

**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

**Investigação do potencial genômico funcional da cianobactéria
Capilliphycus salinus ALCB114379 e sua diversidade microbiana**

Gabriel Schimmelpfeng Passos

Dissertação apresentada para obtenção do título de
Mestre em Ciências. Área de concentração:
Microbiologia Agrícola

**Piracicaba
2023**

Gabriel Schimmelpfeng Passos
Bacharel em Engenharia de Bioprocessos e Biotecnologia

Investigação do potencial genômico funcional da cianobactéria *Capilliphycus salinus* ALCB114379 e sua diversidade microbiana

Orientadora:
Profa. Dra. **MARLI DE FÁTIMA FIORE**

Dissertação apresentada para obtenção do título de
Mestre em Ciências. Área de concentração:
Microbiologia Agrícola

Piracicaba
2023

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Passos, Gabriel Schimmelpfeng

Investigação do potencial genômico funcional da cianobactéria *Capilliphycus salinus* ALCB114379 e sua diversidade microbiana / Gabriel Schimmelpfeng Passos. - - Piracicaba, 2023.

91 p.

Dissertação (Mestrado) - - USP / Escola Superior de Agricultura “Luiz de Queiroz”.

1. Culturas não axênicas 2. Fixação de nitrogênio 3. Metabolismo secundário 4. Microbioma associado 5. Mineração genômica I. Título

AGRADECIMENTOS

À Profa. Dra. Marli de Fátima Fiore por me receber de braços abertos no Laboratório de Ecologia Molecular de Cianobactérias, oportunizando que eu explorasse esses organismos tão excepcionais.

À Profa. Dra. Simone Possedente de Lira pela excelente coordenação baseada em uma gestão humana e afetiva no programa de Pós-Graduação em Microbiologia Agrícola.

Aos professores Fernando Dini Andreote, Elliot Watanabe Kitajima, Claudia Barros Monteiro Vitorello, Diego Mauricio Riaño-Pachón, Ernani Pinto Junior e Marie Anne van Sluys pelas discussões pós-aula que certamente contribuíram para o meu projeto.

Às técnicas do Laboratório de Ecologia Molecular de Cianobactérias Ana Paula Dini Andreote e Renata Beatriz Cruz pela paciência no meu processo de aprendizado e adaptação e por todo apoio e confiança no meu trabalho. Obrigado por sempre estimularem o meu crescimento profissional e claro, pelas conversas de assuntos aleatórios que alegam o meu dia.

Aos meus colegas de laboratório Anderson Miguel Teixeira Feitosa, Maurício Junior Machado e Rafael Barty Dextro. Cada um de vocês não só teve uma participação expressiva no meu projeto, mas, o mais importante, tornaram a minha experiência acadêmica magnífica. Obrigado pelos puxões de orelha, pelas trocas construtivas, pelos ensinamentos e pelas risadas. Vocês fizeram os meus dias coloridos mesmo em tempos nublados.

Aos pesquisadores e técnicos do Laboratório de Toxicologia Ambiental pelas colaborações, em especial Núbia Pereira da Silva, pela parceria profissional e pela amizade que surgiu logo em seguida.

Aos pós-doutorandos, Thierry Alexandre Pellegrinetti e Simone Cotta, por toda ajuda recebida.

À Nathália Cassia Ferreira Dias por todo apoio durante o meu processo de adaptação, pelas saídas em Piracicaba, pelas conversas em que algumas terminaram em risadas e outras em lágrimas e, principalmente, pela amizade de nove anos. Que venham muitos outros.

Aos amigos que fiz em Piracicaba Matheus Ribeiro Barbosa Oliveira, Izabelly Pacifico e Jéssica Chaves Fadul.

Ao serviço de Pós-Graduação, em especial Maria Solizete Granzioi Silva, por toda assistência envolvendo prazos e questões burocráticas do programa.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa fornecida.

Por fim, à minha família e aos meus amigos de Minas Gerais que sempre estiveram ao meu lado e acreditaram em mim.

**“Aqueles que passam por
nós não vão sós. Deixam um pouco de si, levam um pouco de nós”**

Antoine de Saint-Exupéry

SUMÁRIO

RESUMO	8
ABSTRACT.....	9
1. INTRODUÇÃO.....	11
2. REVISÃO BIBLIOGRÁFICA	13
2.1. Cianobactérias supralitorais.....	13
2.2. Biofilmes cianobacterianos	15
2.3. Produtos naturais cianobacterianos e mineração genômica	18
3. OBJETIVO	25
3.1. Objetivo geral.....	25
3.2. Objetivos específicos	25
4. MATERIAIS E MÉTODOS	27
4.1. Linhagem <i>Capilliphyucus salinus</i> ALCB114379	27
4.2. Produção de biomassa	27
4.3. Purificação da cultura, extração do DNA total e sequenciamento	28
4.4. Montagem do genoma cianobacteriano.....	28
4.5. Análise taxonômica	29
4.5.1 Filogenia do gene de 16S rRNA	29
4.5.2. Filogenômica.....	29
4.6. Anotação de genes e potenciais agrupamentos gênicos	30
4.7. Identificação de aminoácidos do tipo micosporina (MAAs).....	31
4.7.1. Espectrofotometria de absorção no ultravioleta	31
4.7.2. Cromatografia Líquida acoplada à espectrometria de massas	31
4.8. Caracterização dos microrganismos associados	32
4.8.1. Verificação da camada mucilaginosa	32
4.8.2. Montagem dos genomas	33
4.8.3. Filogenômica.....	33
4.8.4. Pangenoma.....	33
4.9. Filogenia do gene <i>nifH</i>	34
5. RESULTADOS E DISCUSSÃO.....	35
5.1. Montagem do genoma cianobacteriano	35
5.2. Filogenia de <i>Capilliphyucus salinus</i> ALCB114379	41
5.3. Identificação de agrupamentos gênicos associados ao metabolismo secundário	48
5.3.1. Aminoácidos do tipo micosporina (MAAs)	48
5.3.2. Microciclâmida	54

5.4. Diversidade taxonômica da cianosfera de <i>C. salinus</i> ALCB114379	58
5.5. Avaliação dos genes associados ao metabolismo do nitrogênio.....	68
6. CONCLUSÕES	75
REFERÊNCIAS.....	77

RESUMO

Investigação do potencial genômico funcional da cianobactéria *Capilliphycus salinus* ALCB114379 e sua comunidade microbiana

As cianobactérias, organismos procarióticos de morfologia variada, destacam-se por sua eficiente adaptação a diferentes ambientes, atribuída à produção diversificada de metabólitos secundários e exopolissacarídeos (EPSs). Formas filamentosas, como *Capilliphycus salinus* ALCB114379, em conjunto com a influência da matriz de EPS, possibilitam a associação de microrganismos a esse hospedeiro cianobacteriano. Este organismo marinho, classificado recentemente como linhagem referência e pertencente à ordem Oscillatoriales, não dispõe de genoma em bancos de dados. Este estudo investigou as características genômicas de *C. salinus* ALCB114379, bem como de sua comunidade microbiana associada. A cianobactéria foi sequenciada utilizando a plataforma PacBio HiFi, e os dados foram analisados com ferramentas genômicas. Apesar da escassez de genomas de linhagens de referência em Oscillatoriales, a ALCB114379 agrupou-se dentro da família Serenicapillariaceae, reforçando sua classificação como *Capilliphycus*. A análise do genoma revelou uma quantidade expressiva de genes codificantes de proteínas hipotéticas (57,41%), indicando uma diversidade metabólica desconhecida. A mineração do genoma identificou genes biossintéticos de interesse biotecnológico, como aminoácidos do tipo micosporina e microciclâmica. Enquanto o primeiro foi validado experimentalmente quanto à produção, o segundo permanece desconhecido. A análise taxonômica da comunidade associada confirmou uma predominância de Pseudomonadota, permitindo a montagem do genoma de cinco gêneros distintos para a predição de potenciais interações com seu hospedeiro. A cianobactéria *C. salinus* ALCB114379 apresentou genes *nif* para a produção de nitrogenase de molibdênio, sugerindo um potencial para a realização do importante processo de fixação biológica de nitrogênio atmosférico. O agrupamento taxonômico utilizando *nifH* sugere uma estratégia temporal de fixação de nitrogênio adotada pela cianobactéria. Esses resultados do primeiro genoma de uma linhagem referência de *Capilliphycus* ampliam o conhecimento sobre a filogenia de Oscillatoriales, de sua diversidade metabólica e potenciais interações com bactérias associadas à linhagem *C. salinus* ALCB114379.

Palavras-chave: Culturas não axênicas, Fixação de nitrogênio, Metabolismo secundário, Microbioma associado, Mineração genômica

ABSTRACT

Investigation of the functional genomic potential of the cyanobacterium *Capilliphycus salinus* ALCB114379 and its microbial community

Cyanobacteria, prokaryotic organisms with diverse morphologies, stand out for their efficient adaptation to different environments, attributed to the diversified production of secondary metabolites and exopolysaccharides (EPSs). Filamentous forms, such as *Capilliphycus salinus* ALCB114379, in addition to the influence of the EPS matrix, enable the association of microorganisms with this cyanobacterial host. This marine organism recently classified as a reference strain and belonging to the order Oscillatoriales, lacks a genome in databases. This study investigated the genomic characteristics of *C. salinus* ALCB114379, as well as its associated microbial community. The cyanobacterium was sequenced using the PacBio HiFi platform, and the data were analyzed with genomic tools. Despite the scarcity of reference strains genomes in Oscillatoriales, ALCB114379 clustered within the family Serenicapillariaceae, reinforcing its classification as *Capilliphycus*. Genome analysis revealed a significant amount of genes encoding hypothetical proteins (57.41%), indicating unknown metabolic diversity. Genome mining identified biotechnologically relevant biosynthetic genes, such as mycosporine-like amino acids and microcyclamide. While the first was experimentally validated for production, the second remains unknown. Taxonomic analysis of the associated community confirmed a predominance of Pseudomonadota, allowing the assembly of genomes for five distinct genera to predict potential interactions with its host. The cyanobacterium *C. salinus* ALCB114379 presented *nif* genes for the production of molybdenum nitrogenase, suggesting potential for the crucial process of biological nitrogen fixation. Taxonomic clustering using *nifH* suggests a temporal strategy for nitrogen fixation adopted by the cyanobacterium. These results from the first genome of a reference strain of *Capilliphycus* expand knowledge about the phylogeny of Oscillatoriales, its metabolic diversity, and potential interactions with bacteria associated with the *C. salinus* ALCB114379 strain.

Keywords: Non-axenic cultures, Nitrogen fixation, Secondary metabolism, Associated microbiome, Genomic mining

1. INTRODUÇÃO

As cianobactérias são organismos procariontes que desempenham um papel crucial na história da Terra, sendo pioneiras na fotossíntese oxigênica e contribuindo para a formação da camada de ozônio há cerca de $3,6 \pm 0,2$ bilhões de anos (Garcia-Pichel *et al.*, 2019; Mataloni; Komárek, 2004). Além de suas habilidades fotossintéticas e na fixação de carbono, as cianobactérias exercem funções ecológicas fundamentais, como a fixação de nitrogênio e a precipitação de carbonatos. Podem ser encontradas em diversos ambientes, desde ecossistemas aquáticos, como compartimentos bentônicos e planctônicos, até ambientes terrestres (Capone *et al.*, 2005; Charpy *et al.*, 2012; Dupraz *et al.*, 2009). Além disso, apresentam diversidade morfológica, podendo assumir formas unicelulares, coloniais ou filamentosas (Allaf; Peerhossaini, 2022).

A notável diversidade e história evolutiva das cianobactérias refletem-se em sua capacidade adaptativa diante de variadas condições climáticas, desde os desertos extremamente frios do Ártico até fontes termais (Dadheech *et al.*, 2013; Harding *et al.*, 2018). Essa ampla adaptabilidade é resultado de inúmeras estratégias de aclimação fisiológica, permitindo que esses microrganismos prosperem em nichos ecológicos altamente competitivos (Hirose *et al.*, 2017; Ho *et al.*, 2017).

Essa capacidade pode ser atribuída aos metabólitos secundários que produzem, não sendo essenciais para o crescimento celular, mas importantes para a sobrevivência em condições de estresse (Seyedsayamdost, 2019). A produção de exopolissacarídeos (EPSs), sejam homopolissacarídeos ou heteropolissacarídeos, secretados pelas células cianobacterianas, presume-se também estar relacionada como estratégia de proteção frente às flutuações ambientais. Supõe-se que a matriz de EPS facilite a fixação de elementos essenciais e micronutrientes presentes no meio ambiente, contribuindo para a fixação do microrganismo a um substrato sólido, além de desempenhar um papel na formação de colônias (Laroche, 2022).

Os genomas de cianobactérias têm sido objeto de crescente atenção devido ao potencial destes organismos na produção de metabólitos secundários (Demay *et al.*, 2019). A bioinformática, empregada na mineração genômica, possibilita a vinculação desses produtos aos seus agrupamentos gênicos biossintéticos (BGCs) correspondentes. Essa abordagem não apenas permite a prospecção de novas substâncias ou derivados dentro de famílias de metabólitos de interesse, mas

também destaca o papel fundamental da informação genômica na compreensão dessas características bioquímicas (Dittmann *et al.*, 2015; Machado; Tuttle; Jensen, 2017).

O novo gênero de cianobactéria marinha, coletado na região supralitoral e denominado *Capilliphycus*, foi descrito por meio de uma abordagem polifásica, desvinculando-se do gênero *Lyngbya*. Pertencente à ordem Oscillatoriales, esse gênero apresenta uma estrutura celular filamentosa não heterocitada, incluindo células especializadas chamadas hormogônios, responsáveis pela dispersão (Caires *et al.*, 2018). Dado que esta ordem possui poucas linhagens exploradas na literatura científica e ainda não conta com genomas representativos, o estudo do genoma da linhagem referência *Capilliphycus salinus* ALCB114379 não apenas contribui para o entendimento da diversidade do filo Cyanobacteria, mas também aprimora a compreensão do papel ecológico desse novo gênero, assim como seu potencial na produção de metabólitos secundários.

2. REVISÃO BIBLIOGRÁFICA

2.1. Cianobactérias supralitorais

A região supralitoral é uma área dinâmica e ecologicamente importante situada entre a marca da maré alta e o ambiente terrestre. Caracteriza-se por condições adversas, incluindo exposição ao ar, água doce da chuva, água do mar devido ao respingo das ondas e predação por animais terrestres e marinhos (Figura 1). Esta região abriga uma variedade de espécies especializadas que se adaptaram às mudanças frequentes no nível do mar, temperatura e salinidade (Hurtado; Lee; Mateos, 2013).

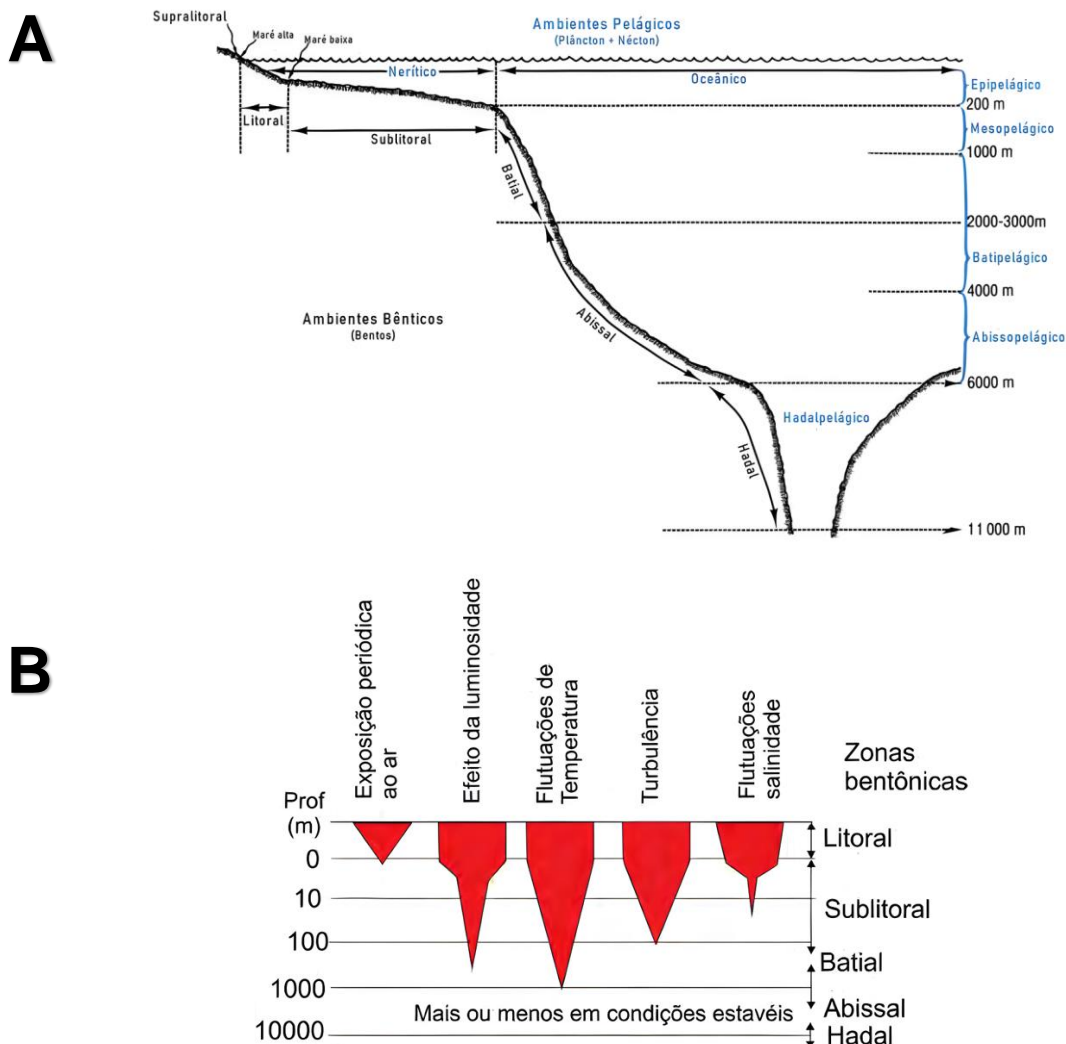


Figura 1. A. Divisões ecológicas básicas dos oceanos (Lalli *et al.*, (1997). B. Fatores abióticos que afetam o bentos marinho, a largura das faixas indicam a influência do fator. Adaptado de Kingston (2009).

O supralitoral é menos estudado em comparação com as biotas entremarés, e a sua biologia e evolução populacional ainda não são totalmente compreendidas.

Pesquisas recentes concentraram-se na estrutura genética de organismos costeiros não marinhos, como os besouros supralitorais na Bacia do Mediterrâneo Ocidental, e na influência das correntes oceânicas na sua dispersão (Villastrigo *et al.*, 2023). Além disso, estudos destacaram o impacto potencial das alterações do nível do mar no ambiente supralitoral, destacando a necessidade de compreender e prever os efeitos das alterações climáticas nos ecossistemas costeiros (Bonello *et al.*, 2022).

Os desafios únicos apresentados pela região supralitoral criam um nicho ecológico distinto, promovendo a adaptação de espécies especializadas às suas condições em constante mudança (Molina-Menor *et al.*, 2019). Conforme explora-se as comunidades microbianas que prosperam nesse espaço dinâmico, destacam-se as cianobactérias. Nesta zona de transição, onde convergem os reinos terrestre e marinho, esses microrganismos desempenham um papel fundamental na formação da paisagem ecológica. Eles são encontrados em relações simbióticas com organismos como protozoários, macroalgas, ervas marinhas e esponjas, até ascídias e outros invertebrados, e produzem uma infinidade de moléculas bioativas, muitas vezes como metabólitos secundários, exibindo várias funções biológicas e ecológicas (Bacci *et al.*, 2015; Burganskaya *et al.*, 2019; Mutalipassi *et al.*, 2021; Pagliara *et al.*, 2020; Sings; Bible; Rinehart, 1996).

Um exemplo dessa relação simbiótica inclui a associação entre cianobactérias e macroalgas, em que as cianobactérias fornecem nutrientes às macroalgas, como nitrogênio, fósforo e ferro, auxiliando no crescimento e reprodução das macroalgas. Em retorno, as macroalgas oferecem um habitat para as cianobactérias e podem protegê-las da predação e de outros estressores ambientais (Cooper; Smith, 2015).

Além disso, as cianobactérias demonstram uma notável capacidade de lidar com a intensa radiação visível e ultravioleta, um desafio comum na região supralitoral. Através do seu metabolismo secundário, produzem moléculas específicas, como os carotenóides, servindo como protetores eficazes contra a radiação (Pentecost, 2014). Esses organismos também desempenham um papel crucial na formação de biofilmes, essenciais para a sobrevivência e crescimento de outros organismos neste ambiente. Os biofilmes são fundamentais para fornecer um substrato para a fixação de vários organismos e promover interações ecológicas. Estudos demonstraram que os EPSs de cianobactérias são componentes-chave na formação e função desses biofilmes. Os EPSs contribuem para a adesão, estrutura, proteção contra estresses abióticos, processos de biointemperismo, motilidade de

deslizamento e armazenamento de nutrientes dentro dos biofilmes (Rossi; De Philippis, 2015).

2.2. Biofilmes cianobacterianos

Os biofilmes de cianobactérias apresentam características distintas que contribuem para o seu significado ecológico na zona supralitoral. Estes sistemas microbianos não estão confinados a um habitat específico; pelo contrário, apresentam uma ampla distribuição, abrangendo diversos ecossistemas, desde ambientes aquáticos a paisagens terrestres, com uma presença notável nas condições desafiadoras da zona supralitoral (Brito *et al.*, 2022).

Os biofilmes de cianobactérias são caracterizados pela incorporação de cianobactérias em uma matriz EPS. Esta matriz tem um duplo propósito: fornecer estabilidade estrutural ao biofilme e criar um ambiente protegido para células microbianas (Figura 2). A intrincada estrutura dos biofilmes de cianobactérias promove ambientes propícios à associação de vários microrganismos com seus hospedeiros cianobacterianos (Stiefelmaier *et al.*, 2020).

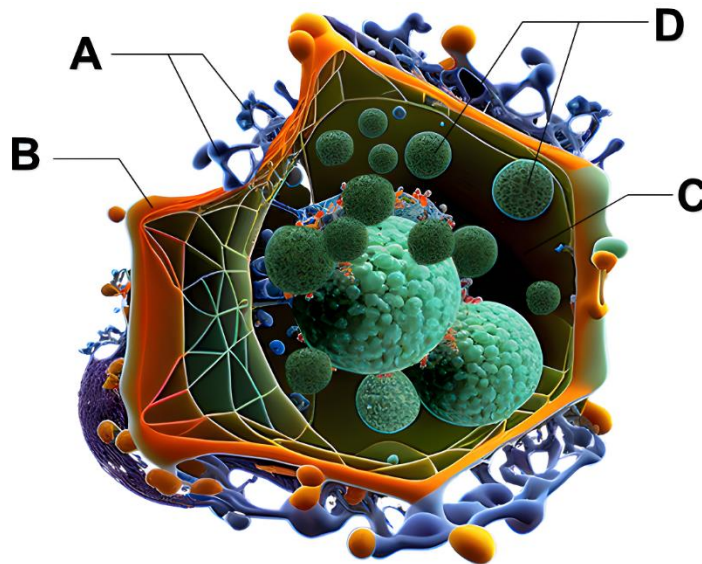


Figura 2. Formação de biofilme cianobacteriano. A. Exopolissacarídeos (EPSs); B. Biofilme; C. Matriz de EPSs; D. Cianobactérias unicelulares.

Uma característica definidora dos biofilmes de cianobactérias é o seu papel ecológico vital nos ecossistemas marinhos. Além do seu significado estrutural, estes biofilmes contribuem ativamente para a ciclagem de nutrientes e para a produção de oxigênio. As cianobactérias, através da fotossíntese, desempenham um papel

fundamental na conversão de dióxido de carbono em matéria orgânica, impactando a disponibilidade de nutrientes no ambiente circundante. Além disso, o oxigênio produzido durante a fotossíntese torna-se um componente crucial do ecossistema marinho local, influenciando tanto os processos microbianos aeróbicos como a biodiversidade global da zona supralitoral (Nagarkar; Williams, 1999).

Uma pesquisa recente revelou uma relação simbiótica até então não observada entre cianobactérias bentônicas e fungos, indicando que os simbiontes fúngicos impactam a durabilidade do biofilme cianobacteriano e a eficiência fotossintética. Esse estudo identificou uma intrigante interação simbiótica entre *Leptolyngbya frigida*, uma cianobactéria bentônica, e *Clonostachys rosea*, um fungo conhecido por suas capacidades de biocontrole contra patógenos de plantas. A interação sugere uma fase inicial de liquenização, e o biofilme resultante apresentou maior resiliência e desempenho fotossintético superior em comparação com aqueles sem a presença de fungos (Swia *et al.*, 2023).

Além disso, foram investigadas a organização espacial e o proteoma de um biofilme de cianobactérias composto por duas espécies: a cianobactéria diazotrófica filamentosa *Tolypothrix* sp. PCC 7712 e microrganismos heterotróficos. O estudo destacou o papel das *Pseudomonas* no processo inicial de fixação, bem como os efeitos de diferentes estratégias de alimentação de nitrogênio e regimes de operação na composição e crescimento do biofilme. Foi revelado que a *Pseudomonas taiwanensis* VLB120 estava utilizando metabólitos fornecidos pelas cianobactérias como únicas fontes de carbono e nitrogênio (Bozan *et al.*, 2023).

A formação de biofilmes de cianobactérias é influenciada por diversos fatores ambientais, como a disponibilidade de água, temperatura, umidade relativa, insolação e condições hidrodinâmicas. A quantidade de água e a umidade relativa são cruciais para o desenvolvimento desses biofilmes, pois a escassez desse elemento pode limitar seu crescimento em certos ambientes. A temperatura é outro fator importante, impactando as taxas de crescimento, eficiência fotossintética e desenvolvimento das cianobactérias. A insolação, que representa a intensidade de luz solar recebida, influencia a atividade fotossintética e as taxas de crescimento, sendo altos níveis benéficos e baixos limitadores para o crescimento dos biofilmes (Reboah *et al.*, 2023).

Condições hidrodinâmicas, como o fluxo de água e a tensão de cisalhamento, exercem influência significativa na formação e desenvolvimento de biofilmes de

cianobactérias. Taxas de cisalhamento e propriedades de superfície diferentes podem impactar o comportamento de formação de biofilme em várias cianobactérias marinhas, afetando seu crescimento e desenvolvimento geral (Romeu *et al.*, 2019).

Em relação às potenciais aplicações tecnológicas com biofilmes cianobacterianos, há um crescente interesse em processos de produção fotobioteecnológicos sustentáveis. Bozan *et al.*, (2022) avaliaram diferentes cianobactérias para aplicação em novos reatores de biofilme capilar, nos quais os organismos se autoimobilizam nas paredes do reator, alcançam alta biomassa e permitem processos de produção longos e robustos. Esta abordagem tem se mostrado promissora no aumento da produção de biomassa e na estabilidade do processo, com linhagens específicas de cianobactérias demonstrando alta produção de biomassa e ótima cobertura de superfície nos reatores.

Além disso, a organização espacial e o proteoma de biofilmes de cianobactérias foram investigados, lançando luz sobre suas potenciais aplicações em processos de produção sustentáveis e em biorrefinarias. Destacando o potencial para a produção sustentável de produtos químicos e combustíveis a granel de baixo valor (Frenkel *et al.*, 2023).

Em decorrência da associação entre as cianobactérias e outros microrganismos, assim como bactérias heterotróficas, o isolamento e a purificação de cianobactérias apresentam vários desafios devido às dificuldades enfrentadas para a obtenção de culturas axênicas a partir de amostras ambientais. Além disso, os EPSs produzidos pelas cianobactérias podem favorecer a formação de agregados de microrganismos, resultando em fortes ligações, sendo possível até mesmo protegê-los contra antibióticos. Essa associação complexa entre cianobactérias e outros microrganismos cria um ambiente desafiador para a obtenção de culturas puras. Os EPSs, além de promoverem a agregação, desempenham um papel significativo na proteção dos microrganismos envolvidos, conferindo uma resistência adicional contra agentes antimicrobianos. Esses fatores tornam essencial o desenvolvimento de métodos aprimorados para o isolamento e purificação eficientes de cianobactérias em estudos ambientais e aplicações biotecnológicas (Vázquez-Martínez *et al.*, 2004; Waterbury, 2006).

2.3. Produtos naturais cianobacterianos e mineração genômica

Os produtos naturais cianobacterianos constituem um grupo diversificado de moléculas sintetizadas, apresentando uma ampla variedade de atividades biológicas. Esses produtos naturais demonstram propriedades antibacterianas, antifúngicas, anticancerígenas, imunossupressoras, anti-inflamatórias e antituberculose, com aplicações potenciais em farmacologia, cosmetologia, agricultura e indústria alimentícia (Demay *et al.*, 2019; Jafari Porzani; Konur; Nowruzi, 2022; Tan, 2007).

