UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CENTRO DE BIOTECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

ALEXANDRA DE AZEVEDO DA ROCHA

"AGRUPAMENTOS GÊNICOS CONSERVADOS EM FUNGOS FILAMENTOSOS ENTOMOPATOGÊNICOS"

Porto Alegre
Março, 2024

ALEXANDRA DE AZEVEDO DA ROCHA

“AGRUPAMENTOS GÊNICOS CONSERVADOS EM FUNGOS FILAMENTOSOS ENTOMOPATOGÊNICOS”

Dissertação submetida ao Programa de Pós-Graduação em Biologia Celular e Molecular do Centro de Biotecnologia da UFRGS como requisito parcial para a obtenção do título de Mestre.

Orientador: Prof. Dr. Charley Christian Staats.

Porto Alegre

Março, 2024

CIP - Catalogação na publicação

Este trabalho foi desenvolvido no Laboratório de Microbiologia Molecular e Celular (LMMC), situado no Centro de Biotecnologia da Universidade Federal do Rio Grande do Sul. Os recursos financeiros para financiamento do projeto foram oriundos da CAPES, CNPq e FAPERGS.

Dedico este trabalho para todos que acreditaram em mim, mesmo nos momentos em que nem eu mesma acreditava. Àqueles que, não só me deixam sonhar alto, como me apoiam até ter cada sonho realizado. Nada seria possível sem vocês.

AGRADECIMENTOS

Primeiramente agradeço a Deus por me fazer tão sonhadora e permitir realizar os sonhos mais profundos que tenho desde muito nova, como a obtenção do título de mestre nesta Universidade. Também agradeço ao meu orientador Prof. Dr. Charley Christian Staats por todos os ensinamentos, por acreditar no meu potencial e aceitar me orientar, todo apoio foi fundamental para o meu desenvolvimento pessoal e acadêmico dentro do PPGBCM - programa que me orgulha fazer parte de minha trajetória.

Agradeço aos meus pais Elisandra e Alexandre pelo apoio constante, e, em especial a minha tia Eva, minha segunda mãe, que mesmo sem entender direito o que todo o mestrado significa, nunca deixou de me apoiar e estar presente em todos os momentos. Aos meus irmãos Lucas, Matheus e Samuel por me incentivarem a cada dia ser uma pessoa melhor e irmã mais velha exemplar. Ao meu padrasto Robison e minha madrasta Valéria por toda disposição e apoio de diferentes maneiras. Ao restante da minha família que sempre torcem por mim.

Ao Lucas Braun, meu noivo, meu incentivador incansável, que foi quem me encorajou a me inscrever neste programa de mestrado, que em todos os problemas encontrou uma solução, que acredita em mim quando ninguém mais acredita, nem mesmo eu. Não existem palavras para agradecer e descrever o quão gratificante é tê-lo comigo e realizar sonhos com a melhor pessoa que eu poderia ter ao meu lado. Eu te amo!

Aos meus cachorros, companheiros diários e grandes amores, meus sinceros agradecimentos. Quando tudo parecia dar errado, poder levantar e ir abraçar qualquer um de vocês, foi como uma terapia. Obrigada por serem exatamente como são.

Agradeço a todo apoio de minhas melhores amigas. A Geovana Nardin, por sonhar este sonho comigo desde o ensino fundamental e ser uma enorme incentivadora. A Talita Pelissioli, por sempre acreditar no meu potencial e ser um porto seguro diário desde a graduação. Vocês significam muito em minha vida, gratidão por estarmos passando por mais esta etapa importante juntas.

Aos meus colegas do laboratório 219 por me acolherem e ajudarem desde o primeiro dia, destacando a Maria Eduarda e a Mauren - vocês duas são um presente na minha vida que levarei para sempre comigo. Ao João Vitor, do laboratório 220, que no final do mestrado apareceu como um amigo fundamental, obrigada! Aos demais colegas dos laboratórios ao redor que de alguma forma compartilharam seus conhecimentos - aprendi que não se faz ciência sem troca de experiências.

A minha cunhada Laura Braun por sempre demonstrar interesse e perguntar como andavam as coisas, tenho grande admiração por ti e sou grata por ter em minha cunhada, também uma amiga.

Aos demais amigos e colegas, que mesmo não mencionados, de alguma forma auxiliaram e/ou torceram por mim, meus sinceros agradecimentos. A todos os professores que compartilharam seus ensinamentos durante a pós-graduação, em especial ao Prof. Dr. Augusto Schrank por, de maneira inesperada, me falar as palavras certas no momento em que eu mais precisava.

Sou muito grata por toda caminhada realizada até aqui e por resultar em mais um sonho realizado. Do fundo do meu coração, muito obrigada!

SUMÁRIO

LISTA DE ABREVIATURAS, SÍMBOLOS E UNIDADES 8
LISTA DE FIGURAS 9
Resumo 10
Abstract 11
1. INTRODUÇÃO 12
1.1. Fungos filamentosos entomopatogênicos 13
1.2. Fungos entomopatogênicos dos gêneros Metarhizium e Beauveria 16
1.3. Ciclo de infecção 18
1.4. Metabólitos secundários de fungos filamentosos 21
1.5. Clusters gênicos biossintéticos (biosynthetic gene clusters (BGCs)) 22
2. OBJETIVOS 25
2.1. Objetivo Geral 25
2.2. Objetivos Específicos 25
3. RESULTADOS 26
4. DISCUSSÃO GERAL 57
5. CONCLUSÕES 60
6. PERSPECTIVAS 61
7. REFERÊNCIAS BIBLIOGRÁFICAS 62
8. CURRICULUM VITAE 66

LISTA DE ABREVIATURAS, SÍMBOLOS E UNIDADES

PAP	Pesticidas Altamente Perigosos
FEP	Fungo Entomopatogênico
MP	Metabólito Primário
MS	Metabólito Secundário
PKS	Policetídeos sintases (*Polyketide Synthases*)
NRPS	Sintetases Peptídicas Não Ribossômicas (*Nonribosomal Peptide Synthetases*)
BGC	*Clusters* Gênicos Biossintéticos (*biosynthetic gene clusters*)
DNA	Ácido desoxirribonucleico
CB	Controle Biológico

LISTA DE TABELAS

Tabela 1: Resumo da biomassa total estimada para os grupos taxonômicos mais abundantes.……14

LISTA DE FIGURAS

Figura 1: Associações filogenéticas dos fungos entomopatogênicos (FEPs) selecionados do filo Ascomycota.……...……...……..16

Figura 2: Aspectos evolutivos de *Metarzhium* sp.…..…….....19

Figura 3: Etapas do processo de infecção de *Metarhizium*…......21

Figura 4: *Clusters* gênicos fúngicos......23

Resumo

Fungos filamentosos entomopatogênicos (FEPs) são espécies cosmopolitas que infectam artrópodes e podem estar associados a plantas. Devido à sua capacidade de controlar pragas, apresentam grande potencial biotecnológico. Algumas espécies de FEPs evoluíram para infectar artrópodes sem prejudicar as plantas com as quais se associam, o que se deve à aquisição de genes de virulência e à perda de genes fitopatogênicos. Neste contexto, alguns destes genes estão agrupados em *clusters* gênicos e seus produtos atuam ativamente e conjuntamente para biossíntese de metabólitos secundários (MSs), pequenas moléculas bioativas. Coletivamente, estes agrupamentos são denominados de *biosynthetic gene clusters* (BGCs). Considerando a biodiversidade fúngica e a potencial quimiodiversidade de MSs associados, objetivamos avaliar a identidade e conservação de BGCs em sequências genômicas de fungos presentes na ordem Hypocreales, visto a presença de fungos entomopatogênicos e fitopatogênicos. Foram recuperadas 338 sequências genômicas de espécies fúngicas do banco de dados NCBI Genome e os BGCs identificados com a ferramenta fungalSMASH. As famílias de BGCs - *biosynthetic gene families* (BGFs) - foram encontradas com a ferramenta BiG-SLiCE. Foram identificados 14.272 BGCs e 1.839 BGFs. Clusters gênicos conservados foram encontrados principalmente entre membros do mesmo gênero. No entanto, quando a comparação foi feita dentro da ordem como um todo, a maioria dos BGCs era órfão, destacando a grande diversidade de BGCs e a quimiodiversidade das espécies fúngicas. O estudo fornece evidências de potencial genético para a descoberta de novas moléculas com aplicações biotecnológicas. Análises refinadas de conservação são essenciais para uma melhor seleção de genes e BGCs para caracterização funcional.

Abstract

Filamentous entomopathogenic fungi (FEPs) present immense biotechnological potential due to their broad host range targeting arthropods and their potential application in pest control. Evolution within certain FEP lineages has led to specialization towards arthropod hosts, often accompanied by the acquisition of virulence genes and the loss of phytopathogenic capabilities. These adaptations are often encoded within clustered biosynthetic gene clusters (BGCs), responsible for the production of diverse secondary metabolites (SMs) with potent bioactivity. To explore the BGC landscape and potential chemodiversity within Hypocreales, encompassing both entomopathogenic and phytopathogenic members, we analyzed 338 fungal genomic sequences retrieved from the NCBI Genome database. Leveraging the fungalSMASH and BiG-SLiCE tools, we identified an extensive repertoire of 14,272 BGCs and 1,839 biosynthetic gene families (BGFs). While conserved gene clusters were observed predominantly within genera, comparative analysis across the order revealed a remarkable abundance of orphan BGCs, highlighting the extensive chemodiversity and untapped potential within Hypocreales. This study provides compelling evidence for the immense genetic potential encoded within these fungi, paving the way for the discovery of novel biomolecules with promising biotechnological applications. However, to fully unlock this potential, refined conservation analyses are crucial for the targeted selection of BGCs and their associated genes for subsequent functional characterization.

1. INTRODUÇÃO

O uso de agrotóxicos resulta no aumento da produtividade na agricultura, mas o seu uso frequente acarreta consequências negativas não só para o meio ambiente como também para os animais e outros seres vivos. Os impactos já aparentes nos seres humanos vão desde sintomas mais simples até problemas mais graves, podendo chegar até no desenvolvimento de diferentes tipos de câncer (MORAES, 2019). Além disso, os impactos ambientais causados pelo uso desenfreado de agrotóxicos também são vários, desde contaminação (da água, do solo e das plantas), encolhimento no número de organismos vivos presentes em determinado ecossistema e o aumento da resistência de diferentes pragas (MORAES, 2019), sequelas preocupantes principalmente à longo prazo.

No Brasil, os pesticidas estão entre os agrotóxicos mais utilizados, sendo também o país que mais se destaca pela compra de Pesticidas Altamente Perigosos (PAP), conforme informado em 2022 pela Organização para a Alimentação e Agricultura (FAO) (F A O, 2022), sendo estes os pesticidas que acarretam maiores riscos, sobretudo ao meio ambiente. Atualmente a regulamentação destes pesticidas acontece através de três órgãos nacionais: Agência Nacional de Vigilância Sanitária (ANVISA), Instituto Brasileiro do Meio Ambiente e Recursos Naturais Renováveis (IBAMA) e Ministério da Agricultura, Pecuária e Abastecimento (MAPA), que juntos são responsáveis pela tripla análise e avaliam respectivamente os diferentes riscos à saúde, ao meio ambiente e riscos de eficiência agronômica que determinado produto pode causar (BRASIL, 2012).

Considerando todos os agravamentos ao meio ambiente, saúde humana e rede alimentar gerados pelo intenso uso de diferentes agrotóxicos nocivos, o Controle Biológico de pragas pode ser utilizado como uma maneira para diminuir o uso indisciplinado de PAPs e reduzir os impactos ambientais comumente gerados por esses produtos fitossanitários (PARRA, et al., 2021). O Controle Biológico (CB) teve termo proposto no início do século XX por Harry Smith, e surgiu como uma forma de “revezamento” em relação ao uso de métodos químicos, já que este controle utiliza micro patógenos como matéria prima (ECKSTEIN et al., 2022). Atualmente, os biopesticidas desenvolvidos por fungos entomopatogênicos apresentam números significativos no mercado mundial da agricultura, principalmente para o controle de insetos-praga, trazendo um cenário otimista considerando que a utilização de fungos como inseticida beneficia não somente o meio ambiente e a saúde em geral, por ser menos nocivo, como também o produtor (já que gera menos resíduos, aumentando o rendimento da produção).

1.1. Fungos filamentosos entomopatogênicos

Diferentes estudos já focam em dados sobre a biomassa da biosfera, e esses números são importantes para que se entenda a estrutura e como funciona a dinâmica do local em que vivemos. É de senso comum que os diferentes reinos da vida estão distribuídos em distintos locais da Terra, mas é curioso perceber que ao analisar os resultados do ranking dos reinos com maiores biomassas no nosso planeta, o reino Animal está longe do primeiro lugar (BAR-ON; PHILLIPS; MILO, 2018). Na **Tabela 1** é possível visualizar o valor de biomassa que cada táxon apresenta, seguido do valor da porcentagem que representa, conforme estudo realizado em 2018 por Bar-On et al.

Tabela 1: Resumo da biomassa total estimada para os grupos taxonômicos mais abundantes. Nesta visualização, um cubo representa 1 milhão de toneladas métricas de carbono e a cada mil cubos, se tem um Gigaton (Gt C). Adaptado de Bar-On et al (BAR-ON; PHILLIPS; MILO, 2018).

Táxon	**Massa (Gt C)**	**% total**
Plantas	450	82.4%
Bactérias	70	12.8%
Fungos	12	2.2%
Archaea	7	1.3%
Protistas	4	0.70%
Animais	2	0.47%
Vírus	0.2	0.04%
Total	**545.8**	**100.0%**

Os fungos, assim como os diferentes táxons existentes no mundo, apresentam uma grande importância ecológica, econômica e biológica, mostrando-se fundamentais nos diferentes setores em que são encaixados. Eles não só são importantes, como também abundantes e, conforme uma revisão publicada em 2017, possuem grande diversidade, apontando a existência de mais 3.8 milhões de espécies fúngicas (HAWKSWORTH; LÜCKING, 2017) - corroborando com o terceiro lugar ocupado no ranking de maiores biomassas do planeta Terra, conforme visualizado na **Tabela 1**.

De modo geral, fungos e insetos são megadiversos, ambos cosmopolitas e é natural que existam relações interespecíficas, mesmo entre reinos distintos. Além de interações

simbióticas benéficas, fungos entomopatogênicos (FEPs) são microrganismos que causam doenças fatais em artrópodes (LITWIN; NOWAK; RÓŻALSKA, 2020), desempenhando um papel fundamental na manutenção do equilíbrio populacional de insetos em diferentes ecossistemas (HONG et al., 2023). Atualmente, existem mais de 1.500 espécies fúngicas descritas como capazes de infectar insetos e esses FEPs estão distribuídos em sete diferentes filos/subfilos, incluindo: Ascomycota, Entomophthoromycotina (Zoopagomycota), Microsporidia, Blastocladiomycota, Chytridiomycota, Basidiomycota e Mucoromycota (HONG et al., 2023). As espécies fúngicas que possuem essa capacidade entomopatogênica apresentam um grande potencial econômico e biotecnológico como agentes de controle biológico, podendo ser utilizado em diferentes formulações para controle de insetos de interesse de distintas formas em todo o mundo (FARIA; WRAIGHT, 2007), com o principal objetivo de ser utilizado como alternativa biológica contra insetos, fazendo com que seja reduzido o uso de pesticidas químicos que são maléficos tanto para o meio ambiente como também para a saúde humana.

Existem FEPs que são em grande parte parasitas obrigatórios de insetos, enquanto outras seguem o estilo de vida patógeno oportunistas. Uma característica interessante dos FEPs é que dependendo da linhagem ou espécie individual que se foca, diferente será a faixa de hospedeiros artrópodes estudada. Os FEPs considerados obrigatórios em grande parte são específicos de um hospedeiro relacionado, já os FEPs ascomicetos, que estão incluídos entre os fungos oportunistas e distribuem-se amplamente em três famílias da ordem Hypocreales: Clavicipitaceae, Ophiocordycipitaceae e Cordycipitaceae (**Figura 1**), são exemplos de gêneros/espécies individuais que possuem uma gama altamente variável de hospedeiros, principalmente da classe insecta (HONG et al., 2023).

A ordem Hypocreales contém nove famílias reconhecidas e mais de 2.600 espécies descritas até então, sendo assim a ordem mais importante dentro do filo Ascomycota, que faz parte do reino Fungi (ZHANG et al., 2018). As espécies que compõem a ordem Hypocreales evoluíram para diferentes estilos de vida além do parasitismo de diferentes seres vivos (plantas, insetos, nematóides e até mesmo outros fungos), de forma que se pode encontrar dentro da mesma ordem organismos com capacidades de saprofitismo (microrganismos recicladores) e também endofitismo (onde o fungo vive como um endossimbionte no interior de uma planta) (BERBEE, 2001).

