UNIVERSIDADE DE MOGI DAS CRUZES

PAULO SALLAROLA TAKAO

CARACTERIZAÇÃO DO PERFIL ZOONÓTICO DO MICROBIOMA SANGUÍNEO, SALIVAR E CUTÂNEO DE CÃES (*Canis lupus familiaris*) ATRAVÉS DE DADOS PÚBLICOS DE SEQUENCIAMENTO DE NOVA GERAÇÃO (NGS)

Mogi das Cruzes, SP

2024

UNIVERSIDADE DE MOGI DAS CRUZES

PAULO SALLAROLA TAKAO

CARACTERIZAÇÃO DO PERFIL ZOONÓTICO DO MICROBIOMA SANGUÍNEO, SALIVAR E CUTÂNEO DE CÃES (*Canis lupus familiaris*) ATRAVÉS DE DADOS PÚBLICOS DE SEQUENCIAMENTO DE NOVA GERAÇÃO (NGS)

Dissertação apresentada ao Programa de Pós-Graduação da Universidade de Mogi das Cruzes (UMC) como parte dos requisitos para a obtenção do grau de Mestre em Engenharia Biomédica.
Área de Concentração: Informática Biomédica e Engenharia Clínica (Bioinformática e Biologia computacional).

Orientador: Prof. Dr. Fabiano Bezerra Menegidio

Coorientador: Prof. Dr. Rubens Pasa

Mogi das Cruzes, SP

2024

FINANCIAMENTO:



Ficha catalográfica
Universidade de Mogi das Cruzes - Biblioteca Central

Takao, Paulo Sallarola
T136c Caracterização do perfil zoonótico do microbioma sanguíneo, salivar e cutâneo de cães (*Canis lúpus familiaris*) através de dados públicos de Sequenciamento de Nova Geração (NGS) / Paulo Sallarola Takao. -- Mogi das Cruzes, SP: Universidade de Mogi das Cruzes, 2024.
179 f. : il.

Dissertação (Mestrado em Engenharia Biomédica) -- Universidade de Mogi das Cruzes, 2024.
Orientação: Prof. Dr. Fabiano Bezerra Menegidio.
Coorientação: Prof. Dr. Rubens Pasa.
Inclui bibliografia.

1. *Canis lúpus familiaris*/classificação 2. Zoonoses 3. Sequenciamento de nova geração (NGS) 4. Sequenciamento de nucleotídeos em larga escala 5. Saúde pública 6. Cães I. Menegidio, Fabiano Bezerra II. Universidade de Mogi das Cruzes III. Título

CDD 616.959

Elaborada por Maria Cristina Alves Silva Xavier – CRB-8/10077

ATAS

ATA DA SESSÃO PÚBLICA DE APRESENTAÇÃO DE DISSERTAÇÃO DE MESTRADO EM ENGENHARIA BIOMÉDICA DA UNIVERSIDADE DE MOGI DAS CRUZES

Às quatorze horas do dia seis de setembro de dois mil e vinte e quatro, pela Universidade de Mogi das Cruzes, realizou-se a defesa de dissertação intitulada "CARACTERIZAÇÃO DO PERFIL ZOONÓTICO DO MICROBIOMA SANGUÍNEO, SALIVAR E CUTÂNEO DE CÃES (CANIS LUPUS FAMILIARIS) ATRAVÉS DE DADOS PÚBLICOS DE SEQUENCIAMENTO DE NOVA GERAÇÃO (NGS)" para obtenção do grau de mestre pelo(a) candidato(a) **Paulo Sallarola Takao.** Tendo sido o número de créditos alcançados pelo(a) mesmo(a) no total de 54 (cinquenta e quatro), a saber: 30 unidades de crédito em disciplinas de pós-graduação e 24 unidades de crédito no preparo da dissertação, o(a) aluno(a) perfaz assim os requisitos para obtenção do grau de mestre. A comissão examinadora estava constituída dos senhores professores doutores: Fabiano Bezerra Menegidio, e Caio Junji Tanaka, da Universidade de Mogi das Cruzes, e Caroline Garcia, da Universidade Estadual do Sudoeste da Bahia, sob a presidência do primeiro, como orientador da dissertação. A sessão pública da defesa de dissertação foi aberta pelo senhor presidente da comissão que apresentou o(a) candidato(a). Em seguida o(a) candidato(a) realizou uma apresentação oral da dissertação. Ao final da apresentação da dissertação, seguiram-se as arguições pelos membros da comissão examinadora. A seguir a comissão, em sessão secreta, conforme julgamento discriminado por cada membro, considerou o(a) candidato(a)

Aprovado por Unanimidade

(aprovado(a) /reprovado(a) — (unanimidade/maioria)

Mogi das Cruzes, 06 de setembro de 2024.

Comissão Examinadora	Julgamento
Documento assinado digitalmente gov.br CAROLINE GARCIA Data: 26/09/2024 16:21:04-0300 Verifique em https://validar.iti.gov.br	
Profª Drª Caroline Garcia	Aprovado (aprovado(a)/reprovado(a))
Documento assinado digitalmente gov.br CAIO JUNJI TANAKA Data: 26/09/2024 13:57:11-0300 Verifique em https://validar.iti.gov.br	
Prof. Dr. Caio Junji Tanaka	Aprovado (aprovado(a)/reprovado(a))
Documento assinado digitalmente gov.br FABIANO BEZERRA MENEGIDIO Data: 25/09/2024 11:53:12-0300 Verifique em https://validar.iti.gov.br	
Prof. Dr. Fabiano Bezerra Menegidio	Aprovado (aprovado(a)/reprovado(a))

Dedico esta dissertação à minha filha, Elóra Sayuri Takao, que, mesmo com seus 11 anos, me inspira a ser um exemplo de dedicação e perseverança, e sonha em seguir os passos do pai como veterinária. Você é a razão do meu esforço e a luz da minha vida.

AGRADECIMENTOS

À Universidade de Mogi das Cruzes, por todos os desafios proporcionados e pela oportunidade de crescimento acadêmico e profissional. Agradeço por fornecer um ambiente estimulante e recursos essenciais para a realização deste trabalho. Ao corpo docente e à coordenação do Programa de Pós-Graduação em Engenharia Biomédica, por contribuírem significativamente para o meu desenvolvimento profissional e pessoal. Agradeço pelo conhecimento compartilhado, pela orientação e pelo apoio constante ao longo desta jornada. Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão da bolsa processo nº 131394/2022-3 que possibilitou a realização desta dissertação. Sem este suporte financeiro, a concretização deste trabalho não teria sido possível.

Ao Professor Doutor Fabiano B. Menegidio, meu orientador, por sua orientação incansável, paciência e incentivo ao longo deste trabalho. Sua expertise, apoio e confiança em meu potencial foram fundamentais para a realização desta pesquisa. Agradeço por me incentivar a ingressar no mestrado e por estar sempre disponível para guiar-me nos momentos de dúvida.

Aos meus pais, Vera Lúcia Sallarola Takao e Katsutoshi Takao, pelo amor, paciência e compreensão que sempre demonstraram. Agradeço por todo o apoio e incentivo em todas as etapas da minha vida acadêmica e profissional. Aos meus colegas de curso e amigos, que compartilharam comigo momentos de aprendizado, desafios e conquistas. Agradeço pelo companheirismo, pelas discussões enriquecedoras e pelo apoio mútuo. E, finalmente, a todos que, direta ou indiretamente, contribuíram para a realização desta dissertação. Meu sincero agradecimento.

“A medicina é a mesma para qualquer espécie. O que muda é simplesmente a abordagem ao paciente”

- Dr. Robert J. Cook

RESUMO

Este estudo investiga o perfil zoonótico do microbioma sanguíneo, salivar e cutâneo de cães (*Canis lupus familiaris*) utilizando dados públicos de sequenciamento de nova geração (NGS) disponíveis no banco de dados *Sequence Read Archive* (SRA) do *National Center for Biotechnology Information* (NCBI). A relevância desta pesquisa reside na crescente interação entre humanos e cães, o que aumenta o risco de transmissão de zoonoses. A dissertação é estruturada em cinco artigos. O primeiro artigo contextualiza a importância do microbioma canino na interface da saúde animal e pública, destacando os desafios metodológicos e a necessidade de padronização em estudos de microbioma. O segundo artigo apresenta uma revisão simplificada das ferramentas de metagenômica para não especialistas, com foco em ferramentas como Kraken2 e Bowtie2, essenciais para a classificação taxonômica e remoção de contaminantes. O terceiro artigo descreve a criação de bancos de dados integrados para Bowtie2 e Kraken2, otimizando a remoção de sequências do hospedeiro (Cão) e contaminantes humanos. O quarto artigo propõe uma nova abordagem combinando Kraken2 e Bowtie2 para a remoção de contaminantes e hospedeiros, além da classificação taxonômica, utilizando bibliotecas simuladas (*Mock*) para avaliar a eficácia das ferramentas. O quinto artigo apresenta um protocolo detalhado para a classificação metagenômica canina, integrando Kraken2, Bowtie2, Bracken e Pavian em um fluxo de trabalho otimizado. Os resultados demonstram a viabilidade dos bancos de dados criados e a eficácia da abordagem combinada de Kraken2 e Bowtie2 na remoção de contaminantes. A análise de amostras reais de sangue, saliva e pele de cães revelou uma diversidade microbiana significativa, com a saliva apresentando a maior diversidade, seguida pela pele e sangue. A presença de patógenos zoonóticos, como *Pasteurella* e *Capnocytophaga* na saliva canina, destaca a importância da pesquisa para a saúde pública. Este estudo contribui para o avanço do conhecimento sobre o microbioma canino e suas implicações zoonóticas, fornecendo ferramentas e protocolos otimizados para análises metagenômicas. A pesquisa ressalta a necessidade de estudos futuros para investigar a relação entre o microbioma canino e a saúde, bem como o impacto de fatores como raça, idade e dieta na composição do microbioma.

Palavras-chave: *Canis lupus familiaris*; zoonoses; microbioma; NGS; saúde pública.

Abstract

This study investigates the zoonotic profile of the blood, saliva, and skin microbiome of dogs (*Canis lupus familiaris*) using publicly available next-generation sequencing (NGS) data from the Sequence Read Archive (SRA) database of the National Center for Biotechnology Information (NCBI). The relevance of this research lies in the increasing interaction between humans and dogs, which elevates the risk of zoonotic transmission. The dissertation is structured in five articles. The first article contextualizes the importance of the canine microbiome at the interface of animal and public health, highlighting methodological challenges and the need for standardization in microbiome studies. The second article presents a simplified review of metagenomics tools for non-specialists, focusing on tools like Kraken2 and Bowtie2, essential for taxonomic classification and contaminant removal. The third article describes the creation of integrated databases for Bowtie2 and Kraken2, optimizing the removal of host (Dog) and human contaminant sequences. The fourth article proposes a novel approach combining Kraken2 and Bowtie2 for contaminant and host removal, as well as taxonomic classification, using simulated (*Mock*) libraries to assess the tools' effectiveness. The fifth article presents a detailed protocol for canine metagenomic classification, integrating Kraken2, Bowtie2, Bracken, and Pavian in an optimized workflow. The results demonstrate the viability of the created databases and the effectiveness of the combined Kraken2 and Bowtie2 approach in contaminant removal. Analysis of real blood, saliva, and skin samples from dogs revealed significant microbial diversity, with saliva exhibiting the highest diversity, followed by skin and blood. The presence of zoonotic pathogens, such as *Pasteurella* and *Capnocytophaga* in canine saliva, highlights the research's public health importance. This study contributes to advancing knowledge about the canine microbiome and its zoonotic implications, providing optimized tools and protocols for metagenomic analyses. The research emphasizes the need for future studies to investigate the relationship between the canine microbiome and health, as well as the impact of factors like breed, age, and diet on microbiome composition.

Keywords: *Canis lupus familiaris*, zoonoses, microbiome, NGS, public health

LISTA DE ILUSTRAÇÕES

Figura 1: **Comparativo entre os dois tipos de *workflow* mais utilizados**.....44
Figura 2: Exemplo de gráfico de qualidade de bases FastQC......46
Figura 3: Parte superior de um relatório MultiQC típico......47
Figura 4: Demonstração de como são realizados os alinhamentos na ferramenta kraken 2, ..50
Figura 5: Primeira linha do arquivo fasta do genoma de *Canis lupus familiaris*.....66
Figura 6: Primeira linha do arquivo fasta do genoma de *Canis lupus familiaris* com o Taxid: 67
Figura 7: *Workflow* 1 utilizado para a avaliação dos bancos de dados e ferramentas nos testes com as bibliotecas *MOCK*......77
Figura 8: *Workflow* 2 utilizado para a avaliação dos bancos de dados e ferramentas nos testes com as bibliotecas *MOCK*......77
Figura 9: *Workflow* 3 utilizado para a avaliação dos bancos de dados e ferramentas nos testes com as bibliotecas *MOCK*......77
Figura 10: Apresenta o relatório gerado pela ferramenta FastQC para a biblioteca *MOCK* de metagenoma.....80
Figura 11: Relatório gerado pela ferramenta fastp para a análise da biblioteca *MOCK* de metagenoma.....81
Figura 12: Passo a passo de uma análise metagenômica. 88
Figura 13: Na página da web da ferramenta Pavian.....97
Figura 14: Gerando arquivo HTML do arquivo *MOCK*_k2_bt2_pf_bracken.kreport, analisado pela ferramenta Pavian. 97
Figura 15: Mensagem exibida após a execução de ferramenta Bowtie2 98
Figura 16: Mensagem exibida após a execução de ferramenta Bowtie2 99
Figura 17: Mensagem exibida após a execução de ferramenta Kraken2 99
Figura 18: Mensagem após a execução da ferramenta Bracken 100
Figura 19: Tabela gerada no Pavian com as bactérias classificadas e número máximo de leituras. 101
Figura 20: Gráfico de Sankey gerado pela ferramenta Pavian.....102
F**igura 2**1: Workflow demonstrando todos os passos para a análise metagenômica completa. 115
Figura 22: Histograma de qualidade de sequencias obtidos com MultiQC. 116

LISTA DE TABELAS

Tabela 1: Dados dos genomas utilizados para a confecção dos bancos de dados de cão e ser humano. 65
Tabela 2: Teste realizado em uma biblioteca de metagenoma (SRR19324877) que possui 250.223.105 leituras. 68
Tabela 3: Composição da biblioteca *MOCK*. 75
Tabela 4: Parâmetros e configurações utilizados nos testes de avaliação dos bancos de dados Takao_K2db-plus_dog_human para Kraken2 e Takao_BT2db-plus_dog_human para Bowtie2. 78
Tabela 5: Arquivos de saída gerados pela ferramenta InSilicoSeq, com seus respectivos tamanhos e formatos. 79
Tabela 6: Relatório de abundância dos organismos presentes na biblioteca *MOCK* de metagenoma. 79
Tabela 7: Resultados dos testes de remoção de hospedeiro e contaminantes. 83
Tabela 8: Composição da biblioteca *MOCK* 92
Tabela 9: Genomas utilizados para a construção dos bancos de dados "Takao-db_dog_human". 93
Tabela 10: Tempo de execução e memória RAM consumida por cada ferramenta. 98
Tabela 11: Arquivos gerados pelas ferramentas após a execução e os arquivos *MOCK* 100
Tabela 12: Organismos classificados e com abundância corrigida por Bracken 100
Tabela 13: Apresenta os metadados detalhados de cada amostra 112
Tabela 14: Demonstra o número de gêneros de organismos presentes por tecid 116
Tabela 15: Esta tabela apresenta as métricas de alfa diversidade microbiana para diferentes amostras de pele de cães.. 118
Tabela 16: Esta tabela apresenta as métricas de alfa diversidade microbiana para diferentes amostras de saliva de cães. 120
Tabela 17 Esta tabela apresenta as métricas de alfa diversidade microbiana para diferentes amostras de sangue de cães.. 122
Tabela 18:A tabela a seguir ilustra a dificuldade de cultivo de alguns desses organismos, que requerem condições específicas ou anaerobiose, dificultando sua identificação por métodos tradicionais 125

LISTA DE GRÁFICOS

Gráfico 1: Comparativo entre as amostras de saliva, sangue e pele,.....117
Gráfico 2: O gráfico mostra a abundância relativa dos gêneros identificados em diferentes amostras de pele.....118
Gráfico 3: O *heatmap* mostra a diversidade beta das amostras de gênero na pele de cães, utilizando o índice de dissimilaridade de Bray-Curtis.....119
Gráfico 4: O gráfico mostra a abundância relativa dos gêneros identificados em diferentes amostras de saliva.....120
Gráfico 5: O *heatmap* mostra a diversidade beta das amostras de gênero da saliva de cães, utilizando o índice de dissimilaridade de Bray-Curtis.....121
Gráfico 6: O gráfico mostra a abundância dos gêneros identificados em diferentes amostras de sangue. 122
Gráfico 7: O *heatmap* mostra a diversidade beta das amostras de gênero de sangue de cães, utilizando o índice de dissimilaridade de Bray-Curtis.....122
Gráfico 8: Esse gráfico demonstras os microrganismos que possuem abundancia relativa maior que 1% do total de amostras e estão destacados por cor de acordo com o complexo bacteriano a qual pertence.....124

LISTA DE ABREVIATURAS E SIGLAS

ABINPET	Associação Brasileira da Indústria de Produtos para Animais de Estimação
CRMV/SP	Conselho Regional de Medicina Veterinária do Estado de São Paulo
BD	Banco de Dados.
bp	Pares de bases
CT	Contaminante
DNA	Ácido Desoxirribonucleico
FN	Falso Negativo
FP	Falso Positivo
GB	Gigabytes
GC	Guanina-Citosina (conteúdo de bases nitrogenadas)
HO	Hospedeiro
k-mers	Subsequências de comprimento k extraídas de uma sequência de DNA.
MOCK	Material de Controle Simulado
MOCK	Biblioteca simulada de metagenômica
NCBI	National Center for Biotechnology Information
NGS	Next-Generation Sequencing (Sequenciamento de Nova Geração)
OMS	Organização Mundial da Saúde
PCR	Reação em Cadeia da Polimerase
Q30	Score de qualidade de base
RAM	Resistência Antimicrobiana
RefSeq	Reference Sequence Database (Banco de Dados de Sequência de Referência)
RNA	Ácido ribonucleico
RNAr 16S	RNA ribossômico 16S
rRNA	RNA ribossômico
SAM	Sequence Alignment/Map
SRA	Sequence Read Archive (Arquivo de Leituras de Sequenciamento)
SUS	Sistema Único de Saúde
TB	Terabytes

SUMÁRIO

1. INTRODUÇÃO 17

2. FUNDAMENTAÇÃO TEÓRICA 20

2.1. Zoonoses: conceitos, classificação e importância para a saúde pública. 20

2.2. Implicações para a saúde pública e para a medicina veterinária: prevenção, diagnóstico e tratamento de zoonoses relacionadas aos cães 20

2.3. Avaliação do potencial zoonótico dos micro-organismos encontrados em cães 21

2.4. Caracterização de microrganismos em amostras de sangue, saliva e pele de cães: importância, métodos e desafios 22

2.5. Metagenômica e Microbioma: definição, diversidade e relação com a saúde do hospedeiro 23

2.6. Tecnologias de sequenciamento de nova geração: princípios e aplicações na análise de microbiomas caninos 25

3. OBJETIVO E ESCOPO DO TRABALHO 27

3.1. Objetivos específicos. 27

Artigo 1: Microbioma Sanguíneo, Salivar e Cutâneo de Cães: Uma Narrativa na Interface da Saúde Animal e Pública 28

1. INTRODUÇÃO 28

2. METODOLOGIA 30

3. DESENVOLVIMENTO 30

3.1. Microbioma Cutâneo Canino 30

3.2. Microbioma Sanguíneo 31

3.3. Microbioma Salivar 32

3.4. Potencial Zoonótico 33

3.5. Importância para a Saúde Pública 34

3.6. Desafios na Pesquisa do Microbioma 35

3.7. Utilização de Dados Públicos na Metagenômica 35

4. CONCLUSÃO 36
REFERÊNCIAS 37
Artigo 2: Ferramentas de Metagenômica: Uma Revisão Simplificada para Não Especialistas 42
1. INTRODUÇÃO 43
2. SEQUENCIAMENTO DE NOVA GERAÇÃO (NGS) 44
3. CONTROLE DE QUALIDADE. 45
3.1. FastQC 45
3.2 MultQC 46
3.3. Fastp 47
4. REMOÇÃO DE HOSPEDEIRO E CONTAMINANTES 48
4.1. Kraken 2 49
4.2. Bowtie 2 50
5. MONTAGEM DE NOVO 51
5.1. MEGAHIT 51
5.2. MetaSPAdes 52
6. ANÁLISE TAXONOMICA 53
6.1. Kraken 2 53
6.2. QIIME 2 53
7. ANÁLISES ESTATISTICAS 54
7.1. Bracken 54
7.2. Kraken Tools 55
7.3. Pavian 55
7.4. Microbiome Analyst 56
8. CONCLUSÃO 56
REFERÊNCIAS: 57

Artigo 3: Desenvolvimento de bancos de dados para análise metagenômica de microbiomas de cães: uma abordagem com kraken2 e bowtie2 61
1. INTRODUÇÃO 62
2. MATERIAIS E MÉTODOS 65
3. RESULTADOS E DISCUSSÃO 68
4. CONCLUSÃO 69
5. REFERÊNCIAS 69
Artigo 4: Uma nova abordagem com Kraken2 e Bowtie2 na remoção de contaminantes, hospedeiros e taxonomia de bibliotecas de metagenômica canina. 71
1. INTRODUÇÃO 72
2. MATERIAIS E MÉTODOS 74
2.1. Criação de biblioteca *MOCK* de metagenômica com InSilicoSeq 74
2.3. Testes realizados e avaliação dos bancos de dados 76
3. RESULTADOS E DISCUSSÃO 78
4. CONCLUSÃO 84
5. REFERÊNCIAS 84
4.3. Artigo 5: Protocolo de Classificação Metagenômica Canina 87
INTRODUÇÃO 87
1.1 Limitações na Remoção de Leituras de Hospedeiro e Contaminantes 89
1.2 Banco de Dados de Referência 90
1.3 Abordagens Baseadas em Montagem 91
2. MATERIAIS E MÉTODOS 91
2.1 Softwares Necessários 92
2.2 Hardware 92
2.3 Dados para Análise Teste 92
2.4 Bancos de Dados 93
2.5 Instalação das Ferramentas 94

2.6 Análise do Microbioma 95
3. RESULTADOS 98
4. DISCUSSÃO 102
4.1. Relevância Científica e Impacto Prático 103
4.2. Limitações e Desafios 103
4.3. Sugestões para Pesquisas Futuras 103
5. CONCLUSÃO 104
6. REFERÊNCIAS: 104
5. RESULTADOS 109
5.1. Artigo 6: Caracterização do perfil zoonóticos do microbioma sanguíneo, salivar e cutâneo de cães (*Canis lupus familiaris*) através de dados públicos de sequenciamento de nova geração (NGS) 109
2. METODOLOGIA 112
2.1 Amostras 112
2.2. Análise de qualidade. 113
2.3. Análises de Bioinformática 114
3. RESULTADOS 115
3.1. Pele 117
3.4. Saliva 119
3.5. Sangue 121
4. DISCUSSÃO 123
5. CONCLUSÃO 126
REFERÊNCIAS 127
ANEXOS 130
Anexo 1: Dados de Pele 130
Anexo 2: Dados de Saliva 131
Anexo 3: Dados de sangue. 137

6. DISCUSSÃO 138

6.1. Limitações e desafios que persistem: 139

6.2. Perspectivas futuras para a continuidade deste trabalho incluem: 139

7.CONCLUSÃO 140

REFERÊNCIAS 141

ANEXOS 150

Anexo1: Artigos publicados e Capítulos de livros 150

Anexo 3: Participações em Congressos e Apresentação de Trabalhos 153

Anexo 4: Cursos 158

Anexo 5: Orientação e Coorientação de alunos de iniciação científica. 162

Anexo 6: Captulo de livro 163

1 INTRODUÇÃO

O sequenciamento de ácidos nucleicos é agora parte integrante da ciência moderna. Usamos rotineiramente o sequenciamento de DNA em muitos campos da microbiologia, incluindo o rastreamento de doenças infecciosas e o estudo da diversidade das comunidades microbianas, como a microbiota humana [1–3]. A aplicação do Sequenciamento de Nova Geração (*Next-Generation Sequencing*; NGS) nas análises de microbioma tem revolucionado a forma como estudamos e compreendemos a diversidade microbiana em diversos ambientes, incluindo o microbioma animal [3].

O sequenciamento de ácidos nucleicos, iniciado com o método de Sanger em 1977 [4], tornou-se uma ferramenta essencial na ciência moderna, impulsionando projetos como o Genoma Humano [5] e o estudo da microbiota. A metagenômica, por sua vez, revolucionou a análise de comunidades microbianas. Desde sua primeira menção em 1998 por Handelsman e colaboradores [6], a metagenômica avançou rapidamente. Em 2004, o primeiro estudo em larga escala investigou a comunidade microbiana do Mar Sargasso [7]. Em 2005, o desenvolvimento do piro sequenciamento aumentou significativamente a capacidade de sequenciamento[8]. A plataforma Illumina, lançada em 2006, tornou a metagenômica mais acessível [9]. Em 2010, o Projeto Microbioma Humano foi iniciado para caracterizar a diversidade microbiana humana [10]. Atualmente, a metagenômica continua a se expandir, explorando microbiomas complexos e impactando áreas como medicina, agricultura e biotecnologia [11,12].

O microbioma, um ecossistema complexo de microrganismos que coexistem em nosso corpo e em outros seres vivos, tem despertado um interesse crescente não apenas na comunidade científica, mas também entre o público em geral. Esta área tem grande promessa para novos tratamentos médicos e veterinários, crescendo significativamente nas últimas duas décadas [13]. Há evidências crescentes de que a variação nas comunidades microbianas está associada à saúde geral do hospedeiro [14]. O risco ou a gravidade de inúmeras doenças e distúrbios em um hospedeiro estão associados ao microbioma. Compreender a diversidade e a função desses micróbios é fundamental não apenas para a saúde dos animais, mas também para a nossa compreensão de saúde pública, especialmente quando consideramos o potencial zoonótico desses microrganismos.

Neste contexto, a análise do potencial zoonótico dos micro-organismos presentes nos microbiomas caninos se torna uma peça fundamental na saúde pública, promovendo intervenções direcionadas, conscientização e estratégias eficazes para mitigar os riscos associados à coexistência entre humanos e seus companheiros caninos[15]. A interação íntima

entre seres humanos e seus cães transcende os limites da convivência diária, estendendo-se ao mundo microbiano compartilhado [15]. Os cães, estando dentro dos nossos lares e tendo livre acesso a diversos ambientes, facilitam uma grande troca de microrganismos, aumentando a possibilidade de transmissão de zoonoses ou genes de resistência bacteriana[15–17]. Portanto, a compreensão do perfil zoonótico do microbioma sanguíneo, salivar e cutâneo de cães (*Canis lupus familiaris*) emerge como um campo de pesquisa crucial, uma vez que não apenas a saúde canina, mas também a saúde humana pode ser influenciada por essa interconexão microbiológica.

De acordo com o relatório de 2023 da Associação Brasileira da Indústria de Produtos para Animais de Estimação (ABINPET), o Brasil é o segundo maior país em população de cães no mundo, com aproximadamente 58,1 milhões e 900 milhões de cães, respectivamente [18]. A pele, o sangue e a saliva dos cães abrigam uma comunidade diversificada de micróbios [19–21]. Enquanto o sangue é tradicionalmente considerado estéril em indivíduos saudáveis, avanços tecnológicos permitiram a detecção de microrganismos circulantes em pequenas quantidades [22]. Esses microrganismos podem desempenhar um papel na modulação da resposta imunológica e na saúde geral do cão. Da mesma forma, a saliva, com sua composição única de enzimas e micróbios, desempenha um papel importante na saúde oral e na proteção contra patógenos invasores [21,23,24]. A pele, sendo o maior órgão do corpo canino, serve como uma barreira vital contra patógenos ambientais, mas também é um habitat ideal para uma variedade de micróbios, incluindo bactérias, fungos, vírus e eucariotos [25,26].

O aspecto mais significativo do microbioma canino, especialmente quando consideramos implicações para a saúde pública, é o seu potencial zoonótico [27–29]. O microbioma do cão pode servir como um reservatório de agentes patogênicos zoonóticos, como bactérias e vírus, que têm o potencial de infectar os humanos através do contato direto ou indireto com os animais [30,31]. Exemplos bem conhecidos de zoonoses associadas ao microbioma canino incluem a Leptospirose e a Raiva [32–34]. Além das doenças infecciosas clássicas, o microbioma canino também pode desempenhar um papel na transmissão de resistência antimicrobiana [17,35]. O uso indiscriminado de antibióticos em animais de estimação pode selecionar cepas bacterianas resistentes, que podem ser transmitidas aos humanos [21,35]. Diante dessas preocupações, é de extrema importância adotar uma abordagem integrada para gerenciar o microbioma canino, promovendo a saúde tanto dos animais quanto dos humanos. Isso inclui medidas como a promoção de práticas de higiene adequadas, vacinação regular, controle de parasitas e o uso responsável de antibióticos [36]. Além disso, uma compreensão mais profunda da dinâmica do

microbioma canino, incluindo sua composição, diversidade e função, é essencial para desenvolver estratégias eficazes de prevenção e controle de doenças zoonóticas.

Portanto, a pesquisa sobre o microbioma da pele, sangue e saliva canina emerge como uma área fascinante, com implicações cruciais tanto para a saúde animal quanto para a saúde pública. Ao investigar a complexidade desses microbiomas e seu potencial zoonótico, podemos fomentar uma convivência mais saudável entre humanos e cães, garantindo o bem-estar de ambas as espécies. Além disso, os avanços científicos e a crescente disponibilidade de dados publicamente acessíveis facilitaram a realização de estudos metagenômicos, aproveitando o vasto volume de dados subutilizados por alguns pesquisadores, que representam uma fonte valiosa para outros estudos [37–39].

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Zoonoses: conceitos, classificação e importância para a saúde pública

As zoonoses são definidas pela Organização Mundial da Saúde (OMS) como doenças ou infecções naturalmente transmissíveis entre animais vertebrados e seres humanos [40]. Este conceito abrange uma ampla gama de patógenos, incluindo bactérias, vírus, parasitas e fungos, que podem atravessar a barreira entre espécies. A importância das zoonoses é destacada pelo fato de que aproximadamente 60% das doenças infecciosas humanas conhecidas têm origem animal, e cerca de 75% das doenças infecciosas emergentes são zoonóticas [41]. A classificação das zoonoses é fundamental para compreender suas dinâmicas e desenvolver estratégias eficazes de prevenção e controle. De acordo com a direção da transmissão, as zoonoses podem ser categorizadas em três grupos principais [42,43].

1. Antropozoonoses: são zoonoses transmitidas dos animais para os humanos. Exemplos incluem a raiva e a brucelose.
2. Zooantroponoses: são zoonoses transmitidas dos humanos para os animais. Um exemplo é a tuberculose humana transmitida para bovinos.
3. Anfixenoses ou ciclozoonoses: são zoonoses que circulam entre humanos e animais, sem um hospedeiro preferencial claro. A influenza é um exemplo desta categoria.

A importância das zoonoses para a saúde pública é multifacetada. Além do impacto direto na saúde humana e animal, as zoonoses têm implicações econômicas significativas, afetando a produção de alimentos, o comércio e o turismo [44]. A globalização, as mudanças climáticas e a crescente proximidade entre humanos e animais em ambientes urbanos e rurais têm contribuído para a emergência e ré emergência de zoonoses [45].

2.2 Implicações para a saúde pública e para a medicina veterinária: prevenção, diagnóstico e tratamento de zoonoses relacionadas aos cães

Os médicos veterinários desempenham um papel crucial como agentes de saúde pública no controle de zoonoses, atuando tanto na clínica de pequenos animais quanto em equipes do Sistema Único de Saúde (SUS) [46]. Esta atuação é fundamental para a implementação da abordagem "*One-Health*" ou "Saúde Única", que reconhece a interconexão entre a saúde

humana, animal e ambiental [47].

Na prática clínica, o médico veterinário não apenas diagnostica e trata zoonoses, mas também promove o vínculo humano-animal de forma segura e saudável. As principais zoonoses emergentes e reemergentes transmitidas por cães incluem [48]:

- Virais: Raiva, Brucelose canina, Influenza.
- Bacterianas: Pasteurelose, *Staphylococcus aureus* resistente à meticilina, Leptospirose, Tuberculose, Bordetella (tosse dos canis).
- Fúngicas e parasitárias: Leishmaniose Visceral Canina, Esporotricose, Micobacterioses.

O diagnóstico preciso e o tratamento adequado são cruciais no manejo das zoonoses. Em alguns casos, como na raiva, medidas drásticas como a eutanásia do animal podem ser necessárias [32]. No entanto, há uma necessidade contínua de desenvolver melhores ferramentas de diagnóstico para ampliar nosso conhecimento sobre patógenos zoonóticos ou potencialmente zoonóticos presentes em nosso ambiente direto.

Os animais de estimação, especialmente os cães, atuam como excelentes sentinelas para identificar patógenos que podem infectar seres humanos, tanto em ambientes domésticos quanto durante viagens [49]. A transmissão de agentes patogênicos entre cães e humanos pode ocorrer através de diversas vias, incluindo contato direto, mordeduras, ingestão de alimentos contaminados ou exposição a ambientes compartilhados [50].

A importância crítica das zoonoses para a saúde pública é evidenciada por sua capacidade de impactar significativamente a morbidade e a mortalidade tanto em populações humanas quanto animais. Estudos que demonstram a presença de patógenos comuns em ambas as espécies enfatizam a necessidade de uma abordagem preventiva e de controle integrada [51].

2.3 Avaliação do potencial zoonótico dos micro-organismos encontrados em cães

Adentrando no cenário da interação entre humanos e cães, a análise do potencial zoonótico dos micro-organismos presentes nestes animais emerge como um campo crítico para a compreensão das dinâmicas de transmissão de doenças e suas implicações para a saúde pública. Ao examinar a transmissibilidade, patogenicidade e interação entre sistemas imunológicos, esta análise busca não apenas identificar riscos iminentes, mas também orientar

estratégias preventivas e intervenções eficazes [52].

A transmissibilidade dos micro-organismos entre cães e humanos é um ponto central de investigação. Mordeduras, contato direto, compartilhamento de objetos e mesmo a presença de vetores como carrapatos, pulgas e mosquitos podem facilitar a transferência de agentes patogênicos [33,52,53]. A compreensão dessas vias de transmissão é crucial para desenvolver medidas preventivas eficazes, reduzindo o potencial de contágio e fortalecendo a barreira entre as espécies.

A patogenicidade dos micro-organismos está intrinsecamente ligada à capacidade de causar doenças em ambos os hospedeiros, cães e humanos. A análise aprofundada desses agentes patogênicos inclui a identificação de fatores de virulência, carga bacteriana e características específicas que influenciam a gravidade da infecção. Essa compreensão é essencial para antecipar o curso clínico das doenças zoonóticas e informar protocolos de tratamento[54,55].

A interação entre os sistemas imunológicos dos cães e dos humanos é um aspecto fundamental na avaliação do potencial zoonótico. A resposta imune, tanto em termos de defesa efetiva quanto de possíveis respostas hiperativas, desempenha um papel crucial na determinação da gravidade da infecção. A identificação de modulações imunológicas em ambos os hospedeiros contribui para a predição de susceptibilidade e resistência a doenças zoonóticas[33,53]. A compreensão desses aspectos também se torna crucial para a tomada de decisões informadas em práticas veterinárias, cuidados com animais de estimação e políticas de saúde pública. Esta análise não apenas preenche lacunas de conhecimento, mas também serve como base para a implementação de estratégias eficazes de prevenção e controle, salvaguardando a saúde tanto dos animais quanto dos humanos que compartilham seu convívio.

2.4 Caracterização de microrganismos em amostras de sangue, saliva e pele de cães: importância, métodos e desafios

A caracterização de microrganismos em amostras de sangue, saliva e pele de cães é crucial na medicina veterinária e pesquisa biomédica, com implicações significativas para a saúde animal e pública. Essa análise é fundamental para compreender a microbiota canina, identificar potenciais patógenos, monitorar a saúde dos cães e contribuir para a detecção de doenças zoonóticas e a prevenção da transmissão de agentes infecciosos [52]. Microrganismos presentes nessas amostras podem influenciar diretamente a saúde canina. Por exemplo, a

presença de certos microrganismos no sangue pode indicar infecções sistêmicas [19,22], enquanto na saliva podem estar associados a doenças bucais ou sistêmicas [23,56]. A análise da microbiota cutânea pode revelar condições dermatológicas e infecções de pele [57,58]

Para a caracterização desses microrganismos, diferentes métodos são empregados, cada um com suas vantagens e limitações. Técnicas moleculares, como a Reação em Cadeia da Polimerase (PCR) e sequenciamento de DNA, são amplamente utilizadas para identificar e quantificar microrganismos específicos com alta sensibilidade e precisão. Técnicas de cultivo microbiológico tradicionais também são empregadas para isolar e caracterizar microrganismos viáveis, fornecendo informações sobre a sua susceptibilidade a antimicrobianos [59–61]. A escolha da abordagem mais adequada depende do objetivo da análise e das características das amostras. Apesar dos avanços tecnológicos, a caracterização de microrganismos ainda enfrenta desafios. A interpretação dos resultados frequentemente requer experiência e conhecimento especializado, especialmente na identificação de microrganismos não convencionais ou emergentes. Questões como resistência antimicrobiana, variabilidade da microbiota e contaminação de amostras podem complicar a análise e a interpretação dos dados [61].

2.5 Metagenômica e Microbioma: definição, diversidade e relação com a saúde do hospedeiro

O termo metagenômica origina-se da união dos conceitos estatísticos de meta-análise e genômica [62]. A meta-análise é uma poderosa ferramenta estatística que correlaciona quantitativamente e resume associações genuínas entre grandes conjuntos de dados e integra estudos independentes, mas relacionados [63]. Assim, como a metagenômica é uma abordagem de alto rendimento que gera um enorme conjunto de dados genômicos, a meta-análise é fundamental para discriminar e correlacionar os dados gerados pelo sequenciamento com dados anotados "*in-silico*", aumentando assim a compreensão da dinâmica da comunidade microbiana em estudo [64]. Também podemos dizer que a metagenômica é o processo de análise do microbioma.

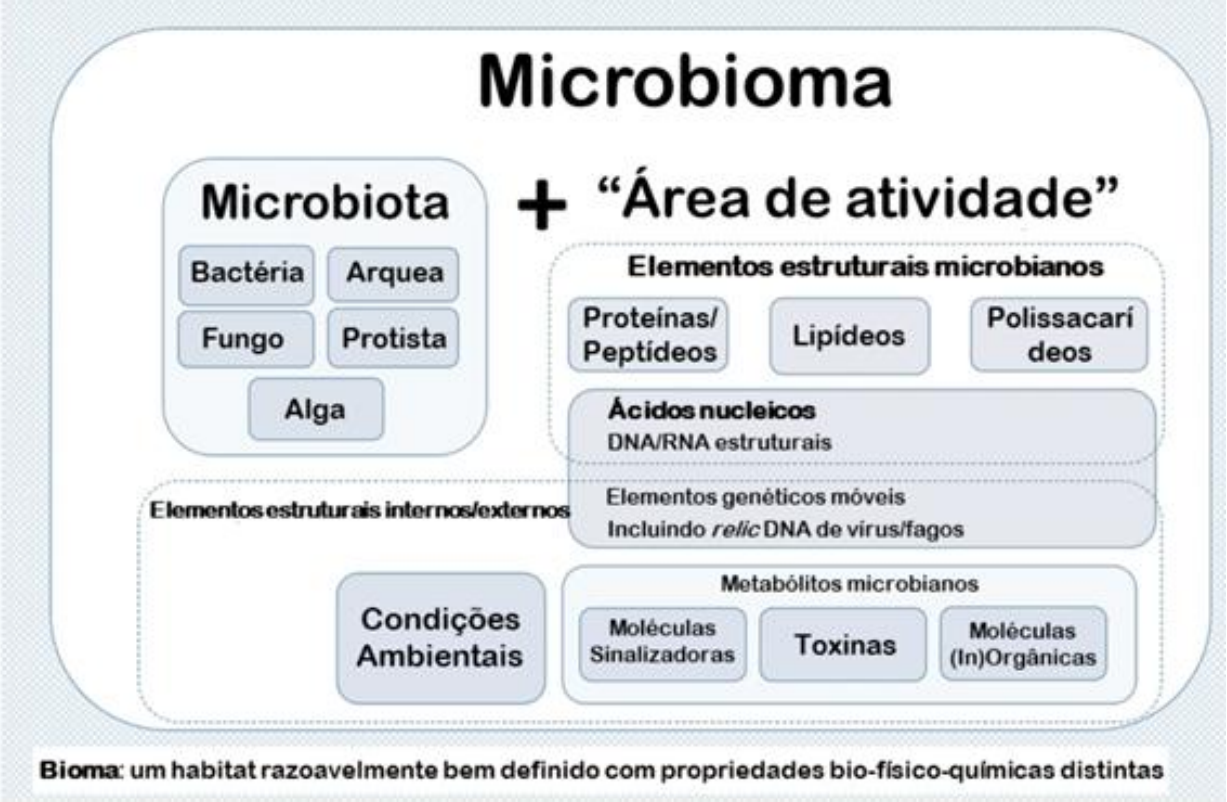
De acordo com Riesenfeld, Schloss e Handelsman (2004), metagenômica pode ser definida como a análise de DNA de comunidades microbianas em uma amostra sem a necessidade de cultura prévia [65] Por outro lado, Garza e Dutilh (2015) sugeriram que o uso do termo metagenômica se aplica apenas ao estudo de material genético recuperado diretamente de uma amostra e submetido à fragmentação aleatória (*shotgun*) e posterior sequenciamento [60].

Esta última definição tem sido preferida porque os métodos baseados em PCR independentes de cultura apenas amplificam as regiões do genoma alvo e podem subestimar a diversidade microbiana de uma amostra [66].

O microbioma de um indivíduo pode ser considerado um "último órgão" [67]. Como qualquer outro órgão, o microbioma tem fisiologia e patologia, e a saúde individual pode ser prejudicada quando sua estrutura populacional coletiva é alterada [67]. O microbioma constitui um ecossistema de microrganismos que coabitam em e sobre diversos organismos hospedeiros, abrangendo desde seres humanos, animais, plantas e até mesmo o ambiente. Essa comunidade microbiana engloba bactérias, arqueias, eucariotos e vírus, formando conjuntos altamente diversificados que desempenham papéis fundamentais na saúde e na funcionalidade do hospedeiro [13,67]. Neste estudo, adotamos a definição de microbioma conforme originalmente proposta por Whipps et all (1988), a qual não se limita apenas à presença dos microrganismos, mas abrange também sua "área de atividade" (Figura 1) [68]. Esta última engloba todo o espectro de moléculas geradas pelos microrganismos, incluindo seus componentes estruturais (ácidos nucleicos, proteínas, lipídios, polissacarídeos), metabólitos (moléculas sinalizadoras, toxinas, compostos orgânicos e inorgânicos) e até mesmo moléculas produzidas por hospedeiros coexistentes, influenciadas pelas condições ambientais circundantes. Assim, todos os elementos genéticos móveis, como fagos, vírus e DNA "remanescente" e extracelular, devem ser considerados como parte integrante do termo microbioma, embora não façam parte da microbiota propriamente dita [13].

Além disso, é essencial ponderar aspectos metodológicos para diferenciar o DNA proveniente de organismos vivos de seu ambiente circundante. É importante ressaltar que o termo microbioma é ocasionalmente confundido com o conceito de metagenoma. No entanto, o metagenoma é claramente definido como uma coleção de genomas e genes dos membros de uma microbiota, representando uma abordagem específica no estudo do material genético presente em uma comunidade microbiana [69]. Em outras palavras, o microbioma refere-se à comunidade de microrganismos e suas interações, enquanto o metagenoma é a soma de todo o material genético presente nessa comunidade [70].

Figura 1: Esquema destacando a composição do termo microbioma



Fonte: Adaptado de Barbosa [69].

Esquema destacando a composição do termo microbioma, contendo tanto a microbiota (comunidade de microrganismos) quanto sua "área de atividade" (elementos estruturais, metabólitos/moléculas sinalizadoras e as condições ambientais circundantes).

2.6 Tecnologias de sequenciamento de nova geração: princípios e aplicações na análise de microbiomas caninos

As tecnologias de sequenciamento de nova geração (NGS) revolucionaram a pesquisa em microbiomas, permitindo a análise aprofundada da composição e função de comunidades microbianas complexas, incluindo o microbioma canino.

O sequenciamento de Sanger, a primeira geração de sequenciamento, era preciso, mas trabalhoso e de baixo rendimento [13,70]. A partir de 2005, o advento do NGS, como o *piro sequenciamento* [71], possibilitou a geração de milhões de leituras de sequenciamento em uma única corrida, tornando a análise de microbiomas mais abrangente e acessível. Atualmente, a plataforma Illumina domina o mercado de NGS, utilizando o sequenciamento por síntese [3,9]. Este método envolve a fragmentação do DNA, ligação de adaptadores, amplificação para formar "aglomerados de DNA" e sequenciamento por incorporação de bases marcadas com

fluorescência. A Illumina oferece uma variedade de instrumentos, como o MiSeq e o HiSeq, com diferentes capacidades de rendimento e comprimento de leitura [9]. Para a análise de microbiomas caninos, dois métodos de sequenciamento NGS são amplamente utilizados:

- Sequenciamento *Amplicon*: Baseia-se na amplificação por PCR de uma região específica do DNA, como o gene 16S rRNA, para identificar e quantificar as diferentes espécies bacterianas presentes na amostra [72].
- Sequenciamento Metagenômico *Shotgun:* Envolve a fragmentação aleatória de todo o DNA da amostra, permitindo a análise de todos os genes presentes, incluindo genes bacterianos, virais e fúngicos [72].

A escolha do método de sequenciamento depende dos objetivos da pesquisa. O sequenciamento *amplicon* é mais econômico e eficiente para a análise da composição da comunidade bacteriana, enquanto o sequenciamento metagenômico *shotgun* fornece uma visão mais completa da diversidade funcional e do potencial metabólico do microbioma.

3 OBJETIVO E ESCOPO DO TRABALHO

A caracterização do perfil zoonótico do microbioma sanguíneo, salivar e cutâneo de cães através de dados públicos de Sequenciamento de Nova Geração é uma tarefa complexa que requer uma abordagem sistemática e reprodutível. Apesar da evidente necessidade de promover ambientes computacionais reprodutíveis neste campo, poucos esforços sistemáticos foram traçados nessa direção. O entendimento incompleto sobre reprodutibilidade, seus requisitos e suposições, associados às definições sobre ambientes computacionais e fluxo de trabalho em um campo heterogêneo como a bioinformática, têm colaborado para o lento desenvolvimento desse conceito. Diante deste cenário, este trabalho se fundamenta no desenvolvimento de soluções envolvendo tecnologias de bioinformática e análise de dados, como um importante recurso capaz de contribuir com a promoção da reprodutibilidade em análises metagenômicas. Estas soluções foram desenvolvidas tomando por base o conceito de análise de dados em larga escala.

3.1 Objetivos específicos.

- O desenvolvimento de uma revisão abrangente sobre o microbioma canino e as ferramentas bioinformáticas atuais, que possibilita a execução de uma grande variedade de análises de bioinformática de maneira reprodutível.
- A criação de um banco de dados específico para remoção de DNA do hospedeiro e contaminantes em análises metagenômicas de cães, visando aprimorar o desempenho das análises em diferentes níveis e camadas computacionais.
- A realização de testes extensivos com bibliotecas de metagenômica simuladas, introduzindo uma estratégia para garantir a validação e otimização das ferramentas bioinformáticas, assegurando resultados mais precisos e reprodutíveis.
- O desenvolvimento de um protocolo padronizado para análises metagenômicas, que servirá como uma plataforma virtual capaz de orientar futuras pesquisas neste campo.
- A identificação da diversidade de espécies microbianas e patógenos zoonóticos nos microbiomas caninos, contribuindo para o avanço do conhecimento na interface entre saúde animal e humana.

.

ARTIGO 1: Microbioma Sanguíneo, Salivar e Cutâneo de Cães: Uma Narrativa na Interface da Saúde Animal e Pública

Artigo a ser submetido:

Microbioma Sanguíneo, Salivar e Cutâneo de Cães: Uma Narrativa na Interface da Saúde Animal e Pública

Paulo Sallarola Takao[1]; Pâmela Souza Corrêa[1]; Rafaela Oliveira Campos[1]; Alessandra de Souza Fernandes[1]; Fabiana Borges[1]; David Aciole[1]; Renan Rodrigues Rocha[1]; Caio Junji Tanaka[1]; Rubens Pasa[2]; Karine Frehner Kavalco[2]; Fabiano Bezerra Menegidio[1]

1.Universidade de Mogi das Cruzes, UMC, Mogi das Cruzes, SP, Brasil

2. Universidade Federal de Viçosa, UFV, Viçosa, MG, Brasil

Resumo

Este artigo aborda os desafios e avanços na pesquisa do microbioma canino, destacando a importância de superar obstáculos metodológicos e a necessidade de padronização em estudos do microbioma. A inconsistência nas técnicas de coleta e análise de amostras, juntamente com a falta de metadados detalhados, limita a interpretação dos resultados e a aplicabilidade clínica dos estudos. A utilização de dados públicos na metagenômica é enfatizada como uma estratégia valiosa para ampliar o entendimento da diversidade microbiana, apesar dos desafios relacionados à qualidade e padronização dos dados. O artigo também explora o potencial dos diagnósticos baseados em microbioma para transformar práticas clínicas, reduzindo o uso precoce de antibióticos e tratamentos inespecíficos. Além disso, discute-se a relevância do estudo do microbioma sanguíneo, salivar e cutâneo de cães na interface da saúde animal e pública, sublinhando a importância da abordagem "One-Health". A pesquisa ressalta a necessidade de uma abordagem multidisciplinar que integre bioinformática, medicina veterinária e saúde pública para uma análise mais precisa e abrangente do microbioma, visando a promoção do bem-estar coletivo e a implementação de estratégias eficazes de prevenção e controle.

Palavras-chaves: Microbioma; Sequenciamento de Próxima Geração; Zoonoses; One-health; Metagenômica; Padronização; Dados Públicos.

1 INTRODUÇÃO

A relação profunda e complexa entre humanos (*Homo sapiens*) e cães (*canis lúpus familiaris*) ultrapassa os limites da estimação, estendendo-se ao microuniverso invisível que

compartilhamos, o microbioma. Nessa interconexão microbiana, a compreensão da estrutura e função dessa população tornou-se uma desafiadora, especialmente em face de ecossistemas microbianos vastos e complexos. Historicamente, a análise de populações bacterianas e fúngicas contava com métodos de cultura convencionais [1]. No entanto, a cultura, apesar de sua utilidade, impõe limitações substanciais na avaliação abrangente da microbiota. A variabilidade na capacidade de cultivar diversas bactérias e fungos e as restrições na profundidade do estudo resultam na subestimação da diversidade e na superestimação da presença de espécies específicas [2]. Essa subestimação, decorrente dos métodos baseados em cultura, motivou a busca por abordagens mais aprofundadas, sendo explorada com maior frequência no trato gastrointestinal [2–6]. Essa exploração desencadeou uma revolução na compreensão da composição microbiana corporal. Embora existam vários métodos independentes da cultura, a ênfase atual recai sobre as técnicas de sequenciamento de próxima geração (NGS) [2]. Essas técnicas têm a capacidade de diferenciar um vasto número de sequências individuais em misturas complexas, proporcionando uma visão mais abrangente das comunidades microbianas [1,2]. NGS surge como uma ferramenta promissora para aprofundar a compreensão da microbiota. Representando uma nova era no diagnóstico de infecções e doenças, oferecendo uma perspectivas inovadoras para práticas clínicas futuras [7]. Contudo, os métodos de diagnóstico padrão-ouro, prevalentes na prática humana e veterinária, ainda ancoram-se em abordagens baseadas em cultura ou PCR, muitas vezes exigindo dias para produzir resultados [8,9]. Essa espera prolongada frequentemente resulta na prescrição precoce e inadequada de antibióticos de amplo espectro e tratamentos inespecíficos, contribuindo significativamente para a crescente ameaça da resistência antimicrobiana (RAM) e tratamentos controversos [8]. Animais sob suspeita de infecções contagiosas aguardam diagnósticos, acarretando custos crescentes, desperdício de recursos e estresse desnecessário durante períodos de quarentena [9].

O entendimento do papel simbiótico da microbiota e microbioma, abrangendo microrganismos e sua união com genes, tem sido aprimorado pelos avanços nas técnicas de sequenciamento de alto rendimento. Este progresso, aliado à redução de custos, possibilitou uma exploração mais abrangente das comunidades microbianas em diversas espécies, incluindo os cães. Em diferentes regiões corporais, os cães hospedam comunidades microbianas distintas, influenciadas por fatores intrínsecos e extrínsecos, como idade, raça, sexo, ambiente e dieta [10]. Este artigo de revisão visa examinar a importância desses avanços na compreensão do microbioma sanguíneo, salivar e cutâneo dos cães, proporcionando uma visão mais abrangente da saúde canina.

2 METODOLOGIA

A pesquisa foi conduzida utilizando-se uma abordagem qualitativa com caráter exploratório, tendo como foco o estudo do microbioma sanguíneo, salivar e cutâneo de cães, no contexto da interface entre saúde animal e pública. Esse enfoque visou aprofundar a compreensão do problema em estudo, tornando-o mais claro e permitindo a formulação de hipóteses relacionadas ao fenômeno investigado[11,12].

A pesquisa foi estruturada com base em técnicas bibliográficas e documentais, a partir de uma seleção criteriosa de documentos publicados em periódicos indexados. Foram consultadas bases científicas como Web of Science, Scopus, PubMed, Google Acadêmico e Scielo para a identificação de artigos, teses, dissertações e monografias que discutem o microbioma em cães e sua relevância para a saúde animal e pública.

