



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

ANA CAROLINA FAVACHO MIRANDA DE OLIVEIRA

**POTENCIAL BIOSINTÉTICO DE CEPAS BACTERIANAS ISOLADAS DE
AMOSTRAS DE SOLO DO PARQUE ESTADUAL UTINGA, PARÁ, BRASIL**

BELÉM, 2024

ANA CAROLINA FAVACHO MIRANDA DE OLIVEIRA

**POTENCIAL BIOSINTÉTICO DE CEPAS BACTERIANAS ISOLADAS DE
AMOSTRAS DE SOLO DO PARQUE ESTADUAL UTINGA, PARÁ, BRASIL**

Tese de Doutorado apresentada ao Programa de Pós-graduação em Genética e Biologia Molecular da Universidade Federal do Pará como parte dos requisitos exigidos para a obtenção do grau de Doutora em Genética e Biologia Molecular.

Orientadora: Dra. Maria Paula Cruz Schneider

Coorientada: Dra. Daniela Barreto Barbosa Trivella

BELÉM, 2024

FICHA CATALOGRÁFICA

**Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)**

F272 Favacho Miranda de Oliveira, Ana Carolina.
POTENCIAL BIOSSINTÉTICO DE CEPAS BACTERIANAS
ISOLADAS DE AMOSTRAS DE SOLO DO PARQUE
ESTADUAL UTINGA, PARÁ, BRASIL / Ana Carolina Favacho
Miranda de Oliveira. — 2024.
88 f. : il. color.

Orientador(a): Prof^a. Dra. Maria Paula Cruz Schneider
Coorientação: Prof^a. Dra. Daniela Barreto Barbosa Trivella
Tese (Doutorado) - Universidade Federal do Pará, Instituto de
Ciências Biológicas, Programa de Pós-Graduação em Genética e
Biologia Molecular, Belém, 2024.

1. Metabolismo secundário. 2. Agrupamentos de genes
biossintéticos. 3. Metabolômica não direcionada. I. Título.

CDD 571.29

ANA CAROLINA FAVACHO MIRANDA DE OLIVEIRA

**POTENCIAL BIOSSINTÉTICO DE CEPAS BACTERIANAS ISOLADAS DE
AMOSTRAS DE SOLO DO PARQUE ESTADUAL UTINGA, PARÁ, BRASIL**

Data de apresentação: 22/04/2024

Belém, PA

BANCA EXAMINADORA

Artur Luiz da Costa da Silva
Universidade Federal do Pará (UFPA)
Membro Interno

Vivian Helena Pellizari
Universidade de São Paulo (USP)
Membro Interno

Agenor Valadares Santos
Universidade Federal do Pará (UFPA)
Membro Externo

Isabel Henriques
Universidade de Coimbra (UC)
Membro Externo

Rafael Azevedo Baraúna
Universidade Federal do Pará (UFPA)
Suplente

DEDICATÓRIA

Ao meu pai, de onde estiver, espero que esteja orgulhoso de mim.

AGRADECIMENTOS

Agradeço aos meus pais, Maria Lucely e Mário, que sempre me incentivaram e apoiaram a seguir o caminho dos estudos. À minha madrinha, Nely, por ter sido uma segunda mãe e ter contribuído ao longo da minha vida com a minha formação. E a todos meus familiares que sempre acreditaram no meu potencial.

Ao meu esposo, Alisson, por está comigo em todos os altos e baixos durante minha jornada acadêmica, por toda a compreensão e apoio em etapas cruciais, por todas as conversas e conselhos e por acreditar em mim, muito mais do que eu mesma. É um privilégio compartilhar mais uma etapa da vida com você.

À minha orientadora, Dra. Maria Paula Schneider, por me acolher como sua aluna em seu laboratório e por toda a confiança depositada em mim e no meu trabalho.

À minha coorientadora, Dra. Daniela Trivella, por me receber em seu grupo no CNPEM, por todas as orientações e ensinamentos e por confiar no meu trabalho.

Ao Dr. Rafael Baraúna, por toda a colaboração neste trabalho e por me receber em seu grupo no Engbio.

À todas as amigas que fiz durante esse período. Especialmente a Bruna Vieira e Débora Munhoz, que além de colegas de trabalho no CNPEM, são grandes amigas que eu levarei para sempre em meu coração. Aos queridos Johnes, Sávio, Silvanira e Soraya que tornaram os dias no Engbio mais divertidos. E a todos os colegas dos laboratórios LNBio, EngBio e CGBS agradeço por todos os conselhos e conhecimentos compartilhados.

Por fim, agradeço a todas as fontes financiadoras que tornaram esse trabalho possível: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Pró-Reitoria de Pesquisa e Pós-Graduação (PROPESP) da UFPA e ao Instituto Serrapilheira.

RESUMO

Os metabólitos secundários produzidos pelas bactérias constituem um rico repertório químico e têm sido amplamente aplicados na medicina, agricultura, indústrias cosméticas e alimentícias. Os avanços nas análises multi-ômicas contribuem para compreensão da complexidade bioquímica. A Amazônia é considerada uma importante fonte de biodiversidade. No entanto, poucos estudos foram realizados para descrever as comunidades microbianas dos habitats amazônicos, e menos ainda foram realizados para explorar o potencial biotecnológico dos produtos naturais sintetizados por essas comunidades microbianas. Diante disto, nosso estudo teve como objetivo isolar bactérias de solo amazônico e analisar o potencial biossintético na produção de metabólitos secundários. Foi utilizada uma abordagem multi-ômica baseada na mineração do genoma e na metabolômica não direcionada. Foram selecionadas cepas pertencentes ao filo Actinobacteria e/ou que apresentassem atividade inibitória durante o isolamento. O sequenciamento do gene 16S rRNA classificou as cepas ACT015, ACT016 e FIR094 nos gêneros *Streptomyces*, *Rhodococcus* e *Brevibacillus*, respectivamente. Um total de 33, 17 e 14 BGCs foram encontrados no genoma de *Streptomyces* sp. ACT015, *Rhodococcus* sp. ACT016 e *Brevibacillus* sp. FIR094, respectivamente. Diversas vias enzimáticas foram detectadas, incluindo BGCs que produzem antibióticos e compostos antitumorais. No entanto, um total de 40 BGCs (62,5%) foram relacionados a metabólitos com função desconhecida. Ao aplicar a abordagem OSMAC (uma cepa, muitos compostos) para o cultivo de cepas, a investigação metabolômica não direcionada revelou uma miríade de metabólitos produzidos por essas cepas em condições de laboratório. *Streptomyces* sp. ACT015 produziu muitos metabólitos de diferentes classes nos meios ISP2 e A1, enquanto *Rhodococcus* sp. ACT016 apresentou maior produção em A1 e TSB, e *Brevibacillus* sp. FIR094 em meio mínimo R2A. Nossos dados demonstraram uma vasta diversidade química inexplorada em bactérias isoladas de regiões conservadas da Amazônia. A abordagem multi-ômica foi eficaz na descrição do metabolismo secundário dos isolados e na prospecção de BGCs de interesse biotecnológico para estudos futuros.

Palavras-chave: Metabolismo secundário; agrupamentos de genes biossintéticos; metabolômica não direcionada; *Streptomyces*; *Rhodococcus*; *Brevibacillus*.

ABSTRACT

The secondary metabolites produced by bacteria constitute a rich chemical repertoire and have been widely applied in medicine, agriculture, cosmetic and food industries. Advances in multi-omics analyzes contribute to understanding biochemical complexity. The Amazon is considered an important source of biodiversity. However, few studies have been carried out to describe the microbial communities of Amazonian habitats, and even fewer have been carried out to explore the biotechnological potential of natural products synthesized by these microbial communities. Therefore, our study aimed to isolate bacteria from Amazonian soil and analyze the biosynthetic potential in the production of secondary metabolites. A multi-omics approach based on genome mining and untargeted metabolomics was used. Strains belonging to the Actinobacteria phyla and/or that presented inhibitory activity during isolation were selected. Sequencing of the 16S rRNA gene classified strains ACT015, ACT016 and FIR094 into the genera *Streptomyces*, *Rhodococcus* and *Brevibacillus*, respectively. A total of 33, 17 and 14 BGCs were found in the genome of *Streptomyces* sp. ACT015, *Rhodococcus* sp. ACT016 and *Brevibacillus* sp. FIR094, respectively. Several enzymatic pathways have been detected, including BGCs that produce antibiotics and antitumor compounds. However, a total of 40 BGCs (62.5%) were related to metabolites with unknown function. By applying the OSMAC (one strain many compounds) approach for strain cultivation, untargeted metabolomics investigation revealed a myriad of metabolites produced by these strains under laboratory conditions. *Streptomyces* sp. ACT015 produced many metabolites of different classes in ISP2 and A1 media, while *Rhodococcus* sp. ACT016 showed higher production in A1 and TSB, and *Brevibacillus* sp. FIR094 in R2A minimal medium. Our data demonstrated a vast unexplored chemical diversity in bacteria isolated from conserved regions of the Amazon. The multi-omics approach was effective in describing the secondary metabolism of isolates and in prospecting BGCs of biotechnological interest for future studies.

Keywords: Secondary metabolism; Biosynthetic gene clusters; Untargeted metabolomics; *Streptomyces*; *Rhodococcus*; *Brevibacillus*.

LISTA DE FIGURAS

Figura 1: Representação geográfica da Amazônia Legal brasileira	11
Figura 2: Representação esquemática da estrutura de agrupamentos de genes biossintéticos	14
Figura 3: Representação esquemática da biossíntese de NRP	16
Figura 4: Representação esquemática da biossíntese de PK	17
Figura 5: Representação esquemática da biossíntese de RiPP	19
Figura 6: Representação gráfica da formação de rede molecular clássica	23

SUMÁRIO

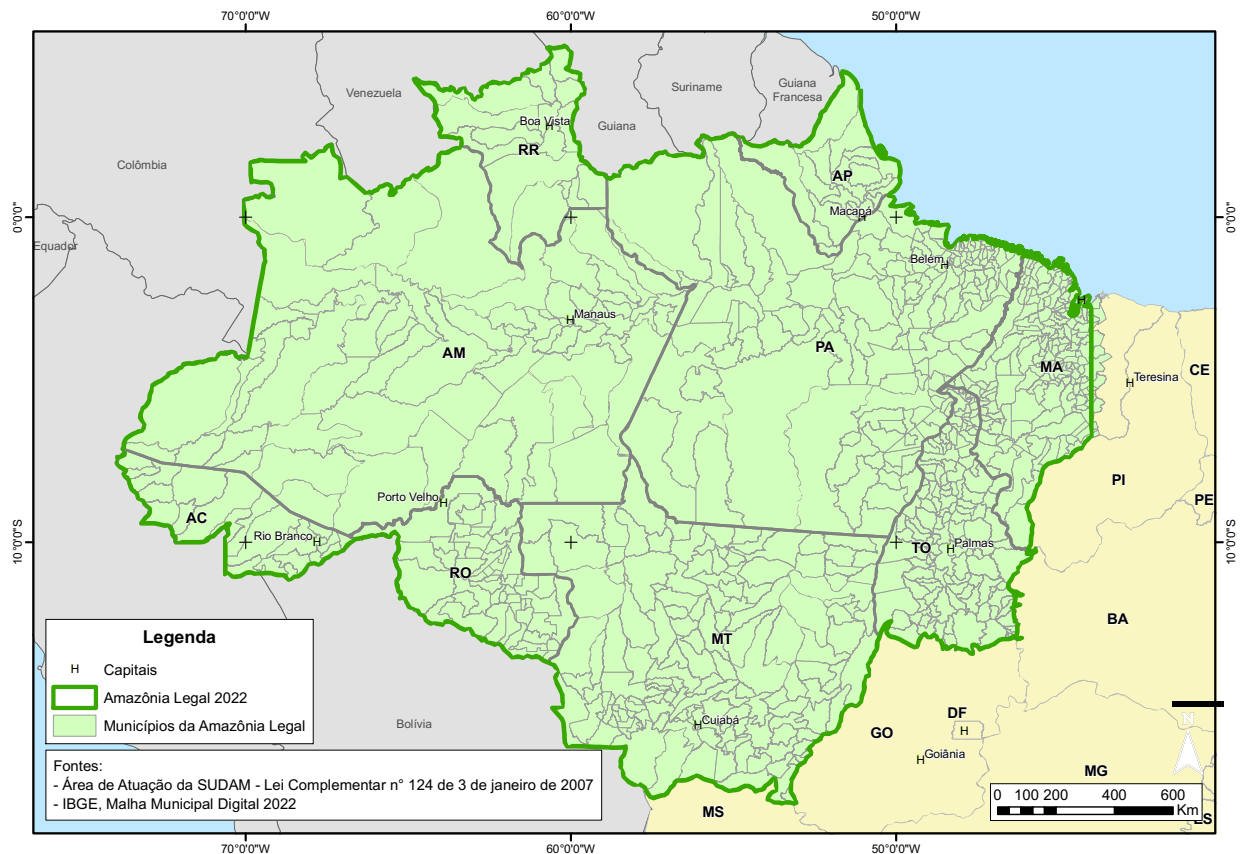
1 INTRODUÇÃO.....	11
1.1 BIODIVERSIDADE AMAZÔNICA	11
1.2 METABÓLITOS SECUNDÁRIOS BACTERIANOS.....	14
1.3 TECNOLOGIAS NA DESCOBERTA DE PRODUTOS NATURAIS BACTERIANOS	21
1.4 GÊNERO STREPTOMYCES	25
1.5 GÊNERO RHODOCOCCLUS	27
1.6 GÊNERO BREVIACILLUS.....	29
2 OBJETIVOS.....	31
2.1 OBJETIVO GERAL	31
2.2 OBJETIVOS ESPECÍFICOS.....	31
3 MATERIAL E MÉTODOS.....	32
3.1 AMOSTRAGEM E CONSIDERAÇÕES ÉTICAS	32
3.2 ISOLAMENTO, CLASSIFICAÇÃO TAXONÔMICA E SELEÇÃO BACTERIANA.....	32
3.3 SEQUENCIAMENTO E MINERAÇÃO DO GENOMA	33
3.4 EXTRAÇÃO DE METABÓLITOS DOS CULTIVOS BACTERIANOS	34
3.5 COLETA E ANÁLISE DE DADOS POR UPLC-MS/MS	35
3.6 PROCESSAMENTO E ANÁLISE DE DADOS NP ³ MS WORKFLOW	35
5 RESULTADOS E DISCUSSÃO.....	37
6 CONCLUSÃO.....	38
REFERÊNCIAS	39

1 INTRODUÇÃO

1.1 BIODIVERSIDADE AMAZÔNICA

A Amazônia Legal abrange cerca de 5 milhões de quilômetros quadrados, dos quais 84% estão dentro do território brasileiro (Figura 1). A biodiversidade da Amazônia está distribuída em diferentes ambientes representados por floresta de terra firme, floresta de igapó e várzeas, tornando o Brasil um dos países com maior biodiversidade do planeta. Há áreas de florestas mais conservadas, áreas em transição e áreas convertidas. Essa imensa diversidade de formas de vida fornece recursos genéticos únicos, adaptações e funções ecológicas que contribuem globalmente em diversos ecossistêmicos (GUAYASAMIN et al., 2021).

Figura 1: Representação geográfica da Amazônia Legal brasileira



Fonte: Instituto Brasileiro de Geografia e Estatística (IBGE), 2024.

A biodiversidade amazônica é amplamente reconhecida como um tesouro medicinal e o maior dispensário de medicamentos do mundo (MCCHESNEY; VENKATARAMAN; HENRI, 2007). Os princípios ativos encontrados em diversas fontes, como extratos de plantas e derivados, são reconhecidos mundialmente. Dentre esses exemplos, destaca-se o óleo de copaíba, extraído da copaibeira (*Copaifera langsdorffii*) originária da Amazônia brasileira e conhecido por suas propriedades anti-inflamatórias e cicatrizantes (OLIVEIRA et al., 2020); e o jaborandi (*Pilocarpus microphyllus*), tradicionalmente utilizado para tratar úlceras na boca, resfriados e gripe. O interesse nas propriedades químicas do jaborandi permitiu a descoberta do alcaloide imidazol, a partir de sais de pilocarpina extraídos das folhas da planta, e empregado na fabricação de um colírio recomendado para o controle do glaucoma (SKIRYCZ et al., 2016).

Uma parte significativa dos medicamentos atualmente em uso clínico são de origem natural. No entanto, a exploração extensiva de plantas medicinais na busca por novos medicamentos levou à preocupação com a possível ameaça ou até mesmo extinção dessa flora (VOLENZO; ODIYO, 2020). Um exemplo que representa esta limitação da descoberta de drogas a partir de produtos naturais é o Paclitaxel (Taxol), um medicamento quimioterápico utilizado no tratamento de cânceres de mama, ovário e pulmão, e desenvolvido a partir do extrato das cascas da árvore do teixo-do-pacífico (*Taxus brevifolia*). A árvore de teixo tem desenvolvimento lento e são necessárias cerca de 3.000 árvores para obtenção de 1 kg deste composto, tornando o processo insustentável do ponto de vista econômico e ambiental (JUNJUA et al., 2015).

Diante desse cenário, a busca por outras fontes de descoberta de fármacos tornou-se essencial. Entre os modelos promissores estão as bactérias ambientais, que poderiam ser fontes de bioprospecção de produtos naturais com características únicas como moléculas ou modelos de medicamentos (SCHNEIDER, 2021). A microbiota associada a Amazônia é abundante, considerando que toda a cobertura vegetal natural do solo está cheia de microrganismos colonizando hospedeiros e outros microhabitats como água e animais (PEREIRA et al., 2017). Além disso, os microrganismos podem ser manipulados com facilidade, tanto fisicamente quanto geneticamente, para a descoberta e desenvolvimento de produtos naturais (ELSEBAI; TEJESVI; PIRTTILÄ, 2014). Apesar do enorme patrimônio genético, pouco se sabe sobre a diversidade microbiana da Amazônia e suas relações ecológicas (PEREIRA et al., 2017). O conhecimento das espécies que compõem os ecossistemas da Amazônia é fundamental para uma melhor compreensão

das suas funções e das propriedades que surgem das interações entre as espécies amazônicas e seus ambientes (MALHI et al., 2008).

A biodiversidade é comumente citada como um importante ativo comercial, pois vários produtos naturais têm potencial para diferentes aplicações, incluindo a descoberta de novos medicamentos. Segundo a Organização das Nações Unidas para Agricultura e Alimentação (FAO), bioeconomia é a produção, utilização e conservação de recursos biológicos, incluindo os conhecimentos relacionados, ciência, tecnologia e inovação, para fornecer informações, produtos, processos e serviços em todos os setores econômicos, visando uma economia sustentável (FAO, 2021).

No Brasil, a bioeconomia também tem sido um importante objeto de estudos por diversas instituições. Estruturas criadas em 2021, como o Observatório de Conhecimento e Inovação em Bioeconomia, da Fundação Getúlio Vargas (FGV), e o Observatório de Bioeconomia, do Centro de Gestão e Estudos Estratégicos (CGEE), têm promovido diversos eventos sobre o tema na Amazônia. Além disso, a elaboração de produtos biotecnológicos baseados na biodiversidade exige uma discussão sobre os direitos de propriedade intelectual. Nesse contexto, em 2015 o governo brasileiro promulgou a Lei nº 13.123 estabelecendo disposições para o acesso ao patrimônio genético, que trata do compartilhamento de benefícios visando à conservação e uso sustentável da biodiversidade. A lei resultou na criação do Sistema Nacional de Gestão do Patrimônio Genético e Conhecimento Tradicional Associado, conhecido como SisGen. A implementação desse arcabouço busca criar um ambiente favorável ao desenvolvimento de produtos derivados da biodiversidade brasileira, garantindo direitos a todos os envolvidos (CASTRO BRAGA, 2021).

A biodiversidade amazônica e a bioeconomia convergem para o desenvolvimento de uma economia de base biológica e sustentável, utilizando biotecnologia e inovação para criar produtos, processos e serviços em diversas indústrias, como a química, farmacêutica e cosmética. Um dos principais objetivos dessa estratégia é a substituição de insumos industriais, agrotóxicos e combustíveis fósseis por matérias-primas de origem biológica. Essa abordagem é amplamente apoiada pelo Ministério da Ciência, Tecnologia e Inovações (MCTI), Confederação Nacional da Indústria (CNI), Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) e Ministério da Agricultura, Pecuária e Abastecimento (MAPA). Assim, espera-se que o setor produtivo seja incentivado a explorar e desenvolver produtos a partir da rica biodiversidade nacional (YANG et al., 2021).

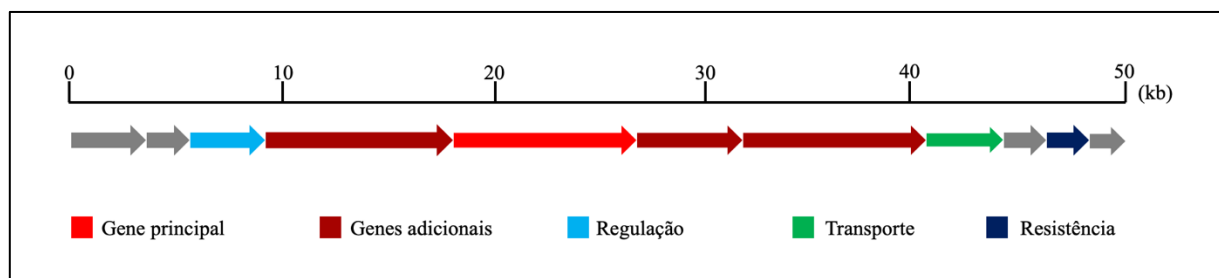
1.2 METABÓLITOS SECUNDÁRIOS BACTERIANOS

O metabolismo pode ser definido como a soma de todas as reações bioquímicas realizadas por um organismo. Os metabólitos são os intermediários e produtos do metabolismo e geralmente são restritos a pequenas moléculas. Diferentes dos metabólitos primários, que são essenciais para a viabilidade do organismo produtor, os metabólitos secundários são moléculas heterogêneas não essenciais, mas que conferem diversas vantagens adaptativas e competitivas (CRANEY; AHMED; NODWELL, 2013). Estas substâncias podem ser sintetizadas de maneira específica por espécies ou gêneros individuais, impulsionadas por razões fisiológicas, como enzimas degradativas, ou por estratégias predatórias específicas, como toxinas e antibióticos (O'BRIEN; WRIGHT, 2011).

Em 1929, a descoberta da penicilina a partir do fungo *Penicillium notatum* marcou o início de uma era revolucionária conhecida como “Era de Ouro dos Antibióticos”. Isto não apenas transformou a medicina, como intensificou a busca por outros microrganismos como fontes promissoras de agentes bioativos. Nesse período, várias classes de medicamentos foram desenvolvidas, como as cefalosporinas, tetraciclínas, aminoglicosídeos, macrolídeos, quinolonas e as sulfonamidas (CRAGG; NEWMAN, 2013).

Metabólitos secundários são codificados por agrupamentos de genes biossintéticos (BGC, do inglês *Biosynthetic Gene Cluster*), que compreendem conjuntos gênicos localmente agrupados que, além do gene de biossíntese principal, possuem genes para regulação, transporte e resistência (Figura 2). O tamanho dos BGCs varia dependendo da complexidade e da maquinaria biossintética do produto da via. A maioria dos BGCs identificados nos genomas das bactérias estão silenciados ou crípticos em condições normais de laboratório e, portanto, tornaram-se os principais alvos na exploração da diversidade química de microrganismos (PETTIT, 2011).

Figura 2: Representação esquemática da estrutura de agrupamentos de genes biossintéticos



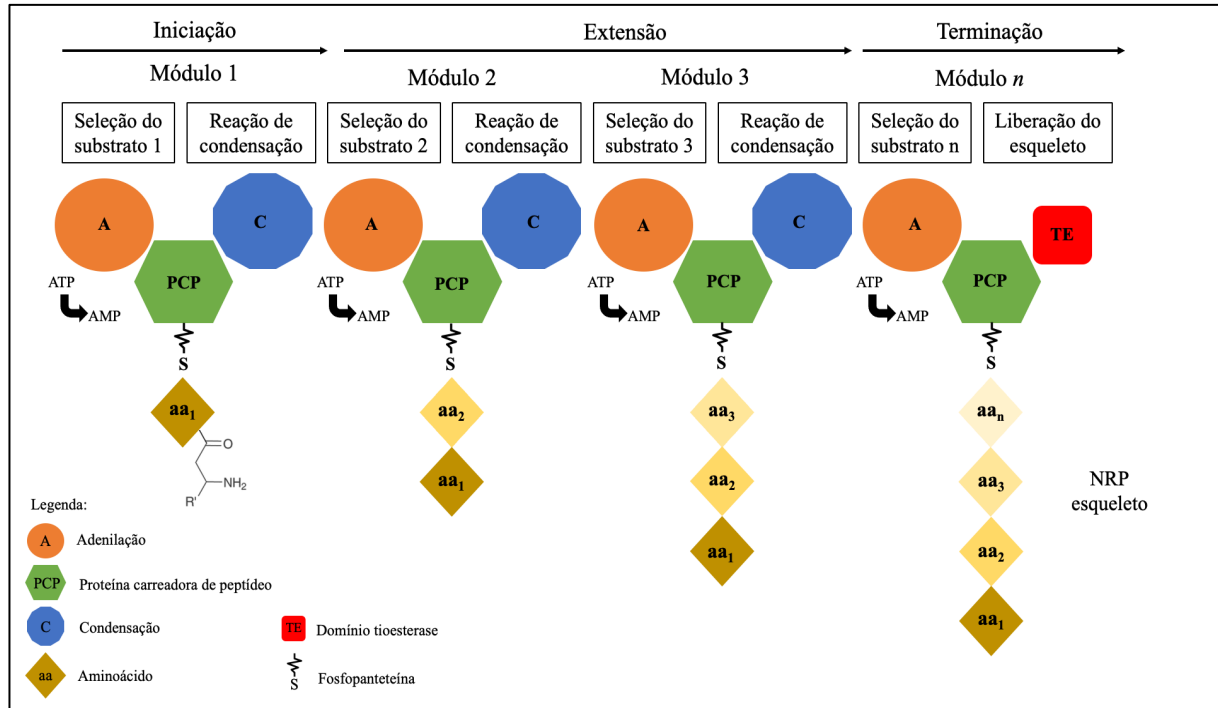
Embora a diversidade de metabólitos secundários seja enorme, as enzimas biossintéticas são altamente conservadas. Entre as principais classes estão as sintetases peptídicas não ribossomais (NRPS – do inglês *Nonribosomal Peptide Synthetase*) e as sintases de policetídeos (PKS – do inglês *Polyketide Synthase*) (CRAGG; NEWMAN, 2013). Os domínios enzimáticos são frequentemente usados como alvos na investigação de BGCs de NRPS e PKS, e sua abundância no genoma pode ser um indicador do potencial biossintético dos microrganismos (REDDY et al., 2012).

A presença de enzimas NRPS foi relatada nos três domínios da vida. No domínio das bactérias, os filos Proteobacteria, Actinobacteria, Firmicutes e Cyanobacteria possuem maior abundância dessas enzimas (WANG et al., 2014a). As NRPSs foram descobertas em 1950 em estudos sobre biossíntese de tirocidina e gramicidina, quando foi observada que sua síntese é independente da maquinaria ribossomal. Desde então, análises enzimológicas de NRPS continuam a ser de interesse significativo para o desenvolvimento de novos medicamentos, dadas as suas inúmeras aplicações, incluindo antimicrobianas (vancomicina e bacitracina), imunossupressoras (ciclosporina) e antitumorais (bleomicina) (IACOVELLI; BOVENBERG; DRIESSEN, 2021).

A arquitetura modular das NRPSs, composta por grupos repetidos de domínios catalíticos facilita a instalação de diversos monômeros durante a montagem da cadeia (IZORÉ et al., 2021). No entanto, enzimas NRPS não modulares são encontradas em vias biossintéticas de sideróforos, como *EntE* e *VibH* na enterobactina ou como uma proteína portadora de peptídeo independente, como *BlmI* da bleomicina (CROSA; WALSH, 2002; MARTÍNEZ-NÚÑEZ; LÓPEZ, 2016).

