



**UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM DESENVOLVIMENTO
AMBIENTE**



DAVI BEZERRA MELQUIADES

**ABORDAGEM METAGENÔMICA PARA ANÁLISE DO MICROBIOMA E
RESISTOMA EM RIOS DO BIOMA AMAZÔNICO**

JOÃO PESSOA

2025

DAVI BEZERRA MELQUIADES

**ABORDAGEM METAGENÔMICA PARA ANÁLISE DO MICROBIOMA E
RESISTOMA EM RIOS DO BIOMA AMAZÔNICO**

Dissertação de Mestrado apresentada ao
Programa de Desenvolvimento e Meio
Ambiente da Universidade Federal da
Paraíba, para obtenção do título de Mestre.

Orientadora: Dra. Eloiza Helena Campana

Coorientadora: Dra. Bruna de Brito Souza

JOÃO PESSOA

2025

Catálogo na publicação
Seção de Catalogação e Classificação

M517a Melquiades, Davi Bezerra.

Abordagem metagenômica para análise do microbioma e resistoma em rios do bioma amazônico / Davi Bezerra Melquiades. - João Pessoa, 2025.

121 f. : il.

Orientação: Eloiza Helena Campana.

Coorientação: Bruna de Brito Souza.

Dissertação (Mestrado) - UFPB/CCEN.

1. Meio ambiente. 2. Sequenciamentos de DNA. 3. Bioinformática. 4. Matrizes aquáticas. 5. DNA ambiental. I. Campana, Eloiza Helena. II. Souza, Bruna de Brito. III. Título.

UFPB/BC

CDU 502(043)

DAVI BEZERRA MELQUIADES

**ABORDAGEM METAGENÔMICA PARA ANÁLISE DO MICROBIOMA E
RESISTOMA EM RIOS DO BIOMA AMAZÔNICO**

Dissertação apresentada ao Programa Regional de Pós-Graduação em Desenvolvimento e Meio Ambiente – PRODEMA – da Universidade Federal da Paraíba, como parte dos requisitos necessários para obtenção do título de Mestre em Desenvolvimento e Meio Ambiente.

Trabalho Aprovado. João Pessoa, 31 de Outubro de 2025.

BANCA EXAMINADORA

Documento assinado digitalmente
 **ELOIZA HELENA CAMPANA**
Data: 19/12/2025 12:20:26-0300
Verifique em <https://validar.iti.gov.br>

Dra. Eloiza Helena Campana
Universidade Federal da Paraíba

Documento assinado digitalmente
 **BRUNA DE BRITO SOUZA**
Data: 19/12/2025 15:57:14-0300
Verifique em <https://validar.iti.gov.br>

Dra. Bruna de Brito Souza
Instituto Federal de Pernambuco

Documento assinado digitalmente
 **MARILIA GABRIELA DOS SANTOS CAVALCANTI**
Data: 19/12/2025 11:07:23-0300
Verifique em <https://validar.iti.gov.br>

Dra. Marília Gabriela Dos Santos Cavalcanti
(Membro interno)
Universidade Federal da Paraíba

Documento assinado digitalmente
 **ARTHUR WILLIAN DE LIMA BRASIL**
Data: 19/12/2025 10:49:15-0300
Verifique em <https://validar.iti.gov.br>

Dr. Arthur Willian de Lima Brasil
(Membro externo)
Universidade Federal da Paraíba

Dedico aos meus tios Zairan (*in memoriam*) e José Iran (*in memoriam*),
que me ensinaram a nunca desistir dos
sonhos, não importa a idade.

AGRADECIMENTOS

Inicialmente à Natureza, pela sua perfeição;

Aos meus pais, Zairlan (Neguinha) e Cláudio, que se esforçaram para me ajudar desde o primeiro dia que sai de casa para fazer Ciências Biológicas. Meus exemplos de força e dedicação (e teimosia);

Às minhas avós, Maria e Fátima, que carregam por gerações a força da mulher nordestina do sertão e que me ensinaram que o amor verdadeiro vem das mães;

Às minhas tias maternas e paternas, que nunca entendem o que eu falo, mas sempre escutam para me deixar melhor. Em especial às tias Zi e Linda, que me passaram a dedicação e o amor ao lecionar;

Aos meus tios, que embora teimosos, sempre escutam meus conselhos e conversam como amigos. Assim, aprendi a ser um adulto responsável;

Aos meus primos que encham meu saco, mas considero a maioria como meus irmãos mais novos. Às vezes ser o certinho da família tem suas vantagens;

Ao meu companheiro Gustavo, que chegou no meio do mestrado e dividiu todos os meus anseios. Com seu amor, veio para me fortalecer e somar na minha vida;

À professora e orientadora Eloiza Helena, por me aceitar depois de eu a perseguir pela universidade. Me propôs um projeto que nunca tive proximidade e me fez apaixonar pela bioinformática. Tem a fama de ser brava, mas tem um coração gigante. Obrigado por “adotar” seus alunos e compartilhar todos os momentos (inclusive as broncas). Agora sempre que vir uma Gram-negativa, lembrarei da senhora.

À minha coorientadora Bruna Brito, que aceitou o convite de me ajudar e mesmo depois de saber que eu estava perdido na área, me guiou para as ômicas. Agradeço pela parceria e pela amizade que surgiu em meio ao caos nas nossas vidas. Virou meu exemplo de bióloga na saúde e na vida.

Aos professores da banca que aceitaram meu convite para enriquecer meu trabalho: à Profa. Dra. Marília Gabriela Dos Santos Cavalcanti, que me acompanhou desde a seleção, ao Prof. Dr. Arthur Willian de Lima Brasil, ao Prof. Dr. Tássio Brito de Oliveira e à Dra. Sandrelli Meridiana de Fátima Ramos dos Santos Medeiros.

À Fabrine e Sandrielli, que me ensinaram como pesquisar e como aguentar todos os dias a sobrecarga da pós-graduação, assim como pela convivência diária, que se transformou em parceria, os cafés, broncas, fofocas e o mais importante, a amizade.

À equipe do laboratório de Biologia Molecular, que permitiu a troca de conhecimentos e o desenvolvimento como pesquisador. Às meninas de Elô, pelas risadas e desabafos. Às minhas quase orientandas Cris e Ray, que aceitavam o que eu falava. Ao Matheus que me acompanhou em quase toda montanha-russa de ser mestrando. Ao Nicolas, o parceiro da bioinformática. À professora Naiara, a pessoa mais doce e gentil que conheci. À Kelly, uma professora que nasceu pra orientar e vai comigo para o doutorado. Às “*Fabrinetes*” que me ajudavam sempre que podiam.

À equipe do projeto 33 do Amazônia +10 e FAPESQ-PB. Em especial ao Professor Emerson, que me acompanhou durante as coletas no Amapá e sua equipe da UNIFAP. À Jemima e à professora Ingrid que me ajudaram no laboratório e Angleron que me apresentou Macapá e ficou até as 23 horas no laboratório, mesmo cheio de carapanã.

A todos que me ajudaram a desenvolver essa pesquisa, seja com coleta, processamento, sequenciamento ou análises. Eloiza, Bruna, Nicolas, Fabrine, Sandrielli, Naiara, Erika, Rayane, Leandro, Emerson, Jemima, Lídia e Gustavo. Esse trabalho não é só meu, mas da colaboração de vocês.

A todos os meus amigos mais próximos pela descontração e desabafos. Ao João Batista que virou meu irmão e depois de 8 anos morando juntos, ainda parecemos grandes irmãos. À minha gêmula, Thalita, que sempre será minha irmã de outra mãe, assim como sua mãe que me adotou como filho, obrigado pelas conexões eternas. Ao meu recém casal oficial Danilo e Mariana, que sempre estiveram comigo desde o início da graduação. A todos amigos que fiz da biologia 2017.1, tento levar o legado, assim como vocês. E à Joery, minha eterna amiga-vizinha, a antissocial mais social que conheço.

À turma do PRODEMA que me adotou de coração e às minhas panelinhas prediletas. Em João Pessoa, em Aracaju, sempre teremos nossos prodeminhas.

Aos meus amigos do Ceará e de São Paulo, que desde a infância vêm me acompanhando e torcendo.

Às minhas professoras-orientadoras que me ajudaram a construir minha vida acadêmica e de pesquisa. Dra. Tatiane Gadelha, Dra. Leonor Alves, Dra. Ângela Menegatti e Dra. Eloiza Campana.

E para finalizar, agradeço ao PRODEMA e a todos os professores que contribuíram para minha formação. À Universidade Federal da Paraíba e à CAPES, que incentivam os pesquisadores e a perpetuação do conhecimento.

O conhecimento é adquirido não só na universidade, mas principalmente na convivência com a sociedade.

RESUMO

Embora o Brasil possua uma rica biodiversidade, o bioma amazônico e suas matrizes aquáticas vêm sofrendo intensa degradação, sendo expostos a poluentes, incluindo antimicrobianos. Ambientes contaminados por esses compostos podem se tornar locais de seleção para genes de resistência aos antimicrobianos (RAM), que podem ser transferidos a bactérias ambientais ou patogênicas. A RAM já constitui grave problema de saúde pública, associada a mais de 4,95 milhões de mortes anuais, reforçando a necessidade de técnicas interdisciplinares para detecção rápida, como análises metagenômicas aliadas à bioinformática para identificação de genes de resistência. Este trabalho teve como objetivo identificar e analisar o microbioma e o resistoma em matrizes aquáticas amazônicas, além de desenvolver *pipelines* para detecção de reservatórios naturais de RAM. O DNA de amostras provenientes de efluentes de Rios do Estado do Amapá foram isoladas com o kit comercial DNeasy® PowerSoil® e em seguida sequenciadas, por meio da tecnologia *Nanopore*. Posteriormente foram aplicados *pipelines* padronizados para avaliação de qualidade, filtragem e montagem. Para validação dos *pipelines*, foram utilizadas amostras pilotos do Rio Jaguaribe (João Pessoa-PB), devido a boa contiguidade das montagens e ótimos valores de N50 e L50. As análises taxonômicas foram realizadas com Kraken2 e o resistoma caracterizado com diferentes bancos de dados, como CARD, ResFinder e NCBI. Foram elaborados materiais visuais abordando RAM e Uma Só Saúde. A aplicação aos dados amazônicos foi eficiente nas duas estratégias de montagem. Na abordagem com Flye, prevaleceram táxons dominantes como o filo *Pseudomonadota*, com destaque para *Alphaproteobacteria* e *Gammaproteobacteria*, além de *Actinomycetota* e *Thermoproteota* em vários pontos. A montagem com MEGAHIT mostrou maior fragmentação, mas possibilitou detectar táxons abundantes e de baixa cobertura, bem como possíveis patógenos como *Pseudomonas* e *Enterobacterales*. A partir das montagens pelo MEGAHIT nos pontos AP2 e AP9. Foram identificados os genes de resistência *blaOXA-60* e *RbpA*, associados aos gêneros *Ralstonia* e *Mycolicibacterium*, respectivamente. Os *pipelines* desenvolvidos mostraram-se eficazes para análise do microbioma e resistoma de DNA metagenômico ambiental de matrizes aquáticas, e a detecção desses genes reforça a importância da vigilância em biomas sob degradação.

Palavras-chave: Meio ambiente; Sequenciamento; Bioinformática; Matrizes aquáticas; DNA ambiental.

ABSTRACT

Although Brazil has a rich biodiversity, the Amazon biome and its aquatic matrices have been undergoing intense degradation, being exposed to pollutants, including antimicrobials. Environments contaminated by these compounds may become sites for the selection of antimicrobial resistance (AMR) genes, which can be transferred to environmental or pathogenic bacteria. AMR already constitutes a serious public health problem, being associated with more than 4.95 million deaths annually, reinforcing the need for interdisciplinary techniques for rapid detection, such as metagenomic analyses combined with bioinformatics for the identification of resistance genes. This study aimed to identify and analyze the microbiome and the resistome in Amazonian aquatic matrices, as well as to develop pipelines for the detection of natural AMR reservoirs. DNA from samples originating from effluents of rivers in the state of Amapá was isolated using the commercial DNeasy® PowerSoil® kit and then sequenced using Nanopore technology. Subsequently, standardized pipelines were applied for quality assessment, filtering, and assembly. For pipeline validation, pilot samples from the Jaguaribe River (João Pessoa, PB) were used, due to the good contiguity of the assemblies and excellent N50 and L50 values. Taxonomic analyses were performed using Kraken2, and the resistome was characterized with different databases, such as CARD, ResFinder, and NCBI. Visual materials addressing AMR and One Health were developed. The application to Amazonian data was efficient in both assembly strategies. In the Flye approach, dominant taxa such as the phylum *Pseudomonadota* prevailed, with emphasis on *Alphaproteobacteria* and *Gammaproteobacteria*, in addition to *Actinomycetota* and *Thermoproteota* at several sites. The MEGAHIT assembly showed greater fragmentation but enabled the detection of abundant and low-coverage taxa, as well as possible pathogens such as *Pseudomonas* and *Enterobacterales*. From the MEGAHIT assemblies at sites AP2 and AP9, the resistance genes *bla*OXA-60 and *RbpA* were identified, associated with the genera *Ralstonia* and *Mycolicibacterium*, respectively. The developed pipelines proved to be effective for the analysis of the microbiome and resistome of environmental metagenomic DNA from aquatic matrices, and the detection of these genes reinforces the importance of surveillance in biomes under degradation.

Key-words: Environment; Sequencing; Bioinformatics; Aquatic matrices; Environmental DNA.

LISTA DE ILUSTRAÇÕES

- Figura 1:** Regiões de coletas do Rio Jaguaribe, João Pessoa.
- Figura 2:** Mapa do território do Amapá e regiões de coletas.
- Figura 03:** Qualidade por sequência base de PJ1.
- Figura 04:** Qualidade por sequência base de PJ2.
- Figura 05:** Qualidade por sequência base de PJ3.
- Figura 07:** Classificação taxonômica PJ1 montado pelo Flye.
- Figura 08:** Classificação taxonômica PJ2 montado pelo Flye.
- Figura 09:** Classificação taxonômica PJ3 montado pelo Flye.
- Figura 10:** Qualidade por sequência base de AP1.
- Figura 11:** Qualidade por sequência base de AP2.
- Figura 12:** Qualidade por sequência base de AP3.
- Figura 13:** Qualidade por sequência base de AP4.
- Figura 14:** Qualidade por sequência base de AP5.
- Figura 15:** Qualidade por sequência base de AP6.
- Figura 16:** Qualidade por sequência base de AP7.
- Figura 17:** Qualidade por sequência base de AP8.
- Figura 18:** Qualidade por sequência base de AP9.
- Figura 19:** Qualidade por sequência base de AP10.
- Figura 20:** Qualidade por sequência base de AP11.
- Figura 21:** Classificação taxonômica da amostra AP1 pelo Flye.
- Figura 22:** Classificação taxonômica da amostra AP2 pelo Flye.
- Figura 23:** Classificação taxonômica da amostra AP3 pelo Flye.
- Figura 24:** Classificação taxonômica da amostra AP4 pelo Flye.
- Figura 25:** Classificação taxonômica da amostra AP6 pelo Flye.
- Figura 26:** Classificação taxonômica da amostra AP7 pelo Flye.
- Figura 27:** Classificação taxonômica da amostra AP8 pelo Flye.
- Figura 28:** Classificação taxonômica da amostra AP9 pelo Flye.
- Figura 29:** Classificação taxonômica da amostra AP10 pelo Flye.
- Figura 30:** Classificação taxonômica da amostra AP11 pelo Flye.
- Figura 31:** Principais classificação taxonômica da amostra AP1 pelo MEGAHIT.
- Figura 32:** Principais classificação taxonômica da amostra AP2 pelo MEGAHIT.
- Figura 33:** Principais classificação taxonômica da amostra AP3 pelo MEGAHIT.
- Figura 34:** Principais classificação taxonômica da amostra AP4 pelo MEGAHIT.
- Figura 35:** Principais classificação taxonômica da amostra AP6 pelo MEGAHIT.
- Figura 36:** Principais classificação taxonômica da amostra AP7 pelo MEGAHIT.
- Figura 37:** Principais classificação taxonômica da amostra AP8 pelo MEGAHIT.

Figura 38: Principais classificação taxonômica da amostra AP9 pelo MEGAHIT.

Figura 39: Principais classificação taxonômica da amostra AP10 pelo MEGAHIT.

Figura 40: Principais classificação taxonômica da amostra AP11 pelo MEGAHIT.

LISTA DE TABELAS E QUADROS

Tabela 01: Caracterização geográfica e físico-química dos pontos de coleta do Amapá.

Tabela 02: Parâmetros de quantificação e pureza do DNA metagenômico extraído das amostras piloto.

Tabela 03: Análise da qualidade do sequenciamento das amostras piloto de saída bruto pelo Nanoplot.

Tabela 04: Análise da qualidade do sequenciamento das amostras piloto de saída após processamento e filtragem pelo Nanoplot.

Tabela 05: Métricas de montagem das amostras piloto pelo Flye.

Tabela 06: Métricas de montagem das amostras piloto pelo MEGAHIT.

Tabela 07: Resistoma do ponto PJ1.

Tabela 08: BLAST NCBI das *contigs* e perfil de resistência dos genes do ponto PJ1.

Tabela 09: Resistoma do ponto PJ2.

Tabela 10: BLAST NCBI das *contigs* e perfil de resistência dos genes da amostra PJ2.

Tabela 11: Resistoma do ponto PJ3.

Tabela 12: BLAST NCBI das *contigs* e perfil de resistência dos genes do ponto PJ3.

Tabela 13: Extração de DNA.

Tabela 14: Análise da qualidade do sequenciamento de saída bruto.

Tabela 15: Análise da qualidade do sequenciamento de saída pós-processamento e filtragem.

Tabela 16: Métricas de montagem Flye.

Tabela 17: Métricas de montagem MEGAHIT.

Tabela 18: Análise do resistoma.

Tabela 19: Análise dos resistomas das amostras AP2 e AP9.

Tabela 20: BLAST NCBI das *contigs* e perfil de resistência dos genes do ponto AP2 e AP9.

LISTA DE ABREVIATURAS E SIGLAS

% – Porcentagem
> – Maior que
AP - Amapá
b – bases nitrogenadas
BIOMOL – Laboratório de Microbiologia e Imunologia
CBiotec – Centro de Biotecnologia
CCM – Centro de Ciências Médicas
CCS – Centro de Ciências da Saúde
cm – centímetros
DCF – Departamento de Ciências Farmacêuticas
DNA – Ácido desoxirribonucleico
EDTA – Ácido etilenodiamino tetra-acético
FAO – Organização das Nações Unidas para Alimentação e Agricultura
g – Gramas
GB - Gigabases
GC – Guanina e Citosina
GPS – Sistema de posicionamento global
GTDB – Banco de Dados de Taxonomia do Genoma
H⁺ – Íon hidrogênio
kb – quilobases
km – Quilômetro
km² – Quilômetro quadrado
L – Litros
LaBiMol – Laboratório de biologia molecular
LAM – Laboratório Analítico Multiusuário
LBCM - Laboratório de Biotecnologia Celular e Molecular
LIMFA – Laboratório de Imunofarmacologia
Mb – Megabases
mL – Mililitros
mm – Milímetros
N – Base indeterminada
N – Norte
NCBI – Centro Nacional de Informação Biotecnológica
ng – Nanograma
NGS – Sequenciamento de nova geração
NUMETROP – Núcleo de Medicina Tropical
O – Oeste
°C – Grau Celsius
ODS – Objetivos de Desenvolvimento Sustentável
OMS – Organização Mundial de Saúde
ONU – Organização das Nações Unidas
pb – Pares de base

PCR – Reação em cadeia da polimerase
pH – Potencial hidrogeniônico
PNUMA – Programa das Nações Unidas para o Meio Ambiente
PRODES – Projeto de Monitoramento do Desmatamento na Amazônia Brasileira por Satélite
RAM – Resistência aos antimicrobianos
RPM – Rotações por minuto
S – Sul
SAD – Sistema de Alerta de Desmatamento
STE – Tris, Saline, EDTA
SUDAM – Superintendência de Desenvolvimento da Amazônia
UFPB – Universidade Federal da Paraíba
UNIFAP – Universidade Federal do Amapá
UV – Ultravioleta
WOAH – Organização Mundial da Saúde Animal
 μm – Micrómetro

SUMÁRIO

Sumário

1. INTRODUÇÃO	17
2. HIPÓTESE	20
3. OBJETIVOS	20
3.1. Objetivo Geral	20
3.2. Objetivos Específicos.....	20
4. FUNDAMENTAÇÃO TEÓRICA.....	20
4.1. Bioma e Degradação	20
4.2. Matrizes Aquáticas	21
4.3. Doenças Zoonóticas.....	22
4.4. Uso de antimicrobianos no ambiente.....	23
4.5. Resistência aos antimicrobianos	24
4.6. <i>One Health</i>	26
4.7. Tecnologias de sequenciamentos de DNA	26
4.8. Metagenômica.....	29
5. MATERIAL E MÉTODOS.....	29
5.1. Estudo Piloto: Padronização das Coletas e Análises de Bioinformática	29
5.2. Coleta das amostras.....	30
5.3. Filtragem das amostras.....	34
5.4. Extração do DNA genômico.....	34
5.5. Preparação da biblioteca Nanopore e sequenciamento	35
5.6. Análise de bioinformática	35
5.6.1. Microbioma	35
5.6.2. Resistoma	36
5.7. Devolutiva social	36
5.8. Divulgação acadêmica e científica	36
5.9. Colaboração do projeto.....	37
6. RESULTADOS E DISCUSSÃO.....	37
6.1. Amostras Piloto	37
6.1.1. Quantificação das Amostras Piloto e Padronização dos <i>Pipelines</i>	37
6.1.2. Controle de Qualidade e Filtragem	38
6.1.3. Montagem e avaliação de qualidade.....	44
6.1.4. Microbioma da amostra piloto.....	46
6.1.5. Análise do resistoma	50

6.2. Amostras das regiões do Amapá	62
6.2.1. Extração, quantificação e avaliação da pureza do DNA metagenômico.....	62
6.2.2. Controle de Qualidade e Processamento dos Dados Brutos.....	64
6.2.3. Montagem	80
6.2.4. Microbioma Flye	83
6.2.5. Microbioma MEGAHIT	91
6.2.6. Montagem Flye vs MEGAHIT	105
6.2.7. Resistoma	106
7. CONCLUSÃO	111
8. PERSPECTIVAS FUTURAS	111
REFERÊNCIAS	113
APÊNDICE	122

1. INTRODUÇÃO

O Brasil possui uma diversidade de biomas, entre eles a maior floresta tropical do mundo, o bioma amazônico, que corresponde a mais da metade do território nacional. Porém, a floresta amazônica vem sofrendo com desmatamento e super exploração. Em 2022 o Sistema de Alerta de Desmatamento (SAD) divulgou um dos maiores registros de perda florestal do bioma, sendo devastados 10.573 km² de área (Fonseca *et al.*, 2022). A perda de cobertura vegetal altera a dinâmica do ecossistema e expõe agentes infecciosos que poderiam ser retidos no solo, além de perturbar os habitats e danificar matrizes aquáticas, que se tornam mais susceptíveis à poluição. Esses fatores se agravam em regiões de vulnerabilidade social e ecológica, pois amplificam os riscos de doenças infecciosas zoonóticas (Winck *et al.*, 2022).

As matrizes aquáticas recebem poluentes diariamente e, com a perda florestal, ocorre um aumento excessivo nessa carga. Dentre os poluentes, destacam-se antimicrobianos que apresentam um risco para o microbioma local e produzem efeitos prejudiciais à saúde humana e ambiental (Branco *et al.*, 2021). O uso de antimicrobianos vai além de ambientes clínicos, sendo também utilizados na agricultura, pecuária e aquicultura e podem ser transportados para matrizes aquáticas por dejetos, poluição ou escoamento de água (Despotovic *et al.*, 2023).

A propagação de antimicrobianos em matrizes aquáticas impulsiona o surgimento de microrganismos resistentes por pressão seletiva, pois, quanto mais os indivíduos são expostos aos antimicrobianos, maior é a probabilidade do desenvolvimento da resistência aos antimicrobianos (RAM) (Merlino; Siarakas, 2022). Microrganismos resistentes presentes nesses ambientes podem causar prejuízos à saúde das pessoas de forma direta, com o uso da água do local para consumo ou recreação, através de alimentos contaminados, ou de forma indireta, por meio de animais contaminados que estão em contato com humanos (Castro; De Oliveira, 2021).

A RAM vem causando um impacto grave para a saúde pública, sendo considerada prioridade para a Organização Mundial de Saúde (OMS), estima-se que houve 4,95 milhões de mortes no mundo associadas a RAM em 2019 e que 1,27 milhão de mortes foram causadas diretamente por bactérias

multirresistentes (Murray *et al.*, 2022; WHO, 2022). O crescente aparecimento de microrganismos resistentes aos antimicrobianos causa preocupação, pois diminui as chances de tratamentos eficazes, aumentando a morbidade e letalidade destas infecções e, consequentemente, os custos na saúde (Chen; Kumar; Wu, 2023).

Uma forma de combater a escalada da RAM foi adotada pelo Programa das Nações Unidas para o Meio Ambiente (PNUMA), utilizando a abordagem que expande a compreensão da RAM e aspectos envolvidos, conhecida como Uma Só Saúde (*One Health*) e representa uma abordagem interdisciplinar que atua como prevenção e vigilância no tratamento e controle de doenças, conectando a saúde humana, animal e ambiental e quando direcionadas para microbiomas de matrizes aquáticas abordam questões significativas na disseminação da RAM (Despotovic *et al.*, 2023; UNEP, 2023). A dinâmica entre a tríade *One Health* demonstra que bactérias resistentes não são exclusivas de um ambiente, sendo necessário a vigilância constante em todas as esferas. A RAM é reconhecida pela Organização das Nações Unidas (ONU) como um fator emergente para zoonose, principalmente em países subdesenvolvidos, uma vez que a relação entre homem e animal é mais próxima e a infraestrutura sanitária e biossegurança são insuficientes (Olaru; Walther; Schaumburg, 2023).

Atualmente, com os avanços nas tecnologias de sequenciamento genético de próxima geração juntamente com o auxílio da bioinformática para análise da grande quantidade de dados gerados provenientes nesse processo, é possível sequenciar e elucidar o genoma de todos os microrganismos presentes em uma amostra. A metagenômica é uma abordagem vantajosa, pois permite sequenciar os genomas de todos os microrganismos cultiváveis e não cultiváveis de maneira mais rápida e eficiente. O estudo de metagenômica *shotgun* fragmenta e sequencia o DNA completo, proporcionando maior cobertura genômica de diversos indivíduos. Assim, por meio dessa abordagem e com auxílio da bioinformática, é possível analisar todo microbioma de um determinado local e identificar genes de resistência nas amostras avaliadas (Akacin *et al.*, 2022; Wensel *et al.*, 2022; Quince *et al.*, 2017).

Visto que os rios atuam como veículos de transmissão, a investigação do microbioma e resistoma desses ambientes é fundamental para prevenir e monitorar infecções por bactérias multirresistentes, já que esses microrganismos

estão associados a altas taxas de mortalidade em todo o mundo, porém, há poucas pesquisas que investigam os efeitos dos impactos da RAM em rios de biomas complexos, como a Amazônia.

Assim, o presente projeto visa analisar o microbioma e resistoma de um rio do bioma amazônico, através da abordagem de metagenômica e análises dos dados por bioinformática, identificando o microbioma e detectando possíveis patógenos e genes de resistência aos antimicrobianos presentes no ambiente que está sendo degradado.

Segundo Domingues (2005), experiências interdisciplinares compartilham conhecimentos de diferentes áreas para solucionar problemas. Este trabalho exemplifica essa interdisciplinaridade ao correlacionar degradação ambiental com riscos zoonóticos, utilizando técnicas de biologia molecular, como microbiologia e genética, para investigar o microbioma e resistoma por meio da metagenômica e bioinformática. Além disso, há uma devolutiva social, com a criação de materiais visuais para as comunidades afetadas e divulgação científica em workshops e congressos.

A pesquisa sobre resistência microbiana abrange vários dos 17 Objetivos de Desenvolvimento Sustentável (ODS). Ela impacta: i) ODS 01, pois países de baixa renda são mais vulneráveis a infecções e têm menor acesso a tratamentos eficazes, afetando a economia e exigindo estratégias financeiras para garantir a saúde de populações carentes; ii) ODS 02, devido ao impacto da RAM nos custos de saúde animal e agropecuária; iii) ODS 03, com o aumento de mortes anuais por RAM, demandando novos medicamentos caros e inacessíveis; iv) ODS 06, já que a falta de tratamento adequado e saneamento transforma as águas em veículos de infecções e microrganismos resistentes; v) ODS 08, pois a RAM ameaça a economia ao aumentar o número de mortes; vi) ODS 12, com o descarte inadequado de antimicrobianos poluindo rios, exigindo controle eficaz; e vii) ODS 17, com a abordagem *One Health* atuando em múltiplos níveis para combater a RAM. A investigação da RAM é crucial para o progresso dos ODS, já que a disseminação de cepas resistentes pode dificultar seu alcance (WHO, 2021).

2. HIPÓTESE

A degradação do bioma amazônico influencia a composição do microbioma e do resistoma dos rios estudados, podendo acarretar desequilíbrios ecológicos e riscos à saúde das comunidades locais.

3. OBJETIVOS

3.1. Objetivo Geral

Analisar o microbioma e resistoma em afluentes de rios em diferentes regiões do estado do Amapá, a fim de identificar a diversidade de bactérias potencialmente de risco e detectar os possíveis reservatórios naturais de genes de resistência aos antimicrobianos.

3.2. Objetivos Específicos

- Identificar bactérias com potencial patogênico e avaliar o resistoma presentes em corpos hídricos do bioma amazônico;
- Analisar os elementos genéticos móveis associados a genes de resistência de importância clínica por meio de ferramentas de bioinformática em cada localidade;
- Identificar possíveis reservatórios naturais de genes de resistência aos antimicrobianos em corpos hídricos do bioma amazônico.
- Desenvolver *pipelines* com recursos de bioinformática para análise do microbioma e resistoma bacterianos;

4. FUNDAMENTAÇÃO TEÓRICA

4.1. Bioma e Degradação

O bioma Amazônia, representa a maior cobertura vegetal do Brasil, abrangendo também a Bolívia, Colômbia, Equador, Guiana, Guiana Francesa, Peru, Suriname e Venezuela. É a maior floresta tropical do mundo, equivalente a 1/3 das reservas florestais tropicais úmidas e assim, fornece diversos serviços ecossistêmicos, como a disponibilidade de 20% de água doce do planeta,

mitigação dos efeitos climáticos e a maior biodiversidade animal e vegetal. A Amazônia possui a maior bacia hidrográfica, com área de 7 milhões de km² abrangendo diversos países da América do Sul, cujo principal rio da bacia é denominado de Amazonas, com extensão de aproximadamente 6.992 km, nascendo na Cordilheira dos Andes, no Peru e desaguando na ilha de Marajó, no estado do Pará, Brasil (Prado, 2021).

A Amazônia Legal, corresponde à floresta dentro do território brasileiro, delimitado geograficamente para atuação política da Superintendência de Desenvolvimento da Amazônia (SUDAM), representa aproximadamente 59% do território brasileiro e engloba os estados do Acre, Roraima, Amazonas, Rondônia, Pará, Amapá, Tocantins, Maranhão e Mato Grosso (IBGE, 2022).

Embora a Amazônia seja a maior floresta tropical, ela enfrenta diversas formas de degradação, principalmente devido à expansão das atividades agropecuárias e atividades ilegais. Entre os anos de 2019-2022 o Projeto de Monitoramento do Desmatamento na Amazônia Brasileira por Satélite (PRODES) registrou a média de desmatamento de 11.396,5 km² de área, considerado alta em comparação com as taxas dos anos anteriores, sendo entre 2015-2018 a média do desmatamento de 7.145,75 km² de área. O último registro de área desmatada acima de 10.000 km², foi em 2008 com taxas de 12.911 km² de área (Brasil, 2022).

As principais causas da degradação do bioma amazônico são atividades pecuárias e agrícolas que exigem grandes áreas para implementação, gerando uma faixa de desmatamento entre propriedades (Rossoni; De Moraes, 2020). Outros fatores que favorecem o desmatamento do bioma são atividades ilegais, como apropriação ilegal de terras, incêndios intencionais e mineração (Bronz, 2023; Assunção; Gandour; Rocha, 2023; Menezes, 2022).

4.2. Matrizes Aquáticas

Além do solo, as matrizes aquáticas do bioma são afetadas com a degradação ambiental, pois por perder sua cobertura vegetal, parte do dano no solo é escoado ou lixiviado para os rios. A presença de florestas próximo aos rios e lagos reduz a carga de poluição, pois atuam como interceptoras e filtradoras, melhorando a qualidade da água (Cecílio; Pimentel; Zanetti, 2019).

Próximo à matriz aquática, de modo natural, encontra-se uma porção vegetal denominada de mata ciliar que atua como obstáculo natural no escoamento de água, minimizando os impactos, evitando a erosão, diminuindo a evaporação e mantendo o equilíbrio do meio. Em consequência da sua degradação, afeta a qualidade da água, aumenta os assoreamentos e compactação do solo (Tolkinen; Vaarala; Aroviita, 2021; Correia *et al.*, 2019). Essas matrizes aquáticas sem mata ciliar, são mais susceptíveis a receberem poluentes e dejetos de animais e assim alterar a turbidez da água, pH e temperatura, que diminuem os efeitos bactericidas da luz UV (Ferreira *et al.*, 2022).

Ademais, os rios possuem alta conectividade com outros ambientes, consequentemente, são afetados por atividades antrópicas, recebem resíduos domésticos, industriais e de outros rios. As matrizes aquáticas são utilizadas diariamente pela população e outros animais, sendo um potencial veículo de transmissão de patógenos. Comunidades de baixa renda e ribeirinhos são os mais suscetíveis a contaminações provenientes de rios contaminados (Porcy *et al.*, 2020; Medeiros; Lima; Guimarães, 2016).

4.3. Doenças Zoonóticas

A degradação ambiental está associada ao aumento de doenças zoonóticas, pois perturbam o equilíbrio dos ecossistemas. A fragmentação das florestas, conversão florestal e consequentemente a degradação de matrizes aquáticas, afetam a dinâmica de transmissão de doenças zoonóticas aumentando o contato humano com vetores ou reservatórios (Morand; Lajaunie, 2021).

Os patógenos zoonóticos são organismos capazes de causar doenças transmitidas entre animais e seres humanos, como bactérias, vírus, fungos e parasitas. A transmissão pode ocorrer de forma direta (mordidas, contato) ou indireta (alimentos, água ou solos contaminados). Segundo a OMS, mais de 60% dos microrganismos patogênicos que podem afetar humanos são considerados zoonóticos. Diversos animais podem atuar como hospedeiros, sendo reservatórios naturais e que exercem um papel central na vida dos patógenos, que acabam infectando outros animais, como os humanos. Muitos desses

agentes infecciosos que estão disseminados entre humanos, tiveram origens zoonóticas (Rahman, 2020; WHO, 2020).

Animais hospedeiros, conhecidos como reservatórios naturais, exercem papel central no ciclo de vida desses patógenos, sendo que muitos agentes infecciosos hoje disseminados entre humanos tiveram origem zoonótica.

Segundo o relatório fronteiras, emitido em 2016 pelo PNUMA sobre questões emergentes de preocupação ambiental, cinco fatores principais favorecem o surgimento e disseminação de doenças zoonóticas, fatores como o desmatamento e mudanças na utilização do solo, comércio irregular ou ilegal de animais silvestres, resistência a antimicrobianos, expansão das atividades agropecuárias em áreas anteriormente preservadas e as mudanças climáticas (UNEP, 2016).

Alguns fatores apresentados impactam diretamente as matrizes aquáticas, como o desmatamento e mudanças na utilização do solo próximos a essas matrizes que aumentam a lixiviação de carga de poluentes e dejetos, o qual pode ocorrer a existência de antimicrobianos ou microrganismos resistentes, ocasionando a proliferação de bactérias resistentes aos antimicrobianos (Reddy *et al.*, 2022).

4.4. Uso de antimicrobianos no ambiente

A ocorrência de antimicrobianos difundidos em matrizes aquáticas, deve-se ao resultado do uso tanto humano, como em ambientes clínicos e hospitalares, quanto veterinário e pecuário, agrícola e aquícola que se torna um fator crucial para a seleção e propagação de RAM. O descarte inadequado desses compostos, que pode ocorrer de maneira direta ou indireta, promove ambientes o qual microrganismos poderão ser expostos a concentrações subinibitórias de antimicrobianos, induzindo ou acelerando a seleção natural por pressão seletiva e adaptação, ocasionando mutações nos indivíduos sobreviventes que adquirem resistência ao agente causador (Sánchez-Baena; Caicedo-Bejarano; Chávez-Vivas, 2021; Zhuang *et al.*, 2021).

Existe uma diversidade de antimicrobianos utilizado para tratar infecções bacterianas, como a isoniazida que é utilizada em terapias para tuberculose ou como os flavofosfolipol e ionóforos que são aplicados para ganho produtivo e

uso profilático em suinocultura, avicultura e bovinocultura (Wesguerber *et al.*, 2024; Aslam *et al.*, 2021). Além disso, existem antimicrobianos que são empregados na agricultura como tetraciclinas e estreptomicinas como usos profiláticos para evitar infecções em frutos (Vidaver, 2002).

