



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA
CURSO DE GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS

ANA CLARICE DA SILVA GADELHA

CHROMOBACTERIACEAE: UMA ANÁLISE FILOGENÉTICA MULTILOCUS

FORTALEZA

2021

ANA CLARICE DA SILVA GADELHA

CHROMOBACTERIACEAE: UMA ANÁLISE FILOGENÉTICA MULTILOCUS

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Ciências Biológicas do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Ciências Biológicas.

Orientador: Prof. Dr. Thalles Barbosa Grangeiro

FORTALEZA

2021

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

- G12c Gadelha, Ana Clarice da Silva.
Chromobacteriaceae: uma análise filogenética multilocus / Ana Clarice da Silva Gadelha. – 2021.
80 f. : il. color.
- Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Ciências,
Curso de Ciências Biológicas, Fortaleza, 2021.
Orientação: Prof. Dr. Thalles Barbosa Grangeiro.
1. Chromobacteriaceae. 2. Chromobacterium violaceum. 3. Bioinformática. 4. Análise Filogenética.
I. Título.

CDD 570

ANA CLARICE DA SILVA GADELHA

CHROMOBACTERIACEAE: UMA ANÁLISE FILOGENÉTICA MULTILOCUS

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Ciências Biológicas do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Ciências Biológicas.

Aprovada em:

BANCA EXAMINADORA

Prof. Dr. Thalles Barbosa Grangeiro (Orientador)
Universidade Federal do Ceará (UFC)

Profa. Dra. Vânia Maria Maciel Melo
Universidade Federal do Ceará (UFC)

Profa. Dra. Suelen Carneiro de Medeiros
Instituto Federal de Educação, Ciência e Tecnologia
do Ceará (IFCE)

À minha família, por sua capacidade de acreditar em mim e investir em mim. Mãe, seu cuidado e dedicação foi que deram, em alguns momentos, a esperança para seguir. Pai, sua presença significou segurança e certeza de que não estou sozinho nessa caminhada.

AGRADECIMENTOS

Ao meu marido, Stalin, pelo amor, paciência, incentivo e companheirismo.

À minha família, por acreditar, investir em mim e tornar possível todas as minhas conquistas.

Ao meu orientador, Prof. Dr. Thalles Barbosa Grangeiro, pelo acompanhamento, estreitamento da relação professor-aluno e exemplo de profissional. Além do incentivo, apoio e sugestões durante a supervisão dos meus estudos.

Ao Doutorando em Engenharia de Teleinformática e grande amigo Bruno Riccelli, pela grande ajuda, apoio e sugestões durante os estudos.

A minha colega de curso Danielly, pela amizade e pelos momentos de descontração e estudo.

À todos os professores do curso de Ciências Biológicas da Universidade Federal do Ceará pela excelência da qualidade técnica de cada um.

Ao Doutor em Engenharia Elétrica, Ednardo Moreira Rodrigues, e seu assistente, Alan Batista de Oliveira, graduado em Engenharia Elétrica, pela adequação do *template* utilizado neste trabalho para que o mesmo ficasse de acordo com as normas da biblioteca da Universidade Federal do Ceará (UFC).

”Por aqui, no entanto, não olhamos para trás por muito tempo. Seguimos avançando, abrindo novas portas e fazendo coisas novas, porque somos curiosos ... e a curiosidade continua nos conduzindo por novos caminhos. Continue seguindo em frente.”

(Walt Disney)

RESUMO

A família *Chromobacteriaceae* (*Betaproteobacteria*) é um grupo que contém espécies de grande importância biotecnológica, incluindo a espécie *Chromobacterium violaceum*. Devido ao avanço das técnicas de bioinformática, constantemente novas espécies são incluídas a essa família. Desde que a família foi proposta, apesar de sua importância e de seu constante crescimento, nenhum novo estudo focou sua filogenia. Neste trabalho, foram utilizadas ferramentas de bioinformática (BLAST, MAFFT e RAxML-NG) para a construção de duas árvores filogenéticas, uma baseada nas sequências de aminoácidos de 19 genes de manutenção (Adk, AtpG, EF-G, FusA, GltX, Gmk, GyrA, GyrB, L5, LipA, MurC, MurE, PgK, RpoB, S2, S3, SecY, ThrRS, UvrD) e outra baseada nas sequências de nucleotídeos dos genes rRNA 16S e 23S. De acordo com as árvores obtidas, a família *Chromobacteriaceae* divide-se em dois grandes ramos, um englobando as espécies dos gêneros *Paludibacterium*, *Gulbenkiania*, *Pseudogulbenkiania*, *Vogesella*, *Aquitalea* e *Chromobacterium* (Clado 1) e outro englobando as espécies dos gêneros *Aquaspirillum*, *Microvirgula*, *Laribacter*, *Leeia*, *Iodobacter*, *Formivibrio*, *Andreprevotia*, *Chitiniphilus*, *Jeongeupia*, *Chitinilyticum*, *Deefgea* e *Chitinibacter* (Clado 2). Foi revelada também, uma possível inconsistência na taxonomia da estirpe *Pseudogulbenkiania ferrooxidans* EGD-HP2 em ambas as árvores.

Palavras-chave: *Chromobacteriaceae*. *Chromobacterium violaceum*. Bioinformática. Análise Filogenética.

ABSTRACT

The family *Chromobacteriaceae* (*Betaproteobacteria*) is a group containing species of great biotechnological importance, including a *Chromobacterium violaceum* species. Required to advance bioinformatics techniques, new species are constantly included in this family. Since the family was proposed, despite its importance and steady growth, no new study has focused on its phylogeny. In this work, bioinformatics tools (BLAST, MAFFT and RAxML-NG) were used to build two phylogenetic trees, one based on amino acid sequences of 19 housekeeping genes (Adk, AtpG, EF-G, FusA, GltX, Gmk, GyrA, GyrB, L5, LipA, MurC, MurE, PgK, RpoB, S2, S3, SecY, ThrRS, UvrD), the other based on nucleotide sequences of 16S/23S rRNA genes. According to the trees, *Chromobacteriaceae* family is divided into two large branches, one containing the genera *Paludibacterium*, *Gulbenkiania*, *Pseudogulbenkiania*, *Vogesella*, *Aquitalea* and *Chromobacterium* (Clado 1), the other containing the genera *Aquaspirillum*, *Microvirgula*, *Laribacter*, *Leeia*, *Iodobacter*, *Formivibrio*, *Andreprevotia*, *Chitiniphilus*, *Jeongeupia*, *Chitinilyticum*, *Deefgea* and *Chitinibacter* (Clado 2). It was also revealed a possible taxonomic inconsistency of the strain *Pseudogulbenkiania ferrooxidans* EGD-HP2 in both trees.

Palavras-chave: *Chromobacteriaceae*. *Chromobacterium violaceum*. Bioinformatics. Phylogenetic Analysis.

LISTA DE FIGURAS

Figura 1 – Dogma central da biologia molecular	18
Figura 2 – Estrutura básica de um nucleotídeo.	18
Figura 3 – Estrutura básica de um aminoácido.	20
Figura 4 – Código genético universal	21
Figura 5 – Exemplo de alinhamento simples local utilizando a ferramenta BLAST . . .	24
Figura 6 – Exemplo de alinhamento múltiplo global utilizando a ferramenta MAFFT .	25
Figura 7 – Matriz de custo MAT70	26
Figura 8 – Matriz de custo BLOSUM80	27
Figura 9 – Comparação entre métodos de inferência filogenética	30
Figura 10 – Visão geral de uma árvore filogenética.	33
Figura 11 – Possíveis representações de uma mesma árvore filogenética.	34
Figura 12 – Tipos de classificações de táxons.	35
Figura 13 – Diagrama de fluxo da metodologia aplicada	50
Figura 14 – Árvore ML baseada nos genes 16/23S rRNA	58
Figura 15 – Árvore ML baseada nos genes de manutenção	59
Figura 16 – Clado I da árvore ML baseada nos genes 16/23S rRNA	61
Figura 17 – Clado II da árvore ML baseada nos genes 16/23S rRNA	62
Figura 18 – Gênero <i>Chromobacterium</i> na árvore ML baseada nos genes 16/23S rRNA .	63
Figura 19 – Gênero <i>Chromobacterium</i> na árvore ML baseada nos genes 16/23S rRNA (continuação)	64
Figura 20 – Árvore ML baseada nos genes de manutenção - Clado I	65
Figura 21 – Árvore ML baseada nos genes de manutenção - Clado II	66
Figura 22 – Árvore ML baseada nos genes de manutenção - <i>Chromobacterium</i>	67
Figura 23 – Árvore ML baseada nos genes de manutenção - <i>Chromobacterium</i> (continuação)	68

LISTA DE TABELAS

Tabela 1 – Relatório de montagem e anotação dos genomas	51
Tabela 2 – Parâmetros utilizados no alinhamento local	54
Tabela 3 – Blocos de genes concatenados para cada arquivo	56
Tabela 4 – Resultados do teste de modelo de substituição	57

LISTA DE ABREVIATURAS E SIGLAS

MSA	Alinhamento Múltiplo de Sequências
CSIs	Deleções e Inserções de Assinatura Conservadas
DNA	Ácido Desoxirribonucleico
RNA	Ácido Ribonucleico
INDELS	inserções e deleções
HSP	Pares de Segmentos de Alta Pontuação
NJ	Vizinho Mais Próximo
UPGMA	Ligação Média entre Grupos
MP	Máxima Parcimônia
ML	Máxima Verossimilhança
BI	Inferência Bayesiana
MCMC	Monte Carlo via Cadeiras de Markov
LPSN	Lista de Nomes Procarióticos com Posição na Nomenclatura
PHB	polihidroxibutirato
hLRT	Teste de Razão de Verossimilhança Hierárquica
AIC	Critério de Informação Akaike
BPP	Probabilidades Bayesianas Posteriores
NCBI	Centro Nacional de Informação Biotecnológica
RefSeq	Banco de Dados de Sequências de Referência
AICc	Critério de Informação Akaike Corrigido
DT	Teoria da Decisão

LISTA DE SÍMBOLOS

mol

molar

μm

micrômetro

SUMÁRIO

1	INTRODUÇÃO	15
1.1	Objetivos	16
1.2	Organização do Trabalho	16
2	FUNDAMENTAÇÃO TEÓRICA	17
2.1	Análise Filogenética	17
2.1.1	Informação Biológica	17
2.1.1.1	Ácidos Nucleicos	18
2.1.1.2	Proteínas	20
2.1.1.3	Mutações	21
2.1.1.4	Sequências Genéticas	22
2.1.2	Alinhamento de sequências	22
2.1.2.1	Programação Dinâmica	25
2.1.2.2	K-Tuple	28
2.1.2.3	Alinhamento Progressivo	28
2.1.2.4	Pontuação Baseada em Consistência	28
2.1.2.5	Refinamento Iterativo	29
2.1.2.6	Significância estatística	29
2.1.3	Inferência Filogenética	30
2.1.3.1	Métodos quantitativos	30
2.1.3.2	Métodos qualitativos	31
2.1.3.3	Confiabilidade	32
2.1.4	Representação das Árvores	32
2.2	Chromobacteriaceae	36
2.2.1	Características Gerais	36
2.2.2	Diversidade de Espécies	36
2.2.2.1	<i>Andreprevotia</i>	36
2.2.2.2	<i>Aquaspirillum</i>	37
2.2.2.3	<i>Aquitalea</i>	37
2.2.2.4	<i>Chitinibacter</i>	38
2.2.2.5	<i>Chitinilyticum</i>	38

2.2.2.6	<i>Chitiniphilus</i>	38
2.2.2.7	<i>Chromobacterium</i>	39
2.2.2.8	<i>Deefgea</i>	39
2.2.2.9	<i>Formivibrio</i>	40
2.2.2.10	<i>Gulbenkiania</i>	40
2.2.2.11	<i>Iodobacter</i>	41
2.2.2.12	<i>Jeongeupia</i>	41
2.2.2.13	<i>Laribacter</i>	41
2.2.2.14	<i>Leeia</i>	42
2.2.2.15	<i>Microvirgula</i>	42
2.2.2.16	<i>Paludibacterium</i>	42
2.2.2.17	<i>Pseudogulbenkiania</i>	43
2.2.2.18	<i>Silvimonas</i>	43
2.2.2.19	<i>Vogesella</i>	43
2.2.3	<i>Importância Biotecnológica</i>	44
2.2.4	<i>Patogenicidade</i>	45
3	TRABALHOS RELACIONADOS	47
4	METODOLOGIA	50
4.1	Coleta de Dados	50
4.2	Alinhamento Local	54
4.3	Alinhamento Múltiplo	55
4.4	Particionamento e Modelo de Substituição	55
4.5	Reconstrução Filogenética	55
5	RESULTADOS E DISCUSSÃO	56
6	CONCLUSÕES E TRABALHOS FUTUROS	70
	REFERÊNCIAS	71

1 INTRODUÇÃO

A bioinformática é uma área multidisciplinar, relativamente nova, que aplica técnicas estatísticas e computacionais na biologia molecular. Na última década, a bioinformática avançou rapidamente e ganhou grande importância como campo de estudo (CATTLEY; ARTHUR, 2007). A bioinformática abriu novas fronteiras para as análises de genômica e proteômica permitindo, dentre outros, criar banco de dados de informações biológicas, reconhecer sequências de genes, organizar e relacionar informação biológica, agrupar proteínas homólogas e estabelecer árvores filogenéticas (ARAÚJO *et al.*, 2008).

Análises filogenéticas são, atualmente, uma parte padrão da análise de sequências de nucleotídeos e aminoácidos, sendo capazes de mostrar relações evolutivas entre organismos (WU; EISEN, 2008). Dessa forma, a utilização das ferramentas de bioinformática nesse estudo é essencial. Com o aprimoramento das técnicas de sequenciamento gênico, os bancos de dados biológicos cresceram de forma explosiva e esse crescimento oferece uma oportunidade para aumentar o uso de marcadores de proteínas para inferência filogenética (WU; EISEN, 2008).

Árvores filogenéticas construídas a partir do gene 16S rRNA se tornaram a base para a classificação de organismos procarióticos (WU; EISEN, 2008), mas apesar do 16S rRNA ser vastamente utilizado em análises filogenéticas, inferir a filogenia de organismos baseando-se em um único gene gera alguns riscos e deve ser corroborado com o uso de outros marcadores filogenéticos (WU; EISEN, 2008). Genes de manutenção são escolhidos preferencialmente em esquemas de Alinhamento Múltiplo de Sequências (MSA) devido seu maior poder de discriminação quando comparado ao 16S rRNA (GLAESER; KÄMPFER, 2015 apud LIU *et al.*, 2017).

A família Chromobacteriaceae é dividida atualmente em 19 gêneros, sendo *Chromobacterium* o gênero que possui a maior quantidade de espécies descritas. A espécie *Chromobacterium violaceum* é uma espécie de grande potencial biotecnológico, sendo a estipe ATCC 12472 um dos organismos a ter seu genoma completamente sequenciado no Projeto Genoma Brasileiro e é alvo de várias pesquisas microbiológicas.

Agrupamentos confiáveis de microrganismos em diferentes grupos e a natureza de seu relacionamento entre si formam o fundamento de toda pesquisa microbiológica (ADEOLU; GUPTA, 2013). Entretanto, as inter-relações entre os organismos que compõem essa família não foram bem estudadas. O agrupamento dos organismos pertencentes a família *Chromobacteriaceae* foi definido por Adelou e Gupta (2013), pela presença de Deleções e Inserções de Assinatura

Conservadas (CSIs) específicas para a ordem *Neiserialles* e ausência de CSIs específicas para a família *Neisseriaceae*. Desde que ele propôs a família, novas espécies foram adicionadas ao grupo, mas nenhum novo estudo foi realizado focando exclusivamente sua filogenia. Assim, com esse estudo pretende-se dar um passo à frente na resolução dessa questão.

1.1 Objetivos

Objetivo Geral

O objetivo geral deste trabalho foi realizar um estudo filogenético da família *Chromobacteriaceae* utilizando ferramentas e técnicas de bioinformática.

Objetivos Específicos

A partir do objetivo geral, podemos traçar os seguintes objetivos específicos:

- Coletar sequências genéticas disponíveis em bancos de dados biológicos.
- Selecionar diferentes grupos de genes para a reconstrução filogenética.
- Realizar e analisar os alinhamentos dos genes utilizados no estudo.
- Representar e analisar as relações evolutivas entre as espécies utilizadas no estudo.

1.2 Organização do Trabalho

Este trabalho está dividido em seis capítulos. No Capítulo 2 está a fundamentação teórica dos temas abordados. No Capítulo 3 são discutidos os trabalhos relacionados. O Capítulo 4 apresenta a metodologia, descrevendo e explicando as análises realizadas nesse trabalho. Os resultados e discussões são abordados no Capítulo 5. Por fim, no Capítulo 6 são sintetizadas as conclusões, assim como as perspectivas de trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

A fim de uma completa compreensão deste trabalho, faz-se necessário expor alguns conceitos básicos que estão relacionados as análises filogenéticas. Logo após, será apresentada uma visão geral da família em estudo.

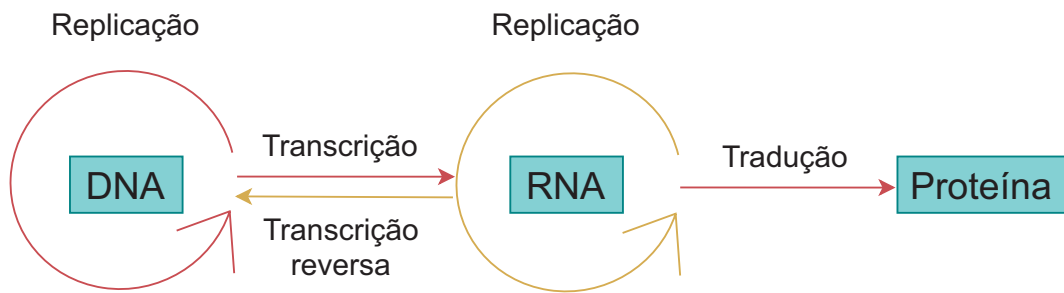
2.1 Análise Filogenética

Nesta seção são apresentados conceitos básicos para se compreender uma análise filogenética. A seção está dividida em quatro partes e seu conteúdo está fortemente baseado nos capítulos 2, 3 e 5, respectivamente, do livro *Bioinformática: da Biologia à Flexibilidade Molecular* de Verli (2014). As informações apresentadas na primeira e quarta parte desta seção se basearam também nos livros *Princípios de Bioquímica* de Lehninger de Nelson e Cox (2018) e *Fundamentos de Filogenética Sistemática* de Wägele (2005). A primeira parte desta seção apresenta conceitos sobre informações biológicas. A segunda descreve conceitos sobre alinhamento de sequências e os principais algoritmos de alinhamento. A terceira descreve conceitos relacionados a inferência filogenética e os principais métodos de inferência.

2.1.1 Informação Biológica

Apesar de apresentar grandes diferenças em suas características gerais, todos os seres vivos são compostos basicamente pelos mesmos tipos de moléculas. As informações que regem como a vida irá se manifestar estão contidas no Ácido Desoxirribonucleico (DNA). No interior das células, porém, essas informações não são utilizadas diretamente. A célula precisa de uma molécula intermediária chamada Ácido Ribonucleico (RNA), que é sintetizada a partir do DNA em um processo chamado transcrição. O RNA funciona como um molde e é a partir dele que a informação vai ser traduzida e expressa na forma de proteínas. Todos esses processos, desde o DNA até a síntese das proteínas, são definidos como o Dogma Central da Biologia Molecular (Figura 1) e seguem um fluxo de informação universal para sistemas biológicos (VERLI, 2014).

Figura 1 – Dogma central da biologia molecular

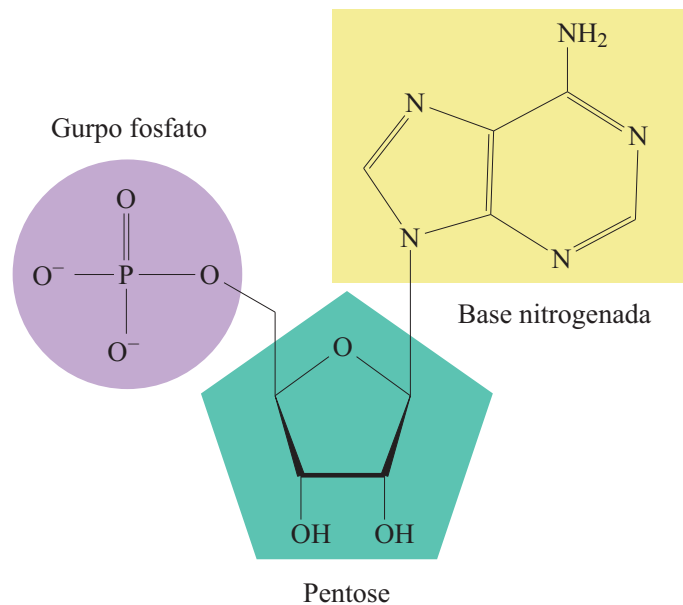


Fonte: adaptado de (NELSON; COX, 2018).

2.1.1.1 Ácidos Nucleicos

Ácidos nucleicos são polímeros formados por unidades menores chamadas nucleotídeos, que se conectam uns aos outros através de ligações fosfodiéster. Os nucleotídeos são formados basicamente por uma pentose, um grupo fosfato e uma base nitrogenada. A Figura 2 mostra a estrutura básica de um nucleotídeo.

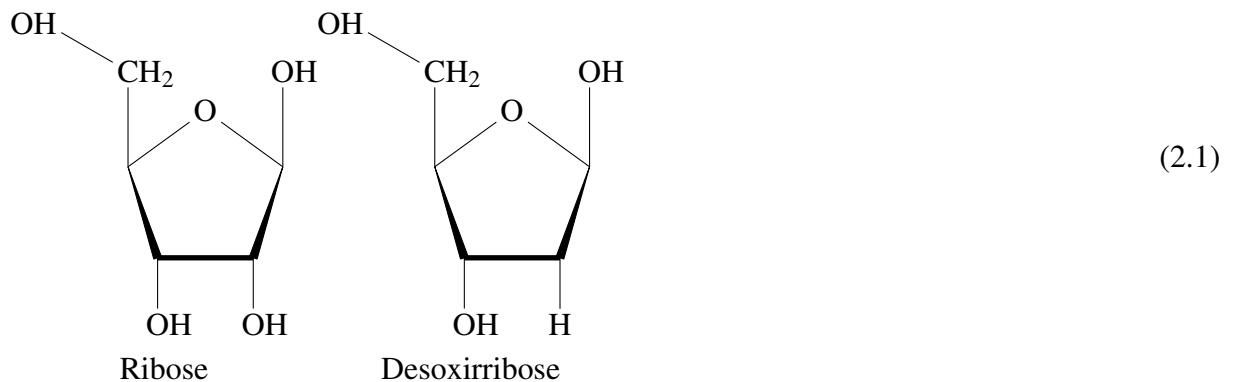
Figura 2 – Estrutura básica de um nucleotídeo.



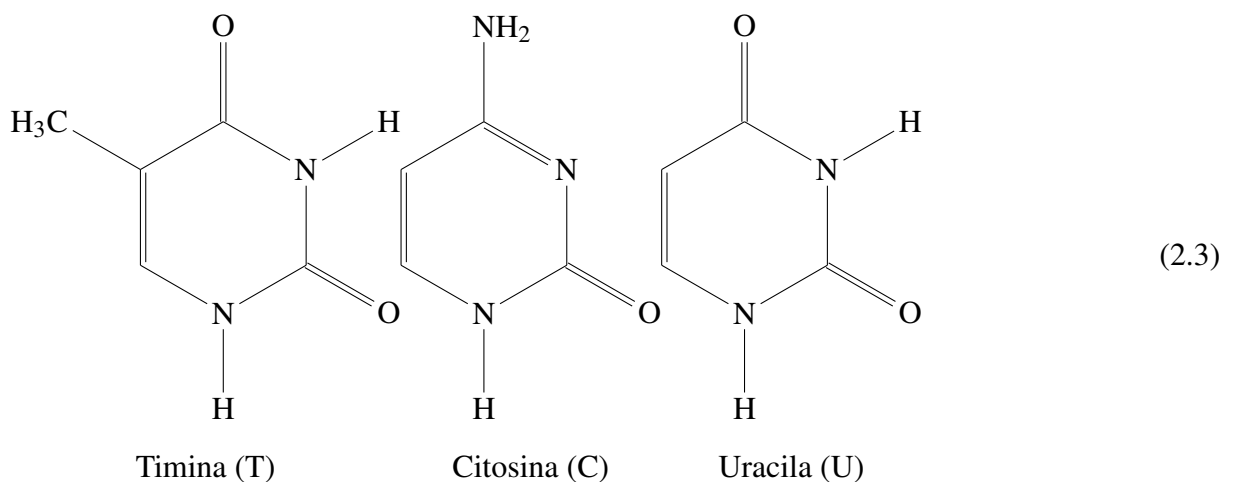
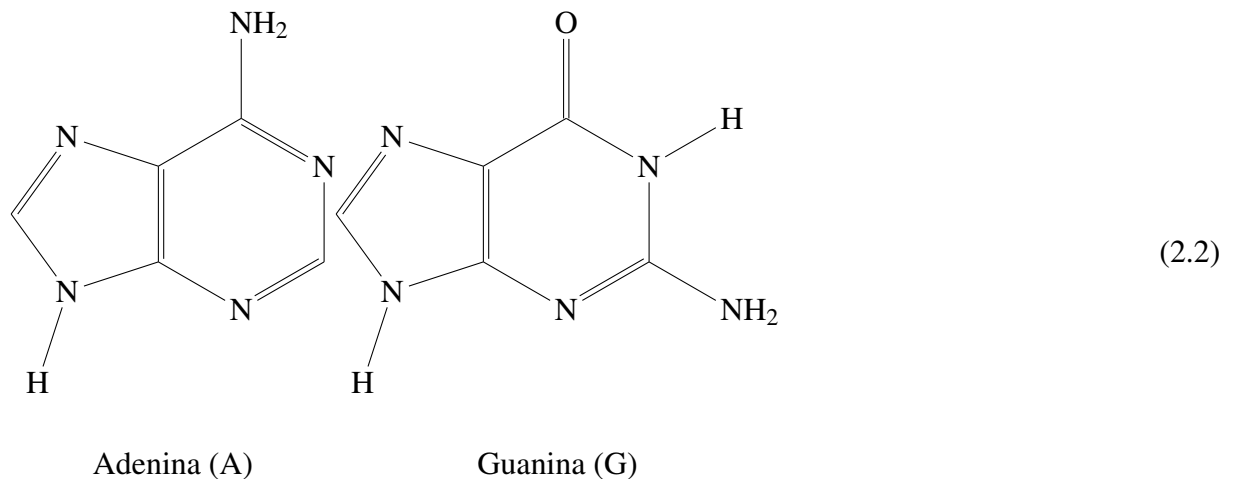
Fonte: adaptado de (NELSON; COX, 2018).