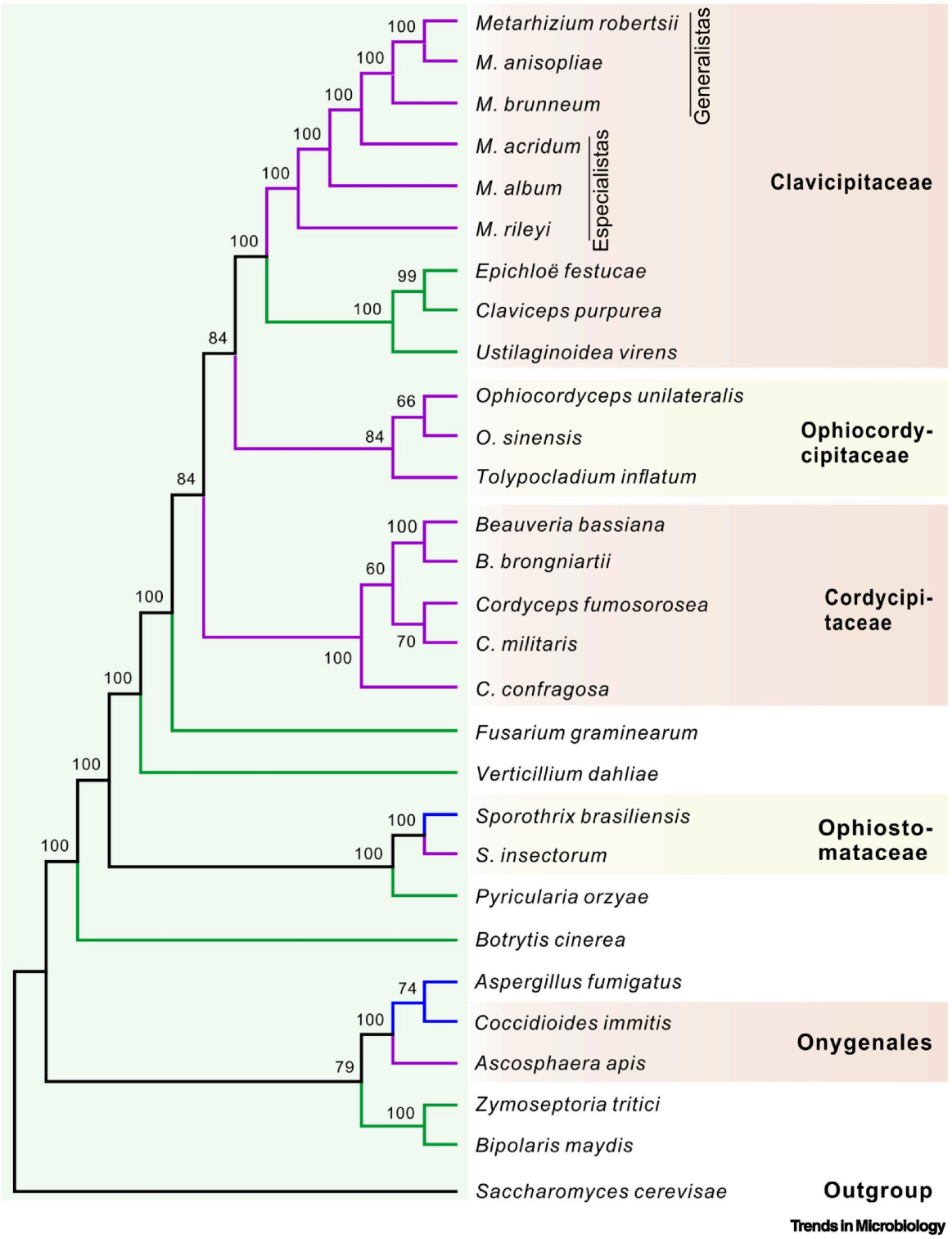


Figura 1: Associações filogenéticas dos fungos entomopatogênicos (FEPs) selecionados do filo Ascomycota. Árvore filogenética de máxima verossimilhança bem suportada inferida concatenando 45 proteínas ortólogas de cada espécie contendo 26610 posições de aminoácidos. Os ramos roxos são para as espécies de FEPs, ramos verdes para fungos patogênicos de plantas e os ramos azuis para fungos patogênicos de mamíferos. Adaptado de Hong et al. 2023. (HONG et al., 2023).

Espécies como *Metarhizium anisopliae* e *Beauveria bassiana*, principais exemplares de fungos entomopatogênicos, são exemplos de fungos generalistas que possuem capacidade de infectar centenas de espécies de artrópodes de diferentes ordens, enquanto fungos do gênero *Ophiocordyceps*, por exemplo, conseguem infectar apenas um hospedeiro específico (as formigas) (HONG et al., 2023).

1.2. Fungos entomopatogênicos dos gêneros *Metarhizium* e *Beauveria*

São descritas mais de 750 espécies diferentes de FEPs, porém a atenção maior vem se direcionando para as espécies pertencentes à ordem Hypocreales, já que inclui gêneros fúngicos que apresentam uma gama ampla de hospedeiros além da possibilidade de produção em massa desses fungos (BUTT et al., 2016). Já existem diferentes estudos que evidenciam os ascomicetos da ordem Hypocreales *Metarhizium anisopliae* e *Beauveria bassiana* como os FEPs mais caracterizados e utilizados no controle biológico de pragas (SCHRANK; VAINSTEIN, 2010) e, a partir disso, dentre os produtos oriundos de FEPs disponíveis comercialmente, em torno de 80% são baseados nos gêneros *Metarhizium* e *Beauveria* (FARIA; WRAIGHT, 2007), o que os faz exemplares principais quando o assunto é biopesticidas fúngicos.

Primeiramente descrita como patógeno do bicho-da-seda há mais de 180 anos, as espécies do gênero *Beauveria* são descritas como cosmopolitas e com capacidade de parasitar mais de 700 espécies de insetos (IMOULAN et al., 2017). Durante grande parte da história, o gênero *Beauveria* foi identificado como parte dos Deuteromicetos, podendo apenas depois do avanço das técnicas taxonômicas moleculares se aliar à sua verdadeira classificação. Embora seja fácil identificar o gênero, a identificação de espécies ainda é um problema até para os tempos atuais por conta da falta de características morfológicas individualizadoras (MASCARIN; JARONSKI, 2016). Diferentes estudos mostram que devido ao estilo de vida saprofítico somado a capacidade de entomopatogenicidade, o desenvolvimento positivo dos ascomicetos em geral permite uma produção em massa desses fungos, diferente de fungos com formas de vida mais restritas (como por exemplo, fungos da ordem Entomophthorales) (MASCARIN; JARONSKI, 2016).

Dentre os insetos que *Beauveria* é capaz de parasitar, a variedade dos hospedeiros inclui pragas agrícolas e florestais com diferentes impactos ambientais e importâncias econômicas, estando entre estes exemplares de cupins, mosca-das-frutas, insetos vetores de doenças, carrapatos e uma maior variedade artrópodes (MASCARIN; JARONSKI, 2016).

Também incluso dentro da ordem Hypocreales, a espécie *Metarhizium anisopliae* foi descrita como primeiro fungo no mundo a ser produzido em massa e utilizado para controle biológico de pragas com foco em insetos, conforme um trabalho publicado em 1888 por Krassilstschik (ROBERTS; ST LEGER, 2004). Nos mais de 100 anos subsequentes houve inúmeras tentativas de uso de *Metarhizium spp.* no controle biológico de pragas, obtendo diferentes resultados e níveis variados de sucesso, historicamente perdendo apenas para *Beauveria spp.* como fungo mais utilizado contra insetos no meio agrícola (SHENG; MCNAMARA; ST LEGER, 2022).

Assim como *Beauveria*, *Metarhizium* também é um gênero cosmopolita, altamente abundante, que abrange diversas espécies, possuindo uma ampla gama de características fisiológicas específicas conforme cada espécie do gênero e também uma ampla gama de hospedeiros, incluindo nematódeos, mamíferos, outros fungos e o mais comum: artrópodes (ROBERTS; ST LEGER, 2004). Existem muitos trabalhos sobre o uso de *Metarhizium* no controle de pragas, assim como também existem muitos dados na literatura que citam as interações deste fungo com seus hospedeiros. Além disso, esse gênero também tem sido objeto de estudo em nível molecular envolvendo temas importantes como virulência fúngica e como funciona a especificidade do hospedeiro (ROBERTS; ST LEGER, 2004) - atualmente ocupando o primeiro lugar quando se trata de gêneros fúngicos entomopatogênicos mais bem compreendidos em níveis básicos.

O gênero *Metarhizium* consegue matar diferentes insetos ao mesmo tempo que maximiza a proteção de plantas, conseguindo se associar a essas sem prejudicá-las (SHENG; MCNAMARA; ST LEGER, 2022). Além disso, a maioria das espécies deste gênero também são saprófitas e comumente aparecem como espécies endofíticas, possuindo capacidade de alternar entre estes diferentes estilos de vida conforme necessidade (ST LEGER; WANG, 2020), mostrando-se um organismo de fácil adaptação e apresentando uma alta divergência entre espécies (conforme pode-se visualizar na **Figura 2**).

Essas divergências interespécies sugerem que a trajetória evolutiva de *Metarhizium* foi caminhando de espécies especialistas, surgindo posteriormente espécies intermediárias e por fim, surgindo as espécies generalistas. A evolução coincidiu com a diversificação e disponibilidade de insetos, sugerindo que quanto mais variado o menu, maior a pressão evolutiva para adaptação e especiação. Essa adaptação foi acoplada a uma série de características genômicas que, interagindo entre si, promoveram a formação de novas espécies - explicando como dentro de um mesmo gênero podem existir espécies com capacidades tão distintas (HU et al., 2014).

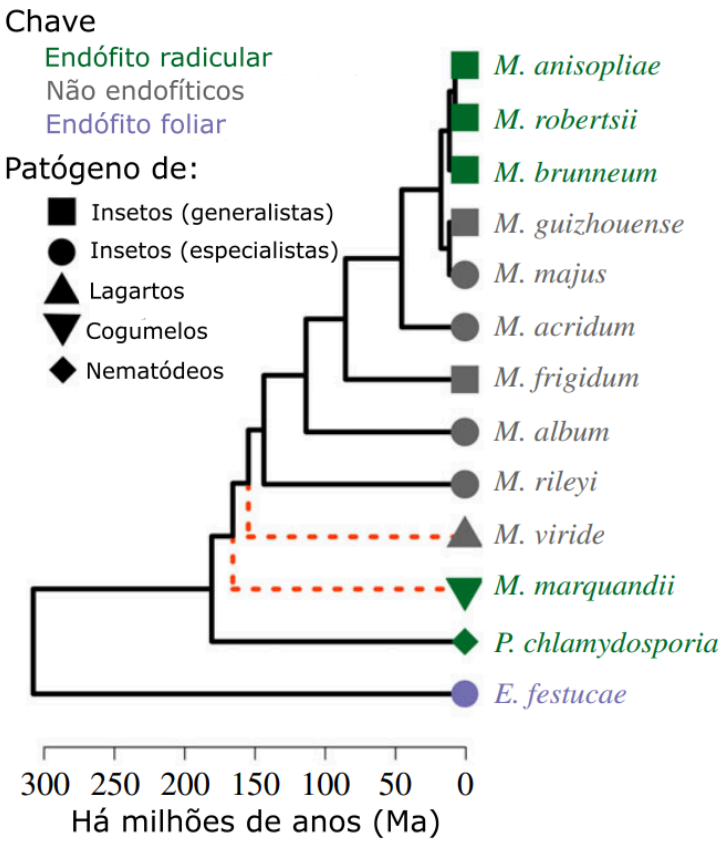


Figura 2: Aspectos evolutivos de *Metarzhium* sp.. Árvore filogenética de espécies sequenciadas de *Metarhizium* e fungos relacionados (linhas sólidas) com tempo estimado de divergência interespécies. O saprófito *M. marquandii* e o patógeno de lagartos *M. viride* também foram incluídos (linhas tracejadas vermelhas). Adaptado de St Leger; Wang, 2020 (ST LEGER; WANG, 2020).

Além das semelhanças que os gêneros *Beauveria* e *Metarhizium* compartilham, esses fungos, assim como a grande maioria dos FEPs, infectam artrópodes funcionando principalmente como um patógeno de contato - diferente de vírus, protozoários e bactérias, que requerem vias específicas de infecção (MASCARIN; JARONSKI, 2016). Por conta de distintas peculiaridades presentes nestes fungos que cada vez mais os FEPs têm sido empregados no controle biológico de pragas.

1.3. Ciclo de infecção

Entre bactérias, vírus e fungos com capacidade de infectar pragas, os fungos estão entre os mais estudados e utilizados nesse ramo, principalmente pela sua eficiência em matar seus hospedeiros, assim como também pela biodiversidade que apresentam conjunta a facilidade de serem acessados na natureza (SCHRANK; VAINSTEIN, 2010). Com isso, pesquisas para compreender melhor a caracterização dos fatores de virulência (visando

melhorar o processo de infecção) dos FEPs, assim como também os mecanismos que permitem especificar os hospedeiros ao qual se associam, tem se tornado cada vez mais importantes ao longo das décadas (HU et al., 2014).

Diferente das bactérias e vírus entomopatogênicos, que infectam através de infecções intestinais, necessitando de ingestão oral, os FEPs iniciam o processo de infecção através da penetração na cutícula do hospedeiro. Normalmente, os esporos assexuados (conídios) são os responsáveis pela infecção e estão no mesmo ambiente em que seus potenciais hospedeiros (HONG et al., 2023).

Os conídios, de espécies de *Metarhizium* por exemplo, após o primeiro contato, germinam na cutícula dos insetos com a finalidade de produzir uma estrutura de infecção, o apressório (**Figura 3**), estrutura que irá aumentar a pressão de turgescência no interior da célula do fungo e secretar enzimas de degradação, fazendo com que a penetração fúngica seja mais eficiente e o fungo chegue até os tecidos do hospedeiro (HONG et al., 2023). Dando continuidade ao processo, um pino de penetração auxilia e permite o seccionamento da penetração do apressório na cutícula do inseto.

As células fúngicas são encapsuladas e melanizadas pelos hemócitos dos insetos quando atingem a cavidade corporal desses animais, porém, como forma de driblar os mecanismos de defesa dos artrópodes, células germinativas fúngicas podem escapar e alterar do crescimento filamentoso comum para o brotamento (como as leveduras), a fim de produzir de forma mais rápida hifas capazes de colonizar a hemocele do hospedeiro (HONG et al., 2023).

As células germinativas dos FEPs só voltam ao crescimento filamentoso após a morte do inseto infectado, e rompem novamente a cutícula do hospedeiro para formar múltiplas camadas de esporos na superfície do animal morto (**Figura 3**). As adaptações das células germinativas dos FEPs lhe permitem escapar da resposta imune dos insetos, apresentando uma infecção eficiente e atingindo o objetivo: ser fatal contra o hospedeiro (HONG et al., 2023).

Conforme estudos publicados, mesmo com as adaptações e demais mecanismos citados anteriormente, todo o sistema de infecção fungo-inseto se encontra consideravelmente conservado na natureza, de forma que o rompimento da cutícula e colonização dos tecidos causada por fungos nos insetos se compare ao processo de infecção dos fungos fitopatogênicos (que infectam plantas), incluindo a formação de tubos germinativos, estrutura apressória e pinos de penetração (HONG et al., 2023).

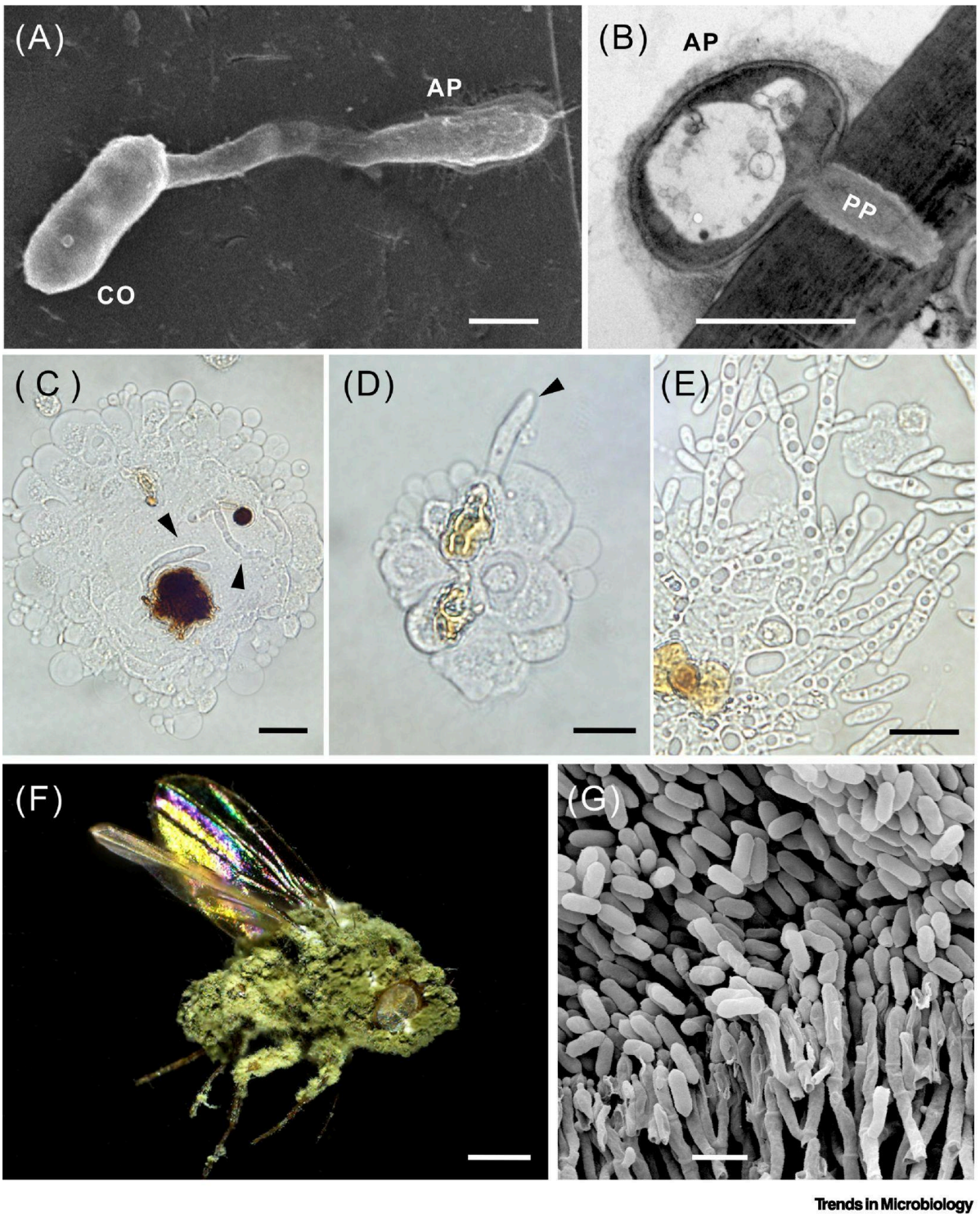


Figura 3: Etapas do processo de infecção de *Metarhizium*. (A) Germinação do conídio (CO) e a formação da estrutura de infecção apressória (AP). (B) Descontinuidade da penetração do apressório na cutícula do hospedeiro através de um pino de penetração (PP). (C) As setas apontam células fúngicas encapsuladas e melanizadas pelos hemócitos dos insetos após invasão dos tecidos do inseto entre 12 a 16 horas. (D) A seta destaca um tubo germinativo escapando de um nódulo de hemócito entre 18 a 20 após a entrada do fungo no inseto. (E) Brotamento do tipo levedura e produção de múltiplas hifas após mais de 30 horas do fungo dentro do hospedeiro. (F) Cadáver de *Drosophila* infectado pelos esporos de *M. robertsii.* (G) Produção de esporos por *M. robertsii* na superfície do corpo do inseto. Adaptado de Hong et al. 2023. (HONG et al., 2023).