Cerca de 80% dos antimicrobianos comercializados globalmente são destinados à produção animal. Alguns possuem usos comuns como a cefalosporina de terceira geração e a colistina que são as mais utilizadas tanto na saúde humana, quanto animal, podendo gerar assim, um aumento na resistência por uso indiscriminado (Wesguerber *et al.*, 2024; Aslam *et al.*, 2021).

4.5. Resistência aos antimicrobianos

As bactérias possuem mecanismos específicos que conferem resistência aos antimicrobianos. Um desses mecanismos é a degradação ou inativação enzimática do fármaco, em que as bactérias produzem, principalmente, a enzima betalactamase, que hidrolisa os anéis betalactâmicos de antibióticos, como as penicilinas e cefalosporinas. Outro mecanismo envolve a modificação do alvo do fármaco, onde alterações no sítio ativo do ribossomo impedem que os antibióticos reconheçam suas moléculas-alvo (Mobarki; Almerabi; Hattan, 2019; Tortora; Funke; Case, 2017).

Além disso, a permeabilidade reduzida do fármaco pode ocorrer quando a bactéria modifica os canais proteicos, como as porinas, reduzindo o fluxo de antibióticos para o interior das células. Bactérias gram-negativas são relativamente mais resistentes devido à presença de uma membrana externa, que, quando modificada, impede a passagem de antibióticos para o periplasma da célula. Outro mecanismo é a exportação ativa do fármaco, em que proteínas presentes na membrana plasmática das bactérias gram-negativas utilizam bombas de efluxo para importar prótons e expelir outras moléculas. Isso impede a concentração efetiva do fármaco dentro da célula, dificultando a ação antimicrobiana (Munita; Arias, 2016).

Os genes de resistência são adquiridos através de mutações mediadas a partir de cromossomos, plasmídeos ou transposons, esses genes podem ser transmitidos para outros indivíduos através do mecanismo de transferência horizontal de genes (Arnold; Huang; Hanage, 2022).

As informações genéticas podem ser transmitidas para outras bactérias de três formas: transformação, em que bactérias incorporam genes de resistência presentes no ambiente deixados por outras bactérias que sofreram morte celular; transdução que ocorre quando bacteriófagos transferem genes de resistência de uma bactéria para outra; e conjugação, que envolve a recombinação do cromossomo ou plasmídeos de uma célula doadora com fator positivo à outra célula receptora com fator negativo, resultando na transferência de novos genes que podem ser transmitidos às células-filhas (Scaldaferri *et al.*, 2020; Tortora; Funke; Case, 2017).

Embora não seja um mecanismo de resistência, é a resposta ao estresse do meio, com a produção de biofilme comumente visto em bactérias gram-negativas, uma substância polimérica excretada no meio que possuem polissacarídeos, DNA extracelular, lipídios, surfactantes, flagelos, proteínas e pili que serve de proteção e dificulta a penetração de antibióticos e favorece o crescimento das colônias de bactérias (Abrantes; Nogueira, 2022).

A RAM é um problema mundial para a saúde pública e que está ganhando destaque atualmente. A OMS emitiu um alerta sobre a gravidade da resistência, detalhando fatos importantes, dentre elas que a RAM causou 1,27 milhão de mortes no mundo em 2019, que o uso excessivo de antimicrobianos em humanos, animais e plantas são os principais responsáveis pelo impulso do desenvolvimento da RAM, além de reforçar que esse problema afeta principalmente países subdesenvolvidos e em desenvolvimento (WHO, 2023).

Além disso, estima-se que se a RAM não for controlada, ocorrerão até 10 milhões de mortes anualmente a partir de 2050 levando a prejuízos econômicos significativos, entre 1 e 3,4 bilhões de dólares por ano, que colocará mais de 24 milhões de pessoas na extrema pobreza (UNEP, 2023).

As bactérias multirresistentes são um dos principais focos das discussões em saúde pública, sendo responsáveis pelo aumento da morbidade e mortalidade. Estima-se que elas causem cerca de 700 mil mortes anuais no mundo devido a infecções, das quais mais de 20 mil ocorrem no Brasil (O'Neill *et al.*, 2016).

4.6. One Health

No entanto, nesses últimos anos, esse problema não se deteve somente aos hospitais e ambientes clínicos, pois os antimicrobianos, bactérias e genes de resistência têm sido detectados em diferentes ambientes como matrizes aquáticas, transportados por dejetos de animais, erosão do solo ou esgotos, promovendo o compartilhamento e disseminação de resistência entre bactérias ambientais. Essa complexa dinâmica de disseminação da RAM é abordada pelo conceito “*One Health*” que relaciona conceitos epidemiológicos moleculares em humanos, animais e o meio ambiente e quando direcionada para a resistência aos antimicrobianos, aborda a vigilância e notificação de RAM, identificando e acompanhando a dinâmica de compartilhamento no ambiente (Serna; Gonzalez-Zorn, 2022; Aslam *et al.*, 2021).

Com o aumento de casos de infecção por microrganismos resistentes, a OMS criou a “Iniciativa *One Health*”, também denominada de “*One Health Quadripartite*”, uma parceria formalizada de quatro organizações internacionais, juntamente com Organização das Nações Unidas para Alimentação e Agricultura (FAO), PNUMA e a Organização Mundial da Saúde Animal (WOAH) que promovem ações relevantes para segurança alimentar e hídrica, nutrição, controle de doenças zoonóticas, poluição e combate à RAM (WHO, 2017).

A RAM é investigada principalmente em bactérias com interesse clínico, porém, diversos trabalhos apontaram que a propagação de genes de resistência pode se dar por microrganismos patogênicos ou não, tornando necessário técnicas que analisam todo o genoma de comunidades microbianas.

4.7. Tecnologias de sequenciamentos de DNA

A investigação de genes de resistência em isolados ambientais enfrenta o desafio de ser limitada aos microrganismos cultiváveis quando realizada por técnicas de cultura. Nesse contexto, o uso de técnicas de sequenciamento de DNA destaca-se como uma abordagem vantajosa para analisar genes presentes em microbiomas (Akacin *et al.*, 2022).

As evoluções nas metodologias de sequenciamento de DNA são compreendidas em diferentes gerações, trazendo inovações que ampliaram as possibilidades de análise genômica. A primeira geração é conhecida como

método de Sanger, desenvolvida em 1977, permitiu que o DNA fragmentado de indivíduos fossem sequenciados com precisão por meio da eletroforese de placa e fluorescência das bases nitrogenadas. Entretanto, sua limitada capacidade de escalonamento e os altos custos restringiram sua utilização em pesquisas mais complexas (Pareek; Smoczynski; Tretyn, 2011; Sanger; Nicklen; Coulson, 1977).

Em 2005 foi desenvolvida a segunda geração de sequenciamento, conhecida como Sequenciamento de Nova Geração (NGS), com a capacidade de sequenciar milhões de fragmentos de DNA ao mesmo tempo, obtendo leituras curtas de maneira rápida e com menos custo. Existem três principais plataformas de NGS, a primeira conhecida como plataforma 454 *Life Sciences*, utiliza a técnica de pirosequenciamento, sendo necessário realizar PCR por emulsão e detectando a luminescência do pirofosfato na reação em cadeia dos nucleotídeos. Embora tenha sido utilizada para abordagens de genômica *De Novo*, ou seja, genomas sem referências, a técnica foi ultrapassada com o surgimento de outras técnicas (Akacin *et al.*, 2022; Pareek; Smoczynski; Tretyn, 2011).

A segunda plataforma de NGS, Illumina, realiza sequenciamento por síntese, sendo necessário um preparo da biblioteca, com técnicas de PCR para amplificar regiões alvo e primers que se ligarão as extremidades dos fragmentos de DNA, com adição de fluorocromos para cada base nitrogenada que serão reconhecidos para gerar os dados. Por fragmentar e amplificar o DNA, os dados gerados exigem uma grande capacidade computacional para serem tratados e analisados (Akacin *et al.*, 2022; McCombie; McPherson; Mardis, 2019).

A plataforma Ion Torrent, lançada em 2010, usa a tecnologia de íons semicondutores, ou seja, identifica o pH gerado pelo acúmulo de H^+ . Durante a preparação da biblioteca, o DNA da amostra é fragmentado e amplificado por PCR por emulsão, posteriormente, o pH para cada ligação é detectado e sequenciado. Esse método não é eficaz para regiões com homopolímero, causando a saturação na detecção do pH e gerando taxas de erros (Akacin *et al.*, 2022; Hu *et al.*, 2021).

A terceira geração de sequenciamento permitiu a leitura de sequenciamentos longos, maiores que 10kb, sem a fragmentação e amplificação do DNA, fornecendo dados precisos, em tempo real e a possibilidade de sequenciar grandes moléculas de DNA (Hu *et al.*, 2021).

Em 2009 a *Pacific Biosciences* lançou a plataforma PacBio® (Menlo Park, CA, EUA), que apresentou uma técnica de sequenciamento denominada de molde *SMRTbell* por um chip “*SMRT Cell*” que contém a DNA polimerase acoplada e assim reconhece a fita de DNA simples. A preparação da biblioteca não necessita de PCR, somente a adição de fluorescentes para reconhecer as bases nitrogenadas. Embora seja uma técnica promissora, de alto rendimento, essa técnica necessita de altos custos financeiros (Rhoads; Au, 2015).

Outro sequenciamento de terceira geração, lançado em 2012 pela *Oxford Nanopore Technologies*, é a tecnologia *Nanopore* que pode gerar leituras longas maiores que 1 Mb sem fragmentar o DNA, pois permite a passagem de fitas únicas por nanoporos de maneira ininterrupta. Para isso, a biblioteca é carregada em uma célula de fluxo, denominada *flow cell* (em inglês), que contém 512 canais com quatro nanoporos por canal, resultando em 2.048 nanoporos por *flow cell*, cada canal está associado a um eletrodo que reconhece a corrente iônica de cada base nitrogenada (Wang *et al.*, 2021).

Existem três plataformas principais para sequenciamento por *Nanopore*: PromethION, um sequenciador de larga escala com capacidade para até 48 *flow cells*; GridION, um equipamento de bancada que suporta até cinco *flow cells*; e MinION, um dispositivo portátil e versátil que acomoda uma *flow cell*. O MinION destaca-se por sua portabilidade, possibilitando análises em locais remotos sem necessidade de infraestrutura laboratorial, desde que haja acesso à rede de dados. Sua capacidade de realizar sequenciamento eficaz, em tempo real e de baixo custo é essencial para estudos que requerem respostas rápidas, como o monitoramento ambiental, a vigilância epidemiológica e a análise de microbiomas. Apesar de apresentar uma taxa de erro maior comparada a outras plataformas, os avanços em bioinformática têm melhorado significativamente a precisão de suas análises (Akacin *et al.*, 2022; Hu *et al.*, 2021; Wang *et al.*, 2021).

Com o avanço das tecnologias de sequenciamento e consequentemente das ômicas (genômica, transcriptômica, proteômica, metabolômica, dentre outros) e com auxílio da bioinformática, é possível obter genomas mistos, identificar e caracterizar genes de resistência em toda amostra sem cultivo e isolamento laboratorial utilizando a metagenômica (Shivaram *et al.*, 2023; Akacin *et al.*, 2022; Gupta; Tiwari; Cytryn, 2020).

4.8. Metagenômica

A abordagem metagenômica está prevalecendo na avaliação de biodiversidade microbiana, descoberta de patógenos humanos, estrutura de comunidades e detecção de resistoma em microbiomas. O sequenciamento por metagenômica fragmenta o DNA e pode ser utilizado para verificar genes específicos, direcionando para regiões do DNA ou rDNA, como amplicons ou gene 16S, que é uma região de conservação do gene, utilizando a técnica PCR ou RT-PCR para amplificação da região selecionada, ou pelo método conhecido como *shotgun*, no qual ocorre a fragmentação do DNA e o sequenciamento de todos os genomas presentes na amostra. Ambas as técnicas variam dependendo do objetivo, recursos e plataforma de sequenciamento (Banar *et al.*, 2024; Srinivas *et al.*, 2022). Com a leitura sequencial do gene, é possível obter informações sobre o perfil de resistência e multirresistência aos antimicrobianos de maneira *in silico*, além de identificar a linhagem da bactéria e se obteve o gene por transferência horizontal de gene, mutação seletiva ou descendência (Liguori *et al.*, 2022; Imchen *et al.*, 2020).

O sequenciamento *shotgun* permite uma cobertura maior para análises de biodiversidade microbiológica, possibilitando a identificação de microrganismos menos abundantes na amostra. Além da composição taxonômica a nível de gênero e espécie, permite identificar os níveis funcionais da comunidade bacteriana e reconstrução do genoma a partir do alinhamento. Contudo, com uma maior resolução do genoma de diversos indivíduos, maior é o custo e o tamanho das leituras (Yasir *et al.*, 2022; Durazzi *et al.*, 2021).

5. MATERIAL E MÉTODOS

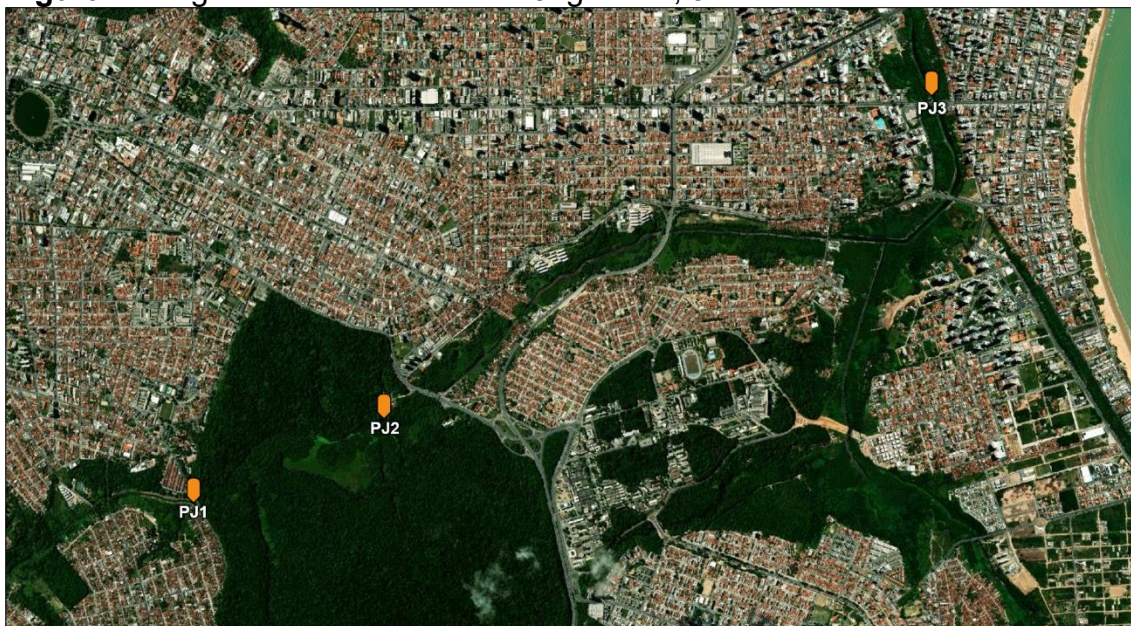
5.1. Estudo Piloto: Padronização das Coletas e Análises de Bioinformática

Previamente às coletas na região amazônica, foi conduzido um estudo piloto com o objetivo de testar e padronizar as metodologias de coleta de água, extração de DNA metagenômico e as etapas da análise de bioinformática.

As coletas foram realizadas em fevereiro e março de 2025 em três pontos distintos ao longo do Rio Jaguaribe (Figura 1), em João Pessoa, Paraíba, as

amostras de água foram coletadas e processadas seguindo o mesmo protocolo planejado para o estudo principal. Ponto 1: PJ1, na região do Varjão (7°08'28.2"S 34°52'16.6"O); PJ2, no Jardim Botânico Benjamin Maranhão (7°08'11.8"S 34°51'39.3"O) e PJ3, próximo a Ponte Tambaú (7°07'08.8"S 34°49'51.8"O), com a temperatura de 30 °C, 27,2 °C e 27,1 °C e o pH 6, 5 e 6, respectivamente. Este piloto foi fundamental para validar a eficácia do fluxo de trabalho e garantir a qualidade dos dados a serem obtidos nas áreas de estudo do bioma amazônico.

Figura 1: Regiões de coletas do Rio Jaguaribe, João Pessoa.

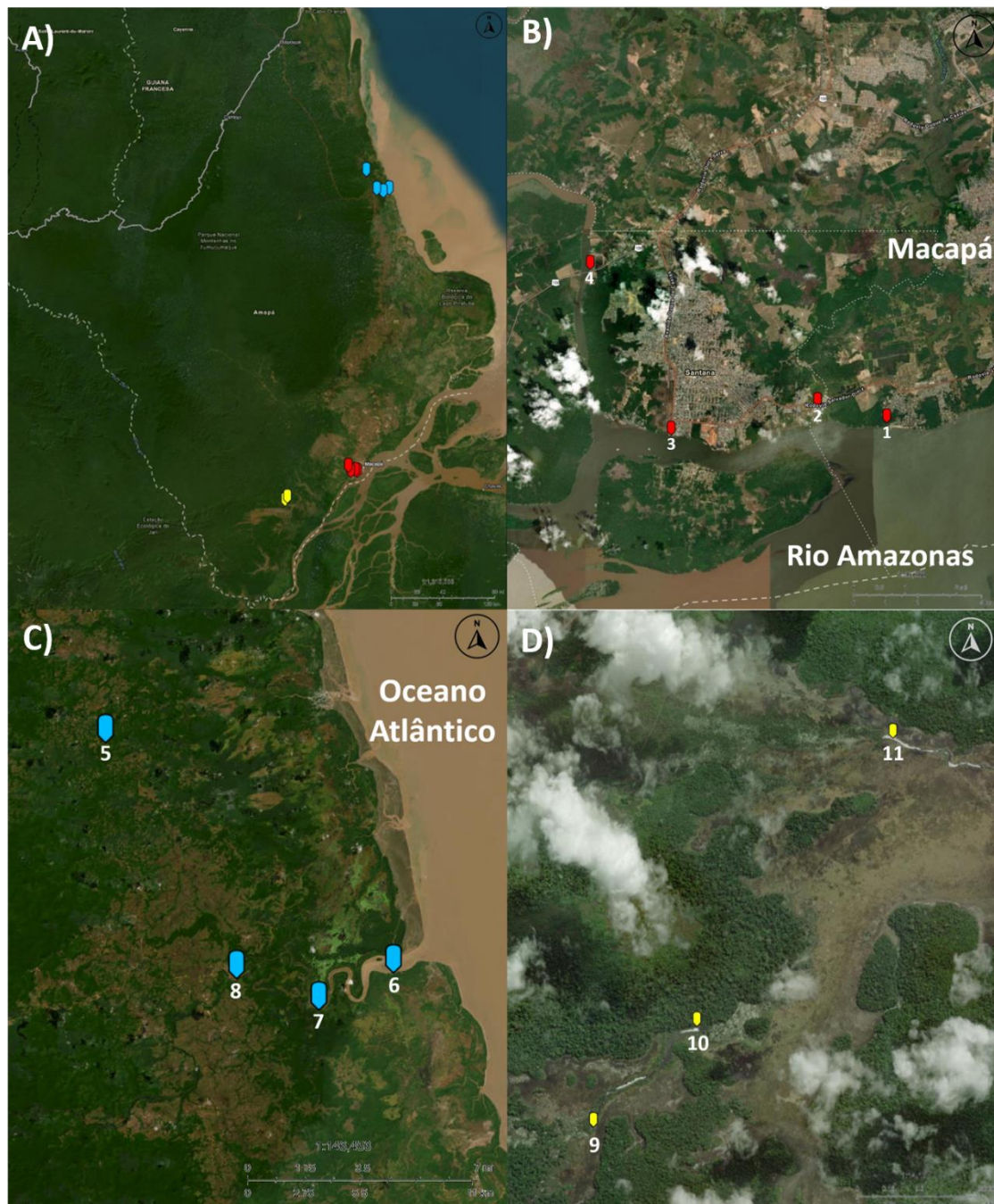


Fonte: Elaborado pelo autor.

5.2. Coleta das amostras

As áreas foram previamente selecionadas, considerando acessibilidade, riscos e profundidade, em regiões distintas no estado do Amapá. Assim, foram selecionadas três regiões: Região 1 – Macapá: localizada na capital e grande Macapá, na porção sudeste do Amapá, próxima à linha do Equador e à margem esquerda do rio Amazonas; Região 2 – Cidade de Calçoene, situada na porção nordeste do estado, abrangendo áreas próximas ao Parque Arqueológico dos Equinócios e zona litorânea; Região 3 – localizada no sudoeste do Amapá, nas imediações do rio Jari, região caracterizada por áreas de floresta densa e influência de afluentes amazônicos.

Figura 2: Mapa do território do Amapá e regiões de coletas.



Fonte: Elaborado pelo autor.

Legenda: A) Mapa do estado do Amapá, com respectivas regiões de coleta. B) Região 1. C) Região 2. D) Região 3.

As coletas foram realizadas entre fevereiro e março de 2025, em colaboração com o Prof. Emerson Augusto Castilho Martins, da Universidade Federal do Amapá (UNIFAP), e os experimentos de Biologia Molecular foram realizados no Laboratório de Biologia Molecular (LaBiMol/CCM/UFPB), sob coordenação da Profa. Eloiza Helena Campana.

Para tanto, aproximadamente um litro de água foi coletada para cada ponto em onze pontos (Tabela 01) das três regiões selecionadas. As amostras foram coletadas utilizando garrafas esterilizadas, que foram submersas a 30 cm de profundidade, a um metro da margem do corpo hídrico, e, em seguida, transportadas para o ambulatório da UNIFAP e armazenadas a 4 °C. Os frascos foram devidamente identificados com a sigla correspondente a cada ponto, incluindo informações de coordenadas no GPS, data e hora da coleta, além de parâmetros físico-químicos de pH e temperatura.

Tabela 01: Caracterização geográfica e físico-química dos pontos de coleta do Amapá.

Ponto	Sigla	Coordenadas de coletas	Matriz aquática	Município	pH	Temperatura (°C)
01	AP1	0°3.2422'S, 51° 7.1689'O	Igarapé APA-Fazendinha	Macapá	7	30,7
02	AP2	0°2.9690'S, 51° 8.3360'O	Igarapé Fortaleza	Santana	6	29,8
03	AP3	0°3.4480'S, 51°10.8010'O	Rio Amazonas	Santana	5	29,3
04	AP4	0°0.6500'S, 51°12.1660'O	Rio Matapi	Santana	5	29,9
05	AP5	2°39.6250'N, 51°2.4100'O	Igarapé Tracajatuba	Calçoene	5	25,8
06	AP6	2°29.6540'N, 50°50.0120'O	Rio Calçoene	Calçoene	5	28,2
07	AP7	2°28.0700'N, 50°53.2210'O	Rio Calçoene	Calçoene	5	26,3
08	AP8	2°29.4340'N, 50°56.7670'O	Rio Calçoene	Calçoene	5	26,1
09	AP9	0°18.5990'S, 51°45.9780'O	Poço Maracá-pacu	Mazagão Velho	5	32,1
10	AP10	0°19.0140'S, 51°46.4080'O	Igarapé Maracá-pacu	Mazagão Velho	5	31,1
11	AP11	0°17.4100'S, 51°45.1630'O	Rio Maracá-pacu	Mazagão Velho	5	31,5

Fonte: Elaborado pelo autor.

5.3. Filtragem das amostras

As amostras foram pré-filtradas no Laboratório de Microbiologia e Imunologia (BIOMOL) da UNIFAP, utilizando filtros de $\geq 0,8$ mm autoclavados, para reter as partículas suspensas. Posteriormente foram filtradas utilizando membranas Millipore Sterivex® de 0,22 μm /poro estéreis. Para cada ponto, foram filtrados, aproximadamente, 150 mL por membrana, o qual dois pares de membranas foram armazenadas com 20 mL de STE (Tris, Saline, EDTA, pH=8) e um par de membrana a seco, em tubos Falcon CRALPLAST® de 50 mL e armazenadas a -20°C .

5.4. Extração do DNA genômico

As membranas foram lavadas utilizando a mesa agitadora orbital Solab® com 180 RPM por 48 horas no Laboratório Analítico Multiusuário (LAM/DCF/UFPB) e os tubos centrifugados a 1.000 RPM por 10 minutos no Laboratório de Imunofarmacologia (LYMFA/IPeFarM/UFPB). O sobrenadante foi descartado e o *pellet* reunido em microtubos Eppendorf Kasvi® de 2 mL no Núcleo de Medicina Tropical (NUMETROP/CCS/UFPB).

A extração de DNA foi realizada no Laboratório de Biologia Molecular (LaBiMol/CCM/UFPB), utilizando o kit comercial DNeasy® PowerSoil® (QUIAGEN®) conforme protocolo do fabricante com modificações (Stojan *et al.*, 2023).

A concentração do DNA obtido nas extrações foi determinada por fluorimetria, utilizando o equipamento Quantus™ (Promega), e confirmada por espectrofotometria (NanoDrop™, Thermo Fisher Scientific) no Laboratório de Biotecnologia Celular e Molecular (LBCM/CBiotec/UFPB). Foram selecionadas amostras com concentração entre 200 ng para o sequenciamento. Como controle negativo da extração, foram incluídas membranas esterilizadas submetidas ao mesmo protocolo experimental, a fim de verificar a ausência de contaminação durante o processo.

5.5. Preparação da biblioteca Nanopore e sequenciamento

A biblioteca *Nanopore* foi obtida no Laboratório de Biologia Molecular (LaBiMol/CCM/UFPB) utilizando o *Kit Rapid Barcod* (SQK-RBK114.24) e *Flow Cell* R10.4.1, com quantidades de nanoporos acima de 1000 antes da realização do sequenciamento, seguindo as recomendações do fabricante. O sequenciamento foi realizado no dispositivo MinION (ONT, UK) e sequenciados usando MinKNOW 1.15.1 com o *script* de execução padrão de 48 horas. O *basecalling* foi transcrito utilizando o *Fast model v4.3.0*, 400bps com o mínimo de Q score 8.

5.6. Análise de bioinformática

5.6.1. Microbioma

Foi gerado e testado dois *pipelines* para a análise de dados metagenômicos após o sequenciamento. O Miniconda (versão 5.1.16) e o Galaxy Server (versão 25.0.3) foram utilizados para gerenciamento de pacotes no Ubuntu (16.04). Após a obtenção dos resultados de saída (arquivos FASTQ), foram avaliadas a qualidade das *reads* (sequências de bases nitrogenadas em formato de texto) por meio das ferramentas selecionadas durante o teste. Os FASTQs foram submetidos para avaliação usando o Falco versão 1.2.4 (De Sena Brandine; Smith, 2021) e NanoPlot versão 1.28.2 (De Coster *et al.* 2018). As *reads* foram tratadas considerando os parâmetros *Phred Quality Score* <10 utilizando o *Porechop.py* versão 0.2.4 (Wick, 2017) e Filtlong versão 0.3.1 (Wick, 2018) para remoção de regiões adaptadoras provenientes do sequenciamento. O alinhamento das sequências obtidas foi feito pelo Blast+ disponível no NCBI (Centro Nacional de Informação Biotecnológica) e a montagem dos genomas por meio das ferramentas Flye (metaFlye) versão 2.9.6 (Kolmogorov *et al.*, 2020) e MEGAHIT versão 1.2.9 (Li *et al.*, 2015), considerando uma abordagem *De Novo*. Para a avaliação da qualidade das montagens, foram utilizados o Quast e Metaquast versão 5.3.0 (Mikheenko; Saveliev; Gurevich, 2016).

Para classificação taxonômica, foi utilizado a ferramenta Kraken versão 1.3.1, Kraken2 versão 2.1.3 e Kraken-report versão 1.3.1 (Wood; Salzberg, 2014) e Krakentools versão 1.2.1 (Lu *et al.*, 2022). A visualização dos arquivos

finais (.report e .txt) foi realizada a partir do software Pavian versão 1.0 (Breitwieser; Salzberg, 2020).

5.6.2. Resistoma

Após a identificação do microbioma nas amostras, foram analisados os resistomas utilizando o banco de dados ResFinder (*Center for Genomic Epidemiology*) versão 4.7.2 (Bortolaia *et al.*, 2020; Camacho *et al.*, 2009), CARD (*The Comprehensive Antibiotic Resistance Database*) (Alcock *et al.*, 2023) e ABRicate versão 1.0.1 (Seemann, 2016) com o banco de dados NCBI-BARRGD (*Bacterial Antimicrobial Resistance Reference Gene Database*). A utilização dessas bases de dados amplia a cobertura de identificação dos determinantes de resistência.

5.7. Devolutiva social

Os dados obtidos estão sendo divulgados para a população através do projeto Iniciativa Amazônia +10 com a confecção de folders (Apêndice 01) e panfletos (Apêndice 02) informativos, abordando os riscos de exposição aos microrganismos resistentes, a dinâmica *One Health*, como se prevenir e combater a perpetuação de genes de resistência e a importância do uso racional clínico e ambiental, para divulgação no site do projeto (<https://www.amazoniamaisdez.org.br/chamada1/projeto-33>) e redes sociais, juntamente com outras unidades onde está sendo realizado o projeto.

5.8. Divulgação acadêmica e científica

Foram realizados palestras e *workshops* para a comunidade acadêmica abordando temas relacionados ao microbioma, resistoma e metagenômica. As atividades foram direcionadas à Universidade Federal da Paraíba, com a participação de professores e estudantes que desenvolvem pesquisas em metagenômica e análises de dados por bioinformática. Os resultados foram divulgados por meios digitais e tradicionais ressaltando a importância da investigação de genes de resistência no bioma amazônico e implicações para o estado da Paraíba e outras regiões do país.

5.9. Colaboração do projeto

Pesquisadores de diferentes estados brasileiros e colaboradores internacionais lançaram a proposta 003/2022 – Iniciativa Amazônia +10 (<https://www.amazoniamaisdez.org.br/chamada1/projeto-33>) e atuam desde então em conjunto no monitoramento de patógenos zoonóticos com potencial de riscos para a saúde humana no bioma amazônico. Amostras biológicas (humanas e animais) e ambientais (água e esgoto) são avaliadas de forma detalhada por meio de técnicas de sequenciamento utilizando abordagem metagenômica e análises de bioinformática.

6. RESULTADOS E DISCUSSÃO

6.1. Amostras Piloto

6.1.1. Quantificação das Amostras Piloto e Padronização dos Pipelines

O protocolo de extração do DNA metagenômico foi validado, garantindo concentrações acima de 20 ng/μl e razões de pureza A260/280 ≥ 1.80 (Tabela 02). Esses parâmetros atendem aos critérios de concentração e qualidade, conforme recomendado pelo protocolo de sequenciamento *Rapid sequencing* DNA V14 - *barcoding* (SQK-RBK114.24).

Tabela 02: Parâmetros de quantificação e pureza do DNA metagenômico extraído das amostras piloto.

Pontos	Pellet (g)	Concentração de DNA (ng/μl)	A260/A280
PJ1	17,6	60,4	1.80
PJ2	20,4	104,7	1.92
PJ3	15,7	40,7	1.90

Fonte: Elaborado pelo autor.

Legenda: PJ1 (Ponto 1 do piloto); PJ2 (Ponto 2 do piloto); PJ3 (Ponto 3 do piloto).

O sequenciamento, realizado no equipamento MinION Mk1C, teve uma execução de 20 horas e 58 minutos, gerando dados brutos em formato de FASTQ. Foram desenvolvidos e padronizados dois *pipelines* integrados com

base nesses dados, para garantir reprodutibilidade e robustez na análise das amostras principais do projeto.

A validação do protocolo com a extração adequada de pureza e concentração é fundamental para o sucesso de análises metagenômicas ambientais, devido à complexidade da amostra, como as de rios amazônicos. A plataforma *Oxford Nanopore* possui alta sensibilidade à qualidade do DNA de entrada, evitando *basecalling* (tradução do sinal elétrico para a respectiva base nitrogenada) errôneos ou indeterminados (N) (Brown *et al.*, 2017). Parâmetros como a concentração de DNA e a razão A260/280 > 1.80 para a biblioteca de entrada, são fatores determinantes para a eficiência do sequenciamento de terceira geração e podem impactar diretamente nos dados de saída (Tyler *et al.*, 2018).

A extração de DNA da amostra piloto apresentou parâmetros de quantidade e qualidade adequados, compatíveis com os dados na literatura para realizar o sequenciamento de metagenômica pela plataforma *Oxford Nanopore*.

6.1.2. Controle de Qualidade e Filtragem

A análise de qualidade do sequenciamento das amostras piloto, foi realizada com o auxílio da ferramenta *Nanoplot* (Tabela 03) e Falco (Figura 3;4;5). No total, foram produzidos 8,9 gigabases (GB) de dados, distribuídos em 1.872.191,0 *reads*. Os comprimentos médios das leituras variaram entre 2.219,5 (PJ3) a 2.558,7 (PJ2), com desvios padrão acima de 2.500 bases, indicando que foram geradas sequências longas e de diferentes tamanhos de *reads*.

O valor de N50, que variou de 3.815,0 bases (PJ3) até 4.477,0 bases (PJ2), também demonstrou a capacidade do sequenciamento em gerar sequências longas, uma vantagem crucial para a montagem e análise metagenômica, uma vez que 50% de todas as bases sequenciadas estão em *reads* com um comprimento próximo ou maior do que 4 kb.

A qualidade das sequências foi adequada, com todos os pontos apresentando um *score* de qualidade de *Phred* > Q10. Este valor indica uma probabilidade de erro de base de 1 em 10 ou inferior para a maioria dos dados gerados.

Após a avaliação inicial, os dados brutos foram submetidos primeiramente ao *Porechop*, especificamente para remover os adaptadores de *barcode* das

extremidades das *reads*, assegurando que as sequências específicas não influenciem nos dados posteriores. Em seguida, aplicou-se ao *Filtlong* para filtragem, sendo descartadas as *reads* com comprimento inferior a 1.000 pares de bases e as 10% com os menores *scores* de qualidade *Phred*.

Com esse processamento, observou-se uma redução no volume total de bases e no número de *reads* em todos os pontos analisados. Houve um aumento no comprimento médio e no desvio padrão, assim como no valor de N50, indicando que 50% das bases estão presentes em leituras de comprimento maior que 4 kb.

Além disso, houve aumento na proporção de *reads* com qualidade *Phred* superior a Q10, com estes representando mais de 86% das *reads* em todas as amostras pós-filtragem. Em conjunto, a comparação das métricas pré e pós-processamento confirmam a eficácia do processo de limpeza e reforça a confiabilidade dos dados para as etapas de análise posteriores.

A análise gráfica da qualidade por posição de base corrobora com os dados numéricos. As Figuras 03 (a), 04 (a) e 05 (a), correspondem a análise de dados brutos dos pontos PJ1, PJ2 e PJ3, respectivamente. As posições iniciais das *reads* apresentam baixa qualidade, acompanhadas de uma ampla variação. Após a filtragem (b), as regiões iniciais apresentaram aumento no *score*, estabilizando-se na região de *score* mediana (aproximadamente *Phred* 20). O aumento da linha média e a uniformidade das barras interquartis ilustram a eficiência dos filtros aplicados e a remoção da região dos adaptadores dos *barcodes*.

O controle de qualidade e filtragem são requisitos que afetam diretamente a precisão da montagem, da classificação taxonômica e da identificação de genes funcionais. O perfil de qualidade observado nos dados brutos, com uma proporção inicial de aproximadamente 62% de *reads* com *Phred score* > Q10, está de acordo com o esperado para a *flowcell* R10.4.1 da tecnologia *Nanopore*, que, embora apresente taxas de erro mais elevadas que as plataformas de *short reads*, oferece a vantagem crucial de gerar *long reads* (Wang et al., 2021).