A pentose pode ser de dois tipos β -D-ribose (ribose) ou a 2-desoxi- β -D-ribose

(desoxirribose), com as seguintes fórmulas estruturais:



As bases nitrogenadas podem ser classificadas em dois tipos, purinas (adenina e guanina) e pirimidinas (timina, citosina e uracila), com as seguintes fórmulas estruturais respectivamente:

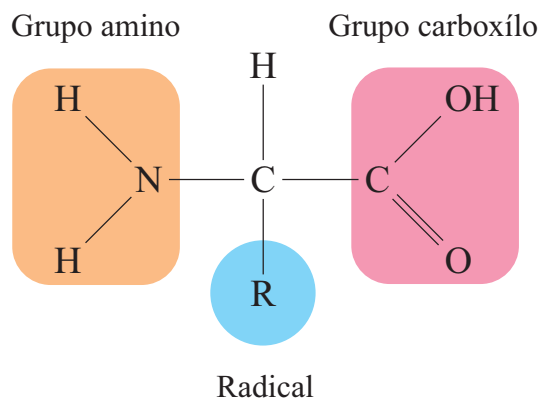


Numa molécula de DNA os desoxirribonucleotídeos formam uma cadeia dupla de aminoácidos, onde guanina se liga a citosina ($G \equiv C$) e adenina se liga a timina ($A \doteq T$). A molécula de RNA, por sua vez, é formada por apenas uma cadeia simples de ribonucleotídeos e apresenta mais uma outra característica diferente: timina (T) é substituída por uracila (U).

2.1.1.2 Proteínas

Como visto anteriormente, as proteínas são o produto direto da informação genética. Esses polímeros são formados por unidades menores chamadas aminoácidos e ligam-se uns aos outros por ligações peptídicas. Todos aminoácidos são compostos por um grupo amina, um grupo carboxílico e um radical ligado ao carbono α . A Figura 3 mostra a estrutura básica de um aminoácido.

Figura 3 – Estrutura básica de um aminoácido.



Fonte: adaptado de (NELSON; COX, 2018).

Durante a síntese proteica cada triplete de nucleotídeos irá codificar para um aminoácido, possibilitando ao todo codificar 20 aminoácidos diferentes. Os códons e seus respectivos aminoácidos podem ser observados na Figura 4.

Figura 4 – Código genético universal

	U	C	A	G	
U	UUU } Phe (F) UUC } UUA } Leu (L) UUG }	UCU } UCC } Ser(S) UCA } UCG }	UAU } Tyr (Y) UAC } UAA } Stop UAG }	UGU } Cys (C) UGC } UGA } Stop UGG } Trp (W)	U C A G
C	CUU } CUC } Leu (L) CUA } CUG }	CCU } CCC } Pro (P) CCA } CCG }	CAU } His(H) CAC } CAA } Gln (Q) CAG }	CCU } CGC } Arg (R) CGA } CGG }	U C A G
A	AUU } AUC } Ile (I) AUA } AUG } Met (M)	ACU } ACC } Thr (T) ACA } ACG }	AAU } Asn (N) AAC } AAA } Lys (K) AAG }	AGU } Ser(S) AGC } AGA } Arg (R) AGG }	U C A G
G	GUU } GUC } Val (V) GUA } GUG }	GCU } GCC } Ala (A) GCA } GCG }	GAU } Asp (D) GAC } GAA } Glu (E) GAG }	GGU } GGC } Gly (G) GGA } GGG }	U C A G

Nota: Ala - Alanina, Arg - Arginina, Asn - Aspargina, Asp - Ácido aspártico, Cys - Cisteína, Gln - Glutamina, Glu - Ácido Glutâmico, Gly - Glicina, His - Histidina, Ile - Isoleucina, Leu - Leucina, Lys - Lisina, Met - Metionina/Códon de início, Phe - Fenilalanina, Pro - 'Prolina, Ser - Serina, Stop - Códon de parada, Thr - Treonina, Trp - Triptofano, Tyr - Tirosina, Val - Valina.

Fonte: adaptado de (NELSON; COX, 2018).

2.1.1.3 Mutações

Durante os processos químicos e biológicos normais do organismo, o DNA, assim como a informação armazenada por ele, pode sofrer modificações. Essas modificações recebem o nome de mutações e são uma das fontes primárias da diversidade genética (CHOUDHURI, 2014).

Uma mutação ocorre quando um par de bases extra é inserido na cadeia polinucleotídica ou quando um par de bases originais é deletado ou quando uma par de bases é substituída por outro. A substituição de uma purina por uma pirimidina, ou vice-versa, é chamada de transversão, enquanto uma substituição de uma purina por uma purina ou uma pirimidina por

pirimidina é chamada de transição.

Como visto anteriormente, cada códon corresponde a um aminoácido. Uma única alteração nos pares de bases pode modificar o aminoácido codificado. As substituições podem gerar três tipos de mutação: mutação silenciosa, mutação de sentido trocado ou mutação sem sentido. Uma mutação silenciosa é quando a troca de um nucleotídeo não altera o aminoácido codificado, geralmente quando a substituição ocorre na primeira ou segunda posição do códon. Uma mutação de sentido trocado é quando a troca de um nucleotídeo altera o aminoácido codificado, geralmente quando a substituição ocorre na terceira posição do códon. Mutação sem sentido é quando a troca de um nucleotídeo gera um códon de parada. Diferente das substituições, as inserções e deleções (INDELS) modificam a fase de leitura do DNA, ou seja, toda a cadeia de aminoácidos será alterada partindo de onde houve a mutação.

2.1.1.4 Sequências Genéticas

As informações contidas nos ácidos nucleicos podem ser representadas por longas séries de letras, possibilitando que sejam analisadas computacionalmente. Em sua maioria, essas séries de letras são provenientes de técnicas de sequenciamento do DNA, devido as regras de pareamento das bases e a facilidade na conversão dos códons (STEVENS; BOUCHER, 2015). As sequências podem ser de dois tipos: sequência de nucleotídeos, onde cada letra corresponde a uma base nitrogenada (A, T, C e G), ou sequência de proteína, onde cada letra corresponde a um aminoácido (A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, e Y).

2.1.2 Alinhamento de sequências

Verli (2014) conceituou os alinhamentos como sendo "técnicas de comparação entre duas ou mais sequências biológicas, que buscam séries de caracteres individuais que se encontram na mesma ordem nas sequências analisadas". De forma mais prática, alinhamentos podem ser definidos como comparações entre sequências afim de observar seus graus de identidade ou similaridade (PROSDOCIMI, 2007).

O alinhamento de sequências é essencial para a maioria das análises de bioinformática. Um bom alinhamento, por exemplo, pode indicar possíveis relações evolutivas entre as sequências ou indicar se duas sequências estão relacionadas ao mesmo tipo de gene (VERLI, 2014; STEVENS; BOUCHER, 2015).

No alinhamento, as sequências são organizadas em linhas e os caracteres biológicos

em colunas. Ao longo das sequências iniciais são adicionados *gaps* (-), representando possíveis INDEL's, de forma que todas as sequências finais possuam o mesmo tamanho. Após alinhados, caracteres idênticos são chamados de *matches* e a substituição de caracteres é chamada de *mismatch*. Após o alinhamento, os graus de identidade e similaridade das sequências podem ser calculados, de acordo com Choudhuri (2014), seguindo as fórmulas abaixo:

$$\% \text{Identidade (PID)} = \left(\frac{\text{aminoácidos idênticos}}{\text{total de aminoácidos}} \right) * 100 \quad (2.4)$$

$$\% \text{Similaridade} = \left(\frac{\text{aminoácidos idênticos} + \text{substituições similares}}{\text{total de aminoácidos}} \right) * 100 \quad (2.5)$$

Para Verli (2014), os alinhamentos podem ser classificados quanto ao número de sequências e quanto a extensão das regiões analisadas. Quanto ao número de sequências, o alinhamento pode ser simples ou múltiplo, onde alinhamentos simples comparam somente duas sequências entre si e o alinhamento múltiplo comprara várias sequências simultaneamente. Quanto a extensão das regiões analisadas, o alinhamento pode ser local ou global, onde o alinhamento local utiliza somente regiões de alta similaridade na busca do melhor alinhamento, não importando as sequências adjacentes, e o global realiza a busca utilizando a sequência em toda a sua extensão. Nas Figuras 5 e 6 podemos ver exemplos de diferentes tipos de alinhamento.

Figura 5 – Exemplo de alinhamento simples local utilizando a ferramenta BLAST

```

BLASTN 2.11.0+

Reference: Zheng Zhang, Scott Schwartz, Lukas Wagner, and Webb
Miller (2000), "A greedy algorithm for aligning DNA sequences", J
Comput Biol 2000; 7(1-2):203-14.

Database: NZ_AUHR01000001.1
          78 sequences; 9,896 total letters

Query= AE016825.1:421225-422698

Length=1474

Sequences producing significant alignments:

          Score      E
          (Bits)    Value

NZ_AUHR01000037.1:107-1651

          2126      0.0

>NZ_AUHR01000037.1:107-1651
Length=1545

Score = 2126 bits (1151), Expect = 0.0
Identities = 1367/1473 (93%), Gaps = 8/1473 (1%)
Strand=Plus/Plus

Query  6      AAGAGTTTGATCCTGGCTCAGATTGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAA 65
      |||
Sbjct 12      AAGAGTTTGATCCTGGCTCAGATTGAACGCTGGCGGCATGCTTTACACATGCAAGTCGAA 71

Query  66      CGGTAACAGGGTGCTT-GCACCGCTGACGAGTGGCGAACGGGTGAGTAATGCGTCGGAAAT 124
      ||| | |
Sbjct  72      CGGTAACAGGG-ACTTCGGTCTGCTGACGAGTGGCGAACGGGTGAGTAATGCATCGGAAC 130

Query 125      GTACCGTGTAATGGGGGATAGCTCGGGCGAAAGCCGGATTAATACCGCATAACGCCCTGAGG 184
      ||| | |
Sbjct 131      GTACCGAGTAATGGGGGATAACGCATCGAAAGGTGTGCTAATACCGCATAACGCCCTGAGG 190
      .
Query 1322     GTGAAGTCGGAATCGCTAGTAATCGCAGATCAGCATGCTGCGGTGAATACGTTCCCGGGT 1381
      |||
Sbjct 1328     GTGAAGTCGGAATCGCTAGTAATCGCGGATCAGCATGTCGCGGTGAATACGTTCCCGGGC 1387

Query 1382     CTTGTACACACCCCGTCACACCATGGGAGTGAGTTTACCAGAAGTGGGTAGGCTAAC 1441
      |||
Sbjct 1388     CTTGTACACACCCCGTCACACCATGGGAGTGAATCCGCCAGAAGTGGGTAGGCTAAC 1447

Query 1442     CGCAAGGAGCCGCTTACCACGGTGGGATTCAT 1474
      || |||
Sbjct 1448     CGTAAGGAGCCGCTTACCACGGTAGGTTTCAT 1480

Lambda      K      H
          1.33  0.621  1.12

Gapped
Lambda      K      H
          1.28  0.460  0.850

Effective search space used: 12608784

Database: NZ_AUHR01000001.1
Posted date: Mar 22, 2021 7:30 PM
Number of letters in database: 9,896
Number of sequences in database: 78

Matrix: blastn matrix 1 -2
Gap Penalties: Existence: 0, Extension: 2.5

```

Fonte: elaborado pela autora.

Figura 6 – Exemplo de alinhamento múltiplo global utilizando a ferramenta MAFFT

```

CLUSTAL format alignment by MAFFT L-INS-i (v7.471)

WP_011136890.1  MRLILLGAPGAGKGTQANYIREKFGIPQISTGDMLEAAVKAGTPLGLEAKAIMDAGGLVR
WP_043632461.1  MRLILLGAPGAGKGTQANFIKEKFGIPQISTGDMLEAAVKAGTPLGLEAKAIMDAGGLVR
WP_147684476.1  MKLILLGAPGAGKGTQANYIKEKFGIPQISTGDMLEAAVKAGTPLGVEAKQIMDAGGLVR
WP_012696587.1  MKLILLGAPGAGKGTQAQFICQKFGIPQISTGDMLEAAVKAGTPLGLEAKVMDAGGLVS
*:*****: * :*****:*** :*****

WP_011136890.1  DDIIIGLVKERIAQDDCANGFLFDGFPRTIPQAEAMIAAGVDIDYVVEIDVPDAAIVERM
WP_043632461.1  DDIIIGLVKERIADADCANGFLFDGFPRTIPQAEAMIAAGVDIDYVVEIDVPDAAIERM
WP_147684476.1  DDIIIGLVKERIAEADCANGFLFDGFPRTIPQAEAMKEAGVDIDYVVEIDVPDENIIDRM
WP_012696587.1  DDIIIGLVKERIAQADCANGFLFDGFPRTIPQAEAMKAAGVNLDFVVEIDVPDSAIIVERM
*****: ***** ***** ***:***** **:**

WP_011136890.1  AGRRVHLASGRTYHVTFNPPKAAGKDDVTGEDLVQRDDKEETVKKRLAVYHEQTAVLVG
WP_043632461.1  AGRRVHLPSTGRTYHIRFNPPKVAGKDDVTGEELVQRDDREDTVKKRLAVYHEQTAVLVG
WP_147684476.1  SGRRVHVASGRTYHVKNPPKAEGVDETEPLVQRDDKAETVKKRLDVEYHEQTEVLVG
WP_012696587.1  SGRRVHVASGRTYHVKNPPKVAGKDDTEGDLIQRADNEETVLKRLAVYHEQTEVLVG
:****:.*****: ****. * ** ** *:* ** .:** ** ***** **

WP_011136890.1  FYGKLAESGSAKAPKYVKIDGTRAVETVRDDVLKALGA
WP_043632461.1  FYGERAASGDAKAPKYVKIDGTRAVETVRDEVLKALGA
WP_147684476.1  FYSQMAASGDAKAPKYVKIDGTQAVESVRDTVLAALGA
WP_012696587.1  YYSNMAASGDKTAPTYVKIPVGSVDGIRDAIFRALGA
*: * ** .**.* ** * :*: **: : **

```

Fonte: elaborado pela autora.

2.1.2.1 Programação Dinâmica

A forma mais comum de se obter um alinhamento simples é utilizando algoritmos de programação dinâmica. A programação dinâmica é um método baseado no princípio de otimização de Bellman, onde é proposto a solução de problemas complexos através da resolução de subproblemas, conferindo uma vantagem funcional e acelerando a resolução do problema complexo (VERLI, 2014). A programação dinâmica pode ser utilizada tanto para alinhamentos globais, utilizando o algoritmo de Needleman-Wunsch, como para alinhamentos locais, utilizando o algoritmo de Smith-Waterman.

O algoritmo de Needleman-Wunsch segue três passos: inicialização, preenchimento e *traceback*. Durante a inicialização o algoritmo transforma a relação entre as sequências em uma matriz $H(n + 1, m + 1)$, onde n e m são os comprimentos das sequências a serem alinhadas. Nessa matriz as linhas são definidas pelos caracteres da sequência a e as colunas pelos caracteres da sequência b . A fim de permitir a inserção de *gaps* no início do alinhamento, o algoritmo

adiciona uma coluna e uma linha vazia no início da matriz. Durante o preenchimento da matriz os subcaminhos associados ao alinhamento serão calculados seguindo a função abaixo:

$$H(i, j) = \max \begin{cases} H(i-1, j-1) + s(a_i, b_i) & \text{mismatch} \\ \max_{k \leq 1} (H(i-k, j) + W_k) & \text{deleção} \\ \max_{l \leq 1} (H(i, j-l) + W_l) & \text{inserção} \end{cases} \quad (2.6)$$

onde $0 \leq i \leq n$, $0 \leq j \leq m$, $s(a, b)$ é a função de *mismatch* e W é a pontuação de *gaps*.

A função de *mismatches* e a pontuação de *gaps* vão receber valores de acordo com uma matriz de custo. Para a pontuação de alinhamentos de nucleotídeos as mais conhecidas são MAT50 e a MAT70, enquanto para sequências protéicas as mais conhecidas são as matrizes PAM e BLOSUM. As Figuras 7 e 8 mostram exemplos de matriz de custo para nucleotídeos e para aminoácidos respectivamente.

Figura 7 – Matriz de custo MAT70

A	2						
C	-2	2					
G	-1	-2	2				
T	-2	-1	-2	2			
Y	-2	0	-2	0	0		
R	0	-2	0	-2	-2	0	
N	0	0	0	0	0	0	0
	A	C	G	T	Y	R	N

Nota: Y representa pirimidinas, R representa purinas e N representa qualquer nucleotídeo.

Fonte: adaptado de (PROSDOCIMI, 2007)

Por fim, no *traceback*, o algoritmo irá escolher o melhor subcaminho começando da célula $H(n, m)$ e indo na direção em que a matriz foi criada.

O algoritmo de Smith-Waterman é uma modificação do algoritmo de Needleman-Wunsch. As principais diferenças consistem em: a e b são subsequências da sequência original e é adicionada uma opção a mais na função de preenchimento da matriz onde todos os valores negativos são zerados.

2.1.2.2 *K-Tuple*

Além da programação dinâmica, outro algoritmo comum para se obter um alinhamento simples é o *K-Tuple*. *K-Tuple* é um método muito útil para buscar similaridade entre uma única sequência e um grande conjunto de dados. Para isso, o algoritmo divide uma sequência alvo e as sequências do banco de dados em pequenas sequências chamadas palavras. As palavras da sequência alvo são comparadas às palavras do banco de dados e após uma busca de identidade, o algoritmo alinha as duas sequências completas a partir das palavras similares, estendendo a análise de similaridade para as regiões vizinhas anteriores e posteriores. Por fim, utilizando uma matriz de pontuação, o algoritmo calcula o melhor alinhamento. Esse método é utilizado por ferramentas como o BLAST.

2.1.2.3 *Alinhamento Progressivo*

O alinhamento progressivo é um método heurístico criado para acelerar os cálculos de um alinhamento global múltiplo, realizando buscas aproximadas pelo resultado ótimo. Primeiramente é realizado um alinhamento simples de todas as possíveis combinações de sequências e verifica-se a distância evolutiva entre elas. A partir dessa distância é criada uma filogenia e então constrói-se progressivamente o alinhamento de acordo com os ramos próximos. O algoritmo construirá um alinhamento através de programação dinâmica para duas sequências e todo o restante das sequências serão alinhadas baseando-se neste primeiro par. Após o alinhamento as duas sequências serão então tratadas como uma e alinhadas à outra sequência filogeneticamente mais próxima. Esse método é utilizado por ferramentas como o CLUSTAL e MATFF.

2.1.2.4 *Pontuação Baseada em Consistência*

Pontuação baseada em consistência é um método baseado no algoritmo de alinhamento progressivo. Seu diferencial é a realização de outros alinhamentos par-a-par para aperfeiçoar as comparações entre as sequências, não levando em consideração apenas o primeiro par de sequências alinhadas. Esse método é utilizado por ferramentas como o T-COFFEE e MATFF.

2.1.2.5 Refinamento Iterativo

Refinamento iterativo também é um método semelhante ao algoritmo de alinhamento progressivo. Seu diferencial é o realinhamento constante dos grupos de sequências ao longo das análises, garantindo que o alinhamento inicial não defina o resultado final. Esse método é utilizado por ferramentas como o MUSCLE e MATFF.

2.1.2.6 Significância estatística

Para determinar o quão bom é um alinhamento, devem ser feitos cálculos de significância estatística. A estatística de Karlin-Altschul (KARLIN; ALTSCHUL, 1990) é usada para determinar a significância de alinhamentos locais e é expressa através do *e-value*. *E-value* é o número esperado de alinhamentos distintos com pontuação igual ou melhor, que seria encontrado em sequências aleatórias similares as originais.

Ao encontrar os Pares de Segmentos de Alta Pontuação (HSP) em um alinhamento local, é necessário analisar se essa pontuação surgiu por acaso. Para isso é gerado um modelo aleatório das sequências e os alinhamentos aleatórios são pontuados de acordo com uma matriz de custo. O cálculo do *E-value* para um HSP de pontuação S é dado pela fórmula:

$$E = Kmn e^{-\lambda S} \quad (2.8)$$

onde m é o comprimento da sequência de entrada, n é o tamanho do banco de dados e K e λ são parâmetros de escala natural para o tamanho do espaço de busca e para o sistema de pontuação, respectivamente. Quanto menor o *e-value* mais significativo é o resultado, ou seja, menor a chance do alinhamento ser aleatório.

Diferente dos alinhamentos locais, a distribuição aleatória de alinhamentos globais ainda não é bem entendida. Um dos métodos para se estimar a significância de alinhamentos globais seria gerar múltiplos pares de sequências aleatórias, de composição e comprimento apropriados, calcular o alinhamento ótimo para cada um e, considerando que se tratasse de uma distribuição normal, expressar um *P-value* de acordo com a média e a variância (ALTSCHUL, 2011).

2.1.3 Inferência Filogenética

A partir de um bom alinhamento é possível aplicar métodos de inferência filogenética para a construção de uma árvore. Uma árvore filogenética permite investigar as relações evolutivas entre os organismos. Como mostra a Figura 9, existem vários métodos de inferência filogenética, cada um com diferentes níveis de poder, eficiência, consistência e robustez. Esses métodos podem ser classificados em dois tipos: quantitativos e qualitativos.

Figura 9 – Comparação entre métodos de inferência filogenética

Tipo	Método	Princípio básico	Exigência Computacional
Quantitativo	Ligação Média Entre Grupos	Agrupar as sequências com menor distância entre si.	Baixa
	Vizinho Mais Próximo	Busca a árvore com menor número de ramos.	Baixa
Qualitativo	Máxima Parcimônia	Busca a filogenia com menor número de eventos evolutivos.	Médio
	Máxima Verossimilhança	Busca a árvore com o valor de maior verossimilhança entre todas as filogenias construídas.	Alta
	Inferência Bayesiana	Amostra um número representativo de filogenias a partir do espaço amostral total de árvores e busca a mais provável.	Alta

Fonte: adaptado de (VERLI, 2014).

2.1.3.1 Métodos quantitativos

Os métodos quantitativos, também chamados métodos de distância, determinam a filogenia baseados na quantidade de diferenças entre os alinhamentos. São métodos de enorme eficiência computacional, geralmente utilizados na construção de árvores iniciais que vão ser utilizadas depois em um método qualitativo (VERLI, 2014). Os métodos quantitativos mais usados são Vizinho Mais Próximo (NJ) e Ligação Média entre Grupos (UPGMA).

O UPGMA (SNEATH *et al.*, 1973) baseia-se na hipótese de que todos os organismos evoluem em taxas constantes. Os alinhamentos são convertidos em matrizes de distância par-a-par para todas as sequências incluídas e o par de sequências com a menor distância é agrupado. Uma nova distância é calculada e o procedimento se repete até que restem apenas duas sequências. Esse método pode gerar múltiplas árvores, caso algumas das distâncias sejam iguais e pode levar a falsas topologias, caso os dados não se encaixem na hipótese do relógio molecular. Esse

método é utilizado por ferramentas como o MEGA e Geneious.

O NJ (SAITOU; NEI, 1987) baseia-se nos algoritmos de evolução mínima, os quais buscam a árvore com menor soma total de ramos. A construção da árvore é similar ao método da Ligação Média entre Grupos, porém a matriz de distâncias é construída a partir da aplicação de um modelo de substituição. Esse método gera somente uma árvore final, uma vez que distâncias iguais são tratadas em ordem ou aleatoriamente. Esse método também é utilizado por ferramentas como o MEGA, Geneious e QuickTree

2.1.3.2 *Métodos qualitativos*

Os métodos qualitativos determinam uma filogenia seguindo um critério, onde, dentre várias árvores, é escolhida a árvore de maior valor associado a esse critério (VERLI, 2014). Os métodos quantitativos mais usados são Máxima Parcimônia (MP), Máxima Verossimilhança (ML) e Inferência Bayesiana (BI).