Entretanto, todas as estratégias e capacidades que os fungos apresentam, incluindo diferenciarem-se entre fungos generalistas ou fungos especialistas, estão atreladas a características individuais que auxiliam no entendimento do surgimento de patógenos e determinantes de virulência, protagonistas quando o assunto é processos de infecção. Dentre os distintos determinantes de virulência que se tem na literatura, os metabólitos secundários (MSs) aparecem como essenciais durante a infecção fúngica (SBARAINI et al., 2016), o que faz com que estudos sobre função de diversos genes, vias e os MSs associados a infecção sejam peças chave para um entendimento genuíno sobre o assunto.

1.4. Metabólitos secundários de fungos filamentosos

Os fungos filamentosos são capazes de produzir inúmeros metabólitos, podendo ser primários ou secundários. Os metabólitos primários (MPs) são moléculas produzidas durante o crescimento do fungo, cruciais para funções vitais e servindo muitas vezes de base para a produção de outros metabólitos (CHROUMPI; MÄKELÄ; DE VRIES, 2020). Já os metabólitos secundários (MSs) podem ser definidos como pequenas moléculas bioativas que não são obrigatórias para a sobrevivência do fungo, porém se mostram essenciais nas interações ecológicas com outros organismos (KWON et al., 2021). Os MSs incluem pigmentos, reguladores de crescimento de plantas, antibióticos, outros produtos aplicados na indústria farmacêutica e toxinas (PROCTOR et al., 2009), apresentando funções naturais que podem variar desde mecanismos de defesa contra predadores ou competidores, agentes de simbiose, agentes transportadores de metais e fatores de virulência ofensivos (característico de fungos entomopatogênicos) (KJÆRBØLLING et al., 2019).

Até então, a maioria dos MSs fúngicos descritos são derivados de três classes enzimáticas: Terpenos, Policetídeos sintases (*polyketide synthases* (PKS)) e Sintetases peptídicas não ribossômicas (*nonribosomal peptide synthetases* (NRPS)). Cada uma dessas classes enzimáticas utiliza metabólitos primários mais simples como substratos e reorganiza-os em moléculas mais complexas estruturalmente, formando então os terpenos, PKS e NRPS, por exemplo. A ação de enzimas sobre moléculas complexas acarreta em uma variedade de MSs fúngicos. Oxigenação, ciclização, isomerização e condensação com distintos metabólitos são algumas das modificações que auxiliam a alta variabilidade química e biológica desses compostos (PROCTOR et al., 2009).

1.5. Clusters gênicos biossintéticos (*biosynthetic gene clusters* (BGCs))

Frequentemente os genes responsáveis pela produção dos MSs são organizados em *clusters* nos genomas fúngicos. Esses agrupamentos normalmente são chamados de *clusters* gênicos biossintéticos (*biosynthetic gene clusters* (BGCs)), incluindo um gene codificador de uma enzima biossintética central (um gene *backbone*) que comumente é representado por classes enzimáticas de NRPS, PKS ou Terpenos, por exemplo, como pode ser visualizado na **Figura 4 (KWON et al., 2021)**. Além disso, os BGCs incluem também genes com diferentes funções, como genes codificadores de enzimas modificadoras de estrutura, proteínas transportadoras de metabólitos e proteínas reguladoras da transcrição de *clusters* gênicos (**Figura 4**) (PROCTOR et al., 2009).

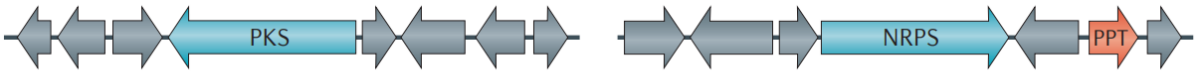


Figura 4: *Clusters* gênicos fúngicos. As setas representam genes. Em azul estão representados os genes *backbone* (indicados como as enzimas PKS e NRPS, respectivamente) e nas setas cinzas estão os genes que apresentam as demais funções. A seta laranja (PPT) representa um domínio normalmente codificado por outro gene dentro do *cluster*. Adaptado de Brakhage, 2013. (BRAKHAGE, 2013).

O sequenciamento genômico permite inúmeras oportunidades para a descoberta de produtos naturais e novos entendimentos sobre a diversidade e como funciona a distribuição de BGCs. Esses BGCs normalmente se mostram muito evoluídos, principalmente por realizarem trocas horizontais de genes entre organismos, o que faz com que abrem-se portas para a curiosidade de entender e testar os diferentes efeitos que suas moléculas podem apresentar (JENSEN, 2016). Pesquisas utilizando modelos para estudo de metabolismo secundário tem se mostrado interessantes, principalmente quando se encontra níveis extraordinários de diversidade de BGCs (como é o caso das bactérias marinhas *Salinispora*), porém, já se sabe que quando BGCs são observados com pouca frequência em um genoma, resultando em uma alta diversidade destes *clusters* em nível populacional, uma maior diversificação, principalmente de capacidades, será encontrada naquele grupo (JENSEN, 2016).

Os MSs descobertos até agora representam apenas uma pequena parte do número total produzido na natureza, considerando a enorme diversidade estrutural de diferentes compostos, desde microbianos até vegetais (CHAVALI; RHEE, 2018). A busca por novos

MSs e os BGCs responsáveis por produzi-los tem se tornado importante, pois novos produtos com diferentes potenciais podem ser descobertos a partir disso, principalmente medicamentos em potencial. Diferentes métodos para a descoberta de novos MSs têm sido explorados, desde modelos experimentais *in vitro*, como também métodos *in silico* - envolvendo técnicas de bioinformática e suas aplicações (CHAVALI; RHEE, 2018).

Um exemplo de método experimental, publicado em trabalhos anteriores, como bem sucedido e estabelecido para a descobertas de novos MSs foi chamado de "cultura/triagem/condição". Neste estudo, o método consistiu em selecionar organismos expostos a diferentes tipos de condições de crescimento e estresse externos, realizando a extração e fracionando os sobrenadantes da cultura, caracterizando as moléculas purificadas utilizando técnicas como ressonância magnética nuclear e espectrometria de massa (CHAVALI; RHEE, 2018). Porém, embora a experimentação seja um processo fundamental do método científico, é importante reconhecer que apresenta limitações e potenciais falhas. Esses erros podem acarretar em perdas significativas, tanto em termos de tempo quanto de recursos financeiros. Assim, uma escolha criteriosa e precisa da metodologia empregada é crucial na mitigação de riscos e também na otimização do processo científico como um todo.

Análises bioinformáticas de diferentes genomas microbianos podem auxiliar bastante na identificação de instruções genéticas de interesse, assim como na predição de BGCs responsáveis por produzir muitos produtos naturais (WALKER; CLARDY, 2021). A análise comparativa de sequências, baseada nas enzimas mais comuns relacionadas (NRPS, PKS, por exemplo), têm sido usadas para identificar grupos de genes potencialmente envolvidos na biossíntese de MSs (INGLIS et al., 2013). Devido a clareza e fácil acesso dessas informações, é possível fazer a predição de BGCs nos genomas fúngicos usando programas computacionais como SMURF e antiSMASH (BLIN et al., 2021; KHALDI et al., 2010). Isso acontece porque, quando se trata de bioinformática, muitos BGCs são altamente informativos sobre a classe de compostos que são capazes de codificar, facilitando o desenvolvimento de ferramentas computacionais que podem ser usadas para prever o potencial do produto natural a partir dos dados de sequências de DNA recuperados (JENSEN, 2016) .

Com a evolução e aumento do poder das metodologias computacionais, métodos de bioinformática que anteriormente tinham como foco identificar BGCs e localizar os genes *backbone* em fungos e bactérias, hoje em dia tem seu foco voltado também para a identificação de outras enzimas exclusivas responsáveis por catalisar diferentes compostos, potencializando a predição de novos MSs, possibilitando a caracterização de diferentes BGCs e criando diferentes oportunidades científicas e comerciais (CHAVALI; RHEE, 2018).

Tendo em vista a relevância da redução do uso de agrotóxicos nocivos ao meio ambiente e à saúde pública, considerando também a biodiversidade fúngica e a quimiodiversidade potencial dos metabólitos secundários associados, este estudo visou identificar e avaliar a conservação de *clusters* gênicos biossintéticos (BGCs) na ordem fúngica Hypocreales, com foco no controle biológico de pragas. O interesse se baseia nas capacidades pesticidas e nas propriedades de virulência associadas a metabólitos secundários presentes em fungos entomopatogênicos desta ordem. Para isso, BGCs foram recuperados e analisados a partir de genomas fúngicos de membros da ordem Hypocreales. A comparação dos BGCs identificados com aqueles das espécies *Metarhizium anisopliae* e *Beauveria bassiana*, reconhecidamente utilizados no controle biológico de pragas, possibilitou avaliar a existência de conservação entre os *clusters* gênicos da mesma ordem. Essa avaliação permitiu inferir sobre o potencial entomopatogênico de outras espécies fúngicas da ordem Hypocreales, considerando os BGCs que possuem.

2. OBJETIVOS

2.1. Objetivo Geral

Avaliar a identidade e conservação de clusters gênicos biossintéticos (*biosynthetic gene clusters* (BGCs)) em sequências genômicas de fungos presentes na ordem Hypocreales, considerando a presença de fungos entomopatogênicos e fitopatogênicos.

2.2. Objetivos Específicos

- Identificar BGCs nos genomas da ordem Hypocreales recuperados de bancos de dados públicos utilizando ferramentas de bioinformática;
- Analisar a diversidade dos BGCs e identificar as famílias gênicas biossintéticas (*biosynthetic gene families* (BGFs)) presentes;
- Avaliar a conservação dos BGCs entre diferentes espécies de Hypocreales em comparação com as espécies entomopatogênicas, principalmente *Metarhizium anisopliae* e *Beauveria bassiana,* considerando as BGFs identificadas;
- Realizar análises de expressão gênica de clusters selecionados para identificar novos potenciais genes envolvidos na virulência de fungos entomopatogênicos, considerando diferentes condições.

3. RESULTADOS

Os resultados apresentados nesta dissertação se encontram disponibilizados na forma de um manuscrito científico, redigido conforme as normas do periódico *Genetics and Molecular Biology*, tendo como base o formato de artigo *SHORTS COMMUNICATION* (https://www.gmb.org.br/instructions-to-authors).

CONSERVED GENE CLUSTERS IN ENTOMOPATHOGENIC FILAMENTOUS FUNGI

Alexandra de Azevedo da Rocha and Charley Christian Staats*

Programa de Pós-graduação em Biologia Celular e Molecular, Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves, 9500 - Agronomia, Porto Alegre, RS, 91501-970, Brasil

* Corresponding author: staats@ufrgs.br

Short title: BGCs in Hypocreales

Keywords: Entomopathogenic fungi, secondary metabolites, biosynthetic gene clusters.

ABSTRACT

Entomopathogenic filamentous fungi from order Hypocreales are studied for biological pest control due to their ability to infect arthropods. Their biotechnological potential lies in the production of secondary metabolites (SMs), encoded by biosynthetic gene clusters (BGCs), which are crucial for the fungal life cycle and infection. Due to the different roles played by SMs in fungi, the in-depth study of these molecules has increased over the last years. Considering the biotechnological importance that species of genus *Beauveria* have already shown, we performed an *in-silico* analysis of BGCs of the species from the order Hypocreales to uncover potential conserved pathways and find new candidates for biological pest control. A total of 295 genome sequences were analyzed using antiSMASH, allowing the identification of 12,968 BGCs. Conservation analysis was performed using BiG-SCAPE, which could group these BGCs into 2,127 biosynthetic gene families. Despite the presence of conserved gene clusters, specially within the same genus, when the comparison of BGCs was performed within the order, a large part of them are orphans, highlighting the great diversity of BGCs and consequently the chemodiversity of fungal species. Our approach shed light to uncover new molecules that may provide new biotechnological tools.

1. INTRODUCTION

Entomopathogenic fungi (EPF) are cosmopolitan species known for their association with plants and by infecting arthropod hosts. These fungi hold immense biotechnological potential, as they are globally employed in biological pest control (Hong et al. 2024; de Sousa et al. 2025; Alviti Kankanamalage et al. 2025). To successfully infect insects and arthropods, EPFs produce an array of virulence determinants. These molecules primarily function to disrupt the host cuticle and inactivate host defenses (Qin et al. 2023; Ying 2024; de Miranda et al. 2024). Additionally, these fungi produce a vast repertoire of secondary metabolites (SMs), which are small-molecule compounds that play crucial roles in their ecological interactions and during fungal infection, even though these metabolites are not generally essential for fungal survival (Zhgun 2023; Shahbaz et al. 2024; Jinna et al. 2025).

The biosynthesis of secondary metabolites (SMs) is regulated by intricate regulatory networks and is widely recognized to be orchestrated by a diverse set of proteins encoded by genes clustered into biosynthetic gene clusters (BGCs). These gene clusters are organized as contiguous regions within the genome and generally harbor all elements for the synthesis of specific SMs (Zhgun 2023). BGCs typically contain backbone genes whose encoded protein domains serve as informative markers for inferring the class of secondary metabolites produced, such as nonribosomal peptide synthetases (NRPS), polyketide synthases (PKS), ribosomally synthesized and post-translationally modified peptides (RiPPs), terpenes, hybrid molecules, and others (Xu et al. 2023).

Considering the ongoing expansion of fungal genome sequences available, is feasible to assume an increasing number of BGCs, paving the way for the discovery of promising new bioactive secondary metabolites with valuable applications (Gu et al. 2020; Riedling et al. 2024). The Ascomycota phylum harbors a diverse array of filamentous fungi with significant ecological and economic importance. Given that the primary source of their remarkable biotechnological potential lies in the production of SMs (Keller 2019; Zhgun 2023), the conservation analysis of BGCs in entomopathogenic filamentous fungi is crucial for uncovering new bioactive compounds and identifying promising candidates for biological pest control (Pedrini 2018). Accordingly, we focused on the order Hypocreales, a rich source of entomopathogenic fungal species, employing a refined *in silico* approach to search

for conserved genes of interest and novel bioactive compounds derived from BGCs with potential applications in agriculture.

2. METHODS

2.1. Genome sequences from Hypocreales

A total of 295 genome sequences from Hypocreales were retrieved from the National Center for Biotechnology Information (NCBI) Genome database using the NCBI datasets (v 14.27.0 (O'Leary et al. 2024)) tool (Table S1). All annotated genomes of the order Hypocreales were retrieved and stored as the gbff files. Completeness of genome sequences was evaluated using BUSCO v. 5 (Manni et al. 2021).

2.2. Identification of biosynthetic gene clusters and gene clusters families in Hypocreales genomes

The corresponding genome files were used for BGC prediction using the antiSMASH v. 6.1.1 (Blin et al. 2021). The resulting gbk files were then used for gene clusters families (GCFs) prediction using the BiG-SCAPE tool v 1.1.2 (Navarro-Muñoz et al. 2020). Summarization of total BGCs, GCFs, as well as GCFs size are performed using R custom made scripts, available at ZENODO (DOI 10.5281/zenodo.15307303). Fine structure conservation of selected BGCs were further analyzed using the CAGECAT server (van den Belt et al. 2023).

2.3. Expression analysis of selected clusters

The Prefetch tool from the SRA Toolkit v 3.0.0 (https://github.com/ncbi/sra-tools) was used to retrieve sequencing data from SRA (Sequence Read Archive) files available on the NCBI for detailed analysis. Runs from the NCBI Bioproject codes PRJNA574771 (Transcriptome sequencing of *Acyrthosiphon pisum* by infecting with *Beauveria bassiana*) and PRJNA1032759 (A dual RNA-seq approach to study oral infection of the red flour beetle *Tribolium castaneum* with the entomopathogenic fungus *Beauveria bassiana*) were retrieved and converted to fastq files using the Fasterq-dump tool with the "--split-files" option. Fastp (Chen et al. 2018) was then used to clean and quality-filter the raw fastq data for further analysis. Processed RNA-seq reads were aligned to *B. bassiana* ARSEF 2860 genome sequence (NCBI accession

code GCA_000280675.1) using the STAR tool with the "--quantMode" option to generate a matrix of read counts per gene. DESeq2 package was used to obtain the normalized expression of each gene in the distinct conditions, which was used for the construction of heatmaps with the 'Pheatmap' package. The R scripts are also available at Zenodo (10.5281/zenodo.15307303).

