

UNIVERSIDADE DE CAXIAS DO SUL
ÁREA DO CONHECIMENTO DE CIÊNCIAS DA VIDA
INSTITUTO DE BIOTECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA

Análise genômica de *Penicillium echinulatum* para predição de agrupamentos de genes envolvidos no metabolismo secundário

NICOLE ANNE MODENA

Caxias do Sul, 2019.

NICOLE ANNE MODENA

*Análise genômica de *Penicillium echinulatum* para predição de agrupamentos de genes envolvidos no metabolismo secundário*

Dissertação apresentada ao Programa de Pós-graduação em Biotecnologia da Universidade de Caxias do Sul, visando à obtenção de grau de Mestre em Biotecnologia

Orientadora: Prof^ª Dra Scheila de Avila e Silva

Co-orientador: Prof Dr Aldo José Pinheiro Dillon

Caxias do Sul, 2019.

Dados Internacionais de Catalogação na Publicação (CIP)
Universidade de Caxias do Sul
Sistema de Bibliotecas UCS - Processamento Técnico

M689a Modena, Nicole Anne

Análise genômica de *Penicillium echinulatum* para predição de agrupamentos de genes envolvidos no metabolismo secundário / Nicole Anne Modena. – 2019.

68 f. : il. ; 30 cm

Dissertação (Mestrado) - Universidade de Caxias do Sul, Programa de Pós-Graduação em Biotecnologia, 2019.

Orientação: Scheila de Avila e Silva.

Coorientação: Aldo José Pinheiro Dillon.

1. *Penicillium*. 2. Genômica. 3. Fungos. 4. Genes. I. Silva, Scheila de Avila e, orient. II. Dillon, Aldo José Pinheiro, coorient. III. Título.

CDU 2. ed.: 582.282.123.2

Catalogação na fonte elaborada pela(o) bibliotecária(o)
Carolina Machado Quadros - CRB 10/2236

NICOLE ANNE MODENA

Análise genômica de *Penicillium echinulatum* para predição de agrupamentos de genes envolvidos no metabolismo secundário

Dissertação apresentada ao Programa de Pós-graduação em Biotecnologia da Universidade de Caxias do Sul, visando à obtenção do título de Mestra em Biotecnologia.

Orientadora: Prof. Dra. Scheila de Avila e Silva

Co-orientador: Prof. Dr. Aldo José Pinheiro Dillon

DISSERTAÇÃO APROVADA EM 11 DE NOVEMBRO DE 2019.

Orientadora: Prof^ª Dra. Scheila de Avila e Silva

Co-orientador: Prof. Dr. Aldo José Pinheiro Dillon

Prof^ª. Dra. Marli Camassola

Prof^ª. Dra. Caroline Dani

Prof. Dr. Rafael Andrade Caceres

AGRADECIMENTOS

Agradeço, em primeiro lugar, pela oportunidade de cursar este PPG. À minha família, pelo incentivo aos estudos desde muito cedo em minha vida. Ao meu querido Gustavo André Manfro, por todo apoio e incentivo, pela compreensão nos momentos de preocupação, por toda ajuda a concluir este trabalho.

Agradeço à Universidade de Caxias do Sul e ao PPGBio, pela estrutura e disponibilidade e à agência CAPES, pelo trabalho de incentivo à pesquisa no país e pelo suporte financeiro. Ao Laboratório de Bioinformática, por ter muito bem me recebido e apoiado nessa caminhada.

Aos colegas, por toda troca de conhecimento que amplia nossa visão de mundo e a todos os meus professores, que deixaram um pouco de si ao compartilhar seus saberes. À Banca de Acompanhamento, Prof. Dr. Daniel Luis Notari e Prof. Dra. Marli Camassola, por todas as sugestões feitas.

Ao meu Co-orientador, Prof. Dr. Aldo José Pinheiro Dillon, pela confiança e carinho depositados em mim.

Em especial, agradeço à minha Orientadora, Prof. Dra. Scheila de Avila e Silva, por ter me recebido em seu grupo de pesquisa, pela oportunidade de concluir este trabalho, pelo apoio incondicional, por todo o conhecimento transmitido. Obrigada por me mostrar os encantos da Bioinformática! Dedico todo este trabalho a você! Muito obrigada!

ÍNDICE

LISTA DE QUADROS.....	iii
LISTA DE FIGURAS.....	iv
NOMENCLATURA	v
RESUMO.....	vi
PALAVRAS-CHAVE:	vi
ABSTRACT	vii
KEYWORDS	vii
1 INTRODUÇÃO	8
2 REVISÃO BIBLIOGRÁFICA	11
2.1 O GÊNERO <i>Penicillium</i>	11
2.2 METABOLISMO SECUNDÁRIO	14
2.2.1 Policetídeos (PKs) e Sintases de Policetídeos (PKSs)	18
2.2.2 Peptídeos Não Ribossômicos (NRPs) e Sintetases de Peptídeos Não Ribossômicos (NRPSs)	22
2.2.3 Peptídeos Sintetizados por Ribossomos e Modificados Pós-Tradução (RiPPs)	28
2.2.4 Terpenos	29
2.2.5 Alcaloides.....	30
2.3 GRUPO DE GENES BIOSINTÉTICOS	32

2.4 <i>Penicillium</i> spp. E METABÓLITOS SECUNDÁRIOS DE INTERESSE	35
2.5 <i>Penicillium echinulatum</i>	37
3 MATERIAL E MÉTODOS	39
3.1 SEQUENCIAMENTO E ANOTAÇÃO GENÔMICA	40
3.2 PREDIÇÃO DE <i>BIOSYNTHETIC GENE CLUSTERS</i>	42
3.3 CURADORIA DE PREDIÇÕES	46
4 RESULTADOS E DISCUSSÃO	47
4.1 ARTIGO CIENTÍFICO A SER SUBMETIDO AO PERIÓDICO “ <i>FUNGAL BIOLOGY</i> ”	47
5 CONCLUSÕES E PERSPECTIVAS	62
6 REFERÊNCIAS BIBLIOGRÁFICAS	63

LISTA DE QUADROS

Quadro 1: Visão geral de classes de metabólitos secundários.....	17
Quadro 2: Domínios de PKSs (adaptado de Guzmán-Chavez <i>et al</i> , 2018).....	21
Quadro 3: Domínios em NRPS (adaptado de Guzmán-Chavez <i>et al</i> , 2018).....	27

LISTA DE FIGURAS

Figura 1: Padrões de ramificação de conidióforos observados em <i>Penicillium</i> (traduzido de Visagie <i>et al</i> , 2014).....	12
Figura 2: Metabólitos secundários encontrados em espécies de <i>Penicillium</i> (traduzido de Nielsen, 2018).....	13
Figura 3: Visão geral do metabolismo (traduzido de Nielsen, 2018).....	15
Figura 4: Biossíntese de NRPS (traduzido de Guzmán-Chavez <i>et al</i> , 2018).	25
Figura 5: Via biossintética geral para os RiPPs (Arnison <i>et al</i> , 2013, modificado).....	29
Figura 6: Visão geral de Agrupamentos de Genes Biossintéticos (BGCs) de metabólitos secundários (adaptado de Nielsen, 2018).	33
Figura 7: Fluxograma de trabalho para a análise do genoma do fungo <i>Penicillium echinulatum</i> 2HH e S1M29.....	39
Figura 8: Recorte de tela de <i>input</i> da ferramenta antiSMASH.....	43
Figura 9: <i>Pipeline</i> de análise genômica de metabólitos secundários realizada por antiSMASH (modificado de Blin <i>et al</i> , (2013).	44

NOMENCLATURA

A	Domínio de Adenilação
ACP	Proteína Carreadora de Acil
AT	Domínio de Acil Transferase
BGC	Grupo de Genes Biossintéticos (<i>Biosynthetic Gene Cluster</i>)
C	Domínio de Condensação
CAZymes	<i>Carbohydrate-Active EnZymes</i>
Cy	Domínio de Ciclização
DH	Domínio Desidratase
DMATS	Sintetase de Dimetilalil Triptofano
E	Domínio de Epimerização
ER	Domínio Enoil-Redutase
FAS	Sintetase de Ácidos Graxos
KR	Domínio Quetoreductase
KS	Domínio Quetossintase
MS	Metabólito Secundário
MT	Domínio Metil-transferase
NRP	Peptídeo Não Ribossômico (<i>Non-Ribosomal Peptide</i>)
NRPS	Sintetase de Peptídeo Não Ribossômico (<i>Non-Ribosomal Peptide Synthetase</i>)
Oxy	Domínio de Oxigenação
PCP	Domínio Proteína Carreadora de Pantotenilação/Peptidilo
PK	Policetídeo
PKS	Sintase de Policetídeo
PT	Preniltransferase
RiPP	Peptídeo Sintetizado por Ribossomos e Modificados Pós-Tradução (<i>Ribosomally Synthesized and Post-Translationally Modified Peptides</i>)
T	Domínio de Tiolação
TC	Terpeno Ciclase
Te	Domínio Tioesterase
TP	Terpenosintetásico
TS	Terpeno Sintase

RESUMO

Fungos filamentosos são conhecidos por sua importância na produção de metabólitos secundários de interesse: são empregados na indústria, na produção de enzimas e fármacos, na alimentação humana, entre outros compostos economicamente relevantes. Devido a este aproveitamento, é crescente o desenvolvimento de métodos de melhoramento para a produção de metabólitos. Um fungo melhorado por mutação ao acaso visando a produção de celulases foi *Penicillium echinulatum*, no qual a linhagem selvagem 2HH resultou, após várias etapas, na linhagem S1M29. A maioria dos compostos fúngicos de interesse biotecnológico são metabólitos secundários, classificados como policetídeos (PK), peptídeos não ribossômicos (NRP), terpenos, entre outros, de acordo com sua enzima estrutural. O objetivo deste trabalho foi prever metabólitos secundários nos genomas previamente sequenciados e anotados de *P. echinulatum* 2HH e S1M29. As sequências genômicas foram analisadas pela ferramenta antiSMASH v.5.0, que as comparou com o banco de dados MIBiG, e os dados obtidos foram verificados por curadoria manual. Os resultados foram semelhantes para as sequências das duas linhagens: cerca de 20% dos fragmentos montados foram identificados em 26 regiões com agrupamentos de genes relacionados a metabólitos secundários, contendo enzimas estruturais para terpenos, sintetases de peptídeos não ribossômicos (NRPS), fragmentos similares a sintetase de peptídeo não ribossômico (NRPS-like), sintases de policetídeo tipo I (PKS), agrupamentos híbridos entre NRPS e PKS tipo I, betalactona contendo inibidor de protease e sideróforo. Foram identificadas sete regiões semelhantes a agrupamentos de genes biossintéticos dos metabólitos ácido clavárico, naftopirona, nidulanina A, dois grupos similares a oxaleimida C, filostictina A e B e 4,4'-piperazina-2,5-diildimetil-bis-fenol. Estes grupos de genes indicam ortologia em outras espécies fúngicas cujos metabólitos secundários foram caracterizados, o que sugere que a mesma condição possa existir para as linhagens estudadas no presente trabalho.

PALAVRAS-CHAVE: Genômica comparativa; Genômica de fungos; Metabólitos secundários; Predição de agrupamentos de genes biossintéticos.

ABSTRACT

Filamentous fungi are known for their importance in the production of secondary metabolites of interest: they are used in industry, in production of enzymes and drugs, in food, among other economically relevant compounds. Due to this use, the development of improvement methods for the production of metabolites is increasing. A fungus enhanced by chance mutation for cellulase production is *Penicillium echinulatum*, in which the wild strain 2HH resulted, after several steps, in the S1M29 strain. Most fungal compounds of biotechnological interest are secondary metabolites, classified as polyketides (PK), non-ribosomal peptides (NRP), terpenes, among others, according to their structural enzyme. The objective of this study was to predict secondary metabolites in previously sequenced and annotated *P. echinulatum* 2HH and S1M29 genomes. The genomic sequences were analyzed by the antiSMASH v.5.0 tool, which compared them with the MIBiG database, and the obtained data were verified by manual curation. Results were similar for the sequences of the two strains: about 20% of assembled fragments were identified in 26 regions with secondary metabolite-related gene clusters containing structural enzymes for terpenes, non-ribosomal peptide synthetases (NRPS), fragments similar to non-ribosomal peptide synthetase (NRPS-like), polyketide type I synthases (PKS), hybrid groupings between NRPS and PKS type I, betalactone containing protease inhibitor and siderophore. Seven regions similar to biosynthetic gene clusters of the metabolites clavarinic acid, naphthopyrone, nidulanine A, two groups similar to oxaleimide C, phyllostictin A and B and 4,4'-piperazine-2,5-diyl dimethyl bis-phenol were identified. These gene groups indicate orthology in other fungal species whose secondary metabolites have been characterized, suggesting that the same condition may exist for the strains studied in the present work.

KEYWORDS: Comparative genomics; Fungal genomics; Secondary metabolites; Prediction of biosynthetic gene clusters.

1 INTRODUÇÃO

Biologia, a ciência que estuda a vida, utiliza tradicionalmente métodos de pesquisa baseados em observação. Desde os primórdios do uso da ciência, visto esta prática metodológica, foi possível desvendar como a vida acontece, quais fenômenos estão relacionados a ela, o que possibilita a existência de seres vivos, como eles se desenvolvem e se relacionam, entre tantos outros aspectos que se possa analisar. Entretanto, com os avanços das descobertas, novas metodologias são necessárias para a organização do conhecimento gerado, dividindo a Biologia em diversas ciências, e relacionando-a com conhecimentos de outras áreas, especialmente com Ciências Exatas.

Nesse contexto, a Bioinformática tem promovido grandes avanços, especialmente na área da Biologia Molecular. O uso de ferramentas computacionais para o armazenamento e análise de dados permitiu o desenvolvimento da Genômica, da Proteômica e da Metabolômica, possibilitando o estudo de sequências de nucleotídeos, aminoácidos, estruturas de metabólitos e trazendo novas descobertas nos mais variados domínios dos organismos.

Com o surgimento dos sequenciadores automáticos de DNA na segunda metade da década de 1990, houve uma explosão na quantidade de sequências estudadas, exigindo recursos computacionais cada vez mais eficientes, tanto para o armazenamento de dados como para a sua análise, tornando indispensável a utilização de plataformas computacionais eficientes para interpretação dos resultados obtidos. Atualmente, após o sequenciamento de um genoma, o foco das pesquisas tem sido a anotação genômica, processo que consiste em adicionar análises e interpretações biológicas sobre as sequências de DNA.

O Laboratório de Enzimas e Biomassas do Instituto de Biotecnologia da Universidade de Caxias do Sul realizou o sequenciamento das linhagens selvagem 2HH e

mutante S1M29 do fungo *Penicillium echinulatum*, considerado um microrganismo com grande potencial para a produção de celulasas. Seu complexo enzimático apresenta proporção equilibrada de atividade sobre papel filtro e β -glicosidase, o que é importante para a hidrólise de celulose. Nesse sentido, torna-se relevante o estudo também de seu metabolismo secundário, informação que pode auxiliar na compreensão da capacidade deste microrganismo de desconstruir os lignocelulósicos.

A produção de enzimas lignocelulósicas por meio deste fungo é estudada há mais de quarenta anos. Sua linhagem selvagem 2HH foi isolada do trato digestivo de larvas do coleóptero *Anobium punctatum* em 1979. A partir disso, o fungo foi submetido a um programa de mutagênese e seleção até a obtenção em 2006 da linhagem 9A02S1, mutante parcialmente desreprimido à glicose, a partir de peróxido de hidrogênio e seleção em meio suplementado com 2-desoxiglicose. Na continuidade do programa de melhoramento foi originado, em 2011, a linhagem S1M29, que possui alta produção de celulasas na presença de sacarose.

O fungo *P. echinulatum* 2HH é um organismo endossimbionte. Possivelmente, ancestrais foram ingeridos pelas larvas de *A. punctatum*, sobreviveram ao seu trato digestivo e passaram a viver de uma dieta de lignocelulósico, que é o alimento da larva, sendo a celulose quebrada pelo sistema enzimático do fungo.

O gênero *Penicillium* tem sido muito estudado por essa sua capacidade de produzir um grande número de produtos naturais com aplicações biotecnológicas e farmacêuticas. O metabólito secundário produzido por *Penicillium* spp. mais conhecido é o antibiótico penicilina, descoberto em 1929 por Alexander Fleming e que hoje é produzido em larga escala por linhagens melhoradas de *Penicillium rubens*. O antifúngico griseofulvina, o ácido micofenólico (um imunossupressor) e a compactina/mevastatina, medicamentos para reduzir o colesterol, estão entre compostos farmacêuticos relevantes produzidos por espécies de

Penicillium. Entretanto, algumas espécies também produzem micotoxinas como citrinina, ocratoxina e patulina, que podem oferecer riscos à saúde de seres humanos e animais. Isto comprova como essas espécies tem grande importância como fontes de produtos bioativos de aplicação médica.

Metabólitos secundários são compostos produzidos por um organismo que não estão diretamente ligados à sobrevivência do indivíduo, mas que trazem grandes benefícios à espécie ou ao organismo em longo prazo. O aproveitamento biotecnológico das espécies de *Penicillium* possui relação com sua capacidade de produzir metabólitos secundários bioativos.

Nesse sentido, é importante avaliar todo o contexto das vias biossintéticas do metabolismo do fungo para melhor compreender como aproveitar a produção desses compostos em processos biotecnológicos ou suprimi-los quando indesejados. O desenvolvimento de ferramentas de análise em bioinformática tem beneficiado a pesquisa de metabólitos secundários, pois possibilita que genomas sejam extraídos para a identificação de genes envolvidos nas vias biossintéticas de metabólitos secundários. As ferramentas desenvolvidas se aproveitam da característica destes genes de se agruparem no genoma em agrupamentos de genes biossintéticos (*Biosynthetic Gene Clusters* – BGC).

Considerando este contexto, o objetivo deste trabalho foi prever e caracterizar agrupamentos gênicos relacionados ao metabolismo secundário de *P. echinulatum*, de forma a compreender mecanismos biológicos do metabolismo secundário do fungo, comparando as linhagens selvagem e mutante utilizadas para um melhor aproveitamento de linhagens com potencial aproveitamento biotecnológico. São objetivos específicos: (a) identificar agrupamentos de genes relacionados ao metabolismo secundário; (b) comparar os agrupamentos preditos entre a linhagem mutante e selvagem com fungos filogeneticamente relacionados; e (c) caracterizar os agrupamentos em relação à enzima principal de suporte e ao metabólito secundário predito.

2 REVISÃO BIBLIOGRÁFICA

2.1 O GÊNERO *Penicillium*

Penicillium é um gênero bastante diverso de fungos filamentosos, com 354 espécies descritas, habitando os mais variados ambientes, com distribuição cosmopolita e grande impacto econômico à vida humana. Em 1809, Link nomeou o gênero com a palavra que significa “pincel”, e desde as primeiras três espécies – *Penicillium candidum*, *Penicillium glaucum* e *Penicillium expansum* – mais de 1000 nomes foram acrescentados, muitos atualmente revisitados por critérios de classificação mais modernos, ou considerados sinônimos de outras espécies, principalmente com o avanço dos sequenciamentos de genoma a partir da década de 1990 (Visagie *et al*, 2014).