A MP (Guilherme de Ockham) baseia-se no princípio de que a natureza opta pelos caminhos mais simples, onde a filogenia mais provável é a que apresentar o menor número de eventos evolutivos (FITCH, 1971). O método calcula as probabilidades de mudança dos nucleotídeos para cada sítio, considerando cada sítio individualmente e utilizando uma matriz de custo. Sítios que não apresentam variação ou autapomorfias são descartados das análises. Os menores custos para cada sítio são somados para encontrar o tamanho dos ramos e a árvore com os ramos mais parcimoniosos será a árvore final. Esse método pode falhar ao estimar a relação evolutiva entre muitos táxons, agrupando incorretamente os táxons com maiores taxas de evolução. Esse método é utilizado por ferramentas como o MEGA, PHYLIP e PAUP.

O ML (FISHER, 1922) baseia-se no princípio de que a verossimilhança de uma árvore é proporcional à probabilidade de explicar os dados do alinhamento. O método cria uma filogenia inicial e calcula a probabilidade de todas as mudanças possíveis para cada sítio do alinhamento, utilizando um modelo evolutivo para atribuir valores e parâmetros às substituições. Esse método considera que cada sítio evolui de forma independente e elimina das análises, sítios que apresentam deleções. As verossimilhanças de cada sítio são multiplicadas para encontrar a verossimilhança da árvore e, após uma amostragem com várias árvores, a árvore de valor máximo de verossimilhança será a árvore final. Esse método é utilizado por ferramentas como o MEGA, PAUP, PAML, PHYLIP e RAxML.

A BI (Thomas Bayes) baseia-se na inclusão de evidências prévias no cálculo da

probabilidade de acontecimento de determinados eventos (STIGLER, 1982). O método calcula as probabilidades posteriores para cada filogenia proposta utilizando a fórmula de Bayes abaixo:

$$L(H | D) = \frac{L(H)L(D | H)}{L(D)} \quad (2.9)$$

onde $L(H | D)$ é a probabilidade posterior, $L(H)$ é a probabilidade a priori, $L(D | H)$ descreve o cálculo da máxima verossimilhança e $L(D)$ é a probabilidade total. Para realizar amostragens a partir do conjunto total de filogenias, esse método se utiliza do Monte Carlo via Cadeiras de Markov (MCMC), relacionando cada filogenia a um valor probabilístico. O algoritmo MCMC mais usado é o Metropolis-Hastings, que realiza perturbações aleatórias em uma filogenia inicial, aceitando ou rejeitando a nova topologia de acordo com o cálculo de razão das probabilidades. Este algoritmo garante que diversas filogenias com probabilidades posteriores crescentes sejam amostradas do espaço total de filogenias. O método de inferência Bayesiana é utilizado em ferramentas como o Mr. Bayes e BEAST.

2.1.3.3 Confiabilidade

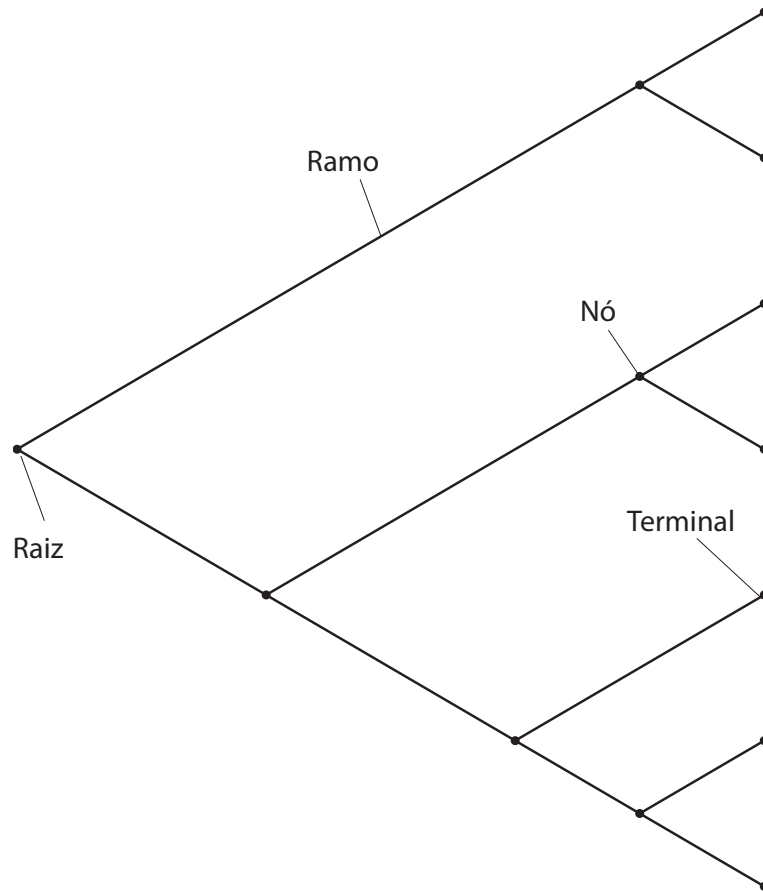
A sustentação de uma filogenia está relacionada ao número de filogenias excluídas, a partir do conjunto total de filogenias. Para determinar essa confiabilidade é necessário aplicar um método de reamostragem chamado *Bootstrap*. A reamostragem é realizada a partir do conjunto de alinhamentos original. Tratando cada sitio de forma independente, são geradas várias pseudoreplicatas com o mesmo tamanho das originais. A partir de cada novo conjunto de alinhamento o algoritmo constrói uma filogenia. Ao final os clados são analisados e é contada a sua presença em todas as filogenias construídas, onde números acima de 70 são considerados de boa sustentação (VERLI, 2014).

2.1.4 Representação das Árvores

As árvores filogenéticas representam o contexto evolutivo dos organismos de forma gráfica. Como pode ser visto na Figura 10, as árvores são formadas por nós, representando ancestrais hipotéticos, conectados por vários ramos, representando as linhagens evolutivas. Os terminais identificam os táxons incluídos na análise filogenética e o tamanho dos ramos geralmente significam a distância evolutiva entre esses táxons.

As árvores podem ser apresentadas de várias maneiras (Figura 11), evidenciando o ancestral mais antigo, através de uma raiz, ou evidenciando apenas as relações evolutivas entre

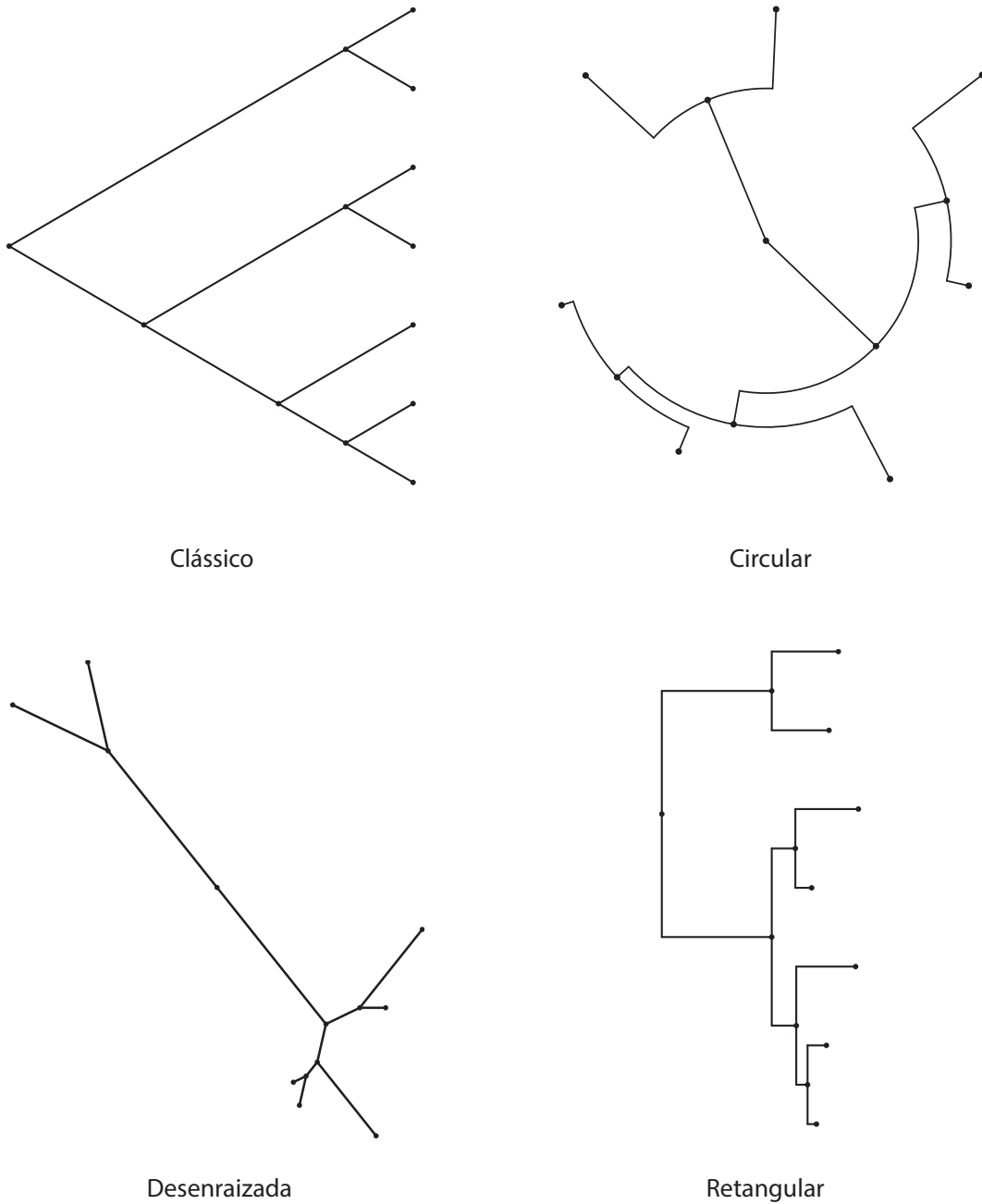
Figura 10 – Visão geral de uma árvore filogenética.



Fonte: elaborado pela autora.

os táxons. Para identificar a raiz de uma filogenia, geralmente é necessário a inclusão de grupos externos, que são grupos próximos ao grupo em estudo, mas que surgiram antes na história evolutiva (LOPES *et al.*, 2014).

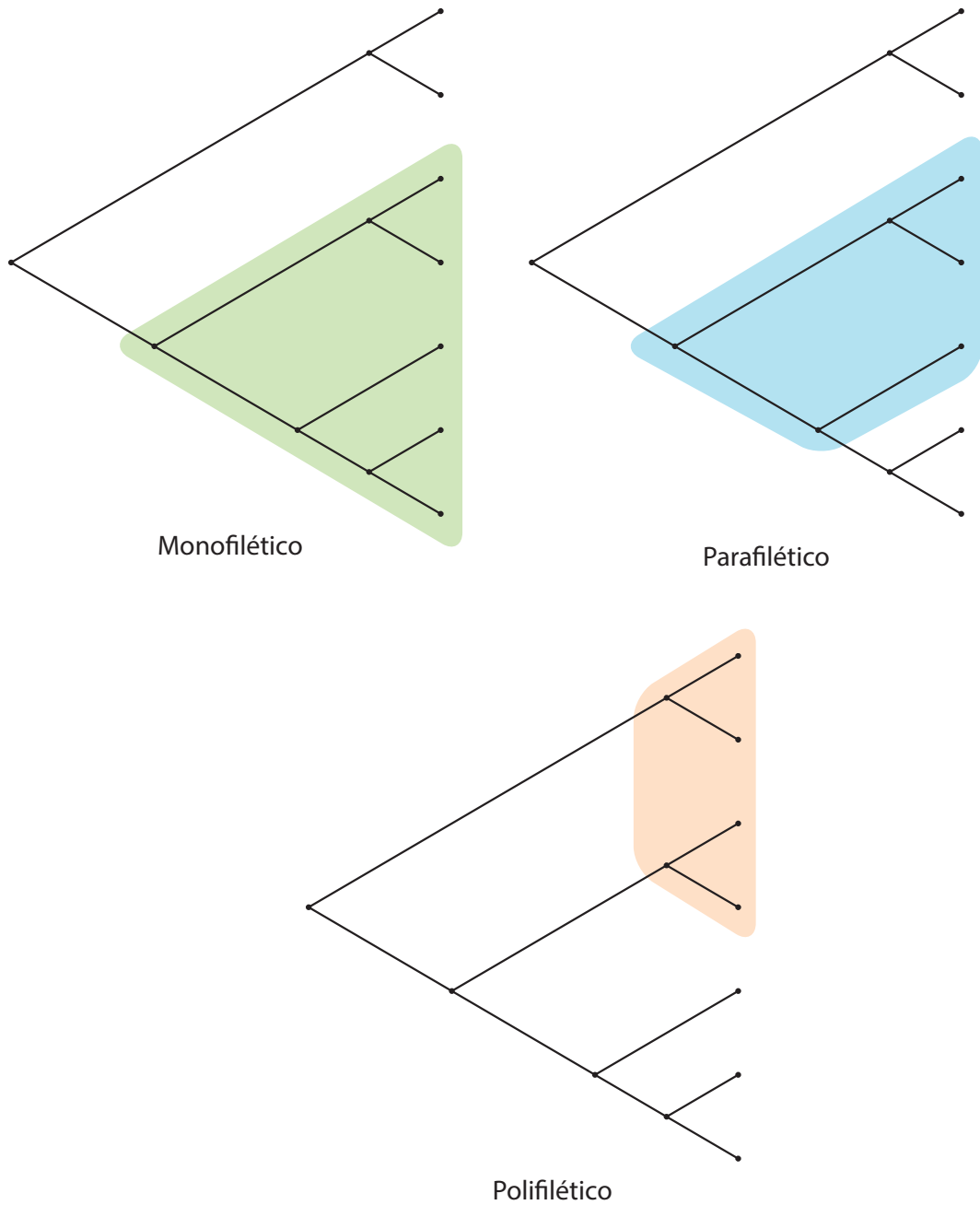
Figura 11 – Possíveis representações de uma mesma árvore filogenética.



Fonte: elaborado pela autora.

Analisando os táxons de uma árvore, podemos reconhecer três tipos de agrupamentos: monofiléticos, parafiléticos e polifiléticos (Figura 12).

Figura 12 – Tipos de classificações de táxons.



Fonte: elaborado pela autora.

Grupos monofiléticos incluem todos os descendentes de um único ancestral comum, podendo ser denominados clados. Grupos parafiléticos também derivam de um único ancestral comum, porém nem todos descendentes fazem parte do grupo. Grupos polifiléticos derivam de dois ou mais ancestrais diferentes. Esses padrões ajudam a descrever características resultantes de homologia, similaridade originada por ancestralidade comum, ou convergência evolutiva, similaridade originada independentemente, sem ancestralidade comum.

2.2 Chromobacteriaceae

Nesta seção é apresentada uma visão geral da família em estudo. A seção está dividida em quatro partes. A primeira parte apresenta algumas características gerais da família. A segunda descreve a diversidade das espécies pertencentes a família. A terceira descreve alguns aspectos sobre a importância econômica desses organismos. Por fim, a quarta parte relata casos de patogenicidade associados a membros da família.

2.2.1 Características Gerais

Chromobacteriaceae é uma família de betaproteobactérias pertencentes a ordem Neisseriales. De maneira geral a família é composta por bactérias de vida livre, gram negativas, ocorrem isoladas, em pares ou em cadeias, possuem formato bacilar e movem-se devido a presença de flagelo, salvo algumas exceções, que apresentam formato espiralar ou movem-se devido a contrações citoplasmáticas. As colônias não são pigmentadas, com exceção de alguns gêneros como *Chromobacterium*, *Iodobacter* e *Vogesella*, que formam colônias de coloração roxa (ADEOLU; GUPTA, 2013).

2.2.2 Diversidade de Espécies

Chromobacteriaceae abriga em torno de 65 espécies com nomenclatura válida. A família está dividida em 19 gêneros, sendo eles: *Andreprevotia*, *Aquaspirillum*, *Aquitalea*, *Chitinibacter*, *Chitinilyticum*, *Chitiniphilus*, *Chromobacterium*, *Deefgea*, *Formivibrio*, *Gulbenkiania*, *Iodobacter*, *Jeongeupia*, *Laribacter*, *Leeia*, *Microvirgula*, *Paludibacterium*, *Pseudogulbenkiania*, *Silvimonas* e *Vogesella*. *Chromobacterium* é o gênero tipo e se caracteriza como o maior gênero da família (MADIGAN *et al.*, 2016).

2.2.2.1 Andreprevotia

O gênero foi descrito por Weon *et al.* (2007) como bactérias em formato de bastão, com tamanho entre 0,7-0,9 x 2,0-3,0 μm , gram-negativas, móveis com um único flagelo polar ou subpolar. As colônias são de cor marfim, redondas e convexas, com margens claras em meio Agar R2A. São positivas para catalase, oxidase e fermentação de glucose. São negativas para redução de nitrato, reação de Voges-Proskauer e desaminação de fenilalanina. O conteúdo G+C do DNA da espécie tipo é 62 mol%. A espécie tipo é *Andreprevotia chitinilytica*.

De acordo com a Lista de Nomes Procarióticos com Posição na Nomenclatura (LPSN)¹ o grupo abriga atualmente duas espécies com taxonomia correta e nomenclatura válida: *Andreprevotia chitinilytica* (WEON *et al.*, 2007) e *Andreprevotia lacus* (SHEU *et al.*, 2009).

2.2.2.2 *Aquaspirillum*

O gênero foi descrito por Hylemon *et al.* (1973) como organismos que possuem células rígidas, helicoidal ou helicoidal curvo, com diâmetro entre 0,2 e 1,5. Móveis, devido a presença de flagelos, predominando um único flagelo bipolar. São gram-negativos, quimiorganotróficos, aeróbicos restritos, oxidase e catalase-positivas. São indol e sulfatase negativos. Não hidrolisam caseína, amido ou hipurato e não fazem a fermentação de carboidratos. Não crescem na presença de 3% NaCl e apresentam polihidroxibutirato (PHB) intracelular. O conteúdo G+C do DNA vai de 49 a 65 mol%. A espécie tipo é *Aquaspirillum serpens*.

De acordo com o LPSN, o grupo abriga atualmente seis espécies com taxonomia correta e nomenclatura válida: *Aquaspirillum arcticum* (BUTLER *et al.*, 1989), *Aquaspirillum bengal* (KUMAR *et al.*, 1974; SKERMAN *et al.*, 1980), *Aquaspirillum polymorphum* (WILLIAMS; RITTENBERG, 1957; HYLEMON *et al.*, 1973; SKERMAN *et al.*, 1980), *Aquaspirillum putridiconchylium* (TERASAKI, 1961; HYLEMON *et al.*, 1973; SKERMAN *et al.*, 1980), *Aquaspirillum serpens* (MULLER, 1786; HYLEMON *et al.*, 1973; SKERMAN *et al.*, 1980) e *Aquaspirillum soli* (MOYA *et al.*, 2017).

2.2.2.3 *Aquitalea*

O gênero foi descrito por Lau *et al.* (2006) como organismos que possuem células de formato bacilar, gram-negativas e sem formação de esporos. Os bacilos são curtos, entre 1 e 7 μm de comprimento, retos ou curvos. Aeróbicos ou anaeróbicos facultativos, quimiorganotróficos e sensíveis a NaCl. As células são móveis, com um flagelo polar e são catalase- e oxidase-positivas. O conteúdo G+c do DNA é em torno de 59,2 mol%. A espécie tipo é *Aquitalea magnusonii*.

De acordo com o LPSN, o grupo abriga atualmente quatro espécies com taxonomia correta e nomenclatura válida: *Aquitalea aquatilis* (NGO *et al.*, 2020), *Aquitalea denitrificans* (LEE *et al.*, 2009), *Aquitalea magnusonii* (LAU *et al.*, 2006) e *Aquitalea pelogenes* (SEDLÁČEK *et al.*, 2016).

¹ <https://lpsn.dsmz.de/>

2.2.2.4 *Chitinibacter*

O gênero foi descrito por Chern *et al.* (2004) como bactérias estritamente aeróbicas, gram-negativas, com formato de bastão, reto ou levemente curvo, com bordas arredondadas, comprimento em torno de 1,3 e 2,6 μm e largura em torno de 0,5 e 0,9 μm . Possuem alta mobilidade, devido a presença de um ou dois flagelos polares. São oxidase- e catalase-positivas. Não formam ácido a partir da glucose. Q-8 é a principal quinona isoprenóide. O conteúdo G+C do DNA é em torno de 56 mol%. A espécie tipo é *Chitinibacter tainanensis*.

De acordo com o LPSN, o grupo abriga atualmente três espécies com taxonomia correta e nomenclatura válida: *Chitinibacter alvei* (YANG *et al.*, 2010), *Chitinibacter fontanus* (SHEU *et al.*, 2016) e *Chitinibacter tainanensis* (CHERN *et al.*, 2004).

2.2.2.5 *Chitinilyticum*

O gênero foi descrito por Chang *et al.* (2007) e corrigido por Chang *et al.* (2009) como bactérias de formato bacilar, gram-negativas e móveis devido a um único flagelo polar. Armazenam reservas de grânulos intracelulares de PHB. Não formam endósporos. Suas colônias, em meio a base de quitina, são envoltas por uma larga zona clara, devido a degradação da quitina. Não produz pigmentos. Crescem bem utilizando quitina como única fonte de carbono, nitrogênio e energia, tanto em condições aeróbicas quanto anaeróbicas. São positivas para oxidase, com atividade variável para diferentes espécies, e positivas para catalase. Q-8 é a principal quinona isoprenóide. O conteúdo G+C do DNA é entre 62,2 e 69,5 mol%. A espécie tipo é *Chitinilyticum aquatile*.

De acordo com o LPSN, o grupo abriga atualmente duas espécies com taxonomia correta e nomenclatura válida: *Chitinilyticum aquatile* (CHANG *et al.*, 2007) e *Chitinilyticum litopenaei* (CHANG *et al.*, 2009).

2.2.2.6 *Chitiniphilus*

O gênero foi descrito por Sato *et al.* (2009) como bactérias de formato bacilar curvo, gram-negativas e sem formação de esporos. Os bacilos tem entre 0,7 e 0,8 μm de largura e entre 1,2 e 1,5 μm de comprimento. As células são móveis devido a um único flagelo polar. São anaeróbicos facultativos e são catalase- e oxidase-positivas. Não produz ácido a partir de glucose, mas é positiva para redução de nitrato. O conteúdo G+c do DNA da espécie tipo é 67,6 mol%. A

espécie tipo é *Chitiniphilus shinanonensis*.

De acordo com o LPSN, o grupo abriga atualmente duas espécies com taxonomia correta e nomenclatura válida: *Chitiniphilus eburneus* (SHENG *et al.*, 2020) e *Chitiniphilus shinanonensis* (SATO *et al.*, 2009).

2.2.2.7 *Chromobacterium*

O gênero surgiu pela primeira vez na publicação de Bergonzini (1880) onde descreveu a espécie *Chromobacterium violaceum*. Em geral, o gênero é composto por bactérias em forma de bastão, gram-negativas, aeróbias ou anaeróbias facultativas, que não formam esporos, com tamanho variando em torno de 0,3-0,8 μm por 1,5-2,6 μm . São móveis devido a um único flagelo polar, podendo em algumas espécies apresentar flagelos laterais ou subpolares. As colônias apresentam coloração violeta devido a um pigmento, solúvel em álcool, denominado violaceína. Algumas espécies, entretanto, não produzem pigmentos. Muitas espécies podem reduzir nitrato para nitrito e não produzem indol. Algumas espécies apresentam Q-8 como a principal quinona isoprenóide e moderadas quantidades de Q-7. Algumas espécies apresentam atividade hemolítica positiva. A espécie tipo é *Chromobacterium violaceum*.

De acordo com o LPSN, o grupo abriga atualmente doze espécies com taxonomia correta e nomenclatura válida: *Chromobacterium alkanivorans* (BAJAJ *et al.*, 2016), *Chromobacterium amazonense* (MENEZES *et al.*, 2015), *Chromobacterium aquaticum* (YOUNG *et al.*, 2008), *Chromobacterium haemolyticum* (HAN *et al.*, 2008), *Chromobacterium paludis* (BLACKBURN *et al.*, 2020), *Chromobacterium phragmitis* (BLACKBURN *et al.*, 2019), *Chromobacterium piscinae* (KÄMPFER *et al.*, 2009), *Chromobacterium rhizoryzae* (ZHOU *et al.*, 2016), *Chromobacterium sphagni* (BLACKBURN *et al.*, 2017), *Chromobacterium subtsugae* (MARTIN *et al.*, 2007), *Chromobacterium vaccinii* (SOBY *et al.*, 2013) e *Chromobacterium violaceum* (BERGONZINI, 1880; SKERMAN *et al.*, 1980).

2.2.2.8 *Deefgea*

O gênero foi descrito Stackebrandt *et al.* (2007) como bactérias gram-negativas, com formato de bastão, com tamanho em torno de 0,7-0,86 μm x 3-4 μm e anaeróbicas facultativas. Em geral ocorrem isoladas e são móveis devido a presença de um ou, raramente, dois flagelos polares. As colônias em meio R2A são circulares, convexas e de cor branco pérola. São catalase- e oxidase-positivas. Possuem grânulos de polihidroxibutirato e polifosfato. A principal quinona

é Q-8. Os lipídios polares são Fosfatidiletanolamina e fosfatidilglicerol. O conteúdo G+C do DNA é em torno de 49mol%. A espécie tipo é *Deefgea rivuli*.