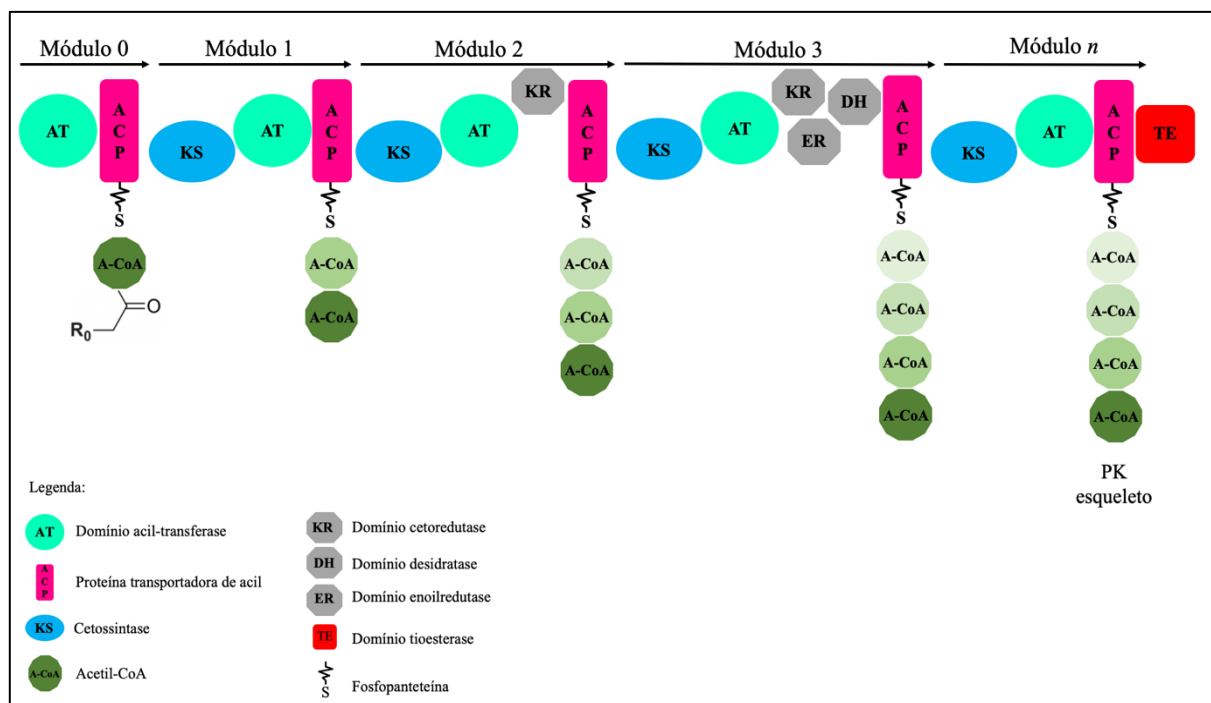
Em um módulo mínimo de extensão de cadeia, o domínio de adenilação (A), dependente de ATP (adenosina trifosfato), ativa e seleciona os blocos de construção de aminoácidos, transfere a molécula para o domínio da proteína carreadora de peptídeo (PCP) no qual ocorre a catálise da ligação peptídica pela ação do domínio de condensação (C) e o esqueleto é liberado da enzima pela ação do domínio tioesterase (TE), presente no módulo de terminação (Figura 3) (SÜSSMUTH; MAINZ, 2017). A síntese prossegue em uma direção N- para C-terminal, produzindo peptídeos que geralmente têm cerca de 3 a 15 aminoácidos de comprimento, e os peptídeos liberados podem ser lineares, cíclicos ou ramificados (MOOTZ; SCHWARZER; MARAHIEL, 2002).

Figura 3: Representação esquemática da biossíntese de NRP



As PKSs são responsáveis pela produção de produtos naturais significativos na clínica médica, como doxorubicina, antraciclina, eritromicina e tetraciclina (DAS; KHOSLA, 2009). A biossíntese inicia com a transferência de um grupo acil (A-CoA) na proteína transportadora de acil (ACP) catalisada pelo domínio acil-transferase (AT) para a cetossintase (KS). As PKS II possuem dois domínios KS (KS α e KS β), que alongam a cadeia de carbono por condensação e modificam estruturas de poliketídeos por domínios adicionais, como domínios cetoreductase (KR), desidratase (DH) e enoilreductase (ER). Por fim, o domínio tioesterase (TE) hidrolisa a cadeia de poliketídeo do domínio ACP para terminar o alongamento (Figura 4) (WANG et al., 2020).

Figura 4: Representação esquemática da biossíntese de PK



Apesar do processo de síntese semelhante, as PKSs são classificadas em PKS tipo I, II e III (PIEL, 2010). As PKS tipo I são peptídeos multifuncionais com domínios lineares e fundidos, usados repetidamente para catalisar várias rodadas de alongamento (KEATINGE-CLAY, 2012); as PKSs Tipo II são complexos multienzimáticos compostos de proteínas monofuncionais independentes. Cada enzima realiza uma função específica, como a condensação de unidades acílicas (HERTWECK et al., 2007); e as PKS tipo III são mais simples em termos de estrutura, consistindo em monômeros individuais ou pequenos complexos. Elas são frequentemente representadas por proteínas homodiméricas (SHIMIZU; OGATA; GOTO, 2017). Essas distinções refletem estratégias adaptativas nos microrganismos para sintetizar policetídeos, proporcionando diversidade de produtos químicos (DAS; KHOSLA, 2009; RIDLEY; LEE; KHOSLA, 2008).

Nas últimas décadas, avanços em pesquisas biossintéticas e enzimológicas contribuem para a identificação de classes adicionais de produtos naturais. Entre essas classes destacam-se os terpenos e peptídeos sintetizados por ribossomo e modificados pós-tradução (RiPP, do inglês *Ribosomally Synthesized and Posttranslationally Modified Peptide*) (ARNISON et al., 2013b).

A diversificação evolutiva das vias biossintéticas dos terpenos levou à formação de uma das classes de metabólitos especializados mais extensa e estruturalmente variada. Geralmente são considerados metabólitos vegetais ou fúngicos, embora hoje se saiba que genes que codificam terpenos sintases mostraram ser amplamente distribuídos em bactérias e representam uma importante fonte para a descoberta de novos produtos naturais (YAMADA et al., 2015). (HELFRICH et al., 2019).

A nomenclatura dos terpenos é baseada no número de unidades de carbono, que compõem seus esqueletos: monoterpeno (C10), sesquiterpeno (C15), diterpeno (C20) triterpeno (C30), tetraterpeno (C40) e politerpeno (> C40) (WANG et al., 2013). Esses compostos desempenham papéis tanto no metabolismo primário quanto no secundário, incluindo esteroides (colesterol), vitaminas (A, D, E e K), aromas e fragrâncias (mentol, limoneno e pineno), hormônios vegetais (giberelinas) e pigmentos fotossintéticos (clorofila), além de fármacos importantes (Taxol e artemisinina) (RUDOLF et al., 2021). Em estreptomicetos, é comum a detecção de terpenos odoríferos voláteis, como a geosmina, o 2-metilisoborneol (2-MIB) e a albaflavenona (WILKINS; SCHÖLLER, 2009).

A terpeno sintetase (TPS) é a enzima que catalisa a formação de terpenóides. Com base nas diferenças nas sequências de aminoácidos, a família de genes TPS pode ser dividida em sete subfamílias: TPS-a, TPS-b, TPS-c, TPS-d, TPS-e/f, TPS-g e TPS-h. Baseado em suas estruturas e mecanismos catalíticos, os TPS podem ser divididos em duas categorias (Classe I e Classe II). A classe I consiste em terpenos sintases dependentes da ionização, e a perda de grupos pirofosfato no substrato depende da sinergia de íons metálicos, como Mg^{2+} e Mn^{2+} ; e a classe II são terpenos sintases dependentes de prótons (LI et al., 2023).

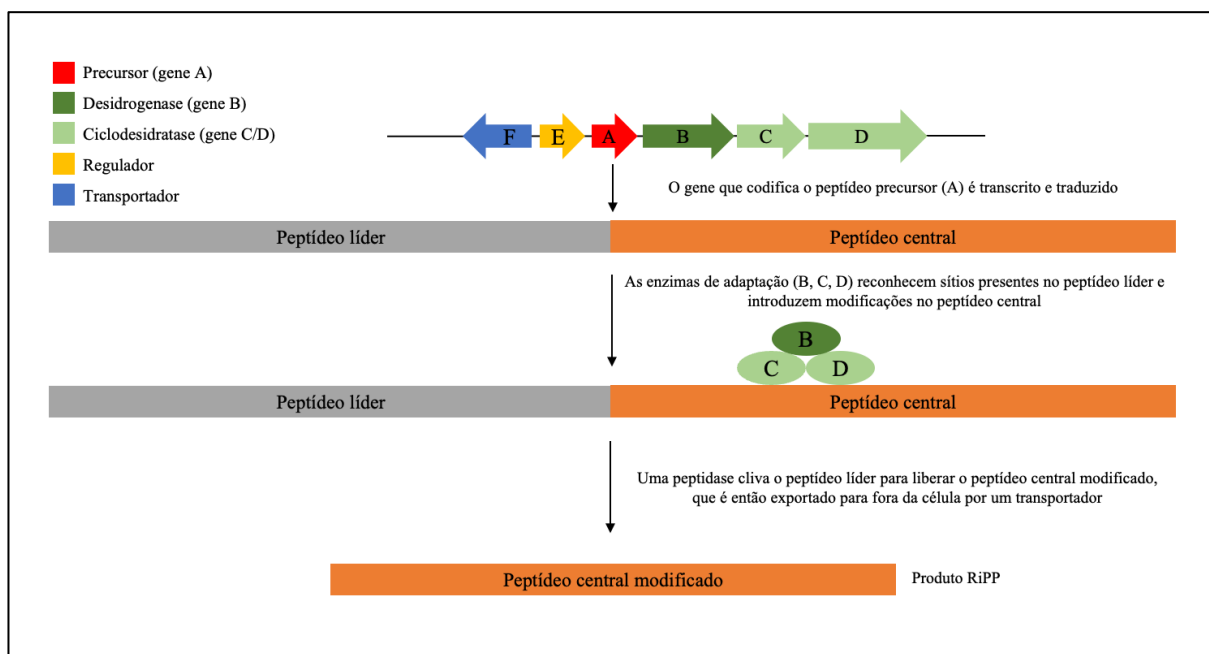
A via biossintética dos terpenos é dividida em três etapas: (i) inicia pela formação dos precursores difosfato de isopreno (IPP) e seu isômero dimetilalil difosfato (DMAPP) a partir da via do ácido mevalônico (MVA) ou via de metileritritol-4-fosfato (MEP), (ii) a condensação de IPP e DMAPP produz difosfato de geranilo (GPP), o precursor do monoterpenóide; a adição sucessiva de IPP ou DMAP resulta na formação de precursores de sesquiterpenóide e diterpenóide, que são farnesil difosfato (FPP) e geranilgeranil difosfato (GGPP); duas moléculas de FPP e GGPP são, respectivamente, condensadas para formar esqualeno e fitoeno, os precursores do triterpenóide e do tetraterpenóide (iii) por fim, esses precursores linearizados podem sofrer ciclizações ou pós-

modificações (oxidação, acetilação), fornecendo terpenos com várias estruturas, propriedades químicas complexas e características funcionais únicas (WANG; QUAN; XIAO, 2019).

Os peptídeos sintetizados por ribossomo e modificados pós-tradução (RiPPs) compõem um grupo de produtos naturais estruturalmente diversos e despertou o interesse por apresentar potentes atividades biológicas e baixa complexidade genética associada à biossíntese tornando-os excelentes candidatos para aplicações na biologia sintética (ORTEGA; VAN DER DONK, 2016). Com base em suas máquinas biossintéticas e características estruturais, os RiPPs são agrupados em subfamílias, como os lantipeptídeos, peptídeos lineares contendo azol (LAP), cianobactinas, tiopeptídeos, microcinas e peptídeos laço (ARNISON et al., 2013a).

A biossíntese de RiPP inicia com a produção de um peptídeo precursor (20 – 110 aa), codificado por um gene estrutural, contendo uma sequência peptídica líder N-terminal, que possui regiões de reconhecimento por enzimas modificadoras pós-traducionais, e uma região central C-terminal que é o local onde ocorrem as modificações pós-traducionais, como macrociclização, modificação da espinha dorsal da amida e epimerização. Ao final, o peptídeo líder é removido por clivagem proteolítica para a maturação do composto maduro (Figura 5) (LEE; VAN DER DONK, 2022).

Figura 5: Representação esquemática da biossíntese de RiPP



A maioria dos genes biossintéticos (BGCs) de RiPPs em genomas microbianos estão inativos sob condições de cultura laboratorial. A mineração genômica e a engenharia biológica permitiram acessar informações cruciais para a descoberta e desenvolvimento de compostos RiPP anteriormente desconhecidos (LI; DING; ZHANG, 2024). Um exemplo na descoberta de novos RiPPs é a darobactina A. Este composto foi inicialmente identificado por meio de uma estratégia guiada pela bioatividade, e muitos de seus análogos e derivados foram posteriormente identificados por abordagens de mineração e engenharia do genoma (IMAI et al., 2019). A partir da mineração genômica por sequências de aminoácidos das partes centrais de peptídeos precursores, também foi possível identificar um BGC previsto para LAP no genoma da bactéria *Rhizobium* Pop5, que posteriormente teve seu produto sintetizado, caracterizado e denominado fazolicina (PHZ – do inglês, *Phazolicin*), um peptídeo que exibe atividade antibacteriana de espectro estreito contra bactérias simbióticas de leguminosas (TRAVIN et al., 2019). Curiosamente, PHZ, usa dois transportadores diferentes, *BacA* e *YejABEF*, para entrar nas células de uma bactéria simbiótica, *Sinorhizobium meliloti*. Isto reduz a probabilidade do aparecimento de mutantes resistentes a PHZ. Como esses transportadores também são cruciais para as associações simbióticas de *S. meliloti* com plantas hospedeiras, a sua transcrição permanece ativa, tornando PHZ um potencial candidato ao desenvolvimento de agentes de biocontrole para a agricultura (TRAVIN et al., 2023).

O metabolismo secundário microbiano é fortemente influenciado pelas vias regulatórias, que geralmente são identificadas e confirmadas através de manipulações genéticas, como mutações pontuais, que afetam a síntese de seus produtos. Apesar da ausência de uma rede regulatória universal, há mecanismos compartilhados entre espécies ou gêneros, como a regulação específica por BGCs (ativação ou repressão) ou reguladores pleiotrópicos, que controlam simultaneamente a expressão de vários BGCs (CRANEY; AHMED; NODWELL, 2013). As moléculas de sinalização butirrolactonas são produzidas por muitos estreptomicetos e seus efeitos podem ser específicos ou globais. Em *Streptomyces griseus*, por exemplo, atua como regulador global durante a esporulação; enquanto em *S. avermitilis*, atua especificamente na produção das avermectinas (HORINOUCI, 2002; KITANI et al., 2011).

Os genes responsáveis por fornecer resistência e transportar os metabólitos secundários para fora das células também estão agrupados dentro de BGCs. Quando o produto da via ocasiona dano contra o hospedeiro produtor, genes de resistência para esse metabólito específico podem ser codificados no *locus* genômico. Da mesma forma, os genes que codificam bombas de efluxo, como

o cassete de ligação ao ATP (ABC), também são responsáveis pela redução da toxicidade e dos efeitos de inibição (MARTÍN; CASQUEIRO; LIRAS, 2005).

As recentes descobertas sugerem que ainda há um extenso repertório de produtos naturais microbianos a serem explorados. Essas moléculas, selecionadas pela evolução ao longo do tempo, têm sido adotadas em diversos âmbitos industriais, refletindo sua importância além do ecológico. A exploração de novos ambientes e as complexas interações entre microrganismos e seus produtos naturais oferece uma oportunidade de recursos inéditos. A compreensão desses detalhes promete expandir nosso conhecimento sobre biodiversidade e sua aplicação em diversas áreas (O'BRIEN; WRIGHT, 2011).

1.3 TECNOLOGIAS NA DESCOBERTA DE PRODUTOS NATURAIS BACTERIANOS

Os produtos naturais a partir de bactérias, fungos e plantas são considerados a principal fonte das classes de medicamentos clinicamente aprovados. Devido aos seus papéis em vários contextos ecológicos, esses compostos naturais possuem estruturas especializadas e otimizadas para agir seletivamente a alvos celulares, especialmente proteínas (BREINBAUER; VETTER; WALDMANN, 2002). Durante anos, medicamentos derivados de bactérias foram desenvolvidos por triagens tradicionais guiadas pela bioatividade. Apesar de diversos produtos bem-sucedidos, essa abordagem perdeu espaço na descoberta de novos medicamentos, principalmente por baixo rendimentos na titulação e redescobertas de compostos (THAKER et al., 2013).

Com os primeiros sequenciamentos de genomas bacterianos, que incluíram os genomas de *Streptomyces coelicolor* e *S. avermitilis*, tornou-se evidente o potencial em produzir uma variedade muito maior de metabólitos secundários do que os obtidos em condições padrão de laboratório, reacendendo o interesse em expandir os esforços baseados em produtos naturais (BALTZ, 2019). Simultaneamente ao desenvolvimento em tecnologias de sequenciamento, foram desenvolvidos novos métodos, algoritmos e ferramentas computacionais para otimizar o uso das informações genômicas adquiridas (SCHERLACH; HERTWECK, 2021).

A mineração genômica abrange análises de informáticas a partir de dados genômicos, incluindo montagem e anotação do genoma, identificação de BGCs, previsão de estruturas e análise genômica comparativa para identificar similaridades e diferenças entre organismos (VAN DER HOOFT et al., 2020). A abordagem clássica de mineração é por vias biossintéticas. Diversas ferramentas foram desenvolvidas para a identificação de BGCs em sequências genômicas. O

antiSMASH é a plataforma mais abrangente para detectar e caracterizar aglomerados de genes biossintéticos (BGCs) em arqueias, bactérias e fungos, com auxílio do banco de referência MIBiG (do inglês, *Minimum Information about a Biosynthetic Gene Cluster*). O principal alvo de mineração desta ferramenta são enzimas biossintéticas conservadas das linhas de montagem molecular, como os peptídeos sintases não ribossômicas (NRPS) e policetídeos sintases (PKS) (ALBARANO et al., 2020; BLIN et al., 2023; KAUTSAR et al., 2019).

Apesar de abordagens direcionadas a vias biossintéticas apresentarem novidades valiosas, ao longo dos anos diferentes estratégias de mineração também surgiram com o intuito de validar novos candidatos a BGCs, como a mineração comparativa de genomas, que prever funções enzimáticas em grande escala no contexto genômico (RUDOLF; YAN; SHEN, 2016); a mineração baseado em filogenia, por meio de árvores filogenéticas para inferir a novidade e o potencial de metabólitos secundários de dados genéticos bacterianos (ZIEMERT et al., 2012); a mineração baseada em resistência, a partir do fato que bactérias produtoras de compostos antibióticos precisam desenvolver mecanismos resistentes (TANG et al., 2015); e a mineração por regulares, que impulsiona o desenvolvimento de ferramentas para prever elementos regulatórios em genomas bacterianos e facilitar a otimização de compostos (SPOHN; WOHLLEBEN; STEGMANN, 2016). Portanto, dependendo dos objetivos e interesses de pesquisa, diferentes técnicas e abordagens de mineração são desenvolvidas para detectar, desreplicar e priorizar BGCs.

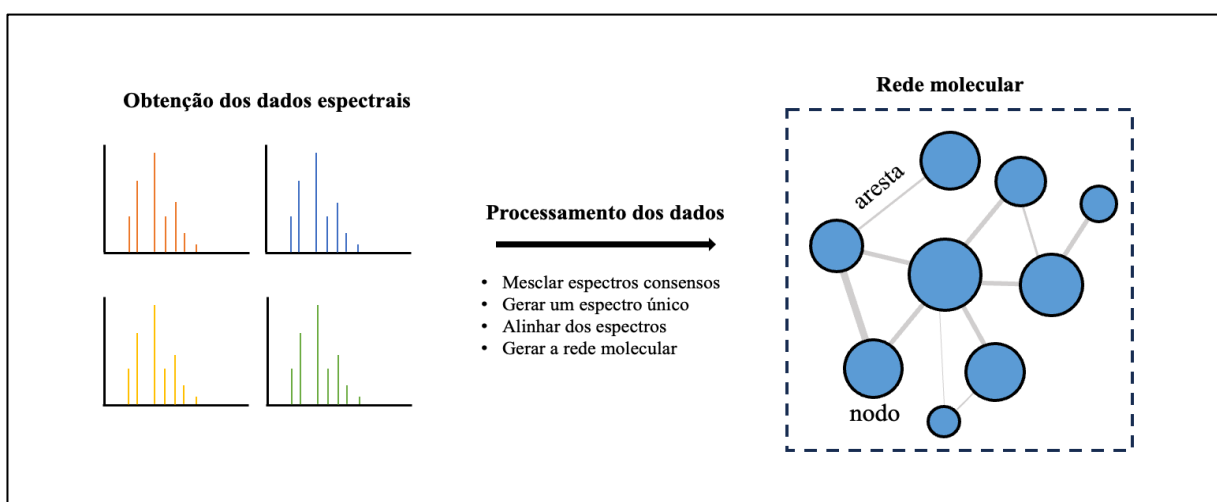
Embora estratégias de mineração tenham identificado diversos novos compostos, as buscas estão limitadas a sequências de enzimas homólogas e regras explicitamente definidas sobre o que é considerado um BGC (ZIEMERT; ALANJARY; WEBER, 2016). Assim, outras abordagens como a proteômica ou a metabolômica, representam uma abordagem promissora, de forma única ou combinatória com a genômica, para a descoberta de novas entidades biológicas que não puderam ser identificadas nas últimas décadas devido à cobertura limitada e resolução dos métodos convencionais (MARCO; ABRAM, 2019).

A espectrometria de massa é um método analítico amplamente aplicado em metabolomas, porém ainda é desafiador vincular o grande volume de dados obtidos a estruturas químicas conhecidas. Acredita-se que apenas uma pequena fração dos espectros em um experimento de metabolômica não direcionado podem ser corretamente anotados, isto é, dentro de todos os parâmetros de confiabilidade inclusos nas análises. Diante disso, a grande maioria das informações coletadas pela metabolômica é “matéria escura”, assinaturas químicas que permanecem não

caracterizadas, ou que merecem avaliações específicas para melhor elucidação de estruturas (DA SILVA; DORRESTEIN; QUINN, 2015). Avanços em métodos e instrumentação analítica de alta resolução, como a cromatografia líquida de ultra eficiência acoplado à espectrometria de massas (UPLC-MS, do inglês *Ultra Performance Liquid Chromatography tandem Mass Spectrometry*) tem contribuído para superar limites técnicos e possibilitado uma melhor separação, identificação e quantificação de compostos, assim como elucidação de propriedades químicas e estruturais (MUKHERJEE, 2019).

Para identificar metabólitos, é comum aplicação de redes moleculares (*Molecular Networking*) que agrupam metabólitos baseado nas semelhanças entre espectros de fragmentação (MS/MS), pois é sugerido que metabólitos com estruturas químicas semelhantes produzirão espectros MS/MS comparáveis (GRIM; LUU; SANCHEZ, 2019). Na formação das redes moleculares, os espectros obtidos são processados para gerar um espectro de consenso e um escore de similaridade. As similaridades são calculadas com uma “pontuação de cosseno” modificada baseada em vetor (variando de 0 a 1; quanto maior a pontuação, mais similar será o resultado), que leva em consideração o número de íons fragmentados correspondentes, as intensidades relativas dos picos e a precisão da massa do íon precursor. Cada entidade química é representada por um círculo ou nó (consenso dos dados MS/MS) e conectadas por arestas (valor do cosseno entre os nós) (Figura 6) (ARON et al., 2020). A plataforma *Global Natural Product Social* (GNPS) fornece instruções de criação de redes moleculares e biblioteca de espectros (WANG et al., 2016a).

Figura 6: Representação gráfica da formação de rede molecular clássica



O processamento e a mineração de dados a partir de dados cromatográficos e de espectrometrias de massas são desafiadores, principalmente devido às propriedades inerentes da eluição de metabólitos e seu comportamento durante a ionização. Se tratamento correto dos dados brutos, a desreplicação de compostos é comprometida, pois a identificação pode ser induzida a erro por semelhanças espectrais com variantes de ionização (GUO et al., 2022). Ferramentas computacionais e plataformas para destacar e identificar metabólitos também tem contribuído na compreensão desses sistemas complexos. O *pipeline NP3 Mass Spectrometry Workflow*, por exemplo, é uma coleção de *scripts* que a partir de dados LC-MS/MS, que quantifica espectros MS/MS e agrupa entidades químicas com base na similaridade espectral e tempo de retenção cromatográfico. Também há a possibilidade da montagem de redes moleculares para manipulação dos resultados e anotação baseada no GNPS e pelo banco de dados espectral *In Silico* (ISDB) de produtos naturais a partir do banco de dados Universal de Produtos Naturais (UNPD) (DE FELÍCIO et al., 2021).

A correlação de dados genômicos com a metabolômica para detectar novos produtos naturais é uma abordagem conhecida como “metabologenômica” (DOROGHAZI et al., 2014). Esta abordagem permitiu a descoberta de um novo produto natural, a tambromicina, a partir dos genomas e da metabolômica baseada das 178 cepas de actinomicetos. Neste estudo, 11.422 BGCs foram previstos em cinco categorias (NPRS, PKS Tipo I, PKS Tipo II, lantipeptídeos e microcinas modificadas com tiazol/oxazol) e 2.520 metabólitos individuais com massas variando de 250.074 a 4538.881 Da em quatro meios de cultura diferentes. Os autores destacam que a abordagem orienta a descoberta de novos produtos naturais, de modo que os dados genômicos contribuem em quase todos os níveis do processo de descoberta. Além disso, permite maior concentrar na identificação características únicas durante a desreplicação, determinação da composição e conectividade do monômero durante a elucidação da estrutura (GOERING et al., 2016).

Uma estratégia comum em programas de biodescoberta são as chamadas estratégias OSMAC (do inglês, *One Strain, Many Compounds*). Esta abordagem é baseada na percepção de que mesmo alterações sutis nas condições de cultivo podem induzir uma cepa a sintetizar uma variedade de metabólitos secundários distintos (BODE et al., 2002). A metodologia metabologenômica combinada com a estratégia OSMAC foi usada para ativar genes crípticos responsáveis pela produção de metabólitos secundários na bactéria *Lacinutrix shetlandiensis* sp. nov. WUR7, encontrada no fundo do mar. A mineração genômica revelou a presença de um

triptofano descarboxilase incomum, direcionando a análise do metabolismo de WUR7 para a produção de alcaloides à base de indol, utilizando L-triptofano como substrato. Isso resultou em uma grande variedade de metabólitos sendo secretados, que não eram observados na ausência de triptofano. A análise quimioinformática identificou 10 alcaloides indólicos, incluindo um novo alcaloide, bisindolo 8,9-diidroscoscinamida B, o primeiro do gênero *Lacinutrix*. Sua estrutura foi elucidada através de análises extensivas de espectroscopia de ressonância magnética nuclear e espectrometria de massa por ionização por electrospray de alta resolução (VITALE et al., 2023).

1.4 GÊNERO *STREPTOMYCES*

Streptomyces (filo Actinobacteria) são bactérias Gram-positivas, aeróbicas e podem ser encontradas em ambientes terrestres, aquáticos e extremófilos (VAN DER MEIJ et al., 2017). Morfologicamente, assemelham-se aos fungos pela formação de micélio, resultante do crescimento em extensão na ponta e ramificação das hifas. Em situações de estresse ambiental, as hifas se tornam hifas aéreas com superfície hidrofóbica e segregam em esporos, que se espalham no ambiente, assegurando a sobrevivência das espécies (BARKA et al., 2016).