A classificação taxonômica de *Penicillium* é: Reino Fungi, Filo Ascomycota, Classe Eurotiomycetes, Ordem Eurotiales, Família Trichomaceae, Gênero *Penicillium*. As colônias geralmente são de rápido crescimento, em tons de verde, às vezes brancas, consistindo principalmente de um feltro denso de conidióforos. Microscopicamente, as cadeias de conídios unicelulares são produzidas em sucessão basipetal de uma célula conidiogênica (fiálides). Em *Penicillium* spp., essas células podem produzir-se isoladamente, em grupos ramificados, dando uma aparência de escova (Figura 1) (Peberdy *et al*, 1987).

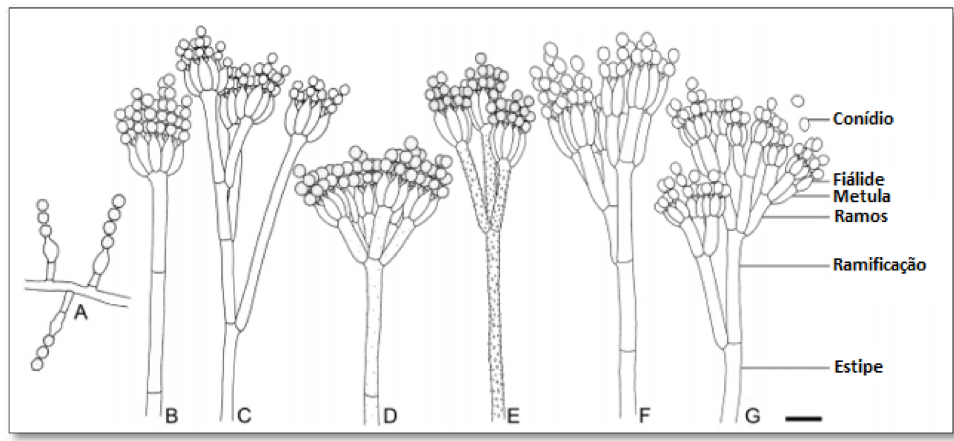


Figura 1: Padrões de ramificação de conidióforos observados em *Penicillium*. A. Conidióforos com fiálides solitárias. B. Monoverticilado. C. Divaricado. D, E. Biverticilado. F. Terverticilado. G. Quaterverticilato, termos utilizados para descrever partes de conidióforos. Barra de escala = 10 µm (traduzido de Visagie *et al*, 2014).

Em 1929, Fleming percebeu que colônias de *Penicillium notatum*, que acidentalmente contaminaram suas placas de Petri, mataram colônias de *Staphylococcus aureus*. Ele então isolou o fungo em meio líquido e separou o fluido das células, verificando que o composto inibia o crescimento de bactérias. Sua descoberta foi nomeada como “penicilina”, mas a pesquisa foi descontinuada. Por décadas, o fato permaneceu como uma curiosidade de laboratório, até que cientistas britânicos conseguiram isolar uma forma segura da penicilina, apesar de que a produção pelo microrganismo fosse bastante lenta. Em busca de melhoramento da produção, as condições de cultivo foram otimizadas e o organismo passou por seleção genética, iniciando com *Penicillium chrysogenum* (Fleming, 1929; Demain & Fang, 2000).

Na década de 1940, devido à II Guerra Mundial, a necessidade de penicilina tornou-se uma realidade, estimulando as pesquisas de seleção de linhagens e de programas de melhoramento, tanto genético como de microbiologia industrial, e consequentemente engajando os estudos sobre metabolismo secundário de fungos de interesse biotecnológico, como o *Penicillium*. Nesse sentido, a penicilina foi o primeiro metabólito secundário utilizado

como agente quimioterapêutico de sucesso produzido por um fungo, e abriu as portas para o desenvolvimento da indústria de antibióticos (Demain & Fang, 2000).

Essa procura por metabólitos secundários bioativos continuou e compostos que inibem o crescimento de bactérias, fungos, protozoários, parasitas, insetos, vírus e até mesmo células tumorais humanas foram descobertos, como exemplificado na Figura 2. Muitas outras moléculas com efeitos citotóxicos, mutagênicos, carcinogênicos, teratogênicos, imunossupressores, inibidores de enzimas, alelopáticas e outros efeitos biológicos também foram encontradas (Keller *et al*, 2005).

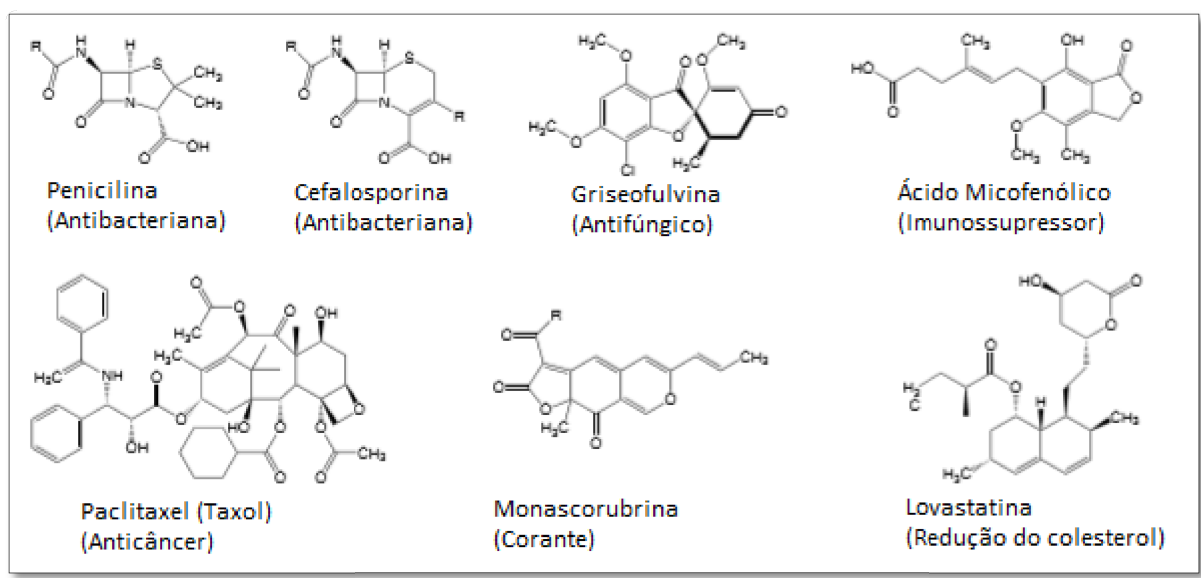


Figura 2: Metabólitos secundários encontrados em espécies de *Penicillium* (traduzido de Nielsen, 2018).

Fungos filamentosos, em geral, são conhecidos pela produção de inúmeros compostos de interesse comercial, especialmente de antibióticos e outros fármacos. Atualmente, o gênero *Penicillium* é amplamente utilizado na indústria, na fabricação de queijos, enzimas lignocelulolíticas e pequenas moléculas bioativas de aplicação farmacêutica (Nielsen, 2018). O estudo do metaboloma secundário de fungos oferece grande potencial nas aplicações industriais (Keller, 2019).

2.2 METABOLISMO SECUNDÁRIO

O desenvolvimento da fermentação da penicilina na década de 1940 marcou o início do que é chamado de era de ouro da microbiologia industrial, resultando em um grande número de metabólitos primários e secundários de importância comercial. O metabolismo primário envolve uma série de reações catabólicas, anfibólicas e anabólicas mediadas por enzimas que fornecem intermediários biossintéticos e energia, e convertem precursores biossintéticos em macromoléculas essenciais, como DNA, RNA, proteínas, lipídios e polissacarídeos. É equilibrado e os intermediários raramente são acumulados. Os principais metabólitos primários da bioindústria são aminoácidos, nucleotídeos, vitaminas e ácidos orgânicos (Demain & Fang, 2000).

Metabólitos primários são fundamentais para a sobrevivência das espécies. Por outro lado, metabólitos secundários são moléculas orgânicas que não estão diretamente envolvidas no crescimento e desenvolvimento de um organismo. Apesar disso, sua ausência não resulta em morte imediata, mas em longo prazo ocorre o comprometimento da capacidade de sobrevivência do organismo (Agostini-Costa *et al*, 2012). A integração entre o metabolismo primário e o secundário é exemplificada na Figura 3.

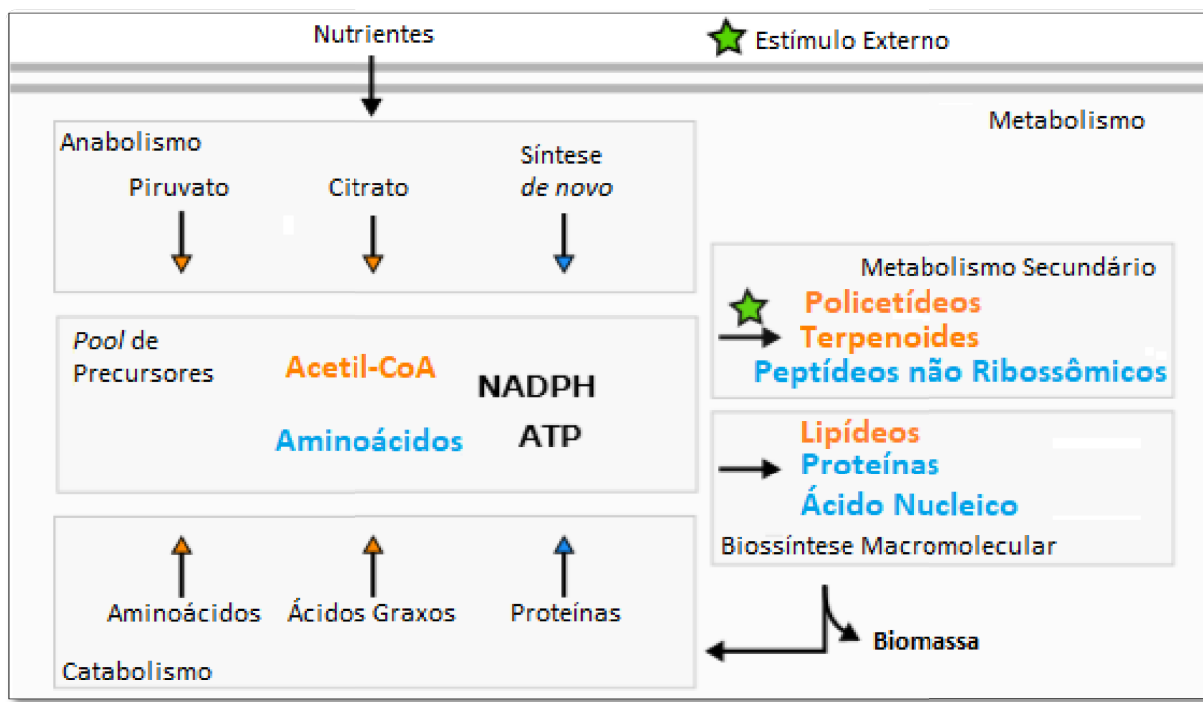


Figura 3: Visão geral do metabolismo. Nutrientes absorvidos são convertidos em precursores de metabólitos por anabolismo. Estes precursores são transformados em macromoléculas para gerar biomassa para o crescimento. Em condições de limitação de nutrientes, os precursores podem ser regenerados por catabolismo a macromoléculas e armazenadas na célula. Estímulos podem induzir o metabolismo secundário e direcionar o fluxo de precursores de metabólitos para a geração de metabólitos secundários. Acetil-CoA e aminoácidos são importantes blocos de construção para macromoléculas e metabólitos secundários (traduzido de Nielsen, 2018).

De todos os produtos tradicionais produzidos por bioprocessos, os mais importantes para a saúde humana são os metabólitos secundários. Estes são metabólitos que não têm função no crescimento, sendo frequentemente produzidos em uma fase de desenvolvimento de cultura subsequente a esta; possuem estruturas químicas incomuns e variadas; e são frequentemente misturas de membros relacionados de uma família química (Demain & Fang, 2000).

Estes compostos são um grupo de produtos naturais bastante diversificados, sintetizados por plantas, fungos, bactérias, algas e animais (Agostini-Costa *et al*, 2012). Eles fornecem vantagens para a sobrevivência do organismo de várias maneiras: melhoram a disponibilidade de nutrientes, como os sideróforos; protegem contra estressores ambientais, como pigmentos e osmoprotetores; estão envolvidos em interações competitivas com outros

organismos, como antibióticos e moléculas de sinalização; ou atuam como um mecanismo de defesa metabólica, como flavonoides vegetais e toxinas alcaloides (Breitling *et al*, 2013).

O metabolismo secundário pode ser um meio para armazenar ou eliminar metabólitos não aproveitados que resultam do metabolismo primário. Outra análise propõe que os genes envolvidos no metabolismo secundário sofreram evolução por mutação e seleção natural que resultou em benefícios para a espécie. Uma terceira observação caracteriza o metabolismo secundário como parte integrante do metabolismo celular, sendo dependente do metabolismo primário para fornecer enzimas, energia, substratos e maquinaria celular que contribuem para a sobrevivência do organismo (Roze *et al*, 2011).

Metabólitos secundários, também chamados de produtos naturais, são: (i) base de fármacos como antibióticos, agentes antitumorais e antivirais, antipiréticos, alucinógenos, e hipolipemiantes; (ii) moléculas importantes para aplicações agrícolas, como herbicidas ou fitotoxinas; (iii) aditivos alimentares para cor, sabor e edulcorantes; (iv) utilizados em fragrâncias; (v) utilizados como precursores para a síntese de plásticos e com relevância na pesquisa científica em biologia celular e molecular (Breitling *et al*, 2013; Blin *et al*, 2017). Ao longo do século XX, os produtos naturais contribuíram no avanço da compreensão humana sobre a biologia e no desenvolvimento da medicina, por exemplo novos compostos descobertos e sua ampla utilização, principalmente na indústria farmacêutica (Arnison *et al*, 2013).

Na natureza, os metabólitos secundários permitem a sobrevivência da linhagem, mas quando os microrganismos produtores são cultivados em cultura pura não há essa necessidade, e a capacidade de produção industrial de um metabólito de interesse é facilmente perdida por mutação. Considerando que o metabolismo primário é basicamente o mesmo para todos os sistemas vivos, o metabolismo secundário é geralmente específico da linhagem (Demain & Fang, 2000).

O metabolismo secundário em plantas, fungos e algas apresenta várias características comuns que promovem a integração com o metabolismo primário nos níveis moleculares e celulares. Primeiro, as vias metabólicas secundárias usam substratos (por exemplo, acil-CoA, aminoácidos, nucleotídeos e carboidratos), cofatores, metabólitos e energia que são formados durante o metabolismo primário. Segundo, supõe-se uma ligação evolucionária entre o metabolismo primário e o secundário, pois as vias metabólicas secundárias utilizam enzimas do metabolismo primário e estabelecem novas funções celulares para essas enzimas por mutação aleatória. Terceiro, os genes para vias metabólicas secundárias frequentemente são agrupados e esses grupos consistem de dois a muitos genes não homólogos. O agrupamento fornece ao organismo um mecanismo para coordenar a regulação temporal da expressão gênica de maneira ordenada (Roze *et al*, 2011).

Pesquisas identificaram várias classes de produtos naturais, sendo quatro grupos os mais relevantes: terpenoides, alcaloides, policetídeos e peptídeos não ribossômicos (Quadro 1). Sequenciamentos de genomas revelaram outra importante classe formada por peptídeos sintetizados por ribossomos e modificados pós-tradução (RiPP). Essas moléculas são produzidas em todos os três domínios da vida: seus genes são onipresentes nos genomas e a diversidade de estruturas é extensa (Arnison *et al*, 2013; Soltani, 2016).

Quadro 1: Visão geral de classes de metabólitos secundários¹.

Produto natural	Sigla	Origem	Enzima	Exemplos
Policetídeos	PK	Ácidos carboxílicos de cadeia curta	Policetídeo Sintase (PKS)	Tetraciclina Eritromicina
Terpenoides	-	Aminoácidos	Terpeno ciclase (TC)	Carotenoides Giberelinas
Alcaloides	-	-	Sem classe específica	Ergotamina
Peptídeos Não Ribossômicos	NRP	-	Sintetase de Peptídeo Não Ribossômico	Vancomicina Bacitracina
Peptídeos Sintetizados por Ribossomos e Modificados Pós-Tradução	RiPP	-	Sem classe específica	Ustiloxina Fomopsina

¹ Para terpenoides e alcaloides, não foi encontrada sigla correspondente. Para alcaloides, NRP e RiPP, não foi encontrada a origem dos metabólitos.

Os principais grupos de metabólitos secundários incluem policetídeos (PKs), peptídeos ribossômicos (RiPPs), peptídeos não ribossômicos (NRPs) e terpenoides. Estes polímeros são construídos pela ação de sintases ou sintetases, com exceção dos peptídeos ribossômicos, e são modificados por (des)hidratases, oxigenases, hidrolases, metilases e outros. Em fungos, como já mencionado, os genes biossintéticos do metabolismo secundário são organizados em *clusters* em torno dos genes da sintase (Andersen *et al*, 2012).

2.2.1 Policetídeos (PKs) e Sintases de Policetídeos (PKSs)

Policetídeos (PKs) são compostos por unidades $\text{CH}_2(\text{C}=\text{O})$, e representam um grande grupo de produtos naturais encontrados em bactérias, fungos e plantas. Sua biossíntese acontece a partir de precursores acil CoA através de enzimas denominadas sintases de policetídeos (PKSs) (Shen, 2003; Andersen *et al*, 2012).

Uma PKS é organizada por pelo menos por três domínios: ceto-sintase (KS) como módulo de partida, aciltransferases (AT) como módulos de alongamento e proteína carreadora de acil (ACP) como módulo de terminação. Após cada alongamento da cadeia, os domínios adicionais, como ceto-redutase (KR), desidratase (DH), enoil-redutase (ER), metiltransferase (MT) e tiosterase (Te), podem alterar a cadeia de policetídeos e produzir uma estrutura saturada de acila (Soltani, 2016).

São conhecidos três tipos de PKS: (a) PKSs do tipo I, que são enzimas multifuncionais arranjadas de forma linear e organizadas em módulos de atividades de ação não iterativa responsáveis pela catálise de um ciclo de alongamento da cadeia de policetídeos; (b) as PKSs do tipo II, que são complexos catalíticos independentes multienzimáticos que transportam um único conjunto de atividades que age de maneira iterativa; e (c) PKSs do tipo III, que são enzimas homodiméricas multifuncionais de condensação de ação iterativa, que

utilizam o malonil-CoA como substrato. Em alguns casos, observa-se que módulos mistos de PKS tipo I e de sintetases de peptídeos não ribossômicos (NRPSs) geram enzimas híbridas (PKS-NRPS), produzindo novas pequenas moléculas de metabólitos secundários (Shen, 2003; Andersen *et al*, 2012; Soltani, 2016).

Os PKs são os metabólitos secundários mais abundantes em fungos. Estes compostos são sintetizados por sintases de policetídeos (PKSs) do tipo I, proteínas de múltiplos domínios que estão relacionadas com as sintases de ácidos graxos de eucariontes e contém estruturas de domínios semelhantes. Para ambas as enzimas, os ácidos carboxílicos de cadeia curta – geralmente acetil-coenzima A (acetil-CoA) e malonil-CoA – são condensados para formar cadeias de carbono de vários comprimentos. A principal diferença entre os policetídeos e os ácidos graxos é a redução total do β -carbono nos ácidos graxos, o que é um evento opcional na síntese de policetídeos. Nas PKSs fúngicas, os domínios da cetoacil-CoA sintase (KS), aciltransferase (AT) e transportador de acila (ACP) são essenciais para a síntese de policetídeos, enquanto os domínios da reductase (KR), desidratase (DH) e enoil redutase (ER) não estão presentes em todas as enzimas PKS fúngicas (Keller *et al*, 2005).