A estratégia de filtragem foi eficaz, pois elevou a porcentagem de *reads* de alta qualidade para mais de 86%. Essa abordagem é similar à utilizada em outros estudos de metagenômica ambiental com sequenciamento *Nanopore*,

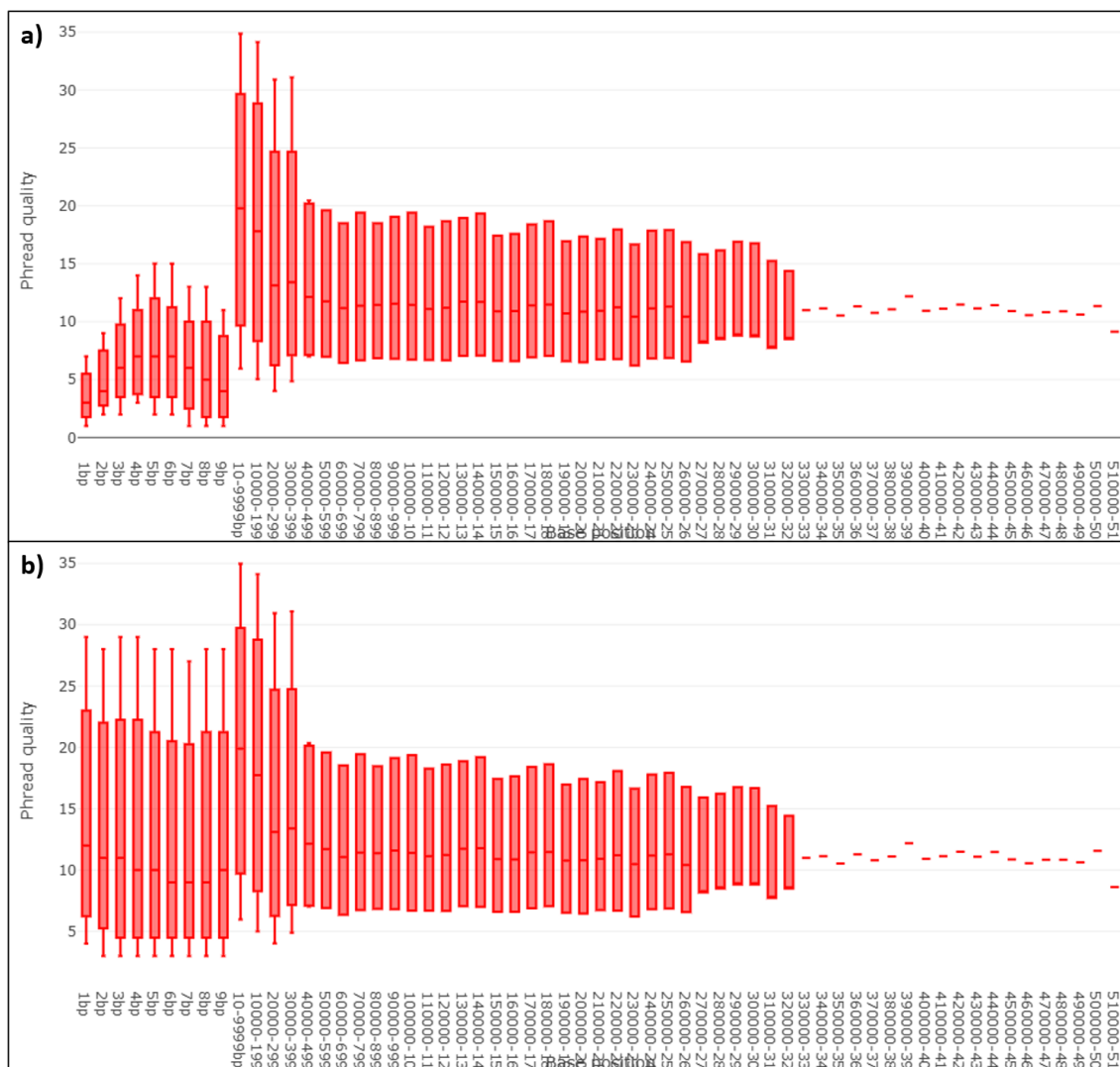
validando a escolha das ferramentas para maximizar a qualidade dos dados sem uma perda excessiva de informação (Martin *et al.*, 2021).

Figura 03: Qualidade por sequência base de PJ1.



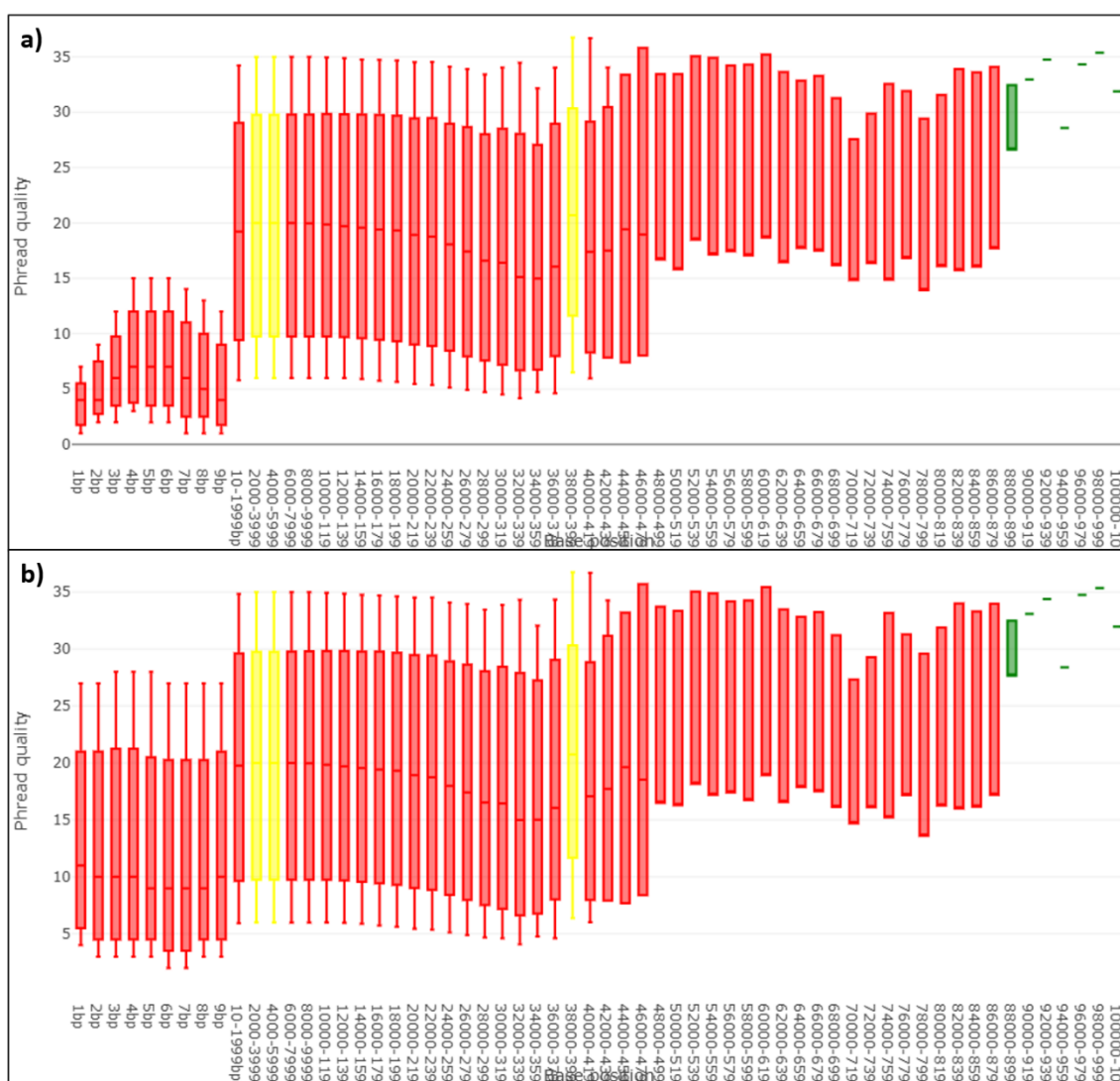
Fonte: Elaborado pelo autor.

Legenda: eixo y) Qualidade Phread; eixo x) Posição dos pares de bases; a) Dados brutos; b) Dados filtrados.

Figura 04: Qualidade por sequência base de PJ2.

Fonte: Elaborado pelo autor.

Legenda: eixo y) Qualidade Phread; eixo x) Posição dos pares de bases; a) Dados brutos; b) Dados filtrados.

Figura 05: Qualidade por sequência base de PJ3.

Fonte: Elaborado pelo autor.

Legenda: eixo y) Qualidade Phread; eixo x) Posição dos pares de bases; a) Dados brutos; b) Dados filtrados.

Tabela 03: Análise da qualidade do sequenciamento das amostras piloto de saída bruto pelo Nanoplot.

Pontos	Fastq	Número de <i>reads</i>	Número de pares de bases	Comprimento médio das <i>reads</i>	Desvio Padrão do comprimento	N50	Phread >Q10	% > Q10
PJ1	3,8 GB	819.007,0	1.900.911.747,0	2.321,0	2.658,0	4.095,0	510.840,0	62,4%
PJ2	2,4 GB	460.363,0	1.177.916.951,0	2.558,7	2.798,6	4.477,0	296.159,0	64,3%
PJ3	2,7 GB	592.821,0	1.315.777.946,0	2.219,5	2.514,3	3.815,0	356.948,0	60,2%

Fonte: Elaborado pelo autor.

Tabela 04: Análise da qualidade do sequenciamento das amostras piloto de saída após processamento e filtragem pelo Nanoplot.

Pontos	Fastq	Número de <i>reads</i>	Número de pares de bases	Comprimento médio das <i>reads</i>	Desvio Padrão do comprimento	N50	Phread >Q10	% > Q10
PJ1	3,3 GB	498.519,0	1.669.780.951,0	3.349,5	2.869,0	4.523,0	430.380,0	86,3%
PJ2	2,1 GB	298.674,0	1.054.440.613,0	3.530,4	2.945,5	4.795,0	256.988,0	86,0%
PJ3	2,3 GB	359.093,0	1.148.084.450,0	3.197,2	2.714,8	4.231,0	311.812,0	86,8%

Fonte: Elaborado pelo autor.

6.1.3. Montagem e avaliação de qualidade

A reconstrução das *contigs* do metagenoma por abordagem *De Novo*, sem genomas de referência, foi realizada utilizando dois *pipelines* distintos: Flye (metaFlye) para *long reads* e MEGAHIT para *short reads*. A montagem Flye passou por polimento com MEDAKA, e a qualidade de ambas foi avaliada pelo QUAST, gerando métricas de continuidade, completude e possíveis erros.

Ambas estratégias apresentaram resultados distintos. O Flye (Tabela 05) produziu um número menor de *contigs* (>500 pb) em todas as amostras (PJ1: 4.152; PJ2: 1.761; PJ3: 2.616), enquanto o MEGAHIT (Tabela 06) gerou centenas de milhares de *contigs* (PJ1: 393.618; PJ2: 199.855; PJ3: 299.667).

Essa diferença se dá pela fragmentação típica de montadores para *short reads*, como o MEGAHIT, o qual o volume total de *contigs* é maior, mas distribuído em fragmentos menores. Diferentemente do Flye, foi capaz de reconstruir *contigs* longas, chegando a 1,9 Mb, como na amostra PJ1.

O comprimento total das montagens pelo MEGAHIT foi muito superior (PJ1: 466.273.810 pb; PJ2: 307.473.377 pb; PJ3: 359.776.912 pb) em relação ao Flye (PJ1: 68.542.316 pb; PJ2: 28.404.943 pb; PJ3: 44.926.799 pb). Isso indica que, mesmo com a fragmentação, o MEGAHIT recupera um volume maior de sequências, incluindo fragmentos de microrganismos de baixa abundância.

A extensão dos *contigs* são verificadas por parâmetros de contiguidade, que verifica a fragmentação das *reads*, o qual maiores contiguidades implicam em *long reads*. O N50 pelo Flye, resultou em *contigs* maiores que 2 kb, ou seja, indica que 50% do metagenoma está montado em *contigs* de tamanho igual ou superior aos valores. O L50 é uma métrica complementar e indica o número de *contigs* necessário para atingir o N50, como por exemplo na montagem pelo Flye que apresentou valores menores, demonstrando a eficiência para montar *contigs* a partir de *long reads*.

A porcentagem de bases GC (Guanina e Citosina) variou entre as amostras e montadores, mas permaneceu dentro do esperado para comunidades microbianas ambientais, sem indicar viés de montagem. Nenhuma das montagens apresentou bases indefinidas (N's), evidenciando alta qualidade e precisão após o polimento com MEDAKA.

Tabela 05: Métricas de montagem das amostras piloto pelo Flye.

Pontos	Número de <i>contigs</i> > 500 pb	Maior contig	Comprimento total > 500 pb	Contiguidade		GC (%)	Bases indefinidas (N's)
				N50	L50		
PJ1	4.152	1.964.404	68.542.316	22.867	684	63,78	0
PJ2	1.761	974.569	28.404.943	24.195	257	57,13	0
PJ3	2.616	668.151	44.926.799	27.765	338	63,89	0

Fonte: Elaborado pelo autor.

Tabela 06: Métricas de montagem das amostras piloto pelo MEGAHIT.

Pontos	Número de <i>contigs</i> > 500 pb	Maior contig	Comprimento total > 500 pb	Contiguidade		GC (%)	Bases indefinidas (N's)
				N50	L50		
PJ1	393.618	35.248	466.273.810	1.337	95.170	59,10	0
PJ2	199.855	21.078	307.473.377	1.958	44.590	57,20	0
PJ3	299.667	26.842	359.776.912	1.362	74.512	61,35	0

Fonte: Elaborado pelo autor.

A comparação direta entre os montadores Flye e MEGAHIT para a montagem de dados de *long reads* de *Nanopore* revelou diferenças de desempenho consideráveis, que são diretamente atribuíveis às suas diferentes abordagens algorítmicas.

O MEGAHIT utiliza um algoritmo baseado em *De Bruijn graphs*, que funciona dividindo as *reads* em pequenos fragmentos de tamanho fixo (*k-mers*) para encontrar sobreposições. As taxas de erro mais altas e a natureza não uniforme das *long reads* criam complexidades no qual o método *De Bruijn graph* não consegue resolver eficientemente, resultando em quebras no grafo e, consequentemente, em uma montagem fragmentada (Tabela 06) (Wick; Judd; Holt, 2019).

Já o Flye possui uma abordagem diferente que emprega o *overlap-layout-consensus* (OLC) e constrói um *repeat graph* para resolver regiões repetitivas complexas. Ao focar nos *overlaps* entre as *reads* inteiras em vez de *k-mers* fixos, o Flye consegue tolerar as imprecisões do sequenciamento e utilizar o comprimento das *reads* para contornar regiões repetitivas resultando na reconstrução de *contigs* longos (Kolmogorov, *et al.*, 2019).

Esses resultados (Tabela 05 e 06) demonstram que o Flye é altamente eficiente para reconstruir *contigs* longas e contínuas, facilitando a recuperação de genomas completos e a análise estrutural de elementos genéticos. Por outro lado, o MEGAHIT gera um grande número de *contigs* que pode ampliar a capacidade de detecção taxonômica, sendo útil em amostras ambientais complexas e fragmentadas, onde a diversidade microbiana é alta e muitos organismos estão presentes em baixa abundância.

6.1.4. Microbioma da amostra piloto

Após realizar a montagem metagenômica, os arquivos resultantes do Flye e MEGAHIT foram submetidos a análise taxonômica pelo Kraken e Kraken2 utilizando o banco de dados: *Prebuilt Refseq indexes: GTDB v220 (Bacterial and archaeal)* (Version: 2024-11-09 - Downloaded: 2025-02-12T205236Z). Foram processados pelo Kraken-report e Krakentools. Posteriormente, os gráficos de sankey foram gerados pelo *Pavian metagenomics data explorer*.

As montagens pelo Flye geraram um número reduzido de dados, distribuídos em 296 linhas para PJ1, 124 linhas para PJ2 e 214 linhas para PJ3,

ou seja, foram atribuídos níveis aos respectivos *contigs* e detalhado por percentual de abundância, refletindo o perfil de baixa fragmentação e alta continuidade.

A análise taxonômica permitiu a classificação de uma alta proporção dos dados, com 76,85%, 68,12% e 80,62% dos *contigs* atribuídos a táxons conhecidos para o PJ1 (Figura 07), PJ2 (Figura 08), e PJ3 (Figura 09), respectivamente. A composição a nível de filo foi majoritariamente por *Pseudomonadota*, que correspondeu a 34,63%, 48,58% e 37,61% para as amostras PJ1, PJ2 e PJ3, respectivamente. Além disso, o filo *Actinomycetota* também foi encontrado nos três pontos.

Os pontos PJ1 e PJ3 apresentaram uma estrutura de comunidade filogenética relativamente similar. Ambos tiveram a predominância do filo *Cyanobacteriota* (6,70% em PJ1; 7,95% em PJ3).

O ponto PJ2, exibiu um perfil taxonômico diferente e de menor diversidade. Ocorreu um aumento do filo *Bacteroidota* (5,78%). Divergindo de filos abundantes nos outros pontos, como *Cyanobacteriota*, apresentaram proporções menores em PJ2 (0,17%).

Essa divergência se tornou ainda mais evidente em nível de gênero. Enquanto PJ1 e PJ3 foram apresentaram gêneros tipicamente aquáticos como *Limnohabitans*, *Acidovorax* e o grupos de cianobactérias, o ponto PJ2 foi dominado por gêneros associados a ambientes impactados e ao trato gastrointestinal, como *Aeromonas* (6,60%) e *Enterobacter* (6,30%).

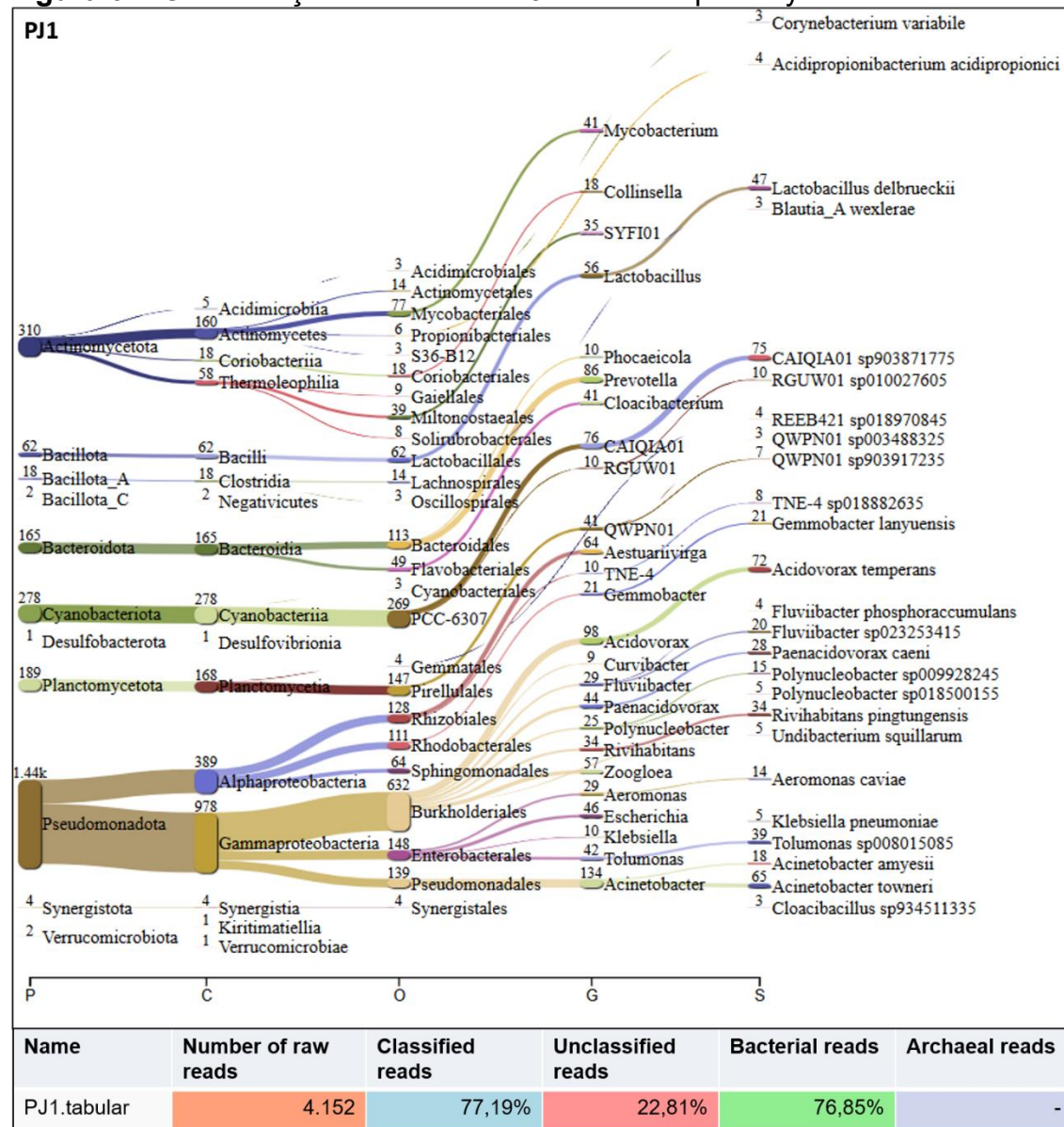
A presença de *Klebsiella* foi de 0,24%, 6,34% e 0,38% para os pontos PJ1, PJ2 e PJ3, respectivamente, com a espécie *K. pneumoniae* detectada. Além disso, os pontos PJ1 e PJ3 apresentaram o gênero *Escherichia*, com 1,11% e 0,04%, com a *E. coli* detectada no PJ1.

No processamento das amostras piloto, o Pavian apresentou limitações ao tentar organizar visualmente os arquivos classificados gerados pelo Kraken2 a partir das montagens com MEGAHIT, que geraram um alto volume de linhas (PJ1: 6.320 linhas; PJ2: 2.272 linhas; PJ3: 2.272 linhas). Este volume de dados excedeu a capacidade da ferramenta Pavian de renderizar uma representação gráfica interativa e coesa, como os diagramas de Sankey.

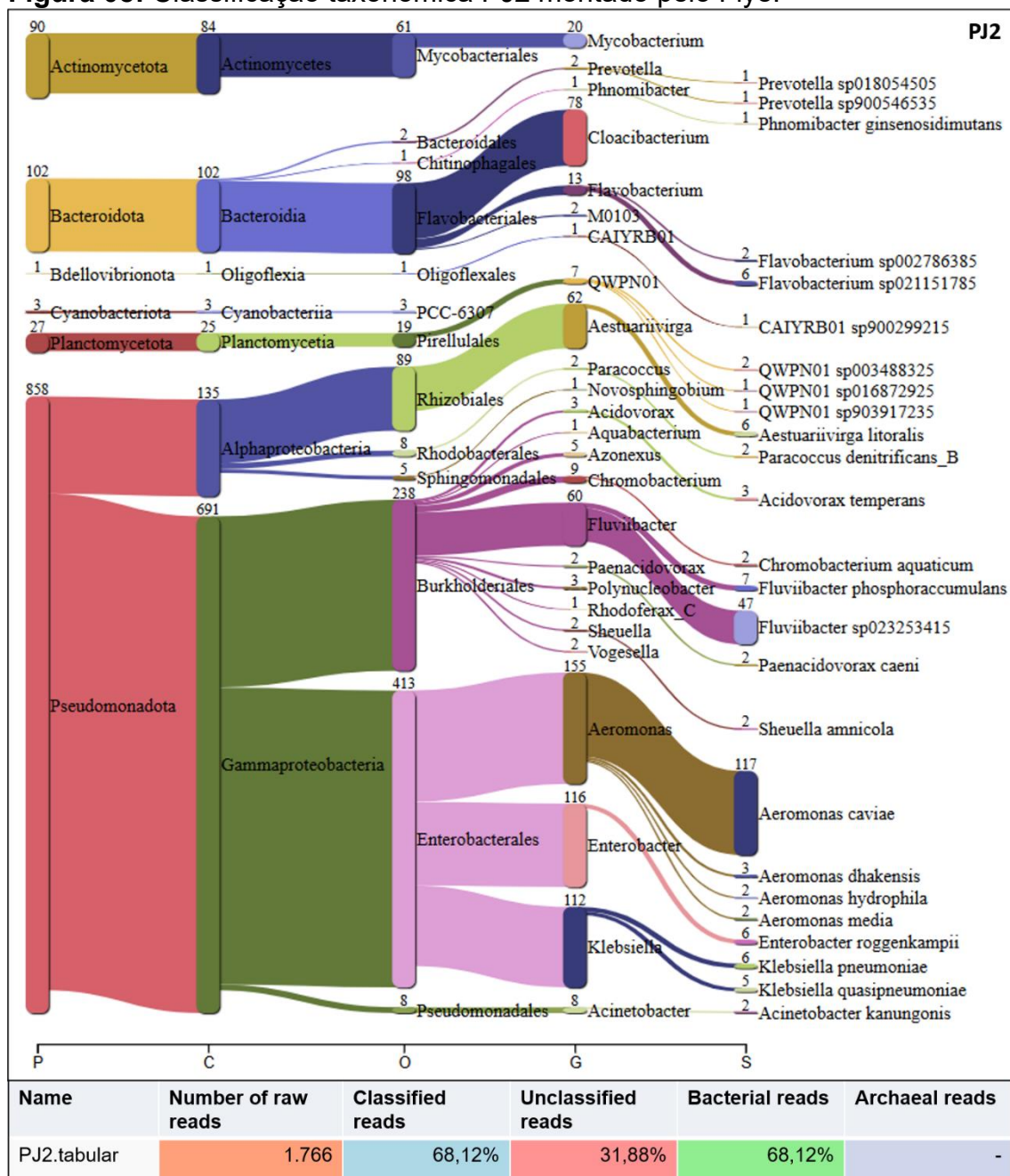
Apesar dessa limitação na visualização, o conjunto de dados permanece útil e foi utilizado nas análises das amostras subsequentes. A decisão de

prosseguir com esses dados baseia-se na robustez da classificação taxonômica pelo Kraken2, que permite uma análise da diversidade microbiana mesmo em montagens fragmentadas.

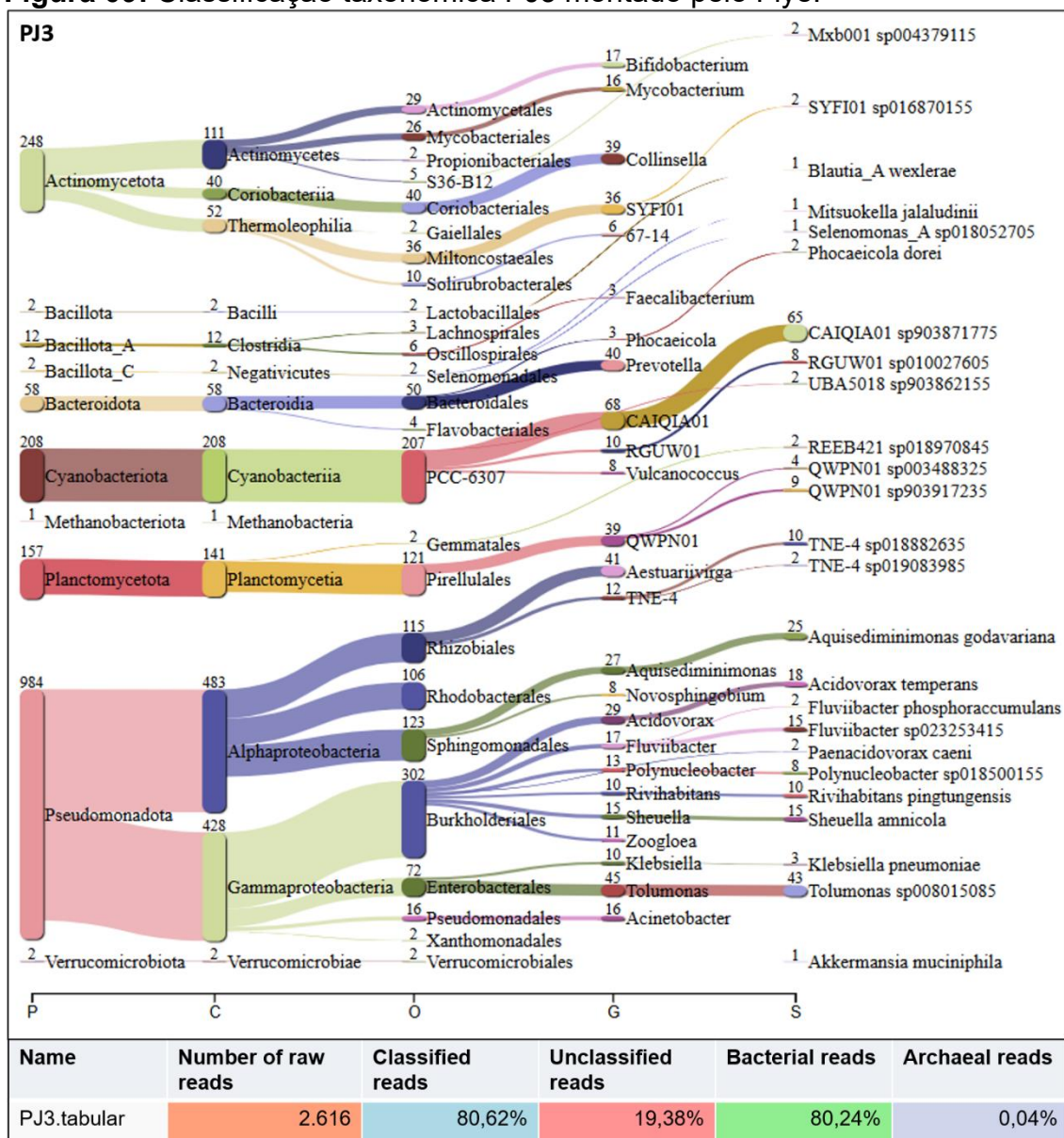
Figura 07: Classificação taxonômica PJ1 montado pelo Flye.



Fonte: Elaborado pelo autor.

Figura 08: Classificação taxonômica PJ2 montado pelo Flye.

Fonte: Elaborado pelo autor.

Figura 09: Classificação taxonômica PJ3 montado pelo Flye.

Fonte: Elaborado pelo autor.

6.1.5. Análise do resistoma

Após a caracterização taxonômica, a investigação foi direcionada ao potencial funcional do metagenoma, com foco na identificação de genes RAM. Assim, os *contigs* da montagem Flye foram analisados utilizando um conjunto os bancos de dados de referência NCBI-BARRGD, ResFinder e CARD.

Para determinar a origem taxonômica dos contigs que continham genes de RAM, foi realizada uma análise de homologia utilizando a ferramenta BLASTn contra o banco de dados de nucleotídeos (nt) do NCBI.

Para o ponto PJ1 (Tabela 07), foram identificados 14 genes de resistência distintos, conferindo resistência a oito classes de antimicrobianos, incluindo betalactâmicos, aminoglicosídeos, trimetoprim, sulfonamidas, tetraciclina e macrolídeos. A qualidade dos alinhamentos foi consistentemente alta, com valores de identidade e cobertura superiores a 94% para a maioria dos genes, indicando a presença de genes completos e potencialmente funcionais.

Dentre os achados, o *contig* 3048 que correspondeu ao plasmídeo do *Citrobacter freundii* (Tabela 08), apresentou um *cluster* com múltiplos genes de RAM, *dfrA14* (resistência ao trimetoprim), *aadA1* (resistência à estreptomicina), *bla*OXA-10 (resistência a betalactâmicos e cefalosporinas), *cmlA5* (resistência ao cloranfenicol) e *tet(C)* (resistência à tetraciclina).

Embora o *C. freundii* seja reconhecida como um patógeno oportunista amplamente distribuído em diversos ambientes, como solo, água e o trato intestinal humano, essa espécie não é atualmente classificada como uma ameaça de emergência global. No entanto, a vigilância sobre esse gênero é essencial, considerando o potencial no surgimento de cepas com elevados níveis de resistência a antimicrobianos. Principalmente em ambientes aquáticos sujeitos a impactos antrópicos, que favorecem a disseminação e seleção de microrganismos resistentes.

O *contig* 4210 continha o sistema de resistência de dupla ação *msr*(E) (bomba de efluxo) e *mph*(E) (enzima modificadora), que conferem resistência de alto nível a macrolídeos. O BLAST associou este *contig* ao plasmídeo de *Acinetobacter baumannii*, um patógeno de alta relevância clínica.

A análise do resistoma do PJ2 (Tabela 09) revelou a maior abundância e diversidade de genes de RAM entre as amostras piloto. O perfil de resistência apresentou 21 genes diferentes, principalmente por genes que conferem resistência a quinolonas e outras classes de antimicrobianos sintéticos, além de diversos genes de bombas de efluxo de multidrogas.

Predominou-se a detecção de múltiplos genes de resistência a quinolonas mediados por plasmídeos (Tabela 10). Foram identificados os genes *qnrS2*, *qnrB19* e *qnrB5* (uma variante de *qnrB*), todos com cobertura e identidade próximas de 100%. O *qnrS2* foi associado ao *contig* 1363 que foi identificado como *Aeromonas* ou *K. pneumoniae* (98,95%), enquanto os genes da família

qnrB foram ligados a *E. coli* e outras bactérias não cultivadas, evidenciando a disseminação destes genes entre diferentes hospedeiros na amostra.

Além dos genes *qnr*, outros genes associados à transferência de genes foram detectados com alta cobertura e identidade. O gene *sul2*, que confere resistência a sulfonamidas, foi identificado com 100%, associado a *A. simiae*. O gene *fosA5*, responsável pela resistência à fosfomicina, foi detectado em um *contig* com alta similaridade ao cromossomo de *K. pneumoniae*.

Foram identificados múltiplos sistemas. Como no *contig* 1140, associado a *Enterobacter roggenkampii*, que apresentou os genes *oqxA* e *oqxB*, complexos da bomba de efluxo *OqxAB* (a posição entre 9444 ao 13794 demonstra essa associação), conhecida por conferir resistência a quinolonas e fenóis. O sistema de efluxo *AcrAB*, também foi detectado no *contig* 1416 e 1720, ambos associados a *Enterobacter*.

A análise do resistoma no ponto PJ3 (tabela 11) revelou a presença de diversos genes de resistência antimicrobiana detectados. Os 7 genes diferentes identificados incluem *cfxA3*, *CRP*, *rpoB2*, *RbpA*, *sul1*, *bla-A*, *mtrA*, cobrindo mecanismos variados de resistência, como betalactâmicos, fluoroquinolonas, macrolídeos, sulfamidas e rifamicinas.

O gene *cfxA3*, identificado na *contig* 1112, está associado à resistência à cefamicina e foi relacionado a *Segatella copri* (Tabela 12), um microrganismo relevante do microbioma intestinal. O *CRP*, presente na *contig* 1712, vinculado a *K. pneumoniae*, é um regulador implicado em resistência a fluoroquinolonas, macrolídeos e penams. Genes como *rpoB2* e *RbpA*, associados a *Mycolicibacterium* e *Mycobacterium*, conferem resistência à rifamicina.

Assim, o uso dos bancos de dados NCBI-BARRGD, ResFinder e CARD possibilitou uma identificação abrangente com alta confiabilidade de genes presentes nos *contigs* gerados pela montagem Flye.

Tabela 07: Resistoma do ponto PJ1.

Número Contig	Banco de Dados	Start	End	Fita	Gene	GAPS	% de cobertura	% de identidade	Acesso
1401	NCBI CARD ResFinder	1248	2061	+	<i>sul2</i>	2/2	99,75	99,75	NG_051852.1
1676	NCBI	5466	6346	+	<i>bla-A</i>	1/1	99,89	99,55	NG_047483.1
3048	NCBI CARD ResFinder	1537	2017	-	<i>dfrA14</i>	7/7	100,00	95,84	NG_056035.1
		2093	2878	-	<i>aadA1</i>	6/8	98,99	96,84	NG_052028.1
		2895	3681	-	<i>blaOXA-10</i>	17/22	97,75	94,04	NG_049393.1
		3948	5205	-	<i>cmlA5</i>	10/10	99,52	98,50	NG_051436.1
		8526	9711	+	<i>tet(C)</i>	5/7	99,50	99,41	NG_048181.1
4210	NCBI CARD ResFinder	215	1688	+	<i>msr(E)</i>	2/2	99,86	99,86	NG_048007.1
		1744	2625	+	<i>mph(E)</i>	9/9	99,32	98,65	NG_064660.1
4236	NCBI CARD ResFinder	11864	12650	-	<i>aadA2</i>	9/13	98,86	97,74	NG_051846.1
4273	NCBI	4645	5430	-	<i>ant(3'')-IIa</i>	4/4	99,37	97,33	NG_054649.1
1419	CARD	3136	3613	-	<i>mphB</i>	1/1	100,00	98,12	AE005174.2: 3397370-3397847
1463	CARD	1860	3491	+	<i>eptA</i>	25/38	98,48	94,81	AP009048: 4340268-4338624
286	CARD	2959	4156	-	<i>acrA</i>	15/16	99,50	91,86	U00096.3: 485619-484425

Fonte: Elaborado pelo autor.

Tabela 08: BLAST NCBI das *contigs* e perfil de resistência dos genes do ponto PJ1.

Número Contig	Identificação taxonômica por BLAST NCBI (% identidade)	Acesso	Gene	Resistência
1401	<i>Thermomonas aquatica</i> (99,85%) <i>Pseudomonas aeruginosa</i> (99,85%)	CP040871.1 CP194327.1	<i>sul2</i>	Sulfonamida
1676	<i>Thauera</i> sp. (97,98%) <i>Klebsiella pneumoniae</i> (98,41%)	CP001281.2 MK649823.1	<i>bla-A</i>	Betalactâmico
3048	<i>Citrobacter freundii</i> (98,62%)	CP056339.1 (plasmídeo)	<i>dfrA14</i> <i>aadA1</i> <i>blaOXA-10</i> <i>cmIA5</i> <i>tet(C)</i>	Trimetoprim Estreptomicina Cefalosporina Cloranfenicol Tetraciclina
4210	<i>Acinetobacter baumannii</i> (98,03%)	CP084298.1 (plasmídeo)	<i>msrE</i> <i>mphE</i>	Macrolídeos Macrolídeos
4236	<i>Aeromonas salmonicida</i> (94,23%) <i>Klebsiella</i> sp. (94,23%)	CP038105.1 AP022205.1	<i>aadA2</i>	Aminoglicosídeos
1419	<i>Escherichia coli</i> (100%)	CP148426.1	<i>mphB</i>	Macrolídeos
4273	<i>Acinetobacter chinensis</i> (98,57%)	CP032128.1	<i>eptA</i>	Peptídeo
286	<i>Enterobacter mori</i> (88,05%)	CP071063.1	<i>acrA</i>	Cefalosporina; Fluoroquinolona; Gliciliclina; Penam; Cloranfenicol; Rifamicina; Tetraciclina; Triclosan

Fonte: Elaborado pelo autor.