A diversidade química desses metabólitos foi extensivamente descrita, abrangendo várias classes químicas, como alcaloides, depsipeptídeos, lipopeptídeos, macrólidos/lactonas, peptídeos, terpenos, polissacarídeos, lipídios, policetídeos e outros. Os produtos naturais de cianobactérias têm sido alvo de estudos devido ao seu potencial na descoberta de medicamentos, incluindo sua aplicação na compreensão e tratamento da dor crônica (Singh *et al.*, 2011).

Essas moléculas são sintetizadas através de agrupamentos gênicos biossintéticos (BGCs – Biosynthetic Gene Clusters), localizados em seus genomas e envolvidos na produção de diversos produtos naturais. Os BGCs são conjuntos intimamente ligados de genes, na maioria das vezes não homólogos, que participam de uma via metabólica comum. Eles podem estar associados a várias vias enzimáticas, como peptídeos sintetases não ribossômicos (NRPS), policetídeo sintases (PKS) e terpênicos ciclases (Kwon *et al.*, 2021). Esses agrupamentos são frequentemente encontrados nos genomas de bactérias, desempenhando um papel fundamental na produção de metabólitos secundários. Esses metabólitos são a base para a maioria dos produtos farmacêuticos, toxinas naturais, comunicação química e adaptação dos organismos às condições de estresse e predatismo. A produção desses metabólitos pelas cianobactérias é frequentemente aumentada em resposta a fatores de estresse abióticos ou bióticos, proporcionando proteção e auxiliando na sobrevivência, oferecendo uma vantagem sobre outras espécies (Popin *et al.*, 2021).

Entre as diversas funções mencionadas dos metabólitos secundários produzidos por cianobactérias, destacam-se as cianotoxinas. Essas toxinas têm uma notável capacidade de adaptação dos organismos produtores, inibindo o crescimento de microrganismos concorrentes e afetando a biossíntese de compostos em invertebrados marinhos (Berry, 2008). Exemplos de cianotoxinas incluem microcistina, cilindrospermopsina, anatoxina e guanitoxina (Fiore *et al.*,

2020; Moreira *et al.*, 2013; Schreidah *et al.*, 2020-; Wood; Biessy; Puddick, 2018). Além disso, outros aleloquímicos podem desempenhar papel na defesa contra predadores e herbívoros, inibindo macrófitas, algas e micróbios concorrentes. (Berry, 2008).

No contexto dessas interações, as cianobactinas, peptídeos cíclicos produzidos por diversas cianobactérias, contribuem para a complexidade desses sistemas. Esses peptídeos estão presentes em diversos gêneros dentro da linhagem cianobacteriana, estimando-se que ocorram em aproximadamente uma em cada cinco (Sivonen *et al.*, 2010). Sintetizadas ribossomicamente, as cianobactinas passam por modificações pós-tradução em vias de múltiplas etapas, envolvendo uma ampla variedade de enzimas. O processo de produção das cianobactinas inclui a clivagem proteolítica e a ciclização de peptídeos precursores, juntamente com outras modificações pós-traducionais, como heterociclização, oxidação ou prenilação de aminoácidos. Essas moléculas podem apresentar atividades antimaláricas, antitumorais e de reversão de múltiplas drogas, apresentando, assim, potenciais produtos farmacêuticos (Gu *et al.*, 2018; Sivonen *et al.*, 2010).

Um exemplo específico é a cianobactina tipo microciclâmida, um hexapeptídeo cíclico citotóxico produzido por *Microcystis aeruginosa*. Sua estrutura química, caracterizada por três anéis azólicos, sugere uma origem na modificação de cisteína e treonina. Sua identificação foi inicialmente relatada na linhagem cianobacteriana *Microcystis aeruginosa* NIES-298 (Ishida; Nakagawa; Murakami, 2000). Esses peptídeos cíclicos são sintetizados pela via de biossíntese ribossômica por meio de um conjunto de enzimas relacionado à biossíntese de patelâmida (Ziemert *et al.*, 2008). Essa molécula tem apresentado citotoxicidade em várias linhas celulares cancerígenas e em organismos aquáticos. Como exemplo, a microciclâmida isolada de *Microcystis aeruginosa* NIES-298 exibiu atividade citotóxica moderada em células de leucemia murina P388 (Ishida; Nakagawa; Murakami, 2000), enquanto a microciclâmida isolada da linhagem *Microcystis aeruginosa* LTPNA 08 apresentou cardiotoxicidade em zebrafish (Freitas *et al.*, 2023). A função ecológica dessa molécula permanece desconhecida.

Ziemert *et al.*, (2008) identificaram os genes que compõem o BGC da microciclâmida em duas linhagens de *Microcystis*, a *M. aeruginosa* NIES-298 e a *M. aeruginosa* PCC 7806. Nesses organismos, agrupamento é composto pelos genes *mcaA*, *mcaB*, *mcaC*, *mcaD*, *mcaE*, *mcaF*, ORF1, ORF2 e *mcaG*, sequencialmente,

sendo que os quadros de leitura abertos (ORFs - Open Reading Frames) 1 e 2 não participam efetivamente da biossíntese da microciclâmica.

Além da adaptação contra o predatismo, as cianobactérias produzem diversos metabólitos secundários com funções fotoprotetoras, permitindo-lhes sobreviver em diversos ambientes sob altas intensidades de luz UV. Um exemplo notável é a escitonemina, um metabólito lipossolúvel que se acumula na bainha polissacarídica extracelular de várias cianobactérias. Possui forte absorção máxima na região UV-A e azul-violeta. Sugere-se que seu acúmulo na bainha de cianobactérias seja uma estratégia adotada por diversas cianobactérias para proteger seus componentes celulares contra os efeitos nocivos da radiação ultravioleta (Pathak *et al.*, 2020). Outra estratégia adotada pelas cianobactérias é a biossíntese de aminoácidos do tipo micosporina (MAAs - Mycosporine-like Amino Acids). Esses metabólitos desempenham um papel crucial na fotoproteção, permitindo que as cianobactérias sobrevivam em uma ampla variedade de ambientes, devido à grande diversidade dessa biomolécula, caracterizada por absorver diferentes comprimentos de onda (Geraldes; Pinto, 2021).

Os MAAs são derivados do metabolismo secundário e são produzidos por uma grande diversidade de organismos encontrados em ambientes com alta incidência de luz solar, geralmente em ambientes aquáticos, como líquens, fungos, algas e cianobactérias. Essas moléculas, além de possuírem ação fotoprotetora ao absorverem a luz UV, também podem ter funções relacionadas à proteção contra o estresse oxidativo e até na regulação osmótica das células (Chrapusta *et al.*, 2017; Geraldes *et al.*, 2020b). Atualmente, um banco de dados sobre MAAs, denominado MYCAS, conta com 74 moléculas, incluindo informações estruturais e dados espectrométricos, e pode ser acessado de forma gratuita e em domínio público, através do seguinte link <http://www.cena.usp.br/ernani-pinto-mycas> (Geraldes; Pinto, 2021).

Na Figura 3, referente a biossíntese de MAAs, a ciclase SH-7P (2-epi-5-epi-valiolona sintase, EVS), em conjunto com uma O-metiltransferase (OMT), desempenham o papel de primeira e segunda enzimas, respectivamente, na via que conduz ao precursor MAA, o 4-desoxigadusol (Dextro *et al.*, 2023; Pope *et al.*, 2015). Em ambas as rotas, o 4-DG serve como a estrutura central parental dos MAAs, e a adição de glicina resulta na formação de micosporina-glicina. Este MAA simples, do tipo ciclohexenona mono-substituída, representa um intermediário comum na

produção de MAAs dissustituídos por meio da adição de um único resíduo de aminoácido. Isso resulta na produção de alguns MAAs comuns, como por exemplo, porphyra-334 e Shinorina. Nesta etapa, é codificada uma proteína que se assemelha a um peptídeo sintase não ribossômica (NRPS) ou a uma D-alanil-D-alanina ligase (D-ala-D-ala ligase) (Gerald *et al.*, 2020c). Outros MAAs são sintetizados por meio de modificações nos grupos laterais anexados, envolvendo substituições de nitrogênio, como esterificação, amidação, desidratação, descarboxilação, hidroxilação, sulfonação e glicosilação (Lawrence; Long; Young, 2017). As variações nas cadeias laterais dos aminoácidos são responsáveis pelas diferenças nos espectros de absorção dos MAAs. Apesar de sugerida a via do chiquimato como uma potencial rota biossintética para MAAs, a mineração do genoma e os dados de expressão gênica heteróloga oferecem evidências convincentes de que a biossíntese de MAAs originada da via das pentoses fosfato é a mais difundida entre as cianobactérias (Carreto; Carignan, 2011; Dextro *et al.*, 2023).

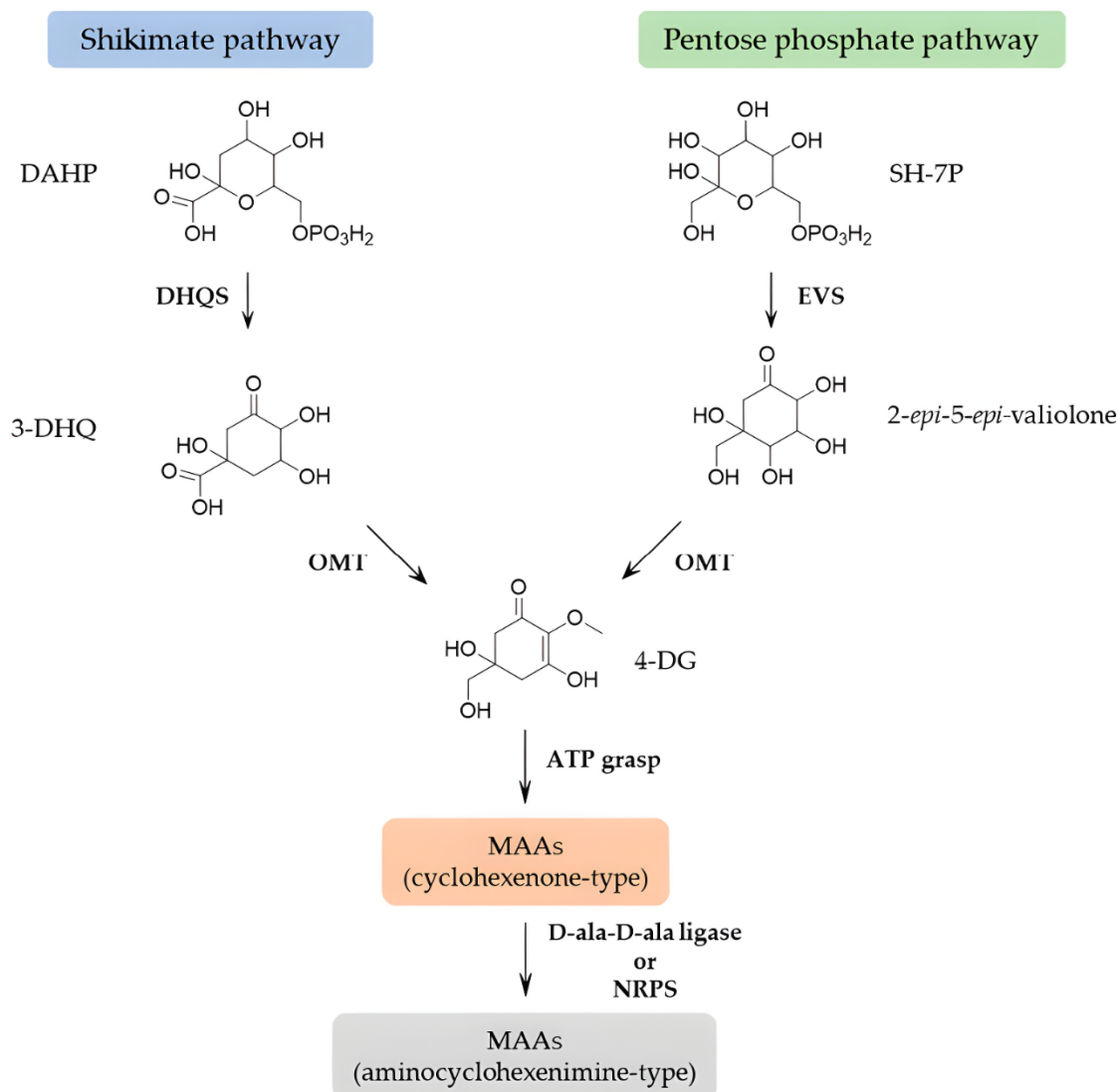


Figura 3. Vias biossintéticas de MAAs. DAHP: fosfato de 3-desoxi-D-arabino-heptulosonato, DHQS: 3-desidroquinato sintase, 3-DHQ: 3-desidroquinato, SH-7P: sedoheptulose-7-fosfato, EVS: ciclase 2-epi-5-epi-valiolona sintase, OMT: O-metiltransferase, 4-DG: 4-desoxigadusol, ATP: trifosfato de adenosina, NRPS: peptídeo sintase não ribossômico, D-ala D-ala-ligase: D-alanil-D-alanina ligase (Geraldes *et al.*, 2020c).

A compreensão dos BGCs é crucial para desvendar os mecanismos de biossíntese desses metabólitos e explorar suas potenciais aplicações. A investigação desses agrupamentos genéticos não apenas contribui para a elucidação de vias metabólicas específicas, mas também abre perspectivas para a descoberta de novas moléculas e a engenharia metabólica para a sua produção. Essa interligação de genes em BGCs é um fenômeno fascinante que revela a complexidade e a sofisticação dos mecanismos adaptativos dos organismos em resposta ao ambiente (Li; Jiang; Lu, 2017; Li; MacIntyre; Brady, 2021).

A pesquisa sobre BGCs foi revolucionada pela análise computacional, que permite a rápida identificação e exploração de BGCs em sequências genômicas. Bancos de dados como o BiG-FAM e o MiBIG fornecem uma plataforma para explorar a diversidade global de BGCs e suas relações, ajudando a mapear sua diversidade arquitetônica e taxonômica e fornecendo previsões sobre potenciais BGCs (Kautsar *et al.*, 2021; Medema *et al.*, 2015).

A evolução dos métodos de sequenciamento teve impactos significativos em diversos campos, incluindo a pesquisa genômica por meio da bioprospecção de BGCs de interesse biotecnológico. Esses avanços trouxeram consigo mudanças substanciais, como o desenvolvimento de tecnologias de sequenciamento de próxima geração (NGS) e de terceira geração. Estas proporcionaram um sequenciamento mais ágil e preciso, viabilizando a análise de milhares a dezenas de milhares de genomas em um intervalo temporal reduzido (Heather; Chain, 2016).

O custo associado ao sequenciamento de um genoma diminuiu expressivamente ao longo do tempo, impulsionando sua adoção em larga escala em diversas áreas de pesquisa e ambientes clínicos (Xiao; Zhou, 2020). Além disso, no âmbito da descoberta de medicamentos, as tecnologias de sequenciamento de terceira geração se destacam ao apresentar potencial para avanços substanciais no processo. Elas possibilitam a identificação mais eficiente e econômica de alvos terapêuticos, bem como a avaliação de candidatos a medicamentos por meio de uma maior diversidade de dados genômicos disponíveis em bancos de dados públicos. Esses progressos promovem uma abordagem mais eficaz na busca por novos tratamentos, contribuindo para avanços significativos no campo da medicina (Van de Sande *et al.*, 2023).

Quanto aos bancos de dados empregados em abordagens genômicas, eles podem ser divididos em dois tipos: bancos de dados biológicos primários e os secundários. Os bancos de dados primários consistem em dados derivados experimentalmente, como sequências de nucleotídeos, proteínas ou estruturas macromoleculares. Esses bancos armazenam dados brutos de sequência e têm uma natureza fundamentalmente arquivística. Uma vez atribuído um número de acesso ao banco de dados, os dados primários permanecem inalterados, tornando-se parte integrante do registro científico. Exemplos de bancos de dados primários incluem o GenBank (National Center for Biotechnology Information, NCBI), o DNA Data Bank of Japan (DDBJ) e o European Nucleotide Archive (ENA) (Burgin *et al.*,

2023; Sayers *et al.*, 2021; Tanizawa *et al.*, 2023). Já os bancos de dados secundários derivam informações dos bancos de dados primários, muitas vezes incluindo dados curados manualmente ou processados computacionalmente. Exemplos de bancos de dados secundários englobam o SWISS-PROT, Uniprot, InterPro, e Ensembl (Bateman *et al.*, 2021; Cunningham *et al.*, 2022; Gasteiger; Jung; Bairoch, 2001; Paysan-Lafosse *et al.*, 2023).

A mineração genômica tem se tornado uma abordagem valiosa para a descoberta de produtos naturais, e identificar BGCs de interesse tornou-se consideravelmente mais fácil devido ao crescente desenvolvimento de ferramentas de bioinformática cada vez mais precisas. No entanto, encontrar os produtos associados permanece um desafio, resultando em gargalos no processo de descoberta. Esses desafios incluem a necessidade de conhecimento em bioinformática para lidar com conjuntos de dados extensos e a dependência de bancos de dados que podem ser insuficientes para análises comparativas. As abordagens automatizadas de mineração de genoma também têm suas limitações na identificação de agrupamentos de genes biossintéticos conhecidos e na previsão das atividades biológicas dos produtos naturais identificados (Atanasov *et al.*, 2021; Albarano *et al.*, 2020).

Em conclusão, o futuro da mineração genômica é bastante promissor para a descoberta de metabolitos secundários e seus genes biossintéticos. Contudo, é importante destacar desafios relacionados à identificação de produtos associados, à expertise em bioinformática e à integralidade do sequenciamento do genoma. Superar esses desafios será crucial para realizar plenamente o potencial da mineração genômica na descoberta de produtos naturais.

3. OBJETIVO

3.1. Objetivo geral

O objetivo geral deste estudo é identificar potenciais agrupamentos gênicos de interesse biotecnológico da cianobactéria *C. salinus* ALCB114379, bem como a caracterização dos microrganismos associados a essa cianobactéria provenientes de uma cultura não axênica.

3.2. Objetivos específicos

- I) Montar e avaliar a qualidade do genoma da linhagem *C. salinus* ALCB114379;
- II) Verificar sua posição evolutiva na ordem Oscillatoriales;
- III) Identificar agrupamentos gênicos envolvidos na biossíntese de metabólitos secundários;
- IV) Caracterizar a diversidade taxonômica e funcional dos microrganismos associados à cianobactéria *C. salinus* ALCB114379;
- V) Identificar possíveis interações entre os microrganismos associados com o seu hospedeiro.

4. MATERIAIS E MÉTODOS

4.1. Linhagem *Capilliphycus salinus* ALCB114379

A cianobactéria *Capilliphycus salinus* ALCB114379 foi coletada na região supralitoral brasileira, durante as marés baixas, na Praia da Pedra do Sal, no município de Salvador (BA) (12°57'06" S, 38°20'42" W) (Figura 4) (Caires *et al.*, 2018). A linhagem está sendo mantida no biobanco do Laboratório de Ecologia Molecular de Cianobactérias do Centro de Energia Nuclear na Agricultura da Universidade de São Paulo (CENA/USP).

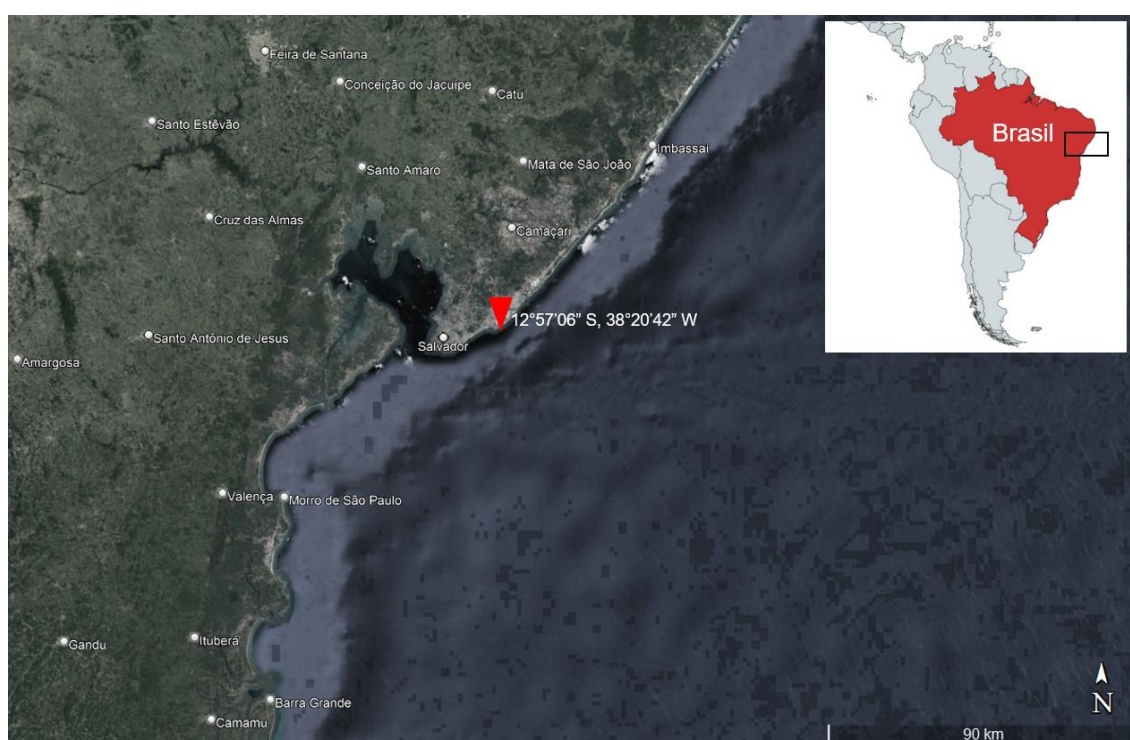


Figura 4. Localização geográfica da origem da linhagem *C. salinus* ALCB114379, na Praia do Sal, Salvador, Estado da Bahia, Brasil.

4.2. Produção de biomassa

A linhagem foi cultivada em meio Z8, sob luz fluorescente (40 a 50 $\mu\text{mol photons}\cdot\mu\text{m}^{-2}\cdot\text{s}^{-1}$) em um fotoperíodo de 14:10 claro/escuro, com temperatura controlada a 22 ± 1 °C e umidade de $\pm 60\%$. Sob as mesmas condições, foi realizada a produção de biomassa de células dessa linhagem em frascos de Erlenmeyer de 125 mL (n=3), contendo 50 mL de meio de cultura Z8.

4.3. Purificação da cultura, extração do DNA total e sequenciamento

Foi utilizado um procedimento de lavagem em série descrito por Heck *et al.*, (2016) antes da extração do DNA genômico utilizando o kit PowerSoil. A qualidade do DNA extraído foi avaliada por eletroforese em gel de agarose a 1% (m/v), e a quantificação foi conduzida utilizando o fluorômetro Qubit® 2.0 Fluorometer da Life Technologies™, EUA. Os DNAs foram enviados ao Joint Genome Institute (JGI) (Berkeley, CA, USA). Uma biblioteca PacBio SMRTbell foi preparada para sequenciamento de consenso circular (CCS) e sequenciado na plataforma PacBio HiFi. As leituras foram filtradas por qualidade, utilizando BMap v38.90 com o pipeline icecreamfinder.sh sob parâmetros de *default* (Kleiner *et al.*, 2020).

4.4. Montagem do genoma cianobacteriano

As leituras do sequenciamento geradas foram identificadas taxonomicamente com a ferramenta Kaiju v1.2.9 utilizando o banco de dados Genomes RefSeq, contendo genomas de referência completamente montados de bactérias, vírus e arqueas (Menzel; Ng; Krogh, 2016). A ausência de sequências de referência levou a uma montagem do genoma *de novo* realizada com parâmetros de *default* usando HiFiASM v0.19.5 (Cheng *et al.*, 2021). A fim de classificar e verificar a integridade do único cromossomo circular resultante, uma busca por genes ortólogos de cópia única de cianobactérias foi feita via BUSCO (Benchmarking Universal Single-Copy Orthologs) v5.3.1 (Simao *et al.*, 2015). A circularidade do genoma montado foi confirmada via Bandage e visualizada com CGView v2.0 (Stothard; Grant; Van Domselaar, 2018; Wick *et al.*, 2015). Os níveis de integridade e contaminação foram medidos usando CheckM v1.0.18 (Parks *et al.*, 2015). O AAI-profiler foi utilizado para análise e controle de qualidade, fazendo uso do proteoma do genoma montado, além de desempenhar um papel na classificação taxonômica preliminar (Medlar; Petri; Holm, 2018). A qualidade da montagem foi acessada através do Quast v5.0.2, informando o tamanho do genoma e o conteúdo do GC (Gurevich *et al.*, 2013). Para verificar a presença de conteúdo plasmidial inserido no genoma, foi utilizado a ferramenta PlasFlow v1.1.0 (Krawczyk; Lipinski; Dziembowski, 2018). O mapa do genoma foi gerado usando Proksee (Grant *et al.*, 2023). A verificação de material gênico exógeno viral no genoma cianobacteriano foi com VirSorter (Roux *et al.*, 2015). A colocação taxonômica preliminar e individual do genoma circular montado foi feita utilizando GTDB-tk Classify v1.6.0. O fluxo de trabalho dessa ferramenta

compreende três etapas: identificar, alinhar e classificar. A identificação dos genes foi feita usando o Prodigal, e modelos HMM, juntamente com o pacote HMMER, sendo selecionados 120 genes marcadores bacterianos e arqueais. No alinhamento desses genes utilizou-se o MAFFT, enquanto a classificação de cada genoma foi feita com base em sua posição na árvore de referência, sua divergência evolutiva relativa e/ou identidade média de nucleotídeos (ANI) em relação a genomas de referência no banco de dados do GTDB (Chaumeil *et al.*, 2020).

4.5. Análise taxonômica

4.5.1. Filogenia do gene de 16S rRNA

Os genes de 16S rRNA foi localizado e extraído do genoma da cianobactéria *C. salinus* ALCB114379 através da plataforma The Microbial Genomes Atlas (MiGA), utilizando a ferramenta BarRNAP (Rodriguez-R *et al.*, 2018). A identidade do gene de 16S rRNA foi comparada com a sequência da mesma linhagem disponível no GenBank (NCBI) (Nº de acesso NR_176504.1) através do programa BLASTn (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome).

Na filogenia do gene de 16S rRNA, a sequência completa do gene obtida do genoma foi alinhada com 117 sequências de 16S rRNA de cianobactérias pertencentes à ordem Oscillatoriales disponíveis no banco de dados NCBI, utilizando MUSCLE sob parâmetros de *default* (Edgar, 2004). O resultado do alinhamento foi introduzido no programa MEGAX, seguido pela seleção do modelo de substituição General Time Reversible (GTR) (Kumar *et al.*, 2018). A construção da árvore filogenética, baseou-se no método de Máxima Verossimilhança, usando um grupo externo *Gloeobacter violaceus* PCC 7421, com uma reamostragem (bootstraps) de 1.000 replicações. A árvore filogenética foi visualizada e editada na plataforma ITOL v6.5.8 (Letunic; Bork, 2019).

4.5.2. Filogenômica

As sequências de aminoácidos de 120 marcadores bacterianos conservados de cópia única foram alinhadas usando GTDB-tk v0.3.2 (Chaumeil *et al.*, 2020). O programa utilizou o Prodigal para a predição de genes e o HMMER para identificar os genes marcadores, empregando o modelo oculto de Markov (perfil HMMs) (Eddy, 2011; Hyatt *et al.*, 2010). Em seguida, o genoma da *C. salinus* ALCB114379 foi

alinhado com outros 35 genomas de referência disponíveis na base de dados NCBI de cianobactérias da ordem Oscillatoriales, além de um grupo externo, o *Gloeobacter violaceus* PCC 7421, utilizando os genes marcadores identificados como referência. Uma árvore filogenômica de máxima verossimilhança com RAxML v8.0.0 foi construída, usando parâmetros de *default* e 1.000 bootstraps (Stamatakis, 2014). A árvore filogenética foi visualizada e editada utilizando a ferramenta FigTree (Theys *et al.*, 2019).

A identidade média de nucleotídeos (ANI) e a identidade média de aminoácidos (AAI), usadas para inferir a proximidade entre genomas e para classificar um novo gênero ou espécie, foram calculadas para *C. salinus* ALCB114379 e cinco outras linhagens de Oscillatoriales dos gêneros *Lyngbya* e *Limnoraphis* de acordo com os métodos descritos por Luo *et al.*, (2014) e Goris *et al.*, (2007).

