



UNIVERSIDADE FEDERAL DO AMAZONAS
PROGRAMA MULTIINSTITUCIONAL DE PÓS-GRADUAÇÃO EM
BIOTECNOLOGIA

Diversidade biossintética de metabólitos secundários de fungos dos gêneros
Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis

ÉDY CARLOS CAVALCANTE DA SILVA

MANAUS
2025

ÉDY CARLOS CAVALCANTE DA SILVA

**Diversidade biossintética de metabólitos secundários de fungos dos gêneros
Pestalotiopsis, *Pseudopestalotiopsis* e *Neopestalotiopsis***

ORIENTADOR: GILVAN FERREIRA DA SILVA

Dissertação de Mestrado apresentado ao Programa de Pós-Graduação Multi-institucional em Biotecnologia da Universidade Federal do Amazonas, como parte dos requisitos para obtenção de título de Mestre em Biotecnologia.

MANAUS

2025

Ficha Catalográfica

Elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

- S586d Silva, Édy Carlos Cavalcante da
Diversidade biossintética de metabólitos secundários de fungos
dos gêneros Pestalotiopsis, Pseudopestalotiopsis e
Neopestalotiopsis / Édy Carlos Cavalcante da Silva. - 2025.
63 f. : il., color. ; 31 cm.
- Orientador(a): Gilvan Ferreira da Silva.
Dissertação (mestrado) - Universidade Federal do Amazonas,
Programa de Pós-Graduação em Biotecnologia, Manaus , 2025.
1. Metabólitos secundários. 2. Biosynthetic Gene Clusters. 3.
Pestalotiopsis. 4. Neopestalotiopsis. 5. Pseudopestalotiopsis. I.
Silva, Gilvan Ferreira da. II. Universidade Federal do Amazonas.
Programa de Pós-Graduação em Biotecnologia. III. Título
-

ÉDY CARLOS CAVALCANTE DA SILVA

**Diversidade biossintética de metabólitos secundários de fungos dos gêneros
Pestalotiopsis, *Pseudopestalotiopsis* e *Neopestalotiopsis***

Dissertação de Mestrado apresentado ao Programa de Pós-Graduação Multi-institucional em Biotecnologia da Universidade Federal do Amazonas, como parte dos requisitos para obtenção de título de Mestre em Biotecnologia.

Aprovada em 19/12/2025

BANCA EXAMINADORA

Prof.(a) Dr.(a) Gilvan Ferreira da Silva (Presidente)
Embrapa Amazônia Ocidental

Prof.(a) Dr.(a) Thiago Fernandes Sousa (Membro)
Northeastern University

Prof.(a) Dr.(a) Ana Francisca Tibúrcia Amorim Ferreira e Ferreira (Membro)
Universidade Federal do Amazonas (UFAM)

AGRADECIMENTOS

À Universidade Federal do Amazonas (UFAM) e ao Programa Multiinstitucional de Pós-Graduação em Biotecnologia, pela oportunidade de realizar este mestrado. Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão da bolsa de mestrado, fundamental para a dedicação integral a este trabalho e para a minha formação como pesquisador.

Ao Prof. **Dr. Gilvan Ferreira da Silva**, meu orientador, pela confiança depositada, pelos ensinamentos valiosos, pela paciência e pelo constante incentivo ao longo desta jornada. Sua orientação foi essencial para o amadurecimento científico e pessoal que alcancei durante o mestrado. À Embrapa Amazônia Ocidental, em especial ao Projeto AmazonMicro, por disponibilizar as linhagens amazônicas que foram fundamentais para o desenvolvimento deste trabalho, ampliando o conhecimento sobre a biodiversidade fúngica da região. Aos professores e pesquisadores do programa de pós-graduação, pelas disciplinas ministradas, discussões científicas e contribuições que enriqueceram minha formação acadêmica. Aos colegas do laboratório e do programa, pela convivência, pelas trocas de experiências, pelos momentos de colaboração e pelo companheirismo que tornaram essa jornada mais leve e produtiva. À minha esposa Alice e à minha filha Maria Alice, pelo amor incondicional, pela compreensão nos momentos de ausência, pelo apoio inabalável e por serem minha fonte de inspiração e força. Vocês são a razão de cada conquista. À minha família, pelo incentivo constante, pela torcida e por sempre acreditarem em mim, mesmo nos momentos mais desafiadores. À memória do meu irmão Francisco de Assis (in memoriam), que partiu em maio de 2020 durante a pandemia. Sua lembrança permanece viva em meu coração e foi fonte de motivação para seguir em frente e honrar sua memória com esta conquista.

A todos que, direta ou indiretamente, contribuíram para a realização deste trabalho, meu sincero agradecimento.

"O temor do Senhor é o princípio da sabedoria..."

(Provérbios 9:10).

"Cada sequência de DNA é um programa rodando há milhões de anos; nós usamos Python para depurar e entender sua lógica"

Autor Desconhecido

RESUMO

Os fungos dos gêneros *Pestalotiopsis*, *Neopestalotiopsis* e *Pseudopestalotiopsis* destacam-se pela capacidade de produzir metabólitos secundários estruturalmente complexos com atividades antiviral, antimicrobiana e antitumoral, sendo reconhecidos como microrganismos de interesse agrícola e biotecnológico. Este estudo teve como objetivo explorar, por meio de mineração genômica, a diversidade de clusters de genes biossintéticos (BGCs) e famílias de clusters (GCFs) em 23 genomas pertencentes aos três gêneros (9 de *Pestalotiopsis*, 10 de *Neopestalotiopsis* e 4 de *Pseudopestalotiopsis*), incluindo 17 linhagens obtidas de bases públicas (NCBI e JGI) e 6 linhagens amazônicas patogênicas da coleção da Embrapa Amazônia Ocidental. A identificação e anotação de BGCs foram realizadas utilizando antiSMASH versão 8.0, seguida de análise de redes de similaridade de sequências e agrupamento em famílias com BiG-SCAPE, possibilitando a construção de redes moleculares e análises de sintenia. Foram identificados 2.113 BGCs distribuídos em sete classes principais: NRPS (795 BGCs), PKS (607 BGCs), Terpene (282 BGCs), Others (596 BGCs), PKS-other (162 BGCs), PKS-NRP Hybrids (138 BGCs) e RiPPs (36 BGCs), evidenciando predominância de vias biossintéticas relacionadas a peptídeos não-ribossômicos e policetídeos tipo I. A análise por gênero revelou que *Neopestalotiopsis* apresentou o maior número de BGCs (989), com composição equilibrada entre PKS (27,40%), NRPS (25,48%) e Terpenos (12,74%), indicando maior versatilidade metabólica, enquanto *Pestalotiopsis* exibiu 666 BGCs com predominância de NRPS (28,98%) e PKS (28,68%), e *Pseudopestalotiopsis* totalizou 458 BGCs, destacando-se pela categoria Others (29,26%). O agrupamento em redes de similaridade resultou na identificação de 632 GCFs, sendo 165 de NRPS, 141 de PKS, 168 de Others, 63 de Terpene, 47 de PKS-other, 36 de PKS-NRP Hybrids e 12 de RiPPs. A análise de compartilhamento revelou predominância de GCFs exclusivas de cada gênero (unique) e singletons, com apenas 9 GCFs classificadas como core (compartilhadas pelos três gêneros), evidenciando elevada especialização metabólica e baixa conservação de vias biossintéticas. A comparação com o banco de dados MIBiG identificou apenas 4 GCFs (0,63%) com correspondência a clusters caracterizados experimentalmente, envolvendo 19 BGCs (0,9% do total), incluindo famílias relacionadas à produção de policetídeos tipo I (T1PKS) em *Nodulisporium sp.*, *Glarea lozoyensis* e *Daldinia eschscholzii*, ácido pesteico em *Pestalotiopsis fici*, derivados de indol e cloropestolides em *Pestalotiopsis fici*, e flavunoidina em *Aspergillus flavus*, indicando que mais de 99% das GCFs constituem clusters órfãos com potencial inexplorado para descoberta de novos compostos bioativos. A abordagem computacional integrada empregada neste estudo reforça a relevância da bioinformática como ferramenta para a prospecção de metabólitos secundários em larga escala e para o entendimento da evolução funcional de BGCs em fungos filamentosos, consolidando o potencial dos fungos *pestalóides*, especialmente linhagens amazônicas de um bioma de megadiversidade, como fontes promissoras de produtos naturais com aplicações biotecnológicas e farmacológicas.

Palavras-chave: Metabólitos secundários, *Biosynthetic Gene Clusters*, antiSMASH, BiG-SCAPE, Biopython, *Pestalotiopsis*, *Neopestalotiopsis*, *Pseudopestalotiopsis*.

ABSTRACT

Fungi from the genera *Pestalotiopsis*, *Neopestalotiopsis*, and *Pseudopestalotiopsis* are recognized for their ability to produce structurally complex secondary metabolites with antiviral, antimicrobial, and antitumoral activities, making them of significant agricultural and biotechnological interest. This study aimed to explore, through genome mining, the diversity of biosynthetic gene clusters (BGCs) and gene cluster families (GCFs) in 23 genomes from these three genera (9 *Pestalotiopsis*, 10 *Neopestalotiopsis*, and 4 *Pseudopestalotiopsis*), including 17 strains from public databases (NCBI and JGI) and 6 pathogenic Amazonian strains from the Embrapa Western Amazon collection. BGC identification and annotation were performed using antiSMASH version 8.0, followed by sequence similarity network analysis and clustering into families using BiG-SCAPE, enabling molecular network construction and synteny analysis. A total of 2,113 BGCs were identified and classified into seven major categories: NRPS (795 BGCs), PKS (607 BGCs), terpenes (282 BGCs), others (596 BGCs), PKSother (162 BGCs), PKS-NRP hybrids (138 BGCs), and RiPPs (36 BGCs), highlighting the predominance of nonribosomal peptide and type I polyketide biosynthetic pathways. Genus-level analysis revealed that *Neopestalotiopsis* exhibited the highest number of BGCs (989), with a balanced composition of PKS (27.40%), NRPS (25.48%), and terpenes (12.74%), suggesting greater metabolic versatility. *Pestalotiopsis* showed 666 BGCs, mainly NRPS (28.98%) and PKS (28.68%), while *Pseudopestalotiopsis* presented 458 BGCs, with a predominance of the “others” category (29.26%). Clustering analysis identified 632 GCFs, including 165 NRPS, 141 PKS, 168 others, 63 terpenes, 47 PKSother, 36 PKS-NRP hybrids, and 12 RiPPs. Most GCFs were genus-specific (unique) or singletons, with only 9 core GCFs shared among the three genera, indicating high metabolic specialization and low conservation of biosynthetic pathways. Comparison with the MIBiG database revealed that only 4 GCFs (0.63%) matched experimentally characterized clusters, encompassing 19 BGCs (0.9% of the total), including type I polyketides, pestheic acid, indole derivatives, chloropestolides, and flavunoidin. These results indicate that more than 99% of GCFs correspond to orphan clusters with unexplored biosynthetic potential. The integrated computational approach employed in this study highlights the relevance of bioinformatics for large-scale secondary metabolite prospecting and for understanding the functional evolution of BGCs in filamentous fungi, reinforcing the potential of pestaloid fungi, particularly Amazonian strains from a megadiverse biome, as promising sources of novel bioactive compounds for biotechnological and pharmaceutical applications.

Keywords: Secondary metabolites; Biosynthetic gene clusters; antiSMASH; BiG-SCAPE; Biopython; *Pestalotiopsis*; *Neopestalotiopsis*; *Pseudopestalotiopsis*.

LISTA DE ABREVIATURAS

antiSMASH	<i>Antibiotics & Secondary Metabolite Analysis SHell</i>
BGC	<i>Biosynthetic Gene Cluster</i> (Cluster de Genes Biossintéticos)
BiG-SCAPE	<i>Biosynthetic Gene Similarity Clustering and Prospecting Engine</i>
BiG-SLiCE	<i>Biosynthetic Gene Similarity Clustering integrative Clustering Engine</i>
BLAST	<i>Basic Local Alignment Search Tool</i>
BLASTn	<i>Basic Local Alignment Search Tool for nucleotides</i>
BUSCO	<i>Benchmarking Universal Single-Copy Orthologs</i>
CORASON	<i>CORe Analysis of Syntenic Orthologues</i>
CPAA	Centro de Pesquisa Agroflorestal da Amazônia
GBK	GenBank (formato de arquivo)
GCF	<i>Gene Cluster Family</i> (Família de Clusters de Genes)
GCFs	<i>Gene Cluster Families</i> (Famílias de Clusters de Genes)
IC₅₀	Concentração Inibitória de 50%
iTOL	<i>Interactive Tree Of Life</i>
ITS	<i>Internal Transcribed Spacer</i> (Espaçador Interno Transcrito)
JGI	<i>Joint Genome Institute</i>
KR	Cetorreductase (<i>Ketoreductase</i>)
LSU	<i>Large Subunit</i> (Subunidade Maior do rRNA)
MAFFT	<i>Multiple Alignment using Fast Fourier Transform</i>
Mb	Megabase
MIBiG	<i>Minimum Information about a Biosynthetic Gene cluster</i>

ML	<i>Maximum Likelihood</i> (Máxima Verossimilhança)
MYST	<i>MOZ, Ybf2/Sas3, Sas2, and Tip60</i> (família de acetiltransferases de histonas)
NCBI	<i>National Center for Biotechnology Information</i>
NR-PKS	<i>Non-Reducing Polyketide Synthase</i>
NRPS	<i>Non-Ribosomal Peptide Synthetase</i> (Sintetase de Peptídeos Não-Ribossomais)
P450	Citocromo P450 monooxygenase
PKS	<i>Polyketide Synthase</i> (Policetídeo Sintase)
PKSI / PKS-I	<i>Polyketide Synthase type I</i> (Policetídeo Sintase tipo I)
PKSother	Outras Policetídeo Sintases
PKS-NRP	Híbrido de <i>Polyketide Synthase</i> e <i>Non-Ribosomal Peptide</i>
PN / PNs	Produto Natural / Produtos Naturais (<i>Natural Product(s)</i>)
PR-PKS	<i>Partially Reducing Polyketide Synthase</i>
RiPPs	<i>Ribosomally synthesized and Post-translationally modified Peptides</i> (Peptídeos Sintetizados Ribossomalmente e Modificados Pós-Traducionalmente)
T1PKS	<i>Type I Polyketide Synthase</i> (Policetídeo Sintase Tipo I)
TC / TCs	<i>Terpene Cyclase</i> (Ciclase de Terpenos)
TEF-1α	<i>Translation Elongation Factor 1-alpha</i> (Fator de Alongamento de Tradução 1-alfa)

SUMÁRIO

1.INTRODUÇÃO	12
2. REVISÃO DA LITERATURA.....	16
2.1 Metabolismo secundário de <i>Pestalotiopsis</i> , <i>Neopestalotiopsis</i> e <i>Pseudopestalotiopsis</i>	16
2.2- Diversidade Biossintética de BGC em larga escala	18
2.2.1-Mineração genômica	18
2.2.2- Biosynthetic Gene Similarity Clustering and Prospecting Engine (BiG-SCAPE) e CORE Analysis of Syntenic Orthologues (CORASON).....	19
3.OBJETIVO	24
3.1Objetivo Geral:	24
3.2Objetivo Específicos:	24
4.MATERIAL E MÉTODOS	25
4.1 Genomas de <i>Pestalotiopsis</i> , <i>Pseudopestalotiopsis</i> e <i>Neopestalotiopsis</i> utilizados neste trabalho	25
4.2 Análise filogenética e identificação filogenômica de linhagens de <i>Pestalotiopsis</i> , <i>Pseudopestalotiopsis</i> e <i>Neopestalotiopsis</i>	30
4.3 Análise com as ferramentas antSMASH e BIG-SCAPE	30
4.5 Análise utilizando bibliotecas PYTHON	31
5. RESULTADOS E DISCUSSÃO	32
5.1 Identificação filogenética e filogenômica de linhagens de <i>Pestalotiopsis</i> , <i>Pseudopestalotiopsis</i> e <i>Neopestalotiopsis</i> com genomas disponíveis	32
5.2. Distribuição de BGCs por gênero	35
5.3. Famílias de Clusters de Genes (GCFs) presentes em <i>Pestalotiopsis</i> , <i>Neopestalotiopsis</i> e <i>Pseudopestalotiopsis</i>	39
5.3.1 GCFs relacionadas à biossíntese de peptídeos não ribossômicos (NRP)	39
5.3.2 GCFs PKS I	40
5.3.3 GCFs Terpeno	41
5.3.4 Panorama quantitativo das famílias de clusters biossintéticos identificadas....	42
6. CONCLUSÃO	53
7. Perspectivas Futuras.....	55
7. CRONOGRAMA DE ATIVIDADES	57
REFERÊNCIAS	58

1.INTRODUÇÃO

Os fungos desempenham um papel significativo nos ecossistemas terrestres, desde simbióticos a patogênicos, e ainda podem produzir uma miríade de metabólitos secundários como é o caso dos fungos conhecidos atualmente como "semelhantes a *Pestalotiopsis*" (WADHWA et al., 2024; ELKHATEEB; DABA, 2021).