3. RESULTS

3.1. Identification of BGCs in Hypocreales genome sequences

We recovered a total of 295 genomic sequences from fungal isolates of the order Hypocreales from NCBI, as of February 2022. To infer the BGCs and analyze their conservation, a stringent pipeline (Figure 1) was constructed and applied for each of these genome sequences. Analysis with antiSMASH resulted in a total of 12,968 BGCs distributed along the genomic sequences. No significant correlation between the genome size and the number of inferred BGCs was observed (Figure 2). Such gene clusters were then categorized based on the class of metabolite they are predicted to produce, following the classification proposed by BiG-SCAPE: Non-Ribosomal Peptides (NRPS), Polyketides (PKSI and PKSOther), Terpenes, Ribosomally synthesized and post-translationally modified Peptides (RiPPs), Hybrids (PKS-NRP_Hybrids), and Others. A total of 4,950 BGCs were classified as NRPS, representing the largest group. PKS accounted for 2,899 BGCs, followed by 2,314 Terpenes, 1,398 Others, 1,156 Hybrids, 211 PKSother, and 40 RiPPs, the smallest group.

The distribution of BGCs along the analyzed genomes was further evaluated. For the four genera containing the higher number of species with genome sequences available (*Claviceps, Fusarium*, *Trichoderma*, and *Metarhizium*), we could detect a large variation in the number of BGCs in each genome (**Figure 3**). This result was even observed with completeness of the genome estimated to be high as revealed by BUSCO analysis (all genomes analyzed displayed over 95% of complete BUSCOs). For instance, there is a variation from 15 BGCs in *Geosmithia morbida* to 89 BGCs in *Hirsutella minnesotensis*, and an average total BGCs per genome of 48.77. Globally, NRPS and PKSI are the most abundant BGC classes, followed by Terpene and PKS-NRP_HYBRIDS. PKSOTHER and RIPPS are the last common.

3.2. Conservation analysis of GCFs in Hypocreales genomes

To evaluate the conservation patterns of the predicted BGCs across the genomes, the antiSMASH output files were further analyzed using BiG-SCAPE (Navarro-Muñoz et al. 2020). BiG-SCAPE employs distance metrics based on the presence and order of conserved protein domains within predicted BGC products to group similar clusters into Gene Cluster Families (GCFs). This analysis was performed independently for BGCs classified into the following categories: NRPS, Others, PKS-NRP Hybrids, PKSI, PKSother, RiPPs, and Terpene. To determine the optimal clustering stringency, the analysis was conducted across a range of cutoff values (from 0.1 to 0.9, with increments of 0.1). The selection of the best cutoff was based on analyzing its impact on the number of resulting GCFs, the number of singletons BGCs (those not assigned to any GCF), and the mean GCF size. Visual inspection of these metrics against increasing cutoff values suggested that, in general, a cutoff value of 0.6 provided a robust clustering result for most classes. Below this threshold, increasing the cutoff led to a steady decrease in the number of GCFs and singletons as more BGCs were grouped (Figures 4A and B, respectively). However, increasing the cutoff beyond 0.6 often resulted in a more drastic increase in the mean GCF size (Figure 4C), indicating the merging of less similar BGCs. Therefore, a cutoff of 0.6 was selected for subsequent analyses. Analysis at this cutoff reveals differential conservation patterns among BGC classes, with PKSI, Terpene, and NRPS BGCs showing greater tendencies to form larger, less singleton-heavy families compared to RiPPs and "Others." This differential clustering reflects inherent differences in the evolutionary dynamics and conservation landscapes of these distinct secondary metabolite types.

The size of each GCF was further evaluated. Distribution of BGC sizes by class revealed a marked heterogeneity (Figure 5A). To quantify disparities in GCF sizes across biosynthetic gene cluster (BGC) classes, we tested whether their distributions followed power-law scaling — a pattern typical of systems dominated by a few large elements among many small ones. GCF sizes were fit to discrete power-law models and compared to lognormal and exponential distributions using Vuong’s likelihood ratio test, with 1,000 bootstrap replicates used to assess parameter stability. While NRPS, PKSI, and RiPPs did not show significant differences between models (p-values 0.83, 0.48, and 0.05, respectively), all other classes (e.g., Others, PKS-NRP hybrids, PKSother, Terpene) were significantly

better fit by a lognormal distribution ($p < 0.02$) – Table S2. These findings, particularly the lack of significant difference between power-law and lognormal fits in major classes like NRPS and PKSI, suggest that while some biosynthetic systems may exhibit heavy-tailed distributions, the overall architecture of fungal GCF sizes is more consistent with lognormal variation than with true power-law scaling. Furthermore, it is noteworthy that when accounting only for GCFs containing BGCs present in the MiBIG database, the distribution of GCF sizes decreases drastically (Figure 5B), suggesting that the unexplored and unidentified portion of fungal chemodiversity is even greater than currently recognized.

3.3. Identification of BGCs and BGFs in entomopathogenic fungi.

Entomopathogenic fungi, especially from the genera *Beauveria* and *Metarhizium*, are widely used for the biological control of agricultural pests, as well as vectors of human diseases. We hypothesized that biosynthetic gene clusters (BGCs) involved in the virulence of entomopathogenic fungi would be conserved across genera that include insect-infecting species. To test this, we first evaluated the BiG-SCAPE output to determine the proportion of gene cluster families (GCFs) that are genus-specific versus those shared among multiple genera. As expected, the majority of GCFs were genus-specific ($\chi^2 = 1888.9$, df = 2, $p < 2.2e\text{-}16$; Figure 6). We then investigated whether this pattern also could be observed within entomopathogenic genera. For this, we filtered GCFs found in *Metarhizium*, *Beauveria*, *Cordyceps*, *Hirsutella*, *Tolypocladium*, *Purpureocillium*, and *Akanthomyces* genera, and examined their distribution. Again, the largest proportion of GCFs was genus-specific ($\chi^2 = 657.07$, df = 2, $p < 2.2e\text{-}16$; Figure 6).

We then analyzed the GCFs that are shared only by entomopathogenic genera. We identified a set of GCFs exclusively shared among genera traditionally associated with entomopathogenic lifestyles. In our dataset, a total of 16 GCFs were present in at least three distinct entomopathogenic genera but absent from non-entomopathogenic taxa. The most frequently co-occurring genera among these shared GCFs were *Beauveria*, *Cordyceps*, and *Akanthomyces*, suggesting conservation of specific biosynthetic pathways across these lineages. Notably, GCF_5048 was found in four genera (*Akanthomyces*, *Beauveria*, *Cordyceps*, and *Hirsutella*), indicating a potentially widespread and conserved function (Table S3).

These observations are consistent with the taxonomic relationships among the genera: *Beauveria*, *Cordyceps*, and *Akanthomyces* all belong to the family Cordycipitaceae, while *Hirsutella* and *Purpureocillium* are members of Ophiocordycipitaceae, and *Metarhizium* belongs to Clavicipitaceae. Fine mapping and comparison of the above genera members confirmed that our approach for setting the BiG-SCAPE cutoff value to 0.6 is not too relaxed, as all genes of a BGC clearly show orthologs in the BGCs of a family Synteny analyses of the BGCs found within these families are shown in Supplementary Figures S1 to S16, which are present in Supplementary File 1. These results support the hypothesis that certain BGCs are conserved among entomopathogenic fungi, potentially reflecting shared virulence strategies or ecological adaptations.

3.4. Expression analysis of entomopathogenic fungi BGCs

Fungi from the genus *Beauveria* are widely used for the biological control of agricultural pests. To assess whether the identified GCFs harbor BGCs potentially associated with virulence, we analyzed the expression of each gene within these BGCs using RNA-Seq data from a study that investigated the infection of the pea aphid (*Acyrthosiphon pisum*) and the red flour beetle (*Tribolium castaneum*) by *B. bassiana* at different time points. After retrieving the data from the SRA and aligning it to the *B. bassiana* ARSEF 2860 genome, gene counts were normalized using the DESeq2 package, and the expression levels of each *B. bassiana* BGC within these GCFs were plotted. Four of the fourteen evaluated BGCs exhibited significant gene upregulation under infection conditions relative to saprophytic and control conditions, suggesting a potential role in virulence in at least one of the hosts analyzed (Figure 7).

The highlighted GCFs, which encode enzymes implicated in virulence such as NRPS (GCF 5045, backbone gene BBA_05020) and Hybrids (GCF 7392, backbone gene BBA_04028), exhibit potential upregulation under infection conditions. This strongly suggests that their metabolic products play a significant role in the fungus-host interaction, making them prime candidates for further analysis to elucidate their precise functions in all this process.

4. Discussion

Numerous studies have demonstrated the significance of filamentous fungi and the potential applications of their secondary metabolites. Consequently, the study of their biosynthetic gene clusters (BGCs) is fundamental, as they are crucial for diverse processes, including biological associations and infections. Our analysis of hundreds of fungal genomes revealed that most gene cluster families (GCFs) in the order Hypocreales are genus-specific or orphans (singletons), a pattern previously observed in bacterial BGCs (Meesil et al. 2023). This genomic distribution reflects a distinctive evolutionary pattern that enables fungi to inhabit and dominate a wide range of specific ecological niches, facilitating unique adaptations to diverse environments (Naranjo-Ortiz and Gabaldón 2019) Coevolutionary studies demonstrate that fungi develop specific adaptations, enabling them to adjust to their environment while also specializing for distinct hosts. This evolutionary pressure can result in diverse interactions, from mutualism to specialized parasitism. For example, some entomopathogenic fungi form mutualistic associations by developing specific genes to interact with various plant species. In contrast, parasitic specialization involves the evolution of toxins specific to different insect species, the development of specialized transmission mechanisms, and the adaptation of strategies to control growth within the host, as exemplified by entomopathogenic fungi when they infect insects (Zaccaron and Stergiopoulos 2025)

The observed lognormal distribution of GCFs points to restricted expansion of conserved BGCs among species, mainly because of strong selective pressure that would allow BGCs diversification and consequently constraining GCF expansion. This finding is consistent with the absence of extreme superfamilies, and suggests that successful fungal evolution selects for specialized biosynthetic capabilities over broad, generalist strategies (Aghdam and Brown 2021). The prevalence of orphan and singleton GCFs indicates a dual adaptative strategy, balancing the stability of conserved genus-level biosynthetic pathways with the flexibility afforded by rapid innovation through the independent creation of GCFs. This combination of stability and flexibility likely underpins the ecological success of fungi in diverse habitats (Muggia et al. 2020).

Although most GCFs were genus-specific, our attention was drawn to 14 GCFs that were present in at least three distinct entomopathogenic genera. The BGCs within these shared families are thus prime candidates for functional analysis, based on the hypothesis that they are crucial for virulence and fungal pathogenesis

(Quesada-Moraga et al. 2023). Among the 14 GCFs of interest, GCF_5048 particularly stands out for containing four distinct entomopathogenic genera, rendering its constituent BGCs high-priority targets for functional characterization. Functional priorization strategies should focus on genes reported as upregulated under relevant conditions in the literature. Since the backbone gene of a BGC determines the core structure of its metabolic product, its high expression during infection renders it a prime target for functional analysis. Such genes are therefore strong candidates for studies aimed at the discovery of new natural products (Riedling et al. 2024). The *B. bassiana* backbone genes BBA_05020 (GCF 5045) and BBA_04028 (GCF 7392) were upregulated during infection of *T. castaneum* at 72h and 48h, respectively. Analysis of the aminoacid sequence encoded by gene BBA_05020 using the NCBI Protein database identified a protein featuring and adenylation domain, a hallmark of Non-Ribosomal Peptide Synthetases (NRPS). In the case of gene BBA_04028, the search revealed domains classifying their backbone gene as Hybrids. The upreglation during infection was not limited to the backbone genes, other genes within the GCFs highlighted in Figure 7 also showed increased expression. This positions them as promising targets for functional characterization of their roles in the fungus-host interaction. Such host-responsive induction of secondary metabolite pathways aligns with previous findings in entomopathogenic fungi (Zhang et al. 2020). Moreover, the differential expression of genes within these clusters, particularly the backbone genes, confirms their metabolic activity and strongly implicates them in host adaptation, marking them as prime candidates for genetic manipulation to improve biocontrol strategies. Notably, the BGCs from *B. bassiana* ARSEF 2860 highlighted in Figure 7 are currently not deposited in the MIBiG repository.

Furthermore, while *B. bassiana* possesses several BGCs that are known to produce various compounds, the overall number of characterized clusters is still limited. This highlights the need for targeted studies to characterize the vast number of unannotated and undeposited BGCs in this species in the MIBiG database. Still about the two backbone genes, this finding, when considered in the context of our 12 other results, underscores the role of fungal specificity in dictating host choice and activating virulence factors. It implies that the fungal response is not uniform but is instead tailored to specific host-derived stimuli (Rinker et al. 2024). This result demonstrates a critical point where the presence of a fungal BGC doesn’t guarantee

its expression or involvement in every pathogen context. This highlights the sophisticated regulatory mechanisms governing these clusters and emphasizes the need for deeper research into their specific activation triggers.

Bioinformatics tools and their applications have advanced the study of fungal BGCs, a field that deserves greater attention. While our *in silico* analyses yielded significant hypotheses, experimental validation is essential. Future research should involve wet-lab techniques (for example, CRISPR-Cas9, metabolomics) and broader taxonomic surveys to assess the entomopathogenic specificity of these clusters. It must be noted that these computational predictions are influenced by numerous parameters and demand meticulously curated, high-quality genomic data for accuracy, underscoring the necessity of experimental verification. Building on this, our study reinforces that orphan and singleton BGCs represent a vast, underexplored reservoir of biosynthetic potential. Focus on investigating them could pave the way for discovering new bioactive compounds of significant biotechnological value.

ACKNOWLEDGEMENTS

This work was supported by grants from the Brazilian funding agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq - 312994/2021-4 and 405015/2023-2), Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS - 19/2551-0001708-1), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). AAR was a recipient of a CAPES scholarships.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest that could be perceived as prejudicial to the impartiality of the reported research.

AUTHORS CONTRIBUTIONS

AAR Conceptualization, Methodology, Investigation, Formal Analysis, Writing and Original Draft Preparation; CCS Supervision, Formal Analysis, Writing – Review and Editing. All authors read and approved the final version.

REFERENCES

Aghdam SA and Brown AMV (2021) Deep learning approaches for natural product discovery from plant endophytic microbiomes. Environ Microbiome 16:6. doi: 10.1186/s40793-021-00375-0

Alviti Kankanamalage HP, Yang J-Y, Karunarathna SC, Tibpromma S, Kumla J, Wei D-P and Lumyong S (2025) Entomopathogenic fungi: insights into recent understanding. World J Microbiol Biotechnol 41:179. doi: 10.1007/s11274-025-04377-9

Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH and Weber T (2021) antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res 49:W29–W35. doi: 10.1093/nar/gkab335

Chen S, Zhou Y, Chen Y and Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. doi: 10.1093/bioinformatics/bty560

de Miranda RPR, Soares TKDA, Castro DP and Genta FA (2024) General aspects, host interaction, and application of Metarhizium sp. in arthropod pest and vector control. Front Fungal Biol 5:1456964. doi: 10.3389/ffunb.2024.1456964

de Sousa TK, Silva AT da and Soares FE de F (2025) Fungi-Based Bioproducts: A Review in the Context of One Health. Pathogens. doi: 10.3390/pathogens14050463

Gu Y, Fujitomo Y, Soeda H, Nakahama C, Hasegawa N, Maesaki S, Maeda M, Matsumoto T, Miyairi I and Ohmagari N (2020) A nationwide questionnaire survey of clinic doctors on antimicrobial stewardship in Japan. J Infect Chemother 26:149–156. doi: 10.1016/j.jiac.2019.12.005

Hong S, Shang J, Sun Y, Tang G and Wang C (2024) Fungal infection of insects: molecular insights and prospects. Trends Microbiol 32:302–316. doi: 10.1016/j.tim.2023.09.005

Jinna Z, Khan T, Naima Kanwal and Wang Zhenji (2025) Diversity and biological functions of fungal secondary metabolites: Biocontrol agents for sustainable

agriculture. A review. Cell Mol Biol (Noisy-le-grand) 71:43–48. doi: 10.14715/cmb/2025.71.5.7

Keller NP (2019) Fungal secondary metabolism: regulation, function and drug discovery. Nat Rev Microbiol 17:167–180. doi: 10.1038/s41579-018-0121-1

Manni M, Berkeley MR, Seppey M, Simão FA and Zdobnov EM (2021) BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. Mol Biol Evol 38:4647–4654. doi: 10.1093/molbev/msab199

Meesil W, Muangpat P, Sitthisak S, Rattanarojpong T, Chantratita N, Machado RAR, Shi Y-M, Bode HB, Vitta A and Thanwisai A (2023) Genome mining reveals novel biosynthetic gene clusters in entomopathogenic bacteria. Sci Rep 13:20764. doi: 10.1038/s41598-023-47121-9

Muggia L, Ametrano CG, Sterflinger K and Tesei D (2020) An overview of genomics, phylogenomics and proteomics approaches in ascomycota. Life (Basel). doi: 10.3390/life10120356

Naranjo-Ortiz MA and Gabaldón T (2019) Fungal evolution: major ecological adaptations and evolutionary transitions. Biol Rev Camb Philos Soc 94:1443–1476. doi: 10.1111/brv.12510

Navarro-Muñoz JC, Selem-Mojica N, Mullowney MW, Kautsar SA, Tryon JH, Parkinson EI, De Los Santos ELC, Yeong M, Cruz-Morales P, Abubucker S et al. (2020) A computational framework to explore large-scale biosynthetic diversity. Nat Chem Biol 16:60–68. doi: 10.1038/s41589-019-0400-9