4.6. Anotação de genes e potenciais agrupamentos gênicos

Realizou-se a anotação genômica utilizando o Prokka v1.14.6 (Seemann, 2014), no qual as sequências de proteínas foram atribuídas às vias metabólicas por meio do BlastKOALA v2.2 (Kanehisa; Sato; Morishima, 2016). A anotação automatizada dos agrupamentos gênicos biossintéticos foi conduzida através da plataforma antiSMASH v7.0 (Blin *et al.*, 2023). As sequências de interesse foram identificadas e passaram por curadoria manual usando o software Artemis v18.2.0 (Carver *et al.*, 2012). A análise de colinearidade e colocação dos agrupamentos gênicos foi realizada através da análise de sintenia com a ferramenta Clinker v0.0.28 (Gilchrist; Chooi, 2021). Cada gene nos agrupamentos de interesse foi comparado com sequências disponíveis no banco de dados do NCBI para análise de homologia, utilizando os valores de identidade e cobertura. As sequências de interesse foram alinhadas com as sequências de maior correspondência do NCBI na plataforma Uniprot (Bateman *et al.*, 2021), e a similaridade foi calculada utilizando a ferramenta Ident and Sim (https://www.bioinformatics.org/sms2/ident_sim.html). Além disso, cada gene passou por predição de domínio conservado através da plataforma InterPro (<https://www.ebi.ac.uk/interpro/>). A identificação dos genes envolvidos no complexo CRISPR, responsável pelo reconhecimento de material genético exógeno, foi realizada com a ferramenta Defense Finder (Tesson *et al.*, 2022). Para o agrupamento de genes da microciclâmica, foi utilizada uma ferramenta adicional

para identificação de viés de códon com o Codon Usage (https://www.bioinformatics.org/sms2/codon_usage.html). Também foram identificados os tRNAs e seus respectivos anticódons com o tRNAscan-SE (<http://lowelab.ucsc.edu/tRNAscan-SE/>). O espaçamento intergênico do BGC da microciclâmica foi analisado com a ferramenta ORFfinder (<https://www.ncbi.nlm.nih.gov/orffinder/>).

4.7. Identificação de aminoácidos do tipo micosporina (MAAs)

4.7.1. Espectrofotometria de absorção no ultravioleta

As três linhagens produtoras de MAAs, *Sphaerospermopsis torques-reginae* ITEP-024, *Nodularia spumigena* CENA 596 e *Brasilonema octagenarum* UFV-OR1, conforme descrito por Dextro *et al.*, (2023), juntamente com a *C. salinus* ALCB114379, foram cultivadas por 45 dias na câmara do tipo BOD, sob um fotoperíodo de 14:10 claro/escuro, com temperatura controlada a 25 ± 1 °C. Utilizou-se $5 \text{ mg} \pm 0,01$ de biomassa liofilizada das quatro linhagens de cianobactérias para a extração, realizada em triplicata, utilizando 2,0 mL de ácido fórmico a 0,1% (v/v) e tampão de formato de amônio a 0,2 mM, ajustado a pH $\sim 2,55$. O processo de extração incluiu um ciclo de sonicação de 1 minuto e agitação em vortex por 5 segundos. Após 3 horas a temperatura ambiente ($\sim 22 \pm 1$ °C), os extratos foram centrifugados a 10.000 rpm por 10 minutos a 5°C, e o sobrenadante foi filtrado através de um filtro de 0,45 μm para remover restos celulares. O filtrado de cada amostra foi colocado em uma cubeta de quartzo e a absorbância na faixa de comprimento de onda de 300 nm a 360 nm foi medida usando um espectrofotômetro Perkin Elmer lambda bio. Os picos de absorbância de interesse para as MAAs shinorina (333 nm), porphyra-334 (334 nm), micosporina-glicina-alanina (332 nm) e palitina (320 nm) encontram-se dentro desse espectro pré-definido.

4.7.2. Cromatografia líquida acoplada à espectrometria de massas

Utilizou-se o método de extração conforme descrito no tópico 4.8.1, com 5,0 mg de biomassa liofilizada em triplicata da cianobactéria *C. salinus* ALCB114379. As análises foram conduzidas em um sistema LC série 1290 equipado com uma bomba 1290 VL e um sistema injetor 1260 HiP ALS acoplado a um espectrômetro de massa Triplo Quadrupolo 6460 (QqQ) da Agilent Technologies, com uma fonte de ionização por eletrospray (ESI). O instrumento QqQ foi operado por meio do software

MassHunter Workstation B.04.01. A seleção de íons precursores e produtos, bem como a otimização das energias de colisão, foi realizada com injeção de fluxo de soluções padrão usando o software MassHunter Optimizer. A separação cromatográfica foi realizada a 30 °C utilizando uma coluna C18 Synergi 4 µ Hydro-RP 80 A (150 × 2,0 mm) (Phenomenex, EUA). Extratos de amostra foram injetados (10 µL) e eluídos a uma vazão de 0,4 mL/min de água + 0,1% (v/v) de ácido fórmico contendo formato de amônio 2 mM (eluente A). O eluente B foi acetonitrila/água (8:2, v/v) + 0,1% (v/v) de ácido fórmico. O gradiente utilizado para a separação foi de 100% A de 0 a 2 minutos, seguido de um gradiente de 0 a 95% entre 2 e 6 minutos, mantendo-se esse gradiente por mais 2 minutos. Posteriormente, houve um período de equilíbrio de 4 minutos a 100% A antes da próxima injeção. O espectrômetro de massa operou no modo de monitoramento de reações múltiplas (MRM), enquanto a fonte foi operada em íon positivo. O tempo total de execução para a cromatografia foi de 12 minutos, conforme descrito por Gerald *et al.*, (2020a).

4.8. Caracterização dos microrganismos associados

4.8.1. Verificação da camada mucilaginosa

O cultivo da cianobactéria foi realizado em meio Z8 ao longo de 15 dias para a preparação de amostras. Utilizando o microscópio óptico de campo claro LEICA ICC40 W, equipado com um sistema óptico de correção infinita e iluminação Kohler, foi possível adquirir imagens precisas e sem distorção. A câmera CMOS colorida de 5 megapixels integrada ao microscópio foi empregada para capturar imagens, permitindo a observação da camada mucilaginosa e sua distribuição ao longo dos filamentos cianobacterianos. Adicionalmente, durante a observação, o corante nanquim foi adicionado para destacar a camada mucilaginosa na *C. salinus* ALCB1143789.

Para a visualização das extensões de EPS da camada mucilaginosa, amostras da cianobactéria *C. salinus* ALCB114379, previamente cultivada em meio Z8 por 15 dias, foram separadas em frascos de eppendorfs e transferidas a pequenas lamínulas de vidro circulares. As amostras foram depositadas em uma placa de Petri selada na presença de ácido ósmio durante 24 horas para fixação e, posteriormente, desidratadas em sílica gel durante duas horas. Após a secagem, as lamínulas foram aderidas sobre os stubs, os quais foram primeiramente envolvidos em papel alumínio e identificados, com fitas carbonatadas duplas. A metalização foi

realizada em um sputter coater Baltec SCD050 e, posteriormente, as amostras foram armazenadas em sílica gel até a sua utilização. As amostras foram visualizadas com um microscópio eletrônico de varredura (MEV) JEOL IT300. As imagens foram registradas digitalmente.

4.8.2. Montagem dos genomas

Devido ao alto percentual de leituras identificadas fora do filo Cyanobacteria com a ferramenta Kaiju, procedeu-se com a montagem dos genomas dos microrganismos associados à cianobactéria utilizando o software Flye v2.9 (Kolmogorov *et al.*, 2019). Dos cinco Genomas Montados de Metagenoma (MAGs - Metagenome-Assembled Genomes) resultantes, compostas por um único cromossomo circular, avaliou-se seus níveis de integridade e contaminação utilizando o CheckM. Para classificar e verificar a integridade de cada cromossomo circular nas MAGs, realizou-se uma busca por genes ortólogos de cópia única usando o BUSCO (Benchmarking Universal Single-Copy Orthologs). A circularidade das MAGs foi confirmada com o auxílio da ferramenta Bandage e sua visualização foi realizada utilizando o CGView. A qualidade das montagens foi avaliada por meio do Quast, fornecendo informações sobre o tamanho das MAGs e seus conteúdos de GC. Para a classificação taxonômica preliminar e individual de cada MAG, empregou-se a ferramenta GTDB-tk Classify.

4.8.3. Filogenômica

Utilizou-se o mesmo procedimento descrito no tópico 4.5.2. No entanto, a amostragem foi constituída por 41 genomas de referência disponíveis na base de dados do NCBI, pertencentes ao filo Pseudomonadota. Além disso, foi incluído um grupo externo representado pelo *Caldovatus aquaticus* SYSU G05006. A árvore filogenética resultante foi visualizada e editada com a ferramenta FigTree.

4.8.4. Pangenoma

As MAGs foram submetidas à anotação gênica usando a ferramenta Prokka. As proteínas preditas foram posteriormente categorizadas em subsistemas biológicos por meio do BlastKOALA para avaliar a diversidade metabólica. Para identificar a presença de genes relacionados ao ciclo biogeoquímico do nitrogênio, os genes categorizados no subsistema de metabolismo energético foram reanotados

com DRAM (Distilled and Refined Annotation of Metabolism) v0.1.2 (Shaffer *et al.*, 2023). Sequências de interesse foram identificadas e passaram por curadoria manual com o auxílio do software Artemis.

Na análise do pangenoma, a ferramenta OrthoVenn2 foi empregada (Xu *et al.*, 2019). Essa análise classificou todos os genes presentes nas cinco MAGs em dois conjuntos distintos: o genoma acessório, representando os genes individuais de cada genoma, e o genoma core, consistindo em genes ortólogos compartilhados entre as MAGs. A partir dessa classificação, foram identificadas as MAGs que potencialmente atuam no ciclo biogeoquímico do nitrogênio, determinando em qual etapa e se há compartilhamento de função.

4.9. Filogenia do gene *nifH*

Com a identificação do BGC responsável pela produção da nitrogenase de molibdênio na cianobactéria *C. salinus* ALCB114379, empregou-se a ferramenta Artemis para extrair a sequência do gene *nifH*. Em seguida, foram selecionadas 77 linhagens de cianobactérias do banco de dados do NCBI, juntamente com um grupo externo contendo o gene *nifH*, representado pelo *Mesorhizobium temperatum* CCNWNX0152. O procedimento para a análise filogenética, assim como a escolha dos parâmetros, foram os mesmos utilizados no tópico 4.5.1. Por fim, a árvore filogenética foi visualizada e editada na plataforma ITOL.

5. RESULTADOS E DISCUSSÃO

5.1. Montagem do genoma cianobacteriano

O sequenciamento do genoma da cianobactéria marinha *C. salinus* ALCB114379 na plataforma PacBio HiFi gerou um total de 326.539 leituras. Classificaram-se 322.185 leituras (98,67%) utilizando o banco de dados de genomas de referência do NCBI. Entre as leituras sequenciadas, 25% foram atribuídas ao filo Cyanobacteria, enquanto a maioria, compreendendo 69% das leituras, foi classificada no filo Proteobacteria, recentemente renomeado para Pseudomonadota.

Realizou-se quatro montagens *de novo* do genoma da cianobactéria com diferentes ferramentas (Canu assembler v2.2, Flye v2.8.3, Flye v2.9 e HiFiASM v0.16.1) para a obtenção de um genoma final consenso (Tabela 1).

Tabela 1. Estimativa e análise de qualidade das montagens individuais e consensus do sequenciamento da cianobactéria *C. salinus* ALCB114379.

	Montagem 1	Montagem 2	Montagem 3	Montagem 4 (consensus)
Programa utilizado	Canu assembler v2.2	Flye v2.8.3	Flye v2.9	HiFiASM v0.16.1
Nº de contigs	1	1	1	1
Tamanho total (pb)	7.282.116	7.712.970	7.712.904	7.712.874
(G+C)%	43,17	43,23	43,23	43,23
Completeness (%)	94,0	99,45	99,67	99,67
Contaminação (%)	0,87	0,87	0,87	0,87
Circularização	Não	Sim	Sim	Sim
Identidade (%) 16S rRNA <i>C. salinus</i> (NCBI: NR_176504.1)	99,8	99,8	99,8	99,8

Com exceção da Montagem 1, todos os outros genomas foram circularizados em um único cromossomo e apresentaram ótimos índices de completude e contaminação aferidos pela ferramenta CheckM. Essa ferramenta utiliza um banco de dados de genes marcadores para identificação e diferenciação de espécies. Os parâmetros de seleção mínima em relação à completude e contaminação foram estabelecidos em >95% e <5%, respectivamente, de acordo com Parks *et al.* (2015).

A baixa contaminação, que foi de 0,87% em todas as montagens, incluindo a Montagem 1, pode ser explicada pela inclusão de sequências de organismos não relacionados, uma vez que se trata de uma amostra não axênica, e pela presença de genes ou sequências específicas de outros organismos inseridos no genoma por transferência horizontal. Outra possível causa para a contaminação observada seria

a inserção de plasmídeos no genoma durante o processo de montagem. No entanto, essa causa, nesse caso, seria pouco provável, uma vez que não foi identificado conteúdo plasmidial no genoma utilizando a ferramenta PlasFlow.

Nas montagens realizadas, o tamanho do genoma se manteve em torno de 7,7 Mbp, enquanto o conteúdo de GC apresentou menos variação, mantendo-se em aproximadamente 43%. Apesar do tamanho de genoma e do conteúdo GC analisado estarem dentro do espectro apresentado nos bancos de dados, é importante ressaltar que as informações sobre os tamanhos dos genomas de cianobactérias, o conteúdo de GC e o número de genomas depositados no NCBI variam amplamente. O tamanho médio dos 516 genomas é de cerca de $4,82 \pm 2,39$ Mbp, com o maior genoma atingindo 15,33 Mbp (*Cyanobacterium* TDX16) e o menor medindo apenas 0,19 Mbp (*Cyanobacteria bacterium* PMG004). O percentual de GC varia de 30,8% (*Cyanobacteria bacterium* UBA6984) a 69,9% (no caso da cianobactéria *Synechococcaceae* WBB10.09), com uma média de cerca de $46,1 \pm 8,23\%$ (Dextro *et al.*, 2021). Dentro das cianobactérias da ordem Oscillatoriales, conforme o banco de dados do NCBI (release 258.0), destaca-se *Microseira wollei* NIES-4236, com o maior genoma registrado de 10,6 Mb, enquanto *Zarconia navalis* LEGE 11467 exibe o menor, com 4,2 Mb. Quanto ao conteúdo de GC, *Phormidium tenue* NIES-30 lidera com 56%, contrastando com o mínimo de 36,5% observado em *Okeania hirsuta* PAP21 Jun06-1.

A única sequência gênica da cianobactéria *C. salinus* ALCB114379 disponível nos bancos de dados é o seu 16S rRNA (NCBI: NR_176504.1). O resultado da busca BLAST desse marcador em relação às quatro montagens apresentou resultados idênticos: duas regiões ao longo do genoma com 100% de cobertura e 99,8% de identidade.

Apesar da Montagem 3 e 4 apresentarem parâmetros de qualidade idênticos, com exceção do tamanho total do genoma variando em apenas 30 pb, a Montagem 4 foi escolhida como genoma consenso (Figura 5). Essa decisão se baseou no uso da ferramenta de montagem genômica HiFiASM, desenvolvida especificamente para montar genomas sequenciados em leituras PacBio HiFi. Essa ferramenta oferece montagens de alta qualidade e tende a gerar contigs mais longos e resolver mais duplicações segmentais do que outros montadores (Gomes-dos-Santos *et al.*, 2023).

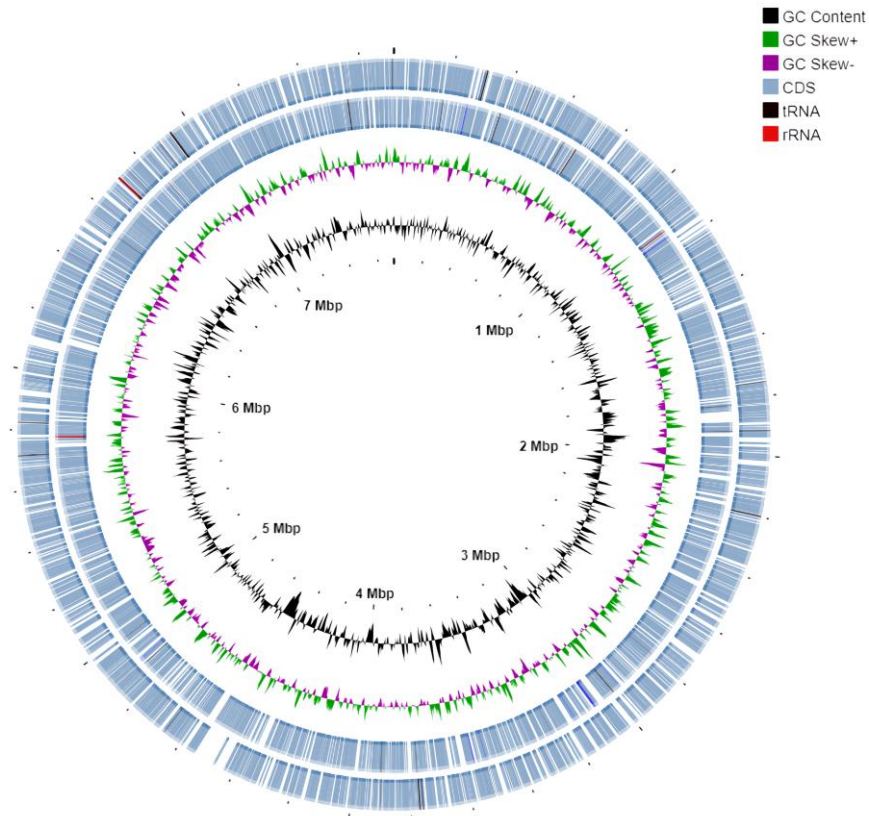


Figura 5. Genoma consenso circular montado de *C. salinus* ALCB114379, mostrando, via esquema de cores, conteúdo e inclinação de GC, regiões de sequência codificante (CDS - Coding DNA Sequences), regiões previstas para RNA de transferência (tRNA) e RNA ribossômico (rRNA).

O genoma obtido a partir da Montagem 4 foi posteriormente submetido a uma análise de qualidade adicional, complementando o CheckM, com a ferramenta BUSCO. Isso permitiu verificar a presença de genes conservados de cópia única (Figura 6).

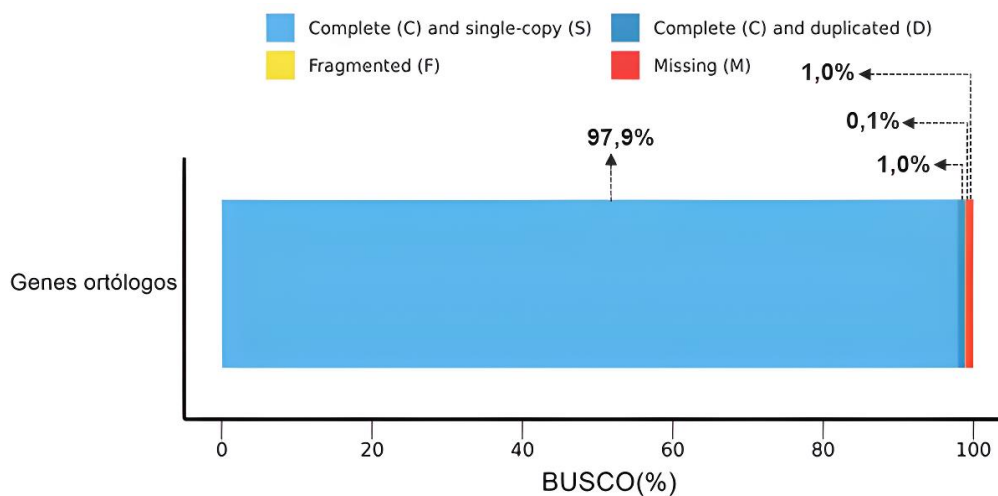


Figura 6. Estimativa percentual correspondendo a 1.462 grupos completos de cópia única, 15 grupos duplicados, 2 grupos fragmentados e 14 grupos ausentes.

Os 23 genomas de linhagens selecionadas automaticamente pelo BUSCO para comparação pertencem à ordem Oscillatoriales, que é a mesma ordem da cianobactéria *C. salinus* ALCB114379. O número limitado de genomas utilizados pela ferramenta pode ser explicado pela escassez de genomas de referência dessa ordem depositados no NCBI (release 258.0), totalizando apenas 32 genomas até o momento.

A análise da Identidade Média de Aminoácidos (AAI) foi conduzida utilizando a ferramenta AAI-profiler que utiliza o banco de dados Uniprot (Figura 7). Essa análise teve o propósito não apenas de complementar as investigações sobre qualidade e contaminação, mas também de determinar se a *C. salinus* ALCB114379 está intimamente relacionada às espécies de referência, por meio de uma estreita relação taxonômica.

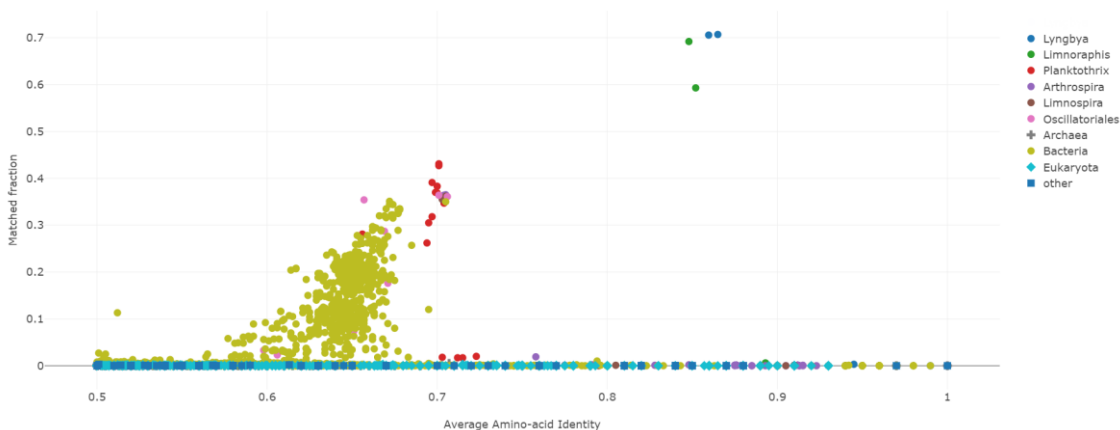


Figura 7. O eixo vertical do gráfico de dispersão mostra a fração correspondente de proteínas de consulta que correspondem às espécies, com espécies com valores de AAI mais altos (à direita) sendo mais estreitamente relacionadas. Os grupos taxonômicos são indicados por cores e estilos de marcadores.

Durante a análise, observou-se uma notável distância entre as espécies que não mantêm uma relação próxima com a espécie de consulta, refletida em baixas frações de correspondência. As contaminações puderam ser identificadas ao examinar valores discrepantes no gráfico de dispersão, que indicavam uma baixa correspondência e ausência de relação próxima com a espécie de consulta. As poucas espécies que apresentaram os valores mais elevados de AAI pertenciam à ordem Oscillatoriales, conforme já era esperado. No entanto, essa análise também

destacou a persistente questão da falta de genomas de referência de cianobactérias dessa ordem, o que limita a realização de análises mais representativas.

Foi identificado o sistema CRISPR-Cas, subtipo I-D, na cianobactéria *C. salinus* ALCB114379, acompanhado por dois fragmentos virais de 32.829 pb e 38.412 pb localizados periféricamente a esse sistema (Figura 8). A proximidade dessas sequências exógenas em relação ao sistema CRISPR-Cas sugere a atuação desse mecanismo de defesa contra bacteriófagos. O sistema CRISPR-Cas subtipo I-D é um sistema CRISPR-Cas classe 1 que foi identificado como um intermediário evolutivo entre os sistemas CRISPR-Cas tipo I e tipo III. Apresenta características de sistemas tipo I e tipo III, tornando-o um sistema híbrido com características únicas. O sistema subtipo I-D contém uma proteína efetora Cas3 e uma proteína efetora única chamada Cas10d, que possui um domínio de nuclease característico não encontrado em outros sistemas CRISPR. O sistema subtipo I-D exibe atividade de clivagem de dsDNA característica de outros sistemas CRISPR-Cas tipo I, mas também mostra clivagem excepcional de ssDNA, semelhante aos padrões de clivagem observados em sistemas tipo III (Lin *et al.*, 2020; Pattharaprachayakul *et al.*, 2020).

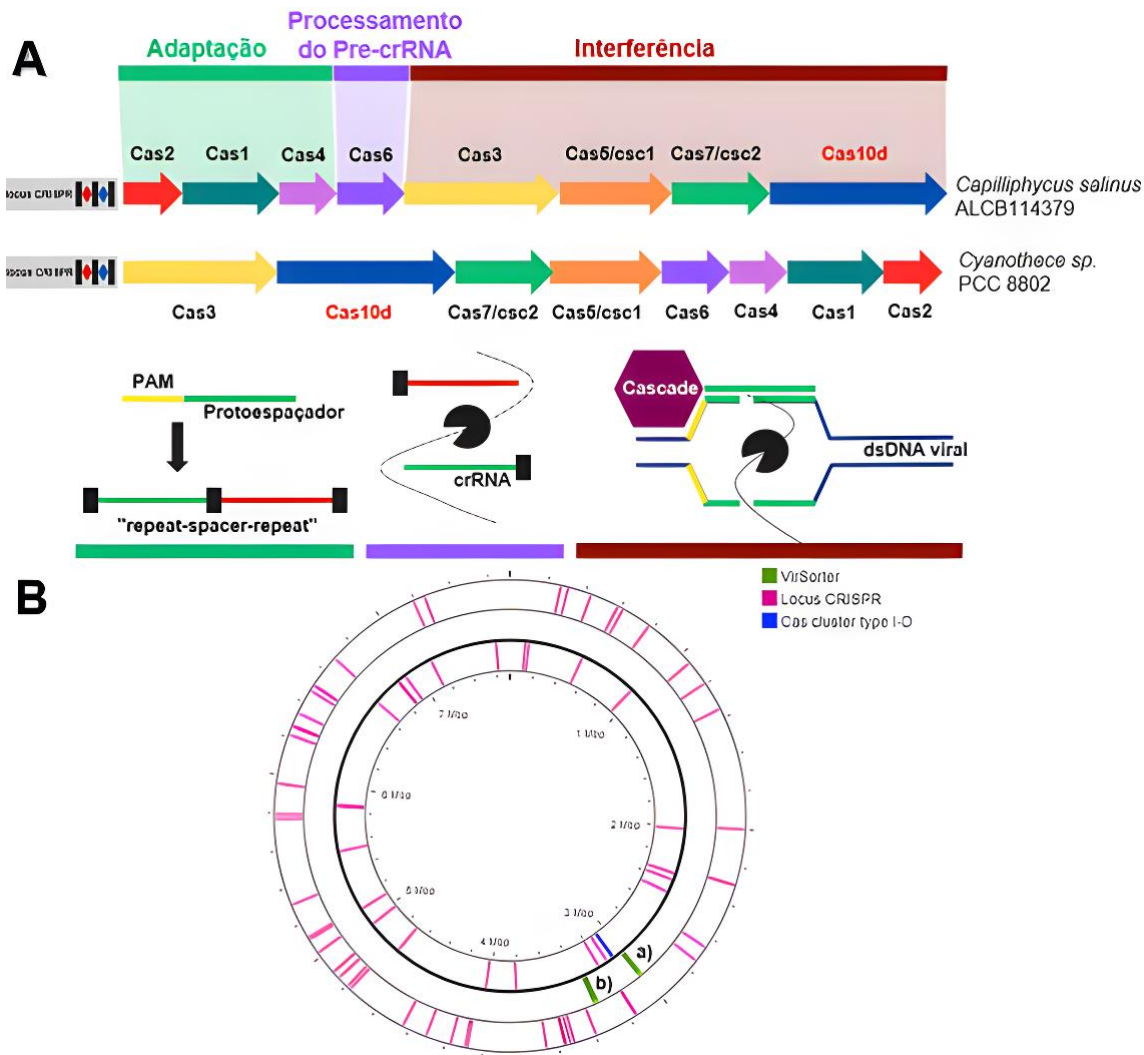


Figura 8. Sistema CRISPR na cianobactéria *C. salinus* ALCB114379. A. O sistema CRISPR-Cas envolve três etapas principais: Adaptação, que consiste na integração do protoespaçador no locus CRISPR; Processamento Pré-crRNA, que se refere à transcrição do pré-crRNA a partir dos espaçadores do locus CRISPR; e Interferência, responsável pelo reconhecimento e/ou clivagem do gene-alvo pelo crRNA, em cooperação com efetores multi-Cas ou efetores Cas únicos. B. Genoma circular da cianobactéria *C. salinus* ALCB114379 destacando, em rosa, os locus CRISPR ao redor do genoma, o sistema CRISPR-cas em azul e dois fragmentos virais, em verde e indicados por a) e b), adjacentemente ao complexo CRISPR-cas.