As espécies do grupo pestalóide compartilham entre si alguns aspectos morfológicos que podem ser utilizados para diferenciá-las, principalmente entre os gênero. Até meados de 2014, todas as espécies deste grupo estavam posicionadas taxonomicamente no gênero *Pestalotiopsis*. Contudo, Maharachchikumbura e colaboradores (2014), ao combinar os dados morfológicos e os de análise filogenética (ITS, LSU, β -*tubulin* e TEF-1 α), propuseram a segregação do gênero em três: *Pestalotiopsis*, *Neopestalotiopsis* e *Pseudopestalotiopsis* (MAHARACHCHIKUMBURA et al., 2012, 2014, 2016b, ARIYAWANSA et al., 2015).

Quanto ao modo de vida, as espécies descritas podem ser endofíticas, saprofíticas e, principalmente, fitopatogênicas, uma vez que causam sintomas como a podridão do fruto, o crestamento foliar, mancha foliar, podridão do caule, além de podridão pós-colheita em diversas famílias de plantas. Pode-se citar como exemplo a doença que causa em *Camellia sinensis* na China, a mancha cinzenta do chá, considerada uma das doenças foliares mais destrutivas nesta cultura (CHEN et al., 2018).

No Brasil, espécies dos gêneros *Pseudopestalotiopsis* e *Neopestalotiopsis* foram identificadas causando mancha foliar semelhante a antracnose em guaranazeiro (*Paullinia cupana*) (GUALBERTO et al., 2021). Ainda que essas espécies sejam associadas a doenças que causam prejuízos à agronomia, a análise química de produtos naturais têm revelado o potencial biotecnológico de espécies do gênero (WU, et al., 2022).

Mesmo antes da reformulação do gênero *Pestalotiopsis*, foram identificados metabólitos secundários produzidos pelo grupo com propriedades interessantes, tais como antivirais, antimicrobianos e antitumorais (Xu et al., 2010). Um exemplo disso, é a classe de policetídeo citosporina W, produzido por *Pseudopestalotiopsis theae*, que mostrou ação contra *Acinetobacter baumannii* multirresistente quando utilizado

como adjuvante da colistina e como atividade citotóxica frente ao linfoma de linhagem celular L5178 com IC₅₀ de 3.0 µM (YU, 2020).

Estudos relacionados a metabólitos secundários no gênero *Pestalotiopsis* identificaram entre 2014 e 2021 um total de 239 novos metabólitos secundários (WU, et al., 2022). Contudo devido a dificuldade de uma correta identificação taxonômica dessas espécies não é possível confirmar se todos esses metabólitos são realmente dos membros do gênero *Pestalotiopsis*, podendo ser também de espécies de *Neopestalotiopsis* ou *Pseudopestalotiopsis*, uma vez que separação dos gêneros ocorreu somente depois de 2014.

Independente do gênero, o grupo pestalóide atrai a atenção da comunidade científica devido aos seus metabólitos secundários estruturalmente complexos e biologicamente ativos. Moléculas com potencial para a cura de patologias e outras aplicações biotecnológicas como, antibióticos, antifúngicos, antioxidantes, anti-inflamatórios, compostos anticâncer e entre outros têm sido identificadas em pestalóide (YANG et al., 2012). Em particular, estudos sobre os *Pseudopestalotiopsis* levaram ao isolamento de novos derivados de policetídeos, com potencial aplicação na luta contra linhagens de câncer e bactérias resistentes a medicamentos (CHU WU et al., 2022).

Além da abordagem química, recentes ferramentas genômicas podem ser utilizadas para explorar a diversidade de clusters de genes biossintéticos (BGCs) que codificam enzimas relacionadas às vias de biossíntese de metabólitos secundários. Plataformas como FungiSMASH, possibilitam a identificação de BGCs e, conseqüentemente, dos metabólitos secundários relacionados a esse conjunto de genes, processo este conhecido como mineração genômica.

A mineração genômica facilita a prospecção de moléculas por meio de predição *in silico* com base em genomas completos, bem como permite identificar potenciais novas vias de biossíntese e novas moléculas por meio da comparação da similaridade entre os BGCs localizados no genoma em investigação e aqueles depositados em bases de dados como MiBIG (*Minimum Information about a Biosynthetic Gene Cluster*) que está interligado ao FungiSMASH, permitindo assim predizer os possíveis produtos naturais derivados desses BGCs (BLIN et al., 2019, TERLOUW et al., 2020).

Nesse contexto, a mineração de clusters biossintéticos (*Genome Mining*) nasce como uma ferramenta moderna e eficaz para a prospecção de metabólitos

secundários. Essa abordagem permite não apenas a descoberta de novos produtos naturais bioativos, mas também a compreensão das vias químicas e biológicas envolvidas na biossíntese desses compostos (Paulo, 2019; Lin, 2020; Mullis, 2019).

As ferramentas como BIG-SCAPE e CORASON que possuem estrutura computacional para explorar a diversidade biossintética em larga escala, usando algoritmos mais complexos para comparar simultaneamente dezenas ou centenas de genomas, e do ponto de vista de computação gráfica, disponibilizam em seus resultados de saída de arquivos, conjuntos de imagens que demonstram comparações de BGCs entre espécies diferentes ou iguais, permitindo avaliar a diversidade biossintética e identificar os BGCs exclusivos ou compartilhados entre eles (NAVARRO-MUÑOZ, JORGE C. et al., 2020).

Os metabólitos secundários preditos via mineração genômica podem ser classificados em vários grupos químicos, sendo que as maiores classes de produtos naturais incluem os terpenoides, policetídeos, peptídeos não-ribossomais e flavonoides. Cada classe possui uma estrutura básica que pode ser modificada por diferentes grupos funcionais, gerando uma enorme variedade de compostos com propriedades físicas, químicas e biológicas distintas (HOOFT *et al.*, 2020).

Em cada classe, a síntese da estrutura principal da molécula é realizada pela enzima codificada pelo que chamamos de "*gene core*" do BGC, podendo ser, por exemplo, uma policetídeo sintase (PKS - *polyketide synthases*), sintetase de peptídeos não ribossomais (NRPS - *non-ribosomal peptide synthetases*) e ciclases de terpenos (TCs - *terpene cyclase*). Algumas enzimas desempenham o papel de realizar modificações pós-traducionais, conhecidas como enzimas de decoração, que incorporam as mais diversas modificações ao esqueleto principal da moléculas através de reações de glicosilação (glicosiltransferase), metilação (metiltransferase), halogenação (halogenase), oxigenação (P450 monoxigenase), dentre outras (SHEN, 2003; SINGH; CHAUDHARY; SAREEN, 2017; WALSH; TANG, 2017; RUSSELL; TRUMAN, 2020 WANG et al., 2020, WANG et al., 2023).

A regulação da expressão gênica dos BGCs é realizada em resposta a estímulos ambientais ou fisiológicos e, embora possam ser identificados dezenas ou centenas de BGCs em cada genoma microbiano, em condições de laboratório a sua maioria está silenciado (SOLDATOU et al., 2019; HOOFT et al., 2020; ZIEMERT, WEBER e MEDEMA, 2020).

A atual disponibilidade em bancos públicos de milhares de genomas 63microbianos são uma fonte primária de BGC para prospecção de produtos naturais bioativos. A maioria da diversidade química codificada ainda permanece inexplorada, exigindo novas ferramentas e estratégias para acessar esses potenciais candidatos a medicamentos ou mediadores químicos. Os principais gargalos na utilização deste potencial incluem o reconhecimento e a priorização de genes biossintéticos interessantes, sua ativação e, finalmente, o estabelecimento da ligação entre os genes e os produtos naturais relacionados (PALAZZOTTO e WEBER, 2018; SOLDATOU et al., 2019).

A identificação e caracterização dos BGCs envolvidos na produção de metabólitos secundários de fungos do grupo pestalóides pode ser de grande interesse para a exploração do potencial biotecnológico desses microrganismos, visto que os dados químicos revelam o potencial das moléculas identificadas até o momento e que correspondem a uma pequena fração da capacidade metabólica desses microrganismos (KUHNERT et al., 2021).

Essas grandes quantidades de informações preditas através dos algoritmos e plataformas computacionais, podem auxiliar na seleção de fungos do grupo pestalóides candidatos a produtores de metabólitos secundários com potencial biotecnológico, e ainda contribuir na elucidação das vias biossintéticas envolvidas na formação desses compostos.

Diante do exposto, e em função da disponibilidade crescente de sequências genômicas que podem estar relacionadas a metabólitos secundários, o objetivo deste trabalho é explorar *in silico* a diversidade de agrupamentos de genes biossintéticos (BGCs) em genomas de *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis* utilizando mineração genômica. Os resultados desta pesquisa contribuirão para caracterização da diversidade de produtores naturais produzidos por pestalóide, contribuindo ainda para o avanço do conhecimento sobre a biossíntese de metabólitos secundários das espécies do grupo e para a descoberta de novas moléculas com potencial biotecnológico.

2. REVISÃO DA LITERATURA

2.1 Metabolismo secundário de *Pestalotiopsis*, *Neopestalotiopsis* e *Pseudopestalotiopsis*

O metabolismo secundário dos gêneros *Pestalotiopsis*, *Neopestalotiopsis* e *Pseudopestalotiopsis* tem recebido crescente atenção devido à sua capacidade de produzir uma diversidade de metabólitos bioativos. Esses compostos desempenham papéis cruciais na ecologia dos fungos, atuando como mecanismos de defesa. Além de suas funções ecológicas, muitos desses metabólitos apresentam potenciais aplicações farmacológicas, exibindo atividades antimicrobianas, antioxidantes, anti-inflamatórias e anticancerígenas (Wu et al., 2020).

Estudos recentes têm demonstrado que fatores epigenéticos exercem influência direta sobre a expressão gênica envolvida na biossíntese de metabólitos secundários em fungos filamentosos. Em *Pestalotiopsis microspora*, por exemplo, Zhang et al. (2018) identificaram o gene *mst2*, pertencente à família MYST de acetiltransferases de histonas, cuja atividade está associada à modificação da cromatina e ao controle transcricional de vias biossintéticas. A deleção desse gene resultou em uma redução significativa na produção de compostos como o Taxol (paclitaxel) e o pestalotiólídeo B, indicando que a acetilação de histonas é um dos mecanismos determinantes na ativação dos genes relacionados à síntese desses metabólitos bioativos. Esses achados reforçam a importância da regulação epigenética como ferramenta biotecnológica para o aumento da produtividade de substâncias de interesse farmacêutico em fungos endofíticos.

Diversas espécies do gênero *Pestalotiopsis* têm sido amplamente estudadas por sua capacidade de sintetizar metabólitos secundários bioativos com relevância farmacológica. Um dos primeiros e mais importantes exemplos foi o isolamento da substância taxol a partir de *Pestalotiopsis microspora*, um fungo endofítico de *Taxus wallachiana*, cuja produção variou entre 50 e 1487 ng/L (WU; WANG; YANG, 2022). Desde então, outras espécies como *P. mangiferae*, *P. pauciseta*, *P. breviseta*, *P. terminaliae* e *P. hainanensis* também demonstraram capacidade de sintetizar taxol e seus análogos, além de outros compostos como pestalotiollide B e melanina, com atividades antitumorais e antioxidantes relatadas (WU; WANG; YANG, 2022).

A diversidade biossintética do gênero é ainda evidenciada em espécies como *Pestalotiopsis heterocornis* e *Pestalotiopsis diploclisia*. A primeira produz compostos como heterocornols A–L, agropyrenol e vaccinol G, que exibem atividades citotóxicas contra diversas linhagens de células cancerígenas, além de ação antibacteriana e antifúngica moderada. Já *P. diploclisia* sintetiza os hidroquinonas pestalotioquinol G e pestalotioquinol H, além de pestalotioquinol A e B, que apresentam efeitos neuroprotetores e antimaláricos, com valores de IC₅₀ de 19 µM para *Plasmodium falciparum* (WU; WANG; YANG, 2022).

Rezazadeh et al. (2024) demonstra que o fungo endofítico *Neopestalotiopsis vitis*, isolado de *Corylus avellana*, é capaz de produzir *taxoides*, incluindo *paclitaxel*, um importante metabólito secundário com reconhecida aplicação no tratamento de diversos tipos de câncer. Por meio de análises por HPLC e LC-MS, foi confirmada a presença desses compostos, o que evidencia o potencial biotecnológico do gênero *Neopestalotiopsis* para a indústria farmacêutica. Essa descoberta amplia a relevância dos fungos endofíticos como fontes alternativas sustentáveis de compostos bioativos originalmente extraídos de plantas como *Taxus spp.*

Os estudos voltados a *Neopestalotiopsis vitis* representa uma promissora fonte para a produção sustentável de *paclitaxel*, reforçando o papel dos fungos endofíticos como plataformas biotecnológicas na síntese de metabólitos com propriedades anticâncer, antibióticas e antioxidantes (REZAZADEH et al., 2024)

Os achados de Xie et al. (2020) descrevem a identificação e caracterização de três metabólitos secundários produzidos por *Neopestalotiopsis clavispora*, isolada de folhas infectadas de *Kadsura coccinea*. Os compostos foram identificados como *afritoxinone A*, *afritoxinone B* e *oxysporone*, com base em análises de RMN e espectrometria de massas. Essas substâncias foram reconhecidas como fitotoxinas, desempenhando papel essencial na patogenicidade da espécie, uma vez que causaram lesões necróticas em folhas tratadas com diferentes concentrações dos compostos.

A investigação sobre a produção dessas fitotoxinas em *Neopestalotiopsis clavispora*, reforça o papel desse gênero como fonte de metabólitos secundários de importância ecológica e biotecnológica. Estudos sugerem que a identificação e o monitoramento de genes associados à biossíntese de *oxysporone* podem contribuir para distinguir entre cepas endofíticas e patogênicas de *N. clavispora*, além de

elucidar os mecanismos moleculares que regulam a síntese de fitotoxinas em fungos do grupo pestalotiíde (XIE et al., 2020).

O gênero *Pseudopestalotiopsis* tem sido reconhecido estruturalmente pelos metabólitos secundários que estão relacionados às vias para produção de policetídeos. *Pseudopestalotiopsis theae*, endofítico isolado de *Rhizophora racemosa*, produziu dezoito novos derivados de policetídeos, denominados pestalothéis I–Q e citosporinas O–W, além de compostos já conhecidos desta mesma classe (Yu et al., 2020). Esses metabólitos apresentam variações estruturais relevantes, como grupos hidroxila, acetila e cadeias laterais insaturadas, o que demonstra a capacidade biossintética altamente versátil dessa espécie. Dentre os compostos isolados, as citosporinas 22 e 23 se destacaram por apresentarem atividade antibacteriana contra linhagens resistentes de *Acinetobacter baumannii* quando combinadas a colistina em concentrações subletais. Além disso, o composto 23 exibiu citotoxicidade significativa ($IC_{50} = 3,0 \mu M$) frente à linhagem de linfoma murino L5178Y. Esses resultados demonstram o potencial farmacológico promissor de metabólitos derivados de *Pseudopestalotiopsis theae*, especialmente no contexto do desenvolvimento de agentes anticâncer e antimicrobianos

2.2- Diversidade Biossintética de BGC em larga escala

2.2.1-Mineração genômica

A descoberta de novos produtos naturais (PNs) microbianos é um passo fundamental na busca por novos antibióticos, antifúngicos e agentes antitumorais. A produção desses metabólitos secundários está intrinsecamente ligada à presença de Clusters de Genes Biossintéticos (BGCs), que são conjuntos de genes adjacentes no genoma, responsáveis por codificar as enzimas e proteínas envolvidas nas rotas metabólicas que levam à síntese de compostos bioativos (ROKAS et al., 2020). Essa organização gênica favorece a coordenação da expressão e a regulação conjunta da biossíntese.