O'Leary NA, Cox E, Holmes JB, Anderson WR, Falk R, Hem V, Tsuchiya MTN, Schuler GD, Zhang X, Torcivia J et al. (2024) Exploring and retrieving sequence and metadata for species across the tree of life with NCBI Datasets. Sci Data 11:732. doi: 10.1038/s41597-024-03571-y

Pedrini N (2018) Molecular interactions between entomopathogenic fungi (Hypocreales) and their insect host: Perspectives from stressful cuticle and

hemolymph battlefields and the potential of dual RNA sequencing for future studies. Fungal Biol 122:538–545. doi: 10.1016/j.funbio.2017.10.003

Qin Y, Liu X, Peng G, Xia Y and Cao Y (2023) Recent advancements in pathogenic mechanisms, applications and strategies for entomopathogenic fungi in mosquito biocontrol. J Fungi (Basel). doi: 10.3390/jof9070746

Quesada-Moraga E, Garrido-Jurado I, González-Mas N and Yousef-Yousef M (2023) Ecosystem services of entomopathogenic ascomycetes. J Invertebr Pathol 201:108015. doi: 10.1016/j.jip.2023.108015

Riedling O, Walker AS and Rokas A (2024) Predicting fungal secondary metabolite activity from biosynthetic gene cluster data using machine learning. Microbiol Spectr 12:e0340023. doi: 10.1128/spectrum.03400-23

Rinker DC, Sauters TJC, Steffen K, Gumilang A, Raja HA, Rangel-Grimaldo M, Pinzan CF, de Castro PA, Dos Reis TF, Delbaje E et al. (2024) Strain heterogeneity in a non-pathogenic Aspergillus fungus highlights factors associated with virulence. Commun Biol 7:1082. doi: 10.1038/s42003-024-06756-8

Shahbaz M, Palaniveloo K, Tan YS, Palasuberniam P, Ilyas N, Wiart C and Seelan JSS (2024) Entomopathogenic fungi in crops protection with an emphasis on bioactive metabolites and biological activities. World J Microbiol Biotechnol 40:217. doi: 10.1007/s11274-024-04022-x

van den Belt M, Gilchrist C, Booth TJ, Chooi Y-H, Medema MH and Alanjary M (2023) CAGECAT: The CompArative GEne Cluster Analysis Toolbox for rapid search and visualisation of homologous gene clusters. BMC Bioinformatics 24:181. doi: 10.1186/s12859-023-05311-2

Xu Z, Park T-J and Cao H (2023) Advances in mining and expressing microbial biosynthetic gene clusters. Crit Rev Microbiol 49:18–37. doi: 10.1080/1040841X.2022.2036099

Ying S-H (2024) Subcellular biochemistry and biology of filamentous entomopathogenic fungi. Adv Appl Microbiol 129:35–58. doi: 10.1016/bs.aambs.2024.04.002

Zaccaron AZ and Stergiopoulos I (2025) The dynamics of fungal genome organization and its impact on host adaptation and antifungal resistance. J Genet Genomics 52:628–640. doi: 10.1016/j.jgg.2024.10.010

Zhang L, Yue Q, Wang C, Xu Y and Molnár I (2020) Secondary metabolites from hypocrealean entomopathogenic fungi: genomics as a tool to elucidate the encoded parvome. Nat Prod Rep 37:1164–1180. doi: 10.1039/d0np00007h

Zhgun AA (2023) Fungal bgcs for production of secondary metabolites: main types, central roles in strain improvement, and regulation according to the piano principle. Int J Mol Sci. doi: 10.3390/ijms241311184

FIGURE LEGENDS

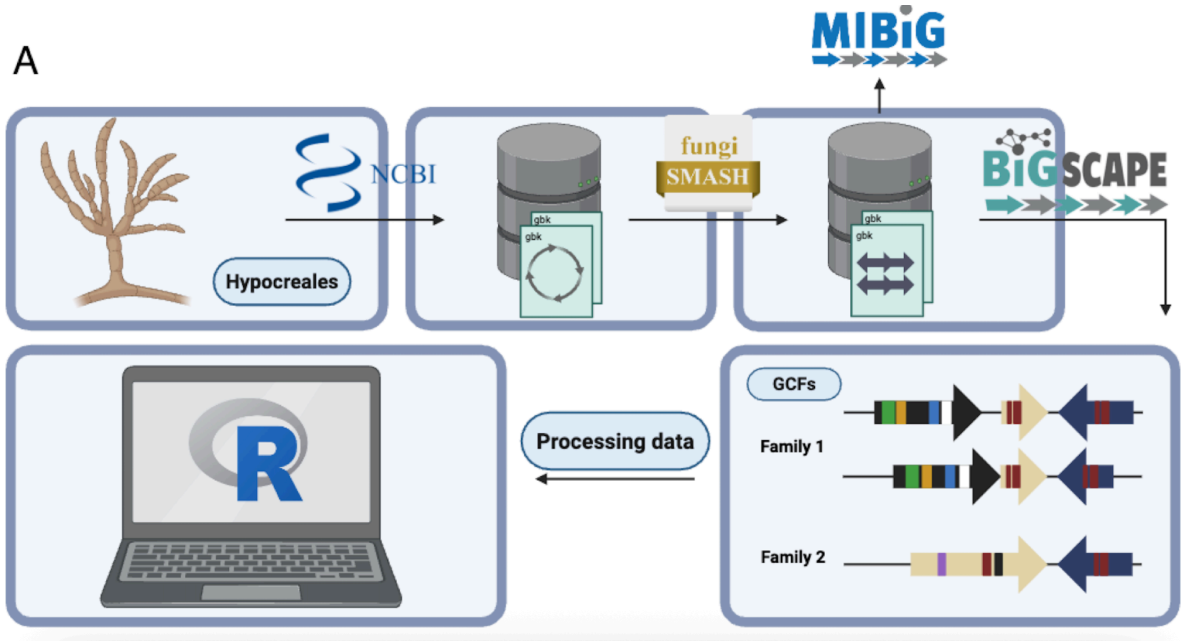


Figure 1: Schematic illustration of the methodology for inferring (A) BGCs and GCFs in Hypocrealean genomes.

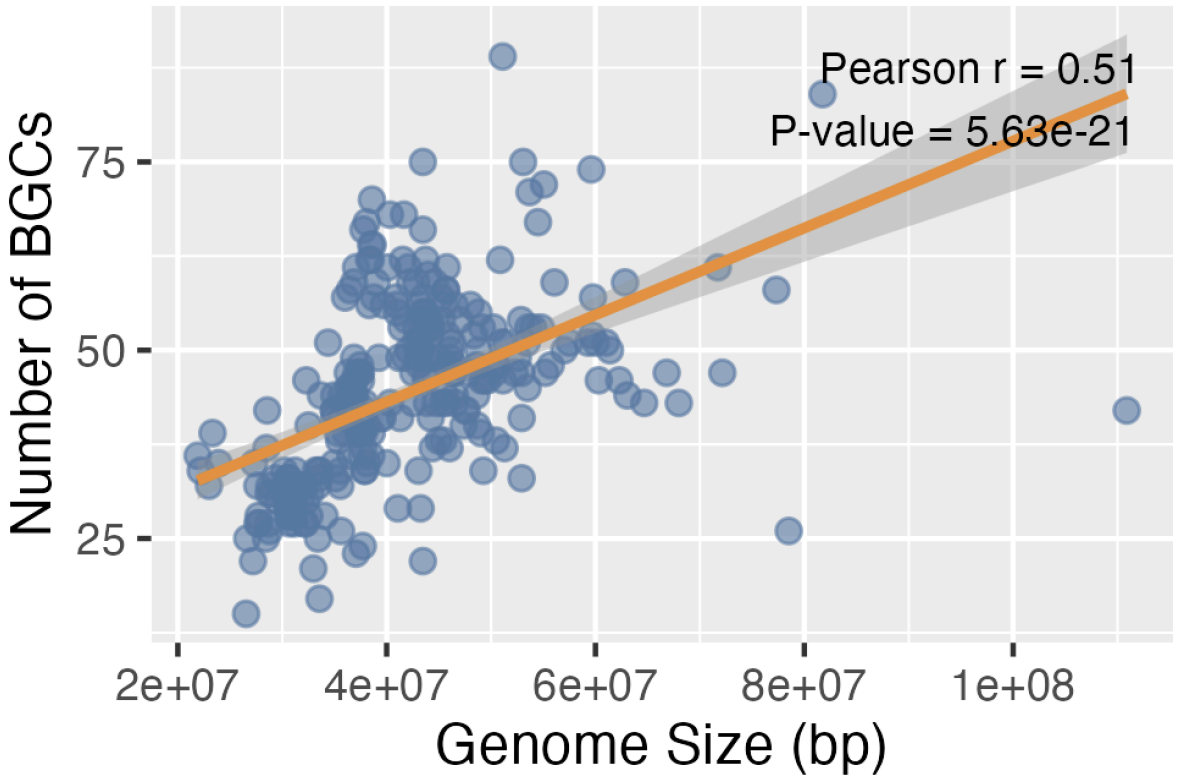


Figure 2: Genome size does not correlate with the number of biosynthetic gene clusters (BGCs) in Hypocrealean genomes. Each dot represents a fungal genome, plotted by its total genome size (Mb) and the number of predicted BGCs. A linear regression model was fitted to assess the relationship between genome size and BGC count. The Person coefficient of correlation (r) and *P*-value were calculated using a simple linear regression analysis.

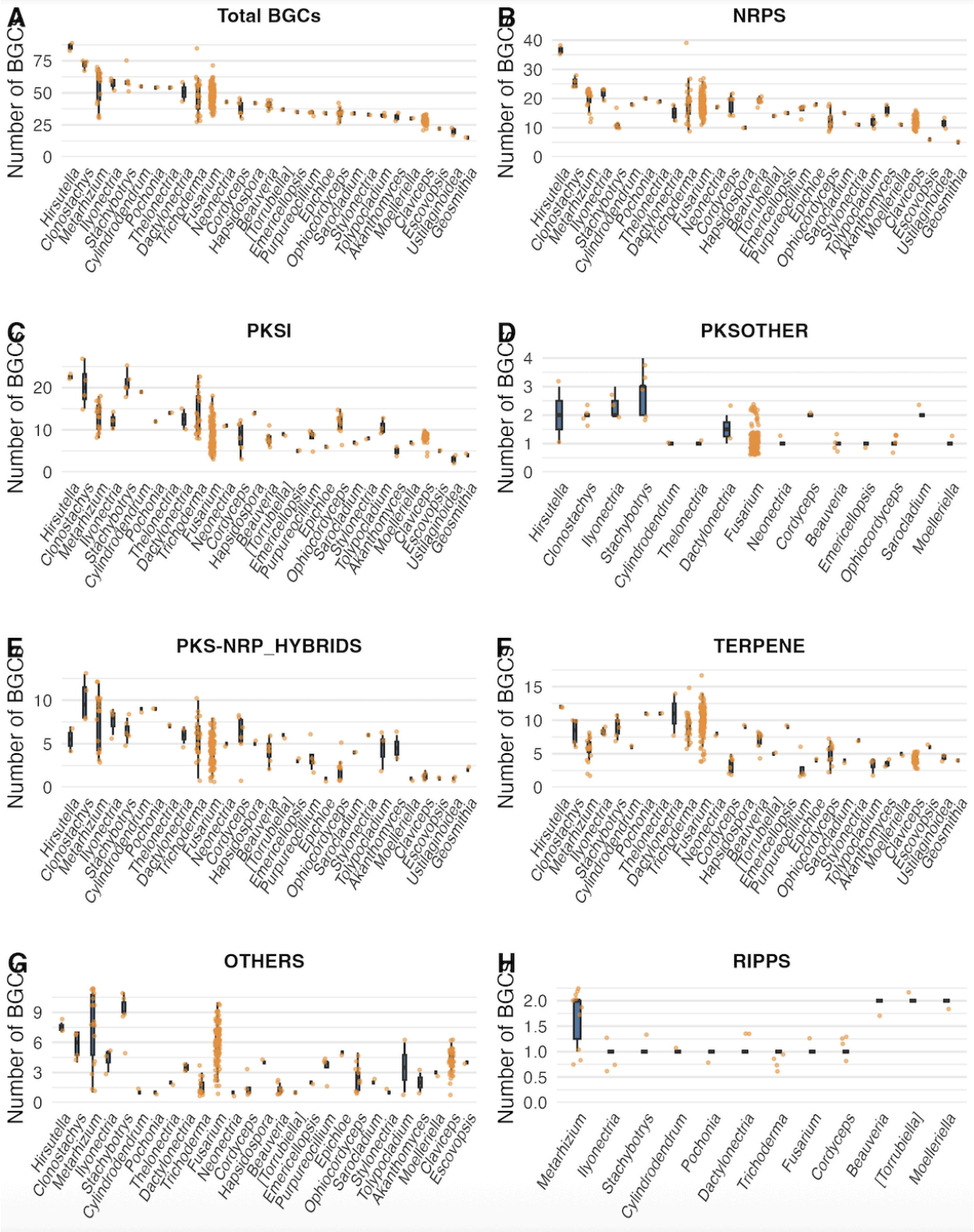


Figure 3: Distribution and abundance of predicted biosynthetic gene clusters (BGCs) across fungal genera in Hypocreales. The number and types of predicted biosynthetic gene clusters (BGCs) identified in the genomes of various fungal species, grouped by genus, within the order Hypocreales and related taxa. Data are presented as box plots, where the box represents the interquartile range (IQR), the

horizontal line within the box indicates the median, and the whiskers extend to 1.5 times the IQR from the box edges. Individual data points (orange circles) represent the number of BGCs detected in each genome analyzed and are jittered horizontally to reduce overlap. Panels (A) through (H) show the distribution for different BGC classes: **(A)** Total BGCs, **(B)** NRPS, **(C)** PKSI, **(D)** PKSOTHER, **(E)**PKS-NRP HYBRIDS, **(F)** TERPENE, **(G)** OTHERS, and **(H)** RIPPS.

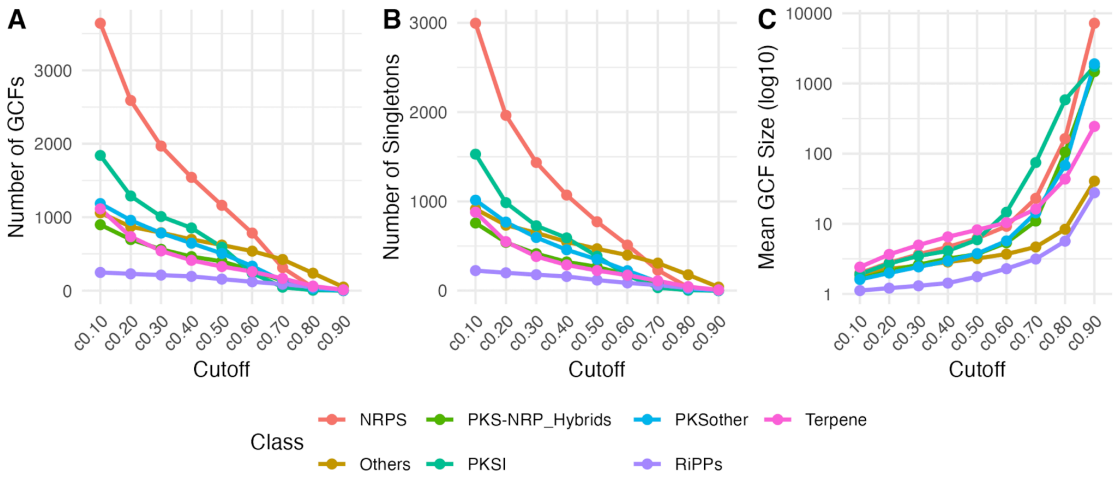


Figure 4. Determination of cutoff value for assignment of BGCs to GCFs. Rarefaction curves were constructed using distinct cutoff values, and the number of GCFs (A), Singletons (B), as well as the average GCF size (C) were determined. Cutoff 0.6 was selected for its optimal balance between sensitivity and specificity for proper BGC classification.

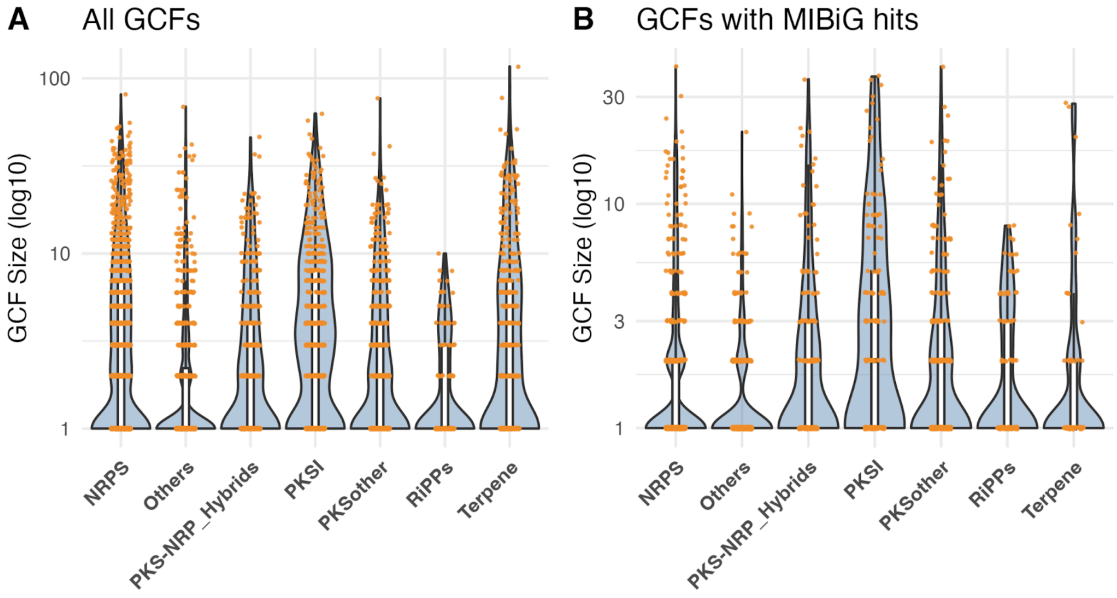


Figure 5. Distribution and comparison of Gene Cluster Family (GCF) sizes across biosynthetic classes at a cutoff of 0.60. (A) GCF Size Distribution of All GCFs: Size distribution for all predicted GCFs identified in the dataset. **(B) GCF Size Distribution of GCFs with MiBIG Entries:** Size distribution specifically for those GCFs that contain at least one BGC that is also present in the MiBIG database.
Violin plots overlaid with box plots, showing the distribution of GCF sizes (on a log10 scale) for each biosynthetic gene cluster (BGC) class, using the BiG-SCAPE clustering results with a cutoff value of 0.60. The box plot inside each violin shows the median (white line), the interquartile range (box), and whiskers extending to 1.5 times the interquartile range. Orange points represent individual GCF sizes, jittered horizontally for visibility.