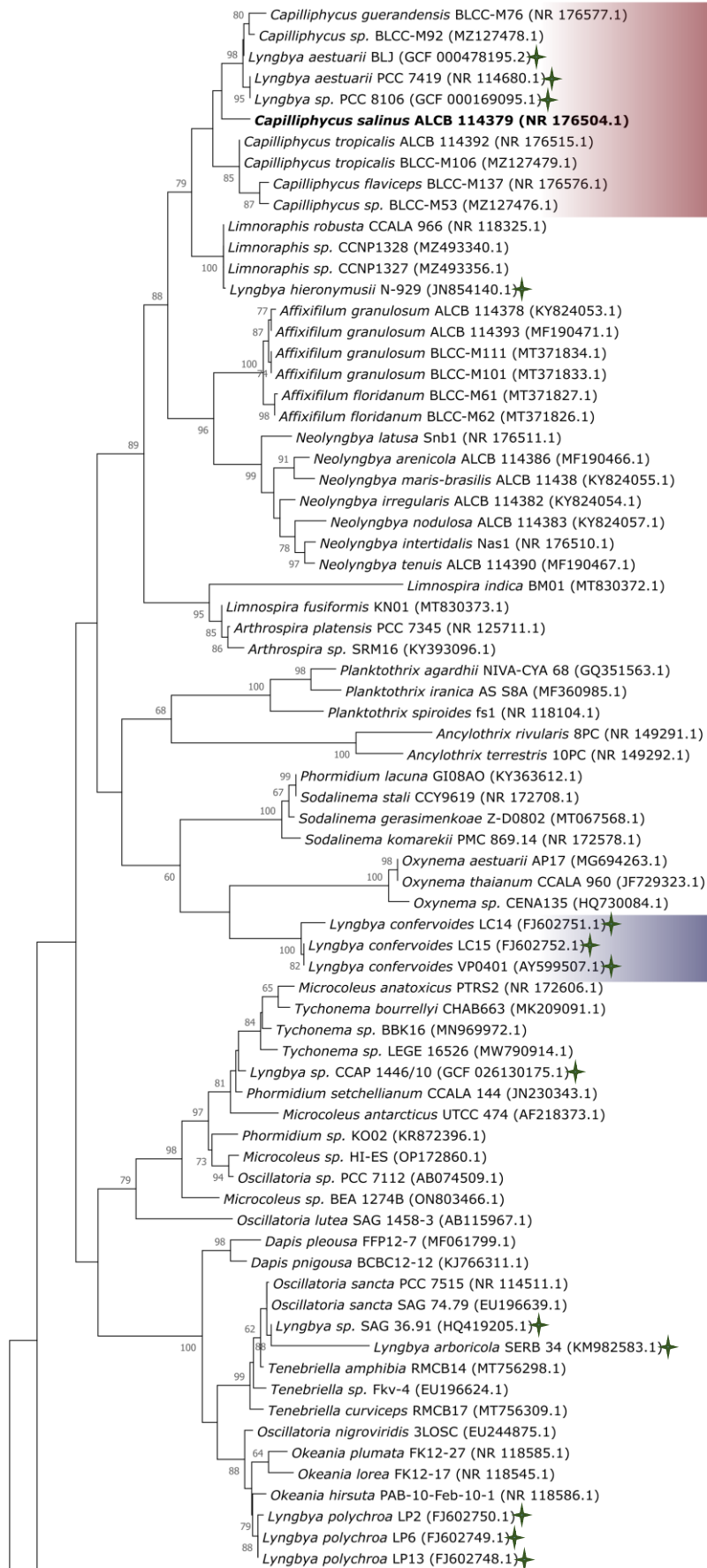
Os fragmentos virais identificados pela ferramenta VirSorter exibiram pontuações significativas de 0,86 e 0,847 para as sequências a) e b), respectivamente, ao serem comparados ao seu próprio banco de dados. Esses fragmentos, originários do genoma da *C. salinus* ALCB114379, foram submetidos a uma análise de BLAST no NCBI vírus (<https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/>). Os resultados sugerem que o fragmento a) pode pertencer a uma linhagem de *Pseudoanabaena* phage, enquanto o fragmento b) pode estar associado a uma linhagem de *Bacteriophage sp.*, sem especificações adicionais. Ambos os fragmentos foram submetidos a uma anotação

funcional usando a ferramenta Prokka, com base em seu banco de dados viral. O propósito desta análise foi identificar proteínas integrase nos fragmentos virais, sendo a integrase uma enzima envolvida na integração de profagos no genoma do hospedeiro (Colavecchio *et al.*, 2017). Diante da ausência de identificação dessa proteína, o genoma da *C. salinus* ALCB114379, com os dois fragmentos virais extraídos, foi novamente submetido a uma análise de qualidade, resultando em uma redução de 24,13% na contaminação do genoma. A presença de fragmentos virais inseridos no genoma, como profagos, ou sequências identificadas pelo sistema CRISPR-cas como parte do processo imunológico adaptativo, pode indicar contaminação no controle de qualidade de genomas. O sistema CRISPR-cas, um mecanismo de imunidade adaptativa específico de sequência, desempenha um papel crucial na identificação e no direcionamento de elementos genéticos estranhos, incluindo sequências virais, para degradação (Barrangou; Marraffini, 2014).

O genoma da *C. salinus* ALCB114379 é o primeiro registrado para o gênero até o momento, o que impossibilita análises comparativas com outras cianobactérias do mesmo gênero devido à falta de dados disponíveis. No entanto, essas análises fornecerão uma base de informações para futuros estudos envolvendo esse táxon.

5.2. Filogenia de *Capilliphycus salinus* ALCB114379

O genoma da cianobactéria *C. salinus* ALCB114379 apresentou duas cópias idênticas do gene de 16S rRNA com 1.485 pb, as quais foram extraídas do genoma com a ferramenta BarRNAP para análise filogenética (Figura 9). O gene de 16S rRNA é um marcador na ecologia microbiana, amplamente usado na identificação e estudos filogenéticos de bactérias. Apesar de sua ampla distribuição, apresenta limitações na resolução e sensibilidade para distinguir táxons próximos. O número de variantes do gene de 16S rRNA é cerca de 2,5 vezes maior do que o de espécies bacterianas, com sequências altamente divergentes em alguns grupos. Mesmo com homogeneização das cópias dos genes de rRNA dentro de um organismo, é essencial considerar que sequências de 16S rRNA da mesma espécie ou genoma podem variar significativamente (Clarridge, 2004; O'Callaghan *et al.*, 2021).



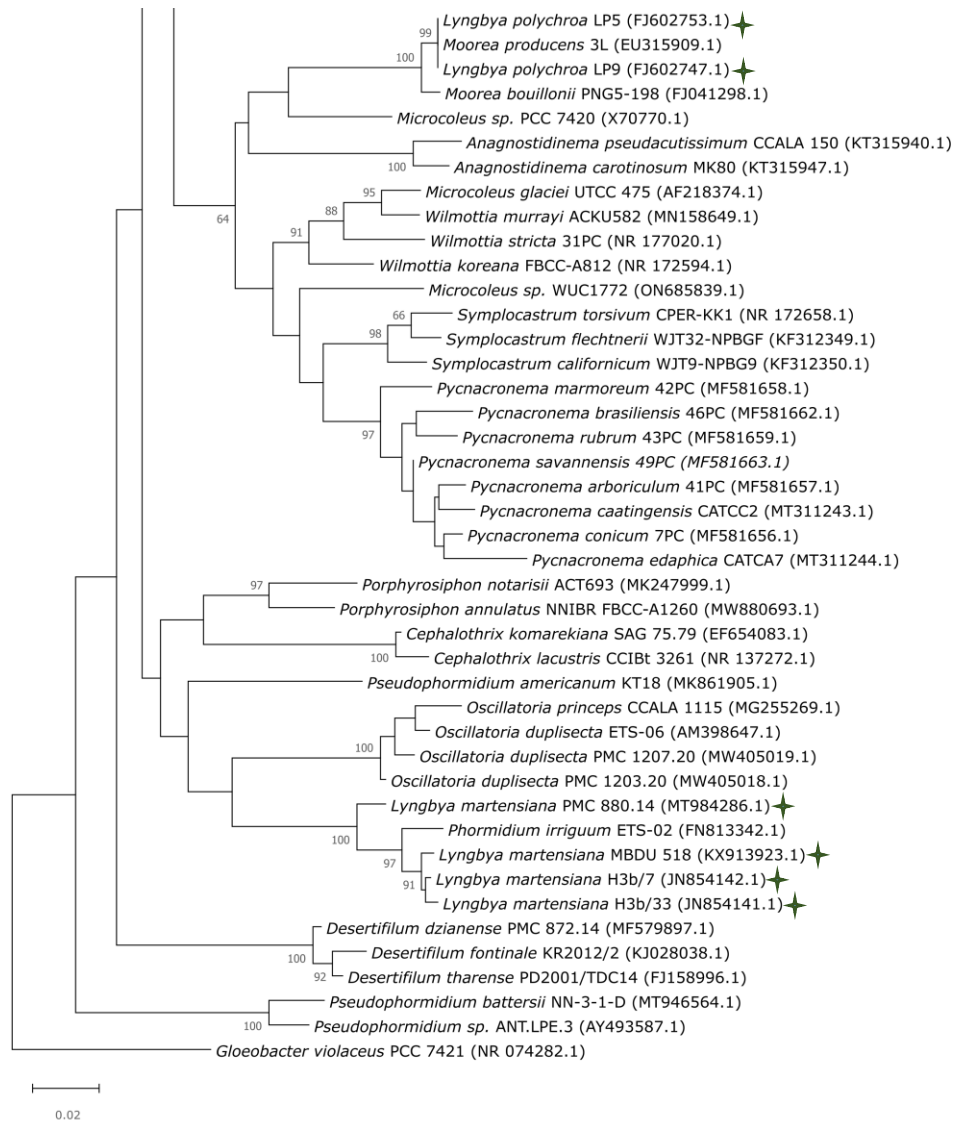


Figura 9. Árvore filogenética do gene de 16S rRNA gerado por Máxima Verossimilhança, demonstrando as relações evolutivas entre 118 linhagens de cianobactérias da ordem Oscillatoriales, distribuídas em diferentes famílias, incluindo 13 da família Colofasciculaceae, 3 da família Desertifilaceae, 26 da família Microcoleaceae, 46 da família Oscillatoriaceae e 29 da família Serenicapillariaceae. O clado da cianobactéria *C. salinus* ALCB114379 está destacado em vermelho, enquanto o clado da espécie tipo da *Lyngbya* está destacado em azul. Todas as linhagens do gênero *Lyngbya* estão sinalizadas com uma estrela de quatro pontas verde. Apenas os valores de reamostragem acima de 60 são indicados nos nós.

A análise filogenética da cianobactéria *C. salinus* ALCB114379 revelou que ela se agrupou com outras linhagens na mesma espécie, formando um clado bem sustentado. Um aspecto notável é que três linhagens de *Lyngbya* também se agruparam dentro do mesmo clado que *C. salinus* ALCB114379, apesar de pertencerem a uma família taxonômica diferente.

A identificação do clado da espécie tipo de *Lyngbya*, representada por *L. confervoides*, revela uma distância evolutiva significativa entre as diversas linhagens do mesmo gênero. Conforme apontado por Engene *et al.*, (2010), este gênero

demonstra ser altamente polifilético, incluindo táxons diversos sem uma relação direta entre si. Para compreender a inclusão de *Lyngbya* no clado da família Serenicapillariaceae, realizou-se uma análise de BLASTN, visando identificar a identidade entre as sequências de 16S rRNA e a quantidade de nucleotídeos divergentes (Tabela 2).

Tabela 2. Identidade do gene de 16S rRNA da família Serenicapillariaceae, com exceção das linhagens do gênero *Lyngbya* que pertencem à família Oscillatoriaceae. A porção superior contém os valores percentuais de identidade e a porção inferior o número de nucleotídeos divergentes. O mapa de calor indica que quanto mais verde, mais próxima de 100% é a identidade, e quanto mais vermelho, maior é a diferença nos nucleotídeos.

Linhagem	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. <i>Capilliphycus salinus</i> ALCB114379		98	98	98	98	97	98	98	98	98	98	98	97	97
2. <i>Capilliphycus tropicalis</i> ALCB14392	26		99	98	98	98	98	99	100	98	98	98	98	98
3. <i>Capilliphycus sp.</i> BLCC-M53	19	10		97	98	98	97	99	99	98	98	98	98	98
4. <i>Capilliphycus sp.</i> BLCC-M92	18	22	30		99	98	99	97	98	100	99	98	98	98
5. <i>Lyngbya aestuarii</i> PCC 7419	28	29	23	7		97	99	98	98	100	100	98	97	97
6. <i>Lyngbya hieronymusii</i> N-929	35	32	22	24	36		97	98	98	98	97	100	100	100
7. <i>Capilliphycus guerandensis</i> BLCC-M76	21	24	32	9	10	28		97	98	99	99	97	97	97
8. <i>Capilliphycus flaviceps</i> BLCC-M137	16	10	6	26	21	21	28		99	98	98	98	98	98
9. <i>Capilliphycus tropicalis</i> BLCC-M106	16	0	10	22	17	18	24	10		98	98	98	98	98
10. <i>Lyngbya aestuarii</i> BLJ	24	28	25	5	5	33	8	21	17		100	98	98	98
11. <i>Lyngbya sp.</i> PCC 8106	29	29	23	7	3	36	10	21	17	5		98	97	97
12. <i>Limnoraphis robusta</i> CCALA 966	26	29	22	25	27	1	28	21	18	25	28		100	100
13. <i>Limnoraphis sp.</i> CCNP1327	34	31	22	25	34	1	28	21	18	32	34	0		100
14. <i>Limnoraphis sp.</i> CCNP1328	34	28	22	25	34	1	28	21	18	32	34	0	0	

A linhagem 1, representada por *C. salinus* ALCB114379, exibe uma identidade notavelmente alta com outras espécies de *Capilliphycus*, variando em torno de 98-99%. Isso sugere uma estreita relação genética entre essas linhagens, apoiando sua

inclusão no mesmo gênero. Os resultados também revelam que *Lyngbya hieronymusii* N-929 exibe 100% de identidade em suas sequências do gene de 16S rRNA com os gêneros *Limnoraphis*. Outras linhagens do gênero *Lyngbya* apresentaram consideráveis valores de identidade em comparação com as de *Capilliphycus*. A cianobactéria *Capilliphycus* sp. BLCC-M92 demonstrou 100% de identidade no gene de 16S rRNA com a *Lyngbya aestuarii* BLJ e 99% de identidade com a *Lyngbya aestuarii* PCC 7419. Nas tentativas de elucidar a natureza polifilética de *Lyngbya*, surgiram vários novos gêneros, como *Moorea*, *Limnoraphis*, *Okeania* e *Microseira* (Cordeiro *et al.*, 2020).

No geral, todos os valores registrados foram maiores ou iguais a 97%, sugerindo a necessidade de uma abordagem taxonômica baseada em marcadores genéticos adicionais. Atualmente, o método mais promissor para resolver essas questões na sistemática bacteriana é a inclusão de análises considerando todo o genoma (Lalucat *et al.*, 2020; Maguvu; Bezuidenhout, 2021).

Na árvore filogenômica gerada pelo GTDB-tk (Figura 10), com base na análise de 120 marcadores genéticos encontrados no domínio Bacteria, observa-se o agrupamento da cianobactéria *C. salinus* ALCB114379 em um clado com as linhagens *Lyngbya* sp. PCC 8106, *Lyngbya aestuarii* BLJ e *Limnoraphis robusta* CS-951. Esse resultado corrobora com a análise filogenética com o gene de 16S rRNA, na qual esse agrupamento também foi observado.

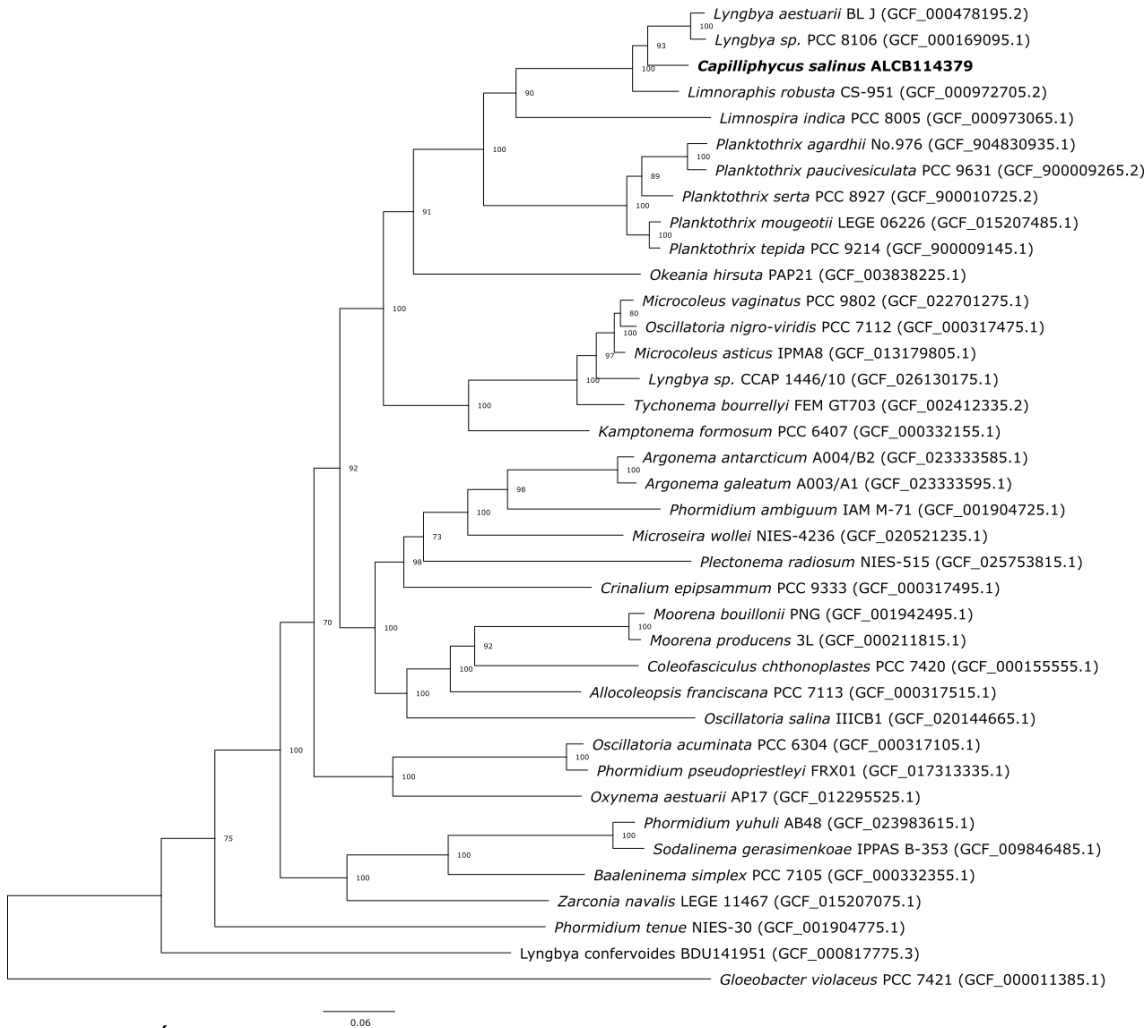


Figura 10. Árvore filogenômica com GTDB-tk, utilizando 120 marcadores para inferir a posição evolutiva da cianobactéria *C. salinus* ALCB114379 dentro da ordem Oscillatoriales. Apenas os valores de reamostragem acima de 60 são indicados nos nós.

Uma análise adicional envolve a Identidade Média de Nucleotídeos (ANI) e a Identidade Média de Aminoácidos (AAI), dois parâmetros usados na identificação taxonômica, especialmente em genomas procarióticos. O ANI mede a similaridade genômica em nível de nucleotídeo entre as regiões codificantes de dois genomas, enquanto o AAI mede a similaridade média das sequências de aminoácidos nos proteomas de dois organismos. Ambos, ANI e AAI, complementam-se na identificação taxonômica, fornecendo diferentes perspectivas sobre a relação genética entre organismos (Rosselló-Mora, 2005).

O ANI é mais indicado para comparar genomas intimamente relacionados, sendo menos afetado por diferenças no conteúdo genético e oferecendo uma medida mais precisa de similaridade. Em contraste, o AAI é mais apropriado para comparar genomas distantes, pois é menos afetado pelo ruído evolutivo, proporcionando uma medida mais precisa de genes ortólogos compartilhados.

Valores de ANI acima de 95% geralmente são utilizados para definir espécies de bactérias e arqueas. Os valores de ANI entre representantes do mesmo gênero variam entre 78% e 95%. Quando os valores de ANI ficam abaixo de 78%, é altamente provável que pertençam a gêneros diferentes. Os valores de AAI seguem esses parâmetros de classificação baseados no ANI (Parks *et al.*, 2019).

A Tabela 3 apresenta os valores de Identidade Média de Nucleotídeos (ANI) e Identidade Média de Aminoácidos (AAI) entre seis linhagens de cianobactérias: *Capilliphyucus salinus* ALCB114379, *Limnoraphis robusta* CS-951, *Lyngbya sp.* PCC 8106, *Lyngbya aestuarii* BLJ, *Lyngbya confervoides* BDU141951 e *Lyngbya sp.* CCAP 144610.

Tabela 3. Similaridade a nível de genoma a partir da Identidade Média de Nucleotídeos (ANI), porção superior em verde, e da Identidade Média de Aminoácidos (AAI), porção inferior em vermelho.

Linhagem	1	2	3	4	5	6
1. <i>Capilliphyucus salinus</i> ALCB114379		0,82	0,83	0,84	0,67	0,75
2. <i>Limnoraphis robusta</i> CS-951	0,83		0,83	0,83	0,66	0,74
3. <i>Lyngbya sp.</i> PCC 8106	0,85	0,83		0,92	0,67	0,70
4. <i>Lyngbya aestuarii</i> BLJ	0,86	0,83	0,92		0,67	0,75
5. <i>Lyngbya confervoides</i> BDU141951	0,54	0,54	0,54	0,54		0,68
6. <i>Lyngbya sp.</i> CCAP 144610	0,62	0,61	0,62	0,61	0,55	

Observa-se que *Capilliphyucus salinus* ALCB114379, *Limnoraphis robusta* CS-951, *Lyngbya sp.* PCC 8106 e *Lyngbya aestuarii* BLJ formam um clado, com valores de ANI e AAI variando de 82% a 92%, sugerindo uma estreita relação taxonômica. No entanto, é importante notar que o gênero *Lyngbya* é descrito como polifilético, refletindo suas diversas origens evolutivas. Isso está de acordo com a distribuição dispersa de *Lyngbya confervoides* BDU141951 e *Lyngbya sp.* CCAP 144610 na árvore filogenômica. Em particular, a espécie do tipo *Lyngbya confervoides* BDU141951 tem um valor de ANI de 68% e AAI de 54% quando comparada com outras linhagens do mesmo gênero.

A amostragem limitada pode comprometer a resolução das classificações taxonômicas, destacando, assim, a necessidade premente de obter dados genômicos mais abrangentes para aprimorar essas classificações. Apesar desse

desafio, os resultados indicam que a família Serenicapillariaceae forma um grupo coeso, sustentando a classificação adequada de *C. salinus* ALCB114379, separando-a do gênero *Lyngbya*.

5.3. Identificação de agrupamentos gênicos associados ao metabolismo secundário

5.3.1. Aminoácidos do tipo micospolina (MAAs)

A anotação funcional do genoma da cianobactéria *C. salinus* ALCB114379 detectou 6.230 potenciais genes codificantes de proteínas, sendo que 3.577 (57,41%) foram classificadas como proteínas hipotéticas (HPs - Hypothetical Proteins). Com base nessa anotação, 2.194 (35,2%) foram categorizados em subsistemas biológicos, enquanto a maioria foi associada a processos biológicos ainda não descritos com base no banco de dados KEGG (*Kyoto Encyclopedia of Genes and Genomes*, release 108.1) (Figura 11). Essa proporção indica uma notável diversidade genética não explorada no genoma da *C. salinus* ALCB114379, abrindo oportunidades significativas para o estudo e descoberta de novas vias metabólicas envolvidas na biossíntese de produtos naturais.

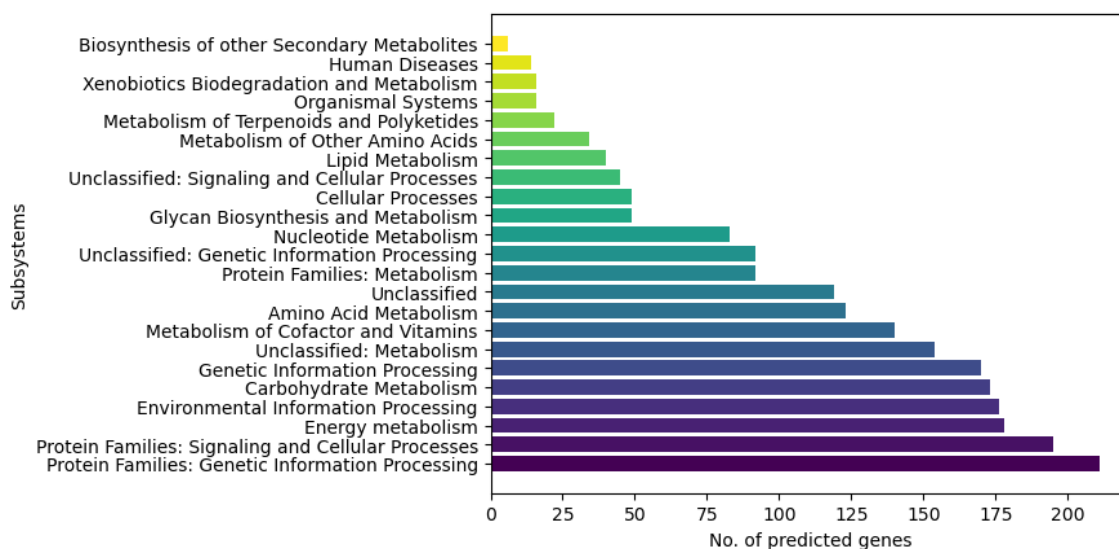


Figura 11. Distribuição dos genes preditos em subsistemas biológicos na cianobactéria *C. salinus* ALCB114379 pela ferramenta BlastKOALA.

A partir dessa categorização, selecionou-se o subsistema relacionado à biossíntese de metabólitos secundários para avaliação. Nesse subsistema, o único BGC que estava completamente anotado e apresentava um conjunto de genes canônicos *mysABCD* bem definidos foi o relacionado aos MAAs.

Os genes *mysA*, *mysB*, *mysC* e *mysD*, pertencentes ao BGC de MAAs na *C. salinus* ALCB114379 passaram por uma curadoria manual utilizando a ferramenta Artemis. Essa análise possibilitou visualizar os genes dispostos em um mesmo agrupamento gênico. A presença de um agrupamento de genes biossintéticos concatenados sugere funcionalidade, uma vez que permite a coexpressão de genes envolvidos na mesma via metabólica secundária (Crits-Christoph *et al.*, 2021; Xia *et al.*, 2022). A organização estruturada de BGCs relacionados a MAAs, contendo *mysD*, em genomas de cianobactérias foi correlacionada com uma elevada probabilidade de produção, uma vez que esse arranjo é incomum no filo, estando presente em apenas 11,8% de todas as sequências genômicas codificadas (Dextro *et al.*, 2023).

Uma comparação usando as sequências de aminoácidos (aa) de cada gene, permitiu identificar organismos que possuem os genes para a biossíntese de MAAs mais próximos evolutivamente da *C. salinus* ALCB114379. Em seguida, as sequências foram submetidas a uma análise de domínios conservados com a ferramenta InterPro (Tabela 4).

Tabela 4. Informações sobre as enzimas codificadoras dos genes *mys* anotados da cianobactéria *C. salinus* ALCB114379.

Gene	Aminoácidos (aa)	Função deduzida	Organismo similar mais próximo (n° acesso NCBI)	Identidade (%)	Similaridade (%)
<i>mysA</i>	412	2-epi-5-epi-valiolone synthase-like	<i>Lyngbya aestuarii</i> (WP_023068561.1)	94,66	97,57
<i>mysB</i>	276	Methyltransferase	<i>Lyngbya sp.</i> PCC 8106 (WP_009783738.1)	92,39	96,38
<i>mysC</i>	455	ATP-grasp ligase	<i>Lyngbya aestuarii</i> (WP_023068529.1)	90,33	93,85
<i>mysD</i>	365	d-Alanine ligase	<i>Lyngbya aestuarii</i> (WP_023068556.1)	82,83	88,83

O gene *mysA* codifica uma fosfato ciclase, que atua sobre a sedoheptulose-7-fosfato da via das pentoses fosfato ou no 3-desoxi-d-arabino-2-heptulosonato-7-fosfato da via do chiquimato para produzir 3-desidroquinato, um precursor na biossíntese de aminoácidos aromáticos (Arsin *et al.*, 2023). A identificação do domínio conservado sedoheptulose-7-fosfato no gene *mysA* da *C. salinus*

ALCB114379 sugere que a biossíntese para a produção de MAAs segue pela via das pentoses fosfato.