Tabela 09: Resistoma do ponto PJ2.

Número Contig	Banco de Dados	Start	End	Fita	Gene	GAPS	% de cobertura	% de identidade	Acesso
1140	NCBI	9444	10614	+	<i>oqxA10</i>	4/5	99,57	87,58	NG_050418.1
	CARD	10639	13794	+	<i>oqxB9</i>	16/17	99,78	93,11	NG_050458.1
1363	NCBI	5705	6361	+	<i>qnrS2</i>	0/0	100,00	100,00	NG_050544.1
	CARD	5705	6361	+	<i>qnrS2</i>	0/0	100,00	100,00	NG_050544.1
269	NCBI	358	1173	+	<i>sul2</i>	0/0	100,00	100,00	NG_051852.1
	CARD	358	1173	+	<i>sul2</i>	0/0	100,00	100,00	NG_051852.1
312	NCBI	35027	35448	-	<i>fosA5</i>	4/4	99,76	93,85	NG_050409.1
	CARD	35027	35448	-	<i>fosA5</i>	4/4	99,76	93,85	NG_050409.1
400	NCBI	3348	3993	+	<i>qnrB19</i>	1/1	100,00	99,84	NG_050479.1
	CARD	3348	3993	+	<i>qnrB19</i>	1/1	100,00	99,84	NG_050479.1
1654	ResFinder	12694	13889	+	<i>mdf(A)_1</i>	8/9	96,43	81,22	Y08743
1012	CARD	4491	5856	+	<i>cpxA</i>	4/5	99,27	82,76	BA000007.3:4905062-4903688
1153	CARD	13752	15498	+	<i>msbA</i>	15/16	99,49	82,57	U00096.3:966620-968369
1275	CARD	15421	16130	-	<i>baeR</i>	1/1	98,20	82,00	AP009048.1:2166412-2167135
1416	CARD	3042	4233	+	<i>acrA</i>	20/27	98,66	91,89	AJ318073.1:793-1990
		4258	7405	+	<i>acrB</i>	11/14	99,75	83,81	U00096.3:484403-481253
1418	CARD	14807	15983	-	<i>KpnG</i>	21/26	99,06	81,23	ACWO01000051.1:22092-23265
		16110	16640	-	<i>emrR</i>	2/2	99,81	84,59	U00096.3:2810769-2811300
1567	CARD	12334	12965	-	<i>CRP</i>	3/3	99,68	87,22	AP009048.1:4154296-4153663
1613	CARD	12400	12770	-	<i>marA</i>	2/2	96,35	83,33	AP009048.1:1621287-1621671
1632	CARD	2888	6024	+	<i>arcD</i>	46/55	99,49	81,29	AP009048.1:2586250-2589364
1688	CARD	2249	3376	+	<i>OmpK37</i>	14/15	99,47	90,92	AJ011502.1:300-1425

1720	CARD	17529	20687	-	<i>acrB</i>	46/55	99,27	81,96	U00096.3:484403-481253
		20710	21910	-	<i>acrA</i>	18/19	99,50	89,81	DQ679966:299-1493
		16764	17274	+	<i>emrR</i>	3/3	95,86	82,03	U00096.3:2810769-2811300
1731	CARD	17427	18597	+	<i>KpnG</i>	8/8	99,57	94,64	ACWO01000051.1:22092-23265
		18613	20152	+	<i>KpnH</i>	3/3	99,94	84,23	ASTU01000063.1:61248-62787
		11479	12844	-	<i>cpxA</i>	13/13	98,98	80,92	BA000007.3:4905062-4903688
399	CARD	11479	12844	-	<i>cpxA</i>	13/13	98,98	80,92	BA000007.3:4905062-4903688
400	CARD	3312	3993	+	<i>QnrB5</i>	1/1	100,00	98,97	DQ303919:0-681
745	CARD	5535	6223	-	<i>mtrA</i>	2/2	100,00	86,36	AL123456.3:3627349-3626662
902	CARD	17436	20519	+	<i>mdtC</i>	21/22	99,74	82,11	U00096:2158385-2161463
		23324	24035	+	<i>baeR</i>	5/5	98,06	82,63	AP009048.1:2166412-2167135
		4190	4817	-	<i>CRP</i>	4/5	99,21	88,31	AP009048.1:4154296-4153663

Fonte: Elaborado pelo autor.

Tabela 10: BLAST NCBI das *contigs* e perfil de resistência dos genes da amostra PJ2.

Número Contig	Identificação taxonômica por BLAST NCBI (% identidade)	Acesso	Gene	Resistência
1140	<i>Enterobacter roggenkampii</i> (94,91%)	CP060733.1 (cromossomo)	<i>oqxA10</i>	Fenicol; Quinolona
			<i>oqxB9</i>	Fenicol; Quinolona
1363	<i>Aeromonas allosaccharophila</i> (98,95%)	KU644676.1	<i>qnrS2</i>	Quinolona
	<i>Klebsiella pneumoniae</i> (98,95%)	KT896501.1		
	<i>Aeromonas</i> sp. (98,95%)	MT231819.1		
269	<i>Aeromonas simiae</i> (100,00%)	CP040449.1 (cromossomo)	<i>sul2</i>	Sulfamida
312	<i>Klebsiella pneumoniae</i> (93,42%)	CP021955.1 (cromossomo)	<i>fosA5</i>	Fosfomicina
400	Bactéria não cultivada (99,36%)	MF554640.1	<i>qnrB19</i>	Quinolona
	<i>Escherichia coli</i> (99,36%)	CP081521.1		
1654	<i>Enterobacter</i> sp. (90,79%)	AP019634.1	<i>Mdf(A)_1</i>	Aminocumarina;
1012	<i>Enterobacter</i> sp. (89,76%)	CP012999.1	<i>cpxA</i>	

	<i>Enterobacter asburiae</i> (89,75%)	CP083815.1		Aminoglicosídeo
1153	<i>Enterobacter roggenkampii</i> (95,13%)	CP143663.1	<i>msbA</i>	Nitroimidazol
1275	<i>Klebsiella quasipneumoniae</i> (94,34%)	CP144152.1	<i>baeR</i>	Aminocumarina; Aminoglicosídeo
1416	<i>Enterobacter roggenkampii</i> (89,68%)	CP060733.1	<i>acrA</i>	Cefalosporina; Fluoroquinolona; Glicilciclina; Penam; Fenicol; Rifamicina; Tetraciclina; Triclosan
			<i>acrB</i>	Cefalosporina; Fluoroquinolona; Glicilciclina; Penam; Fenicol; Rifamicina; Tetraciclina; Triclosan
1418	<i>Enterobacter asburiae</i> (92,76%)	CP103751.1	<i>KpnG</i>	Aminoglicosídeo; Carbapenem; Cefalosporina; Fluoroquinolona; Macrolídeo; Penam; Penem; Peptídeo
1567	<i>Klebsiella quasipneumoniae</i> (9,44%)	CP171486.1	<i>emrR</i> <i>CRP</i>	Fluoroquinolona Fluoroquinolona;

				Macrolídeo; Penam
				Carbapenem; Cefalosporina; Cefamicina; Fluoroquinolona; Glicilciclina; Monobactam; Penam; Penem; Fenicol; Rifamicina; Tetraciclina; Triclosan
1613	<i>Enterobacter roggenkampii</i> (91,39%)	CP128618.1	<i>marA</i>	
1632	<i>Enterobacter roggenkampii</i> (94,93%)	CP143663.1 (cromossomo)	<i>arcD</i>	Aminoglicosídeo
1688	<i>Klebsiella variicola</i> At-22	CP001891.1	<i>OmpK37</i>	Carbapenem; Cefalosporina; Cefamicina; Monobactam; Penam; Penem
1720	<i>Enterobacter roggenkampii</i> (91,49%)	CP086201.1	<i>acrB</i>	Cefalosporina; Fluoroquinolona; Gliociclina; Penam; Fenicol; Rifamicina; Tetraciclina; Triclosan
			<i>acrA</i>	Cefalosporina;

				Fluoroquinolona; Gliociclina; Penam; Fenicol; Rifamicina; Tetraciclina; Triclosan
			<i>emrR</i>	Fluoroquinolona Aminoglicosídeo; Carbapenem; Cefalosporina; Fluoroquinolona; Macrolídeo; Penam; Penem; Peptídeo
			<i>KpnG</i>	Aminoglicosídeo; Carbapenem; Cefalosporina; Fluoroquinolona; Macrolídeo; Penam; Penem; Peptídeo
1731	<i>Enterobacter roggenkampii</i> (92,97%)	CP060733.1 (cromossomo)	<i>KpnH</i>	Aminoglicosídeo; Carbapenem; Cefalosporina; Fluoroquinolona; Macrolídeo; Penam; Penem; Peptídeo
399	<i>Klebsiella quasipneumoniae</i> (95,02%)	CP055010.1 (cromossomo)	<i>cpxA</i>	Aminocumarina; Aminoglicosídeo
745	<i>Mycolicibacterium aichiense</i> (77,33%)	AP022561.1	<i>mtrA</i>	Macrolídeo; Penam
902	<i>Enterobacter cloacae</i> (86,56%)	CP093914.1 (cromossomo)	<i>mdtC</i> <i>baeR</i>	Aminocumarina Aminocumarina;

953	<i>Klebsiella pneumoniae</i> (93,02%)	CP152840.1 (cromossomo)	CRP	Aminoglicosídeo Fluoroquinolona; Macrolídeo; Penam
-----	---------------------------------------	-------------------------	-----	---

Fonte: Elaborado pelo autor.

Tabela 11: Resistoma do ponto PJ3.

Número Contig	Banco de Dados	Start	End	Fita	Gene	GAPS	% de cobertura	% de identidade	Acesso
1112	NCBI CARD ResFinder	538	1483	-	<i>cfxA3</i>	4/4	97,62	98,20	AF472622:52-1018
1712	CARD	6112	6743	+	<i>CRP</i>	3/3	99,68	87,70	AP009048.1:4154296-4153663
2028	CARD	41700	45188	-	<i>rpoB2</i>	21/65	98,62	80,32	AP006618.1:4835199-4838688
2658	CARD	273909	274239	-	<i>RbpA</i>	1/1	95,94	85,54	HQ203032:0-345
643	NCBI CARD ResFinder	28479	29318	-	<i>sul1</i>	0/0	100,00	100,00	JF969163:1053-1893
	NCBI	24864	25746	+	<i>bla-A</i>	1/1	100,00	99,89	NG_047483.1
78	CARD	77853	78538	-	<i>mtrA</i>	1/1	99,85	87,05	AL123456.3:3627349-3626662
2231	NCBI	7591	8465	-	<i>bla-A</i>	20/25	98,19	92,93	NG_047483.1

Fonte: Elaborado pelo autor.

Tabela 12: BLAST NCBI das *contigs* e perfil de resistência dos genes do ponto PJ3.

Número Contig	Identificação taxonômica por BLAST NCBI (% identidade)	Acesso	Gene	Resistência
---------------	---	--------	------	-------------

1112	<i>Segatella copri</i> (91,30%)	CP152352.1	<i>CfxA3</i>	Cefamicina
1712	<i>Klebsiella pneumoniae</i> (96,76%)	CP031792.1	<i>CRP</i>	Fluoroquinolona; Macrolídeo; Penam
2028	<i>Mycolicibacterium</i> (≈82,00%) <i>Mycobacterium</i> (≈82,00%)	CP019882.1 CP107931.1	<i>rpoB2</i>	Rifamicina
2658	<i>Mycolicibacterium</i> (≈80,00%) <i>Mycobacterium</i> (≈80,00%)	CP050192.1 AP043665.1 AP020326.1	<i>RbpA</i>	Rifamicina
643	<i>Acidovorax</i> sp. (98,89%)	CP016447.1	<i>sul1</i> <i>bla-A</i>	Sulfonamida Betalactâmicos
78	<i>Mycolicibacterium</i> (≈79,00%) <i>Mycobacterium</i> (≈79,00%)	CP155796.1 CP192587.1	<i>mtrA</i>	Macrolídeo; Penam
2231	<i>Edwardsiella tarda</i> (95,12%) <i>Aeromonas salmonicida</i> (95,12%)	MF925338.1 ON814113.1	<i>bla-A</i>	Betalactâmicos

Fonte: Elaborado pelo autor.

6.2. Amostras das regiões do Amapá

6.2.1. Extração, quantificação e avaliação da pureza do DNA metagenômico

Os resultados de massa do pellet inicial, concentração final de DNA e a razão de pureza A260/A280 para cada amostra estão detalhados na Tabela 13.

Tabela 13: Extração de DNA.

Pontos	Pellet (g)	Concentração de DNA (ng/μl)	A260/A280
AP1	68,2	31,5	1,87
AP2	86,6	36,2	1,81
AP3	69,0	20,3	1,79
AP4	40,2	54,9	1,83
AP5	15,0	6,1	1,41
AP6	199,1	28,0	1,77
AP7	35,4	14,9	1,65
AP8	16,4	13,5	1,50
AP9	46,5	61,3	1,82
AP10	34,4	21,6	1,66
AP11	23,7	6,6	1,51

Fonte: Elaborado pelo autor.

Observou-se ampla variação entre os pontos, tanto no rendimento quanto na qualidade do DNA; por exemplo, AP1, AP2 e AP6 apresentaram maiores quantidade de pellet, 68,2 g, 86,6 g e 199,1 g, respectivamente, e concentração intermediária de DNA, 31,5 ng/μl, 36,2 ng/μl e 28,0 ng/μl, respectivamente, enquanto AP5, AP8 e AP11 apresentaram os menores valores da quantidade de pellet, 15,0 g, 16,4 g e 23,7 g, respectivamente, e as menores concentração de DNA, com 6,1 ng/μl, 13,5 ng/μl e 6,6 ng/μl, respectivamente.

A pureza do DNA, avaliada pela razão A260/A280, também foi analisada. Valores entre 1.8 e 2.0 são considerados indicativos de alta pureza, com baixa contaminação por proteínas. As amostras AP1, AP2, AP4 e AP9 apresentaram valores ideais (1.87, 1.81, 1.83 e 1.82, respectivamente). Em contrapartida, as amostras AP5, AP8, e AP11 exibiram valores mais baixos (1.41, 1.50, 1.51,

respectivamente) sugerindo a presença de contaminantes residuais ou baixa eficiência no processo de purificação.

Para realizar o sequenciamento, o fabricante recomenda um *input* de 200 ng de DNA em 10 µl de água livre de nuclease com razão de pureza A260/A280 >1.8, para garantir a integridade e quantidade suficiente para a biblioteca na *flowcell*. No entanto, devido aos resultados diferentes da extração de DNA e para garantir o *input* da biblioteca, optou-se por sequenciar as amostras que não atingiram a concentração mínima e grau de pureza recomendada (AP5, AP7, AP8 e AP11) em um momento diferente das demais (1º sequenciamento). Para assim, avaliar a viabilidade e o rendimento do sequenciamento com material de partida subótimo, uma vez que a obtenção de mais material biológico para essas amostras não era viável.

Posteriormente, as sete amostras com concentração e pureza adequadas (AP1, AP2, AP3, AP4, AP6, AP9 e AP10) foram submetidas a outro sequenciamento (2º sequenciamento). Durante este processo, a primeira corrida de sequenciamento com este grupo de amostras foi interrompida prematuramente, resultando em um rendimento de dados inferior ao esperado. Para contornar essa limitação e maximizar a profundidade de sequenciamento, uma outra corrida foi realizada utilizando as mesmas bibliotecas (3º sequenciamento). Assim, obteve-se os dados dos sequenciamentos 1º, 2º e 3º após 20h 58 min, 9h 44 min e 48h, respectivamente. Subsequentemente, os arquivos FASTQ brutos das corridas foram concatenados para cada amostra correspondente, consolidando os dados em um único conjunto para as análises posteriores.

A extração de DNA apresentou variações na massa do pellet, concentração e pureza entre os diferentes pontos amostrados. A variação observada pode ser atribuída à diversidade dos ambientes amostrados, fatores como a densidade de biomassa microbiana, a quantidade de matéria orgânica e inorgânica em suspensão e a presença de sedimentos que influenciam o rendimento da extração. O kit *Qiagen DNeasy PowerSoil Pro*, embora seja recomendado para solos, apresenta alto rendimento de DNA não inibido de amostras de água, tendo um impacto menos significativo na avaliação da comunidade microbiana (Reume *et al.*, 2025).

Embora tenha ocorrido a interrupção da corrida de sequenciamento, os dados gerados puderam ser avaliados sem comprometer ou enviesar as análises. Purushothaman *et al.* (2024) ao testarem diversos kits de extração de DNA para amostras de microbiomas e isolados utilizando o sequenciamento *Oxford Nanopore* para a detecção de resistência e taxonomia, concluiu que após 1,9 horas de sequenciamento, é possível identificar microrganismos de amostras complexas e detectar de genes de resistência.

6.2.2. Controle de Qualidade e Processamento dos Dados Brutos

Os sequenciamentos brutos dos onze pontos foram avaliados para verificar a qualidade utilizando as ferramentas NanoPlot (Tabela 14). e Falco. Foram gerados aproximadamente 9,0 GB de dados, aproximadamente, e distribuídos em 2.420.537 *reads*.

A análise das métricas obteve uma variação esperada entre as amostras. O comprimento médio das leituras e os valores de N50 foram maiores nos pontos AP9 e AP10, que apresentaram médias superiores a 2.000 bases e N50 acima de 4.000 bases. As demais amostras viáveis exibiram valores de N50 entre a faixa de 2.000 a 3.700 bases. Em relação à acurácia, cinco pontos (AP4, AP7, AP8, AP10 e AP11) apresentaram mais de 50% de suas leituras com *score* de qualidade *Phred* superior a Q10.

O ponto AP5 demonstrou um desempenho inferior em todos os parâmetros avaliados, o sequenciamento resultou em 3,0 MB, com 3.281 *reads* geradas, N50 de 443 bases e 5,8% das *reads* com qualidade superior a Q10. Com base nesta análise quantitativa, a amostra AP5 foi excluída das análises subsequentes.

Os dados brutos foram filtrados para remoção dos adaptadores de *barcode* das extremidades das *reads* utilizando a ferramenta *Porechop* e posteriormente ao Filtlong. A eficácia foi confirmada pela análise da qualidade (Tabela 15).

O processo de filtragem resultou na remoção de aproximadamente 52% do total de *reads*, restando um conjunto final de 1.138.666 *reads*. O comprimento médio das leituras apresentou um aumento expressivo em todos os pontos, com valores variando entre 2.619 (AP11) e 3.753 (AP9) bases.

Os valores de N50 aumentaram, os pontos AP2, AP3 e AP4 apresentaram valores acima de 3.500 bases, já AP1, AP7, AP10 acima de 4.000 bases e AP9 apresentando os maiores valores com 5.186 bases. Todos os pontos aumentaram no percentual de *reads* acima do Q10.

As Figuras 10 - 20 apresentam, de maneira comparativa, a avaliação da qualidade das *reads* geradas para cada amostra, utilizando o Falco. Nos gráficos de cada ponto, o painel (a) demonstra os dados brutos, evidenciando que as posições iniciais das sequências apresentam baixas medianas de qualidade e com grande amplitude de variação, frequentemente abaixo de Q10, esse cenário inicial pode indicar a presença dos adaptadores e *barcodes* ainda não removidos.

Após o processamento, como apresentado nos painéis (b), verifica-se um aumento considerável nos *scores* medianos de qualidade em todas as posições das *reads*. Esse perfil pós-filtragem evidencia a remoção eficiente de adaptadores e *barcodes*, além do descarte das *reads* abaixo de 1.000 pb.

A extração de DNA, assim como o sequenciamento metagenômico, está sujeita à divergência entre as amostras, pois a influência de fatores ambientais e do tipo de amostra pode afetar diretamente o rendimento e qualidade do DNA ambiental (Parsley *et al.*, 2024; Demkina *et al.*, 2023; Mathieu *et al.*, 2020). A variabilidade observada em estudos ambientais é esperada, principalmente considerando que diferentes corpos d'água apresentam características físico-químicas distintas que impactam diretamente na biomassa microbiana e na qualidade do material genético extraído (Sahu *et al.*, 2025; Goldberg *et al.*, 2016).

Após a filtragem, observou-se um aumento significativo nos comprimentos médios dos *reads* e nos respectivos valores de N50 de todas as amostras analisadas, com retenção predominante de *reads* de maior comprimento e melhor qualidade. A utilização de ferramentas para filtragem e controle de qualidade em sequenciamentos de dados *long reads* promove uniformidade nos dados e diminui a fragmentação das *reads* que possam gerar taxas de erros (Demkina *et al.*, 2023; Priyam *et al.*, 2021). O processo de filtragem rigoroso é considerado crucial para análises metagenômicas robustas, especialmente quando se utiliza tecnologia *Oxford Nanopore*, que apresenta características únicas de erro que podem ser mitigadas através de estratégias apropriadas de processamento (Szakállas *et al.*, 2024; De La Cerda *et al.*, 2023).

Embora tenha aumentado o N50 após a filtragem, as *reads* das amostras ainda apresentaram fragmentações com variações entre 3.000 a 5.000 bases. A estrutura de erro do sequenciamento *Nanopore* pode resultar em *assemblies* fragmentados, sendo necessário aumentar a profundidade dos dados e aplicar estratégias de correção para otimizar a qualidade do *assembly* (Benoit *et al.*, 2024; Baker, 2022). Alguns estudos demonstram que a profundidade de cobertura adequada é essencial para obter *assemblies* de alta qualidade, pois contorna os desafios de montagem presentes em amostras de comunidades microbianas complexas, como nas análises metagenômica (De La Cerda *et al.*, 2023; Lapidus; Korobeynikov, 2021).

Tabela 14: Análise da qualidade do sequenciamento de saída bruto.

Pontos	Fastq	Número de reads	Número de pares de bases	Comprimento médio das reads	Desvio Padrão do comprimento	N50	Phread >Q10	% > Q10
AP1	758,8 MB	201.362,0	357.116.132,0	1.773,5	2.609,1	3.685,0	73.948,0	36,7 %
AP2	1,6 GB	443.753,0	749.914.899,0	1.689,9	1.927,3	3.234,0	181.823,0	41,0 %
AP3	449.1 MB	145.653,0	206.027.319,0	1.414,5	1.865,3	3.124,0	70.025,0	48,1%
AP4	1,7 GB	452.657,0	818.095.412,0	1.807,3	1.984,4	3.319,0	254.764,0	56,3%
AP5	3,0 MB	3.281,0	933.642,0	284,6	594,6	443,0	190,0	5,8%
AP6	656.2 MB	203.717,0	302.786.062,0	1.486,3	1.594,2	2.695,0	61.713,0	30,3%
AP7	321.0 MB	90.966,0	149.914.958,0	1.648,0	1.982,5	3.256,0	45.976,0	50,5%
AP8	452.5 MB	129.954,0	210.981.457,0	1.623,5	1.927,4	3.126,0	65.927,0	50,7%
AP9	1.0 GB	229.347,0	492.282.767,0	2.146,5	2.860,8	4.547,0	106.269,0	46,3%
AP10	1.3 GB	307.840,0	645.784.623,0	2.097,8	2.417,7	4.031,0	170.124,0	55,3%
AP11	777.2 MB	212.007,0	364.648.986,0	1.720,0	1.664,1	2.653,0	118.045,0	55,7%

Fonte: Elaborado pelo autor.

Tabela 15: Análise da qualidade do sequenciamento de saída pós-processamento e filtragem.

Pontos	Número de reads	Número de pares de bases	Comprimento médio das reads	Desvio Padrão do comprimento	N50	Phread >Q10	% > Q10
AP1	88.493,0	293.801.509,0	3.320,1	3.250,1	4.336,0	58.663,0	66,3%
AP2	201.822,0	611.650.536,0	3.030,6	2.082,8	3.820,0	138.874,0	68,8%
AP3	53.063,0	162.857.971,0	3.069,1	2.196,3	3.946,0	40.269,0	75,9%
AP4	227.589,0	693.310.322,0	3.046,3	2.096,6	3.819,0	174.689,0	76,8%
AP6	85.922,0	232.557.156,0	2.706,6	1.690,1	3.265,0	54.431,0	63,3%
AP7	37.907,0	117.978.463,0	3.112,3	2.243,6	4.009,0	32.506,0	85,8%
AP8	55.130,0	166.427.293,0	3.018,8	2.161,0	3.799,0	47.103,0	85,4%
AP9	113.030,0	424.234.762,0	3.753,3	3.304,0	5.186,0	79.882,0	70,7%
AP10	164.730,0	566.472.000,0	3.438,8	2.573,1	4.495,0	122.597,0	74,4%
AP11	110.980,0	290.710.947,0	2.619,5	1.745,6	3.072,0	91.933,0	82,8%

Fonte: Elaborado pelo autor.

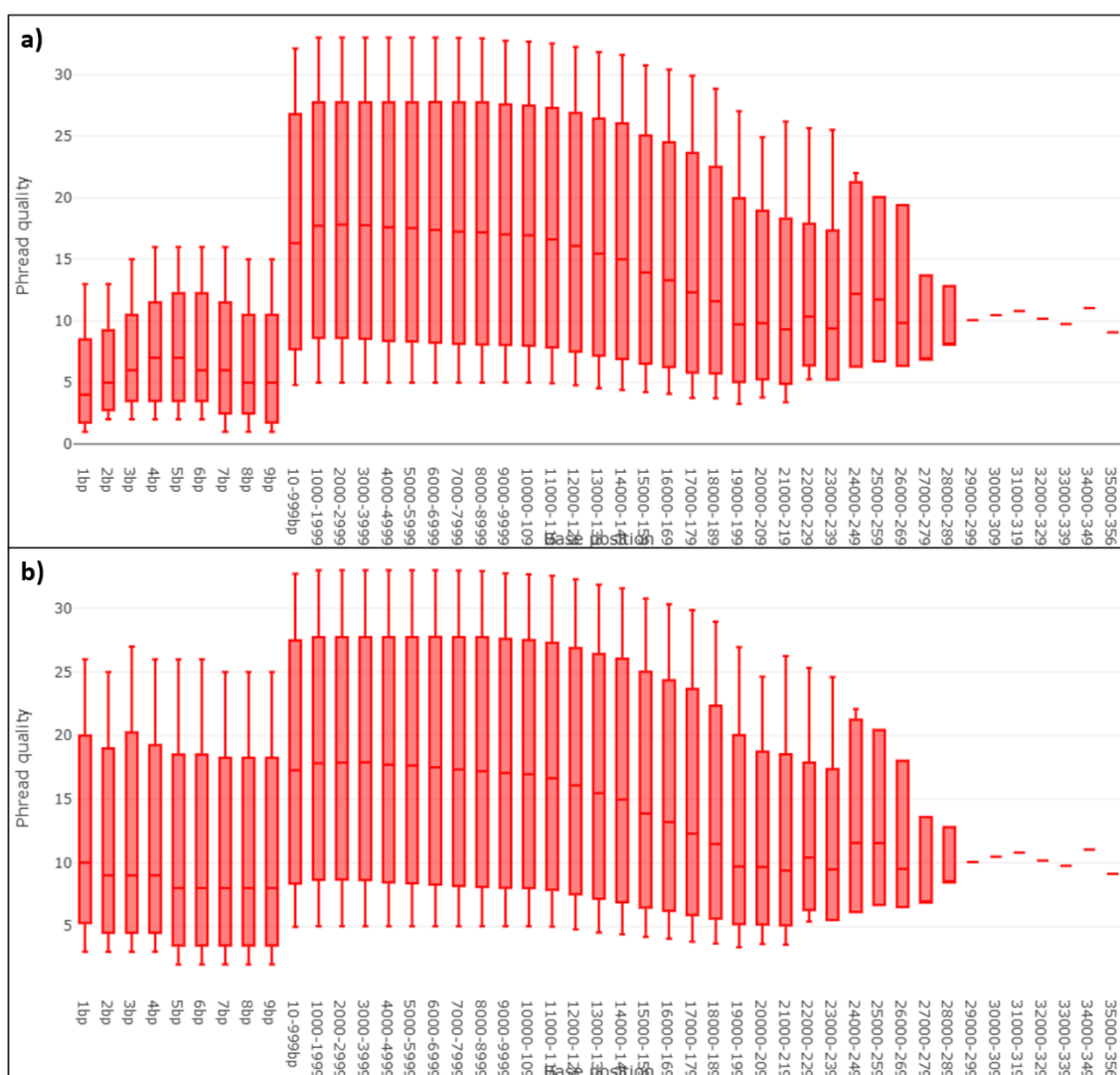
Figura 10: Qualidade por sequência base de AP1.



Fonte: Elaborado pelo autor.

Legenda: a) Qualidade dos dados brutos, b) Qualidade dos dados pós-processamento.

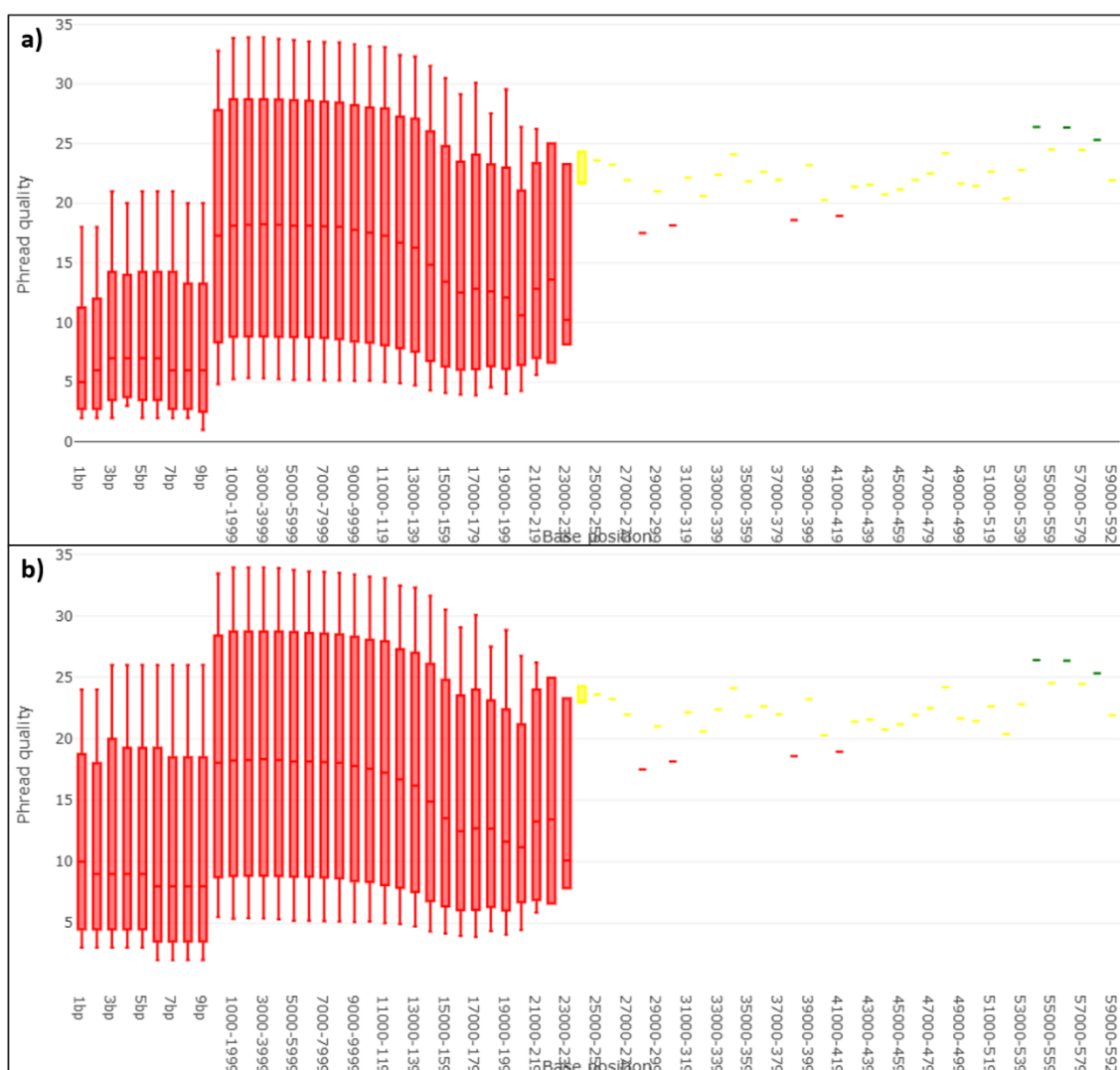
Figura 11: Qualidade por sequência base de AP2.



Fonte: Elaborado pelo autor.

Legenda: a) Qualidade dos dados brutos, b) Qualidade dos dados pós-processamento.

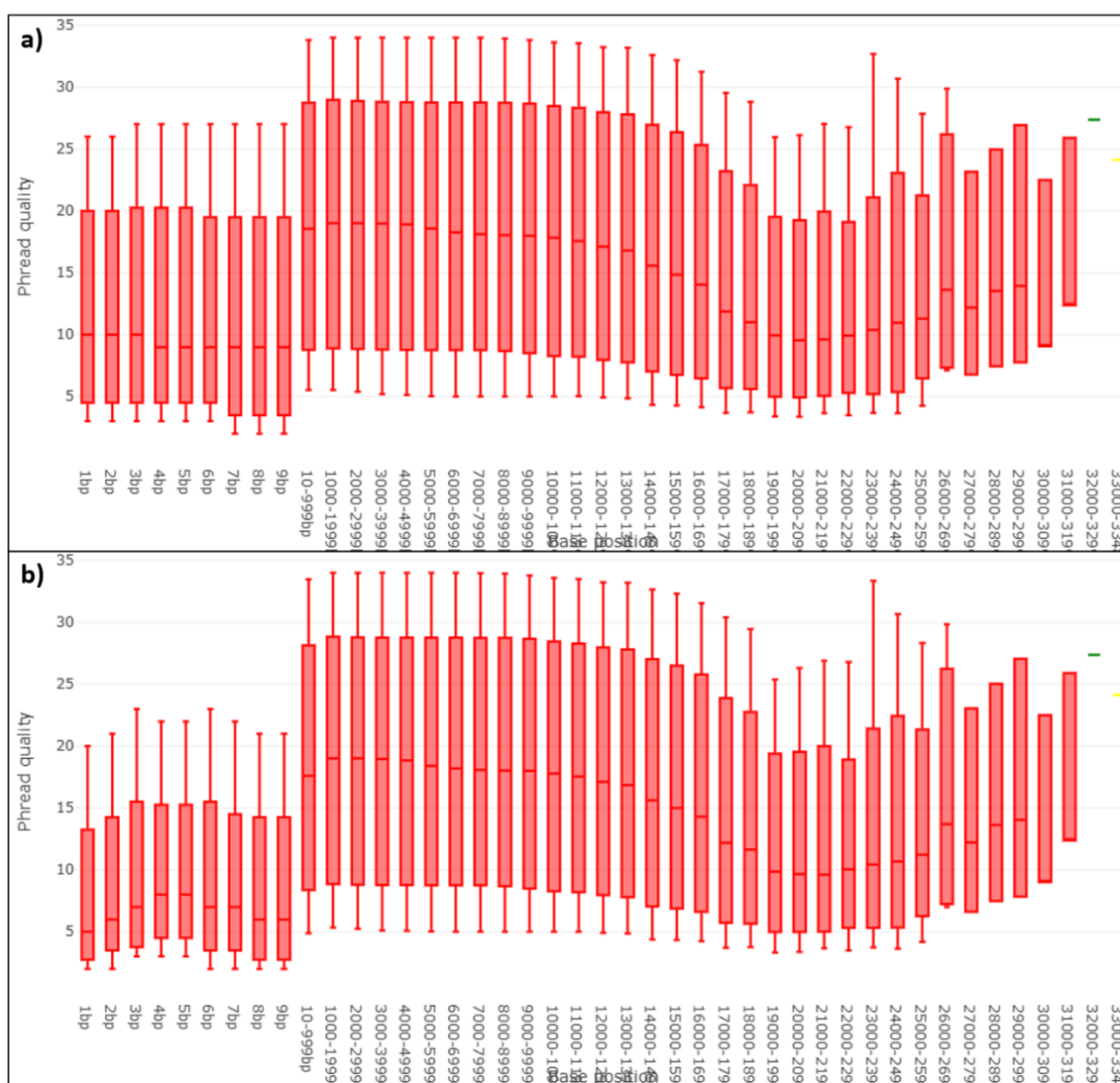
Figura 12: Qualidade por sequência base de AP3.



Fonte: Elaborado pelo autor.

Legenda: a) Qualidade dos dados brutos, b) Qualidade dos dados pós-processamento.