De acordo com o LPSN, o grupo abriga atualmente duas espécies com taxonomia correta e nomenclatura válida: *Deefgea chitinilytica* (CHEN *et al.*, 2010) e *Deefgea rivuli* (STACKEBRANDT *et al.*, 2007).

2.2.2.9 *Formivibrio*

O gênero foi proposto por Tanaka *et al.* (1991) que descreveu a espécie tipo como bactérias gram-negativas, em formato de bastão curvo, com pontas afiladas e arredondadas, tamanho em torno de 1,1-2,5 μm por 0,5-0,6 μm , ocorrendo individualmente ou em pares. Móveis devido a um único flagelo polar. Colônias circulares, convexas incolores e translúcidas, com diâmetro de até 0,2 mm. A célula contém ácido β -hidroximirístico. A parede celular contém peptidoglicano do ácido meso-diaminopimérico (m-Dap) do tipo diretamente ligado. Catalase negativa. Nitrato, sulfato, sulfito, tiosulfato e enxofre elementar não reduzido. Esculina, ureia e gelatina não hidrolisada. Indol não produzido a partir de triptofano. Conteúdo G+C do DNA 59 mol%.

De acordo com o LPSN, o grupo abriga apenas a espécie tipo *Formivibrio citricus* (TANAKA *et al.*, 1991).

2.2.2.10 *Gulbenkiania*

O gênero foi descrito por Vaz-Moreira *et al.* (2007) como bactérias não formadoras de esporos, gram-negativas, móveis e com formato de bastão curvo. Apresentam catalase-, oxidase- e arginina di-hidrolase positivas. Mesofílicas, quimiorganotróficas e aeróbicas. Não requerem nenhum fator de crescimento orgânico específico. O nitrato é reduzido a nitrito. Aminoácidos e ácidos orgânicos, mas não açúcares, são usados como fonte de carbono. A principal quinona é a Q-8, Q-9 é um componente secundário. O conteúdo G+C do DNA da espécie tipo é 63 mol%. A espécie tipo é *Gulbenkiania mobilis*.

De acordo com o LPSN, o grupo abriga atualmente duas espécies com taxonomia correta e nomenclatura válida: *Gulbenkiania indica* (JYOTI *et al.*, 2010) e *Gulbenkiania mobilis* (VAZ-MOREIRA *et al.*, 2007).

2.2.2.11 *Iodobacter*

O gênero foi descrito por LOGAN (1989) como bactérias gram-negativas, com formato de bastão, retas e pontas arredondadas, tamanho 0,7 por 3,0-3,5 μm , ocorrendo individualmente, em pares, às vezes em cadeia ou como filamentos longos. Móveis, com um único flagelo polar e um ou mais flagelos laterais. Cresce em meios de peptona comuns. A maioria das cepas produz o pigmento violaceína. São quimiorganotróficos, anaeróbios facultativos e fermentam carboidratos sem produção de gás. A espécie tipo é *Iodobacter fluviatilis*.

De acordo com o LPSN, o grupo abriga atualmente quatro espécies com taxonomia correta e nomenclatura válida: *Iodobacter arcticus* (SRINIVAS *et al.*, 2013), *Iodobacter ciconiae* (LEE *et al.*, 2019), *Iodobacter fluviatilis* (MOSS *et al.*, 1978; LOGAN, 1989) e *Iodobacter limnosediminis* (SU *et al.*, 2013)

2.2.2.12 *Jeongeupia*

O gênero foi descrito por Yoon *et al.* (2010) como bacterias em forma de bastonete, gram-negativas e anaerobias facultativas. Q-8 como ubiquinona predominante, catalase- e oxidase-positivas. A espécie tipo é *Jeongeupia naejangsanensis*.

De acordo com o LPSN, o grupo abriga atualmente duas espécies com taxonomia correta e nomenclatura válida: *Jeongeupia chitinilytica* (CHEN *et al.*, 2013) e *Jeongeupia naejangsanensis* (YOON *et al.*, 2010).

2.2.2.13 *Laribacter*

O gênero foi proposto por Yuen *et al.* (2001) que descreveu a espécie tipo como bactérias anaeróbicas facultativas, não foram esporos, gram-negativas, em forma de bastão espiralado. Não crescem em NaCl 3%, não são móveis. São oxidase, catalase, urease e arginina dihidrolase positiva e reduzem nitrato. Não fermentam, oxidam ou assimilam açúcares. O conteúdo G+C do DNA da espécie tipo é $68,0 \pm 2,43 \text{ mol\%}$.

De acordo com o LPSN, o grupo abriga atualmente apenas a espécie tipo: *Laribacter hongkongensis* (YUEN *et al.*, 2001).

2.2.2.14 *Leeia*

O gênero foi descrito por Lim *et al.* (2007) como bactérias em formato de bastões curtos, gram-negativos e não formam esporos. Aeróbicas e móveis devido a presença de um único flagelo polar. São oxidase e catalase-positivas. Reduz nitrato para nitrito. A principal quinona isoprenóide é Q-8. Lipídios polares são fosfatidiletanolamina, fosfatidilglicerol e difosfatidilglicerol. A espécie tipo é *Leeia oryzae*.

De acordo com o LPSN, o grupo abriga atualmente duas espécies com taxonomia correta e nomenclatura válida: *Leeia aquatica* (SONG *et al.*, 2020) e *Leeia oryzae* (LIM *et al.*, 2007).

2.2.2.15 *Microvirgula*

O gênero foi descrito por Patureau *et al.* (1998) como bactérias gram-negativas em forma de bastão curvo, muito móveis, ocorrendo individualmente, em pares ou em alguns casos em grupos. Oxidase e catalase-positivas. Aeróbicas e quimiorganotróficas, com grande espectro de fontes de carbono, exceto açúcares. Capaz de crescimento anóxico em nitrato, nitrito e N, O. O conteúdo G+C do DNA em torno de 65 mol%. A espécie tipo é *Microvirgula aerodenitrificans*.

De acordo com o LPSN, o grupo abriga atualmente duas espécies com taxonomia correta e nomenclatura válida: *Microvirgula aerodenitrificans* (PATUREAU *et al.*, 1998) e *Microvirgula curvata* (SUBHASH *et al.*, 2016).

2.2.2.16 *Paludibacterium*

O gênero foi descrito por Kwon *et al.* (2008) como bactérias gram-negativas, com formato de bastão curvo e que não formam esporos. Móveis devido a um único flagelo polar, anaeróbicas facultativas e positivas para catalase e oxidase. Reduzem nitrato, não produzem indol ou acetoina e a quinona isoprenóide predominante é Q-8. A espécie tipo é *Paludibacterium yongneupense*.

De acordo com o LPSN, o grupo abriga atualmente três espécies com taxonomia correta e nomenclatura válida: *Paludibacterium paludis* (SHEU *et al.*, 2014), *Paludibacterium purpuratum* (KANG *et al.*, 2016) e *Paludibacterium yongneupense* (KWON *et al.*, 2008).

2.2.2.17 *Pseudogulbenkiania*

O gênero foi descrito por Lin *et al.* (2008) como bactérias gram-negativas, em formato de bastões e que não formam de esporos, ocorrendo individualmente ou em pares. São aeróbicas, móveis devido a um único flagelo polar, oxidase-positivas e catalase-negativas. São quimiorganotróficas, capazes de utilizar um grande espectro de fontes de carbono. Os principais lípidos polares são difosfatidilglicerol, fosfatidilglicerol e fosfatidiletanolamina e a principal quinona ubiquinona é Q-8. A espécie tipo é *Pseudogulbenkiania subflava*.

De acordo com o LPSN, o grupo abriga atualmente duas espécies com taxonomia correta e nomenclatura válida: *Pseudogulbenkiania gefcensis* (LEE *et al.*, 2013) e *Pseudogulbenkiania subflava* (LIN *et al.*, 2008).

2.2.2.18 *Silvimonas*

O gênero foi descrito por Yang *et al.* (2005) como bactérias gram-negativas, anaeróbicas facultativas e em forma de bastão, tamanho 1,3-2,0 μm por 0,5-0,8 μm . Ocorrem individualmente ou em pares e são móveis devido a um único flagelo polar. As colônias em meio R2A são circulares, convexas e de coloração branco leitoso. São catalase e oxidase positivas e a quinona predominante é Q-8. A espécie tipo é *Silvimonas terrae*.

De acordo com o LPSN, o grupo abriga atualmente três espécies com taxonomia correta e nomenclatura válida: *Silvimonas amylolytica* (MURAMATSU *et al.*, 2010), *Silvimonas iriomotensis* (MURAMATSU *et al.*, 2010) e *Silvimonas terrae* (YANG *et al.*, 2005)

2.2.2.19 *Vogesella*

O gênero foi descrito por Grimes *et al.* (1997) e corrigido por Subhash *et al.* (2013), Sheu *et al.* (2013) e Sheu *et al.* (2016) como bactérias gram-negativas, em forma de bastão. Ocasionalmente bastões vibrióides também são observadas. Aeróbicas ou anaeróbicas facultativas. Têm entre 0,3-0,9 μm de largura e 0,7-4,0 μm de comprimento. Ocorrem individualmente, mas em menores chances podem formar pares ou cadeias curtas. Apresentam extensa formação de grânulos de PHB. Em geral são móveis devido a um único flagelo polar e exibem motilidade de dardo ou ziguezague muito rápida, mas a motilidade pode ausente em algumas espécies. Em algumas espécies as colônias são azul royal profundo devido à produção de indigoidina, em outras as colônias são de incolores a brancas ou acinzentadas. Não são fermentativas, catabolizam

alguns monossacarídeos sob condições aeróbicas, mas não sob condições de baixo oxigênio. Caseína, gelatina e amido são hidrolisado por algumas espécies. vários aminoácidos e peptídeos são utilizados, embora não sejam descarboxilados. São capazes de desnitrificação. Indol é produzido a partir do catabolismo de triptofano em algumas espécies. Citrato, arabinose e Tween 80 são utilizados como fontes únicas de carbono em algumas espécies. O conteúdo G+C do DNA é entre de 61,2-67,4 mol%. A espécie tipo é *Vogesella indigofera*.

De acordo com o LPSN, o grupo abriga atualmente dez espécies com taxonomia correta e nomenclatura válida: *Vogesella alkaliphila* (SUBHASH *et al.*, 2013), *Vogesella amnigena* (CHEN *et al.*, 2015), *Vogesella facilis* (SHEU *et al.*, 2016), *Vogesella fluminis* (SHEU *et al.*, 2013), *Vogesella indigofera* (VOGES, 1893; GRIMES *et al.*, 1997), *Vogesella lacus* (CHOU *et al.*, 2009), *Vogesella mureinivorans* (JØRGENSEN *et al.*, 2010), *Vogesella oryzae* (RAMESHKUMAR *et al.*, 2016), *Vogesella perlucida* (CHOU *et al.*, 2008) e *Vogesella urethralis* (LAN *et al.*, 2020).

2.2.3 Importância Biotecnológica

Chromobacteriaceae é um grupo de grande potencial biotecnológico. A família abriga a espécie *Chromobacterium violaceum* ATCC 12472, o primeiro microrganismo a ter seu genoma completamente sequenciado pelo Projeto Genoma Brasileiro. O genoma desse organismo contém diversas fases de leitura aberta codificando enzimas e metabólitos secundários de grande importância (VASCONCELOS *et al.*, 2003). Dentre essas enzimas estão as β -1,3-glucanases que possuem diversas aplicações biotecnológicas. Há relatos na literatura de aplicações na obtenção de oligossacarídeos bioativos, na caracterização da parede celular de micro organismos, no controle biológico de patógenos de algumas plantas, como aditivos em rações para animais, e na sua adição em vinhos para melhorar as duas características organolépticas (BAUERMEISTER *et al.*, 2010).

Chromobacterium violaceum, assim como outras espécies da família, tem como principal característica a produção de violaceína. A violaceína é um pigmento versátil que exhibe diversas atividades biológicas de importância na indústria farmacêutica, cosmética e têxtil (DURÁN *et al.*, 2012). Suas propriedades biológicas incluem atividade antioxidante, imunomoduladora, antitumoral, antimicrobiana e antiparasitária (DURÁN *et al.*, 2007; KONZEN *et al.*, 2006).

Além disso, várias espécies da família possuem propriedades quitinolíticas. Vainsh-

tein *et al.* (2020), por exemplo, relatou a presença de genes industrialmente relevantes para para modificação de quitina na estirpe *Andreprevotia sp.* IGB-42, sendo eles: 27 quitinases putativas da família de glicosídeo hidrolase 18 (GH18), duas da família GH19 (EC 3.2.14), uma suposta β -N-acetilhexosaminidase da família GH20 (EC 3.2.1.52) e uma quitosanase putativa de a família GH46 (EC 3.2.1.132). Quitinases possuem ampla gama de aplicações biotecnológicas com importância em áreas da medicina, agricultura e indústria (SOUSA *et al.*, 2019). As aplicações incluem isolamento de protoplastos fúngicos, produção de quito-oligômeros, biopesticidas, atividade antifúngica, (RATHORE; GUPTA, 2015). Quito-oligômeros obtidos pela ação das quitinases também possuem ampla gama de aplicações biotecnológicas com importância em indústrias bioquímicas, alimentícias e diversas indústrias químicas (RATHORE; GUPTA, 2015). As aplicações incluem atividade antitumoral, cicatrização de feridas, atividade anti-hipertensiva e anti-inflamatória, bem como atividades antibacteriana e antifúngica (RATHORE; GUPTA, 2015). Trabalhos na literatura relatam a produção de quitinase (CvChi47) (SOUSA *et al.*, 2019) e de quitosanase (CvCsn46) Azevedo *et al.* (2020) recombinante de *Chromobacterium violaceum* ATCC 12472 em *Escherichia coli*.

2.2.4 Patogenicidade

Infecções causadas por *Chromobacteriaceae* são muito raras, no entanto algumas espécies podem comportar-se como patógenos oportunistas. Pelo menos quatro espécies são reconhecidas como potenciais causadoras de infecções em humanos: *Chromobacterium violaceum*, *Chromobacterium haemolyticum*, *Laribacter hongkongensis* e *Vogesella perlucida*.

Infecções por *C. violaceum* podem causar abscessos no fígado, pulmão, pele ou baço e rapidamente progridem para sepse com abscesso metastático (YANG; LI, 2011). Há também relatos na literatura associando *C. violaceum* a doença crônica granulomatosa (MACHER *et al.*, 1982), osteomielite (TUCKER *et al.*, 1979), celulite periorbitária, infecção ocular (FELDMAN *et al.*, 1984) e infecção urinária (SWAIN *et al.*, 2014). Apesar de raras, são infecções sérias e de alta mortalidade. Casos de infecção por *C. violaceum* foram relatados algumas vezes no Brasil (FERNANDES *et al.*, 2014).

Infecções por *C. haemolyticum* podem causar lesões cutâneas com rápido progresso, febre e dor nas áreas infectadas, assim como infecções por *C. violaceum*, podendo ser facilmente confundidas (OKADA *et al.*, 2013). Há relatos na literatura associando *C. haemolyticum* a pneumonia (KANAMORI *et al.*, 2020) e a proctocolite (TANPOWPONG *et al.*, 2014).

Infecção por *L. hongkongensis* foi relatada pela primeira vez após ser isolada do sangue e empiema de um paciente cirrótico, causando febre alta leucocitose e neutrofilia. (YUEN *et al.*, 2001). Há relatos na literatura associando *L. hongkongensis* a infecções gastrointestinais, com diarreia aguda e prolongada (RAJA; GHOSH, 2014; ENGSBRO *et al.*, 2018).

Infecção por *V. perlucida* foi recentemente reportado por Yu *et al.* (2020) causando inflamação nos membros superiores e inferiores, rubor, dor e febre alta em uma paciente idosa de 71 anos.

3 TRABALHOS RELACIONADOS

Durante a realização de uma análise filogenética, os métodos e os parâmetros utilizados devem ser escolhidos cuidadosamente, de forma com que se evite a construção de uma falsa topologia. Trabalhos na literatura tem utilizado diferentes abordagens para a reconstrução de uma filogenia. A seguir são descritos alguns desses trabalhos, enfatizando os métodos e parâmetros utilizados.

Hedlund e Staley (2002) realizou uma análise filogenética envolvendo 16 estirpes de bactérias do gênero *Simosinela*. Em seu trabalho, foi criada uma árvore baseada nos alinhamentos do gene 16S rRNA, utilizando a espécie *Comamonas testosteroni* como grupo externo. Os alinhamentos foram realizados manualmente ou utilizando o RDP Aligner. A checagem foi realizada manualmente no GeneDoc e as posições contendo lacunas foram removidas. O alinhamento foi analisado no DNAPARS usando ordem de entrada de sequência aleatória e 10 repetições. O resultado mais parcimonioso foi importado para o MACCLADE. Foi determinado a razão de transições e transversões e a frequência de base de nucleotídeos. Posteriormente esses valores foram utilizados na análise de verossimilhança. A análise de parcimônia foi realizada no PAUP com ordem de entrada aleatória e 10 replicatas e a análise de distância foi realizada no TREECON.

Christensen *et al.* (2004) realizou uma análise filogenética envolvendo 28 estirpes de bactérias da família *Pasteurellaceae*. Em seu trabalho, foram criadas cinco árvores baseadas nos alinhamentos individuais de 3 genes de manutenção (atpD, infB and rpoB). As sequências foram alinhadas utilizando o PILEUP e as posições com *gaps* foram excluídas. A análise filogenética foi feita utilizando o PROTML usando a matriz de substituição PAM. A análise MP foi realizada utilizando o PROTPARS, com parâmetros padrões, e a análise *bootstrap* foi incluída com o PHYLIP. Uma matriz de distância foi construída utilizando o PROTDIST e uma análise NJ foi realizada no NEIGHBOR usando a matriz de substituição PAM. Uma árvore consenso para cada alinhamento foi criada utilizando o PHYLIP.

Adeolu e Gupta (2013) realizou uma análise filogenética envolvendo 94 estirpes de bactérias da ordem *Neisseriales*. Em seu trabalho, foram criadas duas árvores baseadas no gene 16S rRNA. Além disso, foram criadas duas árvores baseadas no alinhamento concatenado dos genes de 20 proteínas conservadas (UvrD, GyrA, GyrB, RpoB, RpoC, EF-G, EF-Tu, RecA, ArgRS, IleRS, ThrRS, TrpRS, SecY, DnaK, L2, L5, S2, S3, e S9), utilizando as espécies *Bordetella pertussis* e *Burkholderia ambifaria* como grupo externo. Os alinhamentos múltiplos

foram realizados utilizando o Clustal X 1.83 e regiões mal alinhadas foram removidas utilizando o Gblocks 0.91b. As árvores foram construídas no MEGA 5.05 através dos métodos NJ e ML, empregando o modelo de substituição GRT para as árvores baseadas no 16S rRNA e os modelos de substituição propostos por Whelan e Goldman (2001) e de Jones *et al.* (1992), respectivamente para as árvores NJ e ML baseada nas múltiplas proteínas. O suporte das árvores foi feito utilizando 100 réplicas bootstrap.

Adeolu *et al.* (2016) realizou uma análise filogenética envolvendo 179 estirpes de bactérias da ordem *Enterobacteriales*. Em seu trabalho, foi criada uma árvore baseada nas sequências concatenadas de 1548 proteínas core, utilizando quatro membros das famílias *Pasteurellaceae* e *Vibrionaceae* como grupo externo. A identificação das proteínas core com pelo menos 50% de identidade de sequência, 50% de tamanho da sequência e presentes em pelo menos 80% dos genomas foram realizadas utilizando o algoritmo UCLUST. As 53 proteínas ribossomais foram identificadas usando HMMer 3.1 baseado em modelos de Markov obtida na base de dados Pfam. As quatro proteínas MLSA (GyrB, RpoB, AtpD e InfB) foram identificadas usando HMMer 3.1 com base nas sequências de aminoácidos de *Escherichia coli* K12 obtida a partir de Banco de dados UniProt. Cada família de proteínas identificada foi alinhada individualmente usando Clustal Omega. Os alinhamentos foram cortados usando Gblocks 0,91b, com parâmetros relaxados, e concatenados com as outras proteínas do conjunto de dados. As árvores ML foram construídas usando o FastTree 2, empregando o modelo de Whelan e Goldman de evolução da sequência de proteínas, e o RAxML 8, empregando o modelo Le e Gascuel de evolução da sequência de proteínas. Foi realizado o suporte estatístico dos ramos utilizando o RAxML 8, pelo método Shimodaira–Hasegawa.

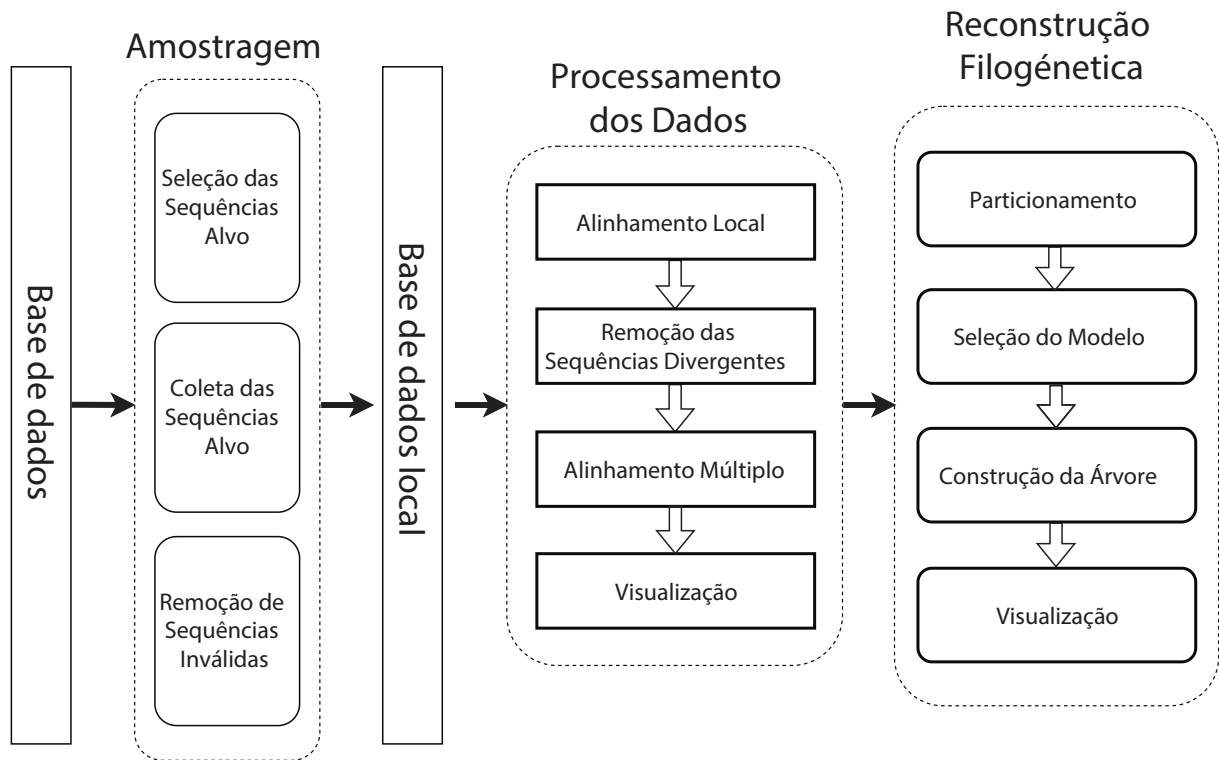
Sun *et al.* (2020) realizou uma análise filogenética envolvendo 179 espécies de fungos da família *Ganodermataceae*. Em seu trabalho, foi criada uma árvore baseada nas sequências concatenadas de 6 proteínas (ITS, nLSU, RPB1, RPB2, TEF e TUB), utilizando *Perenniporiella chaquenia* e *P. pendula* como grupos externos. Os alinhamentos foram realizados no MAFFT 7 e ajustados manualmente no BioEdit. Foram reconstruídas duas matrizes combinadas com um conjunto de 4 genes (ITS + nLSU + RPB1 + TEF) e um conjunto 6 genes (ITS + nLSU + RPB1 + RPB2 + TEF + TUB). A congruência das matrizes foram avaliadas utilizando o PAUP* v.4.0b10 pelo teste de homogeneidade de partição com 1000 replicatas. Foram selecionados os modelos de substituição Teste de Razão de Verossimilhança Hierárquica (hLRT) e Critério de Informação Akaike (AIC) utilizando o MrModeltest 2 v.2.3. As árvores ML foram construídas no

RAxML-HPC v.8.2.3 envolvendo 1.000 buscas no modelo GTRGAMMA. O suporte das árvores foram realizados 1.000 réplicas de *bootstrap* rápido com o modelo GTRCAT. A BI foi realizada no MrBayes v.3.2 com quatro cadeias simultaneas independentes, a cada 100 gerações até o valor $<0,01$ de frequência de desvio de divisão As primeiras 25% das árvores foram descartadas como burn-in. *bootstrap* ML menores que 50% e Probabilidades Bayesianas Posteriores (BPP) menores que 0,95 foram ignorados.