Com essas fontes, foi possível construir um referencial teórico robusto e realizar uma análise bibliográfica detalhada, estabelecendo discussões fundamentadas sobre a temática. As discussões apresentadas na seção de Desenvolvimento foram organizadas de modo a atender aos objetivos de pesquisa, culminando na análise detalhada da interface entre o microbioma canino e a saúde pública, além de propor possíveis soluções para os desafios identificados.

3 DESENVOLVIMENTO

Com base no referencial teórico apresentado na Introdução, esta seção de Desenvolvimento se dedica a apresentar a análise bibliográfica sobre o microbioma sanguíneo, salivar e cutâneo de cães, conforme os documentos recuperados nas buscas realizadas nas bases de dados Web of Science, Scopus, PubMed, Google Acadêmico e Scielo. Adicionalmente, são trazidas nesta seção considerações baseadas na análise dos estudos revisados, com o intuito de aprofundar a discussão sobre a interface entre saúde animal e pública, focando nas implicações e relevâncias desse microbioma.

3.1 Microbioma Cutâneo Canino

A pele, o maior órgão do corpo canino, desempenha uma função vital na proteção contra danos físicos, traumas químicos e na manutenção da homeostase corporal [10]. Este ecossistema complexo é habitado por milhões de microrganismos, incluindo bactérias e fungos, os quais

foram caracterizados na pele de cães saudáveis [10]. No âmbito canino, a pele serve como uma interface crucial entre o indivíduo e o ambiente externo, sendo recoberta por uma rica diversidade de microrganismos. Esses microrganismos, principalmente bactérias, desempenham interações fundamentais com as células do hospedeiro e os receptores imunes, mantendo um equilíbrio dinâmico [13]. O microbioma cutâneo dos cães saudáveis, composto principalmente por bactérias e em menor quantidade, fungos, representa uma população microbiana densamente povoada. Embora as informações sobre protozoários e vírus na pele canina sejam limitadas, a detecção do papilomavírus em cães saudáveis sugere a possível presença de vírus em condições fisiológicas [14].

Comparativamente ao microbioma humano, o microbioma cutâneo canino apresenta semelhanças notáveis, abrigando predominantemente os filos Actinobacteria, Firmicutes, Bacteroidetes e Proteobacteria [10,15]. Diferentes micro-habitats na pele, como sítios sebáceos, úmidos e secos, exibem uma diversidade especializada de táxons, refletindo a complexidade da microbiota cutânea [10,15]. Estudos iniciais indicam que os cães compartilham filos semelhantes aos encontrados na pele humana [15–17].

A disbiose da microbiota cutânea, associada a diversas afecções em humanos, também influencia a saúde dermatológica canina [16–19]. Notavelmente, a coabitação entre cães e humanos afeta a diversidade da microbiota da pele, evidenciando não apenas semelhanças entre donos e cães, mas também o impacto da posse de cães na diversidade da microbiota da pele humana [20].

Essas interações intricadas destacam a importância crítica do microbioma cutâneo canino não apenas na saúde dermatológica dos animais de estimação, mas também nas implicações para a saúde humana. A compreensão aprofundada dessas relações pode abrir portas para abordagens terapêuticas inovadoras e estratégias de manejo que beneficiem tanto os cães quanto seus cuidadores.

3.2 Microbioma Sanguíneo

O entendimento tradicional de que o sangue em indivíduos saudáveis é estéril tem sido desafiado por estudos recentes que identificaram a presença de DNA bacteriano circulante [21]. Essa pesquisa, que inicialmente se concentrava na saúde humana, tem se expandido para incluir a saúde animal, abrindo novas perspectivas na compreensão do microbioma sanguíneo. O sangue, considerado uma parte do corpo antes considerada estéril, agora é reconhecido como um meio onde microrganismos circulam, trazendo à tona conceitos previamente pouco

explorados [22–25]. Avanços tecnológicos, especialmente nas técnicas de sequenciamento, têm revelado que a presença de microrganismos no sangue não é exclusivamente indicativa de infecção, mas pode ser atribuída à colonização natural da corrente sanguínea [21].

A detecção de microrganismos no sangue de humanos saudáveis remonta a estudos pioneiros na década de 1960 [26]. Estudos subsequentes confirmaram a presença de material genético bacteriano no sangue, desafiando a visão convencional da esterilidade sanguínea [21,27–30]. Essa constatação estendeu-se também a animais domésticos e aves, indicando a universalidade desse fenômeno [6,31–34]. A hipótese sugere que as bactérias no sangue podem estar em estado de dormência ou presentes em formas específicas, como as formas L [6].A origem desses microrganismos ainda é objeto de estudo, mas uma hipótese é que a translocação do microbioma de diferentes locais biológicos, procedimentos clínicos ou lesões pode ser uma fonte significativa. A microbiota oral, a microbiota da pele e até mesmo a translocação bacteriana do intestino para o sangue são consideradas possíveis contribuintes [21].

Estudos mais recentes, baseados em diversas técnicas, como microscopia eletrônica, hibridização in situ fluorescente e sequenciamento de genes, consolidaram a compreensão da existência de um microbioma sanguíneo saudável [35]. Essa linha de pesquisa continua a evoluir, abrindo caminho para uma nova compreensão das interações entre microrganismos e o sistema circulatório em diferentes contextos de saúde.

3.3 Microbioma Salivar

A complexidade do microbioma oral canino, assim como seu papel crucial na saúde oral e defesa contra patógenos invasores, emerge como um campo de pesquisa vital. Desde os primórdios da microbiologia, quando Anton van Leeuwenhoek explorou a placa dental com seu rudimentar microscópio em 1683, as bactérias presentes na cavidade oral têm sido objeto de intenso interesse [36,37], no entanto, a pesquisa inicial, que se concentrava principalmente nas espécies humanas, apresentava limitações, especialmente na compreensão do microbioma oral de outros vertebrados[36]. Os estudos iniciais sobre microbiota oral dependiam fortemente de métodos cultiváveis, levando à identificação de aproximadamente 300 espécies na cavidade oral humana [36]. No entanto, com o avanço da tecnologia molecular, especialmente a análise de sequência de RNAr 16S, tornou-se claro que as bactérias em diferentes hospedeiros vertebrados apresentam singularidades, mesmo quando compartilham características fenotípicas semelhantes. Métodos moleculares, ao permitirem a identificação de organismos não

cultiváveis, revelaram uma diversidade bacteriana muito maior do que aquela obtida por métodos baseados em cultura[36]. Enquanto o microbioma oral humano tem sido extensivamente estudado utilizando métodos independentes de cultura [36], o mesmo não se aplica ao microbioma oral canino. Estudos anteriores em cães, muitas vezes baseados em métodos dependentes de cultura, focaram na identificação de espécies comumente encontradas em placas humanas, não refletindo totalmente a complexidade do microbioma oral canino [36].

Este hiato na pesquisa destaca a necessidade premente de explorar o microbioma salivar canino usando abordagens mais avançadas e abrangentes, a fim de compreender plenamente sua composição única de enzimas e micróbios, bem como seu impacto na saúde oral.

3.4 Potencial Zoonótico

Adentrando no cenário da interação entre humanos e cães, a análise do potencial zoonótico dos micro-organismos presentes nestes animais emerge como um campo crítico para a compreensão das dinâmicas de transmissão de doenças e suas implicações para a saúde pública.

Ao examinar a transmissibilidade, patogenicidade e interação entre sistemas imunológicos, esta análise busca não apenas identificar riscos iminentes, mas também orientar estratégias preventivas e intervenções eficazes [38].

A transmissibilidade dos micro-organismos entre cães e humanos é um ponto central de investigação. Mordeduras, contato direto, partilha de objetos e mesmo a presença de vetores como carrapatos, pulgas e mosquitos, podem facilitar a transferência de agentes patogênicos [38–40]. A compreensão destas vias de transmissão é crucial para desenvolver medidas preventivas eficazes, reduzindo o potencial de contágio e fortalecendo a barreira entre as espécies.

A patogenicidade dos micro-organismos é intrinsecamente ligada à capacidade de causar doenças em ambos os hospedeiros, cães e humanos. A análise aprofundada desses agentes patogênicos inclui a identificação de fatores de virulência, carga bacteriana e características específicas que influenciam a gravidade da infecção. Essa compreensão é essencial para antecipar o curso clínico das doenças zoonóticas e informar protocolos de tratamento[7,8,41].

A interação entre os sistemas imunológicos dos cães e dos humanos é um aspecto fundamental na avaliação do potencial zoonótico. A resposta imune, tanto em termos de defesa efetiva quanto de possíveis respostas hiperativas, desempenha um papel crucial na

determinação da gravidade da infecção. A identificação de modulações imunológicas em ambos os hospedeiros contribui para a predição de susceptibilidade e resistência a doenças zoonóticas [39,40].

A compreensão desses aspectos é crucial para a tomada de decisões informadas em práticas veterinárias, cuidados com animais de estimação e políticas de saúde pública. Esta análise não apenas preenche lacunas de conhecimento, mas também serve como base para a implementação de estratégias efetivas de prevenção e controle, salvaguardando a saúde tanto dos animais quanto dos humanos que compartilham seu convívio.

3.5 Importância para a Saúde Pública

Exercendo o papel de agentes de saúde pública, médicos veterinários atuam no controle de zoonoses também na clínica de pequenos animais e integrando equipes do Sistema Único de Saúde (SUS). Sob esta perspectiva, eles são cruciais para a prática dos pilares da Saúde Única (One-Health) [42]. Na área clínica, atuando no diagnóstico das zoonoses o médico-veterinário promove os benefícios do vínculo humano-animal.

Algumas das principais zoonoses emergentes e reemergentes transmitidas pelos cães são as virais como a Raiva a Brucelose canina e o vírus da gripe, as bacterinas como, a Pasteurelose, Staphylococus aureus resistente a meticilina, leptospira sp, tuberculose e tosse dos canis e zoonoses fúngicas e parasitárias como a Leishmaniose Visceral Canina, esporotricose e micobacterium [43]. Com a grande variedade de zoonoses é crucial o diagnóstico e o tratamento se for pertinente, devido alguns casos como o do vírus da raiva, ser indicado a eutanásia do animal. Porém ainda se faz necessário melhores ferramentas de diagnóstico par aumentaram nosso conhecimento dos patógenos zoonóticos ou potencialmente zoonóticos presentes em nosso ambiente direto. Os animais de estimação representam excelentes sentinelas para identificar os patógenos que podem infectar os seres humanos onde vivem ou quando as pessoas viajam com eles [43].

A importância crítica das zoonoses para a saúde pública é evidenciada pela capacidade dessas doenças de impactar significativamente a morbidade e a mortalidade em ambas as populações. A transmissão de agentes patogênicos entre cães e humanos pode ocorrer através de diversas vias, como contato direto, mordeduras, ingestão de alimentos contaminados ou exposição a ambientes compartilhados. A relevância de investigar a transmissibilidade desses

agentes é destacada por estudos que demonstram a presença de patógenos comuns em ambas as espécies, enfatizando a necessidade de uma abordagem preventiva e de controle [44,45]

3.6 Desafios na Pesquisa do Microbioma

A pesquisa do microbioma, apesar de seus notáveis avanços, confronta uma série de desafios intrínsecos que requerem abordagens inovadoras para um progresso eficaz [46].

A complexidade inerente ao microbioma, com sua diversidade abrangente de micro-organismos, emerge como um desafio significativo, dificultando a identificação precisa de espécies e a compreensão de suas interações. Nesse cenário, os avanços nas técnicas de sequenciamento de próxima geração e nas análises bioinformáticas surgem como elementos cruciais para lidar com a complexidade microbiômica [46,47].

Outro desafio crucial reside na falta de padronização nos métodos de coleta, processamento e análise de amostras, o que pode introduzir viés nos resultados e prejudicar comparações entre estudos. Estabelecer diretrizes e protocolos padronizados torna-se, portanto, imperativo para assegurar consistência metodológica e facilitar comparações robustas [46–48].

A ausência de metadados detalhados e contextos clínicos abrangentes surge como um desafio adicional, limitando a interpretação dos resultados e a aplicabilidade clínica. A promoção da coleta de metadados abrangentes, incluindo informações clínicas relevantes, torna-se crucial para correlações significativas [1,49]. Superar esses desafios na pesquisa microbiômica é vital para desvendar completamente o papel do microbioma em campos diversos, desde a saúde humana até a ecologia microbiana. O progresso contínuo nessas áreas não apenas aprimorará nosso entendimento fundamental, mas também abrirá portas para aplicações práticas significativas do conhecimento microbiômico.

3.7 Utilização de Dados Públicos na Metagenômica

Na era da genômica, a abundância de dados e a cultura robusta de compartilhamento de informações proporcionam uma rica base para a prática essencial de reutilização de conjuntos de dados públicos na investigação metagenômica[50–52]. O aumento da disponibilidade desses dados é impulsionado pela queda dos custos do sequenciamento do genoma completo, avanços técnicos rápidos e a expansão das instalações de gerenciamento de dados [53].

A análise de dados públicos de sequenciamento tornou-se uma abordagem poderosa e acessível para pesquisadores interessados em compreender a composição e função de comunidades microbianas em diversos ambientes e hospedeiros [54,55]. O crescimento exponencial de dados de sequenciamento disponíveis em bancos públicos, como o Sequence Read Archive (SRA) do National Center for Biotechnology Information (NCBI), oferece aos pesquisadores conjuntos abrangentes de dados de microbiomas e genomas completos [56–58]. Esses dados abertos permitem uma análise comparativa e abrangente da diversidade microbiana em diferentes contextos, sem a necessidade de gerar novos conjuntos de dados [55].

Ao utilizar dados públicos, os pesquisadores podem explorar uma variedade de ecossistemas, desde microbiomas humanos e animais até ambientes naturais, enriquecendo nossa compreensão da diversidade microbiana em diferentes contextos ambientais [46]. O acesso a esses dados também reduz custos e tempo, evitando a necessidade de gerar novos conjuntos de dados de sequenciamento [46,55].

Para realizar análises, os pesquisadores podem baixar conjuntos de dados públicos e aplicar pipelines de análise bioinformática para identificar e caracterizar as comunidades microbianas [59]. A integração de diferentes conjuntos de dados públicos possibilita a comparação entre estudos, facilitando a identificação de padrões e tendências em larga escala. No entanto, alguns desafios precisam ser superados, como a padronização de métodos de sequenciamento e análise, o controle de qualidade dos dados e a interpretação precisa dos resultados. É fundamental possuir uma compreensão sólida da biologia dos microbiomas e da análise bioinformática para evitar interpretações equivocadas dos dados de sequenciamento público [55,60].

4 CONCLUSÃO

Concluímos que é extrema importância o estudo aprofundado dos microbiomas caninos, não apenas para avançar nosso entendimento da saúde animal, mas também para elucidar as complexas interações entre a saúde humana e animal. Destacamos a necessidade de superar as limitações dos métodos tradicionais de cultura, promovendo o uso de dados de sequenciamento metagenômicos e técnicas de próxima geração para uma análise mais precisa e abrangente da diversidade microbiana. Este estudo destaca a importância da padronização dos métodos de sequenciamento e análise para garantir a qualidade dos dados e a interpretação correta dos

resultados, enfatizando a necessidade de uma abordagem multidisciplinar que integre a bioinformática, a medicina veterinária e a saúde pública.

Além disso, a pesquisa aponta para o potencial transformador dos diagnósticos baseados em microbioma, que podem revolucionar as práticas clínicas ao oferecer diagnósticos mais rápidos e precisos, reduzindo a prescrição precoce de antibióticos e tratamentos inespecíficos. Isso não apenas combate a ameaça da resistência antimicrobiana, mas também reduz o custo, o desperdício de recursos e o estresse desnecessário para os animais.

Finalmente, o artigo enfatiza a importância de compreender a transmissibilidade de micro-organismos entre cães e humanos, a patogenicidade dos agentes infecciosos e a interação entre os sistemas imunológicos de ambas as espécies. Essa compreensão é fundamental para o desenvolvimento de estratégias preventivas eficazes, salvaguardando a saúde de animais e humanos. Portanto, este estudo não apenas preenche lacunas de conhecimento existentes, mas também serve como uma base sólida para futuras pesquisas e a implementação de estratégias de prevenção e controle eficazes, reforçando a importância da abordagem "One Health" na promoção do bem-estar coletivo.

REFERÊNCIAS

1. Berg G, Rybakova D, Fischer D, Cernava T, Vergès MCC, Charles T, et al. Microbiome definition re-visited: old concepts and new challenges. Microbiome. 30 de junho de 2020;8(1):103.

2. Weese JS. The canine and feline skin microbiome in health and disease. Vet Dermatol. fevereiro de 2013;24(1):137-145.e31.

3. Joura MI, Brunner A, Nemes-Nikodém É, Sárdy M, Ostorházi E. Interactions between immune system and the microbiome of skin, blood and gut in pathogenesis of rosacea. Acta Microbiologica et Immunologica Hungarica. 29 de janeiro de 2021;68(1):1–6.

4. Kinross JM, Darzi AW, Nicholson JK. Gut microbiome-host interactions in health and disease. Genome Med, 3 (3), 14. biomedcentral. com/content/pdf; 2011.

5. Rostaher A, Morsy Y, Favrot C, Unterer S, Schnyder M, Scharl M, et al. Comparison of the Gut Microbiome between Atopic and Healthy Dogs-Preliminary Data. Animals (Basel). 12 de setembro de 2022;12(18).

6. Scarsella E, Sandri M, Monego SD, Licastro D, Stefanon B. Blood Microbiome: A New Marker of Gut Microbial Population in Dogs? Vet Sci. 4 de dezembro de 2020;7(4):198.

7. Tang S, Prem A, Tjokrosurjo J, Sary M, Van Bel MA, Rodrigues-Hoffmann A, et al. The canine skin and ear microbiome: A comprehensive survey of pathogens implicated in canine skin and ear infections using a novel next-generation-sequencing-based assay. Vet Microbiol. agosto de 2020;247:108764.

8. Ring N, Low AS, Wee B, Paterson GK, Nuttall T, Gally D, et al. Rapid metagenomic sequencing for diagnosis and antimicrobial sensitivity prediction of canine bacterial

infections. Microbial Genomics [Internet]. 20 de julho de 2023 [citado 2 de novembro de 2023];9(7). Disponível em: https://www.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.001066

9. Samrot AV, Sean TC, Bhavya KS, Sahithya CS, Chan-Drasekaran S, Palanisamy R, et al. Leptospiral Infection, Pathogenesis and Its Diagnosis-A Review. Pathogens. 1º de fevereiro de 2021;10(2):145.

10. Pereira AM, Clemente A. Dogs' Microbiome From Tip to Toe. Top Companion Anim Med. novembro de 2021;45:100584.

11. Merriam SB, Tisdell EJ. Qualitative Research: A Guide to Design and Implementation. John Wiley & Sons; 2015. 375 p.

12. Weyant E. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, 5th Edition: by John W. Creswell and J. David Creswell, Los Angeles, CA: SAGE, 2018, $38.34, 304pp., ISBN: 978-1506386706. Journal of Electronic Resources in Medical Libraries. 3 de abril de 2022;19(1–2):54–5.

13. Naik S, Bouladoux N, Wilhelm C, Molloy MJ, Salcedo R, Kastenmuller W, et al. Compartmentalized control of skin immunity by resident commensals. Science. 31 de agosto de 2012;337(6098):1115–9.

14. Medeiros-Fonseca B, Faustino-Rocha AI, Medeiros R, Oliveira PA, Gil da Costa RM. Canine and feline papillomaviruses: an update. Front Vet Sci. 16 de maio de 2023;10:1174673.

15. Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, et al. Topographical and temporal diversity of the human skin microbiome. Science. 29 de maio de 2009;324(5931):1190–2.

16. Cuscó A, Sánchez A, Altet L, Ferrer L, Francino O. Individual Signatures Define Canine Skin Microbiota Composition and Variability. Front Vet Sci. 2017;4:6.

17. Hoffmann AR, Patterson AP, Diesel A, Lawhon SD, Ly HJ, Stephenson CE, et al. The Skin Microbiome in Healthy and Allergic Dogs. PLOS ONE. 8 de janeiro de 2014;9(1):e83197.

18. Bradley CW, Morris DO, Rankin SC, Cain CL, Misic AM, Houser T, et al. Longitudinal Evaluation of the Skin Microbiome and Association with Microenvironment and Treatment in Canine Atopic Dermatitis. J Invest Dermatol. junho de 2016;136(6):1182–90.

19. Meason-Smith C, Diesel A, Patterson AP, Older CE, Mansell JM, Suchodolski JS, et al. What is living on your dog's skin? Characterization of the canine cutaneous mycobiota and fungal dysbiosis in canine allergic dermatitis. FEMS Microbiol Ecol. dezembro de 2015;91(12):fiv139.

20. Song SJ, Lauber C, Costello EK, Lozupone CA, Humphrey G, Berg-Lyons D, et al. Cohabiting family members share microbiota with one another and with their dogs. Elife. 16 de abril de 2013;2:e00458.

21. Scarsella E, Meineri G, Sandri M, Ganz HH, Stefanon B. Characterization of the Blood Microbiome and Comparison with the Fecal Microbiome in Healthy Dogs and Dogs with Gastrointestinal Disease. Veterinary Sciences. abril de 2023;10(4):277.

22. Ghose C, Ly M, Schwanemann LK, Shin JH, Atab K, Barr JJ, et al. The Virome of Cerebrospinal Fluid: Viruses Where We Once Thought There Were None. Front Microbiol. 2019;10:2061.

23. Kang Y, Ji X, Guo L, Xia H, Yang X, Xie Z, et al. Cerebrospinal Fluid from Healthy Pregnant Women Does Not Harbor a Detectable Microbial Community. Microbiol Spectr. 22 de dezembro de 2021;9(3):e0076921.

24. Lathe R, St Clair D. From conifers to cognition: Microbes, brain and behavior. Genes Brain Behav. novembro de 2020;19(8):e12680.

25. Liao H, Zhang Y, Guo W, Wang X, Wang H, Ye H, et al. Characterization of the Blood and Cerebrospinal Fluid Microbiome in Children with Bacterial Meningitis and Its Potential Correlation with Inflammation. mSystems. 29 de junho de 2021;6(3):e0004921.

26. Tedeschi GG, Amici D, Paparelli M. Incorporation of Nucleosides and Amino-acids in Human Erythrocyte Suspensions: Possible Relation with a Diffuse Infection of Mycoplasms or Bacteria in the L Form. Nature. junho de 1969;222(5200):1285–6.

27. Païssé S, Valle C, Servant F, Courtney M, Burcelin R, Amar J, et al. Comprehensive description of blood microbiome from healthy donors assessed by 16S targeted metagenomic sequencing. Transfusion. 2016;56(5):1138–47.

28. Panaiotov S, Filevski G, Equestre M, Nikolova E, Kalfin R. Cultural Isolation and Characteristics of the Blood Microbiome of Healthy Individuals. Advances in Microbiology. 29 de maio de 2018;8(5):406–21.

29. Velmurugan G, Dinakaran V, Rajendhran J, Swaminathan K. Blood Microbiota and Circulating Microbial Metabolites in Diabetes and Cardiovascular Disease. Trends in Endocrinology & Metabolism. 1º de novembro de 2020;31(11):835–47.

30. Whittle E, Leonard MO, Harrison R, Gant TW, Tonge DP. Multi-Method Characterization of the Human Circulating Microbiome. Frontiers in Microbiology [Internet]. 2019 [citado 3 de março de 2024];9. Disponível em: https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2018.03266

31. Mandal RK, Jiang T, Al-Rubaye AA, Rhoads DD, Wideman RF, Zhao J, et al. An investigation into blood microbiota and its potential association with Bacterial Chondronecrosis with Osteomyelitis (BCO) in Broilers. Sci Rep. 13 de maio de 2016;6(1):25882.

32. Scarsella E, Zecconi A, Cintio M, Stefanon B. Characterization of Microbiome on Feces, Blood and Milk in Dairy Cows with Different Milk Leucocyte Pattern. Animals. maio de 2021;11(5):1463.

33. Tilahun Y, Pinango JQ, Johnson F, Lett C, Smith K, Gipson T, et al. Transcript and blood-microbiome analysis towards a blood diagnostic tool for goats affected by Haemonchus contortus. Sci Rep. 30 de março de 2022;12(1):5362.

34. Vientós-Plotts AI, Ericsson AC, Rindt H, Grobman ME, Graham A, Bishop K, et al. Dynamic changes of the respiratory microbiota and its relationship to fecal and blood microbiota in healthy young cats. PLOS ONE. 9 de março de 2017;12(3):e0173818.

35. Nikkari S, McLaughlin IJ, Bi W, Dodge DE, Relman DA. Does Blood of Healthy Subjects Contain Bacterial Ribosomal DNA? Journal of Clinical Microbiology. maio de 2001;39(5):1956–9.

36. Dewhirst FE, Klein EA, Thompson EC, Blanton JM, Chen T, Milella L, et al. The Canine Oral Microbiome. PLoS One. 27 de abril de 2012;7(4):e36067.

37. Van Leeuwenhoek A. Letter of 17 September 1683 to the Royal Society, London. Royal Society, MS L. 17 de setembro de 1683;1:69.

38. Chomel BB, Sun B. Zoonoses in the Bedroom. Emerg Infect Dis. fevereiro de 2011;17(2):167–72.

39. Belotto A, Leanes LF, Schneider MC, Tamayo H, Correa E. Overview of rabies in the Americas. Virus Res. julho de 2005;111(1):5–12.

40. Cleaveland S, Dye C. Maintenance of a microparasite infecting several host species: rabies in the Serengeti. Parasitology. 1995;111 Suppl:S33-47.

41. Ciszewski M, Szewczyk EM. Potential {Factors} {Enabling} {Human} {Body} {Colonization} by {Animal} {Streptococcus} dysgalactiae subsp. equisimilis {Strains}. Curr Microbiol. 2017;74(5):650–4.

42. CRMV/SP. ZOONOSES CORRESPONDEM A MAIS DE 60% DAS DOENÇAS HUMANAS [Internet]. Conselho Regional de Medicina Veterinária São Paulo. 2022 [citado 11 de março de 2023]. Disponível em: https://crmvsp.gov.br/zoonoses-correspondem-a-mais-de-60-das-doencas-humanas/

43. Chomel B. Emerging and Re-Emerging Zoonoses of Dogs and Cats. Animals. 15 de julho de 2014;4(3):434–45.

44. Cleaveland S, Laurenson MK, Taylor LH. Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. Philos Trans R Soc Lond B Biol Sci. 29 de julho de 2001;356(1411):991–9.

45. One Health [Internet]. [citado 18 de fevereiro de 2024]. Disponível em: https://www.who.int/news-room/questions-and-answers/item/one-health

46. Cullen CM, Aneja KK, Beyhan S, Cho CE, Woloszynek S, Convertino M, et al. Emerging Priorities for Microbiome Research. Frontiers in Microbiology [Internet]. 2020 [citado 4 de março de 2024];11. Disponível em: https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2020.00136

47. Sergaki C, Lagunas B, Lidbury I, Gifford ML, Schäfer P. Challenges and Approaches in Microbiome Research: From Fundamental to Applied. Frontiers in Plant Science [Internet]. 2018 [citado 4 de março de 2024];9. Disponível em: https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2018.01205

48. Waldor MK, Tyson G, Borenstein E, Ochman H, Moeller A, Finlay BB, et al. Where Next for Microbiome Research? PLOS Biology. 20 de janeiro de 2015;13(1):e1002050.

49. Blaser M, Bork P, Fraser C, Knight R, Wang J. The microbiome explored: recent insights and future challenges. Nat Rev Microbiol. março de 2013;11(3):213–7.

50. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, et al. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet. dezembro de 2001;29(4):365–71.

51. Byrd JB, Greene AC, Prasad DV, Jiang X, Greene CS. Responsible, practical genomic data sharing that accelerates research. Nat Rev Genet. outubro de 2020;21(10):615–29.

52. Rung J, Brazma A. Reuse of public genome-wide gene expression data. Nat Rev Genet. fevereiro de 2013;14(2):89–99.

53. Jones KH, Daniels H, Squires E, Ford DV. Public Views on Models for Accessing Genomic and Health Data for Research: Mixed Methods Study. Journal of Medical Internet Research. 21 de agosto de 2019;21(8):e14384.

54. Gilbert JA, Blaser MJ, Caporaso JG, Jansson JK, Lynch SV, Knight R. Current understanding of the human microbiome. Nat Med. abril de 2018;24(4):392–400.

55. Kumar B, Lorusso E, Fosso B, Pesole G. A comprehensive overview of microbiome data in the light of machine learning applications: categorization, accessibility, and future directions. Front Microbiol. 13 de fevereiro de 2024;15:1343572.

56. Bethesda (MD): National Library of Medicine (US). National Center for Biotechnology Information (NCBI) [Internet]. National Center for Biotechnology Information (NCBI). 1988 [citado 2 de novembro de 2023]. Disponível em: https://www.ncbi.nlm.nih.gov/

57. Leinonen R, Sugawara H, Shumway M, on behalf of the International Nucleotide Sequence Database Collaboration. The Sequence Read Archive. Nucleic Acids Research. 1º de janeiro de 2011;39(suppl_1):D19–21.

58. RefSeq: NCBI Reference Sequence Database [Internet]. National Library of Medicine. 2023 [citado 5 de novembro de 2023]. Disponível em: https://www.ncbi.nlm.nih.gov/refseq/

59. Sielemann K, Hafner A, Pucker B. The reuse of public datasets in the life sciences: potential risks and rewards. PeerJ. 22 de setembro de 2020;8:e9954.

60. Amann RI, Baichoo S, Blencowe BJ, Bork P, Borodovsky M, Brooksbank C, et al. Toward unrestricted use of public genomic data. Science. 25 de janeiro de 2019;363(6425):350–2.

ARTIGO 2: Ferramentas de Metagenômica: Uma Revisão Simplificada para Não Especialistas

Ferramentas de Metagenômica: Uma Revisão Simplificada para Não Especialistas

Paulo Sallarola Takao[1]; Pâmela Souza Corrêa[1]; Rafaela Oliveira Campos[1]; Alessandra de Souza Fernandes[1]; Fabiana Borges[1]; David Aciole[1]; Renan Rodrigues Rocha[1]; Rubens Pasa[2]; Karine Frehner Kavalco[2]; Fabiano Bezerra Menegidio[1]

1.Universidade de Mogi das Cruzes, UMC, Mogi das Cruzes, SP, Brasil

2. Universidade Federal de Viçosa, UFV, Viçosa, MG, Brasil

RESUMO

A metagenômica, impulsionada pelo avanço das tecnologias de sequenciamento de nova geração (Next-generation sequencing, NGS) e ferramentas de bioinformática, tem revolucionado o estudo da diversidade microbiana em diversos ambientes. Este artigo de revisão aborda as principais ferramentas utilizadas na análise de dados metagenômicos, destacando a importância de processos como controle de qualidade, anotação taxonômica e análises estatísticas para gerar resultados confiáveis e significativos. Entre as ferramentas discutidas, o Kraken 2 se destaca pela sua eficiência na classificação taxonômica e na remoção de sequências de hospedeiros e contaminantes, utilizando correspondência exatas de k-mer para alcançar alta precisão. O artigo também ressalta a flexibilidade dos fluxos de trabalho metagenômicos, que podem incluir ou não a montagem *de novo* dos genomas, adaptando-se às necessidades específicas de cada projeto. A evolução contínua das tecnologias de sequenciamento de nova geração (NGS) e das ferramentas de bioinformática se apresenta como etapa crucial para aprofundar nosso entendimento sobre a complexidade da vida microbiana e suas funções em diferentes ecossistemas. Este artigo visa simplificar a compreensão dessas ferramentas para pesquisadores não especializados em bioinformática, contribuindo para a expansão do conhecimento na área de metagenômica e suas aplicações em saúde humana, agricultura, conservação ambiental, entre outros campos.

Palavras-chave: Metagenômica; Bioinformática; Sequenciamento de Nova Geração (NGS); Análise de Dados; Diversidade Microbiana

1 INTRODUÇÃO

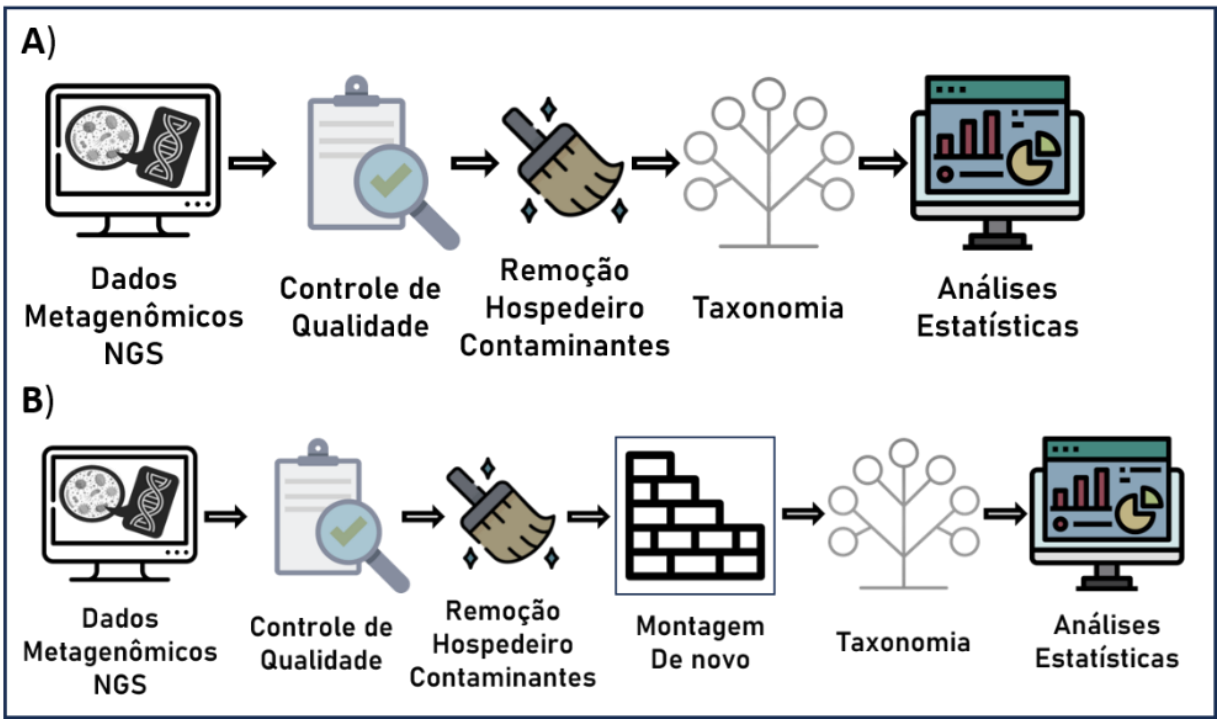
A abordagem da metagenômica transformou significativamente a forma como examinamos amostras de microrganismos. Trata-se de uma metodologia inovadora para investigar microrganismos coletados de um ambiente específico, por meio de triagem genética funcional ou análise de sequenciamento genético[1]. Um dos maiores desafios em microbiologia é a identificação do táxon ou dos táxons microbianos presentes em amostras biológicas e ambientais complexas, desde a determinação da etiologia de uma infecção a partir de uma amostra de sangue de um paciente até o levantamento das bactérias em uma amostra ambiental de solo[2–6]. Com o NGS e a metagenômica, essa identificação tornou-se mais rápida, levando à descoberta de novas espécies ainda não conhecidas[2,7,8].

A metagenômica é a análise do metagenoma que de acordo com Wajid et al.[9], de forma exata, "metagenoma" é uma coleção de genomas independentes de diferentes organismos trabalhando juntos harmoniosamente de maneira complexa, mas necessária, de modo que a estabilidade do ecossistema depende de sua eufonia unida (harmonia)[9].

Nas últimas décadas está ocorrendo um crescimento exponencial no volume de genomas microbianos que vem sendo gerados por NGS e assim o número de novas leituras de sequenciamento vem aumentando junto[2], e, com esse aumento, o desenvolvimento de novas tecnologias para análise desses dados vem surgindo de forma contínua. Diante disso, avaliar continuamente essas ferramentas se faz necessário. Diversos grupos já fizeram essa avaliação [2,10,11], porém a falta de artigos que descrevam essas ferramentas de forma clara e simples se faz presente. Para a análise de dados de metagenomas dois *workflow*s são mais utilizados, ambos descritos na Figura 1 e o que diferencia um do outro é a montagem ou não dos genomas antes de realizar a taxonomia que dependendo do tipo de análise ela não se faz necessária.

Nosso objetivo é descrever, de forma simples e clara, algumas das principais ferramentas de metagenômica, podendo assim elevar o entendimento dos processos para um não bioinformata.

Figura 1: Comparativo entre os dois tipos de *workflow* mais utilizados.



Fonte: Takao et al. (2024)

A) Primeiro realizamos o controle de qualidade para avaliar a amostra, em seguida realizamos a remoção de hospedeiro e contaminantes, assim deixando apenas dados do microbioma para posterior realizamos a analise taxonômica dos microrganismos presentes e posteriormente as análises estatísticas pertinentes

B) Primeiro realizamos o controle de qualidade para avaliar a amostra, em seguida realizamos a remoção de hospedeiro e contaminantes assim deixando apenas dados do microbioma e diferentemente da anterior, nessa etapa realizamos a montagem de novo dos genomas do restantes na amostra, posteriormente realizamos a analise taxonômica dos microrganismos presentes e em seguida as análises estatísticas pertinentes.

2 SEQUENCIAMENTO DE NOVA GERAÇÃO (NGS)

Esta tecnologia de alto rendimento, denominada "Sequenciamento de nova geração", permite a geração de milhares a milhões de leituras curtas de sequenciamento em uma única máquina[12]. Existem várias tecnologias NGS, porém a mais conhecida e utilizada é o sequenciamento por síntese, adotado pelos sequenciadores Illumina desde 2006, dominando atualmente uma parte significativa do mercado e sendo, até hoje, a mais utilizada[13]. O sequenciamento por síntese começa com um método chamado *shotgun* onde ocorre a fragmentação do DNA em pequenos fragmentos de fita simples e ligando adaptadores em suas extremidades, seguida pela amplificação desse DNA aumentar o volume de material genético

da amostra[14]. Esta geração de "aglomerados de DNA" é acompanhada da adição de bases terminadoras reversíveis marcadas com fluorescência (adenina, citosina, guanina e timina) ligadas a um grupo bloqueador [15]. As quatro bases competem, então, pelos locais de ligação no DNA modelo a ser sequenciado, e as moléculas não incorporadas são eliminadas. Após cada ciclo de síntese, um laser é utilizado para excitar os corantes, e uma varredura de alta resolução da base incorporada é realizada. Uma etapa de desbloqueio químico garante a remoção do grupo bloqueador do terminal 3' e do corante em uma única etapa. O processo é repetido até que toda a molécula de DNA seja sequenciada [15].

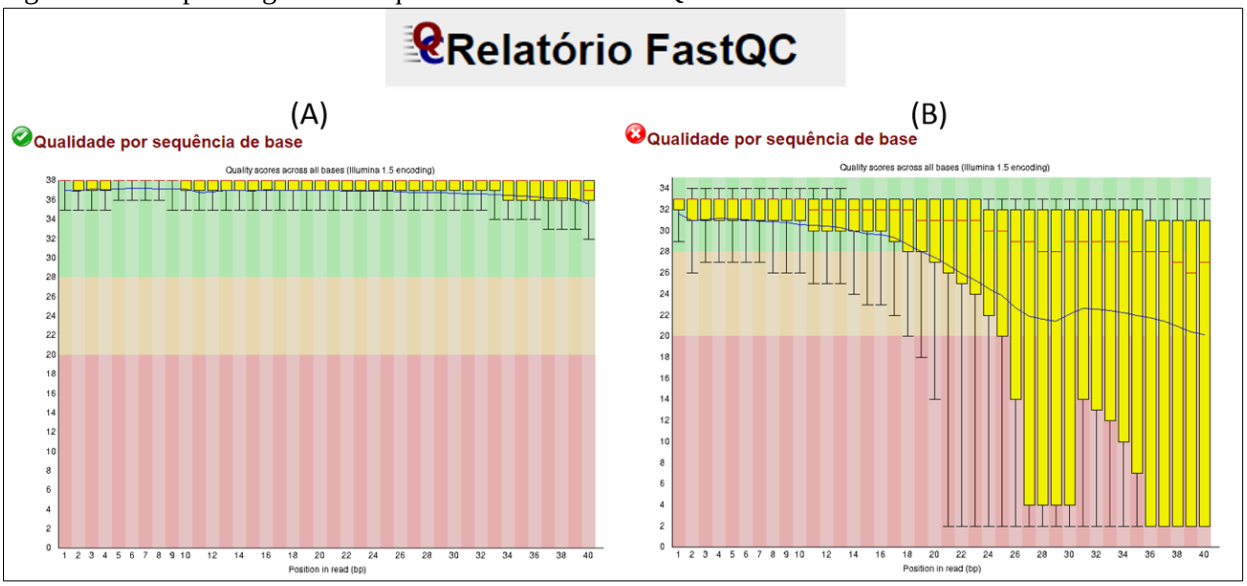
3 CONTROLE DE QUALIDADE.

3.1 FastQC

O *FastQC* [16] é uma ferramenta de bioinformática amplamente utilizada para realizar a análise de qualidade de dados de sequenciamento de DNA. Desenvolvido em Java, o software oferece uma avaliação detalhada e rápida da qualidade dos dados brutos gerados por sequenciadores de DNA, como *Illumina*, Ion Torrent e outros[17,18]. Ao examinar a qualidade das bases em todas as posições ao longo das sequências, o FastQC fornece uma visão geral da distribuição de qualidade utilizando o escore de qualidade Phred para indicar a probabilidade de erro em cada base. Além disso, a ferramenta analisa a composição de nucleotídeos ao longo de toda a sequência, verificando a distribuição de tamanhos das sequencias e identificando possíveis problemas, como a presença de adaptadores ou contaminantes[19]. A capacidade de identificar duplicatas ou sequencias repetitivas também está incorporada, contribuindo para a avaliação global da qualidade dos dados. Uma característica notável do FastQC é a sua capacidade de gerar gráficos (Figura 2) e relatórios interativos, proporcionando uma visualização fácil dos resultados da análise. Sendo altamente configurável, permite que os usuários ajustem os parâmetros conforme necessário. Além disso, a sua integração em pipelines de análise de dados de sequenciamento facilita a automação e a incorporação em fluxos de trabalho mais amplos[19,20].

Comumente utilizado como uma etapa inicial no pré-processamento de dados, o FastQC é essencial para bioinformatas que buscam garantir a qualidade dos dados antes de procederem a análises mais avançadas, como alinhamento de sequências e identificação de variantes20. Esta ferramenta está disponível gratuitamente em: https://github.com/s-andrews/FastQC[16].

Figura 2: Exemplo de gráfico de qualidade de bases FastQC.



Fonte: Babraham Bioinformatics (2024)[17]

(A) Gráfico demonstrando boa qualidade da amostra; (B) Gráfico demonstrando baixa qualidade da amostra.

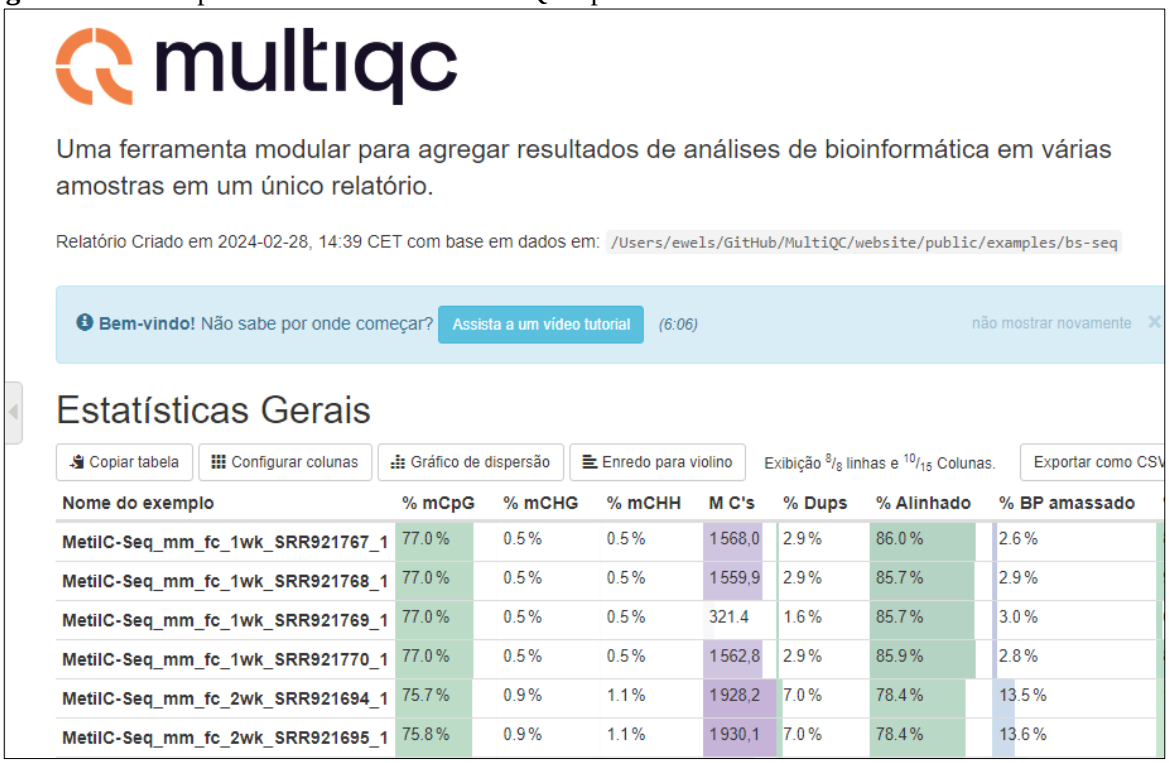
3.2 MultQC

A maioria dos programas de bioinformática utilizados para a verificação de qualidade de bibliotecas de metagenômica, como a ferramenta FastQC[16], produz logs detalhados. Porém, esses logs são produzidos por amostra, gerando assim uma grande quantidade de dados e exigindo que o usuário encontre e compile os resultados de CQ. Esse processo é demorado, repetitivo e complexo, tornando-se propenso a erros[16,19]. O MultiQC resolve esse problema examinando determinados diretórios de análise em busca de arquivos de log e relatórios de CQ, criando um único relatório de resumo que visualiza os resultados em todas as amostras. A coleta de dados em um único relatório fornece uma maneira rápida e fácil de verificar estatísticas importantes (Figura 3). Gráficos compartilhados permitem a comparação precisa entre as amostras, permitindo a detecção de diferenças sutis, não perceptíveis ao alternar entre arquivos diferentes. A visualização dos dados auxilia na detecção do efeito de lote, minimizando o risco de fatores de confusão que afetam os resultados do estudo. O MultiQC tem o potencial de melhorar muito o controle de qualidade e os relatórios para pesquisadores envolvidos no sequenciamento de próxima geração, eliminando a necessidade de scripts comparativos personalizados [19].

O MultiQC é escrito em Python e executado na linha de comando. Os diretórios

especificados são pesquisados recursivamente por quaisquer arquivos reconhecidos. Os submódulos para cada ferramenta suportada são executados, consultando os arquivos de entrada usando configurações de pesquisa configuráveis. Se algum arquivo de log for encontrado, ele é analisado; caso contrário, o módulo é encerrado silenciosamente [19]. O MultiQC gera um único relatório HTML que pode ser compartilhado e aberto em qualquer navegador moderno. Os relatórios renderizam gráficos usando a biblioteca de plotagem JavaScript Highcharts (http://www.highcharts.com)[21]. Os gráficos são redimensionáveis e interativos, alguns com zoom de clique e arraste. Os exemplos podem ser renomeados, ocultos e realçados usando uma caixa de ferramentas de relatório. Os gráficos podem ser exportados em uma variedade de formatos[19,21].

Figura 3: Parte superior de um relatório MultiQC típico.



Nome do exemplo	% mCpG	% mCHG	% mCHH	M C's	% Dups	% Alinhado	% BP amassado
MetilC-Seq_mm_fc_1wk_SRR921767_1	77.0 %	0.5 %	0.5 %	1 568,0	2.9 %	86.0 %	2.6 %
MetilC-Seq_mm_fc_1wk_SRR921768_1	77.0 %	0.5 %	0.5 %	1 559,9	2.9 %	85.7 %	2.9 %
MetilC-Seq_mm_fc_1wk_SRR921769_1	77.0 %	0.5 %	0.5 %	321.4	1.6 %	85.7 %	3.0 %
MetilC-Seq_mm_fc_1wk_SRR921770_1	77.0 %	0.5 %	0.5 %	1 562,8	2.9 %	85.9 %	2.8 %
MetilC-Seq_mm_fc_2wk_SRR921694_1	75.7 %	0.9 %	1.1 %	1 928,2	7.0 %	78.4 %	13.5 %
MetilC-Seq_mm_fc_2wk_SRR921695_1	75.8 %	0.9 %	1.1 %	1 930,1	7.0 %	78.4 %	13.6 %

Fonte: Ewels et al.[19]

A tabela de estatísticas gerais pode ser vista com métricas de várias ferramentas diferentes reunidas para cada amostra.

3.3 Fastp

O *Fastp* [22] é uma ferramenta de pré-processamento de dados de sequenciamento de DNA projetada para aprimorar a qualidade e preparar os dados brutos para análises

subsequentes [22]. Desenvolvido em C++, o *Fastp* oferece eficiência computacional e facilidade de uso. Uma de suas principais funcionalidades é a remoção eficaz de adaptadores contaminantes frequentemente presentes em bibliotecas preparadas para sequenciamento de próxima geração. Além disso, realiza a filtragem de bases de baixa qualidade, contribuindo para a redução de erros em análises posteriores. é uma ferramenta de pré-processamento de dados de sequenciamento de DNA projetada para aprimorar a qualidade e preparar os dados brutos para análises subsequentes. Desenvolvido em C++, o Fastp oferece eficiência computacional e facilidade de uso. Uma de suas principais funcionalidades é a remoção eficaz de adaptadores contaminantes, frequentemente presentes em bibliotecas preparadas para sequenciamento de próxima geração. Além disso, realiza a filtragem de bases de baixa qualidade, contribuindo para a redução de erros em análises posteriores. O Fastp é capaz de remover sequências muito curtas, fornecendo uma avaliação estatística abrangente da qualidade das bases e identificando parâmetros relevantes[22]. Ao empregar algoritmos avançados, o Fastp busca aprimorar a qualidade global dos dados, incluindo a suavização de regiões de baixa qualidade. Sua capacidade de execução paralela contribui para uma eficiência adicional, especialmente em sistemas com vários núcleos. A ferramenta produz resultados em formatos comuns, como FASTQ, e pode gerar relatórios visuais para facilitar a interpretação dos resultados. Sua configuração flexível permite que os usuários personalizem os parâmetros de acordo com as necessidades específicas, adaptando-se a diferentes tipos de dados de sequenciamento. O Fastp é amplamente empregado como uma etapa inicial em pipelines de análise de dados de sequenciamento, desempenhando um papel crucial na preparação dos dados brutos para análises subsequentes, como alinhamento de sequências e identificação de variantes genéticas. Para informações detalhadas e atualizadas, é recomendável consultar a documentação oficial do Fastp (https://github.com/OpenGene/Fastp)[23].

4 REMOÇÃO DE HOSPEDEIRO E CONTAMINANTES

Todas as amostras de sequenciamento metagenômico podem conter não somente o DNA de microrganismos, mas também o material genético indesejado do hospedeiro, caso a amostra tenha sido coletada de um animal ou ser humano, e o material genético contaminante, que na maioria das vezes é do ser humano que coletou, manipulou ou até mesmo estava próximo da amostra. Para evitar erros, utilizamos ferramentas que fazem a identificação desses genomas e os removem, deixando apenas o metagenoma.

4.1 Kraken 2

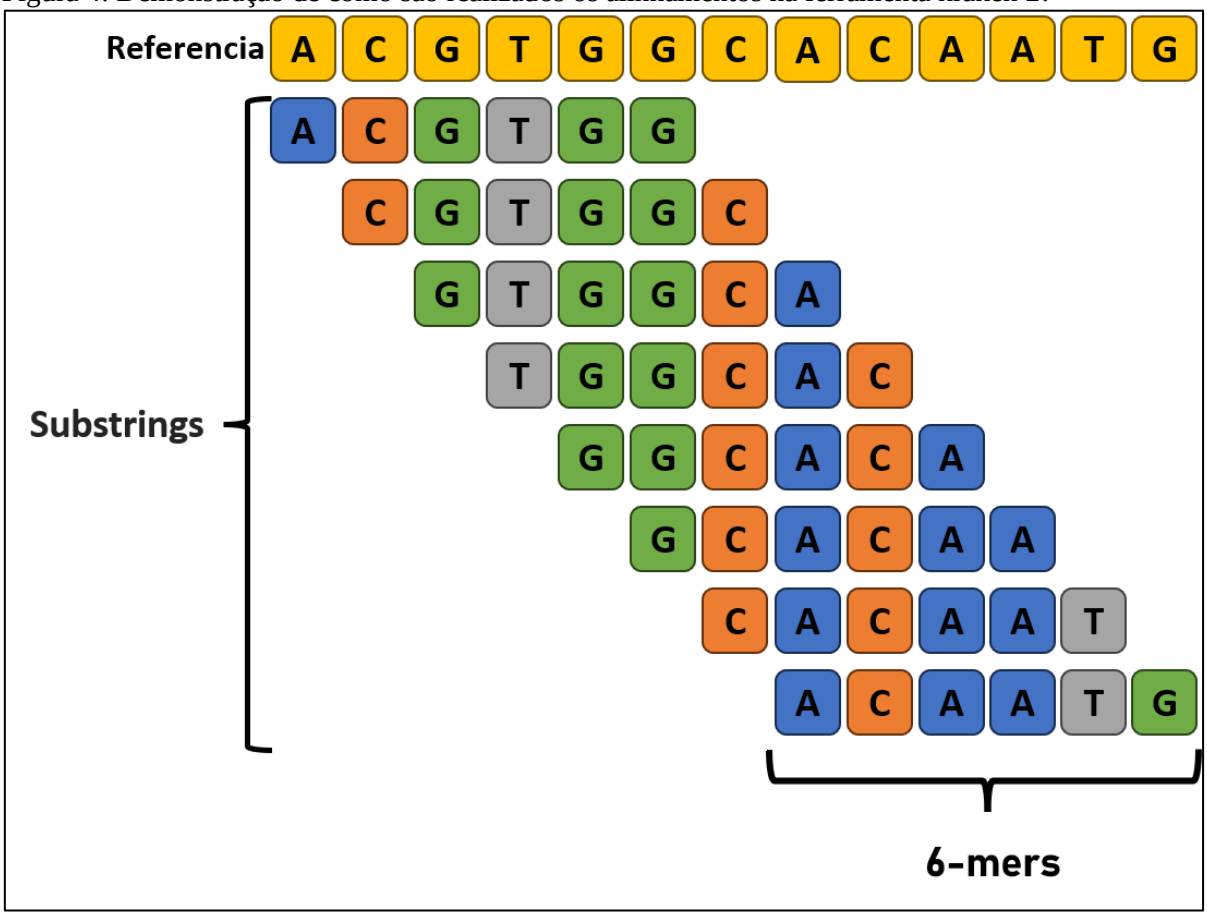
O *Kraken 2* [24] é uma ferramenta rotineiramente utilizada para classificação taxonômica. Com as recentes atualizações, passou a ser uma opção para a remoção de hospedeiros e contaminantes. É uma ferramenta rápida e eficiente, sendo a mais nova versão do Kraken, um sistema de classificação taxonômica que usa partidas K-Mer (Fig. 4) exatas para alcançar alta precisão e altas velocidades de classificação. K-mers são pequenos trechos de sequências genômicas, geralmente de tamanho fixo (por exemplo, 31 nucleotídeos). O Kraken 2 divide as sequências da base de dados e as sequências metagenômicas em k-mers. Esse classificador corresponde cada k-mer, dentro de uma sequência de consulta, ao mais baixo ancestral comum (lowest common ancestor, LCA) de todos os genomas que contêm o k-mer dado. As atribuições k-mer informam o algoritmo de classificação[24–26].

