup>2</sup>Institute of Biological Sciences, Federal University of Pará-UFGPA, Belem, Pará, Brazil<sup>3</sup>Laboratory of Simulation and Computational Biology — SIMBIC, Federal University of Pará, Belém, Pará, Brazil<sup>4</sup>Center of High Performance Computer and Artificial Intelligence — CCAD, Federal University of Pará, Belem. Pará. Brazil

## RNA-SEQ E MICROSATELLITES

### 5. RNAseq-differentiated gene expression profile of people living with HTLV-1 in the Brazilian Amazon region: a pilot study



#### OPEN ACCESS

#### EDITED BY

Juliana Echevarria Lima,  
Federal University of Rio de Janeiro, Brazil

#### REVIEWED BY

Feng Jiang,  
Stanford University, United States  
Maria Grazia Romanelli,  
University of Verona, Italy  
Hidekatsu Iha,  
Faculty of Medicine, Oita University, Japan

#### \*CORRESPONDENCE

Antonio Carlos Rosário Vallinoto,  
✉ vallinoto@ufpa.br

RECEIVED 19 April 2024

ACCEPTED 02 August 2024

PUBLISHED 26 August 2024

#### CITATION

Soares Pereira LM, Ferreira RdN, Azevedo VN, Freitas Queiroz MA, Dias Dantas CW, Ramos RTJ, Patroca da Silva S, da Costa CA, Sousa MSd, Vieira Cayres Vallinoto IM, Cruz ACR and Rosário Vallinoto AC (2024) RNAseq-differentiated gene expression profile of people living with HTLV-1 in the Brazilian Amazon region: a pilot study.  
*Front. RNA Res.* 2:1419833.  
doi: 10.3389/fmar.2024.1419833

## RNAseq-differentiated gene expression profile of people living with HTLV-1 in the Brazilian Amazon region: a pilot study

Leonn Mende Soares Pereira<sup>1,2</sup>,  
Rafaella do Nascimento Ferreira<sup>1,2</sup>, Vânia Nakauth Azevedo<sup>1</sup>,  
Maria Alice Freitas Queiroz<sup>1,2</sup>, Carlos Willian Dias Dantas<sup>3</sup>,  
Rommel Thiago Jucá Ramos<sup>3</sup>, Sandro Patroca da Silva<sup>4</sup>,  
Carlos Araújo da Costa<sup>5</sup>, Maisa Silva de Sousa<sup>5</sup>,  
Izaura Maria Vieira Cayres Vallinoto<sup>1,2</sup>, Ana Cecília Ribeiro Cruz<sup>4</sup>  
and Antonio Carlos Rosário Vallinoto<sup>1,2\*</sup>

<sup>1</sup>Laboratório de Virologia, Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, Pará, Brazil, <sup>2</sup>Programa de Pós-graduação em Agentes Infecciosos e Parasitários, Universidade Federal do Pará, Belém, Pará, Brazil, <sup>3</sup>Faculdade de Biotecnologia, Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, Pará, Brazil, <sup>4</sup>Seção de Arbovirologia do Instituto Evandro Chagas, Ananindeua, Pará, Brazil, <sup>5</sup>Núcleo de Medicina Tropical, Universidade Federal do Pará, Belém, Pará, Brazil

In this study, we aimed investigated the differential gene expression profiles of

### 6. EasySSR: a user-friendly web application with full command-line features for large-scale batch microsatellite mining and samples comparison


**OPEN ACCESS**

EDITED BY  
 Nathan Olson,  
 National Institute of Standards and  
 Technology (NIST), United States

REVIEWED BY  
 Safdar Ali,  
 Aliah University, India  
 Nancy Manchanda,  
 Orna Therapeutics, Inc., United States