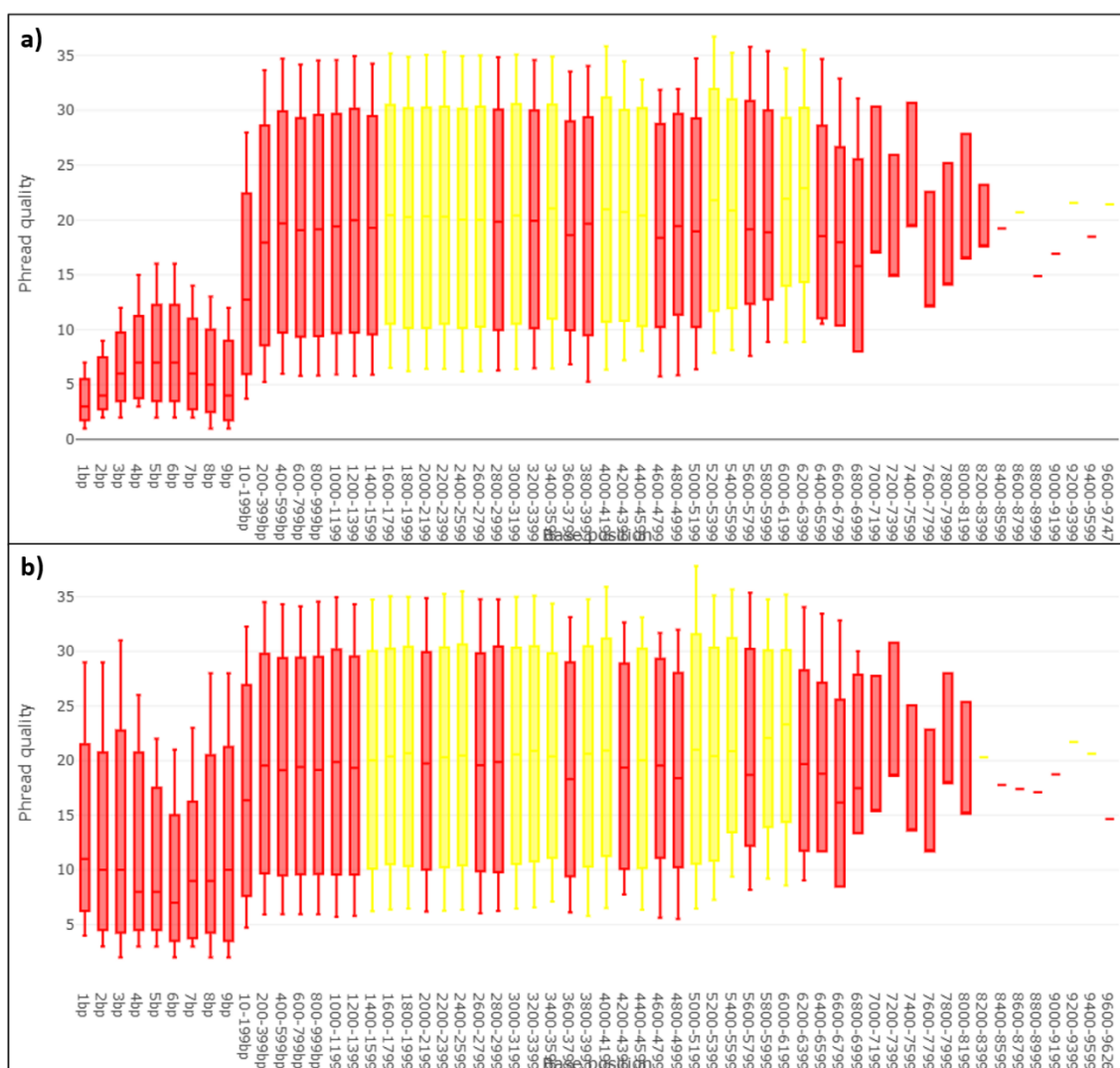
Figura 13: Qualidade por sequência base de AP4.



Fonte: Elaborado pelo autor.

Legenda: a) Qualidade dos dados brutos, b) Qualidade dos dados pós-processamento.

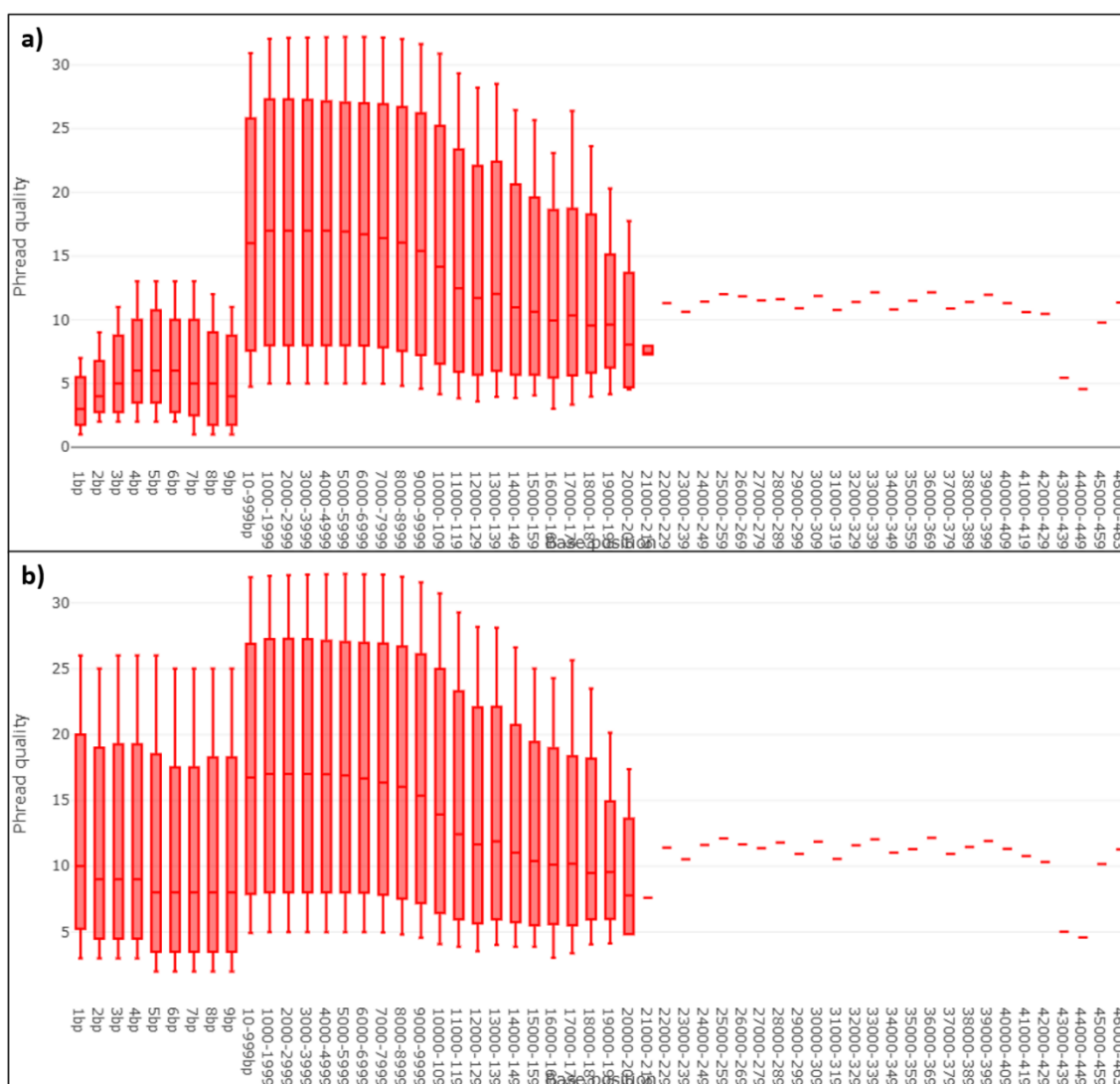
Figura 14: Qualidade por sequência base de AP5.



Fonte: Elaborado pelo autor.

Legenda: a) Qualidade dos dados brutos, b) Qualidade dos dados pós-processamento.

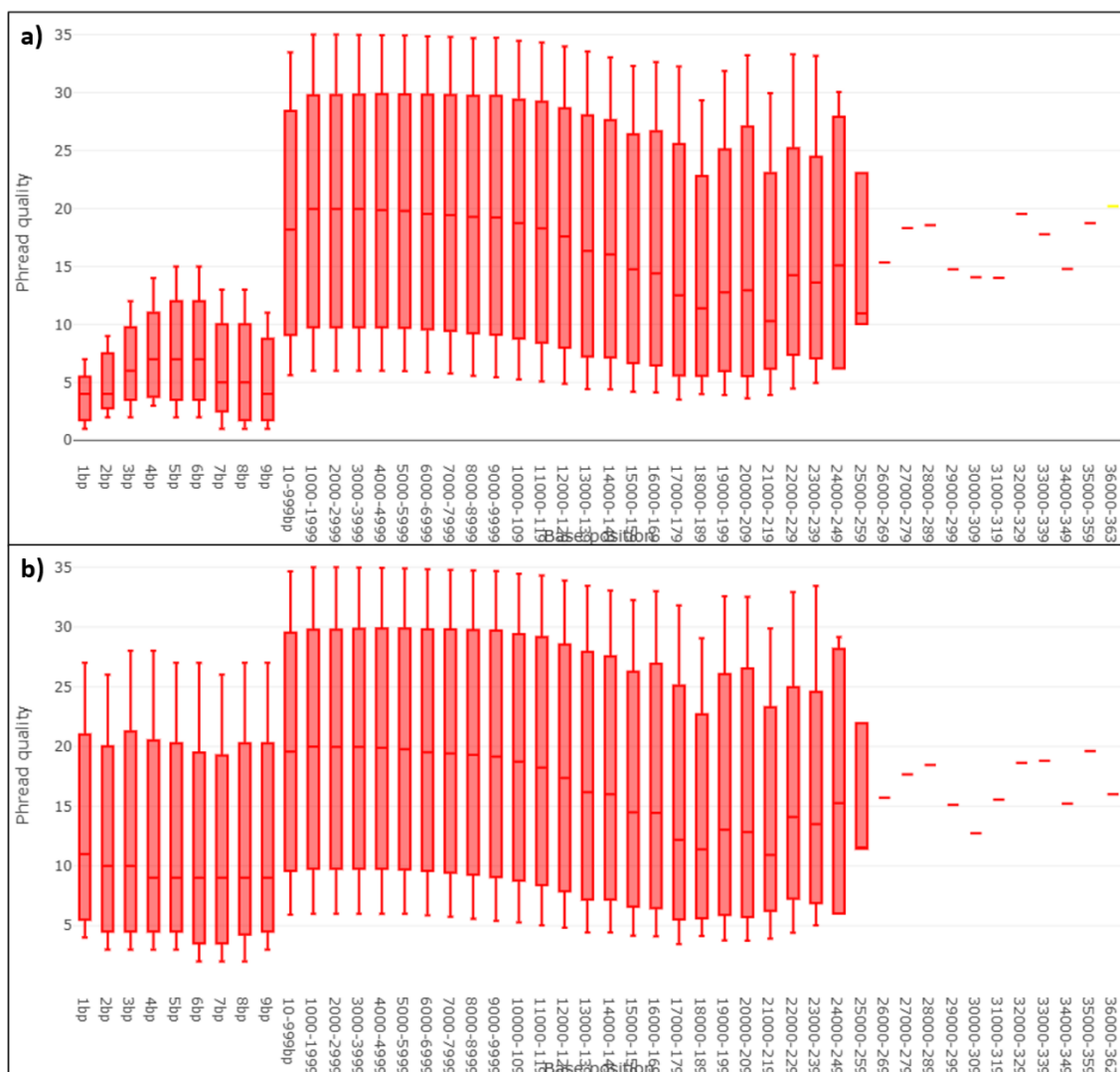
Figura 15: Qualidade por sequência base de AP6.



Fonte: Elaborado pelo autor.

Legenda: a) Qualidade dos dados brutos, b) Qualidade dos dados pós-processamento.

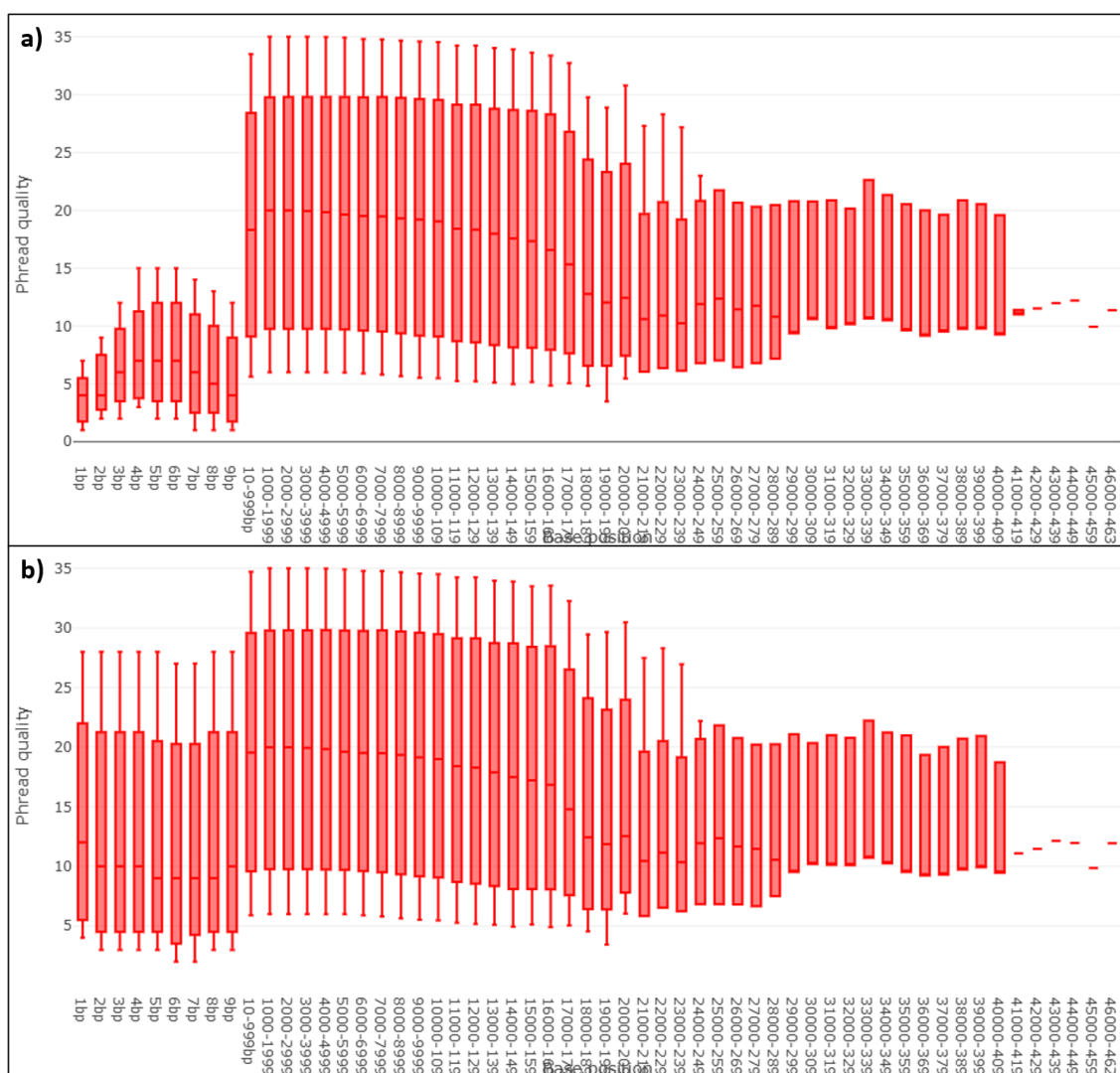
Figura 16: Qualidade por sequência base de AP7.



Fonte: Elaborado pelo autor.

Legenda: a) Qualidade dos dados brutos, b) Qualidade dos dados pós-processamento.

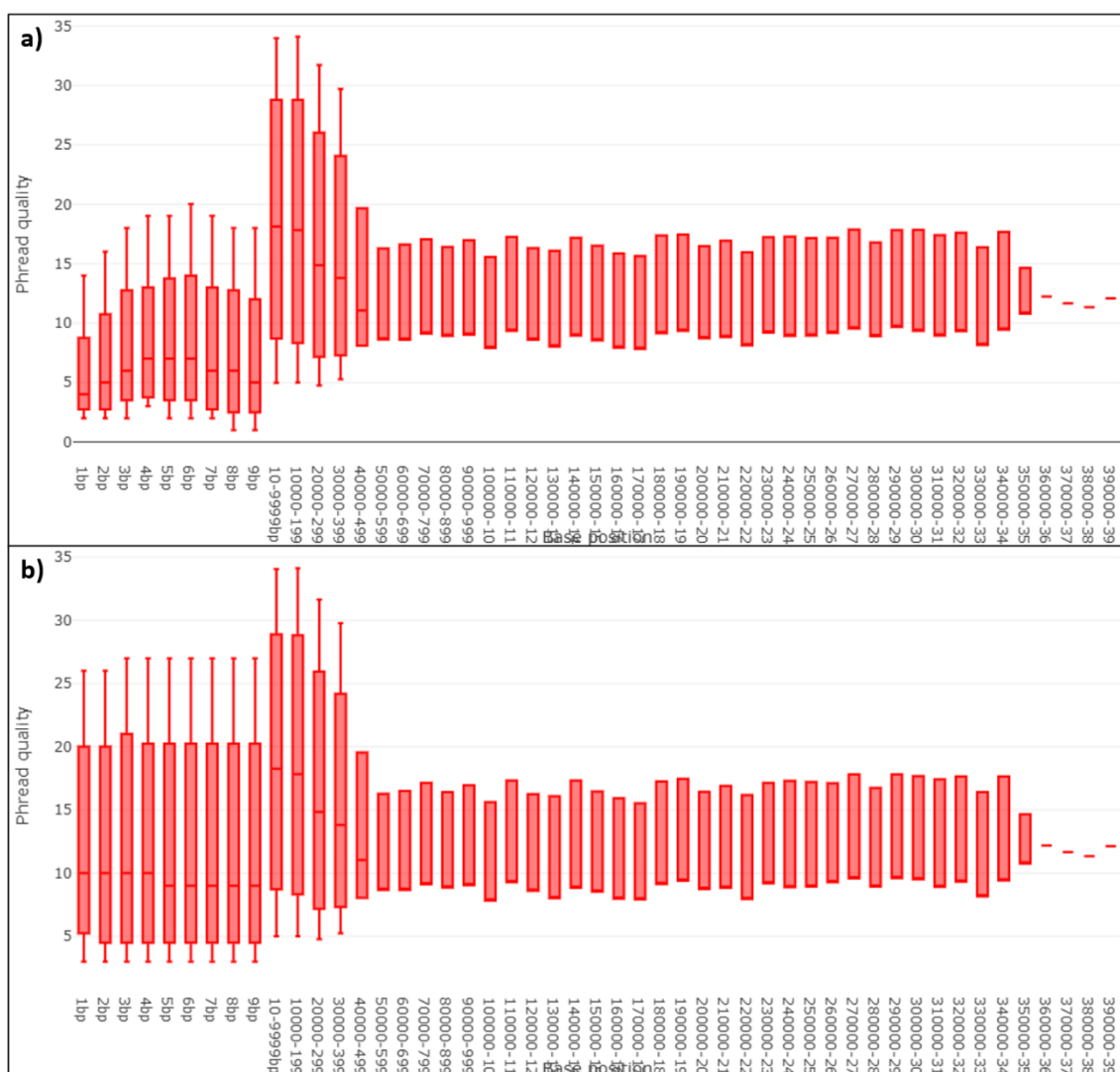
Figura 17: Qualidade por sequência base de AP8.



Fonte: Elaborado pelo autor.

Legenda: a) Qualidade dos dados brutos, b) Qualidade dos dados pós-processamento.

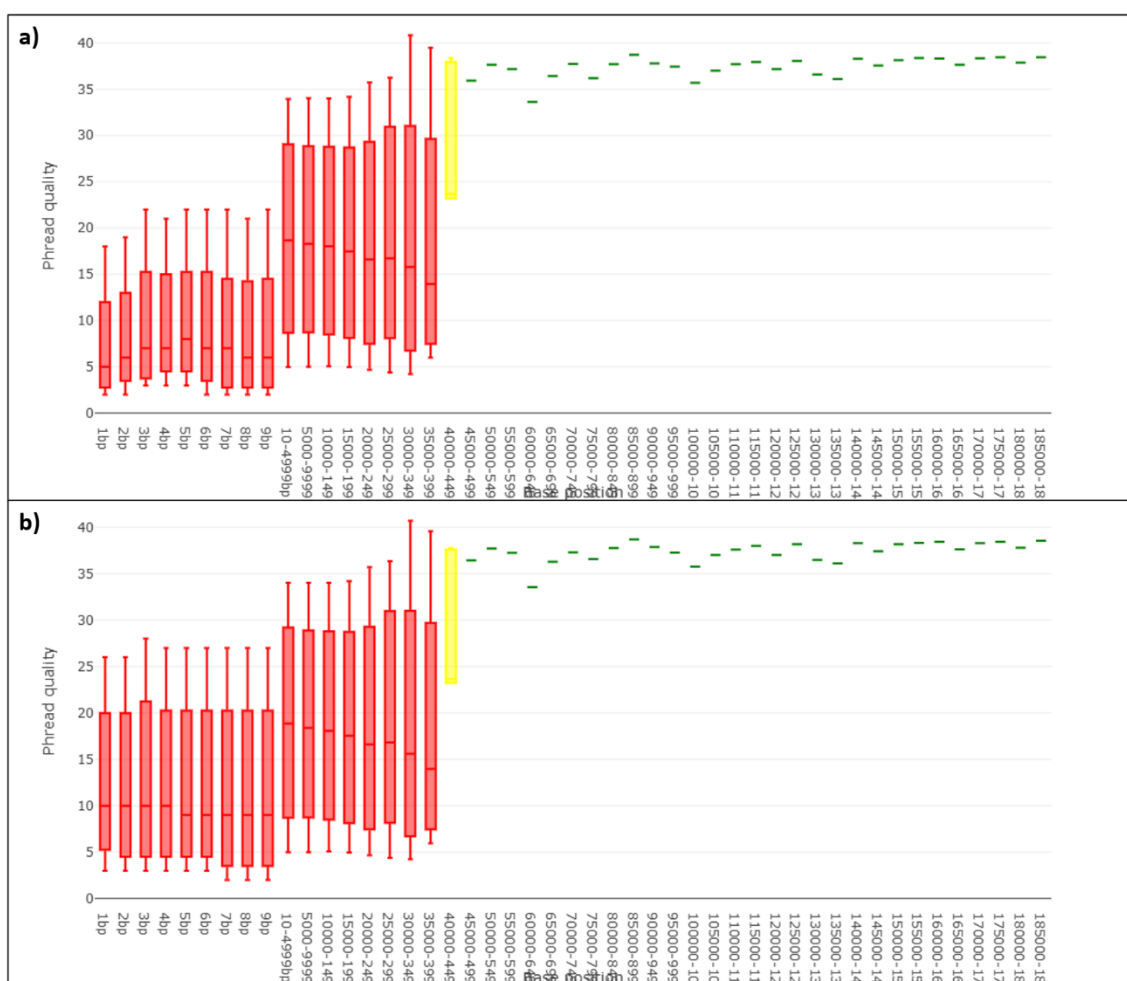
Figura 18: Qualidade por sequência base de AP9.



Fonte: Elaborado pelo autor.

Legenda: a) Qualidade dos dados brutos, b) Qualidade dos dados pós-processamento.

Figura 19: Qualidade por sequência base de AP10.



Fonte: Elaborado pelo autor.

Legenda: a) Qualidade dos dados brutos, b) Qualidade dos dados pós-processamento.

Figura 20: Qualidade por sequência base de AP11.

Fonte: Elaborado pelo autor.

Legenda: a) Qualidade dos dados brutos, b) Qualidade dos dados pós-processamento.

6.2.3. Montagem

A montagem dos dados metagenômicos consistiu na reconstrução de *contigs* a partir das leituras processadas, pela abordagem *De Novo*. As montagens utilizando o Flye (Tabela 16) apresentaram baixos números de *contigs*, variando de 105 a 2.051 bases, indicando alta contiguidade das sequências, sobretudo nas amostras AP1, AP3 e AP6, porém, também obtiveram os menores tamanhos de *contigs*. Isso reflete no L50, que também foram menores, para os pontos AP1 (73) e AP3 (34), ou seja, precisaria desses números de *contigs* para atingir 50% do comprimento total.

O ponto AP10 apresentou os maiores valores de *contigs*, tamanho e comprimento total (2.051, 979.978 e 36.396.380 bases, respectivamente).

Os pontos AP9 e AP10 obtiveram os maiores valores de N50, demonstrando que metade do genoma montado (50% do comprimento total) está contida em *contigs* que têm um tamanho de pelo menos 24.793 e 24.365 bases, respectivamente. Os valores de N50, que variaram de 3.264 a 24.365 bases, refletem *contigs* representativos de grandes regiões do genoma e, em geral, sugerem montagens menos fragmentadas. Adicionalmente, os altos valores de auN, como para os pontos AP10 (66.505) e AP11 (79.093), consolidam a avaliação de alta qualidade, fornecendo uma métrica integrada da contiguidade em toda a extensão da montagem, sendo a amostra AP3 (6.999) com a qualidade de montagem mais baixa. Em todas as amostras, a ausência de bases indefinidas ('N's) apresenta a qualidade no consenso das sequências.

Já nas montagens com MEGAHIT, apresentaram mais fragmentações, com número de *contigs* elevado (de 847 a mais de 122 mil bases), valores de N50 geralmente baixos (entre 665 e 8.788 bases) e maior dispersão nos tamanhos dos maiores *contigs*. Em geral, o MEGAHIT é indicado para amostras metagenômicas de alta complexidade, realizada para sequenciamentos de *short reads*, pois consegue montar sequências de diversas espécies, ainda que com maior fragmentação e menor contiguidade. Também foi observado que as métricas L50 e L90 aumentaram, indicando a necessidade de um maior número de *contigs* para se alcançar 50% ou 90% da montagem total.

Ao comparar as duas ferramentas, observou-se que o Flye favoreceu montagens mais contínuas com menores quantidades de *contigs* e maiores

índices de N50 e auN, mostrando-se eficiente na reconstrução de grandes regiões genômicas, especialmente em genomas mais prevalentes ou menos complexos dentro das comunidades, porém durante o processo de montagem, descartou a maioria das *reads* que não entraram em consenso com as *contigs*. Por sua vez, o MEGAHIT demonstrou vantagens na recuperação da diversidade metagenômica, porém gerou grandes quantidades de *contigs* fragmentadas.

A ausência de bases indefinidas em ambas as abordagens reafirma a alta qualidade dos consensos sequenciais. Esses resultados corroboram com relatos da literatura sobre as diferenças funcionais e estratégias dessas ferramentas em metagenômica aplicada (Wang *et al.*, 2021; Lin *et al.*, 2021).

Assim, as opções entre Flye e MEGAHIT deve considerar o objetivo do estudo: montagens mais contínuas para sequências dominantes ou maior abrangência metagenômica com maior fragmentação. A complementaridade dessas ferramentas pode ser explorada para melhorar a interpretação integral de dados metagenômicos ambientais.

Tabela 16: Métricas de montagem Flye.

Pontos	Número de <i>contigs</i> > 500 pb	Maior contig	Comprimento total	Contiguidade					Bases indefinidas (N's)
				N50	N90	auN	L50	L90	
AP1	245	37.792	2.001.676	9.949	4.669	10.774	73	190	0
AP2	934	46.166	7.832.033	9.980	4.545	11.955	254	709	0
AP3	105	13.070	601.975	6.716	3.264	6.999	34	83	0
AP4	1.140	82.374	10.768.187	11.153	4.836	16.000	277	844	0
AP6	603	40.966	4.378.477	9.429	3.422	11.463	149	443	0
AP7	207	104.485	2.879.265	19.066	5.976	31.375	36	140	0
AP8	583	170.805	6.375.967	15.081	5.150	31.145	104	397	0
AP9	1.178	607.620	21.618.541	24.793	8.744	61.120	228	808	0
AP10	2.051	979.978	36.396.380	24.365	8.684	66.505	423	1.407	0
AP11	1.760	896.985	18.529.749	14.223	4.580	79.093	257	1.175	0

Fonte: Elaborado pelo autor.

Tabela 17: Métricas de montagem MEGAHIT.

Pontos	Número de <i>contigs</i> > 500 pb	Maior contig	Comprimento total	Contiguidade					Bases indefinidas (N's)
				N50	N90	auN	L50	L90	
AP1	69.145	16.725	126.632.440	2.449	857	3.088,3	15.518	50.149	0
AP2	201.307	16.416	281.097.378	1.683	702	2.176,5	50.386	154.053	0
AP3	29.442	18.899	59.949.998	2.801	944	3.483,3	6.447	20.946	0
AP4	194.637	15.719	295.221.283	1.894	739	2.443,0	46.582	146.372	0
AP6	45.980	15.563	79.386.463	2.284	812	2.849,5	10.517	33.683	0
AP7	20.887	22.324	34.521.666	2.215	745	3.085,9	4.344	15.245	0
AP8	30.884	18.812	47.765.353	2.014	717	2.764,6	6.696	22.902	0
AP9	96.552	26.712	148.972.257	2.019	704	2.866,6	20.339	71.424	0
AP10	122.282	32.141	199.220.343	1.772	665	2.519,5	27.096	92.272	0
AP11	50.670	16.986	69.723.827	1.682	668	2.285,9	12.019	38.586	0

Fonte: Elaborado pelo autor

6.2.4. Microbioma Flye

A montagem proveniente do Flye, gerou um arquivo FASTA para cada amostra, que foi classificada taxonomicamente utilizando o Kraken2 (Relatório_completo_:<https://docs.google.com/spreadsheets/d/1IBIT3ySJcliMAxckP0HdsIWuWAQX_3NolCw0_zXL0rg/edit?usp=sharing>) e foram processados pelo Kraken-report e Krakentools. Posteriormente, as informações foram transformadas em dados gráficos através do Pavian.

A análise taxonômica do ponto AP1 (Figura 21), revelou que 60,41% das *contigs* foram classificadas, enquanto 39,59% permaneceram não classificadas, o que pode ser comum em estudos metagenômicos ambientais devido às limitações dos bancos de dados de referência. Entre as leituras classificadas, a grande maioria pertence ao domínio Bactéria (57,96%), com destaque para o filo *Pseudomonadota* (23,67%), e classe *Gammaproteobacteria* (12,65%) e *Alphaproteobacteria* (6,94%). Dentro desse grupo, gêneros como *Polynucleobacter* (1,63%) e *Pseudomonas* (1,22%) foram identificados, sendo que a espécie *P. monteilii* correspondeu a 0,41% das leituras. Também foram detectados representantes de *Actinomycetota* (4,49%), *Planctomycetota* (0,41%) e *Chloroflexota* (0,82%). Além das bactérias, arqueias do filo *Thermoproteota* (0,82%) também foram identificadas, com destaque para o gênero *Nitrosotenuis*.

No ponto AP2 (Figura 22), 69,49% das *contigs* foram classificadas, com 68,20% pertencente ao domínio Bacteria. O filo predominante foi o *Pseudomonadota* (29,23%), com classes *Gammaproteobacteria* (12,42%) e *Alphaproteobacteria* (11,99%). Dentro desse grupo, gêneros como *Polynucleobacter* (2,68%) e *Pseudomonas* (2,03%). Também foram detectados representantes de *Actinomycetota* (7,28%), incluindo gêneros como *Mycobacterium* e *Paenibacillus*, além de filos como *Nitrospirota*, *Chloroflexota*, *Acidobacteriota*, *Verrucomicrobiota*, *Planctomycetota* e *Bacillota*, embora apresente diversidade microbioana, todos apareceram em baixas proporções.

A triagem taxonômica do ponto AP3 (Figura 23) revelou que 74,29% das sequências puderam ser atribuídas a táxons conhecidos, o qual 72,38% foi para o domínio Bactéria. O filo predominante identificado foi *Pseudomonadota*, compreendendo 27,62% do total de leituras classificadas, sendo as classes mais

abundantes *Gammaproteobacteria* (10,48%) e *Alphaproteobacteria* (7,62%). A ordem *Burkholderiales* (6,67%) foi a mais representativa dentro de *Gammaproteobacteria*, seguida por *Pseudomonadales* (0,95%), que inclui o gênero *Pseudomonas*. Outros filos bacterianos relevantes detectados incluem *Actinomycetota* (8,57%) e *Acidobacteriota* (0,95%).

Para o ponto AP4 (Figura 24), 52,19% foram classificadas, sendo 49,56% para o domínio Bactéria, com 29,21% para o filo *Pseudomonadota*, com a classe *Gammaproteobacteria* (19,39%) e *Alphaproteobacteria* (7,28%). Dentro da classe *Gammaproteobacteria*, o gênero predominante foi *Polynucleobacter* (6,84%), incluindo a espécie *P. tropicus*.

Em paralelo, foram identificados uma diversidade microbiana dos filos *Actinomycetota* (3,51%), *Bacteroidota* (0,53%), *Nitrospirota* (0,35%), *Acidobacteriota* (0,09%), *Planctomycetota* (0,09%) e *Chloroflexota* (0,09%). Além dos táxons bacterianos, arqueias do filo *Thermoproteota* (2,02%), sobretudo o gênero *Nitrosotenuis* (2,02%), também foram identificadas.

Foram classificadas 65,67% das *contigs* para AP6 (Figura 25), com 65,34% classificadas para o domínio Bactéria. Dentre o domínio, 56,72% corresponde ao filo *Pseudomonadota*. A classe *Gammaproteobacteria* (53,23%) predominou com as ordens *Burkholderiales* (48,92%) e a família *Burkholderiaceae* (45,94%) concentrado na maior parte das leituras, evidenciando um perfil dominado por gêneros aquáticos como *Polynucleobacter* (19,57%), com detecção da espécie *P. tropicus* (1,00%). Outros gêneros relevantes incluíram *Limnohabitans* e *Fluviibacter*. Além dos *Gammaproteobacteria*, foram identificadas pequenas proporções de *Alphaproteobacteria* (1,00%) e *Actinomycetota* (1,0%), incluindo representantes dos gêneros *Limnohabitans*, *Microbacterium* e *Nanopelagicales*.

O ponto AP7 (Figura 26), obteve 75,85% das *contigs* classificadas, que apresentou um predomínio do filo *Pseudomonadota* (69,57%), com abundância pela classe *Gammaproteobacteria* (67,63%), ordem *Burkholderiales* (64,25%) e família *Burkholderiaceae* (59,42%). O gênero mais abundante foi *Polynucleobacter* (27,54%), predominantemente pela espécie *P. tropicus* (1,93%).

Além disso, identificou-se a presença do gênero *Limnohabitans* (0,97%). Dentro da ordem *Pseudomonadales*, observou-se o gênero *Pseudomonas*

(0,48%). Outros filós detectados incluíram *Actinomycetota* (0,48%) e *Microbacteriaceae* (0,48%).

A amostra AP8 (Figura 27) apresentou 64,84% para Bacteria, com o filo *Pseudomonadota* correspondendo a 57,98% do total, com destaque para a família *Burkholderiaceae* com 47,00%. Que apresentou dois gêneros, o grupo dominante foi *Polynucleobacter*, com 8,92%, sendo principalmente representado pela espécie *P. tropicus* (0,34%; 2 leituras) e *Limnohabitans* (2,57%). Outros grupos bacterianos em menores proporções incluíram a classe *Alphaproteobacteria* (0,51%) e o filo *Actinomycetota* (0,51%), com 0,17% para família *Microbacteriaceae*.

Já o ponto AP9 (Figura 28), apresentou um perfil com predominância maior de organismos não classificados (76,66%). O filo *Pseudomonadota* foi detectado em 11,46%, com destaque para a ordem *Burkholderiales* (7,22%) que apresentou duas famílias, *Burkholderiaceae* (4,92%) com 0,34% para a espécie *P. tropicus* e *Methylophilaceae* (0,82%) com 0,76% para o gênero *Methylophilus*. Outras ordens estão a *Rhizobiales* (0,42%) com o gênero *Methylocystis* (0,25%) e *Shingomonadales* (0,08%).

O filo *Bacteroidota* também foi detectado em maior proporção comparado às amostras anteriores, com 3,06%, representado pela classe *Bacteroidia* (3,06%) e gênero *Sediminibacterium* (2,97%).

Outros filós bacterianos em menor abundância incluíram *Actinomycetota* (0,42%), com a ordem *Actinomycetes* (0,25%) e ordem *Nanopelagicales* (0,25%), e o filo *Verrucomicrobiota* (0,34%), representado pelas ordens *Limisphaerales* (0,25%) e *Chthoniobacterales* (0,08%).

Das 31,01% *contigs* classificadas para o ponto AP10 (Figura 29), 30,57% foram do domínio Bacteria que apresentou uma composição bacteriana com o filo predominante *Pseudomonadota*, representando 17,60% do total, com as classes *Gammaproteobacteria* (13,21%) e *Alphaproteobacteria* (2,39%).