O agrupamento de genes para a biossíntese de MAAs na *C. salinus* ALCB114379 demonstrou identidade e similaridade significativas com o mesmo conjunto de genes encontrado em cianobactérias do gênero *Lyngbya*, apresentando uma cobertura de 100% para todos os genes comparados e número de aminoácidos semelhantes. As cianobactérias do gênero *Lyngbya* apresentam mecanismos de fotoproteção proporcionados por moléculas absorventes/protetores de UV, como os MAAs e escitonemina (Fuentes-Tristan *et al.*, 2019). No entanto, é importante destacar que as cianobactérias da ordem Oscillatoriales são extremamente sub-representadas na avaliação dos MAAs. Apenas algumas linhagens, como *Lyngbya* e *Microcoleus*, foram descritas como produtoras, sendo que o BGC organizado foi encontrado em apenas 21% dos 300 conjuntos de genomas de Oscillatoriales disponíveis e depositados no NCBI (Dextro *et al.*, 2023).

O agrupamento de genes relacionados à produção de MAAs, provenientes de sete cianobactérias reconhecidamente produtoras de algum análogo de micosporina, foi extraído do banco de dados NCBI e posteriormente re-anotado utilizando a ferramenta Prokka. Todas as sequências, incluindo a da *C. salinus* ALCB114379, foram submetidas a uma análise de sintenia para demonstrar, comparativamente, a co-localização e identidade entre os genes (Figura 12). A análise de sintenia é um processo que visa identificar a conservação de genes homólogos e a ordem dos genes entre genomas de diferentes espécies. Este método oferece uma estrutura na qual a conservação de genes homólogos e a ordem dos genes podem ser identificadas entre genomas de espécies distintas. A sintenia pode ser descrita como a co-localização física de loci genéticos no mesmo cromossomo, seja dentro de um indivíduo ou de uma espécie (Liu; Hunt; Tsai, 2018). Em agrupamentos de genes biossintéticos, as relações sintênicas podem revelar a conservação das vias biossintéticas entre diferentes espécies. Os genes sintênicos são ortólogos localizados nesses fragmentos, portanto, frequentemente compartilham funções semelhantes (Ahmad *et al.*, 2023; Cheng *et al.*, 2012; Evdokias *et al.*, 2021).

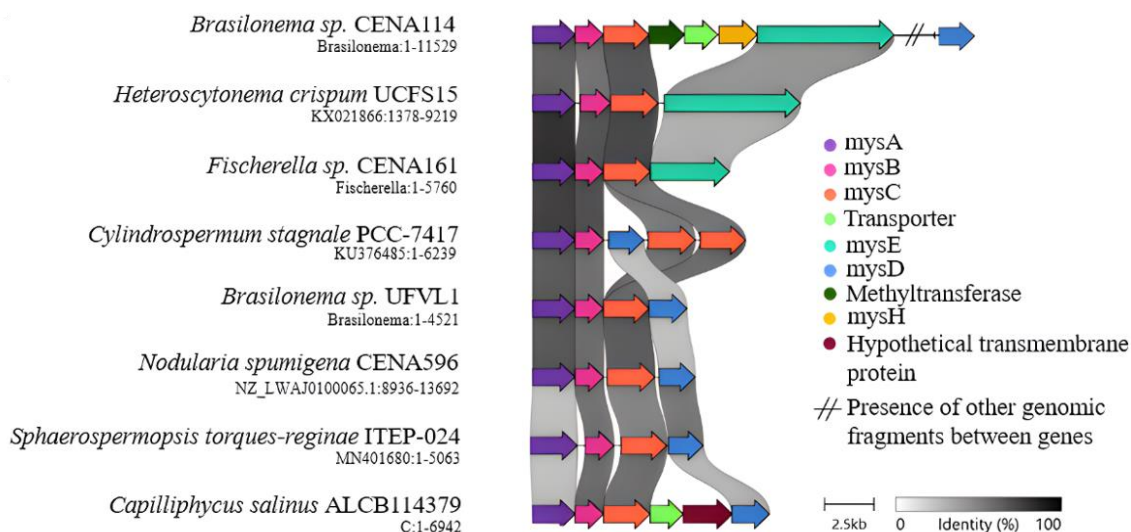


Figura 12. Comparação de identidade do BGC (*mys*) relacionado à produção de MAAs entre *C. salinus* ALCB114379 e linhagens conhecidas de cianobactérias produtoras dessas moléculas.

A análise de sintenia evidenciou os genes *mys* responsáveis pela biossíntese de MAAs, apresentando uma disposição sequencial em diversas linhagens, com exceção da *Cyndrospermum stagnale* PCC-7417. Nesta linhagem, ocorreu uma inversão na ordem dos genes *mysC* e *mysD*, além da presença de uma relação paráloga entre os genes *mysC* no agrupamento. Assim como a cianobactéria *Brasilonema* sp. CENA114, a *C. salinus* ALCB114379 apresentou potenciais genes codificantes entre os genes do agrupamento para a biossíntese de MAAs.

O gene subsequente ao *mysC* da cianobactéria *C. salinus* ALCB114379, teve uma correspondência de 100% de cobertura e 82,77% de identidade com uma família proteica de transportadores DMT da cianobactéria *Lyngbya aestuarii* (Nº de acesso NCBI: WP_040484460.1). Esta família de transportadores DMT, integrante da superfamília DMT, desempenha um papel crucial na exportação de uma ampla gama de substratos, incluindo medicamentos, metabólitos, aminoácidos e açúcares (Jack; Yang; Saier, 2001). A análise de domínio conservado categorizou o gene em uma superfamília de transportadores de drogas/metabólitos (DMT). O gene subsequente ao possível transportador, passou sequencialmente pelas mesmas análises, sendo correspondido por um simportador de sódio/prolina, da linhagem *Lyngbya* sp. PCC 8106 (Nº de acesso NCBI: WP_009783741.1), compartilhando 99% de cobertura e 92,39% de identidade com seu homólogo. O simportador sódio/prolina, também conhecido como PutP, é uma proteína transportadora encontrada em bactérias, desempenhando um papel crucial na captação de L-

prolina, um aminoácido de importância central no metabolismo celular, na osmoproteção e nas interações bactéria-hospedeiro (Olkhova *et al.*, 2011).

Uma vez realizada a caracterização genômica do agrupamento gênico da microsporina, conduziu-se um experimento qualitativo utilizando espectrofotometria de absorção no ultravioleta para avaliar o potencial de produção de MAAs pela cianobactéria *C. salinus* ALCB114379 (Figura 13).

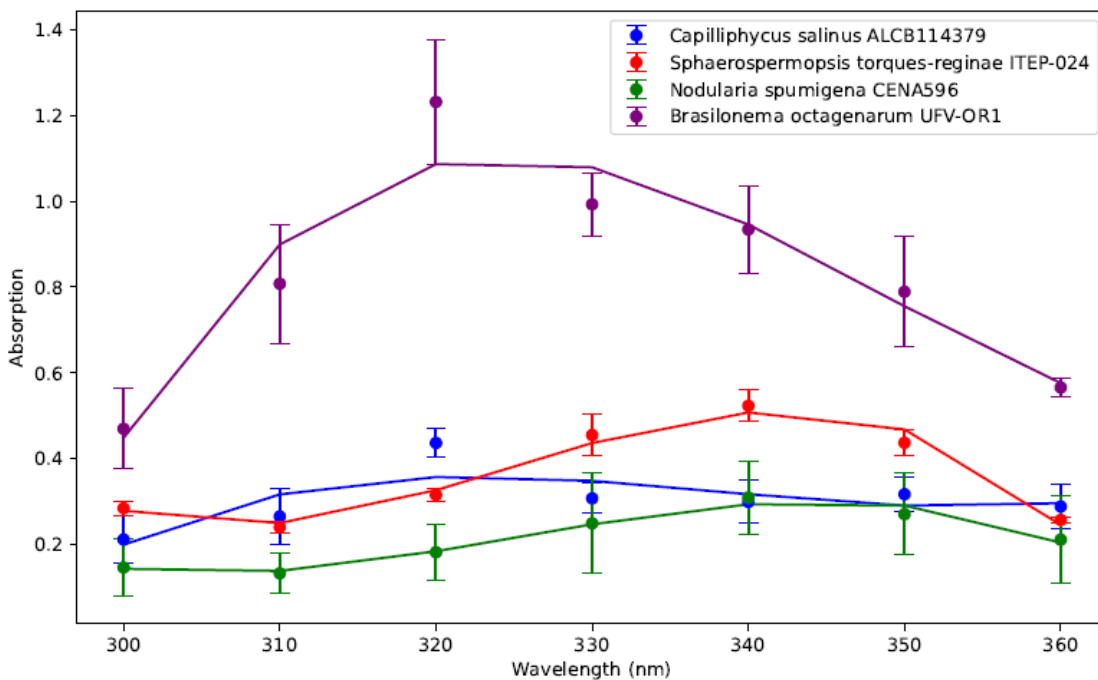


Figura 13. Absorbância detectada no intervalo de comprimento de onda de 300 a 360 nm no extrato da cianobactéria cultivada em luz fluorescente.

A cianobactéria *C. salinus* ALCB114379 apresentou um pico de absorbância em 320 nm, assim como observado em *Brasilonema octagenarum* UFV-OR1, porém em um nível de absorbância significativamente menor. Dextro *et al.* (2023) conduziram uma análise da produção constitutiva e exposta a UV de MAAs em *B. octagenarum* UFV-OR1 por meio de Cromatografia Líquida Acoplada à Espectrometria de Massas (HPLC/MS-MS), identificando a presença de palitina. Este MAA específico demonstrou um pico de absorção ideal em 320 nm. As cianobactérias *Sphaerospermopsis torques-reginae* ITEP-024 e *Nodularia spumigena* CENA596 apresentaram os maiores picos de absorbância na faixa entre 0,3 e 0,5, onde também se encontra o pico de absorbância mais pronunciado da *C. salinus* ALCB114379. Dessa forma, a combinação da análise qualitativa do extrato, juntamente com a presença do agrupamento de genes colineares relacionados à

biossíntese de MAAs, indica o potencial da *C. salinus* ALCB114379 para produzir essas moléculas.

Na análise metabolômica direcionada por HPLC/MS-MS, MAAs específicos foram identificados no extrato de *C. salinus* ALCB114379.

Tabela 5. Informações sobre padrões de fragmentação, massas medidas de íons moleculares $[M+H]^+$ e massas exatas da shinorina e palitina. Os padrões de fragmentação encontrados na *C. salinus* ALCB114379 estão em negrito. Adaptado de Whitehead; Karentz; Hedges, (2001).

MAAs	Shinorina	Palitina
$[M+H]^+$	333	245
Padrão de fragmentação (m/z)	137, 168, 185, 186 , 197, 211, 230 , 241, 255	137 , 150, 155, 162, 184, 186 , 199, 209, 230
Massa	332,12	244,11

A análise confirmou a presença de palitina, identificando os picos de fragmentação com m/z 137 e 186, e também shinorina, com os picos m/z 186 e 230 (Figura 14).

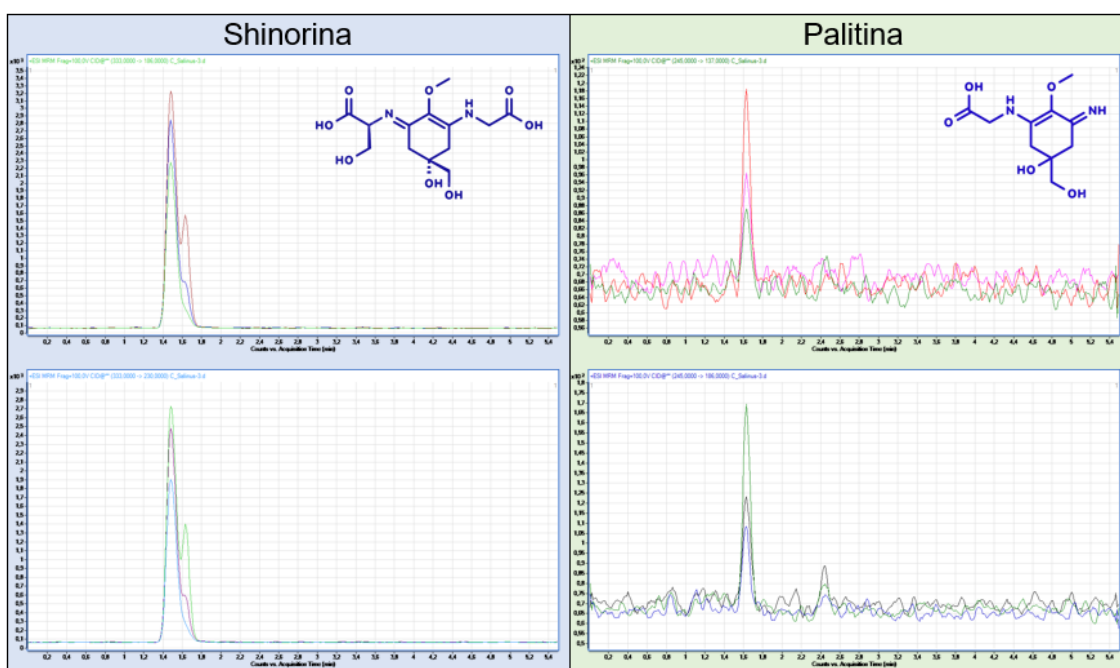


Figura 14. Cromatogramas indicando a presença de shinorina e palitina no extrato da *C. salinus* ALCB114379.

A presença desses MAAs distintos demonstra uma redundância nas estratégias fotoprotetoras adotadas pela *C. salinus* ALCB114379, ressaltando a complexidade da resposta adaptativa das cianobactérias às condições ambientais.

Diferentes MAAs geralmente têm espectros de absorção distintos, permitindo que as cianobactérias absorvam uma variedade de comprimentos de onda ultravioleta (UV). Ao produzir vários tipos, as cianobactérias podem alcançar um espectro mais amplo de proteção UV, aumentando sua capacidade de lidar com diversas condições de luz. Essa diversidade na produção de MAAs destaca a plasticidade adaptativa das cianobactérias, permitindo-lhes otimizar a absorção de luz UV de acordo com as características específicas do ambiente em que estão inseridas (Geraldes; Pinto, 2021).

5.3.2. Microciclâmica

Uma segunda abordagem empregada para a prospecção de possíveis agrupamentos gênicos associados à produção de metabolitos secundários, foi pela ferramenta AntiSMASH. *C. salinus* ALCB114379 apresentou sete possíveis regiões gênicas relacionadas à produção de metabolitos secundários, apresentando 13% de identidade para agrupamentos de policetídeo sintase (T1PKS), 25% para peptídeos sintetases não ribossomais (NRPSs), 33% para peptídeos sintetizados ribossomicamente e modificados pós-tradução (RiPPs), 70% para metalóforos peptídicos não ribossômicos (NRP-metallophore), 88% para cianobactinas, e uma similaridade irrisória para agrupamentos gênicos produtores de terpenos. Foram selecionadas apenas as regiões com potencial para metabolismo secundário acima de 70% para análises de curadoria manual e validação da presença dos agrupamentos preditos. No genoma da *C. salinus* ALCB114379, não foram identificados os agrupamentos gênicos preditos para NRP-metallophore. No entanto, em relação aos BGCs de cianobactinas, observou-se uma correspondência significativa, considerando a presença e co-localização, com os genes biossintéticos da microciclâmica.

O agrupamento gênico de microciclâmica identificado na *C. salinus* ALCB114379 apresentou semelhanças com o descrito em *Microcystis* por Ziemert *et al.*, (2008), com exceção da ausência da ORF2 (Figura 15).

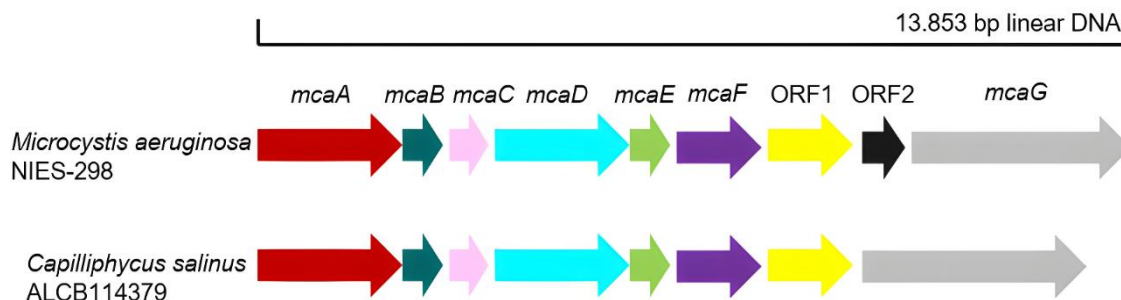


Figura 15. Comparação dos BGCs da microciclâmica entre a *M. aeruginosa* NIES-298 e a *C. salinus* ALCB114379.

Para verificar a falta desse elemento gênico na *C. salinus* ALCB114379 apontada pela anotação no AntiSMASH, utilizou-se a ferramenta ORFfinder. Nessa análise, identificou-se uma ORF com 32 aminoácidos, denominada ORF1C, localizada entre a ORF1 e o gene *mcaG*, sem correspondência à ORF2 descrita em *Microcystis*. Adicionalmente, identificou-se uma ORF chamada ORF3, composta por 219 aminoácidos, presente tanto na *M. aeruginosa* NIES-298 quanto na *M. aeruginosa* PCC 7806. Essa ORF se sobrepõe parcialmente à ORF2 identificada nas mesmas linhagens e precede o gene *mcaG*. Embora a ORF3 não apresente similaridade significativa com o gene *mcaG*, possui 100% de identidade e 49% de cobertura com a ORF2, provavelmente devido à sobreposição observada.

Realizou-se uma anotação funcional com a ferramenta Prokka para identificar esses elementos gênicos, tanto pela anotação do AntiSMASH quanto pela anotação do ORFfinder. O resultado revelou uma classificação de HP para a ORF3, e não foi anotada a ORF2 em *Microcystis*. A ferramenta também não identificou a ORF1C da *C. salinus* ALCB114379.

A não identificação da ORF1C da *C. salinus* ALCB114379 sugere a possibilidade de essa sequência estar associada a um RNA não codificante. Em relação à ORF2 de *Microcystis*, pode ter ocorrido um erro de leitura na ferramenta devido à sobreposição parcial com a ORF3 ou não houve correspondência contra o banco de dados do Prokka. A ORF3, classificada como HP, é um potencial gene codificante com domínio conservado para nucleotidiltransferase.

Com base nesse entendimento sobre as regiões intergênicas entre os genes *mcaF* e *mcaG*, os genes do BGC da microciclâmica na *C. salinus* ALCB114379 e na *M. aeruginosa* NIES-298 foram extraídos e submetidos a uma análise de quantidade de aminoácidos, domínio conservado e relação de homologia por meio de identidade, similaridade e cobertura (Tabela 6).

Tabela 6. Funções deduzidas nos agrupamentos de genes biossintéticos de microciclâmica para as cianobactérias *M. aeruginosa* NIES-298 e a *C. salinus* ALCB114379.

Genes	No. de aminoácidos	Função deduzida	% Identidade/Similaridade/Cobertura
<i>mcaA</i>			
<i>M. aeruginosa</i> NIES-298	657	Subtilisin-like protease	83,99/91,55/100
<i>C. salinus</i> ALCB114379	673		
<i>mcaB</i>			
<i>M. aeruginosa</i> NIES-298	83	Conserved hypothetical protein	95,18/98,80/100
<i>C. salinus</i> ALCB114379	84		
<i>mcaC</i>			
<i>M. aeruginosa</i> NIES-298	80	Conserved hypothetical protein	95/96,25/100
<i>C. salinus</i> ALCB114379	81		
<i>mcaD</i>			
<i>M. aeruginosa</i> NIES-298	776	Adenylation/heterocyclization	97,29/97,81/100
<i>C. salinus</i> ALCB114379	788		
<i>mcaE</i>			
<i>M. aeruginosa</i> NIES-298	74	Microcyclamide precursor protein	68,09/68,09/93
<i>C. salinus</i> ALCB114379	90		
<i>mcaF</i>			
<i>M. aeruginosa</i> NIES-298	321	Conserved hypothetical protein	64,71/76,78/96
<i>C. salinus</i> ALCB114379	321		
ORF1			
<i>M. aeruginosa</i> NIES-298	267	Unknown	83,99/86,48/100
<i>C. salinus</i> ALCB114379	285		
<i>mcaG</i>			
<i>M. aeruginosa</i> NIES-298	1351	Thiazoline oxidase/subtilisin-like protease	77,14/81,53/97
<i>C. salinus</i> ALCB114379	1379		

A quantidade de aminoácidos entre as cianobactérias *M. aeruginosa* NIES-298 e *C. salinus* ALCB114379 não apresentou diferença significativa. Os valores de identidade e similaridade, juntamente com uma cobertura acima de 93%, indicam relações de homologia entre os genes do BGC da microciclâmica, sugerindo a possibilidade de compartilharem a mesma função. Para verificar a ordem dos genes BGC da microciclâmica na *C. salinus* ALCB114379, realizou-se uma análise de sintonia com a *M. aeruginosa* NIES-298 e *M. aeruginosa* PCC 7806, com a adição da linhagem *M. aeruginosa* SPC-777 encontrada no banco de dados NCBI (Figura 16).

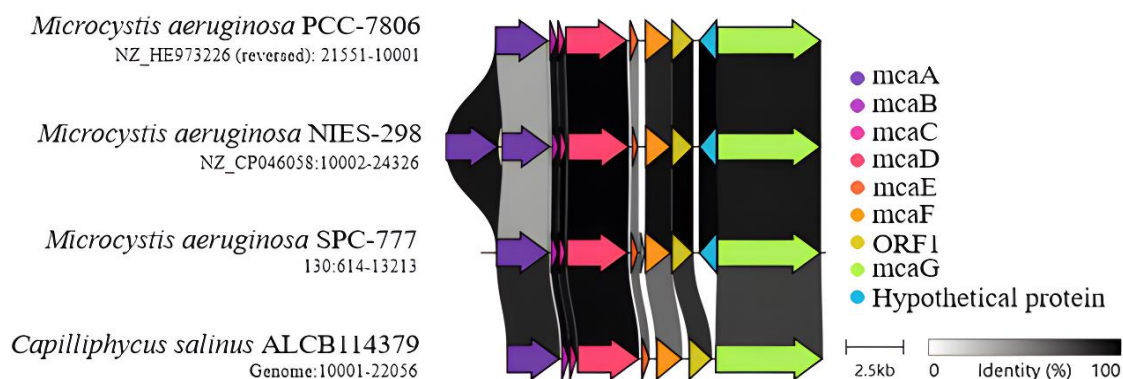


Figura 16. Comparação de identidade do BGC relacionado à produção de microciclâmica (*mca*) entre *C. salinus* ALCB114379 e linhagens de cianobactérias conhecidas produtoras dessa molécula do gênero *Microcystis*.

Com exceção da ORF1C, ORF2 e da ORF3, classificada como HP, que não possuem relevância para a biossíntese da microciclâmica, os demais genes que compõem o BGC dessa molécula foram encontrados sequencialmente e co-localizados quando comparados aos agrupamentos nas *Microcystis*, na cianobactéria *C. salinus* ALCB114379. Apenas a *M. aeruginosa* NIES-298 apresentou uma duplicação do gene *mcaA*. No entanto, a expressiva distância taxonômica entre essas duas cianobactérias, em que *M. aeruginosa* NIES-298 pertence à ordem Chroococcales e a *C. salinus* ALCB114379 pertence à ordem Oscillatoriales, torna desafiador explicar a origem desse agrupamento gênico na cianobactéria *C. salinus* ALCB114379.

A compatibilidade e funcionalidade de um agrupamento de genes biossintéticos no genoma são influenciadas por vários fatores, incluindo padrões de códons. O viés de uso de códon (CUB) refere-se à preferência por certos códons em detrimento de outros para o mesmo aminoácido. O padrão de uso de códons pode impactar a expressão genética, a estrutura e a função das proteínas (Bailly-Bechet *et al.*, 2006; Parvathy; Udayasuriyan; Bhadana, 2022; Zhao *et al.*, 2021). Portanto, foi realizada uma comparação entre a fração total de códons do genoma da cianobactéria *C. salinus* ALCB114379 e seu BGC da microciclâmica para identificar se há aporte gênico para a expressão desse agrupamento (Figura 17).

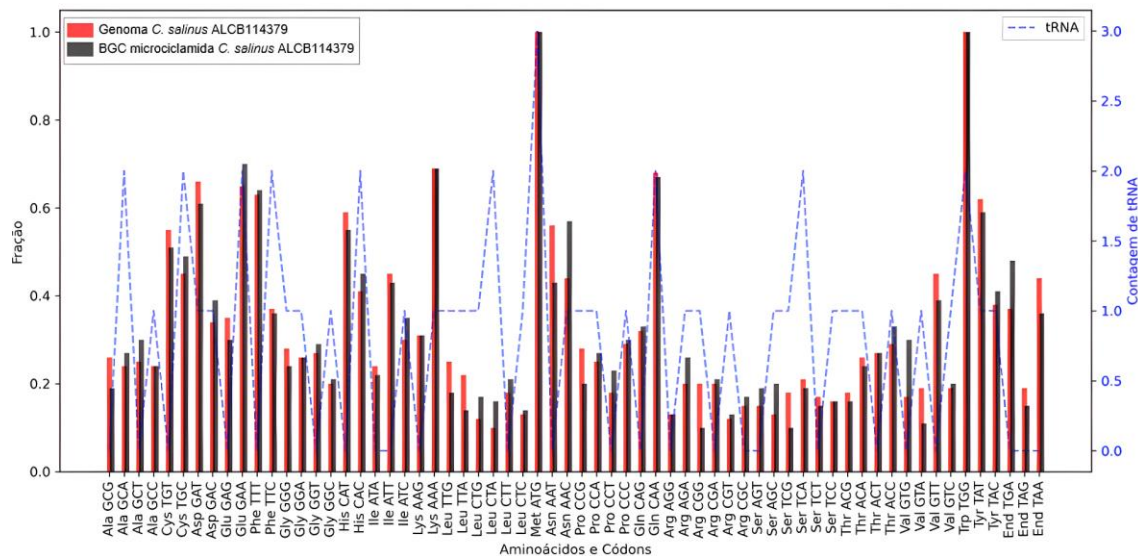


Figura 17. Fração de códon entre o genoma da *C. salinus* ALCB114379 e seu BGC da microciclâmica. As linhas pontilhadas em azul representam os tRNAs pareados com seus respectivos anti- códon.

O padrão de códon no BGC da microciclâmica e em *C. salinus* ALCB114379 é semelhante, sugerindo uma adaptação desse agrupamento ao hospedeiro. Isso pode ser atribuído a um ancestral comum ou a um evento de transferência mais antigo. Embora esse resultado não revele a origem do BGC da microciclâmica, aponta para uma integração bem-sucedida desse agrupamento no genoma do hospedeiro. Essa conclusão, aliada à identificação dos genes biossintéticos da microciclâmica concatenados e às relações de homologia entre os genes de *C. salinus* ALCB114379 e *M. aeruginosa* NIES-298, destaca um potencial funcional significativo desse BGC na *C. salinus* ALCB114379. No entanto, análises complementares, como a metabolômica direcionada, são necessárias para validar a produção dessa molécula.

5.4. Diversidade taxonômica da cianosfera de *C. salinus* ALCB114379

A microscopia eletrônica foi empregada na tentativa de se observar a estrutura física ao redor dos filamentos cianobacterianos. As análises revelaram uma camada de mucilagem ao redor dos filamentos de cianobactérias, com os filamentos interconectados por extensões de EPSs, formando uma espécie de rede. Conforme apontado por Reid *et al.*, (2000), essa configuração tem potencial para criar microambientes favoráveis à retenção de nutrientes e à associação de microrganismos oportunistas, podendo resultar tanto em interações simbióticas benéficas quanto em outras formas de associação (Figura 18).