\*CORRESPONDENCE  
 Rommel Thiago Jucá Ramos,  
 ✉ rommelramos@ufpa.br

<sup>1</sup>PRESENT ADDRESS  
 Carlos Willian Dias Dantas,  
 Laboratory of Biological Engineering,  
 Biological Science Institute, Park of  
 Science and Technology, Federal  
 University of Pará, Belém, Brazil

RECEIVED 25 May 2023  
 ACCEPTED 28 July 2023  
 PUBLISHED 24 August 2023

# EasySSR: a user-friendly web application with full command-line features for large-scale batch microsatellite mining and samples comparison

Sandy Ingrid Aguiar Alves<sup>1</sup>, Victor Benedito Costa Ferreira<sup>1</sup>,  
 Carlos Willian Dias Dantas<sup>2†</sup>, Artur Luiz da Costa da Silva<sup>1</sup> and  
 Rommel Thiago Jucá Ramos<sup>1\*</sup>

<sup>1</sup>Laboratory of Biological Engineering, Biological Science Institute, Park of Science and Technology, Federal University of Pará, Belém, Brazil, <sup>2</sup>Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte, Brazil

Microsatellites, also known as SSRs or STRs, are polymorphic DNA regions with

## 7. What are microsatellites and how to choose the best tool: a user-friendly review of SSR and 74 SSR mining tools



## OPEN ACCESS

EDITED BY  
 Richard R. Rodrigues,  
 Oregon State University, United States

REVIEWED BY  
 Tiyash Parira,  
 Florida International University, United States  
 Kalpana Singh,  
 Guru Angad Dev Veterinary and Animal  
 Sciences University, India  
 Maitreya Das,  
 Sanofi Pasteur, United States

\*CORRESPONDENCE  
 Rommel Thiago Jucá Ramos,  
 ✉ rommelramos@ufpa.br

RECEIVED 01 August 2024  
 ACCEPTED 30 October 2024  
 PUBLISHED 13 November 2024

CITATION  
 Alves SIA, Dantas CWD, Macedo DB and  
 Ramos RTJ (2024) What are microsatellites and  
 how to choose the best tool: a user-friendly  
 review of SSR and 74 SSR mining tools.  
*Front. Genet.* 15:1474611.  
 doi: 10.3389/fgene.2024.1474611

# What are microsatellites and how to choose the best tool: a user-friendly review of SSR and 74 SSR mining tools

Sandy Ingrid Aguiar Alves<sup>1,2,3</sup>, Carlos Willian Dias Dantas<sup>1,2,3</sup>,  
 Daralyns Borges Macedo<sup>2,3</sup> and Rommel Thiago Jucá Ramos<sup>2,3\*</sup>

<sup>1</sup>Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil, <sup>2</sup>Laboratory of Simulation and Computational Biology – SIMBIC, High Performance Computing Center – CCAD, Federal University of Pará, Belém, Pará, Brazil, <sup>3</sup>Laboratory of Bioinformatics and Genomics of Microorganisms, Institute of Biological Sciences, Federal University of Pará, Belém, Pará, Brazil

Microsatellites, also known as SSR or STR, are essential molecular markers in genomic research, playing crucial roles in genetic mapping, population genetics, and evolutionary studies. Their applications range from plant breeding to forensics, highlighting their diverse utility across disciplines. Despite their widespread use, traditional methods for SSR analysis are often laborious and time-consuming, requiring significant resources and expertise. To address these

## 8. Development of microsatellite markers for *Leiarius marmoratus* obtained by partial de novo genome assembly



An Acad Bras Cienc (2025) 97(Suppl. 2): e20241094 DOI 10.1590/0001-3765202520241094  
 Anais da Academia Brasileira de Ciências | *Annals of the Brazilian Academy of Sciences*  
 Printed ISSN 0001-3765 | Online ISSN 1678-2690  
 www.scielo.br/aabc | www.fb.com/aabcjournal