PKSs do tipo I podem ser classificadas em duas categorias: modulares e iterativas, sendo que a primeira é exclusiva de bactérias. PKSs tipo I iterativas utilizam módulos de domínios catalíticos específicos, em que a reação é repetida para produzir o esqueleto completo do policetídeo. As PKSs de fungos possuem um único módulo e os domínios consecutivos atuam em sequência durante a síntese do policetídeo. Eles são equipados com domínios estruturais básicos encontrados em enzimas PKS (domínios ACP, KS e AT), mas também podem conter os domínios KR, DH, ER e Te, sendo que essas enzimas podem ser divididas em PKS altamente redutora, não redutora e parcialmente redutora, dependendo da presença ou ausência de domínios redutores (Guzmán-Chaves *et al*, 2018).

As PKSs do tipo II são exclusivas de bactérias e utilizam mecanismos iterativos semelhantes ao de PKSs tipo I iterativas. As PKSs do tipo III foram descobertas primeiramente em plantas, mas também podem ser encontradas em fungos e bactérias. Elas são formadas por um único domínio KS que catalisa um número de alongamentos, gerando pequenos anéis de fenol ou naftol. A enzima transfere o grupo acil do CoA para o sítio ativo histidina, extremamente conservado. Entretanto, a sequência de aminoácidos não é similar à encontrada em domínios KS de enzimas tipo I e II. Independentemente das diferenças mecânicas ou estruturais, todos os policetídeos sintetizados pelas enzimas PKS seguem o mesmo mecanismo de condensação descarboxilativo dos precursores acil-CoA. No entanto, esses precursores devem ser ativados antes pelo domínio ACP, no caso das enzimas PKS tipo I e II, enquanto as enzimas PKS tipo III atuam independentemente dos domínios ACP (Guzmán-Chávez *et al*, 2018). No Quadro 2 estão descritos os domínios encontrados em enzimas PKSs.

Quadro 2: Domínios de PKSs (adaptado de Guzmán-Chavez *et al*, 2018). A coluna “Código Interpro” indica o código do domínio de proteínas em busca no banco de dados InterPro².

SIGLA	DOMÍNIO	DESCRIÇÃO	CÓDIGO InterPRO
AT	Domínio Aciltransferase	Seleciona a unidade de início (malonil-CoA ou metilmalonil-CoA) antes de sua transferência ao domínio ACP do ciclo de alongamento da cadeia.	IPR020801
ACP	Proteína Carreadora de Acil	Modula três eventos importantes durante a síntese de PK: condensação durante o alongamento da cadeia ao transferir a unidade de início do domínio AT ao domínio KS, transporte da cadeia em formação entre os domínios dos dois extremos, bem como aos domínios opcionais, envolvendo o reconhecimento proteína-proteína entre domínios e prevenção a ciclização prematura e enolização da cadeia de PK.	IPR036736
KS	Domínio Quetossintase	É um domínio de condensação homodimérico que catalisa a extensão do intermediário β -cetoacilo por uma condensação de Claisen descarboxilada. Nessa fase, uma ligação tioéster é formada entre o grupo tiol da cisteína ativa da KS e a PK na fase de alongamento. Portanto, mecanicamente, o domínio KS atua em três estágios: acilação, descarboxilação e condensação.	-
KR	Domínio Quetoreductase	É uma unidade de processamento de carbono β que pertence família de desidrogenase de cadeia curta/redutases. Este domínio reduz o grupo β -ceto, que é formado durante o processo de condensação, em um grupo hidroxila usando NADPH. Além disso, alguns domínios KR estão equipados com atividade de epimerase.	PF08659
DH	Domínio Desidratase	É geralmente acoplado aos domínios KR do tipo B. Este domínio catalisa a eliminação da água na posição da cadeia β -hidroxi acil, produzindo assim ligações duplas trans.	-
ER	Domínio Enoil Redutase	É uma unidade de adaptação opcional envolvida no estado final de oxidação do policetídeo em crescimento. Reduz os grupos α , β -enoíla e, assim, gera ligações α - β saturadas. Esta reação envolve NAD(P)H como doador de hidreto em um mecanismo de tipo de adição de Michael.	-
Te	Domínio Tioesterase	O término da biossíntese do policetídeo envolve o domínio Te, que produz macrolactonas via ciclização intramolecular ou PKs lineares por hidrólise. Em ambos os eventos, um intermediário acil-Te é formado através da transferência da cadeia PK do último ACP para a serina ativa no domínio Te.	IPR001031

² Para os domínios KS, DH e ER, não foi encontrado registro no banco de dados InterPro.

2.2.2 Peptídeos Não Ribossômicos (NRPs) e Sintetases de Peptídeos Não Ribossômicos (NRPSs)

Peptídeos não ribossômicos (NRPs) são metabólitos secundários formados por peptídeos, comumente encontrados em bactérias, fungos ou organismos endossimbiontes de eucariontes superiores. Estes compostos são sintetizados por sintetases de peptídeos não ribossômicos (NRPSs), uma família de enzimas que adicionam substratos acila ao metabólito secundário bioativo (Soltani, 2016; Bloudoff & Schmeing, 2017).

As NRPSs podem utilizar mais de 500 diferentes substratos de monômeros de acila, predominantemente aminoácidos, mas também ácidos graxos, ácidos carboxílicos e outros; podem ser cossinteticamente ou pós-sinteticamente modificados e têm topologias lineares, cíclicas ou ramificadas. Toda essa variedade em estrutura química reflete a diversidade de atividades importantes destes compostos, que podem ser sinalização, comunicação, quelação de íons metálicos, proteção do hospedeiro, e ainda muitos compostos ainda não totalmente caracterizados. Entre produtos farmacêuticos em que os NRPs podem ser empregados estão compostos antifúngicos, antibacterianos, antivirais, antitumorais, sideróforos, imunossupressores e drogas citostáticas (Bloudoff & Schmeing, 2017; Guzmán-Chavez *et al*, 2018).

As sintetases peptídicas não ribossômicas são máquinas enzimáticas altamente estruturadas e complexas, intimamente relacionadas com outras enzimas modulares, tais como PKSs, sintetases híbridas NRPS-PKS e sintetases de ácidos graxos (FASs). Estas sintetases têm certas propriedades em comum, como sua divisão estrutural em domínios e módulos, que se manifesta em sua história evolutiva compartilhada. Cada enzima consiste em um módulo, que é uma unidade funcionalmente distinta que permite o recrutamento e incorporação de um precursor em um produto em crescimento. Domínios, assim como módulos, são claramente definidos e possuem estruturas evolutivas intercambiáveis entre enzimas multimodulares. No

caso de PKS e NRPS, isso levou à ocorrência de uma variedade de híbridos NRPS-PKS (Guzmán-Chavez *et al*, 2018).

Os NRPs são derivados de aminoácidos proteinogênicos e aminoácidos não proteinogênicos polimerizados por enzimas sintetases peptídicas não ribossomais, que são multidomínios e multimodulares. Cada módulo em um NRPS contém vários domínios que permitem o reconhecimento, a ativação e a ligação covalente de um aminoácido específico a um módulo. Subsequentemente, ligações peptídicas são formadas entre os aminoácidos. O peptídeo resultante é então liberado por um domínio semelhante à tioesterase que está normalmente localizado na extremidade do C-terminal do módulo final. Os domínios que realizam estas atividades são denominados A (adenilação), PCP (proteína carreadora de pantotenilação/peptidilo), C (condensação/formação de ligação peptídica) e Te (tioesterase) (Keller *et al*, 2005).

Estruturalmente, cada módulo de NRPS (o módulo de iniciação, n módulos de alongamento e o módulo de terminação) requer um conjunto mínimo de domínios. Os dois domínios essenciais para cada módulo são o domínio de adenilação (A) e o domínio não catalítico de tiolação (T). Esse domínio duplo em *tandem* permite a seleção e ativação específicas de um determinado substrato. No entanto, o domínio T deve primeiro passar pela ativação após expressão da 4'-fosfopanteteinil transferase (PPTase) e da coenzima A (CoA) dependente, transferindo a porção fosfopanteteína da CoA a um resíduo de serina conservado, para entrar no estado ativado. Além disso, os domínios de adenilação (A) possuem fatores de acompanhamento, ou proteínas, denominados proteínas semelhantes a MbtH (MLPs). Além desses domínios, qualquer módulo de alongamento exigirá um domínio de condensação (C), que conecta dois módulos e liga os substratos ativados através de uma ligação peptídica. Os domínios C são estereoespecíficos para ambos os substratos ativados e tornam o composto intermediário resultante ligado ao domínio T. Por fim, o módulo de terminação C-terminal

requer essencialmente um domínio de tioesterase (Te), para liberar cataliticamente o composto covalentemente ligado do NRPS, retornando o complexo NRPS ao estado fundamental para outro ciclo de reação. Além desses domínios essenciais, podemos distinguir uma série de domínios adicionais, realizando epimerização, halogenação, ciclização, macrociclização, multimerização ou metilação. Esses mecanismos enzimáticos podem ser classificados como NRPS tipo I quando os módulos estão dispostos em uma única proteína, e de NRPS tipo II quando as proteínas são independentes e atuam de forma transitória durante a síntese de NRP (Guzmán-Chávez *et al*, 2018).

Uma NRPS pode ser formada por uma única unidade modular de três domínios, ou como a maior e mais complexa estrutura conhecida, contendo 15 módulos com 46 domínios, gerando um complexo proteico de 1,8 MDa (NRPS tipo I). Embora a configuração do NRP seja limitada pelo tamanho e sequência modular da NRPS, esta pode interagir com enzimas de corte, gerando produtos de maior complexidade. Pequenos trechos de até 30 aminoácidos no C- ou N- terminal, que formam um ponto de reconhecimento bastante específico, permitindo a comunicação entre múltiplos NRPS de um *cluster* (NRPS tipo II). Foram descritos três tipos de sistemas NRPS, de acordo com o seu modo de síntese (ou estratégia de biossíntese): Tipo A (linear), tipo B (iterativo) e tipo C (não linear), como está ilustrado na Figura 4 (Guzmán-Chávez *et al*, 2018).

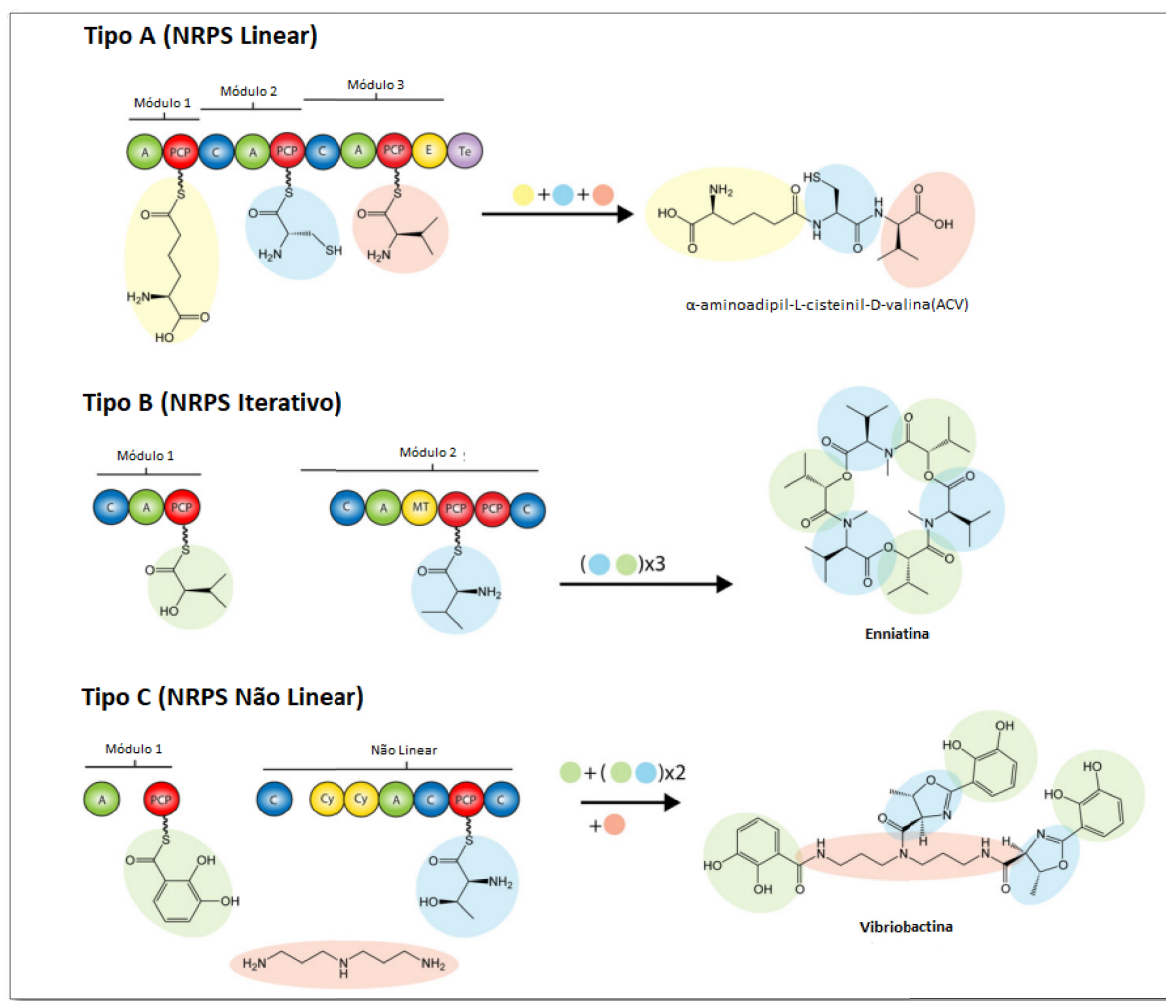


Figura 4: Biossíntese de NRPS (traduzido de Guzmán-Chavez et al, 2018).

O sistema tipo A abriga a típica organização de domínio A-T(C-A-T)_n-1-Te/C, onde n representa o número de aminoácidos no peptídeo. Neste NRPS linear, a ordem e o número de módulos correlacionam-se com a sequência de aminoácidos no NRP e, portanto, é possível prever o produto que será formado. Geralmente, nas NRPSs de fungos, a reação de ciclização é realizada por um domínio C especializado em vez do domínio Te. Uma vez que cada módulo catalisa um ciclo durante o alongamento da cadeia do NRP devido à sua atividade específica, este sistema é considerado análogo ao de PKS tipo I. Nos fungos, os exemplos mais proeminentes deste tipo de NRPS são ACV sintetases (β -lactâmicos), sintetases de ciclosporina (ciclosporina A) e sintetases de peptaibol (peptaibols são uma classe de antibióticos com alto conteúdo de ácido α -aminoisobutírico). O sistema tipo B é

caracterizado por empregar todos os seus módulos ou domínios mais de uma vez durante a síntese de um único NRP, o que permite a montagem de cadeias peptídicas que contêm sequências repetidas ao longo da estrutura. Exemplos de NRPSs do tipo B são as sintetases de sideróforos, que contêm apenas três domínios A que catalisam a biossíntese do ferrocromo. No sistema tipo C, os NRPS não lineares têm pelo menos uma conformação de domínio que se desvia da organização (C-A-T)_{n-1} contida em NRPSs lineares. Da mesma forma, nestas sintetases a disposição do módulo não corresponde à sequência de aminoácidos no NRP. Ao contrário da NRPS tipo A, nas NRPSs tipo C o peptídeo não linear é produzido por uma sintase de ponto de ramificação e contém padrões de ciclização incomuns. Outra diferença importante é que os NRPSs não lineares podem incorporar pequenas estruturas solúveis, como aminas, no NRP crescente através de domínios C especializados. A capreomicina, bleomicina e vibriobactina são exemplos de NRPs produzidos por este tipo de sintetases (Guzmán-Chávez *et al*, 2018). O Quadro 3 apresenta os domínios encontrados em NRPSs.

Quadro 3: Domínios em NRPS (adaptado de Guzmán-Chavez *et al*, 2018).

SIGLA	DOMÍNIO	DESCRIÇÃO
AT	Domínio de Adenilação e Tiolação	Qualquer módulo NRPS consiste minimamente de um domínio A e T (ou proteína carreadora de peptídilo, PCP), permitindo funcionalidade de módulo único e funcionalidade multi-modular após a adição de domínios C. Não existe formação de produtos sem adenilação prévia e tioesterificação de um substrato. As funções do domínio A são caracterizadas primeiro, através da hidrólise de ATP ou adenilação, permitindo a formação de um conjugado de substrato AMP, que é subsequentemente transferido para o grupo tiol livre da porção 40-fosfopantetaininilo (Ppant), que é ancorado a um resíduo de serina conservado no domínio T.
C	Domínio de Condensação	Os domínios C contêm aproximadamente 450 resíduos, uma classe altamente versátil de domínios NRPS. Qualquer NRPS composto por mais de um módulo deve conter, pelo menos, um domínio C. Essencialmente, o alvo primário de um domínio C é a condensação dos substratos ativados através de um ataque nucleofílico, levando principalmente à formação de um n-peptídeo ligado através de uma ligação peptídica. No entanto, vários resíduos do domínio C podem ter o potencial intrínseco para cumprir múltiplas funções.
E	Domínio de Epimerização	Os domínios E estão entre os domínios de modificação mais abundantes intrínsecos ao NRPS. Em contraste com os domínios C estruturalmente semelhantes, eles são responsáveis pela epimerização específica do local de um substrato, desempenhando predominantemente essa função após a formação da ligação peptídica ter ocorrido.
Te	Domínio Tioesterase	A presença do domínio da tioesterase (Te) é comum em NRPS simples e multimodulares. A maioria dos NRPS contém um domínio em seu C-terminal responsável por remover um produto ou suporte de produto da enzima primária. Além disso, esse domínio abriga uma atividade de controle de qualidade (revisão) para verificar a configuração correta do peptídeo criado.
Outros domínios		Além do domínio de epimerização relacionado ao domínio C, existem os domínios de ciclização (Cy), oxigenação (Oxy) e metiltransferase (MT). Os domínios Cy e Oxy substituem especificamente a função clássica dos domínios C, omitindo a condensação de aminoácidos através da formação de ligações peptídicas, resultando em estruturas de tiazolina, oxazolina ou metiloxazolina. Essas reações ocorrem predominantemente em NRPS produtoras de sideróforos e contam com a presença de resíduos de serina, treonina e cisteína. Também os domínios MT seguem a padronização estrutural comum de subdomínio, que também é vista nos domínios A-, C-, E- e Cy-. Os domínios MT são mais restritos em sua função, que abrange a transferência de grupos metil de S-adenosilmetionina para N (N-MT), C (C-MT), O (O-MT) ou para determinados resíduos S (S-MT).

2.2.3 Peptídeos Sintetizados por Ribossomos e Modificados Pós-Tradução (RiPPs)

Peptídeos sintetizados por ribossomos e modificados pós-tradução (RiPPs) são metabólitos secundários mais comuns em bactérias, apesar de terem sido identificados também em fungos basidiomicetos e ascomicetos (Soltani, 2016). Essas moléculas têm estruturas diversas e são onipresentes nos genomas e transcriptomas sequenciados (Arnison *et al*, 2013).

As variadas modificações pós-traducionais ou co-traducionais conferem estruturas não diretamente acessíveis aos peptídeos ribossomais naturais, restringindo sua flexibilidade conformacional para permitir um melhor reconhecimento do alvo, aumentar a estabilidade metabólica e química, e melhorar sua funcionalidade química (Arnison *et al*, 2013).