4 METODOLOGIA

Neste capítulo é apresentada a metodologia do trabalho. As análises foram realizadas de forma que possam ser reproduzidas posteriormente. Para um melhor entendimento a metodologia foi dividida em cinco etapas seguindo a Figura 13.

Figura 13 – Diagrama de fluxo da metodologia aplicada



Fonte: elaborado pela autora.

4.1 Coleta de Dados

A coleta de dados foi realizada a partir do Centro Nacional de Informação Biotecnológica (NCBI)¹. Foram obtidas as sequências de todas as espécies da família *Chromobacteriaceae* disponíveis no Banco de Dados de Sequências de Referência (RefSeq) (O'LEARY *et al.*, 2016) até a data de coleta (Tabela 1). Foram obtidas também as sequências de quatro espécies da família *Neisseriaceae* (*Kingella oralis* ATCC 51147, *Neisseria gonorrhoeae* FA 1090, *Neisseria meningitidis* 8013 e *Simonsiella muelleri* ATCC 29453) para compor o grupo externo.

¹ <https://www.ncbi.nlm.nih.gov/>

Tabela 1 – Relatório de montagem e anotação dos genomas

Estirpe	Montagem	Tamanho (Mb)	GC%	Bioprojeto
<i>Andreprevotia chitinilytica</i> DSM 18519	GCF_000711875.1	5,15	59,9	PRJNA234919
<i>Andreprevotia lacus</i> DSM 23236	GCF_900176275.1	4,63	62,1	PRJEB20317
<i>Andreprevotia</i> sp. IGB-42	GCF_009848685.1	4,71	61,3	PRJNA565654
<i>Aquaspirillum serpens</i> DSM 68	GCF_000420525.1	2,61	49,8	PRJNA185548
<i>Aquaspirillum</i> sp. LM1	GCF_002002905.1	3,96	61,6	PRJNA363092
<i>Aquitalea denitrificans</i> 5YN1-3	GCF_009856625.1	4,50	58,8	PRJNA598530
<i>Aquitalea magnusonii</i> CCM 7607	GCF_001515285.1	4,33	60,4	PRJNA302577
<i>Aquitalea magnusonii</i> DSM 25134	GCF_003202035.1	4,41	60,7	PRJNA463365
<i>Aquitalea magnusonii</i> H3	GCF_002217795.2	4,85	59,3	PRJDB5811
<i>Aquitalea magnusonii</i> HSC-21Su07	GCF_013911085.1	4,46	59,6	PRJNA603836
<i>Aquitalea magnusonii</i> SM6	GCF_000964065.1	4,71	58,9	PRJNA277276
<i>Aquitalea pelogenes</i> CCM 7557	GCF_001515305.1	4,39	59,5	PRJNA302577
<i>Aquitalea</i> sp. FJL05	GCF_003852595.1	4,17	60,0	PRJNA448329
<i>Aquitalea</i> sp. LB_tupeE	GCF_013391415.1	4,47	58,6	PRJNA641464
<i>Aquitalea</i> sp. MWU14-2217	GCF_003693445.1	4,29	60,3	PRJNA497635
<i>Aquitalea</i> sp. THG-DN7.12	GCF_005155025.1	4,36	59,7	PRJNA534005
<i>Aquitalea</i> sp. USM4	GCF_004328945.1	4,35	59,4	PRJNA471805
<i>Chitinibacter fontanus</i> STM-7	GCF_013423785.1	3,63	50,3	PRJNA645016
<i>Chitinibacter</i> sp. 2T18	GCF_013403565.1	3,68	50,6	PRJNA635267
<i>Chitinibacter</i> sp. GC72	GCF_009760905.1	3,46	54,0	PRJNA551952
<i>Chitinibacter</i> sp. ZOR0017	GCF_000799095.1	3,48	55,9	PRJNA205591
<i>Chitinibacter tainanensis</i> DSM 15459	GCF_000429785.1	3,43	56,1	PRJNA185565
<i>Chitinilyticum aquatile</i> DSM 21506	GCF_000430805.1	3,69	60,3	PRJNA188851
<i>Chitinilyticum litopenaei</i> DSM 21440	GCF_000428145.1	3,97	62,6	PRJNA174984
<i>Chitiniphilus shinanonensis</i> DSM 23277	GCF_000374805.1	4,15	66,4	PRJNA175157
<i>Chitiniphilus</i> sp. HX-2-15	GCF_005048205.1	4,34	64,0	PRJNA533946
<i>Chromobacterium amazonense</i> 56AF	GCF_002994485.1	4,57	61,9	PRJNA358091
<i>Chromobacterium amazonense</i> DSM 26508	GCF_001855565.1	4,01	62,6	PRJNA342530
<i>Chromobacterium haemolyticum</i> CH06-BL	GCF_009936155.1	5,31	62,6	PRJDB7639
<i>Chromobacterium haemolyticum</i> DSM 19808	GCF_000711885.1	5,03	62,2	PRJNA234889
<i>Chromobacterium haemolyticum</i> H3973	GCF_002081815.1	5,33	62,7	PRJNA369804
<i>Chromobacterium haemolyticum</i> H4137	GCF_002081825.1	5,30	62,6	PRJNA369804
<i>Chromobacterium haemolyticum</i> H5244	GCF_002081855.1	4,99	62,3	PRJNA369804
<i>Chromobacterium haemolyticum</i> HSC-31F16	GCF_011602385.1	5,72	62,3	PRJNA603836
<i>Chromobacterium haemolyticum</i> IR17	GCF_003332145.1	5,13	62,3	PRJNA308731
<i>Chromobacterium haemolyticum</i> NRRL B-11053	GCF_003052545.1	5,27	62,7	PRJNA448854
<i>Chromobacterium haemolyticum</i> T124	GCF_000758475.1	5,08	62,8	PRJNA253350
<i>Chromobacterium piscinae</i> ND17	GCF_000812585.1	4,09	62,6	PRJNA268264
<i>Chromobacterium rhizoryzae</i> Bb2	GCF_014788985.1	-	-	PRJNA664546
<i>Chromobacterium rhizoryzae</i> JP2-74	GCF_003443515.1	5,17	62,9	PRJNA488117
<i>Chromobacterium</i> sp. 257-1	GCF_008275125.1	4,28	63,6	PRJNA561765
<i>Chromobacterium</i> sp. ATCC 53434	GCF_002848345.1	5,14	65,1	PRJNA421332
<i>Chromobacterium</i> sp. C-61	GCF_000285415.1	5,12	62,7	PRJEA68181

(Continua na próxima página)

Tabela 1 – (Continuação)

Estirpe	Montagem	Tamanho (Mb)	GC%	Fonte
<i>Chromobacterium sp. F49</i>	GCF_001619865.1	4,70	64,8	PRJNA308367
<i>Chromobacterium sp. IIBBL 112-1</i>	GCF_003325475.1	4,81	63,9	PRJNA471525
<i>Chromobacterium sp. IIBBL 274-1</i>	GCF_003325495.1	4,80	63,9	PRJNA471525
<i>Chromobacterium sp. LK1</i>	GCF_001043555.1	4,98	63,6	PRJNA277300
<i>Chromobacterium sp. LK11</i>	GCF_001043705.1	5,07	65,0	PRJNA277403
<i>Chromobacterium sp. MWU13-2610</i>	GCF_002902845.1	4,38	62,4	PRJNA430714
<i>Chromobacterium sp. MWU14-2602</i>	GCF_002924365.1	4,53	63,4	PRJNA430714
<i>Chromobacterium sp. Panama</i>	GCF_003052555.1	5,22	64,0	PRJNA448854
<i>Chromobacterium sphagni IIBBL 14B-1</i>	GCF_001855555.1	4,74	63,4	PRJNA342529
<i>Chromobacterium sphagni IIBBL 37-2</i>	GCF_001855575.1	4,90	63,5	PRJNA342531
<i>Chromobacterium subtsugae Cv017</i>	GCF_001510755.2	4,77	64,5	PRJNA296470
<i>Chromobacterium subtsugae F49</i>	GCF_000812805.1	4,67	64,8	PRJNA270006
<i>Chromobacterium subtsugae MWU2387</i>	GCF_001020525.1	5,23	64,3	PRJNA280628
<i>Chromobacterium subtsugae MWU2576</i>	GCF_001020515.1	4,80	64,8	PRJNA280629
<i>Chromobacterium subtsugae MWU2920</i>	GCF_001020585.1	4,67	64,9	PRJNA280633
<i>Chromobacterium subtsugae MWU3525</i>	GCF_001020505.1	4,69	64,9	PRJNA280634
<i>Chromobacterium subtsugae PRAA4-1</i>	GCF_001676875.1	4,81	64,8	PRJNA262860
<i>Chromobacterium vaccinii 21-1</i>	GCF_001855275.1	5,04	64,3	PRJNA342532
<i>Chromobacterium vaccinii HSC-77S12</i>	GCF_011601295.1	4,94	63,6	PRJNA603836
<i>Chromobacterium vaccinii MWU205</i>	GCF_000971335.1	4,97	64,4	PRJNA277476
<i>Chromobacterium vaccinii MWU328</i>	GCF_000971355.1	4,96	64,3	PRJNA277484
<i>Chromobacterium vaccini NCTC9370</i>	GCF_900446815.1	5,15	63,7	PRJEB6403
<i>Chromobacterium vaccinii XC0014</i>	GCF_002952135.1	4,94	64,3	PRJNA392859
<i>Chromobacterium violaceum 16-419A</i>	GCF_002081735.1	4,76	65,0	PRJNA369804
<i>Chromobacterium violaceum 16-419B</i>	GCF_002081745.1	4,76	65,0	PRJNA369804
<i>Chromobacterium violaceum 16-454</i>	GCF_002081775.1	4,78	64,6	PRJNA369804
<i>Chromobacterium violaceum ATCC 12472</i>	GCF_000007705.1	4,75	64,8	PRJNA444
<i>Chromobacterium violaceum ATCC 315322</i>	GCF_002865685.1	4,76	64,7	PRJNA402426
<i>Chromobacterium violaceum CV1192</i>	GCF_002735945.1	4,39	65,6	PRJNA413993
<i>Chromobacterium violaceum CV1197</i>	GCF_002735965.1	4,39	65,6	PRJNA413993
<i>Chromobacterium violaceum FDAARGOS_635</i>	GCF_012273555.1	4,80	65,0	PRJNA231221
<i>Chromobacterium violaceum GHPS1</i>	GCF_002179535.1	4,82	64,9	PRJNA387362
<i>Chromobacterium violaceum GN5</i>	GCF_000812485.1	4,87	64,7	PRJNA269999
<i>Chromobacterium violaceum H5524</i>	GCF_002081875.1	4,73	64,7	PRJNA369804
<i>Chromobacterium violaceum H5525</i>	GCF_002081895.1	4,72	64,8	PRJNA369804
<i>Chromobacterium violaceum L_IB5_1</i>	GCF_000952105.1	4,86	64,6	PRJNA275561
<i>Chromobacterium violaceum LK15</i>	GCF_001043755.1	4,75	65,0	PRJNA277407
<i>Chromobacterium violaceum LK17</i>	GCF_001043735.1	4,76	65,0	PRJNA277410
<i>Chromobacterium violaceum LK30</i>	GCF_001043815.1	4,75	64,9	PRJNA277426
<i>Chromobacterium violaceum LK6</i>	GCF_001043595.1	4,76	65,0	PRJNA277398
<i>Chromobacterium violaceum LMG 3953</i>	GCF_001953795.1	4,63	64,7	PRJNA353943
<i>Chromobacterium violaceum NCTC8684</i>	GCF_900460965.1	4,86	64,4	PRJEB6403
<i>Chromobacterium violaceum NCTC9695</i>	GCF_900635845.1	4,73	64,9	PRJEB6403

(Continua na próxima página)

Tabela 1 – (Continuação)

Estirpe	Montagem	Tamanho (Mb)	GC%	Fonte
<i>Chromobacterium violaceum</i> NCTC9757	GCF_900446805.1	4,75	64,8	PRJEB6403
<i>Chromobacterium violaceum</i> RDN09	GCF_014117535.1	4,74	64,9	PRJNA640761
<i>Deefgea rivuli</i> DSM 18356	GCF_000620145.1	3,71	49,0	PRJNA220991
<i>Deefgea</i> sp. CFHI-16	GCF_014042245.1	3,45	48,1	PRJNA648006
<i>Deefgea</i> sp. D17	GCF_013284055.1	3,43	48,3	PRJNA634984
<i>Formivibrio citricus</i> DSM 6150	GCF_900115065.1	3,07	58,8	PRJEB17128
<i>Gulbenkiania indica</i> DSM 17901	GCF_001418035.1	2,85	63,6	PRJNA289009
<i>Gulbenkiania indica</i> DSM 17901	GCF_001517245.1	2,85	63,6	PRJEB10555
<i>Gulbenkiania mobilis</i> DSM 18507	GCF_004346645.1	2,87	63,7	PRJNA520311
<i>Gulbenkiania mobilis</i> MB1	GCF_001302325.1	3,42	62,0	PRJNA293922
<i>Iodobacter fluviatilis</i> DSM 3764	GCF_004346505.1	4,96	48,7	PRJNA520324
<i>Iodobacter fluviatilis</i> NCTC11159	GCF_900451195.1	5,10	48,8	PRJEB6403
<i>Iodobacter fluviatilis</i> PCH194	GCF_004194535.1	4,59	47,5	PRJNA428922
<i>Iodobacter</i> sp. BJB302	GCF_002735645.1	4,81	49,6	PRJNA251750
<i>Iodobacter</i> sp. H11R3	GCF_003952345.1	3,88	48,0	PRJNA509323
<i>Iodobacter</i> sp. HSC-16F04	GCF_011601265.1	4,78	49,4	PRJNA603836
<i>Jeongeupia chitinilytica</i> KCTC 23701	GCF_014652315.1	3,91	64,0	PRJDB10508
<i>Jeongeupia</i> sp. USM3	GCF_001808185.1	3,79	65,4	PRJNA345402
<i>Kingella oralis</i> ATCC 51147	GCF_000160435.1	2,41	54,3	PRJNA33145
<i>Laribacter hongkongensis</i> DSM 14985	GCF_000423285.1	3,11	62,4	PRJNA185634
<i>Laribacter hongkongensis</i> HLGZ1	GCF_002215055.1	3,42	62,3	PRJNA389628
<i>Laribacter hongkongensis</i> HLHK9	GCF_000021025.1	3,17	62,4	PRJNA30817
<i>Laribacter hongkongensis</i> MGYG-HGUT-02398	GCF_902386745.1	3,42	62,3	PRJEB33885
<i>Leeia oryzae</i> DSM 17879	GCF_000376945.1	3,79	54,5	PRJNA182391
<i>Leeia</i> sp. IMCC25680	GCF_012641365.1	3,56	60,7	PRJNA623991
<i>Microvirgula aerodenitrificans</i> BE2.4	GCF_003051865.1	4,13	64,3	PRJNA448131
<i>Microvirgula aerodenitrificans</i> DSM 15089	GCF_000620105.1	4,04	64,2	PRJNA221023
<i>Microvirgula</i> sp. AG722	GCF_003269145.1	4,03	64,4	PRJNA463042
<i>Neisseria gonorrhoeae</i> FA 1090	GCF_000006845.1	2,15	52,7	PRJNA23
<i>Neisseria meningitidis</i> 8013	GCF_000026965.1	2,28	51,4	PRJEA34687
<i>Paludibacterium paludis</i> KCTC 32182	GCF_014652495.1	4,00	62,0	PRJDB10508
<i>Paludibacterium purpuratum</i> CECT 8976	GCF_004363805.1	4,04	60,7	PRJNA463291
<i>Paludibacterium</i> sp. dN 18-1	GCA_009707705.1	3,35	57,3	PRJNA589943
<i>Paludibacterium yongneupense</i> DSM 18731	GCF_000422925.1	4,33	62,2	PRJNA185612
<i>Pseudogulbenkiania ferrooxidans</i> 2002	GCF_000174355.1	4,23	64,7	PRJNA30761
<i>Pseudogulbenkiania ferrooxidans</i> EGD-HP2	GCF_000462205.1	4,80	64,2	PRJNA215707
<i>Pseudogulbenkiania</i> sp. MAI-1	GCF_000527175.1	4,34	64,2	PRJNA82805
<i>Pseudogulbenkiania</i> sp. NH8B	GCF_000283535.1	4,33	64,4	PRJDA68325
<i>Pseudogulbenkiania subflava</i> DSM 22618	GCF_900177275.1	4,37	63,4	PRJEB20430
<i>Silvimonas amylolytica</i> CGMCC 1.8860	GCF_014645555.1	4,56	57,0	PRJDB10509
<i>Silvimonas iriomotensis</i> CGMCC 1.8859	GCF_014645535.1	4,59	59,0	PRJDB10509
<i>Silvimonas terrae</i> DSM 18233	GCF_014202765.1	4,83	59,0	PRJNA632214
<i>Simonsiella muelleri</i> ATCC 29453	GCF_002951835.1	2,47	41,5	PRJNA361301

(Continua na próxima página)

Tabela 1 – (Continuação)

Estirpe	Montagem	Tamanho (Mb)	GC%	Fonte
<i>Vogesella alkaliphila</i> KCTC 32041	GCF_014652475.1	3,42	65,0	PRJDB10508
<i>Vogesella fluminis</i> KCTC 23713	GCF_014652335.1	3,30	63,0	PRJDB10508
<i>Vogesella indigofera</i> DSM 3303	GCF_003633895.1	3,62	64,3	PRJNA442674
<i>Vogesella mureinivorans</i> 389	GCF_007644035.1	4,18	61,3	PRJNA552197
<i>Vogesella oryzae</i> Q2602	GCF_902829295.1	3,42	62,3	PRJEB37457
<i>Vogesella perlucida</i> DS-28	GCF_007844155.1	4,33	59,5	PRJNA552196
<i>Vogesella</i> sp. EB	GCF_001037925.1	3,54	64,4	PRJNA286802
<i>Vogesella</i> sp. LIG4	GCF_900090205.1	4,15	63,9	PRJEB14437
<i>Vogesella urethralis</i> YM-1	GCF_007644045.2	4,03	63,2	PRJNA552191

Nota: as informações genômicas foram coletadas de <https://www.ncbi.nlm.nih.gov/genome/browse#!/overview/>.

Fonte: elaborado pela autora

Foram obtidas também as sequências de aminoácidos de 19 genes de manutenção (Adk, AtpG, EF-G, FusA, GltX, Gmk, GyrA, GyrB, L5, LipA, MurC, MurE, PgK, RpoB, S2, S3, SecY, ThrRS, UvrD) e as sequências de nucleotídeos dos genes 16S e 23S rRNA, comumente utilizados em esquemas de MSA. Todos os genes coletados são pertencentes ao *Chromobacterium violaceum* ATCC 12472

4.2 Alinhamento Local

As sequências foram utilizadas para a criação de um banco de dados local. A partir desse banco de dados, foram feitos alinhamentos locais onde os 21 genes coletados foram usados como sequências de consulta afim de obter as sequências dos genes para cada espécie da família Chromobacteriaceae. Os parâmetros utilizados no alinhamento local podem ser conferidos na Tabela 2.

Tabela 2 – Parâmetros utilizados no alinhamento local

Parâmetros	Valor
evaluate	1e-10
max_target_seqs	5
max_hsps	1
best_hit_overhang	0.1
best_hit_score_edge	0.1

Fonte: elaborado pela autora.

Tanto o banco de dados local como as buscas foram realizadas utilizando o BLAST+ (CAMACHO *et al.*, 2009). As sequências obtidas a partir do resultado da busca BLAST foram

agrupadas por gene em arquivos no formato FASTA e analisados visualmente à procura de divergências. A visualização das sequências foram realizadas no MEGA X (KUMAR *et al.*, 2018). Sequências de tamanhos discrepantes e/ou com dados perdidos foram excluídas da análise.

4.3 Alinhamento Múltiplo

Após análise visual, as sequências agrupadas por gene foram alinhadas. Os alinhamentos múltiplos foram feitos utilizando o MAFFT (KATO; STANDLEY, 2013). Os parâmetros do MSA foram definidos automaticamente pelo software de acordo com o tipo e tamanho dos dados. Dois arquivos concatenados foram criados após o alinhamento, um agrupando as sequências 16S e 23S rRNA e outro agrupando os demais genes de manutenção. Todas as manipulações dos arquivos (seleção, concatenação, conversão e identificação) foram feitas utilizando Biopython (COCK *et al.*, 2009).

4.4 Particionamento e Modelo de Substituição

Os alinhamentos múltiplos concatenados foram particionados considerando comprimentos de ramos vinculados e utilizando o algoritmo *greedy* no PartitionFinder 2 (LANFEAR *et al.*, 2012). Para definir o modelo mais apropriado na construção da árvore filogenética cada partição foi submetida a uma ferramenta de teste. Para modelos de aminoácidos foi utilizado Prottest 3 (DARRIBA *et al.*, 2011) e para modelos de nucleotídeos Jmodeltest 2 (DARRIBA *et al.*, 2012). O melhor modelo foi selecionado dentre 24 modelos de acordo com o Critério de Informação Akaike Corrigido (AICc) e a Teoria da Decisão (DT).

4.5 Reconstrução Filogenética

A construção da árvore filogenética foi realizada utilizando RaxML-NG (KOZLOV *et al.*, 2019) disponível no CIPRES² (MILLER *et al.*, 2011). Foram criadas 1000 árvores a partir do método estatístico de máxima verossimilhança para a seleção da melhor árvore. Foi aplicado o teste de confiança *bootstrap* com 1000 replicações para o suporte dos ramos. A árvore foi visualizada utilizando a ferramenta iTOL v4 (LETUNIC; BORK, 2019) e os ramos com suporte abaixo de 70 foram ignorados.

² <http://www.phylo.org/>

5 RESULTADOS E DISCUSSÃO

Nesse capítulo serão apresentados e discutidos os resultados da aplicação da metodologia descrita no Capítulo 4.

Durante a fase de coleta de dados foi possível obter as sequências de 134 estirpes pertencentes a *Chromobacteriaceae*. A busca BLAST retornou 2835 resultados de melhor sucesso. Na fase de visualização, devido a tamanhos discrepantes e/ou dados perdidos, foram excluídas 8 estirpes (*Aquaspirillum serpens* DSM 68, *Chitinilyticum aquatile* DSM 21506, *Chitinilyticum litopenaei* DSM 21440, *Chromobacterium haemolyticum* DSM 19808, *Chromobacterium sp. C-61*, *Chromobacterium sp. MWU13-2610*, *Deefgea rivuli* DSM 18356, *Gulbenkiania indica* DSM 17901, *Paludibacterium yongneupense* DSM 18731 e *Pseudogulbenkiania subflava* DSM 22618) dos arquivos de genes 16S/23S rRNA e 6 estirpes (*Aquitalea magnusonii* CCM 7607, *Chromobacterium sp. C-61*, *Chromobacterium vaccinii* NCTC9370, *Chromobacterium violaceum* NCTC9695, *Deefgea sp. CFH1-16* e *Paludibacterium sp. dN 18-1*) dos arquivos de proteínas conservadas. Após concatenação (Tabela 3), o arquivo 1 agrupou 126 sequências alinhadas com 4609 bases cada e o arquivo 2 agrupou 131 sequências alinhadas com 10779 aminoácidos cada.

Tabela 3 – Blocos de genes concatenados para cada arquivo

Arquivo 1		Arquivo 2	
Gene	Bloco	Gene	Bloco
16S rRNA	1-1561	Adk	1-269
23S rRNA	1562-4609	AtpG	270-564
		EFG	565-1266
		FusA	1267-1968
		GltX	1969-2447
		Gmk	2448-2662
		GyrA	2663-3703
		GyrB	3704-4566
		L5	4567-4746
		LipA	4747-5099
		MurC	5100-5588
		MurE	5589-6586
		PgK	6587-7010
		RpoB	7011-8415
		S2	8416-8665
		S3	8666-8906
		SecY	8907-9363
		ThrRS	9364-10005
		UvrD	10006-10779

Fonte: elaborado pela autora.

O Arquivo 1 foi dividido em duas partições ((16S)(23S)) e o Arquivo 2 foi dividido

em dez partições ((AtpG, Adk, ThrRS, MurC) (FusA, RpoB, EFG) (PgK, GltX) (MurE, UvrD, Gmk) (GyrA) (GyrB) (L5) (LipA) (S3, S2) (SecY)).

O teste de modelo de substituição foi aplicado para cada partição individualmente.

Os modelos selecionados e suas respectivas probabilidades podem ser vistos na Tabela 4.