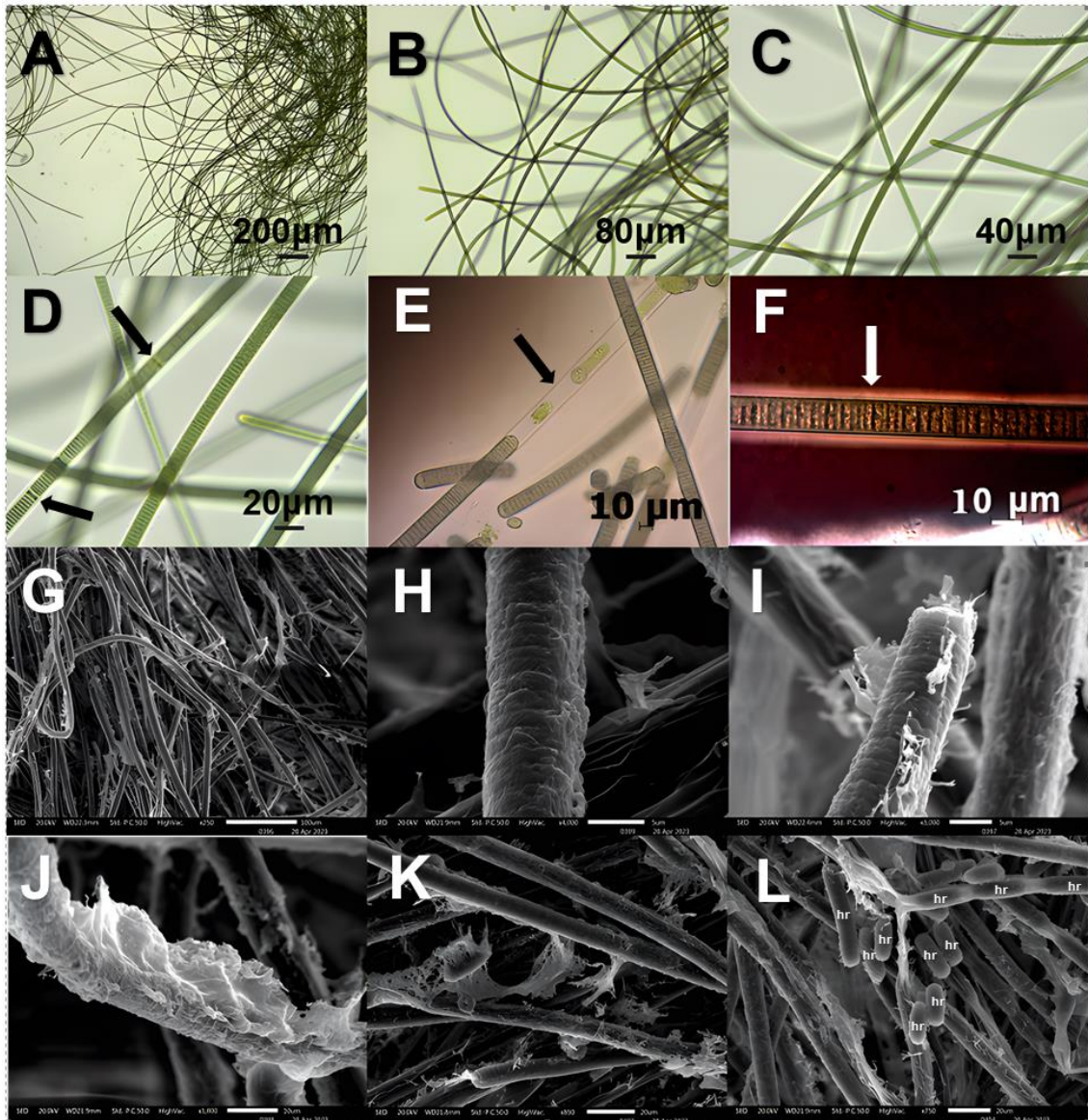


Figura 18. Micrografias da cianobactéria *C. salinus* ALCB114379. A-F. Microscopia óptica de campo claro. A-C. Aspectos gerais dos filamentos; D. Tricoma com necrídios (setas); E. Estrutura da bainha mucilaginosa (seta); F. Tricoma corado com nanquim para ressaltar a camada de EPS (seta). G-L. Microscopia eletrônica de varredura. G. Aspectos gerais dos filamentos; H. Tricoma; I. Formato da estrutura apical; J. Tricoma envolvido por uma espessa camada de EPS; K. Filamentos interligados por extensões de EPSs; L. Tricoma fragmentado expelindo hormogônios (hr).

Das 326.539 leituras geradas no sequenciamento, 244.904 foram afiliadas à Bacteria não Cyanobacterias. Dentre essas leituras, foi observada uma predominância de 92% (225.311 leituras) de conteúdo gênico proveniente do filo Pseudomonadota (Figura 19).

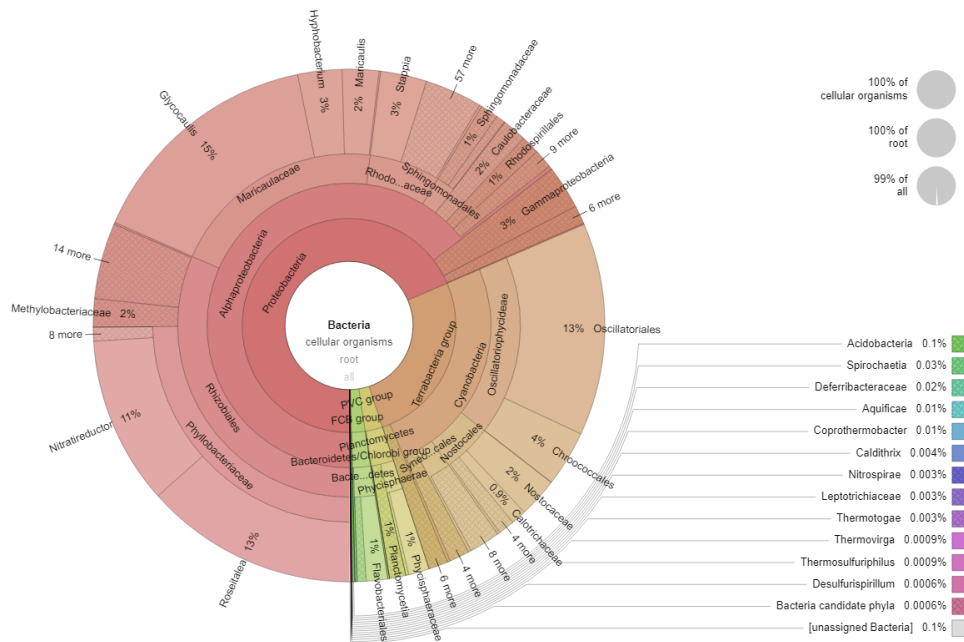


Figura 19. Perfil taxonômico das leituras da cianobactéria *C. salinus* ALCB114379 obtidas por sequenciamento em PacBio HiFi. O gráfico circular mostra a distribuição taxonômica em diferentes níveis, desde o filo até o gênero.

Alguns estudos já relataram a relação simbiótica entre cianobactérias e bactérias do filo Pseudomonadota. Van Goethem *et al.*, (2017) estudaram a co-ocorrência de microrganismos em redes hipolíticas, que são consórcios microbianos encontrados abaixo de rochas translúcidas. Nesse estudo, membros das Cyanobacteria e Alphaproteobacteria exibiram numerosas co-ocorrências positivas com outros microrganismos, sugerindo que esses grupos desempenham papéis importantes na manutenção da estrutura topológica da rede.

De acordo com Wu *et al.*, (2022), em um estudo que buscou identificar as causas da bioerosão em conchas e esqueletos de carbonato de mariscos, mexilhões e caracóis, análises *in situ* da estrutura da comunidade microbiana da camada de bioerosão em diferentes localizações geográficas mostraram que os microrganismos mais predominantes pertenciam a dois filos principais, Cyanobacteria e Pseudomonadota. Isso sugere que essa interação promoveu a função biológica de "bioerosão da casca".

Após a classificação das sequências, procedeu-se à montagem dos genomas desses microrganismos e, subsequentemente, à avaliação da qualidade dessas montagens (Tabela 7).

Tabela 7. Parâmetros de qualidade das montagens individuais dos genomas dos microrganismos associados à cianobactéria *C. salinus* ALCB114379.

	MAG1	MAG2	MAG3	MAG4	MAG5
Programa	Flye v2.9	Flye v2.9	Flye v2.9	Flye v2.9	Flye v2.9
Nº de contigs	1	1	1	1	1
Tamanho total (pb)	3.489.270	4.672.223	5.353.493	4.117.421	3.078.920
(G+C)%	66,82	68,14	71,47	61,53	67,96
Completeness (%)	98,02	99,89	99,37	100	99,51
Contaminação (%)	0,19	1,27	2,25	0,75	0,09
Circularização	Sim	Sim	Sim	Sim	Sim

Cinco genomas foram montados a partir das leituras obtidas (MAGs), sendo que cada um foi circularizado em um cromossomo único. As MAGs passaram por triagem de qualidade, apresentando >95% de completude e <5% de contaminação, conforme indicado pela ferramenta CheckM. Os genomas foram posteriormente submetidos a uma análise de qualidade adicional, complementando o CheckM, utilizando a ferramenta BUSCO que permite identificar genes ortólogos esperados em genomas completos e bem conservados, proporcionando uma avaliação mais abrangente da integridade e confiabilidade dos genomas montados (Figura 20).

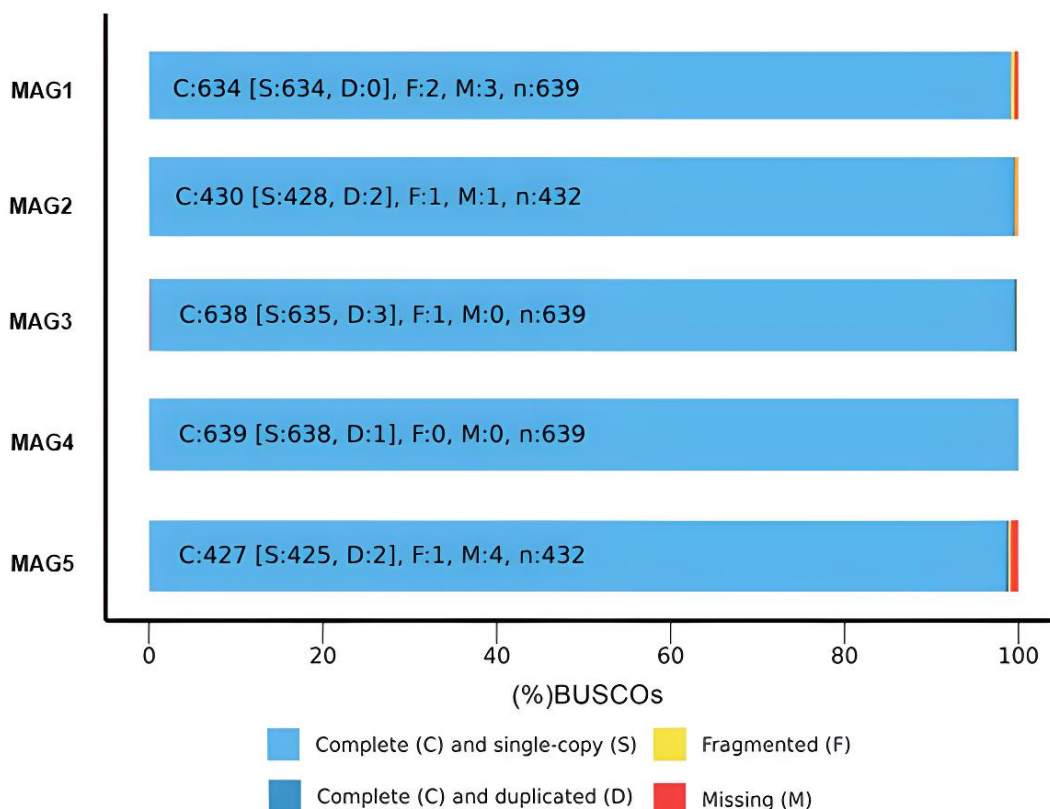


Figura 20. Estimativa percentual correspondente a 634 grupos completos de cópia única, 2 grupos fragmentados e 3 grupos ausentes para a MAG1. Quanto à MAG2, foram identificados 428 grupos completos de cópia única, 2 grupos duplicados, 1 grupo fragmentado e 1 grupo ausente. Em relação à MAG3, observou-se a presença de 665 grupos completos de cópia única, 3 grupos duplicados e 1 grupo fragmentado. No caso da MAG4, foram encontrados 638 grupos completos de cópia única e 1 grupo duplicado. Já para a MAG5, foram identificados 425 grupos completos de cópia única, 2 grupos duplicados, 1 grupo fragmentado e 4 grupos ausentes.

As MAGs geradas apresentam uma alta proporção de genes de cópia única, indicando uma montagem bem-sucedida e uma boa cobertura dos genomas. A presença de alguns genes duplicados e fragmentados, assim como a ausência de alguns genes, pode ser atribuída a variações naturais entre as populações microbianas estudadas. No entanto os resultados sugerem que as MAGs representam de maneira confiável os genomas dos microrganismos associados à cianobactéria *C. salinus* ALCB114379.

As cinco MAGs passaram por uma classificação taxonômica preliminar usando o GTDB-Tk Classify, com o objetivo de orientar a construção de um banco de dados para análise filogenômica posterior, visando elucidar as relações evolutivas entre essas MAGs (Tabela 8).

Tabela 8. Classificação taxonômica das MAGs com GTDB-Tk Classify.

	MAG1	MAG2	MAG3	MAG4	MAG5
Domínio	Bacteria				
Filo	Pseudomonadota				
Classe	Alphaproteobacteria				
Ordem	Rhizobiales				Caulobacterales
Família	Rhizobiaceae	Stappiaceae	Beijerinckiaceae	Rhizobiaceae	Maricaulaceae
Gênero	<i>Roseitalea</i>	<i>Stappia</i>	<i>Salinarimonas</i>	<i>Nitratireductor</i>	<i>Oceanicaulis</i>
Espécie	<i>Roseitalea porphyridii</i>	sp.	sp.	<i>Nitratireductor aquibiodomus</i>	sp.

Todas as MAGs foram classificadas como pertencentes ao filo Pseudomonadota, resultado que corrobora com a classificação taxonômica inicial das leituras pós-sequenciamento. Elas também compartilham a mesma classe Alphaproteobacteria. Em nível de ordem, a MAG1, MAG2, MAG3 e MAG4 se apresentaram como Rhizobiales, ao passo que a MAG5 se mostrou como Caulobacterales. A classificação taxonômica realizada pelo GTDB-Tk Classify resultou em um grau de resolução até o nível de gênero. Apenas a MAG1 e a MAG4 obtiveram correspondência com os bancos de dados a nível de espécie.

Um banco de dados com 41 genomas de referência dentro do filo Pseudomonadota foi construído manualmente a partir da classificação taxonômica realizada pelo GTDB-Tk Classify para a construção de uma árvore filogenômica (Figura 21).

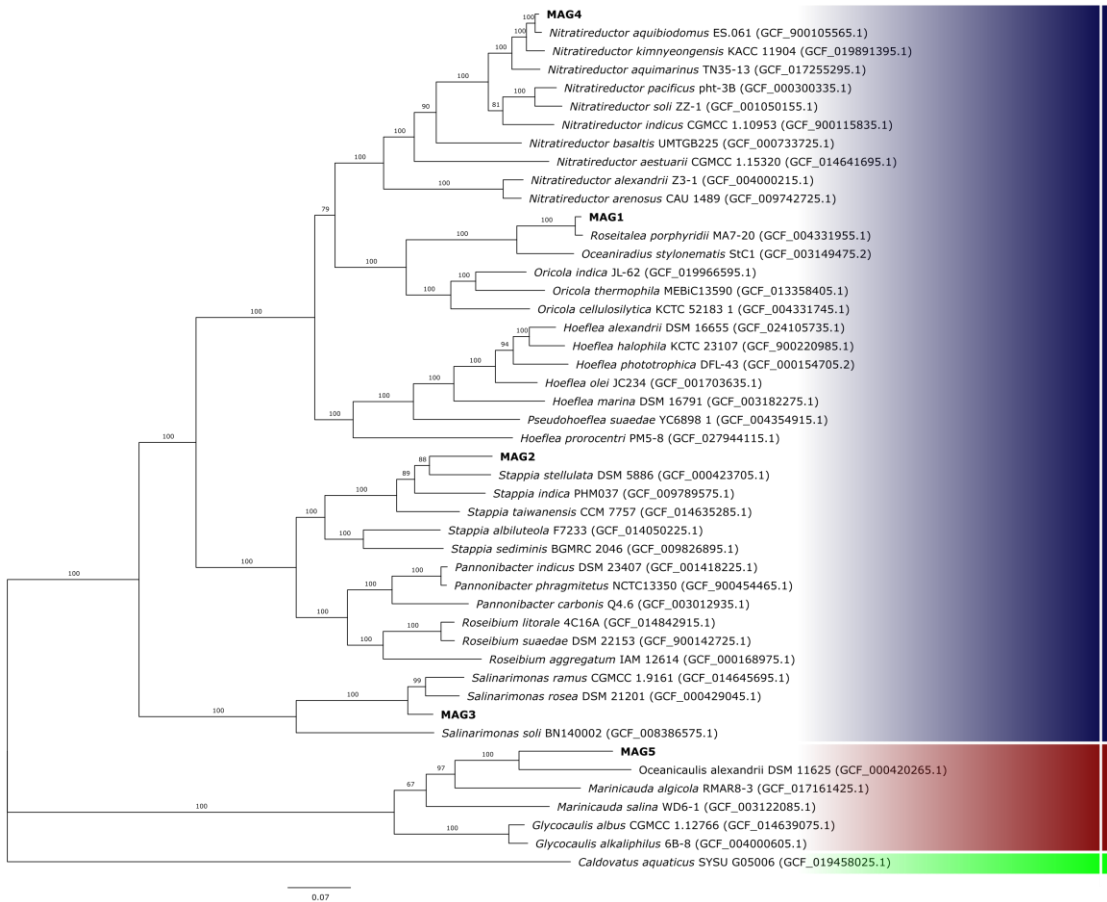


Figura 21. Árvore filogenômica construída com GTDB-tk, utilizando 120 marcadores para inferir a posição evolutiva das MAGs dentro do filo Pseudomonadota. Os clados destacados em azul representam a ordem Rhizobiales, enquanto os clados destacados em vermelho indicam a ordem Caulobacteriales, e a linhagem destacada em verde representa o grupo externo da ordem Rhodospirillales. Apenas os valores de reamostragem acima de 60 são indicados nos nós.

Observa-se que as MAGs se agruparam com linhagens taxonomicamente próximas, de acordo com a classificação inicial realizada com o GTDB-Tk Classify. Os clados também foram segmentados em nível de ordem, sendo possível discriminar as MAGs pertencentes à ordem Rhizobiales em relação à MAG5 da ordem Caulobacteriales. Como análise complementar, foi aplicada a abordagem de ANI entre as linhagens que se agruparam no mesmo clado para compreender a possibilidade de inferir o nível de gênero ou até espécie nas MAGs.

Para as MAGs 1, 2, 4 e 5, foram selecionadas as linhagens que compartilham o mesmo ponto de ramificação, ou seja, o ancestral comum no nó é compartilhado por ambas as linhagens, e a partir desse ponto, elas evoluem de forma independente, levando à formação de ramos distintos na árvore filogenética. Para a MAG3, todas as linhagens de *Salinarimonas* foram utilizadas para a análise de ANI, já que houve um comportamento diferenciado em relação às outras MAGs.

Apenas a MAG1 e a MAG4 apresentaram um valor de ANI próximo a 95%, sendo possível sugerir que esses genomas são das espécies *Roseitalea porphyridii* e *Nitratireductor aquibiodomus*, respectivamente, como já determinado pela ferramenta GTDB-Tk Classify. As demais MAGs obtiveram um valor de ANI entre 78% e 95%, desconsiderando uma pequena variação de 0,15% (77,85%) na MAG5. Esse resultado indica que esses genomas podem compartilhar o mesmo gênero taxonômico com o qual se concatenaram na árvore filogenômica. No entanto, a classificação das MAGs 2, 3 e 5 a nível de espécie permanece indeterminada.

Até o momento, não foram encontradas na literatura científica descrições específicas de microrganismos pertencentes aos mesmos gêneros taxonômicos que estão associados à cianobactéria *C. salinus* ALCB114379, estabelecendo algum tipo de relação simbiótica com outros organismos. No entanto, ampliando para o nível de família, a Rhizobiaceae inclui bactérias simbióticas fixadoras de nitrogênio, e alguns membros desta família são caracterizados por sua capacidade de estabelecer associações simbióticas com plantas hospedeiras, desenvolvendo o processo de fixação biológica de nitrogênio (Velázquez *et al.*, 2005). Sabe-se que a família Stappiaceae inclui o gênero *Stappia*, que engloba espécies quimio-organotróficas do subgrupo α -2 das Pseudomonadotas (Weber; King, 2007). O mesmo ocorreu para a família Beijerinckiaceae. No entanto, alguns membros desta família, como o gênero *Beijerinckia*, são capazes de fixar nitrogênio, o que lhes permite prosperar em habitats onde outras fontes potenciais de azoto são escassas (Verstraete *et al.*, 2013). Além disso, há metanotrofia na família Beijerinckiaceae, que inclui bactérias metanotróficas capazes de oxidar o metano como sua única fonte de carbono e energia (Tamas *et al.*, 2014). Weigel *et al.*, (2022) relataram a descoberta de um novo membro da família Maricaulaceae, *Hellea*, no microbioma de algas. O estudo sugere que *Hellea* pode desempenhar um papel na aquisição de vitamina B12 pelas algas através de uma relação simbiótica com bactérias.

Com o intuito de explorar o potencial metabólico dos microrganismos associados à cianobactéria *C. salinus* ALCB114379, realizou-se a anotação funcional das cinco MAGs por meio da ferramenta Prokka. A MAG1 apresentou 3.337 CDSs anotados, seguida pela MAG2 com 4.099 CDSs, MAG3 com 4.900 CDSs, MAG4 com 3.858 CDSs e MAG5 com 2.984 CDSs. Entretanto, uma proporção significativa dos CDSs anotados pelo Prokka foi classificada como HPs, representando, em média, 55,92% das proteínas considerando os cinco genomas

montados. Os potenciais genes codificantes de proteínas de cada MAG foram categorizados em 23 subsistemas biológicos com base em seus processos metabólicos (Figura 22).

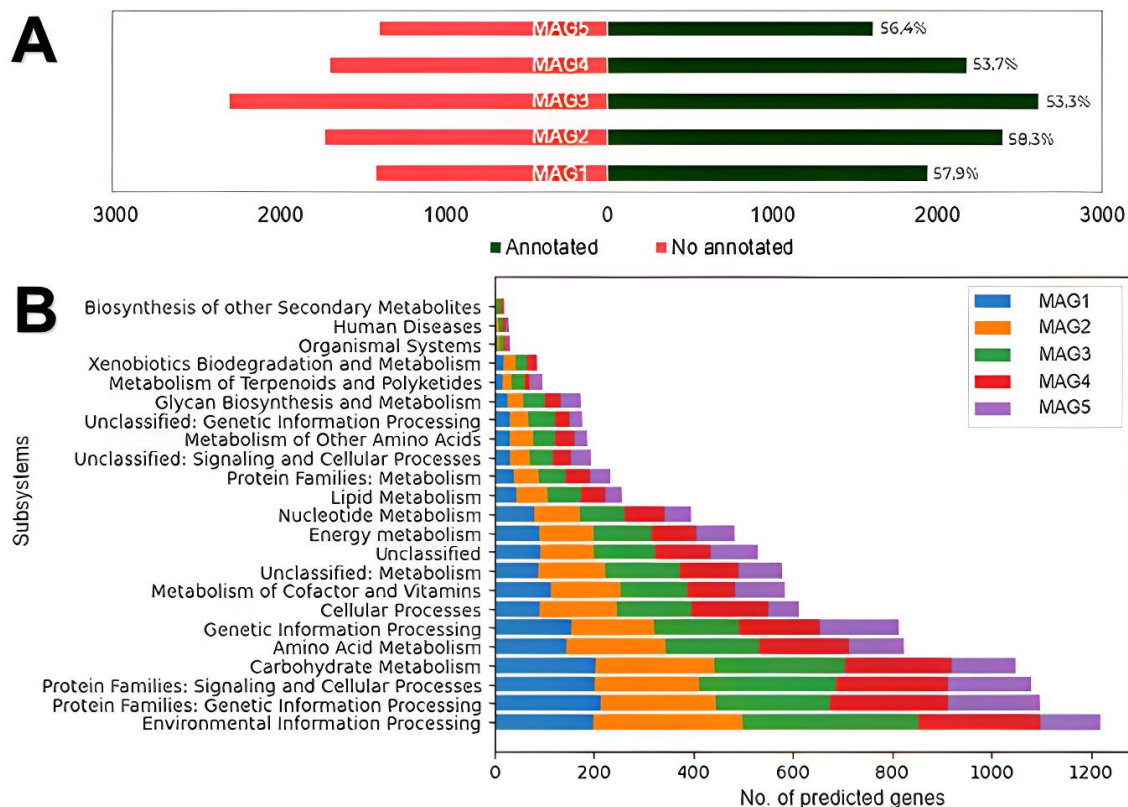


Figura 22. A. Quantidade de potenciais genes codificantes de proteínas categorizados em subsistemas biológicos; B. Distribuição dos genes preditos em subsistemas biológicos pela ferramenta BlastKOALA.

Dos 3.337 CDSs identificados na MAG1, apenas 1.938 (57,9%) foram categorizados em subsistemas biológicos. Em relação à MAG2, somente 2.400 (58,3%) CDSs foram categorizados. Na MAG3, 2.623 (53,3%) CDSs foram categorizados. Na MAG4, 2.189 (53,7%) CDSs e, na MAG5, 1.606 (56,4%) CDSs. No geral, todos os genomas dos microrganismos associados à cianobactéria *C. salinus* ALCB114379 obtiveram uma categorização funcional superior a 53% dos genes previamente anotados. Em contraste, a *C. salinus* ALCB114379 obteve apenas 2.194 (35,2%). Essa diferença pode ser explicada pela escassez de genomas cianobacterianos disponíveis nos bancos de dados, onde atualmente existem 3.846 e 4.740 genomas disponíveis no GTDB (release 214.1) e no NCBI (release 258.0), respectivamente. Em contraste, os genomas disponíveis para o filo

Pseudomonadota no GTDG atingem 174.676, enquanto que no NCBI encontram-se 1.085.306 genomas.

Os CDSs anotados das cinco MAGs foram então analisados em termos de pangenoma, utilizando o agrupamento de proteínas ortólogas (Figura 23). Essa abordagem possibilitou a classificação dos genes como pertencentes ao genoma core e ao genoma acessório.

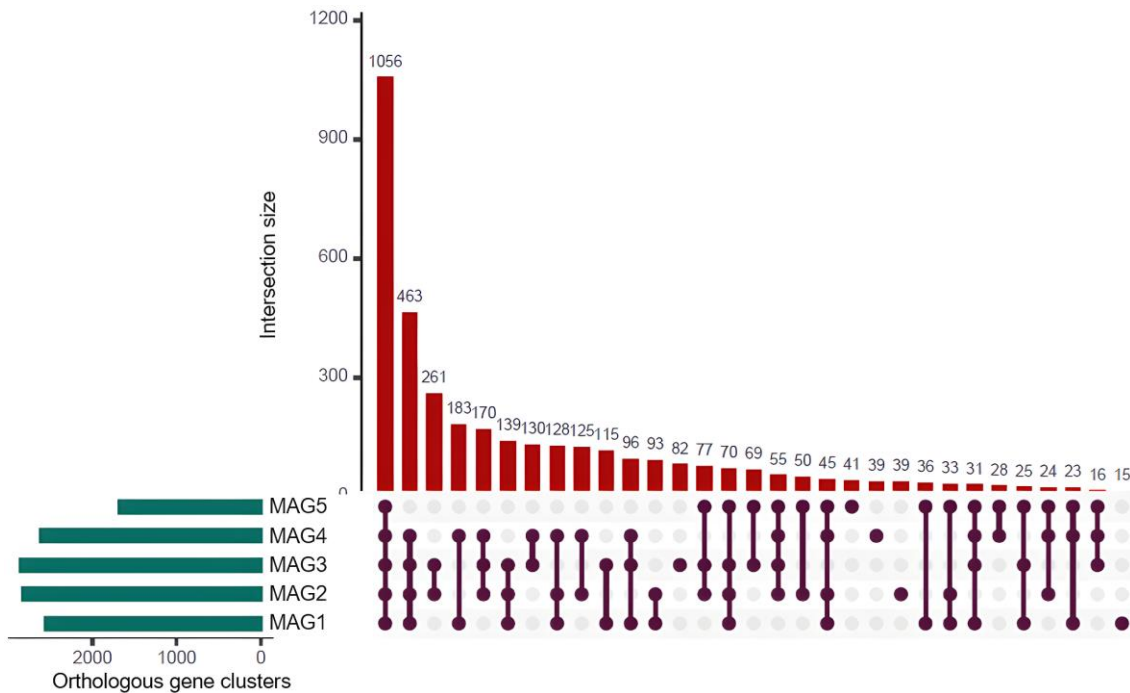


Figura 23. Configuração do genoma core e do genoma acessório entres os microrganismos associados à cianobactéria *C. salinus* ALCB114379.