Quase todos os compostos produzidos por Síntese Peptídica Pós-Ribossômica são inicialmente sintetizados como um peptídeo precursor mais longo, entre 20-110 resíduos de aminoácidos, codificados por um gene estrutural. Pode ser feita uma distinção em relação ao peptídeo central não modificado, no peptídeo precursor e no peptídeo central modificado após as modificações pós-tradução. Na maioria das RiPPs, um peptídeo líder ou uma sequência líder é anexada à porção N-terminal do peptídeo central que é geralmente importante para o reconhecimento de muitas das enzimas de modificação pós-traducional e para exportação (Arnison *et al*, 2013). Na Figura 5 está exemplificado o funcionamento da via para a biossíntese de RiPP.

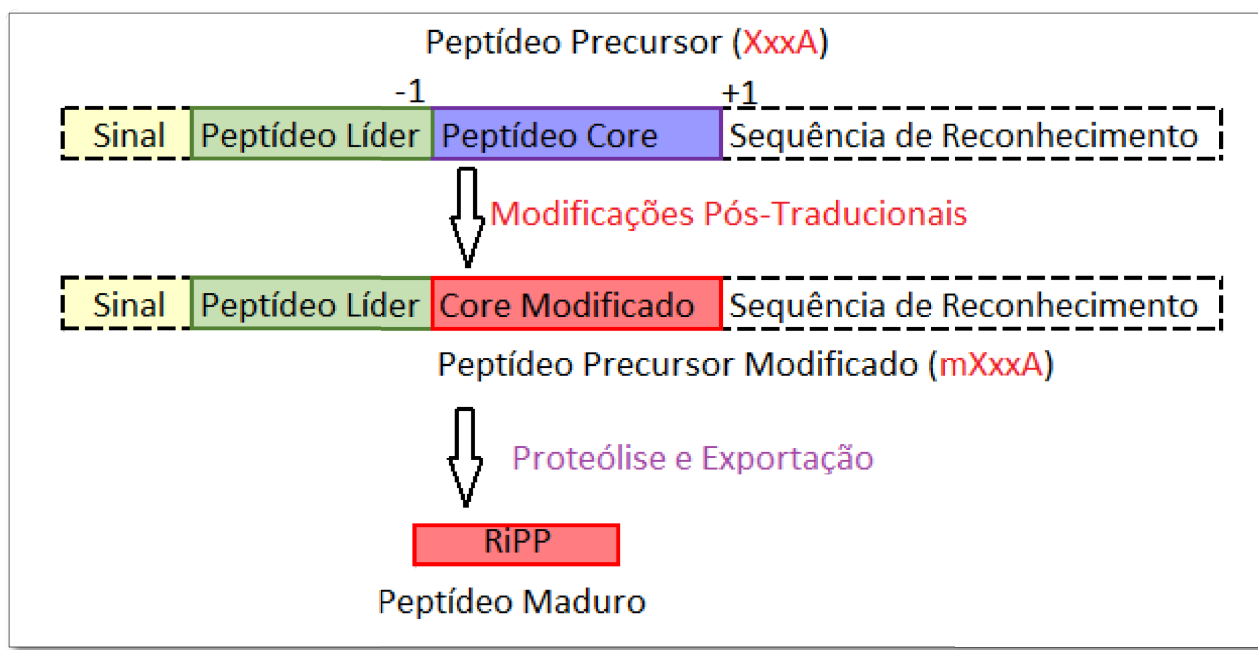


Figura 5: Via biossintética geral para os RiPPs. O peptídeo precursor contém uma região central que é transformada no produto maduro. Muitas das modificações pós-traducionais são guiadas pelo peptídeo líder e pelas sequências de reconhecimento para subclasses de RiPPs. Produtos de origem eucariótica também contêm frequentemente um peptídeo sinal N-terminal que dirige o peptídeo para compartimentos especializados para modificação e secreção. As sequências de reconhecimento do terminal C estão por vezes também presentes para a ciclização de peptídeos (Arnison *et al*, 2013, modificado).

2.2.4 Terpenos

Terpenoides são produtos naturais sintetizados por fungos filamentosos; é o maior e mais diverso grupo de metabólitos secundários derivados de ramificações no carbono C5 da cadeia de isopreno: isopentenil pirofosfato (IPP) e dimetilalil pirofosfato (DMAPP), que são sintetizados na via mevalonato de acetil-CoA. Variam de cadeias lineares a policíclicas e de hemiterpenos de cinco carbonos a milhares de unidades de isopreno (Agostini-Costa *et al*, 2012; Soltani, 2016; Guzmán-Chávez *et al*, 2018).

Os terpenos mais conhecidos são os metabólitos de plantas odoríferas, como a cânfora e a terebintina, mas os fungos também sintetizam vários terpenos importantes, incluindo os aristoloquenos, carotenoides, giberelinas, diterpenos indólicos e tricotecenos (Keller *et al*, 2005).

São exemplos de terpenoides alguns sabores e moléculas aromáticas, como mentol, linalol, geraniol e cariofileno, formados por monoterpenos (C10, que são gerados a partir do geranyl pirofosfato), com duas unidades de isopreno, e sesquiterpenos (C15, que são gerados a partir de farnesil pirofosfato), com três unidades de isopreno, além de outros compostos bioativos com propriedades especiais, como diterpenos (C20, que são gerados a partir do geranyl-geranyl-pirofosfato), triterpenos (C30) e tetraterpenos ou carotenoides (C40) (Keller *et al*, 2005; Agostini-Costa *et al*, 2012; Soltani, 2016; Guzmán-Chávez *et al*, 2018).

Eventualmente, estas cadeias lineares de diferentes comprimentos são posteriormente modificadas por ciclases, terpenosintetásicos (TPs) e preniltransferases (PTs), produzindo diferentes subclasses de terpenoides (Guzman-Chávez *et al*, 2018). A enzima que atua na síntese de terpenos é a terpeno ciclase, que é essencial para a produção de diferentes terpenos a partir de diferentes difosfatos. Embora as terpeno ciclases tenham homologia estrutural, elas têm pouca similaridade na sequência primária e parecem ter divergido de forma relativamente rápida de um ancestral comum (Keller *et al*, 2005).

2.2.5 Alcaloides

Os alcaloides são um grupo de produtos naturais altamente diversificados que contém um ou mais átomos de nitrogênio básicos em um anel heterocíclico (Soltani, 2016). Alcaloides indólicos são geralmente derivados do triptofano e do dimetilalil pirofosfato, embora às vezes outros aminoácidos além do triptofano sejam usados como precursores (Keller *et al*, 2005). Devido à sua grande diversidade estrutural, não há classificação uniforme de alcaloides e nenhuma classe específica de enzimas para sua biossíntese. No entanto, de acordo com as suas origens biogenéticas, os alcaloides podem ser agrupados em quatro classes: alcaloides derivados de aminoácidos, alcaloides purínicos, terpenos aminados e alcaloides policetidos (Soltani, 2016).

A via metabólica de alcaloides mais bem compreendida é a síntese de ergotamina em *Claviceps purpurea* e espécies relacionadas: primeiro acontece a prenilação do triptofano pela dimetilalil triptofano sintetase (DMATS). Após a metilação do dimetilalil triptofano, inicia uma série de etapas de oxidação através da agroclavina para o ácido lisérgico. O ácido lisérgico é então ativado por uma NRPS de módulo único, condensado com um tripeptídeo que é produzido por uma segunda NRPS e liberado como ergotamina (Keller *et al*, 2005).

Embora muitos alcaloides bem conhecidos (como hipocianamina, fisostigmina, cocaína, morfina, codeína, vimblastina) sejam produzidos por plantas, os alcaloides também podem ser produzidos por vários outros organismos, como os fungos. Estes compostos são considerados parte da defesa química de um organismo, mas apresentam também aplicações na medicina humana como analgésicos, antiarrítmicos, anti-hipertensivos, antiprotozoários, antitumorais, estimulantes e vasodilatadores e na agricultura como inseticidas (Soltani, 2016).

Devido ao fato de que os genomas de muitos microrganismos codificam a produção de metabólitos secundários a partir de múltiplos conjuntos de genes (*biosynthetic gene clusters*), a mineração de genoma tem sido estabelecida como importante ferramenta que complementa bioensaios no estudo de produtos naturais, o que foi amparado pelo desenvolvimento e disponibilidade de várias ferramentas de mineração de dados (Blin *et al*, 2017).

O rápido desenvolvimento da genômica nos últimos anos possibilitou entender que os microrganismos codificam o potencial para produzir muito mais metabólitos secundários do que o inicialmente esperado. Muitas destas moléculas são preditas por análise bioinformática do suposto agrupamento genético do metabólito secundário em genomas sequenciados, mas não são produzidas naturalmente em condições de laboratório ou estão presentes em níveis que são demasiado baixos para serem detectados por métodos padrão. Em

alguns casos, a produção desses metabólitos pode ser induzida por manipulações genéticas (Breitling *et al*, 2013).

2.3 GRUPO DE GENES BIOSSINTÉTICOS

Ao contrário dos genes necessários para a síntese de um metabólito primário, que são dispersos pelo genoma do fungo, os genes que codificam as atividades enzimáticas para produzir um metabólito secundário são organizados de forma agrupada, formando um grupo de genes biossintéticos (Keller, 2019). Um grupo de genes biossintéticos (BGC) é um agrupamento de dois ou mais genes em um cromossomo de um dado genoma que juntos codificam uma via biossintética para a produção de um metabólito secundário e suas variantes químicas (Medema *et al*, 2015; Weber *et al*, 2015). Estes genes são corregulados de acordo com a função ecológica do metabólito secundário codificado (Keller, 2019).

O agrupamento de genes de vias metabólicas não é comum em eucariotos, possivelmente por causa do risco de perder toda uma rota essencial caso parte de um cromossomo seja perdida. No entanto, nas bactérias, os genes das vias são frequentemente agrupados em *operons*, propriedade em que os genes são mais propensos a sobreviver à transferência horizontal se todos os compostos da via forem adquiridos de uma só vez. Dessa forma, entende-se que BGCs de fungos tenham origem de bactérias. Como os metabólitos secundários não são essenciais, pode não haver uma forte pressão seletiva para distribuir os genes da via no genoma. O agrupamento de genes também pode servir como uma característica vantajosa para novas vias metabólicas secundárias em rápida evolução. Mutações podem gerar compostos novos e úteis, sem o risco de prejudicar o hospedeiro por

afetar genes essenciais, o que se confirma pelo fato de geralmente os BGC localizarem-se em regiões teloméricas dos cromossomos (Nielsen, 2018).

A síntese de metabólitos secundários envolve principalmente a polimerização de metabólitos primários por enzimas dedicadas (muitas vezes referidas como enzimas centrais ou *core*) (Figura 6). Os metabólitos gerados por estas enzimas também possuem enzimas adicionais que podem alterar suas bioatividades. A enzima central define a classe química do metabólito secundário gerado: poliketídeos sintases (PKSs) produzem poliketídeos a partir de acil-CoAs, peptídeos não ribossômicos sintetases (NRPSs) geram peptídeos não ribossômicos a partir de aminoácidos e terpenos sintases e terpeno ciclases (TSs e TCs) geram terpenos de unidades de isopreno ativadas. Alguns metabólitos secundários são híbridos – são sintetizados a partir de duas sintases e/ou sintetases (Keller, 2019).

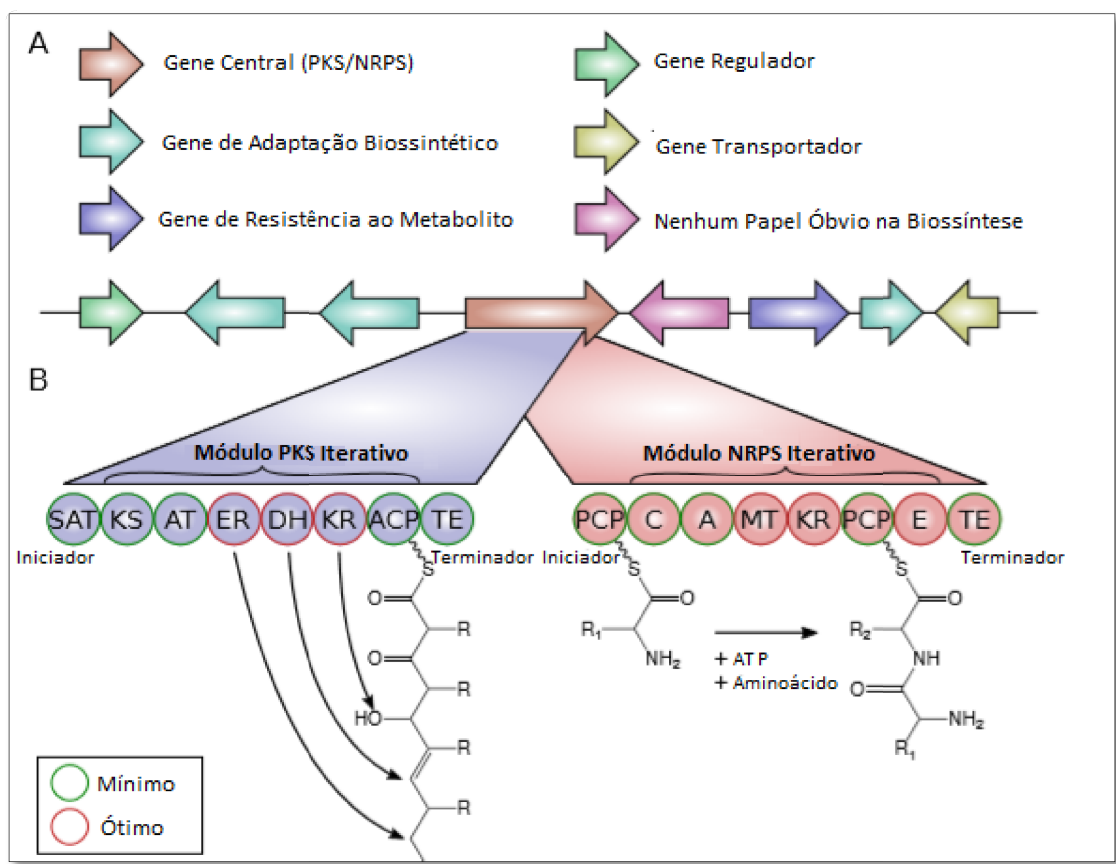


Figura 6: Visão geral de Agrupamentos de Genes Biossintéticos (BGCs) de metabólitos secundários. (A) BGC típico e as classes de genes mais comuns. (B) Arquitetura de domínios proteicos e princípio catalítico de genes estruturais de PKS e NRPS catalisando a formação de PKs e NRPs, respectivamente (adaptado de Nielsen, 2018).

Devido ao seu grande potencial na indústria farmacêutica e bioquímica sintética, as vias biossintéticas dos metabólitos secundários dos fungos são foco de estudos para elucidar a base genética para estes compostos. O método mais utilizado identifica as enzimas de suporte para sintases-chave, um grupo de cada vez (Andersen *et al*, 2012).

Muitas vias enzimáticas de metabólitos secundários em bactérias, fungos e plantas são codificadas em BGCs. As informações sobre esses grupos, vias e metabólitos é de difícil exploração (Medema *et al*, 2015).

A atribuição e predição de genes dos agrupamentos a partir da sequência do genoma não são precisas, apesar da disponibilidade de algoritmos eficientes para a identificação de possíveis genes biossintéticos de metabólitos secundários, especialmente PKSs, NRPSs e DMATs. Domínios proteicos relevantes podem ser previstos para alguns dos genes, mas os genes em grupos identificados geralmente têm funções desconhecidas, o que torna impossível prever sua inclusão. Além disso, os grupos de genes do metabolismo secundário frequentemente localizam os cromossomos, o que dificulta a separação de aglomerados baseados apenas em predições da função gênica (Andersen *et al*, 2012).

Atualmente, entende-se que apenas uma pequena parte do potencial de moléculas possíveis foi descoberta. Um grande número de BGCs ainda não foi caracterizado, sendo que muitos desses genes não são expressos. Sendo assim, são muitos os desafios a serem superados para o aproveitamento da diversidade de produtos naturais, desde cultivar linhagens em laboratório a efetivamente ativar a expressão de BGCs. Três abordagens podem ser utilizadas: manipulação de condições de cultivo, engenharia de NRPS e PKS e interferência genética (Guzmán-Chávez *et al*, 2018).

2.4 *Penicillium* spp. E METABÓLITOS SECUNDÁRIOS DE INTERESSE

O gênero *Penicillium* é uma fonte rica e diversificada de compostos bioativos, que variam em estrutura e são sintetizados via diferentes vias biossintéticas. Diferentes espécies foram isoladas de uma variedade de fontes ambientais coletadas em diferentes áreas ao redor do mundo. O metabolismo secundário destes isolados foi estudado e diferentes compostos foram identificados. Os próximos parágrafos apresentam uma visão geral dessa diversidade.

Bazioli *et al* (2017) reuniram um total de 40 metabólitos secundários biossintetizados por *Penicillium brasilianum*, incluindo metabólitos de natureza alcaloide, dicetopiperazinas, meroterpenoides, policetídeos e ciclodepsipeptídeos. Schürmann *et al* (2010) detectaram austin, dehidroaustin, D-manitol e ácido penicílico em inóculo desta mesma espécie em meio aquoso contendo dextrose. Fill *et al* (2010) encontraram cinco tipos de brasiliamidas, compostos potencialmente convulsivantes e bacteriostáticos, em cultivo em meio BDA de *P. brasilianum* isolado de *Melia azedarach*.

Em *Penicillium decumbens*, Lin *et al* (2011) encontraram um composto chamado ciclopenicilone, que apresentou efeito inibitório moderado contra a produção de óxido nítrico (NO) induzido por lipopolissacarídeos (LPS). Mamma *et al* (2004) detectaram atividades de β -glicosidase, α -ramnosidase e quercetase em cultivo em um biorreator usando rutina como única fonte de carbono e energia, para induzir a atividade glicosídica.

Para *Penicillium griseofulvum*, Banani *et al* (2016) estudaram a sequência do genoma de PG3 (29,3 Mb), mostrando que, filogeneticamente, *P. griseofulvum* se ramificou após a divergência de *Penicillium oxalicum*, mas antes da divergência de *P. chrysogenum*. A análise em todo o genoma revelou agrupamentos de genes putativos para biossíntese de patulina, griseofulvina e roquefortina C. Além disso, realizaram quantificação da produção de

metabólitos secundários *in vitro* e em maçãs, revelando a cinética de expressão dos principais genes de metabólitos secundários produzidos na maçã infectada. Nessa análise, encontraram *clusters* adicionais de metabólitos secundários, incluindo genes potencialmente responsáveis pela síntese de penicilina, yanuthone D, ácido ciclopiazônico e chanoclavina I.

Hautbergue *et al* (2017) encontraram 98 metabólitos secundários em *Penicillium verrucosum* em análise com HPLC-HPRS de seu cultivo. Ghanbari *et al* (2014) estudaram a produção de metabólitos secundários em diversas espécies de *Penicillium* por meio de análises de cromatografia gasosa, obtendo os seguintes resultados: (i) em *Penicillium jenseii*, encontraram ácido octadecanoico, undecano, antraceno, ácido oxálico, 2,4-nonadiene, phenanthreno, dodecano, dibenzotiofeno, ácido hexadecanoico, entre outros; (ii) em *Penicillium pusillum*, foram encontrados os compostos ácido decanedioico, uvidina B, hexadecano, tridecano, decano, camazuleno, entre outros; (iii) em *Penicillium purpurogenum*, encontraram azetidina, decano, timol, dibenzotiofeno, sarcocapnidina, antraceno, ácido octadecanoico, entre outros; (iv) e em *Penicillium canescens* foram encontrados os compostos dodecano, ácido oxálico, ácido hexadecanoico e tridecano, entre outros.