Os estreptomicetos possuem cromossomos lineares maiores e com alto teor de citocina e guanina (CG) (CHOULET et al., 2006). O genoma é compartimentado, com genes conservados no centro do cromossomo e genes variáveis nas partes terminais. Os genes biossintéticos (BGCs) são consistentemente enriquecidos na parte variável do cromossomo (KIM et al., 2015). A plasticidade cromossômica garante o alto nível de variabilidade genética entre as espécies, e que também pode ser impulsionada pela transferência horizontal de genes por elementos conjugativos (ZHOU et al., 2012).

A diversidade de metabólitos especializados produzidos por este gênero oferece vantagens competitivas e evolutivas (SOTTORFF et al., 2019). Esses metabólitos atuam principalmente no sequestro de metais como ferro (sideróforos), proteção contra a luz UV (através da pigmentação) e inibição de cepas concorrentes (antimicrobianos) (BIBB, 2005). Outro processo importante é a simbiose com plantas e invertebrados, pois atuam na proteção contra patógenos, e os exsudatos dos hospedeiros contribuem para o desenvolvimento das cepas (BOSSO; MAULDIN; SALGADO, 2010).

Para além de sua função ecológica, *Streptomyces* é um dos gêneros mais versáteis para aplicações biotecnológicas, como agentes de biocontrole na agricultura (WANG et al., 2018),

produção de enzimas para bioconversão industrial (SCHMID et al., 2001) e metabólitos especializados para produção de medicamentos (ALAM et al., 2022). Os produtos naturais bioativos derivados de *Streptomyces* foram responsáveis por cerca de 70% dos antibióticos desenvolvidos durante a “Era de Ouro dos Antibióticos”, como a estreptomicina, primeiro antibiótico eficaz contra a tuberculose isolado de *S. griseus* (MARTENS; DEMAIN, 2011). Entre outros antibióticos identificados na época, destacam-se as tetraciclina, como clotetraciclina e oxitetraciclina, a partir de *S. aureofaciens* e *S. rimosus*, respectivamente (CHOPRA; ROBERTS, 2001); o cloranfenicol, um antibiótico oftalmológico de amplo espectro isolado de *Streptomyces venezuelae* (BALE et al., 2023) e a vancomicina, um antibiótico derivado de *S. orientalis* com ação contra bactérias Gram-positivas, incluindo *Staphylococcus aureus* resistente à meticilina (MRSA) (PATEL; PREUSS; BERNICE, 2023). Seus metabólitos secundários também são importantes agentes quimioterápicos, incluindo a doxorubicina e daunorubicina, a partir de *S. peucetius* (JOHNSON-ARBOR; DUBEY, 2023); compostos antifúngicos, como a anfotericina B produzido por *S. nodosus* e a nistatina produzido por *S. noursei* (CAFFREY et al., 2001; FJÆRVIK; ZOTCHEV, 2005); imunossupressores, incluindo rapamicina produzida por *S. hygroscopicus* e a FK506 produzido por *S. tsukubaensis* (KIM et al., 2014; KINO et al., 1987) e agentes antiparasitários, como as avermectinas isoladas de *S. avermitilis* (GUO et al., 2016).

Desde o início dos anos 2000, com os novos métodos de sequenciamento e ferramentas de bioinformática, análises dos genomas em *Streptomyces* revelaram que esses organismos contêm em média 30 BGCs em seu genoma, que podem variar de 20 a 100 kilobases de tamanho. Essas estruturas de biossíntese são essenciais para a caracterização de compostos, dos quais a maioria ainda é desconhecido sob condições de cultura padrão em laboratório (BENTLEY et al., 2002; NETT; IKEDA; MOORE, 2009). Com a crescente demanda para lidar com resistência a medicamentos, vários métodos sistemáticos de engenharia metabólica, como expressão heteróloga e otimização da biossíntese de precursores, foram desenvolvidos para permitir a produção eficiente desses compostos (HOBSON; CHAN; WRIGHT, 2021).

A partir da mineração do genoma da cepa *Streptomyces* sp. BRA-346, isolada do tunicado endêmico brasileiro *Euherdmania* sp., pesquisadores brasileiros identificaram as estruturas responsáveis e reproduziram por recombinação heteróloga peptídeos de epoxicetona, como didroeponemicina (DHE) e análogos estruturalmente relacionados, que inibem a atividade semelhante à quimotripsina do proteossoma e apresenta alta citotoxicidade para células de glioma

(DOMINGUES VIEIRA et al., 2022; FURTADO et al., 2021). A edição do genoma por CRISPR-Cas9 para espécies de *Streptomyces* também beneficiou a manipulação genética para caracterização, engenharia e produção de produtos naturais (COBB; WANG; ZHAO, 2015). Inserções de genes promotores estrategicamente introduzidos usando CRISPR-Cas9 foram eficientes na ativação de BGCs de diferentes classes em várias espécies de *Streptomyces*, incluindo PKS, NRPS e híbridos PKS-NRPS (ZHANG et al., 2017).

1.5 GÊNERO *RHODOCOCCLUS*

O gênero *Rhodococcus* (filo Actinobacteria) inclui bactérias Gram-positivas, aeróbicas, não esporulantes e possuem ácidos micólicos em sua parede celular, os quais facilitam a absorção de compostos hidrofóbicos (GÜRTLER; MAYALL; SEVIOUR, 2004). São amplamente distribuídos no solo, água, ambientes extremófilos, microbiota de animais e plantas (ALVAREZ, 2010). Membros do gênero *Rhodococcus* estão em constante descrições de novas espécies (GARRIDO-SANZ et al., 2020). Diante disto, é comum que haja longos históricos e atualização entre cepas. Como no caso da espécie originalmente conhecida como *Corynebacterium equi*, com história taxonômica conturbada, resultando em diversas reclassificações como *R. equi*, *Prescottella equi* e *R. hoagii* (GOODFELLOW; ALDERSON, 1977; GUO et al., 2015; JONES; SUTCLIFFE; GOODFELLOW, 2013).

Espécies de *Rhodococcus* têm sido investigadas para aplicações em biodegradação, biocatálise e biorremediação devido à sua versatilidade catabólica em degradar compostos orgânicos complexos, como hidrocarbonetos, fenóis, compostos halogenados e heterocíclicos (VANDERGEIZE; DIJKHUIZEN, 2004). Em estudo, *Rhodococcus* spp. formou o grupo com maior potencial metabólico entre todas as cepas obtidas com preferências catabólicas estendidas para biorremediação (cresceram na presença de alcanos e compostos aromáticos). Curiosamente, algumas dessas cepas foram obtidas de locais nunca expostos à contaminação por petróleo, assim o potencial genético para a degradação de hidrocarbonetos entre *Rhodococcus* pode ser amplamente distribuído e conservado entre espécies (BRZESZCZ et al., 2023).

Nos últimos anos, estudos mostram que os rodococos não possuem apenas potencial catabólico especializado, mas também um grande repertório anabólico, principalmente em relação à biossíntese de lipídios. Espécies de *Rhodococcus* são considerados microrganismos oleaginosos pela capacidade de produzir quantidades significativas de lipídeos na forma de triacilgliceróis

(acima de 20% do peso seco da célula) como uma resposta de adaptação a condições e nutrientes específicos de crescimento, tornando-as fontes alternativas de óleos de célula única, com possíveis aplicações na indústria, como biocombustíveis, biolubrificantes, produtos cosméticos, nutricionais e farmacêuticos (ALVAREZ et al., 2021). Cepas de *Rhodococcus* spp. também produzem biossurfactantes em resposta à substratos insolúveis em água (hidrocarbonetos) e em alguns casos também sobre substâncias solúveis em água (etanol e glicose). Biossurfactantes são classificados como glicolipídios, fosfolipídios, lipídios neutros, poliméricos e lipopeptídeos e possuem diversas aplicações comerciais, como em cosméticos, fábrica têxtil, indústria de alimentos e remediação do solo (HABIB et al., 2020).

Apesar de pertencer ao Actinobacteria, o conhecimento sobre a produção de compostos antimicrobianos é muito recente. Até o momento, aproximadamente 900 genomas de *Rhodococcus* estão depositados no NCBI, porém há escassez de conhecimento sobre a análise comparativa de BGCs nesse gênero. A grande maioria dos estudos em *Rhodococcus* se concentram na elucidação da filogenia e análise funcional do potencial catabólico (ORRO et al., 2015). As raras análises genômicas comparativas do gênero destacam o potencial surpreendentemente grande e inexplorado para a produção de metabólitos secundários (CENICEROS et al., 2017).

Entre os metabólitos especializados anteriormente descritos a partir de *Rhodococcus* estão sideróforos, como a requiquelina, heterobactina, rodoquelina e requibactina (BOSELLO et al., 2011, 2013; MIRANDA-CASOLUENGO et al., 2008, 2012) e compostos antimicrobianos, como as lariatinas, auraquinas, rodopeptinas e as humimicinas (CHIBA et al., 1999; CHU et al., 2016; IWATSUKI et al., 2006; KITAGAWA; TAMURA, 2008; KUNZE; HÖFLE; REICHENBACH, 1987). *Rhodococcus* também parecem ser um potencial produtor de bacteriocinas. As bacteriocinas são peptídeos sintetizados por ribossomos produzidos para inibir organismos taxonomicamente próximos (espectro estreito) (GILLOR; ETZION; RILEY, 2008). Em *Rhodococcus erythropolis* JCM 2895, um plasmídeo circular de 5,4 kb contém genes que codificam as proteínas RapA e RapB. RapA é uma proteína pequena, solúvel em água, estável ao calor e possui atividade antimicrobiana contra outras cepas de *R. erythropolis*, enquanto RapB é uma proteína de imunidade contra RapA que pode estar localizada na membrana celular (KITAGAWA et al., 2018).

O extenso número de BGCs altamente conservados em cepas de *Rhodococcus* as tornam alvos promissores na descoberta de produtos naturais, tendo em vista principalmente seu repertório

biossintético não caracterizado, seu crescimento relativamente rápido e a disponibilidade de ferramentas genéticas eficazes para sua modificação genômica (THOMPSON et al., 2020).

1.6 GÊNERO *BREVIBACILLUS*

O gênero *Brevibacillus* abrange o aglomerado de *Bacillus brevis*, pertence à família *Bacillaceae* do filo *Firmicutes*. São bactérias Gram-positivas ou Gram-variáveis, aeróbicas e formadoras de esporos (SHIDA et al., 1996). Atualmente, o gênero inclui 31 espécies validadas de acordo com a Lista de Nomes Procarióticos com Posição na Nomenclatura (LPSN) (PARTE et al., 2020). *Brevibacillus* spp. podem ser encontrados em diversos habitats ambientais, incluindo plantas, trato intestinal animal, ambientes aquático e solo (RUIU, 2013).

Raramente cepas deste gênero estão associadas a casos de infecções humanas, os quadros relacionados até o momento incluem um caso de peritonite em paciente com carcinoma hepatocelular, possivelmente causado pela ingestão de alimentos fermentados contendo esporos de *Brevibacillus brevis* e um quadro de bacteremia causado por *Brevibacillus laterosporus* (BUKHARI; TABOADA, 2017; PARVEZ; CORNELIUS; FADER, 2009).

Nos últimos anos, os destaques foram para o seu potencial biológico e aplicação em diversas áreas. A cepa *Brevibacillus agri* 5-2 isolada da água de formação de um reservatório de petróleo profundo no campo petrolífero na China foi considerada responsável pela biodegradação de compostos de hidrocarbonetos e organossulfurados, como tetradecano, hexadecano e o alcanosulfonato. Por análises genômicas as vias metabólicas responsáveis pelos mecanismos de adaptação e biodegradação dos microrganismos foram elucidadas, gerando informações sobre suas aplicabilidades industriais (SHE et al., 2014). Estudos anteriores também documentaram a capacidade de biodegradação de hidrocarbonetos em *Brevibacillus borstelensis* 707 (HADAD; GERESH; SIVAN, 2005), *Brevibacillus* sp. PDM-3 (REDDY et al., 2010) e *Brevibacillus panacihumi* W25 (WANG et al., 2014b). A sua aplicabilidade em tratamentos microbianos também foi eficaz na melhoria da qualidade das águas residuais de lixiviado bruto e primário tratado (JAIN; MAJUMDAR; DEVI, 2023).

Os membros de *Brevibacillus* também são considerados uma rica fonte de antimicrobianos peptídicos e lipopeptídicos. Em 1939, foi isolado o primeiro complexo peptídico do caldo de fermentação de *B. brevis* (ATCC 8185), definido como tirotricina. Composto por uma fração básica de decapeptídeos cíclicos (as tirocidinas) e uma fração neutra composta por peptídeos lineares (as

gramicidinas). Esta preparação demonstrou ter propriedades antibióticas e, embora descoberta cerca de 10 anos após a penicilina, foi o primeiro antibiótico utilizado na prática clínica (TANG; THIBAUT; BOYD, 1992).

As gramicidinas A-C são decapeptídeos lineares, enquanto a gramicidina S e a tirocidina A e suas isoformas B-D são decapeptídeos cíclicos (YANG; YOUSEF, 2018). Esses compostos são produzidos como precursores lineares pela síntese de peptídeos não ribossômicos em bactérias, que posteriormente são ciclizadas pela ação das enzimas tioesterase (LOLL et al., 2014). Estruturalmente, a tirocidina A e a gramicidina S diferem apenas em relação à sua sequência primária, e formam uma estrutura secundária do tipo β , levando a uma molécula anfipática. Peptídeos anfipáticos cíclicos são conhecidos por atuarem nas membranas bacterianas e desestabilizarem a integridade estrutural (ABRAHAM et al., 2014). As tirocidinas e seus derivados apresentam atividade antibacteriana, antifúngica, antimalárica e antiamebíca (AKBAR et al., 2022; RAUTENBACH et al., 2007; TROSKIE et al., 2014; YANG; YOUSEF, 2018). As gramicidinas lineares possuem atividade contra bactérias gram-positivas, incluindo *Streptococcus pyogenes*, *Staphylococcus aureus* e *Bacillus subtilis* (WANG et al., 2012).

As edeínas, um grupo bem caracterizado de peptídeos antimicrobianos catiônicos, inicialmente isoladas de *Brevibacillus brevis* Vm4, têm amplos efeitos inibitórios como bactérias Gram-positivas e Gram-negativas, fungos e micoplasmas (CZAJGUCKI; ANDRUSZKIEWICZ; KAMYSZ, 2006; JOHNSON; BOWMAN; DUNLAP, 2020).

Brevibacillus sp. Leaf182, isolada da folha de *Arabidopsis thaliana*, revelou produtos naturais que contribuíram para a atividade antibiótica observada, entre eles o novo policetídeo trans-AT macrobrevina (HELFRICH et al., 2018). Análogos de macrobrevina identificados em *Bacillus amyloliquefaciens* isolado de macroalgas também exibiram atividades antibacterianas. O BGC candidato PKS trans-AT é 46% semelhante a macrobrevina derivada de Leaf182, destacando a complexidade genética envolvida na produção de compostos semelhantes de diferentes fontes (CHAKRABORTY; KIZHAKKEKALAM; JOY, 2022). Assim, pesquisa em novos ambientes é essencial para a descoberta de novas espécies e amplia nosso entendimento sobre a influência da diversidade ambiental na síntese de antibióticos.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Avaliar o potencial biossintéticos de bactérias isoladas de amostras de solo amazônico a partir de abordagens integradas de genoma e metaboloma.

2.2 OBJETIVOS ESPECÍFICOS

- a) Coletar amostras de solo de região de conservação ambiental amazônica;
- b) Cultivar, isolar e identificar a similaridade filogenética de bactérias ambientais;
- c) Sequenciar e minerar o genoma das bactérias;
- d) Cultivar em meios de culturas diferentes e extrair os metabólitos secundários;
- e) Avaliar por similaridade de espectros agrupamentos de metabólitos secundários;
- f) Correlacionar o potencial genômico e a produção de metabólitos secundários.

3 MATERIAL E MÉTODOS

3.1 AMOSTRAGEM E CONSIDERAÇÕES ÉTICAS

Amostras de solo foram coletadas em três pontos da Área de Proteção Ambiental do Parque Estadual do Utinga, na cidade de Belém, Amazônia brasileira. A amostragem foi realizada durante o verão amazônico em 7 de outubro de 2021. Foram coletados 10 g de solo a uma profundidade de 10 cm e armazenados a 4°C em tubos de polipropileno previamente esterilizados. A amostragem foi autorizada pelo Instituto de Desenvolvimento Florestal e da Biodiversidade do Estado do Pará – IDEFLOR-BIO (nº 2021/830676) e pelo Instituto Chico Mendes de Conservação da Biodiversidade – ICMBio por meio do sistema online SISBIO (nº 78518-1). O acesso ao patrimônio genético brasileiro também foi registrado no Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado – SisGen (Nº A8DA4E2).

3.2 ISOLAMENTO, CLASSIFICAÇÃO TAXONÔMICA E SELEÇÃO BACTERIANA

O isolamento bacteriano foi realizado por diluição seriada em solução salina (NaCl 0,9%) a partir de 1 g de solo com posterior inoculação de alíquotas de até 10^{-4} em Ágar Amido Caseína pelo método de propagação em placa. O meio foi suplementado com cicloheximida 100 µg mL⁻¹ para inibir o crescimento de fungos. As placas foram incubadas a 28°C±2°C por até 72 h. Colônias sugestivas de actinomicetos (textura dura com aspecto seco e filamentos com ou sem micélios aéreos) foram isoladas pelo método *strake plate* em SCA e armazenadas em glicerol 20% a -80°C.

As bactérias foram classificadas taxonomicamente com base na sequência do seu gene 16S rRNA. Os isolados foram cultivados em *Tryptic Soy Broth* (TSB) e o DNA genômico foi extraído usando o kit *DNeasy Blood and Tissue* (Qiagen) de acordo com o protocolo do fabricante. O Fluorômetro *Qubit* (Thermo Fisher Scientific) foi utilizado para quantificação de DNA enquanto a integridade do DNA foi avaliada em gel de agarose a 1%. O gene 16S rRNA foi amplificado utilizando o par de primers universais 8F (5'-AGAGTTTGATCCTGGCTCAG-3') e 1492R (5'-GGTTACCTTGTTACGACTT-3'). A reação foi realizada utilizando o kit *GoTaq Green Master Mix* (Promega) com 10ng de DNA e 200 nM de cada primer. Os parâmetros de ciclagem da PCR foram os seguintes: uma etapa inicial a 95°C por 5 min, seguida de 30 ciclos a 95°C por 1 min,

55°C por 1 min, 72°C por 1 min e uma etapa final de extensão a 72°C por 10 min. Os *amplicons* foram purificados usando o reagente de limpeza de produto de PCR *ExoSAP* (Thermo Fisher Scientific). O sequenciamento de Sanger foi realizado no ABI Prism 3500 *Genetic Analyzer* (Thermo Fisher Scientific) usando o kit *BigDye Terminator v3.1*. As sequências foram analisadas e alinhadas no BioEdit v.7.2 e comparadas no banco de dados GenBank usando BLASTn.

Após o isolamento, foram adotados dois critérios para seleção de cepas com potencial biossintético: I) deveriam pertencer ao filo Actinobacteria, que são grupos bacterianos conhecidos por produzirem metabólitos secundários de interesse biotecnológico; e II) deveriam apresentar atividade inibitória contra cepas controle. A atividade inibitória foi avaliada pelo método de estrias cruzadas (BALOUIRI; SADIKI; IBNSOUDA, 2016). Resumidamente, uma única linha do isolado foi inoculada em *Tryptic Soy Agar* (TSA) e incubada a 28°C±2°C por até 48 h. Posteriormente, duas cepas de controle, *Staphylococcus aureus* ATCC 29213 e *Escherichia coli* ATCC 25922, foram cruzadas perto da linha isolada. Após a incubação, a formação de zonas de inibição próximas ao crescimento do isolado foi considerada um indicio de bioatividade.

3.3 SEQUENCIAMENTO E MINERAÇÃO DO GENOMA

O sequenciamento genômico das bactérias foi realizado no Laboratório de Engenharia Biológica (EngBio/PCT-Guamá/UFGA). O sequenciamento de leituras longas foi executado no sequenciador *PromethION 2 Solo* (Oxford Nanopore Technologies), kit *Rapid Sequencing DNA V14, native barcoding* kit 96V14, seguindo as instruções do fabricante. O processo de montagem do conjunto de dados empregou *Flye* versão 2.9.2 (KOLMOGOROV et al., 2020). O sequenciamento genômico de leituras curtas foi realizado na plataforma Ion GeneStudio S5 (Thermo Fisher Scientific) utilizando o chip 550 de acordo com o protocolo do fabricante. A montagem do genoma foi realizada com SPAdes (BANKEVICH et al., 2012). A anotação do genoma foi feita com a ferramenta de software Prokka (SEEMANN, 2014). A completude da montagem do genoma e a anotação com ortólogos de cópias únicas foram avaliadas com BUSCO (SIMÃO et al., 2015). O sistema MUMmer v.4.0.0 foi utilizado para realizar o alinhamento do genoma contra referências depositadas no GenBank. Representações esquemáticas dos genomas foram criadas com a assistência do Circos (KRZYWINSKI et al., 2009). Os *Clusters of Orthologous Groups of proteins* (COGs) foram anotados usando o eggNOG-mapper com o banco de dados eggNOG v4.5, que prevê ORFs diretamente a partir de *contigs* montados, para classificar

os genes codificados nos genomas completos de bactérias (TATUSOV, 2000) e os bancos de dados *Kyoto Encyclopedia of Genes and Genomes* (KEGG) para obter informações de anotação funcional (KANEHISA, 2000).

Os genomas foram submetidos na plataforma antiSMASH versão 7.0.0 (BLIN et al., 2023). Foram analisados cada um dos genes biossintéticos propostos e comparados com genes homólogos para termos análises mais detalhadas do cluster gênico candidato. Esta etapa foi realizada com o auxílio da plataforma MIBiG (do inglês, *Minimum Information about a Biosynthetic Gene cluster*) (MEDEMA et al., 2015) e BLAST (do inglês, *Basic Local Alignment Search Tool*) (JOHNSON et al., 2008). Análises mais aprofundadas foram realizadas em banco de dados, como o GenBank (BENSON et al., 2017), o Uniprot (THE UNIPROT CONSORTIUM, 2017), o Pfam (FINN et al., 2016) e o PDB (do inglês *Protein Data Bank*) (BERMAN, 2008).

3.4 EXTRAÇÃO DE METABÓLITOS DOS CULTIVOS BACTERIANOS

As extrações de metabólitos secundários de cultivos bacterianos foram realizadas com a técnica de separação líquido-líquido utilizando o solvente de polaridade intermediária acetato de etila. Os cultivos foram realizados em meios de cultura diferentes: ACT015 e ACT016 em meios A1, TSB, TSBY e ISP2 e a FIR094 em meios LB, TSB e R2A. Os cultivos foram preparados em Erlenmeyer de 2 L com 500 mL de meio mantidos a 30 °C e 200 RPM, por 48h (tempo de saturação da curva de crescimento). Ao final do período, os cultivos foram transferidos para béquer de plásticos de 1 L e submetidos ao sonicador por 10 ciclos de 30 segundos cada. As amostras lisadas foram transferidas para tubos e centrifugadas a 37.000 xg por 30 minutos. Os sobrenadantes foram utilizados para realização das extrações.

Em 500 mL de cultura foi adicionado 400 mL (80% do volume) de acetato de etila dividido em duas etapas (200 mL cada etapa). Após a primeira separação da fase orgânica da fase aquosa, retornou a fase aquosa ao funil de separação e foi adicionado a segunda parte de acetato de etila. As partes da fase orgânica foram evaporadas no rotaevaporador a 40 °C, 100 rpm. Após evaporação total do solvente, foram ressuspensas com metanol (até 10 mL) e transferidas para tubos falcon de 50 mL previamente pesados. A secagem das amostras foi realizada no *SpeedVac* e pesadas novamente para calcular a quantidade de massa obtida. Para injeção em UPLC-MS/MS, as amostras foram ressuspensas em DMSO para concentração de 10 mg/mL.

3.5 COLETA E ANÁLISE DE DADOS POR UPLC-MS/MS

Para o perfil químico UPLC-MS/MS, alíquotas de 2 µl das amostras foram injetadas em uma coluna de fase reversa BEH C18 (1,7 µm, 2,1 × 100 mm) acoplada a uma pré-coluna compatível, usando um *Acquity H-Class UPLC Waters* (Waters, Milford, MA, Estados Unidos) acoplado a um espectrômetro de massa Bruker Impact II UHR-ESI-QqTOF (Bruker Daltonics, Billerica, MA, Estados Unidos). Um sistema solvente de água (A), acetonitrila (B), ácido fórmico 2% (C) metanol (D), foi utilizado para compor o seguinte método analítico: vazão de 0,5 ml/min; temperatura da coluna 40°C; 0–1 min, 5% B; 1–6 min, 5% B a 35% B (curva 6); 6–10 min, 35% B a 95% B (curva 1); 10–12 min, 95% B; 12–15 min, 95% D (curva 1); 15–18 min para equilíbrio da coluna na fase inicial. C foi mantido constante, a 5% (concentração final de 0,1%). O espectrômetro de massa funcionou no modo de íon positivo varrendo massa na faixa de 30 a 2.000 Da, com taxa de aquisição de 8 Hz. *Offset* da placa final = 500 Volts (V); Vcap 4.500 V; nebulizador 4,0 bar; fluxo de gás de secagem (N₂) 10 L/min; temperatura do gás de secagem 200°C, seguido por uma varredura MS/MS para os íons mais intensos em um tempo de ciclo de 1 s, limite absoluto (por 1.000 somas) de 1.500 cts. Como regras do MS₂, a razão de massa (m/z) abaixo de 200 Da foi excluída e a função “exclusão ativa” foi habilitada. Cada execução foi calibrada automaticamente usando HCOONa (10 mM). Os espectros calibrados foram convertidos em arquivos mzXML por meio do Data Analysis 4.3 e incluíram ferramentas bio CompassXport (Bruker Daltonics, versão 4.0.0.8).

3.6 PROCESSAMENTO E ANÁLISE DE DADOS NP³ MS WORKFLOW

O processamento e análise de dados foram realizados utilizando a plataforma NP³ *Mass Spectrometry Workflow* (DE FELÍCIO et al., 2021). A versão beta está disponível no repositório (https://bitbucket.org/cnpemlqpn/np3_ms_workflow/src/master/). A rede molecular de anotação de variantes de ionização (IVAMN, do inglês *Ionization Variant Annotation Molecular Networking*) foi usada como base para selecionar e excluir os nodos presentes em amostras de espaços em branco cromatográficos (*blanks*) e seus primeiros vizinhos imediatos. Em seguida, os nodos remanescentes marcados como representante protonado (coluna *protonated_representative* igual a 1 na tabela de contagem de nodos) foram selecionados e filtrados na rede molecular de

similaridade de espectros (SSMN, do inglês *Spectra Similarity Molecular Networking*). A SSMN bruta foi montada com valor de cosseno de 0,6, número máximo de 200 componentes por cluster e número máximo de 15 conexões para cada nodo. A construção da IVAMN e da SSMN no Cytoscape (SHANNON et al., 2003) foi realizada seguindo um *script* interno *python* que automatiza tarefas com a biblioteca *py2cytoscape* (<https://github.com/cytoscape/py2cytoscape>). Após essas etapas automatizadas, os nodos marcados como meio de cultura (*bed*), nodos únicos (*selfloop*) e sem conexão foram removidos manualmente. A anotação da estrutura química MS/MS foi realizada usando correspondências de espectros com os bancos de dados GNPS (WANG et al., 2016b) e UNPD-ISDB (ALLARD et al., 2016). As anotações foram revisadas manualmente e os clusters MS/MS relacionados foram agrupados e nomeados pelos grupos químicos encontrados no SSMN.

5 RESULTADOS E DISCUSSÃO

A metabologenomics approach reveals the unexplored biosynthetic potential of bacteria isolated from pristine Amazonian soil

A metabologenomics approach reveals the unexplored biosynthetic potential of bacteria isolated from pristine Amazonian soil

Ana Carolina Favacho Miranda de Oliveira^{1,2}; Bruna Domingues Vieira³; Rafael de Felício³; Lucas da Silva e Silva²; Adonney Allan de Oliveira Veras¹; Diego Assis das Graças¹; Artur Silva¹; Rafael Azevedo Baraúna^{1*}; Daniela Barretto Barbosa Trivella³; Maria Paula Cruz Schneider^{1,2}

¹ Biological Engineering Laboratory, Innovation Space, Guamá Science and Technology Park, 66075-110, Belém, PA, Brazil.