O avanço das técnicas de sequenciamento genômico revelou uma diversidade metabólica muito superior ao número de compostos já isolados e descritos

experimentalmente, indicando que a maior parte do potencial biossintético microbiano permanece latente ou silencioso (HAUTBERGUE et al., 2018) (GUZMÁN-CHÁVEZ et al., 2018). Neste cenário, a mineração genômica emergiu como uma estratégia central para explorar esse potencial inexplorado.

A mineração genômica se inicia com a identificação de BGCs, sendo o antiSMASH (Antibiotics & Secondary Metabolite Analysis SHell) a ferramenta mais conhecida e utilizada para prever e identificar esses clusters (BLIN et al., 2025). O antiSMASH complementa essa análise ao fornecer uma anotação detalhada das sequências genômicas, facilitando a previsão das vias biossintéticas. A integração com o banco de dados MiBIG (Minimum Information about a Biosynthetic Gene cluster) permite que as predições sejam comparadas a BGCs já conhecidos, reduzindo as taxas de redescoberta (ZDOUC et al., 2025).

No entanto, o volume exponencial de dados genômicos impôs um desafio de escalabilidade. A análise comparativa de milhares de BGCs identificados pelo antiSMASH exige uma solução que vá além da simples anotação. É neste ponto que o pipeline BiG-SCAPE/CORASON (Biosynthetic Gene Similarity Clustering and Prospecting Engine / CORE Analysis of Syntenic Orthologues) se estabelece como uma estrutura computacional essencial, fornecendo a capacidade de análise comparativa e evolutiva em larga escala, transformando a mineração genômica em uma disciplina verdadeiramente escalável (NAVARRO-MUÑOZ et al., 2020).

2.2.2- Biosynthetic Gene Similarity Clustering and Prospecting Engine (BiG-SCAPE) e CORE Analysis of Syntenic Orthologues (CORASON)

O BiG-SCAPE reside na sua abordagem integrada e tecnicamente sofisticada para a comparação de BGCs, utilizando a saída do antiSMASH como dado de entrada. O algoritmo utiliza uma métrica de similaridade de BGCs que é robusta a diferentes modos de evolução, como a recombinação e a transferência horizontal de genes (HGT). A similaridade entre dois BGCs é determinada pela comparação do seu conteúdo de domínios proteicos, ordem e número de cópias dos genes constituintes. Isso é realizado através de uma métrica de distância que pondera a similaridade de sequência de genes centrais e a conservação da arquitetura do cluster. Com base

nessa métrica, os BGCs são conectados em uma Rede de Similaridade de Sequência (SSN), onde os nós representam BGCs e as arestas refletem a similaridade acima de um limiar definido. O algoritmo agrupa os BGCs em Famílias de Clusters de Genes (GCFs), que são sub-redes altamente conectadas (NAVARRO-MUÑOZ et al., 2019)..

Os GCFs representam agrupamentos de BGCs que partilham similaridade em conteúdo de domínios proteicos, ordem gênica e identidade de sequência, sendo presumivelmente responsáveis pela biossíntese de compostos naturais estruturalmente ou funcionalmente relacionados (NAVARRO-MUÑOZ et al., 2019). A validação do BiG-SCAPE demonstrou que os GCFs se alinham consistentemente com grupos de PNs quimicamente relacionados, estabelecendo uma ponte crucial entre o espaço genômico e o espaço químico (NAVARRO-MUÑOZ et al., 2020). O delineamento de GCFs pelo BiG-SCAPE é o primeiro passo para a descoberta de PNs, permitindo a identificação de BGCs órfãos e a priorização de BGCs com maior probabilidade de codificar novos PNs, com base em sua posição na rede e sua divergência em relação a BGCs conhecidos.

A filogenia de ortólogos sintênicos obtida pelo O CORASON complementa o BiG-SCAPE ao fornecer um contexto evolutivo detalhado para os GCFs. Para um GCF de interesse, o CORASON identifica um conjunto central de ortólogos sintênicos (enzimas) presentes nos BGCs, a partir dos quais uma árvore filogenética é construída. Esta filogenia permite rastrear a divergência evolutiva dos BGCs dentro de um GCF, auxiliando na identificação de BGCs que podem ter evoluído para produzir análogos estruturais com variações químicas sutis, mas biologicamente significativas. A combinação das SSNs do BiG-SCAPE com as árvores filogenéticas do CORASON fornece uma visão poderosa sobre como a diversidade química surgiu através de eventos de recombinação, HGT e duplicação/perda de genes. A integração entre genômica e metabolômica, reforçada pelo uso do BiG-SCAPE e ferramentas complementares como antiSMASH e MiBIG, permite mapear a diversidade metabólica de maneira mais eficiente, facilitando a identificação de vias biossintéticas responsáveis pela produção de novos compostos bioativos (VAN DER HOOFT et al., 2020).

Nos últimos cinco anos, o pipeline BiG-SCAPE/CORASON tem sido uma ferramenta muito utilizada para inúmeros estudos de mineração de genomas, demonstrando sua versatilidade em diversos contextos biológicos (Tabela 01), desde a análise de genomas isolados até a exploração de metagenomas complexos. Nos últimos anos alguns diversos estudos têm sido realizados por meio da análise de redes moleculares (NAVARRO-MUÑOZ et al., 2019)..

O futuro da mineração de genomas está intrinsecamente ligado à capacidade de processar o volume crescente de dados. A introdução do BiG-SCAPE 2.0 e do BiG-SLiCE 2.0 aborda essa necessidade de forma proativa (DRAISMA et al., 2025). O BiG-SLiCE foi inicialmente introduzido para lidar com a escalabilidade, sendo capaz de agrupar milhões de BGCs de forma eficiente (KAUTSAR et al., 2021). O BiG-SCAPE 2.0 representa uma atualização significativa, incorporando melhorias na classificação, métodos de alinhamento e visualizações, resultando em análises mais precisas e tempos de execução até oito vezes mais rápidos (DRAISMA et al., 2025). Essa evolução contínua garante que o pipeline BiG-SCAPE/CORASON permaneça na vanguarda da genômica de produtos naturais, permitindo a exploração de conjuntos de dados cada vez maiores com precisão e velocidade. Além disso, o BiG-SCAPE/CORASON tem sido fundamental para a validação de novas ferramentas de mineração de genomas, como o GATOR-GC que demonstrou maior habilidade para detecção de BGCs do que o antiSMASH (CEDIEL-BECERRA et al., 2025), estabelecendo-se como um padrão ouro para a análise de BGCs em larga escala.

Tabela 1: Estudos de caso recentes (2021-2025) que demonstram a aplicação do pipeline BiG-SCAPE/CORASON na mineração de genomas e os achados obtidos através da análise de redes moleculares.

Estudo de Caso	Organismo/Amostra	Foco da Análise	Achados Chave da Rede Molecular (BiG-SCAPE/CORASON)
<i>Rhodococcus</i>	Gênero <i>Rhodococcus</i>	Diversidade de NRPS e Evolução Filogenômica	A SSN revelou 44 GCFs de NRPS não identificados, com um padrão de evolução dependente da filogenômica, sugerindo eventos de HGT ou perda de genes (UNDABARRENA et al., 2021).
<i>Hypoxylaceae</i>	Família fúngica Hypoxylaceae e <i>Xylaria hypoxylon</i>	Diversidade de Metabólitos Secundários em Fungos	O BiG-SCAPE organizou BGCs em 375 GCFs. Apenas 10 GCFs foram conservados, indicando que a especiação é acompanhada por mudanças no metabolismo secundário (KUHNERT et al., 2021).
<i>Pseudomonas</i>	<i>Pseudomonas</i>	Potencial Genômico para Novos PNs	O pipeline avaliou a diversidade e distribuição de BGCs, resultando na identificação de 800 GCFs, muitos dos quais eram novos, destacando o potencial inexplorado do gênero (SAATI-SANTAMARÍA et al., 2022).
<i>Micromonosporaceae</i>	Micromonosporaceae de origem marinha	Diversidade Genética e Comparação com Banco de Dados	A análise BiG-SCAPE revelou diversidade genética e BGCs únicos em cepas marinhas, comparando-as com o banco de dados BiG-SLiCE (ALAS et al., 2024).
Microbioma Intestinal	Organismos longevos	Potencial Biossintético em Metagenomas	O BiG-SCAPE/CORASON analisou BGCs de terpenos em contigs montados, com o CORASON visualizando o maior GCF para análise detalhada da divergência (LIU et al., 2024).

<i>Streptomyces</i> sp. VB1	<i>Streptomyces</i> sp. VB1	Insights Metabolo-Genômicos	A análise BiG-SCAPE de GCFs confirmou a singularidade da cepa, identificando dois BGCs únicos em comparação com cepas relacionadas, sugerindo um potencial de descoberta (SERNA-CARDONA et al., 2024).
<i>Burkholderia</i>	Gênero <i>Burkholderia</i>	Padrão de Cluster de Metabólitos Secundários	A análise de rede BiG-SCAPE/CORASON identificou padrões específicos de grupo na distribuição de BGCs de metabólitos secundários, auxiliando na análise de componentes principais (KIM et al., 2024).
Microbioma Pulmonar	Microbioma Pulmonar Humano	Potencial Metabólico	O BiG-SCAPE identificou 1831 BGCs, revelando o potencial metabólico do microbioma pulmonar humano, um nicho pouco explorado (SEMMLER et al., 2024).
<i>Vibrio</i> Marinho	Bactérias Marinhas (<i>Vibrio</i>)	Diversidade e Variabilidade Estrutural de BGCs	O BiG-SCAPE foi usado para clustering de BGCs de sideróforos, avaliando variações genéticas e estruturais em espécies de <i>Vibrio</i> , destacando a plasticidade genômica (AL-SIYABI et al., 2025).

3.OBJETIVO

3.1Objetivo Geral:

Explorar a diversidade metabolômica por meio da análise de agrupamentos de genes biossintéticos (BGCs) e famílias de clusters de genes (GCFs) em genomas de *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis*

3.2Objetivo Específicos:

1- Identificar BGCs relacionados a síntese de produtos naturais com base na análise de 23 genomas de espécies dos gêneros *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis*;

2- Construir redes de similaridade de sequência de clusters de genes biossintéticos (BGCs) e agrupá-los em famílias de clusters de genes (GCFs) visando explorar a biossíntese de produtos naturais nas espécies de *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis*;

3-Realizar correção de redes de similaridade de BGCs e famílias de agrupamentos com produtos naturais já caracterizado conectando os BGCs aos respectivas moléculas;

4.MATERIAL E MÉTODOS

4.1 Genomas de *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis* utilizados neste trabalho

Para a identificação e caracterização de (BGCs) e (GCFs), foram analisados os genomas completos de 23 linhagens pertencentes aos gêneros *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis*. Entre esses genomas, 17 foram obtidos de bancos de dados públicos como o NCBI e o JGI. Os 6 genomas restantes são provenientes da coleção da Embrapa Amazônia Ocidental e correspondem a linhagens isoladas na região amazônica, atuando como patógenos em plantações de guaraná e açaí, conforme detalhado na (Tabela 2).

Tabela 2. Genomas completos de *Pestalotiopsis*, *Neopestalotiopsis* e *Pseudopestalotiopsis* disponíveis nos bancos de dados públicos ou depositados no banco de dados do Projeto AmazonMicro Embrapa.

Identificação no NCBI	Banco	Nº de acesso	Tamanho (Mb)	Scaffold	Nº de Proteínas	Modo de Vida
<i>Pestalotiopsis</i> sp. CR013	NCBI	CR013	46.69	12	12827	Não informado
<i>Pestalotiopsis</i> sp. 9143B	NCBI	9143B	46,26	104	18111	Não Informado
<i>Pestalotiopsis</i> sp. JCM 9685	NCBI	BCGF01	48.23	13	12781	Não Informado
<i>Pestalotiopsis</i> sp. LF1912	NCBI	LF1912	48,14	320	19613	Não Informado
<i>Pestalotiopsis</i> sp. DP-2020a	NCBI	DP-2020a	44.92	1,119	12021	Não informado
<i>Pestalotiopsis</i> sp. MUOB 44551	NCBI	JOB 440515	47.67	10	14284	Não informado

<i>Pestalotiopsis kenyana</i>	NCBI	PG52	57.99	331	16873	Patógeno
<i>Pestalotiopsis</i> sp. NC0098	NCBI	NC0098	46.41	24	15175	Não informado
<i>Pestalotiopsis fici</i>	NCBI	W106-1	51.91	50	15413	Endofítico
<i>Pseudopestalotiopsis theae</i>	NCBI	CYF27	50.41	61	14786	Patógeno
<i>Pseudopestalotiopsis</i> sp. 1.26	mbrapa	CPAA-1.26	49.73	162	19909	Patógeno
<i>Pseudopestalotiopsis</i> sp. 1.2	mbrapa	CPAA-1.2	50.52	228	20317	Patógeno
<i>Pseudopestalotiopsis chinensis</i>	NCBI	AH-3-21	51.22	36	14737	Patógeno
<i>Neopestalotiopsis</i> sp. 1.3	mbrapa	CPAA-1.3	52.38	399	21475	Patógeno

<i>Neopestalotiopsis sp. 1.5</i>	mbrapa	CPAA-1.5	51.16	228	21011	Patógeno
<i>Neopestalotiopsis formicidarum</i>	mbrapa	NPA 2916	50.8	196	14722	Patógeno
<i>Neopestalotiopsis sp.</i>	mbrapa	PAA-AÇAÍ	50.66	162	20899	Patógeno
<i>Neopestalotiopsis sp. 37M</i>	NCBI	37M	50.5	229	14867	Não informado
<i>Neopestalotiopsis sp. M2102B</i>	NCBI	M2102B	50,62	314	17365	Não informado
<i>Neopestalotiopsis sp. PM2101</i>	NCBI	PM2101	50,05	220	18928	Não informado
<i>Neopestalotiopsis sp. PM2101E</i>	NCBI	PM2101E	49,95	197	17325	Não informado
<i>Neopestalotiopsis clavispora</i>	NCBI	HI 201606	50.46	661	15199	Patógeno

<i>Neopestalotiopsis rosae</i>	NCBI	ML1664	53.78	18	18402	Endofítico
--------------------------------	------	--------	-------	----	-------	------------

4.2 Análise filogenética e identificação filogenômica de linhagens de *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis*

As sequências dos barcodes ITS, *tef1- α* e β -tubulina dos isolados que possuem genoma completo disponíveis no banco de dados do GenBank foram processadas para obtenção de dados por meio de BLASTn. Para a construção do dataset, selecionamos espécies próximas, resultando em 23 linhagens de *Pestalotiopsis* (incluindo 15 isolados tipo), 16 linhagens de *Pseudopestalotiopsis* (com 11 isolados tipo) e 20 linhagens de *Neopestalotiopsis* (com 10 isolados tipo). Utilizamos *Truncatella angustata* como grupo externo para a análise filogenética.

O alinhamento das sequências foi guiado individualmente para cada locus (ITS, *tef1- α* e β -tubulina) utilizando script MAFFT usando o software UGENE na versão 51.0. Após o alinhamento, as sequências concatenadas foram submetidas a uma análise de Máxima Verossimilhança (ML), com 1000 reamostragens de bootstrap, utilizando a plataforma IQTREE. A topologia resultante da árvore filogenética, assim como os valores de suporte dos ramos, foi visualizada e editada na plataforma iTOL. Para edição final da figura, utilizou-se o software CorelDraw versão 2020.

4.3 Análise com as ferramentas antiSMASH e BIG-SCAPE

A identificação e caracterização dos BGCs foram realizadas utilizando a ferramenta antiSMASH (*Antibiotics and Secondary Metabolite Analysis Shell*), empregada na predição automática de clusters biossintéticos em genomas microbianos. Foram analisados 23 genomas completos pertencentes aos gêneros *Pestalotiopsis*, *Neopestalotiopsis* e *Pseudopestalotiopsis*.