Ele utiliza uma extensa base de dados que contém sequências genômicas de uma variedade de organismos. Essa base de dados é essencial para atribuir taxonomias às sequências metagenômicas. A ferramenta possui bancos de dados já indexados; porém, há a possibilidade de se criarem bancos de dados personalizados para cada tipo de estudo. A ferramenta cria uma tabela de hash que mapeia cada k-mer para o identificador do organismo correspondente na base de dados. Isso permite uma busca eficiente durante a fase de classificação. Durante a fase de classificação, as sequências metagenômicas são divididas em k-mers, e o Kraken 2 utiliza a tabela de hash para identificar a qual organismo cada k-mer pertence. Com base na maioria dos k-mers associados a um organismo específico, o Kraken 2 classifica a sequência metagenômica como pertencente a esse organismo. Uma vez que as sequências são classificadas, o Kraken 2 atribui taxonomias aos organismos identificados, o que significa que o resultado fornece informações sobre a composição taxonômica da amostra metagenômica, indicando quais microrganismos estão presentes e em quais proporções. Após a taxonomia, são gerados relatórios que incluem informações sobre a abundância relativa de cada organismo identificado, permitindo uma análise detalhada da composição da comunidade microbiana na amostra[24–27].

Para a utilização da ferramenta para a remoção de hospedeiro e contaminantes da amostra, basta executá-la contra os genomas do hospedeiro e contaminantes, e todo o material não classificado é considerado parte do microbioma. A ferramenta está disponível gratuitamente em: https://github.com/DerrickWood/kraken2 [28]. Uma das principais vantagens do Kraken 2 é a sua velocidade. Ele é capaz de classificar milhões de sequências em poucos minutos, tornando-o adequado para análises metagenômicas em larga escala[25]. Além disso, o

Kraken 2 oferece alta precisão de classificação, especialmente em níveis taxonômicos mais altos[25].

Figura 4: Demonstração de como são realizados os alinhamentos na ferramenta kraken 2.



Fonte: Takao et al. (2024)

Aqui utilizando como exemplo o tamanho de "K" como sendo 6.

4.2 Bowtie 2

O Bowtie 2 é uma ferramenta de alinhamento de sequências de DNA/RNA em genomas de referência, utilizada principalmente em análises genômicas e metagenômicas. Desenvolvido como uma evolução do Bowtie original, oferece melhorias significativas em termos de velocidade, sensibilidade e flexibilidade[29,30]. Essa ferramenta foi projetada para realizar alinhamentos rápidos e eficientes de sequências de leituras de DNA/RNA em genomas de referência, utilizando uma técnica de indexação sofisticada para facilitar a busca e o mapeamento eficiente das leituras. Sendo capaz de operar em modos sensíveis, prioriza a sensibilidade no alinhamento, o que é crucial para detectar regiões genômicas altamente semelhantes ou para análises que exigem correspondências precisas. Ao contrário de seu

antecessor, oferece suporte a leituras mais longas, tornando-se adequado para as demandas de sequenciamento de próxima geração (NGS), que frequentemente geram leituras de tamanho significativo. É capaz de lidar com leituras pareadas, comuns em tecnologias de sequenciamento modernas. Isso é essencial para análises metagenômicas, em que a identificação e a classificação de leituras pareadas podem fornecer informações mais precisas sobre a comunidade microbiana. Oferece modos de alinhamento local e global, o que permite adaptar a análise conforme necessário[29,30].

O Bowtie 2 gera resultados detalhados sobre os alinhamentos, permitindo uma análise aprofundada da qualidade e da distribuição das leituras no genoma de referência. É comumente usados em conjunto com outras ferramentas e pipelines de análise, sendo uma peça essencial em fluxos de trabalho mais amplos de bioinformática. No caso da metagenômica, é uma ferramenta essencial para a remoção de hospedeiro e contaminantes da amostra. Está disponível gratuitamente em: https://github.com/BenLangmead/bowtie2 [31].

5 MONTAGEM DE NOVO

A montagem *de novo* consiste em montar todos os genomas presentes na amostra de metagenoma. No entanto, essa montagem nem sempre é utilizada, já que ferramentas como o Kraken 2 conseguem identificar e alinhar os fragmentos do genoma sem a necessidade de uma montagem prévia.

5.1 MEGAHIT

O MEGAHIT é uma ferramenta de montagem metagenômica de novo desenvolvida por Li et al. (2015) para lidar com conjuntos de dados metagenômicos grandes e complexos de forma eficiente em termos de tempo e uso de memória. Ele emprega uma abordagem de grafo de Bruijn multi-k-mer para a montagem, construindo e mesclando gradualmente grafos de Bruijn com diferentes tamanhos de k-mer[32]. Essa estratégia permite capturar informações de diferentes níveis de especificidade, o que resulta em montagens de alta qualidade com contigs longos e precisos. Uma das principais vantagens do MEGAHIT é a sua eficiência computacional, pois ele é projetado para ser executado em computadores com recursos limitados de memória, tornando-se acessível a uma ampla gama de usuários[32]. Além disso, o MEGAHIT oferece vários parâmetros ajustáveis que podem ser otimizados com base nas

características do conjunto de dados, o que permite uma montagem personalizada e aprimorada[32]. Ele tem sido amplamente aplicado em estudos metagenômicos para a montagem de novo de genomas microbianos a partir de amostras ambientais diversas, como solo, água e intestino humano[33–35]. Esses estudos demonstram a eficácia do MEGAHIT na descoberta de novos genes, vias metabólicas e espécies microbianas, contribuindo para a compreensão da diversidade e da função dos ecossistemas microbianos. Comparações com outras ferramentas de montagem metagenômica, como o MetaSPAdes[36] e o IDBA-UD[37], mostram que o MEGAHIT apresenta um desempenho competitivo em termos de qualidade da montagem e eficiência computacional[32,38]. Sua capacidade de lidar com conjuntos de dados metagenômicos desafiadores e produzir montagens de alta qualidade o torna uma escolha popular entre os pesquisadores.

5.2 MetaSPAdes

O MetaSPAdes é uma ferramenta de montagem *de novo* de metagenomas desenvolvida pela equipe responsável pelo SPAdes, um montador de genoma único amplamente utilizado [36]. Ele emprega uma abordagem de grafo de Bruijn multi-k-mer para a montagem, construindo grafos de Bruijn com diferentes tamanhos de k-mer e mesclando-os em um grafo de montagem por consenso[36]. Uma das principais características do MetaSPAdes é a sua capacidade de utilizar informações de pareamento de *reads* para melhorar a qualidade da montagem. Ele suporta bibliotecas pareadas e mate-pairs, permitindo a resolução de repetições e a obtenção de contigs mais longos e precisos[36]. Além disso, o MetaSPAdes inclui um módulo de correção de erros chamado BayesHammer, que corrige erros de sequenciamento nos *reads* antes da montagem, contribuindo para a precisão da montagem fina[36]. O MetaSPAdes emprega uma abordagem iterativa para a montagem, em que os *contigs* montados em uma iteração são usados como referência para a próxima[36]. Esse processo iterativo ajuda a estender e melhorar os contigs, resultando em uma montagem mais completa e contígua. Comparações com outras ferramentas de montagem metagenômica destacam o desempenho do MetaSPAdes em termos de qualidade da montagem. Forouzan et al. (2018) avaliaram nove montadores populares e constataram que o MetaSPAdes produziu montagens de alta qualidade, especialmente quando informações de pareamento de reads estavam disponíveis[38]. Nurk et al. (2017) também demonstraram a capacidade do MetaSPAdes de recuperar genomas microbianos a partir de conjuntos de dados metagenômicos complexos[36]. No entanto, é importante observar que o

MetaSPAdes pode ter requisitos de memória mais altos em comparação com outras ferramentas, como o MEGAHIT[32]. Isso pode ser uma consideração importante dependendo dos recursos computacionais disponíveis.

6 ANÁLISE TAXONOMICA

A análise taxonômica é uma etapa fundamental nos estudos metagenômicos, permitindo a identificação e classificação dos organismos presentes em amostras ambientais complexas. Essa etapa é crucial para compreender a diversidade microbiana, a estrutura das comunidades e a função dos ecossistemas. Neste capítulo, discutiremos a importância da análise taxonômica em metagenômica e apresentaremos duas ferramentas amplamente utilizadas: Kraken 2 e QIIME 2[25,39]. A análise taxonômica em metagenômica pode ser realizada utilizando diferentes abordagens, como a baseada em similaridade e a baseada em composição[40]. A abordagem baseada em similaridade compara as sequências metagenômicas com bancos de dados de referência, enquanto a abordagem baseada em composição utiliza características intrínsecas das sequências, como frequência de k-mers[40]. Os bancos de dados de referência, como o NCBI RefSeq, o SILVA e o Greengenes, desempenham um papel fundamental na análise taxonômica, fornecendo sequências anotadas de genes ou genomas de organismos conhecidos[40]. Várias ferramentas foram desenvolvidas para a classificação taxonômica de sequências metagenômicas, cada uma com suas próprias características e abordagens. Neste capítulo, destacaremos duas ferramentas populares: Kraken 2 e QIIME 2[25,39].

6.1 Kraken 2

Já descrito na subseção 4.1.

O Kraken 2 tem sido amplamente aplicado em estudos metagenômicos para a montagem de novo de genomas microbianos a partir de amostras ambientais diversas, como solo, água e intestino humano[33–35]. A sua velocidade e a sua precisão o tornam uma escolha popular entre os pesquisadores.

6.2 QIIME 2

O QIIME 2 (*Quantitative Insights Into Microbial Ecology* 2) é uma plataforma de análise metagenômica desenvolvida por Bolyen et al. (2019)[39], que oferece um conjunto abrangente de ferramentas para o processamento, a análise e a visualização de dados metagenômicos, incluindo a classificação taxonômica. O QIIME 2 suporta várias abordagens para a classificação taxonômica, como a baseada em similaridade, usando o classificador Naive Bayes[41], e a baseada em composição, usando o classificador q2-feature-classifier[41]. Ele também permite a utilização de diferentes bancos de dados de referência, como o Greengenes e o SILVA[39]. Uma das principais características do QIIME 2 é a sua flexibilidade e a sua modularidade. Ele oferece uma ampla gama de plugins e ferramentas que podem ser combinados em pipelines de análise personalizados[39]. Além disso, o QIIME 2 possui uma interface gráfica do usuário (q2studio) que facilita a configuração e execução das análises[39]. O QIIME 2 tem sido amplamente adotado pela comunidade científica para a análise metagenômica de diversos ambientes, como o microbioma humano[42], solo[43] e água. Sua abrangência e flexibilidade o tornam uma ferramenta poderosa para a análise taxonômica em metagenômica.

O Kraken 2 destaca-se por sua velocidade e precisão, sendo adequado para análises em larga escala. O QIIME 2, por sua vez, oferece uma plataforma abrangente e modular para a análise metagenômica, incluindo a classificação taxonômica[25,39].

7 ANÁLISES ESTATISTICAS

Após a anotação taxonômica de todos os genomas contidos na amostra, algumas ferramentas são utilizadas para a realização de correções e análises estatísticas dos dados gerados pelas ferramentas de taxonomia.

7.1 Bracken

O Bracken (Bayesian Reestimation of Abundance with Kraken)[39] é uma ferramenta bioinformática projetada para melhorar a precisão na estimativa da abundância taxonômica em análises metagenômicas[25,44]. Desenvolvido como uma extensão do Kraken, o Bracken atua especificamente na correção de estimativas de abundância, levando em consideração a distribuição da profundidade de cobertura. Foi desenvolvido para abordar um desafio comum em análises metagenômicas, que é a estimação imprecisa da abundância taxonômica devido a

variações na profundidade de cobertura das sequências, ele utiliza um modelo probabilístico para ponderar as estimativas de abundância. Ao considerar a distribuição da profundidade de cobertura, o Bracken calcula as probabilidades de cada organismo ser responsável por uma determinada leitura[25,44]. Ele se baseia suas estimativas em k-mers, assim como o Kraken 2[25,26,44]. No entanto, ao contrário do Kraken 2, que gera estimativas brutas, o Bracken realiza uma correção considerando a confiabilidade das estimativas para cada organismo ao levar em conta a distribuição da profundidade de cobertura, melhora a precisão na identificação e quantificação de organismos presentes em amostras metagenômicas, especialmente em regiões do genoma onde a cobertura pode variar. Ele oferece flexibilidade na configuração de parâmetros, permitindo aos usuários ajustarem o comportamento de acordo com as características específicas de suas amostras e experimentos[44]. Ao proporcionar uma abordagem mais refinada para a estimação da abundância taxonômica, o *Bracken* contribui para uma interpretação mais precisa e confiável dos resultados em análises metagenômicas, sendo uma ferramenta valiosa para pesquisadores envolvidos em estudos de microbioma e genômica ambiental. A ferramenta está disponível gratuitamente em https://github.com/jenniferlu717/Bracken[45].

7.2 Kraken Tools

Kraken Tools são um conjunto de utilitários associados à ferramenta de análise metagenômica *Kraken* e sua versão aprimorada, o *Kraken 2*. Essas ferramentas adicionais foram desenvolvidas para fornecer funcionalidades complementares e facilitar a manipulação, interpretação e análise dos dados gerados pelo *Kraken* e *Kraken 2*[25,46]. Elas oferecem uma gama de funcionalidades para otimizar o uso dessas ferramentas em diferentes contextos e cenários de análise metagenômica, contribuindo para uma compreensão mais aprofundada da diversidade taxonômica em amostras microbiológicas, a ferramenta está disponível gratuitamente em https://github.com/DerrickWood/kraken2[28].

7.3 Pavian

Pavian é uma aplicação web para explorar resultados de classificação de experimentos metagenômicos[47]. Com o *Pavian*, os pesquisadores podem analisar, visualizar e transformar resultados de vários classificadores como *Kraken*, *Centrifuge* e *MethaPhlAn* usando tabelas de

dados interativas, mapas de calor e diagramas de fluxo Sankey. Um visualizador interativo de cobertura de alinhamento pode ajudar na validação de correspondências para um genoma específico, o que pode ser crucial ao usar experimentos de metagenômica para detecção de patógenos. *Pavian* implementa uma interface direta para analisar e comparar conjuntos de dados metagenômicos complexos. É escrito em R (https://www.R-project.org)[48] como um aplicativo modular Shiny. Ele incorpora várias tabelas interativas de Javascript e gráficos D3 (Bostock et al., 2011), como o diagrama de Sankey que é baseado no código sankey de D3 e a biblioteca networkD3. O visualizador de alinhamento usa a biblioteca Rsam*Tools* para carregar o arquivo bam. O back-end pode ser iniciado a partir de um ambiente R ou instalado em um servidor, e a interface é acessada por meio de um navegador da Web. Suas funções ajudam os pesquisadores do microbioma, bem como os microbiologistas clínicos a obter uma melhor compreensão de seus dados por meio de diagramas de fluxo Sankey, tabelas de comparação múltipla e um visualizador de alinhamento do genoma [47,49].

7.4 Microbiome Analyst

O *MicrobiomeAnalyst* é uma plataforma fácil de usar, baseada na web, para análise abrangente de saídas de dados comuns geradas a partir de estudos atuais de microbioma. Ele permite que pesquisadores e clínicos com pouco ou nenhum treinamento em bioinformática explorem uma ampla variedade de métodos bem estabelecidos para processamento de dados de microbioma, análise estatística, perfil funcional e comparação com conjuntos de dados públicos ou assinaturas microbianas conhecidas[50]. Atualmente, o *MicrobiomeAnalyst* contém quatro módulos: Perfil de dados de genes marcadores (Marker-gene Data Profiling, MDP), Perfil de dados de espingarda (Shotgun Data Profiling, SDP), Projeção com Dados Públicos (Projection with Public Data, PPD) e Análise de enriquecimento de conjunto de táxons (Taxon Set Enrichment Analysis, TSEA). Em resumo, o *MicrobiomeAnalyst* oferece um balcão único que permite que os pesquisadores de microbioma explorem completamente seus dados de microbioma pré-processados por meio de interfaces web intuitivas. [50]

8 CONCLUSÃO

A evolução das tecnologias de sequenciamento de nova geração (NGS) e as ferramentas de bioinformática revolucionaram o campo da metagenômica, permitindo uma compreensão

mais profunda e detalhada da diversidade microbiana presente em diferentes ambientes. Este artigo revisou as principais ferramentas e os *workflow*s utilizados na análise de dados de metagenoma, destacando a importância do controle de qualidade, da anotação taxonômica e das análises estatísticas para a obtenção de resultados confiáveis e significativos. As tecnologias NGS oferecem uma capacidade sem precedentes de gerar dados de sequenciamento em larga escala, o que, combinado com ferramentas de bioinformática avançadas, permite uma análise detalhada e abrangente da composição microbiana de amostras complexas. A escolha entre *workflow*s que incluem ou não a montagem de novo dos genomas destaca a flexibilidade e a adaptabilidade necessárias para atender às demandas específicas de cada projeto de metagenômica. Ressalta-se a necessidade contínua de desenvolvimento e avaliação de novas tecnologias e ferramentas para a análise de dados de metagenoma. Apesar dos avanços significativos, desafios permanecem, especialmente em relação à interpretação dos vastos volumes de dados gerados e à necessidade de ferramentas que simplifiquem essa interpretação para pesquisadores não especializados em bioinformática.

Em conclusão, a metagenômica, apoiada por tecnologias NGS e ferramentas de bioinformática robustas, desempenha um papel crucial na expansão do nosso entendimento sobre a diversidade microbiana e sua função em diferentes ecossistemas. À medida que continuamos a explorar e a entender a complexidade da vida microbiana, é essencial que as ferramentas de análise evoluam para fornecer insights mais precisos e profundos, facilitando as descobertas que podem ter implicações significativas na saúde humana, na agricultura, na conservação ambiental e em muitos outros campos.

REFERÊNCIAS:

1. Zhang L, Chen F, Zeng Z, Sun F, Yang L, Bi X, et al. Advances in Metagenomics and Its Application in Environmental Microorganisms. Front Microbiol [Internet]. 17 de dezembro de 2021 [citado 7 de março de 2024];12. Disponível em: https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2021.766364/full

2. Ye SH, Siddle KJ, Park DJ, Sabeti PC. Benchmarking Metagenomics Tools for Taxonomic Classification. Cell. Agosto de 2019;178(4):779–94.

3. Jones S, Baizan-Edge A, MacFarlane S. Viral Diagnostics in Plants Using Next Generation Sequencing: Computational Analysis in Practice. Front Plant Sci [Internet]. 24 de outubro de 2017 [citado 7 de março de 2024];8. Disponível em: https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2017.01770/full

4. Pedersen HK, Gudmundsdottir V, Nielsen HB, Hyotylainen T, Nielsen T, Jensen BAH, et al. Human gut microbes impact host serum metabolome and insulin sensitivity. Nature. julho de 2016;535(7612):376–81.

5. Somasekar S, Lee D, Rule J, Naccache SN, Stone M, Busch MP, et al. Viral Surveillance in Serum Samples From Patients With Acute Liver Failure By Metagenomic Next-Generation Sequencing. Clinical Infectious Diseases. 16 de outubro de 2017;65(9):1477–85.

6. Zhang W, Li L, Deng X, Blümel J, Nübling CM, Hunfeld A, et al. Viral nucleic acids in human plasma pools. Transfusion. 2016;56(9):2248–55.

7. Knights D, Kuczynski J, Charlson ES, Zaneveld J, Mozer MC, Collman RG, et al. Bayesian community-wide culture-independent microbial source tracking. Nat Methods. setembro de 2011;8(9):761–3.

8. Loman NJ, Constantinidou C, Christner M, Rohde H, Chan JZM, Quick J, et al. A Culture-Independent Sequence-Based Metagenomics Approach to the Investigation of an Outbreak of Shiga-Toxigenic Escherichia coli O104:H4. JAMA. 10 de abril de 2013;309(14):1502–10.

9. Wajid B, Anwar F, Wajid I, Nisar H, Meraj S, Zafar A, et al. Correction to: Music of metagenomics—a review of its applications, analysis pipeline, and associated tools. Funct Integr Genomics. 1º de fevereiro de 2022;22(1):137–137.

10. Mavromatis K, Ivanova N, Barry K, Shapiro H, Goltsman E, McHardy AC, et al. Use of simulated data sets to evaluate the fidelity of metagenomic processing methods. Nat Methods. junho de 2007;4(6):495–500.

11. Sczyrba A, Hofmann P, Belmann P, Koslicki D, Janssen S, Dröge J, et al. Critical Assessment of Metagenome Interpretation—a benchmark of metagenomics software. Nat Methods. novembro de 2017;14(11):1063–71.

12. Oulas A, Pavloudi C, Polymenakou P, Pavlopoulos GA, Papanikolaou N, Kotoulas G, et al. Metagenomics: Tools and Insights for Analyzing Next-Generation Sequencing Data Derived from Biodiversity Studies. Bioinform Biol Insights. 1º de janeiro de 2015;9:BBI.S12462.

13. Vincent AT, Derome N, Boyle B, Culley AI, Charette SJ. Next-generation sequencing (NGS) in the microbiological world: How to make the most of your money. Journal of Microbiological Methods. 1º de julho de 2017;138:60–71.

14. Mardis ER. Next-generation DNA sequencing methods. Annu Rev Genomics Hum Genet. 2008;9:387–402.

15. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 6 de novembro de 2008;456(7218):53–9.

16. Andrews S. s-andrews/FastQC [Internet]. 2024 [citado 1º de março de 2024]. Disponível em: https://github.com/s-andrews/FastQC

17. Andrews S. A quality control tool for high throughput sequence data. [Internet]. Babraham Bioinformatics. [citado 11 de março de 2023]. Disponível em: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

18. Leggett RM, Ramirez-Gonzalez RH, Clavijo BJ, Waite D, Davey RP. Sequencing quality assessment tools to enable data-driven informatics for high throughput genomics. Front Genet. 17 de dezembro de 2013;4:288.

19. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 1º de outubro de 2016;32(19):3047–8.

20. Lo CC, Chain PSG. Rapid evaluation and quality control of next generation sequencing data with FaQCs. BMC Bioinformatics. 19 de novembro de 2014;15(1):366.

21. Interactive charting library | Highcharts [Internet]. Highcharts Blog | Highcharts. [citado 1º de março de 2024]. Disponível em: https://www.highcharts.com/blog/homepage21may/

22. Chen S, Zhou Y, Chen Y, Gu J. fastp : an ultra-fast all-in-one FASTQ preprocessor [Internet]. Bioinformatics; 2018 mar [citado 11 de março de 2023]. Disponível em: http://biorxiv.org/lookup/doi/10.1101/274100

23. OpenGene/fastp [Internet]. OpenGene - Open Source Genomics Toolbox; 2024 [citado 7 de março de 2024]. Disponível em: https://github.com/OpenGene/fastp

24. lu jen. Manual Kraken 2 [Internet]. github. 2020 [citado 6 de novembro de 2023]. Disponível em: https://github.com/DerrickWood/kraken2/wiki/Manual

25. Lu J, Rincon N, Wood DE, Breitwieser FP, Pockrandt C, Langmead B, et al. Metagenome analysis using the Kraken software suite. Nat Protoc. dezembro de 2022;17(12):2815–39.

26. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. 3 de março de 2014;15(3):R46.

27. Breitwieser FP, Baker DN, Salzberg SL. KrakenUniq: confident and fast metagenomics classification using unique k-mer counts. Genome Biol. 16 de novembro de 2018;19(1):198.

28. Wood D. DerrickWood/kraken2 [Internet]. 2024 [citado 7 de março de 2024]. Disponível em: https://github.com/DerrickWood/kraken2

29. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. abril de 2012;9(4):357–9.

30. Langmead B, Kin D, Charles R, Chen NC, Wilks C, Antonescu Va. Bowtie 2 Fast and sensitive read alignment [Internet]. Bowtie 2. 2023 [citado 5 de novembro de 2023]. Disponível em: https://bowtie-bio.sourceforge.net/bowtie2/manual.shtml#the-bowtie2-build-indexer

31. Langmead B. BenLangmead/bowtie2 [Internet]. 2024 [citado 7 de março de 2024]. Disponível em: https://github.com/BenLangmead/bowtie2

32. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 15 de maio de 2015;31(10):1674–6.

33. Howe AC, Jansson JK, Malfatti SA, Tringe SG, Tiedje JM, Brown CT. Tackling soil diversity with the assembly of large, complex metagenomes. Proc Natl Acad Sci U S A. 1º de abril de 2014;111(13):4904–9.

34. Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G, et al. Ocean plankton. Structure and function of the global ocean microbiome. Science. 22 de maio de 2015;348(6237):1261359.

35. MetaHIT Consortium, Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, et al. A human gut microbial gene catalogue established by metagenomic sequencing. Nature. março de 2010;464(7285):59–65.

36. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. Genome Res. maio de 2017;27(5):824–34.

37. Peng Y, Leung HCM, Yiu SM, Chin FYL. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics. 1º de junho de 2012;28(11):1420–8.

38. Forouzan E, Shariati P, Mousavi Maleki MS, Karkhane AA, Yakhchali B. Practical evaluation of 11 de novo assemblers in metagenome assembly. J Microbiol Methods. agosto de 2018;151:99–105.

39. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. agosto de 2019;37(8):852–7.

40. Breitwieser FP, Lu J, Salzberg SL. A review of methods and databases for metagenomic classification and assembly. Brief Bioinform. 19 de julho de 2019;20(4):1125–36.

41. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, et al. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. Microbiome. 17 de maio de 2018;6(1):90.

42. Estaki M, Jiang L, Bokulich NA, McDonald D, González A, Kosciolek T, et al. QIIME 2 Enables Comprehensive End-to-End Analysis of Diverse Microbiome Data and Comparative Studies with Publicly Available Data. Curr Protoc Bioinformatics. junho de 2020;70(1):e100.

43. Naylor D, Sadler N, Bhattacharjee A, Graham E, Anderton C, McClure R, et al. Soil Microbiomes Under Climate Change and Implications for Carbon Cycling. Annual Review of Environment and Resources. 17 de outubro de 2020;45.

44. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: estimating species abundance in metagenomics data. PeerJ Computer Science. 2 de janeiro de 2017;3:e104.

45. Lu J. jenniferlu717/Bracken [Internet]. 2024 [citado 7 de março de 2024]. Disponível em: https://github.com/jenniferlu717/Bracken

46. KrakenTools [Internet]. [citado 7 de março de 2024]. Disponível em: https://ccb.jhu.edu/software/krakentools/

47. Breitwieser FP, Salzberg SL. Pavian: interactive analysis of metagenomics data for microbiome studies and pathogen identification. Bioinformatics. 15 de fevereiro de 2020;36(4):1303–4.

48. R: The R Project for Statistical Computing [Internet]. [citado 1º de março de 2024]. Disponível em: https://www.r-project.org/

49. Breitwieser F. fbreitwieser/pavian [Internet]. 2024 [citado 7 de março de 2024]. Disponível em: https://github.com/fbreitwieser/pavian

50. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data | Nature Protocols [Internet]. [citado 7 de março de 2024]. Disponível em: https://www.nature.com/articles/s41596-019-0264-1

ARTIGO 3: Desenvolvimento de bancos de dados para análise metagenômica de microbiomas de cães: uma abordagem com kraken2 e bowtie2

DESENVOLVIMENTO DE BANCOS DE DADOS PARA ANÁLISE METAGENÔMICA DE MICROBIOMAS DE CÃES: UMA ABORDAGEM COM KRAKEN2 E BOWTIE2

Paulo sallarola takao; Julio franz moura; Pamela souza corrêa; David aciole barbosa; Fabiano bezerra menegidio

1.Universidade de Mogi das Cruzes, UMC, Mogi das Cruzes, SP, Brasil

Artigo ganhador do IV CONBIOTEC

Takao PS, Moura JF, Correa PS, Barbosa DA, Menegidio FB. DESENVOLVIMENTO DE BANCOS DE DADOS PARA ANÁLISE METAGENÔMICA DE MICROBIOMAS DE CÃES: UMA ABORDAGEM COM KRAKEN2 E BOWTIE2. Revista Multidisciplinar de Educação e Meio Ambiente; 18 de março de 2024. DOI: 10.51189/conbiotec2024/33826

RESUMO

A análise metagenômica desempenha um papel crucial na compreensão das comunidades microbianas e seu impacto ambiental, desfrutando de uma relevância especial na medicina veterinária para diagnóstico, tratamento e prevenção de doenças em animais de estimação. Este estudo teve como objetivo principal criar bancos de dados integrados para as ferramentas Bowtie2 e Kraken2, consolidando os genomas do cão e do humano em um recurso atualizado e de fácil acesso. Com o avanço tecnológico, a análise metagenômica do microbioma emerge como uma ferramenta promissora na rotina veterinária, especialmente considerando a alta prevalência de cães nas clínicas veterinárias, sabendo-se que apenas no Brasil temos em torno de 58,1 milhos de cães, sendo o segundo animal de estimação mais presente nos lares, perdendo apenas para as aves. No entanto, as etapas bioinformáticas necessárias, como a remoção do genoma do hospedeiro e de contaminantes humanos, demandam tempo e recursos computacionais consideráveis. Para superar esse desafio, foram desenvolvidos bancos de dados específicos para cada ferramenta, reduzindo significativamente o tempo de análise e assegurando a atualização contínua dos genomas. Enquanto o Bowtie2 realiza o alinhamento exato de sequências, o Kraken2 utiliza sequências menores (k-mers) para uma classificação taxonômica mais rápida e eficiente. Após a construção dos bancos de dados, foram conduzidos testes em uma biblioteca de metagenoma canino, evidenciando uma alta porcentagem de alinhamento e uma remoção eficaz de leituras pertinentes ao cão ou humano. Apesar dos desafios enfrentados devido à falta de documentação detalhada, os bancos de dados criados se mostraram viáveis e reproduzíveis, oferecendo uma contribuição significativa para estudos futuros em análise metagenômica veterinária.

Palavras-chave: *Canis lupus familiaris*, Metagenoma, Kraken2-build, Bowtie2-build, Contaminantes.

1 INTRODUÇÃO

A análise metagenômica é o processo utilizado para classificar o metagenoma (MARCHESI; RAVEL, 2015), sendo o metagenoma uma derivação da genômica que é o estudo dos genes de um organismo, então na tradução literal metagenômica seria "além do genoma", ou seja, na metagenômica não analisamos apenas um genoma e sim todo os genomas contidos em uma amostra (GILBERT; DUPONT, 2011), podendo esses genomas serem de microrganismos e até fragmentos de DNA livres no ambiente (amostra), genes de elementos estruturais microbianos, vírus, fagos, toxinas e tudo mais que possua material genético (BERG et al., 2020; HANDELSMAN et al., 1998; MERRIAM-WEBSTER, 2023; WHIPPS; LEWIS; COOKE, 1988), então podemos descrever a análise metagenômica como a análise de uma coleção de genomas de microrganismos e suas condições ambientais em um determinado ambiente por uma técnica de sequenciamento genético independente da utilização de cultura para crescimento e análise (HANDELSMAN et al., 1998; MARCHESI; RAVEL, 2015; MERRIAM-WEBSTER, 2023; WHIPPS; LEWIS; COOKE, 1988). Portanto com a tecnologia de análise metagenômica podemos descobrir um mundo totalmente novo, já que se sabe que apenas pouco mais de 1% de todos os microrganismos existentes são passiveis de serem cultivados (COWAN, 2000; HANDELSMAN, 2004; MERRIAM-WEBSTER, 2023). Uma das principais aplicações da análise metagenômica é o estudo do microbioma que foi definido pela primeira vês por Whipps et al. (1988), como sendo uma comunidade de microrganismos (WHIPPS; LEWIS; COOKE, 1988), mas essa definição foi se alterando ao longo dos anos onde chegou-se a definição de microbioma como todo o material genético (DNA) da microbiota de um determinado ambiente e sua área de atividade, incluindo elementos estruturais interno e externos dos microrganismo ou seja é quando se analisa não apenas a comunidade de micróbios que colonizam este ambiente, mas também a expressão dos genes desta microbiota e como eles influenciam aquele ambiente (BERG et al., 2020; MERRIAM-WEBSTER, 2023).

Neste contexto, a análise metagenômica do microbioma emerge como uma ferramenta promissora na rotina veterinária, especialmente considerando a alta prevalência de cães (*Canis lúpus familiaris*) nas clínicas veterinárias e de acordo com o relatório de 2023 da AMBIPET (Associação brasileira da indústria de produtos para animais de estimação) o Brasil é o segundo maior pais em população de cães no mundo, ficando atrás apenas do Estados Unidos, tendo aproximadamente de 58,1 milhões e 900 milhões de cães respectivamente ("Mercado Pet Brasil 2023", 2023).

Com o avanço das tecnologias de sequenciamento genético e da diminuição dos custos para a sua execução (BHARTI; GRIMM, 2021), a análise metagenômica do microbioma está cada dia mais perto da utilização na rotina da clínica veterinária de pequenos animais sendo uma promissora ferramenta de diagnóstico, tratamento e prevenção de doenças, sendo que na medicina humana já está bastante difundida principalmente para a análise do microbioma de intestino, boca e pele (THE INTEGRATIVE HMP (IHMP) RESEARCH NETWORK CONSORTIUM, 2019), como o cão é o animal com a maior presença na rotina do médico veterinário de pequenos animais o conhecimento do seu microbioma é de fundamental importância ("Mercado Pet Brasil 2023", 2023; "Número de cães e gatos no Brasil deve chegar a mais de 100 milhões em 10 anos", 2023). Porem para a realização de uma análise metagenômica várias etapas de bioinformática são necessários, sendo uma das principais etapas a retirada do genoma do hospedeiro, sendo nesse caso o cão (*Canis lupus familiaris*) e de genomas contaminantes do ser humano (*Homo sapiens*), assim conseguimos diminuir consideravelmente o volume de dados a serem processados, já que se sabe que em uma amostra de sequenciamento metagenômico a quantidade de material genético do hospedeiro e do contaminante pode ser mais de 99% da amostra.

Para a retirada do genoma do hospedeiro e de contaminantes a ferramenta de bioinformática a ser utilizada possui um banco de dados (BD) previamente indexado específico para cada organismo que muitas vezes é disponibilizado gratuitamente pela ferramenta, no entanto esses BD são disponibilizados de forma individual, no nosso caso um BD para cão e um BD para ser humano fazendo com que a ferramenta tenha que ser utilizada duas vezes tendo uma perda de tempo desnecessária, já que algumas ferramentas podem demorar muitas horas para realizar apena uma análise fora isso muitas vezes esses banco de dados já estão defasados sem atualização do genoma e de acordo com o estudo conduzido por Gihawi et al (2023). O uso de apenas 1 genoma de referência por espécie pode ocorrer erros e uma remoção de hospedeiro e contaminantes falha, pode acarretar e em erros na taxonomia, portanto 2 genomas ou mais por espécie é o mais indicado. Duas das ferramentas de uso corriqueiro para essa função, são o Bowtie2 e Kraken2. (LANGMEAD; SALZBERG, 2012; LU et al., 2022; WOOD; SALZBERG, 2014)

Bowtie2, realiza o alinhamento de sequencias exatas de acordo com o seu BD de referência, esse alinhamento ocorre em 4 etapas onde na etapa 1 ele irá coletar sementes (inicio) de sequencias a serem analisadas, na etapa 2 ele irá alinhara as sementes são extraídas da amostra e alinhadas de forma não lacunar, na etapa 3 as sementes são posicionadas no genoma

de referência e na etapa 4 essas sementes são estendidas para um alinhamento completo (LANGMEAD; SALZBERG, 2012; LU et al., 2022; RUMBAVICIUS; ROUNGE; ROGNES, 2023). As sementes não alinhadas e suas respectivas sequencias são arquivadas separadamente para posterior análise do microbioma (LANGMEAD; SALZBERG, 2012). Devido ser dividido em 4 etapas torna a ferramenta mais lenta, porém com um baixo consumo de memória ram (RUMBAVICIUS; ROUNGE; ROGNES, 2023). Essa ferramenta possui os bancos de dados já indexados do cão e do Homem disponíveis publicamente, porém com genomas defasados e de forma individual, sendo um genoma por banco de dados, acabando que temos que executar duas vezes a ferramenta, primeiro com um banco de dados e depois com o outro (LANGMEAD et al., 2023).

Para criação do banco de dados Bowtie2, é utilizado o bowtie2-build que constrói um índice Bowtie a partir de sequencias de DNA contidos em um arquivo do tipo ".fasta", produzindo um conjunto de 6 arquivos com o sufixo ".bt2" (LANGMEAD; SALZBERG, 2012).

Kraken2 é uma ferramenta de classificação taxonômica muito utilizada na metagenômica que diferente do Bowtie2 que realiza o alinhamento completo com uma sequência de DNA, ela utiliza sequencias menores com tamanho k (k-mers) sendo o tamanho padrão pré-definido de "k" igual a 50 nucleotídeos, que pode ser alterado de acordo com a necessidade, assim deixando a ferramenta muito mais ágil podendo analisar até 1,5 milhões de leituras (reads) por minuto com muita precisão (LANGMEAD et al., 2023; LANGMEAD; SALZBERG, 2012; LU et al., 2022; RUMBAVICIUS; ROUNGE; ROGNES, 2023; WOOD; SALZBERG, 2014). Essa ferramenta até a pouco tempo não era utilizada para a retirada de hospedeiro e contaminantes de amostras de metagenomas, somente após a sua última atualização foi implementado essa funcionalidade e vem demonstrando uma ótima opção para isso, sendo mais rápida e com uma boa sensibilidade após testes realizados por nossa equipe. Diferente da ferramenta anterior, essa não possui bancos de dados com apenas um ou com os dois genomas juntos, por ser uma ferramenta utilizada corriqueiramente apenas para análise taxonômica de metagenomas, ela tem apenas BD com genomas de múltiplos organismos. Para a criação de BD Kraken2 é utilizado o kraken2-build que criará um diretório com 3 arquivos com o sufixo ".k2d", que é criado a partir de arquivos .fasta contendo os genomas de referência.

O Objetivo desse trabalho foi criar bancos de dados específicos para cada ferramenta sendo elas Bowtie2 e Kraken2 que são disponibilizadas gratuitamente por seus criadores, com

o genoma do hospedeiro e contaminante em um único BD com os genomas atualizados para as duas espécies *Canis lupus familiaris* e *Homo sapiens*.

2 MATERIAIS E MÉTODOS

Para a confecção do banco de dados foi utilizado um notebook Acer Nitro 5 com processador Intel core I5 11ª geração com 6 núcleos e 12 threads, 42 Gb de memória ram, placa de vídeo GeForce GTX 1650 e sistema operacional Linux Mint 21.2 Cinnamon, baseado em Ubuntu.

Para que as ferramentas pudessem ser executadas, foi instalado um pacote de softwares de pesquisa biomédicas o Bioconda através da linha de comando e em seguida foram instaladas as ferramentas Bowtie2 e Kraken2.

Iniciamos a confecção dos BD, o primeiro passo foi realizado o download dos arquivos dos genomas de referência atualizados no formato fasta (.fna) do cão e do Homem, disponíveis publicamente no banco de dados do *RefSeq* (Banco de Dados de Sequência de Referência) do NCBI (National Center for Biotechnology Information) (tabela 1) (BETHESDA (MD): NATIONAL LIBRARY OF MEDICINE (US), 1988; "RefSeq: NCBI Reference Sequence Database", 2023).

Tabela 1: Dados dos genomas utilizados para a confecção dos bancos de dados de cão e ser humano, extraídos do banco de dados público do RefSeq NCBI (https://www.ncbi.nlm.nih.gov/refseq/).

Nome científico	**Biblioteca**	**RefSeq**	**Nível**	**Data**
Canis lupus familiais	Dog10k_Boxer_Tasha	GCF_000002285.5	Cromossomo	Outubro 2020
Canis lupus familiais	UMICH_Zoey_3.1	GCF_005444595.1	Cromossomo	Maio 2019
Canis lupus familiais	UU_Cfam_GSD_1.0	GCF_011100685.1	Cromossomo	Março 2020
Canis lupus familiais	ROS_Cfam_1.0	GCF_014441545.1	Cromossomo	Setembro 2020
Homo sapiens	GRCh38.p14	GCF_000001405.40	Cromossomo	Fevereiro 2022
Homo sapiens	T2T-CHM13v2.0	GCF_009914755.1	Completo	Janeiro 2022

Fonte: Takao et al. (2024)

Legenda: Foram utilizados 4 genomas de referência canino e dois genomas de referência humano de acordo com o sugerido por Gihawi et al. (2023). (BETHESDA (MD): NATIONAL LIBRARY OF MEDICINE (US), 1988; "RefSeq: NCBI Reference Sequence Database", 2023).

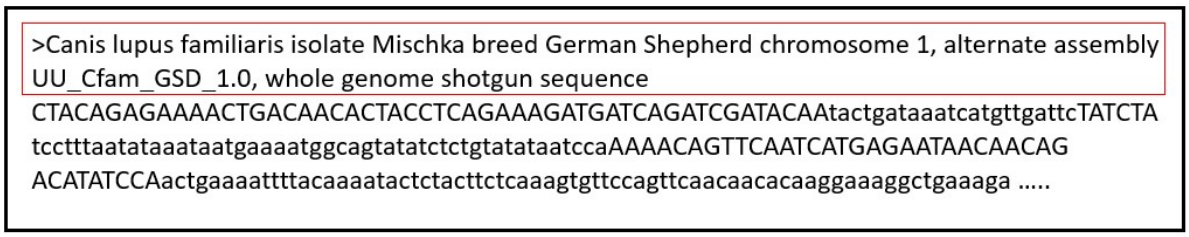
Após o download dos arquivos, iniciamos a confecção dos BD pelo Bowtie2 pela linha de comando, no ambiente conda, utilizando os parâmetros de 12 threads (-- threads 12) forçando a utilização da capacidade máxima do computador para que fosse executado com o menor tempo possível e -f (local do arquivo fasta com o genoma) separando os arquivos com os respectivos genomas com uma virgula (,) e em seguida separado por um espaço o local de destino do novo banco de dados, comando completo a seguir:

```
$ bowtie2-build --threads 12 -f \
          pasta_origem/genoma_cao.fna,pasta_origem/genoma_humano.fna \
          pasta_destino_BD/Takao_BT2db-Plus_dog_human/bt2
```

Para a confecção do banco de dado Kraken2, tivemos que realizar uma alteração no arquivo fasta para que o programa kraken2-build consiga identificar corretamente cada sequência de acordo com o sua taxonomia, então colocamos o número de identificação da taxonomia da espécie (taxid) na linha de identificação de cada sequência contida no arquivo (figura 5), o número de identificação taxonômica do *Canis lupus familiaris* (taxid: 9615) e do *Homo sapiens* (taxid: 9606) foram prospectados da base de dados pública de taxonomia do NCBI (BETHESDA (MD): NATIONAL LIBRARY OF MEDICINE (US), 1988; "Taxonomy Database", [s.d.]) e como cada arquivo pode ter milhares de linhas de identificação, utilizamos o software SeqFu na linha de comando, que é utilizado para manipulação de arquivos fasta, foi acrescentado a linha de identificação o seguinte texto "|kraken:taxid|000000" (figura 6), para isso utilizamos o seguinte comando:

```
$ seqfu cat --append "|kraken:taxid|(número_da_taxonomia)" \
          ~/local_arquivo fasta/arquivo_fna \
          > ~/destino_arquivo_editado/arquivo_taxid.fna
```

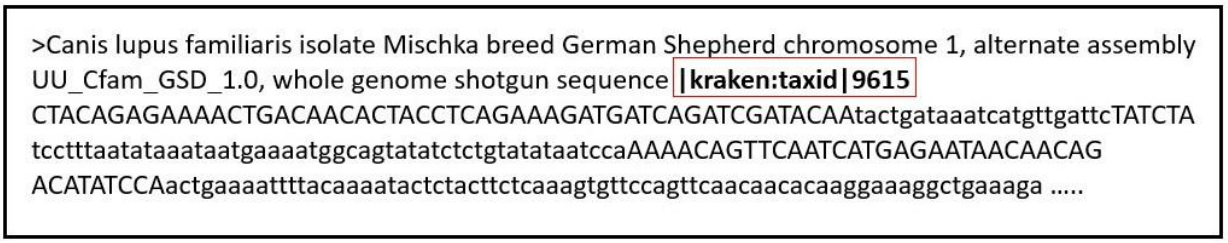
Figura 5: Primeira linha do arquivo fasta do genoma de *Canis lupus familiaris.*



Fonte: Takao et al. (2024)

Legenda: Destaca a linha de identificação onde deve conter a identificação da taxonomia para ferramenta kraken2-build.

Figura 6: Primeira linha do arquivo fasta do genoma de *Canis lupus familiaris* com o Taxid: 9615 em seu devido local para identificação pela ferramenta kraken2-build.



Fonte: Takao et al. (2024)

Com os arquivos fastas já preparados com os respectivos taxid iniciamos a construção da biblioteca para Kraken2 com os 2 genomas. A criação desse banco de dados consiste em 4 etapas, onde iremos criar uma biblioteca com os dois bancos de dados, em seguida será realizado o download da arvore filogenética para os dois genomas com a identificação de todas as suas taxonomias em seguida é realizada a compilação do banco de dados e por último realizamos a limpeza dos arquivos desnecessários criados durante o processo.

Etapa 1: Criação de uma biblioteca com os dois genomas, deve ser realizado os comandos 2x um para cada genoma e arquivá-los na mesma pasta de saída:

```
$ kraken2-build --add-to-library ~/local_arquivo_fasta/cao_taxid.fasta \
      --db ~/Takao_K2db-Plus_dog_human --threads 12

$ kraken2-build --add-to-library ~/local_arquivo_fasta/humano_taxid.fasta\
      --db ~/Takao_K2db-Plus_dog_human --threads 12
```

Etapa 2: Download da taxonomia:

```
$ kraken2-build --download-taxonomy --db ~/ Takao_K2db-Plus_dog_human
```

Etapa 3: Compilação do banco de dados final para Kraken2:

```
$ kraken2-build --build --db ~/Takao_K2db-Plus_dog_human
```

Etapa 4: Realização da limpeza de arquivos desnecessários utilizados para a criação do banco de dados:

```
$ kraken2-build --clean --db ~/Takao_K2db-Plus_dog_human
```

3 RESULTADOS E DISCUSSÃO

Após a compilação dos bancos de dados para ambas tivemos a criação de dois arquivos contendo os bancos de dados, sendo eles Takao_BT2db-plus_dog_human para Bowtie2 e Takao_K2db-plus_dog_human para Kraken2, para a confirmação da qualidade das ferramentas nós realizamos a inspeção "*inspect"* das mesmas onde foi possível identificar que os bancos de dados estavam completos e contendo todos os genomas e aptos a serem utilizados para a retirada de hospedeiro e contaminantes de bibliotecas de metagenoma canina e como teste complementar os dois bancos de dados foram testadas em uma biblioteca de metagenoma de cão com o SRA ID SRR19324877 disponíveis publicamente no *sequence reads arquive* (SRA) do NCBI.

Tabela 2: Teste realizado em uma biblioteca de metagenoma (SRR19324877) que possui 250.223.105 leituras.

Ferramenta	Leituras Brutas	Leituras alinhadas após análise	% de alinhamento	Leituras para análise do microbioma
Bowtie2	250223105	234749227	93,81%	15473878
Kraken2	250223105	250014348	99,91%	208757

Fonte: Takao et al. (2024)

Foi possível observar que os dois bancos de dados executados nas ferramentas Bowtie2 e Kraken2 com a configuração padrão das ferramentas, foi capas de ter uma boa porcentagem de alinhamentos e deixando apenas leituras de genomas não identificados como cão e ser humano.

A maior dificuldade para a construção dos bancos de dados foi a falta de artigos disponíveis sobre o tema e os manuais das ferramentas são muito sucintos nas explicações, faltando alguns dados fundamentais como a colocação da identificação taxonômica (taxid) no arquivo fasta para a criação do BD Kraken-2, a maioria das informações que tivemos foram a partir de fóruns e blogs do gênero, mas mesmo assim com alguns dados não tão precisos, a criação desses bancos de dados só foi possível após muitos testes de diferentes configurações para ambas as ferramentas.

Comparamos o nosso banco de dados utilizando o protocolo proposto por Lu et al. (2022) e obtivemos uma melhora significativa na remoção de sequências do hospedeiro e

contaminantes. Com o banco de dados e protocolo propostos por eles, obtivemos 54,28% de leituras não classificadas, sendo estas, em sua maioria, genomas do hospedeiro e contaminantes remanescentes. No entanto, com a utilização do nosso banco de dados, tivemos apenas 1,36% de leituras não classificadas, demonstrando que a utilização de um banco de dados mais completo e a combinação de Kraken2 com Bowtie2 resultam em um desempenho expressivamente melhor (LU et al., 2022).

4 CONCLUSÃO

Concluímos que os bancos de dados criados são viáveis e demonstraram êxito na realização do que foi proposto, sendo reprodutíveis para as mais diversas finalidades e genomas variados. O banco de dados está disponível em https://github.com/Paulotakao/Database-Kraken2-and-Bowtie2-canine-metagenomics

5 REFERÊNCIAS

BERG, G. et al. Microbiome definition re-visited: old concepts and new challenges. **Microbiome**, v. 8, n. 1, p. 103, 30 jun. 2020.

BETHESDA (MD): NATIONAL LIBRARY OF MEDICINE (US). **National Center for Biotechnology Information (NCBI)**. Disponível em: <https://www.ncbi.nlm.nih.gov/>. Acesso em: 2 nov. 2023.

BHARTI, R.; GRIMM, D. G. Current challenges and best-practice protocols for microbiome analysis. **Briefings in Bioinformatics**, v. 22, n. 1, p. 178–193, 18 jan. 2021.

COWAN, D. A. Microbial genomes – the untapped resource. **Trends in Biotechnology**, v. 18, n. 1, p. 14–16, jan. 2000.

GILBERT, J. A.; DUPONT, C. L. Microbial Metagenomics: Beyond the Genome. **Annual Review of Marine Science**, v. 3, n. 1, p. 347–371, 15 jan. 2011.

HANDELSMAN, J. et al. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. **Chemistry & Biology**, v. 5, n. 10, p. R245-249, out. 1998.

HANDELSMAN, J. Metagenomics: Application of Genomics to Uncultured Microorganisms. **Microbiology and Molecular Biology Reviews**, v. 68, n. 4, p. 669–685, dez. 2004.

LANGMEAD, B. et al. **Bowtie 2 Fast and sensitive read alignment**. **Bowtie 2**, 13 out. 2023. Disponível em: <https://bowtie-bio.sourceforge.net/bowtie2/manual.shtml#the-bowtie2-build-indexer>. Acesso em: 5 nov. 2023

LANGMEAD, B.; SALZBERG, S. L. Fast gapped-read alignment with Bowtie 2. **Nature Methods**, v. 9, n. 4, p. 357–359, abr. 2012.

LU, J. et al. Metagenome analysis using the Kraken software suite. **Nature Protocols**, v. 17, n. 12, p. 2815–2839, dez. 2022.

MARCHESI, J. R.; RAVEL, J. The vocabulary of microbiome research: a proposal. **Microbiome**, v. 3, n. 1, p. 31, 30 jul. 2015.

Mercado Pet Brasil 2023. **AMBIPET associação brasileira da industria de produtos para naimais de estimação**, 2023. Disponível em: <https://abinpet.org.br/dados-de-mercado/>. Acesso em: 2 nov. 2023

MERRIAM-WEBSTER. **Definition of MICROBIOME**. Disponível em: <https://www.merriam-webster.com/dictionary/microbiome>. Acesso em: 17 nov. 2023.

RefSeq: NCBI Reference Sequence Database. **National Library of Medicine**, 11 set. 2023. Disponível em: <https://www.ncbi.nlm.nih.gov/refseq/>. Acesso em: 5 nov. 2023

RUMBAVICIUS, I.; ROUNGE, T. B.; ROGNES, T. HoCoRT: host contamination removal tool. **BMC Bioinformatics**, v. 24, n. 1, p. 371, 2 out. 2023.

Taxonomy Database. **National Library of Medicine**, [s.d.]. Disponível em: <https://www.ncbi.nlm.nih.gov/taxonomy>. Acesso em: 5 nov. 2023

THE INTEGRATIVE HMP (IHMP) RESEARCH NETWORK CONSORTIUM. The Integrative Human Microbiome Project. **Nature**, v. 569, n. 7758, p. 641–648, maio 2019.