² Center of Genomics and Systems Biology, Institute of Biological Sciences, Federal University of Pará, 66075-110, Belém, PA, Brazil.

³ Brazilian Biosciences National Laboratory (LNBio), Brazilian Center for Research in Energy and Materials (CNPEM), 13083-100, Campinas, SP, Brazil.

* Correspondence author: Rafael A. Baraúna. E-mail: rabarauna@ufpa.br. Address: Laboratório de Engenharia Biológica, Parque de Ciência e Tecnologia Guamá, Av Perimetral, 66075-110, Belém, PA, Brazil.

ABSTRACT

The Amazon is considered an important source of biodiversity. However, few studies have been conducted to describe the microbial communities of Amazonian habitats, and even fewer have been carried out to explore the biotechnological potential of natural products synthesized by these microbial communities. To help fill this knowledge gap, our study aimed to characterize the metabolic potential of Gram-positive strains affiliated to the Actinomycetes and Bacilli classes, isolated from pristine Amazonian soils. A multi-omics approach based on genome mining and untargeted metabolomics was used. The 16S rRNA gene sequencing classified the strains ACT015, ACT016, and FIR094 within the genera *Streptomyces*, *Rhodococcus*, and *Brevibacillus*, respectively. A total of 33, 17, and 14 biosynthetic gene clusters (BGCs) were found in the genomes of *Streptomyces* sp. ACT015, *Rhodococcus* sp. ACT016, and *Brevibacillus* sp. FIR094. A wide variety of enzymatic pathways were detected including BGCs that produce antibiotics and antitumor compounds. However, a total of 40 BGCs (62,5%) were related to metabolites with unknown functions. By applying the OSMAC (one strain many compounds) approach for strain cultivation, untargeted metabolomics investigation revealed a myriad of metabolites produced by these strains under laboratory conditions. Our data demonstrated a vast unexplored chemical diversity in bacteria isolated from conserved regions of the Amazon. The metabologenomics approach was effective in describing the secondary metabolism of the isolates and in prospecting BGCs of biotechnological interest for future studies.

KEYWORDS: Secondary metabolism; Biosynthetic gene clusters; Untargeted metabolomics; *Streptomyces*; *Rhodococcus*; *Brevibacillus*.

INTRODUCTION

The Amazon is the largest rainforest in the world and is considered an important source of biodiversity. The bioeconomy has been one of the strategic solutions for the social and economic development of the region. It recognizes biodiversity as an important commercial asset and, therefore, emphasizes the need for ecosystem conservation (1). Natural products (secondary metabolites, or specialized metabolites) are produced by ribosomal or non-ribosomal synthesis and play a crucial role in the adaptation of microorganisms to their natural habitat (2,3). Beyond their ecological importance, their chemically diverse structures have been extensively explored for the development of biotechnological products for a

plethora of applications, such as antibiotics for human health (4). Traditionally, most natural product discovery has been driven by top-down approaches, with bioactivity-guided fractionation of extracts from microbes and plants. However, the laborious steps of this kind of analysis and its technical limitations compromise the efficiency of discovery rates (5).

The omics technologies have enabled the increasing number of complete genomes deposited in public databases, which increased considerably since the emergence of high-throughput DNA sequencers approximately 20 years ago (6). Through genome mining techniques, it is possible to observe many Biosynthetic Gene Clusters (BGCs) that encode secondary metabolites with unknown functions (7). Additionally, advances in untargeted metabolomics due to increasingly sensitive tandem mass spectrometry (MS/MS) systems have allowed a more in-depth analysis of the metabolic components of interesting extracts (8). The integration of omics data has complemented traditional discovery methods, allowing for the identification of new enzymatic pathways involved in the biosynthetic production of promising metabolites (9).

Metabologenomics (genome sequencing combined with mass spectrometry-based metabolomics) combined with the OSMAC strategy (One Strain Many Compounds) is a powerful multi-omics approach for discovering new secondary metabolites, being successfully employed in the discovery of novel alkaloids from *Lacinutrix shetlandiensis* isolated from the southern sea (10–12). Same strategy was used by Iacovelli and colleagues (2024) to describe the sesquiterpene- and lactone-producing fungus *Anthostomella pinea*, including its enzymes with biotechnological relevance such as carbohydrate-active enzymes (CAZymes) and non-specific peroxygenases (UPOs) (13). This study shows the unexplored biosynthetic potential of three free-living bacterial strains isolated from soil samples collected at an environmental protection area near the city of Belém (PA, Brazil) at the Brazilian Amazon. The metabologenomic approach enabled the identification of the main classes of secondary metabolites produced by *Streptomyces* sp. ACT015, *Rhodococcus* sp. ACT016, and *Brevibacillus* sp. FIR094, and revealed a significant number of molecules that have not yet been explored.

MATERIALS AND METHODS

Sampling and ethical considerations

Soil samples were collected from three points in the Utinga State Park, an environmental protection area near the city of Belém, Brazilian Amazon (Sample A:

1°25'35.6"S 48°25'28.2"W; Sample B: 1°25'30.6"S 48°25'35.0"W; Sample C: 1°25'28.3"S 48°25'42.9"W) (Figure S1). Sampling was conducted during Amazon summer on October 7, 2021. A total of 10 g of soil was collected from a depth of 10 cm and stored at 4°C in previously sterilized polypropylene tubes. Sampling was authorized by environmental authorities, the Instituto de Desenvolvimento Florestal e da Biodiversidade do Estado do Pará – IDEFLOR-BIO (N° 2021/830676), and the Instituto Chico Mendes de Conservação da Biodiversidade – ICMBio through the online system SISBIO (N° 78518-1). Access to the Brazilian genetic heritage was also registered in the Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado – SisGen (N° A8DA4E2).

Bacterial isolation, taxonomic classification, and selection of strains

Bacterial isolation was performed by serial dilution in saline solution (NaCl 0,9%) from 1 g of soil with subsequent inoculation of aliquots of up to 10^{-4} on Starch Casein Agar (SCA) by spread plate method. The medium was supplemented with $100 \mu\text{g}/\text{mL}^{-1}$ cycloheximide to inhibit fungal growth. Plates were incubated at $28^{\circ}\text{C} \pm 2^{\circ}\text{C}$ for up to 72 h. The suggestive actinomycete colonies (with a tough texture and a dry appearance and filaments with or without aerial mycelia) were isolated by strake plate method on SCA and stored in 20% glycerol at -80°C in Biological Engineering Laboratory biobank.

The strains were taxonomically classified based on the sequence of their 16S rRNA gene. Isolates were grown on Tryptic Soy Broth (TSB) and genomic DNA was extracted using DNeasy Blood and Tissue kit (Qiagen) according to the manufacture's protocol. The Qubit Fluorometer (Thermo Fisher Scientific) was used for DNA quantification, and DNA integrity was evaluated on 1% agarose gel. The 16S rRNA gene was amplified using the universal primer pair 8F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-GGTTACCTTGTACGACTT-3'). The PCR reaction was performed using the GoTaq Green Master Mix (Promega) with 10 ng of DNA and both primers at 200 nM. Amplicons were purified using ExoSAP PCR Product Cleanup Reagent (Thermo Fisher Scientific) and sequenced on ABI Prism 3500 Genetic Analyzer (Thermo Fisher Scientific) using the BigDye Terminator v3.1 kit. The sequences were analyzed in BioEdit v.7.2 and compared to the NCBI GenBank database using web BLASTn available at <https://ncbi.nlm.nih.gov/blast>.

The criteria adopted to select strains with biosynthetic potential were I) they should belong to the Actinomycetes, which are known to produce secondary metabolites of biotechnological interest; and II) they should present inhibitory activity against control strains. Inhibitory activity was evaluated using the cross-streak method (14). Briefly, a single streak of

the isolate was inoculated on Tryptic Soy Agar (TSA) and incubated at 28°C±2°C for up to 48 h. Subsequently, two control strains, *Staphylococcus aureus* ATCC 29213 and *Escherichia coli* ATCC 25922, were cross-streaked near the isolate streak. After incubation, the formation of inhibition zones near the growth of the isolate was considered an indication of bioactivity, allowing the selection of metabolically promising strains.

Genome mining

Genome sequencing was performed using short- and long-read approaches to the three selected strain. Short-read sequencing was performed on Ion GeneStudio S5 Plus (Thermo Fisher Scientific) using the Ion 550 Chip kit. Long-read sequencing was performed on PromethION P2 Solo (Oxford Nanopore Technologies) using a R10.4.1 flow cell. Genome assembly was performed using SPAdes v.3.15.5 (15) and Flye v.2.9.2 (16) for short and long-read, respectively. Prokka v.1.14.5 (17) was used for genome annotation and BUSCO (18) was used to evaluate genome completeness. The MUMmer v.4.0.0 systems was used to perform genome alignment against references deposited in GenBank and CIRCOS v.0.69-9 (19) was used to construct circular maps. COG (Cluster of Orthologous Groups) assignment was performed in eggNOG-mapper v. 2.1.12 with the eggNOG database v. 5.0.2 database (20). Additionally, the Kyoto Encyclopedia of Genes and Genomes (KEGG) was used to obtain functional information of the predicted genes (21).

AntiSMASH v.7.0.0 software (22) was used to identify BGCs in the sequenced genomes (22). The following parameters were used: detection strictness “relaxed” and “all on” extra features. Manual curation was performed after AntiSMASH prediction using the MIBiG platform (Minimum Information about a Biosynthetic Gene cluster) (23) and BLAST (24). Additional analyses were performed using GenBank (25), UniProt (26), Pfam (27), and the Protein Data Bank (PDB) (28).

Sample preparation and metabolite extraction

The genomically characterized isolates were grown in different culture media to stimulate the expression of different classes of secondary metabolites (OSMAC strategy): *Streptomyces* sp. ACT015 and *Rhodococcus* sp. ACT016 were grown in A1, TSB, TSBY, and ISP2 media while *Brevibacillus* sp. FIR094 was grown in LB, TSB and R2A media. Erlenmeyer flasks containing 500 mL of medium were incubated at 30°C and 200 rpm for 48 h, and subsequently sonicated in an ice bath for 10 cycles of 30 s at 30 s intervals. After lysis,

the samples were centrifuged at 37,000 g for 30 min and the supernatant was used for secondary metabolites extraction.

To 500 mL of culture, 400 mL (80% of the volume) of ethyl acetate was added in two stages (200 mL each stage). After the first separation of the organic phase from the aqueous phase, the aqueous phase was returned to the separation funnel, and the second portion of ethyl acetate was added. The organic phase portions were evaporated on a rotary evaporator at 40 °C, 100 rpm. After complete solvent evaporation, they were resuspended with methanol (up to 10 mL) and transferred to pre-weighed 50 mL falcon tubes. The sample was dried using a SpeedVac (Thermo Scientific), and the solids were reweighed to calculate the obtained mass. For UPLC-MS/MS injection, the samples were resuspended in DMSO to a concentration of 10 mg mL⁻¹.

Collection and analysis of data by UPLC-MS/MS

For the UPLC-MS/MS chemical profile, 2 µL aliquots from the samples were injected into a BEH C18 reversed-phase column (1.7 µm, 2.1 × 100 mm) coupled to a compatible pre-column, using a Waters Acquity H-Class UPLC system (Waters, Milford, MA, USA) coupled to a Bruker Impact II UHR-ESI-QqTOF mass spectrometer (Bruker Daltonics, Billerica, MA, USA). A solvent system of water (A), acetonitrile (B), 2% formic acid (C), and methanol (D) was used to compose the following analytical method: flow rate of 0.5 mL/min; column temperature 40°C; 0–1 min, 5% B; 1–6 min, 5% B to 35% B (curve 6); 6–10 min, 35% B to 95% B (curve 1); 10–12 min, 95% B; 12–15 min, 95% D (curve 1); and 15–18 min for column equilibration to the initial phase. C was kept constant at 5% (final concentration of 0.1%). The mass spectrometer was operated in positive ion mode, scanning mass in the range of 30 to 2,000 Da, with an acquisition rate of 8 Hz. End plate offset = 500 Volts (V); Vcap 4,500 V; nebulizer 4.0 bar; drying gas flow (N₂) 10 L/min; drying gas temperature 200°C, followed by an MS/MS scan for the most intense ions in a 1 s cycle time, absolute threshold (per 1,000 sums) of 1,500 cts. For MS², the mass-to-charge ratio (*m/z*) below 200 Da was excluded, and the "active exclusion" function was used. Each run was automatically calibrated using HCOONa (10 mM). The calibrated spectra were converted into mzXML files through Data Analysis 4.3 and included bio CompassXport tools (Bruker Daltonics, version 4.0.0.8).

NP³ MS Workflow data processing and analysis

Data processing and analysis were carried out using the NP³ Mass Spectrometry Workflow platform (29). The Spectra Similarity Molecular Networking (SSMN) was useful for organizing spectra in terms of similarity, providing an overview of metabolite distribution, and visualizing clusters of similar metabolites (possible analogs or compound families). The molecular network Ionization Variant Annotation Molecular Networking (IVAMN) was used as the basis to select and exclude nodes present in chromatographic blank samples and their immediate first neighbors. Then, the remaining nodes marked as protonated representatives (protonated_representative column equal to 1 in the node count table) were selected and filtered. The construction of IVAMN and SSMN in Cytoscape (30) was carried out following a Python script that automates some tasks in the software using the py2cytoscape library (<https://github.com/cytoscape/py2cytoscape>). After these automated steps, nodes marked as culture media (bed), self-loop nodes, and nodes without connections were manually removed. The MS/MS chemical structure was performed using spectrum matches with the Global Natural Products Social Molecular Networking (GNPS) (31) and In Silico Spectral Databases of Natural Products (UNPD-ISDB) (32) databases. Annotations were manually reviewed, and related MS/MS clusters were grouped and named based on the chemical groups found in SSMN.

RESULTS

Selection of strains and general genomic characteristics

A total of 50 bacterial strains were isolated from pristine Amazonian soils and maintained in the biobank of the Biological Engineering Laboratory (Belém, PA, Brazil). The strains ACT015, ACT016, and FIR094 were selected according to the criteria presented in the methods section and were classified by their 16S rRNA gene sequences in the genera *Streptomyces*, *Rhodococcus*, and *Brevibacillus*, respectively (Figure 1a, 1b and 1c). *Streptomyces* sp. ACT015 presented a genome size of 8.6 Mbp and 7,161 coding sequences (CDSs) while *Rhodococcus* sp. ACT016 and *Brevibacillus* sp. FIR094 presented genome sizes of 5.8 and 6.4 Mbp, and 5,910 and 3,303 CDSs, respectively. General genomic characteristics were summarized in table 1 and were in accordance with the average values for each genus. The genome of *Streptomyces* sp. ACT015 and *Rhodococcus* sp. ACT016 were assembled using long-reads, while the *Brevibacillus* sp. FIR094 genome was assembled using short-reads (Table 1).

The *Rhodococcus* sp. ACT016 genome was assembled into two contigs. The smaller contig (642,065 Kbp) showed a very low similarity to the 13 *Rhodococcus* complete genomes used as reference (Figure S2). Some genes found in this contig were related to plasmid functions such as the *tra* genes, which are directly involved in conjugative DNA transfer. One of these genes had a conjugative relaxase domain TrwC (Figure S3a). Additionally, the plasmid-encoded enzyme AtzC was partially detected, presenting 97% similarity to the *Rhodococcus* sp. SGAir0479 gene (Figure S3b). However, it was not possible to identify the plasmid incompatibility group using the PlasmidFinder v. 2.1.6 database (33).

Genomic rings were drawn to compare the genomes of ACT015, ACT017 and FIR094 against other reference strains with complete genomes deposited in GenBank (Figure 2a, 3a, and 4a) (Table S1). CDSs containing COGs of unknown function (S) were the most represented in the three genomes (1,085 CDSs in *Streptomyces* sp. ACT015; 1,208 CDSs in *Rhodococcus* sp. ACT016; and 1,227 CDSs in *Brevibacillus* sp. FIR094) followed by transcription (K) (914 CDSs in *Streptomyces* sp. ACT015; 653 CDSs in *Rhodococcus* sp. ACT016; and 589 CDSs in *Brevibacillus* sp. FIR094). The transcription category encompasses activator or repressor proteins and sigma factors; therefore, they are key genes for environmental adaptation. The most abundant COG pathways were related to primary cellular functions such as carbohydrate metabolism, protein metabolism, and inorganic ion metabolism (Figure 2b, 3b, and 4b). Other primary metabolic pathways were also detected such as those involved in carboxylic acid metabolism, nucleotide metabolism, and biosynthesis of amino acids and cofactors (Figure 5).

The COG analysis identified 388 CDSs related to secondary metabolism in *Streptomyces* sp. ACT015, 422 in *Rhodococcus* sp. ACT016, and 178 in *Brevibacillus* sp. FIR094 (Figure 2b, 3b, and 4b). The KEEG enrichment, in contrast, found 370 CDSs related to secondary metabolism in *Streptomyces* sp. ACT015, 327 CDSs in *Rhodococcus* sp. ACT016, and 286 in *Brevibacillus* sp. FIR094 and distributed in three heterogeneous profiles to the strains, especially for ACT015, the more external clade in Figure 5. According to the genomic enrichment data (Figure 5), the most abundant biosynthetic pathways in *Streptomyces* sp. ACT015 were related to the synthesis of terpenoids, polyketides, aminoglycosides, alkaloids, and beta lactams. *Rhodococcus* sp. ACT016 presented genes for the synthesis of terpenes (including carotenoids), siderophores, polyketides, aminoglycosides, and beta lactams, while *Brevibacillus* sp. FIR094 presented a significant number of pathways for the synthesis of terpenes, non-ribosomal peptides, and polyketides.

Genome mining and biosynthetic potential

A total of 64 BGCs were identified in the genomes, most of which had low or no similarity to known clusters (62,5%). Occasionally, a BGC may contain one or more biosynthetic genes and therefore synthesize, from the same cluster, different types of metabolites (34). This type of cluster was mainly observed in the *Streptomyces* sp. ACT015 and *Brevibacillus* sp. FIR094 genomes.

Thirty-three BGCs were detected in the genome of *Streptomyces* sp. ACT015. The main classes of biosynthetic genes were Polyketide Synthases (PKS), Ribosomally Synthesized and Post-translationally modified Peptide (RiPP), Non-Ribosomal Peptide Synthetase (NRPS), terpenes, and hybrid NRPS-PKS. Eighteen of the thirty-three BGCs showed 70% homology with the MIBiG database, including BGCs responsible for the biosynthesis of carotenoids, mycelin, candicidin, flavolin, e-Polylysine, ectoin, melanin, deferoxamine, albaflavenone, indigoidine, minimycin, geosmin, antimycin, splenocin C, neoantimycin, trioxacarcin A, hopene, streptazone E, neocarziline A/B, SapB, birimostide, and ethylenediaminesuccine acid hydroxyarginine (EDHA). Among the 15 BGCs with low or no homology, 6 were related to RiPP, 4 NRPS, 4 PKS, 2 NRPS-PKS, and 2 terpenes (Table 2).

In *Rhodococcus* sp. ACT016 genome, 17 BGCs were identified. The main classes were NRPS and terpenes. Only two had >70% similarity with MIBiG database and were predicted to be involved in the synthesis of e-Polylysine and ectoine. BGCs with similarity between 30% and 70% were related to the production of isorenieratene, corynecin, phosphoramidon, and thermochelin. The remaining 11 BGCs presented low or no similarity with the database, including 7 NRPS, 1 terpene, 1 RiPP, and 1 PKS (Table 3). Additionally, we were able to find 2 BGCs in the accessory contig, a putative plasmid, being an uncharacterized NRPS and a Butyrolactone 63% similar to Heterobactin (Table 3).

Finally, 14 BGCs were identified in *Brevibacillus* sp. FIR094 genome. The main BGC classes were RiPP, NRPS, and hybrid NRPS-PKS. Only 4 showing >70% homology with the MiBiG database, including BGCs for the synthesis of tyrocidine, gramicidin, macrobrevin and petrobactin. The other 10 BGCs showed low or no similarity, including 3 NRPS-PKS, 6 RiPPs, 1 PKS, and 1 terpene (Table 4).

Untargeted metabolomics

To stimulate the production of as many secondary metabolites as possible, strains were grown in different culture media (29,35). Molecular network analysis also facilitated the identification of variations caused by the OSMAC strategy (Figure 6a). Considering all culture

media, *Streptomyces* sp. ACT015 presented the richer raw molecular network containing 335 nodes, while *Brevibacillus* sp. FIR094 presented 312 nodes, and *Rhodococcus* sp. ACT016 presented only 100 nodes, totalizing 747 nodes. The strains shared few nodes, for example, *Streptomyces* sp. ACT015 and *Rhodococcus* sp. ACT016 shared only 4 nodes related to primary metabolism (Figure 6b).

Streptomyces sp. ACT015 produced more metabolites in ISP2 and A1 culture media (Figure 6c). In contrast, *Rhodococcus* sp. ACT016 produced more metabolites in A1 and TSB media (Figure 6d), while R2A was the best medium for *Brevibacillus* sp. FIR 094 (Figure 6e). Approximately 135, 40, and 119 nodes of *Streptomyces* sp. ACT015, *Rhodococcus* sp. ACT016 e *Brevibacillus* sp. FIR 094, respectively, were annotated by GNPS and/or UNPD-USDB databases. In contrast, 89, 23, and 105 nodes were not identified. When we applied restrictive parameters for spectra annotation (GNPS: Score >0.85; UNPS-ISDB Score >0.4; and Error < 20 ppm), the number of identified spectra decreased drastically (Figura 6f).

One hundred and twenty nodes were annotated in the *Streptomyces* sp. ACT015 metabolome, 85 of which were annotated using the UNPD-ISDB database and 50 using the GNPS database. Eighty-nine nodes were not annotated, corresponding to 42% of the metabolome. *Rhodococcus* sp. ACT016 produced 32 secondary metabolites that were annotated using UNPD-ISDB and GNPS databases, and 23 metabolites had no match (approximately 41% of the metabolome). *Brevibacillus* sp. FIR094 produced the largest number of unidentified metabolites (105 nodes), corresponding to 54% of the metabolome. Of the 88 secondary metabolites annotated, 41 had matches with the GNPS database and 78 were annotated with the UNPD-ISDB database.

Among the main superclasses, the *Streptomyces* sp. ACT015 metabolome highlights the presence of organic acids, organic oxygen compounds, organoheterocyclic compounds, lipids, alkaloids, benzenoids, macrolides, peptides and oligopeptides, polyketides, and terpenoids (Figure 7a). In the *Rhodococcus* sp. ACT016 metabolome the following super classes were the most prominent: organic acids, organic oxygen compounds, organoheterocyclic compounds, lipids, alkaloids, and terpenoids (Figure 7a). Finally, in the *Brevibacillus* sp. FIR094 metabolome showed mostly: organoheterocyclic compounds, lipids, alkaloids, benzenoids, oligopeptides, polyketides, and terpenoids (Figure 7a).

Some BGCs predicted by genome mining were possible to be confirmed with the metabolites annotated by LC-MS/MS. For example, BGC 20 (Table 2) from *Streptomyces* sp. ACT015 showed 100% similarity with the NRPS-PKS gene cluster that produces Antimycin (**compound 1**). This metabolite was annotated in its metabolome by using the GNPS database

(**1**, m/z 506.3116) (Figure 7b). Another example was the BGC 13 from *Streptomyces* sp. ACT015 that presented 100% similarity with a gene cluster that produces Deferoxamine E (**compound 2**), a molecule that was further annotated in the metabolome of the bacteria grown on ISP2 medium (**2**, m/z 601.355) (Figure 7b).

Despite correlations with this degree of specificity could not be made for another BGCs, some metabolites were annotated in the same chemical classes as some BGCs predicted in the genomes. For example, the ion with m/z 219.1741 in SSMN from *Streptomyces* sp. ACT015 metabolome was annotated as a sesquiterpene class and likely corresponds to the product of BGC 15, which presented 100% similarity to a gene cluster involved in the synthesis of the sesquiterpene Albaflavenone ($218.33 \text{ g mol}^{-1}$). Comparisons between the genome and metabolome of *Brevibacillus* sp. FIR094 and *Rhodococcus* sp. ACT016 also revealed several correlations between BGCs and metabolites at the chemical class level. However, in most cases no direct correlation could be made.

DISCUSSION

Streptomyces sp. ACT015 is a resource of known and unknown NPs genetically rich and metabolically active

Genomic parameters such as genome size, number of CDSs, and GC content were in accordance with what is expected for isolates of the genus *Streptomyces* (36). PKS, RiPPs such as lanthipeptides, terpenes, NRPS, and hybrids NRPS-PKS were the most predominant BGCs. Belknap and colleagues (2020) analyzed 1,110 *Streptomyces* genomes and identified these same BGCs as the most common of the genus (7). Some BGCs contained NRPS-PKS chimeric modules (BGCs 3, 8, 20, and 33 - Table 2). These clusters are commonly known as hybrids BGCs producing important specialized metabolites, as salinosporamide A and eponemycin or TMC-86A. Although the origin and role of hybrid BGCs are poorly understood, these systems produce bioactive chimeric molecules that contribute significantly to the diversity of secondary metabolites observed in nature (37). One of the 33 BGCs found, the BGC 20, is a hybrid cluster containing PKS and NRPS-PKS genes unique to 8 of the 10 reference genomes in comparative genomics (Figure 2a; Table 2). It is the most complex BGC of *Streptomyces* sp. ACT015, composed of 3 NRPS genes and 4 PKS genes, showing 65% similarity with a *S. griseoviridis* F1-27 BGC.

The most of BGCs in ACT015 genome had some similarity to previously described NPs, with 18 of the 33 BGCs showing more than 70% (Table 2). Comparative genomics

identified 5 BGCs of *Streptomyces* sp. ACT015 (Figure 2a). BGC 7, predicted as T1PKS, was detected in a genomic region with low level of similarity (<50%) compared to the other genomes of the reference strains (Figure 2a). Additionally, the AntiSMASH result indicates a low level of homology (25%) with a cluster from *Streptomyces* sp. CS081A. Another example was BGC 14 (Table 2; Figure 2a): despite showing 88% homology with *S. purpureus* JCM 3172 in AntiSMASH, this BGC is responsible for synthesizing a molecule with no similarity to the MIBiG database.

Regarding the chemical classes of metabolites, *Streptomyces* sp. ACT015 produced essential metabolites for environmental adaptation such as carotenoids, melanin, geosmin, ectoin, hopene, and the siderophores amychelin and deferoxamines. These last two molecules are widely used in medicine. Amychelin was able to reduce the pathogenicity of *Pseudomonas aeruginosa* under infectious condition (38). Deferoxamine was used in the treatment of iron overload, such as in beta-thalassemia disease (39) and cases of iron or aluminum poisoning (40). Other BGCs synthesized molecules with antibiotic potential, such as albaflavenone, indigoidine, minimycin, trioxacarcin A, and birimositide (41–45); antifungals such as candicidin (46); and chemotherapeutics such as minimycin, antimycin, neoantimycin and neocarzilines (47–49).

The considerable number of BGCs with unknown functions (15 of 33) and unannotated metabolic network nodes draws attention to a still unexplored biosynthetic potential. Despite *Streptomyces* being the taxon with the greatest number of entries in the MIBiG database (approximately 800), 15 BGCs of the ACT015 strain showed low or no similarity with the available data. Therefore, ACT015 carries new gene clusters with unknown functions potentially involved in adaptation to the Amazon environment, a biome with a rich biodiversity, high average annual temperatures, and intense rainfall.