Tabela 4 – Resultados do teste de modelo de substituição

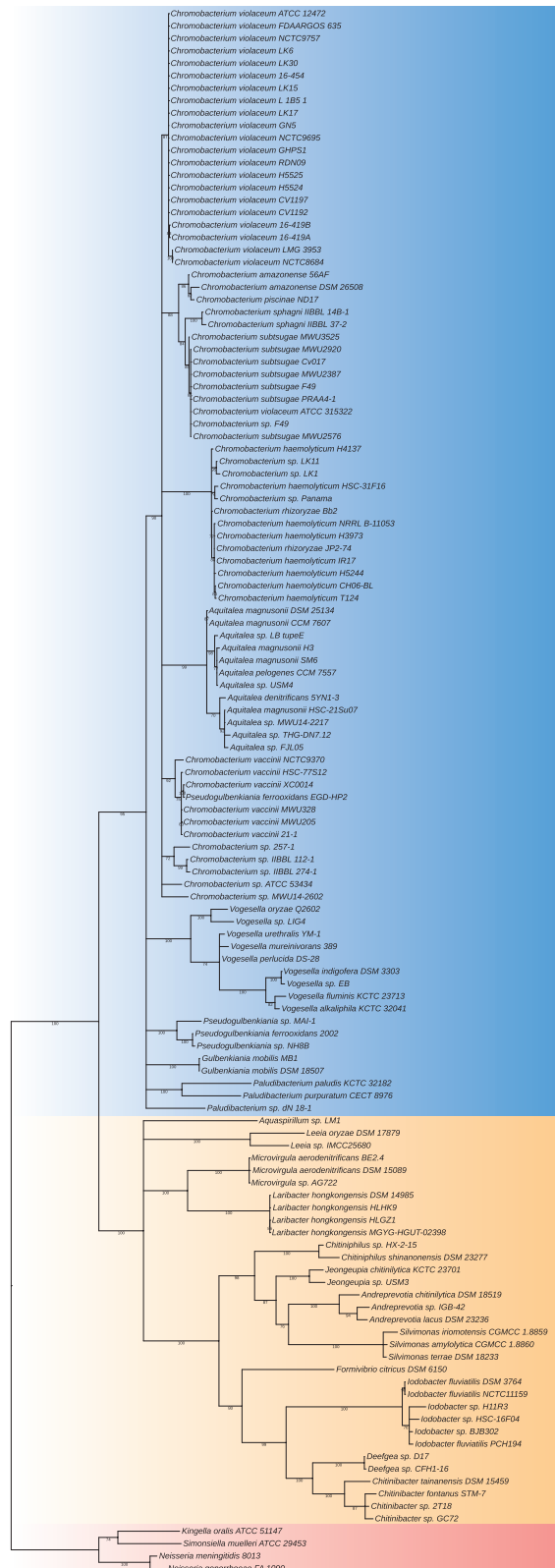
Partição	Modelo	AICc	DT
(16S rRNA)	TIM1+I+G	23495.51	3.01
(23S rRNA)	GTR+I+G	47195.47	0.02
(AtpG, Adk, ThrRS, MurC)	LG+I+G+F	82760.88	0.00
(FusA, RpoB, EFG)	LG+I+G+F	94020.99	0.00
(PgK, GltX)	LG+I+G+F	51332.50	0.00
(MurE, UvrD, Gmk)	LG+I+G+F	116042.69	0.00
(GyrA)	LG+I+G+F	46775.89	0.00
(GyrB)	LG+I+G	41101.21	0.00
(L5)	LG+G	3094.16	0.00
(LipA)	LG+I+G	15892.77	0.00
(S3, S2)	LG+I+G	16190.29	0.00
(SecY)	LG+G	19140.85	0.09

Fonte: elaborado pela autora.

As árvores ML podem ser conferidas nas Figuras 14 e 15. Para uma melhor visualização, as árvores tiveram alguns dos ramos colapsados e divididos em sub-árvores nas Figuras 16, 17, 18, 19, 20, 21, 22 e 23.

Figura 14 – Árvore ML baseada nos genes 16/23S rRNA

Escala: 0.1



Nota – Clado I em azul, Clado II em laranja e Grupo externo em vermelho. O suporte *bootstrap* pode ser visualizado próximo a cada ramificação. A barra localizada no canto superior esquerdo define a escala dos comprimentos dos ramos de acordo com a distância filogenética.

Fonte: elaborado pela autora.

Figura 15 – Árvore ML baseada nos genes de manutenção

Escala: 1



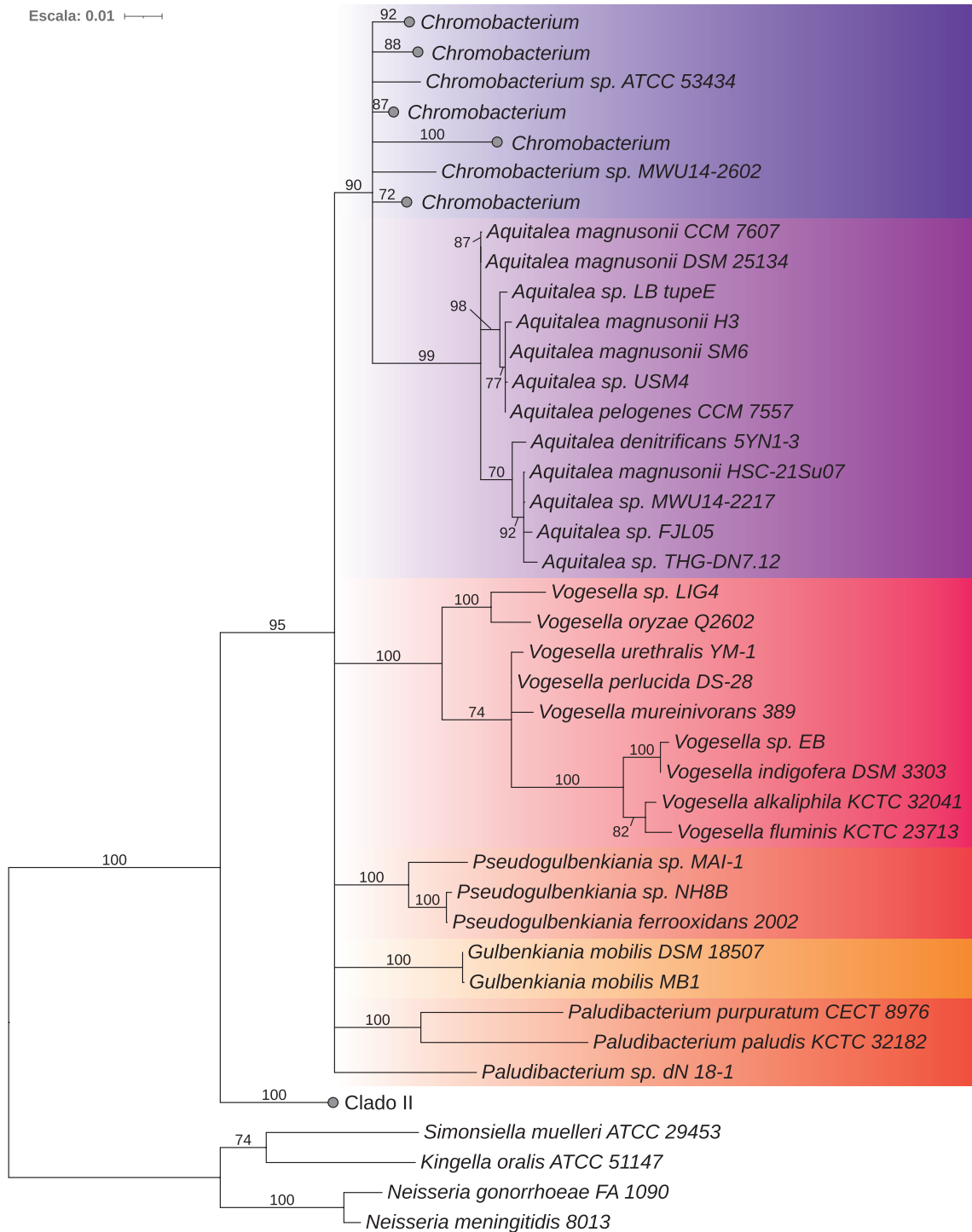
Nota: Clado I em azul, Clado II em laranja e Grupo externo em vermelho. O suporte *bootstrap* pode ser visualizado próximo a cada ramificação. A barra localizada no canto superior esquerdo define a escala dos comprimentos dos ramos de acordo com a distância filogenética.

Fonte: elaborado pela autora.

A árvore baseada no gene 16S/23S rRNA mostra dois grupos monofiléticos distintos dentro da família *Chromobacteriaceae*. Clado I (Figura 16) agrupando as estirpes de 12 gêneros: *Aquaspirillum*, *Microvirgula*, *Laribacter*, *Leeia*, *Iodobacter*, *Formivibrio*, *Andreprevotia*, *Chitiniphilus*, *Jeongeupia*, *Chitinilyticum*, *Deefgea* e *Chitinibacter*. Clado II (Figura 17) agrupando as estirpes de 6 gêneros: *Paludibacterium*, *Gulbenkiania*, *Pseudogulbenkiania*, *Vogesella*, *Aquitalea* e *Chromobacterium*. A presença de dois grandes clados bem definidos, somado a futura identificação de CSIs para cada clado, abre a possibilidade para uma divisão dessa família em duas famílias distintas, similar ao trabalho de Adeolu e Gupta (2013), que propôs a divisão da família *Neisseriaceae*.

Analisando a árvore mais de perto, é possível perceber que todos os gêneros que compõem a família formam grupos monofiléticos, com exceção do gênero *Chromobacterium* (Figuras 18 e 19), que apresenta-se formando politomias junto ao ramo do gênero *Aquitalea*, e *Paludibacterium*. Essas politomias podem ocorrer devido a falta de informação suficiente determinar a ordem de ramificação dos grupos ou devido a vários eventos de divisão simultâneos (DESALLE *et al.*, 1994).

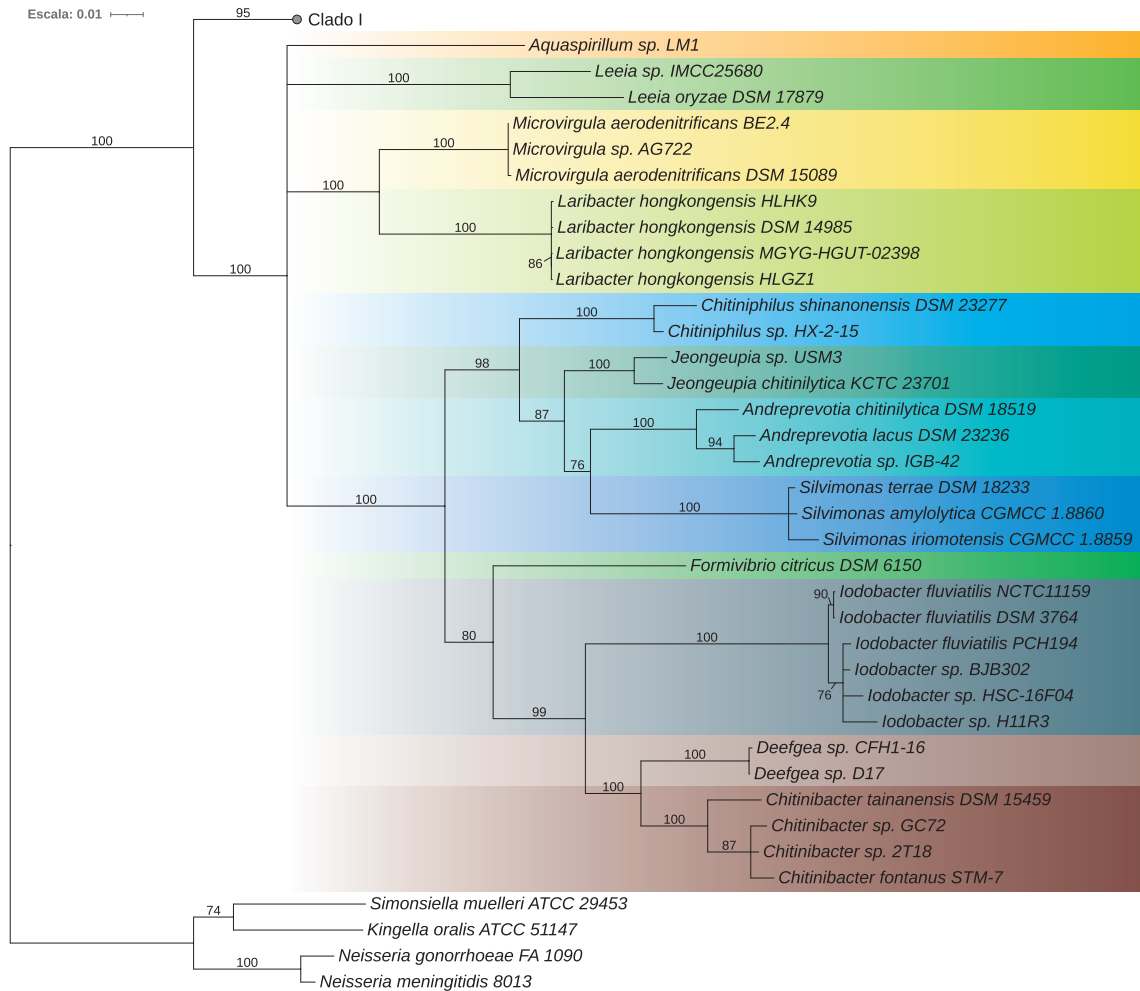
Figura 16 – Clado I da árvore ML baseada nos genes 16/23S rRNA



Nota: O suporte *bootstrap* pode ser visualizado próximo a cada ramificação. A barra localizada no canto superior esquerdo define a escala dos comprimentos dos ramos de acordo com a distância filogenética.

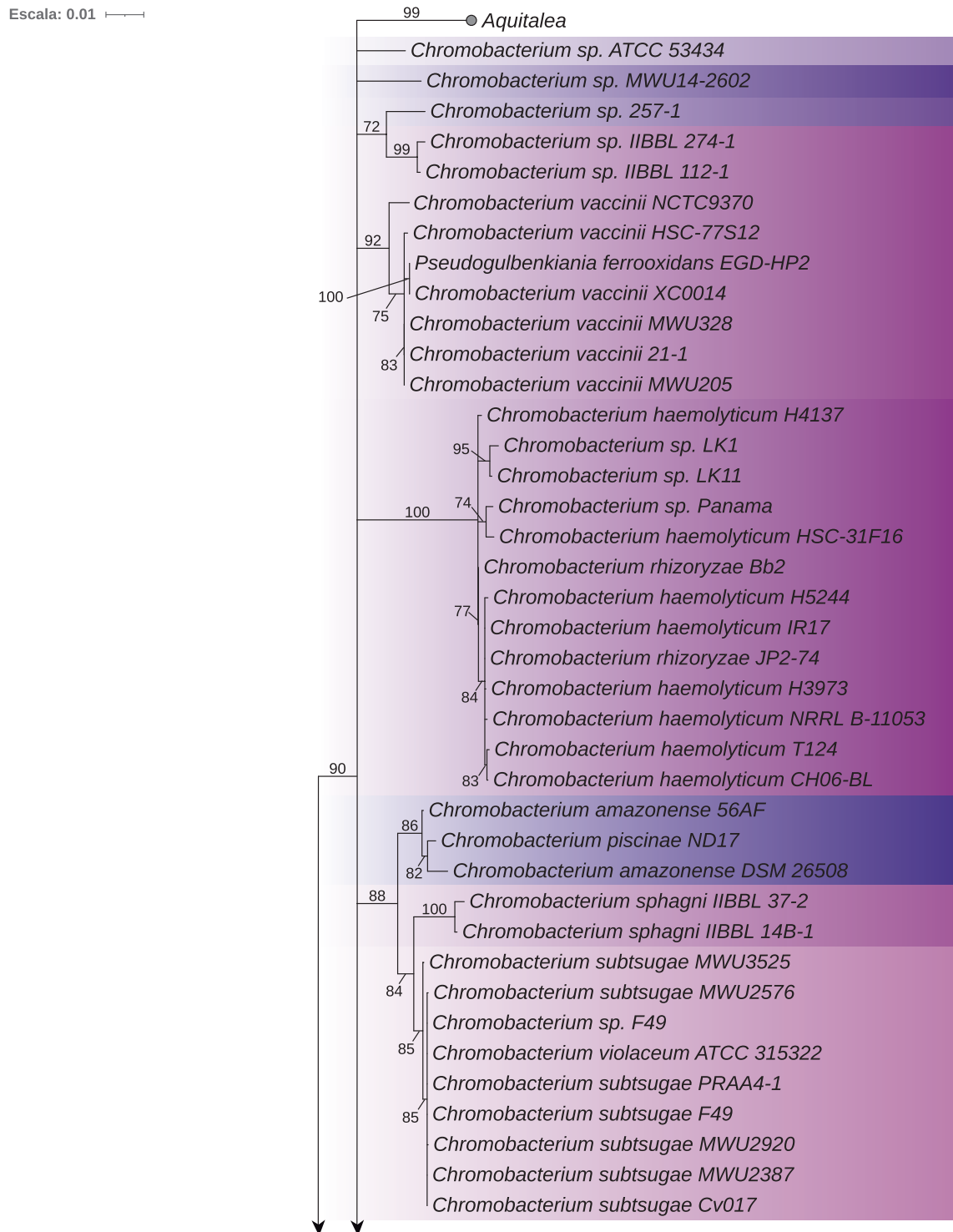
Fonte: elaborado pela autora.

Figura 17 – Clado II da árvore ML baseada nos genes 16/23S rRNA



Nota: O suporte *bootstrap* pode ser visualizado próximo a cada ramificação. A barra localizada no canto superior esquerdo define a escala dos comprimentos dos ramos de acordo com a distância filogenética. Fonte: elaborado pela autora.

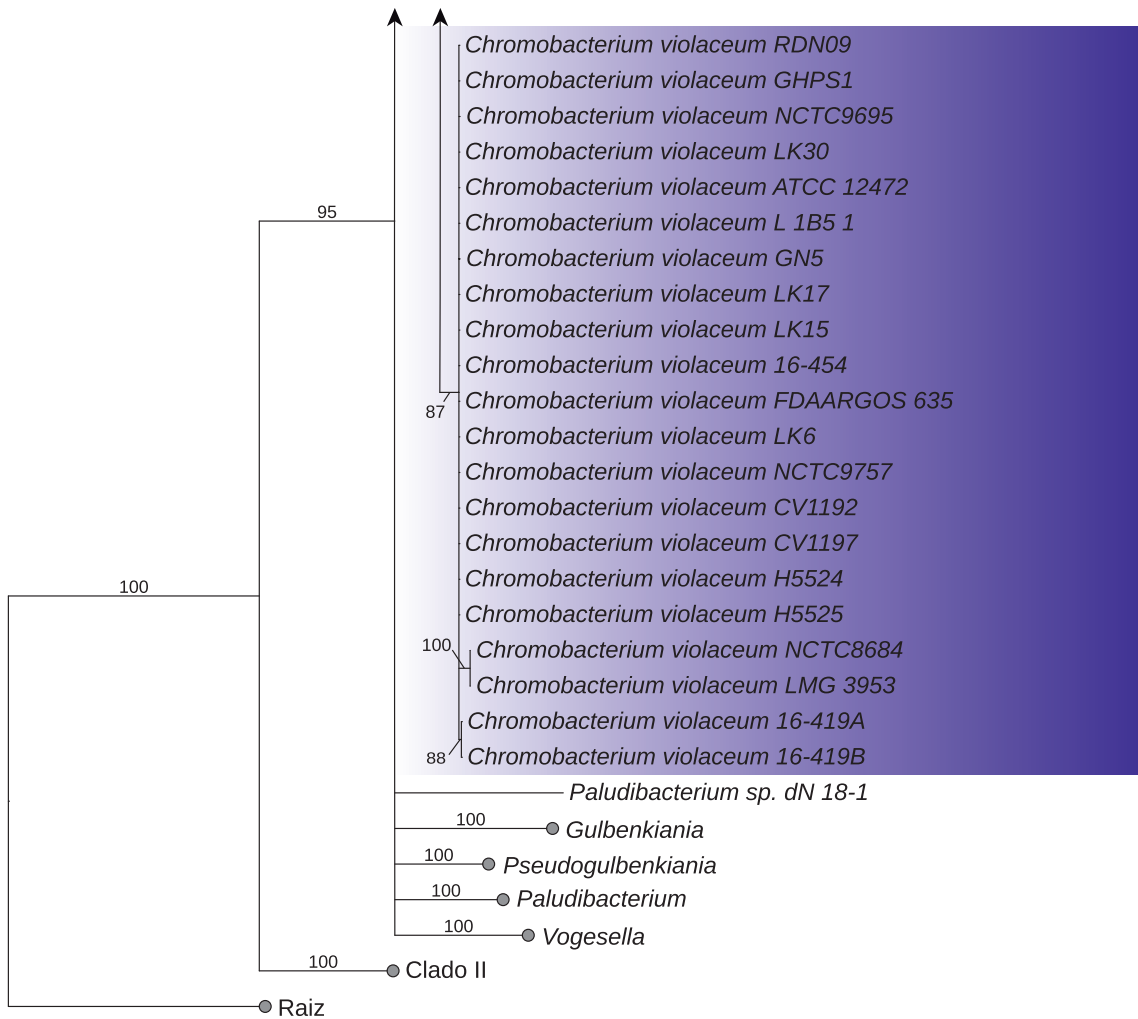
Figura 18 – Gênero *Chromobacterium* na árvore ML baseada nos genes 16/23S rRNA



Nota: O suporte *bootstrap* pode ser visualizado próximo a cada ramificação. A barra localizada no canto superior esquerdo define a escala dos comprimentos dos ramos de acordo com a distância filogenética.

Fonte: elaborado pela autora.

Figura 19 – Gênero *Chromobacterium* na árvore ML baseada nos genes 16/23S rRNA (continuação)

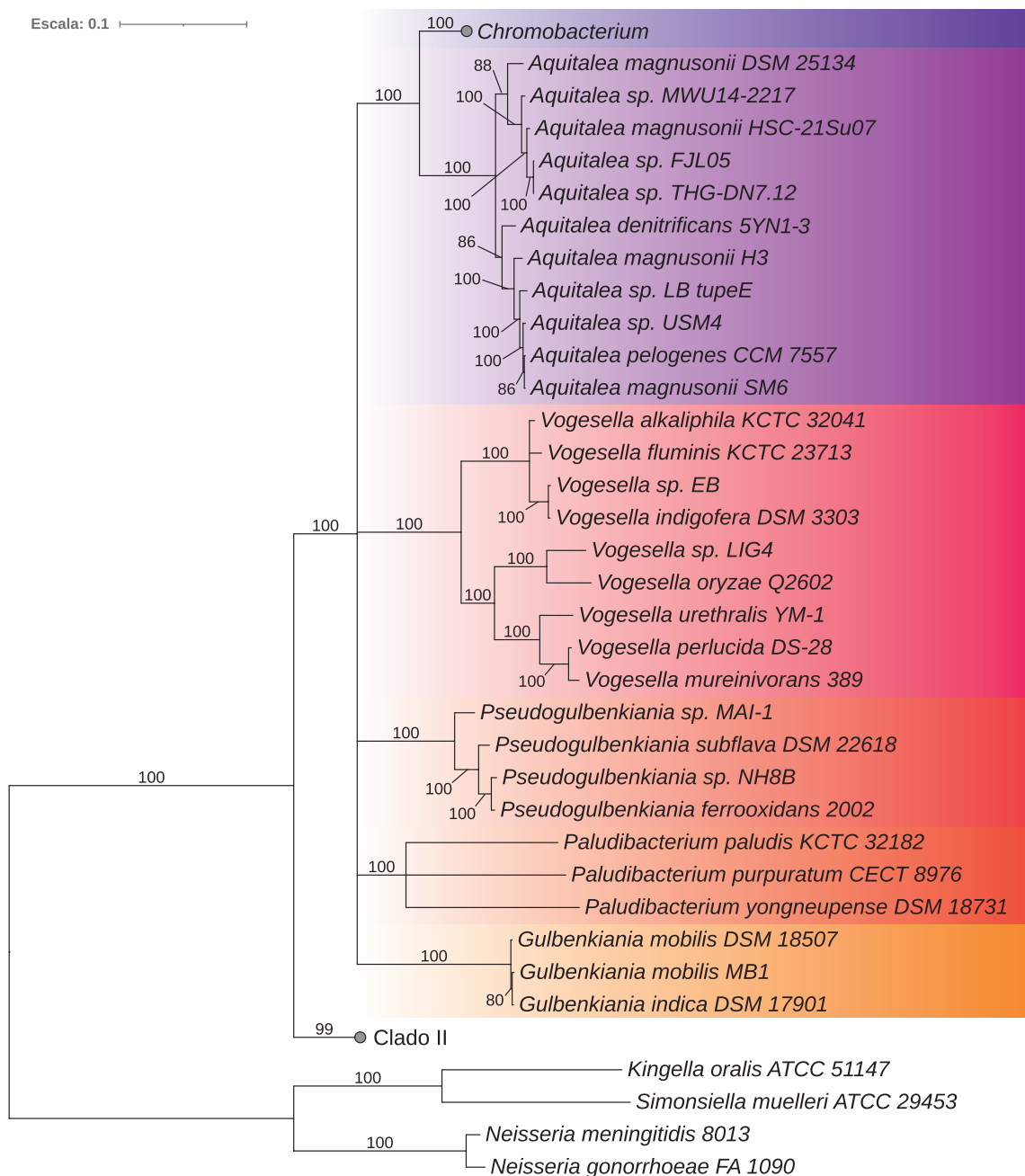


Nota: O suporte *bootstrap* pode ser visualizado próximo a cada ramificação. A barra localizada no canto superior esquerdo define a escala dos comprimentos dos ramos de acordo com a distância filogenética.

Fonte: elaborado pela autora.

A árvore baseada nos genes de manutenção (Figuras 20 e 21) corroborou a árvore baseada no RNA ribossômico, apresentando uma topologia muito similar. Entretanto, ao comparar as topologias, é possível perceber que a árvore baseada nos genes de manutenção mostra as relações evolutivas com uma melhor resolução. A maior capacidade de resolução das filogenias baseadas em múltiplos genes, quando comparadas a filogenias baseadas em um único gene, também foi demonstrada em outros trabalhos na literatura (GLAESER; KÄMPFER, 2015).