WHIPPS, J.; LEWIS, K.; COOKE, R. Mycoparasitism and plant disease control. **Fungi Biol Control Syst.**, p. 161–187, 1988.

WOOD, D. E.; SALZBERG, S. L. Kraken: ultrafast metagenomic sequence classification using exact alignments. **Genome Biology**, v. 15, n. 3, p. R46, 3 mar. 2014.

ARTIGO 4: Uma nova abordagem com Kraken2 e Bowtie2 na remoção de contaminantes, hospedeiros e taxonomia de bibliotecas de metagenômica canina.

Artigo a ser submetido:

Uma nova abordagem com Kraken2 e Bowtie2 na remoção de contaminantes, hospedeiros e taxonomia de bibliotecas de metagenômica canina.

Paulo Sallarola Takao[1]; Pâmela Souza Corrêa[1]; Rafaela Oliveira Campos[1]; Alessandra de Souza Fernandes[1]; Fabiana Borges[1]; David Aciole[1]; Renan Rodrigues Rocha[1]; Caio Junji Tanaka[1]; Rubens Pasa[2]; Karine Frehner Kavalco[2]; Fabiano Bezerra Menegidio[1]

1.Universidade de Mogi das Cruzes, UMC, Mogi das Cruzes, SP, Brasil

2. Universidade Federal de Viçosa, UFV, Viçosa, MG, Brasil

RESUMO

Neste estudo, abordamos uma metodologia inovadora para a análise metagenômica, focando no microbioma canino, utilizando ferramentas avançadas como Kraken2, Bowtie2 e InSilicoSeq para aprimorar a eliminação de sequências contaminantes e de hospedeiros e aumentar a precisão na classificação taxonômica. A metagenômica, que explora o DNA de amostras ambientais sem a necessidade de cultivo, abre janelas para a compreensão da complexidade das comunidades microbianas. Ao construir bancos de dados específicos e realizar testes extensivos, destacamos a importância de garantir a qualidade e a confiabilidade das análises. O uso de bibliotecas *MOCK*, simulações conhecidas de microrganismos, permitiu avaliar a eficácia das ferramentas de análise na remoção de contaminantes e na classificação taxonômica correta, mostrando uma melhoria significativa em relação aos métodos tradicionais e minimizando falsos positivos. A necessidade de bancos de dados abrangentes e bem curados tornou-se evidente, sendo crucial para a qualidade dos resultados. A combinação de uma seleção cuidadosa de ferramentas e configurações, junto com o uso de bibliotecas *MOCK*, forneceu uma base sólida para futuras pesquisas e aplicações em metagenômica, contribuindo para um entendimento mais profundo da diversidade microbiana e seu impacto na saúde. Este estudo demonstra o potencial das técnicas avançadas de análise metagenômica para melhorar significativamente a precisão e confiabilidade das análises do microbioma canino, marcando um avanço no campo da metagenômica e abrindo caminho para novas descobertas e aplicações.

Palavras-chave: Kraken2; InSilicoSeq; *MOCK*; Microbioma canino; Bowtie2;

1 INTRODUÇÃO

A metagenômica é um campo de estudo que surgiu como uma abordagem revolucionária para explorar a diversidade e complexidade das comunidades microbianas em diversos ecossistemas, incluindo o microbioma canino[1]. Essa abordagem se baseia no sequenciamento e análise do DNA diretamente extraído de amostras ambientais, sem a necessidade de cultivo prévio dos microrganismos[2]. Ao contrário das técnicas tradicionais de microbiologia, que se limitam ao estudo de microrganismos cultiváveis, a metagenômica permite a caracterização abrangente da composição taxonômica e do potencial funcional das comunidades microbianas, incluindo a vasta maioria dos microrganismos que não podem ser cultivados em laboratório[3].

A metagenômica é amplamente utilizada para investigar a diversidade e função dos microbiomas em diversos ambientes, incluindo solos, oceanos e o trato gastrointestinal de humanos e animais[4]. No contexto do microbioma canino, essa abordagem tem revelado informações valiosas sobre a composição e dinâmica das comunidades microbianas presentes no trato gastrointestinal, na pele e em outros sítios anatômicos dos cães[1]. Esses estudos têm destacado a importância do microbioma para a saúde e bem-estar dos cães, com implicações na nutrição, imunidade e prevenção de doenças[5].

Além de desvendar a diversidade taxonômica dos microbiomas, a metagenômica também permite a exploração do potencial funcional das comunidades microbianas, por meio da análise de sequências de genes e vias metabólicas[6]. Essa abordagem tem levado à descoberta de novos genes, enzimas e produtos naturais de interesse biotecnológico e farmacêutico[2]. Ademais, a metagenômica tem contribuído para a compreensão das interações complexas entre o microbioma e o hospedeiro, incluindo a modulação da resposta imune, a regulação do metabolismo e a proteção contra patógenos[7].

No entanto, a análise de dados metagenômicos apresenta desafios significativos devido à complexidade e ao volume dos dados gerados. Um dos principais desafios é a presença de sequências contaminantes e do hospedeiro misturadas às sequências microbianas de interesse[8]. Essas sequências indesejadas podem afetar a caracterização precisa da comunidade microbiana e levar a conclusões errôneas sobre a composição e função do microbioma[9]. Portanto, a remoção eficiente de sequências contaminantes e do hospedeiro é uma etapa crucial no processamento de dados metagenômicos.

Para lidar com esses desafios, várias ferramentas computacionais têm sido

desenvolvidas, entre as quais se destacam o Kraken2 e o Bowtie2. O Kraken2 é uma ferramenta de classificação taxonômica que utiliza uma abordagem baseada em k-mers para atribuir sequências metagenômicas a níveis taxonômicos específicos, desde o domínio até a espécie[10,11]. Essa ferramenta é conhecida por sua velocidade e precisão na classificação de grandes conjuntos de dados metagenômicos. Por outro lado, o Bowtie2 é um algoritmo de mapeamento de sequências que permite alinhar leituras metagenômicos a um conjunto de sequências de referência, como genomas de microrganismos ou sequências contaminantes[12,13]. O Bowtie2 é amplamente utilizado para a remoção de sequências do hospedeiro e contaminantes antes da análise taxonômica e funcional.

O desempenho do Kraken2 e do Bowtie2 depende fortemente da qualidade e abrangência dos bancos de dados de referência utilizados[14]. Bancos de dados genéricos, como os disponibilizados pelas ferramentas, podem não ser adequados para todos os contextos de estudo, pois não incluir sequências específicas do ambiente ou do hospedeiro em questão. Além disso, a presença de sequências mal anotadas ou de baixa qualidade nesses bancos de dados pode afetar a precisão da classificação taxonômica e da remoção de contaminantes[15]. Portanto, a criação de bancos de dados personalizados, específicos para o contexto do estudo, pode melhorar significativamente o desempenho dessas ferramentas e a confiabilidade dos resultados obtidos. Pensando nisso utilizamos os bancos de dados criados por Takao et al. (2024)[16], Takao_K2db-plus_dog_human para Kraken2 e Takao_BT2db-plus_dog_human para Bowtie2. Esses bancos de dados possuem maior eficiência na remoção de sequências de hospedeiro e contaminantes, pois são compostos por mais de um genoma de referência para o cão (*Canis lupus familiaris*) como hospedeiro e para o ser humano (Homo sapiens) como contaminante, conforme também indicado por Gihawi et al. (2023)[10,12,16,17].

Outra estratégia importante para avaliar a qualidade e ajustar os parâmetros das ferramentas de análise é o uso de bibliotecas simuladas (*MOCK*), que são misturas conhecidas de microrganismos em proporções definidas[18]. Essas bibliotecas podem ser criadas in vitro, por meio da mistura de culturas puras de microrganismos, ou in sílico, utilizando ferramentas como o InSilicoSeq[19]. O InSilicoSeq é uma ferramenta que permite gerar dados de sequenciamento sintéticos a partir de genomas de referência, simulando diferentes tecnologias de sequenciamento e parâmetros experimentais. Ao analisar uma biblioteca *MOCK* com os bancos de dados e ferramentas de interesse, é possível avaliar a precisão da classificação taxonômica e identificar possíveis vieses ou erros sistemáticos[18].

Além das ferramentas de classificação taxonômica e remoção de contaminantes, a

análise estatística e a visualização dos resultados são etapas essenciais na interpretação dos dados metagenômicos. Nesse contexto, a ferramenta Bracken que realiza o cálculo de abundancia e a ferramenta online Pavian que se mostra uma solução abrangente e interativa para a exploração e visualização de resultados de classificação taxonômica[20–22].

A integração do Kraken2, Bowtie2, Bracken e Pavian em um pipeline de análise metagenômica permite uma abordagem abrangente e robusta para a caracterização do microbioma canino. A criação de bancos de dados personalizados e o uso de bibliotecas *MOCK* garantem a qualidade e assertividade das ferramentas[18].

Nesse contexto, o presente estudo tem como objetivo criar uma biblioteca *MOCK* de metagenômica e juntamente com os bancos de dados criados por Takao et al.[16], avaliar e otimizar a utilização das ferramentas Kraken2 e Bowtie2 na remoção de hospedeiro e contaminantes e Kraken2 para a taxonomia de amostras de metagenômica incrementando o protocolo proposto por Lu et al. (2022)[10]. Além disso, será criada uma biblioteca *MOCK* usando o InSilicoSeq para que possamos ter uma amostra conhecida a fim de otimizar a configuração das ferramentas de análise. Essa abordagem permitirá uma caracterização mais precisa e confiável do microbioma canino, contribuindo para a compreensão de sua diversidade e papel na saúde e doença desses animais. Os resultados deste estudo têm o potencial de aprimorar as práticas de análise metagenômica em pesquisas sobre o microbioma canino e fornecer diretrizes para a criação de bibliotecas *MOCK* e o uso adequado de ferramentas computacionais nesse campo de pesquisa.

2 MATERIAIS E MÉTODOS

2.1 Criação de biblioteca *MOCK* de metagenômica com InSilicoSeq

A criação da biblioteca *MOCK* com a ferramenta InSilicoSeq começou com o download de todos os genomas necessários no formato FASTA (Tabela 3) da plataforma *RefSeq* (Banco de Dados de Sequência de Referência) do *NCBI* (National Center for Biotechnology Information)[14,23]. Os arquivos FASTA foram armazenados em uma pasta chamada 'genomas'. Em seguida, todos os genomas foram concatenados em um único arquivo FASTA usando o comando cat do Linux, resultando no arquivo 'all_genomes.fna'.

```
$ cat pasta_com_genomas/*.fna > pasta_destino/all_genomes.fna
```

Depois de concatenar os arquivos, usamos a ferramenta InSilicoSeq para criar a biblioteca de metagenoma sintético, especificando o modelo de erro do sequenciador metagenômico *Illumina NovaSeq*.

```
$ iss generate --draft pasta_destino/all_genomes.fna \
               --model novaseq --output pasta_destino/MOCK
```

O comando acima gerou os arquivos forward e reverse, totalizando 1 milhão de leituras (*reads*) simuladas. Após a compilação da biblioteca *MOCK*, utilizamos as ferramentas FastQC e fastp, também pela linha de comando, para realizar uma análise de similaridade com uma biblioteca de metagenoma real. Essa análise teve como objetivo identificar os erros esperados em um sequenciamento metagenômico realizado pelo sequenciador *Illumina NovaSeq*, tornando a biblioteca *MOCK* o mais próxima possível de dados reais.

A ferramenta FastQC foi utilizada para avaliar a qualidade das sequências geradas, fornecendo informações sobre a distribuição de tamanho dos reads, qualidade das bases, conteúdo de GC e presença de adaptadores. Já a ferramenta fastp foi empregada para realizar o pré-processamento dos dados, incluindo a remoção de adaptadores, filtragem de reads de baixa qualidade e trimagem de bases de baixa qualidade.

Tabela 3: Composição da biblioteca *MOCK*.

	Organismo	**NCBI TaxID**	**RefSeq**	**%**	**Reads**
1	*Canis lupus familiaris*	9615	GCF_011100685.1	59,1996	295998
2	*Homo sapiens*	9606	GCF_000001405.40	18,9816	94908
3	*Babesia bigemina*	5866	GCF_000981445.1	13,4758	67379
4	*Trichomonas vaginalis*	5722	GCF_026262505.1	5,4171	27085
5	*Leishmania panamensis*	5679	GCF_000755165.1	1,2657	6328
6	*Fusarium falciforme*	195108	GCF_026873545.1	0,5723	2861
7	*Fusarium graminearum*	5518	GCF_000240135.3	0,4002	2001
8	*Clostridium perfringens*	1502	GCF_020138775.1	0,2211	1105
9	*Pasteurella multocida*	747	GCF_002073255.2	0,0967	483
10	*Staphylococcus epidermidis*	1282	GCF_006094375.1	0,0941	470
11	*Klebsiella aerogenes*	548	GCF_007632255.1	0,0655	328
12	*Staphylococcus aureus*	1280	GCF_000013425.1	0,0594	297
13	*Serratia marcescens*	615	GCF_003516165.1	0,0437	219
14	*Human immunodeficiency virus*	12721	GCA_003190765.1	0,0291	145
15	*Staphylococcus felis*	46127	GCF_003012915.1	0,0220	110

16	*Tupanvirus soda lake*	2126985	GCF_002966485.1	0,0215	108
17	*Mycobacterium canettii*	78331	GCF_000253375.1	0,0144	72
18	*Aureococcus anophagefferens virus*	1474867	GCF_000922335.1	0,0085	43
19	*Escherichia coli*	562	GCF_000005845.2	0,0079	39
20	*Mycobacterium tuberculosis*	1773	GCF_000195955.2	0,0037	19

Fonte: Takao et al. (2024)
Legenda: Tabela inclui o nome científico de cada organismo, sua identificação taxonômica, código de referência no RefSeq do NCBI, porcentagem e quantidade de leituras presentes na biblioteca.

2.3 Testes realizados e avaliação dos bancos de dados

Foram realizados testes abrangentes utilizando a biblioteca *MOCK* gerada pelo InSilicoSeq para avaliar o desempenho e a eficácia na utilização conjunta e isolada das ferramentas Kraken2 e Bowtie2 na remoção de contaminantes e hospedeiros. Esses testes tiveram como objetivo identificar as melhores configurações e parâmetros das ferramentas.

Um total de 22 testes distintos foram conduzidos, sendo 12 com a utilização conjunta de Kraken2 e Bowtie2 na remoção de contaminantes e hospedeiro *workflow* 1 (figura 7) e 10 testes com a utilização sozinha de Kraken2 ou Bowtie2 *workflow* 2 e 3 (figura 8 e 9) e para todos os testes utilizamos Kraken2 para a taxonomia com valores de confiança de 0,0 e 0,1 para cada teste.

Os parâmetros testados estão descritos na Tabela 4.

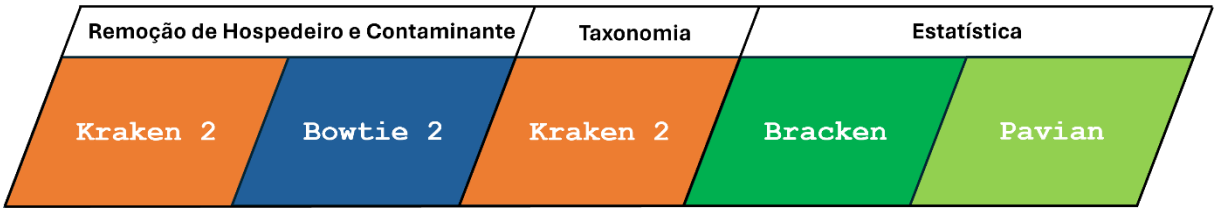
Durante os testes, foram avaliados diferentes aspectos do desempenho dos bancos de dados e das ferramentas, incluindo a precisão na identificação e remoção de sequências do hospedeiro e contaminantes, a sensibilidade na detecção de microrganismos presentes na biblioteca *MOCK*, a especificidade na atribuição taxonômica, o tempo de execução e os recursos computacionais necessários.

Diferentes combinações de parâmetros foram testadas, como valores de confiança, algoritmos de alinhamento e filtros adicionais, a fim de identificar as configurações ideais para cada banco de dados e ferramenta. Os resultados obtidos foram comparados com a composição conhecida da biblioteca *MOCK*, permitindo o cálculo de métricas de desempenho, como precisão, sensibilidade e especificidade.

Quando necessário, foram realizados ajustes nos parâmetros do Kraken2 e do Bowtie2 com base nos resultados obtidos, buscando otimizar o desempenho e a acurácia das análises. Esses ajustes foram documentados e justificados, visando estabelecer as melhores práticas para a análise de dados metagenômicos.

Os resultados dos testes foram tabulados e são apresentados de forma clara na seção de resultados, destacando as configurações que obtiveram o melhor desempenho. São discutidas as implicações desses resultados para a análise de dados metagenômicos reais e as possíveis limitações e desafios encontrados durante os testes.

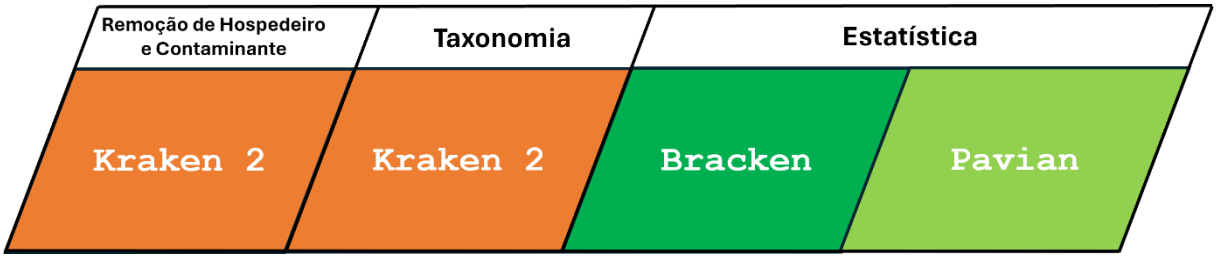
Figura 7: *Workflow* 1 utilizado para a avaliação dos bancos de dados e ferramentas nos testes com as bibliotecas *MOCK*.



Fonte: Takao et al. (2024)

O processo inicia-se com a remoção de sequências do hospedeiro e contaminantes utilizando o Kraken2. Em seguida, as leituras não classificadas pelo Kraken2 são analisadas com o Bowtie2 para identificar e remover possíveis sequências remanescentes do hospedeiro e contaminantes. As leituras restantes são então submetidas à classificação taxonômica utilizando o Kraken2 com o banco de dados PlusPF-16, que contém genomas de bactérias, vírus e eucariotos. Por fim, as análises estatísticas dos resultados são realizadas com as ferramentas Bracken e Pavian.

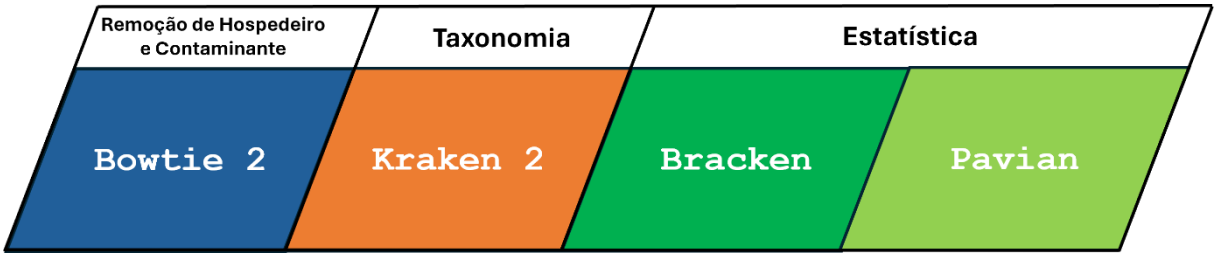
Figura 8: *Workflow* 2 utilizado para a avaliação dos bancos de dados e ferramentas nos testes com as bibliotecas *MOCK*.



Fonte: Takao et al. (2024)

O processo inicia-se com a remoção de sequências do hospedeiro e contaminantes utilizando o Kraken2. As leituras não classificadas são então submetidas à classificação taxonômica utilizando o Kraken2 com o banco de dados PlusPF-16, que contém genomas de bactérias, vírus e eucariotos. Por fim, as análises estatísticas dos resultados são realizadas com as ferramentas Bracken e Pavian.

Figura 9: *Workflow* 3 utilizado para a avaliação dos bancos de dados e ferramentas nos testes com as bibliotecas *MOCK*.



Fonte: Takao et al. (2024)

O processo inicia-se com a remoção de sequências do hospedeiro e contaminantes utilizando o Bowtie2. As leituras não classificadas são então submetidas à classificação taxonômica utilizando o Kraken2 com o banco de dados PlusPF-16, que contém genomas de bactérias, vírus e eucariotos. Por

fim, as análises estatísticas dos resultados são realizadas com as ferramentas Bracken e Pavian.

Tabela 4: Parâmetros e configurações utilizados nos testes de avaliação dos bancos de dados, Takao_K2db-plus_dog_human para Kraken2 e Takao_BT2db-plus_dog_human para Bowtie2.

Workflow	**Kraken2**	**Bowtie2**	**Kraken2 (taxonomia)**
2	Padrão	Não executado	Padrão
2	Padrão	Não executado	confiança de 0,1
2	confiança de 0,1	Não executado	Padrão
2	confiança de 0,1	Não executado	confiança de 0,1
3	Não executado	Padrão	Padrão
3	Não executado	Padrão	confiança de 0,1
3	Não executado	sensível	Padrão
3	Não executado	sensível	confiança de 0,1
3	Não executado	muito-sensível	Padrão
3	Não executado	muito-sensível	confiança de 0,1
1	Padrão	Padrão	Padrão
1	Padrão	Padrão	confiança de 0,1
1	Padrão	sensível	Padrão
1	Padrão	sensível	confiança de 0,1
1	Padrão	muito-sensível	Padrão
1	Padrão	muito-sensível	confiança de 0,1
1	confiança de 0,1	Padrão	Padrão
1	confiança de 0,1	Padrão	confiança de 0,1
1	confiança de 0,1	sensível	Padrão
1	confiança de 0,1	sensível	confiança de 0,1*
1	confiança de 0,1	muito-sensível	Padrão
1	confiança de 0,1	muito-sensível	confiança de 0,1

Fonte: Takao et al. (2024)

Foram testadas diferentes combinações de parâmetros, incluindo valores de confiança (confidence scores) e filtros adicionais. O objetivo foi identificar as configurações ideais para a remoção eficiente de sequências do hospedeiro e contaminantes, bem como para a classificação taxonômica precisa. Os testes abrangeram variações nos valores de confiança, como 0,0 e 0,1 para Kraken2, para determinar o limite mínimo de confiabilidade na classificação das sequências. Além disso, foram avaliados diferentes algoritmos de alinhamento, como end-to-end e local para Bowtie2 a fim de verificar qual abordagem resultava em uma melhor correspondência entre as sequências e os bancos de dados.
Padrão: Configurações padrão da ferramenta; **Confiança 0,1**: Ferramenta Kraken2 com valor de confiança de 0,1; **Sensível:** Ferramenta Bowtie2 com configuração de maior sensibilidade de ponta a ponta; **Muito sensível**: Ferramenta Bowtie2 com configuração de maior sensibilidade de ponta a ponta e local.

3 RESULTADOS E DISCUSSÃO

Após a montagem da biblioteca *MOCK* de metagenoma, obtivemos como resultado três arquivos: um arquivo forward (*MOCK*_R1.fastq) e um arquivo reverse (*MOCK*_R2.fastq) no formato FASTQ, além de um arquivo de abundância (*MOCK*_abundance.txt) no formato TXT (Tabelas 5 e 6). A análise desses arquivos utilizando as ferramentas FastQC (Figura 10) e fastp

(Figura 11) nos permitiu verificar a semelhança entre a biblioteca gerada e uma biblioteca real de metagenômica. O relatório do FastQC indicou que a biblioteca *MOCK* apresenta características típicas de um sequenciamento Illumina, como a distribuição dos scores de qualidade por base, o conteúdo de GC e a distribuição do tamanho das leituras. Já o relatório do fastp forneceu métricas detalhadas sobre a qualidade das sequências, como o número total de leituras, a porcentagem de bases com qualidade acima de Q30 e a presença de adaptadores.

Foi observada uma maior perda de qualidade nas leituras do arquivo reverse em comparação com o arquivo forward. No entanto, essa perda de qualidade é comum em bibliotecas reais de metagenômica e não prejudica a utilização dos dados. O relatório de abundância (Tabela 7) destaca a maior quantidade de leituras atribuídas ao cão (*Canis lupus familiaris*) e ao homem (*Homo sapiens*), mimetizando uma biblioteca metagenômica típica de amostras caninas. Essa distribuição de abundância permite avaliar a capacidade das ferramentas de análise em identificar e quantificar corretamente os organismos presentes, especialmente em cenários com alta contaminação por sequências do hospedeiro e de outros contaminantes comuns. Essas observações indicam que a biblioteca *MOCK* de metagenoma gerada apresenta características similares às encontradas em bibliotecas reais, validando assim a qualidade e a representatividade dos dados obtidos. Essa semelhança é essencial para garantir que as análises subsequentes, como a classificação taxonômica e a avaliação da diversidade microbiana, sejam confiáveis e reflitam adequadamente a composição da comunidade microbiana simulada.

Tabela 5: Arquivos de saída gerados pela ferramenta InSilicoSeq.

Arquivo	Tamanho
*MOCK*_R1.fastq	160 Mb
*MOCK*_R2.fastq	160 Mb
*MOCK*_abundance.txt	112 Kb

Fonte: Takao et al. (2024).

Com seus respectivos tamanhos e formatos. A ferramenta InSilicoSeq foi utilizada para simular dados de sequenciamento de uma biblioteca metagenômica, gerando três arquivos principais. O primeiro arquivo, denominado "*MOCK*_R1.fastq", contém as leituras forward no formato FASTQ e possui um tamanho de 160 Mb. O segundo arquivo, "*MOCK*_R2.fastq", contém as leituras reverse também no formato FASTQ e possui um tamanho de 160 Mb. Além disso, a ferramenta InSilicoSeq também gerou um arquivo de abundância, "*MOCK*_abundance.txt", no formato TXT, com um tamanho de 112 KB. Esse arquivo contém informações sobre a abundância relativa de cada organismo simulado na biblioteca metagenômica.

Tabela 6: Relatório de abundância dos organismos presentes na biblioteca *MOCK* de metagenoma, com um total de 500.000 leituras.

Organismo	%	Leituras
Aureococcus anophagefferens virus	0,0085	43

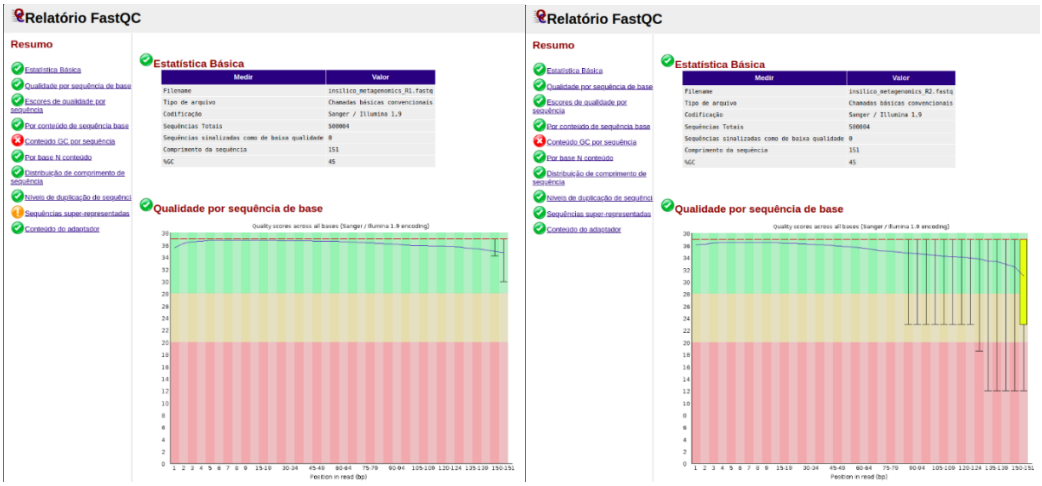
Babesia bigemina	13,4758	67379
Clostridium perfringens	0,2211	1105
Canis lupus familiaris	**59,1996**	**295998**
Escherichia coli	0,0079	39
Fusarium falciforme	0,5723	2861
Fusarium graminearum	0,4002	2001
Human immunodeficiency virus	0,0291	145
Homo sapiens	**18,9816**	**94908**
Klebsiella aerogenes	0,0655	328
Leishmania panamensis	1,2657	6328
Mycobacterium canettii	0,0144	72
Mycobacterium tuberculosis	0,0037	19
Pasteurella multocida	0,0967	483
Serratia marcescens	0,0437	219
Staphylococcus aureus	0,0594	297
Staphylococcus epidermidis	0,0941	470
Staphylococcus felis	0,0220	110
Tricomonas vaginalis	5,4171	27085
Tupanvirus soda lake	0,0215	108

Fonte: Takao et al. (2024)

Destaca-se a maior quantidade de leituras atribuídas ao cão e ao homem, mimetizando uma biblioteca metagenômica típica de amostras caninas. Essa distribuição de abundância permite avaliar a capacidade das ferramentas de análise em identificar e quantificar corretamente os organismos presentes, especialmente em cenários com alta contaminação por sequências do hospedeiro e de outros contaminantes comuns.

Figura 10: Apresenta o relatório gerado pela ferramenta FastQC para a biblioteca *MOCK* de metagenoma.

***MOCK*_R1.fastq** ***MOCK*_R2.fastq**



Fonte: Takao et al. (2024)

O relatório inclui métricas importantes, como a distribuição dos scores de qualidade por base, o conteúdo de GC, a presença de sequências super-representadas e a distribuição do tamanho das leituras. Através dessa análise, é possível verificar se a biblioteca *MOCK* apresenta características similares às

encontradas em dados reais de sequenciamento, validando assim a adequação dos parâmetros utilizados na simulação.

Figura 11: Relatório gerado pela ferramenta fastp para a análise da biblioteca *MOCK* de metagenoma.



relatório rápido

Sumário

General (Geral)

versão fastp:	0.20.1 (https://github.com/OpenGene/fastp)
de sequenciamento:	extremidade emparelhada (151 ciclos + 151 ciclos)
Comprimento médio antes de filtrar:	151bp, 151bp
comprimento médio após a filtragem:	150bp, 150bp, em pé
Taxa de duplicação:	0.007993%
Insira o pico do tamanho:	151 anos

Antes de filtrar

leituras totais:	1.000008 M
as bases totais:	151.001208 M (m)
Bases do Q20:	Endereço: 144.456094 M (95.665522%)
Bases do Q30:	140.925492 M (93.327394%)
Conteúdo do GC:	45.164616%

Após a filtragem

leituras totais:	Endereço: 977.384000 K
as bases totais:	Endereço de 147.534050 M
Bases do Q20:	Endereço: 141.531875 M (95.931668%)
Bases do Q30:	138.187002 M (93.664481%)
Conteúdo do GC:	45.274898%

Resultado de filtragem

As leituras passaram dos filtros:	977.384000 K (97.737618%)
leituras com baixa qualidade:	19.350000 K (1.934985%)
Leia sobre muitos N:	Baixar K (0.327397%)
É muito curto:	0 (0.0000%)

Fonte: Takao et al. (2024)

O relatório apresenta métricas detalhadas sobre a qualidade das sequências, incluindo o número total de leituras, a porcentagem de bases com qualidade acima de Q30, a distribuição dos tamanhos das leituras e a presença de adaptadores.

Após as análises de qualidade da amostra, iniciamos a remoção de hospedeiro e contaminantes, seguida da classificação taxonômica com o Kraken2 e análise estatística com Bracken e Pavian. Observamos que a utilização isolada das ferramentas Kraken2 ou Bowtie2, como proposto por Lu et a. (2022) para a remoção de contaminantes e hospedeiros é altamente ineficiente, deixando ainda muitas leituras não classificadas na amostra. Em nossos testes, a melhor configuração das ferramentas para remover hospedeiro e contaminantes foi a utilização do Kraken2 com suas configurações padrão, seguida da aplicação do Bowtie2 com as configurações de alinhamento end-to-end e local, aumentando a sensibilidade da ferramenta ao máximo. Com essa abordagem, alcançamos uma remoção de quase 99,93% das leituras de hospedeiro e contaminantes, diferentemente da utilização isolada do Bowtie2, que ainda manteve 44,76% de contaminantes na amostra, conforme demonstrado na Tabela 7. Quanto

menor a quantidade de hospedeiro e contaminantes na amostra a ser analisada, menor a probabilidade de se obter falsos positivos na análise taxonômica, evitando assim erros. Isso pode ser observado nas tabelas que foram geradas com o Pavian utilizando os arquivos *report* do Kraken2 após a taxonomia e análise de abundância da biblioteca *MOCK*, onde constatamos que, quanto maior a quantidade de leituras de contaminantes, maior a quantidade de falsos positivos. No entanto, mesmo com praticamente 100% dos contaminantes removidos, ainda se observa uma pequena quantidade de falsos positivos, reforçando a necessidade de melhorias nas ferramentas. Além disso, pudemos observar que a utilização de um banco de dados com mais genomas, conforme descrito por Gihawi et al., resultou em uma melhora na remoção de contaminantes, diminuindo assim os erros na classificação taxonômica.

Esses resultados destacam a importância de uma abordagem combinada e cuidadosa na remoção de hospedeiro e contaminantes, utilizando ferramentas adequadas e configurações otimizadas. A redução eficiente de leituras não desejadas é fundamental para minimizar a ocorrência de falsos positivos e garantir uma análise taxonômica mais precisa e confiável. Além disso, a escolha de bancos de dados abrangentes e bem curados também desempenha um papel crucial na qualidade dos resultados obtidos.

Tabela 7: Resultados dos testes de remoção de hospedeiro e contaminantes. O destacado em verde, apresentou a maior eficácia na remoção de leituras indesejadas, alcançando uma precisão de 78,26%. Os demais testes, utilizando diferentes configurações e ferramentas, demonstraram uma menor eficiência na remoção dessas leituras, com destaque para a utilização isolada do Bowtie2, que obteve 31,44%% de contaminantes removidos da amostra.

Especificações	Kraken2			Bowtie2			Removido HO CT	% removido HO CT	Kraken2 PlusPf			FP	FN
	% Clas.	% Ñ clas.	Total ñ clas.	% Clas.	% Ñ clas.	Total ñ clas.			% Clas.	% Ñ clas.	Total ñ clas.		
K2 + PlusPf	76,59%	23,41%	117051	x	x	0	382953	76,59	92,55%	7,45%	8720	7%	
K2 + PlusPf 0,1	76,59%	23,41%	117051	x	x	0	382953	76,59	87,16%	12,84%	15029		15,78%
K2 0,1 + PlusPf	75,67%	24,33%	121651	x	x	0	378353	75,67	89,97%	10,03%	12202	7%	
K2 0,1 + PlusPf 0,1	75,67%	24,33%	121651	x	x	0	378353	75,67	84,63%	15,37%	18698		15,78%
BT2 + PlusPf	x	x	0	31,44%	68,56%	342803	157201	**31,44**	45,72%	54,28%	186073	**17%**	
BT2 + PlusPf 0,1	x	x	0	31,44%	68,56%	342803	157201	31,44	42,02%	57,98%	198757		15,78%
BT2 sensível + PlusPf	x	x	0	18,41%	81,59%	407953	92051	18,41	45,69%	54,31%	221559	21%	
BT2 sensível + PlusPf 0,1	x	x	0	18,41%	81,59%	407953	92051	18,41	42,17%	57,83%	235919		15,78%
BT2 muito sensível + PlusPf	x	x	0	31,44%	68,56%	342803	157201	31,44	51,06%	48,94%	167768	19%	
BT2 muito sensível + PlusPf 0,1	x	x	0	31,44%	68,56%	342803	157201	31,44	46,98%	53,02%	181754		15,78%
K2 + BT2 + PlusPF	76,59%	23,41%	117051	1,57%	98,43%	115213	384791	76,96	93,71%	6,29%	7247	7%	
K2 + BT2 + PlusPF 0,1	76,59%	23,41%	117051	1,57%	98,43%	115213	384791	76,96	88,26%	11,74%	13526		15,78%
K2 + BT2 sensível + PlusPF	76,59%	23,41%	117051	1,64%	98,36%	115131	384873	76,97	93,79%	6,21%	7152	7%	
K2 + BT2 sensível + PlusPF 0,1	76,59%	23,41%	117051	1,64%	98,36%	115131	384873	76,97	88,33%	11,67%	13436		15,78%
K2 + BT2 muito sensível + PlusPF	76,59%	23,41%	117051	7,15%	92,85%	108681	391322	78,26	98,61%	1,39%	1511	8%	
K2 + BT2 muito sensível + PlusPF 0,1	76,59%	23,41%	117051	7,15%	92,85%	108681	391322	**78,26**	93,02%	6,98%	7586		**15,78%**
K2 0,1 + BT2 + PlusPF	75,67%	24,33%	121651	2,27%	97,73%	118889	381115	76,22	91,64%	8,36%	9939	8%	
K2 0,1 + BT2 + PlusPF 0,1	75,67%	24,33%	121651	2,27%	97,73%	118889	381115	76,22	86,23%	13,77%	16371		15,78%
K2 0,1 + BT2 sensível + PlusPF	75,67%	24,33%	121651	2,36%	97,64%	118780	381224	76,24	91,72%	8,28%	9835	8%	
K2 0,1 + BT2 sensível + PlusPF 0,1	75,67%	24,33%	121651	2,36%	97,64%	118780	381224	76,24	86,31%	13,69%	16261		15,78%
K2 0,1 + BT2muito sensível + PlusPF	75,67%	24,33%	121651	10,07%	89,93%	109400	390603	78,12	98,56%	1,44%	1575	8%	
K2 0,1 + BT2 muito sensível + PlusPF 0,1	75,67%	24,33%	121651	10,07%	89,93%	109400	390603	78,12	92,94%	7,06%	7724		15,78%

K2 = Kraken2 com configurações padrão; **K2 0,1** = Kraken2 com 0,1 de confiança; **BT2** = Bowtie2 com configurações padrão; **BT2 sensível** = configuração para aumentar sensibilidade de ponta a ponta**; BT2 muito sensível**: configurado para aumentar a sensibilidade local e de ponta a ponta; **PlusPf**: Ferramenta Kraken2 com banco de dados PlusPf, na configuração padrão; **PlusPf 0,1**: Ferramenta Kraken2 com banco de dados PlusPf, na configuração com 0,1 de confiança;

4 CONCLUSÃO

A construção de um banco de dados específico e abrangente é fundamental para a eficiente remoção de sequências de contaminantes e hospedeiro em análises metagenômicas. Os resultados obtidos neste estudo evidenciam a importância de incluir não apenas os genomas de referência mais recentes, mas também versões anteriores, a fim de garantir uma cobertura mais completa e precisa. Além disso, ficou evidente que a realização de testes extensivos utilizando amostras simuladas in silico, conhecidas como "*MOCK*", é uma etapa crucial no desenvolvimento de um *workflow* robusto para análises metagenômicas. Esses testes permitem avaliar o desempenho e a precisão das ferramentas utilizadas, identificando as melhores configurações e parâmetros para cada banco de dados e conjunto de dados. Embora as ferramentas atualmente disponíveis ainda apresentem limitações e não sejam perfeitas, os resultados obtidos neste estudo contribuem significativamente para aprimorar a qualidade e a precisão das análises metagenômicas. A combinação de bancos de dados abrangentes e específicos, juntamente com a realização de testes exaustivos, nos permite nos aproximar de um *workflow* ideal para a análise de dados metagenômicos. No entanto, é importante ressaltar que ainda há espaço para melhorias contínuas nas ferramentas e abordagens utilizadas. A comunidade científica deve se empenhar em desenvolver algoritmos mais precisos, bancos de dados mais completos e métodos de análise mais robustos. A colaboração entre pesquisadores, o compartilhamento de conhecimentos e a validação constante dos resultados são essenciais para impulsionar avanços nesse campo. Em suma, este estudo destaca a importância de uma abordagem cuidadosa e abrangente na construção de bancos de dados específicos e na realização de testes extensivos para garantir a qualidade e a confiabilidade das análises metagenômicas. Os resultados obtidos fornecem uma base sólida para futuros estudos e aplicações nessa área, contribuindo para uma melhor compreensão da complexidade e diversidade dos ecossistemas microbianos.

5 REFERÊNCIAS

1. Swanson KS, Dowd SE, Suchodolski JS, Middelbos IS, Vester BM, Barry KA, et al. Phylogenetic and gene-centric metagenomics of the canine intestinal microbiome reveals similarities with humans and mice. ISME J. abril de 2011;5(4):639–49.

2. Handelsman J. Metagenomics: Application of Genomics to Uncultured Microorganisms. Microbiol Mol Biol Rev. dezembro de 2004;68(4):669–85.

3. Hugenholtz P, Tyson GW. Metagenomics. Nature. setembro de 2008;455(7212):481–3.

4. Gilbert JA, Jansson JK, Knight R. The Earth Microbiome project: successes and aspirations. BMC Biology. 22 de agosto de 2014;12(1):69.

5. Deng P, Swanson KS. Gut microbiota of humans, dogs and cats: current knowledge and future opportunities and challenges. Br J Nutr. janeiro de 2015;113 Suppl:S6-17.

6. Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol. setembro de 2013;31(9):814–21.

7. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. Nature. 18 de outubro de 2007;449(7164):804–10.

8. Schmieder R, Edwards R. Fast Identification and Removal of Sequence Contamination from Genomic and Metagenomic Datasets. Rodriguez-Valera F, organizador. PLoS ONE. 9 de março de 2011;6(3):e17288.

9. Knight R, Jansson J, Field D, Fierer N, Desai N, Fuhrman JA, et al. Unlocking the potential of metagenomics through replicated experimental design. Nature Biotechnology. junho de 2012;30(6):513–20.

10. Lu J, Rincon N, Wood DE, Breitwieser FP, Pockrandt C, Langmead B, et al. Metagenome analysis using the Kraken software suite. Nat Protoc. dezembro de 2022;17(12):2815–39.

11. Wood D. DerrickWood/kraken2 [Internet]. 2024 [citado 7 de março de 2024]. Disponível em: https://github.com/DerrickWood/kraken2

12. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 4 de março de 2012;9(4):357–9.

13. Langmead B, Kin D, Charles R, Chen NC, Wilks C, Antonescu Va. Bowtie 2 Fast and sensitive read alignment [Internet]. Bowtie 2. 2023 [citado 5 de novembro de 2023]. Disponível em: https://bowtie-bio.sourceforge.net/bowtie2/manual.shtml#the-bowtie2-build-indexer

14. Nasko DJ, Koren S, Phillippy AM, Treangen TJ. RefSeq database growth influences the accuracy of k-mer-based lowest common ancestor species identification. Genome Biology. 30 de outubro de 2018;19(1):165.

15. Méric G, Wick R, Watts S, Holt K, Inouye M. Correcting index databases improves metagenomic studies. 2019.

16. Takao PS, Moura JF, Correa PS, Barbosa DA, Menegidio FB. DESENVOLVIMENTO DE BANCOS DE DADOS PARA ANÁLISE METAGENÔMICA DE MICROBIOMAS DE CÃES: UMA ABORDAGEM COM KRAKEN2 E BOWTIE2. Em: Anais do IV Congresso Brasileiro de Biotecnologia On-line [Internet]. Revista Multidisciplinar de Educação e Meio Ambiente; 2024 [citado 13 de julho de 2024]. Disponível em: https://ime.events/conbiotec2024/pdf/33826

17. Gihawi A, Ge Y, Lu J, Puiu D, Xu A, Cooper CS, et al. Major data analysis errors invalidate cancer microbiome findings. mBio. 31 de outubro de 2023;14(5):e0160723.

18. Bokulich NA, Rideout JR, Mercurio WG, Shiffer A, Wolfe B, Maurice CF, et al. *MOCK*robiota: a Public Resource for Microbiome Bioinformatics Benchmarking. mSystems. 2016;1(5):e00062-16.

19. Gourlé H, Karlsson-Lindsjö O, Hayer J, Bongcam-Rudloff E. Simulating Illumina metagenomic data with InSilicoSeq. Bioinformatics. 1º de fevereiro de 2019;35(3):521–2.

20. Breitwieser FP, Salzberg SL. Pavian: Interactive analysis of metagenomics data for microbiomics and pathogen identification [Internet]. bioRxiv; 2016 [citado 14 de maio de 2024]. p. 084715. Disponível em: https://www.biorxiv.org/content/10.1101/084715v1

21. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: estimating species abundance in metagenomics data. PeerJ Computer Science. 2 de janeiro de 2017;3:e104.

22. Lu J. jenniferlu717/Bracken [Internet]. 2024 [citado 7 de março de 2024]. Disponível em: https://github.com/jenniferlu717/Bracken

23. Bethesda (MD): National Library of Medicine (US). National Center for Biotechnology Information (NCBI) [Internet]. National Center for Biotechnology Information (NCBI). 1988 [citado 2 de novembro de 2023]. Disponível em: https://www.ncbi.nlm.nih.gov/

4.3 ARTIGO 5: Protocolo de Classificação Metagenômica Canina

Artigo a ser submetido:

Protocolo de Classificação Metagenômica Canina

Paulo Sallarola Takao[1]; Pâmela Souza Corrêa[1]; Rafaela Oliveira Campos[1]; Alessandra de Souza Fernandes[1]; Fabiana Borges[1]; David Aciole[1]; Renan Rodrigues Rocha[1]; Caio Junji Tanaka[1]; Rubens Pasa[2]; Karine Frehner Kavalco[2]; Fabiano Bezerra Menegidio[1]

1.Universidade de Mogi das Cruzes, UMC, Mogi das Cruzes, SP, Brasil

2. Universidade Federal de Viçosa, UFV, Viçosa, MG, Brasil

RESUMO

Este artigo apresenta um protocolo avançado para a análise metagenômica, focando na otimização da classificação taxonômica e na remoção eficiente de contaminantes, com ênfase particular em amostras caninas. A metodologia proposta destaca a importância de manter bancos de dados de referência atualizados e robustos para garantir a precisão das análises metagenômicas, diante do crescente volume de genomas disponíveis. O estudo demonstra como a implementação estratégica de ferramentas, como Kraken2 seguido por Bowtie2, e o uso de bancos de dados extensivos, são essenciais para melhorar a velocidade e a acurácia da classificação, apesar do aumento no uso de recursos computacionais. O artigo detalha cada etapa do protocolo, fornecendo comandos específicos e parâmetros para cada ferramenta utilizada, e exemplifica sua aplicação em um conjunto de dados metagenômicos caninos reais. Além disso, discute as vantagens e limitações da abordagem proposta, esperando que o protocolo sirva como um recurso valioso para pesquisadores na área de análises metagenômicas robustas e reproduzíveis. A implementação do novo protocolo demonstrou uma eficiência temporal significativa, reduzindo o tempo de execução das ferramentas utilizadas e permitindo a realização de estudos mais amplos de forma eficiente. Apesar dos avanços, o artigo reconhece a necessidade de melhorias contínuas, especialmente na minimização de taxas de erro para aumentar a acurácia das análises. A evolução contínua das técnicas de sequenciamento e análise metagenômica é vista como fundamental para superar desafios existentes e explorar plenamente o potencial da metagenômica, com o objetivo de expandir nosso entendimento sobre a complexidade dos ecossistemas microbianos.

Palavras-chave: Metagenômica; Análise Taxonômica; Kraken2; Bowtie2; Bancos de Dados

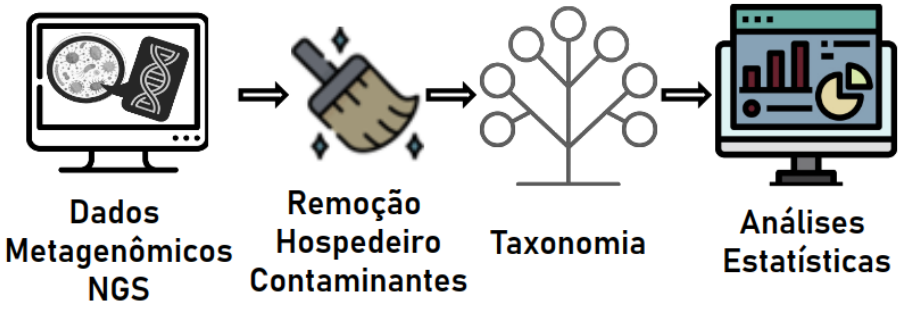
1 INTRODUÇÃO

Os avanços nas tecnologias de sequenciamento de nova geração (NGS) revolucionaram o campo da metagenômica, permitindo a caracterização abrangente de comunidades

microbianas complexas [1]. A metagenômica, que envolve o sequenciamento direto do DNA total de uma amostra ambiental, oferece uma visão sem precedentes da diversidade, estrutura e função dos microbiomas [2]. Essa abordagem tem sido amplamente aplicada em diversos campos, desde a ecologia microbiana até a medicina personalizada, revelando a importância crucial das comunidades microbianas em ecossistemas e na saúde humana[3,4]. A metagenômica canina, em particular, tem ganhado crescente atenção devido ao papel fundamental que a microbiota desempenha na saúde e nas doenças dos cães[5]. Estudos recentes destacaram a influência da microbiota intestinal canina no metabolismo, imunidade e comportamento dos cães, bem como sua associação com várias condições patológicas[6]. No entanto, a análise de dados metagenômicos apresenta desafios significativos, incluindo a presença de sequências contaminantes do hospedeiro e a necessidade de uma classificação taxonômica precisa[7,8].

O fluxo de trabalho típico em experimentos de microbioma envolve várias etapas cruciais, conforme descrito por[9,10]. Inicialmente, é realizada a remoção do DNA do hospedeiro, seguida pela classificação das leituras de sequenciamento, atribuindo cada leitura a uma categoria taxonômica específica, como espécie, gênero ou táxon de nível superior. Em seguida, é calculada a abundância relativa das diferentes espécies presentes na amostra (Fig. 12). Embora alguns estudos se concentrem em um conjunto limitado de genes marcadores, como o *RNA* ribossômico ou genes codificadores de proteínas, para a classificação taxonômica, muitos estudos optam por analisar todas as leituras de sequenciamento obtidas da amostra, abordagem que é o foco deste protocolo. Após a caracterização dos dados brutos, análises downstream podem ser realizadas, incluindo métodos estatísticos para comparar as composições microbianas de diferentes ambientes e métodos de visualização para compreender as composições microbianas.

Figura 12: Passo a passo de uma análise metagenômica.



Fonte: Takao et al. (2024)

1ª obtenção da amostra, 2ª remoção de hospedeiro e contaminantes, 3ª análise da taxonomia e 4ª análises estatísticas.

Neste artigo, apresentamos um novo protocolo para a análise metagenômica de amostras caninas, que aborda essas questões através da integração de várias ferramentas bioinformáticas state-of-the-art. Este protocolo foi baseado e aperfeiçoado a partir do artigo "*Metagenome analysis using the Kraken software suite"* de Lu et al.[10], que fornece uma estrutura abrangente para a análise metagenômica utilizando o conjunto de ferramentas Kraken [10]. Nosso protocolo começa com a utilização do Kraken2 e do Bowtie2 para remover eficientemente as sequências do hospedeiro (*Canis lupus familiaris*) e dos contaminantes humanos (*Homo sapiens*) indesejados. O Kraken2 é um classificador taxonômico ultrarrápido e preciso[11], enquanto o Bowtie2 é um alinhador de sequências amplamente utilizado[12]. Após a remoção de contaminantes, aplicamos novamente o Kraken2 para atribuir taxonomia às sequências metagenômicas da microbiota. Em seguida, utilizamos o Bracken, uma ferramenta complementar ao Kraken2, para estimar a abundância de espécies microbianas na amostra[10,13]. O Bracken utiliza a contagem de k-mers e a classificação do Kraken2 para fornecer estimativas mais precisas da abundância em nível de espécie [10,13]. Os resultados são então visualizados utilizando o Pavian, uma interface web interativa que permite a exploração intuitiva dos dados taxonômicos e de abundância[14].

A combinação estratégica dessas ferramentas em um fluxo de trabalho integrado oferece uma abordagem abrangente e eficiente para a análise metagenômica de amostras caninas. O protocolo proposto não apenas melhora a precisão da classificação taxonômica e a remoção de contaminantes, mas também fornece uma visualização intuitiva dos resultados e garante a acessibilidade dos dados. Ao longo deste artigo, descreveremos detalhadamente cada etapa do protocolo, fornecendo comandos e parâmetros específicos para cada ferramenta. Demonstraremos a aplicação do protocolo em um conjunto de dados metagenômicos caninos reais e discutiremos as vantagens e limitações da abordagem proposta. Esperamos que este protocolo sirva como um recurso valioso para pesquisadores que buscam realizar análises metagenômicas robustas e reproduzíveis em amostras caninas, contribuindo para uma melhor compreensão da microbiota canina e seu papel na saúde e na doença. Este protocolo foi concebido com foco em leituras de sequenciamento geradas pela tecnologia Illumina, notáveis por sua alta precisão, com taxas de erro inferiores a 0,5%, embora limitadas a um comprimento de 100 a 250 pares de bases (pb)[15].

1.1 Limitações na Remoção de Leituras de Hospedeiro e Contaminantes

Em estudos metagenômicos voltados para amostras caninas, a etapa de remoção de DNA proveniente do hospedeiro e de contaminantes, predominantemente humanos. Essa remoção, geralmente realizada como uma fase preliminar de pré-processamento, emprega a ferramenta Bowtie2 [12] para o alinhamento de sequências ao genoma do hospedeiro, porém nesse protocolo acrescentamos o Kraken2 [11], que se vale de um banco de dados específico para identificar e excluir tais contaminantes. A adoção dessa estratégia de pré-processamento é frequentemente justificada pela significativa redução no volume de leituras, o que, por sua vez, facilita e agiliza as análises subsequentes [10]. Contudo, a aplicação desse pré-processamento não está isenta de desafios. Há o risco de exclusão inadvertida de sequências potencialmente relevantes para o estudo, particularmente aquelas que, por sua raridade ou peculiaridade, como é o caso de sequências virais análogas a retrovírus endógenos humanos, poderiam oferecer descobertas valiosas [16]. Em situações assim, o mapeamento direto dessas leituras em um banco de dados de genomas virais, sem a prévia filtragem de material humano, revela-se uma alternativa promissora para a detecção desses elementos.