As MAGs apresentaram um número significativo de genes do genoma core, ou seja, proteínas únicas que não possuem homólogos próximos nos demais genomas analisados. A MAG1 apresentou 724, seguida pela MAG2 com 1.129, MAG3 com 1.693, MAG4 com 1.128 e, por fim, a MAG5 com 1.215. Esses genes exclusivos são potencialmente associados a processos de adaptação e à produção de metabólitos especializados, indicando a presença de estratégias metabólicas únicas em cada genoma.

Por outro lado, foi possível observar que a maioria dos genes classificados como ortólogos é compartilhada entre todos os microrganismos associados à cianobactéria *C. salinus* ALCB114379, configurando essa organização como um pangenoma fechado. Um pangenoma fechado é definido quando o número de novos genes atinge um patamar, e a adição de novos genomas não leva à descoberta de

novas capacidades de codificação. Em contraste, um pangenoma aberto ocorre quando o número de novas famílias de genes continua aumentando à medida que novos genomas são adicionados, e esse incremento não parece ser assintótico, independentemente de quantos novos genomas são adicionados ao pangenoma (Rouli *et al.*, 2015). Esses genes estão relacionados ao metabolismo basal da comunidade microbiana, envolvendo processos essenciais como divisão celular e metabolismo de aminoácidos e carboidratos. No entanto, à medida que analisamos outros compartilhamentos de genes ortólogos, o número de proteínas em comum entre as MAGs tende a diminuir consideravelmente. Esses resultados sugerem que, embora haja uma base metabólica comum compartilhada por todos os microrganismos, cada MAG possui um conjunto único de genes que pode desempenhar papéis importantes na adaptação a condições específicas do ambiente ou na interação com a cianobactéria hospedeira.

5.5. Avaliação dos genes associados ao metabolismo do nitrogênio

Tanto para o genoma cianobacteriano, quanto para as 5 MAGs geradas, os CDSs categorizados no subsistema de metabolismo energético, que inclui os genes para o processo biológico de fixação de nitrogênio, foram extraídos manualmente e re-annotados com a ferramenta DRAM.

No genoma cianobacteriano verificou-se a presença de um agrupamento contendo dez genes *nif* para a fixação do nitrogênio, sendo que os genes *nifB*, *nifS*, *nifU*, *nifH*, *nifD*, *nifK*, *nifN* e *nifE* estavam localizados sequencialmente no genoma, enquanto o *nifW* estava a uma distância intergênica de quatro CDSs a partir do gene *nifE*. O gene *nifJ* foi identificado a uma distância de 3.356.004 pb a partir do gene *nifB*. A localização do gene *nifJ* nos genomas procarióticos pode ser variável e nem sempre pode estar intimamente ligada ao agrupamento principal do gene *nif* (Dai *et al.*, 2014).

As sequências de aminoácidos (aa) de cada gene *nif* foram extraídas, quantificadas em termos de aminoácidos e submetidas a um BLASTP. Este processo teve como objetivo identificar organismos que possuem os genes mais próximos evolutivamente da *C. salinus* ALCB114379, considerando níveis de identidade e cobertura para inferência de homologia (Tabela 9).

Tabela 9. Informações sobre as enzimas codificadoras dos genes envolvidos na fixação de nitrogênio anotados da cianobactéria *C. salinus* ALCB114379.

Gene	Aminoácidos (aa)	Função deduzida	Organismo similar mais próximo (n° acesso NCBI)	% Identidade/ Cobertura
<i>nifW</i>	112	Nitrogenase-stabilizing/protective protein	<i>Limnoraphis robusta</i> (WP_046276706.1)	95,54/100
<i>nifE</i>	505	Nitrogenase iron-molybdenum cofactor biosynthesis protein	<i>Limnoraphis sp.</i> WC205 (MCG5061633.1)	91,53/93
<i>nifN</i>	451	Nitrogenase iron-molybdenum cofactor biosynthesis protein	<i>Lyngbya aestuarii</i> (WP_023065321.1)	93,13/100
<i>nifK</i>	511	Nitrogenase molybdenum-iron protein subunit beta	<i>Lyngbya sp.</i> PCC 8106 (WP_009784197.1)	93,93/100
<i>nifD</i>	478	Nitrogenase molybdenum-iron protein alpha	<i>Limnoraphis sp.</i> WC205 (MCG5061630.1)	96,23/100
<i>nifH</i>	299	Nitrogenase iron protein	<i>Limnoraphis sp.</i> WC205 (MCG5061629.1)	96,32/100
<i>nifU</i>	298	Fe-S cluster assembly protein	<i>Limnoraphis robusta</i> (WP_046276697.1)	91,95/100
<i>nifS</i>	400	Cysteine desulfurase	<i>Limnoraphis sp.</i> WC205 (MCG5061627.1)	94,44/99
<i>nifB</i>	486	Nitrogenase cofactor biosynthesis protein	<i>Limnoraphis sp.</i> WC205 (MCG5061626.1)	93,84/100
<i>nifJ</i>	1.204	Pyruvate:ferredoxin (flavodoxin) oxidoreductase	<i>Lyngbya aestuarii</i> (WP_023068000.1)	93,77/100

Os resultados sugerem a provável presença de um sistema de nitrogenase de molibdênio. A identificação de genes-chave associados à biossíntese do cofator ferro-molibdênio (*nifE* e *nifN*) e das subunidades da nitrogenase do molibdênio (*nifK* e *nifD*), juntamente com a presença do componente proteico de ferro (*nifH*), indica fortemente a utilização de uma nitrogenase contendo molibdênio para a fixação de nitrogênio. O gene *nifW*, com sua função de proteína Nitrogenase-stabilizing/protective, apresenta indícios de atuar na proteção da nitrogenase de molibdênio contra danos oxidativos (Lee *et al.*, 1998). Essa função torna-se

significativamente relevante ao considerar que a cianobactéria *C. salinus* ALCB114379 não possui uma célula especializada para fixação de nitrogênio, denominada heterócito. Portanto, presume-se que haja outra estratégia para a fixação de nitrogênio, caso o agrupamento de genes *nif* seja efetivamente funcional.

As cianobactérias empregam diversas estratégias para a fixação de nitrogênio. Uma dessas estratégias é a separação espacial, onde a conversão de N_2 em NH_3 ocorre em locais separados para evitar a inativação da enzima nitrogenase pelo oxigênio. Células especializadas, denominadas heterocistos, são responsáveis por essa fixação de nitrogênio, segregando-se espacialmente das células vegetativas onde ocorre a fotossíntese (Esteves-Ferreira *et al.*, 2017; Mohr *et al.*, 2013). Além disso, as cianobactérias também adotam a separação temporal, realizando atividades de fixação de nitrogênio e fotossíntese em momentos distintos para evitar a inativação da enzima pelo oxigênio. Por exemplo, em cianobactérias como *Oscillatoria sp.*, a atividade da nitrogenase é predominantemente presente durante o período escuro, quando a fotossíntese oxigenada é limitada (Stal; Krumbein, 1987). Já a cianobactéria *Trichodesmium* combina estratégias espaciais e temporais, segregando as células especializadas chamadas diazócitos espacialmente e regulando a fixação de nitrogênio predominantemente durante o dia (Gardner; Hodge; Boyle, 2023).

A análise taxonômica, utilizando o gene *nifH*, foi empregada para propor uma possível estratégia de fixação de nitrogênio adotada pela cianobactéria *C. salinus* ALCB114379 (Figura 24). A amostragem incluiu 77 linhagens de cianobactérias pertencentes a diferentes táxons, que utilizam estratégias espaciais e temporais distintas para a fixação de nitrogênio. O gene *nifH*, responsável pela codificação da enzima nitrogenase redutase, foi escolhido para a análise taxonômica, pois é um gene conservado presente em todos os organismos diazotróficos, abrangendo bactérias e arqueias (Angel *et al.*, 2018). Essa escolha se deve à sua ampla presença e conservação, tornando-o um marcador funcional amplamente sequenciado, cuja sequência é valiosa para inferir a identidade taxonômica de diazotróficos em amostras ambientais (Gaby; Buckley, 2014).

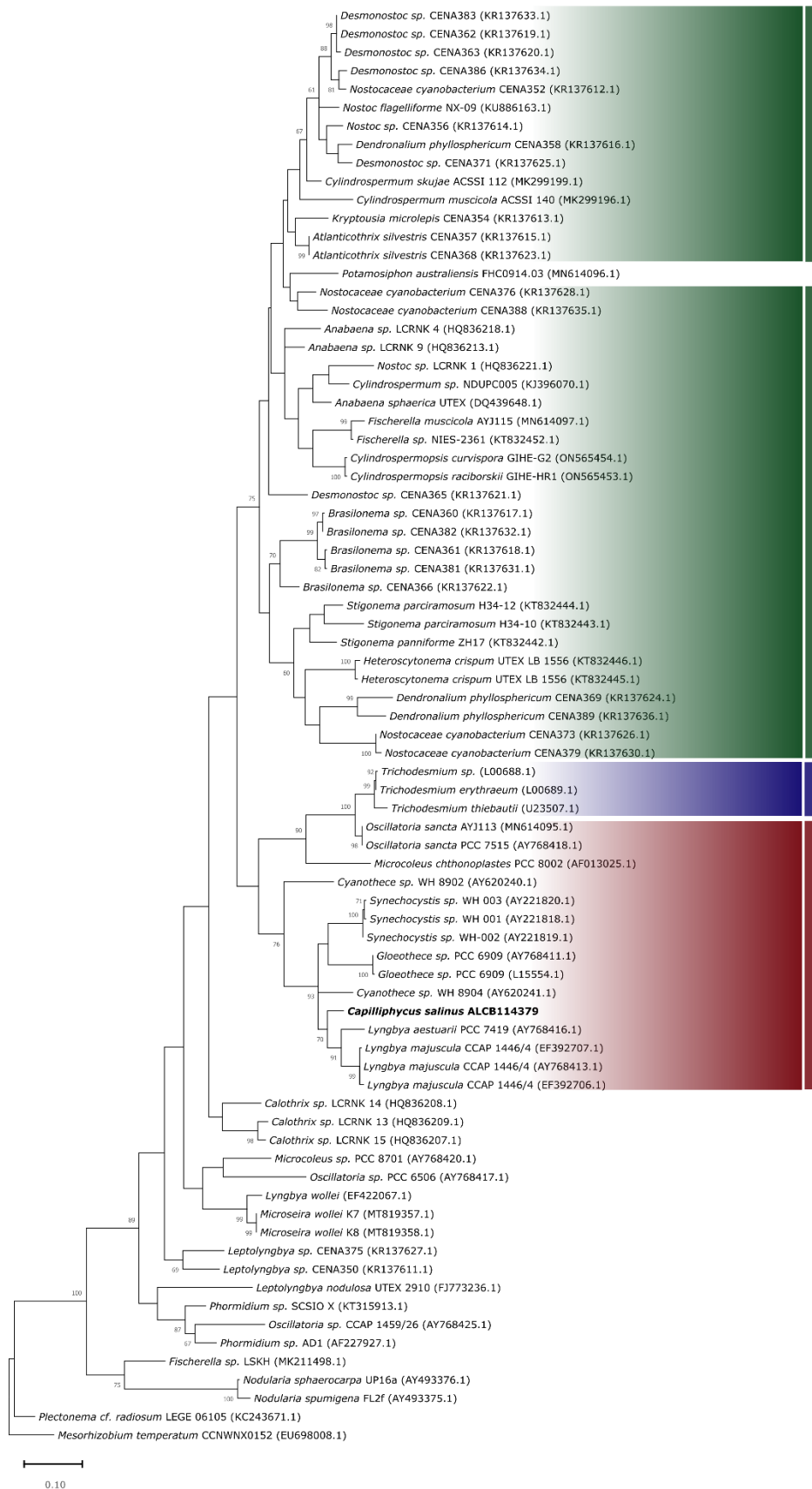


Figura 24. Filogenia do gene *nifH* gerada por Máxima Verossimilhança, destacando cianobactérias diazotróficas em verde (fixação espacial), em vermelho (fixação temporal) e em azul (fixação espacial e temporal). Apenas os valores de reamostragem acima de 60 são indicados nos nós.

A Figura 25 evidencia um agrupamento consistente das cianobactérias com base nas diferentes estratégias adotadas para a fixação de nitrogênio. Destacam-se três grupos principais: cianobactérias heterocitadas, que realizam a fixação espacial de N_2 ; cianobactérias não heterocitadas, que adotam a fixação temporal de N_2 ; e a cianobactéria *Trichodesmium*, que apresenta células modificadas chamadas diazófitos, realizando assim a fixação tanto espacial quanto temporal. Uma exceção notável a essa organização aparente é a inclusão da cianobactéria não heterocitada *Potamosiphon australiensis* FHC0914.03 no grupo das heterocitadas.

A inferência decorrente desses agrupamentos sugere que *C. salinus* ALCB114379 emprega a estratégia de fixação temporal de nitrogênio. Essa conclusão é corroborada pela observação de que essa cianobactéria se agrupou com outras linhagens conhecidas por adotarem a estratégia de fixação temporal de nitrogênio. No entanto, para fundamentar essa suposição e confirmar a funcionalidade dos genes *nif* identificados, análises complementares adicionais, como a PCR quantitativa (qPCR) e ensaio de redução de acetileno (ARA), são consideradas necessárias. Tais análises proporcionarão evidências mais diretas sobre a expressão e funcionalidade dos genes *nif*, consolidando assim a compreensão da estratégia de fixação de nitrogênio adotada por *C. salinus* ALCB114379.

Nenhuma das MAGs apresentou agrupamento de genes biossintéticos, contendo os genes *nif*, para a fixação de nitrogênio. Os genes identificados compõem outras etapas específicas dentro do processo de fixação, em que a MAG1 e a MAG4 apresentaram potencial na conversão de nitrito em óxido nítrico. Enquanto isso, tanto a MAG2 quanto a MAG3, assim como a MAG4, podem estar envolvidas em dois processos do ciclo: a conversão de nitrato em nitrito e a conversão de nitrito para nitrato. Para a MAG5, não foi identificado nenhum aporte genético para atuação no ciclo do nitrogênio.

Para a fixação de nitrogênio, é necessário um conjunto de genes mínimos essenciais para a biossíntese da nitrogenase, os quais podem variar entre diferentes organismos. Dos Santos *et al.*, (2012) propuseram um conjunto mínimo de genes para a fixação de nitrogênio que pode ser utilizado como uma ferramenta de busca *in silico* para a identificação de diazotróficos adicionais. Identificaram-se seis genes *nif* que são conhecidos por serem essenciais para a fixação de nitrogênio em sistemas caracterizados. Esses genes são *nifH*, *nifD*, *nifK*, *nifE*, *nifN* e *nifB*. A co-

ocorrência desses seis genes *nif* levou à proposta de um conjunto mínimo de genes para a fixação de nitrogênio.

A partir da predição desses genes mínimos essenciais, tanto na linhagem cianobacteriana, assim como nos microorganismos da cianosfera, mapeou-se a possível ação desses organismos no ciclo desse elemento (Figura 25).

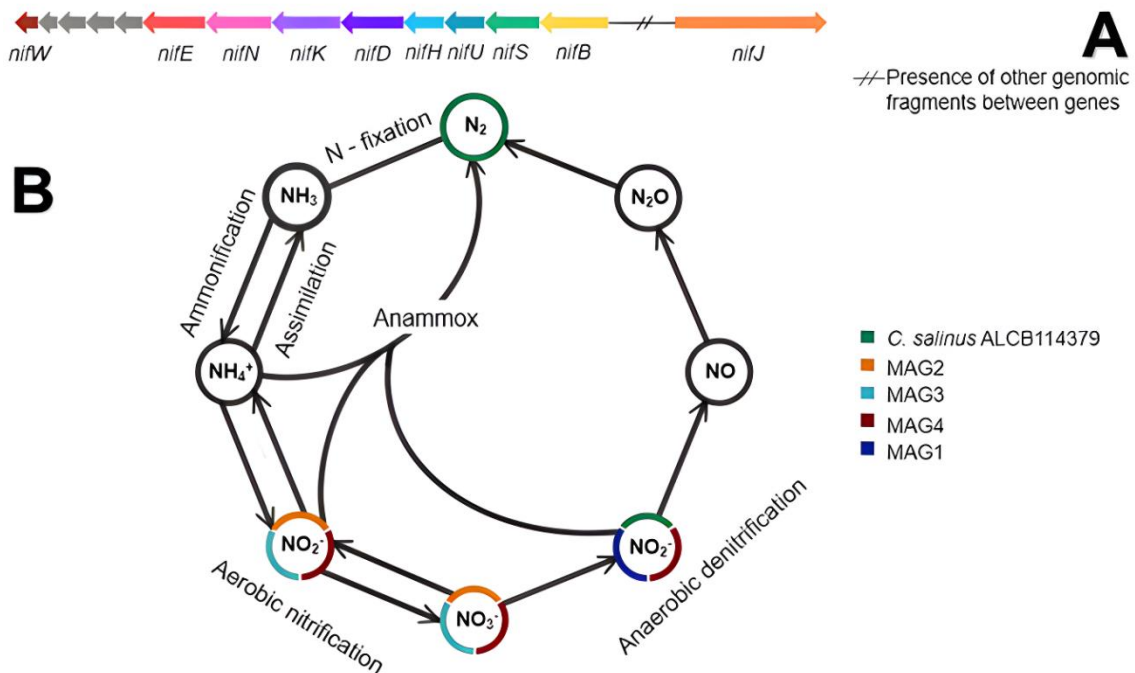


Figura 25. A. Agrupamento de genes biossintéticos da nitrogenase na cianobactéria *C. salinus* ALCB114379; B. Potencial ação da cianobactéria *C. salinus* e de seus microorganismos associados no ciclo do nitrogênio.

Embora a ausência de um conjunto de genes biossintéticos para fixação de nitrogênio nos microorganismos associados sugira que eles podem não contribuir diretamente para a fixação de nitrogênio pela via tradicional da nitrogenase, genes específicos em MAG1 e MAG4 indicam envolvimento na conversão de nitrito em óxido nítrico. MAG2 e MAG3 exibem versatilidade ao participarem potencialmente na conversão de nitrato em nitrito e nitrito em nitrato.

6. CONCLUSÕES

Devido à escassez de genomas de referência para cianobactérias da ordem Oscillatoriales, a resolução taxonômica e análises comparativas tornam-se desafiadoras. O agrupamento da *C. salinus* ALCB114379 na família Serenicapillariaceae fortalece sua classificação. Análises de anotação gênica revelaram diversidade genética não explorada, indicando oportunidades para a descoberta de novas vias metabólicas. Dois agrupamentos de genes biossintéticos foram identificados; o BGC relacionado a MAAs teve sua produção validada, enquanto o da microciclâmida requer análises adicionais. Moléculas como shinorina e palitina sugerem adaptações ao ambiente supralitoral, exposto à alta intensidade de luz UV. A microciclâmida pode ter papel antipredatório na transição entre ambientes aquáticos e terrestres. Quanto ao ciclo do nitrogênio, a cianobactéria mostrou potencial de fixação, enquanto os microrganismos associados podem atuar no ciclo biológico desse elemento em outras etapas. No entanto, as interações entre os microrganismos com seu hospedeiro cianobacteriano ainda precisam ser exploradas. A estratégia de fixação temporal de nitrogênio pela adotada pela *C. salinus* ALCB114379 foi inferida, embora análises adicionais sejam requeridas.

REFERÊNCIAS

AHMAD, Nadim *et al.* Biosynthetic gene cluster synteny: Orthologous polyketide synthases in *Hypogymnia physodes*, *Hypogymnia tubulosa*, and *Parmelia sulcata*. **MicrobiologyOpen**, [s. l.], v. 12, n. 5, 2023.

ALLAF, Malihe Mehdizadeh; PEERHOSSAINI, Hassan. Cyanobacteria: Model Microorganisms and Beyond. **Microorganisms**, [s. l.], v. 10, n. 4, 2022.

ANGEL, Roey *et al.* Evaluation of primers targeting the diazotroph functional gene and development of NifMAP - A bioinformatics pipeline for analyzing nifH amplicon data. **Frontiers in Microbiology**, [s. l.], v. 9, n. APR, p. 1–15, 2018.

ATANASOV, Atanas G. *et al.* Natural products in drug discovery: advances and opportunities. **Nature Reviews Drug Discovery**, [s. l.], v. 20, n. 3, p. 200–216, 2021.

BACCI, Giovanni *et al.* Composition of supralittoral sediments bacterial communities in a Mediterranean island. **Annals of Microbiology**, [s. l.], v. 65, n. 1, p. 1–13, 2015.

BAILLY-BECHET, Marc *et al.* Codon usage domains over bacterial chromosomes. **PLoS Computational Biology**, [s. l.], v. 2, n. 4, p. 263–275, 2006.

BARRANGOU, Rodolphe; MARRAFFINI, Luciano A. CRISPR-cas systems: Prokaryotes upgrade to adaptive immunity. **Molecular Cell**, [s. l.], v. 54, n. 2, p. 234–244, 2014.

BATEMAN, Alex *et al.* UniProt: the universal protein knowledgebase in 2021. **Nucleic Acids Research**, [s. l.], v. 49, n. D1, p. D480–D489, 2021.

BERRY, John P. Cyanobacterial Toxins as Allelochemicals with Potential Applications as Algaeicides, Herbicides and Insecticides. **Marine Drugs**, [s. l.], v. 6, n. 2, p. 117–146, 2008.

BLIN, Kai *et al.* antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. **Nucleic Acids Research**, [s. l.], v. 51, n. May, p. 46–50, 2023.

BONELLO, Guido *et al.* Sea-level change and the supralittoral environment: Potential impact on a splashpool habitat on the Ligurian coast (NW Mediterranean). **Journal of Biological Research (Italy)**, [s. l.], v. 95, n. 2, 2022.

BOZAN, Mahir *et al.* Spatial organization and proteome of a dual-species cyanobacterial biofilm alter among N₂-fixing and non-fixing conditions. **mSystems**, [s. l.], v. 8, n. 3, 2023.

BOZAN, Mahir; SCHMID, Andreas; BÜHLER, Katja. Evaluation of self-sustaining cyanobacterial biofilms for technical applications. **Biofilm**, [s. l.], v. 4, n. February, p. 100073, 2022. Disponível em: <https://doi.org/10.1016/j.bioflm.2022.100073>.

BRITO, Ângela *et al.* **Cyanobacterial biofilms: Formation, distribution, and applications.** [S. l.]: Elsevier Inc., 2022. Disponível em: <https://doi.org/10.1016/B978-0-323-91202-0.00004-X>.

BURGANSKAYA, E. I. *et al.* Bacterial Communities of Microbial Mats of the White Sea Supralittoral and of the Littoral of the Lakes Separated from the Sea. **Microbiology (Russian Federation)**, [s. l.], v. 88, n. 5, p. 600–612, 2019.

BURGIN, Josephine *et al.* The European Nucleotide Archive in 2022. **Nucleic Acids Research**, [s. l.], v. 51, n. D1, p. D121–D125, 2023.

CAIRES, Taiara A. *et al.* Polyphasic delimitation of a filamentous marine genus, *capillus* gen. Nov. (cyanobacteria, oscillatoriaceae) with the description of two Brazilian species. **Algae**, [s. l.], v. 33, n. 4, p. 291–304, 2018.

CAPONE, Douglas G. *et al.* Nitrogen fixation by *Trichodesmium* spp.: An important source of new nitrogen to the tropical and subtropical North Atlantic Ocean. **Global Biogeochemical Cycles**, [s. l.], v. 19, n. 2, p. 1–17, 2005.

CARRETO, Jose I.; CARIGNAN, Mario O. Mycosporine-like amino acids: Relevant secondary metabolites. chemical and ecological aspects. **Marine Drugs**, [s. l.], v. 9, n. 3, p. 387–446, 2011.

CARVER, Tim *et al.* Artemis: An integrated platform for visualization and analysis of high-throughput sequence-based experimental data. **Bioinformatics**, [s. l.], v. 28, n. 4, p. 464–469, 2012.

CHARPY, L. *et al.* Cyanobacteria in Coral Reef Ecosystems: A Review. **Journal of Marine Biology**, [s. l.], v. 2012, p. 1–9, 2012.

CHAUMEIL, Pierre Alain *et al.* GTDB-Tk: A toolkit to classify genomes with the genome taxonomy database. **Bioinformatics**, [s. l.], v. 36, n. 6, p. 1925–1927, 2020.

CHENG, Haoyu *et al.* Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. **Nature Methods**, [s. l.], v. 18, n. 2, p. 170–175, 2021. Disponível em: <http://dx.doi.org/10.1038/s41592-020-01056-5>.

CHENG, Feng *et al.* Syntenic gene analysis between *Brassica rapa* and other Brassicaceae species. **Frontiers in Plant Science**, [s. l.], v. 3, n. AUG, p. 1–6, 2012.

CHRAPUSTA, Ewelina *et al.* Mycosporine-Like Amino Acids: Potential health and beauty ingredients. **Marine Drugs**, [s. l.], v. 15, n. 10, p. 1–29, 2017.

CLARRIDGE, Jill E. Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. **Clinical Microbiology Reviews**, [s. l.], v. 17, n. 4, p. 840–862, 2004.

COLAVECCHIO, Anna *et al.* Prophage integrase typing is a useful indicator of genomic diversity in *Salmonella enterica*. **Frontiers in Microbiology**, [s. l.], v. 8, n. JUL, p. 1–11, 2017.

COOPER, Matthew B.; SMITH, Alison G. Exploring mutualistic interactions between microalgae and bacteria in the omics age. **Current Opinion in Plant Biology**, [s. l.], v. 26, n. Figure 1, p. 147–153, 2015. Disponível em: <http://dx.doi.org/10.1016/j.pbi.2015.07.003>.

CORDEIRO, Rita *et al.* Cyanobacteria phylogenetic studies reveal evidence for polyphyletic genera from thermal and freshwater habitats. **Diversity**, [s. l.], v. 12, n. 8, 2020.

CRITS-CHRISTOPH, Alexander *et al.* Transporter genes in biosynthetic gene clusters predict metabolite characteristics and siderophore activity. **Genome Research**, [s. l.], v. 31, n. 2, p. 239–250, 2021.

CUNNINGHAM, Fiona *et al.* Ensembl 2022. **Nucleic Acids Research**, [s. l.], v. 50, n. D1, p. D988–D995, 2022.

DADHEECH, Pawan K. *et al.* Cyanobacterial diversity in the hot spring, pelagic and benthic habitats of a tropical soda lake. **FEMS Microbiology Ecology**, [s. l.], v. 85, n. 2, p. 389–401, 2013.

DAI, Zhimin *et al.* Identification of nitrogen-fixing genes and gene clusters from metagenomic library of acid mine drainage. **PLoS ONE**, [s. l.], v. 9, n. 2, p. 1–9, 2014.

DEMARY, Justine *et al.* **Natural products from cyanobacteria: Focus on beneficial activities.** [S. l.]: MDPI AG, 2019.

DEXTRO, Rafael B. *et al.* Exploring the Relationship between Biosynthetic Gene Clusters and Constitutive Production of Mycosporine-like Amino Acids in Brazilian Cyanobacteria. **Molecules**, [s. l.], v. 28, n. 3, 2023.

DEXTRO, Rafael B. *et al.* **Trends in Free-access Genomic Data Accelerate Advances in Cyanobacteria Taxonomy.** [S. l.]: John Wiley and Sons Inc, 2021.

DITTMANN, Elke *et al.* **Natural Product Biosynthetic Diversity and Comparative Genomics of the Cyanobacteria.** [S. l.]: Elsevier Ltd, 2015.

DOS SANTOS, Patricia C. *et al.* Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. **BMC Genomics**, [s. l.], v. 13, n. 1, p. 1–12, 2012.

DUPRAZ, Christophe *et al.* **Processes of carbonate precipitation in modern microbial mats.** [S. l.: s. n.], 2009.

EDDY, Sean R. Accelerated profile HMM searches. **PLoS Computational Biology**, [s. l.], v. 7, n. 10, 2011.

EDGAR, Robert C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, [s. l.], v. 32, n. 5, p. 1792–1797, 2004.

ENGINE, Niclas; CAMERON COATES, R.; GERWICK, William H. 16S Rrna gene heterogeneity in the filamentous marine cyanobacterial genus *lyngbya*. **Journal of Phycology**, [s. l.], v. 46, n. 3, p. 591–601, 2010.