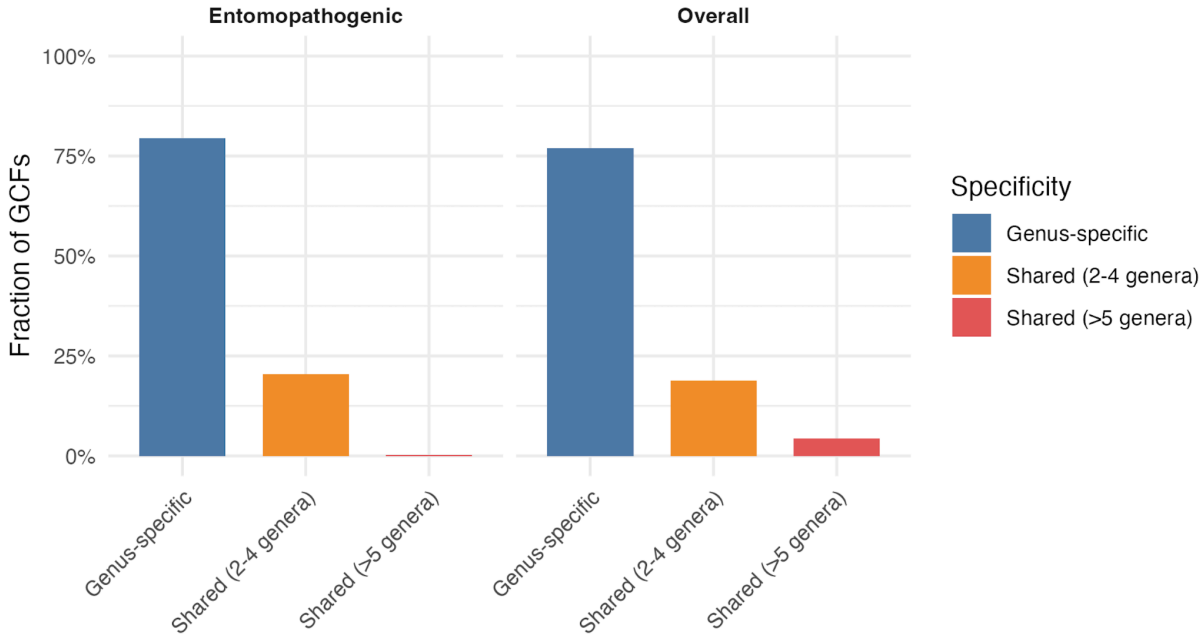


Figure 6: Distribution of Gene Cluster Family (GCF) taxonomic specificity in entomopathogenic and overall datasets. Bar plots show the fraction of GCFs classified according to their taxonomic specificity in entomopathogenic species (left) and across all species (right).

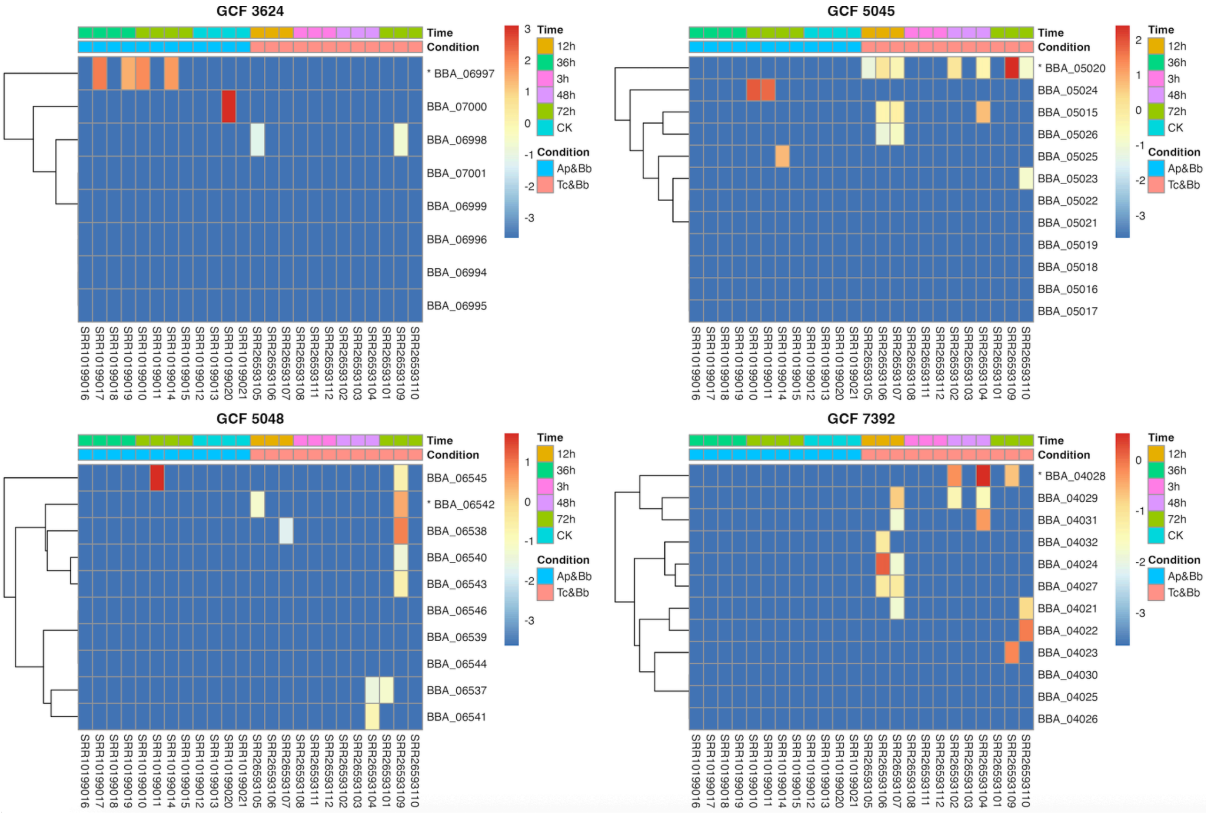


Figure 7. Expression patterns of Beauveria bassiana biosynthetic gene clusters (BGCs) within gene cluster families (GCFs) during insect infection. Heatmaps display normalized gene expression levels (DESeq2-normalized counts) for genes within *B. bassiana* (Bb) BGCs across different time points of infection in two insect hosts: the pea aphid (*Acyrthosiphon pisum* - Ac) and the red flour beetle (*Tribolium castaneum* - Tc). RNA-Seq data were retrieved from the Sequence Read Archive (SRA) and aligned to the *B. bassiana* ARSEF 2860 genome. Each panel represents one of the evaluated BGCs from each GCF. The columns indicate samples grouped by time point and host, and rows represent individual genes within the BGC.

SUPPLEMENTARY MATERIAL

Table S1: Genome sequences analyzed, size, and BGC counts.

Table S2: Power law analysis results.

Table S3: Exclusive GCFs from entomopathogenic genera.

File S1: Supplementary figures file.

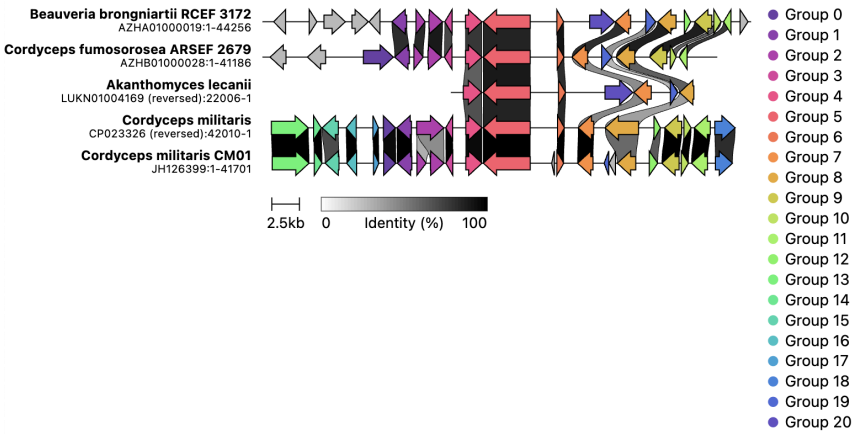


Figure S1: Identity analysis of GCF 2292. Identity analysis of 5 members of GCF 2292.

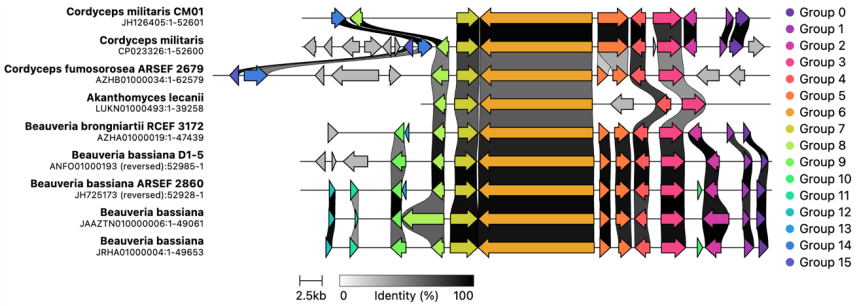


Figure S2: Identity analysis of GCF 3615. Identity analysis of 9 members of GCF 3615.

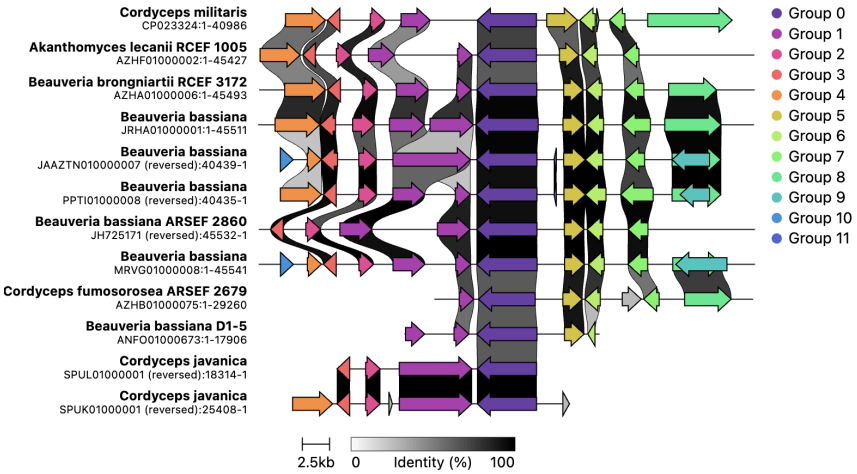


Figure S3: Identity analysis of GCF 3624. Identity analysis of 12 members of GCF 3624.

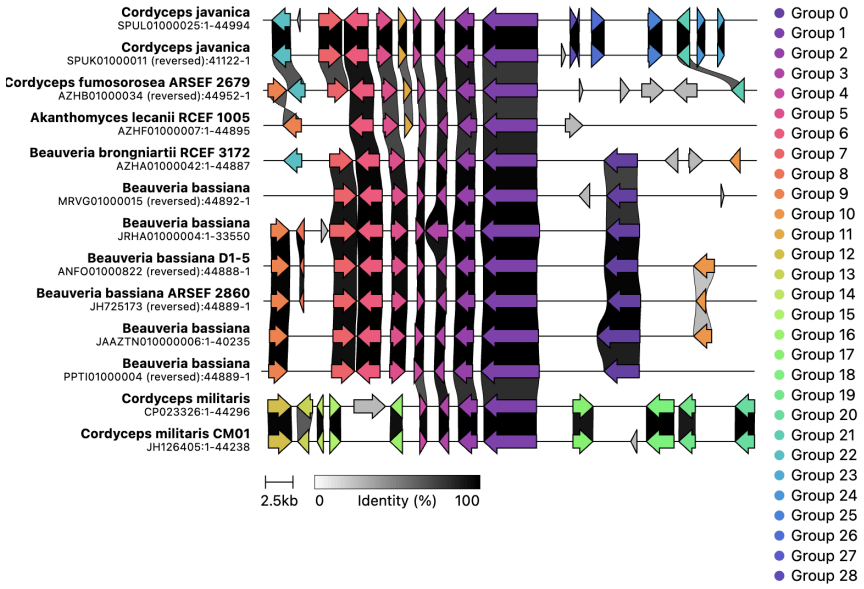


Figure S4: Identity analysis of GCF 3634. Identity analysis of 13 members of GCF 3634.

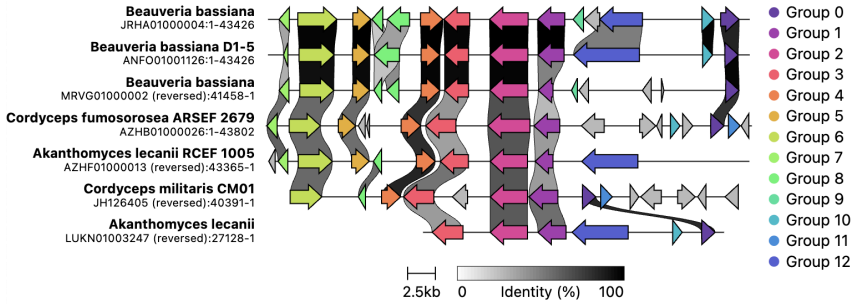


Figure S5: Identity analysis of GCF 3644. Identity analysis of 7 members of GCF 3644.

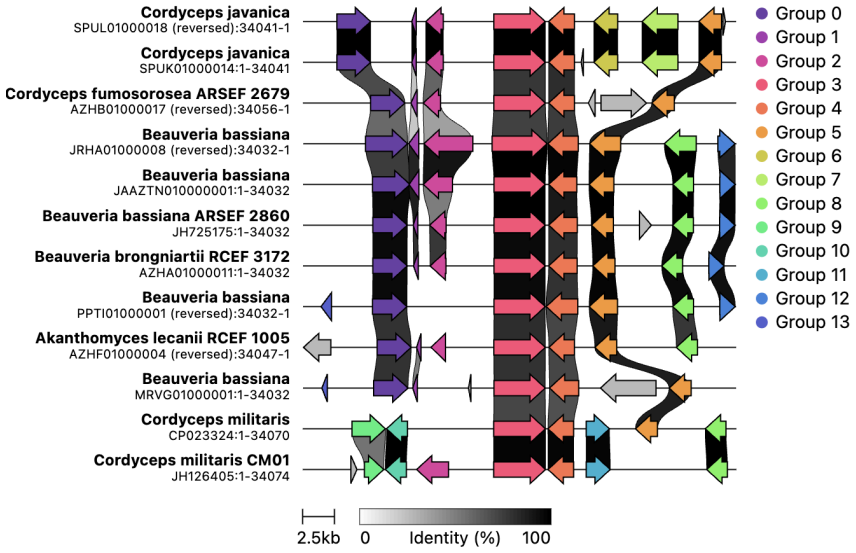


Figure S6: Identity analysis of GCF 4693. Identity analysis of 12 members of GCF 4693.

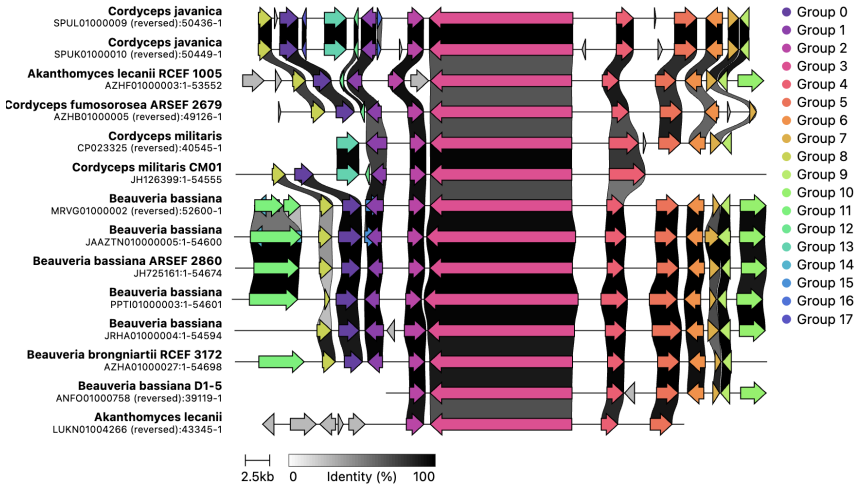


Figure S7: Identity analysis of GCF 5045. Identity analysis of 14 members of GCF 5045.