CELLULAR AND MOLECULAR BIOLOGY

### Development of microsatellite markers for *Leiaris marmoratus* obtained by partial *de novo* genome assembly

DÁRCIA GABRIELA B. DA COSTA, DARALYNS B. MACEDO, CARLOS WILLIAN D. DANTAS, REGIANNE M. DOS SANTOS CORREA, ÁDRIA D. DOS SANTOS, HENDRYA JULIANNY P. COELHO, CARLA GIOVANNI Á. MOREIRA, HEDEN LUIZ M. MOREIRA, ROMMEL THIAGO J. RAMOS & MARÍLIA DANYELLE N. RODRIGUES

**Abstract:** This study addresses the identification of repetitive regions called microsatellites in the genome of the Amazonian species *Leiaris marmoratus*. An efficient methodology was used with sequencing and genome assembly to obtain and map specific markers. The programs MISA and EasySSR identified 221,273 and 145,601 markers, respectively. In addition, primers were developed, of which 590 were considered successful, according to in-silico testing with the Batchprime3 program. The results confirm the quality of the methodology, which can be applied in studies of genetic improvement, variability control, and genomic and comparative analyses.

**Key words:** Catfish, Molecular markers, Next-generation sequencing, SSR.



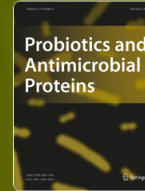
## GENOMICA BACTERIANA

### 9. Probiogenomics of *Leuconostoc Mesenteroides* Strains F-21 and F-22 Isolated from Human Breast Milk Reveal Beneficial Properties

Home &gt; Probiotics and Antimicrobial Proteins &gt; Article


## Probiogenomics of *Leuconostoc Mesenteroides* Strains F-21 and F-22 Isolated from Human Breast Milk Reveal Beneficial Properties


Research | Published: 07 October 2023

Volume 17, pages 500–515, (2025) [Cite this article](#)Download PDF  Access provided by Federal University of Pará 66075-110

Probiotics and Antimicrobial Proteins

[Aims and scope](#) →[Submit manuscript](#) →

Juan Carlos Ariute, Nina Dias Coelho–Rocha, Carlos Willian Dias Dantas, Larissa Amorim Tourinho de Vasconcelos, Rodrigo Profeta, Thiago de Jesus Sousa, Ane de Souza Novaes, Bruno Galotti, Lucas Gabriel Gomes, Enrico Giovanelli Toccani Gimenez, Carlos Diniz, Mariana Vieira Dias, Luís Cláudio Lima de Jesus, Arun Kumar Jaiswal, Sandeep Tiwari, Rodrigo Carvalho, Ana Maria Benko–Iseppon, Bertram Brenig, Vasco Azevedo, Debmalya Barh, Flaviano S. Martins & Flavia Aburjaile 

[Use our pre-submission checklist](#) →Avoid common mistakes on your manuscript. 

Sections

Figures

References


Abstract

## 10. Whole-Genome Sequence Reveals Genetic Determinants of the Colonization of Shiga Toxin-Producing *Escherichia coli* O22:H8 Isolates in Cattle

Home &gt; Current Microbiology &gt; Article

## Whole-Genome Sequence Reveals Genetic Determinants of the Colonization of Shiga Toxin-Producing *Escherichia coli* O22:H8 Isolates in Cattle


Published: 25 March 2025

Volume 82, article number 207, (2025) [Cite this article](#)Download PDF  Access provided by Federal University of Pará 66075-110

Current Microbiology

[Aims and scope](#) →[Submit manuscript](#) →



Wanderson Marques Da Silva , Líbia Yael Smith, Flavia Figueira Aburjaile, Mariano Larzabal, Sandy Ingrid Aguiar Alves, Pablo Farace, Carlos Willian Dantas, Mateus Matiuuzzi Costa, Bertram Brenig, Rommel Thiago Jucá Ramos, Vasco Azevedo & Angel Cataldi

[Use our pre-submission checklist](#) →Avoid common mistakes on your manuscript. 