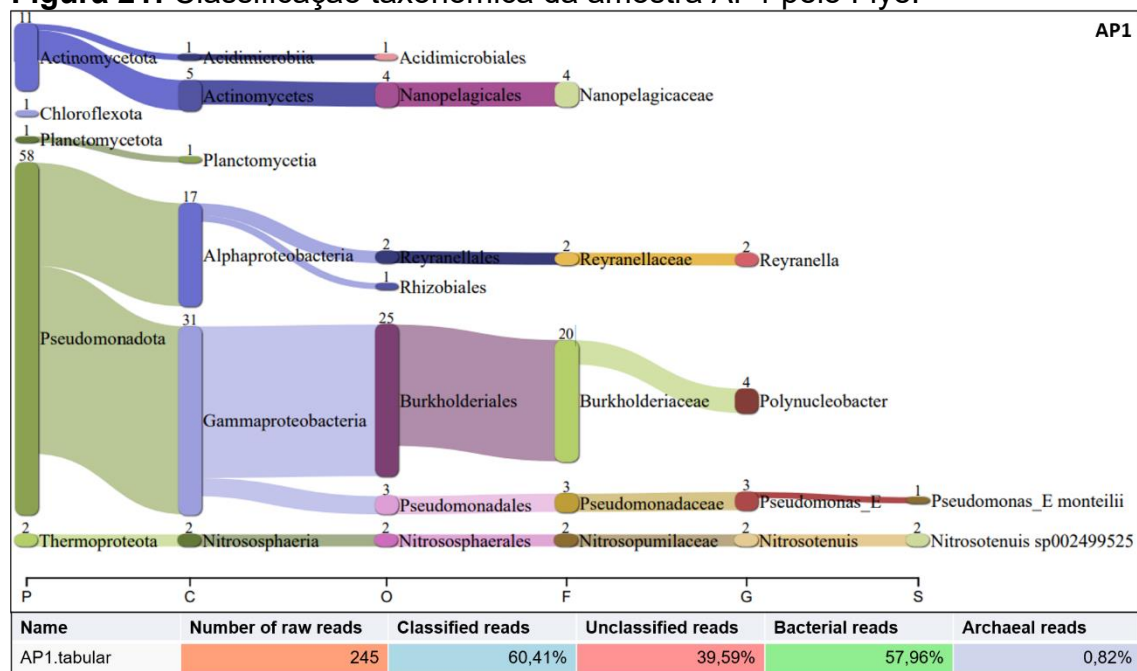
A ordem mais representativa foi *Burkholderiales* (11,07%), seguida da família *Burkholderiaceae* com 8,48%. O gênero predominante foi *Polynucleobacter* (2,39%), sendo a espécie *P. tropicus* detectada em menor proporção (0,24%). Outros gêneros importantes incluíram *Limnohabitans* (0,20%) e a família *Methylophilaceae* (0,05%).

O filo *Bacteroidota* esteve presente em 2,68%, sendo representado pelo gênero *Sediminibacterium* (2,63%). Outro filo bacteriano em menor abundância foi o *Actinomycetota* (0,49%), representados pela classe *Actinomycetes* (0,24%) e *Acidimicrobiia* (0,05%).

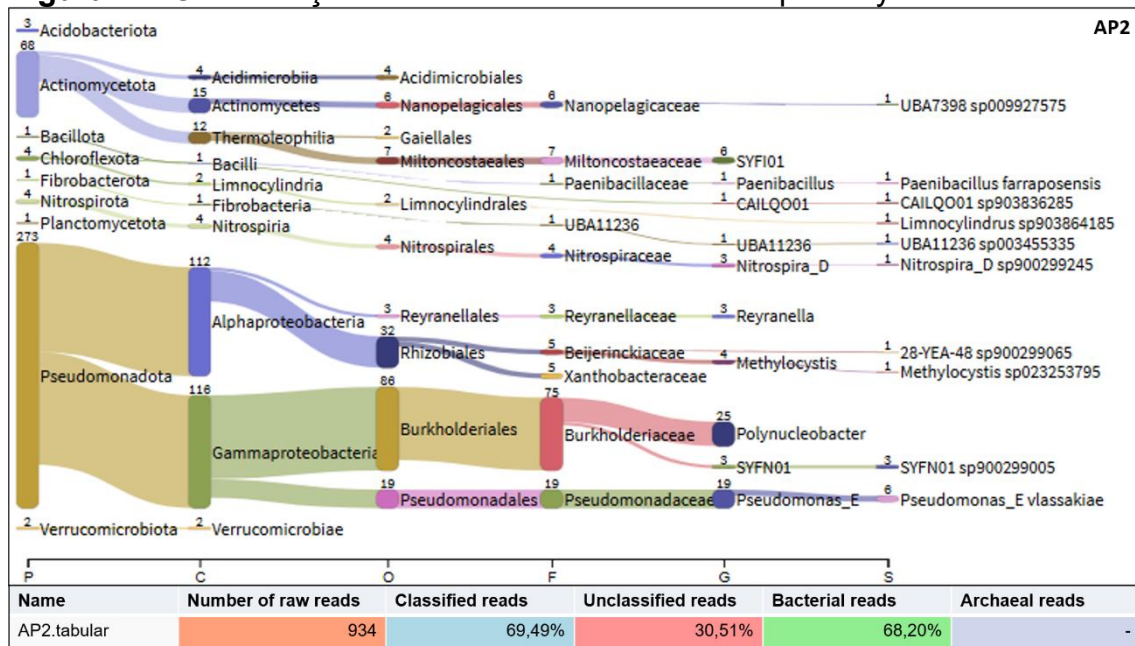
Já para o ponto AP11 (Figura 30), houve predominância do filo *Pseudomonadota* (23,85%), dividido em dois gêneros, *Limnohabitans* (9,82%) e *Polynucleobacter* (3,01%), com a espécie *P. tropicus* (0,34%).

O filo *Actinomycetota* (0,57%), apresentou duas ordens, *Acidimicrobiales* (0,28%) e *Nanopelagicales* (0,17%), com a família *Nanopelagicaceae* (0,17%). Grupos menos representados incluíram os filis *Verrucomicrobiota* (0,45%), e *Bacillota* (0,06%).

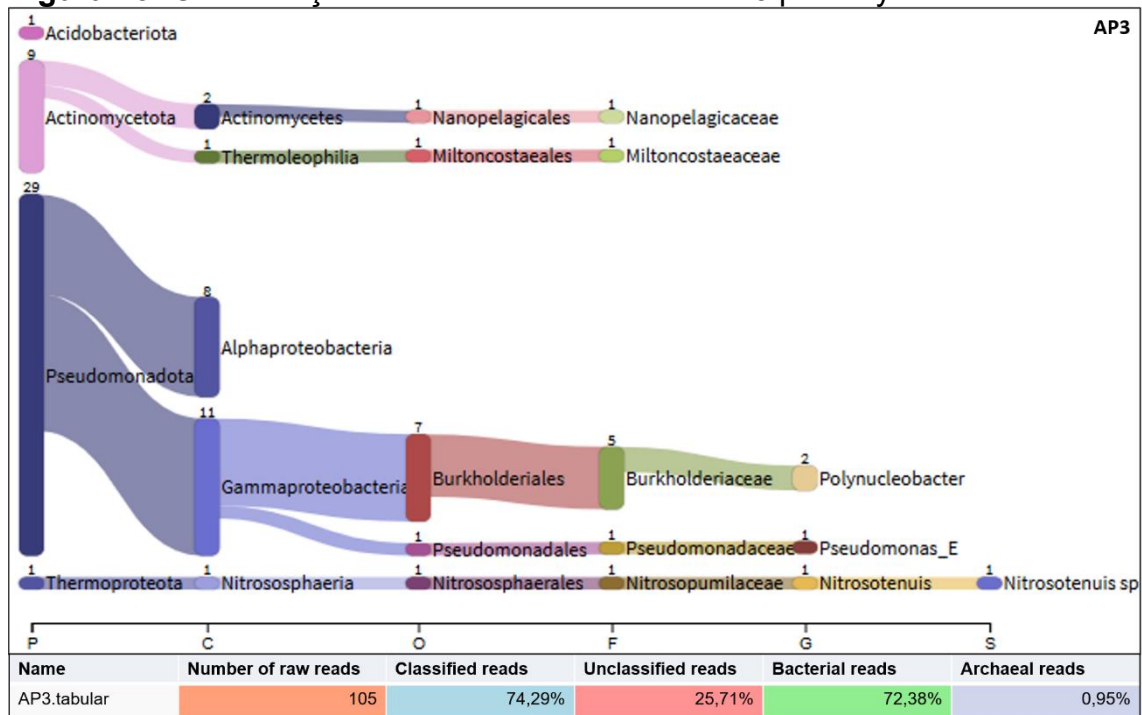
Figura 21: Classificação taxonômica da amostra AP1 pelo Flye.



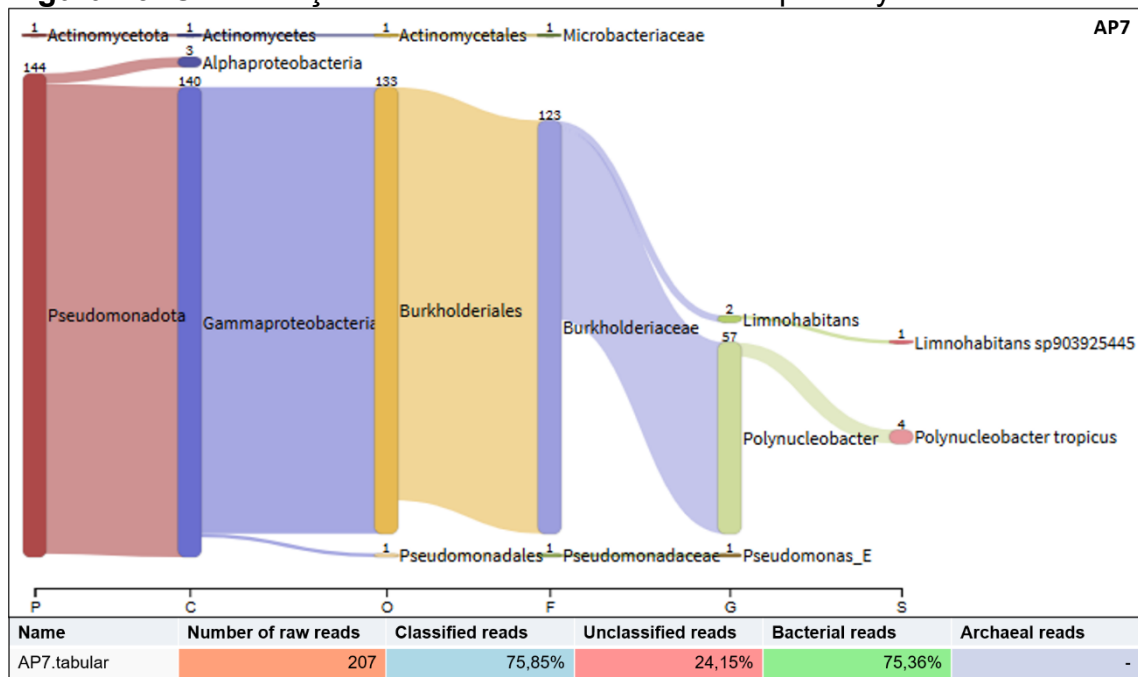
Fonte: Elaborado pelo autor.

Figura 22: Classificação taxonômica da amostra AP2 pelo Flye.

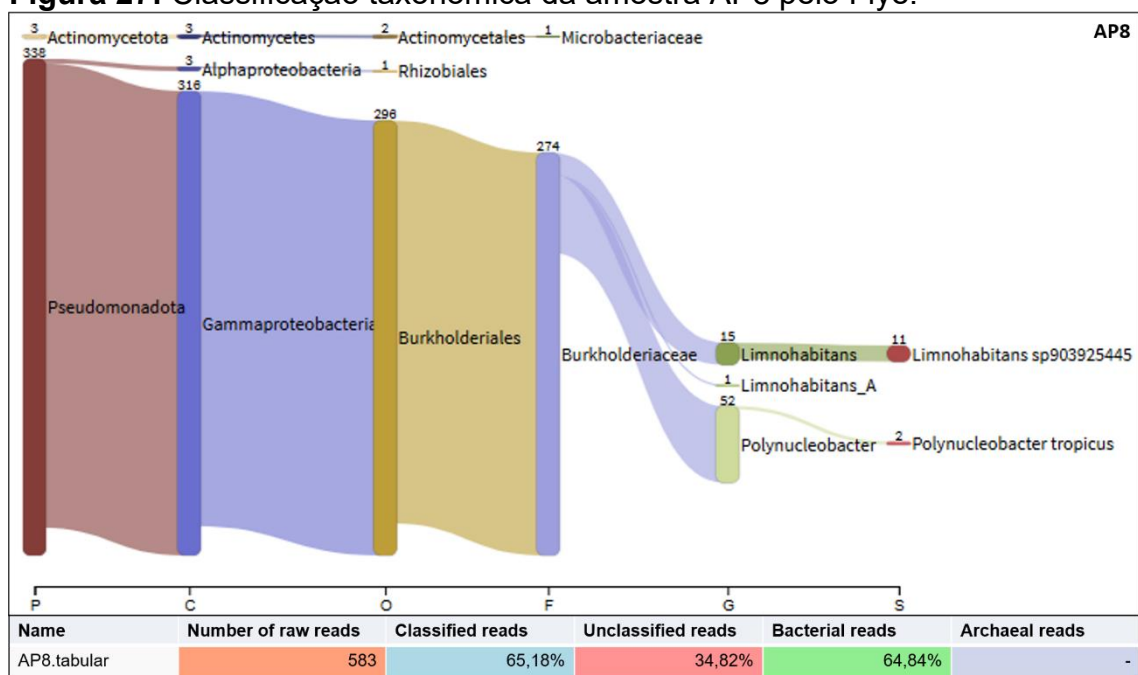
Fonte: Elaborado pelo autor.

Figura 23: Classificação taxonômica da amostra AP3 pelo Flye.

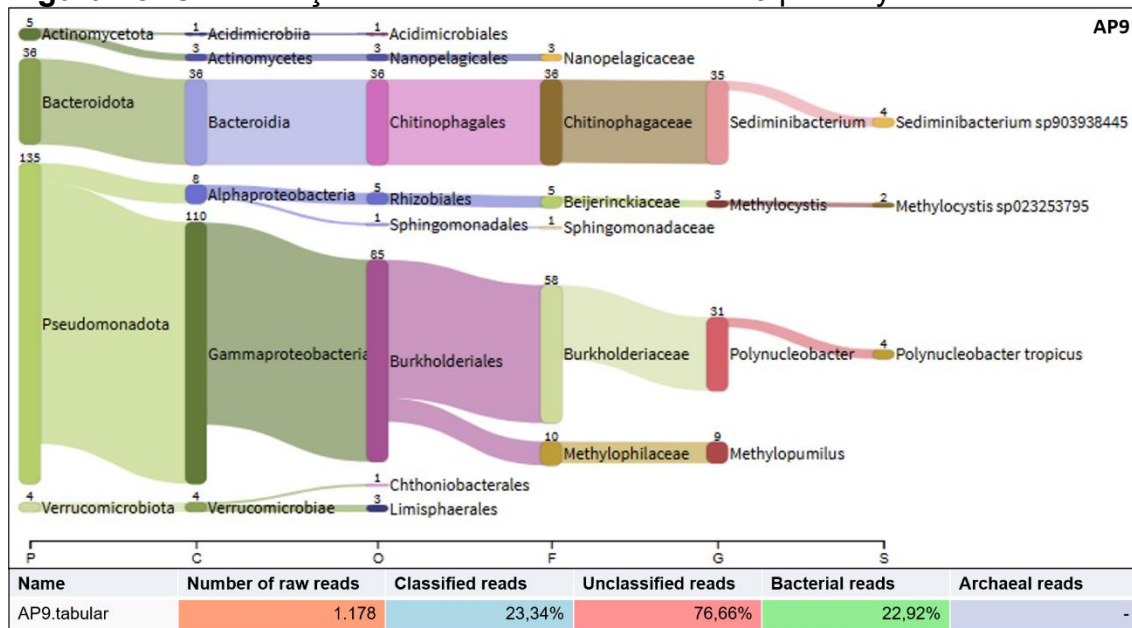
Fonte: Elaborado pelo autor.

Figura 26: Classificação taxonômica da amostra AP7 pelo Flye.

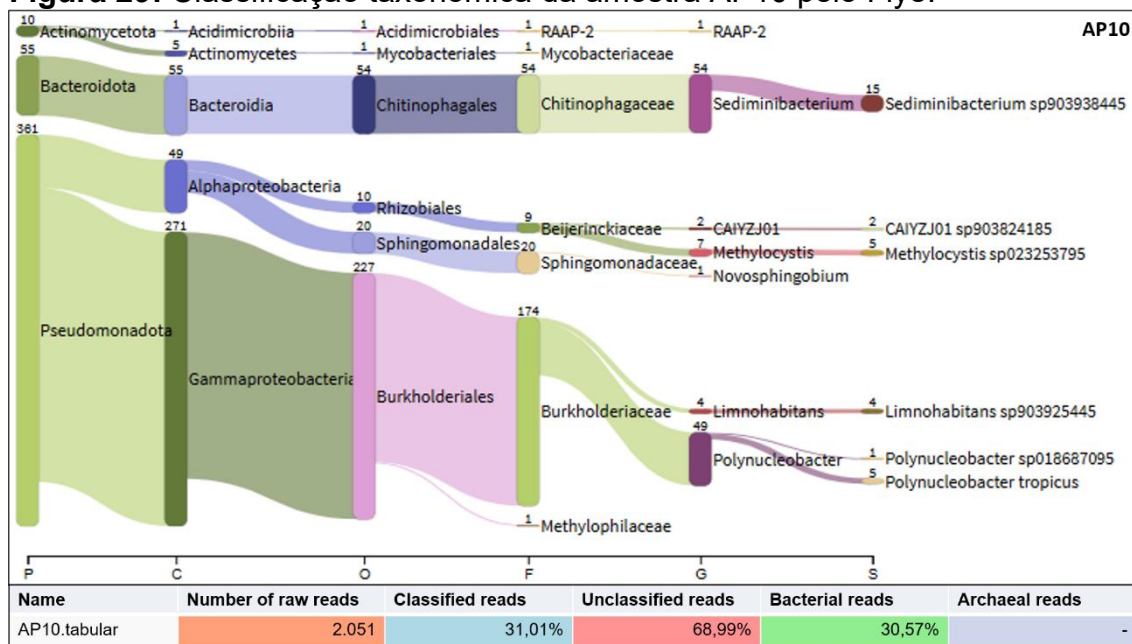
Fonte: Elaborado pelo autor.

Figura 27: Classificação taxonômica da amostra AP8 pelo Flye.

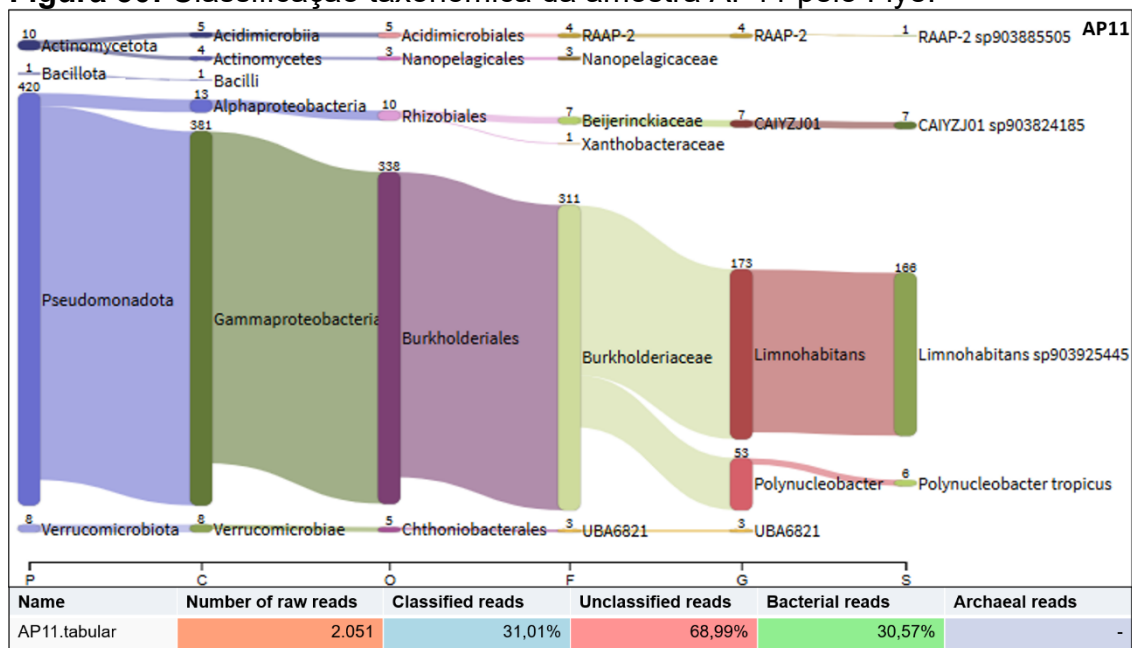
Fonte: Elaborado pelo autor.

Figura 28: Classificação taxonômica da amostra AP9 pelo Flye.

Fonte: Elaborado pelo autor.

Figura 29: Classificação taxonômica da amostra AP10 pelo Flye.

Fonte: Elaborado pelo autor.

Figura 30: Classificação taxonômica da amostra AP11 pelo Flye.

Fonte: Elaborado pelo autor.

6.2.5. Microbioma MEGAHIT

Para obter uma visão abrangente da diversidade taxonômica presente nas amostras, os dados metagenômicos foram processados e montados com o software MEGAHIT. Este montador, otimizado para metagenomas, utiliza uma estratégia de múltiplos *k-mers* para reconstruir sequências genômicas a partir de leituras curtas de forma eficiente.

(Relatório_completo_:<https://docs.google.com/spreadsheets/d/1pB_nJM4DVS_eTVEOdIW4n8rjcaYfbgJIYuzAi0F0jHUo/edit?usp=sharing>).

Para o ponto AP1 (Figura 31), foram classificados 28,41% das *contigs*, sendo 27,27% como Bacteria e 0,21% como Archaea (0,21%). Os principais filos classificados são *Pseudomonadota* (8,39%), *Actinomycetota* (1,72%), *Chloroflexota* (0,30%) e *Nitrospirota* (0,13%), com gêneros relevantes como *Pseudomonas* (1,09%), *Polynucleobacter* (0,26%), *Nitrosotenuis* (0,20%) e *Microbacterium* (0,02%).

No ponto AP2 (Figura 32), 25,77% foram classificadas. Entre as sequências classificadas, 25,02% corresponderam a Bacteria e apenas 0,04% a Archaea. A análise revelou que a composição microbiana foi marcada pela predominância dos filos *Pseudomonadota* (8,30%), *Actinomycetota* (1,69%), *Chloroflexota* (0,31%) e *Nitrospirota* (0,20%). *Pseudomonadota* se destacou com

grande abundância, principalmente através da classe *Gammaproteobacteria* (4,20%) e das ordens *Burkholderiales* (2,66%) e *Pseudomonadales* (0,91%). *Alphaproteobacteria* (2,43%) também apresentou alta diversidade, distribuindo-se por ordens como *Rhizobiales* (0,72%), *Reyranellales* (0,24%) e *Sphingomonadales* (0,05%), com gêneros relevantes como *Limnohabitans* (0,10%), *Polynucleobacter* (0,26%) e *Reyranella* (0,24%). O filo *Actinomycetota* (1,69%) apresentou grupos como *Actinomycetes* (0,70%), *Thermoleophilia* (0,26%), *Acidimicrobiia* (0,14%) e *Mycobacteriales* (0,11%), evidenciando diversidade funcional.

Outros grupos bacterianos importantes incluíram *Chloroflexota* (0,31%), *Nitrospirota* (0,20%), *Planctomycetota* (0,13%) e *Verrucomicrobiota* (0,17%), indicando uma comunidade ambiental complexa e funcionalmente diversa.

O ponto AP3 (Figura 33) revelou que 66,79% dos *contigs* não puderam ser classificados, enquanto 33,21% receberam uma atribuição taxonômica. Dentro da fração classificada, a comunidade foi composta por Bacteria (31,98%), e por Archaea (0,05%), representada principalmente pelo gênero *Nitrosotenuis*. O filo mais abundante foi *Pseudomonadota* (8,40%), distribuído entre as classes *Alphaproteobacteria* (3,30%) e *Gammaproteobacteria* (2,79%), seguido por *Actinomycetota* (2,04%).

Também apresentou alta diversidade de filos secundários em baixa abundância, como *Chloroflexota* (0,32%), *Planctomycetota* (0,19%) e *Verrucomicrobiota* (0,17%). Além disso, foi detectada múltiplos gêneros que contêm espécies patogênicas oportunistas. Foram identificados *contigs* atribuídos aos gêneros *Pseudomonas* (0,42%), *Ralstonia* e *Stenotrophomonas*, além da ordem *Enterobacterales*. No filo *Actinomycetota* (2,04%), foi detectado o gênero *Mycobacterium* (0,05%), e foram encontrados também gêneros frequentemente associados à microbiota humana, como *Cutibacterium* e *Bifidobacterium*.

Para o ponto AP4 (Figura 34), a montagem resultou em 23,8% de *contigs* classificados. Composta por 22,9% de Bacteria e 0,37% de Archaea. A estrutura da comunidade bacteriana seguiu o padrão observado nas outras amostras da região, com *Pseudomonadota* (8,79%) e *Actinomycetota* (1,61%).

Foram identificados *contigs* pertencentes aos gêneros *Pseudomonas*, *Burkholderia*, *Ralstonia*, *Acinetobacter*, *Stenotrophomonas*, *Aeromonas* e

Legionella. A ordem *Enterobacterales*, que inclui patógenos como *Escherichia* e *Plesiomonas* (ambos também detectados) assim como os gêneros *Mycobacterium* e *Prevotella*.

A análise da montagem para a amostra AP6 (Figura 35) indicou que 76,3% dos *contigs* não foram classificados, enquanto 23,7% receberam uma atribuição taxonômica, sendo a quase totalidade pertencente ao domínio Bacteria (23,2%). O filo *Pseudomonadota* compôs 12,0% do total de *contigs* classificados, com a classe *Gammaproteobacteria* (9,1%) e a ordem *Burkholderiales* (7,4%) sendo as mais representativas. O segundo filo mais abundante foi *Actinomycetota* (0,41%). Esta amostra também revelou a presença de diversos gêneros bacterianos de interesse para a saúde. Foram identificados *contigs* atribuídos aos gêneros *Pseudomonas*, *Burkholderia*, *Ralstonia* e *Achromobacter*. De forma notável, foi detectado o gênero *Klebsiella*, um importante patógeno oportunista pertencente à família *Enterobacteriaceae*. Adicionalmente, foram encontrados os gêneros *Mycobacterium* e *Prevotella*.

Para a amostra AP7 (Figura 36), a análise da montagem MEGAHIT classificou 34,2% dos *contigs*, deixando 65,8% sem atribuição taxonômica. A comunidade microbiana identificada foi do domínio Bactéria (33,9%). O filo *Pseudomonadota* demonstrou uma dominância, correspondendo a 24,1% do total de *contigs* e sendo majoritariamente representado pela classe *Gammaproteobacteria* (20,0%) e pela ordem *Burkholderiales* (17,2%). O segundo filo mais representativo foi *Actinomycetota* (0,50%). A triagem por gêneros de relevância para saúde identificou a presença de *Pseudomonas*, *Ralstonia*, *Aeromonas* e *Herbaspirillum*. Notavelmente, foi detectado o gênero *Serratia*, um membro da família *Enterobacteriaceae*, e também o gênero *Mycobacterium*.

O ponto AP8 (Figura 37), apresentou um perfil similar às outras amostras do mesmo rio, com 67,7% dos *contigs* da montagem não classificados. Da porção classificada (32,3%), a grande maioria foi atribuída ao domínio Bactéria (32,1%). A comunidade foi dominada pelo filo *Pseudomonadota*, que representou 22,9% do total de *contigs*, com a classe *Gammaproteobacteria* (18,4%) e a ordem *Burkholderiales* (15,9%) sendo os táxons mais abundantes. O filo *Actinomycetota* (0,82%) foi o segundo mais representativo. Os gêneros de importância para a saúde foram de *Pseudomonas*, *Burkholderia*, *Ralstonia* e

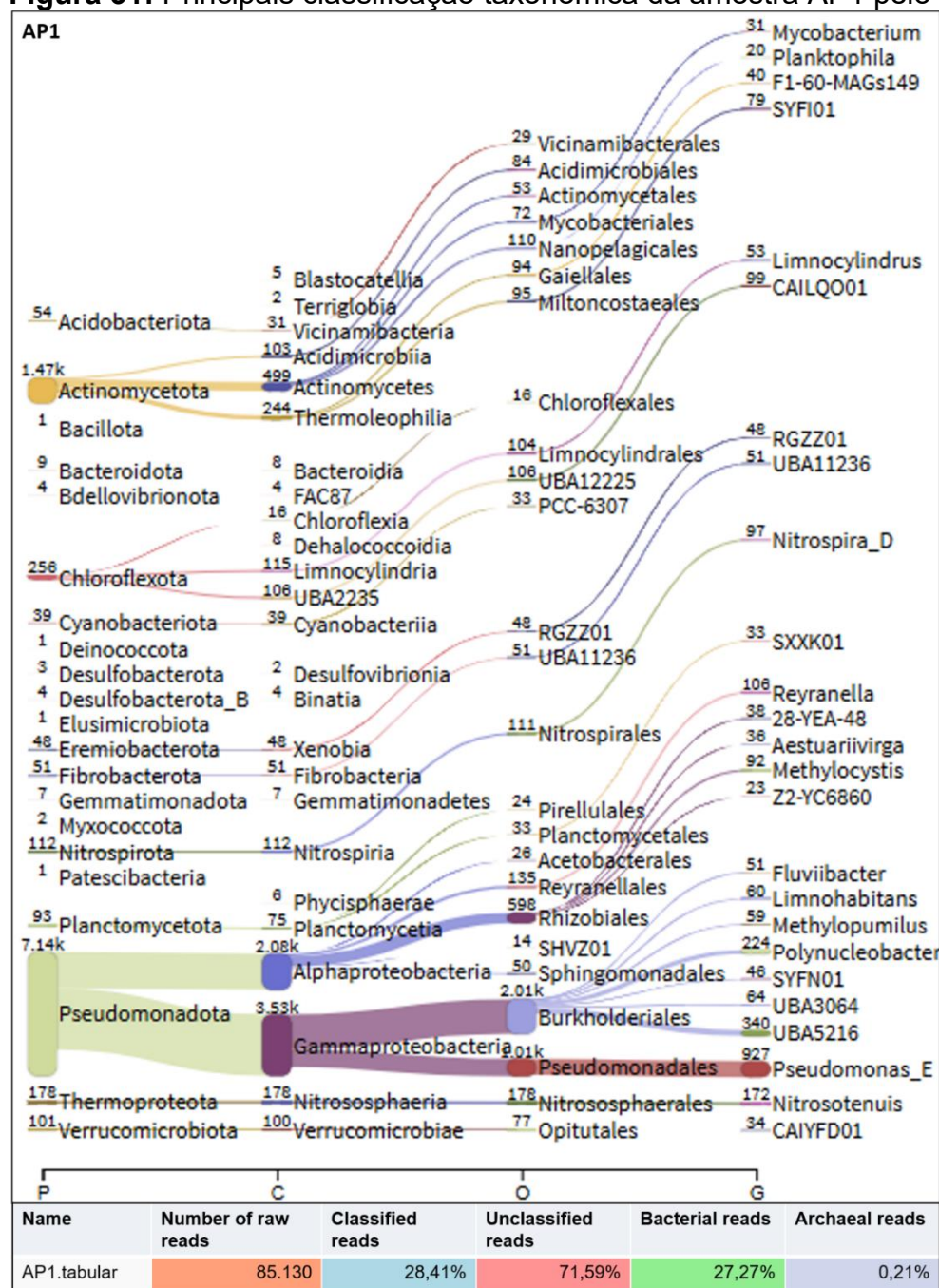
Cupriavidus. A ordem *Enterobacterales*, que inclui diversos patógenos, também foi detectada. Adicionalmente, foram identificados os gêneros *Mycobacterium* e *Clostridium*.

A amostra AP9 (Figura 38), apresentou a maior proporção de sequências não classificadas, com 89,6% dos *contigs* da montagem sem correspondência nos bancos de dados. Da fração classificada (10,4%), foi atribuída ao domínio Bactéria (10,2%). Diferentemente das outras amostras, a comunidade bacteriana, embora ainda liderada pelo filo *Pseudomonadota* (4,9%), mostrou uma codominância do filo *Bacteroidota* (1,25%). Outros fillos relevantes incluíram *Verrucomicrobiota* (0,26%) e *Actinomycetota* (0,27%). A análise de gêneros com potencial patogênico identificou a presença de *Pseudomonas*, *Ralstonia*, *Burkholderia*, *Mycobacterium*, *Bacillus* e *Paraclostridium*. A ordem *Legionellales* também foi detectada.

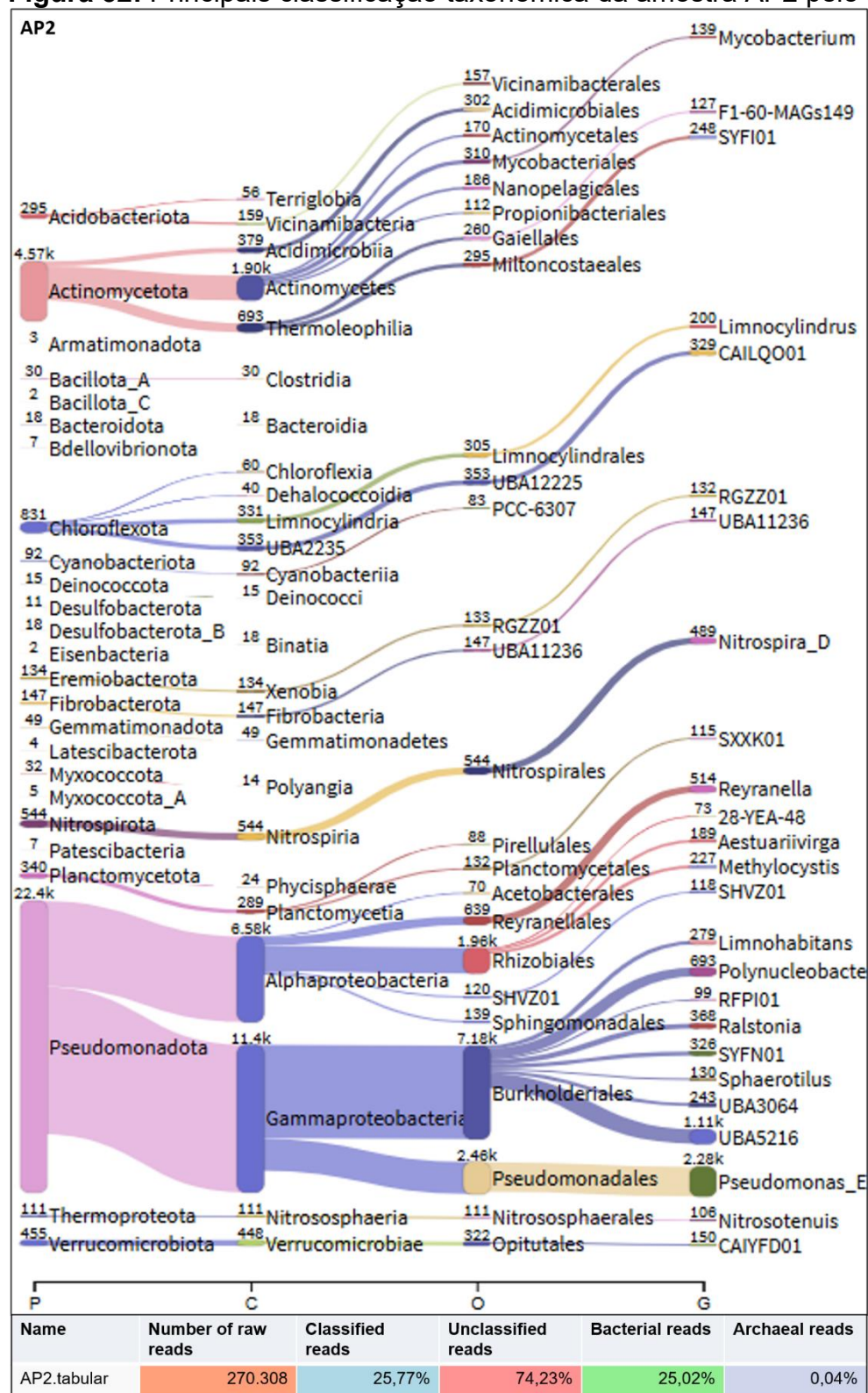
Para o ponto AP10 (Figura 39), da mesma região do poço, apresentou 86,9% de seus *contigs* não classificados. Na porção classificada o domínio Bactéria representou 13,10%. O perfil taxonômico mostrou-se um intermediário entre o poço (AP9) e as amostras de rios, com o filo *Pseudomonadota* (6,80%) sendo o mais abundante, mas aproximado pelo filo *Bacteroidota* (1,10%), com destaque para o gênero *Sediminibacterium*. O filo *Actinomycetota* (0,52%) foi o terceiro mais representativo. Ocorreu mais gêneros de importância clínica nesta amostra, identificando *Pseudomonas*, *Burkholderia*, *Ralstonia*, *Acinetobacter*, *Aeromonas*, *Herbaspirillum*, *Chryseobacterium* e *Clostridium*. A ordem *Enterobacterales* também foi detectada, com a identificação dos gêneros *Pantoea*, *Serratia* e *Enterobacter*. Além destes, foram encontrados os gêneros *Mycobacterium* e *Cutibacterium*.