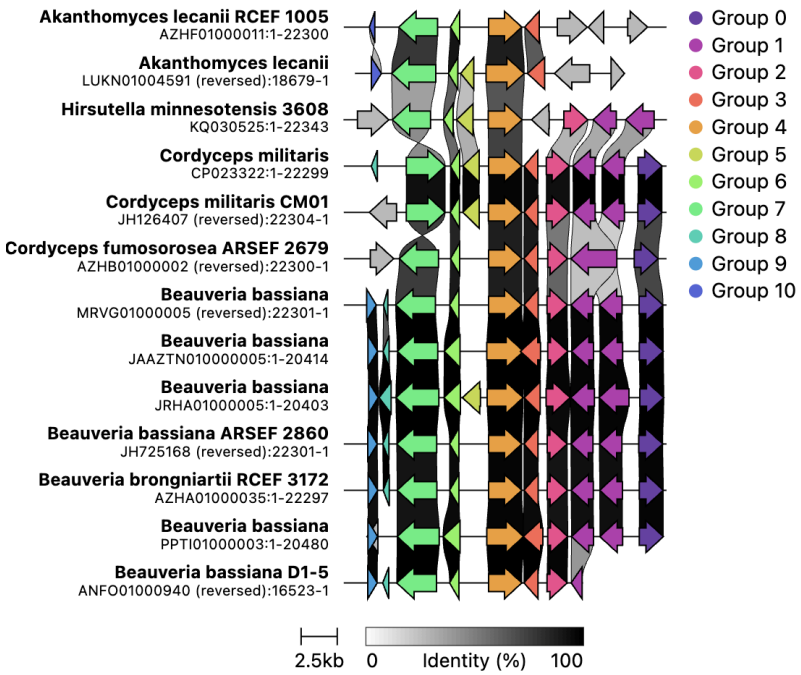


Figure S8: Identity analysis of GCF 5048. Identity analysis of 13 members of GCF 5048.

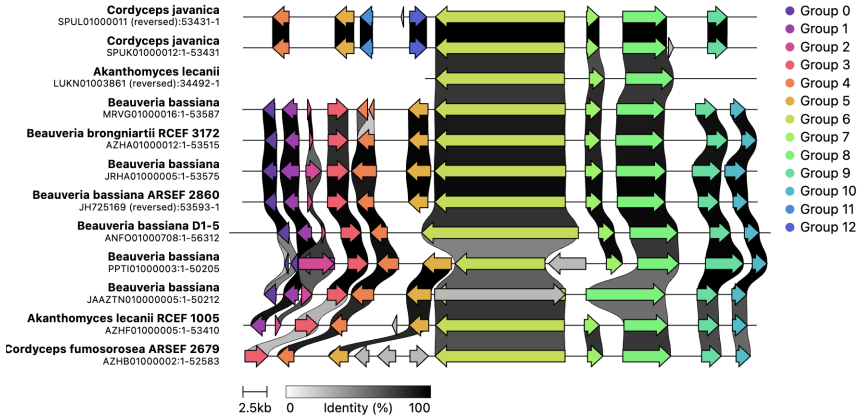


Figure S9: Identity analysis of GCF 5052. Identity analysis of 12 members of GCF 5052.

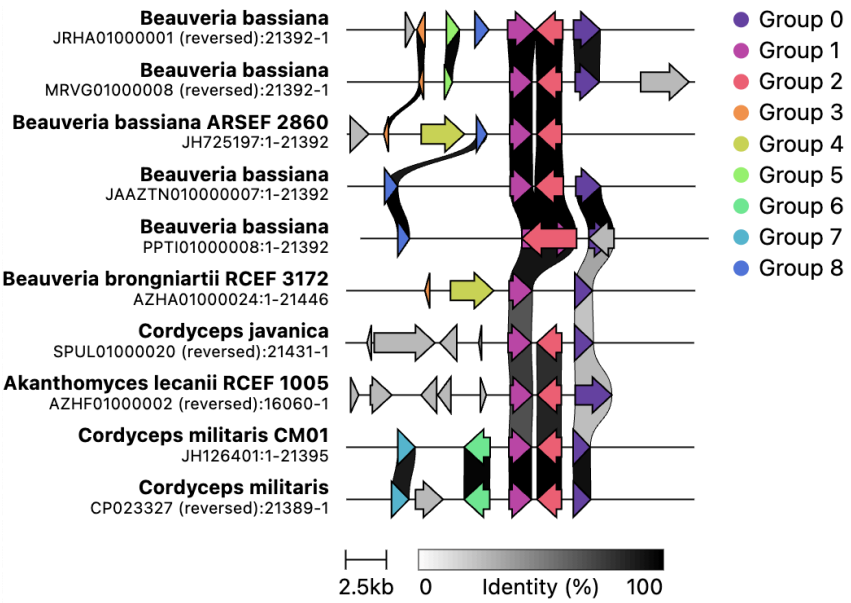


Figure S10: Identity analysis of GCF 5402. Identity analysis of 10 members of GCF 5402.

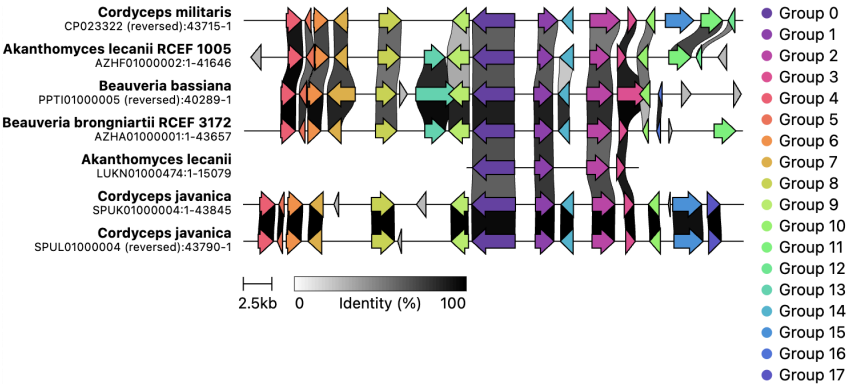


Figure S11: Identity analysis of GCF 7082. Identity analysis of 7 members of GCF 7082.

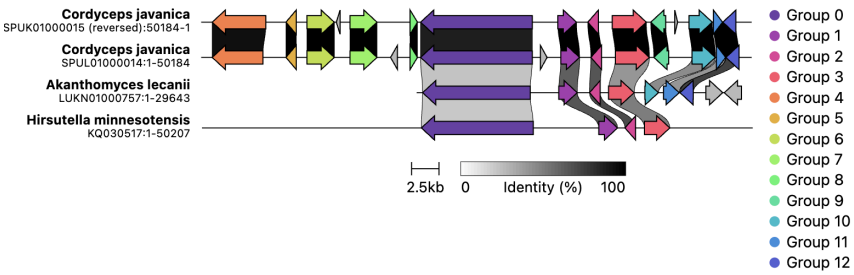


Figure S12: Identity analysis of GCF 7102. Identity analysis of 4 members of GCF 7102.

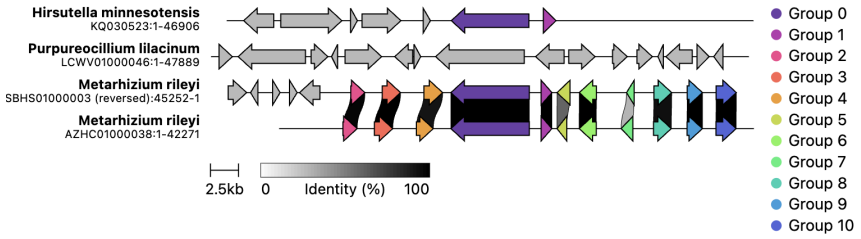


Figure S13: Identity analysis of GCF 7120. Identity analysis of 4 members of GCF 7120.

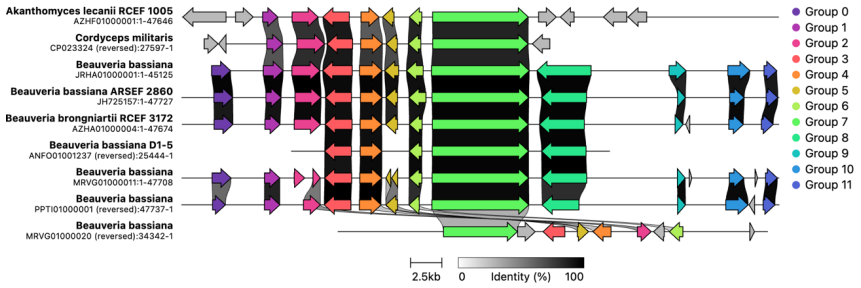


Figure S14: Identity analysis of GCF 7366. Identity analysis of 9 members of GCF 7366.

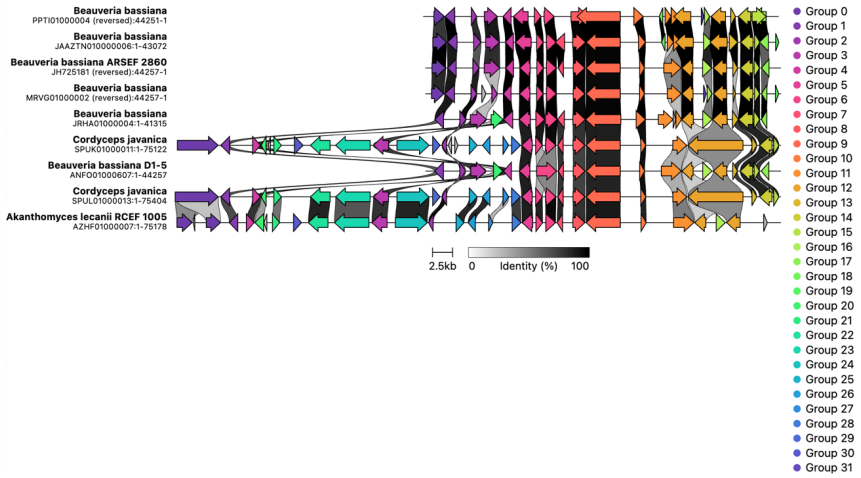


Figure S15: Identity analysis of GCF 7384. Identity analysis of 9 members of GCF 7384.

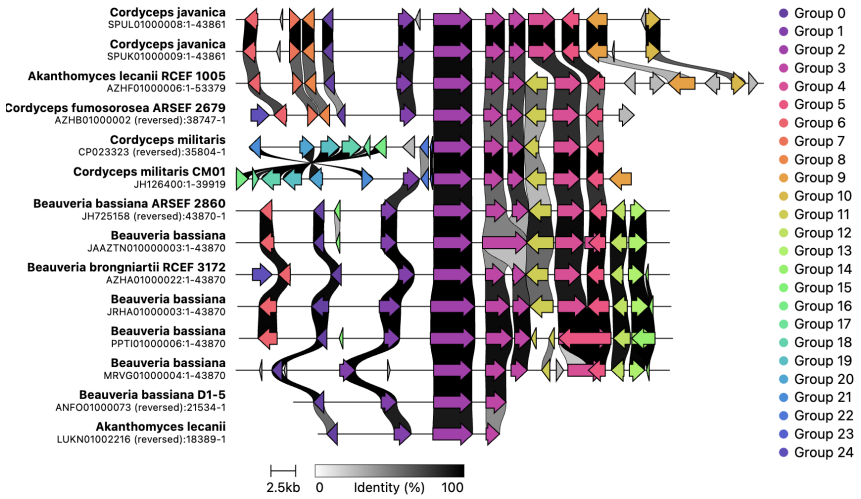


Figure S16: Identity analysis of GCF 7392. Identity analysis of 14 members of GCF 7392.

4. DISCUSSÃO GERAL

A necessidade de uma diminuição no uso de agrotóxicos nocivos na agricultura, considerando os danos causados ao meio ambiente e à saúde coletiva, faz com que o uso do controle biológico (CB) de pragas seja uma interessante alternativa e que estudos na área, visando um melhor aprimoramento da técnica, aumente. Também é muito importante que a escolha do micro patógeno designado para atuar no CB traga bons resultados, de forma que o "revezamento" em relação aos métodos químicos acarrete resultados tão eficazes, e menos danosos, quanto os alegados por quem usa pesticidas químicos (ECKSTEIN et al., 2022).

Os biopesticidas que possuem fungos entomopatogênicos (FEPs), fungos com capacidade de infectar artrópodes, como matéria prima, vem apresentando resultados significativos no mercado mundial da agricultura, se mostrando uma alternativa sustentável e econômica, gerando menos resíduos e também aumentando o rendimento da produção (MANTZOUKAS et al., 2022). Apesar disso, é importante ressaltar a necessidade de atenção no manejo desses patógenos, já que em condições não adequadas, podem ter sua eficiência reduzida (DONG et al., 2016), e por apresentarem um processo de infecção complexo, qualquer condição não favorável pode interferir negativamente nos resultados esperados (HONG et al., 2023).

Assim como *Beauveria*, *Metarhizium* é um dos gêneros de fungos entomopatogênicos mais estudados quando o assunto é controle biológico de pragas (ROBERTS; ST LEGER, 2004). Ambos fazem parte da ordem Hypocreales, que é reconhecida principalmente por incluir fungos com diferentes estilos de vida, caracterizada como a ordem mais importante dentro do filo Ascomycota (ZHANG et al., 2018). São essas peculiaridades que tornam estudos sobre Hypocreales importantes, principalmente para um maior entendimento de como fungos tão diferentes podem compartilhar uma mesma classificação (HONG et al., 2023). Essas diferenças podem ter distintas raízes, porém, já se sabe que os metabólitos secundários (MSs), moléculas que os fungos filamentosos são altamente capazes de produzir, estão atrelados com a presença de características individuais apresentadas por diferentes espécies (KJÆRBØLLING et al., 2019).

Além de se mostrarem moléculas importantes para distintas interações ecológicas entre outros organismos, já existem dados na literatura que mencionam os MSs produzidos por FEPs como determinantes de virulência, característica principal destes fungos (AW; HUE, 2017). Já se sabe que os MSs são produzidos por *clusters* gênicos biossintéticos (BGCs) (KWON et al., 2021), porém ainda pouco se sabe sobre como e quais BGCs estão

ligados diretamente com as capacidades entomopatogênicas dos fungos em questão, e se a presença de determinados BGCs em um genoma fúngico, tornará àquele fungo capaz de infectar insetos também.

A partir disso, investigar *in silico* evidências de conservação de BGCs entre genomas fúngicos da ordem Hypocreales se torna processo crucial para a busca de novas espécies com potencial entomopatogênico, assim como o já descrito em gêneros como *Beauveria* e *Metarhizium.* Porém, com os resultados obtidos e análises realizadas ao longo deste trabalho, além de não ser encontrada conservação de BGCs de diferentes espécies com os principais gêneros relatados como entomopatogênicos, também não encontrou-se conservação entre os dois gêneros modelo. Essas divergências interespécies corroboram com o que já se tem na literatura sobre *Metarhizium,* alegando que houveram pressões evolutivas para adaptação e especiação, que ao ser acoplada a uma série de características genômicas, promoveram a formação de novas espécies. Esses eventos coincidiram com a disponibilidade e diversificação de insetos, de maneira que o gênero se sentiu pressionado a adaptar-se para diferentes estilos de vida, buscando sucesso na sobrevivência. Desta maneira, acredita-se que essas evidências expliquem como dentro de um mesmo gênero possam existir espécies com capacidades tão distintas (HU et al., 2014), assim como a observação de uma alta diversidade de BGCs em uma população, seguida de uma baixa frequência dos mesmos, gera uma maior diversificação - explicando como dentro de uma mesma ordem possa haver gêneros com habilidades tão diferentes (JENSEN, 2016).

Considerando os resultados obtidos através das análises de conversação realizadas, direcionamos nosso trabalho conforme estudos anteriores e interesse do grupo ao que faço parte (STAATS et al., 2014). Através dos genes obtidos dos genomas das espécies de *M. anisopliae* E6 recuperados, selecionamos aqueles que demonstraram bons resultados na análise de expressão gênica quando comparados às amostras publicadas anteriormente por Li et al, considerando as condições de nosso interesse (LI et al., 2022). Nossa análise resultou em 10 genes de importância, identificados como códigos MANI, que foram determinados através dos resultados das análises de expressão gênica e presença na tabela de genes *backbone* publicada em 2016 por Sbaraini et al (SBARAINI et al., 2016).

Desta forma, este estudo mostra que, apesar de não haver conservação entre os BGCs analisados dos FEPs presentes na ordem Hypocreales, é alto o potencial de diferentes genes de espécies modelo (como *Metarhizium* e *Beauveria*) estarem envolvidos nos fatores de virulência desses organismos, mostrando que estudos a fundo de espécies já muito relatadas também são essenciais para um aprimoramento do CB. Além disso, o aumento da

disponibilidade de sequenciamento genômico de distintas espécies fúngicas é fundamental para eficiência deste tipo de pesquisa, de maneira que a busca por novos MSs ativos úteis *in silico* seja muito mais preciso, trazendo bons resultados para a agricultura, economia e também para diferentes indústrias biotecnológicas (KJÆRBØLLING et al., 2019).

5. CONCLUSÕES

- Diferentes análises *in silico* evidenciaram não haver conservação de BGCs em espécies fúngicas da ordem Hypocreales, nem mesmo entre os gêneros modelo *Metarhizium* e *Beauveria*, que compartilham capacidades entomopatogênicas.
- A identificação de 14272 BGCs em 338 sequências genômicas fúngicas analisadas corroboram com a literatura que afirma que quando há uma elevada diversidade de BGCs, uma diversificação considerada de capacidades será encontrada naquele grupo taxonômico.
- Em *M. anisopliae* E6, foram identificados através de análises de expressão gênica, dez genes com potencial de codificar proteínas associadas à virulência deste fungo, mostrando que pesquisas a fundo em espécies já muito estudadas sempre serão relevantes para obtenção de resultados mais precisos.

6. PERSPECTIVAS

- Realizar análises *in silico* refinadas entre os BGCs presentes nas espécies de *M. anisopliae* E6 que apresentaram genes de interesse, a fim de encontrar conservação e justificativas nestas moléculas que expliquem as capacidades entomopatogênicas preditas;
- Confirmar, *in vitro*, se o potencial de virulência predito através das análises de expressão gênica se mantém.