Os arquivos dos genomas, no formato .fasta e .fna foram processados individualmente no antiSMASH versão 8.0, com todas as opções de detecção de clusters ativadas, incluindo módulos de NRPS (Nonribosomal Peptide Synthetase), PKS (Polyketide Synthase), terpenos, RiPPs e híbridos. O modo “*KnownClusterBlast*” foi habilitado, permitindo a comparação dos BGCs identificados com o banco de dados MIBiG (*Minimum Information about a Biosynthetic Gene Cluster*), o que possibilita avaliar a similaridade entre clusters novos e compostos previamente descritos.

Como saída, o antiSMASH gerou um conjunto de arquivos “.GBK”, “.JSON” e “.SVG”, organizados em pastas específicas para cada genoma. Esses arquivos contêm as anotações dos genes, domínios catalíticos e predições de produtos secundários associados. Posteriormente, todos os arquivos. GBK referentes aos BGCs de cada genoma foram agrupados em uma pasta geral, utilizada como entrada para a análise de agrupamento com o BiG-SCAPE.

O BiG-SCAPE foi executado em modo local utilizando os arquivos GBK gerados pelo antiSMASH para os 23 genomas, com a opção de comparação com o banco MIBiG ativada, permitindo o alinhamento dos BGCs dos isolados com clusters conhecidos de metabólitos bioativos.

4.5 Análise utilizando bibliotecas PYTHON

Neste estudo, Python foi empregada para processar os arquivos de saída do antiSMASH e BiG-SCAPE, integrando pacotes como Pandas para tratamento tabular, Matplotlib e Seaborn para geração de gráficos e heatmaps. Essa abordagem programática permitiu não apenas maior reprodutibilidade e rastreabilidade dos passos de análise, mas também a criação de visualizações personalizadas (por exemplo, gráficos empilhados e mapas de calor) que elucidam a diversidade de clusters biossintéticos nos 23 genomas analisados. O módulo Biopython, especificamente projetado para tarefas em biologia computacional como leitura de formatos de sequência (FASTA, GenBank), análise de alinhamentos e extração de informações de metadados também foi incorporado nesta metodologia para auxiliar na pré-processamento de arquivos anotados e extração de IDs de BGCs. Ao combinar Biopython com scripts em Python para cruzamento de tabelas, cálculo de proporções e geração gráfica, foi possível estabelecer um fluxo integrado (“pipeline”) que vai desde a identificação de clusters biossintéticos até a produção de relatórios visuais.

Essa cadeia computacional assegura que os resultados sejam replicáveis, escaláveis e aptos para discussão em termos de diversidade biossintética em nível genômico.

5. RESULTADOS E DISCUSSÃO

5.1 Identificação filogenética e filogenômica de linhagens de *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis* com genomas disponíveis

A acurada identificação e delimitação taxonômica das espécies fúngicas dos gêneros *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis* são passos cruciais para a compreensão de seu potencial biossintético e para a correta atribuição de quaisquer metabólitos secundários descobertos. Para tanto a análise filogenética das 23 linhagens em foco foi realizada (Figura 1).

A topologia da árvore filogenética, resultante da análise combinada dos marcadores moleculares ITS, *tef1- α* e β -tubulina, demonstrou uma clara e bem suportada segregação em clados distintos e monofiléticos para cada um dos gêneros investigados. Essa distinção corrobora plenamente as reclassificações taxonômicas modernas, que fragmentaram o antigo gênero *Pestalotiopsis* em entidades filogeneticamente mais coerentes. A eficácia dos barcodes empregados para uma precisa identificação taxonômica das espécies dos gêneros analisados apresentou baixa resolução.

A análise filogenética permitiu resolver e retificar ambiguidades e incorreções taxonômicas previamente existentes, tanto em identificações de bancos de dados quanto em designações genéricas. As 17 linhagens oriundas de genomas públicos apresentaram três cenários distintos. Primeiramente, observou-se a reclassificação de espécies previamente nomeadas para designações de "sp. nov.", indicando que a análise filogenética sugeriu que se tratavam de novas espécies dentro de seus respectivos gêneros, apesar de uma identificação anterior. Exemplos notáveis incluem linhagens originalmente identificadas como *Neopestalotiopsis clavispora* e *Neopestalotiopsis rosae*, que foram reavaliadas como *Neopestalotiopsis* sp. nov., e *Pseudopestalotiopsis chinensis* e *Pseudopestalotiopsis theae*, também reclassificadas para *Pseudopestalotiopsis* sp. nov.. Tais reavaliações sublinham a importância da confirmação filogenética em depósitos do NCBI mostrando a imprecisão taxonômica de muitos acessos do banco público.

Em segundo lugar, a análise filogenética foi crucial para a identificação precisa de linhagens que estavam genericamente designadas como "sp." (espécie indeterminada) para o nível de "sp. nov.". Este grupo inclui várias linhagens de *Neopestalotiopsis*, como *Neopestalotiopsis* sp. 37M, *Neopestalotiopsis* sp. M2102B, *Neopestalotiopsis* sp. PM2101 e *Neopestalotiopsis* sp. PM2101E, além de linhagens de *Pestalotiopsis* como *Pestalotiopsis* sp. JCM 9685, *Pestalotiopsis* sp. DP-2020a e *Pestalotiopsis* sp. LF1912. A elevação dessas designações genéricas para "sp. nov." não apenas confere maior especificidade às linhagens, mas também direciona o foco para sua potencial singularidade dentro do gênero, possivelmente refletindo perfis biossintéticos ainda não explorados.

Por fim, e de forma igualmente significativa, o estudo demonstrou a capacidade de identificar linhagens "sp." a espécies conhecidas e, em um caso, a reclassificar uma espécie nomeada para um novo gênero e espécie. Linhagens como *Pestalotiopsis* sp. 9143B e *Pestalotiopsis* sp. CR013 foram identificadas como *Pestalotiopsis jangxiensis*, enquanto *Pestalotiopsis* sp. MUOB 44551 foi designada como *Pestalotiopsis kenya*, e *Pestalotiopsis* sp. NC0098 como *Pestalotiopsis photinicola*. Adicionalmente, uma correção notável foi a linhagem *Pestalotiopsis fici*, que foi reavaliada e corretamente atribuída a *Pseudopestalotiopsis fici*, evidenciando que a imprecisão taxonômica por esta inclusive em nível de gênero. Essas correções e novas identificações de linhagens que anteriormente careciam de uma atribuição taxonômica clara são importantes para a bioprospecção, pois cada espécie, ou mesmo cada nova espécie, pode possuir um perfil metabólico distinto e único. A superação dessas incertezas taxonômicas, proporcionando um nível de resolução elevado e confiável para as linhagens analisadas, demonstra a robustez da abordagem e estabelece uma base sólida para as investigações subsequentes sobre a diversidade de clusters de genes biossintéticos.



Figura 1- Estudo filogenético de espécies de *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis*. Destacam-se em verde 17 linhagens cujas sequências genômicas estão disponíveis em repositórios de dados públicos, e em magenta, sequências genômicas de isolados patogênicos de plantas de guaraná e açai apresentadas neste estudo. Os símbolos * indicam espécies cuja identificação prévia está incorreta, e o símbolo # denota espécies anteriormente classificadas como "sp" e que foram corretamente identificadas a nível de espécie através desta pesquisa.

5.2. Distribuição de BGCs por gênero

A análise quantitativa dos Clusters de Genes Biossintéticos (BGCs) revelou variações marcantes no potencial metabólico entre os gêneros *Pestalotiopsis*, *Neopestalotiopsis* e *Pseudopestalotiopsis*. A contagem total de BGCs por gênero evidenciou que *Neopestalotiopsis* se destaca com o maior número de clusters, totalizando 989 BGCs, seguido por *Pestalotiopsis* com 666 BGCs e *Pseudopestalotiopsis* com 458 BGCs. Essa hierarquia na abundância de BGCs sugere uma maior versatilidade biossintética em *Neopestalotiopsis*, o que pode estar associado a estratégias de adaptação ecológica mais amplas ou a uma amplificação de vias de defesa e competição.

A distribuição individual por genoma reforça essa observação, com destaque para linhagens como *Neopestalotiopsis* sp. 1.5 (111 BGCs) e *Neopestalotiopsis rosae* ML1664 (104 BGCs), que exibem os maiores números de clusters (Figura 2). As diferenças observadas na quantidade total de BGCs entre os gêneros indicam uma variação evolutiva na capacidade de geração de metabólitos secundários, refletindo possíveis especializações químicas dentro da família.

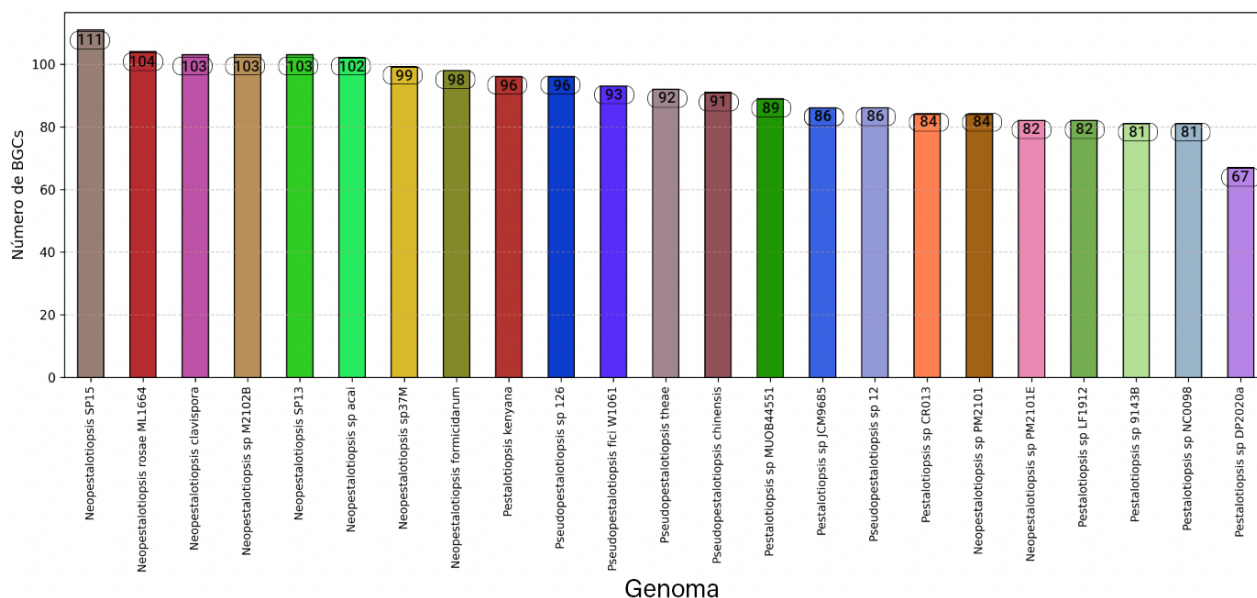


Figura 2 - Número total de *Biosynthetic Gene Clusters* (BGCs) por genoma (23 genomas analisados).

Ao examinar a composição das classes de BGCs, a análise percentual revelou perfis distintos e complementares entre os gêneros (Figuras 3 e 4). As classes NRPS (*Nonribosomal Peptide Synthetase*) e PKS (*Polyketide Synthase*) tipo I foram as mais representativas em todos os gêneros, reafirmando sua importância central na produção de metabólitos secundários nessas espécies fúngicas. Em *Neopestalotiopsis*, observou-se uma distribuição notavelmente equilibrada, com PKS representando 27,40% do total de BGCs, NRPS com 25,48% e Terpenos com 12,74%. Essa composição equilibrada sugere que *Neopestalotiopsis* pode possuir rotas metabólicas mais diversificadas para a síntese de compostos bioativos, especialmente peptídeos e terpenos, o que pode refletir uma maior especialização ecológica e versatilidade química, permitindo a adaptação a múltiplos nichos e a produção de compostos bioativos com variadas funções fisiológicas.

Para *Pestalotiopsis*, o perfil apresentou uma ligeira predominância das classes NRPS (28,98%) e PKS (28,68%), indicando um foco biossintético mais concentrado na produção de policetídeos e peptídeos não ribossomais. Já em *Pseudopestalotiopsis*, embora NRPS e PKS continuem sendo classes relevantes, houve um destaque para a categoria "Others" (29,26%) e para Terpenos (10,26%). A presença de uma proporção significativa da classe "Others", embora indique diversidade, sugere que *Pseudopestalotiopsis* pode estar envolvido na biossíntese de metabólitos menos categorizados ou com vias biossintéticas menos comuns.

Adicionalmente, os dados destacam a consistente presença de híbridos PKS-NRP em todos os gêneros, indicando a existência de vias biossintéticas complexas que integram mecanismos de diferentes famílias enzimáticas, expandindo ainda mais a diversidade estrutural dos metabólitos produzidos. Em contraste com o perfil de *Neopestalotiopsis*, a composição mais concentrada de *Pestalotiopsis* e *Pseudopestalotiopsis* sugere a conservação de vias metabólicas específicas, que podem estar relacionadas à evolução divergente e à especialização ambiental dentro da família Sporocadaceae.

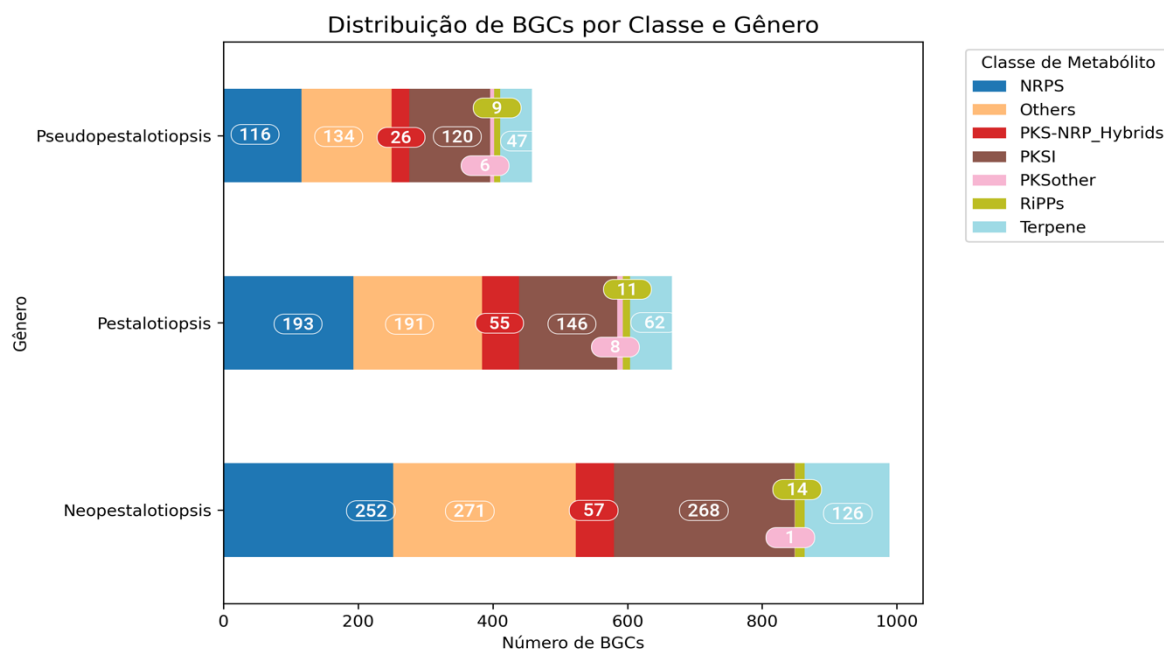


Figura 3: Distribuição das classes de *Biosynthetic Gene Clusters* (BGCs) identificadas por gênero entre os 23 genomas analisados, evidenciando diferenças nas vias metabólicas predominantes.