Tannous *et al* (2017) realizaram uma revisão na literatura sobre os metabólitos secundários produzidos por *Penicillium expansum*, dando ênfase à patulina. Além dessa micotoxina, também são produzidos outros policetídeos, como a citrinina, peptídeos não ribossômicos, como a communesina e roquefortina C, metabólitos híbridos entre policetídeos e peptídeos não ribossômicos, como chaetoglobosina A e C e citochalasina, terpenos como a geosmina e expansolida A e B e as andrastinas A, B, e C, além de meroterpenoides.

Shen *et al* (2013) estudaram extratos de *P. oxalicum* 0312F₁, que exibiu atividade inibidora contra o vírus do mosaico do tabaco e na replicação e proliferação de células de câncer gástrico humano e células de câncer hepático, buscando identificar os compostos produzidos por este fungo, que tenha potencial atividade inibitória contra fitovírus e células

tumorosas. Neste trabalho, foram identificados por espectroscopia de ressonância magnética nuclear os compostos 2-(4-hidroxibenzil)quinazolin-4(3H)-ona, rubinafina A, citreorosein e 4-hidroxifenilacetato de metila, além da descoberta do composto 2-(4-hidroxibenzoil)quinazolin-4(3H)-ona.

Até o presente, dados sobre a genômica de *P. echinulatum* relacionada a metabólitos secundários ainda é incipiente.

2.5 *Penicillium echinulatum*

O fungo filamentoso *Penicillium echinulatum* é estudado desde 1979, com o isolamento da linhagem selvagem 2HH a partir de fungos descobertos no trato digestivo de *Anobium punctatum*, uma espécie de coleóptero (Carrau *et al*, 1981). A partir disso, o fungo foi submetido a um programa de mutagênese e seleção até a obtenção em 2006 da linhagem 9A02S1, mutante parcialmente desreprimido à glicose, a partir de peróxido de hidrogênio e seleção em meio suplementado com 2-desoxiglicose (Dillon *et al*, 2006). Em 2009, o gene *EGL1* (endo-1,4-beta-glicanase-GH5-5) foi clonado e caracterizado (Rubini *et al*, 2009). Na continuidade do programa de melhoramento foi originado, em 2011, a linhagem S1M29, que possui alta produção de celulasas na presença de sacarose (Dillon *et al*, 2011).

As linhagens 2HH e S1M29 tiveram seus genomas sequenciados em 2013 (dados ainda não publicados) e em 2014 sua fisiologia foi estudada quanto à capacidade de metabolizar diferentes fontes de carbono (Schneider *et al*, 2014). Em 2015, foi realizada a caracterização e análise da expressão gênica de oito enzimas do complexo envolvendo celulasas e hemicelulasas da linhagem S1M29 (Zampieri, 2015). No ano seguinte, as linhagens selvagem e mutante foram comparadas, identificando 165 proteínas, sendo 111 as proteínas em comum; 36 exclusivas da linhagem selvagem e 18 que por sua vez tornaram-se

expressas no mutante. Foi observado que em torno de 44% das enzimas secretadas pelo mutante são *CAZymes* (*Carbohydrate-Active EnZymes*) (Schneider *et al*, 2016). Por fim, em 2019, acontece a montagem e anotação dos genomas das linhagens 2HH e S1M29.

3 MATERIAL E MÉTODOS

Neste capítulo, é explicada a origem dos dados utilizados e a metodologia empregada neste trabalho. O fluxograma da Figura 7 resume os procedimentos para o estudo dos metabólitos secundários de *P. echinulatum* 2HH e S1M29: etapas e ferramentas utilizadas em trabalhos anteriores até a obtenção das sequências utilizadas neste trabalho, que foram analisadas com o auxílio da ferramenta antiSMASH. Maiores detalhes experimentais estão na sessão 4 (Resultados), no formato de artigo científico.

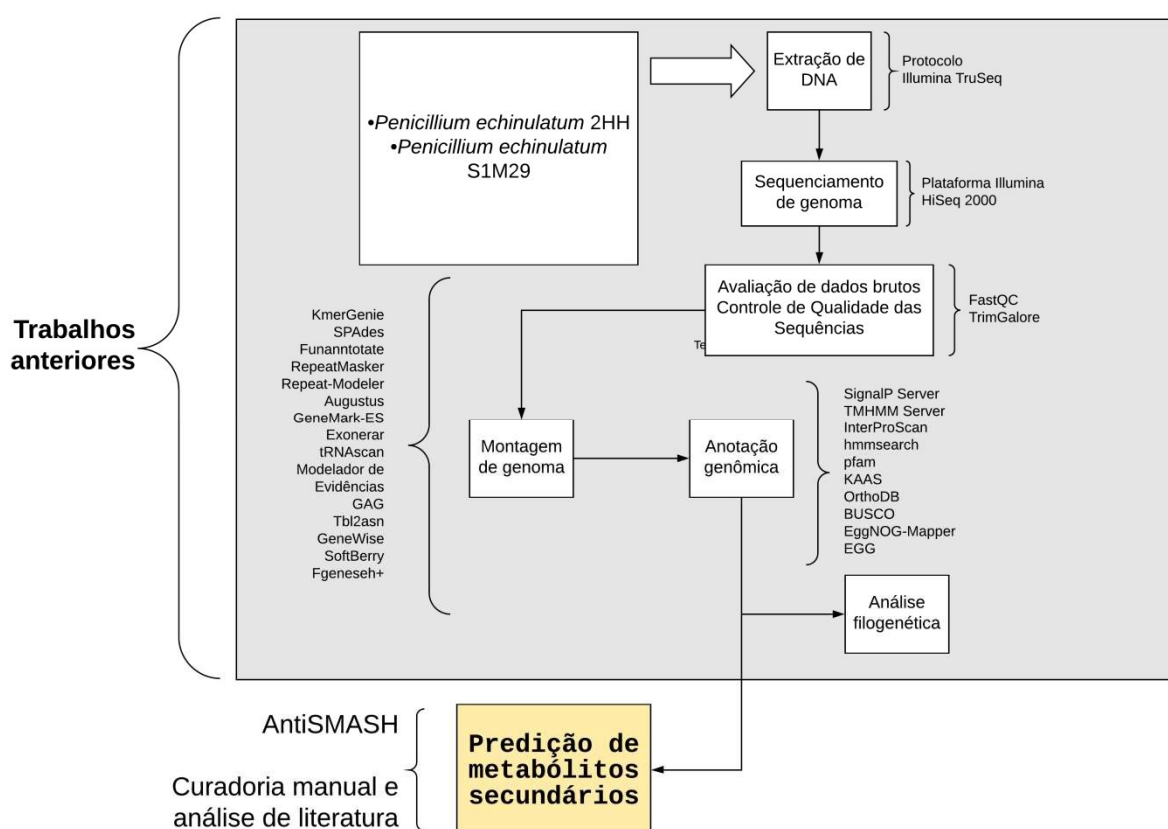


Figura 7: Fluxograma de trabalho para a análise do genoma do fungo *Penicillium echinulatum* 2HH e S1M29. Em cinza, a lista de ferramentas utilizadas em cada etapa de análise em trabalhos anteriores. Em amarelo, as etapas realizadas e apresentadas nesta dissertação.

3.1 SEQUENCIAMENTO E ANOTAÇÃO GENÔMICA

As linhagens selvagem (2HH) e mutante (S1M29) do fungo *P. echinulatum* empregadas neste trabalho pertencem à coleção de microrganismos do Laboratório de Enzimas e Biomassa do Instituto de Biotecnologia da Universidade de Caxias do Sul. Os dados utilizados foram obtidos em estudos anteriores (ainda não publicados), em que as 8.750 proteínas previstas para cada cepa foram funcionalmente anotadas usando o protocolo descrito por De Vries *et al* (2018). Para a anotação genômica, primeiramente foi realizado o sequenciamento do genoma completo das linhagens selvagem 2HH e mutante S1M29 de *P. echinulatum*, seguidas de procedimentos para verificação e controle de qualidade, montagem de leituras e predição de genes.

O DNA genômico foi extraído de ambas as linhagens 2HH e S1M29 utilizando-se protocolo *Illumina TruSeq*, seguido de sequenciamento usando a plataforma *Illumina HiSeq* 2000. Foram geradas bibliotecas de 100 pares de bases para cada linhagem. A avaliação de dados brutos usando FastQC (v.0.11.5) foi realizada para verificar a qualidade geral da sequência, a distribuição percentual do GC e a presença ou ausência de sequências super-representadas. Usando o TrimGalore (v.0.4.4), as leituras foram adaptadas, bases de baixa qualidade das extremidades 3' e 5' foram removidas antes de serem cortadas e qualidade aparada (valores de qualidade de sequenciamento, Q28) e pequenos fragmentos (30 pb) abandonado após o corte do adaptador e corte de qualidade.

A montagem do genoma foi realizada com a ferramenta KmerGenie (v.1.7044), para predizer o melhor k e tamanho de genoma e SPAdes (v. 3.11.0), um *multi-k assembler*, utilizando-se todos valores ímpares ($21 \leq k \leq 83$). *Scaffolds* com menos de 500 pb foram removidos utilizando Funannotate (v.0.7.2). As sequências de repetição foram mascaradas usando RepeatMasker (v.4.0.7 2), Repeat-Modeler (v.1.0.8 1) e a biblioteca RepBase (2017-

01-27). Os conjuntos mascarados de repetições foram usados para codificar a predição de genes executada independentemente com um conjunto de localizadores de genes incluindo Augustus (v.3.2.2), GeneMark-ES (v.4.33), Exonerar (v.2.2.0) e tRNAscan (v.1.23). Funannotate (v.0.7.2) foram utilizados para tomar várias predições diferentes de genes e produzir modelos de genes consensuais e mesclá-los, usando o Modelador de Evidências (v.0.1.3) como núcleo, então GAG (v.2.0.1) e Tbl2asn (v.25.3) foram acionados para gerar os arquivos contendo o conjunto de genes previsto. *Penicillium* sp. foi usado para revisar e complementar os genes previstos por pesquisas de homologia usando versões *online* do GeneWise, SoftBerry Fgenesh+. Este conjunto representativo foi sujeito a análise e anotação adicionais.

As proteínas preditas foram funcionalmente anotadas utilizando genômica de fungos. Todos os modelos genéticos previstos foram anotados funcionalmente usando o SignalP Server (v.5.0) para prever a presença e localização de sítios de clivagem de peptídeo sinal, TMHMM Server (v.2.0) foi utilizado para identificar hélices transmembranares e InterProScan (v.5.25-64.0) foi usado para mapear famílias Interpro, domínios e termos GO, hmmsearch (v.3.2.1) foi usada para identificar os domínios do PFAM sobre o banco de dados do pfam (v.32.0 - 2018-06) usando pontos de corte de coleta. As atribuições de KO (KOGG Orthology) e as vias KEGG geradas automaticamente foram atribuídas às proteínas previstas usando KAAS (Servidor de anotação automático KEGG). Ortólogos conservados selecionados de OrthoDB v.9 usando conjuntos de dados: Eucariota, Fungi, Dikarya, Ascomycota e Eurotiomycetes foram designados usando BUSCO (v.3.1.0) por avaliação de completude de anotação. EggNOG-Mapper (v.0.99.3) foi usado para mapear categorias funcionais gerais a partir de clusters de COG e aglomerados de EGG com precursores ortogênicos a partir do EGG v.4.5.1.

3.2 PREDIÇÃO DE *BIOSYNTHETIC GENE CLUSTERS*

Neste trabalho de análise de metabólitos secundários do fungo, foi utilizado o *software* antiSMASH v.5.0 (*antibiotics & Secondary Metabolite Analysis Shell*), uma ferramenta de identificação, anotação e análise de agrupamentos de genes de metabolismo secundário de bactérias, fungos e plantas (Medema *et al*, 2011; Blin *et al*, 2019). A sequência *.gbk* anotada foi inserida na ferramenta, mantendo-se as configurações padrões, como mostra a Figura 8, para que seja comparada com os bancos de dados utilizados, especialmente o *Minimum Information about a Biosynthetic Gene Cluster* (MIBiG) (Medema *et al*, 2015). Concluída a análise, os resultados obtidos foram analisados manualmente a fim de que sejam criadas as relações e inferências necessárias para o estudo de metabólitos secundários preditos, incluindo comparações de revisão de literatura.

Nucleotide input Results for existing job

Search a genome sequence for secondary metabolite biosynthetic gene clusters Load sample input Open example output

Notification settings

your@email.com Email address (optional)

Data input

Upload file Get from NCBI NCBI acc # NCBI accession number of desired sequence

Detection strictness: relaxed

strict relaxed loose

- Detects well-defined clusters containing all required parts.
- Detects partial clusters missing one or more functional parts.

Extra features All off All on

☒ KnownClusterBlast ☐ ClusterBlast ☒ SubClusterBlast

☒ ActiveSiteFinder ☐ Cluster Pfam analysis ☐ Pfam-based GO term annotation

☐ Cluster-border prediction based on transcription factor binding sites (CASSIS)

Submit

fungi
SMASH

If you have found fungal antiSMASH useful, please [cite us](#).

Figura 8: Recorte de tela de *input* da ferramenta antiSMASH.

Enquanto outras ferramentas permitem a identificação apenas de NRP e de PK tipo I, antiSMASH é capaz de detectar todas as classes de agrupamentos de genes para a biossíntese de metabólitos secundários. Esta ferramenta permite a rápida identificação, anotação e análise em todo o genoma de agrupamentos de genes de biossíntese de metabólitos secundários em genomas de bactérias e fungos. Ele integra e cruza ligações com um grande número de ferramentas de análise de metabólitos secundários *in silico* que foram publicadas anteriormente, sendo alimentado por várias ferramentas de código aberto: NCBI BLAST+, HMMer 3, Muscle 3, FastTree, PySVG e JQuery SVG (Medema *et al*, 2011).

AntiSMASH pode ser executado a partir de um servidor web (<http://antismash.secondarymetabolites.org/>) ou como uma versão *desktop* independente em um computador padrão. Ele é capaz de detectar rapidamente todas as classes conhecidas de agrupamentos de genes de biossíntese de metabólitos secundários, fornecer anotações funcionais detalhadas de NRPS / PKS e prever a estrutura química dos produtos NRPS / PKS. Os genes são extraídos ou previstos a partir da sequência de nucleotídeos de entrada e os aglomerados de genes são identificados com o gene de assinatura pHMMs. Posteriormente, várias análises podem ser realizadas: análise e anotação do domínio NRPS / PKS, previsão da estrutura química do núcleo de PKSs e NRPSs, análise comparativa de agrupamentos de genes *ClusterBlast* e análise da família de proteínas do metabolismo secundário *smCOG*. A saída é visualizada em uma página da web XHTML interativa e todos os detalhes são armazenados em um arquivo EMBL para análise e edição adicional em um navegador de genoma. Um arquivo do Microsoft Excel com uma visão geral de todos os agrupamentos de genes detectados e seus detalhes também é gerado (Figura 9) (Blin *et al*, 2013).

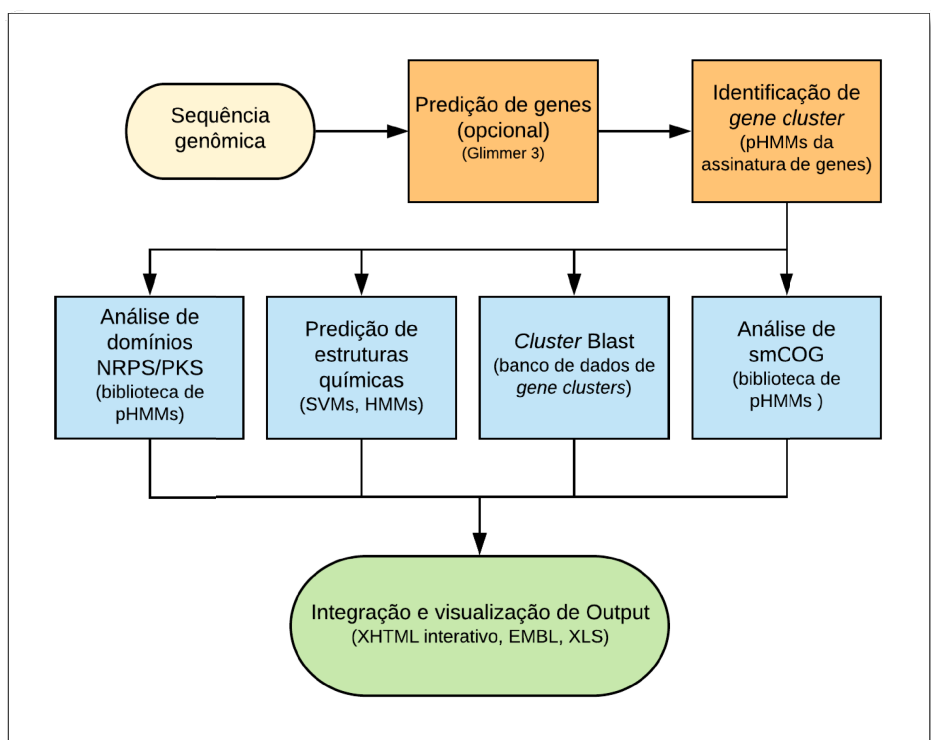


Figura 9: Pipeline de análise genômica de metabólitos secundários realizada por antiSMASH (modificado de Blin *et al*, (2013).

Semelhanças evolutivas entre um agrupamento de genes consultado e outros agrupamentos de genes são detectadas e visualizadas para poder inferir rapidamente funções de genes e *operons* com base na homologia. A partir desses genes, agrupamentos de genes ortólogos de metabolismo secundário (smCOGs) são construídos, utilizados para prever e categorizar as funções dos genes acessórios e para calcular as árvores filogenéticas para cada gene (Blin *et al*, 2013).

Recentemente, o *software* passou por atualização (v.5.0) que agora conta com regras para a detecção de 52 diferentes tipos de BGC, incluindo a biossíntese de acil-aminoácidos, β -lactonas, RiPPs de fungos, RaS-RiPPs, éteres difenólicos polibromados, C-nucleosídeos, cetonas semelhantes a PPY e lipolantinas, além de previsões mais detalhadas para *clusters* de genes codificadores da sintase de policetídeo tipo II. Além disso, uma nova regra '*nrps-like*' foi definida para fragmentos de NRPS, ou seja, NRPSs atípicos que não possuem a arquitetura típica do módulo C-A-T (Blin *et al*, 2019).

Alguns conceitos são importantes para compreender os resultados apresentados pela ferramenta: (i) Núcleo (*core*): a área mínima que contém um ou mais genes que codificam enzimas para um único tipo de BGC que são detectados pelas regras de detecção selecionadas manualmente. Esses genes não precisam ser contíguos, mas podem estar a uma certa distância de corte, conforme definido pela regra de detecção para o tipo de BGC em questão; (ii) Vizinhaça: distância para cima e para baixo do núcleo do *cluster* usado para encontrar genes ou enzimas adaptados; (iii) *Protocluster*: Contém núcleo e vizinhaça em ambos os lados do núcleo; cada *protocluster* sempre terá um único tipo de produto (por exemplo, NRPS); (iv) *Cluster* candidato: Contém um ou mais *protoclusters*. Essas definições permitem melhor a modelagem de *clusters* híbridos, como híbridos PKS/NRPS, que combinam duas ou mais classes biossintéticas diferentes (conforme identificado nas regras de detecção) ou casos em

que uma classe é usada para biossintetizar um precursor para uma segunda classe; e (v)

Região: contém um ou mais agrupamentos de *clusters* candidatos (Blin *et al*, 2019).