Figura 20 – Árvore ML baseada nos genes de manutenção - Clado I

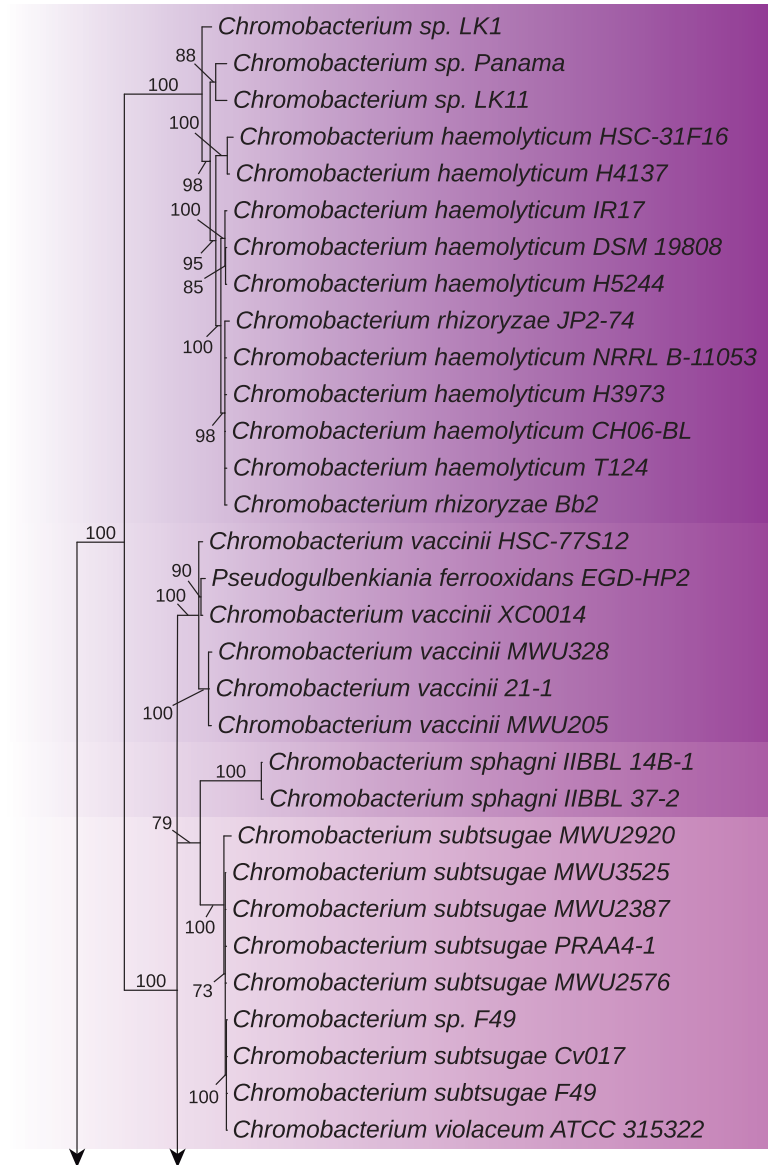


Nota: O suporte *bootstrap* pode ser visualizado próximo a cada ramificação. A barra localizada no canto superior esquerdo define a escala dos comprimentos dos ramos de acordo com a distância filogenética.

Fonte: elaborado pela autora.

Figura 22 – Árvore ML baseada nos genes de manutenção - Chromobacterium

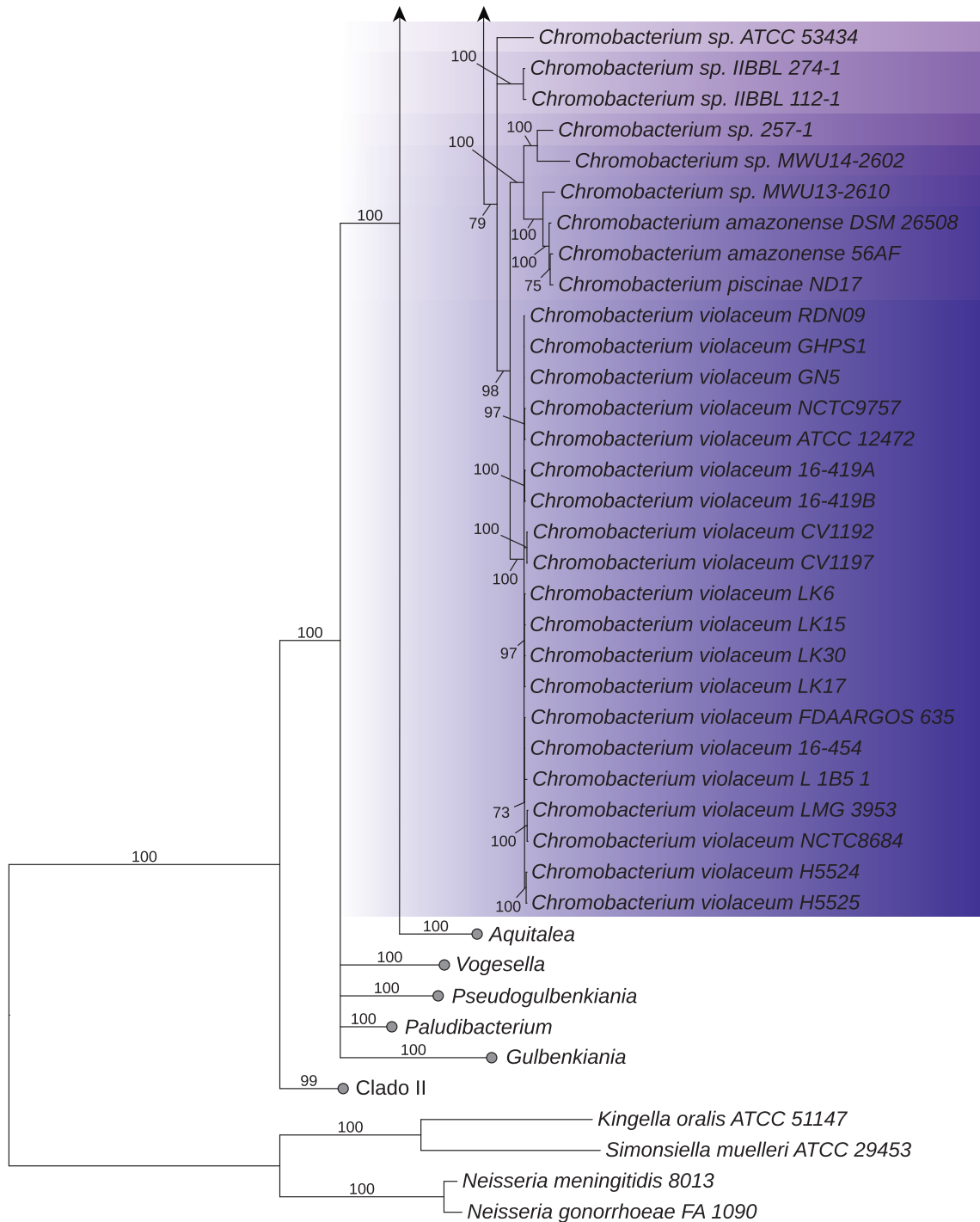
Escala: 0.1



Nota: O suporte *bootstrap* pode ser visualizado próximo a cada ramificação. A barra localizada no canto superior esquerdo define a escala dos comprimentos dos ramos de acordo com a distância filogenética.

Fonte: elaborado pela autora.

Figura 23 – Árvore ML baseada nos genes de manutenção - *Chromobacterium* (continuação)



Nota: O suporte *bootstrap* pode ser visualizado próximo a cada ramificação. A barra localizada no canto superior esquerdo define a escala dos comprimentos dos ramos de acordo com a distância filogenética.

Fonte: elaborado pela autora.

A principal diferença entre as duas topologias deve-se aos gêneros *Chromobacterium* (Figuras 22 e 23) e *Paludibacterium* formarem grupos monofiléticos na segunda árvore, indicando que a politomia da primeira árvore provavelmente ocorreu devido a falta de informação suficiente.

Assim, podemos afirmar que os resultados desse estudo são consistentes com a atual classificação desses táxons como gêneros independentes.

Comparando os resultados deste trabalho aos resultados do trabalho de Adeolu e Gupta (2013), podemos fazer algumas observações. Apesar da árvore 16S rRNA, apresentada por Adeolu e Gupta, não incluir grupo externo, ainda é possível perceber a formação de dois cladogramas além do clado da família *Neisseriaceae*. Um clado incluindo os gêneros *Gulbenkiania*, *Pseudogulbenkiania*, *Paludibacterium*, *Aquitalea* e *Chromobacterium* e outro incluindo os gêneros *Chitiniphilus*, *Formivibrio*, *Chitinilyticum*, *Iodobacter*, *Chitinibacter*, *Deefgea*, *Jeongeupia*, *Andreprevotia* e *Silvimonas*. A ausência do gênero *Vogesella* no primeiro clado e dos gêneros *Laribacter*, *Aquaspirillum*, *Leeia* e *Microvirgula* no segundo, pode ser explicada pela dificuldade de distinção confiável de qualquer subgrupo dentro da classe *Neisseriales*, baseando-se somente nas ramificações da árvore 16S rRNA, ponto que foi evidenciado no trabalho Adeolu e Gupta. A árvore baseada em múltiplas proteínas de manutenção apresentada por Adeolu e Gupta incluiu somente quatro estirpes de *Chromobacteriaceae*. Por esse motivo, não ficou clara a presença de um terceiro clado dentro da classe *Neisseriales* em seu estudo.

Voltando para a análise das árvores apresentadas nesse trabalho, percebe-se que as estirpes *Pseudogulbenkiania ferrooxidans* EGD-HP2 e *Chromobacterium violaceum* ATCC 31532 foram posicionadas longe dos demais membros de suas espécies e agrupadas junto as estirpes de *Chromobacterium vaccinni* e *Chromobacterium subtsugae*, respectivamente. As posições inesperadas dessas estirpes indicam possíveis inconsistências nas classificações das mesmas.

A estirpe *P. ferrooxidans* EGD-HP2 foi sequenciada por Puranik *et al.* (2013). Nesse trabalho, os autores afirmaram que a estirpe possui, além de vias para o metabolismo de ferro e compostos aromáticos, uma via de produção de violaceína homóloga a de *Chromobacterium violaceum*. Entretanto, outras espécies de *Pseudogulbenkiania* apresentam colônias de coloração amarela ou branca (LIN *et al.*, 2008; LEE *et al.*, 2013) e, em adição, vias de produção de violaceína podem ser encontradas nas estirpes de *Chromobacterium vaccinii* (SOBY *et al.*, 2013). Além disso, os métodos utilizados para a classificação da estirpe no gênero *Pseudogulbenkiania* não foram citados. Assim, possivelmente esta estirpe foi mal classificada.

A inconsistência na classificação da estirpe *Chromobacterium violaceum* ATCC 31532, foi esclarecida por Harrison e Soby (2020), que publicou um trabalho reclassificando a estirpe como um membro da espécie *Chromobacterium subtsugae*.

6 CONCLUSÕES E TRABALHOS FUTUROS

Nesse trabalho, foram avaliadas as relações filogenéticas entre as espécies da família *Chromobacteriaceae* utilizando ferramentas e técnicas de bioinformática. Foram realizados a coleta de sequências genéticas disponíveis em bancos de dados biológicos, a seleção de diferentes grupos de genes para a reconstrução filogenética, o alinhamento dos grupos de genes e a análise das sequências alinhadas.

Na reconstrução filogenética, foram apresentadas duas árvores baseadas, respectivamente, em dois grupos de genes, um contendo 19 proteínas de manutenção (Adk, AtpG, EF-G, FusA, GltX, Gmk, GyrA, GyrB, L5, LipA, MurC, MurE, PgK, RpoB, S2, S3, SecY, ThrRS, UvrD) e outro contendo RNA ribossômico (16S e 23S).

As análises realizadas indicaram a existência de dois grande grupos monofiléticos compondo a família *Chromobacteriaceae*. A nível de gênero, o estudo se mostrou consistente com a atual classificação, os quais todos os táxons formam grupos monofiléticos. A nível de espécie, houve indicativo de uma possível inconsistência na classificação da estirpe *Pseudogulbenkiania ferrooxidans* EGD HP2.

Um próximo passo para outros estudos seria a identificação de CSIs específicos para os clados da família *Chromobacteriaceae* e uma revisão da taxonomia da estirpe *P. ferrooxidans* EGD-HP2.

REFERÊNCIAS

- ADEOLU, M.; ALNAJAR, S.; NAUSHAD, S.; GUPTA, R. S. Genome-based phylogeny and taxonomy of the ‘enterobacteriales’: proposal for enterobacterales ord. nov. divided into the families enterobacteriaceae, erwiniaceae fam. nov., pectobacteriaceae fam. nov., yersiniaceae fam. nov., hafniaceae fam. nov., morganellaceae fam. nov., and budviciaceae fam. nov. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 66, n. 12, p. 5575–5599, 2016.
- ADEOLU, M.; GUPTA, R. S. Phylogenomics and molecular signatures for the order neisseriales: proposal for division of the order neisseriales into the emended family neisseriaceae and chromobacteriaceae fam. nov. **Antonie van Leeuwenhoek**, Springer, v. 104, n. 1, p. 1–24, 2013.
- ALTSCHUL, S. **The statistics of sequence similarity scores**. 2011. Disponível em: <https://www.ncbi.nlm.nih.gov/BLAST/tutorial/Altschul-1.html>. Acessado em 20/02/2021.
- ARAÚJO, N. D. de; FARIAS, R. P. de; PEREIRA, P. B.; FIGUEIRÊDO, F. M. de; MORAIS, A. M. B. de; SALDANHA, L. C.; GABRIEL, J. E. A era da bioinformática: seu potencial e suas implicações para as ciências da saúde. **Estudos de biologia**, v. 30, n. 70/72, 2008.
- AZEVEDO, M. I.; OLIVEIRA, S. T.; SILVA, C. F.; CARNEIRO, R. F.; NAGANO, C. S.; GADELHA, A. C.; TORRES, D. C.; MONTEIRO-JÚNIOR, J. E.; GIRÃO, M. S.; MUNIZ, C. R. *et al.* Secretory production in escherichia coli of a gh46 chitosanase from chromobacterium violaceum, suitable to generate antifungal chitoooligosaccharides. **International Journal of Biological Macromolecules**, Elsevier, v. 165, p. 1482–1495, 2020.
- BAJAJ, A.; KUMAR, A.; YADAV, S.; KAUR, G.; BALA, M.; SINGH, N. K.; KUMAR, R. M.; MANICKAM, N.; MAYILRAJ, S. Isolation and characterization of a novel gram-negative bacterium chromobacterium alkanivorans sp. nov., strain iitr-71t degrading halogenated alkanes. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 66, n. 12, p. 5228–5235, 2016.
- BAUERMEISTER, A.; REZENDE, M. I.; GIESE, E. C.; DEKKER, R. F. H.; BARBOSA, A. de M. beta-(1, 3)-glucanases fúngicas: Produção e aplicações biotecnológicas. **Semina: Ciências Exatas e Tecnológicas**, v. 31, n. 2, p. 75–86, 2010.
- BERGONZINI, C. Sopra un nuovo bacterio colorato. **Annuar Soc. Nat. Modena**, v. 14, n. 2, p. 149–158, 1880.
- BLACKBURN, M. B.; JR, R. R. F.; SPARKS, M. E.; KUHAR, D.; MITCHELL, A.; GUNDERSEN-RINDAL, D. E. Chromobacterium sphagni sp. nov., an insecticidal bacterium isolated from sphagnum bogs. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 67, n. 9, p. 3417–3422, 2017.
- BLACKBURN, M. B.; JR, R. R. F.; SPARKS, M. E.; KUHAR, D.; MOWERY, J. D.; MITCHELL, A.; GUNDERSEN-RINDAL, D. E. Chromobacterium phragmitis sp. nov., isolated from estuarine marshes. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 69, n. 9, p. 2681–2686, 2019.
- BLACKBURN, M. B.; JR, R. R. F.; SPARKS, M. E.; KUHAR, D.; MOWERY, J. D.; MITCHELL, A.; GUNDERSEN-RINDAL, D. E. Chromobacterium paludis sp. nov., a novel bacterium isolated from a chesapeake bay marsh. **International Journal of Systematic and Evolutionary Microbiology**, Microbiology Society, v. 70, n. 12, p. 6142–6146, 2020.

- BUTLER, B. J.; MCCALLUM, K. L.; INNIS, W. E. Characterization of *aquaspirillum arcticum* sp. nov., a new psychrophilic bacterium. **Systematic and applied microbiology**, Elsevier, v. 12, n. 3, p. 263–266, 1989.
- CAMACHO, C.; COULOURIS, G.; AVAGYAN, V.; MA, N.; PAPADOPOULOS, J.; BEALER, K.; MADDEN, T. L. Blast+: architecture and applications. **BMC bioinformatics**, Springer, v. 10, n. 1, p. 1–9, 2009.
- CATTLE, S.; ARTHUR, J. W. Biomanager: the use of a bioinformatics web application as a teaching tool in undergraduate bioinformatics training. **Briefings in bioinformatics**, Oxford University Press, v. 8, n. 6, p. 457–465, 2007.
- CHANG, S.-C.; CHEN, W.-M.; WANG, J.-T.; WU, M.-C. *Chitinilyticum aquatile* gen. nov., sp. nov., a chitinolytic bacterium isolated from a freshwater pond used for pacific white shrimp culture. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 57, n. 12, p. 2854–2860, 2007.
- CHANG, S.-C.; WU, M.-C.; CHEN, W.-M.; TSAI, Y.-H.; LEE, T.-M. *Chitinilyticum litopenaei* sp. nov., isolated from a freshwater shrimp pond, and emended description of the genus *chitinilyticum*. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 59, n. 11, p. 2651–2655, 2009.
- CHEN, W.-M.; CHANG, R.-C.; CHENG, C.-Y.; SHIAU, Y.-W.; SHEU, S.-Y. *Jeongeupia chitinilytica* sp. nov., a chitinolytic bacterium isolated from soil. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 63, n. 3, p. 934–938, 2013.
- CHEN, W.-M.; CHEN, J.-C.; WANG, C.; HUANG, C.-W.; SHEU, S.-Y. *Vogesella amnigena* sp. nov., isolated from a freshwater river. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 65, n. 10, p. 3634–3640, 2015.
- CHEN, W.-M.; CHUNG, Y.-N.; CHIU, T.-F.; CHENG, C.-Y.; ARUN, A.; SHEU, S.-Y. *Deefgea chitinilytica* sp. nov., isolated from a wetland. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 60, n. 6, p. 1450–1453, 2010.
- CHERN, L.-L.; STACKEBRANDT, E.; LEE, S.-F.; LEE, F.-L.; CHEN, J.-K.; FU, H.-M. *Chitinibacter tainanensis* gen. nov., sp. nov., a chitin-degrading aerobe from soil in taiwan. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 54, n. 4, p. 1387–1391, 2004.
- CHOU, J.-H.; CHOU, Y.-J.; ARUN, A.; YOUNG, C.-C.; CHEN, C. A.; WANG, J.-T.; CHEN, W.-M. *Vogesella lacus* sp. nov., isolated from a soft-shell turtle culture pond. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 59, n. 10, p. 2629–2632, 2009.
- CHOU, Y.-J.; CHOU, J.-H.; LIN, M.-C.; ARUN, A.; YOUNG, C.-C.; CHEN, W.-M. *Vogesella perlucida* sp. nov., a non-pigmented bacterium isolated from spring water. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 58, n. 12, p. 2677–2681, 2008.
- CHOUDHURI, S. **Bioinformatics for beginners: genes, genomes, molecular evolution, databases and analytical tools**. [S. l.]: Elsevier, 2014.

CHRISTENSEN, H.; KUHNERT, P.; OLSEN, J. E.; BISGAARD, M. Comparative phylogenies of the housekeeping genes *atpD*, *infB* and *rpoB* and the 16s rna gene within the pasteurellaceae. **International Journal of Systematic and Evolutionary Microbiology**, Microbiology Society, v. 54, n. 5, p. 1601–1609, 2004.

COCK, P. J.; ANTAO, T.; CHANG, J. T.; CHAPMAN, B. A.; COX, C. J.; DALKE, A.; FRIEDBERG, I.; HAMELRYCK, T.; KAUFF, F.; WILCZYNSKI, B. *et al.* Biopython: freely available python tools for computational molecular biology and bioinformatics. **Bioinformatics**, Oxford University Press, v. 25, n. 11, p. 1422–1423, 2009.

DARRIBA, D.; TABOADA, G. L.; DOALLO, R.; POSADA, D. Protest 3: fast selection of best-fit models of protein evolution. **Bioinformatics**, Oxford University Press, v. 27, n. 8, p. 1164–1165, 2011.

DARRIBA, D.; TABOADA, G. L.; DOALLO, R.; POSADA, D. jmodeltest 2: more models, new heuristics and parallel computing. **Nature methods**, Nature Publishing Group, v. 9, n. 8, p. 772–772, 2012.

DESALLE, R.; ABSHER, R.; AMATO, G. Speciation and phylogenetic resolution. **Trends in ecology & evolution**, v. 9, n. 8, p. 297–298, 1994.

DURÁN, M.; PONEZI, A. N.; FALJONI-ALARIO, A.; TEIXEIRA, M. F.; JUSTO, G. Z.; DURÁN, N. Potential applications of violacein: a microbial pigment. **Medicinal Chemistry Research**, Springer, v. 21, n. 7, p. 1524–1532, 2012.

DURÁN, N.; JUSTO, G. Z.; FERREIRA, C. V.; MELO, P. S.; CORDI, L.; MARTINS, D. Violacein: properties and biological activities. **Biotechnology and applied biochemistry**, Wiley Online Library, v. 48, n. 3, p. 127–133, 2007.

ENGSBRO, A. L.; NIELSEN, K. L.; HORNUM, M.; ANDERSEN, L. P. *Laribacter hongkongensis*: clinical presentation, epidemiology and treatment. a review of the literature and report of the first case in denmark. **Infectious Diseases**, Taylor & Francis, v. 50, n. 6, p. 417–422, 2018.

FELDMAN, R. B.; STERN, G. A.; HOOD, C. I. *Chromobacterium violaceum* infection of the eye: a report of two cases. **Archives of Ophthalmology**, American Medical Association, v. 102, n. 5, p. 711–713, 1984.

FERNANDES, M. J. d. B. C.; LUZ, K. G.; DANTAS, L. d. A.; MELO, M. C. N. d.; ALMEIDA, D. *Chromobacterium violaceum*: a fatal case in the northeast of the brazil. **Jornal Brasileiro de Patologia e Medicina Laboratorial**, SciELO Brasil, v. 50, n. 4, p. 278–279, 2014.

FISHER, R. A. On the mathematical foundations of theoretical statistics. **Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character**, The Royal Society London, v. 222, n. 594–604, p. 309–368, 1922.

FITCH, W. M. Toward defining the course of evolution: minimum change for a specific tree topology. **Systematic Biology**, Society of Systematic Zoology, v. 20, n. 4, p. 406–416, 1971.

GLAESER, S. P.; KÄMPFER, P. Multilocus sequence analysis (mlsa) in prokaryotic taxonomy. **Systematic and applied microbiology**, Elsevier, v. 38, n. 4, p. 237–245, 2015.

GRIMES, D. J.; WOESE, C. R.; MACDONELL, M. T.; COLWELL, R. R. Systematic study of the genus *Vogesella* gen. nov. and its type species, *Vogesella indigofera* comb. nov. **International Journal of Systematic and Evolutionary Microbiology**, Microbiology Society, v. 47, n. 1, p. 19–27, 1997.

HAN, X. Y.; HAN, F. S.; SEGAL, J. *Chromobacterium haemolyticum* sp. nov., a strongly haemolytic species. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 58, n. 6, p. 1398–1403, 2008.

HARRISON, A. M.; SOBY, S. D. Reclassification of *Chromobacterium violaceum* ATCC 31532 and its quorum biosensor mutant cv026 to *Chromobacterium subtsugae*. **AMB Express**, SpringerOpen, v. 10, n. 1, p. 1–7, 2020.

HEDLUND, B. P.; STALEY, J. T. Phylogeny of the genus *Simonsiella* and other members of the Neisseriaceae. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 52, n. 4, p. 1377–1382, 2002.

HYLEMON, P.; JR, J. W.; KRIEG, N.; JANNASCH, H. The genus *Spirillum*: a taxonomic study 1. **International Journal of Systematic and Evolutionary Microbiology**, Microbiology Society, v. 23, n. 4, p. 340–380, 1973.

JONES, D. T.; TAYLOR, W. R.; THORNTON, J. M. The rapid generation of mutation data matrices from protein sequences. **Bioinformatics**, Oxford University Press, v. 8, n. 3, p. 275–282, 1992.

JØRGENSEN, N. O.; BRANDT, K. K.; NYBROE, O.; HANSEN, M. *Vogesella mureinivorans* sp. nov., a peptidoglycan-degrading bacterium from lake water. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 60, n. 10, p. 2467–2472, 2010.

JYOTI, V.; NARAYAN, K. D.; DAS, S. K. *Gulbenkiania indica* sp. nov., isolated from a sulfur spring. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 60, n. 5, p. 1052–1055, 2010.