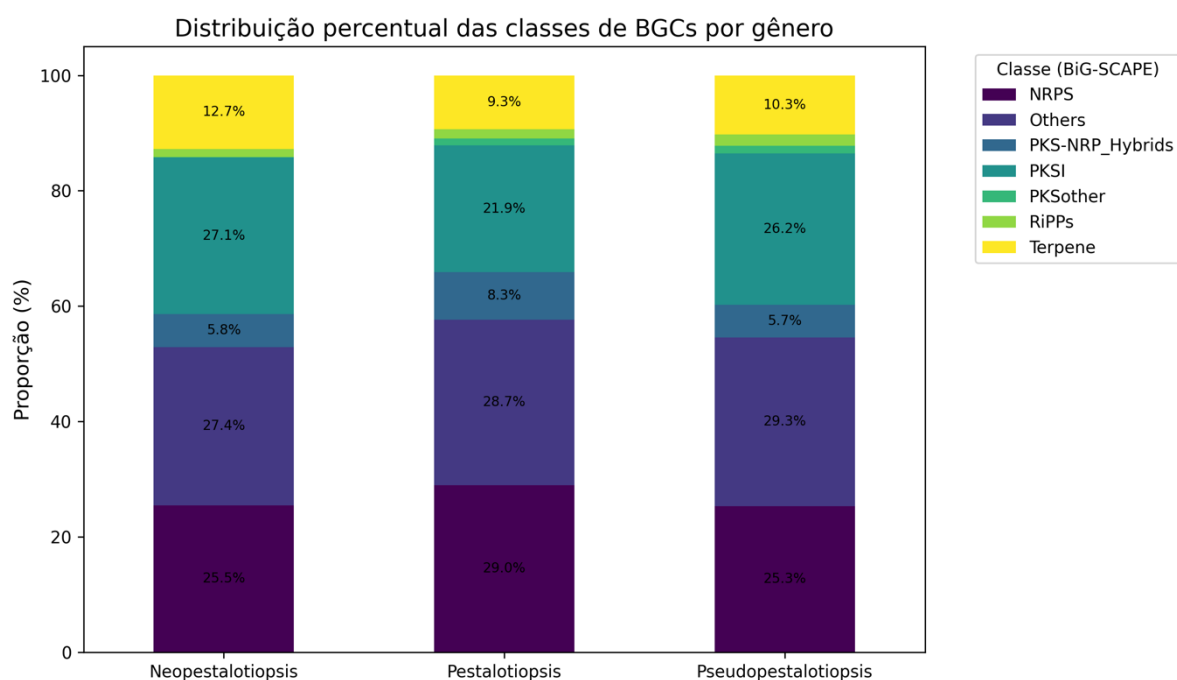


Figura 4 – Distribuição percentual das classes de *Biosynthetic Gene Clusters* (BGCs) identificadas em cada gênero (*Pestalotiopsis*, *Neopestalotiopsis* e *Pseudopestalotiopsis*).

A análise global dos 23 genomas pertencentes a espécies dos gêneros *Pestalotiopsis*, *Pseudoestalotiopsis* e *Neopestalotiopsis* permitiu a identificação de um total de 2.113 BGCs. Este volume de BGCs, para a biossíntese de metabólitos secundários, evidencia um numeroso conjunto de metabólitos secundários que essas espécies fúngicas possuem.

A distribuição desses BGCs entre as diversas classes de metabólitos secundários revelou um padrão predominante de NRPS (*Nonribosomal Peptide Synthetase*), PKS (*Polyketide Synthase*) e cluster híbridos PKS-NRP (Figura 5).

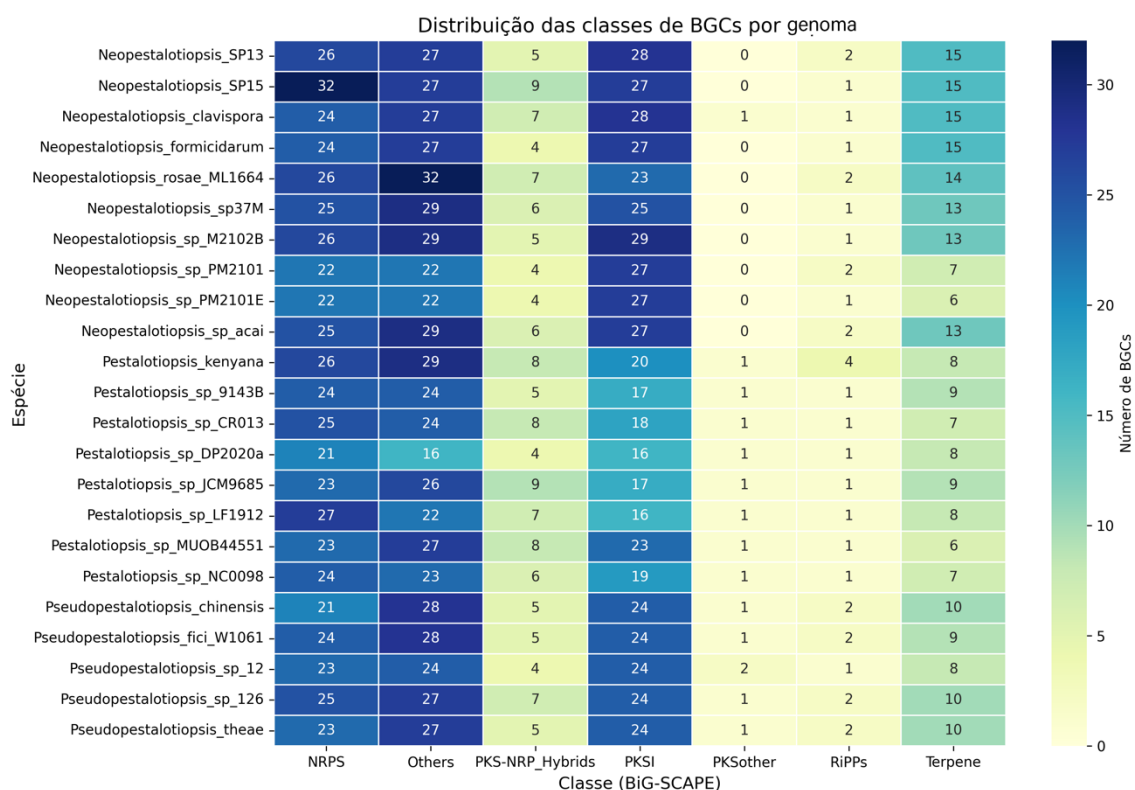


Figura 5. Heatmap representando a distribuição das classes de Clusters de Genes Biossintéticos (BGCs) entre os 23 genomas dos gêneros *Pestalotiopsis*, *Neopestalotiopsis* e *Pseudoestalotiopsis*. Cada célula indica o número de BGCs por classe (NRPS, PKSI, híbridos PKS-NRP, PKS other, RiPPs, Terpene e Others) identificado em cada genoma. A intensidade da coloração varia conforme a quantidade de BGCs, onde tons mais escuros representam maior abundância.

A alta representatividade de NRPS e PKS é particularmente interessantes do ponto de vista biotecnológico, uma vez que peptídeos não ribossomais e policetídeos são reconhecidos como compostos bioativos, incluindo antibióticos, antifúngicos, citotoxinas e imunomoduladores e no caso de patógenos de plantas como muitas das

linhagens analisadas a produção de fitotoxinas que podem ser utilizadas no desenvolvimento de herbicidas para a agricultura. Essa preponderância de BGCs ressalta o potencial dos fungos estudados como fontes de moléculas com interesse farmacológico e biotecnológico, consolidando o crescente interesse nesses gêneros na busca por novos produtos naturais. Classes como terpenos também foram consistentemente representadas, enquanto RiPPs (peptídeos ribossomalmente sintetizados e pós-traducionalmente modificados) apresentaram uma frequência menor, um padrão comum em comparação com outros fungos e bactérias.

5.3. Famílias de Clusters de Genes (GCFs) presentes em *Pestalotiopsis*, *Neopestalotiopsis* e *Pseudopestalotiopsis*

5.3.1 GCFs relacionadas à biossíntese de peptídeos não ribossomais (NRP)

Foram identificadas 165 famílias de clusters biossintéticos (GCFs) entre 795 NRPS identificados. Destes, 9 correspondem a “core GCFs” (presentes simultaneamente nos três gêneros analisados). A distribuição das famílias denominadas “accessory GCFs” (compartilhados entre dois gêneros) corresponde a 18 GCFs principalmente entre *Neopestalotiopsis*/*Pseudopestalotiopsis* (12 GCFs), *Neopestalotiopsis*/*Pestalotiopsis* (4 GCFs) e *Pestalotiopsis*/*Pseudopestalotiopsis* (2 GCFs). As GCFs gênero específicas (“unique GCFs”) são 95, sendo 35 exclusivos de *Pestalotiopsis*, 43 de *Neopestalotiopsis* e 17 de *Pseudopestalotiopsis*. E 43 representam “singletons”, contendo apenas um membro na família (Figura 6). Os dados evidenciam que a maioria das famílias NRPS são específicas de cada gênero, especialmente para *Neopestalotiopsis*. As interseções destacam o compartilhamento metabólico acessório e o núcleo biossintético conservado entre todos os grupos (core GCFs). Esse panorama revela uma ampla especialização metabólica, apontando tanto para potenciais vias centrais quanto acessórias envolvidas na produção de metabólitos secundários nesses fungos.

Compartilhamento de famílias de clusters NRPS entre gêneros

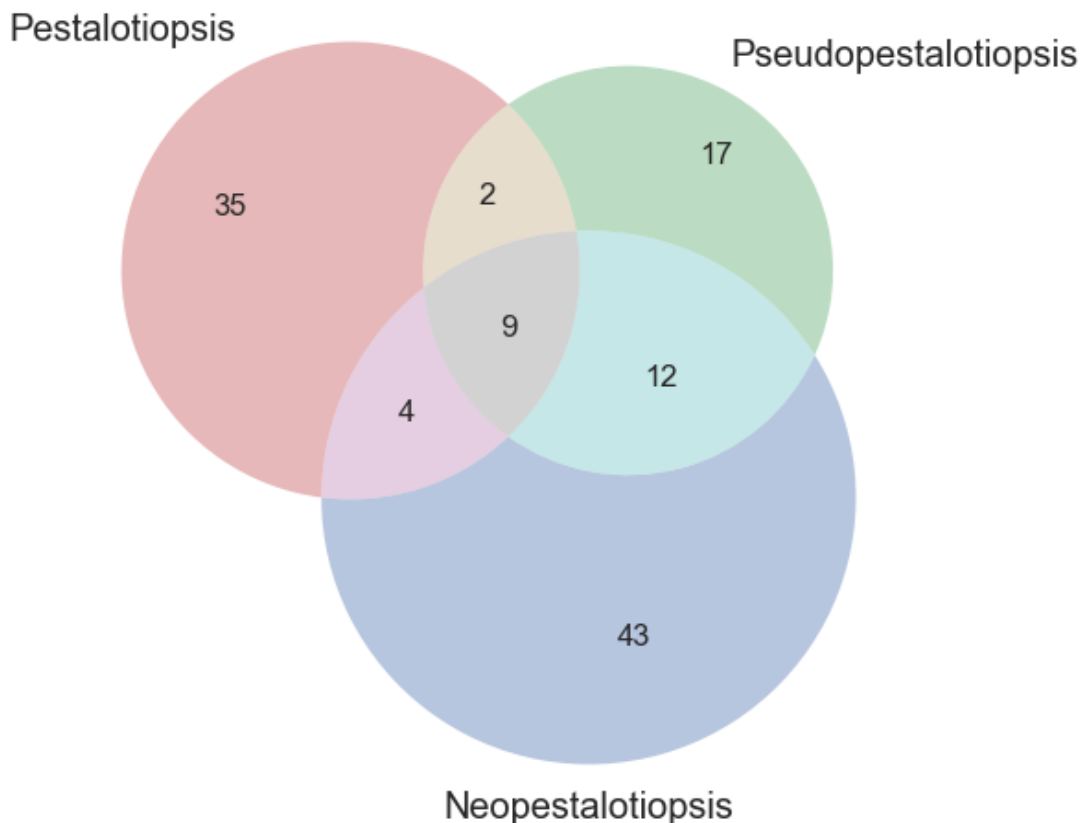


Figura 06. Diagrama de Venn ilustrando a distribuição e o clusters NRPS (GCFs) entre os gêneros *Pestalotiopsis*, *Pseudopestalotiopsis* e compartilhamento de famílias de *Neopestalotiopsis*.

5.3.2 GCFs PKS1

Entre as policetídeo sintase tipo 1 (PKS1) foram identificadas 141 famílias de clusters biossintéticos entre os 607 BGCs identificados. Destes, apenas 3 GCFs são compartilhadas simultaneamente entre os três gêneros. Foram encontradas 14 GCFs, compartilhada entre dois gêneros *Neopestalotiopsis*/*Pseudopestalotiopsis* (7 GCFs), *Neopestalotiopsis*/*Pestalotiopsis* (5 GCFs) e *Pestalotiopsis*/*Pseudopestalotiopsis* (2 GCFs). A maior parte das famílias são exclusivas para cada gênero: *Neopestalotiopsis* (39), *Pestalotiopsis* (22), *Pseudopestalotiopsis* (18). Este resultado evidencia forte especialização metabólica para PKS1 entre os três gêneros, com predominância de famílias únicas e limitada presença de vias conservadas (core) (Figura 7).

Compartilhamento de famílias de clusters PKS1 entre gêneros

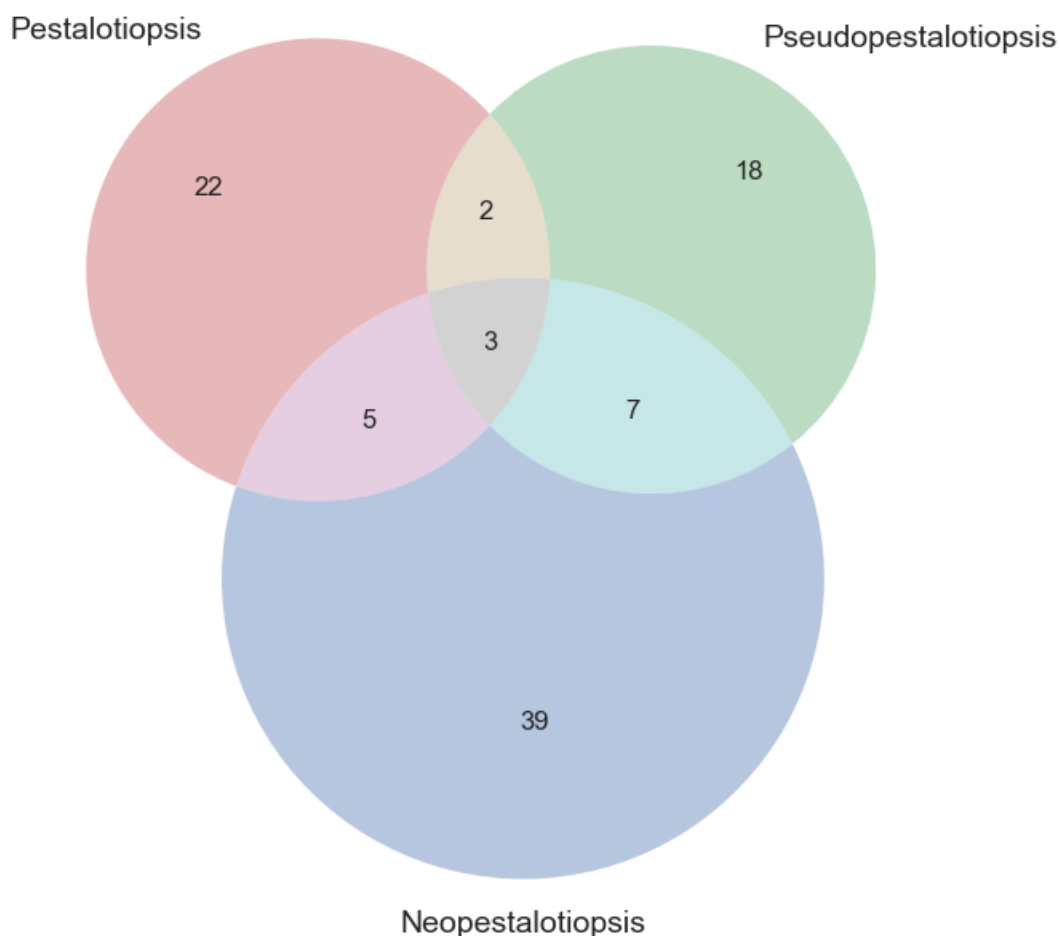


Figura 07. Compartilhamento de famílias de clusters PKS1 entre os gêneros *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis*.