The largest number of secondary metabolites were identified in ISP2 and A1 media. Bacterial growth and metabolism are directly impacted by the carbon source (50). ISP2 is routinely used for *Streptomyces* growth and for antimicrobial activity assays (51). It contains glucose as carbon source while A1 medium contains soluble starch and TSBY, sucrose. The last medium used was TSB, which is composed by peptone digest and other components of lower concentration. The Carbon Catabolite Repression (CCR) is a mechanism highly conserved in *Streptomyces* that regulates carbon metabolism. CCR also influences cell morphology and synthesis of some secondary metabolites (52).

Approximately 27% of the nodes in the *Streptomyces* sp. ACT015 metabolic network were not annotated. Identification of these metabolites is limited by the availability of

reference spectra (50). The GNPS platform provides a large library of natural compounds for similarity analysis (53). However, annotation of small molecule mass spectra remains a challenging problem. Therefore, *in silico* tools with a larger spectral database can be used to complement annotation. We used ISDB (32) for *in silico* annotation of spectra. In fact, for the *Streptomyces* sp. ACT015 metabolome, the UNPD-ISDB annotated a greater number of spectra (85 nodes) compared to the GNPS (50 nodes). The tool confirmed the annotation of Antimycin A4 and Deferoxamine E, as well as classified several other spectra into alkaloids, oligopeptides, peptides, polyketides, and terpenes. Desferoxamine E, also known as Nocardamine, was annotated into oligopeptides superclass (Figure 7b). This molecule is a cyclic hydroxamic acid siderophore with antitumor activity that is produced by several species of bacteria (54). Amino acids in cyclic conformation in the molecule structure may have induced its classification into the oligopeptides superclass.

Interestingly, all clusters of the molecular network containing annotated compounds also contained unidentified adduct ion nodes that possibly have a similar chemical structure (Figure 7a). These unidentified nodes may represent different adducts of the same known compound or may characterize potential new analogues of predicted compounds. *In silico* tools are generally more accurate in predicting molecular classes and, therefore, special attention should be taken for more specific identifications. In this work, annotations were based on standard reliability parameters available in the molecular networks (GNPS: Score >0.85; UNPD-ISDB Score >0.4 e Error < 20 ppm). Many annotations based on these criteria were discarded to reduce false positive annotations or were limited to the identification of the molecular class. Exploratory studies can take advantage of these *in silico* tools to better describe the chemical potential of promising isolates.

***Rhodococcus* sp. ACT016 is an environmental strain genetically rich and active with a highly shapable metabolism**

The genome size of *Rhodococcus* species varies between 4 and 10 Mbp. This large difference in genome size is related to the presence of large plasmids, extensive genome instability, and the diversity of habitats that are colonized by *Rhodococcus* species (55). Environmental strains such as *R. opacus* B4 usually have larger genomes (8.2 Mbp) compared to pathogenic strains such as *R. equi* (5.4 Mbp) (56,57). Some exceptions have been observed, such as *Rhodococcus* sp. Br-6 (5.5 Mbp) isolated from soil samples and *Rhodococcus defluvii* Ca-11 (5.1 Mbp) isolated from a wastewater treatment bioreactor (58,59). The genome size

of the strain ACT016 isolated in our study was 5.8 Mb, which is smaller than that expected for other environmental strains.

Although approximately 900 *Rhodococcus* genomes have been deposited at the GenBank, there is a lack of knowledge about the BGCs of this genus. Genome mining of 20 strains revealed an extensive repertoire of uncharacterized BGCs (60). In another study, Doroghazi and Metcalf (61) demonstrated that *Rhodococcus* isolates presented a greater proportion of NRPS than actinomycetes. Interestingly, this was the main type of cluster found in the ACT016 genome, which possesses 17 BGCs with 9 of it being NRPS. Additionally, 88% of the detected BGCs were not similar to BGCs previously described (Table 3). Some ACT016 clusters found as absent in other *Rhodococcus* reference genomes suggest this strain as harboring exclusive BGCs: the clusters 4, 6, 7, 8, and 13 (Table 3; Figure 3), which were of particular interest because they were all NRPS. Among the BGCs identified using the AntiSMASH database, those with the highest identity were involved in the synthesis of ϵ -Polylysine, an antibiotic commonly produced by *Streptomyces* and *Kitasatospora*; and ectoine, a metabolite related to osmoprotection and survival in saline environments (62).

According to the untargeted metabolomics results, *Rhodococcus* sp. ACT016 showed higher production of metabolites in A1 and TSB media, despite the lower nodes number when compared to ACT015, which was grown in the same culture media. The same metabolite was rarely synthesized in more than one culture medium, showing that the OSMAC approach was a success. Other metabolites from the classes of heterocyclic compounds, organic acids, lipids, terpenoids, polyketides, and alkaloids were also detected. Using the same approach as our study, Gamaleldin and colleagues (63) identified molecules with potential antimalarial activity produced by *Rhodococcus* strains. Palma-Esposito and colleagues (64) identified glycopeptides with potential antiviral activity.

Eleven nodes (10%) were related to lipid production, with no significant annotation. Some *Rhodococcus* species such as *R. opacus*, *R. jostii*, and *R. wratislaviensis*, are known to be oleaginous microorganisms, as they are capable of accumulate triacylglycerols (TAG) at more than 20% of their cell dry weight. Oleaginous rhodococci are source of lipids with industrial application (65). High TAG production is obtained through growth under nitrogen starvation conditions or abundance of carbon sources. An imbalance between carbon and nitrogen availability redirects carbon metabolism toward lipogenesis (66). Half of the nodes identified as lipids were produced in TSBY, which is composed of 10% sucrose and 3% soybean digest. The other half of the nodes were produced in A1 medium, which contains 1%

soluble starch and 0.2% casein peptone, confirming that lipid production is strongly influenced by the carbon/nitrogen availability ratio.

***Brevibacillus* sp. FIR094 is a metabolically active strain with conserved and still unknown NPs to be explored**

Brevibacillus sp. FIR094 presented a genome with a size of 6.5 Mbp and 14 BGCs. Most of these BGCs are widely distributed and conserved in the genus that synthesizes molecules with antimicrobial activity, including tyrocidine, gramicidin and macrobrevin, but only 4 showed >70% similarity to the AntiSMASH database, suggesting an underexplored potential in FIR094 close-related strains.

Prospecting bacteria of biotechnological interest in underexplored environments is essential for the discovery of new molecules and to expand our understanding of the ecological context of the natural antibiotic production. In *Brevibacillus* sp. Leaf182, which was isolated from an *Arabidopsis thaliana* leaf, several natural antibiotics, including a new trans-AT polyketide from the macrobrevin class, have been described (67). Macrobrevin analogs with antimicrobial activity were also identified by Chakraborty and colleagues (68) in *Bacillus amyloliquefaciens* isolated from macroalgae. These two macrobrevin molecules presented only 46% similarity, demonstrating the genetic complexity involved in the production of similar compounds by different isolates.

Brevibacillus is a genus widely recognized as an Antimicrobial Peptide (AMP) producer, especially for AMPs synthesized by NRPS clusters. They also synthesize AMPs from RiPPs, especially bacteriocins, such as laterosporulin (69), laterosporulin10 (70), and Bac-GM100 (71). In *Brevibacillus* sp. FIR094, 6 RiPPs were detected, four of which showed low identity with known BGCs (Table 4). BGC 1 synthesizes a metabolite from the subfamily of linear azol(in)e-containing peptides (LAPs). LAPs comprise a subclass of RiPPs that share heterocyclic thiazol(in)e and (methyl)oxazol(in)e structures, promoting the cyclization of cysteine, serine, or threonine side chains during the post-translational modification step (72). This BGC was composed by the precursor peptide (gene A), the enzymes dehydrogenase (gene B) and cyclodehydratase (gene C/D), the maturation protein (gene F), and regulatory and transport genes (Figure 8a). A metabolite of approximately 5,000 Da was calculated based on the 54 amino acid sequence of the core peptide. The peptide contains 14 cysteines (C) in sequence that can form seven disulfide bonds (S-S). These disulfide bonds play a crucial role in the stabilization of the molecule, giving it a specific conformation and increasing its resistance to environmental factors. In addition, the sequence also included

aromatic amino acids such as phenylalanine (F), which may contribute to the structural and functional properties of the peptide. The BGC also contained the transcription factor AbrC3, an inducer of antibiotic production (Figure 9a). AbrC3 is a positive transcriptional regulator of specialized metabolites and morphological differentiation in *S. coelicolor* (73). The BGC of *Brevibacillus* sp. FIR 094 showed high similarity to other BGCs from the antiSMASH database, suggesting that the region is highly conserved within the genus (Figure 9b). However, the metabolic network cluster showed low similarity with known gene clusters from MIBiG (Figure 9c). BGC 10 synthesizes a lanthipeptide of class II. It contains all genes necessary for the synthesis of regulatory elements and carries three genes for auto resistance. Identification of antimicrobial resistance genes allows insights into self-defense mechanisms and suggests production of specialized metabolites with biological activity (74). These BGCs are strong candidates for cloning and biosynthesis by heterologous expression to analyze the structure and functionality of the AMPs.

Metabolically, *Brevibacillus* sp. FIR094 showed a considerable nodes number synthesizing more metabolites when grown in R2A minimal medium, probably because this type of medium creates a nutritional stress that enhances secondary metabolite production (75). Despite the good results in metabolites production, most of the metabolome was not possible to be annotated, the poorest characterized metabolome of herein studied strains. BGCs for the synthesis of important antimicrobial compounds such as tyrocidine, gramicidin, and macrobrevin were detected in the genome, but we were unable to find these molecules in the metabolome. This may have occurred because of the metabolite extraction and detection methods that were not focused on large peptides, or the tight regulation of their production, which expression could be hampered under the laboratory conditions. The similarity network of *Brevibacillus* sp. FIR094 revealed metabolites with molecular weights of up to 1,200 Da. Tyrocidine and gramicidin have a molecular weight of 1,270.5 and 1,882.3 g/mol, respectively, a value above the threshold of detection.

CONCLUSIONS

In this research we evaluated the biosynthetic potential of three free-living bacterial strains isolated from pristine soil samples from the Brazilian Amazon using a multi-omics approach based on genome mining and untargeted metabolomics. Genome mining identified several BGCs for the synthesis of antimicrobial and antitumor compounds. However, a significant number of unidentified BGCs were found, corresponding to 45%, 88%, and 71%

of the predicted clusters in *Streptomyces* sp. ACT015, *Rhodococcus* sp. ACT016, and *Brevibacillus* sp. FIR094, respectively, including relevant biosynthetic classes such as NRPS, PKS, and RiPPs, indicating the untapped potential of Amazonian microorganisms.

The biosynthesis of secondary metabolites was modulated by the type of culture media. Most of the acquired spectra were not annotated by the consulted databases (GNPS and UNPD-ISDB), revealing a significant gap of information of the available databases and/or an untapped chemical novelty of the assessed Amazonian strains, even for exhaustively studied taxa such as *Streptomyces* spp. The Amazonian microorganisms are, therefore, an important source for the registry of natural compounds and a promising source for discovering new natural molecules.

Data integration in the metabologenomics approach allows for a more accurate search of new secondary metabolites. This type of approach is especially useful for samples that have been poorly explored, such as the pristine soils of Amazon. In this study, were rare the cases in which we could correlate the predicted BGC with its metabolite, possibly due to the novelty of the BGCs and spectra detected. These results indicate a high genetic and metabolic variability in Amazonian microorganisms, that could be explored for the detection and production of metabolites of medical and biotechnological utility.

ACKNOWLEDGMENTS

Authors would like to thank Pró-Reitoria de Pesquisa e Pós-Graduação (PROPESP) from Universidade Federal do Pará (UFPA) for the payment of the Article Processing Charges through the PAPQ 2024 Call. This research was funded by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) under the grant number 88887.465291/2019-00 and Instituto Serrapilheira under the grant number Serra-1709-19681. Authors thank the Brazilian Bioscience National Laboratory (LNBio-CNPq) and the EngBio Laboratory (UFPA) for infrastructure and personal dedicated to the project.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

DATA AVAILABILITY

Genome sequences were submitted to GenBank under the BioProject ID [PRJNA1089999](#) and accession numbers [SAMN40553816](#), [SAMN40553817](#), and

[SAMN40553818](https://massive.ucsd.edu/). Metabolomics data can be found at: <https://massive.ucsd.edu/>, and accession numbers MSV000094460, MSV000094461 and MSV000094462.

AUTHOR CONTRIBUTION

ACMFO: designed and executed the experiments. Performed the field expeditions. Wrote the manuscript and compiled the data for deposition; BDV: designed the cultivation experiments and assisted on their execution. Assisted on the metabologenomics investigation; RF: designed and assisted on the metabolomics experiments; LSS: assistance in field expeditions. Performed genome assemblies, enrichment, and comparative genomics analysis. Paper review and support in data management; AAOV: bioinformatics assistance in genome assemblies. Support in server access and data demands; DAG: support in NGS data post-processing; AS: funding and coordination of the project; RAB: Conceptualization, discussions based on experiments, funding, and co-coordination of the project; DBBT: Conceptualization, discussions based on experiments, funding, and co-coordination of the project; MPCs: Conceptualization, discussions based on experiments, funding, and co-coordination of the project. All authors contributed to the text and agreed with the final version of the paper.

REFERENCES

1. Guayasamin JM, Ribas CC, Carnaval AC, Carrillo JD, Hoorn C, Lohmann LG, et al. Chapter 2: Evolution of Amazonian biodiversity. In: Amazon Assessment Report 2021. UN Sustainable Development Solutions Network (SDSN); 2021.
2. Davies J. Specialized microbial metabolites: functions and origins. *J Antibiot (Tokyo)*. 2013 Jul 12;66(7):361–4.
3. O'Brien J, Wright GD. An ecological perspective of microbial secondary metabolism. *Curr Opin Biotechnol*. 2011 Aug;22(4):552–8.
4. Sharrar AM, Crits-Christoph A, Méheust R, Diamond S, Starr EP, Banfield JF. Bacterial Secondary Metabolite Biosynthetic Potential in Soil Varies with Phylum, Depth, and Vegetation Type. *mBio*. 2020 Jun 30;11(3).
5. van der Hooft JJJ, Mohimani H, Bauermeister A, Dorrestein PC, Duncan KR, Medema MH. Linking genomics and metabolomics to chart specialized metabolic diversity. *Chem Soc Rev*. 2020;49(11):3297–314.
6. Reuter JA, Spacek D V., Snyder MP. High-Throughput Sequencing Technologies. *Mol Cell*. 2015 May;58(4):586–97.
7. Belknap KC, Park CJ, Barth BM, Andam CP. Genome mining of biosynthetic and chemotherapeutic gene clusters in *Streptomyces* bacteria. *Sci Rep*. 2020 Feb 6;10(1):2003.
8. Mohd Kamal K, Mahamad Maifiah MH, Abdul Rahim N, Hashim YZHY, Abdullah Sani MS, Azizan KA. Bacterial Metabolomics: Sample Preparation Methods. *Biochem Res Int*. 2022 Apr 12;2022:1–14.

- 598 9. Trivella DBB, de Felicio R. The Tripod for Bacterial Natural Product Discovery: Genome
599 Mining, Silent Pathway Induction, and Mass Spectrometry-Based Molecular
600 Networking. *mSystems*. 2018 Apr 24;3(2).
- 601 10. Bode HB, Bethe B, Höfs R, Zeeck A. Big Effects from Small Changes: Possible Ways to
602 Explore Nature's Chemical Diversity. *ChemBioChem*. 2002 Jul 3;3(7):619.
- 603 11. Goering AW, McClure RA, Doroghazi JR, Albright JC, Haverland NA, Zhang Y, et al.
604 Metabologenomics: Correlation of Microbial Gene Clusters with Metabolites Drives
605 Discovery of a Nonribosomal Peptide with an Unusual Amino Acid Monomer. *ACS Cent*
606 *Sci*. 2016 Feb 24;2(2):99–108.
- 607 12. Vitale GA, January GG, Oppong-Danquah E, Della Sala G, Palma Esposito F, Tasdemir D,
608 et al. A metabologenomics approach to unlock the metabolome of the novel Antarctic
609 deep-sea isolate *Lacinutrix shetlandiensis* sp. nov. WUR7. *PNAS Nexus*. 2023 Jul 3;2(7).
- 610 13. Iacovelli R, He T, Allen JL, Hackl T, Haslinger K. Genome sequencing and molecular
611 networking analysis of the wild fungus *Anthostomella pinea* reveal its ability to
612 produce a diverse range of secondary metabolites. *Fungal Biol Biotechnol*. 2024 Jan
613 3;11(1):1.
- 614 14. Balouiri M, Sadiki M, Ibnsouda SK. Methods for in vitro evaluating antimicrobial
615 activity: A review. *J Pharm Anal*. 2016 Apr;6(2):71–9.
- 616 15. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A
617 New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing.
618 *Journal of Computational Biology*. 2012 May;19(5):455–77.
- 619 16. Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, Shin SB, et al. metaFlye:
620 scalable long-read metagenome assembly using repeat graphs. *Nat Methods*. 2020
621 Nov 5;17(11):1103–10.
- 622 17. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014 Jul
623 15;30(14):2068–9.
- 624 18. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva E V, Zdobnov EM. BUSCO:
625 assessing genome assembly and annotation completeness with single-copy orthologs.
626 *Bioinformatics*. 2015 Oct 1;31(19):3210–2.
- 627 19. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: An
628 information aesthetic for comparative genomics. *Genome Res*. 2009 Sep;19(9):1639–
629 45.
- 630 20. Tatusov RL. The COG database: a tool for genome-scale analysis of protein functions
631 and evolution. *Nucleic Acids Res*. 2000 Jan 1;28(1):33–6.
- 632 21. Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*.
633 2000 Jan 1;28(1):27–30.
- 634 22. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, et al. antiSMASH 7.0:
635 new and improved predictions for detection, regulation, chemical structures and
636 visualisation. *Nucleic Acids Res*. 2023 Jul 5;51(W1):W46–50.
- 637 23. Medema MH, Kottmann R, Yilmaz P, Cummings M, Biggins JB, Blin K, et al. Minimum
638 Information about a Biosynthetic Gene cluster. *Nat Chem Biol*. 2015 Sep 18;11(9):625–
639 31.
- 640 24. Johnson M, Zaretskaya I, Raytselis Y, Merezuk Y, McGinnis S, Madden TL. NCBI BLAST:
641 a better web interface. *Nucleic Acids Res*. 2008 May 19;36(Web Server):W5–9.
- 642 25. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, et al.
643 GenBank. *Nucleic Acids Res*. 2017 Jan 4;45(D1):D37–42.

- 644 26. The UniProt Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids*
645 *Res.* 2017 Jan 4;45(D1):D158–69.
- 646 27. Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. The Pfam protein
647 families database: towards a more sustainable future. *Nucleic Acids Res.* 2016 Jan
648 4;44(D1):D279–85.
- 649 28. Berman HM. The Protein Data Bank: a historical perspective. *Acta Crystallogr A.* 2008
650 Jan 1;64(1):88–95.
- 651 29. de Felício R, Ballone P, Bazzano CF, Alves LFG, Sigrist R, Infante GP, et al. Chemical
652 Elicitors Induce Rare Bioactive Secondary Metabolites in Deep-Sea Bacteria under
653 Laboratory Conditions. *Metabolites.* 2021 Feb 12;11(2):107.
- 654 30. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a
655 software environment for integrated models of biomolecular interaction networks.
656 *Genome Res.* 2003 Nov;13(11):2498–504.
- 657 31. Wang M, Carver JJ, Phelan V V, Sanchez LM, Garg N, Peng Y, et al. Sharing and
658 community curation of mass spectrometry data with Global Natural Products Social
659 Molecular Networking. *Nat Biotechnol.* 2016 Aug 9;34(8):828–37.
- 660 32. Allard PM, Péresse T, Bisson J, Gindro K, Marcourt L, Pham VC, et al. Integration of
661 Molecular Networking and *In-Silico* MS/MS Fragmentation for Natural Products
662 Dereplication. *Anal Chem.* 2016 Mar 15;88(6):3317–23.
- 663 33. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, et al. *In*
664 *Silico* Detection and Typing of Plasmids using PlasmidFinder and Plasmid Multilocus
665 Sequence Typing. *Antimicrob Agents Chemother.* 2014 Jul;58(7):3895–903.
- 666 34. Martinet L, Naômé A, Deflandre B, Maciejewska M, Tellatin D, Tenconi E, et al. A Single
667 Biosynthetic Gene Cluster Is Responsible for the Production of Bagremycin Antibiotics
668 and Ferroverdin Iron Chelators. *mBio.* 2019 Aug 27;10(4).
- 669 35. Domingues Vieira B, Niero H, de Felício R, Giolo Alves LF, Freitas Bazzano C, Sigrist R, et
670 al. Production of Epoxyketone Peptide-Based Proteasome Inhibitors by *Streptomyces*
671 *sp.* BRA-346: Regulation and Biosynthesis. *Front Microbiol.* 2022 Mar 24;13.
- 672 36. Lee N, Kim W, Hwang S, Lee Y, Cho S, Palsson B, et al. Thirty complete *Streptomyces*
673 genome sequences for mining novel secondary metabolite biosynthetic gene clusters.
674 *Sci Data.* 2020 Feb 13;7(1):55.
- 675 37. Gallagher KA, Jensen PR. Genomic insights into the evolution of hybrid isoprenoid
676 biosynthetic gene clusters in the MAR4 marine streptomycete clade. *BMC Genomics.*
677 2015 Dec 17;16(1):960.
- 678 38. Xie F, Dai S, Zhao Y, Huang P, Yu S, Ren B, et al. Generation of Fluorinated Amychelin
679 Siderophores against *Pseudomonas aeruginosa* Infections by a Combination of
680 Genome Mining and Mutasynthesis. *Cell Chem Biol.* 2020 Dec;27(12):1532-1543.e6.
- 681 39. Ballas SK, Zeidan AM, Duong VH, DeVeaux M, Heeney MM. The effect of iron chelation
682 therapy on overall survival in sickle cell disease and β -thalassemia: A systematic
683 review. *Am J Hematol.* 2018 Jul 28;93(7):943–52.
- 684 40. Wright SW, Valento M, Mazor SS, Chen BC. Severe iron poisoning treated with
685 prolonged deferoxamine infusion: a case report. *Toxicol Commun.* 2018 Jan 25;2(1):6–
686 9.
- 687 41. Cude WN, Mooney J, Tavanaei AA, Hadden MK, Frank AM, Gulvik CA, et al. Production
688 of the Antimicrobial Secondary Metabolite Indigoidine Contributes to Competitive
689 Surface Colonization by the Marine Roseobacter *Phaeobacter sp.* Strain Y4I. *Appl*
690 *Environ Microbiol.* 2012 Jul 15;78(14):4771–80.

- 691 42. GÜRTLER H, PEDERSEN R, ANTHONI U, CHRISTOPHERSEN C, NIELSEN PH, WELLINGTON
692 EMH, et al. Albaflavenone, a sesquiterpene ketone with a zizaene skeleton produced
693 by a streptomycete with a new rope morphology. *J Antibiot (Tokyo)*. 1994;47(4):434–9.
- 694 43. MASKEY RP, HELMKE E, KAYSER O, FIEBIG HH, MAIER A, BUSCHE A, et al. Anti-cancer
695 and Antibacterial Trioxacarcins with High Anti-malaria Activity from a Marine
696 Streptomycete and their Absolute Stereochemistry. *J Antibiot (Tokyo)*.
697 2004;57(12):771–9.
- 698 44. Takahashi H, Kumagai T, Kitani K, Mori M, Matoba Y, Sugiyama M. Cloning and
699 Characterization of a Streptomyces Single Module Type Non-ribosomal Peptide
700 Synthetase Catalyzing a Blue Pigment Synthesis. *Journal of Biological Chemistry*. 2007
701 Mar;282(12):9073–81.
- 702 45. Walker MC, Eslami SM, Hetrick KJ, Ackenhusen SE, Mitchell DA, van der Donk WA.
703 Precursor peptide-targeted mining of more than one hundred thousand genomes
704 expands the lanthipeptide natural product family. *BMC Genomics*. 2020 Dec
705 3;21(1):387.
- 706 46. Barke J, Seipke RF, Grüşchow S, Heavens D, Drou N, Bibb MJ, et al. A mixed community
707 of actinomycetes produce multiple antibiotics for the fungus farming ant *Acromyrmex*
708 *octospinosus*. *BMC Biol*. 2010 Dec 26;8(1):109.
- 709 47. Gleissner CML, Pyka CL, Heydenreuter W, Gronauer TF, Atzberger C, Korotkov VS, et al.
710 Neocarzin A Is a Potent Inhibitor of Cancer Cell Motility Targeting VAT-1 Controlled
711 Pathways. *ACS Cent Sci*. 2019 Jul 24;5(7):1170–8.
- 712 48. Guha G, Mandal T, Rajkumar V, Ashok Kumar R. Antimycin A-induced mitochondrial
713 apoptotic cascade is mitigated by phenolic constituents of *Phyllanthus amarus*
714 aqueous extract in Hep3B cells. *Food and Chemical Toxicology*. 2010 Dec;48(12):3449–
715 57.
- 716 49. Izumikawa M, Ueda J ya, Chijiwa S, Takagi M, Shin-ya K. Novel GRP78 Molecular
717 Chaperone Expression Down-regulators JBIR-04 and -05 Isolated from *Streptomyces*
718 *violaceoniger*. *J Antibiot (Tokyo)*. 2007 Oct;60(10):640–4.
- 719 50. Ruiz-Villafán B, Cruz-Bautista R, Manzo-Ruiz M, Passari AK, Villarreal-Gómez K,
720 Rodríguez-Sanoja R, et al. Carbon catabolite regulation of secondary metabolite
721 formation, an old but not well-established regulatory system. *Microb Biotechnol*. 2022
722 Apr 6;15(4):1058–72.
- 723 51. Messaoudi O, Bendahou M, Benamar I, Abdelwouhid DE. Identification and
724 preliminary characterization of non-polyene antibiotics secreted by new strain of
725 actinomycete isolated from sebkha of Kenadsa, Algeria. *Asian Pac J Trop Biomed*. 2015
726 Jun;5(6):438–45.
- 727 52. Romero-Rodríguez A, Rocha D, Ruiz-Villafán B, Guzmán-Trampe S, Maldonado-
728 Carmona N, Vázquez-Hernández M, et al. Carbon catabolite regulation in
729 *Streptomyces*: new insights and lessons learned. *World J Microbiol Biotechnol*. 2017
730 Sep 2;33(9):162.
- 731 53. Wang M, Carver JJ, Phelan V V, Sanchez LM, Garg N, Peng Y, et al. Sharing and
732 community curation of mass spectrometry data with Global Natural Products Social
733 Molecular Networking. *Nat Biotechnol*. 2016 Aug 9;34(8):828–37.
- 734 54. Yu Y, Wong J, Lovejoy DB, Kalinowski DS, Richardson DR. Chelators at the Cancer
735 Coalface: Desferrioxamine to Triapine and Beyond. *Clinical Cancer Research*. 2006 Dec
736 1;12(23):6876–83.