Neste protocolo, propomos a inclusão do pré-processamento com Kraken2 e Bowtie2 como uma etapa facultativa, permitindo aos usuários do protocolo a liberdade de escolher incorporá-la ou não, conforme as necessidades específicas de sua investigação. Esta abordagem flexível busca equilibrar a eficiência na análise de dados metagenômicos com a preservação da integridade e da riqueza informativa do conjunto de leituras, assegurando que decisões metodológicas não comprometam a capacidade de descoberta de novos agentes ou sequências de interesse.

1.2 Banco de Dados de Referência

A acurácia nas análises realizadas por ferramentas como Kraken2 e Bowtie2 está intrinsecamente atrelada à qualidade e abrangência dos bancos de dados de referência utilizados. Ainda estamos distantes de um cenário onde a totalidade da biosfera microbiana esteja mapeada e sequenciada, o que significa que a ausência ou incompletude de genomas no banco de dados pode introduzir vieses e comprometer o desempenho da classificação metagenômica. Leituras que não encontram correspondência no banco de dados são invariavelmente categorizadas como não classificadas ou, erroneamente, associadas ao táxon mais similar disponível. Adicionalmente, a contaminação de genomas por DNA de outros organismos, um fenômeno não raro segundo estudos como o que identificou milhares de

proteínas espúrias devido à contaminação humana em genomas bacterianos, pode levar à atribuição incorreta de rótulos taxonômicos pelo Kraken, afetando, ainda que minimamente, a precisão da classificação [17,18].

Este protocolo, utilizamos os bancos de dados para taxonomia disponibilizados pela ferramenta Kraken2 [19], que adota uma estratégia criteriosa na seleção dos genomas para o banco de dados de referência, priorizando genomas completos do *RefSeq* (*Reference Sequence Database*) [20] do NCBI (*National Center for Biotechnology Information*) [21]. Esta escolha visa minimizar a possibilidade de contaminação que poderia advir de genomas incompletos. Para a remoção de DNA contaminante e de hospedeiro, recorremos à criação de um banco de dados que inclui múltiplos genomas de referência tanto para cães quanto para humanos, uma prática que, conforme demonstrado por Gihawi et al.[19], se mostra mais eficiente na eliminação de contaminantes. Além disso, conforme os bancos de dados de referência expandem, novos desafios emergem, como a distribuição desigual de genomas por táxon. Nasko et al.[20] observaram que o crescimento do banco de dados *RefSeq* pode influenciar a tendência do Kraken em classificar uma proporção maior de leituras em níveis taxonômicos superiores, sobretudo no nível de gênero, em detrimento da precisão em níveis inferiores, como espécie [20]. Este fenômeno sublinha a importância de um monitoramento contínuo e ajustes nos bancos de dados de referência, para garantir que o aumento quantitativo de genomas disponíveis não prejudique a acuidade das classificações metagenômicas.

Portanto, a construção e manutenção de bancos de dados de referência robustos e atualizados são fundamentais para o sucesso das análises metagenômicas, exigindo um equilíbrio cuidadoso entre abrangência, precisão e a constante evolução do conhecimento genético [10].

1.3 Abordagens Baseadas em Montagem

Quando uma espécie está presente em quantidades suficientemente elevadas, pode-se desejar montar as leituras de novo para criar sequências grandes e contíguas daquela espécie. Os procedimentos aqui descritos não realizam a montagem do genoma, o que requer outros métodos e adição de mais softwares como o Megahit [10,20,22,23].

2 MATERIAIS E MÉTODOS

2.1 Softwares Necessários

- Kraken2 (https://github.com/DerrickWood/kraken2/, versão 2.1.2 ou posterior) [19]
- Bracken (https://github.com/jenniferlu717/Bracken, versão 2.6.2 ou posterior) [24].
- Pavian (https://fbreitwieser.shinyapps.io/pavian/) [14].
- Bowtie2 (https://github.com/BenLangmead/bowtie2, versão 2.4.4 ou posterior) [25].

2.2 Hardware

- Linux 64-bit
- 40Gb de RAM
- 2Tb armazenamento
- Core I5 12 threads

2.3 Dados para Análise Teste

Para a análise teste foi utilizada uma biblioteca simulada de metagenômica (*MOCK*) criada com InSilicoSeq [26], contendo o genoma do cão e do homem como hospedeiro e contaminante, respectivamente, e mais 18 genomas de fungos, bactérias, protozoários e vírus (tabela 8) mimetizando uma biblioteca de metagenômica, gerada por um aparelho *Illumina NovaSeq* [27] onde tivemos como resultado dois arquivos sendo 1 *forward* (*MOCK*_R1.fastq) e 1 *reverse* (*MOCK*_R2.fastq) que foram armazenados na pasta ~/*MOCK*. O passo a passo para a confecção do *MOCK* está disponível no anexo 1.

Tabela 8: Composição da biblioteca *MOCK*.

	Organismo	**NCBI TaxID**	***RefSeq***	**%**	**Reads**
1	*Canis lupus familiaris*	9615	GCF_011100685.1	59,1996	295998
2	Homo sapiens	9606	GCF_000001405.40	18,9816	94908
3	Babesia bigemina	5866	GCF_000981445.1	13,4758	67379
4	Trichomonas vaginalis	5722	GCF_026262505.1	5,4171	27085
5	Leishmania panamensis	5679	GCF_000755165.1	1,2657	6328
6	Fusarium falciforme	195108	GCF_026873545.1	0,5723	2861
7	Fusarium graminearum	5518	GCF_000240135.3	0,4002	2001
8	Clostridium perfringens	1502	GCF_020138775.1	0,2211	1105
9	Pasteurella multocida	747	GCF_002073255.2	0,0967	483
10	Staphylococcus epidermidis	1282	GCF_006094375.1	0,0941	470
11	Klebsiella aerogenes	548	GCF_007632255.1	0,0655	328
12	Staphylococcus aureus	1280	GCF_000013425.1	0,0594	297
13	Serratia marcescens	615	GCF_003516165.1	0,0437	219
14	Human immunodeficiency virus	12721	GCA_003190765.1	0,0291	145
15	Staphylococcus felis	46127	GCF_003012915.1	0,0220	110

16	Tupanvirus soda lake	2126985	GCF_002966485.1	0,0215	108
17	Mycobacterium canettii	78331	GCF_000253375.1	0,0144	72
18	Aureococcus anophagefferens virus	1474867	GCF_000922335.1	0,0085	43
19	Escherichia coli	562	GCF_000005845.2	0,0079	39
20	Mycobacterium tuberculosis	1773	GCF_000195955.2	0,0037	19

Fonte: Takao et al. (2024)

Nome científico de cada organismo, sua identificação taxonômica, código de referência no *RefSeq*, porcentagem presente e quantidade de leituras presentes na biblioteca.

2.4 Bancos de Dados

Para a remoção de hospedeiro e contaminantes, criamos dois bancos de dados (BD) específicos, sendo um para Kraken2 (Takao-K2db-Plus_dog_human) [28] e um para Bowtie2 (Takao-BT2db-Plus_dog_human) [28], ambos os bancos de dados com 4 genomas de referência para cão e 2 genomas de referência para o ser humano (Tabela 9). O BD para Kraken2 foi armazenado em ~/index/k2/Takao-K2db-Plus_dog_human e o BD para Bowtie2 foi armazenado em ~/index/bt2/Takao-BT2db-Plus_dog_human. Já para a taxonomia utilizamos o BD disponibilizado pela ferramenta Kraken2 (https://benlangmead.github.io/aws-indexes/k2) [25], PlusPf-16 com última atualização em 12/01/2024 com tamanho de 16Gb contendo genomas completos de fungos, bactérias, eucariotos e vírus, que estão disponíveis no *RefSeq* [29]. O banco de dados foi baixado diretamente do site https://benlangmead.github.io/aws-indexes/k2 [25] para a pasta ~/index/k2 no formato tar.gz e descompactado logo em seguida criando o arquivo ~/index/k2/pluspf_16gb.

Tabela 9: Genomas utilizados para a construção dos bancos de dados "Takao-db_dog_human".

Nome científico	Biblioteca	RefSeq	Nível	Data
Canis lupus familiais	Dog10k_Boxer_Tasha	GCF_000002285.5	Cromossomo	Outubro 2020
Canis lupus familiais	UMICH_Zoey_3.1	GCF_005444595.1	Cromossomo	Maio 2019
Canis lupus familiais	UU_Cfam_GSD_1.0	GCF_011100685.1	Cromossomo	Março 2020
Canis lupus familiais	ROS_Cfam_1.0	GCF_014441545.1	Cromossomo	Setembro 2020
Homo Sapiens	GRCh38.p14	GCF_000001405.40	Cromossomo	Fevereiro 2022
Homo Sapiens	T2T-CHM13v2.0	GCF_009914755.1	Completo	Janeiro 2022

Fonte: Takao et al. [28]

A tabela contendo quatro genomas do hospedeiro canino e dois genomas contaminante humano, obtidos a partir do banco de dados público *RefSeq* (https://www.ncbi.nlm.nih.gov/refseq/) [21].

2.5 Instalação das Ferramentas

Todas as ferramentas utilizadas, exceto o Pavian [14], foram instaladas no computador através da linha de comando Linux. As ferramentas instaladas estão disponíveis no gerenciador de pacotes de código aberto Conda [30,31]. Essas ferramentas são armazenadas no Bioconda [30], que é um repositório específico para ferramentas de bioinformática.

Antes da instalação das ferramentas, realizamos a instalação do pacote Conda no computador, realizando o download do instalador para Linux diretamente do site (https://docs.conda.io/projects/conda/en/latest/user-guide/install/linux.html) [31] e, em seguida, utilizamos o seguinte comando via terminal:

```
$ bash anaconda-latest-Linux-x86_64.sh
```

Após a instalação do pacote conda, criamos dois ambientes diferentes para a execução dos programas, devido a conflitos de funcionamento entre Kraken2 e Bracken, já que o Bracken é baseado no Kraken (1).

Ambiente 1:

```
$ conda create -n metagenomics
```

Ambiente 2:

```
$ conda create -n bracken
```

Para trocar os ambientes, utilizamos o comando:

```
$ conda activate metagenomics
```

ou

```
$ conda activate bracken
```

Agora, com os dois ambientes criados, iniciamos a instalação das ferramentas.

- **Kraken2 (ambiente metagenomics):**

```
$ conda install bioconda::kraken2
$ conda update kraken2
```

- **Bowtie2 (ambiente metagenomics):**

```
$ conda install bioconda::bowtie2
$ conda update bowtie2
```

- **Bracken (ambiente bracken):**

```
$ conda install bioconda::bracken
$ conda update bracken
```

2.6 Análise do Microbioma

2.6.1 Remoção de Hospedeiro e Contaminante

Antes de iniciar as leituras, criamos a pasta ~/cao_metagenomica, onde iremos realizar todas as análises. A primeira ferramenta a ser utilizada é o Kraken2 com o seguinte comando (~6 segs.):

```
$ kraken2 \
      --use-names --threads 10 \
      --db ~/index/k2/Takao-K2db-Plus_dog_human \
      --paired ~/MOCK/MOCK_R1.fastq  ~/MOCK/MOCK_R2.fastq \
      --unclassified-out \
      ~/cao_metagenomica/MOCK_k2_no_dog-human_un#.fastq \
      --report ~/cao_metagenomica/MOCK_k2_no_dog-human.k2report \
      > ~/cao_metagenomica/MOCK_k2_no_dog-human_k2.kraken
```

OBS: Colocar o símbolo (#) antes do .fastq faz com que sejam gerados dois arquivos de saída, um forward e um reverse das leituras não classificadas.

Em seguida, utilizamos a ferramenta Bowtie2 para fazer o alinhamento do genoma restante de cão e humano não classificado pela ferramenta Kraken2, utilizando o arquivo gerado

pelo comando --unclassified-out (*MOCK*_k2_no_dog-human_un_1.fastq e *MOCK*_k2_no_dog-human_un_2.fastq) da ferramenta Kraken2. Para o alinhamento, utilizamos os comandos a seguir (~40 segs.):

```
$ bowtie2 -p 10 --very-sensitive --very-sensitive-local \
        -x ~/index/bt2/Takao-BT2db-Plus_dog_human/bt2 \
        -1 ~/cao_metagenomica/MOCK_k2_no_dog-human_1.fastq \
        -2 ~/cao_metagenomica/MOCK_k2_no_dog-human_2.fastq \
        --un-conc ~/cao_metagenomica/MOCK_k2_bt2_no_dog-human_un.fastq \
        -S ~/cao_metagenomica/MOCK_k2_bt2_no_dog-human_un.sam
```

Obs: Para o banco de dados ser identificado pela ferramenta, obrigatoriamente deve ser colocado o primeiro nome dos arquivos presentes na pasta do banco de dados, sendo, nesse caso, “bt2”.

2.6.2 Taxonomia com Kraken2

Para a taxonomia com Kraken2, utilizamos o banco de dados PlusPf-16 com os seguintes comandos (~8 segs.):

```
$ kraken2 \
        --use-names \
        --threads 12 \
        --db ~/index/k2/pluspf_16gb \
        --minimum-hit-groups 3 --paired \
       ~/cao_metagenomica/MOCK_k2_bt2_no_dog-human_un.1.fastq \
       ~/cao_metagenomica/MOCK_k2_bt2_no_dog-human_un.2.fastq \
        --output ~/cao_metagenomica/MOCK_k2_bt2_pf.kraken \
        --report ~/cao_metagenomica/MOCK_k2_bt2_pf.kreport
```

2.7 Análises Estatísticas

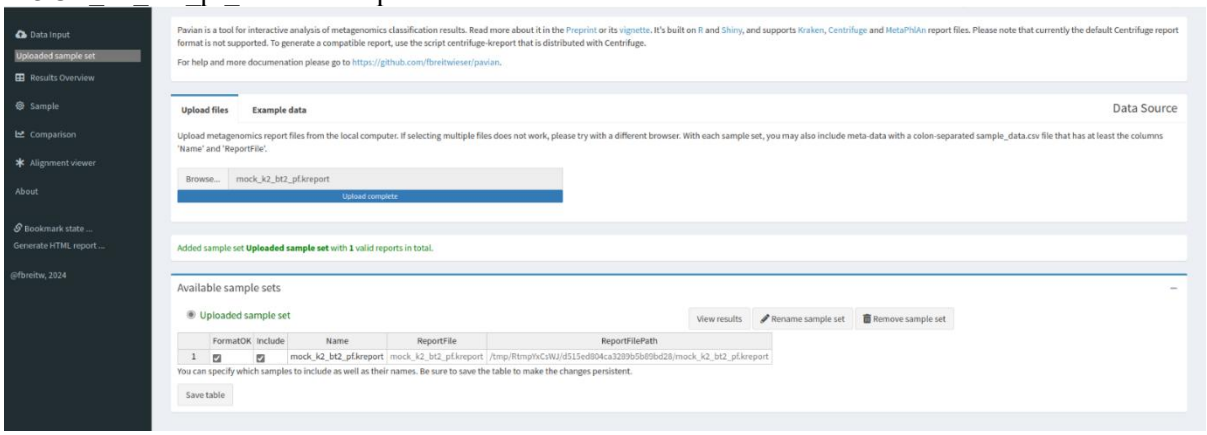
Para as análises estatísticas, iniciamos com o Bracken com os seguintes comandos (~2 segs.):

```
$ bracken \
```

```
-d ~/index/k2/pluspf_16gb \
-i ~/cao_metagenomica/MOCK_k2_bt2_pf.kreport \
-o ~/cao_metagenomica/MOCK_k2_bt2_pf.bracken \
-r 100 -l S -t 12
```

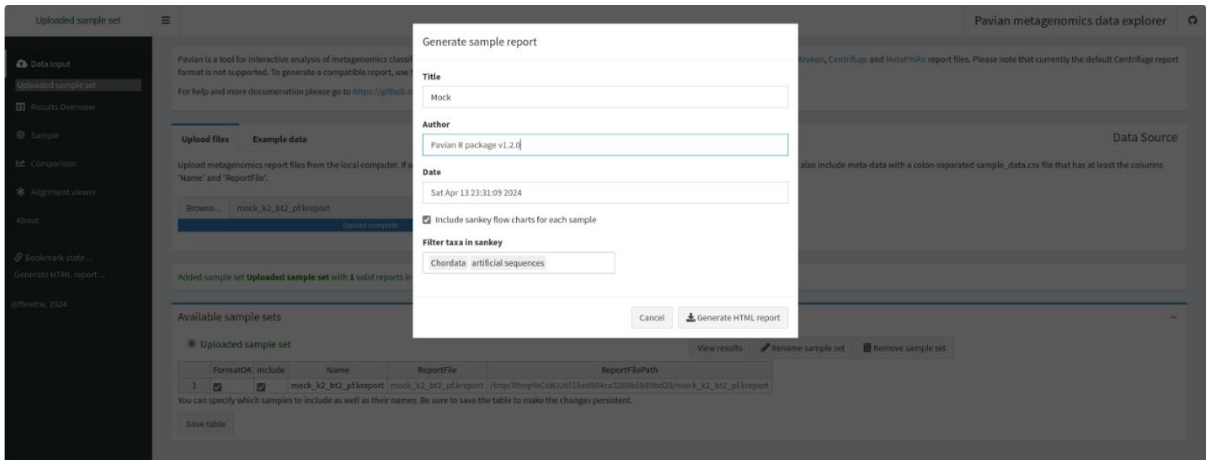
Após as análises com Bracken, realizamos o upload do arquivo *MOCK*_k2_bt2_pf_bracken.kreport para o Pavian, seguindo os passos disponíveis nas figuras 13 e 14. (< 1 min)

Figura 13: Na página da web da ferramenta Pavian, foi realizado o upload do arquivo *MOCK*_k2_bt2_pf_bracken.kreport.



Fonte: Takao et al. (2024)

Figura 14: Gerando arquivo HTML do arquivo *MOCK*_k2_bt2_pf_bracken.kreport, analisado pela ferramenta Pavian**.**



Fonte: Takao et al. (2024)

3 RESULTADOS

Todos os processos até a obtenção dos dados finais levaram menos de 1 minuto para o seu processamento completo e com um pico de uso de memória RAM de 23Gb (tabela 10).

Tabela 10: Tempo de execução e memória RAM consumida por cada ferramenta.

Etapa	Ferramenta	Tempo	Memória Ram
Remoção de Hospedeiro e Contaminante	Kraken 2	~ 6 segundos	10 Gb
	Bowtie 2	~ 40 segundos	23 Gb
Taxonomia	Kraken 2	~ 8 segundos	16 Gb
Estatísticas	Bracken	~ 2 segundos	2 Gb

Fonte: Takao et al. (2024)

Na etapa de remoção de hospedeiro e contaminantes, a ferramenta Kraken2 teve como resultado quatro arquivos, sendo dois arquivos .fastq (*MOCK*_k2_no_dog-human_un_1.fastq e *MOCK*_k2_no_dog-human_un_2.fastq), um arquivo .report e um arquivo .kraken (tabela 11). Após o processamento, é exibida na tela a mensagem contida na figura 15.

Figura 15: Mensagem exibida após a execução de ferramenta Bowtie2.

```
Loading database information... done.
500004 sequences (151.00 Mbp) processed in 1.218s (24633.8 Kseq/m, 7439.42 Mbp/m).
  386406 sequences classified (77.28%)
  113598 sequences unclassified (22.72%)
```

Fonte: Takao et al. (2024)

Demonstrando que foram analisadas 500004 sequências, onde 386406 (77,28%) foram classificadas e 113598 (22,72%) não foram classificadas.

Ainda na etapa de remoção de hospedeiro e contaminantes, tivemos a execução da ferramenta Bowtie2, onde foram criados três arquivos, sendo dois arquivos com as sequências não classificadas (*MOCK*_k2_bt2_no_dog-human_un.1.fastq e *MOCK*_k2_bt2_no_dog-human_un.2.fastq) e um arquivo .sam (tabela 11). Após a execução da ferramenta, foi exibida a mensagem contida na figura 16.

Figura 16: Mensagem exibida após a execução de ferramenta Bowtie2.

```
113598 reads; of these:
  113598 (100.00%) were paired; of these:
    108920 (95.88%) aligned concordantly 0 times
    58 (0.05%) aligned concordantly exactly 1 time
    4620 (4.07%) aligned concordantly >1 times
    ----
    108920 pairs aligned concordantly 0 times; of these:
      39 (0.04%) aligned discordantly 1 time
    ----
    108881 pairs aligned 0 times concordantly or discordantly; of these:
      217762 mates make up the pairs; of these:
        216249 (99.31%) aligned 0 times
        263 (0.12%) aligned exactly 1 time
        1250 (0.57%) aligned >1 times
4.82% overall alignment rate
```

Fonte: Takao et al. (2024)

Demonstrando que foram analisadas 113598 leituras, onde 108920 (95,88%) não foram classificadas, 58 (0,05%) tiveram o alinhamento pelo menos uma vez e 4620 (4,07%) tiveram o alinhamento mais de uma vez.

Na fase de taxonomia com a ferramenta Kraken2, foram criados dois arquivos, sendo um .kreport (*MOCK*_k2_bt2_pf.kreport), que contém toda a filogenia dos organismos encontrados, e um arquivo .kraken (*MOCK*_k2_bt2_pf.kraken), que contém todas as leituras classificadas (tabela 11). Após a execução, foi exibida a mensagem contida na figura 17.

Figura 17: Mensagem exibida após a execução de ferramenta Kraken2.

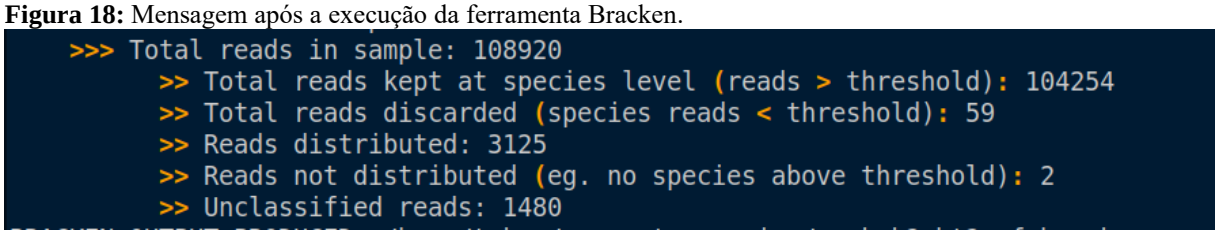
```
Loading database information... done.
108920 sequences (32.89 Mbp) processed in 0.284s (23021.6 Kseq/m, 6952.54 Mbp/m).
  107440 sequences classified (98.64%)
  1480 sequences unclassified (1.36%)
```

Fonte: Takao et al. (2024)

Demonstrando que foram analisadas 108920 sequências, onde 107440 (98,64%) foram classificadas e 1480 (1,36%) não foram classificadas.

Após a taxonomia, iniciamos as análises estatísticas, primeiro com a ferramenta Bracken, onde foi realizada a análise de abundância dos organismos classificados pelo Kraken2 (tabela 12). Após a execução da ferramenta, foi exibida a mensagem contida na figura 18.

Figura 18: Mensagem após a execução da ferramenta Bracken.



Fonte: Takao et al. (2024)

Mostrando a quantidade de leituras analisadas (108920), leituras a nível de espécie (104254), leituras descartadas (59) e não classificadas (1480).

Tabela 11: Arquivos gerados pelas ferramentas após a execução e os arquivos *MOCK* brutos a fim de comparação.

Ferramenta	Arquivo criado	Tamanho
MOCK	*MOCK*_R1.fastq	164,1 Mb
	*MOCK*_R2.fastq	164,1 Mb
Kraken 2 Remoção de contaminante e hospedeiro	*MOCK*_k2_no_dog-human_un_1.fastq	38,2 Mb
	*MOCK*_k2_no_dog-human_un_2.fastq	38,2 Mb
	*MOCK*_k2_no_dog-human.kreport	2,8 Kb
	*MOCK*_k2_no_dog-human.kraken	97,4 Mb
Bowtie 2 Remoção de contaminante e hospedeiro	*MOCK*_k2_bt2_no_dog-human_un.1.fastq	35,5 Mb
	*MOCK*_k2_bt2_no_dog-human_un.2.fastq	35,5 Mb
	*MOCK*_k2_bt2_no_dog-human_un.sam	84,1 Mb
Kraken 2 Taxonomia	*MOCK*_k2_bt2_pf.kreport	12,8 Kb
	*MOCK*_k2_bt2_pf.kraken	22,9 Mb
Bracken	*MOCK*_k2_bt2_pf_bracken.report	7,7 Kb
	*MOCK*_k2_bt2_pf.bracken	1,1kb

Fonte: Takao et al. (2024)

Tabela 12: Organismos classificados e com abundância corrigida por Bracken.

Nome	Id	Nível	Leituras K2	Leituras Add.	Novas Leitura	%
Babesia bigemina	5866	S	65399	381	65780	0.6126
Trichomonas vaginalis	5722	S	26898	23	26921	0.2507
Leishmania panamensis	5679	S	5593	519	6112	0.0569
Fusarium graminearum	5518	S	1719	497	2216	0.0206
Fusarium falciforme	195108	S	1909	77	1986	0.0185
Clostridium perfringens	1502	S	992	98	1090	0.0101
Escherichia coli	562	S	126	625	751	0.0070
Staphylococcus aureus	1280	S	63	646	709	0.0066
Pasteurella multocida	747	S	441	51	492	0.0045
Klebsiella aerogenes	548	S	218	43	261	0.0024
Serratia marcescens	615	S	95	78	173	0.0016
Staphylococcus epidermidis	1282	S	149	15	164	0.0015
Human immunodeficiency virus 1	11676	S	145	0	145	0.0013
Homo sapiens	9606	S	126	2	128	0.0011
Staphylococcus felis	46127	S	110	2	112	0.0010
Tupanvirus soda lake	2126985	S	106	0	106	0.0009

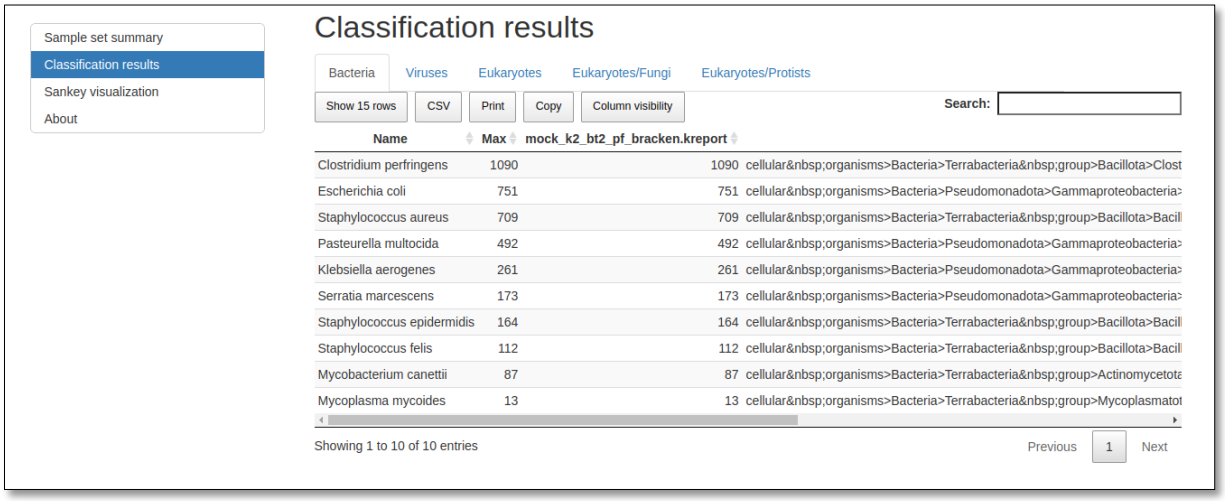
Mycobacterium canettii	78331	S	34	53	87	0.0008
Leishmania braziliensis	5660	S	76	6	82	0.0007
Aureococcus anophagefferens virus	147486	S	42	0	42	0.0003
Mycoplasma mycoides	2102	S	13	0	13	0.0001

Fonte: Takao et al. (2024)

A tabela com o nome dos organismos, número de identificação taxonômica, nível de classificação (S) espécie, leituras alinhadas por Kraken2, leituras adicionadas pelo cálculo de abundância, total de leituras e porcentagem presente no total da amostra.

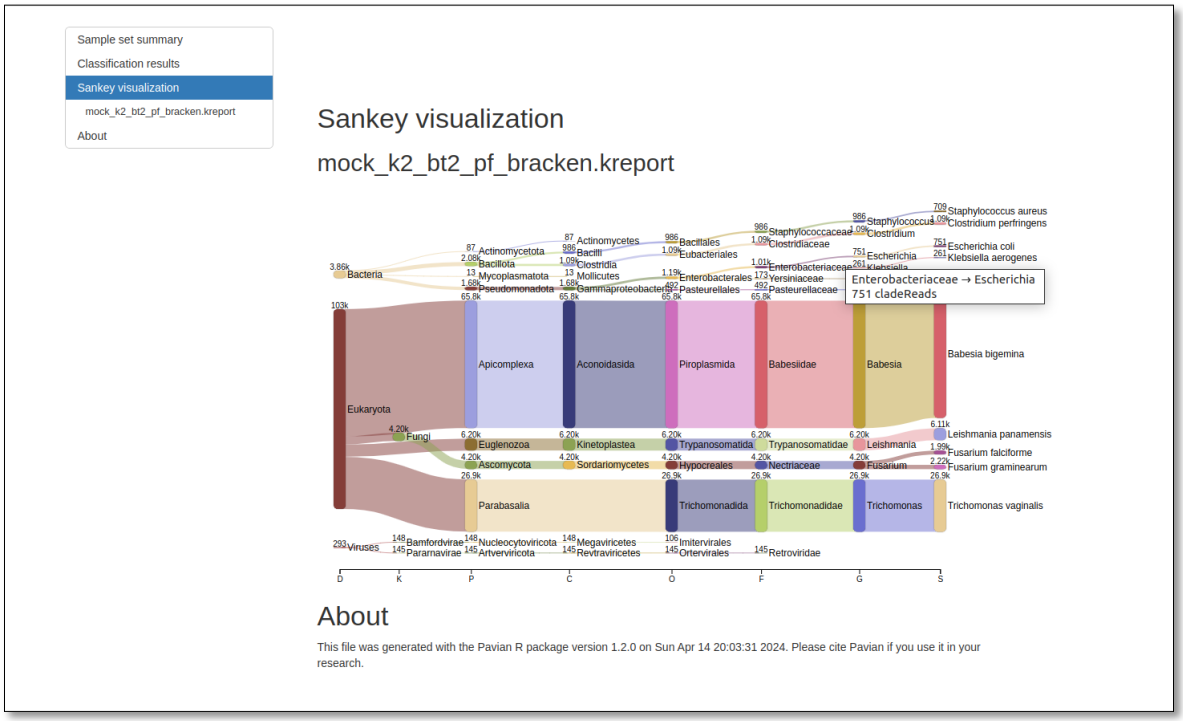
Ainda nas análises estatísticas, os dados gerados pelo Bracken, já com a abundância corrigida, foram enviados para o Pavian para que fosse feita a visualização da taxonomia (fig. 19) e da visualização gráfica (fig. 20).

Figura 19: Tabela gerada no Pavian com as bactérias classificadas e número máximo de leituras.



Fonte: Takao et al. (2024)

Figura 20: Gráfico de Sankey gerado pela ferramenta Pavian demonstrando a filogenia do arquivo *MOCK*.



Fonte: Takao et al. (2024)

4 DISCUSSÃO

O protocolo que desenvolvemos apresentou resultados comparáveis aos do protocolo descrito por Lu et al.[10], mas com uma notável melhoria no tempo de execução, sendo até cinco vezes mais rápido. Essa eficiência aprimorada pode ser atribuída principalmente ao uso preliminar da ferramenta Kraken2, antes da aplicação do Bowtie2. Interessantemente, o Bowtie2 demonstrou ser quase seis vezes mais lento que o Kraken2. Além disso, a redução no volume de dados analisados pelo Bowtie2 contribuiu significativamente para acelerar o processo. Observamos também uma diminuição na quantidade de leituras não classificadas, caindo para 1,36% em nosso protocolo, em comparação com 54,28% no protocolo proposto por Lu et al.[10] . Essa melhoria foi possível graças à utilização de bancos de dados mais abrangentes, contendo uma variedade maior de genomas de hospedeiros e contaminantes, conforme sugerido por Gihawi et al.[28,32]. Isso indica que a inclusão de bancos de dados mais completos pode aumentar significativamente a precisão da classificação taxonômica.

4.1 Relevância Científica e Impacto Prático

Os resultados deste estudo são particularmente relevantes para a metagenômica canina, uma área que tem ganhado atenção devido à importância da microbiota na saúde e nas doenças dos cães. A capacidade de remover contaminantes de forma eficiente e de classificar com precisão as sequências metagenômicas pode permitir uma melhor compreensão da microbiota canina e seu papel em diversas condições patológicas. Além disso, o protocolo proposto pode ser facilmente adaptado para outras espécies, ampliando seu impacto prático.

4.2 Limitações e Desafios

Apesar das melhorias observadas, a precisão da análise taxonômica ainda não alcançou a idealidade, com a ocorrência de alguns falsos positivos e negativos. Isso pode ser atribuído à incompletude dos bancos de dados de referência [20] e à presença de sequências altamente similares entre diferentes táxons. Futuras pesquisas devem focar na expansão e no aprimoramento dos bancos de dados de referência, bem como no desenvolvimento de algoritmos mais robustos para a classificação taxonômica.

4.3 Sugestões para Pesquisas Futuras

Para futuras investigações, sugerimos a inclusão de tecnologias de sequenciamento mais avançadas, como o sequenciamento de terceira geração, que oferece leituras mais longas e pode melhorar a resolução taxonômica. Além disso, a integração com outras técnicas ômicas, como metatranscriptômica e proteômica, pode fornecer uma visão mais abrangente das funções microbiotas e suas interações com o hospedeiro.

4.4 Integração com Outras Tecnologias

A combinação estratégica de ferramentas como Kraken2 e Bowtie2, juntamente com a utilização de bancos de dados extensivos, demonstrou ser decisiva para melhorar tanto a velocidade quanto a acurácia da classificação. No entanto, a integração com novas tecnologias de sequenciamento e análise de dados pode oferecer ainda mais precisão e eficiência. Por exemplo, a utilização de técnicas de aprendizado de máquina para a análise de dados

metagenômicos pode abrir novas possibilidades para a identificação de padrões e a predição de interações microbiotas-hospedeiro.

5 CONCLUSÃO

Neste estudo, demonstramos que a implementação do novo protocolo resultou em um avanço significativo na eficiência temporal das análises metagenômicas, evidenciando uma redução substancial no tempo de execução das ferramentas utilizadas. Tal ganho expressivo não apenas otimiza o processo de análise como um todo, mas também viabiliza a realização de estudos mais amplos e complexos com uma alocação de tempo mais eficiente. No entanto, apesar dos avanços observados, reconhecemos a existência de áreas que demandam atenção e aprimoramento contínuo. Especificamente, a questão dos falsos positivos e negativos permanece como um desafio significativo, impactando a precisão e a confiabilidade dos resultados obtidos. Portanto, enfatizamos a necessidade de esforços colaborativos entre os desenvolvedores das ferramentas para aprimorar os algoritmos e as bases de dados subjacentes. Essas melhorias são cruciais para minimizar as taxas de erro e, por consequência, aumentar a acurácia das análises metagenômicas. Além disso, a evolução das técnicas de sequenciamento e análise metagenômica continua a ser uma área de intensa pesquisa e desenvolvimento. A adoção de abordagens inovadoras e a integração de novas tecnologias serão fundamentais para superar os obstáculos existentes e explorar plenamente o potencial da metagenômica.

Em suma, o protocolo proposto representa um passo importante em direção à otimização do processamento de dados metagenômicos. Contudo, a busca por soluções para os desafios identificados deve prosseguir, com o objetivo de aprimorar continuamente as metodologias de análise e, consequentemente, expandir nosso entendimento sobre a complexidade dos ecossistemas microbianos.

6 REFERÊNCIAS:

1. Metzker ML. Sequencing technologies — the next generation. Nat Rev Genet. janeiro de 2010;11(1):31–46.

2. Handelsman J. Metagenomics: Application of Genomics to Uncultured Microorganisms. Microbiol Mol Biol Rev. dezembro de 2004;68(4):669–85.

3. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. Nature. 18 de outubro de 2007;449(7164):804–10.

4. Gilbert JA, Jansson JK, Knight R. The Earth Microbiome project: successes and aspirations. BMC Biology. 22 de agosto de 2014;12(1):69.

5. Suchodolski JS. Companion animals symposium: microbes and gastrointestinal health of dogs and cats. J Anim Sci. maio de 2011;89(5):1520–30.

6. Honneffer JB, Minamoto Y, Suchodolski JS. Microbiota alterations in acute and chronic gastrointestinal inflammation of cats and dogs. World J Gastroenterol. 28 de novembro de 2014;20(44):16489–97.

7. Kunin V, Copeland A, Lapidus A, Mavromatis K, Hugenholtz P. A Bioinformatician's Guide to Metagenomics. Microbiol Mol Biol Rev. dezembro de 2008;72(4):557–78.

8. Schmieder R, Edwards R. Fast Identification and Removal of Sequence Contamination from Genomic and Metagenomic Datasets. Rodriguez-Valera F, organizador. PLoS ONE. 9 de março de 2011;6(3):e17288.

9. Zhou R, Ng SK, Sung JJY, Goh WWB, Wong SH. Data pre-processing for analyzing microbiome data – A mini review. Computational and Structural Biotechnology Journal. 1º de janeiro de 2023;21:4804–15.

10. Lu J, Rincon N, Wood DE, Breitwieser FP, Pockrandt C, Langmead B, et al. Metagenome analysis using the Kraken software suite. Nat Protoc. dezembro de 2022;17(12):2815–39.

11. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. 3 de março de 2014;15(3):R46.

12. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 4 de março de 2012;9(4):357–9.

13. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: estimating species abundance in metagenomics data. PeerJ Computer Science. 2 de janeiro de 2017;3:e104.

14. Breitwieser FP, Salzberg SL. Pavian: interactive analysis of metagenomics data for microbiome studies and pathogen identification. Bioinformatics. 15 de fevereiro de 2020;36(4):1303–4.

15. Cock PJA, Fields CJ, Goto N, Heuer ML, Rice PM. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. Nucleic Acids Research. abril de 2010;38(6):1767–71.

16. Exogene: A performant workflow for detecting viral integrations from paired-end next-generation sequencing data | PLOS ONE [Internet]. [citado 13 de maio de 2024]. Disponível em: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0250915

17. Breitwieser FP, Pertea M, Zimin AV, Salzberg SL. Human contamination in bacterial genomes has created thousands of spurious proteins. Genome Res. junho de 2019;29(6):954–60.

18. Steinegger M, Salzberg SL. Terminating contamination: large-scale search identifies more than 2,000,000 contaminated entries in GenBank. Genome Biology. 12 de maio de 2020;21(1):115.

19. Wood D. DerrickWood/kraken2 [Internet]. 2024 [citado 7 de março de 2024]. Disponível em: https://github.com/DerrickWood/kraken2

20. Nasko DJ, Koren S, Phillippy AM, Treangen TJ. *RefSeq* database growth influences the accuracy of k-mer-based lowest common ancestor species identification. Genome Biology. 30 de outubro de 2018;19(1):165.

21. Bethesda (MD): National Library of Medicine (US). National Center for Biotechnology Information (NCBI) [Internet]. National Center for Biotechnology Information (NCBI). 1988 [citado 2 de novembro de 2023]. Disponível em: https://www.ncbi.nlm.nih.gov/

22. Yang C, Chowdhury D, Zhang Z, Cheung WK, Lu A, Bian Z, et al. A review of computational tools for generating metagenome-assembled genomes from metagenomic sequencing data. Computational and Structural Biotechnology Journal. 2021;19:6301–14.

23. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 15 de maio de 2015;31(10):1674–6.

24. Lu J. jenniferlu717/Bracken [Internet]. 2024 [citado 7 de março de 2024]. Disponível em: https://github.com/jenniferlu717/Bracken

25. Langmead B. BenLangmead/bowtie2 [Internet]. 2024 [citado 7 de março de 2024]. Disponível em: https://github.com/BenLangmead/bowtie2

26. Gourlé H, Karlsson-Lindsjö O, Hayer J, Bongcam-Rudloff E. Simulating Illumina metagenomic data with InSilicoSeq. Bioinformatics. 1º de fevereiro de 2019;35(3):521–2.

27. Sequencing Platforms | Illumina NGS platforms [Internet]. [citado 13 de maio de 2024]. Disponível em: https://www.illumina.com/systems/sequencing-platforms.html

28. Takao PS, Moura JF, Correa PS, Barbosa DA, Menegidio FB. DESENVOLVIMENTO DE BANCOS DE DADOS PARA ANÁLISE METAGENÔMICA DE MICROBIOMAS DE CÃES: UMA ABORDAGEM COM KRAKEN2 E BOWTIE2. Em: Anais do IV Congresso Brasileiro de Biotecnologia On-line [Internet]. Revista Multidisciplinar de Educação e Meio Ambiente; 2024 [citado 13 de julho de 2024]. Disponível em: https://ime.events/conbiotec2024/pdf/33826

29. *RefSeq*: NCBI Reference Sequence Database [Internet]. National Library of Medicine. 2023 [citado 5 de novembro de 2023]. Disponível em: https://www.ncbi.nlm.nih.gov/*RefSeq*/

30. The Bioconda Team, Grüning B, Dale R, Sjödin A, Chapman BA, Rowe J, et al. Bioconda: sustainable and comprehensive software distribution for the life sciences. Nat Methods. julho de 2018;15(7):475–6.

31. Installing on Linux — conda 24.5.1.dev19 documentation [Internet]. [citado 14 de maio de 2024]. Disponível em: https://docs.conda.io/projects/conda/en/latest/user-guide/install/linux.html

32. Gihawi A, Ge Y, Lu J, Puiu D, Xu A, Cooper CS, et al. Major data analysis errors invalidate cancer microbiome findings. mBio. 31 de outubro de 2023;14(5):e0160723.

ANEXO

Anexo 1: Passo a Passo para a Confecção da Biblioteca *MOCK*

Para criar a biblioteca simulada de metagenômica (*MOCK*), utilizamos a ferramenta InSilicoSeq. A seguir, apresentamos o passo a passo detalhado para a confecção da biblioteca:

1. Instalação do InSilicoSeq:

```
$ pip install insilicoseq
```

2. Criação da Biblioteca *MOCK*:

```
$ iss generate --genomes genomes.fasta --output MOCK –n reads 500000
```

3. Configuração dos Genomas:

- Os genomas utilizados foram obtidos do banco de dados RefSeq do NCBI.
- A tabela 1 lista os genomas e suas respectivas porcentagens e quantidades de leituras na biblioteca *MOCK*.

5. RESULTADOS

5.1. Artigo 6: Caracterização do perfil zoonóticos do microbioma sanguíneo, salivar e cutâneo de cães (*Canis lupus familiaris*) através de dados públicos de sequenciamento de nova geração (NGS)

CARACTERIZAÇÃO DO PERFIL ZOONÓTICO DO MICROBIOMA SANGUÍNEO, SALIVAR E CUTÂNEO DE CÃES (*Canis lupus familiaris*) ATRAVÉS DE DADOS PÚBLICOS DE SEQUENCIAMENTO DE NOVA GERAÇÃO (NGS)

Paulo Sallarola Takao[1]; Pâmela Souza Corrêa[1]; Rafaela Oliveira Campos[1]; Alessandra de Souza Fernandes[1]; Fabiana Borges[1]; David Aciole[1]; Renan Rodrigues Rocha[1]; Caio Junji Tanaka[1]; Rubens Pasa[2]; Karine Frehner Kavalco[2]; Fabiano Bezerra Menegidio[1]

1.Universidade de Mogi das Cruzes, UMC, Mogi das Cruzes, SP, Brasil

2. Universidade Federal de Viçosa, UFV, Viçosa, MG, Brasil

RESUMO

O microbioma canino, composto por comunidades complexas de microrganismos, desempenha um papel crucial na saúde animal e humana, especialmente devido ao seu potencial zoonótico. A pesquisa utilizou 36 amostras de cães, divididas entre sangue, saliva e pele, disponíveis no banco de dados Sequence Read Archive (SRA) do *NCBI*. As análises foram realizadas com ferramentas como Kraken2 e Bowtie2 para remoção de contaminantes e análise taxonômica, enquanto Bracken e Pavian foram usados para avaliação de abundância e visualização dos dados. Os resultados revelaram uma diversidade microbiana significativa, com a saliva apresentando a maior diversidade (353 gêneros), seguida pela pele (69 gêneros) e sangue (9 gêneros). A presença de gêneros como *Porphyromonas* e *Treponema* na saliva sugere um papel potencial na saúde oral e doenças periodontais em cães, similar ao observado em humanos. A análise do microbioma dérmico destacou a presença de gêneros como *Streptomyces* e Pseudomonas, sugerindo sua importância na saúde da pele canina. A baixa diversidade microbiana no sangue era esperada, refletindo a baixa ocorrência de microrganismos nesse ambiente. A presença de alguns gêneros bacterianos e protozoários pode indicar infecções subclínicas ou contaminação durante a coleta. Do ponto de vista zoonótico, a identificação de gêneros como *Pasteurella* e *Capnocytophaga* na saliva canina é relevante, dado seu potencial de causar infecções em humanos. O estudo destaca a importância de considerar fatores ambientais e comportamentais na análise do microbioma canino. Em conclusão, este estudo fornece uma visão abrangente dos microbiomas da pele, saliva e sangue em cães, com implicações significativas para a saúde animal e humana. Futuros estudos podem explorar a relação entre o microbioma canino e a saúde, bem como seu impacto na interface homem-animal.

Palavras-chave: Microbioma canino; Zoonose; Sequenciamento metagenômico; Diversidade microbiana; Saúde animal

1 INTRODUÇÃO

O microbioma canino, composto por comunidades complexas de microrganismos, desempenha um papel fundamental na saúde e na patogênese de doenças, assim como observado em humanos [1]. A compreensão do microbioma em cães (*Canis lupus familiaris*) tem implicações significativas para a saúde animal e humana, particularmente devido à sua relevância zoonótica [2]. A crescente conscientização sobre a interação dinâmica entre o microbioma e a saúde do hospedeiro levou a um aumento significativo nas pesquisas sobre microbiomas em várias espécies, incluindo cães. A relação intrincada entre o microbioma canino e a saúde do hospedeiro é evidente em diversos locais do corpo, incluindo o sangue, a saliva e a pele. O microbioma sanguíneo, embora menos explorado, oferece insights valiosos sobre infecções sistêmicas e a presença de patógenos [3]. Da mesma forma, o microbioma salivar, caracterizado por sua vasta diversidade microbiana, reflete a complexidade do ambiente oral e sua influência na saúde geral [4]. O microbioma cutâneo, conhecido por sua rica diversidade, desempenha um papel crucial na proteção contra patógenos externos e na manutenção da homeostase cutânea [5]. É importante ressaltar que, neste estudo, o foco do microbioma cutâneo reside na derme, a camada mais profunda da pele, e não na epiderme, a camada superficial. A investigação do microbioma dérmico é particularmente relevante, considerando a crescente evidência de que as comunidades microbianas residentes na derme podem influenciar a saúde da pele e contribuir para o desenvolvimento de doenças dermatológicas [6]. Em cães, doenças de pele são comuns e frequentemente desafiadoras de tratar, e a compreensão do papel do microbioma dérmico na patogênese dessas condições pode levar a novas estratégias terapêuticas.

Avanços no sequenciamento metagenômico revolucionaram a caracterização do microbioma, permitindo a identificação abrangente de comunidades microbianas. Ferramentas como Kraken2 e Bowtie2 tornaram-se componentes essenciais na análise metagenômica, facilitando a remoção de contaminantes e a avaliação taxonômica [7,8]. Essas ferramentas demonstraram eficácia na identificação de uma ampla gama de microrganismos, incluindo bactérias, vírus e fungos, em várias amostras biológicas [9].

Estudos anteriores utilizando essas técnicas destacaram a diversidade e o papel do microbioma na saúde e na doença periodontal, tanto em humanos quanto em cães. Por exemplo,

Socransky et al. (1998) identificaram complexos microbianos específicos, como o complexo vermelho, associados à periodontite em humanos [10]. Esse conceito de complexos microbianos, representando grupos de microrganismos que coexistem e interagem dentro de um nicho específico, revolucionou a compreensão da ecologia microbiana na saúde e na doença [10]. Em cães, a periodontite é uma doença oral comum e frequentemente grave que pode levar à perda de dentes e afetar a saúde sistêmica do animal. A caracterização do microbioma oral em cães com periodontite pode ajudar a identificar alvos terapêuticos e desenvolver estratégias preventivas [11].

Pesquisas mais recentes utilizando sequenciamento de rRNA 16S revelaram uma complexa interação entre o microbioma oral e fatores do hospedeiro, como sexo e status de saúde periodontal. Henne et al.[12] relataram diferenças específicas do sexo na composição do microbioma oral humano, particularmente em pacientes com periodontite grave, destacando a potencial influência dos hormônios sexuais na composição e função do microbioma[12]. Essas descobertas são particularmente relevantes considerando as diferenças observadas na prevalência, gravidade e progressão da periodontite entre homens e mulheres[13]. Em cães, a influência do sexo na saúde periodontal não foi tão extensivamente estudada quanto em humanos, mas há evidências de que fatores hormonais podem desempenhar um papel na suscetibilidade à doença periodontal em cães[14].

A influência do sexo nas respostas ao tratamento periodontal também tem sido objeto de investigação. Castro dos Santos et al.[13] conduziram uma análise retrospectiva de oito ensaios clínicos randomizados, avaliando a influência do sexo nos resultados do tratamento periodontal. Seus resultados indicaram que, embora homens tendessem a apresentar maior profundidade de sondagem e nível clínico de inserção no início do estudo, as diferenças na resposta ao tratamento não foram clinicamente relevantes [13]. Em cães, a influência do sexo na resposta ao tratamento periodontal não foi especificamente avaliada, mas é um tópico que merece investigação adicional. Além disso, a influência das condições ambientais na composição e no comportamento do microbioma oral foi reconhecida [15]. Tanaka et al.[15] demonstraram que diferentes meios de cultura e condições atmosféricas influenciaram significativamente a densidade e a diversidade bacteriana de biofilmes ex vivo de indivíduos com periodontite avançada. Essas descobertas enfatizam a adaptabilidade do microbioma oral e a necessidade de considerar fatores ambientais ao estudar comunidades microbianas[15]. A capacidade do microbioma oral de se adaptar a diferentes condições ambientais destaca a complexa interação entre fatores do hospedeiro, fatores ambientais e a composição da comunidade microbiana. Em

cães, a influência de fatores ambientais, como dieta, estilo de vida e exposição a antibióticos, no microbioma oral é um tópico de crescente interesse [16,17].

Este estudo tem como objetivo caracterizar o perfil zoonótico do microbioma sanguíneo, salivar e dérmico de cães utilizando dados de sequenciamento metagenômico disponíveis no *NCBI*. Empregando Kraken2 e Bowtie2 para remoção de contaminantes, Kraken2 para taxonomia, Bracken para análise de abundância e Pavian para visualização, pretendemos fornecer uma análise abrangente da composição e do potencial zoonótico do microbioma canino. A análise do microbioma dérmico, com foco na derme, fornecerá informações valiosas sobre as comunidades microbianas que residem nessa camada específica da pele, que é distinta da microbiota da superfície da pele. Este estudo contribuirá para a crescente base de conhecimento sobre o microbioma canino e suas implicações para a saúde animal e humana.

2 METODOLOGIA

2.1 Amostras

Este estudo utilizou dados de sequenciamento metagenômico de 36 amostras de cães (*Canis lupus familiaris*) disponíveis no banco de dados *Sequence Read Archive* (SRA) do *National Center for Biotechnology Information* (*NCBI*) (tabela 13). As amostras foram selecionadas com base na disponibilidade de metadados, incluindo sexo, tipo de amostra e raça.

Foram incluídas 12 amostras de sangue coletadas com EDTA, 12 amostras de saliva e 12 amostras de pele, ambas pertencentes ao BioProjeto PRJNA648123 referente ao programa Dog 10k. As amostras de sangue foram coletadas de cães machos e fêmeas, predominantemente sem raça definida. As amostras de saliva também foram coletadas de cães machos e fêmeas, todos sem raça definida. As amostras de pele são pertencentes ao BioProjeto PRJNA839753 on as amostras foram coletadas de fêmeas de diversas raças, incluindo Labrador Retriver, Yorkshire Terrier, Golden Retriver e Dálmata (tabela 14). É importante destacar que as amostras de pele foram coletadas da derme, a camada mais profunda da pele, e não da superfície, durante procedimentos cirúrgicos de mastectomia.

Tabela 13: Apresenta os metadados detalhados de cada amostra.