ESTEVES-FERREIRA, Alberto A. *et al.* Cyanobacterial nitrogenases: Phylogenetic diversity, regulation and functional predictions. **Genetics and Molecular Biology**, [s. l.], v. 40, n. 1, p. 261–275, 2017.

EVDOKIAS, Gregory *et al.* Identification of a novel biosynthetic gene cluster in *aspergillus niger* using comparative genomics. **Journal of Fungi**, [s. l.], v. 7, n. 5, 2021.

FIORE, Marli Fátima *et al.* Guanitoxin, re-naming a cyanobacterial organophosphate toxin. **Harmful Algae**, [s. l.], v. 92, n. December 2019, p. 101737, 2020. Disponível em: <https://doi.org/10.1016/j.hal.2019.101737>.

FRENKEL, Alona *et al.* Cell specialization in cyanobacterial biofilm development revealed by expression of a cell-surface and extracellular matrix protein. **npj Biofilms and Microbiomes**, [s. l.], v. 9, n. 1, p. 1–10, 2023.

FUENTES-TRISTAN, Susana *et al.* Bioinspired biomolecules: Mycosporine-like amino acids and scytonemin from *Lyngbya* sp. with UV-protection potentialities. **Journal of Photochemistry and Photobiology B: Biology**, [s. l.], v. 201, n. August, p. 111684, 2019. Disponível em: <https://doi.org/10.1016/j.jphotobiol.2019.111684>.

GABY, John Christian; BUCKLEY, Daniel H. A comprehensive aligned nifH gene database: A multipurpose tool for studies of nitrogen-fixing bacteria. **Database**, [s. l.], v. 2014, p. 1–8, 2014.

GARCIA-PICHEL, Ferran *et al.* Timing the evolutionary advent of cyanobacteria and the later great oxidation event using gene phylogenies of a sunscreen. **mBio**, [s. l.], v. 10, n. 3, 2019.

GARDNER, Joseph J.; HODGE, Bri-Mathias S.; BOYLE, Nanette R. Investigating the Unique Ability of *Trichodesmium* To Fix Carbon and Nitrogen Simultaneously Using MiMoSA. **mSystems**, [s. l.], v. 8, n. 1, p. 1–14, 2023.

GASTEIGER, E.; JUNG, E.; BAIROCH, A. SWISS-PROT: Connecting biomolecular knowledge via a protein database. **Current Issues in Molecular Biology**, [s. l.], v. 3, n. 3, p. 47–55, 2001.

GERALDES, Vanessa *et al.* Development and validation of a rapid LC-MS/MS method for the quantification of mycosporines and mycosporine-like amino acids (MAAs) from cyanobacteria. **Algal Research**, [s. l.], v. 46, n. October 2019, p. 101796, 2020a. Disponível em: <https://doi.org/10.1016/j.algal.2020.101796>.

GERALDES, Vanessa *et al.* Genetic and biochemical evidence for redundant pathways leading to mycosporine-like amino acid biosynthesis in the cyanobacterium *Sphaerospermopsis torques-reginae* ITP-024. **Algae**, [s. l.], v. 35, n. 2, p. 177–187, 2020b.

GERALDES, Vanessa *et al.* Identification and distribution of mycosporine-like amino acids in Brazilian cyanobacteria using ultrahigh-performance liquid chromatography with diode array detection coupled to quadrupole time-of-flight mass spectrometry. **Rapid Communications in Mass Spectrometry**, [s. l.], v. 34, n. S3, p. 1–10, 2020c.

GERALDES, Vanessa; PINTO, Ernani. Mycosporine-like amino acids (Maas): Biology, chemistry and identification features. **Pharmaceuticals**, [s. l.], v. 14, n. 1, p. 1–17, 2021.

GILCHRIST, Cameron L.M.; CHOOI, Yit Heng. Clinker & clustermap.js: Automatic generation of gene cluster comparison figures. **Bioinformatics**, [s. l.], v. 37, n. 16, p. 2473–2475, 2021.

GOMES-DOS-SANTOS, André *et al.* PacBio Hi-Fi genome assembly of the Iberian dolphin freshwater mussel *Unio delphinus* Spengler, 1793. **Scientific Data**, [s. l.], v. 10, n. 1, p. 1–11, 2023.

GORIS, Johan *et al.* DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. **International Journal of Systematic and Evolutionary Microbiology**, [s. l.], v. 57, n. 1, p. 81–91, 2007.

GRANT, Jason R. *et al.* Proksee: In-depth characterization and visualization of bacterial genomes. **Nucleic Acids Research**, [s. l.], v. 51, n. W1, p. W484–W492, 2023.

GU, Wenjia *et al.* The biochemistry and structural biology of cyanobactin biosynthetic enzymes. **Methods in enzymology**, [s. l.], v. 604, n. 3, p. 113, 2018. Disponível em: [/pmc/articles/PMC6463883/%0A/pmc/articles/PMC6463883/?report=abstract%0Ahttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC6463883/](https://pmc/articles/PMC6463883/%0A/pmc/articles/PMC6463883/?report=abstract%0Ahttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC6463883/).

GUREVICH, Alexey *et al.* QUASt: Quality assessment tool for genome assemblies. **Bioinformatics**, [s. l.], v. 29, n. 8, p. 1072–1075, 2013.

HARDING, Katie *et al.* Symbiotic unicellular cyanobacteria fix nitrogen in the Arctic Ocean. **Proceedings of the National Academy of Sciences of the United States of America**, [s. l.], v. 115, n. 52, p. 13371–13375, 2018.

HEATHER, James M.; CHAIN, Benjamin. The sequence of sequencers: The history of sequencing DNA. **Genomics**, [s. l.], v. 107, n. 1, p. 1–8, 2016. Disponível em: <http://dx.doi.org/10.1016/j.ygeno.2015.11.003>.

HECK, Karina *et al.* Evaluating methods for purifying cyanobacterial cultures by qPCR and high-throughput Illumina sequencing. **Journal of Microbiological Methods**, [s. l.], v. 129, p. 55–60, 2016.

HIROSE, Yuu *et al.* Characterization of the genuine type 2 chromatic acclimation in the two *Geminocystis* cyanobacteria. **DNA Research**, [s. l.], v. 24, n. 4, p. 387–396, 2017.

HO, Ming Yang *et al.* Far-red light photoacclimation (FaRLiP) in *Synechococcus* sp. PCC 7335. II. Characterization of phycobiliproteins produced during acclimation to far-red light. **Photosynthesis Research**, [s. l.], v. 131, n. 2, p. 187–202, 2017.

HURTADO, Luis A.; LEE, Eun Jung; MATEOS, Mariana. Contrasting Phylogeography of Sandy vs. Rocky Supralittoral Isopods in the Megadiverse and Geologically Dynamic Gulf of California and Adjacent Areas. **PLoS ONE**, [s. l.], v. 8, n. 7, 2013.

HYATT, Doug *et al.* **Prodigal: prokaryotic gene recognition and translation initiation site identification**. [S. l.: s. n.], 2010. Disponível em: <http://www.biomedcentral.com/1471-2105/11/119>.

ISHIDA, K.; NAKAGAWA, H.; MURAKAMI, M. Microcyclamide, a cytotoxic cyclic hexapeptide from the cyanobacterium *Microcystis aeruginosa*. **Journal of Natural Products**, [s. l.], v. 63, n. 9, p. 1315–1317, 2000.

JACK, Donald L.; YANG, Nelson M.; SAIER, Milton H. The drug/metabolite transporter superfamily. **European Journal of Biochemistry**, [s. l.], v. 268, n. 13, p. 3620–3639, 2001.

JAFARI PORZANI, Samaneh; KONUR, Ozcan; NOWRUZI, Bahareh. Cyanobacterial natural products as sources for antiviral drug discovery against COVID-19. **Journal of Biomolecular Structure and Dynamics**, [s. l.], v. 40, n. 16, p. 7629–7644, 2022. Disponível em: <https://doi.org/10.1080/07391102.2021.1899050>.

KANEHISA, Minoru; SATO, Yoko; MORISHIMA, Kanae. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. **Journal of Molecular Biology**, [s. l.], v. 428, n. 4, p. 726–731, 2016. Disponível em: <http://dx.doi.org/10.1016/j.jmb.2015.11.006>.

KAUTSAR, Satria A. *et al.* BiG-FAM: The biosynthetic gene cluster families database. **Nucleic Acids Research**, [s. l.], v. 49, n. D1, p. D490–D497, 2021.

KINGSTON, P. F. Benthic marine overview. In: STEELE, J. H. (Org.). **Encyclopedia of Marine Sciences: Marine Biology**. London: Academic Press, 2009. p. 123-131.

KLEINER, Manuel *et al.* Transductomics: sequencing-based detection and analysis of transduced DNA in pure cultures and microbial communities. **Microbiome**, [s. l.], v. 8, n. 1, 2020.

KOLMOGOROV, Mikhail *et al.* Assembly of long, error-prone reads using repeat graphs. **Nature Biotechnology**, [s. l.], v. 37, n. 5, p. 540–546, 2019.

KRAWCZYK, Pawel S.; LIPINSKI, Leszek; DZIEMBOWSKI, Andrzej. PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. **Nucleic Acids Research**, [s. l.], v. 46, n. 6, p. E35, 2018.

KUMAR, Sudhir *et al.* MEGA X: Molecular evolutionary genetics analysis across computing platforms. **Molecular Biology and Evolution**, [s. l.], v. 35, n. 6, p. 1547–1549, 2018.

KWON, Min Jin *et al.* Beyond the Biosynthetic Gene Cluster Paradigm: Genome-Wide Coexpression Networks Connect Clustered and Unclustered Transcription Factors to Secondary Metabolic Pathways. **Microbiology Spectrum**, [s. l.], v. 9, n. 2, p. 1–16, 2021.

LALLI, CM; PARSONS, T.R. **Biological oceanography: an introduction**. 2nd ed. Oxford: Elsevier, 1997.

LALUCAT, Jorge *et al.* Genomics in bacterial taxonomy: Impact on the genus *pseudomonas*. **Genes**, [s. l.], v. 11, n. 2, 2020.

LAROCHE, Céline. Exopolysaccharides from Microalgae and Cyanobacteria: Diversity of Strains, Production Strategies, and Applications. **Marine Drugs**, [s. l.], v. 20, n. 5, 2022.

LAWRENCE, Karl P.; LONG, Paul F.; YOUNG, Antony R. Mycosporine-Like Amino Acids for Skin Photoprotection. **Current Medicinal Chemistry**, [s. l.], v. 25, n. 40, p. 5512–5527, 2017.

LEE, Sae Hong *et al.* Genetic analysis on the NifW by utilizing the yeast two-hybrid system revealed that the NifW of *Azotobacter vinelandii* interacts with the NifZ to form higher-order complexes. **Biochemical and Biophysical Research Communications**, [s. l.], v. 244, n. 2, p. 498–504, 1998.

LETUNIC, Ivica; BORK, Peer. Interactive Tree of Life (iTOL) v4: Recent updates and new developments. **Nucleic Acids Research**, [s. l.], v. 47, n. W1, 2019.

LI, Lei; JIANG, Weihong; LU, Yinhua. New strategies and approaches for engineering biosynthetic gene clusters of microbial natural products. **Biotechnology Advances**, [s. l.], v. 35, n. 8, p. 936–949, 2017. Disponível em: <https://doi.org/10.1016/j.biotechadv.2017.03.007>.

LI, Lei; MACLINTYRE, Logan W.; BRADY, Sean F. Refactoring biosynthetic gene clusters for heterologous production of microbial natural products. **Current Opinion in Biotechnology**, [s. l.], v. 69, p. 145–152, 2021. Disponível em: <https://doi.org/10.1016/j.copbio.2020.12.011>.

LIN, Jinzhong *et al.* DNA targeting by subtype I-D CRISPR-Cas shows type I and type III features. **Nucleic Acids Research**, [s. l.], v. 48, n. 18, p. 10470–10478, 2020.

LIU, Dang; HUNT, Martin; TSAI, Isheng J. Inferring synteny between genome assemblies: A systematic evaluation. **BMC Bioinformatics**, [s. l.], v. 19, n. 1, p. 1–13, 2018.

LUISA ALBARANO, ROBERTA ESPOSITO, NADIA RUOCCO, Maria Costantini. Genome Mining as New Challenge in Natural Products Discovery. **Marine drugs**, [s. l.], p. 1–17, 2020.

LUO, Chengwei; RODRIGUEZ-R, Luis M.; KONSTANTINIDIS, Konstantinos T. MyTaxa: An advanced taxonomic classifier for genomic and metagenomic sequences. **Nucleic Acids Research**, [s. l.], v. 42, n. 8, p. 1–12, 2014.

MACHADO, Henrique; TUTTLE, Robert N.; JENSEN, Paul R. **Omics-based natural product discovery and the lexicon of genome mining**. [S. l.]: Elsevier Ltd, 2017.

MAGUVU, Tawanda E.; BEZUIDENHOUT, Cornelius C. Whole genome sequencing based taxonomic classification, and comparative genomic analysis of potentially human pathogenic enterobacter spp. Isolated from chlorinated wastewater in the North West Province, South Africa. **Microorganisms**, [s. l.], v. 9, n. 9, 2021.

MATALONI, Gabriela; KOMÁREK, Jiří. *Gloeocapsopsis aurea*, a new subaerophytic cyanobacterium from maritime Antarctica. **Polar Biology**, [s. l.], v. 27, n. 10, p. 623–628, 2004.

MEDEMA, Marnix H. *et al.* Minimum Information about a Biosynthetic Gene cluster. **Nature Chemical Biology**, [s. l.], v. 11, n. 9, p. 625–631, 2015.

MEDLAR, Alan J; PETRI, T; HOLM, Liisa. AAI-profiler: fast proteome-wide exploratory analysis reveals taxonomic identity, misclassification and. [s. l.], v. 46, n. May, p. 479–485, 2018.

MENZEL, Peter; NG, Kim Lee; KROGH, Anders. Fast and sensitive taxonomic classification for metagenomics with Kaiju. **Nature Communications**, [s. l.], v. 7, p. 1–9, 2016. Disponível em: <http://dx.doi.org/10.1038/ncomms11257>.

MOHR, Wiebke *et al.* Resolution of Conflicting Signals at the Single-Cell Level in the Regulation of Cyanobacterial Photosynthesis and Nitrogen Fixation. **PLoS ONE**, [s. l.], v. 8, n. 6, p. 1–7, 2013.

MOLINA-MENOR, Esther *et al.* Microbial communities of the Mediterranean rocky shore: ecology and biotechnological potential of the sea-land transition. **Microbial Biotechnology**, [s. l.], v. 12, n. 6, p. 1359–1370, 2019.

MOREIRA, C. *et al.* Cylindrospermopsin: Occurrence, methods of detection and toxicology. **Journal of Applied Microbiology**, [s. l.], v. 114, n. 3, p. 605–620, 2013.

MUTALIPASSI, Mirko *et al.* Symbioses of cyanobacteria in marine environments: Ecological insights and biotechnological perspectives. **Marine Drugs**, [s. l.], v. 19, n. 4, p. 1–29, 2021.

NAGARKAR, Sanjay; WILLIAMS, Gray A. Spatial and temporal variation of cyanobacteria-dominated epilithic communities on a tropical shore in Hong Kong. **Phycologia**, [s. l.], v. 38, n. 5, p. 385–393, 1999.

NATHANE NUNES DE FREITAS, Paloma *et al.* Evaluation of the Toxicity of Microcyclamide Produced by *Microcystis aeruginosa* in *Danio rerio* Embryos. **Toxics**, [s. l.], v. 11, n. 2, p. 1–11, 2023.

O'CALLAGHAN, Jessica L. *et al.* Limitations of 16S rRNA Gene Sequencing to Characterize *Lactobacillus* Species in the Upper Genital Tract. **Frontiers in Cell and Developmental Biology**, [s. l.], v. 9, n. July, 2021.

OLKHOVA, Elena *et al.* Homology model of the Na⁺/proline transporter PutP of *Escherichia coli* and its functional implications. **Journal of Molecular Biology**, [s. l.], v. 406, n. 1, p. 59–74, 2011.

PAGLIARA, Patrizia *et al.* The marine sponge *Petrosia ficiformis* harbors different cyanobacteria strains with potential biotechnological application. **Journal of Marine Science and Engineering**, [s. l.], v. 8, n. 9, 2020.

PARKS, Donovan H. *et al.* CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. **Genome Research**, [s. l.], v. 25, n. 7, p. 1043–1055, 2015.

PARKS, D.H. *et al.* Selection of representative genomes for 24,706 bacterial and archaeal species clusters provide a complete genome-based taxonomy. **bioRxiv**, [s. l.], 2019.

PARVATHY, Sujatha Thankeswaran; UDAYASURIYAN, Varatharajalu; BHADANA, Vijaipal. Codon usage bias. **Molecular Biology Reports**, [s. l.], v. 49, n. 1, p. 539–565, 2022. Disponível em: <https://doi.org/10.1007/s11033-021-06749-4>.

PATHAK, Jainendra *et al.* Cyanobacterial Secondary Metabolite Scytonemin: A Potential Photoprotective and Pharmaceutical Compound. **Proceedings of the National Academy of Sciences India Section B - Biological Sciences**, [s. l.], v. 90, n. 3, p. 467–481, 2020. Disponível em: <https://doi.org/10.1007/s40011-019-01134-5>.

PATTHARAPRACHAYAKUL, Napisa *et al.* Current understanding of the cyanobacterial CRISPR-Cas systems and development of the synthetic CRISPR-Cas systems for cyanobacteria. **Enzyme and Microbial Technology**, [s. l.], v. 140, n. February, p. 109619, 2020. Disponível em: <https://doi.org/10.1016/j.enzmictec.2020.109619>.

PAYSAN-LAFOSSE, Typhaine *et al.* InterPro in 2022. **Nucleic Acids Research**, [s. l.], v. 51, n. D1, p. D418–D427, 2023.

PENTECOST, Allan. Distribution and ecology of cyanobacteria in the rocky littoral of an English Lake District Water Body, Devoke water. **Life**, [s. l.], v. 4, n. 4, p. 1026–1037, 2014.

POPE, Matthew A. *et al.* O-methyltransferase is shared between the pentose phosphate and shikimate pathways and is essential for mycosporine-like amino acid biosynthesis in *Anabaena variabilis* ATCC 29413. **ChemBioChem**, [s. l.], v. 16, n. 2, p. 320–327, 2015.

POPIN, Rafael Vicentini *et al.* Mining of Cyanobacterial Genomes Indicates Natural Product Biosynthetic Gene Clusters Located in Conjugative Plasmids. **Frontiers in Microbiology**, [s. l.], v. 12, n. November, p. 1–12, 2021.

REBOAH, Paloma *et al.* Influence of climatic factors on cyanobacteria and green algae development on building surface. **PLoS ONE**, [s. l.], v. 18, n. 3 March, p. 1–19, 2023.

REID, R. P. *et al.* The role of microbes in accretion, lamination and early lithification of modern marine stromatolites. **Nature**, [s. l.], v. 406, n. 6799, p. 989–992, 2000.

RODRIGUEZ-R, Luis M. *et al.* The Microbial Genomes Atlas (MiGA) webserver: Taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. **Nucleic Acids Research**, [s. l.], v. 46, n. W1, p. W282–W288, 2018.

ROMEU, Maria J. *et al.* Biofilm formation behaviour of marine filamentous cyanobacterial strains in controlled hydrodynamic conditions. **Environmental Microbiology**, [s. l.], v. 21, n. 11, p. 4411–4424, 2019.

ROSSELLÓ-MORA, Ramon. Updating prokaryotic taxonomy. **Journal of Bacteriology**, [s. l.], v. 187, n. 18, p. 6255–6257, 2005.

ROSSI, Federico; DE PHILIPPIS, Roberto. Role of cyanobacterial exopolysaccharides in phototrophic biofilms and in complex microbial mats. **Life**, [s. l.], v. 5, n. 2, p. 1218–1238, 2015.

ROULI, L. *et al.* The bacterial pangenome as a new tool for analysing pathogenic bacteria. **New Microbes and New Infections**, [s. l.], v. 7, p. 72–85, 2015. Disponível em: <http://dx.doi.org/10.1016/j.nmni.2015.06.005>.

ROUX, Simon *et al.* VirSorter: Mining viral signal from microbial genomic data. **PeerJ**, [s. l.], v. 2015, n. 5, p. 1–20, 2015.

SAYERS, Eric W. *et al.* Database resources of the National Center for Biotechnology Information. **Nucleic Acids Research**, [s. l.], v. 49, n. D1, p. D10–D17, 2021.

SCHREIDAH, Celine M. *et al.* **Microcystins: Biogenesis, Toxicity, Analysis, and Control**. [S. l.: s. n.], 2020-. ISSN 15205010.v. 33.

SEEMANN, Torsten. Prokka: Rapid prokaryotic genome annotation. **Bioinformatics**, [s. l.], v. 30, n. 14, p. 2068–2069, 2014.

SEYEDSAYAMDOST, Mohammad R. Toward a global picture of bacterial secondary metabolism. **Journal of Industrial Microbiology and Biotechnology**, [s. l.], v. 46, n. 3–4, p. 301–311, 2019. Disponível em: <https://doi.org/10.1007/s10295-019-02136-y>.

SHAFFER, Michael *et al.* kb_DRAM: annotation and metabolic profiling of genomes with DRAM in KBase. **Bioinformatics**, [s. l.], v. 39, n. 4, p. 1–3, 2023.

SIMAO, F A *et al.* BUSCO online supplementary information: assessing genome assembly and annotation completeness with single-copy orthologs. **Bioinformatics**, [s. l.], v. 31, n. 19, p. 1–2, 2015.

SINGH, Rahul Kunwar *et al.* Cyanobacteria: An emerging source for drug discovery. **Journal of Antibiotics**, [s. l.], v. 64, n. 6, p. 401–412, 2011.

SINGS, Heather L.; BIBLE, Keith C.; RINEHART, Kenneth L. Acyl tunichlorins: A new class of nickel chlorins isolated from the Caribbean tunicate *Trididemnum solidum*. **Proceedings of the National Academy of Sciences of the United States of America**, [s. l.], v. 93, n. 20, p. 10560–10565, 1996.

SIVONEN, Kaarina *et al.* Cyanobactins-ribosomal cyclic peptides produced by cyanobacteria. **Applied Microbiology and Biotechnology**, [s. l.], v. 86, n. 5, p. 1213–1225, 2010.

STAL, L. J.; KRUMBEIN, W. E. Temporal separation of nitrogen fixation and photosynthesis in the filamentous, non-heterocystous cyanobacterium *Oscillatoria* sp. **Archives of Microbiology**, [s. l.], v. 149, n. 1, p. 76–80, 1987.

STAMATAKIS, Alexandros. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. **Bioinformatics**, [s. l.], v. 30, n. 9, p. 1312–1313, 2014.

STIEFELMAIER, Judith *et al.* Characterization of terrestrial phototrophic biofilms of cyanobacterial species. **Algal Research**, [s. l.], v. 50, n. February, 2020.

STOTHARD, Paul; GRANT, Jason R.; VAN DOMSELAAR, Gary. Visualizing and comparing circular genomes using the CGView family of tools. **Briefings in Bioinformatics**, [s. l.], v. 20, n. 4, p. 1576–1582, 2018.

SWIA, Piotr *et al.* Report Fungal symbionts impact cyanobacterial biofilm durability and photosynthetic efficiency Report Fungal symbionts impact cyanobacterial biofilm durability and photosynthetic efficiency. [s. l.], p. 1–6, 2023.

TAMAS, Ivica *et al.* The (d)evolution of methanotrophy in the Beijerinckiaceae—a comparative genomics analysis. **ISME Journal**, [s. l.], v. 8, n. 2, p. 369–382, 2014. Disponível em: <http://dx.doi.org/10.1038/ismej.2013.145>.

TAN, Lik Tong. Bioactive natural products from marine cyanobacteria for drug discovery. **Phytochemistry**, [s. l.], v. 68, n. 7, p. 954–979, 2007.

TANIZAWA, Yasuhiro *et al.* DNA Data Bank of Japan (DDBJ) update report 2022. **Nucleic Acids Research**, [s. l.], v. 51, n. D1, p. D101–D105, 2023.

TESSON, Florian *et al.* Systematic and quantitative view of the antiviral arsenal of prokaryotes. **Nature Communications**, [s. l.], v. 13, n. 1, 2022.

THEYS, Kristof *et al.* Advances in Visualization Tools for Phylogenomic and Phylodynamic Studies of Viral Diseases. **Frontiers in Public Health**, [s. l.], v. 7, n. August, p. 1–18, 2019.

VAN DE SANDE, Bram *et al.* Applications of single-cell RNA sequencing in drug discovery and development. **Nature Reviews Drug Discovery**, [s. l.], v. 22, n. 6, p. 496–520, 2023.

VAN GOETHEM, Marc W. *et al.* Cyanobacteria and Alphaproteobacteria may facilitate cooperative interactions in niche communities. **Frontiers in Microbiology**, [s. l.], v. 8, n. OCT, p. 1–11, 2017.

VÁZQUEZ-MARTÍNEZ, Guadalupe *et al.* Strategy to obtain axenic cultures from field-collected samples of the cyanobacterium *Phormidium animalis*. **Journal of Microbiological Methods**, [s. l.], v. 57, n. 1, p. 115–121, 2004.

VELÁZQUEZ, Encarna *et al.* The coexistence of symbiosis and pathogenicity-determining genes in *Rhizobium rhizogenes* strains enables them to induce nodules and tumors or hairy roots in plants. **Molecular Plant-Microbe Interactions**, [s. l.], v. 18, n. 12, p. 1325–1332, 2005.

VERSTRAETE, Brecht *et al.* Symbiotic β -Proteobacteria beyond Legumes: Burkholderia in Rubiaceae. **PLoS ONE**, [s. l.], v. 8, n. 1, 2013.

VILLASTRIGO, Adrián *et al.* Oceanic currents maintain the genetic structure of non-marine coastal taxa in the western Mediterranean Sea. [s. l.], p. 1–9.

WATERBURY, John B. The Cyanobacteria—Isolation, Purification and Identification. **The Prokaryotes**, [s. l.], p. 1053–1073, 2006.

WEBER, Carolyn F.; KING, Gary M. Physiological, ecological, and phylogenetic characterization of *Stappia*, a marine CO-oxidizing bacterial genus. **Applied and Environmental Microbiology**, [s. l.], v. 73, n. 4, p. 1266–1276, 2007.

WEIGEL, Brooke L. *et al.* Functional Insights into the Kelp Microbiome from Metagenome-Assembled Genomes. **mSystems**, [s. l.], v. 7, n. 3, p. 1–18, 2022.

WHITEHEAD, K.; KARENTZ, D.; HEDGES, J. Mycosporine-like amino acids (MAAs) in phytoplankton, a herbivorous pteropod (*Limacina helicina*), and its pteropod predator (*Clione antarctica*) in McMurdo Bay, Antarctica. **Marine Biology**, [s. l.], v. 139, n. 5, p. 1013–1019, 2001.

WICK, Ryan R. *et al.* Bandage: Interactive visualization of de novo genome assemblies. **Bioinformatics**, [s. l.], v. 31, n. 20, p. 3350–3352, 2015.

WOOD, Susanna A.; BIESSY, Laura; PUDDICK, Jonathan. Anatoxins are consistently released into the water of streams with *Microcoleus autumnalis*-dominated (cyanobacteria) proliferations. **Harmful Algae**, [s. l.], v. 80, n. October, p. 88–95, 2018. Disponível em: <https://doi.org/10.1016/j.hal.2018.10.001>.

WU, Guimei *et al.* Euendolithic Cyanobacteria and Proteobacteria Together Contribute to Trigger Bioerosion in Aquatic Environments. **Frontiers in Microbiology**, [s. l.], v. 13, n. July, p. 1–12, 2022.

XIA, Liming *et al.* Biosynthetic gene cluster profiling predicts the positive association between antagonism and phylogeny in *Bacillus*. **Nature Communications**, [s. l.], v. 13, n. 1, p. 1–11, 2022.

XIAO, Tiantian; ZHOU, Wenhao. The third generation sequencing: The advanced approach to genetic diseases. **Translational Pediatrics**, [s. l.], v. 9, n. 2, p. 163–173, 2020.

XU, Ling *et al.* OrthoVenn2: A web server for whole-genome comparison and annotation of orthologous clusters across multiple species. **Nucleic Acids Research**, [s. l.], v. 47, n. W1, p. W52–W58, 2019.

ZHAO, Fangzhou *et al.* Genome-wide role of codon usage on transcription and identification of potential regulators. **Proceedings of the National Academy of Sciences of the United States of America**, [s. l.], v. 118, n. 6, 2021.

ZIEMERT, Nadine *et al.* Microcyclamide biosynthesis in two strains of *Microcystis aeruginosa*: From structure to genes and vice versa. **Applied and Environmental Microbiology**, [s. l.], v. 74, n. 6, p. 1791–1797, 2008.