- 737 55. Larkin MJ, Kulakov LA, Allen CCR. Genomes and Plasmids in *Rhodococcus*. In 2010. p.
738 73–90.
- 739 56. Letek M, González P, MacArthur I, Rodríguez H, Freeman TC, Valero-Rello A, et al. The
740 Genome of a Pathogenic *Rhodococcus*: Cooptive Virulence Underpinned by Key Gene
741 Acquisitions. *PLoS Genet*. 2010 Sep 30;6(9):e1001145.
- 742 57. Song Y, Xu X, Huang Z, Xiao Y, Yu K, Jiang M, et al. Genomic Characteristics and Pan-
743 Genome Analysis of *Rhodococcus equi*. *Front Cell Infect Microbiol*. 2022 Feb 16;12.
- 744 58. Guo QQ, Ming H, Meng XL, Duan YY, Gao R, Zhang JX, et al. *Rhodococcus agglutinans*
745 sp. nov., an actinobacterium isolated from a soil sample. *Antonie Van Leeuwenhoek*.
746 2015 May 10;107(5):1271–80.
- 747 59. Tamai N, Ishii T, Sato Y, Fujiya H, Muramatsu Y, Okabe N, et al. Bromate Reduction by
748 *Rhodococcus* sp. Br-6 in the Presence of Multiple Redox Mediators. *Environ Sci*
749 *Technol*. 2016 Oct 4;50(19):10527–34.
- 750 60. Cenicerós A, Dijkhuizen L, Petrusma M, Medema MH. Genome-based exploration of
751 the specialized metabolic capacities of the genus *Rhodococcus*. *BMC Genomics*. 2017
752 Dec 9;18(1):593.
- 753 61. Doroghazi JR, Metcalf WW. Comparative genomics of actinomycetes with a focus on
754 natural product biosynthetic genes. *BMC Genomics*. 2013;14(1):611.
- 755 62. Nazari MT, Simon V, Machado BS, Crestani L, Marchezi G, Concolato G, et al.
756 *Rhodococcus*: A promising genus of actinomycetes for the bioremediation of organic
757 and inorganic contaminants. *J Environ Manage*. 2022 Dec;323:116220.
- 758 63. Gamaleldin NM, Bahr HS, Mostafa YA, McAllister BF, El Zawily A, Ngwa CJ, et al.
759 Metabolomic Profiling, In Vitro Antimalarial Investigation and In Silico Modeling of the
760 Marine Actinobacterium Strain *Rhodococcus* sp. UR111 Associated with the Soft Coral
761 *Nephthea* sp. *Antibiotics*. 2022 Nov 15;11(11):1631.
- 762 64. Palma Esposito F, Giugliano R, Della Sala G, Vitale GA, Buonocore C, Ausuri J, et al.
763 Combining OSMAC Approach and Untargeted Metabolomics for the Identification of
764 New Glycolipids with Potent Antiviral Activity Produced by a Marine *Rhodococcus*. *Int J*
765 *Mol Sci*. 2021 Aug 22;22(16):9055.
- 766 65. Alvarez HM, Herrero OM, Silva RA, Hernández MA, Lanfranconi MP, Villalba MS.
767 Insights into the Metabolism of Oleaginous *Rhodococcus* spp. *Appl Environ Microbiol*.
768 2019 Sep 15;85(18).
- 769 66. Alvarez HM, Kalscheuer R, Steinbüchel A. Accumulation and mobilization of storage
770 lipids by *Rhodococcus opacus* PD630 and *Rhodococcus ruber* NCIMB 40126. *Appl*
771 *Microbiol Biotechnol*. 2000 Aug 15;54(2):218–23.
- 772 67. Helfrich EJN, Vogel CM, Ueoka R, Schäfer M, Ryffel F, Müller DB, et al. Bipartite
773 interactions, antibiotic production and biosynthetic potential of the *Arabidopsis* leaf
774 microbiome. *Nat Microbiol*. 2018 Aug 1;3(8):909–19.
- 775 68. Chakraborty K, Kizhakkekalam VK, Joy M. Polyketide-derived macrobrevins from
776 marine macroalga-associated *Bacillus amyloliquefaciens* as promising antibacterial
777 agents against pathogens causing nosocomial infections. *Phytochemistry*. 2022
778 Jan;193:112983.
- 779 69. Singh PK, Chittipurna, Ashish, Sharma V, Patil PB, Korpole S. Identification, Purification
780 and Characterization of Laterosporulin, a Novel Bacteriocin Produced by *Brevibacillus*
781 sp. Strain GI-9. *PLoS One*. 2012 Mar 5;7(3):e31498.
- 782 70. Baidara P, Singh N, Ranjan M, Nallabelli N, Chaudhry V, Pathania GL, et al.
783 Laterosporulin10: a novel defensin like Class IId bacteriocin from *Brevibacillus* sp.

- 784 strain SKDU10 with inhibitory activity against microbial pathogens. *Microbiology* (N Y).
785 2016 Aug 1;162(8):1286–99.
- 786 71. GHADBANE M, HARZALLAH D, LARIBI AI, JAOUADI B, BELHADJ H. Purification and
787 Biochemical Characterization of a Highly Thermostable Bacteriocin Isolated from
788 *Brevibacillus brevis* Strain GM100. *Biosci Biotechnol Biochem*. 2013 Jan 23;77(1):151–
789 60.
- 790 72. Mordhorst S, Ruijne F, Vagstad AL, Kuipers OP, Piel J. Emulating nonribosomal peptides
791 with ribosomal biosynthetic strategies. *RSC Chem Biol*. 2023;4(1):7–36.
- 792 73. Rico S, Santamaría RI, Yepes A, Rodríguez H, Laing E, Bucca G, et al. Deciphering the
793 Regulon of *Streptomyces coelicolor* AbrC3, a Positive Response Regulator of Antibiotic
794 Production. *Appl Environ Microbiol*. 2014 Apr 15;80(8):2417–28.
- 795 74. Dong H, Ming D. A Comprehensive Self-Resistance Gene Database for Natural-Product
796 Discovery with an Application to Marine Bacterial Genome Mining. *Int J Mol Sci*. 2023
797 Aug 4;24(15):12446.
- 798 75. Mohamed OG, Dorandish S, Lindow R, Steltz M, Shoukat I, Shoukat M, et al.
799 Identification of a New Antimicrobial, Desertomycin H, Utilizing a Modified Crowded
800 Plate Technique. *Mar Drugs*. 2021 Jul 27;19(8):424.
801

Figure 01. Visual aspect of colonies of (a) *Streptomyces* sp. ACT015, (b) *Rhodococcus* sp. ACT016, and (c) *Brevibacillus* sp. FIR094, grown on Tryptone Soy Agar (TSA).

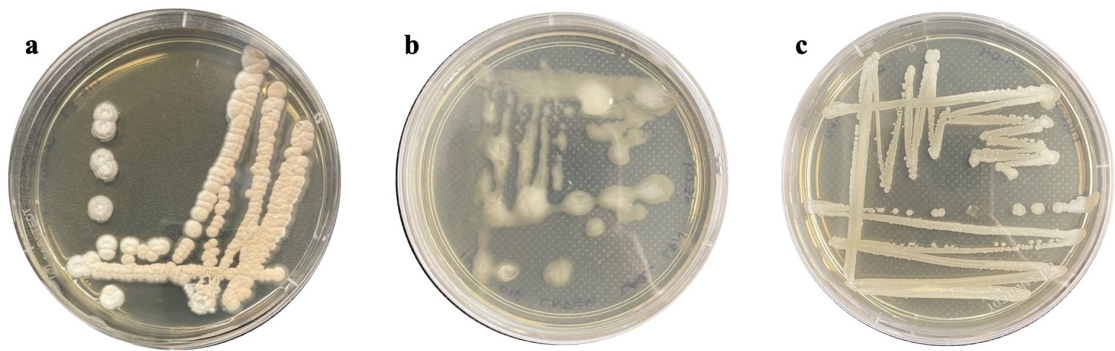


Figure 02. (a) Circular map of *Streptomyces* sp. ACT015 genome. From the outermost ring to the innermost: locations of the 33 Biosynthetic Gene Clusters; genome size; Coding Sequences (forward and reverse); the genomes of *S. cadmiisoli* ZFG47; *S. griseoviridis* F1-27; *S. griseoviridis* K61; *S. koyangensis* SCSIO 5802; *S. koyangensis* VK-A60T; *Streptomyces* sp. HF10; *Streptomyces* sp. NEAU-sy36; *Streptomyces* sp. SUK 48; *Streptomyces* sp. WA1-19; and *Streptomyces* sp. WP-1); G+C skew; and G+C content. (b) Cluster of Orthologous Groups functional categories.

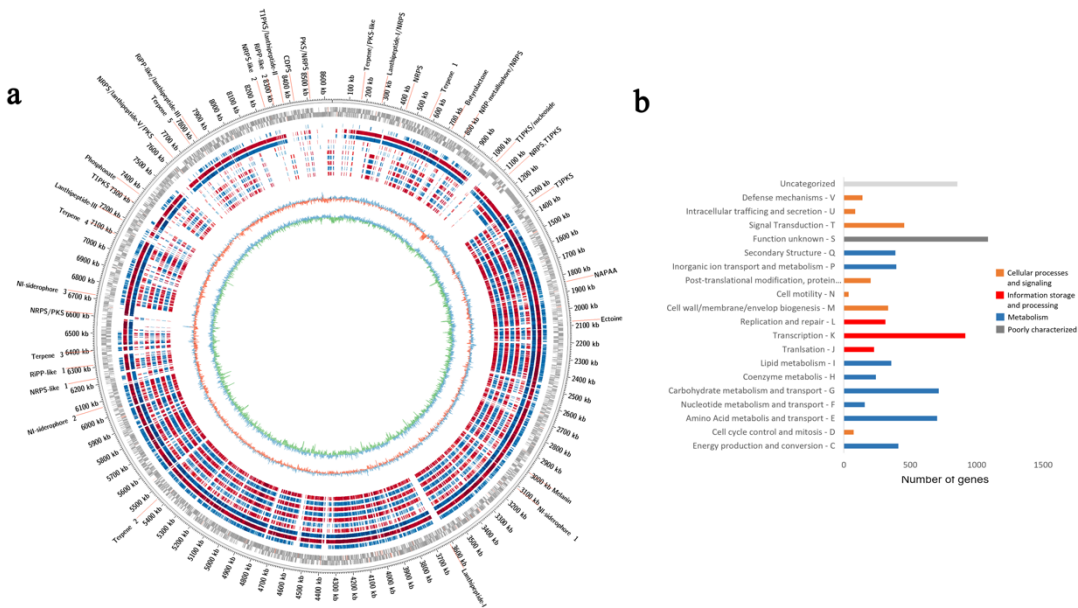


Figure 03. (a) Circular map of *Rhodococcus* sp. ACT016 genome. From the outermost ring to the innermost: locations of the 17 Biosynthetic Gene Clusters; genome size; Coding Sequences (forward and reverse); the genomes of *Rhodococcus* sp. W8901; *Prescottella equi* ATCC 33701; *P. equi* BJ13; *P. equi* DSSKP-R-001; *P. equi* FDAARGOS 952; *P. equi* JCM94-14; *P. equi* JCM94-27; *P. equi* JCM94-31; *P. equi* JCM94-3; *P. equi* JID03-46; *P. equi* JID03-56; *P. equi* PAM2287; and *P. equi* U19; G+C skew; and G+C content; **(b)** Cluster of Orthologous Groups functional categories; **(c)** Putative plasmid with 642,065 bp, 65% GC content and 2 BGCs.

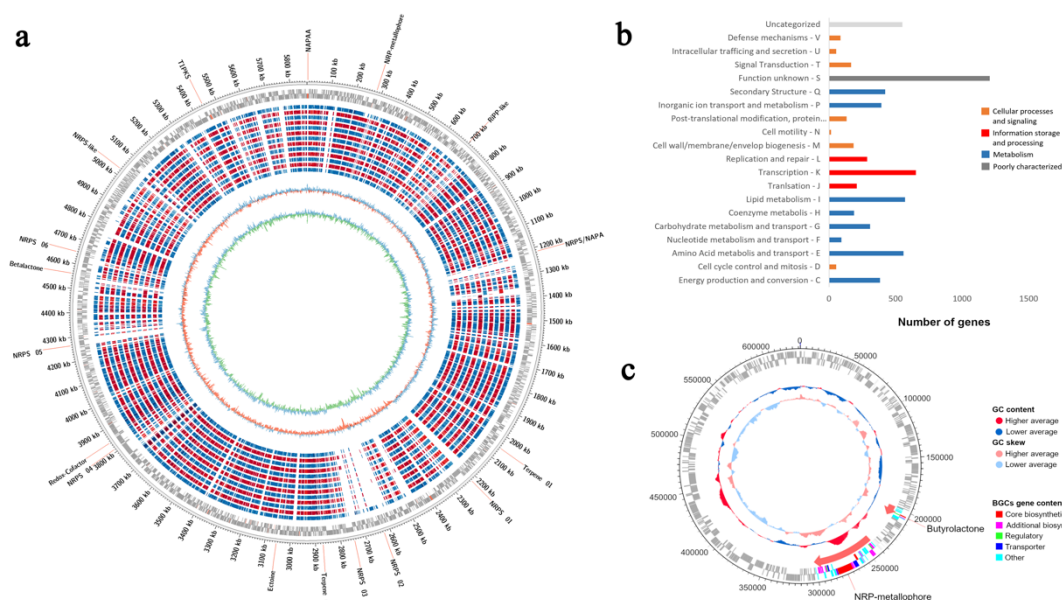


Figure 04. (a) Circular map of *Brevibacillus* sp. FIR094 genome. From the outermost ring to the innermost: locations of the 14 Biosynthetic Gene Clusters; genome size; Coding Sequences (forward and reverse); the genomes of *B. brevis* B011; *B. brevis* DZQ7; *B. brevis* HK544; *B. brevis* HNCS-1; *B. brevis* NBRC-100599; *B. brevis* NCTC2611-1; *B. choshinensis* HPD31-SP3; *B. formosus* NF2; *B. parabrevis* B3; *B. parabrevis* BCP-09; *Brevibacillus* sp. DP1.3A; and *Brevibacillus* sp. HD3.3A; G+C skew; and G+C content. **(b)** Cluster of Orthologous Groups functional categories.

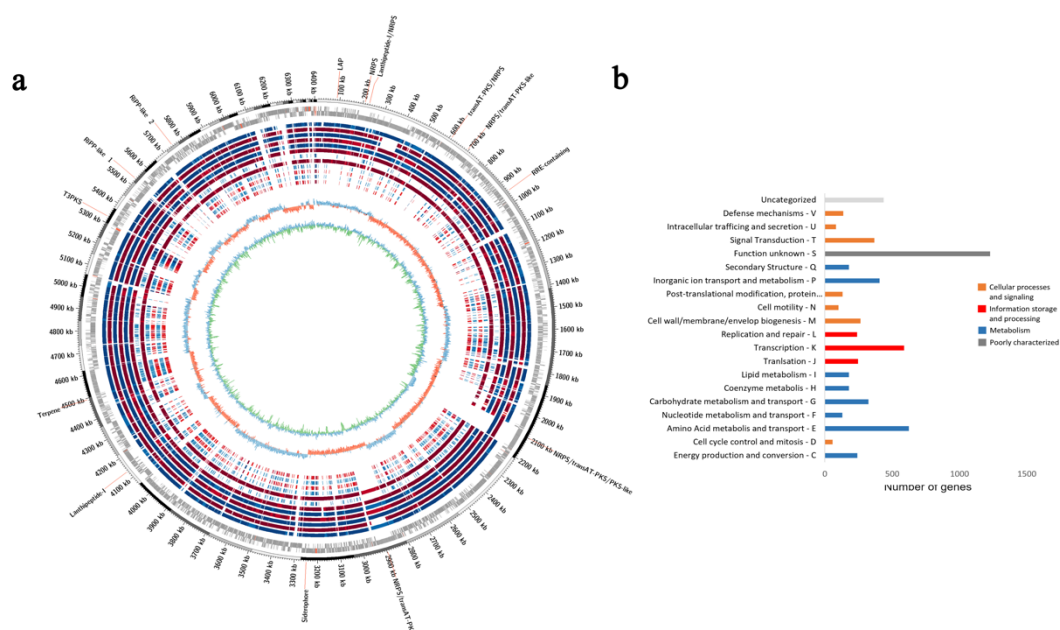


Figure 05. Heatmap of enriched KEGG metabolism pathways. The color scale (right side) corresponds to the relative value of metabolites enriched in each KEGG pathway, being yellow the “highest” enrichment and blue the “lower” enrichment. The upper dendrogram was generated using hierarchical clustering and represents the distance across the predicted metabolic profiles.

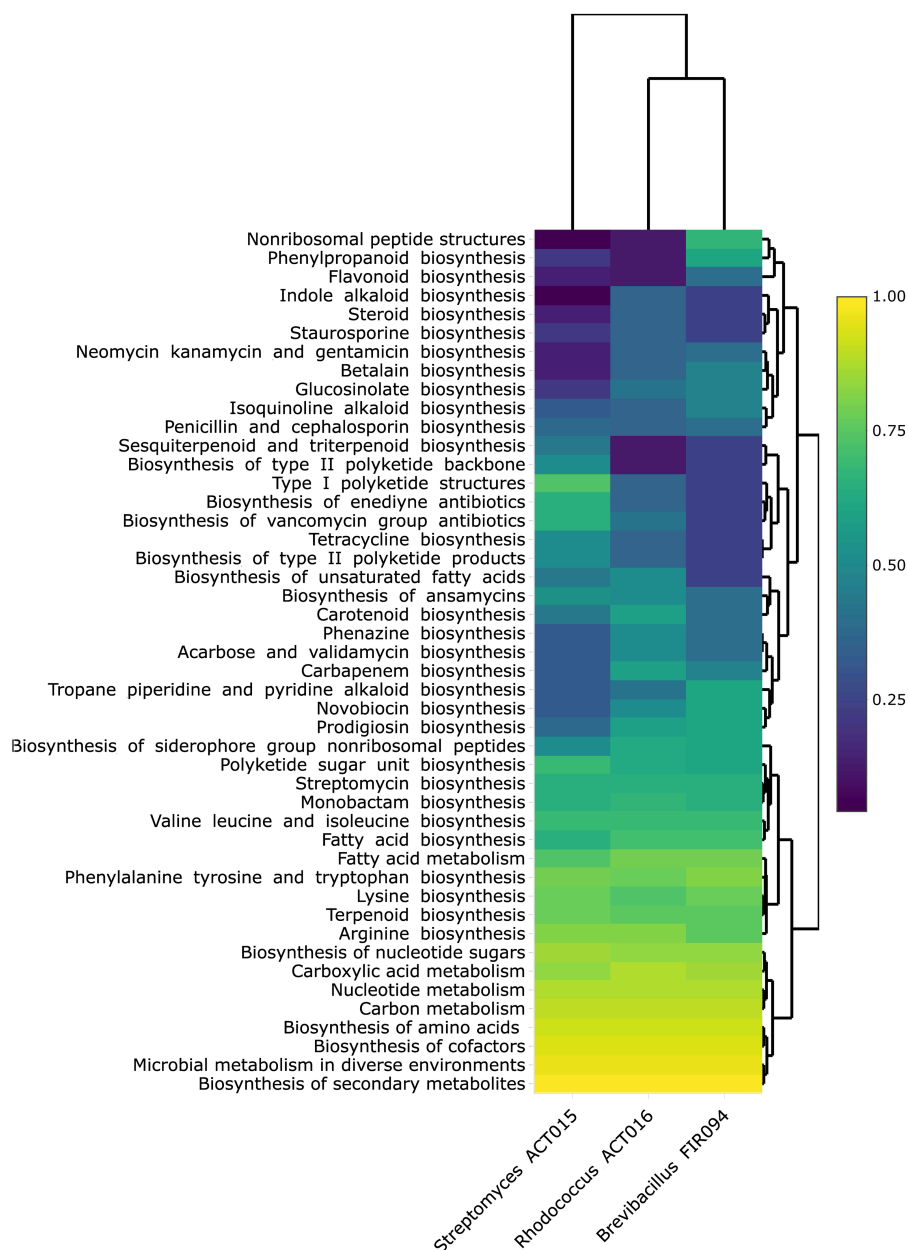


Figure 06. (a) Spectra Similarity Molecular Networking (SSMN) representing the secondary metabolites synthesized by the strains grown in different culture media. Each spectrum is represented by a node and the level of similarity between nodes is directly proportional to the size of the edge that connects those nodes. Orange, purple, and green nodes represent spectra acquired from the metabolome of *Brevibacillus* sp. FIR094, *Streptomyces* sp. ACT015, and *Rhodococcus* sp. ACT016, respectively. The size of the node indicates the percentage of ion contribution (peak area). Pink circles (UNPD-ISDB), black squares (GNPS), or pink squares (for both) identify the database responsible for the metabolite annotation. (b) Venn diagram with the total number of spectra detected in the crude extracts of the three strains. (c) Venn diagram illustrating the distribution of metabolite production in each of the culture media in *Streptomyces* sp. ACT015. (d) Venn diagram illustrating the distribution of metabolite production in each of the culture media in *Rhodococcus* sp. ACT016. (e) Venn diagram illustrating the distribution of metabolite production in each of the culture media in *Brevibacillus* sp. FIR094.

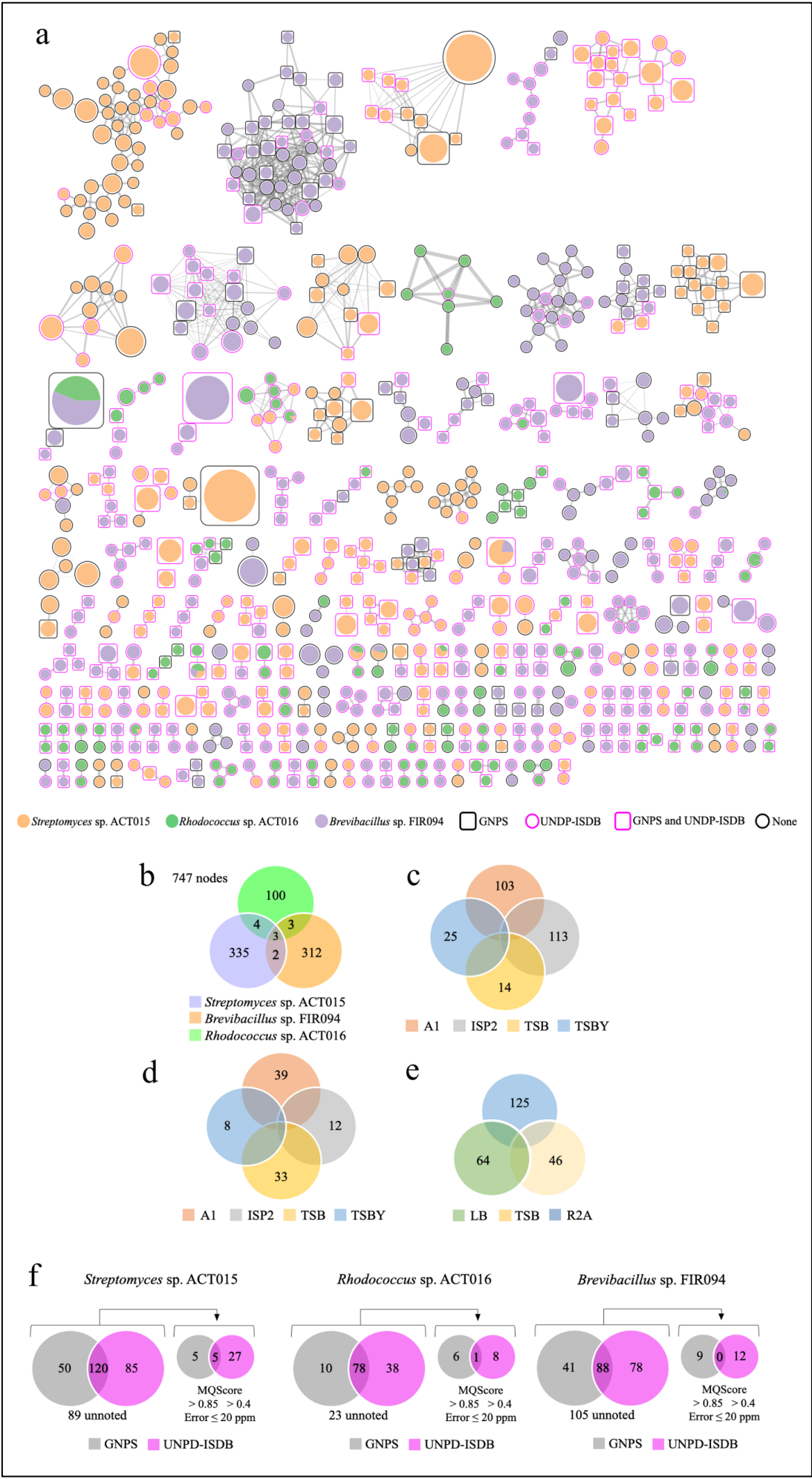


Figure 7. (a) Spectra Similarity Molecular Network (SSMN) representing the chemical superclass of each node. Each superclass is represented by a different color identified by the legend. Unidentified spectra are represented by white nodes. **(b)** LC-MS/MS analysis highlights the presence of Antimycin A4 (1) and Deferoxamine E (2) in *Streptomyces* sp. ACT015 crude extract, validating BGCs 20 and 13 respectively. The clusters are highlighted in the SSMN by dotted circle.

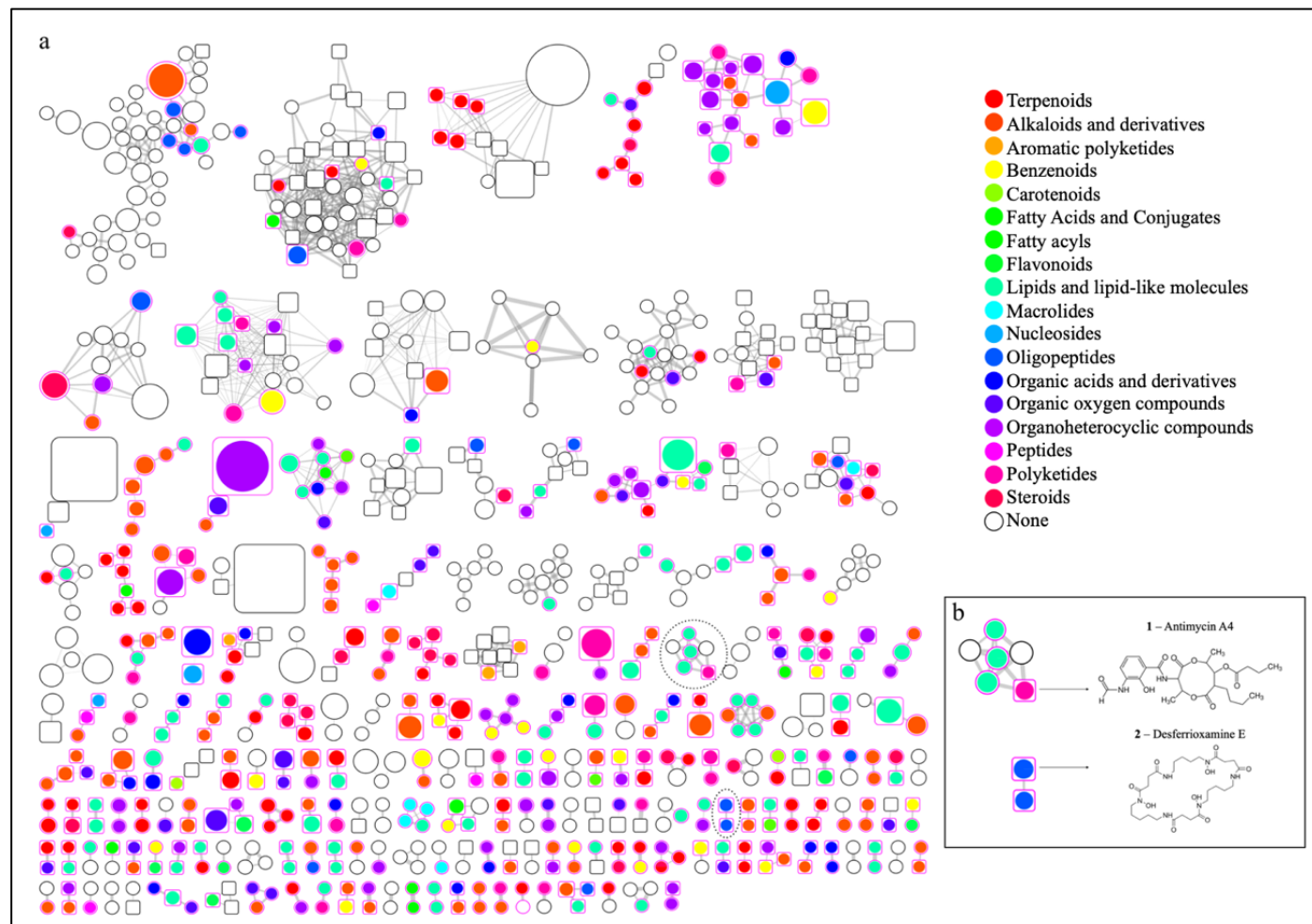


Figure 8. (a) BGC from *Brevibacillus* sp. FIR094 suggested as a transcriptional region for RiPP subclass of linear peptides containing azole(s) (LAPs). The essential genes of the BGCs are depicted as colored arrows, including the precursor peptide, maturation protein, cyclodehydratase and dehydratase enzymes, the regulator, and the transporter. The precursor peptide is represented by the leader region (dark gray) and the core region (light gray). The predicted amino acid sequence is rich in cysteine and serine. The leader peptide is cleaved after maturation. **(b)** The chemical structure of the molecule predicted from the amino acid sequence of the core peptide.