KÄMPFER, P.; BUSSE, H.-J.; SCHOLZ, H. C. *Chromobacterium piscinae* sp. nov. and *Chromobacterium pseudoviolaceum* sp. nov., from environmental samples. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 59, n. 10, p. 2486–2490, 2009.

KANAMORI, H.; AOYAGI, T.; KURODA, M.; SEKIZUKA, T.; KATSUMI, M.; ISHIKAWA, K.; HOSAKA, T.; BABA, H.; OSHIMA, K.; TOKUDA, K. *et al.* *Chromobacterium haemolyticum* pneumonia associated with near-drowning and river water, Japan. **Emerging infectious diseases**, Centers for Disease Control and Prevention, v. 26, n. 9, p. 2186, 2020.

KANG, H.; KIM, H.; JOUNG, Y.; KIM, K.-J.; JOH, K. *Paludibacterium purpuratum* sp. nov., isolated from wetland soil. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 66, n. 7, p. 2711–2716, 2016.

KARLIN, S.; ALTSCHUL, S. F. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 87, n. 6, p. 2264–2268, 1990.

- KATOH, K.; STANDLEY, D. M. Mafft multiple sequence alignment software version 7: improvements in performance and usability. **Molecular biology and evolution**, Society for Molecular Biology and Evolution, v. 30, n. 4, p. 772–780, 2013.
- KONZEN, M.; MARCO, D. D.; CORDOVA, C. A.; VIEIRA, T. O.; ANTONIO, R. V.; CRECZYNSKI-PASA, T. B. Antioxidant properties of violacein: possible relation on its biological function. **Bioorganic & medicinal chemistry**, Elsevier, v. 14, n. 24, p. 8307–8313, 2006.
- KOZLOV, A. M.; DARRIBA, D.; FLOURI, T.; MOREL, B.; STAMATAKIS, A. Raxml-ng: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. **Bioinformatics**, Oxford University Press, v. 35, n. 21, p. 4453–4455, 2019.
- KUMAR, R.; BANERJEE, A.; BOWDRE, J.; MCELROY, L.; KRIEG, N. Isolation, characterization, and taxonomy of aquaspirillum bengal sp. nov. **International Journal of Systematic and Evolutionary Microbiology**, Microbiology Society, v. 24, n. 4, p. 453–458, 1974.
- KUMAR, S.; STECHER, G.; LI, M.; KNYAZ, C.; TAMURA, K. Mega x: molecular evolutionary genetics analysis across computing platforms. **Molecular biology and evolution**, Oxford University Press, v. 35, n. 6, p. 1547–1549, 2018.
- KWON, S.-W.; KIM, B.-Y.; KIM, W.-G.; YOO, K.-H.; YOO, S.-H.; SON, J.-A.; WEON, H.-Y. Paludibacterium yongneupense gen. nov., sp. nov., isolated from a wetland, yongneup, in korea. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 58, n. 1, p. 190–194, 2008.
- LAN, K.; CAI, Y.-M.; LI, L.-H.; ZENG, J.-M.; YU, X.-G.; QU, P.-H.; LI, H.-L.; LIU, Y.-Y.; CHEN, L.; CHEN, C. *et al.* Vogesella urethralis sp. nov., isolated from human urine, and emended descriptions of vogesella perlucida and vogesella mureinivorans. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 70, n. 1, p. 624–630, 2020.
- LANFEAR, R.; CALCOTT, B.; HO, S. Y.; GUINDON, S. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. **Molecular biology and evolution**, Oxford University Press, v. 29, n. 6, p. 1695–1701, 2012.
- LAU, H.-T.; FARYNA, J.; TRIPLETT, E. W. Aquitalea magnusonii gen. nov., sp. nov., a novel gram-negative bacterium isolated from a humic lake. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 56, n. 4, p. 867–871, 2006.
- LEE, C.-M.; WEON, H.-Y.; KIM, Y.-J.; SON, J.-A.; YOON, S.-H.; KOO, B.-S.; KWON, S.-W. Aquitalea denitrificans sp. nov., isolated from a korean wetland. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 59, n. 5, p. 1045–1048, 2009.
- LEE, D.-G.; IM, D.-M.; KANG, H.; YUN, P.; PARK, S.-K.; HYUN, S.-S.; HWANG, D.-Y. Pseudogulbenkiania gefcensis sp. nov., isolated from soil. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 63, n. 1, p. 187–191, 2013.
- LEE, J.-Y.; KANG, W.; KIM, P. S.; LEE, S.-Y.; SHIN, N.-R.; SUNG, H.; LEE, J.-Y.; YUN, J.-H.; JEONG, Y.-S.; HAN, J. E.; JUNG, M.-J.; HYUN, D.-W.; KIM, H. S.; TAK, E. J.; KANG,

- M.-S.; LEE, K.-E.; LEE, B.-H.; BAE, J.-W. *Iodobacter ciconiae* sp. nov., a bacterium isolated from feces of oriental stork, *ciconia boyciana*. **International Journal of Systematic and Evolutionary Microbiology**, Microbiology Society, v. 69, n. 9, p. 2948–2953, 2019.
- LETUNIC, I.; BORK, P. Interactive tree of life (itol) v4: recent updates and new developments. **Nucleic acids research**, Oxford University Press, v. 47, n. W1, p. W256–W259, 2019.
- LIM, J.-M.; JEON, C. O.; LEE, G. S.; PARK, D.-J.; KANG, U.-G.; PARK, C.-Y.; KIM, C.-J. *Leeia oryzae* gen. nov., sp. nov., isolated from a rice field in korea. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 57, n. 6, p. 1204–1208, 2007.
- LIN, M.-C.; CHOU, J.-H.; ARUN, A.; YOUNG, C.-C.; CHEN, W.-M. *Pseudogulbenkiania subflava* gen. nov., sp. nov., isolated from a cold spring. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 58, n. 10, p. 2384–2388, 2008.
- LIU, Y.; LAI, Q.; SHAO, Z. A multilocus sequence analysis scheme for phylogeny of thioclava bacteria and proposal of two novel species. **Frontiers in microbiology**, Frontiers, v. 8, p. 1321, 2017.
- LOGAN, N. A. Numerical taxonomy of violet-pigmented, gram-negative bacteria and description of *iodobacter fluviatile* gen. nov., comb. nov. **International Journal of Systematic and Evolutionary Microbiology**, Microbiology Society, v. 39, n. 4, p. 450–456, 1989.
- LOPES, S. G. B. C.; CHOW, F.; LAHR, D. J. G.; TURRINI, P. Diversidade biológica e filogenia. In: _____. [S. l.]: USP/Univesp/Edusp, 2014.
- MACHER, A. M.; CASALE, T. B.; FAUCI, A. S. Chronic granulomatous disease of childhood and *chromobacterium violaceum* infections in the southeastern united states. **Annals of internal medicine**, American College of Physicians, v. 97, n. 1, p. 51–54, 1982.
- MADIGAN, M. T.; MARTINKO, J. M.; BENDER, K. S.; BUCKLEY, D. H.; STAHL, D. A. **Microbiologia de Brock-14ª Edição**. [S. l.]: Artmed Editora, 2016.
- MARTIN, P. A.; GUNDERSEN-RINDAL, D.; BLACKBURN, M.; BUYER, J. *Chromobacterium subtsugae* sp. nov., a betaproteobacterium toxic to colorado potato beetle and other insect pests. **International Journal of Systematic and Evolutionary Microbiology**, Microbiology Society, v. 57, n. 5, p. 993–999, 2007.
- MENEZES, C. B. A.; TONIN, M. F.; CORRÊA, D. B. A.; PARMA, M.; MELO, I. S. de; ZUCCHI, T. D.; DESTÉFANO, S. A. L.; FANTINATTI-GARBOGGINI, F. *Chromobacterium amazonense* sp. nov. isolated from water samples from the rio negro, amazon, brazil. **Antonie van Leeuwenhoek**, Springer, v. 107, n. 4, p. 1057–1063, 2015.
- MILLER, M. A.; PFEIFFER, W.; SCHWARTZ, T. The cipres science gateway: a community resource for phylogenetic analyses. In: **Proceedings of the 2011 TeraGrid Conference: extreme digital discovery**. [S. l.: s. n.], 2011. p. 1–8.
- MOSS, M. O.; RYALL, C.; LOGAN, N. A. The classification and characterization of *chromobacteria* from a lowland river. **Microbiology**, Microbiology Society, v. 105, n. 1, p. 11–21, 1978.

MOYA, G.; TRINH, H.; YAN, Z.-F.; WON, K.-H.; YANG, J.-E.; WANG, Q.-J.; KOOK, M.-C.; YI, T.-H. *Aquaspirillum soli* sp. nov., isolated from a soil sample. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 67, n. 5, p. 1312–1317, 2017.

MULLER, O. F. **Animalcula infusoria fluviatilia et marina, quae detexit, systematice descripsit et ad vivum delineari curavit Otho Fridericus Muller... sistit opus hoc posthumum quod... in lucem tradit vidua ejus nobilissima, cura Othonis Fabricii.** [S. l.: s. n.], 1786.

MURAMATSU, Y.; SUZUKI, K.-i.; NAKAGAWA, Y. *Silvimonas iriomotensis* sp. nov. and *silvimonas amylolytica* sp. nov., new members of the class betaproteobacteria isolated from the subtropical zone in japan. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 60, n. 1, p. 174–178, 2010.

NCBI C Toolkit Cross Reference. 2021. Disponível em: https://www.ncbi.nlm.nih.gov/IEB/ToolBox/C_DOC/lxr/source/data. Acessado em 03/25/2021.

NELSON, D. L.; COX, M. M. **Princípios de Bioquímica de Lehninger-7.** [S. l.]: Artmed Editora, 2018.

NGO, H. T.; KIM, H.; TRINH, H.; YI, T.-H. *Aquitalea aquatilis* sp. nov., isolated from jungwon waterfall. **International Journal of Systematic and Evolutionary Microbiology**, Microbiology Society, v. 70, n. 9, p. 4903–4907, 2020.

OKADA, M.; INOKUCHI, R.; SHINOHARA, K.; MATSUMOTO, A.; ONO, Y.; NARITA, M.; ISHIDA, T.; KAZUKI, C.; NAKAJIMA, S.; YAHAGI, N. *Chromobacterium haemolyticum*-induced bacteremia in a healthy young man. **BMC infectious diseases**, BioMed Central, v. 13, n. 1, p. 1–4, 2013.

O'LEARY, N. A.; WRIGHT, M. W.; BRISTER, J. R.; CIUFO, S.; HADDAD, D.; MCVEIGH, R.; RAJPUT, B.; ROBBERTSE, B.; SMITH-WHITE, B.; AKO-ADJEI, D. *et al.* Reference sequence (refseq) database at ncbi: current status, taxonomic expansion, and functional annotation. **Nucleic acids research**, Oxford University Press, v. 44, n. D1, p. D733–D745, 2016.

PATUREAU, D.; GODON, J.-J.; DABERT, P.; BOUCHEZ, T.; BERNET, N.; DELGENES, J. P.; MOLETTA, R. *Microvirgula aerodenitrificans* gen. nov., sp. nov., a new gram-negative bacterium exhibiting co-respiration of oxygen and nitrogen oxides up to oxygen-saturated conditions. **International Journal of Systematic and Evolutionary Microbiology**, Microbiology Society, v. 48, n. 3, p. 775–782, 1998.

PROSDOCIMI, F. Introdução à bioinformática. **Instituto de Química**, 2007.

PURANIK, S.; TALKAL, R.; QURESHI, A.; KHARDENAVIS, A.; KAPLEY, A.; PUROHIT, H. J. Genome sequence of the pigment-producing bacterium *pseudogulbenkiania ferrooxidans*, isolated from loktak lake. **Genome announcements**, Am Soc Microbiol, v. 1, n. 6, 2013.

RAJA, M. K.; GHOSH, A. R. *Laribacter hongkongensis*: an emerging pathogen of infectious diarrhea. **Folia microbiologica**, Springer, v. 59, n. 4, p. 341–347, 2014.

RAMESHKUMAR, N.; LANG, E.; TANAKA, N. Description of *vogesella oryzae* sp. nov., isolated from the rhizosphere of saline tolerant pokkali rice. **Systematic and applied microbiology**, Elsevier, v. 39, n. 1, p. 20–24, 2016.

RATHORE, A. S.; GUPTA, R. D. Chitinases from bacteria to human: properties, applications, and future perspectives. **Enzyme research**, Hindawi, v. 2015, 2015.

SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular biology and evolution**, v. 4, n. 4, p. 406–425, 1987.

SATO, K.; KATO, Y.; TAGUCHI, G.; NOGAWA, M.; YOKOTA, A.; SHIMOSAKA, M. *Chitiniphilus shinanonensis* gen. nov., sp. nov., a novel chitin-degrading bacterium belonging to betaproteobacteria. **The Journal of general and applied microbiology**, Applied Microbiology, Molecular and Cellular Biosciences Research Foundation, v. 55, n. 2, p. 147–153, 2009.

SEDLÁČEK, I.; KWON, S.-W.; ŠVEC, P.; MAŠLANOVÁ, I.; KÝROVÁ, K.; HOLOCHOVÁ, P.; ČERNOHLÁVKOVÁ, J.; BUSSE, H.-J. *Aquitalea pelogenes* sp. nov., isolated from mineral peloid. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 66, n. 2, p. 962–967, 2016.

SHENG, M.; XU, J.; YANG, Z.; ZHU, J.; QIAN, M.; HE, Q.; HE, J. *Chitiniphilus eburneus* sp. nov., a novel chitinolytic bacterium isolated from sludge. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 70, n. 1, p. 352–357, 2020.

SHEU, S.-Y.; CHEN, J.-C.; YOUNG, C.-C.; CHEN, W.-M. *Vogesellafluminis* sp. nov., isolated from a freshwater river, and emended description of the genus *vogesella*. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 63, n. 8, p. 3043–3049, 2013.

SHEU, S.-Y.; CHEN, Y.-L.; YOUNG, C.-C.; CHEN, W.-M. *Vogesella facilis* sp. nov., isolated from a freshwater river, and emended description of the genus *vogesella*. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 66, n. 2, p. 817–823, 2016.

SHEU, S.-Y.; CHEN, Z.-H.; YOUNG, C.-C.; CHEN, W.-M. *Paludibacterium paludis* sp. nov., isolated from a marsh. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 64, n. 7, p. 2497–2502, 2014.

SHEU, S.-Y.; CHIU, T. F.; CHOU, J.-H.; SHEU, D.-S.; ARUN, A.; YOUNG, C.-C.; CHEN, C. A.; WANG, J.-T.; CHEN, W.-M. *Andreprevotia lacus* sp. nov., isolated from a fish-culture pond. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 59, n. 10, p. 2482–2485, 2009.

SHEU, S.-Y.; LI, Y.-S.; YOUNG, C.-C.; CHEN, W.-M. *Chitinibacter fontanus* sp. nov., isolated from a spring. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 66, n. 10, p. 4262–4268, 2016.

SKERMAN, V. B. D.; MCGOWAN, V.; SNEATH, P. H. A. Approved lists of bacterial names. **International Journal of Systematic and Evolutionary Microbiology**, Microbiology Society, v. 30, n. 1, p. 225–420, 1980.

SNEATH, P. H.; SOKAL, R. R. *et al.* **Numerical taxonomy. The principles and practice of numerical classification.** [S. l.: s. n.], 1973.

SOBY, S. D.; GADAGKAR, S. R.; CONTRERAS, C.; CARUSO, F. L. *Chromobacterium vaccinii* sp. nov., isolated from native and cultivated cranberry (*vaccinium macrocarpon* ait.) bogs

- and irrigation ponds. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 63, n. 5, p. 1840–1846, 2013.
- SONG, J.; JOUNG, Y.; LI, S.-H.; HWANG, J.; CHO, J.-C. *Leeia aquatica* sp. nov., isolated from freshwater. **International Journal of Systematic and Evolutionary Microbiology**, Microbiology Society, v. 70, n. 11, p. 5848–5853, 2020.
- SOUSA, A. J.; SILVA, C. F.; SOUSA, J. S.; JÚNIOR, J. E. M.; FREIRE, J. E.; SOUSA, B. L.; LOBO, M. D.; MONTEIRO-MOREIRA, A. C.; GRANGEIRO, T. B. A thermostable chitinase from the antagonistic chromobacterium violaceum that inhibits the development of phytopathogenic fungi. **Enzyme and microbial technology**, Elsevier, v. 126, p. 50–61, 2019.
- SRINIVAS, T.; MANASA, P.; BEGUM, Z.; SUNIL, B.; SAILAJA, B.; SINGH, S.; PRASAD, S.; SHIVAJI, S. *Iodobacter arcticus* sp. nov., a psychrotolerant bacterium isolated from meltwater stream sediment of an arctic glacier. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 63, n. 8, p. 2800–2805, 2013.
- STACKEBRANDT, E.; LANG, E.; COUSIN, S.; PÄUKER, O.; BRAMBILLA, E.; KROPPESTEDT, R.; LÜNSDORF, H. *Deefgea rivuli* gen. nov., sp. nov., a member of the class betaproteobacteria. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 57, n. 3, p. 639–645, 2007.
- STEVENS, T. J.; BOUCHER, W. **Python programming for biology**. [S. l.]: Cambridge University Press, 2015.
- STIGLER, S. M. Thomas bayes's bayesian inference. **Journal of the Royal Statistical Society: Series A (General)**, Wiley Online Library, v. 145, n. 2, p. 250–258, 1982.
- SU, W.; ZHOU, Z.; JIANG, F.; CHANG, X.; LIU, Y.; WANG, S.; KAN, W.; XIAO, M.; SHAO, M.; PENG, F.; FANG, C. *Iodobacter limnosediminis* sp. nov., isolated from arctic lake sediment. **International Journal of Systematic and Evolutionary Microbiology**, Microbiology Society, v. 63, n. Pt4, p. 1464 – 1470, 2013.
- SUBHASH, Y.; PARK, M.-J.; LEE, S.-S. *Microvirgula curvata* sp. nov., isolated from hydrocarbon-contaminated soil, and emended description of the genus microvirgula. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 66, n. 12, p. 5309–5313, 2016.
- SUBHASH, Y.; TUSHAR, L.; SASIKALA, C.; RAMANA, C. V. *Vogesella alkaliphila* sp. nov., isolated from an alkaline soil, and emended description of the genus vogesella. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 63, n. 6, p. 2338–2343, 2013.
- SUN, Y.-F.; COSTA-REZENDE, D.; XING, J.-H.; ZHOU, J.-L.; ZHANG, B.; GIBERTONI, T.; GATES, G.; GLEN, M.; DAI, Y.-C.; CUI, B.-K. Multi-gene phylogeny and taxonomy of amauroderma s. lat.(ganodermataceae). **Persoonia: Molecular Phylogeny and Evolution of Fungi**, Westerdijk Fungal Biodiversity Institute, v. 44, p. 206, 2020.
- SWAIN, B.; OTTA, S.; SAHU, K. K.; PANDA, K.; ROUT, S. Urinary tract infection by chromobacterium violaceum. **Journal of clinical and diagnostic research: JCDR**, JCDR Research & Publications Private Limited, v. 8, n. 8, p. DD01, 2014.

- TANAKA, K.; NAKAMURA, K.; MIKAMI, E. Fermentation of s-citramalate, citrate, mesaconate, and pyruvate by a gram-negative strictly anaerobic non-spore-former, formivibrio citricus gen. nov., sp. nov. **Archives of microbiology**, Springer, v. 155, n. 5, p. 491–495, 1991.
- TANPOWONG, P.; CHAROENMUANG, R.; APIWATTANAKUL, N. First pediatric case of chromobacterium haemolyticum causing proctocolitis. **Pediatrics International**, Wiley Online Library, v. 56, n. 4, p. 615–617, 2014.
- TERASAKI, Y. On spirillum putridiconchylium nov. sp. **Bot. Mag.(Tokyo)**, v. 74, p. 79–85, 1961.
- TUCKER, R. E.; JR, W. W.; WILSON, H. Osteomyelitis associated with chromobacterium violaceum sepsis. a case report. **JBJJ**, LWW, v. 61, n. 6, p. 949–951, 1979.
- VAINSHTEIN, Y.; WERNER, N.; KIRSTAHLER, P.; GLANZ, K.; GRUMAZ, C.; HAHN, T.; ZIBEK, S.; SOHN, K. Draft genome sequence of andreprevotia sp. strain igb-42, a chitinolytic bacterium isolated from a soil sample of an anthill in stuttgart, germany. **Microbiology resource announcements**, Am Soc Microbiol, v. 9, n. 10, 2020.
- VASCONCELOS, A. T. R. de; ALMEIDA, D. F. D.; HUNGRIA, M.; GUIMARAES, C. T.; ANTÔNIO, R. V.; ALMEIDA, F. C.; ALMEIDA, L. G. D.; ALMEIDA, R. D.; ALVES-GOMES, J. A.; ANDRADE, E. M. *et al.* The complete genome sequence of chromobacterium violaceum reveals remarkable and exploitable bacterial adaptability. **Proceedings of the national academy of sciences of the United States of America**, JSTOR, p. 11660–11665, 2003.
- VAZ-MOREIRA, I.; NOBRE, M. F.; NUNES, O. C.; MANAIA, C. M. Gulbenkiania mobilis gen. nov., sp. nov., isolated from treated municipal wastewater. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 57, n. 5, p. 1108–1112, 2007.
- VERLI, H. Bioinformática: da biologia à flexibilidade molecular. Sociedade Brasileira de Bioquímica e Biologia Molecular, 2014.
- VOGES, O. Ueber einige im wasser vorkommende pigmentbakterien. **Zentralbl. Bakteriol. Parasitenk. Infektionskr. Hyg. Abt. I**, v. 14, p. 301–314, 1893.
- WÄGELE, J. Foundations of phylogenetic systematics. münchen: Dr. F. Pfei, 2005.
- WEON, H.-Y.; KIM, B.-Y.; YOO, S.-H.; JOA, J.-H.; KWON, S.-W.; KIM, W.-G. Andreprevotia chitinilytica gen. nov., sp. nov., isolated from forest soil from halla mountain, jeju island, korea. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 57, n. 7, p. 1572–1575, 2007.
- WHELAN, S.; GOLDMAN, N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. **Molecular biology and evolution**, Oxford University Press, v. 18, n. 5, p. 691–699, 2001.
- WILLIAMS, M. A.; RITTENBERG, S. A taxonomic study of the genus spirillum ehrenberg. **International Journal of Systematic and Evolutionary Microbiology**, Microbiology Society, v. 7, n. 2, p. 49–112, 1957.
- WU, M.; EISEN, J. A. A simple, fast, and accurate method of phylogenomic inference. **Genome biology**, Springer, v. 9, n. 10, p. 1–11, 2008.

- YANG, C.-H.; LI, Y.-H. *Chromobacterium violaceum* infection: a clinical review of an important but neglected infection. **Journal of the Chinese Medical Association**, Elsevier, v. 74, n. 10, p. 435–441, 2011.
- YANG, C.-S.; SHEU, S.-Y.; YOUNG, C.-C.; ARUN, A.; CHENG, C.-Y.; CHEN, W.-M. *Chitinibacter alvei* sp. nov., isolated from stream water. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 60, n. 8, p. 1760–1764, 2010.
- YANG, H.-C.; IM, W.-T.; AN, D.-S.; PARK, W.-s.; KIM, I. S.; LEE, S.-T. *Silvimonas terrae* gen. nov., sp. nov., a novel chitin-degrading facultative anaerobe belonging to the ‘betaproteobacteria’. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 55, n. 6, p. 2329–2332, 2005.
- YOON, J.-H.; CHOI, J. H.; KANG, S.-J.; CHOI, N.-S.; LEE, J.-S.; SONG, J. J. *Jeongeupia naejangsanensis* gen. nov., sp. nov., a cellulose-degrading bacterium isolated from forest soil from naejang mountain in korea. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 60, n. 3, p. 615–619, 2010.
- YOUNG, C.-C.; ARUN, A.; LAI, W.-A.; CHEN, W.-M.; CHAO, J.-H.; SHEN, F.-T.; REKHA, P.; KÄMPFER, P. *Chromobacterium aquaticum* sp. nov., isolated from spring water samples. **International Journal of Systematic and Evolutionary Microbiology**, Microbiology Society, v. 58, n. 4, p. 877–880, 2008.
- YU, Z.; ZHU, F.; TAO, X.; ZHANG, L.; WU, S.; DONG, C.; DONG, Y.; ZHOU, X.; FANG, Y.; XU, K. *et al.* *Vogesella perlucida*-induced bacteremia in an advanced-age patient: first case report. **BMC infectious diseases**, Springer, v. 20, n. 1, p. 1–5, 2020.
- YUEN, K.-Y.; WOO, P. C.; TENG, J. L.; LEUNG, K.-W.; WONG, M. K.; LAU, S. K. *Laribacter hongkongensis* gen. nov., sp. nov., a novel gram-negative bacterium isolated from a cirrhotic patient with bacteremia and empyema. **Journal of clinical microbiology**, Am Soc Microbiol, v. 39, n. 12, p. 4227–4232, 2001.
- ZHOU, S.; GUO, X.; WANG, H.; KONG, D.; WANG, Y.; ZHU, J.; DONG, W.; HE, M.; HU, G.; ZHAO, B. *et al.* *Chromobacterium rhizoryzae* sp. nov., isolated from rice roots. **International journal of systematic and evolutionary microbiology**, Microbiology Society, v. 66, n. 10, p. 3890–3896, 2016.