Para o ponto AP11 (Figura 40) apresentou 18,0% de seus *contigs* classificados. O domínio Bactéria (17,8%) foi predominante. A comunidade bacteriana foi predominante para o filo *Pseudomonadota* (11,90%), com a classe *Gammaproteobacteria* (9,90%) e a ordem *Burkholderiales* (9,00%) sendo as mais abundantes, e os gêneros *Limnohabitans* e *Polynucleobacter* como os principais representantes. O segundo filo mais relevante foi *Actinomycetota* (0,77%). A busca por gêneros de interesse clínico identificou a presença de *Pseudomonas*, *Burkholderia*, *Ralstonia*, *Stenotrophomonas* e *Herbaspirillum*. A ordem *Enterobacterales* foi detectada, incluindo o gênero *Pantoea*.

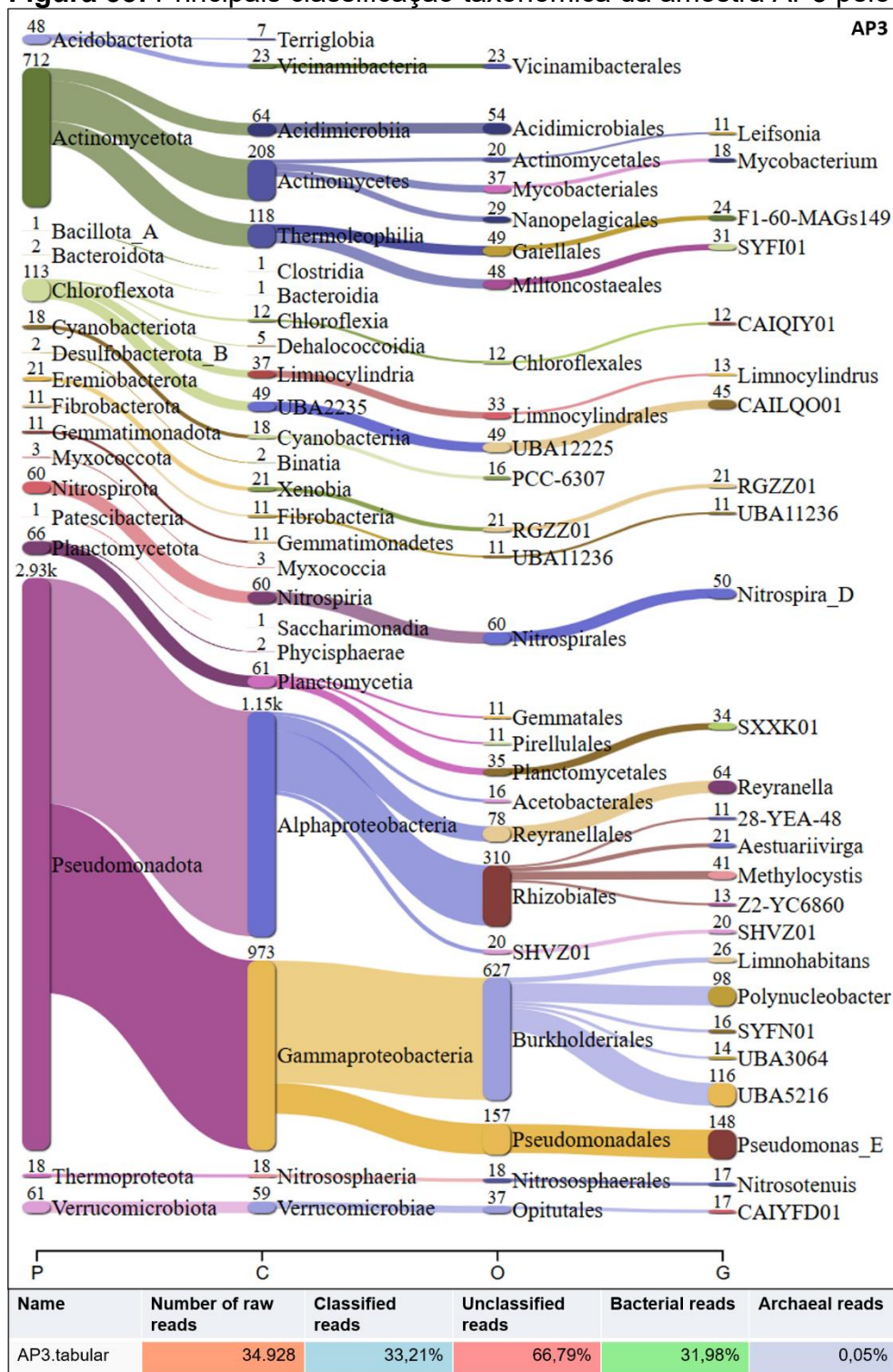
Figura 31: Principais classificação taxonômica da amostra AP1 pelo MEGAHIT.



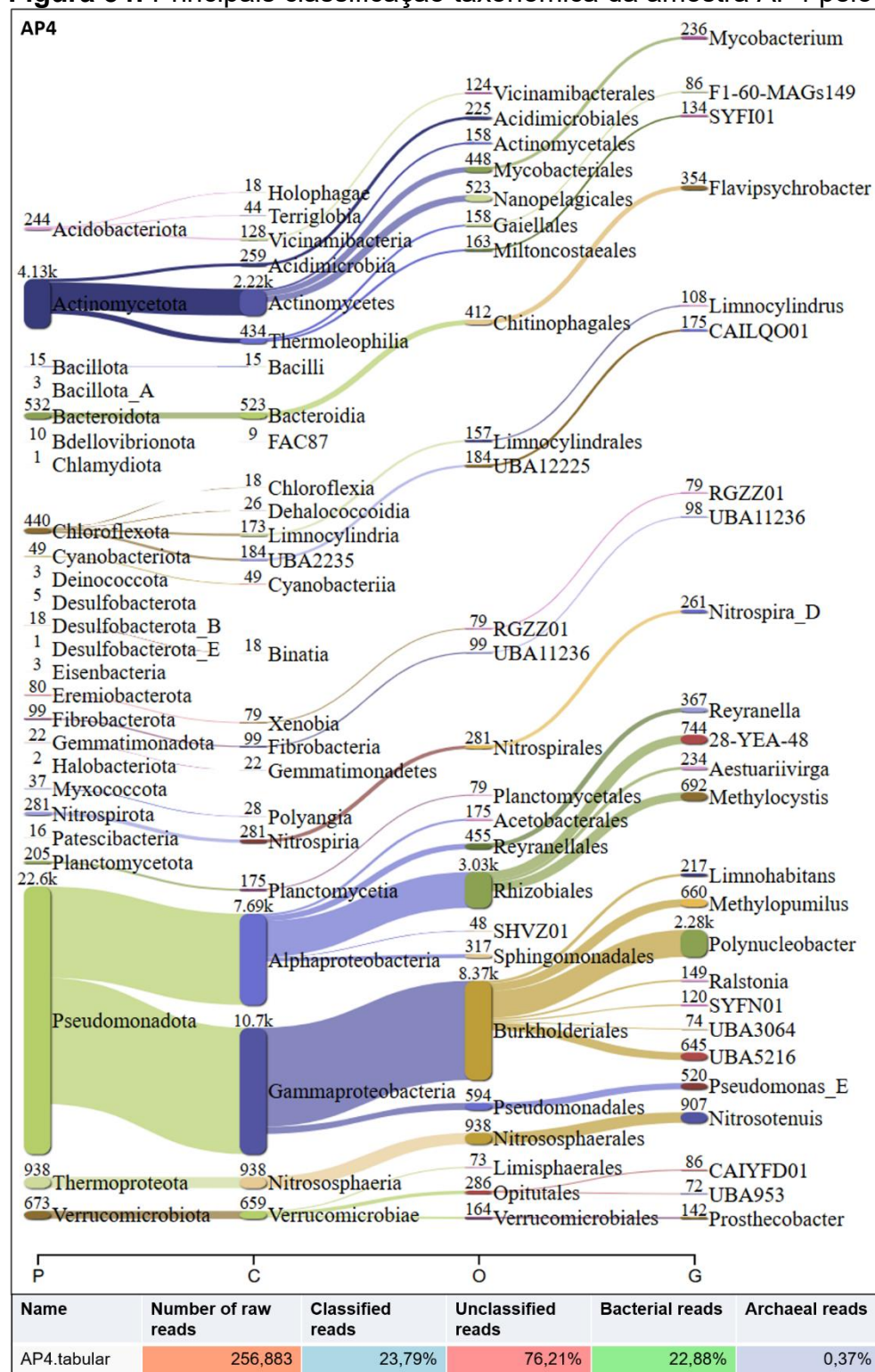
Fonte: Elaborado pelo autor.

Figura 32: Principais classificação taxonômica da amostra AP2 pelo MEGAHIT.

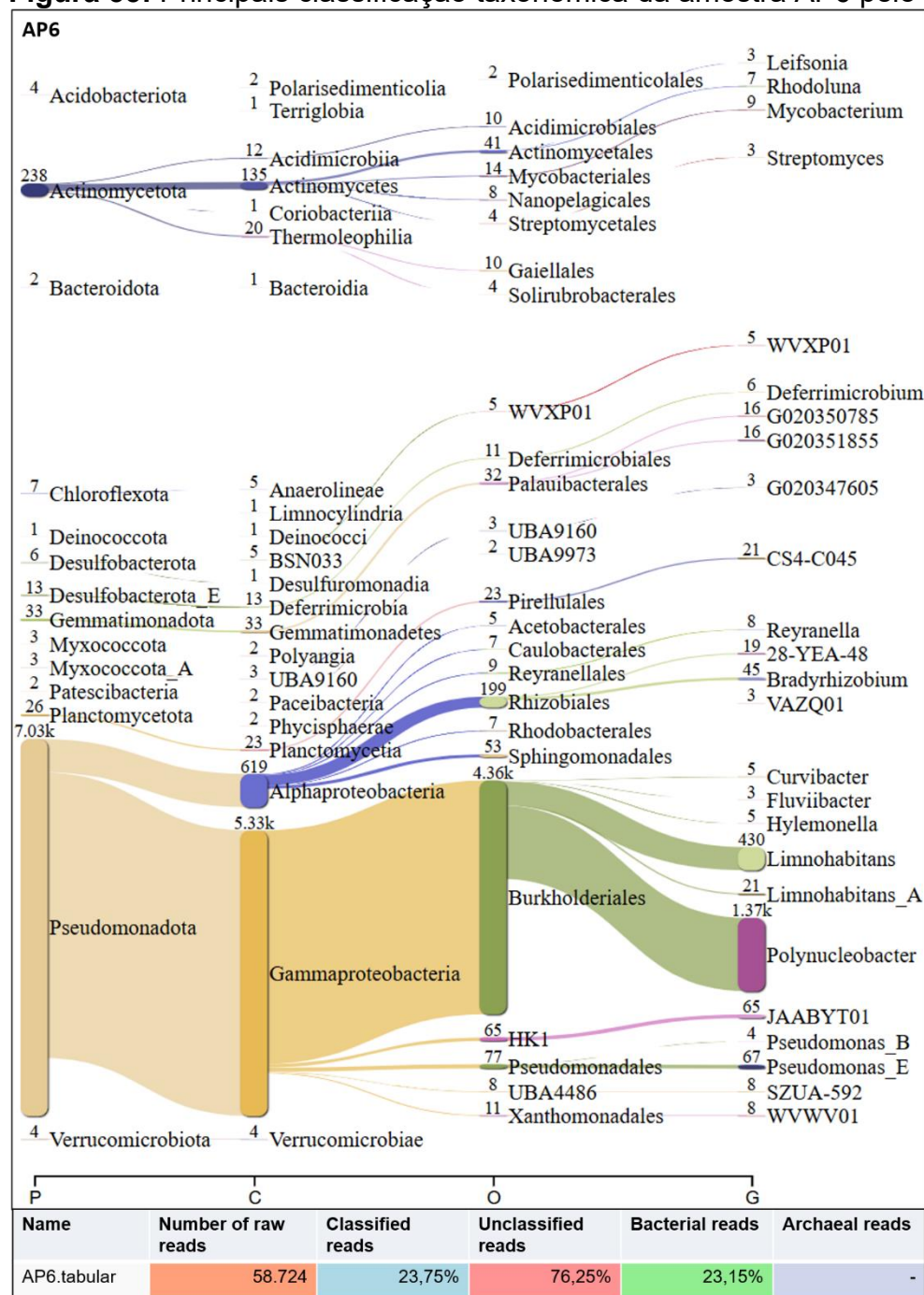
Fonte: Elaborado pelo autor.

Figura 33: Principais classificação taxonômica da amostra AP3 pelo MEGAHIT.

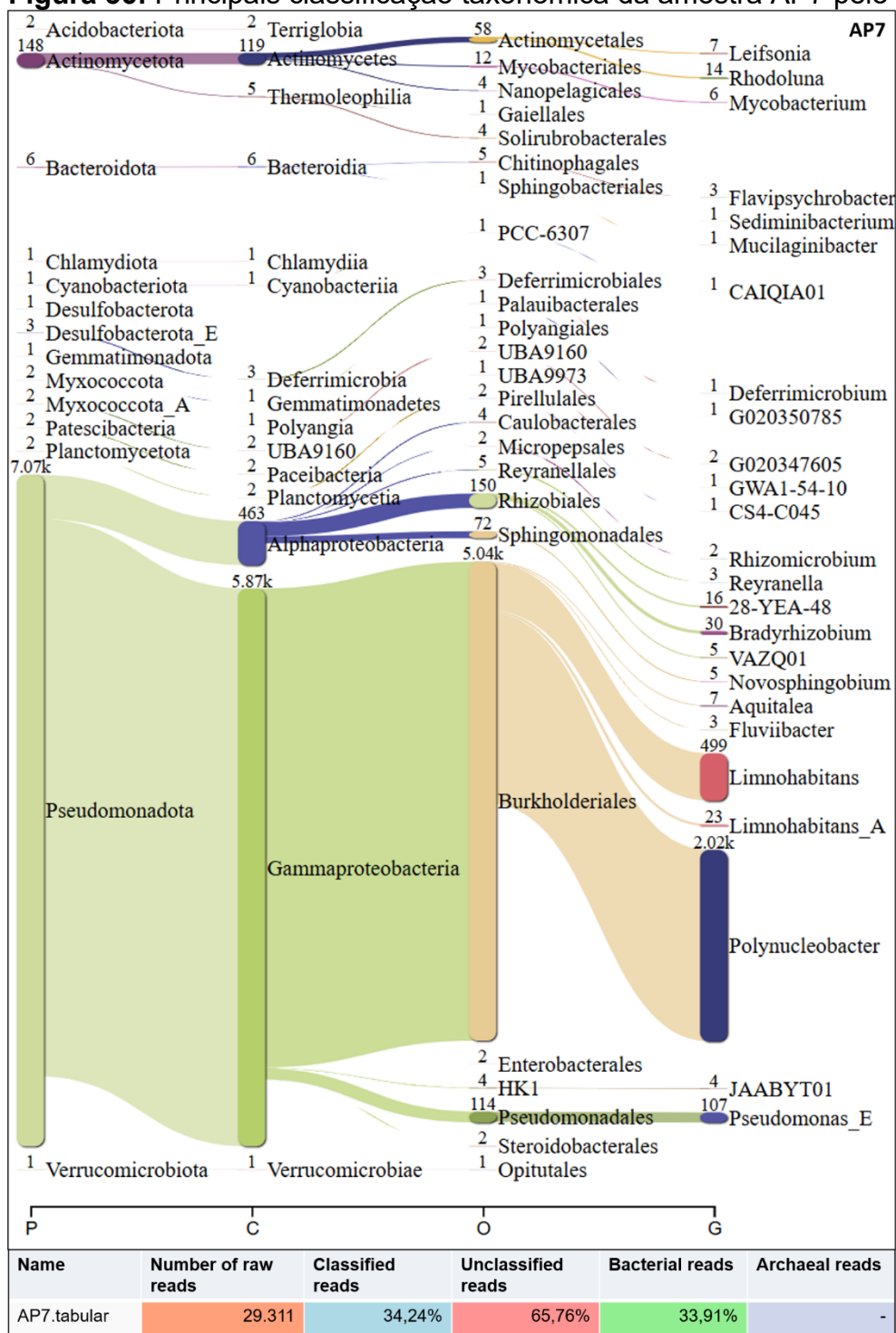
Fonte: Elaborado pelo autor.

Figura 34: Principais classificação taxonômica da amostra AP4 pelo MEGAHIT.

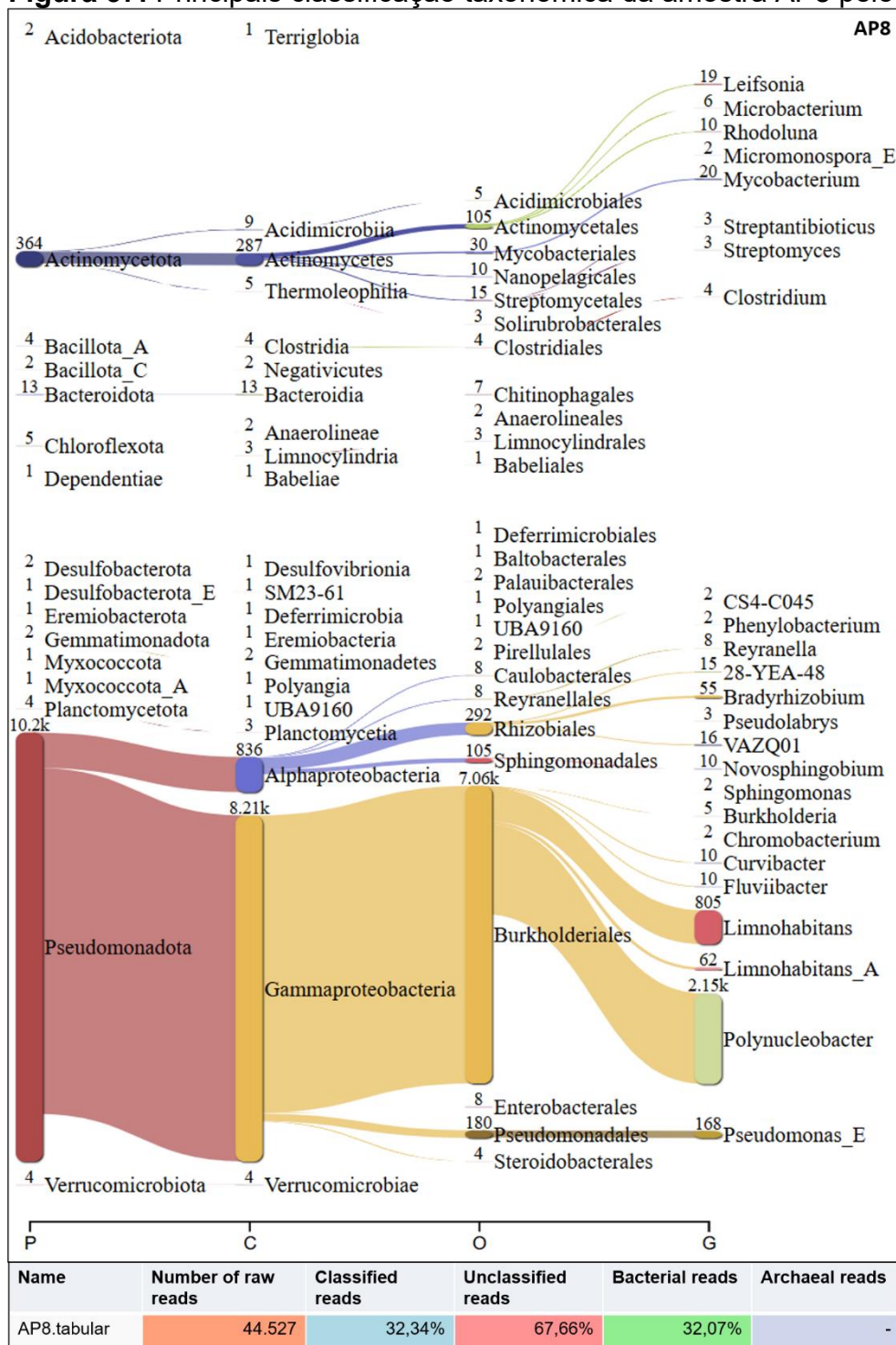
Fonte: Elaborado pelo autor.

Figura 35: Principais classificação taxonômica da amostra AP6 pelo MEGAHIT.

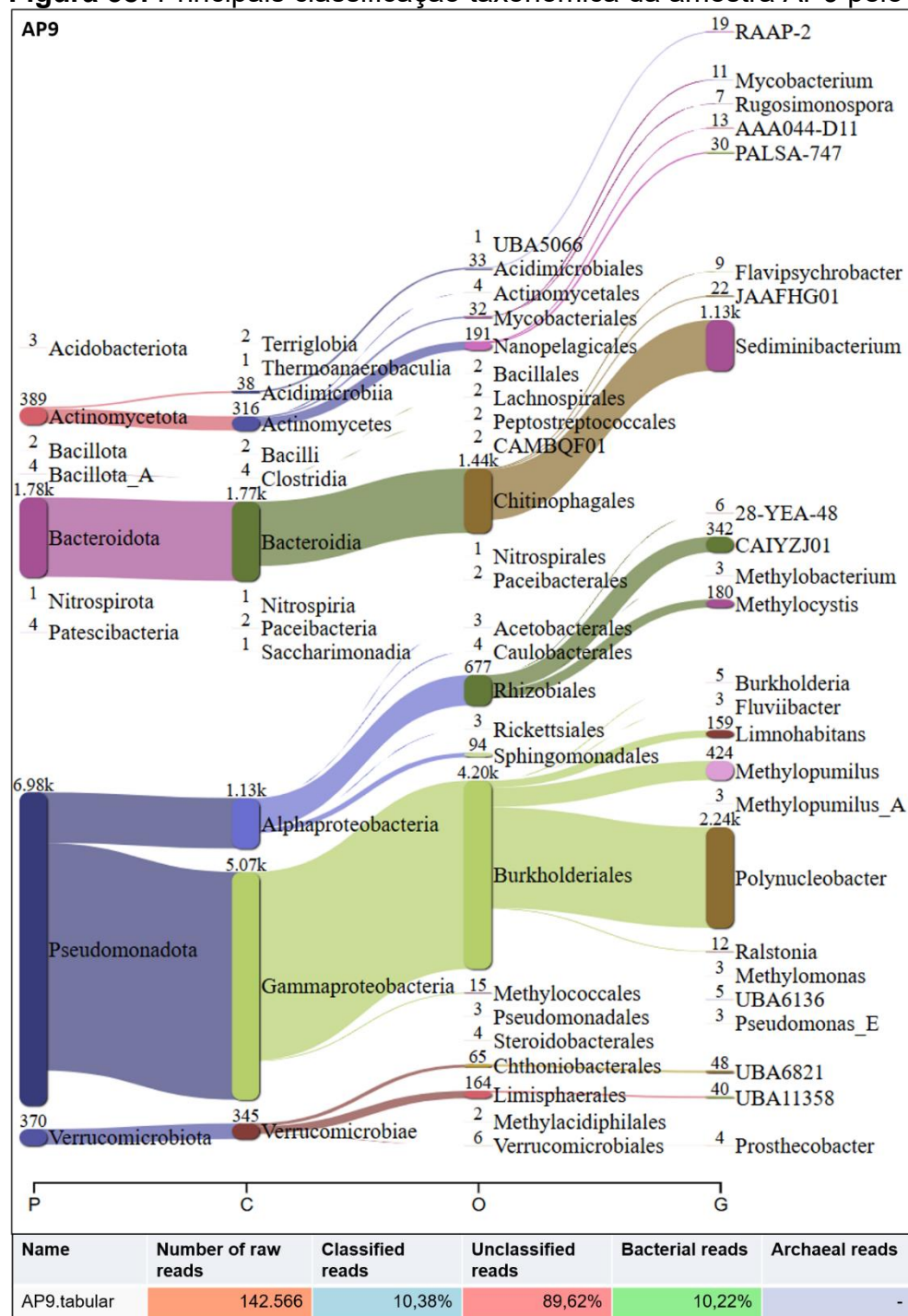
Fonte: Elaborado pelo autor.

Figura 36: Principais classificação taxonômica da amostra AP7 pelo MEGAHIT.

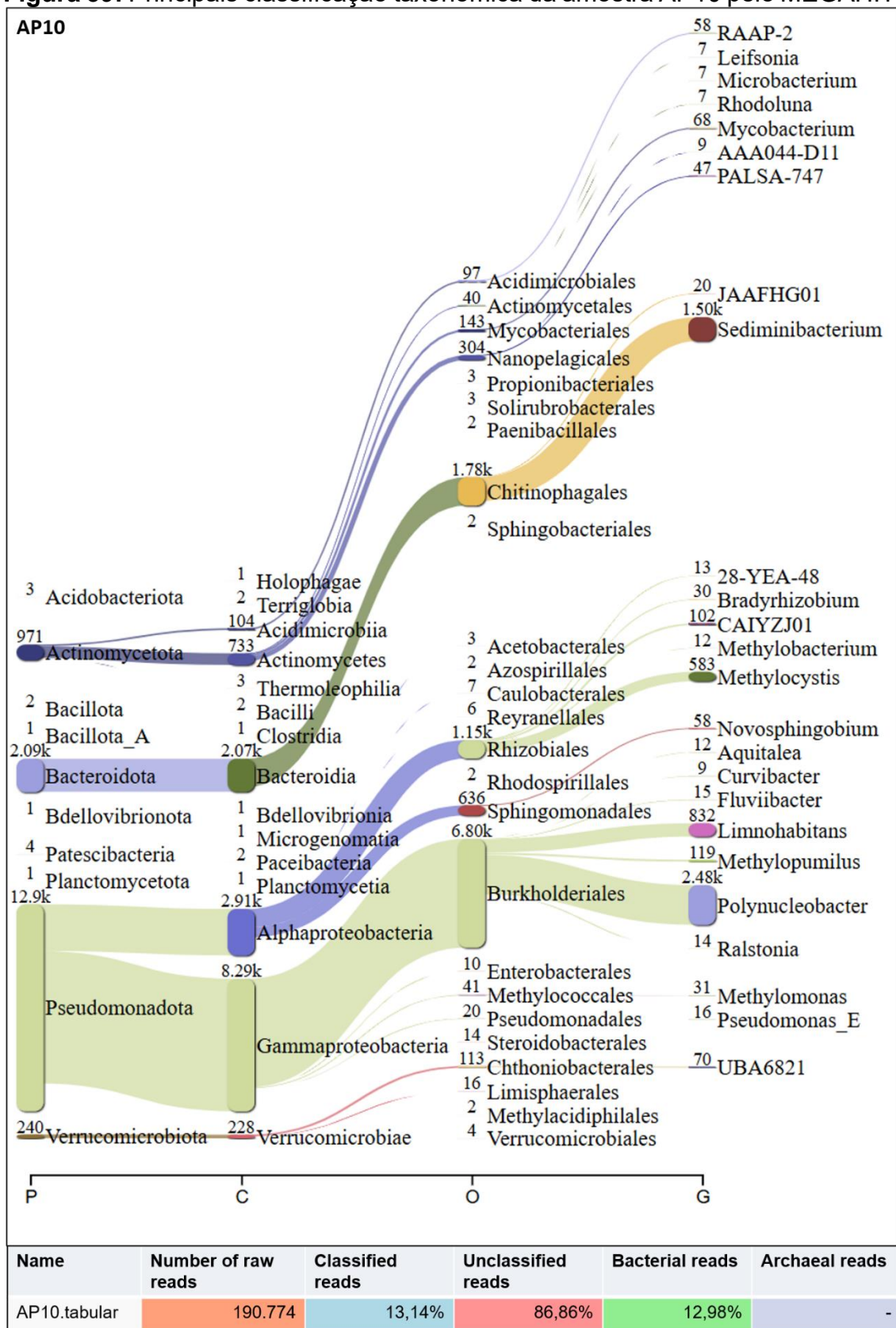
Fonte: Elaborado pelo autor.

Figura 37: Principais classificação taxonômica da amostra AP8 pelo MEGAHIT.

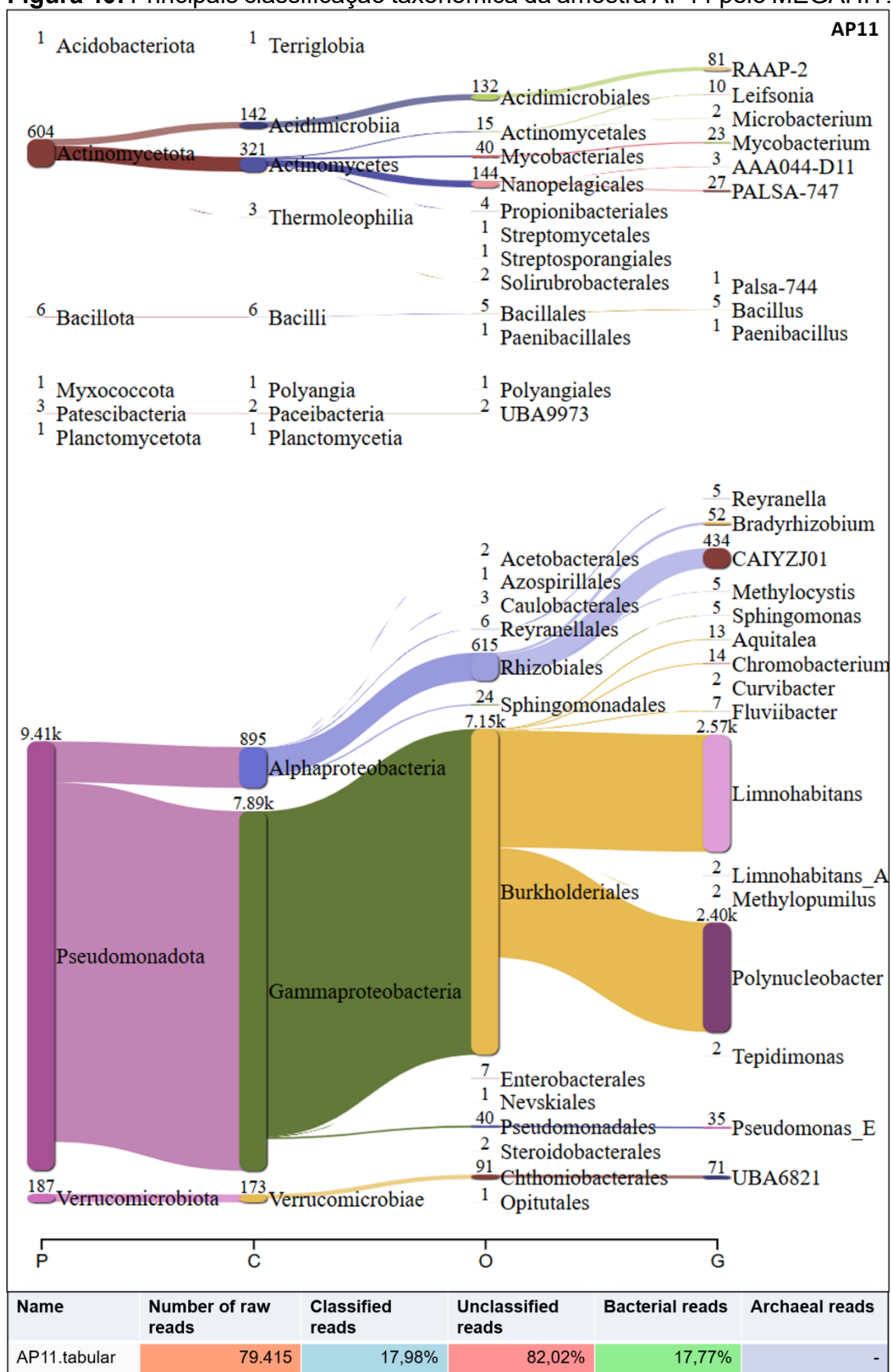
Fonte: Elaborado pelo autor.

Figura 38: Principais classificação taxonômica da amostra AP9 pelo MEGAHIT.

Fonte: Elaborado pelo autor.

Figura 39: Principais classificação taxonômica da amostra AP10 pelo MEGAHIT.

Fonte: Elaborado pelo autor.

Figura 40: Principais classificação taxonômica da amostra AP11 pelo MEGAHIT.

Fonte: Elaborado pelo autor.

6.2.6. Montagem Flye vs MEGAHIT

A montagem realizada com o Flye, ferramenta desenvolvida para *long reads*, apresentou resultados evidenciando maior contiguidade e continuidade das sequências, como apresentado pelos baixos números de *contigs* e maiores valores de N50 e auN nas amostras avaliadas. Amostras consideradas complexas, devido ao número de microrganismos diferentes, necessita de abordagens que consiga montar *contigs* completas, menos fragmentadas e com maior contiguidade. O modo metaFlye, utilizado pelo Flye, proporciona resultados eficazes para a reconstrução genômica, gerando *assemblers* de *longs reads* com organismos em maior abundância (Kolmogorov *et al.*, 2020; Lapidus; Korobeynikov, 2021).

Ademais, a ausência de bases indefinidas (Ns) nas montagens demonstra qualidade no consenso da sequência e a eficiência do processo para correção de erros típicos do sequenciamento *Nanopore*. No entanto, é importante destacar que dados de *long reads* apresentam maior taxa de erro intrínseca em comparação com dados de *short reads* (*error rates* entre 5-15%), o que pode impactar diretamente na qualidade do *assembly* e pode levar a artefatos, como *contigs* quiméricos, principalmente em regiões de alta complexidade e heterogeneidade (Goussarov *et al.*, 2024).

Por outro lado, o MEGAHIT foi projetado para montagem eficiente de dados de *short reads* e apresentou montagens com maior fragmentação, evidenciada pelo maior número de *contigs* e menores valores de N50.

O MEGAHIT maximiza a recuperação da diversidade microbiana mesmo em comunidades altamente complexas, o que resulta em montagens mais fragmentadas, porém com representatividade dos táxons presentes (Li *et al.*, 2015; Yang *et al.*, 2021).

Apesar da fragmentação, os resultados taxonômicos do MEGAHIT mostraram concordância com os obtidos pela montagem do Flye no que diz respeito aos táxons dominantes nas amostras ambientais, como o filo *Pseudomonadota* e os gêneros *Polynucleobacter*, *Pseudomonas* e outros integrantes das classes *Alphaproteobacteria* e *Gammaproteobacteria*. Além disso, o MEGAHIT conseguiu detectar uma maior diversidade de táxons distintos e fragmentados, que, embora em baixa proporção, contribuem para o panorama

completo da comunidade microbiana e podem representar diversidade funcional relevante.

A similaridade relativa entre os resultados taxonômicos dominantes indica que ambos os métodos são eficazes para a caracterização ecológica das comunidades. O Flye apresenta montagens precisas com longas *contigs*, porém descartou diversas *reads*. Já o MEGAHIT acrescenta informações complementares importantes sobre o *pool* e a complexidade da comunidade. Essa complementaridade tem sido reconhecida em análises comparativas recentes de montagem metagenômica, recomendando a utilização conjunta de *assemblers* para obter uma representação mais abrangente (Han *et al.*, 2025; Zhang *et al.*, 2024).

Estudos metagenômicos demonstrando o papel ativo das comunidades na ciclagem biogeoquímica e reciclagem de nutrientes nos rios amazônicos têm reportado uma rica diversidade microbiana dominada por táxons similares aos observados neste trabalho, como *Pseudomonadota*, *Actinomycetota* e arqueias nitrificantes do filo *Thermoproteota* (Ramos *et al.*, 2024; Ghai *et al.*, 2011). A presença desses grupos em alta abundância consolida a atuação das ferramentas para detecção de microrganismos presentes nas amostras. Os filos detectados frequentemente dominam a comunidade bacteriana devido à sua vasta diversidade metabólica (Ghai *et al.*, 2011).

6.2.7. Resistoma

Após a montagem dos dados metagenômicos com Flye e MEGAHIT, as amostras foram submetidas à análise para detecção de genes de resistência a antimicrobianos utilizando os bancos de dados NCBI-BARRGD, ResFinder e CARD (Tabela 18).

Tabela 18: Análise do resistoma.

Pontos	NCBI - BARRGD		ResFinder		CARD	
	Flye	MEGAHIT	Flye	MEGAHIT	Flye	MEGAHIT
AP1	-	-	-	-	-	-
AP2	-	<i>bla</i> OXA60	-	<i>bla</i> OXA60	-	<i>bla</i> OXA60
AP3	-	-	-	-	-	-
AP4	-	-	-	-	-	-
AP6	-	-	-	-	-	-
AP7	-	-	-	-	-	-
AP8	-	-	-	-	-	-
AP9	-	-	-	<i>RbpA</i>	-	<i>RbpA</i>
AP10	-	-	-	-	-	-
AP11	-	-	-	-	-	-

Fonte: Elaborado pelo autor.

A análise revelou a presença de genes de RAM em duas amostras: AP2 e AP9 (Tabela 19). Em AP2, o gene *bla*OXA-60, associado à resistência a antibióticos betalactâmicos, foi detectado na montagem MEGAHIT. Esse gene está relacionado à produção de betalactamase do tipo oxacilinase, que confere resistência a cefalosporinas e carbapenêmicos (Alam *et al.*, 2018; Ibrahim *et al.*, 2021; Flores *et al.*, 2016). A *contig* k141_18965 que apresentou o gene, foi atribuída à espécie *Ralstonia pickettii* pelo Blast no NCBI.