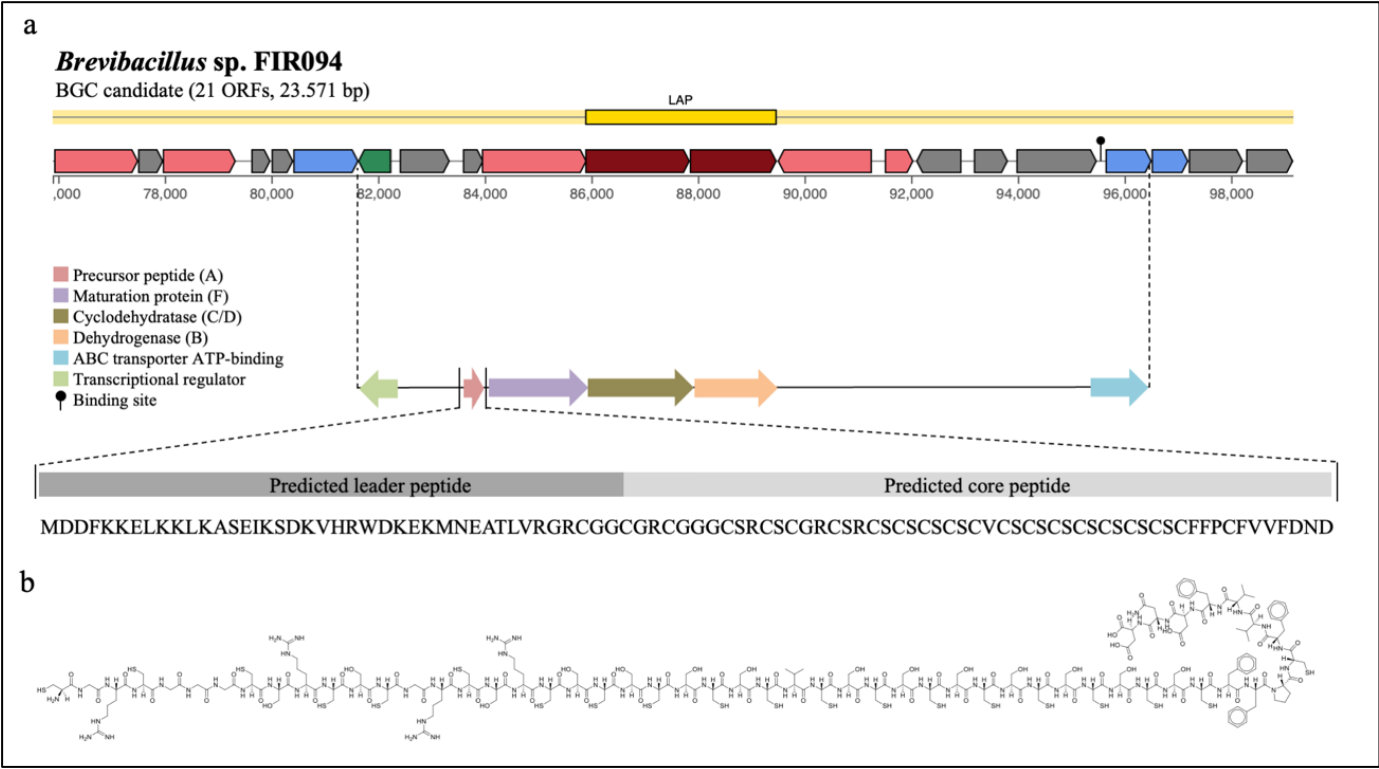


Figure 9. (a) Binding sites to the transcription factor AbrC3. **(b)** ClusterBlast analysis with regions from the AntiSMASH database that are similar to the LAP BGC sequence. Genes marked with the same color are interrelated. Genes in white are unrelated. **(c)** Similar known gene clusters from MIBiG 3.1. Genes marked with the same color are interrelated. Genes in white are unrelated.

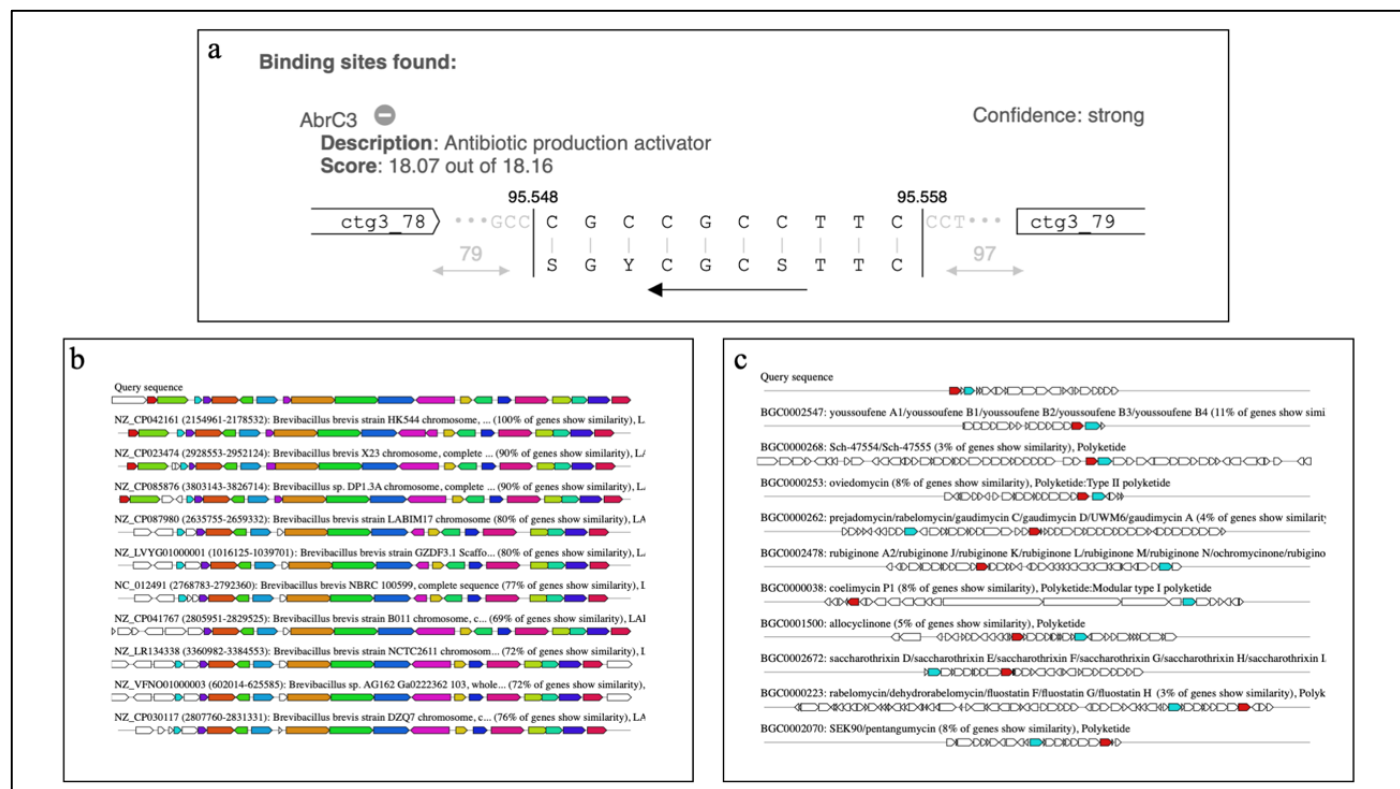


Table 1. General genomic characteristics of the three isolated strains.

Strains	<i>Streptomyces</i> sp. ACT015	<i>Rhodococcus</i> sp. ACT016	<i>Brevibacillus</i> sp. FIR094
Assembly			
Genome size (pb)	8,643,132	5,868,135	6,348,900
Number of contigs	1	1	117
N50 (bp)	8,643,132	5,868,135	116,003
G+C (%)	73	68,4	47,2
Coverage (x)	329	82	507
BUSCO			
Complete and single copy	1,562	730	437
Complete and duplicate	6	11	4
Fragmented	6	2	6
Missing	5	0	3
Total BUSCO genes	1,579	743	450
BUSCO completeness (%)	99,4	99,8	98,0
Annotation			
Number of CDS	7,161	5,919	3,309
Number of hypothetical genes	1,579	743	2,333

Table 2. Biosynthetic Gene Clusters predicted by AntiSMASH in the *Streptomyces* sp. ACT015 genome.

BGC	Biosynthetic pathway	Similar cluster	Similarity
BGC 1	Terpene	Isorenieratene	100%
		Carotenoid	87%
		Rustimicin	33%
BGC 2	RiPP-Lanthipeptide I	-	0%
	RiPP-Lanthipeptide II		
	NRPS		
BGC 3	NRPS-PKS	Detoxin P1/P2/P3	50%
		Detoxin N2/N3	40%
BGC 4	Terpene	-	0%
BGC 5	Butyrolactone	-	0%
BGC 6	NRPS	Amychelin A/B	81%
		Cahuitamycin A/B/C	62%
BGC 7	T1PKS	Chlorothricin	32%
	Nucleoside	-	0%
BGC 8	NRPS-T1PKS	Candididin	100%
BGC 9	T3PKS	Flaviolin	100%
BGC 10	Other	ϵ -Poli-L-lysine	100%
BGC 11	Other	Ectoine	100%
BGC 12	Other	Melanin	100%
BGC 13	Other	Deferoxamine E	100%
		Deferoxamine A/B	83%
		FW0622	62%
BGC 14	RiPP-Lanthipeptide	-	0%
BGC 15	Terpene	Albaflavenone	100%
BGC 16	PKS	Kinamycin	13%
BGC 17	NRPS	Indigoidine	100%
		Minimycin	100%
BGC 18	RiPP	-	0%
BGC 19	Terpene	Geosmin	100%
		Antimycin	100%
BGC 20	NRPS-PKS	Splenocin C	93%
		Neoantimycin	76%
		Trioxacarcin A	71%
BGC 21	Siderophore	Paulomycin	13%
BGC 22	Terpene	Hopene	92%
BGC 23	RiPP-Lanthipeptide	-	0%
BGC 24	T1PKS	Streptazone E	75%
BGC 25	Phosphonate	-	0%

	T1PKS	Neocarzilin A/B	85%
BGC 26	NRPS	-	0%
	RiPP-Lanthipeptide V	-	0%
BGC 27	Terpene	-	0%
BGC 28	RiPP-Lanthipeptide III	SapB	75%
BGC 29	NRPS	Azalomycin	13%
BGC 30	RiPP	-	0%
	RiPP-Lanthipeptide	Birimositide	100%
BGC 31	T1PKS	Amicomycin	25%
	Betalactone	-	0%
BGC 32	Other	EDHA	100%
	NRPS-T1PKS	Totopotensamide A/B	58%
BGC 33	T3PKS	-	0%
	NRPS	-	0%

Table 3: Biosynthetic Gene Clusters predicted by AntiSMASH in the *Rhodococcus* sp. ACT016 genome.

BGC	Biosynthetic pathway	Similar cluster	Similarity
BGC 1.1	Other	ϵ -Poly-L-lysine	100%
BGC 1.2	NRPS	Thermochelin	38%
BGC 1.3	RiPP	-	0%
BGC 1.4	NRPS	Calicheamicin	8%
BGC 1.5	Terpene	-	0%
BGC 1.6	NRPS	-	0%
BGC 1.7	NRPS	-	0%
BGC 1.8	NRPS	-	0%
BGC 1.9	Terpene	Isorenieratene	42%
BGC 1.10	Other	Ectoine	75%
BGC 1.11	NRPS	Corynecin	53%
BGC 1.12	Redox cofactor	-	0%
BGC 1.13	NRPS	-	0%
BGC 1.14	Betalactone	-	0%
BGC 1.15	NRPS	-	0%
BGC 1.16	NRPS	-	0%
	Other	Phosphoramidon	40%
BGC 1.17	T1PKS	-	0%
BGC 2.1	Butyrolactone	-	0%
BGC 2.1	NRPS	Heterobactin	63%

Table 4. Biosynthetic Gene Clusters predicted by AntiSMASH in the *Brevibacillus* sp. FIR094 genome.

BGC	Biosynthetic pathway	Similar cluster	Similarity
BGC 1	RiPP-LAP	Youssoufene	11%
BGC 2	NRPS	Tyrocidine	81%
BGC 3	NRPS	Gramicidin	91%
	RiPP-Lanthipeptide I	-	0%
BGC 4	NRPS-PKS	Zwittermicin A	11%
BGC 5	NRPS-PKS	-	0%
BGC 6	RiPP	-	0%
BGC 7	T1PKS Trans-AT	Macrobrevin	100%
	NRPS	-	0%
BGC 8	NRPS-PKS Trans-AT	Aurantinin	25%
BGC 9	Siderophore	Petrobactina	83%
BGC 10	RiPP	TVA-YJ-2	4%
BGC11	Terpene	-	0%
BGC 12	PKS	Bacitracin	22%
BGC 13	RiPP	-	0%
BGC 14	RiPP	-	0%

The figure consists of two maps. The left map shows the state of Pará, Brazil, with the Brazilian Amazon region highlighted in green and other biomes in white. A dashed line indicates the location of Belém. The right map is a detailed view of the City of Belém, showing the Utinga State Park (green area) and the surrounding urban area (grey area). Three sample collection points are marked: A (black triangle), B (black circle), and C (black square). A legend at the bottom left identifies the green area as 'Brazilian Amazon' and the white area as 'Others biomes'. A legend at the bottom right identifies the black triangle as 'Sample collection points'.

Supplementary figure 3: Analysis carried out on contig 2 of the Rhodococcus sp. ACT016 genome.



Supplementary table 1: References used in comparative genomics analysis for the three strains.

Reference genome	Access Number	Identity (%)*
<i>Streptomyces cadmiisoli</i> ZFG47	CP030073.1	99.15
<i>Streptomyces griseoviridis</i> F1-27	CP034687.1	99.87
<i>Streptomyces griseoviridis</i> K61	CP029078.1	99.87
<i>Streptomyces koyangensis</i> SCSIO 5802	CP049945.1	99.21
<i>Streptomyces koyangensis</i> VK-A60T	CP031742.1	99.28
<i>Streptomyces</i> sp. HF10	CP047144.1	99.21
<i>Streptomyces</i> sp. NEAU-sy36	CP058977.1	99.28
<i>Streptomyces</i> sp. SUK 48	CP045740.1	99.02
<i>Streptomyces</i> sp. WA1-19	CP085039.1	99.08
<i>Streptomyces</i> sp. WP-1	CP123923.1	99.21
<i>Rhodococcus</i> sp. W8901	CP054690.1	99.14
<i>Prescottella equi</i> ATCC 33701	AP025268.1	98.68
<i>Prescottella equi</i> BJ13	CP118697.1	98.68
<i>Prescottella equi</i> DSSKP-R-001	CP027793.1	98.68
<i>Prescottella equi</i> FDAARGOS 952	CP065594.1	98.68
<i>Prescottella equi</i> JCM94-14	AP024181.1	98.68
<i>Prescottella equi</i> JCM94-27	AP024187.1	98.68
<i>Prescottella equi</i> JCM94-31	AP024189.1	98.68
<i>Prescottella equi</i> JCM94-3	AP024192.1	98.68
<i>Prescottella equi</i> JID03-46	AP024196.1	98.68
<i>Prescottella equi</i> JID03-56	AP024198.1	98.68
<i>Prescottella equi</i> PAM2287	CP095477.1	98.68
<i>Prescottella equi</i> U19	AP025544.1	98.68
<i>Brevibacillus brevis</i> B011	CP041767.1	99.74
<i>Brevibacillus brevis</i> DZQ7	CP030117.1	99.61
<i>Brevibacillus brevis</i> HK544	CP042161.1	100.00
<i>Brevibacillus brevis</i> HNCS-1	CP128411.1	99.61
<i>Brevibacillus brevis</i> NBRC-100599	AP008955.1	99.67
<i>Brevibacillus brevis</i> NCTC2611-1	LR134338.1	99.80
<i>Brevibacillus choshinensis</i> HPD31-SP3	CP069127.1	99.15
<i>Brevibacillus formosus</i> NF2	CP018145.1	99.74
<i>Brevibacillus parabrevis</i> B3	CP064090.1	99.22
<i>Brevibacillus parabrevis</i> BCP-09	CP118544.1	99.41
<i>Brevibacillus</i> sp. DP1.3A	CP085876.1	99.74
<i>Brevibacillus</i> sp. HD3.3A	CP085874.1	99.35

*Identity based on strains 16s BLAST against nr database (NCBI). All alignments had a sequence cover >99%.

6 CONCLUSÃO

Aqui, buscamos avaliar o potencial biossintético de três cepas de bactérias ambientais de espécies diferentes isoladas da Amazônia brasileira com abordagens integradas em genômica e metabolômica. A mineração do genoma permitiu avaliar a capacidade das bactérias em carrear genes responsáveis por uma variedade de compostos antimicrobianos e antitumorais. As classes biossintéticas predominantes foram PKS, RiPP e NRPS em *Streptomyces* sp. ACT015, NRPS em *Rhodococcus* sp. ACT016 e RiPP, NRPS e NRPS-PKS em *Brevibacillus* sp. FIR094. Além disso, foi possível observar que muitos BGCs não possuem anotação no principal repositório atualmente disponível. Em *Streptomyces* sp. ACT015, 15 BGCs (45%) não foram identificados, em *Rhodococcus* sp. ACT016 15 BGCs (88%) e em *Brevibacillus* sp. FIR094 10 BGCs (71%). Isto sugere que o inventário completo de metabólitos especializados de cepas bacterianas de solo amazônico permanece ainda pouco explorado.

O metaboloma não direcionado a partir de cultivos com meios diferentes revelou a significativa manipulação dos meios de cultura e a produção de moléculas com espectros correspondentes a importantes classes químicas de interesse biotecnológico, como alcaloides, policetídeos e peptídeos. Houve predominância de espectros sem anotação nos bancos de dados nas três cepas envolvidas no estudo, com 89 nodos, 23 nodos e 105 nodos em *Streptomyces* sp. ACT015, *Rhodococcus* sp. ACT016 e *Brevibacillus* sp. FIR094, respectivamente. Esses resultados indicam uma alta variabilidade metabólica nos microrganismos amazônicos, que poderiam ser explorados para a produção de metabólitos de utilidade médica e biotecnológica.

A partir da integração de análises multi-ômicas foi possível comprovar a enorme possibilidade de metabólitos secundários produzidos por bactérias de solo amazônico. Neste estudo, foram raros os casos em que pudemos correlacionar o BGC previsto com seu metabólito, possivelmente devido à novidade dos BGCs e espectros detectados. A diversidade química que ainda aguarda descoberta é vasta, e a capacidade de manipular o uso de tecnologias para desencadear a produção e caracterizar compostos específicos é promissora. Ao mesmo tempo, os bancos de dados continuam a evoluir, mas ainda há muito trabalho a ser feito para preencher as lacunas na anotação de metabólitos secundários microbianos. À medida que avançamos nessa pesquisa, desvendar os metabólitos secundários promete revelar novas aplicações em áreas como a biotecnologia, a medicina e a indústria farmacêutica.

REFERÊNCIAS

- ABRAHAM, T. et al. Structure–activity relationships of the antimicrobial peptide gramicidin S and its analogs: Aqueous solubility, self-association, conformation, antimicrobial activity and interaction with model lipid membranes. **Biochimica et Biophysica Acta (BBA) - Biomembranes**, v. 1838, n. 5, p. 1420–1429, maio 2014.
- AKBAR, N. et al. Novel Antiamoebic Tyrocidine-Derived Peptide against Brain-Eating Amoebae. **ACS Omega**, v. 7, n. 33, p. 28797–28805, 23 ago. 2022.
- ALAM, K. et al. Streptomyces: The biofactory of secondary metabolites. **Frontiers in Microbiology**, v. 13, 29 set. 2022.
- ALBARANO, L. et al. Genome Mining as New Challenge in Natural Products Discovery. **Marine Drugs**, v. 18, n. 4, p. 199, 9 abr. 2020.
- ALLARD, P.-M. et al. Integration of Molecular Networking and *In-Silico* MS/MS Fragmentation for Natural Products Dereplication. **Analytical Chemistry**, v. 88, n. 6, p. 3317–3323, 15 mar. 2016.
- ALVAREZ, H. **Biology of Rhodococcus**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. v. 16
- ALVAREZ, H. M. et al. Rhodococcus as Biofactories for Microbial Oil Production. **Molecules**, v. 26, n. 16, p. 4871, 11 ago. 2021.
- ARNISON, P. G. et al. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. **Nat. Prod. Rep.**, v. 30, n. 1, p. 108–160, 2013a.
- ARNISON, P. G. et al. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. **Natural product reports**, v. 30, n. 1, p. 108–60, jan. 2013b.
- ARON, A. T. et al. Reproducible molecular networking of untargeted mass spectrometry data using GNPS. **Nature Protocols**, v. 15, n. 6, p. 1954–1991, 13 jun. 2020.
- BALE, B. I. et al. Antibiotic resistance in ocular bacterial infections: an integrative review of ophthalmic chloramphenicol. **Tropical Medicine and Health**, v. 51, n. 1, p. 15, 9 mar. 2023.
- BALOUIRI, M.; SADIKI, M.; IBNSOUDA, S. K. Methods for in vitro evaluating antimicrobial activity: A review. **Journal of Pharmaceutical Analysis**, v. 6, n. 2, p. 71–79, abr. 2016.
- BALTZ, R. H. Natural product drug discovery in the genomic era: realities, conjectures, misconceptions, and opportunities. **Journal of Industrial Microbiology and Biotechnology**, v. 46, n. 3–4, p. 281–299, 1 mar. 2019.
- BANKEVICH, A. et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. **Journal of Computational Biology**, v. 19, n. 5, p. 455–477, maio 2012.
- BARKA, E. A. et al. Taxonomy, Physiology, and Natural Products of Actinobacteria. **Microbiology and Molecular Biology Reviews**, v. 80, n. 1, p. 1–43, mar. 2016.
- BENSON, D. A. et al. GenBank. **Nucleic Acids Research**, v. 45, n. D1, p. D37–D42, 4 jan. 2017.

- BENTLEY, S. D. et al. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). **Nature**, v. 417, n. 6885, p. 141–147, maio 2002.
- BERMAN, H. M. The Protein Data Bank: a historical perspective. **Acta Crystallographica Section A Foundations of Crystallography**, v. 64, n. 1, p. 88–95, 1 jan. 2008.
- BIBB, M. J. Regulation of secondary metabolism in streptomycetes. **Current Opinion in Microbiology**, v. 8, n. 2, p. 208–215, abr. 2005.
- BLIN, K. et al. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. **Nucleic Acids Research**, v. 51, n. W1, p. W46–W50, 5 jul. 2023.
- BODE, H. B. et al. Big Effects from Small Changes: Possible Ways to Explore Nature's Chemical Diversity. **ChemBioChem**, v. 3, n. 7, p. 619, 3 jul. 2002.
- BOSELLO, M. et al. Biosynthesis of the Siderophore Rhodochelin Requires the Coordinated Expression of Three Independent Gene Clusters in *Rhodococcus jostii* RHA1. **Journal of the American Chemical Society**, v. 133, n. 12, p. 4587–4595, 30 mar. 2011.
- BOSELLO, M. et al. Structural Characterization of the Heterobactin Siderophores from *Rhodococcus erythropolis* PR4 and Elucidation of Their Biosynthetic Machinery. **Journal of Natural Products**, v. 76, n. 12, p. 2282–2290, 27 dez. 2013.
- BOSSO, J. A.; MAULDIN, P. D.; SALGADO, C. D. The association between antibiotic use and resistance: the role of secondary antibiotics. **European Journal of Clinical Microbiology & Infectious Diseases**, v. 29, n. 9, p. 1125–1129, 10 set. 2010.
- BREINBAUER, R.; VETTER, I. R.; WALDMANN, H. From protein domains to drug candidates-natural products as guiding principles in the design and synthesis of compound libraries. **Angewandte Chemie (International ed. in English)**, v. 41, n. 16, p. 2879–90, 16 ago. 2002.
- BRZESZCZ, J. et al. Bacteria degrading both n-alkanes and aromatic hydrocarbons are prevalent in soils. **Environmental Science and Pollution Research**, 21 dez. 2023.
- BUKHARI, S.; TABOADA, J. A Case of Miller Fisher Syndrome and Literature Review. **Cureus**, 22 fev. 2017.
- CAFFREY, P. et al. Amphotericin biosynthesis in *Streptomyces nodosus*: deductions from analysis of polyketide synthase and late genes. **Chemistry & Biology**, v. 8, n. 7, p. 713–723, jul. 2001.
- CASTRO BRAGA, F. Brazilian traditional medicine: Historical basis, features and potentialities for pharmaceutical development. **Journal of Traditional Chinese Medical Sciences**, v. 8, p. S44–S50, nov. 2021.
- CENICEROS, A. et al. Genome-based exploration of the specialized metabolic capacities of the genus *Rhodococcus*. **BMC Genomics**, v. 18, n. 1, p. 593, 9 dez. 2017.
- CHAKRABORTY, K.; KIZHAKKEKALAM, V. K.; JOY, M. Polyketide-derived macrobrevins from marine macroalga-associated *Bacillus amyloliquefaciens* as promising antibacterial agents against pathogens causing nosocomial infections. **Phytochemistry**, v. 193, p. 112983, jan. 2022.

- CHIBA, H. et al. Rhodopeptins, Novel Cyclic Tetrapeptides with Antifungal Activities from *Rhodococcus* sp. III. Synthetic Study of Rhodopeptins. **The Journal of Antibiotics**, v. 52, n. 8, p. 710–720, 1999.
- CHOPRA, I.; ROBERTS, M. Tetracycline Antibiotics: Mode of Action, Applications, Molecular Biology, and Epidemiology of Bacterial Resistance. **Microbiology and Molecular Biology Reviews**, v. 65, n. 2, p. 232–260, jun. 2001.
- CHOULET, F. et al. Intraspecific Variability of the Terminal Inverted Repeats of the Linear Chromosome of *Streptomyces ambofaciens*. **Journal of Bacteriology**, v. 188, n. 18, p. 6599–6610, 15 set. 2006.
- CHU, J. et al. Discovery of MRSA active antibiotics using primary sequence from the human microbiome. **Nature Chemical Biology**, v. 12, n. 12, p. 1004–1006, 17 dez. 2016.
- COBB, R. E.; WANG, Y.; ZHAO, H. High-Efficiency Multiplex Genome Editing of *Streptomyces* Species Using an Engineered CRISPR/Cas System. **ACS Synthetic Biology**, v. 4, n. 6, p. 723–728, 19 jun. 2015.
- CRAGG, G. M.; NEWMAN, D. J. Natural products: A continuing source of novel drug leads. **Biochimica et Biophysica Acta (BBA) - General Subjects**, v. 1830, n. 6, p. 3670–3695, jun. 2013.
- CRANEY, A.; AHMED, S.; NODWELL, J. Towards a new science of secondary metabolism. **The Journal of Antibiotics**, v. 66, n. 7, p. 387–400, 24 jul. 2013.
- CROSA, J. H.; WALSH, C. T. Genetics and Assembly Line Enzymology of Siderophore Biosynthesis in Bacteria. **Microbiology and Molecular Biology Reviews**, v. 66, n. 2, p. 223–249, jun. 2002.
- CZAJGUCKI, Z.; ANDRUSZKIEWICZ, R.; KAMYSZ, W. Structure activity relationship studies on the antimicrobial activity of novel edeine A and D analogues. **Journal of Peptide Science**, v. 12, n. 10, p. 653–662, out. 2006.
- DA SILVA, R. R.; DORRESTEIN, P. C.; QUINN, R. A. Illuminating the dark matter in metabolomics. **Proceedings of the National Academy of Sciences**, v. 112, n. 41, p. 12549–12550, 13 out. 2015.
- DAS, A.; KHOSLA, C. Biosynthesis of aromatic polyketides in bacteria. **Accounts of chemical research**, v. 42, n. 5, p. 631–9, 19 maio 2009.
- DE FELÍCIO, R. et al. Chemical Elicitors Induce Rare Bioactive Secondary Metabolites in Deep-Sea Bacteria under Laboratory Conditions. **Metabolites**, v. 11, n. 2, p. 107, 12 fev. 2021.
- DOMINGUES VIEIRA, B. et al. Production of Epoxyketone Peptide-Based Proteasome Inhibitors by *Streptomyces* sp. BRA-346: Regulation and Biosynthesis. **Frontiers in Microbiology**, v. 13, 24 mar. 2022.
- DOROGHAZI, J. R. et al. A roadmap for natural product discovery based on large-scale genomics and metabolomics. **Nature Chemical Biology**, v. 10, n. 11, p. 963–968, 28 nov. 2014.
- ELSEBAI, M. F.; TEJESVI, M. V.; PIRTILÄ, A. M. Endophytes as a Novel Source of Bioactive New Structures. Em: **Advances in Endophytic Research**. New Delhi: Springer India, 2014. p. 191–202.