Sections

Figures

References

 310 Accesses  1 Altmetric [Explore all metrics](#) →

## PRIMEIRA AUTORIA

### 11. SATIN: a micro and mini satellite mining tool of total genome and coding regions with analysis of perfect repeats polymorphism in coding regions

**BMC** Part of Springer Nature

Search  Explore journals Get published About BMC Login

**BMC Bioinformatics**

Home About Articles Submission Guidelines Collections Join The Board [Submit manuscript](#)

Software | [Open access](#) | Published: 18 June 2024

### SATIN: a micro and mini satellite mining tool of total genome and coding regions with analysis of perfect repeats polymorphism in coding regions

[Carlos Willian Dias Dantas](#) , [Sebastião Rodrigues da Costa Neto](#), [Sandy Ingrid Aguiar Alves](#), [Kenny da Costa Pinheiro](#), [Edian Franklin Franco De Los Santos](#) & [Rommel Thiago Jucá Ramos](#)

*BMC Bioinformatics* **25**, Article number: 217 (2024) | [Cite this article](#)

1859 Accesses | [Metrics](#)

**Abstract**

**Background**

Tandem repeats are specific sequences in genomic DNA repeated in tandem that are present in all organisms. Among the subcategories of TRs we have Satellite repeats, that is divided into macrosatellites, minisatellites, and microsatellites, being the last two of specific interest because they can identify polymorphisms between organisms due to their instability.

[Download PDF](#)

**Collection**

[Evolutionary studies](#)

**Sections** **Figures** **References**

[Abstract](#)

[Background](#)

[Implementation](#)

[Tool overview](#)

[Results and discussion](#)

[Conclusion](#)

[Availability and requirements](#)

Availability of data and materials

## ARTIGO DOUTORADO 1

### 12. Tools and methodology to *in silico* phage discovery in freshwater environments

**frontiers** | Frontiers in Microbiology

Sections  Articles Research Topics Editorial board About journal  Submit your research  Search

**MINI REVIEW article**

Front. Microbiol., 30 May 2024  
Sec. Phage Biology  
Volume 15 - 2024 | <https://doi.org/10.3389/fmicb.2024.1390726>

### Tools and methodology to *in silico* phage discovery in freshwater environments

[Carlos Willian Dias Dantas](#)<sup>1,2,3\*</sup> [David Tavares Martins](#)<sup>2,3†</sup> [Wylerson Guimarães Nogueira](#)<sup>1†</sup>

[Oscar Victor Cardenas Alegria](#)<sup>2,3</sup> [Rommel Thiago Jucá Ramos](#)<sup>2,3\*</sup>

<sup>1</sup> Department of Biochemistry and Immunology, Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil  
<sup>2</sup> Laboratory of Simulation and Computational Biology — SIMBIC, High Performance Computing Center — CCAD, Federal University of Pará, Belém, Pará, Brazil  
<sup>3</sup> Laboratory of Bioinformatics and Genomics of Microorganisms, Institute of Biological Sciences, Federal University of Pará, Belém, Pará, Brazil

Freshwater availability is essential, and its maintenance has become an enormous challenge. Due to population growth and climate changes, freshwater sources are becoming scarce, imposing the need for strategies for its reuse. Currently, the constant discharge of waste into water bodies from human activities leads to the dissemination of pathogenic bacteria, negatively impacting water quality from the source to the infrastructure required for treatment, such as the accumulation of biofilms. Current water treatment methods cannot keep pace with bacterial evolution, which increasingly exhibits a profile of multidrug resistance to antibiotics. Furthermore, using more powerful disinfectants may

[Download article](#)

**2,5K** Total views **1,1K** Downloads **4** Citations

[View article impact](#)

[View altmetric score](#)

Share on

**Edited by**

[Marcin Łoś](#)  
University of Gdansk, Poland

**Reviewed by**

[Przemysław Decewicz](#)  
Faculty of Biology, University of Warsaw, Poland