BioProjeto	SRA	Sexo	Amostra	Raça
PRJNA648123	SRR20326628	Macho	Sangue EDTA	Sem Raça Definida
PRJNA648123	SRR20326573	Macho	Sangue EDTA	Sem Raça Definida
PRJNA648123	SRR20326528	Macho	Sangue EDTA	Sem Raça Definida

PRJNA648123	SRR20326274	Macho	Sangue EDTA	Sem Raça Definida
PRJNA648123	SRR20326506	Macho	Sangue EDTA	Sem Raça Definida
PRJNA648123	SRR20326503	Fêmea	Sangue EDTA	Sem Raça Definida
PRJNA648123	SRR20326276	Macho	Sangue EDTA	Sem Raça Definida
PRJNA648123	SRR20326277	Fêmea	Sangue EDTA	Sem Raça Definida
PRJNA648123	SRR20326597	Fêmea	Sangue EDTA	Sem Raça Definida
PRJNA648123	SRR20326596	Fêmea	Sangue EDTA	Sem Raça Definida
PRJNA648123	SRR20326270	Fêmea	Sangue EDTA	Sem Raça Definida
PRJNA648123	SRR20326604	Fêmea	Sangue EDTA	Sem Raça Definida
PRJNA648123	SRR20326433	Fêmea	Saliva	Sem Raça Definida
PRJNA648123	SRR20326258	Macho	Saliva	Sem Raça Definida
PRJNA648123	SRR20326501	Macho	Saliva	Sem Raça Definida
PRJNA648123	SRR20326586	Macho	Saliva	Sem Raça Definida
PRJNA648123	SRR20326255	Macho	Saliva	Sem Raça Definida
PRJNA648123	SRR20326606	Fêmea	Saliva	Sem Raça Definida
PRJNA648123	SRR20326253	Macho	Saliva	Sem Raça Definida
PRJNA648123	SRR20326271	Fêmea	Saliva	Sem Raça Definida
PRJNA648123	SRR20326617	Fêmea	Saliva	Sem Raça Definida
PRJNA648123	SRR20326500	Macho	Saliva	Sem Raça Definida
PRJNA648123	SRR20326502	Fêmea	Saliva	Sem Raça Definida
PRJNA648123	SRR20326257	Fêmea	Saliva	Sem Raça Definida
PRJNA839753	SRR19324871	Fêmea	Pele	Sem Raça Definida
PRJNA839753	SRR19324893	Fêmea	Pele	Yorkshire Terrier
PRJNA839753	SRR19324868	Fêmea	Pele	Labrador Retriver
PRJNA839753	SRR19324861	Fêmea	Pele	Jack Russel Terrier
PRJNA839753	SRR19324870	Fêmea	Pele	Buldogue Frances
PRJNA839753	SRR19324857	Fêmea	Pele	Sheltie
PRJNA839753	SRR19324874	Fêmea	Pele	Cocker spaniel
PRJNA839753	SRR19324879	Fêmea	Pele	Labrador Retriver
PRJNA839753	SRR19324859	Fêmea	Pele	Yorkshire Terrier
PRJNA839753	SRR19324891	Fêmea	Pele	Golden Retriver
PRJNA839753	SRR19324886	Fêmea	Pele	Dálmata
PRJNA839753	SRR19324888	Fêmea	Pele	Golden Retriver

Fonte: Takao et al. (2024)

Tabela mostra o BioProjeto, código de acesso SRA, sexo, tipo de amostra e raça.

2.2 Análise de qualidade.

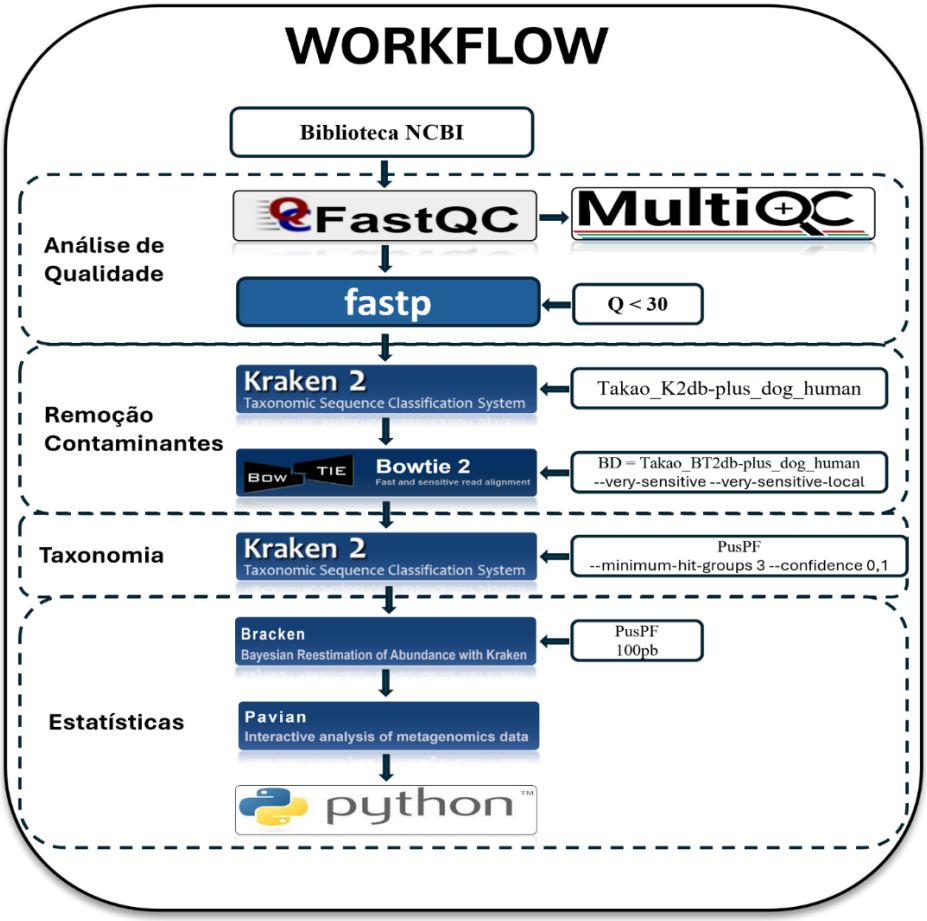
Antes do início das análises de bioinformática realizamos uma análise de qualidade das amostras, já que elas foram coletadas de dados públicos e nem sempre são de boa qualidade, para isso analisamos com a ferramenta FastQC e para visualizar utilizamos a ferramenta

MultiQC e para garantia de não haver erro realizamos o corte de 30 primeiros pares de bases com a ferramenta fastp.

2.3 Análises de Bioinformática

Após a confirmação da qualidade das amostras, iniciamos com a remoção de contaminantes e hospedeiros com as ferramentas Kraken2 e Bowtie 2 respectivamente, sendo que com Kraken2 utilizamos as configurações padrões da ferramenta e com bowtie2 utilizamos a configuração com maior sensibilidade aumentar a sensibilidade local e de ponta a ponta e ambas utilizando o banco de dados criado por Takao et al.[18] contendo múltiplos genomas do cão e do ser humano como indicado por Gihawi et al.[19].Após remover os contaminantes e hospedeiro, prosseguimos para a análise taxonômica do material contido na amostra utilizando a ferramenta Kraken2 com as configurações de sensibilidade 0,1 e mínimo de 3 grupos se sequencias e o banco de dados disponibilizado pela ferramenta PlusPF[20], que é composto por bactérias, virus arqueas e eucariotos. Após a taxonomia, para a avaliação dos dados utilizamos a ferramenta Bracken com comprimento mínimo de 100 Pb (pares de bases) para o cálculo de abundância e posteriormente a ferramenta Pavian para visualização dos dados (Figura 21). As leituras obtidas foram convertidas para porcentagem para termos uma abundância relativa, onde posteriormente realizamos o corte de leituras com presença de menos de 0,01% (ruídos) em cada amostra, então prosseguimos para análise da alfa diversidade utilizando o índice de Shannon, Características observadas e índice de Chao1 e beta diversidade utilizando a distância de Bray-Curtis, todas amostras foram analisadas a nível de Filo, Família e Gênero (figura 21).

Figura 21: Workflow demonstrando todos os passos para a análise metagenômica completa.



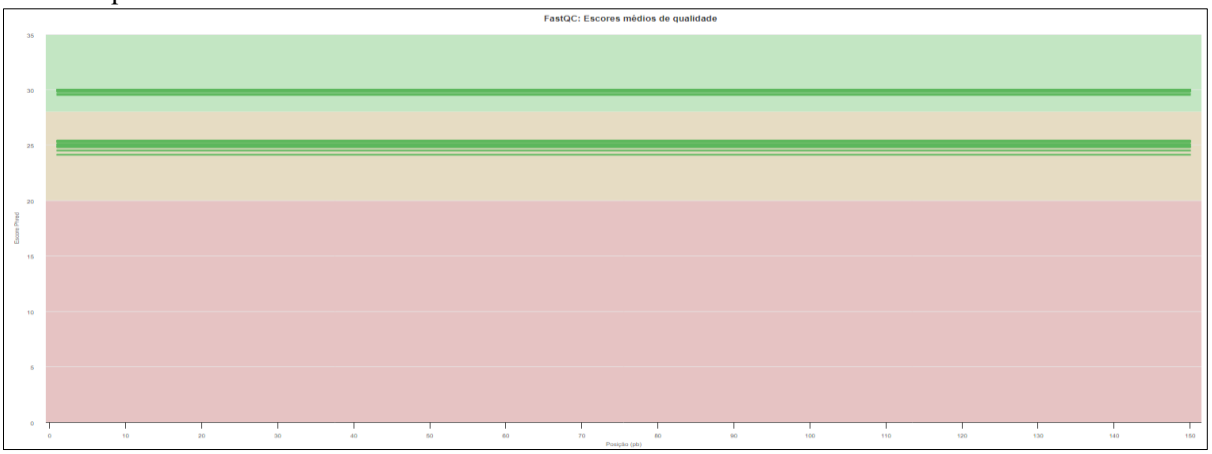
Fonte: Takao et al. (2024)

3 RESULTADOS

Todas as amostras utilizadas se mostraram com uma boa qualidade com índices satisfatórios para serem utilizadas (figura 21). Após a confirmação da qualidade das amostras, prosseguimos para as análises, onde obtivemos 69 gêneros de organismos para a pele, 353 gêneros na saliva e 9 gêneros para as amostras de sangue, demonstrados na tabela 15, proporções essas, já esperadas devido a variabilidade de organismos para cada tipo de tecido, esses resultados foram obtidos após o corte, onde eliminamos todos os organismo que estavam com uma abundância relativa menor que 0,01%, após esse corte tivemos uma redução significativa de microrganismos onde nas amostras de saliva tivemos uma redução de 81,61% do número de microrganismos, considerados ruídos ou erros de leitura. No gráfico 1 demonstramos uma comparação entre pele, saliva e sangue onde observamos a similaridade de organismos entre os tipos de amostras, onde temos apenas 4 organismos presentes nos 3 tipos de amostras, sendo elas: *Streptomyces*, *Bradyrhizobium*, *Mycoplasma* e *Plasmodium*, 47 entre

pele e saliva e 4 entre pele e sangue.

Figura 22: Histograma de qualidade de sequencias obtidos com MultiQC onde observamos um bom score de qualidade das bibliotecas obtidas no *NCBI*.



Fonte: Takao et al (2024)

Tabela 14: Demonstra o número de gêneros de organismos presentes por tecido, após a eliminação de organismos com abundância menor que 0,01% em cada amostra.

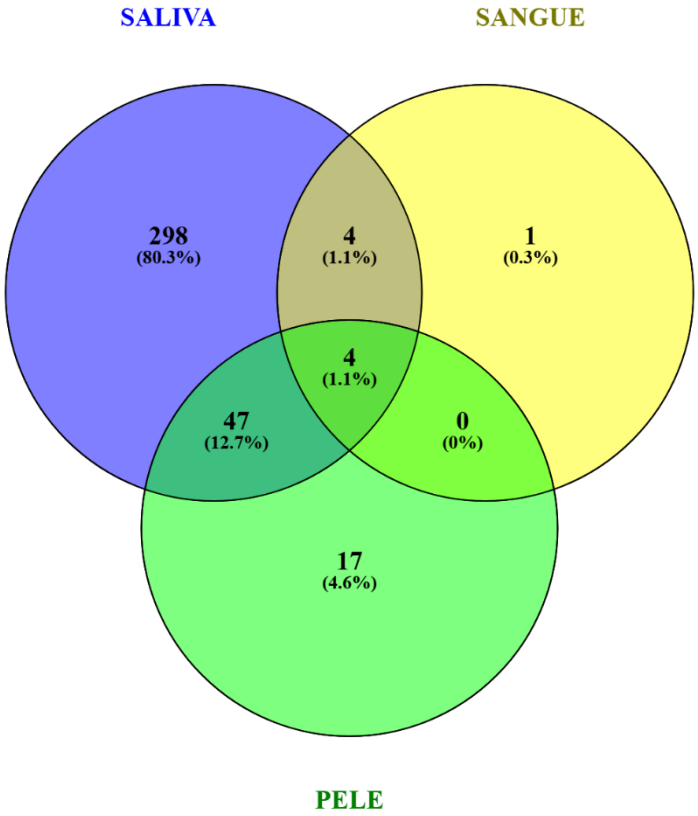
TECIDO	ABUNDÂNCIA BRUTA	ABUNDÂNCIA CORRIGIDA	REDUÇÃO
PELE	74	68	6,76%
SALIVA	1919	353	81,61%
SANGUE	9	9	0%

Tabela 3: Quantidade de microrganismos (taxas) por nível taxonômico.

Tecido	*Filo*	*Família*	*Gênero*
PELE	13	53	68
SALIVA	23	162	353
SANGUE	6	9	9

Fonte: Takao et al. (2024)

Gráfico 1: Comparativo entre as amostras de saliva, sangue e pele, onde temos apenas 4 organismos presentes nos 3 tipos de amostras, 47 entre pele e saliva e 4 entre pele e sangue



Fonte: Takao et al (2024)

3.1 Pele

Foram encontrados 13 filos, 53 famílias e 68 gêneros de organismos (anexo 1), entre os gêneros estava presente, 2 vírus (*Lentivirus* e *Thetatorquevirus*), 1 fungo (*Malassezia*), 4 protozoários (*Plasmodium, Leishimania, Dictyoestelium e Trichomonas),* sendo o plasmodium encontrado em 11 das 12 amostras e 61 gêneros bacterianos, no entanto somente 5 dessas bactérias estavam presentes em todas as amostras analisadas exceto amostra SRR1932487, sendo elas: *Streptomyces, Mycobacterium, Clostridium, Pseudomonas e Flavobacterium* . No gráfico 2 temos a comparação das amostras com os 10 gêneros mais presentes, distribuídos entre as amostras e com a análises de α-diversidade (tabela 15) e β- diversidade (gráfico 3), podemos observar a variação que ocorre em cada amostra e entre as amostras, onde 4 amostras se destacam por uma maior variação, sendo a amostra SRR19324871 a mais distinta das outras.

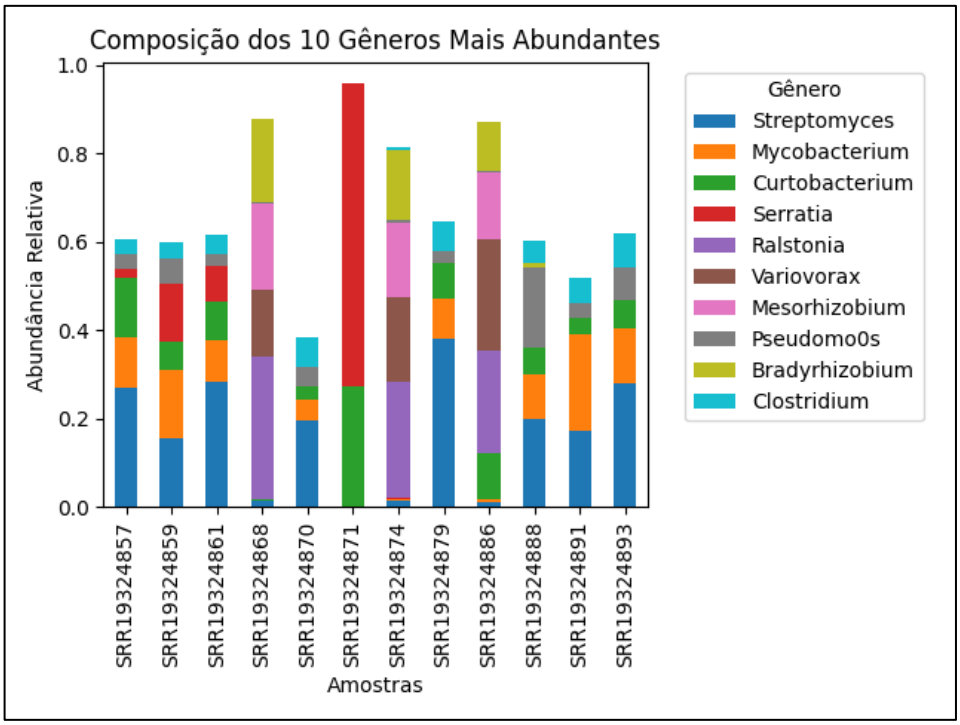
Tabela 15: Esta tabela apresenta as métricas de alfa diversidade microbiana para diferentes amostras de pele de cães.

SRR	*Shannon*	*Chao1*
SRR19324857	2,527760311	20
SRR19324859	2,623149824	19
SRR19324861	2,718654482	27
SRR19324868	1,873384258	35
SRR19324870	2,511814014	18
SRR19324871	0,745092287	3
SRR19324874	2,193598805	40
SRR19324879	2,02548564	12
SRR19324886	2,170970154	45
SRR19324888	2,667328641	25
SRR19324891	2,502471787	17
SRR19324893	2,423642522	16

Fonte: Takao et al (2024)

As métricas incluídas são o Índice de Shannon, onde quanto maior o número maior a diversidade da amostra, e o índice estimado de riqueza de espécies (Chao1). Cada linha da tabela corresponde a uma amostra identificada por um código único (SRR).

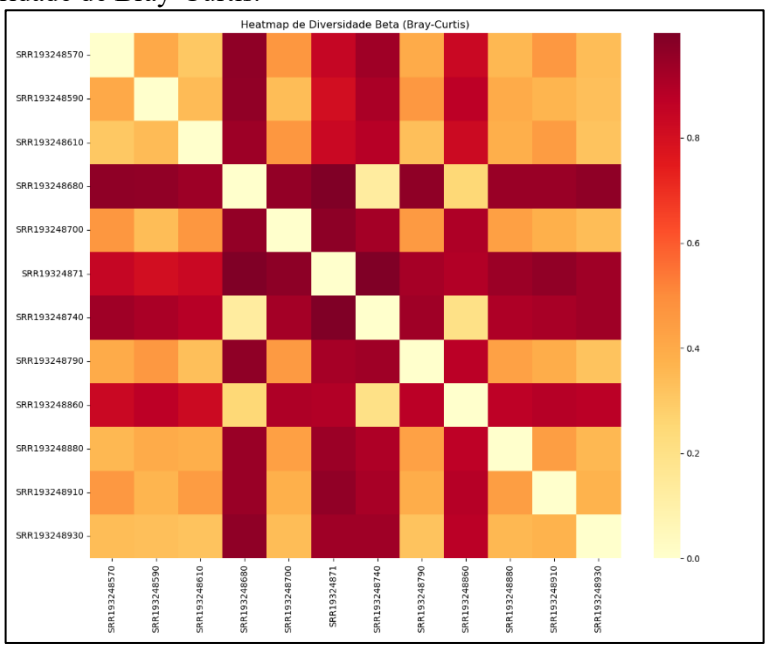
Gráfico 2: O gráfico mostra a abundância relativa dos gêneros identificados em diferentes amostras de pele.



Fonte: Takao et al (2024)

As amostras são identificadas pelos seus respectivos códigos SRR, e a abundância relativa é expressa em uma escala de 0 a 1.

Gráfico 3: O *heatmap* mostra a diversidade beta das amostras de gênero na pele de cães, utilizando o índice de dissimilaridade de Bray-Curtis.



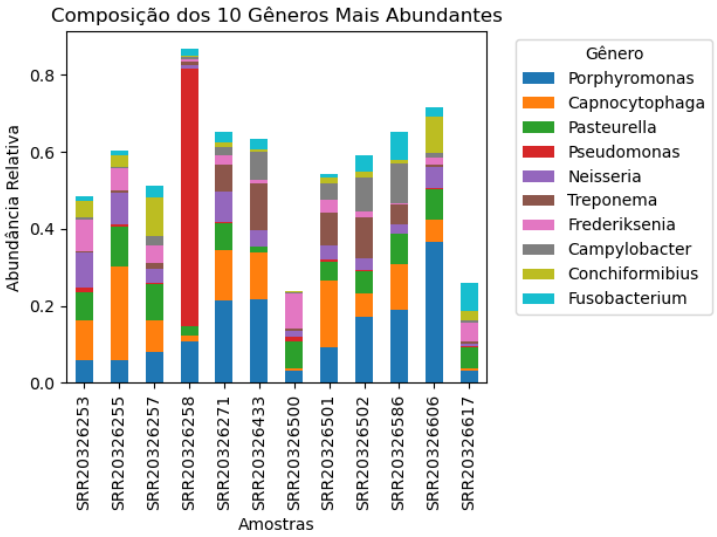
Fonte: Takao et al (2024)

Cada célula do *heatmap* representa a dissimilaridade entre um par de amostras, com cores variando de amarelo claro (alta similaridade) a vermelho (baixa similaridade).

3.4 Saliva

Foram encontrados 23 filos, 162 famílias e 353 gêneros de organismos (anexo 2), sendo 6 fungos (*Fusarium, Colletotrichum, Pyricularia, Thermothelomyces, Aspergillus, Sporisorium*), 4 protozoários (*Dictyostelium, Trichomonas, Leishmania e Plasmodium*), 2 arqueas (*Nanosynbacter* e *Natrialba*), 1 vírus (*Taupapillomavirus*) e 342 bactérias sendo 68 delas, presentes em 100% das amostras analisadas sendo os 10 gêneros mais presentes: *Porphyromonas* (14%), *Capnocytophaga* (10%), *Pasteurella* (7%), *Pseudomonas* (6%), Neisseria (4%), Treponema (4%), *Frederiksenia* (4%), *Campylobacter* (3%), *Conchiformibius* (3%) e *Fusobacterium* (3%). No gráfico 4 temos a comparação das amostras com os 10 gêneros mais presentes, distribuídos entre as amostras e com a análises de α-diversidade (tabela 16) e β- diversidade (gráfico 5), podemos observar a variação que ocorre em cada amostra e entre as amostras, onde 2 amostras se destacam por uma maior variação em relação as outras.

Gráfico 4: O gráfico mostra a abundância relativa dos gêneros identificados em diferentes amostras de saliva.



Fonte: Takao et al (2024)

As amostras são identificadas pelos seus respectivos códigos SRR, e a abundância relativa é expressa em uma escala de 0 a 1.

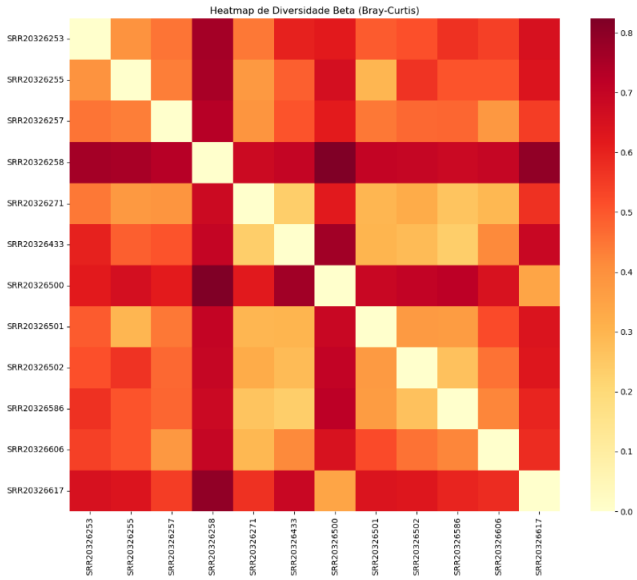
Tabela 16: Esta tabela apresenta as métricas de alfa diversidade microbiana para diferentes amostras de saliva de cães.

	SHANNON	CHAO1
SRR20326253	3,581201	136
SRR20326255	3,323539	348
SRR20326257	3,416411	158
SRR20326258	1,674938	349
SRR20326271	3,398752	349
SRR20326433	3,295347	350
SRR20326500	3,632701	258
SRR20326501	3,697189	351
SRR20326502	3,477164	182
SRR20326586	3,285013	348
SRR20326606	2,908107	347
SRR20326617	4,053148	304

Fonte: Takao et al (2024)

As métricas incluídas são o Índice de Shannon e o índice estimado de riqueza de espécies (Chao1). Cada linha da tabela corresponde a uma amostra identificada por um código único (SRR).

Gráfico 5: O *heatmap* mostra a diversidade beta das amostras de gênero da saliva de cães, utilizando o índice de dissimilaridade de Bray-Curtis.



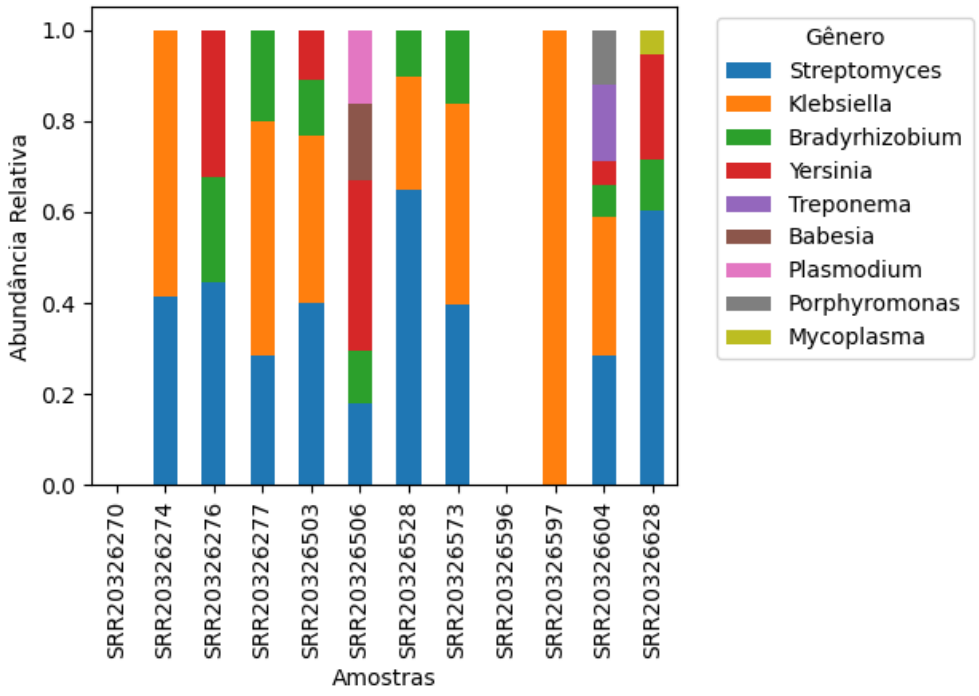
Fonte: Takao et al (2024)

Cada célula do *heatmap* representa a dissimilaridade entre um par de amostras, com cores variando de amarelo claro (alta similaridade) a vermelho (baixa similaridade).

3.5 Sangue

Foram encontrados 6 filos, 9 famílias e 9 gêneros de organismos (anexo 3), sendo 7 generos de bactérias (*Klebsiella, Streptomyces, Yersinia, Bradyrhizobium, Treponema, Porphyromonas e Mycoplasma*) e 2 protozoários (*Babesia e Plasmodium*), porém em 2 amostras não se teve nenhuma leitura e em uma amostra apenas 1 leitura, dados esses, que já eram esperados, devido à baixa ocorrência de microrganismos no sangue. No gráfico 6 temos a comparação das amostras com seus respectivos organismos presentes, a análises de α-diversidade (tabela 17) e β- diversidade (gráfico 7), podemos observar a variação que ocorre em cada amostra e entre as amostras, onde 2 amostras se destacam por uma maior variação em relação as outras.

Gráfico 6: O gráfico mostra a abundância dos gêneros identificados em diferentes amostras de sangue.



Fonte: Takao et al (2024)

As amostras são identificadas pelos seus respectivos códigos SRR, e a abundância relativa é expressa em uma escala de 0 a 1.

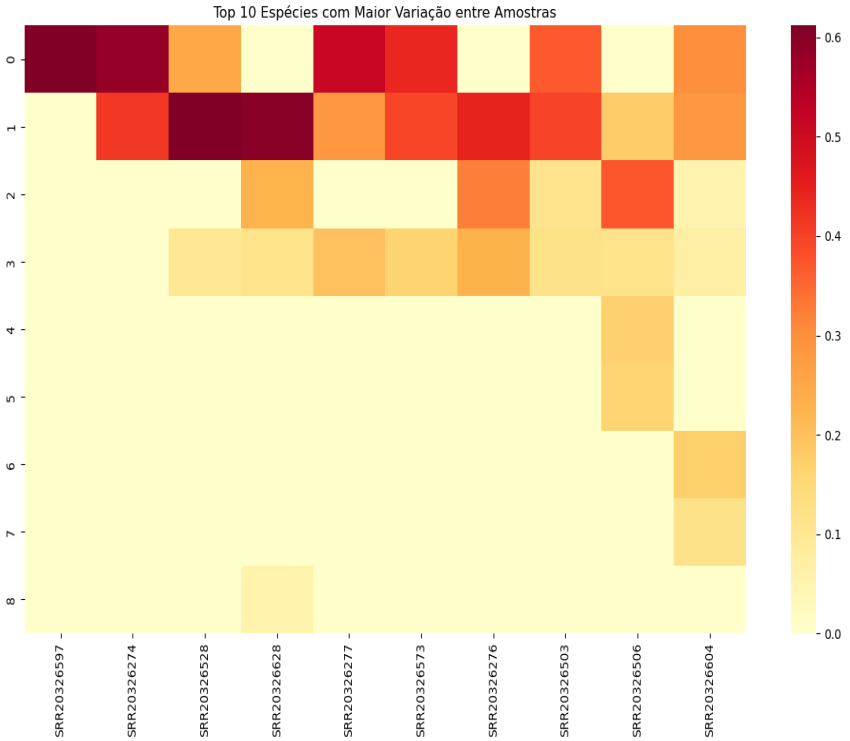
Tabela 17 Esta tabela apresenta as métricas de alfa diversidade microbiana para diferentes amostras de sangue de cães.

SRR	*Shannon*	*Chao1*
SRR20326270	0,00	0
SRR20326274	0,68	2
SRR20326276	1,06	3
SRR20326277	1,02	3
SRR20326503	1,23	4
SRR20326506	1,52	5
SRR20326528	0,86	3
SRR20326573	1,02	3
SRR20326596	0,00	0
SRR20326597	0,00	1
SRR20326604	1,61	6
SRR20326628	1,05	4

Fonte: Takao et al (2024)

As métricas incluídas são o Índice de Shannon e o índice estimado de riqueza de espécies (Chao1). Cada linha da tabela corresponde a uma amostra identificada por um código único (SRR).

Gráfico 7: O *heatmap* mostra a diversidade beta das amostras de gênero de sangue de cães.



Fonte: Takao et al (2024)

Utiliza índice de dissimilaridade de Bray-Curtis. Cada célula do *heatmap* representa a dissimilaridade entre um par de amostras, com cores variando de amarelo claro (alta similaridade) a vermelho (baixa similaridade).

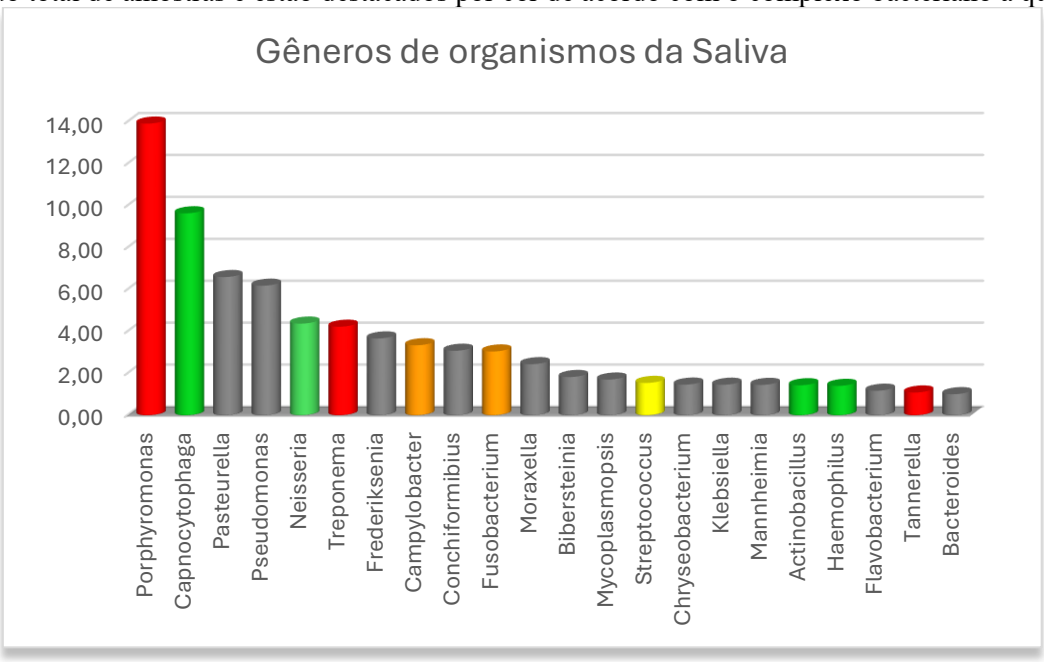
4.DISCUSSÃO

Os resultados deste estudo revelam diferenças significativas na diversidade e composição dos microbiomas da pele, saliva e sangue em cães. A saliva apresentou a maior diversidade microbiana, com 354 gêneros identificados, seguida pela pele com 69 gêneros e o sangue com apenas 9 gêneros. Essa variação na diversidade microbiana entre os diferentes locais do corpo é consistente com estudos anteriores em humanos e animais, refletindo as diferentes condições ambientais e funções de cada local.

Comparando com o estudo de Socransky et al. (1998), que identificou complexos microbianos específicos associados à periodontite em humanos, como o complexo vermelho composto por *Porphyromonas gingivalis*, *Tannerella forsythia* e *Treponema denticola*, observamos que gêneros relacionados, como *Porphyromonas* e *Treponema*, também estão presentes na saliva canina. Isso sugere que, assim como em humanos, esses gêneros podem desempenhar um papel importante na saúde oral e na doença periodontal em cães.

A análise metagenômica das amostras de saliva revelou uma distribuição interessante dos gêneros bacterianos mais abundantes (gráfico 8). Os gêneros com abundância relativa superior a 1% incluem *Porphyromonas* (13,91%), *Capnocytophaga* (9,63%), *Pasteurella* (6,59%), *Pseudomonas* (6,18%), *Neisseria* (4,37%), *Treponema* (4,22%), *Frederiksenia* (3,66%), *Campylobacter* (3,34%), *Conchiformibius* (3,08%), e *Fusobacterium* (3,04%). Esta distribuição é particularmente interessante quando comparada aos complexos microbianos descritos por Socransky et al. (1998) em humanos. Por exemplo, *Porphyromonas* e *Treponema*, que são membros do complexo vermelho associado à periodontite em humanos, também são encontrados em abundância significativa na saliva canina. Da mesma forma, *Fusobacterium*, parte do complexo laranja em humanos, também está presente em níveis notáveis. A presença de *Capnocytophaga* e *Neisseria*, que são frequentemente associadas ao complexo verde em humanos, sugere que esses gêneros também podem desempenhar papéis importantes na ecologia microbiana oral canina. É importante notar a presença significativa de *Pasteurella* e *Pseudomonas*, que não são tipicamente associadas aos complexos microbianos orais humanos, mas podem ter implicações zoonóticas importantes. Estes resultados destacam tanto as semelhanças quanto as diferenças entre os microbiomas orais canino e humano, e sugerem a necessidade de mais pesquisas para entender as implicações dessas comunidades microbianas para a saúde oral canina e o potencial de transmissão zoonótica.

Gráfico 8: Esse gráfico demonstras os microrganismos que possuem abundancia relativa maior que 1% do total de amostras e estão destacados por cor de acordo com o complexo bacteriano a qual pertence.



Fonte: Takao et al. (2024)

A metagenômica desempenhou um papel crucial na identificação de uma ampla gama de microrganismos, muitos dos quais são de difícil cultivo. Na análise de amostras de saliva, identificamos 354 gêneros de microrganismos, incluindo fungos, protozoários, arqueas, vírus e bactérias. Entre os gêneros bacterianos, 68 estavam presentes em 100% das amostras analisadas. Esses dados são consistentes com estudos realizados por Esturjão, Costa e Weese[21], Ruparell et al.[22], Tóth et al.[23].

A metagenômica, portanto, oferece uma vantagem significativa ao permitir a identificação de microrganismos que são difíceis ou impossíveis de cultivar em laboratório. Isso é particularmente importante para a compreensão da ecologia microbiana e das interações complexas entre microrganismos e seus hospedeiros, sendo isso demonstrado na tabela 19.

Tabela 18:A tabela a seguir ilustra a dificuldade de cultivo de alguns desses organismos, que requerem condições específicas ou anaerobiose, dificultando sua identificação por métodos tradicionais

Organismo	***Meio de Cultivo***	***Condições***	***Nível de Dificuldade***
Pasteurella	Ágar sangue ou ágar chocolate	Aeróbicas, 35-37°C	Baixo
Pseudomonas	Ágar cetrimida ou ágar MacConkey	Aeróbicas, 35-37°C	Baixo
Neisseria	Ágar Thayer-Martin ou ágar chocolate	5-10% CO2, 35-37°C	Médio
Capnocytophaga	Ágar chocolate ou ágar sangue	5-10% CO2, 35-37°C	Médio
Frederiksenia	Ágar sangue ou ágar chocolate	Microaerofílicas (5-10% CO2), 35-37°C	Médio
Campylobacter	Ágar Skirrow ou ágar CCDA	Microaerofílicas (5% O2, 10% CO2, 85% N2), 42°C	Médio
Porphyromonas	Ágar sangue enriquecido com hemina e menadiona	Anaeróbicas	Alto
Fusobacterium	Ágar sangue anaeróbico ou ágar Brucella suplementado	Anaeróbicas, 35-37°C	Alto
Treponema	Meio de Barbour-Stoenner-Kelly (BSK)	Microaerofílicas, 30-34°C	Alto
Conchiformibius	Ágar sangue ou ágar chocolate	Microaerofílicas (5-10% CO2), 35-37°C	Alto

Fonte: Takao et al (2024)

O foco na derme para a análise do microbioma cutâneo revelou a presença de gêneros como *Streptomyces*, *Mycobacterium* e *Pseudomonas* em todas as amostras de pele (exceto uma), sugerindo que esses microrganismos podem ser componentes importantes do microbioma dérmico canino. Isso contrasta com estudos da microbiota da superfície da pele, que geralmente encontram uma predominância de gêneros como *Staphylococcus* e *Corynebacterium*.

A baixa diversidade encontrada no sangue era esperada, dado que em estudos com ser humanos o baixa quantidade de microrganismos foram encontradas no sangue[24–27] . A presença de alguns gêneros bacterianos e protozoários no sangue pode indicar infecções subclínicas ou contaminação durante a coleta. Estes resultados são semelhantes aos relatados por Scarsella et al.[24], que também encontraram uma baixa diversidade microbiana no sangue de cães saudáveis

[24]. Em um estudo comandado por Tan et al.[27], onde foram analisados o microbioma de 9770 pessoas saldáveis, puderam identificar que não existe um microbioma comum entre os seres humanos, então podemos presumir que o mesmo ocorre com os cães [27].

Apenas quatro gêneros (*Streptomyces*, *Bradyrhizobium*, *Mycoplasma* e *Plasmodium*) foram encontrados nos três tipos de amostras, destacando a especificidade dos microbiomas em diferentes locais do corpo. A maior sobreposição foi observada entre pele e saliva, o que pode ser explicado pela exposição compartilhada a fatores ambientais e comportamentos como lambedura.

Do ponto de vista zoonótico, a presença de gêneros como *Pasteurella*, *Pseudomonas* e *Capnocytophaga* na saliva canina é particularmente relevante. *Pasteurella multocida*, por exemplo, é conhecida por causar infecções em humanos após mordidas de cães[28]. *Pseudomonas aeruginosa* é um patógeno oportunista que pode afetar tanto humanos quanto animais[29]. *Capnocytophaga canimorsus*, associada a mordidas de cães, pode causar septicemia em humanos[30].

É importante notar que o estudo utilizou dados públicos de sequenciamento, o que pode introduzir variabilidade devido a diferentes protocolos de coleta e processamento de amostras. Além disso, o foco na derme para as amostras de pele limita a comparação direta com estudos que analisam a microbiota da superfície da pele.

Este estudo fornece uma base para futuras investigações sobre o microbioma canino e suas implicações zoonóticas. Pesquisas adicionais poderiam explorar a relação entre o microbioma oral e a saúde periodontal em cães, expandindo o trabalho de Socransky et al. (1998) para o contexto canino, bem como o papel do microbioma dérmico na saúde da pele canina e em doenças dermatológicas. Além disso, a influência de fatores como raça, idade e dieta na composição do microbioma canino e o potencial zoonótico de microrganismos específicos identificados neste estudo são áreas promissoras para futuras pesquisas.

5 CONCLUSÃO

Em conclusão, este estudo fornece uma visão abrangente dos microbiomas da pele, saliva e sangue em cães, destacando a diversidade e especificidade desses ecossistemas microbianos. Os resultados têm implicações importantes para a saúde animal e humana, especialmente considerando o potencial zoonótico de certos microrganismos identificados.

Futuros estudos podem se basear nesses achados para aprofundar nossa compreensão da relação entre o microbioma canino e a saúde, bem como seu impacto na interface homem-animal.

REFERÊNCIAS

1. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. Nature. 18 de outubro de 2007;449(7164):804–10.

2. Song SJ, Lauber C, Costello EK, Lozupone CA, Humphrey G, Berg-Lyons D, et al. Cohabiting family members share microbiota with one another and with their dogs. Weigel D, organizador. eLife. 16 de abril de 2013;2:e00458.

3. Petersen C, Round JL. Defining dysbiosis and its influence on host immunity and disease. Cellular Microbiology. julho de 2014;16(7):1024–33.

4. FE D, EA K, EC T, JM B, T C, L M, et al. The canine oral microbiome. PloS one. 2012;7(4):e36067.

5. Grice EA, Segre JA. The skin microbiome. Nat Rev Microbiol. abril de 2011;9(4):244–53.

6. Fitz-Gibbon S, Tomida S, Chiu BH, Nguyen L, Du C, Liu M, et al. Propionibacterium acnes strain populations in the human skin microbiome associated with acne. The Journal of investigative dermatology. 2013;133(9):2152.

7. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. 3 de março de 2014;15(3):R46.

8. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. abril de 2012;9(4):357–9.

9. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: estimating species abundance in metagenomics data. PeerJ Computer Science. 2 de janeiro de 2017;3:e104.

10. Socransky SS, Haffajee AD, Cugini MA, Smith C, Kent RL. Microbial complexes in subgingival plaque. Journal of Clinical Periodontology. fevereiro de 1998;25(2):134–44.

11. Cunha E, Tavares L, Oliveira M. Revisiting Periodontal Disease in Dogs: How to Manage This New Old Problem? Antibiotics 2022, Vol 11, Page 1729. dezembro de 2022;11(12):1729.

12. Henne K, Schilling H, Stoneking M, Conrads G, Horz HP. Sex-specific differences in the occurrence of Fusobacterium nucleatum subspecies and Fusobacterium periodonticum in the oral cavity. Oncotarget. abril de 2018;9(29):20631.

13. Santos NC dos, Westphal MR, Retamal-Valdes B, Duarte PM, Figueiredo LC, Faveri M, et al. Influence of gender on periodontal outcomes: A retrospective analysis of eight randomized clinical trials. Journal of Periodontal Research [Internet]. 2024; Disponível em: https://onlinelibrary.wiley.com/doi/full/10.1111/jre.13272

https://onlinelibrary.wiley.com/doi/abs/10.1111/jre.13272
https://onlinelibrary.wiley.com/doi/10.1111/jre.13272

14. Steffens JP, Herrera BS, Coimbra LS, Stephens DN, Rossa C, Spolidorio LC, et al. Testosterone regulates bone response to inflammation. Hormone and Metabolic Research. 2014;46(3):193–200.

15. Junji Tanaka C, Figueiredo L, Trovisco P, Teles F, Lisboa T, Feres M, et al. Evaluation of the influence of Different Culture Media and Atmospheric Conditions on ex vivo Biofilms of Individuals with Advanced Periodontitis. Journal of the International Academy of Periodontology. junho de 2021;23:246.

16. Bell SE, Nash AK, Zanghi BM, Otto CM, Perry EB. An Assessment of the Stability of the Canine Oral Microbiota After Probiotic Administration in Healthy Dogs Over Time. Frontiers in Veterinary Science. setembro de 2020;7:565881.

17. Sinkko H, Lehtimäki J, Lohi H, Ruokolainen L, Hielm-Björkman A. Distinct healthy and atopic canine gut microbiota is influenced by diet and antibiotics. Royal Society Open Science [Internet]. abril de 2023;10(4). Disponível em: /pmc/articles/PMC10130713/ /pmc/articles/PMC10130713/?report=abstract https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10130713/

18. Takao PS, Moura JF, Correa PS, Barbosa DA, Menegidio FB. DESENVOLVIMENTO DE BANCOS DE DADOS PARA ANÁLISE METAGENÔMICA DE MICROBIOMAS DE CÃES: UMA ABORDAGEM COM KRAKEN2 E BOWTIE2. Anais do IV Congresso Brasileiro de Biotecnologia On-line. 18 de março de 2024;5(2):170.

19. Gihawi A, Ge Y, Lu J, Puiu D, Xu A, Cooper CS, et al. Major data analysis errors invalidate cancer microbiome findings. mBio. 31 de outubro de 2023;14(5):e0160723.

20. Lu J, Salzberg SL. Removing contaminants from databases of draft genomes. Sun F, organizador. PLoS Comput Biol. 25 de junho de 2018;14(6):e1006277.

21. A S, JW S, MC C, JS W. Metagenomic analysis of the canine oral cavity as revealed by high-throughput pyrosequencing of the {16S} {rRNA} gene. Veterinary microbiology. 2013;162(2):891–8.

22. Ruparell A, Inui T, Staunton R, Wallis C, Deusch O, Holcombe LJ. The canine oral microbiome: variation in bacterial populations across different niches. BMC Microbiol. 28 de fevereiro de 2020;20(1):42.

23. Tóth AG, Tóth I, Rózsa B, Dubecz A, Patai ÁV, Németh T, et al. Canine Saliva as a Possible Source of Antimicrobial Resistance Genes. Antibiotics. novembro de 2022;11(11):1490.

24. Scarsella E, Meineri G, Sandri M, Ganz HH, Stefanon B. Characterization of the {Blood} {Microbiome} and {Comparison} with the {Fecal} {Microbiome} in {Healthy} {Dogs} and {Dogs} with {Gastrointestinal} {Disease}. Veterinary Sciences. 2023;10(4):277.

25. Castillo DJ, Rifkin RF, Cowan DA, Potgieter M. The {Healthy} {Human} {Blood} {Microbiome}: {Fact} or {Fiction}? Front Cell Infect Microbiol [Internet]. 2019;9. Disponível em: https://www.frontiersin.org/journals/cellular-and-infection-microbiology/articles/10.3389/fcimb.2019.00148/full

26. Liao H, Zhang Y, Guo W, Wang X, Wang H, Ye H, et al. Characterization of the Blood and Cerebrospinal Fluid Microbiome in Children with Bacterial Meningitis and Its Potential Correlation with Inflammation. mSystems. 29 de junho de 2021;6(3):e0004921.

27. Tan CCS, Ko KKK, Chen H, Liu J, Loh M, Chia M, et al. No evidence for a common blood microbiome based on a population study of 9,770 healthy humans. Nat Microbiol. maio de 2023;8(5):973–85.

28. Wilson BA, Ho M. Pasteurella multocida: from zoonosis to cellular microbiology. Clin Microbiol Rev. julho de 2013;26(3):631–55.

29. Gellatly SL, Hancock REW. Pseudomonas aeruginosa: new insights into pathogenesis and host defenses. Pathog Dis. abril de 2013;67(3):159–73.

30. Brenner DJ, Hollis DG, Fanning GR, Weaver RE. Capnocytophaga canimorsus sp. nov. (formerly CDC group DF-2), a cause of septicemia following dog bite, and C. cynodegmi sp. nov., a cause of localized wound infection following dog bite. J Clin Microbiol. fevereiro de 1989;27(2):231–5.