- FAO. **Aspirational principles and criteria for a sustainable bioeconomy**. Roma, Italy: [s.n.].
- FINN, R. D. et al. The Pfam protein families database: towards a more sustainable future. **Nucleic Acids Research**, v. 44, n. D1, p. D279–D285, 4 jan. 2016.
- FJÆRVIK, E.; ZOTCHEV, S. B. Biosynthesis of the polyene macrolide antibiotic nystatin in *Streptomyces noursei*. **Applied Microbiology and Biotechnology**, v. 67, n. 4, p. 436–443, 8 jun. 2005.
- FURTADO, L. C. et al. Marine *Streptomyces* sp. Isolated From the Brazilian Endemic Tunicate *Euherdmania* sp. Produces Dihydroeponemycin and Analogs With Potent Antiglioma Activity. **Frontiers in Marine Science**, v. 8, 3 maio 2021.
- GARRIDO-SANZ, D. et al. Comparative Genomics of the *Rhodococcus* Genus Shows Wide Distribution of Biodegradation Traits. **Microorganisms**, v. 8, n. 5, p. 774, 21 maio 2020.
- GILLOR, O.; ETZION, A.; RILEY, M. A. The dual role of bacteriocins as anti- and probiotics. **Applied Microbiology and Biotechnology**, v. 81, n. 4, p. 591–606, 1 dez. 2008.
- GOERING, A. W. et al. Metabologenomics: Correlation of Microbial Gene Clusters with Metabolites Drives Discovery of a Nonribosomal Peptide with an Unusual Amino Acid Monomer. **ACS Central Science**, v. 2, n. 2, p. 99–108, 24 fev. 2016.
- GOODFELLOW, M.; ALDERSON, G. The Actinomycete-genus *Rhodococcus*: A Home for the “rhodochrous” Complex. **Journal of General Microbiology**, v. 100, n. 1, p. 99–122, 1 maio 1977.
- GRIM, C. M.; LUU, G. T.; SANCHEZ, L. M. Staring into the void: demystifying microbial metabolomics. **FEMS Microbiology Letters**, v. 366, n. 11, 1 jun. 2019.
- GUAYASAMIN, J. M. et al. Chapter 2: Evolution of Amazonian biodiversity. Em: **Amazon Assessment Report 2021**. [s.l.] UN Sustainable Development Solutions Network (SDSN), 2021.
- GUO, J. et al. Raising the avermectins production in *Streptomyces avermitilis* by utilizing nanosecond pulsed electric fields (nsPEFs). **Scientific Reports**, v. 6, n. 1, p. 25949, 16 maio 2016.
- GUO, J. et al. Addressing big data challenges in mass spectrometry-based metabolomics. **Chemical Communications**, v. 58, n. 72, p. 9979–9990, 2022.
- GUO, Q.-Q. et al. *Rhodococcus agglutinans* sp. nov., an actinobacterium isolated from a soil sample. **Antonie van Leeuwenhoek**, v. 107, n. 5, p. 1271–1280, 10 maio 2015.
- GÜRTLER, V.; MAYALL, B. C.; SEVIOUR, R. Can whole genome analysis refine the taxonomy of the genus *Rhodococcus*? **FEMS Microbiology Reviews**, v. 28, n. 3, p. 377–403, jun. 2004.
- HABIB, S. et al. Production of Lipopeptide Biosurfactant by a Hydrocarbon-Degrading Antarctic *Rhodococcus*. **International Journal of Molecular Sciences**, v. 21, n. 17, p. 6138, 26 ago. 2020.
- HADAD, D.; GERESH, S.; SIVAN, A. Biodegradation of polyethylene by the thermophilic bacterium *Brevibacillus borstelensis*. **Journal of Applied Microbiology**, v. 98, n. 5, p. 1093–1100, maio 2005.
- HELFRICH, E. J. N. et al. Bipartite interactions, antibiotic production and biosynthetic potential of the *Arabidopsis* leaf microbiome. **Nature Microbiology**, v. 3, n. 8, p. 909–919, 1 ago. 2018.

- HELFRICH, E. J. N. et al. Bacterial terpene biosynthesis: challenges and opportunities for pathway engineering. **Beilstein Journal of Organic Chemistry**, v. 15, p. 2889–2906, 29 nov. 2019.
- HOBSON, C.; CHAN, A. N.; WRIGHT, G. D. The Antibiotic Resistome: A Guide for the Discovery of Natural Products as Antimicrobial Agents. **Chemical Reviews**, v. 121, n. 6, p. 3464–3494, 24 mar. 2021.
- HORINOUCI, S. A microbial hormone A-factor as a master switch for morphological differentiation and secondary metabolism in streptomyces griseus. **Frontiers in Bioscience**, v. 7, n. 4, p. A897, 2002.
- IACOVELLI, R.; BOVENBERG, R. A. L.; DRIESSEN, A. J. M. Nonribosomal peptide synthetases and their biotechnological potential in *Penicillium rubens*. **Journal of Industrial Microbiology and Biotechnology**, v. 48, n. 7–8, 24 ago. 2021.
- IMAI, Y. et al. A new antibiotic selectively kills Gram-negative pathogens. **Nature**, v. 576, n. 7787, p. 459–464, 19 dez. 2019.
- IWATSUKI, M. et al. Lariatins, Antimycobacterial Peptides Produced by *Rhodococcus* sp. K01–B0171, Have a Lasso Structure. **Journal of the American Chemical Society**, v. 128, n. 23, p. 7486–7491, 1 jun. 2006.
- IZORÉ, T. et al. Structures of a non-ribosomal peptide synthetase condensation domain suggest the basis of substrate selectivity. **Nature Communications**, v. 12, n. 1, p. 2511, 4 maio 2021.
- JAIN, R.; MAJUMDAR, D.; DEVI, S. Microbial Treatment of Raw and Primary Treated Sanitary Landfill Leachate by Indigenous Strain *Brevibacillus agri*. **Applied Biochemistry and Biotechnology**, v. 195, n. 4, p. 2317–2331, 18 abr. 2023.
- JOHNSON, E. T.; BOWMAN, M. J.; DUNLAP, C. A. *Brevibacillus fortis* NRS-1210 produces edeines that inhibit the in vitro growth of conidia and chlamydospores of the onion pathogen *Fusarium oxysporum* f. sp. *cepae*. **Antonie van Leeuwenhoek**, v. 113, n. 7, p. 973–987, 11 jul. 2020.
- JOHNSON, M. et al. NCBI BLAST: a better web interface. **Nucleic Acids Research**, v. 36, n. Web Server, p. W5–W9, 19 maio 2008.
- JOHNSON-ARBOR, K.; DUBEY, R. **Doxorubicin**. [s.l: s.n.].
- JONES, A. L.; SUTCLIFFE, I. C.; GOODFELLOW, M. Proposal to replace the illegitimate genus name *Prescottia* Jones et al. 2013 with the genus name *Prescottella* gen. nov. and to replace the illegitimate combination *Prescottia equi* Jones et al. 2013 with *Prescottella equi* comb. nov. **Antonie van Leeuwenhoek**, v. 103, n. 6, p. 1405–1407, 25 jun. 2013.
- JUNJUA, M. et al. From Lab to Pharmacy Shelves: The Story of a Plant Derived Anticancer Drug, “Paclitaxel”. **Journal of Bioresource Management**, v. 2, n. 4, 22 dez. 2015.
- KANEHISA, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. **Nucleic Acids Research**, v. 28, n. 1, p. 27–30, 1 jan. 2000.
- KAUTSAR, S. A. et al. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. **Nucleic Acids Research**, 15 out. 2019.

- KEATINGE-CLAY, A. T. The structures of type I polyketide synthases. **Natural Product Reports**, v. 29, n. 10, p. 1050, 2012.
- KIM, J.-N. et al. Comparative Genomics Reveals the Core and Accessory Genomes of *Streptomyces* Species. **Journal of Microbiology and Biotechnology**, v. 25, n. 10, p. 1599–1605, 28 out. 2015.
- KIM, Y. H. et al. Production of Rapamycin in *Streptomyces hygroscopicus* from Glycerol-Based Media Optimized by Systemic Methodology. **Journal of Microbiology and Biotechnology**, v. 24, n. 10, p. 1319–1326, 28 out. 2014.
- KINO, T. et al. FK-506, a novel immunosuppressant isolated from a *Streptomyces*. I. Fermentation, isolation, and physico-chemical and biological characteristics. **The Journal of Antibiotics**, v. 40, n. 9, p. 1249–1255, 1987.
- KITAGAWA, W. et al. Identification of a novel bacteriocin-like protein and structural gene from *Rhodococcus erythropolis* JCM 2895, using suppression-subtractive hybridization. **The Journal of Antibiotics**, v. 71, n. 10, p. 872–879, 6 out. 2018.
- KITAGAWA, W.; TAMURA, T. A Quinoline Antibiotic from *Rhodococcus erythropolis* JCM 6824. **The Journal of Antibiotics**, v. 61, n. 11, p. 680–682, nov. 2008.
- KITANI, S. et al. Avenolide, a *Streptomyces* hormone controlling antibiotic production in *Streptomyces avermitilis*. **Proceedings of the National Academy of Sciences**, v. 108, n. 39, p. 16410–16415, 27 set. 2011.
- KOLMOGOROV, M. et al. metaFlye: scalable long-read metagenome assembly using repeat graphs. **Nature Methods**, v. 17, n. 11, p. 1103–1110, 5 nov. 2020.
- KRZYWINSKI, M. et al. Circos: An information aesthetic for comparative genomics. **Genome Research**, v. 19, n. 9, p. 1639–1645, set. 2009.
- KUNZE, B.; HÖFLE, G.; REICHENBACH, H. The aurachins, new quinoline antibiotics from myxobacteria : Production, physico-chemical and biological properties. **The Journal of Antibiotics**, v. 40, n. 3, p. 258–265, 1987.
- LEE, H.; VAN DER DONK, W. A. Macrocyclization and Backbone Modification in RiPP Biosynthesis. **Annual Review of Biochemistry**, v. 91, n. 1, p. 269–294, 21 jun. 2022.
- LI, C. et al. Advances in the Biosynthesis of Terpenoids and Their Ecological Functions in Plant Resistance. **International Journal of Molecular Sciences**, v. 24, n. 14, p. 11561, 17 jul. 2023.
- LI, H.; DING, W.; ZHANG, Q. Discovery and engineering of ribosomally synthesized and post-translationally modified peptide (RiPP) natural products. **RSC Chemical Biology**, v. 5, n. 2, p. 90–108, 2024.
- LOLL, P. J. et al. The high resolution structure of tyrocidine A reveals an amphipathic dimer. **Biochimica et Biophysica Acta (BBA) - Biomembranes**, v. 1838, n. 5, p. 1199–1207, maio 2014.
- MALHI, Y. et al. Climate Change, Deforestation, and the Fate of the Amazon. **Science**, v. 319, n. 5860, p. 169–172, 11 jan. 2008.

- MARCO, D. E.; ABRAM, F. Editorial: Using Genomics, Metagenomics and Other “Omics” to Assess Valuable Microbial Ecosystem Services and Novel Biotechnological Applications. **Frontiers in Microbiology**, v. 10, 12 fev. 2019.
- MARTENS, E.; DEMAÏN, A. L. Platensimycin and platencin: promising antibiotics for future application in human medicine. **The Journal of Antibiotics**, v. 64, n. 11, p. 705–710, 14 nov. 2011.
- MARTÍN, J. F.; CASQUEIRO, J.; LIRAS, P. Secretion systems for secondary metabolites: how producer cells send out messages of intercellular communication. **Current Opinion in Microbiology**, v. 8, n. 3, p. 282–293, jun. 2005.
- MARTÍNEZ-NÚÑEZ, M. A.; LÓPEZ, V. E. L. Y. Nonribosomal peptides synthetases and their applications in industry. **Sustainable Chemical Processes**, v. 4, n. 1, p. 13, 2 dez. 2016.
- MCCHESNEY, J. D.; VENKATARAMAN, S. K.; HENRI, J. T. Plant natural products: Back to the future or into extinction? **Phytochemistry**, v. 68, n. 14, p. 2015–2022, jul. 2007.
- MEDEMA, M. H. et al. Minimum Information about a Biosynthetic Gene cluster. **Nature Chemical Biology**, v. 11, n. 9, p. 625–631, 18 set. 2015.
- MIRANDA-CASOLUENGO, R. et al. The Intracellular Pathogen *Rhodococcus equi* Produces a Catecholate Siderophore Required for Saprophytic Growth. **Journal of Bacteriology**, v. 190, n. 5, p. 1631–1637, mar. 2008.
- MIRANDA-CASOLUENGO, R. et al. The Hydroxamate Siderophore Rhequichelin Is Required for Virulence of the Pathogenic Actinomycete *Rhodococcus equi*. **Infection and Immunity**, v. 80, n. 12, p. 4106–4114, dez. 2012.
- MOOTZ, H. D.; SCHWARZER, D.; MARAHIEL, M. A. Ways of Assembling Complex Natural Products on Modular Nonribosomal Peptide Synthetases A list of abbreviations can be found at the end of the text. **ChemBioChem**, v. 3, n. 6, p. 490, 3 jun. 2002.
- MUKHERJEE, P. K. LC–MS: A Rapid Technique for Understanding the Plant Metabolite Analysis. Em: **Quality Control and Evaluation of Herbal Drugs**. [s.l.] Elsevier, 2019. p. 459–479.
- NETT, M.; IKEDA, H.; MOORE, B. S. Genomic basis for natural product biosynthetic diversity in the actinomycetes. **Natural Product Reports**, v. 26, n. 11, p. 1362, 2009.
- O'BRIEN, J.; WRIGHT, G. D. An ecological perspective of microbial secondary metabolism. **Current Opinion in Biotechnology**, v. 22, n. 4, p. 552–558, ago. 2011.
- OLIVEIRA, D. F. F. DE et al. Antimicrobial potential of Copaiba Oil (*Copaifera multijuga* Hayne-Leguminosae) against bubaline mastitis multiresistant isolates. **Anais da Academia Brasileira de Ciências**, v. 92, n. 4, 2020.
- ORRO, A. et al. Genome and Phenotype Microarray Analyses of *Rhodococcus* sp. BCP1 and *Rhodococcus opacus* R7: Genetic Determinants and Metabolic Abilities with Environmental Relevance. **PLOS ONE**, v. 10, n. 10, p. e0139467, 1 out. 2015.
- ORTEGA, M. A.; VAN DER DONK, W. A. New Insights into the Biosynthetic Logic of Ribosomally Synthesized and Post-translationally Modified Peptide Natural Products. **Cell Chemical Biology**, v. 23, n. 1, p. 31–44, jan. 2016.

- PARTE, A. C. et al. List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. **International Journal of Systematic and Evolutionary Microbiology**, v. 70, n. 11, p. 5607–5612, 1 nov. 2020.
- PARVEZ, N.; CORNELIUS, L. K.; FADER, R. *Brevibacillus brevis* Peritonitis. **The American Journal of the Medical Sciences**, v. 337, n. 4, p. 297–299, abr. 2009.
- PATEL, S.; PREUSS, C.; BERNICE, F. **Vancomycin**. Disponível em: <<https://www.ncbi.nlm.nih.gov/books/NBK459263/>>. Acesso em: 15 nov. 2023.
- PEREIRA, J. O. et al. Overview on Biodiversity, Chemistry, and Biotechnological Potential of Microorganisms from the Brazilian Amazon. Em: **Diversity and Benefits of Microorganisms from the Tropics**. Cham: Springer International Publishing, 2017. p. 71–103.
- PETTIT, R. K. Small-molecule elicitation of microbial secondary metabolites. **Microbial Biotechnology**, v. 4, n. 4, p. 471–478, 6 jul. 2011.
- PIEL, J. Biosynthesis of polyketides by trans-AT polyketide synthases. **Natural Product Reports**, v. 27, n. 7, p. 996, 2010.
- RAUTENBACH, M. et al. Inhibition of malaria parasite blood stages by tyrocidines, membrane-active cyclic peptide antibiotics from *Bacillus brevis*. **Biochimica et Biophysica Acta (BBA) - Biomembranes**, v. 1768, n. 6, p. 1488–1497, jun. 2007.
- REDDY, B. V. B. et al. Natural Product Biosynthetic Gene Diversity in Geographically Distinct Soil Microbiomes. **Applied and Environmental Microbiology**, v. 78, n. 10, p. 3744–3752, 15 maio 2012.
- REDDY, M. S. et al. Biodegradation of phenanthrene with biosurfactant production by a new strain of *Brevibacillus* sp. **Bioresource Technology**, v. 101, n. 20, p. 7980–7983, out. 2010.
- RIDLEY, C. P.; LEE, H. Y.; KHOSLA, C. Evolution of polyketide synthases in bacteria. **Proceedings of the National Academy of Sciences**, v. 105, n. 12, p. 4595–4600, 25 mar. 2008.
- RUDOLF, J. D. et al. Bacterial terpenome. **Natural Product Reports**, v. 38, n. 5, p. 905–980, 2021.
- RUDOLF, J. D.; YAN, X.; SHEN, B. Genome neighborhood network reveals insights into enediyne biosynthesis and facilitates prediction and prioritization for discovery. **Journal of Industrial Microbiology and Biotechnology**, v. 43, n. 2–3, p. 261–276, 1 mar. 2016.
- RUIU, L. *Brevibacillus laterosporus*, a Pathogen of Invertebrates and a Broad-Spectrum Antimicrobial Species. **Insects**, v. 4, n. 3, p. 476–92, 5 set. 2013.
- SCHERLACH, K.; HERTWECK, C. Mining and unearthing hidden biosynthetic potential. **Nature Communications**, v. 12, n. 1, p. 3864, 23 jun. 2021.
- SCHMID, A. et al. Industrial biocatalysis today and tomorrow. **Nature**, v. 409, n. 6817, p. 258–268, 11 jan. 2001.
- SCHNEIDER, Y. K. Bacterial Natural Product Drug Discovery for New Antibiotics: Strategies for Tackling the Problem of Antibiotic Resistance by Efficient Bioprospecting. **Antibiotics**, v. 10, n. 7, p. 842, 10 jul. 2021.
- SEEMANN, T. Prokka: rapid prokaryotic genome annotation. **Bioinformatics (Oxford, England)**, v. 30, n. 14, p. 2068–9, 15 jul. 2014.

- SHANNON, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. **Genome research**, v. 13, n. 11, p. 2498–504, nov. 2003.
- SHE, Y. et al. Genome sequence of *Brevibacillus agri* strain 5-2, isolated from the formation water of petroleum reservoir. **Marine Genomics**, v. 18, p. 123–125, dez. 2014.
- SHIDA, O. et al. Proposal for Two New Genera, *Brevibacillus* gen. nov. and *Aneurinibacillus* gen. nov. **International Journal of Systematic Bacteriology**, v. 46, n. 4, p. 939–946, 1 out. 1996.
- SHIMIZU, Y.; OGATA, H.; GOTO, S. Type III Polyketide Synthases: Functional Classification and Phylogenomics. **ChemBioChem**, v. 18, n. 1, p. 50–65, 3 jan. 2017.
- SIMÃO, F. A. et al. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. **Bioinformatics**, v. 31, n. 19, p. 3210–3212, 1 out. 2015.
- SKIRYCH, A. et al. Medicinal Bioprospecting of the Amazon Rainforest: A Modern Eldorado? **Trends in Biotechnology**, v. 34, n. 10, p. 781–790, out. 2016.
- SOTTORFF, I. et al. Different Secondary Metabolite Profiles of Phylogenetically almost Identical *Streptomyces griseus* Strains Originating from Geographically Remote Locations. **Microorganisms**, v. 7, n. 6, p. 166, 6 jun. 2019.
- SPOHN, M.; WOHLLEBEN, W.; STEGMANN, E. Elucidation of the zinc-dependent regulation in *Amycolatopsis japonicum* enabled the identification of the ethylenediamine-disuccinate ([*S,S*]-EDDS) genes. **Environmental Microbiology**, v. 18, n. 4, p. 1249–1263, 18 abr. 2016.
- SÜSSMUTH, R. D.; MAINZ, A. Nonribosomal Peptide Synthesis—Principles and Prospects. **Angewandte Chemie International Edition**, v. 56, n. 14, p. 3770–3821, 27 mar. 2017.
- TANG, X. et al. Identification of Thiotetronic Acid Antibiotic Biosynthetic Pathways by Target-directed Genome Mining. **ACS Chemical Biology**, v. 10, n. 12, p. 2841–2849, 18 dez. 2015.
- TANG, X.-J.; THIBAUT, P.; BOYD, R. K. Characterisation of the tyrocidine and gramicidin fractions of the tyrothricin complex from *Bacillus brevis* using liquid chromatography and mass spectrometry. **International Journal of Mass Spectrometry and Ion Processes**, v. 122, p. 153–179, dez. 1992.
- TATUSOV, R. L. The COG database: a tool for genome-scale analysis of protein functions and evolution. **Nucleic Acids Research**, v. 28, n. 1, p. 33–36, 1 jan. 2000.
- THAKER, M. N. et al. Identifying producers of antibacterial compounds by screening for antibiotic resistance. **Nature Biotechnology**, v. 31, n. 10, p. 922–927, 22 out. 2013.
- THE UNIPROT CONSORTIUM. UniProt: the universal protein knowledgebase. **Nucleic Acids Research**, v. 45, n. D1, p. D158–D169, 4 jan. 2017.
- THOMPSON, D. et al. Phylogenomic Classification and Biosynthetic Potential of the Fossil Fuel-Biodesulfurizing *Rhodococcus* Strain IGTS8. **Frontiers in Microbiology**, v. 11, 7 jul. 2020.
- TRAVIN, D. Y. et al. Structure of ribosome-bound azole-modified peptide phazolicin rationalizes its species-specific mode of bacterial translation inhibition. **Nature Communications**, v. 10, n. 1, p. 4563, 8 out. 2019.

- TRAVIN, D. Y. et al. Dual-Uptake Mode of the Antibiotic Phazolicin Prevents Resistance Acquisition by Gram-Negative Bacteria. **mBio**, v. 14, n. 2, 25 abr. 2023.
- TROSKIE, A. M. et al. Synergistic Activity of the Tyrocidines, Antimicrobial Cyclodecapeptides from *Bacillus aneurinolyticus*, with Amphotericin B and Caspofungin against *Candida albicans* Biofilms. **Antimicrobial Agents and Chemotherapy**, v. 58, n. 7, p. 3697–3707, jul. 2014.
- VAN DER HOOFT, J. J. J. et al. Linking genomics and metabolomics to chart specialized metabolic diversity. **Chemical Society Reviews**, v. 49, n. 11, p. 3297–3314, 2020.
- VAN DER MEIJ, A. et al. Chemical ecology of antibiotic production by actinomycetes. **FEMS Microbiology Reviews**, v. 41, n. 3, p. 392–416, 1 maio 2017.
- VANDERGEIZE, R.; DIJKHUIZEN, L. Harnessing the catabolic diversity of rhodococci for environmental and biotechnological applications. **Current Opinion in Microbiology**, v. 7, n. 3, p. 255–261, jun. 2004.
- VITALE, G. A. et al. A metabologenomics approach to unlock the metabolome of the novel Antarctic deep-sea isolate *Lacinutrix shetlandiensis* sp. nov. WUR7. **PNAS Nexus**, v. 2, n. 7, 3 jul. 2023.
- VOLENZO, T.; ODIYO, J. Integrating endemic medicinal plants into the global value chains: the ecological degradation challenges and opportunities. **Heliyon**, v. 6, n. 9, p. e04970, set. 2020.
- WANG, C. et al. Screening and Whole-Genome Sequencing of Two *Streptomyces* Species from the Rhizosphere Soil of Peony Reveal Their Characteristics as Plant Growth-Promoting Rhizobacteria. **BioMed Research International**, v. 2018, p. 1–11, 5 set. 2018.
- WANG, F. et al. Solubilized Gramicidin A as Potential Systemic Antibiotics. **ChemBioChem**, v. 13, n. 1, p. 51–55, 2 jan. 2012.
- WANG, H. et al. Atlas of nonribosomal peptide and polyketide biosynthetic pathways reveals common occurrence of nonmodular enzymes. **Proceedings of the National Academy of Sciences**, v. 111, n. 25, p. 9259–9264, 24 jun. 2014a.
- WANG, J. et al. Biosynthesis of aromatic polyketides in microorganisms using type II polyketide synthases. **Microbial Cell Factories**, v. 19, n. 1, p. 110, 24 dez. 2020.
- WANG, L. et al. Sesterterpenoids. **Natural Product Reports**, v. 30, n. 3, p. 455, 2013.
- WANG, M. et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. **Nature Biotechnology**, v. 34, n. 8, p. 828–837, 9 ago. 2016a.
- WANG, M. et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. **Nature Biotechnology**, v. 34, n. 8, p. 828–837, 9 ago. 2016b.
- WANG, Q.; QUAN, S.; XIAO, H. Towards efficient terpenoid biosynthesis: manipulating IPP and DMAPP supply. **Bioresources and Bioprocessing**, v. 6, n. 1, p. 6, 6 dez. 2019.
- WANG, X. et al. Draft Genome Sequence of *Brevibacillus panacihumi* Strain W25, a Halotolerant Hydrocarbon-Degrading Bacterium. **Genome Announcements**, v. 2, n. 1, 27 fev. 2014b.

WILKINS, K.; SCHÖLLER, C. Volatile Organic Metabolites from Selected Streptomyces Strains. **Actinomycetologica**, v. 23, n. 2, p. 27–33, 2009.

YAMADA, Y. et al. Terpene synthases are widely distributed in bacteria. **Proceedings of the National Academy of Sciences**, v. 112, n. 3, p. 857–862, 20 jan. 2015.

YANG, A. et al. **Rethinking the Brazilian Amazon Sustainable development for a thriving future**. [s.l: s.n.].

YANG, X.; YOUSEF, A. E. **Antimicrobial peptides produced by Brevibacillus spp.: structure, classification and bioactivity: a mini review**. **World Journal of Microbiology and Biotechnology** Springer Netherlands, , 1 abr. 2018.

ZHANG, M. M. et al. CRISPR–Cas9 strategy for activation of silent Streptomyces biosynthetic gene clusters. **Nature Chemical Biology**, v. 13, n. 6, p. 607–609, 10 jun. 2017.

ZHOU, Z. et al. Genome plasticity and systems evolution in Streptomyces. **BMC Bioinformatics**, v. 13, n. S10, p. S8, 25 jun. 2012.

ZIEMERT, N. et al. The Natural Product Domain Seeker NaPDoS: A Phylogeny Based Bioinformatic Tool to Classify Secondary Metabolite Gene Diversity. **PLoS ONE**, v. 7, n. 3, p. e34064, 29 mar. 2012.

ZIEMERT, N.; ALANJARY, M.; WEBER, T. The evolution of genome mining in microbes – a review. **Natural Product Reports**, v. 33, n. 8, p. 988–1005, 2016.