3.3 CURADORIA DE PREDIÇÕES

Ainda que as ferramentas disponíveis atualmente para as análises genômicas, filogenéticas e de metabolismo secundário sejam confiáveis, é necessário revisar manualmente os resultados para garantir a qualidade das predições realizadas. Além disso, o metabolismo secundário, por ser característico de cada espécie e por evidenciar similaridades filogenéticas entre os seres vivos, é estudado a partir de inferências com espécies relacionadas. Sendo assim, as predições foram verificadas manualmente e comparadas com dados oriundos de revisão de literatura.

4 RESULTADOS E DISCUSSÃO

4.1 ARTIGO CIENTÍFICO A SER SUBMETIDO AO PERIÓDICO “*FUNGAL BIOLOGY*”

Os resultados estão apresentados no formato do artigo científico “*Overview of Penicillium echinulatum biosynthetic gene cluster: an in silico approach for secondary metabolism potential*”. O artigo apresenta os resultados obtidos com a análise feita pela ferramenta antiSMASH: os agrupamentos em regiões e os *clusters* de genes biossintéticos encontrados. Também está descrita uma análise comparada de micotoxinas preditas no genoma do fungo, a partir de alinhamento de sequências com o auxílio da ferramenta BLAST.

Overview of *Penicillium echinulatum* biosynthetic gene cluster: an *in silico* approach for secondary metabolism potential

Nicole Anne Modena^a, Alexandre Rafael Lenz^a, Fernanda Pessi de Abreu^a, Scheila de Avila e Silva^a, Aldo José Pinheiro Dillon^a

^aInstituto de Biotecnologia, Universidade de Caxias do Sul, Caxias do Sul, Rio Grande do Sul, Brazil

Abstract

Filamentous fungi are known for their production of secondary metabolites of interest, used in industry, in enzymes production and other relevant compounds. Due to this, the development of production improvement methods is increasing. One of these improved fungi is *Penicillium echinulatum* 2HH resulting in its optimal cellulase producing strain S1M29. The objective of this work was to predict secondary metabolites and mycotoxins in *P. echinulatum*. The annotated genomic sequence was analyzed by antiSMASH tool, followed by manual curation. Mycotoxins were analyzed by BLAST tool, compared to known genes and checked manually. AntiSMASH analysis identified families of terpene, non-ribosomal peptide synthetase, similar to non-ribosomal peptide synthetase fragment, type I PKS, hybrid cluster between non-ribosomal peptide synthetase and type I PKS, betalactone containing protease inhibitor and a siderophore cluster. Regions containing biosynthetic gene clusters were identified as clavarinic acid, naphthopyrone, nidulanin A, oxaleimide C, phyllostictine and 4,4'-piperazine-2,5-diyl dimethyl-bis-phenol. For mycotoxin analysis, only HC-toxin has potential for expression by the fungus genome. The results do not indicate that these enzymes are actually secreted, but indicate orthologous gene groups. These are prediction data, which might be verified by experimental analysis of gene expression and identification of secreted secondary metabolites.

Keywords: fungal secondary metabolites, biosynthetic gene clusters, comparative genomics

1. Introduction

Filamentous ascomycetes fungi are known to have important roles in natural cycles. Due to their metabolically versatile biocatalysis, they have been employed in the industrial production of enzymes, antibiotics, organic acids and human food [1]. The development of methods to improve enzymatic hydrolysis and fermentation is also important to increase the production of biomolecules, liquid fuels and chemical supplies by lignocellulosic materials,

Email addresses: namodena@ucs.br (Nicole Anne Modena), arlenz@ucs.br (Alexandre Rafael Lenz), fpabreu1@ucs.br (Fernanda Pessi de Abreu), sasilva6@ucs.br (Scheila de Avila e Silva), ajpdillo@ucs.br (Aldo José Pinheiro Dillon)

present in the cell wall of superior plants [2]. The economically efficient production of second generation ethanol could be obtained by the development of more productive strains of microorganisms and processes with increased enzyme levels. This fact has motivated the implementation of genetic programs for cellulolytic filamentous fungi and the development of fermentation processes that could be more efficient [3–6].

Cellulases are mainly produced using filamentous fungi. Several species in *Penicillium* genus might be alternatives for the production of second generation biofuels, due to their high β -glucosidase activity [5]. In China, *Penicillium oxalicum* (previously named as *Penicillium decumbens*) has been used for industrial high-scale production since the 1990s, after its isolation from decayed straw-covered soil and a long-term strain improvement process to increase the cellulase production ability [7].

In Brazil, cellulases production is studied using the fungus *P. echinulatum*. A wild strain, 2HH, was isolated from the furniture beetle *Anobium punctatum* [8]. The strain S1M29 is a genetic variant isolated from strain 9A02S1 (a partially glucose-derepressed mutant) by mutagenesis with hydrogen peroxide [9, 10]. This mutant shows higher cellulase productivity and these enzymes are faster than the parental strain [9]. Significant efforts to understand the mechanisms to regulate the expression of genes related to cellulases expression have been developed, according to obtaining genetic variants carrying specific markers of interest [4].

These mutants of *P. echinulatum* are able to secrete high cellulase titers in both submerged and solid-state fermentation [5]. Their Filter Paper Activity (FPA) and β -glucosidases enzymes have good thermal stability at 50°C, justifying its application in enzymatic hydrolysis of cellulose and lignocelluloses processes [11].

Cellulolytic microorganisms present several genes related to enzyme production. Most of them are still not identified, which could operate jointly to achieve an efficient cellulase secretion. Due to this, the known cellulase genes are only a part of all genes involved in this process. Additionally, these genes could vary in different species or genus of cellulase producer microorganisms [4].

Penicillium genus is known as an important producer of secondary metabolites of interest for human activities since the 1930s following Alexander Fleming's findings on penicillin [12, 13]. After the development of penicillin fermentation methodologies, a large number of primary and secondary metabolites of commercial interest were discovered [14]. The primary metabolism of an organism involves a series of biochemical reactions that provides essentials macromolecules for organism survival, such as DNA, RNA, proteins, polysaccharides and lipids. However, products derived from biological processes relevant to human activities are mostly secondary metabolites, which have no function in culture growing. Considering that the primary metabolism is basically the same for all living systems, the secondary metabolism is usually specific to a particular lineage of plants and microorganisms, being the most well-known secondary metabolites are antibiotics [14, 15].

Secondary or Specialized Metabolites (SM), also called Natural Products (NP), are the base of drugs such as antibiotics, antitumor and antiviral agents, antipyretics, hallucinogens, and lipid lowering agents; molecules important for agricultural applications, such as herbicides or phytotoxins; as food additives for color, flavor and sweeteners; in fragrances; as precursors for the synthesis of plastics and with relevance in scientific research in cellular

and molecular biology [16, 17]. Throughout the twentieth century, natural products have contributed to the advancement of human understanding of biology and the development of medicine [18].

SMs are usually classified in four chemical categories: polyketides, non-ribosomal peptides, terpenes and prenylated tryptophan derivatives. Each class of secondary metabolite has a specific backbone enzyme required to its production, named polyketide synthases (PKSs), non-ribosomal peptide synthetases (NRPSs), terpene synthases/cyclases (TCs) and dimethylallyl tryptophan synthases (DMATSs), respectively [19]. The predicted secondary metabolite clusters are defined according to their backbone enzyme [20].

Unlike genes required for the synthesis of a primary metabolite, which are dispersed throughout the fungal genome, genes that encode enzymatic activities to produce any secondary metabolites are organized in a cluster, forming a group of biosynthetic genes [21]. Biosynthetic gene cluster (BGC) can be defined as a clustering of genes in a given genome that together encode a pathway for the production of a metabolite, often biosynthesized by multienzymatic pathways [22, 23]. In view of the fact that the genomes of many microorganisms encode the production of secondary metabolites from multiple BGCs, genome mining has been established as an important tool complementing bioassays in the study of natural products, supported by development and availability of various data mining tools [24].

The aim of this paper was to identify BGCs related to secondary metabolism, including mycotoxins in *P. echinulatum* wild and mutant strain. For this, both genome annotation of the wild strain 2HH and its mutant S1M29 of *P. echinulatum* were used and the data were compared with other genomes of fungi described as cellulases-producing strains.

2. Material and Methods

2.1. Penicillium echinulatum 2HH and S1M29 genome

In previous works, the fungus *P. echinulatum* was isolated, cultivated and studied for the production of lignocellulosic enzymes [3, 11, 25]. To optimize this production, the fungus underwent mutagenesis and selection of best producing mutants [4, 9]. These strains had their genetic material extracted, sequenced, assembled and annotated to relate the genes present to their function, in yet to be published studies.

2.2. Secondary metabolites prediction

The annotated .gff sequence has been inserted into the antiSMASH software v.5.0 (<http://fungismash.secondarymetabolites.org>), a tool for identification, annotation and analysis of groups of secondary metabolism genes [26], which compares the query sequence with MIBiG — Minimum Information about a Biosynthetic Gene Cluster database [22]. After the analysis, the results obtained were manually curated by comparison with related literature in order to create the necessary relationships and inferences for the prediction of secondary metabolites.

2.3. Mycotoxins analysis

In order to gauge the antiSMASH assay to mycotoxins BGCs, the annotated genome of *P. echinulatum* 2HH was compared to known mycotoxins gene clusters deposited in MIBiG database using the BLAST tool, followed by manual curation in literature.

3. Results and Discussion

3.1. Secondary metabolites clusters according to backbone enzymes

The 2HH strain assembly obtained 708 scaffolds in a 30.44 Mb draft genome, while the assembly of S1M29 strain identified 682 scaffolds in a 30.41 Mb draft genome. The secondary metabolites cluster prediction in *P. echinulatum* using antiSMASH software package has predicted 26 regions for both wild and mutant strains, being the same number of cluster an expected information due the strain S1M29 be a mutant of 2HH. In *P. echinulatum* 2HH, 588 scaffolds were not identified as a backbone enzyme region. In *P. echinulatum* S1M29, 568 scaffolds couldn't be clustered in any region.

In order to compare these strains of *P. echinulatum* with other filamentous fungi usually used in enzymatic hydrolysis of cellulose and lignocellulose processes, *P. oxalicum* 114-2, a similar annotated genome sequence available in GenBank database, was submitted to antiSMASH assay. It has identified 46 regions in its genome. A single scaffold was not classified as a region.

The most abundant families of secondary metabolite gene clusters identified in *P. echinulatum* genome are terpene, non-ribosomal peptide synthetase cluster (NRPS) and similar to non-ribosomal peptide synthetase fragment (NRPS-like). It was also found type I PKS (polyketide synthase), hybrid cluster between non-ribosomal peptide synthetase and type I PKS, a beta-lactone containing protease inhibitor and a siderophore cluster (Figure 1).

The Figure 1 allows to compare the amount of predict BGC in the three genomes. In *P. echinulatum*, there is no difference between the wild (2HH) and the mutant (S1M29) strains. On the other hand, the number of predicted BCG is different when compared to *P. oxalicum* 114-2. This can be explained by the fact that secondary metabolites are very specific for a given organism.

Polyketides are NP containing multiple β -hydroxyketone or β -hydroxyaldehyde ($-H_2C(=O)CH_2CH(OH)CH_2C(=O)-$) functional groups [27]. They are the most abundant fungal secondary metabolites. One of the genetically best-characterized polyketides is the yellow spore-pigment intermediate naphthopyrone in *Aspergillus nidulans*, the carcinogen aflatoxin and the cholesterol-lowering compound lovastatin. Non-ribosomal peptides are derived from proteinogenic amino acids and non-proteinogenic amino acids by non-ribosomal peptide synthetases (NRPSs). Examples of NRP are the peptaibol synthetase of *Trichoderma virens*; in *Tolypocladium niveum*, a NRPS synthesizes cyclosporin, an immunosuppressive drug, and in *A. nidulans*, there is a NRPS gene involved in the biosynthesis of the siderophore ferricrocin. Terpenes are known as odoriferous plant metabolites, but fungi also synthesize them, as example, aristolochenes, carotenoids, gibberellins, indole-diterpenes and trichothecenes. All terpenes are composed of several isoprene units, can be linear or cyclic, saturated or unsaturated, and can be modified in various ways [28].

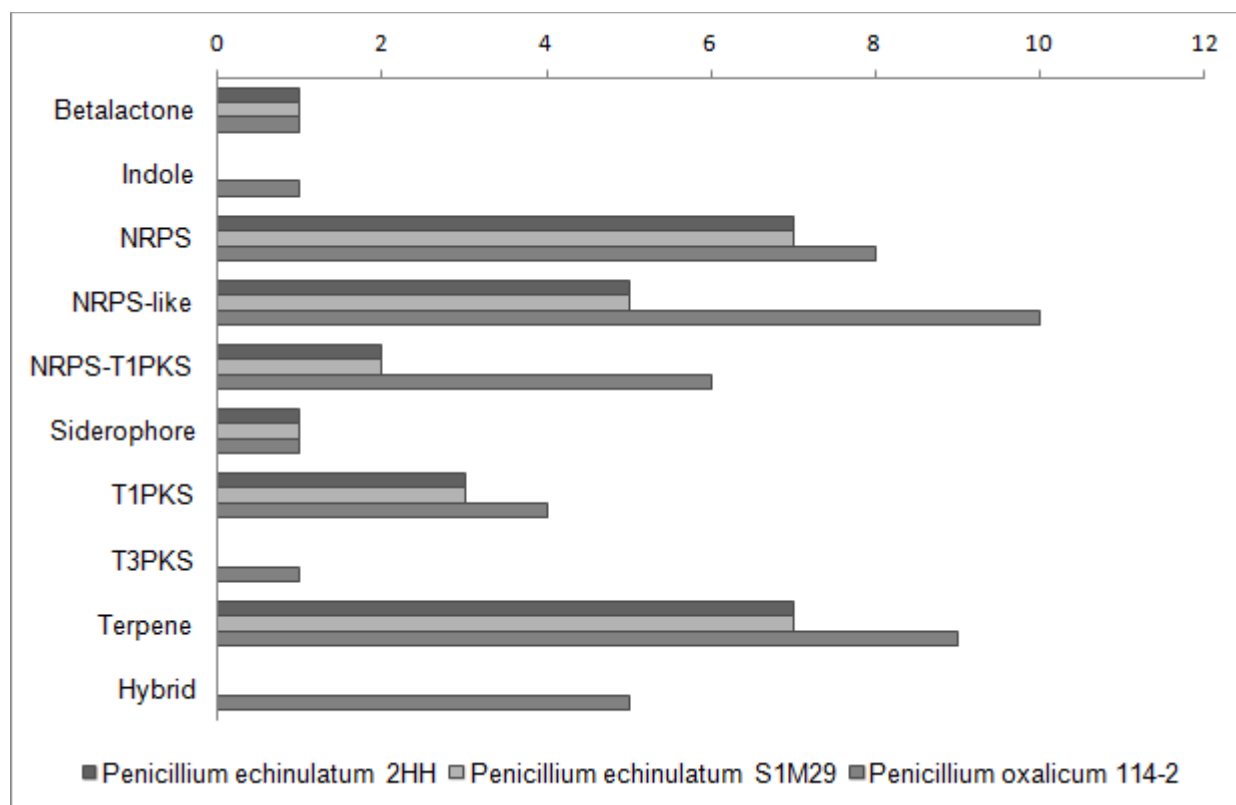


Figure 1: Distribution of predicted secondary metabolite clusters, according to their backbone enzymes. “Hybrid” refers to clusters with multiple core genes belonging to different secondary metabolite families, NRPS (non-ribosomal peptide synthetase), PKS (polyketide synthase).

3.2. Biosynthetic Gene Clusters

As expected due to their phylogenetic proximity, the antiSMASH assay identified similar known biosynthetic gene clusters (a percent of genes that shows similarity) in seven regions for 2HH and S1M29 strains: clavarinic acid, naphthopyrone, nidulanin A, two clusters as oxaleimide C, phyllostictine A / phyllostictine B and 4,4’-piperazine-2,5-diyl dimethyl-bis-phenol (Figure 2).

3.2.1. 4,4’-piperazine-2,5-diyl dimethyl-bis-phenol

The 4,4’-piperazine-2,5-diyl dimethyl-bis-phenol BGC was predicted with 50% of similarity, but the MIBiG information is no longer available, so it was not possible to relate the orthologous genes identified (Figure 3). No literature review was found for this compound until the date of this article was written.

3.2.2. Clavarinic acid

The clavarinic acid BGC was identified based on the comparison with the *Hypholoma sublateritium* cluster hosted on MIBiG database. The ACF70484.1 gene is similar to the PECH008628 gene in *P. echinulatum* 2HH and PECH005203 in *P. echinulatum* S1M29.

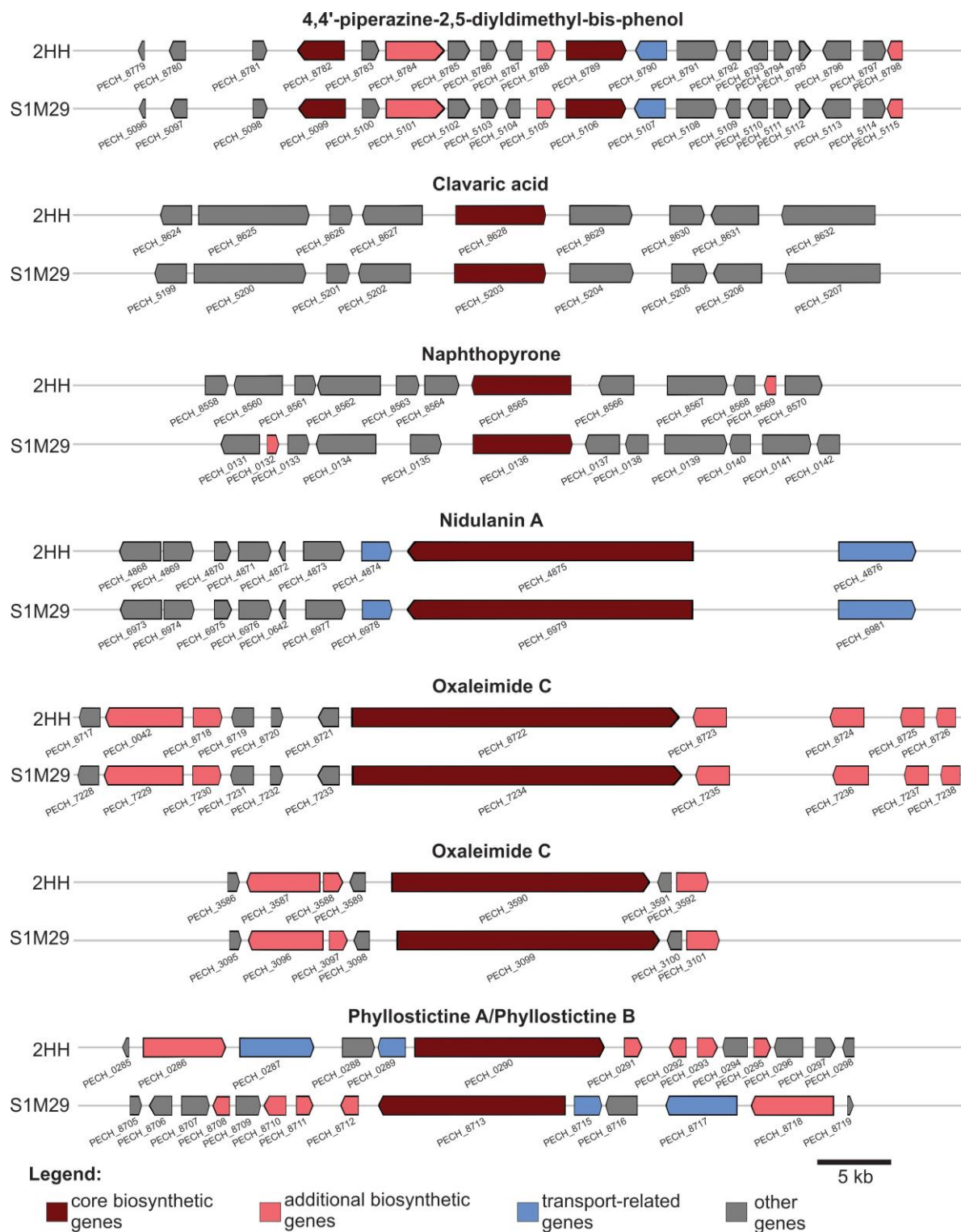


Figure 2: Biosynthetic gene cluster predicted by antiSMASH assay in *P. echinulatum* 2HH and S1M29 annotated genome.