ANEXOS

Anexo 1: Dados de Pele.

nome	SRR193 24857	SRR193 24859	SRR193 24861	SRR193 24868	SRR193 24870	SRR193 24871	SRR193 24874	SRR193 24879	SRR193 24886	SRR193 24888	SRR193 24891	SRR193 24893
Acidovorax	0,000	0,000	0,000	0,001	0,000	0,000	0,001	0,000	0,001	0,000	0,000	0,000
Acinetobacter	0,020	0,033	0,020	0,000	0,042	0,000	0,006	0,027	0,002	0,022	0,000	0,056
Actinomyces	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000
Aliarcobacter	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000
Aromatoleum	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,005	0,000	0,000	0,000
Bacillus	0,019	0,031	0,036	0,000	0,066	0,000	0,003	0,000	0,012	0,028	0,072	0,047
Bacteroides	0,000	0,015	0,011	0,001	0,017	0,000	0,000	0,022	0,000	0,010	0,000	0,000
Brachybacterium	0,000	0,000	0,000	0,000	0,000	0,000	0,003	0,000	0,000	0,000	0,000	0,000
Bradyrhizobium	0,000	0,000	0,000	0,187	0,000	0,000	0,159	0,000	0,110	0,011	0,000	0,000
Brevundimo0s	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000
Burkholderia	0,000	0,000	0,009	0,003	0,000	0,000	0,003	0,000	0,002	0,000	0,000	0,000
Campylobacter	0,000	0,016	0,000	0,001	0,019	0,000	0,002	0,000	0,001	0,000	0,020	0,000
Chryseobacterium	0,016	0,018	0,010	0,001	0,016	0,000	0,002	0,000	0,001	0,010	0,017	0,027
Clostridium	0,032	0,038	0,044	0,002	0,069	0,000	0,007	0,065	0,002	0,049	0,056	0,079
Comamo0s	0,000	0,000	0,000	0,002	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000
Corynebacterium	0,000	0,015	0,013	0,001	0,000	0,000	0,022	0,000	0,002	0,010	0,000	0,000
Cupriavidus	0,000	0,000	0,000	0,004	0,000	0,000	0,003	0,000	0,004	0,000	0,000	0,000
Curtobacterium	0,136	0,062	0,088	0,001	0,031	0,274	0,000	0,082	0,102	0,058	0,039	0,064
Cutibacterium	0,037	0,094	0,020	0,003	0,228	0,000	0,003	0,000	0,010	0,000	0,000	0,035
Delftia	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000
Dictyosteliu m	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000
Dietzia	0,000	0,013	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Escherichia	0,000	0,000	0,000	0,003	0,000	0,000	0,000	0,000	0,002	0,000	0,000	0,000
Flavobacterium	0,020	0,038	0,026	0,002	0,040	0,000	0,003	0,049	0,002	0,030	0,056	0,043
Fusobacterium	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,023	0,000
Haemophilus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,026	0,000
Hydrogenophaga	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,001	0,000	0,000	0,000
Lactobacillus	0,015	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Leishmania	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,009	0,000	0,000
Lentivirus	0,000	0,000	0,000	0,000	0,000	0,041	0,000	0,000	0,000	0,000	0,000	0,000
Malassezia	0,000	0,000	0,000	0,001	0,024	0,000	0,000	0,000	0,001	0,000	0,000	0,000
Mesorhizobium	0,000	0,000	0,000	0,195	0,000	0,000	0,169	0,000	0,150	0,000	0,000	0,000
Methylobacterium	0,000	0,000	0,016	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Microbacterium	0,017	0,000	0,011	0,000	0,000	0,000	0,002	0,000	0,000	0,008	0,000	0,000
Micromonospora	0,000	0,013	0,012	0,000	0,000	0,000	0,001	0,000	0,001	0,011	0,000	0,000

Moraxella	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,001	0,000	0,000	0,000
Mycobacterium	0,112	0,156	0,094	0,003	0,050	0,000	0,004	0,092	0,009	0,101	0,217	0,126
Mycolicibacterium	0,000	0,000	0,012	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Mycoplasma	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,017
Mycoplasmopsis	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,019
Neisseria	0,019	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000
Nocardia	0,000	0,000	0,009	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000
Nocardioides	0,031	0,000	0,045	0,000	0,000	0,000	0,002	0,000	0,001	0,011	0,000	0,000
Novosphingobium	0,000	0,000	0,000	0,001	0,000	0,000	0,006	0,000	0,001	0,000	0,000	0,000
Paenibacillus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,011	0,000	0,013
Parabacteroides	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,024
Paraburkholderia	0,000	0,000	0,000	0,001	0,000	0,000	0,002	0,000	0,002	0,000	0,000	0,000
Plasmodium	0,036	0,040	0,033	0,001	0,062	0,000	0,005	0,044	0,002	0,042	0,056	0,061
Polaribacter	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,010	0,000	0,000
Pseudomo0s	0,033	0,056	0,027	0,004	0,041	0,000	0,005	0,027	0,005	0,183	0,033	0,073
Pseudonocardia	0,000	0,000	0,014	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Ralstonia	0,000	0,000	0,000	0,324	0,000	0,000	0,262	0,000	0,233	0,000	0,000	0,000
Rhizobium	0,000	0,000	0,000	0,001	0,000	0,000	0,002	0,000	0,001	0,000	0,000	0,000
Rothia	0,024	0,000	0,000	0,002	0,018	0,000	0,003	0,020	0,002	0,024	0,020	0,000
Segatella	0,021	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Serratia	0,020	0,132	0,081	0,000	0,000	0,685	0,002	0,000	0,001	0,000	0,000	0,000
Sphingobium	0,000	0,000	0,000	0,005	0,000	0,000	0,009	0,000	0,004	0,000	0,000	0,000
Sphingomo0s	0,000	0,000	0,020	0,072	0,000	0,000	0,071	0,000	0,045	0,000	0,018	0,000
Sphingopyxis	0,000	0,000	0,000	0,002	0,000	0,000	0,002	0,000	0,002	0,000	0,000	0,000
Staphylococcus	0,000	0,043	0,025	0,006	0,033	0,000	0,015	0,154	0,007	0,018	0,083	0,037
Stenotrophomo0s	0,000	0,000	0,023	0,000	0,000	0,000	0,003	0,039	0,001	0,000	0,024	0,000
Streptococcus	0,036	0,032	0,000	0,001	0,028	0,000	0,003	0,000	0,002	0,015	0,069	0,000
Streptomyces	0,270	0,154	0,283	0,012	0,194	0,000	0,014	0,379	0,010	0,199	0,173	0,279
Thetatorquevirus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,065	0,000	0,000
Trichomo0s	0,000	0,000	0,012	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Variovorax	0,000	0,000	0,000	0,151	0,000	0,000	0,192	0,000	0,251	0,000	0,000	0,000
Vibrio	0,000	0,000	0,010	0,001	0,023	0,000	0,002	0,000	0,000	0,009	0,000	0,000
Xanthomo0s	0,084	0,000	0,000	0,003	0,000	0,000	0,000	0,000	0,004	0,055	0,000	0,000

Anexo 2: Dados de Saliva

NOME	SRR20326253	SRR20326255	SRR20326257	SRR20326258	SRR20326271	SRR2032 6433	SRR2032 6500	SRR2032 6501	SRR2032 6502	SRR2032 6586	SRR2032 6606	SRR2032 6617	SRR2032 6618	SRR2032 6619
Porphyromonas	0,061	0,060	0,079	0,107	0,215	0,218	0,031	0,092	0,172	0,189	0,366	0,030	1,620	0,135
Capnocytophaga	0,103	0,243	0,084	0,016	0,130	0,120	0,007	0,175	0,062	0,119	0,057	0,007	1,122	0,093

Pasteurella	0,071	0,103	0,094	0,024	0,071	0,014	0,069	0,049	0,056	0,078	0,081	0,057	0,768	0,064
Pseudomonas	0,012	0,005	0,002	0,669	0,002	0,003	0,011	0,006	0,003	0,002	0,001	0,001	0,720	0,060
Neisseria	0,092	0,085	0,037	0,008	0,079	0,040	0,017	0,035	0,031	0,022	0,055	0,007	0,509	0,042
Treponema	0,002	0,006	0,015	0,009	0,070	0,124	0,005	0,087	0,107	0,053	0,007	0,007	0,492	0,041
Frederiksenia	0,084	0,057	0,045	0,005	0,025	0,007	0,091	0,033	0,015	0,002	0,017	0,047	0,427	0,036
Campylobacter	0,004	0,003	0,026	0,006	0,020	0,073	0,002	0,042	0,087	0,105	0,013	0,007	0,389	0,032
Conchiformibius	0,043	0,031	0,099	0,003	0,014	0,007	0,003	0,016	0,016	0,008	0,093	0,024	0,358	0,030
Fusobacterium	0,013	0,010	0,031	0,020	0,027	0,028	0,002	0,009	0,043	0,073	0,025	0,073	0,354	0,030
Moraxella	0,018	0,057	0,033	0,014	0,019	0,038	0,002	0,032	0,014	0,028	0,014	0,015	0,284	0,024
Bibersteinia	0,000	0,004	0,019	0,003	0,013	0,004	0,089	0,004	0,000	0,003	0,016	0,058	0,212	0,018
Mycoplasmopsis	0,003	0,003	0,160	0,000	0,003	0,003	0,003	0,001	0,005	0,001	0,008	0,007	0,198	0,017
Streptococcus	0,012	0,006	0,021	0,008	0,009	0,011	0,023	0,008	0,016	0,011	0,013	0,041	0,180	0,015
Chryseobacterium	0,021	0,041	0,014	0,001	0,014	0,009	0,012	0,033	0,010	0,003	0,007	0,004	0,170	0,014
Klebsiella	0,097	0,002	0,012	0,001	0,001	0,003	0,002	0,003	0,042	0,003	0,002	0,003	0,169	0,014
Mannheimia	0,004	0,002	0,009	0,002	0,007	0,001	0,083	0,002	0,003	0,001	0,012	0,043	0,169	0,014
Actinobacillus	0,006	0,002	0,010	0,002	0,008	0,001	0,078	0,002	0,002	0,001	0,012	0,043	0,167	0,014
Haemophilus	0,012	0,003	0,014	0,003	0,011	0,002	0,048	0,004	0,003	0,002	0,011	0,052	0,163	0,014
Flavobacterium	0,003	0,031	0,006	0,002	0,010	0,017	0,004	0,028	0,005	0,006	0,017	0,008	0,136	0,011
Tannerella	0,001	0,002	0,010	0,006	0,007	0,012	0,001	0,006	0,031	0,039	0,011	0,001	0,126	0,011
Bacteroides	0,004	0,005	0,006	0,003	0,011	0,024	0,005	0,012	0,014	0,019	0,006	0,008	0,117	0,010
Histophilus	0,003	0,002	0,006	0,002	0,013	0,001	0,014	0,001	0,002	0,002	0,008	0,058	0,111	0,009
Glaesserella	0,002	0,001	0,006	0,001	0,005	0,001	0,056	0,001	0,001	0,001	0,008	0,028	0,111	0,009
Prevotella	0,004	0,005	0,003	0,003	0,012	0,007	0,009	0,010	0,018	0,012	0,007	0,012	0,100	0,008
Serratia	0,018	0,000	0,000	0,004	0,009	0,000	0,044	0,003	0,012	0,006	0,000	0,000	0,098	0,008
Actinomyces	0,007	0,001	0,005	0,008	0,004	0,001	0,000	0,008	0,050	0,007	0,001	0,001	0,092	0,008
Gemella	0,007	0,002	0,004	0,002	0,004	0,002	0,040	0,001	0,004	0,004	0,004	0,013	0,085	0,007
Acinetobacter	0,004	0,007	0,006	0,002	0,006	0,017	0,004	0,010	0,005	0,007	0,002	0,010	0,080	0,007
Clostridium	0,004	0,003	0,007	0,002	0,004	0,004	0,005	0,004	0,006	0,008	0,004	0,014	0,066	0,006
Avibacterium	0,001	0,001	0,005	0,001	0,004	0,001	0,016	0,001	0,001	0,001	0,004	0,019	0,054	0,004
Aliarcobacter	0,000	0,000	0,001	0,000	0,000	0,005	0,000	0,029	0,005	0,006	0,000	0,001	0,048	0,004
Vibrio	0,001	0,002	0,003	0,001	0,003	0,003	0,012	0,002	0,006	0,002	0,003	0,010	0,048	0,004
Riemerella	0,018	0,008	0,002	0,000	0,002	0,001	0,003	0,006	0,002	0,000	0,001	0,001	0,045	0,004
Streptomyces	0,023	0,002	0,004	0,001	0,001	0,001	0,000	0,002	0,005	0,001	0,001	0,000	0,043	0,004
Psychrobacter	0,002	0,005	0,003	0,001	0,003	0,010	0,000	0,005	0,003	0,005	0,001	0,002	0,040	0,003
Aggregatibacter	0,001	0,001	0,003	0,001	0,003	0,001	0,011	0,001	0,001	0,001	0,002	0,012	0,037	0,003
Bradyrhizobium	0,020	0,001	0,004	0,000	0,000	0,001	0,000	0,001	0,007	0,001	0,000	0,001	0,036	0,003
Comamonas	0,004	0,004	0,002	0,001	0,003	0,004	0,000	0,010	0,005	0,001	0,001	0,000	0,035	0,003
Parvimonas	0,002	0,001	0,001	0,002	0,002	0,001	0,000	0,001	0,001	0,012	0,005	0,007	0,035	0,003
Elizabethkingia	0,018	0,003	0,002	0,000	0,001	0,001	0,003	0,003	0,001	0,000	0,001	0,001	0,034	0,003
Staphylococcus	0,006	0,002	0,002	0,001	0,002	0,002	0,006	0,002	0,003	0,002	0,002	0,005	0,033	0,003
Taupapillomavirus	0,030	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,031	0,003
Gallibacterium	0,000	0,000	0,002	0,000	0,001	0,000	0,013	0,001	0,000	0,000	0,002	0,008	0,030	0,002
Vespertiliibacter	0,000	0,000	0,001	0,000	0,001	0,000	0,015	0,000	0,000	0,000	0,002	0,007	0,028	0,002
Bacillus	0,003	0,002	0,004	0,001	0,002	0,002	0,002	0,002	0,003	0,003	0,001	0,004	0,027	0,002
Otariodibacter	0,001	0,000	0,001	0,000	0,001	0,000	0,014	0,000	0,000	0,000	0,002	0,007	0,027	0,002
Ottowia	0,005	0,004	0,001	0,000	0,002	0,002	0,000	0,007	0,004	0,001	0,001	0,000	0,027	0,002
Yersinia	0,010	0,001	0,003	0,000	0,001	0,001	0,002	0,001	0,003	0,001	0,001	0,002	0,024	0,002
Streptobacillus	0,001	0,003	0,002	0,001	0,005	0,000	0,000	0,000	0,001	0,002	0,001	0,008	0,024	0,002
Acidovorax	0,003	0,003	0,002	0,000	0,002	0,002	0,000	0,006	0,003	0,001	0,000	0,000	0,022	0,002
Corynebacterium	0,002	0,003	0,003	0,001	0,002	0,002	0,000	0,002	0,003	0,003	0,001	0,000	0,022	0,002
Parabacteroides	0,005	0,001	0,001	0,000	0,001	0,002	0,001	0,002	0,002	0,003	0,001	0,002	0,021	0,002
Rodentibacter	0,001	0,000	0,001	0,000	0,001	0,000	0,007	0,000	0,000	0,000	0,001	0,007	0,020	0,002
Desulfomicrobium	0,000	0,000	0,002	0,001	0,001	0,008	0,000	0,002	0,004	0,000	0,001	0,000	0,019	0,002
Xanthomonas	0,003	0,006	0,003	0,000	0,002	0,001	0,000	0,002	0,002	0,000	0,000	0,000	0,019	0,002
Enterococcus	0,004	0,001	0,001	0,000	0,001	0,001	0,002	0,001	0,001	0,002	0,001	0,004	0,019	0,002
Leptotrichia	0,001	0,001	0,002	0,001	0,001	0,001	0,000	0,001	0,001	0,002	0,001	0,006	0,017	0,001
Arcobacter	0,000	0,000	0,001	0,000	0,000	0,002	0,000	0,009	0,002	0,002	0,000	0,001	0,017	0,001
Shewanella	0,002	0,001	0,002	0,000	0,001	0,001	0,002	0,001	0,001	0,001	0,001	0,003	0,016	0,001
Filifactor	0,000	0,001	0,002	0,001	0,002	0,003	0,000	0,002	0,002	0,001	0,001	0,001	0,016	0,001
Bisgaardia	0,000	0,000	0,001	0,000	0,001	0,000	0,003	0,000	0,000	0,000	0,001	0,007	0,015	0,001
Lysobacter	0,001	0,007	0,001	0,000	0,002	0,001	0,000	0,002	0,000	0,000	0,000	0,000	0,015	0,001
Hydrogenophaga	0,003	0,002	0,001	0,000	0,001	0,001	0,000	0,004	0,002	0,001	0,000	0,000	0,015	0,001
Kaistella	0,001	0,004	0,001	0,000	0,001	0,001	0,001	0,003	0,001	0,000	0,001	0,001	0,014	0,001
Polaribacter	0,000	0,002	0,001	0,000	0,001	0,001	0,001	0,002	0,001	0,001	0,001	0,002	0,014	0,001

Enterobacter	0,007	0,001	0,000	0,000	0,001	0,001	0,001	0,001	0,001	0,000	0,000	0,001	0,014	0,001
Variovorax	0,002	0,002	0,001	0,000	0,001	0,001	0,000	0,003	0,002	0,000	0,000	0,000	0,014	0,001
Paenibacillus	0,001	0,001	0,001	0,000	0,001	0,001	0,001	0,001	0,002	0,001	0,001	0,002	0,013	0,001
Segatella	0,001	0,000	0,001	0,000	0,001	0,001	0,004	0,001	0,001	0,001	0,001	0,001	0,013	0,001
Cloacibacterium	0,001	0,004	0,001	0,000	0,001	0,001	0,001	0,003	0,000	0,000	0,001	0,001	0,013	0,001
Diaphorobacter	0,001	0,002	0,001	0,000	0,001	0,002	0,000	0,004	0,002	0,000	0,000	0,000	0,013	0,001
Stenotrophomonas	0,001	0,005	0,001	0,000	0,002	0,001	0,000	0,001	0,001	0,000	0,000	0,000	0,013	0,001
Lautropia	0,006	0,002	0,001	0,000	0,001	0,000	0,000	0,001	0,001	0,000	0,000	0,000	0,012	0,001
Blautia	0,000	0,000	0,001	0,000	0,001	0,001	0,004	0,001	0,001	0,001	0,001	0,002	0,012	0,001
Escherichia	0,002	0,001	0,001	0,000	0,001	0,001	0,003	0,001	0,000	0,001	0,000	0,001	0,012	0,001
Lactobacillus	0,002	0,000	0,000	0,000	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,003	0,011	0,001
Pseudoalteromonas	0,000	0,001	0,001	0,000	0,001	0,001	0,002	0,001	0,001	0,001	0,001	0,003	0,011	0,001
Providencia	0,000	0,000	0,001	0,000	0,001	0,001	0,002	0,001	0,000	0,000	0,001	0,003	0,011	0,001
Roseburia	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,001	0,007	0,000	0,000	0,000	0,010	0,001
Schaalia	0,002	0,001	0,001	0,000	0,001	0,001	0,000	0,001	0,001	0,002	0,001	0,000	0,010	0,001
Myroides	0,000	0,002	0,001	0,000	0,001	0,001	0,000	0,002	0,001	0,001	0,001	0,001	0,010	0,001
Proteus	0,000	0,000	0,001	0,000	0,001	0,000	0,003	0,000	0,000	0,000	0,001	0,003	0,010	0,001
Phocaeicola	0,000	0,000	0,000	0,000	0,001	0,002	0,001	0,001	0,001	0,002	0,000	0,001	0,010	0,001
Nanosynbacter	0,001	0,001	0,000	0,000	0,001	0,001	0,000	0,000	0,001	0,004	0,000	0,000	0,010	0,001
Mesomycoplasma	0,000	0,001	0,001	0,000	0,000	0,000	0,001	0,000	0,003	0,000	0,000	0,002	0,009	0,001
Ligilactobacillus	0,000	0,000	0,000	0,000	0,000	0,000	0,007	0,000	0,000	0,000	0,000	0,001	0,009	0,001
Epilithonimonas	0,000	0,003	0,001	0,000	0,001	0,001	0,001	0,002	0,000	0,000	0,000	0,000	0,009	0,001
Granulicatella	0,000	0,000	0,000	0,000	0,001	0,002	0,000	0,000	0,000	0,001	0,001	0,003	0,009	0,001
Delftia	0,000	0,001	0,001	0,000	0,001	0,001	0,000	0,003	0,001	0,000	0,000	0,000	0,009	0,001
Rhodoferax	0,000	0,001	0,001	0,000	0,001	0,001	0,000	0,003	0,001	0,001	0,000	0,000	0,009	0,001
Sphingobacterium	0,000	0,001	0,000	0,000	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,009	0,001
Bulleidia	0,001	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,001	0,000	0,000	0,004	0,008	0,001
Photobacterium	0,001	0,000	0,001	0,000	0,000	0,000	0,002	0,000	0,000	0,000	0,000	0,002	0,008	0,001
Burkholderia	0,002	0,001	0,001	0,000	0,000	0,001	0,000	0,001	0,001	0,000	0,000	0,000	0,008	0,001
Salmonella	0,000	0,001	0,000	0,000	0,001	0,001	0,002	0,001	0,000	0,001	0,000	0,002	0,008	0,001
Mycobacterium	0,004	0,000	0,001	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,008	0,001
Paracidovorax	0,000	0,001	0,001	0,000	0,001	0,001	0,000	0,002	0,000	0,000	0,000	0,000	0,007	0,001
Bergeyella	0,001	0,002	0,001	0,000	0,001	0,000	0,001	0,001	0,000	0,000	0,000	0,000	0,007	0,001
Mergibacter	0,000	0,000	0,000	0,000	0,001	0,000	0,003	0,000	0,000	0,000	0,001	0,002	0,007	0,001
Paenimyroides	0,000	0,002	0,000	0,000	0,000	0,001	0,000	0,002	0,000	0,000	0,001	0,000	0,007	0,001
Basfia	0,000	0,000	0,001	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,003	0,007	0,001
Mogibacterium	0,001	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,001	0,001	0,000	0,001	0,007	0,001
Tenacibaculum	0,000	0,001	0,001	0,000	0,000	0,001	0,001	0,001	0,000	0,000	0,000	0,001	0,007	0,001
Rothia	0,000	0,001	0,000	0,000	0,000	0,001	0,000	0,001	0,001	0,001	0,000	0,001	0,007	0,001
Hoylesella	0,000	0,000	0,000	0,000	0,001	0,001	0,001	0,001	0,001	0,001	0,000	0,001	0,007	0,001
Lysinibacillus	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,001	0,001	0,000	0,001	0,006	0,001
Mycoplasma	0,000	0,001	0,001	0,000	0,001	0,000	0,000	0,000	0,001	0,001	0,000	0,001	0,006	0,001
Vagococcus	0,000	0,000	0,001	0,000	0,001	0,001	0,001	0,000	0,000	0,001	0,001	0,001	0,006	0,001
Limnohabitans	0,002	0,000	0,001	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,000	0,000	0,006	0,001
Pedobacter	0,000	0,001	0,000	0,000	0,000	0,001	0,000	0,001	0,001	0,000	0,000	0,001	0,006	0,001
Mucilaginibacter	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,006	0,001
Alistipes	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,001	0,001	0,001	0,001	0,001	0,006	0,001
Alicycliphilus	0,002	0,001	0,000	0,000	0,000	0,001	0,000	0,002	0,000	0,000	0,000	0,000	0,006	0,001
Arthrobacter	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,000	0,006	0,000
Methylophaga	0,004	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,006	0,000
Kingella	0,001	0,001	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,000	0,001	0,001	0,006	0,000
Lactococcus	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,001	0,006	0,000
Cupriavidus	0,001	0,001	0,001	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,000	0,000	0,006	0,000
Microbacterium	0,001	0,000	0,001	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,006	0,000
Polynucleobacter	0,000	0,001	0,000	0,000	0,000	0,001	0,000	0,001	0,001	0,001	0,000	0,001	0,005	0,000
Winogradskyella	0,000	0,001	0,000	0,000	0,000	0,001	0,000	0,001	0,000	0,000	0,000	0,001	0,005	0,000
Clostridioides	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,001	0,001	0,001	0,005	0,000
Spiroplasma	0,000	0,001	0,000	0,000	0,001	0,000	0,001	0,000	0,000	0,001	0,000	0,001	0,005	0,000
Christiangramia	0,001	0,001	0,001	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,001	0,005	0,000
Plasmodium	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,001	0,005	0,000
Buchnera	0,000	0,000	0,001	0,000	0,001	0,001	0,000	0,001	0,000	0,000	0,000	0,001	0,005	0,000
Nocardioides	0,001	0,001	0,001	0,001	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,005	0,000

Amniculibacterium	0,000	0,001	0,001	0,000	0,001	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,005	0,000
Blattabacterium	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,001	0,005	0,000
Neobacillus	0,000	0,000	0,001	0,000	0,000	0,000	0,001	0,000	0,000	0,001	0,000	0,001	0,005	0,000
Halomonas	0,001	0,001	0,001	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,000	0,000	0,005	0,000
Lachnoanaerobaculum	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,001	0,001	0,001	0,000	0,001	0,005	0,000
Gleimia	0,000	0,001	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,003	0,000	0,000	0,005	0,000
Ornithobacterium	0,000	0,001	0,000	0,000	0,000	0,001	0,000	0,001	0,001	0,000	0,000	0,000	0,005	0,000
Nonlabens	0,000	0,001	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,000	0,000	0,001	0,005	0,000
Cellulophaga	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,001	0,005	0,000
Luteimonas	0,001	0,002	0,001	0,000	0,001	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,005	0,000
Mycoplasmoides	0,003	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,005	0,000
Gilliamella	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,001	0,005	0,000
Aerococcus	0,003	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,004	0,000
Allomuricauda	0,000	0,001	0,000	0,000	0,000	0,001	0,000	0,001	0,000	0,000	0,000	0,000	0,004	0,000
Melaminivora	0,001	0,001	0,000	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,000	0,000	0,004	0,000
Colwellia	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,001	0,000	0,000	0,001	0,004	0,000
Butyrivibrio	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,004	0,000
Paraprevotella	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,001	0,004	0,000
Legionella	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,001	0,000	0,000	0,000	0,001	0,004	0,000
Fusarium	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,004	0,000
Roseateles	0,001	0,001	0,000	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,000	0,000	0,004	0,000
Malaciobacter	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,001	0,000	0,000	0,000	0,004	0,000
Thermomonas	0,001	0,002	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,004	0,000
Sneathia	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,004	0,000
Hymenobacter	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,004	0,000
Pantoea	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,001	0,004	0,000
Pseudoxanthomonas	0,000	0,002	0,000	0,000	0,001	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,004	0,000
Leishmania	0,000	0,000	0,003	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,004	0,000
Paracoccus	0,000	0,001	0,001	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,004	0,000
Aeromonas	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,000	0,000	0,004	0,000
Mediterraneibacter	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,001	0,000	0,001	0,004	0,000
Anaerococcus	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,001	0,004	0,000
Empedobacter	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,000	0,000	0,004	0,000
Pseudoprevotella	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,001	0,000	0,000	0,001	0,004	0,000
Massilia	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,000	0,000	0,004	0,000
Achromobacter	0,001	0,001	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,004	0,000
Sulfurimonas	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,004	0,000
Aquimarina	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,004	0,000
Listeria	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,004	0,000
Mycolicibacterium	0,001	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,004	0,000
Coprobacter	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,001	0,000	0,000	0,001	0,004	0,000
Mageeibacillus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,003	0,004	0,000
Anaerocolumna	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,004	0,000
Bordetella	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,001	0,003	0,000
Romboutsia	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,001	0,003	0,000
Crassaminicella	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,003	0,000
Helicobacter	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,000	0,000	0,001	0,003	0,000
Citrobacter	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,001	0,003	0,000
Halarcobacter	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000	0,000	0,000	0,000	0,003	0,000
Erysipelothrix	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,003	0,000
Proteiniphilum	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,001	0,001	0,000	0,000	0,003	0,000
Alysiella	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,003	0,000
Sphingomonas	0,001	0,000	0,001	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,003	0,000
Maribacter	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,003	0,000
Coprococcus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,003	0,000
Paraburkholderia	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,003	0,000
Formosa	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,003	0,000
Erwinia	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,003	0,000
Mammaliicoccus	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,001	0,003	0,000
Cytobacillus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,003	0,000
Draconibacterium	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,001	0,000	0,000	0,000	0,003	0,000

Limosilactobacillus	0,000	0,000	0,000	0,000	0,001	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,003	0,000
Synechococcus	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,003	0,000
Pectobacterium	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,001	0,003	0,000
Xenorhabdus	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,001	0,003	0,000
Finegoldia	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,001	0,003	0,000
Serpentinimonas	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,000	0,000	0,003	0,000
Amygdalobacter	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,003	0,000
Cardiobacterium	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,003	0,000
Ramlibacter	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,003	0,000
Marinomonas	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,003	0,000
Faecalibacterium	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,003	0,000
Polaromonas	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,003	0,000
Duncaniella	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,003	0,000
Francisella	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,003	0,000
Priestia	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,003	0,000
Muribaculum	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,003	0,000
Virgibacillus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,003	0,000
Wielerella	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,003	0,000
Metamycoplasma	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,003	0,000
Thomasclavelia	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Echinicola	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,002	0,000
Carnobacterium	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Micromonospora	0,001	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Thauera	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,002	0,000
Peptoniphilus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,002	0,000
Kocuria	0,000	0,001	0,001	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Thalassotalea	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Pandoraea	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,002	0,000
Ezakiella	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,002	0,000
Lacinutrix	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Alteromonas	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,002	0,000
Mesorhizobium	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Chitinophaga	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Tessaracoccus	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Paenacidovorax	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Macrococcus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,002	0,000
Pseudoleptotrichia	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,002	0,000
Natrialba	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Helcococcus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,001	0,002	0,000
Colletotrichum	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Rhizobium	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Dietzia	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Agrobacterium	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Arsenophonus	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Jeotgalibaca	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Barnesiella	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,002	0,000
Lutibacter	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Brachyspira	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,002	0,000
Caloramator	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,002	0,000
Cyanobacterium	0,000	0,000	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Butyricimonas	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Bartonella	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Anaerostipes	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,002	0,000
Brevibacillus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Nocardia	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Spirosoma	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Dorea	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,002	0,000
Apibacter	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Eubacterium	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,002	0,000
Pontibacter	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Lachnoclostridium	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Brachybacterium	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,002	0,000
Ralstonia	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,002	0,000

Tepidibacter	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Flammeovirga	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Nostoc	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Dokdonia	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Algibacter	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Ruminococcus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Aspergillus	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,002	0,000
Eikenella	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Planococcus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Trueperella	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,002	0,000
Salegentibacter	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,002	0,000
Macellibacteroides	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Fenollaria	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,002	0,000
Metabacillus	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Deinococcus	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Peptacetobacter	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Faecalibacter	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Peribacillus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Selenomonas	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Niallia	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Labilibaculum	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Moritella	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Ilyobacter	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,002	0,000
Facklamia	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,002	0,000
Pyricularia	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,002	0,000
Cruoricaptor	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Xylanibacter	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Aliivibrio	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,001	0,002	0,000
Petrimonas	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Cetobacterium	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,002	0,000
Pelistega	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Carboxylicivirga	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Sporisorium	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Vallitalea	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Glutamicibacter	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Curtobacterium	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Rhodococcus	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Dolichospermum	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Amedibacterium	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Antarcticibacterium	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Tissierella	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Entomomonas	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Ureaplasma	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Thermoanaerobacterium	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,002	0,000
Cohaesibacter	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Anoxybacillus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Aequorivita	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Agarivorans	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Kosakonia	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Algoriphagus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Peptostreptococcus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Trichomonas	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Marinobacter	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Thiomicrorhabdus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Arcanobacterium	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,002	0,000
Leptospira	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000
Arachnia	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Pulveribacter	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,001	0,000
Marinifilum	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Marivirga	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Dickeya	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Bifidobacterium	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000

Wenzhouxiangella	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,001	0,000
Dysgonomonas	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Roseovarius	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Alkaliphilus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Globicatella	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Herbaspirillum	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Frondihabitans	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Photorhabdus	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Ancylobacter	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Seonamhaeicola	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Sporosarcina	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Pseudoduganella	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Companilactobacillus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Zunongwangia	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Agathobacter	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Paeniclostridium	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Gardnerella	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Stutzerimonas	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Terrisporobacter	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Veillonella	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,001	0,000
Cedecea	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Thermothelomyces	0,001	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Mesoplasma	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Prochlorococcus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Paraglaciecola	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Megamonas	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Caminicella	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Leuconostoc	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Olleya	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Uruburuella	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Dialister	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,001	0,000
Zobellia	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Devosia	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Janthinobacterium	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Sodaliphilus	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Weissella	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Methylomonas	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Brevibacterium	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Dictyostelium	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000

Anexo 3: Dados de sangue.

name	SRR20326270	SRR20326274	SRR20326276	SRR20326277	SRR20326503	SRR20326506	SRR20326528	SRR20326573	SRR20326596	SRR20326597	SRR20326604	SRR20326628
Klebsiella	0	144	0	146	77	0	49	38	0	10	107	0
Streptomyces	0	102	66	81	83	20	127	34	0	0	101	121
Yersinia	0	0	48	0	23	42	0	0	0	0	18	46
Bradyrhizobium	0	0	34	57	25	13	20	14	0	0	25	23
Babesia	0	0	0	0	0	19	0	0	0	0	0	0
Plasmodium	0	0	0	0	0	18	0	0	0	0	0	0
Treponema	0	0	0	0	0	0	0	0	0	0	60	0
Porphyromonas	0	0	0	0	0	0	0	0	0	0	42	0
Mycoplasma	0	0	0	0	0	0	0	0	0	0	0	11

6 DISCUSSÃO

Esta dissertação apresentou um novo protocolo para análises metagenômicas, visando aprimorar a eficiência, precisão e reprodutibilidade neste campo de pesquisa em rápida evolução. O estudo focou na caracterização de microrganismos em amostras de sangue, saliva e pele de cães, abordando diversos aspectos críticos das análises metagenômicas, desde a construção de bancos de dados de referência até a implementação de workflows otimizados.

Os principais resultados e contribuições deste trabalho incluem:

1. **Desenvolvimento de um protocolo inovador**: Foi proposto um novo protocolo que integra ferramentas e abordagens metodológicas avançadas para análises metagenômicas. Este protocolo demonstrou uma eficiência temporal significativa, reduzindo substancialmente o tempo de execução das ferramentas utilizadas e permitindo a realização de estudos mais amplos e complexos de forma eficaz.

2. **Otimização de bancos de dados**: A combinação de bancos de dados abrangentes e específicos foi fundamental para melhorar a precisão das análises taxonômicas. Conforme sugerido por Gihawi et al. (2023), a inclusão de bancos de dados mais completos aumentou significativamente a acurácia da classificação, como evidenciado pelos resultados obtidos.

3. **Remoção eficiente de contaminantes**: Os testes realizados demonstraram a importância de uma abordagem combinada na remoção de hospedeiro e contaminantes. O método mais eficaz alcançou uma precisão de 78,26% na remoção de leituras indesejadas, destacando a importância de utilizar ferramentas adequadas e configurações otimizadas para minimizar falsos positivos e garantir análises taxonômicas mais precisas.

4. **Caracterização da diversidade microbiana**: As análises revelaram uma variação significativa na diversidade microbiana entre os diferentes tipos de amostras:

    - Pele: 69 gêneros de organismos identificados
    - Saliva: 353 gêneros identificados
    - Sangue: 9 gêneros identificados Esta variação reflete as diferentes condições ambientais e funções de cada local no corpo canino.

5. **Identificação de microrganismos comuns**: O estudo identificou 4 organismos presentes nos três tipos de amostras (pele, saliva e sangue), destacando a interconexão entre diferentes nichos microbianos no corpo canino.

6. **Redução de ruídos e erros de leitura**: A aplicação de um corte de abundância relativa (< 0,01%) resultou em uma redução significativa de microrganismos considerados ruídos ou erros de leitura, particularmente nas amostras de saliva,

onde houve uma redução de 81,61% no número de microrganismos identificados.

7. **Implicações para saúde animal e pública**: A identificação de patógenos zoonóticos, como Pasteurella e Capnocytophaga na saliva canina, ressalta a importância deste estudo para a saúde pública e a compreensão das interações entre humanos e animais de estimação.

6.1. Limitações e desafios que persistem:

- A necessidade de melhorias contínuas, especialmente na minimização de taxas de erro para aumentar ainda mais a acurácia das análises.
- A questão dos falsos positivos e negativos permanece como um desafio significativo, impactando a precisão e a confiabilidade dos resultados obtidos.
- A constante evolução das técnicas de sequenciamento e análise metagenômica, que exige atualizações frequentes nos protocolos e ferramentas utilizados.
- A complexidade inerente às comunidades microbianas, que continua a desafiar os métodos atuais de análise e interpretação de dados.
- A necessidade de estudos adicionais para compreender a influência de fatores como raça, idade e dieta na composição do microbioma canino.

6.2. Perspectivas futuras para a continuidade deste trabalho incluem:

1. Aprimoramento contínuo do protocolo, incorporando novas ferramentas e métodos à medida que surgem no campo.
2. Expansão dos testes para incluir uma gama mais ampla de ambientes microbianos e tipos de amostras.
3. Desenvolvimento de abordagens para lidar com desafios específicos, como a detecção de espécies raras ou a análise de interações microbianas complexas.
4. Exploração de técnicas de aprendizado de máquina e inteligência artificial para melhorar a interpretação dos dados metagenômicos.
5. Colaboração entre desenvolvedores de ferramentas para aprimorar algoritmos e bases de dados, visando minimizar taxas de erro e aumentar a acurácia das análises.
6. Investigação aprofundada da relação entre o microbioma canino e a saúde, tanto dos animais quanto dos humanos que convivem com eles.
7. Desenvolvimento de ferramentas diagnósticas baseadas no perfil microbiano para detecção precoce de doenças em cães.
8. Exploração do potencial terapêutico da modulação do microbioma canino.

7 CONCLUSÃO

Em conclusão, este trabalho representa um passo significativo no avanço das análises metagenômicas, fornecendo um protocolo robusto e eficiente que pode ser amplamente aplicado na pesquisa microbiológica. A abordagem desenvolvida não apenas melhora a qualidade e confiabilidade das análises, mas também estabelece uma base sólida para futuros desenvolvimentos neste campo dinâmico e crucial para nossa compreensão dos ecossistemas microbianos.

Os resultados obtidos neste estudo contribuem significativamente para aprimorar a qualidade e a precisão das análises metagenômicas, fornecendo uma base sólida para futuros estudos e aplicações nessa área. A combinação de bancos de dados abrangentes e específicos, juntamente com a realização de testes exaustivos, nos permite nos aproximar de um workflow ideal para a análise de dados metagenômicos, contribuindo para uma melhor compreensão da complexidade e diversidade dos ecossistemas microbianos.

A caracterização detalhada dos microbiomas da pele, saliva e sangue de cães oferece insights valiosos para a medicina veterinária, a saúde pública e a compreensão das interações homem-animal. Os resultados obtidos destacam a complexidade e especificidade dos ecossistemas microbianos em diferentes tecidos caninos, fornecendo uma base para estudos futuros sobre o papel desses microrganismos na saúde e doença. A identificação de potenciais patógenos zoonóticos ressalta a importância contínua da vigilância e da pesquisa na interface entre saúde animal e humana.

Esperamos que o protocolo proposto sirva como um recurso valioso para pesquisadores, contribuindo para análises metagenômicas mais robustas, reproduzíveis e informativas. À medida que o campo continua a evoluir, a colaboração contínua entre pesquisadores, o compartilhamento de conhecimentos e a validação constante dos resultados serão essenciais para impulsionar avanços e desvendar os mistérios do mundo microbiano. O protocolo e os insights gerados por este estudo têm o potencial de servir como recursos valiosos para pesquisadores, veterinários e profissionais de saúde pública, contribuindo para o desenvolvimento de abordagens mais eficazes para o diagnóstico, tratamento e prevenção de doenças em animais de estimação e, por extensão, na promoção da saúde humana.

REFERÊNCIAS

[1] Gire SK, Goba A, Andersen KG, Sealfon RSG, Park DJ, Kanneh L, et al. Genomic surveillance elucidates {Ebola} virus origin and transmission during the 2014 outbreak. Science (1979) 2014;345:1369–72. https://doi.org/10.1126/science.1259657.

[2] Guttman DS, McHardy AC, Schulze-Lefert P. Microbial genome-enabled insights into plant-microorganism interactions. Nat Rev Genet 2014;15:797–813. https://doi.org/10.1038/nrg3748.

[3] Vincent AT, Derome N, Boyle B, Culley AI, Charette SJ. Next-generation sequencing ({NGS}) in the microbiological world: {How} to make the most of your money. J Microbiol Methods 2017;138:60–71. https://doi.org/10.1016/j.mimet.2016.02.016.

[4] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences 1977;74:5463–7. https://doi.org/10.1073/pnas.74.12.5463.

[5] Sawicki MP, Samara G, Hurwitz M, Passaro E. Human Genome Project. The American Journal of Surgery 1993;165:258–64. https://doi.org/10.1016/S0002-9610(05)80522-7.

[6] Handelsmanl J, Rondon' MR, Brady2 SF, Clardy2 J, Goodman' RM. Molecular biological access to the chemistry of unknown 05SP 09 soil microbes: a new frontier for natural products E4 + qfs. [cited 2024 Jul 9]; Available from: http://biomednet.com/elecref/10745521005R0245

[7] Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, et al. Environmental {Genome} {Shotgun} {Sequencing} of the {Sargasso} {Sea}. Science (1979) 2004;304:66–74. https://doi.org/10.1126/science.1093857.

[8] Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature 2005 437:7057 2005;437:376–80. https://doi.org/10.1038/nature03959.

[9] Fan JB, Gunderson KL, Bibikova M, Yeakley JM, Chen J, Wickham Garcia E, et al. [3] Illumina Universal Bead Arrays. Methods Enzymol 2006;410:57–73. https://doi.org/10.1016/S0076-6879(06)10003-8.

[10] Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The Human Microbiome Project. Nature 2007 449:7164 2007;449:804–10. https://doi.org/10.1038/nature06244.

[11] Buckley RM, Harris AC, Wang G-D, Whitaker DT, Zhang Y-P, Ostrander EA. Best practices for analyzing imputed genotypes from low-pass sequencing in dogs. Mamm Genome 2022;33:213–29. https://doi.org/10.1007/s00335-021-09914-z.

[12] Stewart RD, Auffret MD, Warr A, Wiser AH, Press MO, Langford KW, et al. Assembly of 913 microbial genomes from metagenomic sequencing of the cow rumen. Nature Communications 2018 9:1 2018;9:1–11. https://doi.org/10.1038/s41467-018-03317-6.

[13] Berg G, Rybakova D, Fischer D, Cernava T, Vergès M-CC, Charles T, et al. Microbiome definition re-visited: old concepts and new challenges. Microbiome 2020;8:103. https://doi.org/10.1186/s40168-020-00875-0.

[14] Zhou Y-H, Gallins P. A {Review} and {Tutorial} of {Machine} {Learning} {Methods} for {Microbiome} {Host} {Trait} {Prediction}. Front Genet 2019;10:579. https://doi.org/10.3389/fgene.2019.00579.

[15] Jin M, Osman M, Green BA, Yang Y, Ahuja A, Lu Z, et al. Evidence for the transmission of antimicrobial resistant bacteria between humans and companion animals: A scoping review. One Health 2023;17:100593. https://doi.org/10.1016/J.ONEHLT.2023.100593.

[16] Røken M, Forfang K, Wasteson Y, Haaland AH, Eiken HG, Hagen SB, et al. Antimicrobial resistance—Do we share more than companionship with our dogs? J Appl Microbiol 2022;133:1027–39. https://doi.org/10.1111/JAM.15629.

[17] Tóth AG, Csabai I, Krikó E, Tőzsér D, Maróti G, Patai Á V, et al. Antimicrobial resistance genes in raw milk for human consumption. Sci Rep 2020;10:7464. https://doi.org/10.1038/s41598-020-63675-4.

[18] Abinpet. Mercado {Pet} {Brasil} 2022. Associação Brasileira Da Industria e Produtos Para Animais de Estimação 2022.

[19] Tilahun Y, Pinango JQ, Johnson F, Lett C, Smith K, Gipson T, et al. Transcript and blood-microbiome analysis towards a blood diagnostic tool for goats affected by {Haemonchus} contortus. Sci Rep 2022;12:5362. https://doi.org/10.1038/s41598-022-08939-x.

[20] García-Fonticoba R, Ferrer L, Francino O, Cuscó A. The microbiota of the surface, dermis and subcutaneous tissue of dog skin. Anim Microbiome 2020;2:34. https://doi.org/10.1186/s42523-020-00050-8.

[21] Tóth AG, Tóth I, Rózsa B, Kovács EG, Dubecz A, Patai Á V, et al. Canine saliva is a source of interspecies antimicrobial resistance gene transfer 2022. https://doi.org/10.1101/2022.03.07.483304.

[22] Castillo DJ, Rifkin RF, Cowan DA, Potgieter M. The healthy human blood microbiome: Fact or fiction? Front Cell Infect Microbiol 2019;9:449041. https://doi.org/10.3389/FCIMB.2019.00148/BIBTEX.

[23] Socransky SS, Haffajee AD, Cugini MA, Smith C, Kent RL. Microbial complexes in subgingival plaque. J Clin Periodontol 1998;25:134–44. https://doi.org/10.1111/J.1600-051X.1998.TB02419.X.

[24] Oba PM, Carroll MQ, Alexander C, Valentine H, Somrak AJ, Keating SCJ, et al. Microbiota populations in supragingival plaque, subgingival plaque, and saliva habitats of adult dogs. Anim Microbiome 2021;3:38. https://doi.org/10.1186/s42523-021-00100-9.

[25] Cogen AL, Nizet V, Gallo RL. Skin microbiota: a source of disease or defence? Br J Dermatol 2008;158:442–55. https://doi.org/10.1111/j.1365-2133.2008.08437.x.

[26] Grice EA, Segre JA. The skin microbiome. Nature Reviews Microbiology 2011 9:4 2011;9:244–53. https://doi.org/10.1038/nrmicro2537.

[27] Occurrence of zoonotic enteric parasites in fecal samples from dogs in shelters, parks, squares and public roads, and the dog guardians' perception of zoonoses as for the risk to public health in the city of {Guarapuava}, {Paraná}, {Brazil}. Top Companion Anim Med 2024;58:100826. https://doi.org/10.1016/j.tcam.2023.100826.

[28] Kmieciak W, Szewczyk EM. Are zoonotic {Staphylococcus} pseudintermedius strains a growing threat for humans? Folia Microbiol (Praha) 2018;63:743–7. https://doi.org/10.1007/s12223-018-0615-2.

[29] Grace D, Gilbert J, Randolph T, Kang'ethe E. The multiple burdens of zoonotic disease and an ecohealth approach to their assessment. Trop Anim Health Prod 2012;44:67–73. https://doi.org/10.1007/s11250-012-0209-y.

[30] FE D, EA K, EC T, JM B, T C, L M, et al. The canine oral microbiome. PLoS One 2012;7:e36067. https://doi.org/10.1371/journal.pone.0036067.

[31] Cleaveland S, Laurenson MK, Taylor LH. Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. Philos Trans R Soc Lond B Biol Sci 2001;356:991–9. https://doi.org/10.1098/rstb.2001.0889.

[32] Burgos-Cáceres S. Canine Rabies: A Looming Threat to Public Health. Animals 2011;1:326–42. https://doi.org/10.3390/ani1040326.

[33] Cleaveland S, Dye C. Maintenance of a microparasite infecting several host species: rabies in the {Serengeti}. Parasitology 1995;111 Suppl:S33--47. https://doi.org/10.1017/s0031182000075806.

[34] Samrot A V, Sean TC, Bhavya KS, Sahithya CS, Chan-Drasekaran S, Palanisamy R, et al. Leptospiral {Infection}, {Pathogenesis} and {Its} {Diagnosis}-{A} {Review}. Pathogens 2021;10:145. https://doi.org/10.3390/pathogens10020145.

[35] Røken M, Forfang K, Wasteson Y, Haaland AH, Eiken HG, Hagen SB, et al. Antimicrobial resistance—{Do} we share more than companionship with our dogs? J Appl Microbiol 2022;133:1027–39. https://doi.org/10.1111/jam.15629.

[36] Røken M, Forfang K, Wasteson Y, Haaland AH, Eiken HG, Hagen SB, et al. Antimicrobial resistance—Do we share more than companionship with our dogs? J Appl Microbiol 2022;133:1027–39. https://doi.org/10.1111/JAM.15629.

[37] Rung J, Brazma A. Reuse of public genome-wide gene expression data. Nat Rev Genet 2013;14:89–99. https://doi.org/10.1038/nrg3394.

[38] Sielemann K, Hafner A, Pucker B. The reuse of public datasets in the life sciences: Potential risks and rewards. PeerJ 2020;8:e9954. https://doi.org/10.7717/PEERJ.9954/SUPP-1.

[39] Thomas T, Gilbert J, Meyer F. Metagenomics - a guide from sampling to data analysis. Microb Inform Exp 2012;2:1–12. https://doi.org/10.1186/2042-5783-2-3/FIGURES/1.

[40] Zoonoses. World Health Organization 2020. [Internet]. 2020 [cited 2024 Jul 10]. Zoonoses. Available from: https://www.who.int/news-room/fact-sheets/detail/zoonoses

[41] Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, et al. Global trends in emerging infectious diseases. Nature 2008 451:7181 2008;451:990–3. https://doi.org/10.1038/nature06536.

[42] Hubálek Z. Emerging Human Infectious Diseases: Anthroponoses, Zoonoses, and Sapronoses - Volume 9, Number 3—March 2003 - Emerging Infectious Diseases journal - CDC. Emerg Infect Dis 2003;9:403–4. https://doi.org/10.3201/EID0903.020208.

[43] Bidaisee S, Macpherson CNL. Zoonoses and One Health: A Review of the Literature. J Parasitol Res 2014;2014:874345. https://doi.org/10.1155/2014/874345.

[44] Volume. PEOPLE, PATHOGENS AND OUR PLANET 2012. [cited 2024 Jul 10]; Available from: www.worldbank.org } [Internet]. Conselho Regional de Medicina Veterinária São Paulo.

2022 [cited 2023 Mar 10]. Available from: https://crmvsp.gov.br/zoonoses-correspondem-a-mais-de-60-das-doencas-humanas/

[45] CRMV/SP. {ZOONOSES} {CORRESPONDEM} {A} {MAIS} {DE} 60\% {DAS} {DOENÇAS} {HUMANAS}. Conselho Regional de Medicina Veterinária São Paulo 2022.

[46] O médico-veterinário é profissional de Saúde Pública. Conselho Federal de Medicina Veterinária 2022. https://www.cfmv.gov.br/o-medico-veterinario-e-profissional-de-saude-publica/comunicacao/noticias/2020/09/01/ (accessed July 10, 2024).

[47] One {Health}. World Health Organization (WHO) 2017. [cited 2023 Mar 9]. Available from: https://www.who.int/news-room/questions-and-answers/item/one-health

[48] Day MJ. One health: The importance of companion animal vector-borne diseases. Parasit Vectors 2011;4:1–6. https://doi.org/10.1186/1756-3305-4-49/TABLES/1.

[49] Chomel B. Emerging and {Re}-{Emerging} {Zoonoses} of {Dogs} and {Cats}. Animals 2014;4:434–45. https://doi.org/10.3390/ani4030434.

[50] Macpherson CNL. Human behaviour and the epidemiology of parasitic zoonoses. Int J Parasitol 2005;35:1319–31. https://doi.org/10.1016/J.IJPARA.2005.06.004.

[51] Zinsstag J, Schelling E, Waltner-Toews D, Tanner M. From “one medicine” to “one health” and systemic approaches to health and well-being. Prev Vet Med 2011;101:148–56. https://doi.org/10.1016/J.PREVETMED.2010.07.003.

[52] Chomel BB, Sun B. Zoonoses in the {Bedroom}. Emerg Infect Dis 2011;17:167–72. https://doi.org/10.3201/eid1702.101070.

[53] Belotto A, Leanes LF, Schneider MC, Tamayo H, Correa E. Overview of rabies in the {Americas}. Virus Res 2005;111:5–12. https://doi.org/10.1016/j.virusres.2005.03.006.

[54] Fatores facilitadores da invasão microbiana - Doenças infecciosas - Manuais MSD edição para profissionais n.d. https://www.msdmanuals.com/pt/profissional/doen%C3%A7as-infecciosas/biologia-das-doen%C3%A7as-infecciosas/fatores-facilitadores-da-invas%C3%A3o-microbiana (accessed August 23, 2024).

[55] Soni J, Sinha S, Pandey R. Understanding bacterial pathogenicity: a closer look at the journey of harmful microbes. Front Microbiol 2024;15. https://doi.org/10.3389/FMICB.2024.1370818.

[56] PM O, MQ C, C A, H V, AJ S, SCJ K, et al. Microbiota populations in supragingival plaque, subgingival plaque, and saliva habitats of adult dogs. Anim Microbiome 2021;3:38. https://doi.org/10.1186/s42523-021-00100-9.

[57] Pierezan F, Olivry T, Paps JS, Lawhon SD, Wu J, Steiner JM, et al. The skin microbiome in allergen-induced canine atopic dermatitis. Vet Dermatol 2016;27:332--e82. https://doi.org/10.1111/vde.12366.

[58] Rostaher A, Rodriguez-Campos S, Deplazes P, Zwickl L, Akdis AC, Urwyler A, et al. Atopic dermatitis in {West} {Highland} white terriers - {Part} {III}: early life peripheral blood regulatory {T} cells are reduced in atopic dermatitis. Vet Dermatol 2021;32:239--e63. https://doi.org/10.1111/vde.12939.

[59] Knights D, Kuczynski J, Charlson ES, Zaneveld J, Mozer MC, Collman RG, et al. Bayesian community-wide culture-independent microbial source tracking. Nat Methods 2011;8:761–3. https://doi.org/10.1038/nmeth.1650.

[60] Garza DR, Dutilh BE. From cultured to uncultured genome sequences: metagenomics and modeling microbial ecosystems. Cell Mol Life Sci 2015;72:4287–308. https://doi.org/10.1007/s00018-015-2004-1.

[61] Galhano BSP, Ferrari RG, Panzenhagen P, de Jesus ACS, Conte-Junior CA. Antimicrobial {Resistance} {Gene} {Detection} {Methods} for {Bacteria} in {Animal}-{Based} {Foods}: {A} {Brief} {Review} of {Highlights} and {Advantages}. Microorganisms 2021;9:923. https://doi.org/10.3390/microorganisms9050923.

[62] Schloss PD, Handelsman J. Biotechnological prospects from metagenomics. Curr Opin Biotechnol 2003;14:303–10. https://doi.org/10.1016/s0958-1669(03)00067-3.

[63] Pabalan N, Jarjanazi H, Steiner TS. Meta-analysis in microbiology. Indian J Med Microbiol 2014;32:229–35. https://doi.org/10.4103/0255-0857.136547.

[64] Bioinformatics tools to assess metagenomic data for applied microbiology {\textbar} {Applied} {Microbiology} and {Biotechnology} n.d.. [Internet]. [cited 2024 May 13]. Available from: https://link.springer.com/article/10.1007/s00253-018-9464-9

[65] Riesenfeld CS, Schloss PD, Handelsman J. Metagenomics: {Genomic} {Analysis} of {Microbial} {Communities}. Annu Rev Genet 2004;38:525–52. https://doi.org/10.1146/annurev.genet.38.072902.091216.

[66] Berg G, Rybakova D, Fischer D, Cernava T, Vergès M-CC, Charles T, et al. Microbiome definition re-visited: old concepts and new challenges. Microbiome 2020;8:103. https://doi.org/10.1186/s40168-020-00875-0.

[67] Baquero F, Nombela C. The microbiome as a human organ. Clinical Microbiology and Infection 2012;18:2–4. https://doi.org/10.1111/j.1469-0691.2012.03916.x.

[68] Whipps J, Lewis K, Cooke R. Mycoparasitism and plant disease control. Fungi Biol Control Syst 1988:161–87.

[69] Barbosa DA. {ANÁLISE} {DA} {COMPOSIÇÃO} {E} {DA} {INFLUÊNCIA} {EXERCIDA} {PELA} {MICROBIOTA} {INTESTINAL} {FÚNGICA} {DE} {CAMUNDONGOS} {DURANTE} {O} {DESENVOLVIMENTO} {DE} {CAQUEXIA} {INDUZIDA} {POR} {TRANSPLANTE} {DE} {CÉLULAS} {DE} {CÂNCER} {PULMONAR} ({LLC}). Dissertação de mestrado. Universidade Mogi das Cruzes, 2022.

[70] Liu L, Li Y, Li S, Hu N, He Y, Pong R, et al. Comparison of {Next}-{Generation} {Sequencing} {Systems}. Biomed Res Int 2012;2012:e251364. https://doi.org/10.1155/2012/251364.

[71] Oulas A, Pavloudi C, Polymenakou P, Pavlopoulos GA, Papanikolaou N, Kotoulas G, et al. Metagenomics: {Tools} and {Insights} for {Analyzing} {Next}-{Generation} {Sequencing} {Data} {Derived} from {Biodiversity} {Studies}. Bioinform Biol Insights 2015;9:BBI.S12462. https://doi.org/10.4137/BBI.S12462.

[72] Wensel CR, Pluznick JL, Salzberg SL, Sears CL. Next-generation sequencing: insights to advance clinical investigations of the microbiome. J Clin Invest 2022;132:e154944. https://doi.org/10.1172/JCI154944.

ANEXOS

Anexo1: Artigos publicados e Capítulos de livros

1. Takao PS, Moura JF, Corrêa PS, Oliveira R de C, Rocha RR, Silva FF da, et al. Tenebrio AI Detection: Um modelo para identificação e classificação precisa de Tenebrio molitor por meio de visão computacional. Revista Científica UMC. 11 de dezembro de 2023;8(3):e080300001–e080300001.

2. Takao PS, Moura JF, Correa PS, Barbosa DA, Menegidio FB. DESENVOLVIMENTO DE BANCOS DE DADOS PARA ANÁLISE METAGENÔMICA DE MICROBIOMAS DE CÃES: UMA ABORDAGEM COM KRAKEN2 E BOWTIE2. Anais do IV Congresso Brasileiro de Biotecnologia On-line. 18 de março de 2024;5(2):170.

3. Berlofa MVC. Interações entre malassezia restricta e o micobioma humano: uma perspectiva abrangente. Revista Científica Multidisciplinar Núcleo do Conhecimento. 8 de fevereiro de 2024;01(08):21–34.

4. De Almeida Moretti Pinna R, Aguiar FL, Filyppo Ferreira Porto Decaria, Naufal GGA, Pereira VS, Takao PS, et al. AUTOMEDICAR-SE OU NÃO? ANÁLISE DO COMPORTAMENTO DE AUTOMEDICAÇÃO EM ESTUDANTES DE MEDICINA E ODONTOLOGIA. 5 de outubro de 2023 [citado 27 de agosto de 2024]; Disponível em: https://zenodo.org/record/8409063

5. De Almeida Pinna Moretti M, Zarate GC, De Souza IV, Carvallo GT, Soler J, Takao PS, et al. ALÉM DAS MARCAS VISÍVEIS: UMA ANÁLISE DA VIOLÊNCIA DOMÉSTICA CONTRA A MULHER. 9 de outubro de 2023 [citado 27 de agosto de 2024]; Disponível em: https://zenodo.org/record/8422970

6. Takao PS, Correa PS, Oliveira R de C, et al (2024) RELAÇÃO ENTRE O MICROBIOMA BACTERIANO GASTROINTESTINAL COM A DOENÇA INFLAMATÓRIA INTESTINAL (DII) EM CÃES. Editora Científica Digital, pp 80–94

Anexo 2: Registro de Softwares



REPÚBLICA FEDERATIVA DO BRASIL
MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA, COMÉRCIO E SERVIÇOS
INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL
DIRETORIA DE PATENTES, PROGRAMAS DE COMPUTADOR E TOPOGRAFIAS DE CIRCUITOS

Certificado de Registro de Programa de Computador

Processo Nº: **BR512024002892-7**

O Instituto Nacional da Propriedade Industrial expede o presente certificado de registro de programa de computador, válido por 50 anos a partir de 1° de janeiro subsequente à data de 01/08/2024, em conformidade com o §2°, art. 2° da Lei 9.609, de 19 de Fevereiro de 1998.