A ocorrência dessa espécie foi relatada em águas na região do bioma amazônico, associado à sua capacidade de sintetizar compostos como poli-hidroxiclcanoato (Castro *et al.*, 2021). Embora não tenha sido encontrado trabalhos relacionando o gene específico ocorrendo na espécie em regiões do bioma amazônico, há relatos sobre a resistência a carbapenêmicos por *R. pickettii* e outras espécies (Dos Santos *et al.*, 2024). Além disso, Tejera e colaboradores (2016) relataram dois casos ocorridos no Brasil por infecção de *R. pickettii* resistente em humanos, com a presença de duas betalactamases, *bla*OXA-60 e *bla*OXA-22.

Na amostra AP9, o gene *RbpA*, implicado na resistência à rifampicina, foi identificado exclusivamente na montagem MEGAHIT, relacionado a *contigs* atribuídos ao gênero *Mycobacterium*. *RbpA* é uma proteína associada à modulação da RNA polimerase e tem sido descrita como um mecanismo relevante de resistência em micobactérias ambientais e patogênicas (Ge *et al.*, 2012; Alraey *et al.*, 2025).

A classificação do gênero *Mycolicibacterium* é recente e possui uma divergência taxonômica, sendo considerado sinônimo do gênero *Mycobacterium* (Gupta; Brian; Son, 2018). Existe ocorrência desses gêneros associados a microbiota de animais presentes no rio Amazonas, como encontrado em *Serrasalmus rhombeus*, espécie de piranha endêmica da bacia amazônica (Da Silva *et al.*, 2024) ou infecções de espécies ocorrida em bovinos e búfalos presentes na região do Amazonas e Pará (Carneiro *et al.*, 2020).

A detecção do gene *blaOXA-60* em *R. pickettii* no ponto AP2, e do gene *RbpA* no gênero *Mycolicibacterium* no AP9, ilustra como esses genes circulam em espécies ambientais resistentes e a presença em ambientes aquáticos pode indicar o potencial de se tornar reservatórios e pontos críticos para a evolução e dispersão de RAM.

O ponto AP2 correspondente ao Igarapé Fortaleza, Santana - AP, está localizado em uma área comercial intensamente movimentada, conectada diretamente ao rio Amazonas, com intenso tráfego de embarcações e atividades de comércio fluvial, o que pode favorecer a entrada e recirculação de microrganismos resistentes provenientes de múltiplas fontes. Em contraste, o ponto AP9 é um poço do Rio Maracá-Pacu, Mazagão Velho – AP, e situa-se em uma área conservada e com baixo impacto humano, onde a detecção de genes como *RbpA* pode estar relacionada à resistência natural de bactérias ambientais ou o contato com animais silvestres.

A ausência de detecção de genes de RAM nas montagens Flye pode indicar a perda de *reads* durante a formação das *contigs* que dificultou a detecção de genes presente nas porções fragmentadas, além da própria menor profundidade de montagem em certos níveis de resolução (Li *et al.*, 2021). Por outro lado, o MEGAHIT, apesar de gerar montagens fragmentadas foi capaz de apresentar uma maior diversidade de genomas, incluindo genes de resistência, demonstrando a importância da utilização de *assemblers* complementares para caracterização abrangente do microbioma e resistoma em amostras ambientais complexas (Zhang *et al.*, 2024).

Por outro lado, a baixa presença dos genes de RAM identificados pode refletir boas condições ambientais ou baixa pressão seletiva por antimicrobianos nas áreas de levantamento, mas também reforça a necessidade de monitoramento contínuo, visto o potencial impacto desses genes em

ecossistemas aquáticos e na saúde pública, especialmente considerando a disseminação global da resistência a antibióticos (Lupo *et al.*, 2021; Chen *et al.*, 2023). Ainda assim, não elimina o risco potencial associado à disseminação desses genes pela água, que pode atingir diversos nichos e populações, humanos e animais, conforme enfatizado pela abordagem Uma Só Saúde (Resende; Silva; Diniz, 2020).

Nesse contexto, a maneira como as comunidades ribeirinhas utilizam os corpos hídricos, intensifica o potencial impacto da circulação de genes de RAM, já que os rios, igarapés e poços são empregados cotidianamente para consumo, higiene, pesca, transporte e outras atividades essenciais à subsistência local. A dependência direta desses ambientes aquáticos, muitas vezes em cenários de déficit de saneamento básico e vulnerabilidade hídrica, aumenta a exposição humana a microrganismos resistentes e reforça a necessidade de estratégias de monitoramento e gestão que integrem as dimensões socioambientais e de saúde pública corroborando com a abordagem Uma Só Saúde.

Essa interação entre uso humano do território e circulação de microrganismos resistentes também é observada em cadeias produtivas animais na região amazônica. No Amapá, Creão (2025) demonstrou elevada ocorrência de enfermidades zoonóticas, como tuberculose e brucelose, em bovídeos abatidos sob inspeção estadual, evidenciando falhas sanitárias estruturais ao longo da cadeia produtiva pecuária. Considerando que a criação extensiva de bovinos e bubalinos na Amazônia frequentemente ocorre em áreas ribeirinhas, com acesso direto a cursos d'água utilizados simultaneamente por comunidades humanas e animais, esses sistemas podem se tornar pontos críticos de interface entre saúde animal, ambiental e humana. A circulação de agentes zoonóticos em rebanhos, associada ao descarte inadequado de resíduos orgânicos e efluentes, pode contribuir para a introdução de microrganismos e genes de resistência em ambientes aquáticos compartilhados. Assim, os achados de Creão (2025) reforçam que problemas sanitários na produção animal não devem ser analisados de forma isolada, mas integrados à dinâmica socioambiental local, pois a água atua como vetor de conexão entre atividades pecuárias, populações ribeirinhas e ecossistemas naturais, ampliando o risco de disseminação de patógenos e genes de RAM nesses ambientes.

Tabela 19: Análise dos resistomas das amostras AP2 e AP9.

Ponto	Número Contig	Banco de Dados	Start	End	Fita	Gene	GAPS	% de cobertura	% de identidade	Acesso
AP2	k141_18965	NCBI CARD ResFinder	1067	1880	+	blaOXA-60	10/12	99,14	97,32	NG_049800.1
AP9	k141_138929	CARD ResFinder	420	755	-	RbpA	11/12	95,94	83,38	HQ203032:0-345

Fonte: Elaborado pelo autor.

Tabela 20: BLAST NCBI das *contigs* e perfil de resistência dos genes do ponto AP2 e AP9.

Ponto	Número Contig	Identificação taxonômica por BLAST NCBI (% identidade)	Acesso	Gene	Resistência
AP2	k141_18965	<i>Ralstonia pickettii</i> (95,25%)	AF525303.2	blaOXA-60	Cefalosporinas; Carbapenêmicos
AP9	k141_138929	<i>Mycolicibacterium</i> (78,86%)	CP050192.1	RbpA	Rifamicina

Fonte: Elaborado pelo autor.

7. CONCLUSÃO

Através do desenvolvimento, testes e aplicação de *pipelines* de bioinformática, foi possível realizar a análise do microbioma e resistoma em corpos hídricos, utilizando dois montadores distintos (Flye e MEGAHIT). A aplicação dos *pipelines*, principalmente do Flye, demonstrou a eficácia de caracterizar a diversidade microbiana.

A identificação de bactérias com potencial patogênico, como membros dos gêneros *Pseudomonas*, *Burkholderia*, *Mycobacterium*, e *Ralstonia*, reforça a presença de agentes que podem representar riscos ambientais e à saúde pública nos corpos hídricos estudados.

A análise do resistoma evidenciou a presença de genes de resistência antimicrobianos, tanto para as amostras piloto quanto para as amostras principais, sendo os genes *blaOXA-60* e *RbpA*, localizados em montagens feitas pelo MEGAHIT. A ausência dessas detecções em montagens Flye sugere que estratégias complementares de montagem são importantes para a captura precisa dos genes de RAM, incluindo aqueles em baixa abundância ou em fragmentos.

Os genes encontrados em microrganismos ambientais, reforça a importância de monitorar toda a biodiversidade presente no ecossistema, uma vez que pode ocorrer a disseminação de genes via elementos genéticos móveis.

8. PERSPECTIVAS FUTURAS

A presença de elementos genéticos móveis associados à resistência ainda demanda investigação mais aprofundada. Os resultados obtidos neste trabalho indicam que tais elementos podem desempenhar papel relevante na persistência e disseminação da resistência antimicrobiana em matrizes aquáticas, abordagens mais aprofundadas são necessárias para compreender seus mecanismos de mobilidade e transferência horizontal.

Os *pipelines* desenvolvidos mostraram-se eficazes nas análises-piloto, demonstrando potencial para aplicação em diferentes contextos ambientais. As amostras-piloto analisadas validaram o desempenho das ferramentas utilizadas, evidenciando a consistência dos resultados metagenômicos e a capacidade do método em detectar genes de resistência com alta sensibilidade.

Como continuidade deste estudo, pretende-se reprocessar os sinais brutos gerados pelo sequenciador MinION Mk1C (formato POD5) utilizando diferentes programas de *basecalling*, a fim de aprimorar a qualidade dos dados e reanalisar as amostras por meio dos *pipelines* desenvolvidos. Essa etapa permitirá verificar a reprodutibilidade dos resultados e aprimorar a acurácia das análises.

Além disso, a complementação e otimização dos *pipelines* de bioinformática são etapas fundamentais nas análises metagenômicas. Processos de montagem exigem calibração criteriosa de parâmetros, de modo a minimizar vieses e evitar a geração de *contigs* quiméricos. Para alcançar maior confiabilidade nos resultados, recomenda-se uma investigação aprofundada dos metadados, permitindo o rastreamento de potenciais fontes de erro e a interpretação contextualizada da variabilidade ambiental observada.

REFERÊNCIAS

ABRANTES, J. A.; NOGUEIRA, J. M. da R. Biofilme e células persisters: da persistência à resistência microbiana. **RBAC**, v. 54, n. 3, p. 228-234, 2022.

AKACIN, I. *et al.* Comparing the significance of the utilization of next generation and third generation technologies in microbial metagenomics. **Microbiological Research**, v. 264, p. 127154, 2022.

ALCOCK, Brian P. *et al.* CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. **Nucleic acids research**, v. 51, n. D1, p. D690-D699, 2023.

ARNOLD, Brian J.; HUANG, I.-Ting; HANAGE, William P. Horizontal gene transfer and adaptive evolution in bacteria. **Nature Reviews Microbiology**, v. 20, n. 4, p. 206-218, 2022.

ASLAM, B. *et al.* Antibiotic resistance: one health one world outlook. **Frontiers in cellular and infection microbiology**, v. 11, p. 771510, 2021.

ASSUNÇÃO, J.; GANDOUR, C.; ROCHA, R. DETER-ing Deforestation in the Amazon: Environmental Monitoring and Law Enforcement. **American Economic Journal: Applied Economics**, v. 15, n. 2, p. 125-156, 2023.

BAKER, Jonathon L. Using nanopore sequencing to obtain complete bacterial genomes from saliva samples. **Msystems**, v. 7, n. 5, p. e00491-22, 2022.

BANAR, M. *et al.* Oral bacteriophages: metagenomic clues to interpret microbiomes. **PeerJ**, v. 12, p. e16947, 2024.

BENOIT, Gaëtan *et al.* High-quality metagenome assembly from long accurate reads with metaMDBG. **Nature Biotechnology**, v. 42, n. 9, p. 1378-1383, 2024.

BLANKENBERG, Daniel *et al.* Galaxy: a web-based genome analysis tool for experimentalists. **Current protocols in molecular biology**, v. 89, n. 1, p. 19.10. 1-19.10. 21, 2010.

BORTOLAIA, Valeria *et al.* ResFinder 4.0 for predictions of phenotypes from genotypes. **Journal of Antimicrobial Chemotherapy**, v. 75, n. 12, p. 3491-3500, 2020.

Bracing for Superbugs: Strengthening environmental action in the One Health response to antimicrobial resistance – UNEP/OMS | United Nations Environment Programme. Disponível em: < https://www.unep.org/resources/superbugs/environmental-action?gad_source=1&gclid=Cj0KCQjwhb60BhCIARIsABGGtw_cgtrhSgdhJJuAd7Sf7INTIGD2yEr9tw3mInEuFTBr5WS1Pqq96RUaAjjkEALw_wcB >. Acesso em: 16 mar. 2024.

BRANCO, N. M. C. *et al.* Ocorrência de antimicrobianos em águas superficiais e residuais do Município do Rio de Janeiro: uma questão de vulnerabilidade

ambiental e de saúde pública. **Research, Society and Development**, v. 10, n. 10, p. e415101019000-e415101019000, 2021.

BRASIL. **Taxas PRODES Amazônia - 1988 a 2021 (km²)**. 30 nov. 2022. Disponível em: <<http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes>>. Acesso em: 03 nov. 2023.

BREITWIESER, Florian P.; SALZBERG, Steven L. Pavian: interactive analysis of metagenomics data for microbiome studies and pathogen identification. **Bioinformatics**, v. 36, n. 4, p. 1303-1304, 2020.

BRONZ, D. O desmonte ambiental pela via dos incêndios florestais na Amazônia brasileira. **Horizontes Antropológicos**, v. 29, p. e660401, 2023.

BROWN, Bonnie L. *et al.* MinION™ nanopore sequencing of environmental metagenomes: a synthetic approach. **Gigascience**, v. 6, n. 3, p. gix007, 2017.

CAMACHO, C. *et al.* BLAST+: architecture and applications **BMC Bioinformatics** 10: 421 doi: 10.1186. 2009.

CREÃO, Camilo Pantoja. **Principais causas de condenação total de carcaças bovídeas em abatedouro frigorífico sob inspeção estadual no Amapá, Brasil**. 2025.

CARNEIRO, Paulo AM *et al.* Molecular characterization of *Mycobacterium bovis* infection in cattle and buffalo in Amazon Region, Brazil. **Veterinary medicine and science**, v. 6, n. 1, p. 133-141, 2020.

CASTRO, A. B. L.; DE OLIVEIRA, M. A. Saneamento básico de Manaus e as doenças provenientes Basic sanitation of Manaus and diseases from. **Brazilian Journal of Health Review**, v. 4, n. 6, p. 26792-26807, 2021.

CASTRO, Lorena M. *et al.* Microbial prospection of an Amazonian blackwater lake and whole-genome sequencing of bacteria capable of polyhydroxyalkanoate synthesis. **Polymer Journal**, v. 53, n. 1, p. 191-202, 2021.

CECÍLIO, R. A.; PIMENTEL, S. M.; ZANETTI, S. S. Modeling the influence of forest cover on streamflows by different approaches. **Catena**, v. 178, p. 49-58, 2019.

CHEN, L.; KUMAR, S.; WU, H. A review of current antibiotic resistance and promising antibiotics with novel modes of action to combat antibiotic resistance. **Archives of Microbiology**, v. 205, n. 11, p. 356, 2023.

CORREIA, I. M. G. *et al.* Mata ciliar, conservação e sustentabilidade, fundamentos da importância para o semiárido paraibano: estudo de caso no alto curso do Rio Paraíba. **Revista de Geociências do Nordeste**, v. 5, n. 2, p. 41-60, 2019.

DA SILVA, Sheila *et al.* The piranha gut microbiome provides a selective lens into river water biodiversity. **Scientific Reports**, v. 14, n. 1, p. 21518, 2024.

DE COSTER, Wouter *et al.* NanoPack: visualizing and processing long-read sequencing data. **Bioinformatics**, v. 34, n. 15, p. 2666-2669, 2018.

DE LA CERDA, Gisel Y. *et al.* Balancing read length and sequencing depth: Optimizing Nanopore long-read sequencing for monocots with an emphasis on the Liliales. **Applications in Plant Sciences**, v. 11, n. 3, p. e11524, 2023.

DE SENA BRANDINE, Guilherme; SMITH, Andrew D. Falco: high-speed FastQC emulation for quality control of sequencing data. **F1000Research**, v. 8, p. 1874, 2021.

DEMKINA, Alina *et al.* Benchmarking DNA isolation methods for marine metagenomics. **Scientific reports**, v. 13, n. 1, p. 22138, 2023.

DESPOTOVIC, M. *et al.* Reservoirs of antimicrobial resistance in the context of One Health. **Current opinion in microbiology**, v. 73, p. 102291, 2023.

DOMINGUES, Ivan. **Conhecimento e transdisciplinaridade II: aspectos metodológicos**. Editora UFMG, 2005.

DOS SANTOS, Levy Assis *et al.* Biodiversity of carbapenem-resistant bacteria in clinical samples from the Southwest Amazon region (Rondônia/Brazil). **Scientific Reports**, v. 14, n. 1, p. 9383, 2024.

DURAZZI, Francesco *et al.* Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota. **Scientific reports**, v. 11, n. 1, p. 3030, 2021.

FERREIRA, V. *et al.* Occurrence of fecal bacteria and zoonotic pathogens in different water bodies: Supporting water quality management. **Water**, v. 14, n. 5, p. 780, 2022.

FONSECA, A. *et al.* **Sistema de Alerta de Desmatamento (SAD)** – Janeiro de 2022. Belém: Imazon, 2022. Disponível em: <<https://imazon.org.br/imprensa/desmatamento-em-areas-protegidas-cai-quase-quatro-vezes-na-amazonia-em-2023/>>. Acesso em: 13 mar. 2024.

GOLDBERG, Caren S. *et al.* Critical considerations for the application of environmental DNA methods to detect aquatic species. **Methods in ecology and evolution**, v. 7, n. 11, p. 1299-1307, 2016.

GOUSSAROV, Gleb *et al.* Benchmarking short-, long-and hybrid-read assemblers for metagenome sequencing of complex microbial communities. **Microbiology**, v. 170, n. 6, p. 001469, 2024.

GUPTA, C. L.; TIWARI, R. K.; CYTRYN, Eddie. Platforms for elucidating antibiotic resistance in single genomes and complex metagenomes. **Environment international**, v. 138, p. 105667, 2020.

GUPTA, Radhey S.; LO, Brian; SON, Jeen. Phylogenomics and comparative genomic studies robustly support division of the genus *Mycobacterium* into an emended genus *Mycobacterium* and four novel genera. **Frontiers in microbiology**, v. 9, p. 67, 2018.

- HAN, Haitao; WANG, Ziyue; ZHU, Shanfeng. Benchmarking metagenomic binning tools on real datasets across sequencing platforms and binning modes. **Nature Communications**, v. 16, n. 1, p. 2865, 2025.
- HU, Taishan *et al.* Next-generation sequencing technologies: An overview. **Human Immunology**, v. 82, n. 11, p. 801-811, 2021.
- IBGE. **Amazônia Legal**. 2022. Disponível em: <<https://www.ibge.gov.br/geociencias/cartas-e-mapas/mapas-regionais/15819-amazonia-legal.html>>. Acesso em: 03 nov. 2023.
- IMCHEN, M. *et al.* Current trends in experimental and computational approaches to combat antimicrobial resistance. **Frontiers in Genetics**, v. 11, p. 563975, 2020.
- KOLMOGOROV, Mikhail *et al.* Assembly of long, error-prone reads using repeat graphs. **Nature biotechnology**, v. 37, n. 5, p. 540-546, 2019.
- KOLMOGOROV, Mikhail *et al.* metaFlye: scalable long-read metagenome assembly using repeat graphs. **Nature methods**, v. 17, n. 11, p. 1103-1110, 2020.
- LAPIDUS, Alla L.; KOROBEYNIKOV, Anton I. Metagenomic data assembly—the way of decoding unknown microorganisms. **Frontiers in Microbiology**, v. 12, p. 613791, 2021.
- LI, Dinghua *et al.* MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. **Bioinformatics**, v. 31, n. 10, p. 1674-1676, 2015.
- LIGUORI, K. *et al.* Antimicrobial resistance monitoring of water environments: a framework for standardized methods and quality control. **Environmental science & technology**, v. 56, n. 13, p. 9149-9160, 2022.
- LIN, Bo; HUI, Jianan; MAO, Hongju. Nanopore technology and its applications in gene sequencing. **Biosensors**, v. 11, n. 7, p. 214, 2021.
- LU, Jennifer *et al.* Metagenome analysis using the Kraken software suite. **Nature protocols**, v. 17, n. 12, p. 2815-2839, 2022.
- MARTIN, Cristina *et al.* Nanopore-based metagenomics analysis reveals prevalence of mobile antibiotic and heavy metal resistome in wastewater. **Ecotoxicology**, v. 30, n. 8, p. 1572-1585, 2021.
- MATHIEU, Chloé *et al.* A systematic review of sources of variability and uncertainty in eDNA data for environmental monitoring. **Frontiers in Ecology and Evolution**, v. 8, p. 135, 2020.
- MCCOMBIE, W. Richard; MCPHERSON, John D.; MARDIS, Elaine R. Next-generation sequencing technologies. **Cold Spring Harbor perspectives in medicine**, v. 9, n. 11, p. a036798, 2019.

MEDEIROS, A. C.; LIMA, M. O.; GUIMARÃES, R. M. Avaliação da qualidade da água de consumo por comunidades ribeirinhas em áreas de exposição a poluentes urbanos e industriais nos municípios de Abaetetuba e Barcarena no estado do Pará, Brasil. **Ciência & Saúde Coletiva**, v. 21, p. 695-708, 2016.

MENEZES, T. C. C. Colunas de fogo, cortinas de fumaça e narrativas inflamáveis: multiplicação de incêndios florestais e as novas dinâmicas sociais da expansão da fronteira agropecuária amazônica. **Revista Trabalho Necessário**, v. 20, n. 41, p. 01-26, 2022.

MERLINO, J.; SIARAKAS, S. Antibiotic prescribing and antimicrobial resistance from an Australian perspective. **Microbial Drug Resistance**, v. 28, n. 5, p. 536-538, 2022.

MIKHEENKO, A; SAVELIEV, V; GUREVICH, A. MetaQUAST: avaliação de conjuntos de metagenoma, **Bioinformática** (2016) 32 (7): 1088-1090. doi: 10.1093/bioinformatics/btv697

MOBARKI, N.; ALMERABI, B.; HATTAN, A. Antibiotic resistance crisis. **Int. J. Med. Dev. Ctries**, v. 40, n. 4, p. 561-564, 2019.

MORAND, S.; LAJAUNIE, C. Outbreaks of vector-borne and zoonotic diseases are associated with changes in forest cover and oil palm expansion at global scale. **Frontiers in veterinary science**, v. 8, p. 230, 2021.

MUNITA, Jose M.; ARIAS, Cesar A. Mechanisms of antibiotic resistance. **Virulence mechanisms of bacterial pathogens**, p. 481-511, 2016.

MURRAY, C. J. L. *et al.* Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. **The lancet**, v. 399, n. 10325, p. 629-655, 2022.

NAVARRO, A. *et al.* Study of antibiotic resistance in freshwater ecosystems with low anthropogenic impact. **Science of The Total Environment**, v. 857, p. 159378, 2023.

O'NEILL, J *et al.* Tackling Drug-Resistant Infections Globally: Final Report and Recommendations. London: **The Review on Antimicrobial Resistance**. 2016. Disponível em: https://amrreview.org/sites/default/files/160525_Final%20paper_with%20cover.pdf. Acesso em: 15 mar. 2024.

OLARU, I. D.; WALTHER, B.; SCHAUMBURG, F. Zoonotic sources and the spread of antimicrobial resistance from the perspective of low and middle-income countries. **Infectious Diseases of Poverty**, v. 12, n. 03, p. 73-87, 2023.

PAREEK, Chandra Shekhar; SMOCZYNSKI, Rafal; TRETYN, Andrzej. Sequencing technologies and genome sequencing. **Journal of applied genetics**, v. 52, p. 413-435, 2011.

PARSLEY, Meghan B. *et al.* Environmental DNA concentrations vary greatly across productive and degradative conditions, with implications for the precision of population estimates. **Scientific Reports**, v. 14, n. 1, p. 17392, 2024.

PORCY, C. *et al.* Avaliação microbiológica da água de consumo de casas localizadas em área alagada em um município do estado Amapá. **Revista Eletrônica Acervo Saúde**, v. 12, n. 4, p. e2938-e2938, 2020.

PRADO, R. B. Serviços ecossistêmicos: estado atual e desafios para a pesquisa na Amazônia. **Terceira Margem Amazônia**, v. 6, n. 16, p. 11-22, 2021.

PRIYAM, Anurag *et al.* Parameter exploration improves the accuracy of long-read genome assembly. **bioRxiv**, p. 2021.05. 28.446135, 2021.

PURUSHOTHAMAN, Srinithi *et al.* Evaluation of DNA extraction kits for long-read shotgun metagenomics using Oxford Nanopore sequencing for rapid taxonomic and antimicrobial resistance detection. **Scientific Reports**, v. 14, n. 1, p. 29531, 2024.

QUINCE, C. *et al.* Shotgun metagenomics, from sampling to analysis. **Nature biotechnology**, v. 35, n. 9, p. 833-844, 2017.

RAHMAN, Md Tanvir *et al.* Zoonotic diseases: etiology, impact, and control. **Microorganisms**, v. 8, n. 9, p. 1405, 2020.

RAMOS, Sérgio *et al.* Metagenomics insights into bacterial diversity and antibiotic resistome of the sewage in the city of Belém, Pará, Brazil. **Frontiers in Microbiology**, v. 15, p. 1466353, 2024.

REAUME, Ashley M. *et al.* Testing the Waters: A Comparison of DNA Extraction Methods for the Metabarcoding of Microbial Communities. **Available at SSRN 5347215**, 2025.

REDDY, S. *et al.* Antimicrobial resistance in urban river ecosystems. **Microbiological Research**, v. 263, p. 127135, 2022.

Relatório sinaliza aumento da resistência a antibióticos em infecções bacterianas em humanos - OPAS/OMS | Organização Pan-Americana da Saúde. Disponível em: <<https://www.paho.org/pt/noticias/9-12-2022-relatorio-sinaliza-aumento-da-resistencia-antibioticos-em-infeccoes-bacterianas>>. Acesso em: 15 mar. 2024.

REN, Z.; LUO, W. Metagenomic analysis reveals the diversity and distribution of antibiotic resistance genes in thermokarst lakes of the Yellow River Source Area. **Environmental Pollution**, v. 313, p. 120102, 2022.

RESENDE, Juliana Alves; SILVA, Vânia Lúcia da; DINIZ, Claudio Galuppo. Aquatic environments in the One Health context: modulating the antimicrobial resistance phenomenon. **Acta Limnologica Brasiliensia**, v. 32, p. e102, 2020.

RHOADS, Anthony; AU, Kin Fai. PacBio sequencing and its applications. **Genomics, Proteomics and Bioinformatics**, v. 13, n. 5, p. 278-289, 2015.

ROSSONI, R. A.; DE MORAES, M. L. Agropecuária e desmatamento na Amazônia Legal Brasileira: uma análise espacial entre 2007 e 2017. **Geografia em Questão**, v. 13, n. 3, 2020.

SAHU, Ashish *et al.* A systematic review on environmental DNA (eDNA) Science: An eco-friendly survey method for conservation and restoration of fragile ecosystems. **Ecological Indicators**, v. 173, p. 113441, 2025.

SÁNCHEZ-BAENA, A. M.; CAICEDO-BEJARANO, L. D.; CHÁVEZ-VIVAS, M. Structure of bacterial community with resistance to antibiotics in aquatic environments. A systematic review. **International journal of environmental research and public health**, v. 18, n. 5, p. 2348, 2021.

SANGER, Frederick; NICKLEN, Steven; COULSON, Alan R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the national academy of sciences**, v. 74, n. 12, p. 5463-5467, 1977.

SCALDAFERRI, L. G. *et al.* Formas de resistência microbiana e estratégias para minimizar sua ocorrência na terapia antimicrobiana: Revisão. **Pubvet**, v. 14, p. 163, 2020.

SEEMANN, T. 2016. **ABRicate: mass screening of contigs for antibiotic resistance genes**. Disponível em <<https://github.com/tseemann/abricate>>. Acesso em: 01 out. 2025.

SERNA, C.; GONZALEZ-ZORN, B. Antimicrobial resistance and One Health. **Revista Española de Quimioterapia**, v. 35, n. Supl 3, p. 37, 2022.

SHIVARAM, K. B. *et al.* Bacteriophage-based biosensors for detection of pathogenic microbes in wastewater. **Science of The Total Environment**, p. 165859, 2023.

SRINIVAS, M. *et al.* The Application of Metagenomics to Study Microbial Communities and Develop Desirable Traits in Fermented Foods. **Foods**, v. 11, n. 20, p. 3297, 2022.

STOJAN, Iva *et al.* Evaluation of DNA extraction methods and direct PCR in metabarcoding of mock and marine bacterial communities. **Frontiers in microbiology**, v. 14, p. 1151907, 2023.

SZAKÁLLAS, Nikolett *et al.* Can long-read sequencing tackle the barriers, which the next-generation could not? A review. **Pathology and Oncology Research**, v. 30, p. 1611676, 2024.

TEJERA, Darwin *et al.* *Ralstonia pickettii* bacteremia in hemodialysis patients: a report of two cases. **Revista Brasileira de terapia intensiva**, v. 28, p. 195-198, 2016.

TOLKKINEN, M.; VAARALA, S.; AROVIITA, J. The importance of riparian forest cover to the ecological status of agricultural streams in a nationwide assessment. **Water Resources Management**, v. 35, p. 4009-4020, 2021.

TORTORA, G. J.; FUNKE, B. R; CASE, C. L. **Microbiologia**. 12 ed, Porto Alegre: Artmed, 2017, 935 p.

TYLER, Andrea D. *et al.* Evaluation of Oxford Nanopore's MinION sequencing device for microbial whole genome sequencing applications. **Scientific reports**, v. 8, n. 1, p. 10931, 2018.

UNEP, UNITED NATIONS ENVIRONMENT PROGRAMME. **Antimicrobial resistance: a global threat**. Disponível em: <<https://www.unep.org/topics/chemicals-and-pollution-action/pollution-and-health/antimicrobial-resistance-global-threat>>. Acesso em: 16 mar. 2024.

UNITED NATIONS ENVIRONMENT PROGRAMME. **UNEP Frontiers 2016 Report; Emerging Issues Of Environmental Concern**. UN, 2016.

VIDAVER, Anne K. Uses of antimicrobials in plant agriculture. **Clinical infectious diseases**, v. 34, n. Supplement_3, p. S107-S110, 2002.

WANG, Yunhao *et al.* Nanopore sequencing technology, bioinformatics and applications. **Nature biotechnology**, v. 39, n. 11, p. 1348-1365, 2021.

WENSEL, C. R. *et al.* Next-generation sequencing: insights to advance clinical investigations of the microbiome. **The Journal of clinical investigation**, v. 132, n. 7, 2022.

WESGUERBER, Tiago *et al.* A influência do uso de antimicrobianos na produção animal e o impacto nos produtos derivados destinados à alimentação humana e animal. **Brazilian Journal of Animal and Environmental Research**, v. 7, n. 3, p. e71369-e71369, 2024.

WHO *et al.* Antimicrobial resistance and the United Nations sustainable development cooperation framework: Guidance for United Nations Country teams. **World Health Organization**, 2021.

WHO, World Health Organization, 2020. **Zoonoses**. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/zoonoses>> Acesso em: 12 mar. 2024.

WHO, World Health Organization. **Antimicrobial resistance**, 2023. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance>>. Acesso em: 17 mar. 2024.

WHO, World Health Organization. **One Health**, 2017. Disponível em: <<https://www.who.int/features/qa/one-health/en/>> Acesso em: 12 mar. 2024.

WICK R. R. 2017. **Porechop: adapter trimmer for Oxford Nanopore reads**. Disponível em <<https://github.com/rrwick/Porechop/>>. Acesso em: 01 out. 2025.

WICK, R. R. 2018. **Filtlong**. Disponível em: <<https://github.com/rrwick>>. Acesso em: 01 out. 2025.

WICK, R. R.; JUDD, Louise M.; HOLT, Kathryn E. Performance of neural network basecalling tools for Oxford Nanopore sequencing. **Genome biology**, v. 20, n. 1, p. 129, 2019.

WINCK, G. R. *et al.* Socioecological vulnerability and the risk of zoonotic disease emergence in Brazil. **Science Advances**, v. 8, n. 26, p. eabo5774, 2022.

WOOD, Derrick E.; SALZBERG, Steven L. Kraken: ultrafast metagenomic sequence classification using exact alignments. **Genome biology**, v. 15, n. 3, p. R46, 2014.

YANG, Chao *et al.* A review of computational tools for generating metagenome-assembled genomes from metagenomic sequencing data. **Computational and structural biotechnology journal**, v. 19, p. 6301-6314, 2021.

YASIR, Muhammad *et al.* New insights of bacterial communities in fermented vegetables from shotgun metagenomics and identification of antibiotic resistance genes and probiotic bacteria. **Food Research International**, v. 157, p. 111190, 2022.

ZHANG, Zhenmiao *et al.* Benchmarking genome assembly methods on metagenomic sequencing data. **Briefings in Bioinformatics**, v. 24, n. 2, p. bbad087, 2023.

ZHUANG, M. *et al.* Distribution of antibiotic resistance genes in the environment. **Environmental pollution**, v. 285, p. 117402, 2021.

APÊNDICE

Apêndice 01: Folder desenvolvido para a metodologia 5.6 do trabalho.

A RESISTÊNCIA É UM RISCO PARA A SAÚDE DE TODOS

Mortalidade
 Causa **1,27 milhão de mortes** anuais e contribui para mais 4,95 milhões. No Brasil são 221 mil óbitos.

Aumento no custo
 O tratamento da infecção por bactérias resistentes é mais **demorado, caro e de difícil acesso**.

Efeito econômico
24 milhões de pessoas podem ser levadas à extrema pobreza e chegar a 39 milhões de mortes até 2050.

SAÚDE DE TODOS

A resistência afeta a saúde Humana, Animal e Ambiental

Universidade Federal da Paraíba
 Programa de Pós-Graduação em Desenvolvimento e Meio Ambiente
 Laboratório de Biologia Molecular - LabiMol

Elaborado por:
 Discente: Davi Bezerra Melquiades
 Orientadora: Eloiza Helena Campana

Referências:
 Resistência aos antimicrobianos - Ministério da Saúde. Disponível: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-1/ram>. Acesso: jan, 2025.
 RAM no Brasil - Ministério da Saúde. Disponível: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-1/ram>. Acesso: jan, 2025.
 Resistência antimicrobiana - OPAS/OMS Organização Pan-Americana da Saúde. Disponível: <https://www.pia.org.br/pt/temas/resistencia-antimicrobiana/info>. Acesso: jan, 2025.

RESISTÊNCIA BACTERIANA AOS ANTIMICROBIANOS

Um risco para a saúde de **TODOS**

IMAGEM: Bernardo Rodrigues, 2023.

CONHECENDO BACTÉRIAS

As bactérias são seres vivos microscópicos, compostos por uma única célula, que não conseguimos ver a olho nu. Mesmo pequenas, elas se multiplicam rapidamente por divisão celular. Esses organismos estão presentes em diversos lugares, como no nosso corpo, na água, no solo e no ar. Algumas bactérias, conhecidas como patogênicas, podem causar doenças em humanos, animais e plantas.

ANTIMICROBIANOS

Antimicrobianos combatem bactérias, fungos e protozoários, sendo usados na saúde humana, animal e na agricultura. Podem ser bactericidas (eliminam bactérias) ou bacteriostáticos (inibem o crescimento das bactérias).

RESISTÊNCIA AOS ANTIMICROBIANOS

A resistência aos antimicrobianos é a capacidade de microrganismos sobreviverem aos efeitos de medicamentos, como antibióticos e antifúngicos, devido a mutações genéticas. Pode ocorrer naturalmente ou ser agravada por ações humanas, como uso inadequado de antibióticos, mudanças climáticas e problemas de saneamento.

O uso excessivo de antimicrobianos e o descarte inadequado de resíduos aceleram a resistência, impactando a saúde humana, animal e ambiental.

COMO COMBATER?

- Conscientização pública a partir de campanhas sobre o problema
- Utilizar antimicrobianos de acordo com o recomendado pelo seu médico
- Descarte correto dos antimicrobianos em farmácias e posto de coleta
- Melhorias na higienização diária e prevenção da propagação de infecções
- Redução do uso indiscriminado na indústria pecuária, veterinária e agricultura
- Diagnósticos mais rápidos e tratamentos específicos
- Ampliar equipes profissionais capacitados em doenças infecciosas
- Incentivar a pesquisa para melhorias e desenvolvimentos de novos medicamentos
- Manter a vigilância constante na disseminação dos genes de resistência aos antimicrobianos

Rocha, GNM, 2023.

Apêndice 02: Panfleto desenvolvido para a metodologia 5.6 do trabalho.

ANTIBIÓTICOS:

Usar com **Cuidado**,
Viver com **Saúde**!

O uso em animais também afeta você!



O que é o uso racional de antibióticos em animais?

É o uso **correto** de antibióticos para tratar infecções, com a **dose certa**, pelo **tempo certo** e somente quando realmente **necessário**.



Uma Só Saúde!
A saúde dos animais, humanos e do ambiente está conectada.

Por que isso importa?

 O uso exagerado de antibióticos em animais pode gerar bactérias resistentes.

 Essas bactérias podem chegar até as pessoas por meio da carne, do leite, do solo, da água e até do ar!

 Isso afeta a saúde dos animais, das pessoas e do meio ambiente.

Como contribuir

- ✓ Não exija antibióticos sem necessidade.
- ✓ Siga sempre a orientação do médico veterinário.
- ✓ Respeite o tempo de tratamento.

Vamos juntos cuidar da saúde de todos!

Elaborado por: Discente: Davi Bezerra Melquiades
Orientadora: Eloiza Helena Campana