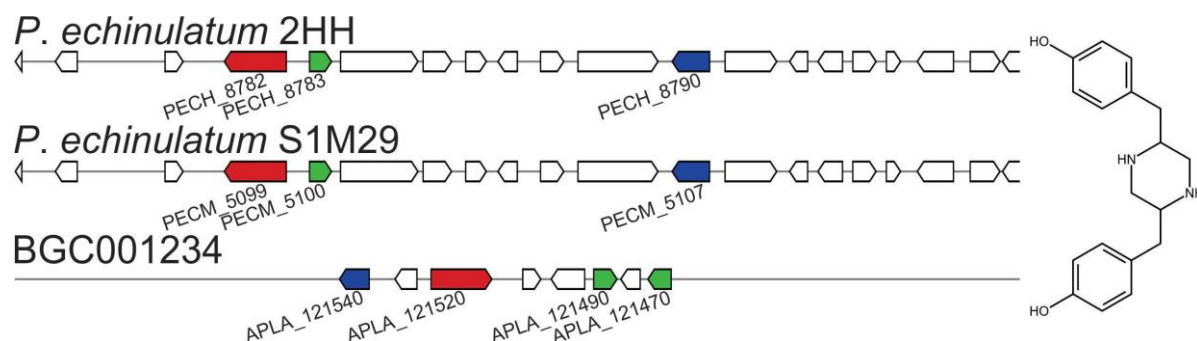


Figure 3: 4,4'-piperazine-2,5-diyl dimethyl-bis-phenol BGC and molecular formula. The genes in same the color means they are orthologous.

The antiSMASH assay showed 100% of similarity in the query sequence and the database sequence, but there is only one gene in the *H. sublateritium* BGC similar to one gene in nine genes present in that pathway in *P. echinulatum* (Figure 4).

Clavarinic acid is a triterpenoid described as an antitumor compound due to its ability of inhibit the Farnesyl-protein transferase (FPTase), an essential catalytic for oncogenic activity in oncogene-mediated tumors [29]. There is a minor chance that *P. echinulatum* might secrete this compound, due to its position in the metabolic pathway — only one in nine genes in these fungi showed similarity to the related pathway.

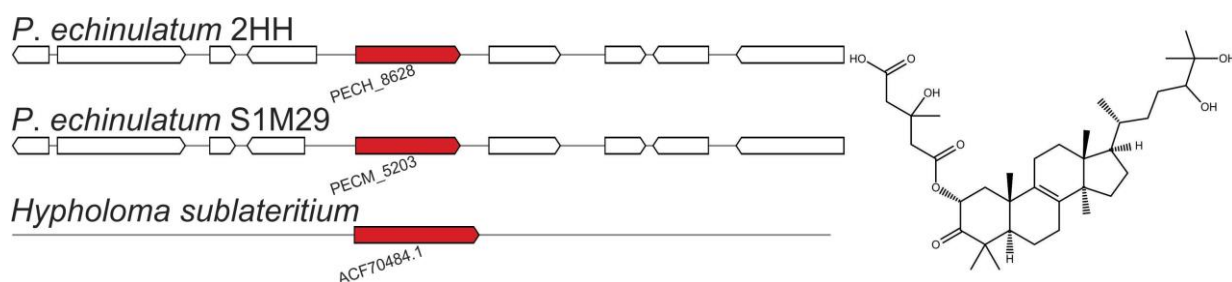


Figure 4: Clavarinic acid BGC and molecular formula. The genes in same the color means they are orthologous.

3.2.3. Naphthopyrone

Naphthopyrone is a heptaketide involved in the conidial pigmentation [30], a precursor in the synthesis of 1,8-dihydroxynaphtalene (1,8-DHN) melanin [31]. AntiSMASH identified the naphthopyrone BGC from *A. nidulans* FGSC A4. The gene ANIA_08209 from *A. nidulans* showed 100% of similarity with the gene PECH_008565 in *P. echinulatum* 2HH and the gene PECM_000136 in *P. echinulatum* S1M29 (Figure 5). Despite, these results doesn't necessary indicate that *P. echinulatum* would be able to secrete this compound, because of the other genes present in this pathway: *A. nidulans* has a single gene in its role, while *P. echinulatum* has 12.

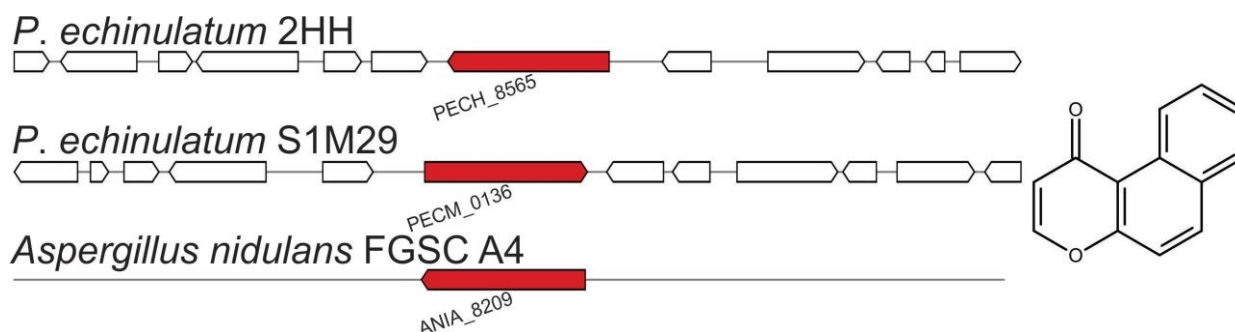


Figure 5: Naphtopyrone BGC and molecular formula. The genes in same the color means they are orthologous.

3.2.4. Nidulanin A

AntiSMASH identified a nidulanin A BGC from *A. nidulans* FGSC A4 with 75% of genes showing similarity. Three in four genes in *A. nidulans* pathway showed similarity with three in nine genes in *P. echinulatum*: the gene ANIA_01243 from *A. nidulans* is similar to the gene PECH_004874 in *P. echinulatum* 2HH and the gene PECM_006978 in *P. echinulatum* S1M29; the gene ANIA_01242 from *A. nidulans* is similar to the gene PECH_004875 in *P. echinulatum* 2HH and the gene PECM_006979 in *P. echinulatum* S1M29 and the gene ANIA_01240 from *A. nidulans* is similar to the gene PECH_004876 in *P. echinulatum* 2HH and the gene PECM_006981 in *P. echinulatum* S1M29 (Figure 6).

Nidulanin A is a tetracyclopeptide/isoprene isolated from *A. nidulans* [32]. Its gene cluster is conserved in all *Aspergillus* spp., and has yet to be tested for pharmaceutical properties [33].

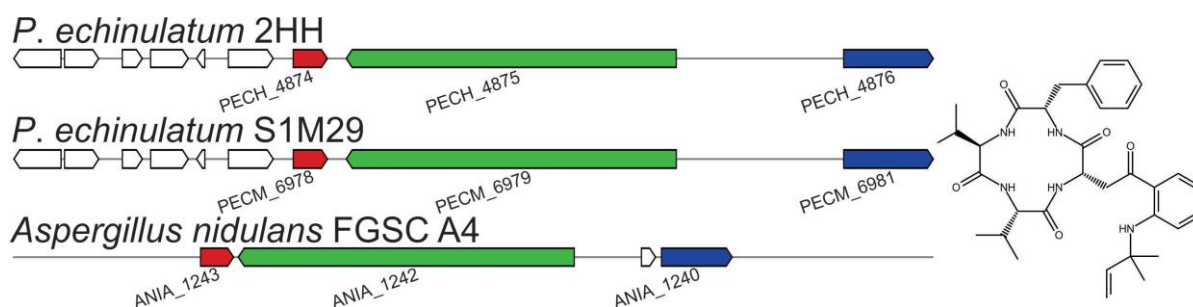


Figure 6: Nidulanin A BGC and molecular formula. The genes in same the color means they are orthologous.

3.2.5. Oxaleimide C

The oxaleimide C BGC was identified in two genomic regions. In one of these regions, the genes showed 10% of similarity with oxaleimide C BGC from *P. oxalicum* 114-2: the gene PDE_04012 is an orthologous with PECH_003589 and PECM_003098 and the gene PDE_04019 is an orthologous with PECH_003591 and PECM_003100. In the second region, the genes showed 15% of similarity with oxaleimide C BGC from *P. oxalicum* 114-2:

the gene PDE_04021 is an orthologous with PECH_008717 and PECM_007228; the gene PDE_04024 showed similarity with PECH_8719 and PECM_007231; the gene PDE_04025 showed similarity with PECH_008718 and PECM_007230 (Figure 7). No literature review was found for this compound until the date of this article was written.

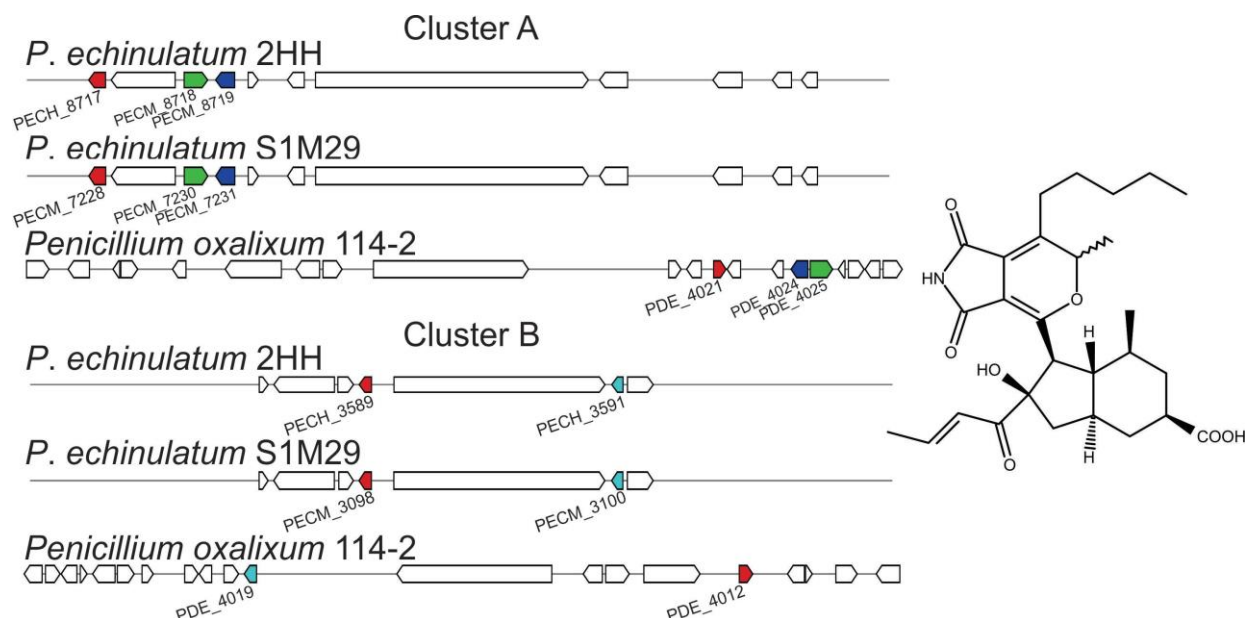


Figure 7: Oxaleimide C BGC and molecular formula. The genes in same the color means they are orthologous.

3.2.6. Phyllostictine A / phyllostictine B

The phyllostictine A biosynthetic gene cluster from *Phyllosticta cirsii* was compared to *Penicillium echinulatum*, as four in fourteen genes in the query sequence has showed similarity to four in ten genes from *P. cirsii*. The gene ARP51720.1 has shown similarity to PECH_000294 and PECM_008709; the gene ARP517109.1 has shown similarity to PECH_000295 and PECM_008708; the gene ARP51714.1 has shown similarity to PECH_000293 and PECM_008710 and the gene ARP51711.1 has shown similarity to PECH_000290 and PECM_008713 (Figure 8).

Phyllostictine is an oxazatricycloalkenone, produced by *Phyllosticta cirsii*, a phytotoxic metabolite with potential herbicidal activity. Four compounds were identified, named Phyllostictine A-D in decreasing order of phytotoxicity. Beside its phytotoxic properties, phyllostictine A has no antifungal activity, an antibiotic activity only against Gram+ bacteria, and a zootoxic activity when tested at high concentrations in *Artemia salina* [34].

3.3. Mycotoxin analysis

The mycotoxins analysis compared the number of potential genes found in *P. echinulatum* 2HH with other known genes in filamentous fungi, and its results are shown in Figure 9.

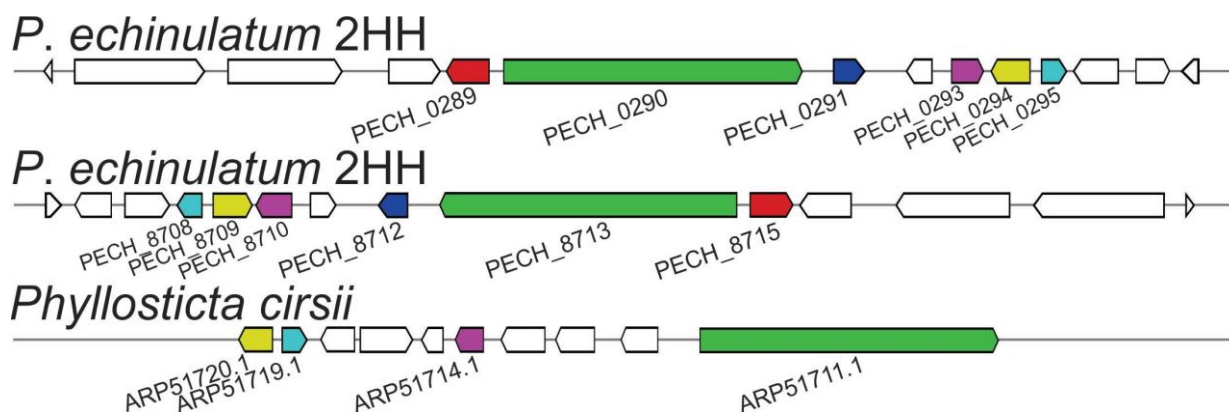


Figure 8: Phyllostictine A and B BGC and molecular formula. The genes in same the color means they are orthologous.

For the known mycotoxins present in filamentous fungi, the analysis of genes in *P. echinulatum* with orthology does not indicate that these mycotoxins are expressed, since the number of genes in the cluster is greater than the number of genes that have orthology with any known gene to mycotoxins.

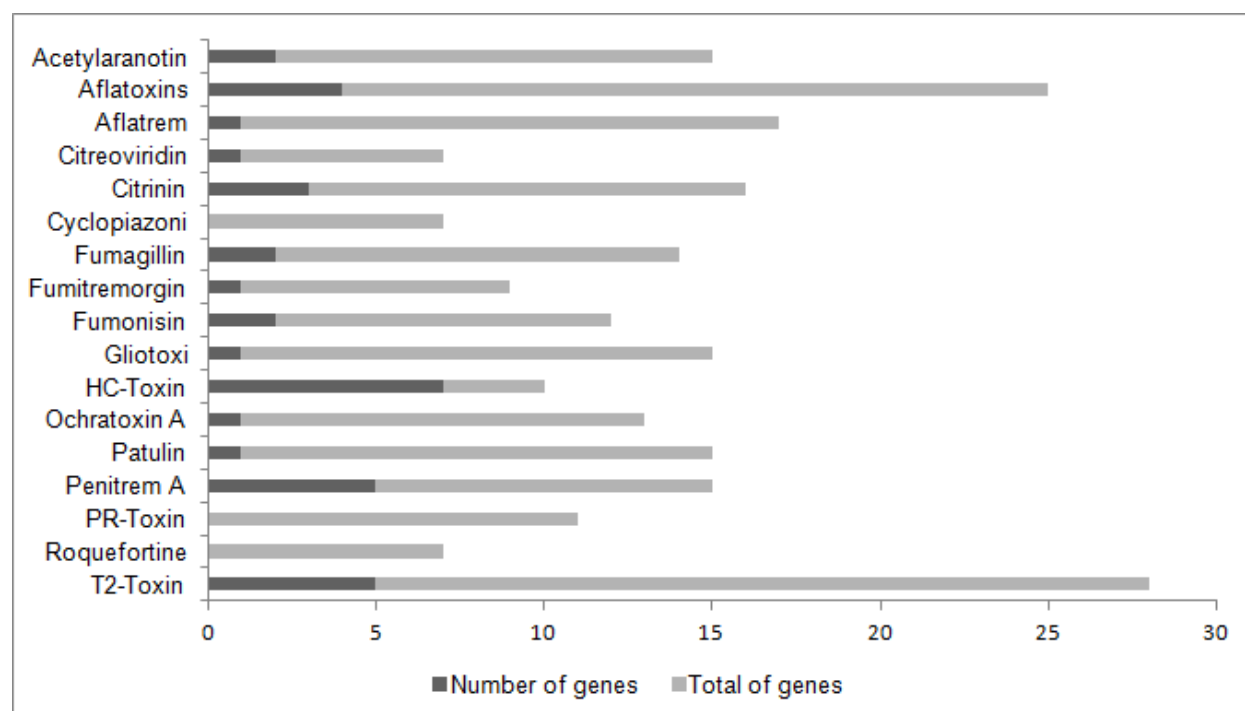


Figure 9: Comparison between the number of genes demonstrating orthology with a known mycotoxin gene in filamentous fungi and the total number of genes in *P. echinulatum* in its gene clusters.

For cyclopiazoni, PR toxin and roquefortine, no cluster gene demonstrated orthology. For acetylaranotin, aflatoxins, aflatrem, citreoviridin, citrinin, fumagillin, fumitremorgin,

fumonisin, gliotoxin, ochratoxin A, patulin, penitrem A and T2-toxin, the amount of orthologous cluster genes is less than half of the total genes present in the cluster. For HC-toxin, 70% of the genes present in the cluster demonstrate orthology, thus being the only mycotoxin studied that shows expression potential. HC-toxin is a cyclic tetrapeptide, an inhibitor of histone deacetylases in many organisms, and has antitumor properties [35, 36].

4. Conclusions

The present study by antiSMASH analysis showed that *P. echinulatum*, at least 2HH and its mutants, such as in other fungi, had carried a set of gene cluster to a high diversity of secondary metabolites. With a total number of 26 gene clusters, with featured to 7 gene clusters, found so much to NRPS (non-ribosomal peptide synthetase) cluster as also to terpene. Although were found various genes relates with mycotoxins, the antiSMASH assay did not predict any mycotoxins in the fungus genome, which corroborates BLAST analysis. It is also worth pointing out that the same also can be told to other putative secondary metabolites.