5.3.3 GCFs Terpeno

Para a classe de terpeno, foram identificadas famílias de clusters cuja distribuição evidencia alta especificidade entre os gêneros analisados. Destacam-se 20 famílias exclusivas de *Neopestalotiopsis*, 9 exclusivas de *Pestalotiopsis* e 8 exclusivas de *Pseudopestalotiopsis*. Foram observadas ainda 4 core GCFs compartilhadas entre todos os gêneros, além de interseções entre os diferentes gêneros: 1 entre *Neopestalotiopsis*/*Pseudopestalotiopsis*, 1 entre *Pestalotiopsis*/*Pseudopestalotiopsis* e nenhuma entre *Pestalotiopsis*/*Neopestalotiopsis* (Figura 8). Assim como para as demais classes, os resultados indicam predominância de vias metabólicas específicas para cada gênero, com limitada presença de GCFs terpênicos

compartilhados entre os gêneros. Dessa forma, a análise revela que a diversidade biossintética de terpenos é majoritariamente impulsionada pela especialização metabólica.

Compartilhamento de famílias de clusters Terpene entre gêneros

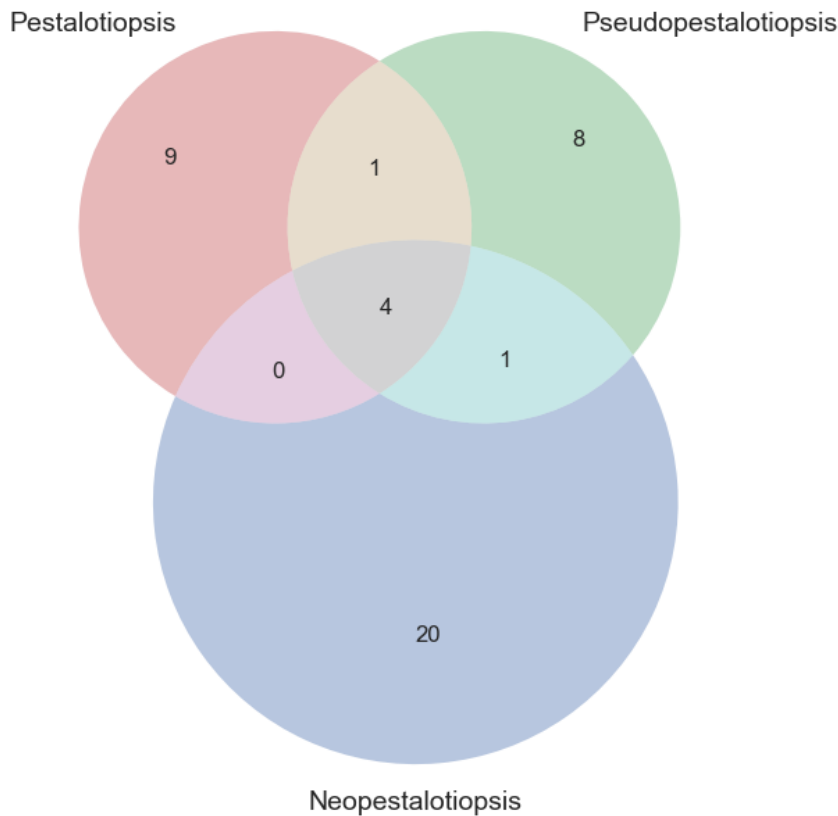


Figura 08. Diagrama de Venn referente ao compartilhamento de famílias de clusters Terpene entre os gêneros *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis*.

5.3.4 Panorama quantitativo das famílias de clusters biossintéticos identificadas

Número total de famílias de clusters (GCFs) identificadas, número de BGCs agrupados, e distribuição das categorias “core” compartilhada pelo três gêneros, “accessory” compartilhados por dois dos três gêneros, “unique” exclusivos de cada gênero e “singleton” para cada classe de BGC analisada nos gêneros *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis*.

Tabela 3. Distribuição das famílias de clusters biossintéticos (GCFs) identificadas em cada classe entre espécies dos gêneros *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis*.

Classe	N° GCFs	N° BGCs	Core	Accessory	Unique	Singleton
NRPS	165	795	9	18	95	43
PKS1	141	607	3	14	79	45
Terpene	63	282	4	2	37	20
PKS other	47	162	0	6	25	16
PKS-NRP	36	138	0	7	19	10
RiPPs	12	36	2	0	3	7
Others	168	596	4	12	86	66

Observa-se que as classes NRPS e Others reúnem o maior número de famílias e de BGCs. O número de GCFs consideradas “core” é restrito em todas as classes, destacando que apenas uma pequena fração das vias metabólicas identificadas é conservada e compartilhada pelos três gêneros analisados. Em contrapartida, predomina a ocorrência de GCFs exclusivas (“unique”) ou com apenas um membro

(“singleton”), especialmente em terpene, PKS-other, PKS-NRP e Others, reforçando a especialização química de cada gênero. Classes como PKS1 e NRPS apresentam valores relativamente maiores para GCFs “accessory”, sugerindo certa plasticidade evolutiva e possibilidade de compartilhamento facultativo de vias biossintéticas entre dois gêneros. Esses resultados sustentam a hipótese de que a diversidade de metabólitos secundários desses fungos se deve principalmente à aquisição de famílias exclusivas, contribuindo para perfis metabólicos singulares.

A comparação das 632 famílias de clusters biossintéticos (GCFs) identificadas nos 23 genomas analisados com o banco de dados MIBiG revelou correspondência com clusters caracterizados experimentalmente em apenas quatro famílias, conforme apresentado na Tabela 4. Essas quatro GCFs agrupam um total de 29 BGCs provenientes dos genomas de *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis*, representando aproximadamente 0,9% do repertório biossintético total de 2.113 BGCs identificados.

A GCF 3199, pertencente à classe PKS1, constitui a família mais robusta, reunindo 13 BGCs, dos quais 10 são originários dos genomas estudados e três correspondem a clusters do MIBiG (BGC0001257.1, BGC0001258.1 e BGC0001906.1) previamente caracterizados em fungos endofíticos e fitopatogênicos como *Nodulisporium* sp. ATCC74245, *Glarea lozoyensis* e *Daldinia eschscholzii* IFB-TL01. No contexto de *Daldinia eschscholzii*, o cluster BGC0001906.1 (ou os clusters gênicos subjacentes a ele) está especificamente associado à biossíntese dos imunossupressores dalmanol A e acetodalmanol A. Conforme detalhado por Zhou e colaboradores (2019), a produção dessas moléculas envolve o acoplamento de dois aglomerados gênicos de policetídeos (PGCs) distintos: um que codifica naftaleno (com a NR-PKS pksTL) e outro que codifica cromano (com a PR-PKS ChrA e a parceira KR ChrB). Esta via, que combina precursores de naftaleno e cromano através de um mecanismo de acoplamento catalisado por uma monooxigenase inespecífica (provavelmente do tipo P450), resulta na formação de produtos híbridos, destacando um mecanismo de orquestração inter-cluster que amplia a diversidade de policetídeos fúngicos e tem relevância no desenvolvimento de novos compostos bioativos.

A GCF 3482, também da classe PKS1, agrupa seis BGCs, sendo cinco provenientes dos genomas analisados e um do MIBiG (BGC0000121.1) caracterizado

em *Pestalotiopsis fici*, sugerindo que os isolados estudados podem produzir compostos relacionados (Figura 9).

A GCF 3705, classificada como "Others", reúne cinco BGCs. Dentro deste contexto, a pesquisa realizada por Pan et al. (2018) detalha a biossíntese de importantes metabólitos em *Pestalotiopsis fici*. Este fungo endofítico é a fonte de moléculas como o iso-A82775C (10), um prenilepoxiciclohexano, e os cloropestolides. A via biossintética do iso-A82775C envolveu a identificação do BGC *iac* e a caracterização da prenilttransferase *lacE* como uma enzima chave, cuja deleção resultou na acumulação de sicayne (2) e na geração de quatro novos cloropestolides: cloropestolides H-K (3-6). As aplicações dessas moléculas são notáveis: cloropupukeananina e os cloropestolides A-D (11-14), co-isolados do mesmo fungo, já eram conhecidos por suas propriedades antimicrobianas, antitumorais e anti-HIV. O estudo aprofundou essas descobertas, revelando que os recém-gerados cloropestolides 5 e 6 demonstraram efeitos antibacterianos contra *Staphylococcus aureus* e *Bacillus subtilis*. Adicionalmente, o cloropestolide 5 exibiu citotoxicidade contra linhagens de células tumorais humanas HeLa, MCF-7 e SW480, destacando o potencial farmacológico dessas moléculas complexas.

A baixa proporção de GCFs com correspondência ao MIBiG (4 de 632, ou 0,63%) e de BGCs individuais (19 de 2.113, ou 0,9%) indica que a vasta maioria das famílias biossintéticas identificadas constitui clusters órfãos, sem rotas metabólicas caracterizadas ou produtos conhecidos na literatura. Esse panorama reforça o elevado potencial dos fungos estudados para a descoberta de novos metabólitos secundários com estruturas químicas inéditas e possíveis aplicações biotecnológicas e farmacêuticas. A identificação de BGCs com correspondência ao MIBiG facilita a priorização de famílias para caracterização funcional e estudos de expressão gênica, enquanto as GCFs órfãs (mais de 99% do total) representam alvos promissores para abordagens de prospecção de compostos novos por meio de técnicas de genômica, transcriptômica e metabolômica integradas.

Tabela 4. Famílias de clusters biossintéticos (GCFs) com correspondência ao banco de dados MIBiG identificadas nos 23 genomas de *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis*

Classe	GCF_ID	Total BGCs	BGCs MIBiG	BGCs MIBiG	Genomas	TAXO	Produto	Organismos MIBiG
PKSI	3199	13	3	BGC0001257.1; BGC0001258.1; BGC0001906.1	10	<i>Neopestalotiopsis_sp_PM2101</i> ; <i>Pestalotiopsis_sp_JCM9685</i> ; <i>Pestalotiopsis_sp_9143B</i> ; <i>Pestalotiopsis_sp_CR013</i> ; <i>Pestalotiopsis_kenyana</i> ; <i>Pseudopestalotiopsis_sp_12</i> ; <i>Pestalotiopsis_sp_DP2020a</i> <i>Pestalotiopsis_sp_MUOB44551</i> ; <i>Pestalotiopsis_sp_LF1912</i> ; <i>Pestalotiopsis_sp_NC0098</i> ;	T1PKS	<i>Nodulisporium sp. ATCC74245</i> ; <i>Glarea lozoyensis</i> ; <i>Daldinia eschschoizii IFB-TL01</i>
PKSI	3482	6	1	BGC0000121.1	5	<i>Pseudopestalotiopsis_chinensis</i> ; <i>Pseudopestalotiopsis_fici_W1061</i> ; <i>Pseudopestalotiopsis_theae</i> ; <i>Pseudopestalotiopsis_sp_12</i> ; <i>Pseudopestalotiopsis_sp_126</i>	T1PKS	<i>Pestalotiopsis_fici</i>

Others	3705	5	1	BGC0001624.1	4	<i>Pseudoestalotipsis_chinensis</i> ; <i>Pseudoestalotipsis_fici_W1061</i> ; <i>Pseudoestalotipsis_theae</i> ; <i>Pseudoestalotipsis_sp_126</i> ;	Indole iso- A82775C	<i>Pestalotipsis_fici</i>
NRPS	3843	11	1	BGC002248	10	<i>Neopestalotipsis sp PM21011.1</i> ; <i>Neopestalotipsis sp. PM2101E</i> ; <i>Neopestalotipsis sp. Açai</i> ; <i>Neopestalotipsis sp. 37M</i> ; <i>Neopestalotipsis sp. M2102B</i> ; <i>Neopestalotipsis rosae ML 16641.1</i> ; <i>Pestalotipsis sp. CR013</i> ; <i>Pestalotipsis sp. NC0098</i> ; <i>Pestalotipsis sp. MUOB 44551</i> ; <i>Pestalotipsis kenyana pg52</i>	NRPS/Ter pno flavunoidin e	<i>Aspergillus falvus NRRL3357</i>

5.4 Caracterização de redes moleculares

A comparação entre as topologias das redes de PKS e NRPS revela padrões evolutivos fundamentalmente distintos na organização dos clusters biossintéticos destes fungos. As redes de PKS apresentam dois padrões predominantes: (1) clusters exclusivos de um único gênero, ou (2) famílias gênicas com mistura dos três gêneros, onde nodos laranjas (*Pseudopestalotiopsis*), azuis (*Neopestalotiopsis*) e verdes (*Pestalotiopsis*) aparecem interconectados na mesma família, indicando conservação evolutiva e compartilhamento destas vias biossintéticas.

Em contraste marcante, as redes de NRPS exibem um padrão de segregação intragênica: dentro de uma mesma família gênica, os nodos se agrupam de forma gênero-específica, formando sub-redes distintas (Figura 10A). Ou seja, embora pertençam à mesma família de NRPS e compartilhem ancestralidade comum, cada gênero desenvolveu variantes próprias destes clusters, com os nodos laranjas, azuis e amarelos/verdes formando agrupamentos separados dentro da rede maior. Esta segregação intrafamiliar sugere evolução divergente acelerada nas NRPS, onde cada linhagem modificou e especializou seu repertório biossintético de forma independente, mesmo mantendo o arcabouço gênico básico.

É relevante o fato de que a vasta maioria destas famílias de NRPS que apresentam segregação por gênero não possuem similaridade com BGCs caracterizados, indicando que representam vias biossintéticas completamente novas e não descritas. Este padrão contrasta com o cluster de flavunoidina, uma das poucas famílias onde foi possível estabelecer correspondência com um BGC conhecido, mas que mesmo assim mantém a segregação gênero-específica entre *Neopestalotiopsis* e *Pestalotiopsis*.

Esta dicotomia entre conservação em PKS versus diversificação em NRPS sugere que os peptídeos não-ribossomais representam um hotspot de inovação metabólica, com cada gênero explorando o espaço químico de forma única. O potencial inexplorado desta diversidade química críptica representa uma fronteira promissora para a bioprospecção de compostos com atividades farmacológicas e agrônômicas inovadoras.

A análise da rede de similaridade de sequências de genes relacionados a policetídeo sintases (Figura 9A) revelou uma distribuição diferencial entre os três gêneros avaliados. Notavelmente, a família de genes associada à biossíntese do ácido pesteico (destacada na rede delimitada) foi identificada exclusivamente em isolados de *Pseudopestalotiopsis* (nodos laranjas), sugerindo uma restrição filogenética deste cluster biossintético a este gênero específico.

A análise de sintenia (Figura 9B) demonstrou alta conservação estrutural do cluster de ácido pesteico entre seis linhagens de *Pseudopestalotiopsis* analisadas (1.2, AH-3-21, CYF271, 1.26, W10611, e *P. fici* BGC0000121). O cluster caracterizado, composto por 13 genes essenciais (ptaA-ptaM) e três reguladores transcricionais (ptaR1-ptaR3), apresentou arquitetura gênica altamente sintênica entre todas as linhagens avaliadas. A presença consistente de genes de *tailoring* (ptaH, ptaI, ptaJ) e metiltransferases (ptaK, ptaL) sugere a conservação do potencial biossintético completo para produção de ácido pesteico neste gênero.

Já a rede de similaridade de sequências de NRPS (Figura 10A) revelou um padrão de distribuição contrastante em relação aos clusters de PKS. A família gênica associada à biossíntese de flavunoidina (painel ampliado) foi identificada predominantemente em espécies de *Neopestalotiopsis* (nodos azuis) e *Pestalotiopsis* (nodos laranjas), com ausência conspícua em *Pseudopestalotiopsis* (nodos verdes). Esta segregação taxonômica sugere eventos evolutivos distintos na aquisição ou perda deste cluster biossintético entre os três gêneros.

A análise de sintenia comparativa (Figura 10B) entre os BGCs tipo-flavunoidina identificados em pestalotiídeos e o cluster caracterizado em *Aspergillus flavus* NRRL3357 (BGC0002248) revelou características notáveis. Dez linhagens de *Neopestalotiopsis* (*N. sp.* açai, *N. sp.* 37M, *N. sp.* M2102B1, *N. rosae* ML16641.1, *N. sp.* PM2101E1.1, *N. sp.* PM21011.1) e quatro linhagens de *Pestalotiopsis* (*P. sp.* NC00981.1, *P. sp.* M4405151, *P. kenya* PG52, *P. sp.* CR0131) apresentaram clusters com arquitetura gênica parcialmente conservada em relação ao BGC de referência.