Título: ResearchReflect

Data de publicação: 01/08/2024

Data de criação: 30/07/2024

Titular(es): DAVID ACIOLE BARBOSA; RAFAELA DE CAMPOS OLIVEIRA

Autor(es): FABIANO BEZERRA MENEGIDIO; LUIZ ROBERTO NUNES; DANIELA LEITE JABES; KARINE FREHNER KAVALCO; RUBENS PASA; RENAN RODRIGUES ROCHA; PAMELA SOUZA CORRÊA; FABIANA BORGES DOS SANTOS; ALEXANDRE SANTOS SIMEONE; GLEDISON ERIC TEIXEIRA; RAQUEL BARBOSA CINTRA; PAULO SALLAROLA TAKAO

Linguagem: PYTHON

Campo de aplicação: CO-02; IF-01

Tipo de programa: TC-01

Algoritmo hash: SHA-256

Resumo digital hash: 3A8AC283004B64054C8180C224F063E957812944C644EC311473F9DF38FBA1B4

Expedido em: 20/08/2024

Aprovado por:
Carlos Alexandre Fernandes Silva
Chefe da DIPTO

REPÚBLICA FEDERATIVA DO BRASIL
MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA, COMÉRCIO E SERVIÇOS
INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL
DIRETORIA DE PATENTES, PROGRAMAS DE COMPUTADOR E TOPOGRAFIAS DE CIRCUITOS

Certificado de Registro de Programa de Computador

Processo Nº: **BR512024002955-9**

O Instituto Nacional da Propriedade Industrial expede o presente certificado de registro de programa de computador, válido por 50 anos a partir de 1° de janeiro subsequente à data de 01/08/2024, em conformidade com o §2°, art. 2° da Lei 9.609, de 19 de Fevereiro de 1998.

Título: MetagenomeCleaner

Data de publicação: 01/08/2024

Data de criação: 30/07/2024

Titular(es): DAVID ACIOLE BARBOSA; PAULO SALLAROLA TAKAO

Autor(es): FABIANO BEZERRA MENEGIDIO; LUIZ ROBERTO NUNES; DANIELA LEITE JABES; KARINE FREHNER KAVALCO; RUBENS PASA; RENAN RODRIGUES ROCHA; PAMELA SOUZA CORRÊA; FABIANA BORGES DOS SANTOS; ALEXANDRE SANTOS SIMEONE; GLEDISON ERIC TEIXEIRA; RAQUEL BARBOSA CINTRA; RAFAELA DE CAMPOS OLIVEIRA

Linguagem: PYTHON

Campo de aplicação: CO-02; IF-01

Tipo de programa: TC-01

Algoritmo hash: SHA-256

Resumo digital hash: 41815AD29D85BDEBE9EAC991876F29E60B7459749CA0A9ADCDC68D028628E813

Expedido em: 20/08/2024

Aprovado por:
Carlos Alexandre Fernandes Silva
Chefe da DIPTO

Anexo 3: Participações em Congressos e Apresentação de Trabalhos

Premiados.



IV CONGRESSO BRASILEIRO DE BIOTECNOLOGIA ON-LINE

Certificado

Certificamos para os devidos fins que o trabalho intitulado "**DESENVOLVIMENTO DE BANCOS DE DADOS PARA ANÁLISE METAGENÔMICA DE MICROBIOMAS DE CÃES: UMA ABORDAGEM COM KRAKEN2 E BOWTIE2"** de autoria de **PAULO SALLAROLA TAKAO, JULIO FRANZ MOURA, PAMELA SOUZA CORREA, DAVID ACIOLE BARBOSA E FABIANO BEZERRA MENEGIDIO** recebeu menção honrosa no IV Congresso Brasileiro de Biotecnologia On-line - CONBIOTEC, realizados nos dias 18 a 21 de Março de 2024.

FORTALEZA/CE, 12 de abril de 2024.

IV CONGRESSO BRASILEIRO DE BIOTECNOLOGIA ON-LINE

REALIZAÇÃO Instituto | PARCEIRO EDITORA INTEGRAR | APOIO ABED

Prof. Dr. Vandbergue Santos Pereira
Coordenador do evento
Instituto Multiprofissional de Ensino - IME
CNPJ: 36.773.074/0001-08

CÓDIGO DO CERTIFICADO: 3KH6D-7PUEP-0388C-W70RN

VERIFIQUE AUTENTICIDADE EM: https://ime.events/certificado/validar/3KH6D-7PUEP-0388C-W70RN



UMC UNIVERSIDADE

CERTIFICADO

Certificamos que o trabalho intitulado "**TENEBRIO AI DETECTION: UM MODELO PARA IDENTIFICAÇÃO E CLASSIFICAÇÃO PRECISA DE TENEBRIO MOLITOR POR MEIO DE VISÃO COMPUTACIONAL**" foi apresentado no 1º SUMMIT - UMC na categoria "**PESQUISA**", realizado no período de 30 de novembro de 2023 a 01 de dezembro de 2023, na Universidade de Mogi das Cruzes – UMC, tendo como autores "PAULO SALLAROLA TAKAO, JULIO FRANZ MOURA, PÂMELA SOUZA CORRÊA, RAFAELA DE CAMPOS OLIVEIRA, RENAN RODRIGUES ROCHA, FLÁVIO FRANCISCO DA SILVA, FABIANO BEZERRA MENEGIDIO".

Mogi das Cruzes, 01 de dezembro de 2023.

Prof. Dr. Cláudio José Alves de Brito
Pró-reitor Acadêmico

Profa. Dra. Tatiana Ribeiro de Campos Mello
Coordenadora de Pesquisa

UMC UNIVERSIDADE



II INSETEC
CONGRESSO BRASILEIRO DE INSETOS ALIMENTÍCIOS E TECNOLOGIAS ASSOCIADAS

Certificamos que

o trabalho intitulado

Desenvolvimento de bancos de dados para análise metagenômica de microbiomas de insetos: uma abordagem com Kraken2 e Bowtie2

de autoria de

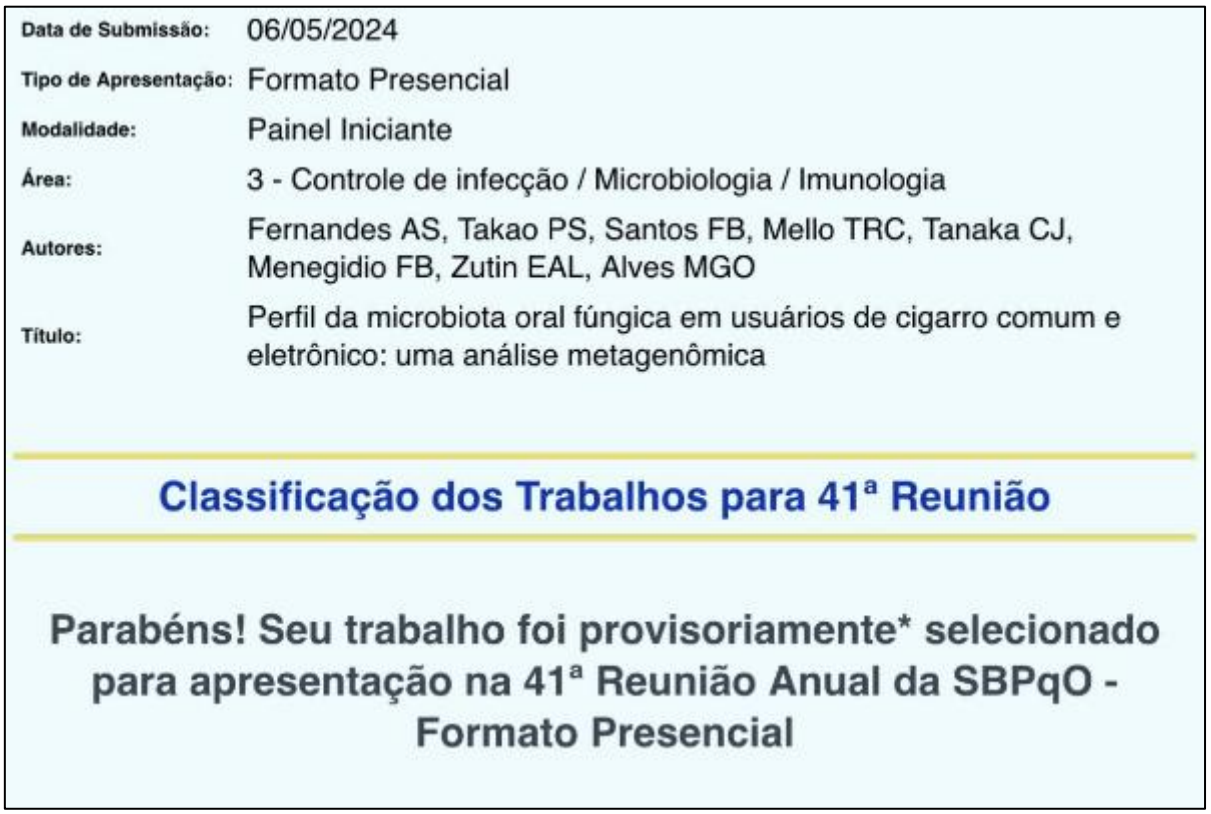
Takao, Paulo S.; Barbosa, David A.; Moura, Julio F.; Menegidio, Fabiano B.

recebeu a premiação de MENÇÃO HONROSA pela Apresentação na Sessão de Comunicação Oral do II INSETEC no dia 18 de Outubro de 2023 no Centro de Convenções da UNICAMP, Campinas, SP.

Patricia J. Thyssen, PhD.
Presidente do II INSETEC

Case Oliveira
Casé Oliveira
Presidente da ASBRACIA

Participação.



Data de Submissão:	06/05/2024
Tipo de Apresentação:	Formato Presencial
Modalidade:	Painel Iniciante
Área:	3 - Controle de infecção / Microbiologia / Imunologia
Autores:	Fernandes AS, Takao PS, Santos FB, Mello TRC, Tanaka CJ, Menegidio FB, Zutin EAL, Alves MGO
Título:	Perfil da microbiota oral fúngica em usuários de cigarro comum e eletrônico: uma análise metagenômica

Classificação dos Trabalhos para 41ª Reunião

Parabéns! Seu trabalho foi provisoriamente* selecionado para apresentação na 41ª Reunião Anual da SBPqO - Formato Presencial

Verifique o código de autenticidade **2349192.0886691.442927.8.5162680773362118684 7** em https://www.even3.com.br/documentos

IV COBICET

CERTIFICADO DE TRABALHO

Certificamos que o trabalho intitulado **A importância da Bioinformática nas práticas educativas para profissionais das áreas de informática e saúde: o caso dos mitogenomas**, de autoria de MARIA VITORIA CAVALHEIRO BERLOFA, Mariana da Palma Valério, Paulo Sallarola Takao, Alessandro Pereira da Silva, Daniela Leite Jabes e Fabiano Bezerra Menegidio, foi apresentado no formato e-banner - anais online, no IV Congresso Brasileiro Interdisciplinar em Ciência e Tecnologia - "Educação inovadora: ensino, pesquisa e extensão interdisciplinar". O trabalho foi avaliado por pares em avaliação duplo cega e integra os anais online do evento no formato Resumo, área temática ENSI - Ensino em Ciência e Tecnologia.

Brasil /Online, 01 de setembro de 2023.

Profª Dra. Poliana Mendes de Souza
Presidente do IV COBICET

Profº Dr. Ademar Nogueira do Nascimento
Coordenador do Comitê Científico do IV COBICET

Profº Dr. Gabriel Gomes de Oliveira
Vice-Coordenador de Comitê Científico do IV COBICET

Instituições organizadoras e apoiadoras:

Verifique o código de autenticidade **2349192.0886691.445552.8.5162680773362119998 7** em https://www.even3.com.br/documentos

IV COBICET

CERTIFICADO DE TRABALHO

Certificamos que o trabalho intitulado **MONTAGEM DO GENOMA MITOCONDRIAL COMPLETO DO Gryllus bimaculatus**, de autoria de Paulo Sallarola Takao, Lurye Silva Theodoro, Fabiana Borges dos Santos, Karine Frehner Kavalco, Rubens Pasa e Fabiano Bezerra Menegidio, foi apresentado no formato e-banner - anais online, no IV Congresso Brasileiro Interdisciplinar em Ciência e Tecnologia - "Educação inovadora: ensino, pesquisa e extensão interdisciplinar". O trabalho foi avaliado por pares em avaliação duplo cega e integra os anais online do evento no formato Resumo, área temática CBIO - Ciências Biológicas.

Brasil /Online, 01 de setembro de 2023.

Profª Dra. Poliana Mendes de Souza
Presidente do IV COBICET

Profº Dr. Ademar Nogueira do Nascimento
Coordenador do Comitê Científico do IV COBICET

Profº Dr. Gabriel Gomes de Oliveira
Vice-Coordenador de Comitê Científico do IV COBICET

Instituições organizadoras e apoiadoras:

Verifique o código de autenticidade **2124532.0886691.464992.8.52374210099653936620** em https://www.even3.com.br//documentos

X-meeting 2023

Brazilian **Symposium** on Bioinformatics 2023

BSB 2023

Certificate of DNA and Genomics presentation

This certifies that the work entitled **Unveiling the Microbial Diversity in a Labrador's Skin: Implications for Disease Transmission**, authored by **Paulo Sallarola Takao, Rafaela de Campos Oliveira, David Aciole Barbosa, YARA NATERCIA LIMA FAUSTINO DE MARIA, Regina Costa de Oliveira, Daniela Leite Jabes, Luiz Nunes and Fabiano Bezerra Menegidio** was presented by Paulo Sallarola Takao during the Poster session of the X-Meeting / BSB 2023, held in Curitiba - Brazil between June 13th and june 16th, 2023.

Curitiba, June, 16th 2023

Marcelo Mendes Brandão
Presidente da AB3C

Fabricio Martins Lopes
General Chair do BSB

Maithê Magalhães
Co-Chair Sessão de posters

Felipe Ciamponi
Co-Chair Sessão de posters

Verifique o código de autenticidade **2124532.0886691.464997.8.52374210099653936680** em https://www.even3.com.br//documentos

X-meeting 2023

Brazilian **Symposium** on Bioinformatics 2023

BSB 2023

Certificate of DNA and Genomics presentation

This certifies that the work entitled **Mycobacterium tuberculosis in the Saliva Microbiome of Dogs: Implications for Public Health**, authored by **Paulo Sallarola Takao, Rafaela de Campos Oliveira, David Aciole Barbosa, YARA NATERCIA LIMA FAUSTINO DE MARIA, Daniela Leite Jabes, Luiz Nunes and Fabiano Bezerra Menegidio** was presented by Paulo Sallarola Takao during the Poster session of the X-Meeting / BSB 2023, held in Curitiba - Brazil between June 13th and june 16th, 2023.

Curitiba, June, 16th 2023

Marcelo Mendes Brandão
Presidente da AB3C

Fabricio Martins Lopes
General Chair do BSB

Maithê Magalhães
Co-Chair Sessão de posters

Felipe Ciamponi
Co-Chair Sessão de posters


IV COMDEV
2023 | CURITIBA
Certificado
Medvep Cursos e Eventos
A Medvep confere a
PAULO SALLARO TAKAO
o certificado de participação no COMDEV 2023 - Congresso Medvep Internacional de Dermatologia Veterinária realizado entre os dias 13 e 15 de Abril de 2023, em Curitiba - PR na modalidade de congressista com carga horária equivalente à 24 horas
Dilso Henrique Bloot
Diretor Medvep
Dalmar Moreira
Diretor Medvep
PATROCÍNIO
VETNIL
ROYAL CANIN
PremieRpet
Boehringer Ingelheim
Vet Allergenic
zoetis
COORDENAÇÃO CIENTÍFICA
DERMATOVET
REALIZAÇÃO
Medvep

Anexo 4: Cursos



CERTIFICADO

O Reitor do Centro Universitário União das Américas Descomplica, Instituição de Ensino Superior recredenciado pela portaria MEC nº 1842 de 14/09/2023, no uso de suas atribuições, certifica que

Paulo Sallarola Takao

portador(a) do CPF nº 326.407.318-82, concluiu o curso de Pós-Graduação Lato Sensu em **Microbiologia Avançada**, com carga horária de 390 horas, realizado no período de 06 de fevereiro de 2023 a 08 de janeiro de 2024, estando apto(a) a gozar de todos os direitos e prerrogativas legais.

Foz do Iguaçu, 08 de Janeiro de 2024.

Marina Araujo
Secretária Acadêmica

Dieter Sergei Sardeli de Paiva
Reitor

UniAmérica Centro Universitário
+descomplica



ACS Chemistry for Life®

ACS Publications
Most Trusted. Most Cited. Most Read.

A ACS Publications certifica que

Paulo Sallarola Takao

participou do ***Curso de Comunicação e Escrita Científica***, ministrado pelo Prof. Dr. Osvaldo Novais de Oliveira Junior nos dias 25 de Outubro, 01 e 09 de Novembro de 2023, por transmissão online, totalizando a carga horária de 4h30.

REGIANE BRACCHI
ACS Publications Manager, LATAM
American Chemical Society

Verifique o código de autenticidade **2124538.0886691.469521.8.52374010099653967250** em https://www.even3.com.br//documentos

X-meeting 2023

Brazilian **Symposium** on Bioinformatics 2023

BSB 2023

Course certificate

This certifies that **Paulo Sallarola Takao** has attended the Course entitled "Course: Introduction to biological databases" during the X-Meeting / BSB 2023 held in Curitiba - Brazil.

Curitiba, June 16th, 2023.

Marcelo Mendes Brandão
Presidente da AB3C

Fabrício Martins Lopes
General Chair do BSB

Verifique o código de autenticidade **2124538.0886691.469538.8.52374010099653967400** em https://www.even3.com.br//documentos

X-meeting 2023

Brazilian **Symposium** on Bioinformatics 2023

BSB 2023

Course certificate

This certifies that **Paulo Sallarola Takao** has attended the Course entitled "Course: Bioinformatics in Personalized medicine (MGI)" during the X-Meeting / BSB 2023 held in Curitiba - Brazil.

Curitiba, June 16th, 2023.

Marcelo Mendes Brandão
Presidente da AB3C

Fabrício Martins Lopes
General Chair do BSB

Wolters Kluwer

.periodicos.

CERTIFICADO

A Wolters Kluwer em parceria com
o Portal de Periódicos CAPES confere a

PAULO SALLAROLA TAKAO

o certificado pela participação na sessão "Tipos de Estudos de Revisão de Literatura: definições e aplicabilidades" ministrado por Anna Carolina Alencar no dia 27 de abril de 2023 por videoconferência com carga horária de 2 (duas) horas

Enf.ª Me. Anna Carolina Alencar
Training Manager, Brasil
Wolters Kluwer Health

Sarah Carpenter
Director Sales Enablement & Train
Wolters Kluwer Health

W20215587
R. Bela Cintra, 768 - Consolação, São Paulo - SP, 01415-000 Brasil
+55 11 3508-7200 - https://www.wolterskluwer.com/en/health




CERTIFICADO
Certificamos que Paulo Sallarola Takao, participou com êxito do minicurso PROGRAMAÇÃO EM LINGUAGEM R APLICADA EM BIOINFORMÁTICA: anotação de transcritos, expressão diferencial e enriquecimento de vias, realizado em 14 e 15 de Setembro de 2022, contabilizando carga horária total de 6 horas.
BIOINFORMÁTICA PARA NOVATOS
Even3

Anexo 5: Orientação e Coorientação de alunos de iniciação científica.

1. ODILON GONÇALO PINHO MENDES BATISTA (PIBIC 2023-2024)
   Caracterização do microbioma dermal de cães (*Canis lúpus familiaris*) através de dados públicos de sequenciamento de nova geração

2. Larissa Pinto de Medeiros (PIBIC 2024-2025)
   Bioprospecção in silico de bacteriocinas em dados de Metagenômica canino de saliva

3. Larissa Palmeira dos Santos da Costa (PIBIC 2024-2025)
   Bioprospecção in silico de bacteriocinas em dados de Metagenômica canino de pele

4. Leonardo Carlos Neiva (PIBIC 2024-2025)
   Desenvolvimento de um simulador de ausculta pulmonar e avaliação da frequência respiratória em cães como ferramenta auxiliar para o treinamento na graduação em medicina veterinária

5. Maria Suzana Freire Martins (PIBIC 2024-2025)
   Bioprospecção in silico de bacteriocinas em dados metagenômicos intestinais de bicho-da-farinha (*Tenebrio molitor)*

Anexo 6**: Capítulo de livro**

RELAÇÃO ENTRE O MICROBIOMA BACTERIANO GASTROINTESTINAL COM A DOENÇA INFLAMATÓRIA INTESTINAL (DII) EM CÃES

Paulo Sallarola Takao
Universidade de Mogi das Cruzes (UMC)
Pamela Souza Correa
Universidade de Mogi das Cruzes (UMC)
Rafaela de Campos Oliveira
Universidade de Mogi das Cruzes (UMC)
Renan Rodrigues Rocha
Universidade de Mogi das Cruzes (UMC)
Daniela Sampaio Silva
Universidade de Mogi das Cruzes (UMC)
Carolina Garcia
Universidade Federal do Sul da Bahia (UFSB)
Rubens Pasa
Universidade Federal de Viçosa (UFV)
Karine F. Kavalco
Universidade Federal de Viçosa (UFV)
Daniela Leite Jabes
Universidade de Mogi das Cruzes (UMC)
Fabiano Bezerra Menegidio
Universidade de Mogi das Cruzes (UMC)

RESUMO

Introdução: O microbioma gastrointestinal dos cães é crucial para a homeostase, impactando a digestão e a saúde imunológica. **Objetivo:** Este artigo revisa a relação entre o microbioma e as doenças gastrointestinais, com foco na Doença Inflamatória Intestinal (IBD). **Metodologia:** Realizou-se uma revisão da literatura, empregando sequenciamento de nova geração para avaliar a diversidade microbiana. **Resultados:** Desequilíbrios na microbiota estão associados à IBD, com diminuição de bactérias benéficas (*Bacteroides*, *Prevotella*) e aumento de patógenas (*Escherichia coli*). **Discussão:** A interação entre microbioma e hospedeiro é essencial para a saúde gastrointestinal, oferecendo oportunidades para novas abordagens terapêuticas. **Conclusão:** Entender o microbioma é vital para o manejo de doenças gastrointestinais, podendo levar à identificação de biomarcadores e intervenções mais eficazes.

Palavras-chave: microbioma gastrointestinal, doença inflamatória intestinal, diversidade microbiana, ciências ômicas, cães.

INTRODUÇÃO

Os cães têm sido companheiros dos humanos por milhares de anos, desempenhando papéis diversos que vão desde a companhia até o trabalho. Atualmente, compartilham lares tanto em áreas urbanas quanto rurais, sendo considerados parte integrante da família por muitos tutores (GRZEŚKOWIAK et al., 2015). Essa mudança de ambiente foi gradual para os humanos, porém abrupta para os cães, que passaram a conviver com os humanos em um novo contexto, proporcionando benefícios mútuos em uma dinâmica diferente daquela a que estavam habituados. Devido a essa proximidade, o interesse na saúde e no bem-estar dos cães tem crescido nas últimas décadas, juntamente com a pesquisa científica voltada para a saúde e prevenção de doenças desses animais. Nesse contexto, a saúde gastrointestinal emerge como uma área de interesse significativo, abrangendo uma variedade de doenças que podem afetar os cães. Uma das áreas de pesquisa relevantes para a saúde animal é o estudo da composição microbiana do trato gastrointestinal, com o objetivo de promover e manter a saúde do hospedeiro (ALESSANDRI et al., 2020; MONDO et al., 2019).

Ao longo do tempo, os cães têm passado por mudanças em sua dieta, historicamente consumindo uma maior quantidade de proteína em comparação com as dietas ricas em carboidratos mais comuns nos dias de hoje. A saúde e o bem-estar desses animais dependem do microbioma gastrointestinal, cuja composição e atividade têm sido associadas a diversas doenças tanto nos próprios cães quanto em seus tutores (GRZEŚKOWIAK et al., 2015).

A relação entre a comunidade microbiana e o hospedeiro é de grande importância para a saúde dos cães. Assim, o microbioma gastrointestinal nos cães tem ganhado destaque tanto na prática clínica veterinária quanto na pesquisa científica, devido à sua conexão inseparável com a saúde animal. A alimentação desempenha um papel crucial na manutenção, composição e regulação do microbioma gastrointestinal de cães, servindo como substrato para sua atividade (WERNIMONT et al., 2020). A importância de uma população microbiana equilibrada reside na capacidade de fornecer diversos benefícios nutricionais e gastrointestinais, proteger o intestino contra agentes patogênicos e promover uma estrutura intestinal saudável, assim como uma função intestinal otimizada (LYU et al., 2020).

Estudos sugerem que o microbioma, além de sua influência no sistema gastrointestinal, também desempenha um papel no desenvolvimento e na regulação de vários outros sistemas do corpo, incluindo o nervoso, renal, digestivo, dérmico, endócrino, imunológico e respiratório (TIZARD & JONES, 2018; WERNIMONT et al., 2020). Assim, os tutores desempenham um

papel essencial na modulação do microbioma de seus cães, principalmente por meio da seleção de alimentos apropriados (WERNIMONT et al., 2020).

DESENVOLVIMENTO

Microbioma e Microbioma Intestinal

O microbioma intestinal nos cães consiste no conjunto de microrganismos que habitam o sistema gastrointestinal, incluindo bactérias, protozoários, fungos e vírus, sendo considerado um componente essencial do sistema digestivo (HONNEFFER et al., 2014; LYU et al., 2020; SUCHODOLSKI, 2011b; Wernimont et al., 2020). Esta definição também engloba o material genético desses microrganismos. Embora tradicionalmente fosse considerado como um "agregado de material genético da totalidade dos microrganismos que habitam um habitat definido" (WERNIMONT et al., 2020), hoje em dia, devido aos seus potenciais efeitos na saúde humana e animal, o microbioma gastrointestinal é reconhecido como um órgão com funções metabólicas únicas (POSSEMIERS et al., 2011).

Estudos recentes indicam que o trato gastrointestinal dos mamíferos abriga aproximadamente 10^10 a 10^14 microrganismos, um número cerca de 10 vezes maior do que o de células do próprio hospedeiro (LYU et al., 2020; SUCHODOLSKI, 2011a). Nesse artigo, abordaremos principalmente o componente bacteriano do microbioma, uma vez que ele representa a maior parte da população gastrointestinal dos mamíferos (SUCHODOLSKI, 2011b, 2016). A predominância das bactérias se deve principalmente aos benefícios que proporcionam ao hospedeiro, como a fermentação de carboidratos complexos em ácidos graxos de cadeia curta, a síntese de vitaminas como vitamina A, K2 e B12, e a conversão de ácidos biliares, entre outros processos metabólicos (SUCHODOLSKI, 2016).

Atualmente, reconhece-se que o microbioma tem um desenvolvimento durante a gestação (STINSON et al., 2019) e exerce influência ao longo da vida do cão, afetando sua fisiologia, anatomia, reprodução e atividade física (WERNIMONT et al., 2020). Em humanos e ratos, foi demonstrado que o microbioma começa a se formar antes do nascimento. Embora não haja estudos que confirmem a colonização microbiana gastrointestinal antes do nascimento em cães e gatos, considera-se essa possibilidade devido às semelhanças fisiológicas e ao fato de serem mamíferos (GRZEŚKOWIAK et al., 2015).

O microbioma coexiste em simbiose com o cão hospedeiro. Assim, o hospedeiro fornece um ambiente propício à sobrevivência do microbioma, incluindo temperatura e níveis

adequados de oxigênio para os microrganismos, além de defesa contra agentes patogênicos. Por sua vez, o microbioma desempenha seu papel, influenciando vários aspectos da função gastrointestinal, como tamanho das vilosidades, renovação de enterócitos e motilidade. A presença de um microbioma estável e saudável é essencial para o desenvolvimento e equilíbrio contínuo do sistema imunológico, incluindo a mucosa gastrointestinal (ETTINGER et al., 2017). Além disso, o microbioma desempenha um papel importante na prevenção da colonização por agentes patogênicos, liberando nutrientes e metabólitos que influenciam as células do sistema imunológico (TIZARD & JONES, 2018). Essa ação é fundamental para o desenvolvimento fisiológico adequado do intestino, sendo moldada pela influência do microbioma (MONDO et al., 2019; TIZARD & JONES, 2018).

Em casos de doença gastrointestinal, a composição do microbioma é alterada e sua diversidade é reduzida. Essa alteração na composição resulta em mudanças nas espécies que compõem a flora de um cão saudável, o que pode levar a um aumento de espécies com propriedades prejudiciais (ETTINGER et al., 2017).

A composição microbiana do intestino dos cães não é uniforme, variando ao longo do trato gastrointestinal, com comunidades distintas identificadas em cada região anatômica (WERNIMONT et al., 2020). A diversidade e o número de bactérias aumentam gradualmente ao longo do trato gastrointestinal, e essa complexa composição pode ser influenciada por fatores externos (HONNEFFER et al., 2017; RITCHIE et al., 2008; SUCHODOLSKI et al., 2008). A dieta, a presença ou ausência de muco, alterações metabólicas, doenças gastrointestinais agudas ou crônicas e o sistema imunológico do cão são os principais fatores que afetam e moldam a localização e a composição da população microbiana gastrointestinal (ALESSANDRI et al., 2020; DONALDSON et al., 2015).

Em cães, assim como em humanos, há evidências de que a administração de antibióticos pode alterar a composição de certos grupos de bactérias, resultando em mudanças significativas na composição microbiana. Essas alterações podem persistir por várias semanas (LYU et al., 2020). Estudos mostram que mudanças relevantes na composição da microbiota intestinal foram observadas em casos de doenças gastrointestinais crônicas em humanos e animais (INNESS et al., 2007; JANECZKO et al., 2008).

Constituição do microbioma canino

Os cães, como animais de companhia, não dependem principalmente de seu microbioma para obtenção de energia. Eles evoluíram como carnívoros, e seus tratos gastrointestinais são relativamente simples em comparação com os seres humanos ou animais de criação (DENG & SWANSON, 2015). Apesar de compartilharem muitas semelhanças anatômicas com os gatos, os cães são omnívoros e podem usar carboidratos como fonte de energia, sendo capazes de digerir, absorver e metabolizar esses compostos (DENG & SWANSON, 2015). Por outro lado, os gatos domésticos têm uma maior demanda por proteína, obtendo sua nutrição principalmente de tecidos animais. Eles estão metabolicamente adaptados, com uma menor dependência de glicose e um metabolismo proteico mais proeminente para obtenção de energia (DENG & SWANSON, 2015; WERNIMONT et al., 2020).

Embora existam diferenças mencionadas anteriormente, a composição do microbioma gastrointestinal em cães e gatos saudáveis é fundamentalmente semelhante (HOFFMANN et al., 2016). Isso não significa que não haja distinções entre eles, especialmente entre raças, idades, gêneros, composição genética, doenças, tratamentos e regiões geográficas. Esses fatores contribuem para uma diversidade individual entre espécies e entre indivíduos da mesma espécie, resultando em composições individuais para cada animal (DEUSCH et al., 2015; WERNIMONT et al., 2020).

Embora haja variações ao longo do trato gastrointestinal dos cães, tanto em termos de sua taxonomia quanto de seu número, a maioria dos estudos publicados até agora foi conduzida usando amostras fecais. No entanto, é importante notar que essas amostras fecais podem não refletir completamente as diferenças nas diferentes regiões do trato gastrointestinal. A coleta de amostras fecais oferece vantagens, como maior facilidade de obtenção e menores questões éticas em comparação com a coleta de amostras de múltiplas porções do trato gastrointestinal (ALESSANDRI et al., 2020).

A biodiversidade e a quantidade de microrganismos no trato gastrointestinal de cães saudáveis aumentam gradualmente ao longo do sistema digestivo (HONNEFFER et al., 2017; RITCHIE et al., 2008; SUCHODOLSKI et al., 2008). A composição varia, com números que variam de 10^2 a 10^10 unidades formadoras de colônias (UFC/g) de conteúdo luminal, sendo menores no estômago e intestino delgado em comparação com o trato gastrointestinal distal (RITCHIE et al., 2008). No intestino delgado, estão presentes tanto bactérias aeróbias quanto anaeróbias (ALESSANDRI et al., 2020), enquanto no cólon e no ceco, locais de maior

fermentação, predominam as bactérias anaeróbias devido à ausência de oxigênio (HONNEFFER et al., 2017; SUCHODOLSKI, 2016; SUCHODOLSKI et al., 2008).

Apesar das diferenças no microbioma entre os cães, é comum encontrar os filos *Fusobacteria*, *Bacteroidetes*, *Firmicutes*, *Proteobacteria* e *Actinobacteria* (ALESSANDRI et al., 2020; HANDL et al., 2011; RITCHIE et al., 2008; WERNIMONT et al., 2020). No filo *Firmicutes*, as bactérias dominantes pertencem às classes *Bacilli*, *Clostridia* e *Erysipelotrichi*. Embora esses números sejam semelhantes em cães e gatos, há especificidades dentro dessas classes, como a predominância do gênero *Streptococcus* e *Lactobacillus* em cães na classe *Bacilli* (HANDL et al., 2011).

O filo Fusobacteria é um componente significativo do microbioma gastrointestinal de cães, embora sua abundância relativa seja menor em comparação com outros filos (ALESSANDRI et al., 2020; SUCHODOLSKI et al., 2008). Ele é encontrado em animais carnívoros saudáveis e não está associado a doenças gastrointestinais (DENG & SWANSON, 2015). Os filos *Proteobacteria* e *Actinobacteria* são menos abundantes em comparação com outros filos, sendo mais comuns em cães (ALESSANDRI et al., 2020; HANDL et al., 2011).

Embora o microbioma intestinal de cães compartilhe muitos microrganismos semelhantes, sua composição pode variar devido a uma série de fatores, incluindo dieta, idade, sexo, genética, doenças e região geográfica (HANDL et al., 2011; JHA et al., 2020). A saúde dos cães está estreitamente ligada à interação com seu microbioma gastrointestinal. Desiquilíbrios nesse microbioma podem levar à proliferação de agentes patogênicos ou à alteração de mecanismos fisiológicos e metabólicos essenciais, resultando em doenças gastrointestinais (DENG & SWANSON, 2015).

Métodos de estudo do microbioma gastrointestinal

Nos primeiros estudos que buscavam identificar os componentes do microbioma em cães, foram empregados métodos tradicionais baseados em culturas microbianas. No entanto, esses métodos apresentavam algumas limitações, como baixas contagens, entre 10^2 a 10^5 unidades formadoras de colônias (UFC/g) (HONNEFFER et al., 2014). Eles dependiam da cultura, isolamento e identificação posterior dos microrganismos presentes nas amostras. Embora fossem úteis para determinar a presença de uma infecção ativa ou realizar testes de sensibilidade a antibióticos, eles não eram eficazes para avaliar a diversidade total do microbioma, pois

muitas espécies não podem ser cultivadas em laboratório (HOFFMANN et al., 2016; SUCHODOLSKI, 2016).

Atualmente, as limitações dos métodos de cultura foram superadas por técnicas de identificação mais avançadas, como as tecnologias de sequenciamento molecular. Estas permitem uma compreensão mais profunda e precisa do microbioma gastrointestinal, através de métodos independentes de cultura, como a metagenômica e metatranscriptômica (ALESSANDRI et al., 2020).

A avaliação metagenômica, por exemplo, permite estudar a composição e as funções da comunidade microbiana intestinal, incluindo a identificação de bactérias que não podem ser cultivadas *in vitro*, bem como de outros microrganismos como protozoários, fungos, archaea e vírus (BLAKE & SUCHODOLSKI, 2016; HAMADY & KNIGHT, 2009).

O método inicialmente utilizado foi o sequenciamento do gene 16S rRNA, que amplifica e sequência regiões hipervariáveis do gene 16S rRNA através da técnica de PCR (DAMS et al., 1988; HAMADY & KNIGHT, 2009). No entanto, este método apresenta algumas limitações, como a identificação taxonômica limitada até o nível do gênero e dificuldades na identificação de grupos bacterianos sub-representados (ALESSANDRI et al., 2020).

Uma alternativa mais precisa é o sequenciamento do espaço interno transcrito (ITS), que permite uma identificação taxonômica mais específica até a nível de espécie e subespécie, embora esteja disponível apenas para alguns gêneros microbianos específicos (MILANI et al., 2014, 2018). Além disso, o sequenciamento metagenômico *shotgun* oferece uma visão mais abrangente, identificando não apenas a composição taxonômica, mas também as funções dos microrganismos presentes (ALESSANDRI et al., 2020).

Outro método, a hibridização in situ fluorescente (FISH), permite estudar a população bacteriana nos tecidos, preservando o biofilme e fornecendo informações sobre sua distribuição e quantidade (AMANN et al., 2001; GIARETTA et al., 2020). Essa técnica utiliza sondas de moléculas de rRNA marcadas com fluorescência para identificar as bactérias nos tecidos (AMANN et al., 2001; GARRAWAY et al., 2018; GIARETTA et al., 2020).

Microbioma na doença gastrointestinal

Nas doenças gastrointestinais em cães, ocorrem alterações na barreira intestinal, afetando a sua permeabilidade e desencadeando desequilíbrios na homeostase do intestino (HONNEFFER et al., 2014). Essas mudanças repentinas na composição e na atividade do microbioma podem desregular a resposta imunológica, levando à inflamação e aumentando a suscetibilidade a infecções oportunistas. Tais desequilíbrios estão associados a condições como enteropatias crônicas, diarreia aguda hemorrágica ou não-hemorrágica e até mesmo neoplasias em cães (ALESSANDRI et al., 2020; KIELER et al., 2016; MINAMOTO et al., 2012; SUCHODOLSKI & MARKEL, et al., 2012).

A inflamação intestinal resulta em mudanças na permeabilidade da membrana, levando a um aumento da disponibilidade de oxigênio no intestino, favorecendo o crescimento de bactérias anaeróbicas facultativas, como Enterobacteriaceae (ALESSANDRI et al., 2020; JANECZKO et al., 2008). Além disso, a inflamação interfere no metabolismo dos ácidos biliares, essenciais para a digestão de lipídios e com propriedades anti-inflamatórias. A disbiose intestinal resulta em concentrações reduzidas de ácidos biliares de cadeia curta, afetando a diferenciação de células T reguladoras (*Treg*) e contribuindo para o processo inflamatório (ALESSANDRI et al., 2020; SUCHODOLSKI, 2016).

Embora os estudos sobre as alterações do microbioma em doenças gastrointestinais em cães sejam abundantes, ainda não está claro se essas alterações são a causa ou o resultado das doenças inflamatórias intestinais. O diagnóstico dessas condições em cães geralmente ocorre após a exclusão de outras doenças, às vezes após tentativas de tratamento com antibióticos. A identificação de biomarcadores precoces e o desenvolvimento de intervenções terapêuticas são áreas importantes de pesquisa para o manejo dessas doenças (PILLA & SUCHODOLSKI, 2020; BLAKE & SUCHODOLSKI, 2016).

Estudos recentes têm se concentrado no papel do microbioma gastrointestinal em doenças crônicas gastrointestinais, como a doença inflamatória intestinal (do inglês *Inflammatory Bowel Disease*, IBD), principalmente em cães. Comparados com cães saudáveis, aqueles com IBD apresentam níveis reduzidos de ácidos graxos de cadeia curta e alterações na composição do microbioma fecal (MINAMOTO et al., 2019).

Doença inflamatória intestinal

A Doença Inflamatória Intestinal (IBD) em cães é caracterizada por sintomas gastrointestinais persistentes ou recorrentes, acompanhados de inflamação intestinal evidente nas análises histológicas (ETTINGER et al., 2017). A causa exata da IBD em cães ainda não é totalmente compreendida, mas fatores genéticos, composição do microbioma intestinal e respostas imunológicas desreguladas têm sido apontados como possíveis influências no desenvolvimento da doença (ETTINGER et al., 2017; JERGENS, 2012). Raças específicas, como *Basenjis, Soft coated Wheaton Terriers* e *Boxers*, demonstram predisposição para diferentes formas de IBD (JERGENS, 2012; HANDL et al., 2013; MAGRO et al., 2013). Além disso, a disbiose intestinal, uma alteração na composição do microbioma, está correlacionada com a inflamação gastrointestinal e o desenvolvimento subsequente da IBD (CASSMANN et al., 2016; HONNEFFER et al., 2014; LYU et al., 2020).

Os sintomas da IBD em cães podem variar e estão relacionados com os órgãos afetados, muitas vezes apresentando-se de forma intermitente devido às diferentes fases da doença, que podem estar ativas ou inativas. Os sinais clínicos mais comuns incluem diarreia, vômitos, anorexia e perda de peso (JERGENS, 2012). Em gatos, o vômito é o sintoma mais frequente, e a inflamação intestinal pode estar associada à inflamação pancreática e hepática, formando uma condição conhecida como triadite (JERGENS, 2012).

O diagnóstico da IBD geralmente é feito por exclusão, sendo necessário realizar exames clínicos e complementares para descartar outras doenças com sintomas semelhantes (JERGENS, 2012). A avaliação histopatológica das amostras de tecido gastrointestinal continua sendo o padrão ouro para o diagnóstico da IBD (COUËTIL et al., 2007). Biópsias obtidas por meio de endoscopia ou laparotomia exploratória são comuns para confirmar a inflamação intestinal (CERQUETELLA et al., 2010). Na endoscopia, as alterações visuais mais frequentes incluem mucosa friável, tecido de granulação e erosões (JERGENS, 2012). No entanto, é importante notar que biópsias inconclusivas não descartam a possibilidade de IBD em outras partes do trato gastrointestinal, e a interpretação histopatológica pode variar entre patologistas (CERQUETELLA et al., 2010; JERGENS, 2012). A padronização dos critérios de avaliação histológica é essencial para reduzir essas variações (COUËTIL et al., 2007; DAY et al., 2008).

Alterações do microbioma gastrointestinal em IBD

Os cães afetados pela Doença Inflamatória Intestinal (IBD) exibem mudanças notáveis na composição do microbioma gastrointestinal, incluindo um desequilíbrio geral conhecido como disbiose, juntamente com uma diminuição na presença de espécies que produzem ácidos graxos de cadeia curta (OMORI et al., 2017; WERNIMONT et al., 2020). Estudos destacam que cães diagnosticados com IBD frequentemente apresentam uma redução na abundância relativa de *Fusobacteria*, *Bacteroidetes* (especificamente *Bacterioidaceae* e *Prevotellaceae*), *Firmicutes* (incluindo *Megamonas*, *Ruminococcus*, *Faecalibacterium*, *Blautia*, *Turicibascter* e *Lachnospiracae*), em comparação com cães saudáveis (MINAMOTO et al., 2015; RAMADAN et al., 2014; SUCHODOLSKI, DOWD, et al., 2012; SUCHODOLSKI, MARKEL, et al., 2012).

Além disso, há um aumento observado no número de bactérias do gênero *Proteobacteria*, incluindo *Acinetobacter* e *Diaphorobacter*, em cães com IBD, em comparação com animais saudáveis. Especificamente, a família *Enterobacteriaceae*, particularmente *E. coli*, é mais prevalente na mucosa intestinal de cães com IBD (ALESSANDRI et al., 2020; MINAMOTO et al., 2015; SUCHODOLSKI, DOWD, et al., 2012). Evidências também sugerem que *E. coli* pode ser encontrada aderida ao epitélio celular ou dentro da mucosa intestinal, indicando uma associação direta com a doença (CASSMANN et al., 2016). Estudos recentes destacam que E. coli está aumentada em cães com enteropatias crônicas, enquanto *Helicobacter spp.* e *Akkermansia spp.* estão diminuídas (GIARETTA et al., 2020). *Akkermansia spp.*, embora comum no trato gastrointestinal humano, não mostra uma presença significativa em animais de estimação, e é associada à degradação do muco gastrointestinal em humanos (GARCIA-MAZCORRO et al., 2020).

A análise dessas populações microbianas pode variar dependendo dos métodos de identificação utilizados. Estudos demonstraram diferenças nos resultados entre métodos, como a identificação baseada em sequenciamento do gene 16S rRNA e análise de bibliotecas genômicas de clones 16S rRNA. Por exemplo, enquanto alguns métodos mostram maior abundância de Proteobacteria, outros revelam uma menor presença de *Clostridia*, *Fusobacteria*, *Bacteroidaceae* e *Prevotellaceae* (SUCHODOLSKI et al., 2010; SUCHODOLSKI, DOWD, et al., 2012).

Embora a maioria dos estudos se baseie em amostras fecais, a informação sobre a distribuição espacial da microbiota na mucosa é limitada. Estudos utilizando a técnica de

hibridização in situ fluorescente (FISH) mostraram diferenças na composição da microbiota entre cães com IBD e cães saudáveis. Observou-se um aumento de certas bactérias na mucosa do íleo e ceco em cães com IBD, incluindo *Bacteroides* e *Enterobacteriaceae*, enquanto amostras do cólon revelaram um aumento de E. coli (CASSMANN et al., 2016). Essas variações nos resultados podem ser atribuídas não apenas a diferenças nos métodos de análise, mas também à diversidade natural do microbioma, que varia entre indivíduos e é influenciada por fatores como idade, ambiente e uso de antibióticos (ALESSANDRI et al., 2020).

CONSIDERAÇÕES FINAIS

A análise do microbioma gastrointestinal revela sua importância fundamental para a saúde do trato gastrointestinal dos cães. Em particular, condições crônicas como a Doença Inflamatória Intestinal (IBD) destacam-se pela sua associação com alterações marcantes nesse ecossistema microbiano. A aplicação de técnicas avançadas, como PCR, Hibridização in situ por Fluorescência (FISH) e Sequenciamento de Nova Geração (NGS), oferece uma compreensão mais profunda e precisa das comunidades bacterianas presentes no intestino dos cães. Esses métodos permitem identificar não apenas as espécies de bactérias, mas também sua abundância relativa e diversidade, facilitando comparações entre cães saudáveis e aqueles com condições gastrointestinais.

Estudos recentes demonstraram diferenças significativas na composição microbiana entre cães saudáveis e aqueles diagnosticados com IBD. Em casos de IBD, observa-se um desequilíbrio na microbiota intestinal, com redução de certas espécies benéficas, como *Bacteroides* e *Prevotella*, e um aumento de bactérias potencialmente prejudiciais, como *E. coli* e *Proteobacteria*. Essas descobertas não são exclusivas de cães; pesquisas em outras espécies, como gatos, humanos e ratos de laboratório, também destacam a importância do microbioma na saúde gastrointestinal.

Portanto, compreender a composição e a dinâmica do microbioma gastrointestinal dos cães não apenas pode guiar estratégias terapêuticas mais eficazes para condições gastrointestinais, como IBD, mas também pode fornecer insights valiosos para o prognóstico dessas doenças em animais de companhia.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e da Fundação de Amparo ao Ensino e Pesquisa (FAEP) da Universidade de Mogi das Cruzes (UMC).

REFERÊNCIAS

ALESSANDRI, M. et al. The role of the gut microbiome in the health and disease of dogs: A review. **Microbiome**, v. 8, p. 1-14, 2020.

AMANN, R. I. et al. Hibridização in situ fluorescente para estudos de microbiota. **Nature Reviews Microbiology**, v. 5, n. 10, p. 873-883, 2001.

BLAKE, A. et al. The gut microbiome and its role in gastrointestinal disease in dogs. **Veterinary Clinics of North America: Small Animal Practice**, v. 46, n. 3, p. 363-377, 2016.

Cassmann, E. et al. Alterations of the Ileal and Colonic Mucosal Microbiota in Canine Chronic Enteropathies. **PloS one**, v. 11 n. 2 p. e0147321, 2016. https://doi.org/10.1371/journal.pone.0147321

CERQUETELLA, M. et al. Diagnostic approach to inflammatory bowel disease in dogs. **Veterinary Clinics of North America: Small Animal Practice**, v. 40, n. 6, p. 1097-1111, 2010.

COUËTIL, L. et al. Histopathological diagnosis of inflammatory bowel disease. **Journal of Veterinary Internal Medicine**, v. 21, n. 6, p. 1341-1348, 2007.

DENG, P. et al. Nutrition and the gut microbiome in dogs. **Journal of Animal Science**, v. 93, n. 5, p. 2515-2527, 2015.

DEUSCH, O. et al. Microbiome diversity in dogs and cats: Influence of diet and environmental factors. **Veterinary Microbiology**, v. 181, n. 3-4, p. 215-220, 2015.

ETTINGER, S. J. et al. **Textbook of veterinary internal medicine**. W.B. Saunders, 2017.

GARRAWAY, T. et al. Advances in FISH for the study of gastrointestinal microbiomes. **Frontiers in Microbiology**, v. 9, p. 1973, 2018.

GIARETTA, S. et al. The role of the gut microbiome in inflammatory bowel disease. **Journal of Animal Science and Technology**, v. 62, n. 1, p. 1-10, 2020.

GARCIA-MAZCORRO, J. et al. Akkermansia spp. in human and veterinary medicine. **Microbial Ecology in Health and Disease**, v. 31, p. 1832061, 2020.

GRZEŚKOWIAK, L. et al. The impact of diet on gut microbiota in dogs. **Journal of Animal Science and Technology**, v. 57, n. 1, p. 1-8, 2015.

HAMADY, M.; KNIGHT, R. Microbial community profiling for environmental microbiology. **Current Protocols in Microbiology**, v. 1, n. 1, p. 1-24, 2009.

HANDL, S. et al. Comparative analysis of gut microbiota in dogs and cats. **Veterinary Microbiology**, v. 151, n. 3-4, p. 156-163, 2011.

HONNEFFER, J. et al. Gut microbiota in dogs: Influence of diet and health status. **Veterinary Microbiology**, v. 170, n. 3-4, p. 131-140, 2014.

HONNEFFER, J. et al. Composition of the gut microbiota in healthy dogs: A metagenomic study. **Veterinary Research**, v. 48, n. 1, p. 10, 2017.

JHA, A. et al. The gut microbiome of dogs: Implications for health and disease. **Frontiers in Veterinary Science**, v. 7, p. 212, 2020.

JANECZKO, S. et al. Gastrointestinal diseases in dogs: The role of microbiota. **Journal of Veterinary Internal Medicine**, v. 22, n. 3, p. 635-642, 2008.

JERGENS, A. E. Inflammatory bowel disease in dogs: A review. **Veterinary Clinics of North America: Small Animal Practice**, v. 42, n. 4, p. 727-740, 2012.

KIELER, I. N. et al. The role of gut microbiota in canine gastrointestinal disease. **Journal of Animal Physiology and Animal Nutrition**, v. 100, n. 5, p. 883-890, 2016.

LYU, Z. et al. Canine gut microbiome: An overview of its role in health and disease. **Journal of Animal Science**, v. 98, n. 2, p. 1-15, 2020.

MAGRO, F. et al. IBD in dogs and cats: Clinical aspects and treatment. **Journal of Veterinary Internal Medicine**, v. 27, n. 5, p. 1194-1203, 2013.

MINAMOTO, Y. et al. Microbiome in dogs with inflammatory bowel disease. **Veterinary Research**, v. 43, n. 1, p. 1-9, 2012.

MINAMOTO, Y. et al. Dysbiosis of gut microbiota in dogs with inflammatory bowel disease. **Veterinary Microbiology**, v. 179, n. 1-2, p. 43-51, 2015.

MINAMOTO, Y. et al. Changes in the gut microbiota of dogs with inflammatory bowel disease. **Journal of Veterinary Internal Medicine**, v. 33, n. 6, p. 2183-2191, 2019.

MILANI, C. et al. Comparative genomics of the human gut microbiome. **Microbiome**, v. 2, p. 1-10, 2014.

MILANI, C. et al. The human gut microbiome: A new frontier for human health. **Nature Reviews Microbiology**, v. 16, n. 2, p. 55-64, 2018.

OMORI, Y. et al. Gut microbiota and inflammatory bowel disease: Insights from canine studies. **Journal of Animal Science**, v. 95, n. 1, p. 335-345, 2017.

POSSEMIERS, T. et al. The gut microbiome as a metabolic organ. **Nature Reviews Gastroenterology & Hepatology**, v. 8, n. 5, p. 241-248, 2011.

PILLA, R. et al. Therapeutic approaches for managing gut dysbiosis in dogs. **Veterinary Clinics of North America: Small Animal Practice**, v. 50, n. 5, p. 1123-1137, 2020.

RITCHIE, K. et al. The gastrointestinal microbiota of dogs and its implications for health. **Veterinary Clinics of North America: Small Animal Practice**, v. 38, n. 4, p. 673-690, 2008.

RAMADAN, A. et al. The role of the gut microbiome in canine IBD. **Veterinary Record**, v. 175, n. 7, p. 172, 2014.

SUCHODOLSKI, J. S. The gastrointestinal microbiome of dogs: A review. **Veterinary Clinics of North America: Small Animal Practice**, v. 41, n. 4, p. 671-688, 2011a.

SUCHODOLSKI, J. S. The role of the gut microbiota in health and disease. **Journal of Animal Science and Technology**, v. 53, n. 1, p. 1-8, 2011b.

SUCHODOLSKI, J. et al. Characterization of the fecal microbiota in dogs with diarrhea. **Veterinary Microbiology**, v. 128, n. 1-2, p. 35-43, 2008.

SUCHODOLSKI, J. et al. The role of gut microbiota in canine inflammatory bowel disease. **Veterinary Journal**, v. 192, n. 1, p. 18-24, 2012.

TIZARD, I.; JONES, G. The microbiome and its impact on the immune system of dogs. **Journal of Veterinary Internal Medicine**, v. 32, n. 1, p. 13-25, 2018.

WERNIMONT, T. et al. Nutrition and gut health in dogs: A review. **Veterinary Clinics of North America: Small Animal Practice**, v. 50, n. 6, p. 1027-1041, 2020.

YANG, D. et al. The impact of gut microbiome on health and disease. **Nature Reviews Microbiology**, v. 13, n. 8, p. 1-16, 2015.