7. REFERÊNCIAS BIBLIOGRÁFICAS

AW, K. M. S.; HUE, S. M. Mode of Infection of Metarhizium spp. Fungus and Their Potential as Biological Control Agents. **Journal of fungi (Basel, Switzerland)**, v. 3, n. 2, 7 Jun. 2017.

BAR-ON, Y. M.; PHILLIPS, R.; MILO, R. The biomass distribution on Earth. **Proceedings of the National Academy of Sciences of the United States of America**, v. 115, n. 25, p. 6506–6511, 19 Jun. 2018.

BERBEE, M. L. The phylogeny of plant and animal pathogens in the Ascomycota. **Physiological and Molecular Plant Pathology**, v. 59, n. 4, p. 165–187, Oct. 2001.

BLIN, K. et al. antiSMASH 6.0: improving cluster detection and comparison capabilities. **Nucleic Acids Research**, v. 49, n. W1, p. W29–W35, 2 Jul. 2021.

BRAKHAGE, A. A. Regulation of fungal secondary metabolism. **Nature Reviews. Microbiology**, v. 11, n. 1, p. 21–32, Jan. 2013.

BRASIL. Manual de procedimentos para o registro de agrotóxicos. **Coordenação Geral de Agrotóxico e Afins**, v. 1, p. 68, 2012.

BUTT, T. M. et al. Entomopathogenic Fungi: New Insights into Host-Pathogen Interactions. **Advances in genetics**, v. 94, p. 307–364, 9 Mar. 2016.

CHAVALI, A. K.; RHEE, S. Y. Bioinformatics tools for the identification of gene clusters that biosynthesize specialized metabolites. **Briefings in Bioinformatics**, v. 19, n. 5, p. 1022–1034, 28 Sep. 2018.

CHROUMPI, T.; MÄKELÄ, M. R.; DE VRIES, R. P. Engineering of primary carbon metabolism in filamentous fungi. **Biotechnology advances**, v. 43, p. 107551, 1 Nov. 2020.

DONG, T. et al. Isolation and Classification of Fungal Whitefly Entomopathogens from Soils of Qinghai-Tibet Plateau and Gansu Corridor in China. **Plos One**, v. 11, n. 5, p. e0156087, 26 May 2016.

ECKSTEIN, D. et al. Who Suffers Most from Extreme Weather Events? Weather-Related Loss Events in 2018 and 1999 to 2018. **GERMANWATCH**, 2022.

FARIA, M. R. DE; WRAIGHT, S. P. Mycoinsecticides and Mycoacaricides: A comprehensive list with worldwide coverage and international classification of formulation types. **Biological Control**, v. 43, n. 3, p. 237–256, Dec. 2007.

F A O. **Food and Agriculture Organization of the United Nations**. Disponível em: <https://www.fao.org/soils-portal/about/all-definitions/en/>. Acesso em: 29 dec. 2023.

HAWKSWORTH, D. L.; LÜCKING, R. Fungal diversity revisited: 2.2 to 3.8 million species. **Microbiology spectrum**, v. 5, n. 4, Jul. 2017.

HONG, S. et al. Fungal infection of insects: molecular insights and prospects. **Trends in Microbiology**, 29 Sep. 2023.

HU, X. et al. Trajectory and genomic determinants of fungal-pathogen speciation and host adaptation. **Proceedings of the National Academy of Sciences of the United States of America**, v. 111, n. 47, p. 16796–16801, 25 Nov. 2014.

IMOULAN, A. et al. Entomopathogenic fungus Beauveria: Host specificity, ecology and significance of morpho-molecular characterization in accurate taxonomic classification. **Journal of Asia-Pacific entomology**, v. 20, n. 4, p. 1204–1212, Dec. 2017.

INGLIS, D. O. et al. Comprehensive annotation of secondary metabolite biosynthetic genes and gene clusters of Aspergillus nidulans, A. fumigatus, A. niger and A. oryzae. **BMC Microbiology**, v. 13, p. 91, 26 Apr. 2013.

JENSEN, P. R. Natural products and the gene cluster revolution. **Trends in Microbiology**, v. 24, n. 12, p. 968–977, Dec. 2016.

KHALDI, N. et al. SMURF: Genomic mapping of fungal secondary metabolite clusters. **Fungal Genetics and Biology**, v. 47, n. 9, p. 736–741, Sep. 2010.

KJÆRBØLLING, I. et al. Strategies to establish the link between biosynthetic gene clusters and secondary metabolites. **Fungal Genetics and Biology**, v. 130, p. 107–121, Sep. 2019.

KWON, M. J. et al. Beyond the Biosynthetic Gene Cluster Paradigm: Genome-Wide Coexpression Networks Connect Clustered and Unclustered Transcription Factors to Secondary Metabolic Pathways. **Microbiology spectrum**, v. 9, n. 2, p. e0089821, 31 Oct. 2021.

LI, C. et al. Mutation of a prenyltransferase results in accumulation of subglutinols and destruxins and enhanced virulence in the insect pathogen, Metarhizium anisopliae. **Environmental Microbiology**, v. 24, n. 3, p. 1362–1379, Mar. 2022.

LITWIN, A.; NOWAK, M.; RÓŻALSKA, S. Entomopathogenic fungi: unconventional applications. **Reviews in Environmental Science and Bio/Technology**, v. 19, n. 1, p. 23–42,

Mar. 2020.

MANTZOUKAS, S. et al. Entomopathogenic fungi: interactions and applications. **Encyclopedia**, v. 2, n. 2, p. 646–656, 24 Mar. 2022.

MASCARIN, G. M.; JARONSKI, S. T. The production and uses of *Beauveria bassiana* as a microbial insecticide. **World journal of microbiology & biotechnology**, v. 32, n. 11, p. 177, Nov. 2016.

MORAES, R. F. DE ET AL. Agrotóxicos no Brasil: padrões de uso,política da regulação e prevenção da captura regulatória. 2019.

PARRA, J. R. P. et al. **Controle Biológico no Brasil com parasitoides e predadores naagricultura brasileira.** 1. ed. Piracicaba: [s.n.]. v. 1

PROCTOR, R. H. et al. Evidence that a secondary metabolic biosynthetic gene cluster has grown by gene relocation during evolution of the filamentous fungus Fusarium. **Molecular Microbiology**, v. 74, n. 5, p. 1128–1142, Dec. 2009.

ROBERTS, D. W.; ST LEGER, R. J. Metarhizium spp., cosmopolitan insect-pathogenic fungi: mycological aspects. **Advances in Applied Microbiology**, v. 54, p. 1–70, 2004.

SBARAINI, N. et al. Secondary metabolite gene clusters in the entomopathogen fungus Metarhizium anisopliae: genome identification and patterns of expression in a cuticle infection model. **BMC Genomics**, v. 17, n. Suppl 8, p. 736, 25 Oct. 2016.

SCHRANK, A.; VAINSTEIN, M. H. Metarhizium anisopliae enzymes and toxins. **Toxicon**, v. 56, n. 7, p. 1267–1274, 15 Dec. 2010.

SHENG, H.; MCNAMARA, P. J.; ST LEGER, R. J. Metarhizium: an opportunistic middleman for multitrophic lifestyles. **Current Opinion in Microbiology**, v. 69, p. 102176, Oct. 2022.

STAATS, C. C. et al. Comparative genome analysis of entomopathogenic fungi reveals a complex set of secreted proteins. **BMC Genomics**, v. 15, p. 822, 29 Sep. 2014.

ST LEGER, R. J.; WANG, J. B. Metarhizium: jack of all trades, master of many. **Open biology**, v. 10, n. 12, p. 200307, 9 Dec. 2020.

WALKER, A. S.; CLARDY, J. A Machine Learning Bioinformatics Method to Predict Biological Activity from Biosynthetic Gene Clusters. **Journal of Chemical Information and Modeling**, v. 61, n. 6, p. 2560–2571, 28 Jun. 2021.

ZHANG, W. et al. Introgression and gene family contraction drive the evolution of lifestyle and host shifts of hypocrealean fungi. **Mycology**, v. 9, n. 3, p. 176–188, 24 May 2018.

8. CURRICULUM VITAE

ROCHA, A. A.

1. DADOS PESSOAIS

Nome: Alexandra de Azevedo da Rocha

Cidade Natal: Porto Alegre/RS **Data de nascimento:** 25/07/1999

Endereço Profissional: Universidade Federal do Rio Grande do Sul, Centro de Biotecnologia Avenida Bento Gonçalves, 9500 – Campus do Vale Agronomia – Porto Alegre CEP 91500-970.

E-mail: aarocha1999@hotmail.com.

2. FORMAÇÃO

2022 - 2024: Pós-Graduação nível Stricto Sensu (mestrado) na área de Biologia Celular e Molecular na Universidade Federal do Rio Grande do Sul.

Título: CONSERVED GENE CLUSTERS IN ENTOMOPATHOGENIC FILAMENTOUS FUNGI.

Orientador: Prof. Dr. Charley Christian Staats.

2022: Pós-Graduação nível Lato Sensu (especialização) em Microbiologia Avançada no Centro Universitário Uniamérica.

2022: Pós-Graduação nível Lato Sensu (especialização) em Análises Clínicas e Toxicológicas no Centro Universitário Uniamérica.

2017 - 2021: Ensino superior completo em Ciências Biológicas pela Universidade La Salle (UniLaSalle).

Título: IMUNOGENICIDADE CITOTÓXICA ESTRUTURAL EM *CHLAMYDIA TRACHOMATIS.*

Orientador: Prof. Dr. Gustavo Fioravanti Vieira.

3. EXPERIÊNCIA PROFISSIONAL

2022: Bióloga Júnior. UNICONTROL CONTROLE DE PRAGAS LTDA.

Área: Departamento técnico/operacional.

Funções realizadas: Início da implementação de um laboratório de análises de água, desde as normas técnicas, conforme a NBR ABNT 17025:2017, a fim de conseguir a acreditação dos órgãos responsáveis - compras de material de laboratório, avaliação de técnicas, organização comercial, entre outros. Vínculo interrompido por conta do mestrado iniciado na UFRGS em março/22.

2018 - 2022: Bolsista de iniciação científica pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (Cnpq). UNIVERSIDADE LA SALLE.

Área: Pesquisa científica.

Funções realizadas: Pesquisa realizada na área de imunologia e bioinformática, a fim de obter as sequências de proteínas (epitopos) melhores ranqueadas da bactéria Chlamydia trachomatis, para posteriores testes de modelagem tridimensional de MHC.

2019 - 2021: Estagiária de Ciências Biológicas. UNICONTROL CONTROLE DE PRAGAS LTDA.

Área: Departamento técnico/operacional.

Funções realizadas: Auxiliar do departamento técnico e operacional, elaborando laudos, certificados e relatórios técnicos. Identificação e coleta de pragas conforme requerimento da empresa. Elaboração de uma coleção de pragas na empresa, além do início da implementação de um laboratório de análises de água, desde as normas técnicas, conforme a NBR ABNT 17025:2017, a fim de conseguir a acreditação dos órgãos responsáveis - processo ainda em andamento, não finalizado até o término do estágio.

2019: Estagiária de Ciências Biológicas. UNIVERSIDADE LA SALLE.

Área: Auxiliar de laboratório.

Funções realizadas: Auxiliar de laboratório no Herbário, ajudando na organização de documentos, fichas de controle, fichas de atualização, elaboração de exsicatas, manuseio e identificação de plantas, organização de pranchas com material da coleção presente no local, etc.

2019: Estagiária de Ciências Biológicas. UNIVERSIDADE LA SALLE.

Área: Auxiliar de laboratório.

Funções realizadas: Auxiliar de laboratório de microbiologia da Universidade. Preparo de meios de cultura, placas para colônias de bactérias, lavagem adequada de vidrarias, manuseio de autoclave e esterilização de material, preparo de aulas práticas no laboratório, etc.

2017 - 2018: Estagiária de Ciências Biológicas. ANTINSECT DESINSETIZADORA E PRODUTOS QUÍMICOS.

Área: Auxiliar do departamento técnico.

Funções realizadas: Auxiliar do departamento técnico e da qualidade, elaborando laudos, certificados e relatórios técnicos. Identificação, coleta e criação de pragas conforme

requerimento da empresa. Realização de trabalhos conforme sistema utilizado pela empresa, entre outros.

2016 - 2017: Bolsista júnior de iniciação científica pelo Cnpq. UNIVERSIDADE LUTERANA DO BRASIL.

Área: Pesquisa científica em laboratórios.

Funções realizadas: Manipulação de instrumentos em autoclave, preparação de meio de cultura complexo, cultivo de bactérias, contagem de eficácia por meio de Teste Cometa, entre outros.

4. RESUMOS PUBLICADOS EM ANAIS DE CONGRESSOS

ROCHA, A. A.; STAATS, C. C.; Conserved gene clusters in entomopathogenic filamentous fungi. In: 32º CONGRESSO DE MICROBIOLOGIA 2023 - Brazilian Congress of Microbiology, 2023, Foz do Iguaçu/PR.

ROCHA, A. A.; MEIRELES, M. R. ; VIEIRA, G. F. . Patterns of *Chlamydia trachomatis* epitopes for MHC-I allotypes HLA-A*02:01 and HLA-B*27:05: a cross-reactivity hypothesis. In: GENÉTICA 2021 - Brazilian Congress of Genetics, 2021, SAO PAULO.

ROCHA, A. A.; VIEIRA, G. F. . Imunogenicidade citotóxica estrutural em Chlamydia trachomatis. In: SEFIC 2020, Canoas. ANAIS SEFIC 2020.

ROCHA, A. A.; VIEIRA, G. F. . Imunogenicidade citotóxica estrutural em Chlamydia trachomatis. In: SEFIC 2019, Canoas. ANAIS SEFIC 2019.

DUTRA, C. A. B.; ROCHA, A. A. ; PELISSIOLI, T. ; WOLFARTH, M. . Lendas e Mitos: a influência etnocultural na conservação do Boto-cor-de-r. In: SEFIC 2019, Canoas. ANAIS SEFIC 2019.

ROCHA, A. A.; VIEIRA, G. F. Compreensão do mecanismo estrutural de alvos de células T em *Chlamydia trachomatis* que contribuem para proteção e os que favorecem a patogênese causada por este patógeno intracelular. In: SEFIC 2018, Canoas. ANAIS SEFIC 2018.

5. PRÊMIOS E TÍTULOS

2021 - Menção Honrosa, UNILASALLE - SEFIC.

2020 - Menção Honrosa, UNILASALLE - SEFIC.

6. FORMAÇÃO COMPLEMENTAR

2021: TREINAMENTO ABNT NBR ISO/IEC 17025:2017 (24 horas). IQM - Instituto de Qualidade e Metrologia, IQM, Brasil.

2021: TREINAMENTO RECICLAGEM NA NORMA ABNT NBR ISO/IEC 17025:2017 (8 horas). InnVitro - Suporte e Gestão em Toxicologia, Brasil.

2021: TREINAMENTO ENSAIOS MICROBIOLÓGICOS (3 horas). InnVitro - Suporte e Gestão em Toxicologia, Brasil.

2021: TREINAMENTO BIOSSEGURANÇA E TRABALHO ESTÉRIL PARA LABORATÓRIOS DE MICROBIOLOGIA (4 horas). InnVitro - Suporte e Gestão em Toxicologia, Brasil.

2021: TREINAMENTO GESTÃO DE RISCOS E OPORTUNIDADES (4 horas). InnVitro - Suporte e Gestão em Toxicologia, Brasil.

2021: TREINAMENTO VERIFICAÇÃO DE MÉTODOS EM MICROBIOLOGIA (4 horas). InnVitro - Suporte e Gestão em Toxicologia, Brasil.

2021: TREINAMENTO INTERPRETAÇÃO E UTILIZAÇÃO DE CERTIFICADOS DE CALIBRAÇÃO (4 horas). InnVitro - Suporte e Gestão em Toxicologia, Brasil.

2021: TREINAMENTO ENSAIOS EM FÍSICO-QUÍMICA - TEORIA (4 horas). InnVitro - Suporte e Gestão em Toxicologia, Brasil.

2021: TREINAMENTO VERIFICAÇÃO DE MÉTODOS EM FÍSICO-QUÍMICA (4 horas). InnVitro - Suporte e Gestão em Toxicologia, Brasil.

2018: CIÊNCIA EM FOCO - A INFLUÊNCIA DE DARWIN NO PENSAMENTO OCIDENTAL (3 horas). Universidade La Salle, Brasil.

2018: ACIDENTES TÓXICOS POR ANIMAIS PEÇONHENTOS (60 horas). Universidade Federal do Rio Grande do Sul, Brasil.

2018: 2° SEMANA ACADÊMICA DA BIOLOGIA: ÁREAS DE ATUAÇÃO DO PROFISSIONAL BIÓLOGO E O PAPEL DOS CONSELHOS FEDERAL E REGIONAL DE BIOLOGIA (1,5 horas). Universidade La Salle, Brasil.

2018: 2° SEMANA ACADÊMICA DA BIOLOGIA - VISITAÇÃO FLORAL DE VESPAS E ABELHAS EM PLANTAS: COMO VESPAS PODEM AUXILIAR NA MANUTENÇÃO DA DIVERSIDADE VEGETAL (1,5 horas). Universidade La Salle, Brasil.

7. INFORMAÇÕES ADICIONAIS

2017: Incluída voluntariamente em projetos de pesquisa da UniLasalle, a fim de se encaixar na área laboratorial da biologia e aprimorar os conhecimentos.

2019 - 2020: Monitora de Biologia Celular e Molecular.

2021: Monitora de Genética e Genética Humana.

2021: Monitora de Biologia Molecular e Bioquímica.

2021: Participação no Programa de Apadrinhamento na Universidade La Salle, atuando como madrinha do curso de Ciências Biológicas.