Particularmente relevante é a presença de genes de *tailoring* adicionais (destacados com bordas vermelhas) nos clusters de pestalotiídeos que estão

ausentes no BGC de *A. flavus*, tais como metiltransferases, p450 e hidrolases que são conhecidas por fazerem adiões de grupos funcionais em metabólitos secundários. Esta expansão do repertório enzimático sugere que os pestalotiídes podem produzir derivados estruturais de flavunoidina com modificações pós-biossintéticas adicionais, potencialmente gerando uma diversidade química ampliada em relação ao composto descrito em *Aspergillus*.

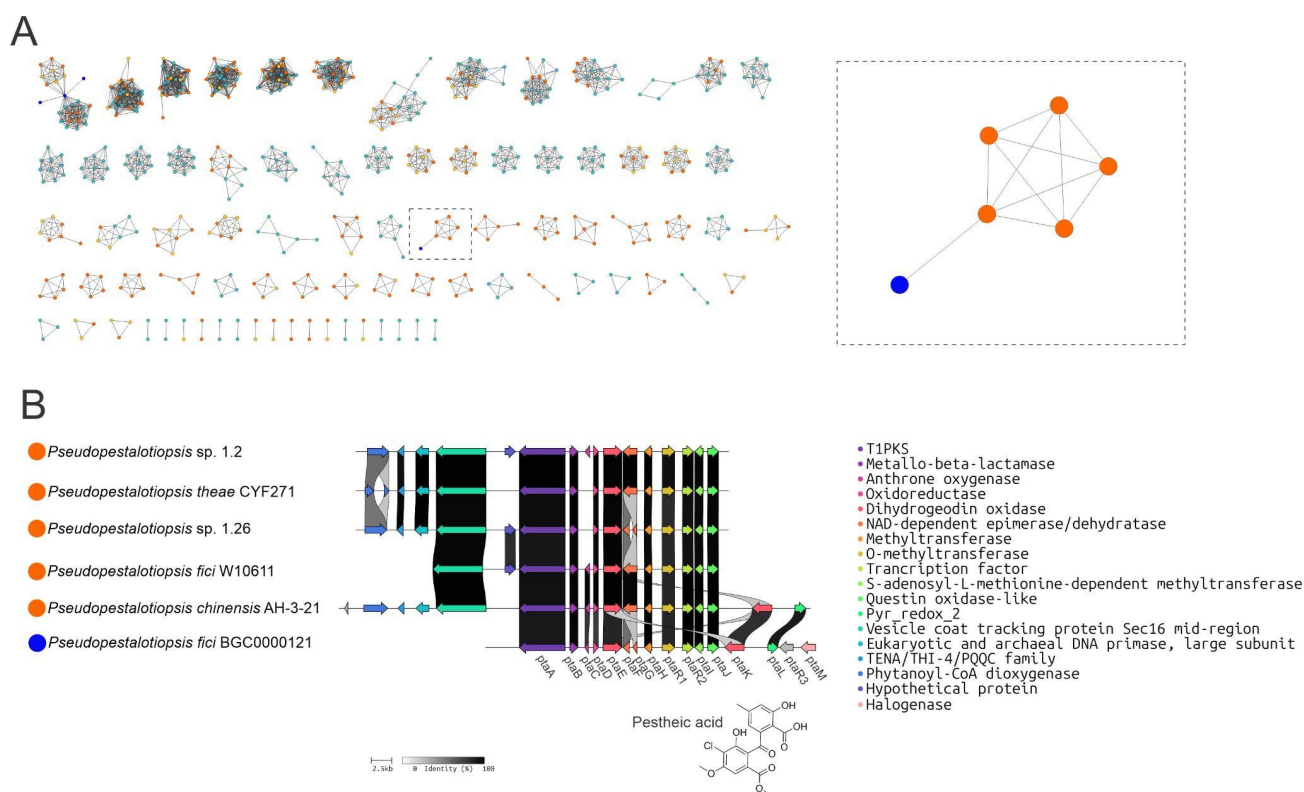


Figura 9. Rede de similaridade de seqüências e análise de sintenia de clusters gênicos biossintéticos de policetídeo sintases (PKS) em genomas de *Neopestalotiopsis*, *Pestalotiopsis* e *Pseudopestalotiopsis*. **A.** Rede de similaridade de seqüências mostrando famílias de genes relacionados a policetídeo sintases. Cada nó representa um gene, e as conexões indicam similaridade de seqüência. A rede delimitada (tracejada) destaca a família de genes associada à biossíntese do ácido pesteico. **B.** Análise de sintenia comparativa entre clusters gênicos biossintéticos identificados em diferentes espécies de *Pseudopestalotiopsis* e o BGC caracterizado para produção de ácido pesteico (*P. fici* BGC0000121). Genes homólogos são

conectados por regiões sombreadas, e a orientação das setas indica a direção transcricional.

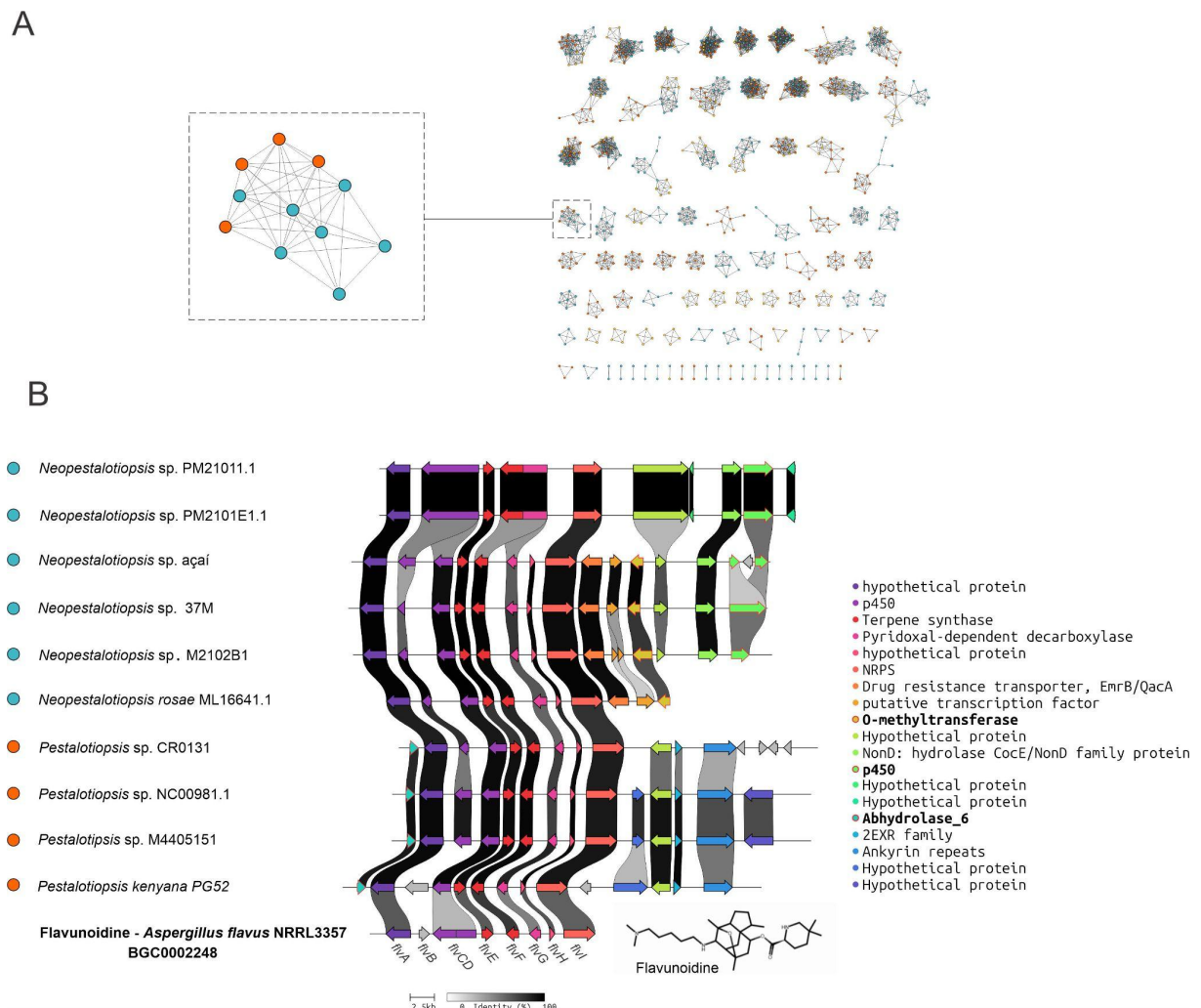


Figura 10: (A) Rede de similaridade de sequências mostrando famílias de genes relacionados a NRPS identificados nos genomas analisados. Cada nó representa um gene, e as conexões indicam similaridade de sequência. O painel ampliado (tracejado) destaca a família gênica associada à biossíntese de flavunoidina, com nós laranjas representando genes de *Pestalotiopsis* spp. e nós azuis representando genes de *Neopestalotiopsis* spp. (B) Análise de sintenia comparativa entre clusters gênicos biossintéticos identificados em diferentes espécies de pestalotiídes e o BGC caracterizado para produção de flavunoidina em *Aspergillus flavus* NRRL3357 (BGC0002248). As regiões sombreadas conectam genes homólogos entre os clusters, e a orientação das setas indica a direção transcricional. A intensidade do sombreamento reflete o percentual de identidade entre as sequências (escala de 0-100%). Genes destacados com bordas vermelhas correspondem a *tailoring genes*

ausentes no BGC de referência, indicando possíveis modificações estruturais adicionais nas moléculas produzidas por esses clusters.

6. CONCLUSÃO

Este estudo de mineração genômica revelou a diversidade biossintética de metabólitos secundários nos fungos dos gêneros *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis*, com base na análise de 23 genomas. Os resultados obtidos consolidam o potencial desses microrganismos como fontes promissoras de produtos naturais e fornecem uma base robusta para futuras investigações biotecnológicas. Abaixo apresentamos as conclusões com base em cada objetivo específico:

6.1 Diversidade metabolômica dos gêneros *Pestalotiopsis*, *Pseudopestalotiopsis* e *Neopestalotiopsis*

A análise demonstrou uma expressiva diversidade metabolômica entre os gêneros estudados. Foram identificados um total de 2.113 Clusters de Genes Biossintéticos (BGCs) distribuídos em sete classes principais, com predominância notável de NRPS (795 BGCs, 37,6%) e PKS (607 BGCs, 28,7%). Essa abundância sugere um elevado potencial para a produção de peptídeos não ribossomais e policetídeos, compostos conhecidos por suas diversas atividades biológicas. Observou-se que *Neopestalotiopsis* apresentou o maior número de BGCs (989), indicando uma maior versatilidade biossintética em comparação a *Pestalotiopsis* (666 BGCs) e *Pseudopestalotiopsis* (458 BGCs), com linhagens como *Neopestalotiopsis* sp. 1.5 (111 BGCs) e *Neopestalotiopsis rosae* ML1664 (104 BGCs) destacando-se na capacidade de geração de metabólitos secundários.

6.2 Identificação e agrupamento de BGCs em Famílias de Clusters de Genes (GCFs)

A aplicação das ferramentas antiSMASH e BiG-SCAPE permitiu a identificação e agrupamento dos BGCs em 632 Famílias de Clusters de Genes (GCFs). A distribuição dessas GCFs revelou 165 de NRPS, 141 de PKS, 168 de Others, 63 de Terpene, 47 de PKSother, 36 de PKS-NRP Hybrids e 12 de RiPPs. A análise do compartilhamento indicou uma predominância de GCFs exclusivas ("unique") e singletons, refletindo uma alta especialização metabólica e baixa conservação das vias biossintéticas entre os gêneros. Apenas 22 GCFs (3,5% do total) foram classificadas como "core", compartilhadas pelos três gêneros. As redes de similaridade também evidenciaram padrões evolutivos distintos, com as redes de

NRPS apresentando uma segregação gênero-específica, sugerindo uma diversificação acelerada e inovação metabólica para esta classe.

6.3 Conexão entre BGCs e Produtos Naturais Caracterizados

A comparação das 632 GCFs com o banco de dados MIBiG identificou **4 famílias (0,63%)** com correspondência a clusters caracterizados experimentalmente, envolvendo 19 BGCs dos genomas estudados (0,9% do total). As GCFs identificadas e seus respectivos produtos caracterizados são:

1- GCF 3199 (classe PKS1): Esta família, que agrupa 13 BGCs, inclui homologia com clusters previamente caracterizados em *Nodulisporium sp.*, *Glarea lozoyensis* (BGC0001257.1, BGC0001258.1) e *Daldinia eschscholzii* IFB-TL01 (BGC0001906.1). No contexto de *D. eschscholzii*, o cluster está associado à biossíntese dos imunossupressores dalmanol A e acetodalmanol A, cuja produção envolve o acoplamento de dois aglomerados gênicos de policetídeos (PGCs) distintos (naphthalene-encoding e chromane-encoding) via um mecanismo catalisado por uma monooxigenase inespecífica (ZHOU et al., 2019).

2- GCF 3482 (classe PKS1): Esta família agrupa seis BGCs, sendo cinco provenientes dos genomas analisados e um do MIBiG (BGC0000121.1) caracterizado em *Pestalotiopsis fici*. Este cluster é homologado à biossíntese do ácido pestheic, um regulador de crescimento vegetal (XU et al., 2014).

3-GCF 3705 (classe Others): Esta família reúne cinco BGCs e apresenta correspondência com o cluster MIBiG BGC0001624.1, caracterizado em *Pestalotiopsis fici*. Este cluster está envolvido na biossíntese do iso-A82775C, um prenilepoxiciclohexano, e de cloropestolides (como H-K). Compostos como cloropupukeananina e cloropestolides A-D, co-isolados do mesmo fungo, já demonstraram atividades antimicrobianas, antitumorais e anti-HIV (PAN et al., 2018).

4- GCF 3843 (classe NRPS): Esta família foi associada ao cluster caracterizado em *Aspergillus flavus* NRRL3357 (BGC0002248) para a biossíntese de flavunoidina. A presença de genes de *tailoring* adicionais nos clusters de pestalotiídes homologados a essa família sugere a capacidade de produzir derivados estruturais de flavunoidina com modificações pós-biossintéticas, potencialmente gerando uma diversidade química ampliada em relação ao composto descrito em *Aspergillus* (Figura 10).

Esta baixa proporção de correspondência com o MIBiG é um achado de grande impacto, indicando que mais de 99% das GCFs identificadas representam "clusters órfãos", sem rotas metabólicas caracterizadas ou produtos conhecidos na literatura científica. Esses clusters inexplorados representam um vasto potencial para a descoberta de novos metabólitos secundários com estruturas químicas inéditas e potenciais aplicações biotecnológicas e farmacêuticas.

7. Perspectivas Futuras

Os resultados desta dissertação abrem múltiplas avenidas para futuras investigações, visando desvendar plenamente o potencial biotecnológico dos fungos estudados:

1-Validação experimental e caracterização funcional: priorizar a validação laboratorial dos "clusters órfãos" identificados. Isso pode envolver estudos de genômica funcional para ativar BGCs silenciosos, empregando técnicas de edição gênica ou engenharia de cepas para superprodução.

2-Integração ômica: combinar abordagens de genômica, transcriptômica e metabolômica para estabelecer a ligação direta entre os BGCs previstos e os produtos naturais reais. A análise metabolômica aprofundada pode revelar a estrutura química dos metabólitos codificados pelos clusters órfãos.

3-Estudos químicos e farmacológicos: uma vez que novos metabólitos sejam isolados, caracterizar suas estruturas químicas detalhadamente e investigar suas atividades biológicas em ensaios *in vitro* e *in vivo*, com foco em aplicações antimicrobianas, antitumorais, imunomoduladoras ou agrícolas.

4-Engenharia biossintética: utilizar plataformas de expressão heteróloga para produzir metabólitos de interesse em sistemas microbianos mais manejáveis, permitindo a otimização da produção e a geração de análogos semi-sintéticos.

5-Análise filogenética e ecológica aprofundada: expandir a análise filogenômica para incluir um número maior de linhagens, correlacionando a presença de GCFs únicas e acessórias com nichos ecológicos específicos e padrões de adaptação. Aprofundar a

investigação dos isolados amazônicos pode revelar metabólitos com características únicas relacionadas à biodiversidade da região.

6-Exploração de mecanismos de regulação: investigar os mecanismos regulatórios que controlam a expressão dos BGCs, incluindo fatores epigenéticos e ambientais, para desenvolver estratégias que maximizem a produção de metabólitos secundários em condições de laboratório.

7. CRONOGRAMA DE ATIVIDADES

ATIVIDADE	MESES (2025)											
	J	F	M	A	M	J	J	A	S	O	N	D
Revisão de literatura	X	X	X	X	X	X	X	X	X	X	X	X
Aquisição de genomas públicos (NCBI e JGI)	X	X										
Obtenção dos 6 genomas (Embrapa Amazônia Ocidental)	X	X										
Processamento dos genomas no <i>antiSMASH</i>		X	X	X	X	X	X	X				
Instalação e testes <i>BiG-SCAPE</i>		X	X	X	X	X	X	X				
Execução das análises no <i>BiG-SCAPE</i>		X	X	X	X	X	X	X				
Desenvolvimento dos algoritmos (Python + Biopython)		X	X	X	X	X	X	X	X			
Tratamento de dados e visualização (Python, Matplotlib, Seaborn)		X	X	X	X	X	X	X	X	X		
Interpretação dos resultados e redação da dissertação		X	X	X	X	X	X	X	X	X		
Qualificação e Defesa da Dissertação											X	X

REFERÊNCIAS

1. ALAS, I. et al. Micromonosporaceae biosynthetic gene cluster diversity highlights the need for broad-spectrum investigations. *Microbial Genomics*, v. 10, n. 1, p. 001167, 2024.
2. AL-SIYABI, N. et al. Genomic insights into biosynthetic gene cluster diversity and structural variability in marine bacteria. *Scientific Reports*, v. 15, n. 1, p. 37644, 2025.
3. ARAD, N. et al. Biosynthetic potential of the culturable foliar fungi associated with field-grown lettuce. *Applied Microbiology and Biotechnology*, v. 109, n. 1, p. 197, 2025.
4. ARIYAWANSA, H. A. et al. Fungal diversity notes 111–252—taxonomic and phylogenetic contributions to fungal taxa. *Fungal Divers*, v. 75, p. 27–274, 2015.
5. AUGUSTIJN, H. E. et al. Genome mining based on transcriptional regulatory elements reveals a new class of desferrioxamine-like siderophores. *ACS Chemical Biology*, 2025. (No prelo).
6. BLIN, K. et al. antiSMASH 5.0: updates, new features, and a major overhaul for secondary metabolite genome mining. *Nucleic Acids Research*, v. 47, n. W1, p. W81-W87, 2019.
7. BLIN, K. et al. antiSMASH 8.0: extended gene cluster detection capabilities and analyses of chemistry, enzymology, and regulation. *Nucleic Acids Research*, p. gkaf334, 2025.
8. CEDIEL-BECERRA, J. D. D. et al. Targeted genome mining with GATOR-GC maps the biosynthetic potential of microbial genomes. *Nucleic Acids Research*, v. 53, n. 13, p. gkaf606, 2025.
9. CHEN, Yingjuan et al. *Pestalotiopsis*-like species causing gray blight disease on *Camellia sinensis* in China. *Plant Disease*, v. 102, n. 1, p. 98-106, 2018.
10. DE CARVALHO, L. C. et al. Unveiling Microbial Chemical Interactions Based on Metabolomics Approaches. *Advances in Experimental Medicine and Biology*, v. 1439, p. 51–70, 2023.
11. DRAISMA, A. A. et al. BiG-SCAPE 2.0 and BiG-SLiCE 2.0: scalable, accurate and interactive gene cluster analyses. *bioRxiv*, 2025.08.20.671210, 2025.

12. ELKHATEEB, Waill A.; DABA, Ghoson M. The endophytic fungi *Pestalotiopsis* what's for it and what's on it. *Journal of Pharmaceutics and Pharmacology Research*, v. 4, n. 1, p. 1-5, 2021.
13. FIGUEIREDO, G. et al. The Unexplored Wealth of Microbial Secondary Metabolites: the Sphingobacteriaceae Case Study. *Microbial Ecology*, v. 83, n. 2, p. 470–481, 2022.
14. GUALBERTO, Gilvana F. et al. *Pseudopestalotiopsis gilvanii* sp. nov. and *Neopestalotiopsis formicarum* leaves spot pathogens from guarana plant: a new threat to global tropical hosts. *Phytotaxa*, v. 489, n. 2, p. 121-139, 2021.
15. GUZMÁN-CHÁVEZ, F. et al. Genome mining for natural products: a historical perspective. *Natural Product Reports*, v. 35, n. 1, p. 19-35, 2018.
16. GUZMÁN-CHÁVEZ, Fernando et al. Engineering of the filamentous fungus *Penicillium chrysogenum* as cell factory for natural products. *Frontiers in Microbiology*, v. 9, p. 2768, 2018.
17. HAUTBERGUE, G. et al. Genome mining: a powerful tool for the discovery of novel natural products. *Natural Product Reports*, v. 35, n. 1, p. 1-18, 2018.
18. HAUTBERGUE, Thaïs et al. From genomics to metabolomics, moving toward an integrated strategy for the discovery of fungal secondary metabolites. *Natural Product Reports*, v. 35, n. 2, p. 147-173, 2018.
19. HOOFT, J. J. J. VAN DER et al. Linking genomics and metabolomics to chart specialized metabolic diversity. *Chem. Soc. Rev.*, v. 49, p. 3297–3314, 2020. doi: 10.1039/d0cs00162g.
20. JAARSMA, A. H. et al. The undiscovered biosynthetic potential of the Greenland Ice Sheet microbiome. *Frontiers in Microbiology*, v. 14, p. 1285791, 2023.
21. KAUTSAR, S. A. et al. BiG-SLiCE: A highly scalable tool maps the diversity of 1.2 million biosynthetic gene clusters. *Gigascience*, v. 10, n. 1, p. giaa154, 2021.
22. KAUTSAR, Satria A. et al. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic acids research*, v. 48, n. D1, p. D454-D458, 2020.
23. KIM, B. et al. Insights into group-specific pattern of secondary metabolite gene cluster in *Burkholderia* genus. *Frontiers in Microbiology*, v. 14, p. 1302236, 2024.
24. KUHNERT, E. et al. Secondary metabolite biosynthetic diversity in the fungal family Hypoxylaceae and *Xylaria hypoxylon*. *Studies in Mycology*, v. 99, p. 100118, 2021.

25. LEÃO, T. et al. A Multi-Omics Characterization of the Natural Product Potential of Tropical Filamentous Marine Cyanobacteria. *Marine Drugs*, v. 19, n. 1, p. 20, 2021.
26. LEI, H. et al. New cytotoxic natural products from the marine sponge-derived fungus *Pestalotiopsis* sp. by epigenetic modification. *RSC Advances*, v. 10, p. 37982-37988, 2020.
27. LIN, Zhenquan; NIELSEN, Jens; LIU, Zihe. Bioprospecting through cloning of whole natural product biosynthetic gene clusters. *Frontiers in Bioengineering and Biotechnology*, v. 8, p. 526, 2020.
28. LIU, Gaoran et al. New bioactive sesquiterpenoids from the plant endophytic fungus *Pestalotiopsis theae*. *Frontiers in Microbiology*, v. 12, p. 641504, 2021.
29. LIU, S. et al. Biosynthetic potential of the gut microbiome in longevous organisms. *Gut Microbes*, v. 16, n. 1, p. 2426623, 2024.
30. LOPES, Fernanda Cortez. *Produção e análise de metabólitos secundários de fungos filamentosos*. Dissertação (Mestrado em Ciências) – Universidade Federal de Minas Gerais, Belo Horizonte, 2011.
31. MAHARACHCHIKUMBURA, S. S. N. et al. *Pestalotiopsis* species associated with *Camellia sinensis* (tea). *Mycotaxon*, v. 123, p. 47–61, 2013.
32. MEDEMA, M. H. et al. Minimum Information about a Biosynthetic Gene cluster (MIBiG) project. *Nature Chemical Biology*, v. 11, n. 9, p. 625-631, 2015.
33. MULLIS, Megan M. et al. Diversity, ecology, and prevalence of antimicrobials in nature. *Frontiers in Microbiology*, v. 10, p. 2518, 2019.
34. NAVARRO-MUÑOZ, J. C. et al. A computational framework to explore large-scale biosynthetic diversity. *Nature Chemical Biology*, v. 16, n. 1, p. 60–68, 2020.
35. PALAZZOTTO, E.; WEBER, T. Omics and multi-omics approaches to study the biosynthesis of secondary metabolites in microorganisms. *Current Opinion in Microbiology*, v. 45, p. 109–116, 2018. doi: 10.1016/j.mib.2018.03.004.
36. PAULO, Bruno S.; SIGRIST, Renata; OLIVEIRA, Luciana G. de. Avanços recentes em biossíntese combinatória de policetídeos: perspectivas e desafios. *Química Nova*, v. 42, n. 1, p. 71-83, 2019.
37. PAN, Y. et al. Characterization of a Prenyltransferase for Iso-A82775C Biosynthesis and Generation of New Congeners of Chloropestolides. **ACS**

- Chemical Biology**, Washington, v. 13, n. 3, p. 703-711, 31 jan. 2018. DOI: 10.1021/acscchembio.7b01059.
38. ROKAS, A. et al. The biosynthetic gene cluster concept in the age of genomics. *Nature Chemical Biology*, v. 16, n. 1, p. 1–3, 2020.
 39. ROKAS, Antonis; WISECAVER, Jennifer H.; LIND, Abigail L. The birth, evolution and death of metabolic gene clusters in fungi. *Nature Reviews Microbiology*, v. 16, n. 12, p. 731-744, 2018.
 40. RUSSELL, A. H.; TRUMAN, A. W. Genome mining strategies for ribosomally synthesised and post-translationally modified peptides. *Computational and Structural Biotechnology Journal*, v. 18, p. 1838–1851, 2020. doi: 10.1016/j.csbj.2020.06.032.
 41. SAATI-SANTAMARÍA, Z. et al. Unveiling the genomic potential of *Pseudomonas* type strains for discovering new natural products. *Microbial Genomics*, v. 8, n. 2, p. 000758, 2022.
 42. SALOMON, C. E.; MAGARVEY, N. A.; SHERMAN, D. H. Merging the potential of microbial genetics with biological and chemical diversity: an even brighter future for marine natural product drug discovery. *Natural Product Reports*, v. 21, n. 1, p. 105–121, 2004. doi: 10.1039/b301384g.
 43. SEMMLER, F. et al. The Metabolic Potential of the Human Lung Microbiome. *Microorganisms*, v. 12, n. 7, p. 1448, 2024.
 44. SERNA-CARDONA, N. et al. Unveiling metabolo-genomic insights of potent antitumoral and antibiotic activity in *Streptomyces* sp. VB1 from Valparaíso Bay. *Frontiers in Microbiology*, v. 15, p. 1463911, 2024.
 45. SERRANO, S. et al. Haloarchaea have a high genomic diversity for the biosynthesis of carotenoids of biotechnological interest. *Research in Microbiology*, v. 173, n. 3, p. 103919, 2022.
 46. SHEN, B. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Current Opinion in Chemical Biology*, v. 7, n. 2, p. 285–295, 2003. doi: 10.1016/S1367-5931(03)00020-6.
 47. SIMUNOVIĆ, V.; GRUBIŠIĆ, I. Amino acid (acyl carrier protein) ligase-associated biosynthetic gene clusters reveal unexplored biosynthetic potential. *Molecular Genetics and Genomics*, v. 298, n. 1, p. 49–65, 2023.

48. SINGH, G. et al. Are there conserved biosynthetic genes in lichens? Genome-wide assessment of terpene biosynthetic genes suggests ubiquitous distribution of the squalene synthase cluster. *BMC Genomics*, v. 25, n. 1, p. 936, 2024.
49. SOLDATOU, S. et al. Linking biosynthetic and chemical space to accelerate microbial secondary metabolite discovery. *FEMS Microbiology Letters*, v. 366, n. 13, p. 1–8, 2019. doi: 10.1093/femsle/fnz142.
50. TERLOUW, Barbara R. et al. MIBiG 3.0: a community-driven effort to annotate experimentally validated biosynthetic gene clusters. *Nucleic acids research*, v. 51, n. D1, p. D603-D610, 2023.
51. UNDABARRENA, A. et al. *Rhodococcus* comparative genomics reveals a phylogenomic-dependent non-ribosomal peptide synthetase distribution: insights into biosynthetic gene cluster connection to an orphan metabolite. *Microbial Genomics*, v. 7, n. 7, p. 000621, 2021.
52. VAN DER HOOFT, J. J. et al. Molecular networking as a data analysis tool for integrating metabolomics and genomics. *Nature Protocols*, v. 15, n. 10, p. 3048-3071, 2020.
53. VILLA-RODRIGUEZ, E. et al. Integrated omics approaches for deciphering antifungal metabolites produced by a novel *Bacillus* species, *B. cabrialesii* TE3T, against the spot blotch disease of wheat (*Triticum turgidum* L. subsp. *durum*). *Microbiological Research*, v. 251, p. 126826, 2021.
54. WADHWA, Khushbu et al. A comprehensive review of the diversity of fungal secondary metabolites and their emerging applications in healthcare and environment. *Mycobiology*, v. 52, n. 6, p. 335-387, 2024.
55. WALSH, C. T.; TANG, Y. *Natural Product Biosynthesis: chemical logic and enzymatic machinery*. Cambridge: The Royal Society of Chemistry, 2017.
56. WANG, Jia et al. Biosynthesis of aromatic polyketides in microorganisms using type II polyketide synthases. *Microbial Cell Factories*, v. 19, p. 1-11, 2020.
57. WANG, Xiuna et al. Genomic and transcriptomic analysis of the endophytic fungus *Pestalotiopsis fici* reveals its lifestyle and high potential for synthesis of natural products. *BMC Genomics*, v. 16, p. 1-13, 2015.
58. WU, Chu; WANG, Yun; YANG, Yujie. *Pestalotiopsis* Diversity: Species, Dispositions, Secondary Metabolites, and Bioactivities. *Molecules*, v. 27, n. 22, p. 8088, 2022.

59. XIE, J. et al. Three phytotoxins produced by *Neopestalotiopsis clavispora*, the causal agent of ring spot on *Kadsura coccinea*. *Microbiological Research*, v. 238, p. 126531, 2020.
60. XU, Jing; EBADA, Sherif S.; PROKSCH, Peter. *Pestalotiopsis* a highly creative genus: chemistry and bioactivity of secondary metabolites. *Fungal Diversity*, v. 44, n. 1, p. 15, 2010.
61. YANG, X.-L.; ZHANG, J.-Z.; LUO, D.-Q. The taxonomy, biology and chemistry of the fungal *Pestalotiopsis* genus. *Nat. Prod. Rep.*, v. 29, p. 622–641, 2012.
62. YU, Xiaoqin et al. Polyketide derivatives from mangrove derived endophytic fungus *Pseudopestalotiopsis theae*. *Marine Drugs*, v. 18, n. 2, p. 129, 2020.
63. ZDOUC, Mitja M. et al. MIBiG 4.0: advancing biosynthetic gene cluster curation through global collaboration. *Nucleic acids research*, v. 53, n. D1, p. D678-D690, 2025.
64. ZHANG, Huawei et al. Cytotoxic Natural Products from Marine Sponge-Derived Microorganisms. *Marine Drugs*, v. 15, 2017.
65. ZHANG, Qian et al. A MYST histone acetyltransferase modulates conidia development and secondary metabolism in *Pestalotiopsis microspora*, a taxol producer. *Scientific Reports*, v. 8, n. 1, p. 8199, 2018.
66. ZHOU, Z. Z. et al. Dalmanol biosyntheses require coupling of two separate polyketide gene clusters. **Chemical Science**, v. 10, p. 73-82, 2019. DOI: 10.1039/c8sc03697g.
67. ZIEMERT, N.; WEBER, T.; MEDEMA, M. H. Genome Mining Approaches to Bacterial Natural Product Discovery. In: *Comprehensive Natural Products III*. 3. ed. Elsevier Inc., 2020. doi: 10.1016/b978-0-12-409547-2.14627-x.