By observing the metabolic pathways presented by the antiSMASH analysis, it is clear that the related genes are smaller part of the pathway and not part of the core, which indicates that they will hardly be expressed. The same goes for mycotoxin analysis, except for what has already been cited for HC-toxin. The results presented are prediction data, does not mean that *Penicillium echinulatum* actually secretes the secondary metabolites or the mycotoxins previously related, but evidences the orthology between other fungi. In addition, it is interesting to note that the antiSMASH assay did not predict any mycotoxins in the fungus genome, which corroborates BLAST analysis. However, many scaffolds were unrelated to any secondary metabolite region (588 scaffolds for the wild strain and 568 for the mutant strain). This demonstrates the potential for further analysis of these genes, whether they may be related to metabolism or mycotoxins or just residues. Thus, in future is important that this analysis *in silico* would be strengthened by experimental data about gene expression and secretion of secondary metabolites.

References

- [1] J. A. Ferreira, A. Mahboubi, P. R. Lennartsson, M. J. Taherzadeh, Waste biorefineries using filamentous ascomycetes fungi: Present status and future prospects, *Bioresource Technology* 215 (2016) 334–345. doi:10.1016/j.biortech.2016.03.018.
- [2] L. F. Martins, D. Kolling, M. Camassola, A. J. P. Dillon, L. P. Ramos, Comparison of *Penicillium echinulatum* and *Trichoderma reesei* cellulases in relation to their activity against various cellulosic substrates, *Bioresource Technology* 99 (5) (2008) 1417–1424. doi:10.1016/j.biortech.2007.01.060.
- [3] M. Camassola, A. J. Dillon, Production of cellulases and hemicellulases by *Penicillium echinulatum* grown on pretreated sugar cane bagasse and wheat bran in solid-state fermentation, *Journal of Applied Microbiology* 103 (6) (2007) 2196–2204. doi:10.1111/j.1365-2672.2007.03458.x.
- [4] A. J. P. Dillon, M. Camassola, J. A. P. Henriques, M. H. P. Fungaro, A. C. S. Azevedo, T. A. F. Velho, S. E. Laguna, Generation of recombinants strains to cellulases production by protoplast fusion between *Penicillium echinulatum* and *Trichoderma harzianum*, *Enzyme and Microbial Technology* 43 (6) (2008) 403–409. doi:10.1016/j.enzmictec.2008.07.009.

- [5] L. dos Reis, R. C. Fontana, P. da Silva Delabona, D. J. da Silva Lima, M. Camassola, J. G. da Cruz Pradella, A. J. P. Dillon, Increased production of cellulases and xylanases by *Penicillium echinulatum* S1M29 in batch and fed-batch culture, *Bioresource Technology* 146 (2013) 597–603. doi:10.1016/j.biortech.2013.07.124.
- [6] M. R. Rubini, A. J. Dillon, C. M. Kyaw, F. P. Faria, M. J. Poças-Fonseca, I. Silva-Pereira, Cloning, characterization and heterologous expression of the first *Penicillium echinulatum* cellulase gene, *Journal of Applied Microbiology* 108 (4) (2010) 1187–1198. doi:10.1111/j.1365-2672.2009.04528.x.
- [7] G. Liu, L. Zhang, X. Wei, G. Zou, Y. Qin, L. Ma, J. Li, H. Zheng, S. Wang, C. Wang, L. Xun, G. P. Zhao, Z. Zhou, Y. Qu, Genomic and Secretomic Analyses Reveal Unique Features of the Lignocellulolytic Enzyme System of *Penicillium decumbens*, *PLoS ONE* 8 (2) (2013). doi:10.1371/journal.pone.0055185.
- [8] A. J. Dillon, C. Zorgi, M. Camassola, J. A. P. Henriques, Use of 2-deoxyglucose in liquid media for the selection of mutant strains of *Penicillium echinulatum* producing increased cellulase and β -glucosidase activities, *Applied Microbiology and Biotechnology* 70 (6) (2006) 740–746. doi:10.1007/s00253-005-0122-7.
- [9] A. J. Dillon, M. Bettio, F. G. Pozzan, T. Andrighetti, M. Camassola, A new *Penicillium echinulatum* strain with faster cellulase secretion obtained using hydrogen peroxide mutagenesis and screening with 2-deoxyglucose, *Journal of Applied Microbiology* 111 (1) (2011) 48–53. doi:10.1111/j.1365-2672.2011.05026.x.
- [10] D. Zampieri, L. C. Nora, V. Basso, M. Camassola, A. J. Dillon, Validation of reference genes in *Penicillium echinulatum* to enable gene expression study using real-time quantitative RT-PCR, *Current Genetics* 60 (3) (2014) 231–236. doi:10.1007/s00294-014-0421-6.
- [11] M. Camassola, A. J. Dillon, Effect of methylxanthines on production of cellulases by *Penicillium echinulatum*, *Journal of Applied Microbiology* 102 (2) (2007) 478–485. doi:10.1111/j.1365-2672.2006.03098.x.
- [12] A. Fleming, On the Antibacterial Action of Cultures of a *Penicillium*, with Special Reference to their Use in the Isolation of *B. influenzae*, *British journal of experimental pathology* (1929).
- [13] C. M. Visagie, J. Houbraken, J. C. Frisvad, S. B. Hong, C. H. Klaassen, G. Perrone, K. A. Seifert, J. Varga, T. Yaguchi, R. A. Samson, Identification and nomenclature of the genus *Penicillium*, *Studies in Mycology* 78 (1) (2014) 343–371. doi:10.1016/j.simyco.2014.09.001.
- [14] A. L. Demain, A. Fang, The natural functions of secondary metabolites., *Advances in biochemical engineering/biotechnology* 69 (2000) 1–39.
- [15] T. da S. Agostini-Costa, R. F. Vieira, H. R. Bizzo, D. Silveira, M. A. Gimenes, Secondary metabolites (2012).
- [16] K. Blin, M. H. Medema, R. Kottmann, S. Y. Lee, T. Weber, The antiSMASH database, a comprehensive database of microbial secondary metabolite biosynthetic gene clusters, *Nucleic Acids Research* 45 (D1) (2017) 555–559. doi:10.1093/nar/gkw960.
- [17] R. Breitling, A. Cenicerós, A. Jankevics, E. Takano, Metabolomics for Secondary Metabolite Research, *Metabolites* 3 (4) (2013) 1076–1083. doi:10.3390/metabo3041076.
- [18] P. G. Arnison, M. J. Bibb, G. Bierbaum, A. A. Bowers, T. S. Bugni, G. Bulaj, J. A. Camarero, D. J. Campopiano, G. L. Challis, J. Clardy, P. D. Cotter, D. J. Craik, M. Dawson, E. Dittmann, S. Donadio, P. C. Dorrestein, K.-D. Entian, M. A. Fischbach, J. S. Garavelli, U. Göransson, C. W. Gruber, D. H. Haft, T. K. Hemscheidt, C. Hertweck, C. Hill, A. R. Horswill, M. Jaspars, W. L. Kelly, J. P. Klinman, O. P. Kuipers, A. J. Link, W. Liu, M. A. Marahiel, D. A. Mitchell, G. N. Moll, B. S. Moore, R. Müller, S. K. Nair, I. F. Nes, G. E. Norris, B. M. Olivera, H. Onaka, M. L. Patchett, J. Pielar, M. J. T. Reaney, S. Rebuffat, R. P. R. H.-G. Sahl, E. W. Schmidt, M. E. Selsted, K. Severinov, B. Shen, K. Sivonen, L. Smith, T. Tein, R. D. Süssmuth, J. R. Tagg, G.-L. Tang, A. W. Truman, J. C. Vederas, C. T. Walsh, J. D. Walton, S. C. Wenzel, J. M. Willey, W. A. van der Donk, Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature, *Nat Prod Rep* 30 (1) (2014) 108–160. doi:10.1039/c2np20085f.
- [19] P. Wiemann, N. P. Keller, Strategies for mining fungal natural products, *Journal of Industrial Microbiology and Biotechnology* 41 (2) (2014) 301–313. doi:10.1007/s10295-013-1366-3.
- [20] E. Choque, C. Klopp, S. Valiere, J. Raynal, F. Mathieu, Whole-genome sequencing of *Aspergillus*

- tubingensis G131 and overview of its secondary metabolism potential, *BMC Genomics* 19 (2018) 1–16. doi:10.1186/s12864-018-4574-4.
- [21] N. P. Keller, Fungal secondary metabolism: regulation, function and drug discovery, *Nature Reviews Microbiology* 17 (3) (2019) 167–180. doi:10.1038/s41579-018-0121-1.
- [22] M. H. Medema, R. Kottmann, P. Yilmaz, M. Cummings, J. B. Biggins, K. Blin, I. D. Bruijn, Y. H. Chooi, J. Claesen, R. C. Coates, P. Cruz-Morales, S. Duddela, S. Düsterhus, D. J. Edwards, D. P. Fewer, N. Garg, C. Geiger, J. P. Gomez-Escribano, A. Greule, M. Hadjithomas, A. S. Haines, E. J. Helfrich, M. L. Hillwig, K. Ishida, A. C. Jones, C. S. Jones, K. Jungmann, C. Kegler, H. U. Kim, P. Kötter, D. Krug, J. Masschelein, A. V. Melnik, S. M. Mantovani, E. A. Monroe, M. Moore, N. Moss, H. W. Nützmann, G. Pan, A. Pati, D. Petras, F. J. Reen, F. Rosconi, Z. Rui, Z. Tian, N. J. Tobias, Y. Tsunematsu, P. Wiemann, E. Wyckoff, X. Yan, G. Yim, F. Yu, Y. Xie, B. Aigle, A. K. Apel, C. J. Balibar, E. P. Balskus, F. Barona-Gómez, A. Bechthold, H. B. Bode, R. Borriss, S. F. Brady, A. A. Brakhage, P. Caffrey, Y. Q. Cheng, J. Clardy, R. J. Cox, R. D. Mot, S. Donadio, M. S. Donia, W. A. V. D. Donk, P. C. Dorrestein, S. Doyle, A. J. Driessen, M. Ehling-Schulz, K. D. Entian, M. A. Fischbach, L. Gerwick, W. H. Gerwick, H. Gross, B. Gust, C. Hertweck, M. Höfte, S. E. Jensen, J. Ju, L. Katz, L. Kaysser, J. L. Klassen, N. P. Keller, J. Kormanec, O. P. Kuipers, T. Kuzuyama, N. C. Kyrpides, H. J. Kwon, S. Lautru, R. Lavigne, C. Y. Lee, B. Linqun, X. Liu, W. Liu, A. Luzhetskyy, T. Mahmud, Y. Mast, C. Méndez, M. Metsä-Ketelä, J. Micklefield, D. A. Mitchell, B. S. Moore, L. M. Moreira, R. Müller, B. A. Neilan, M. Nett, J. Nielsen, F. O’Gara, H. Oikawa, A. Osbourn, M. S. Osburne, B. Ostash, S. M. Payne, J. L. Pernodet, M. Petricek, J. Piel, O. Ploux, J. M. Raaijmakers, J. A. Salas, E. K. Schmitt, B. Scott, R. F. Seipke, B. Shen, D. H. Sherman, K. Sivonen, M. J. Smanski, M. Sosio, E. Stegmann, R. D. Süßmuth, K. Tahlan, C. M. Thomas, Y. Tang, A. W. Truman, M. Vi-
aud, J. D. Walton, C. T. Walsh, T. Weber, G. P. V. Wezel, B. Wilkinson, J. M. Willey, W. Wohlleben, G. D. Wright, N. Ziemert, C. Zhang, S. B. Zotchev, R. Breitling, E. Takano, F. O. Glöckner, Minimum Information about a Biosynthetic Gene cluster, *Nature Chemical Biology* 11 (9) (2015) 625–631. doi:10.1038/nchembio.1890.
- [23] Y. F. Li, K. J. Tsai, C. J. Harvey, J. J. Li, B. E. Ary, E. E. Berlew, B. L. Boehman, D. M. Findley, A. G. Friant, C. A. Gardner, M. P. Gould, J. H. Ha, B. K. Lilley, E. L. McKinstry, S. Nawal, R. C. Parry, K. W. Rothchild, S. D. Silbert, M. D. Tentilucci, A. M. Thurston, R. B. Wai, Y. Yoon, R. S. Aiyar, M. H. Medema, M. E. Hillenmeyer, L. K. Charkoudian, Comprehensive curation and analysis of fungal biosynthetic gene clusters of published natural products, *Fungal Genetics and Biology* 89 (2016) 18–28. doi:10.1016/j.fgb.2016.01.012.
- [24] K. Blin, T. Wolf, M. G. Chevrete, X. Lu, C. J. Schwalen, S. A. Kautsar, H. G. S. Duran, E. L. D. L. Santos, H. U. Kim, M. Nave, J. S. Dickschat, D. A. Mitchell, E. Shelest, R. Breitling, E. Takano, S. Y. Lee, T. Weber, M. H. Medema, AntiSMASH 4.0 - improvements in chemistry prediction and gene cluster boundary identification, *Nucleic Acids Research* 45 (W1) (2017) 36–41. doi:10.1093/nar/gkx319.
- [25] M. Camassola, A. J. Dillon, Biological pretreatment of sugar cane bagasse for the production of cellulases and xylanases by *Penicillium echinulatum*, *Industrial Crops and Products* 29 (2-3) (2009) 642–647. doi:10.1016/j.indcrop.2008.09.008.
- [26] K. Blin, S. Shaw, K. Steinke, R. Villebro, N. Ziemert, S. Y. Lee, M. H. Medema, T. Weber, antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline, *Nucleic Acids Research* 47 (W1) (2019) 81–87. doi:10.1093/nar/gkz310.
- [27] E. S. Gomes, V. Schuch, E. G. de Macedo Lemos, Biotechnology of polyketides: New breath of life for the novel antibiotic genetic pathways discovery through metagenomics, *Brazilian Journal of Microbiology* 44 (4) (2013) 1007–1034. doi:10.1590/S1517-83822013000400002.
- [28] N. P. Keller, G. Turner, J. W. Bennett, Fungal secondary metabolism - From biochemistry to genomics, *Nature Reviews Microbiology* 3 (12) (2005) 937–947. doi:10.1038/nrmicro1286.
- [29] H. Jayasuriya, K. C. Silverman, D. L. Zink, R. G. Jenkins, M. P. Sánchez, F. Peláez, D. Vilella, R. B. Lingham, S. B. Singh, Clavaric acid: a triterpenoid inhibitor of farnesyl-protein transferase from *Clavariadelphus truncatus*, *Journal of natural products* 61 12 (1998) 1568–70.
- [30] A. Watanabe, *Aspergillus fumigatus* alb1 encodes naphthopyrone synthase when expressed in *As-*

- pergillusoryzae, FEMS Microbiology Letters 192 (2000) 39–44. doi:10.1016/S0378-1097(00)00402-X.
- [31] D. Gerin, L. González-Candelas, A.-R. Ballester, S. Pollastro, R. D. M. Angelini, F. Faretra, Functional Characterization of the alb1 Orthologue Gene in the Ochratoxigenic Fungus *Aspergillus carbonarius* (AC49 strain), Toxins 10 (3) (2018) 120.
 - [32] M. R. Andersen, J. B. Nielsen, A. Klitgaard, L. M. Petersen, M. Zachariasen, T. J. Hansen, L. H. Blicher, C. H. Gotfredsen, T. O. Larsen, K. F. Nielsen, U. H. Mortensen, Accurate prediction of secondary metabolite gene clusters in filamentous fungi, Proceedings of the National Academy of Sciences of the United States of America 110 (1) (jan 2013). doi:10.1073/pnas.1205532110.
 - [33] N. Raffa, N. P. Keller, A call to arms: Mustering secondary metabolites for success and survival of an opportunistic pathogen, PLoS pathogens 15 (4) (2019) e1007606.
 - [34] A. Evidente, A. Cimmino, A. Andolfi, M. Vurro, M. Zonno, C. Cantrell, A. Motta, Phyllostictines A–D, oxazatricycloalkenones produced by *Phyllosticta cirsii*, a potential mycoherbicide for *Cirsium arvense* biocontrol, Tetrahedron 64 (2008) 1612–1619. doi:10.1016/j.tet.2007.12.010.
 - [35] J. D. Walton, HC-toxin, Phytochemistry 67 (14) (2006) 1406–1413. doi:https://doi.org/10.1016/j.phytochem.2006.05.033.
 - [36] W. Zhou, X. Chen, K. He, J. Xiao, X. Duan, R. . . Huang, G. Xiang, Histone deacetylase inhibitor screening identifies HC toxin as the most effective in intrahepatic cholangiocarcinoma cells, Oncology Reports 35 (2016) 2535–2542.

5 CONCLUSÕES E PERSPECTIVAS

Este trabalho visava prever e caracterizar agrupamentos gênicos relacionados ao metabolismo secundário de *Penicillium echinulatum*. Apenas 20% dos *scaffolds* montados foram agrupados em regiões a partir de ortologia com genes já anotados depositados no banco de dados. Foi possível analisar mecanismos biológicos do metabolismo secundário do fungo, realizando inferências quanto ao potencial biotecnológico de aproveitamento de metabólitos a partir do que está no genoma de *P. echinulatum*. A comparação entre as linhagens selvagem e mutante demonstrou que as diferenças entre os genomas são insignificantes no que se refere ao metabolismo secundário dos compostos avaliados.

Ao observar os dados obtidos, é importante ter claro que as análises *in silico* representam uma predição, uma visão geral sobre o potencial metabolismo secundário de *P. echinulatum*. Os dados demonstram a ortologia de genes conhecidos em outros fungos filamentosos. Entretanto, é conhecido que são inúmeros os fatores que influenciam na produção de determinada substância, desde condições ambientais a fatores genéticos, epigenéticos e de regulação, o que torna interessante, no futuro, complementar este estudo com a análise de expressão gênica das linhagens, relacionando fatores ambientais de cultivo e a genética do fungo para se obter resultados quanto à metabolômica. Além disso, também é interessante observar que, nas análises com o *software* antiSMASH, a maioria genes anotados não foram agrupados em regiões, pois não foi encontrada nenhuma enzima relacionada a aquela sequência. Desse modo, também é interessante verificar essa situação, conferindo em que parte do metabolismo do fungo atuam estes genes.

6 REFERÊNCIAS BIBLIOGRÁFICAS

- Agostini-Costa, T Da S.; Vieira, R F.; Bizzo, H R.; *et al.* (2012) Secondary metabolites. **Chromatography and Its Applications**. p. 131–164.
- Andersen, M R.; Nielsen, J B.; Klitgaard, A; *et al.* (2013) Accurate prediction of secondary metabolite gene clusters in filamentous fungi. **Proceedings of the National Academy of Sciences of the United States of America**, v. 110, n. 1, p. E99–E107.
- Arnison, P G.; Bibb, M J.; Bierbaum, G; *et al.* (2014) Ribossomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. **Natural Product Reports**, v. 30, n. 1, p. 108–160.
- Banani, H; Marcet-Houben, M; Ballester, A R; *et al.* (2016) Genome sequencing and secondary metabolism of the postharvest pathogen *Penicillium griseofulvum*. **BMC Genomics**, v. 17, n. 1, p. 1–14.
- Bazioli, J M; Amaral, L S; Fill, T P; *et al.* (2017) Insights into *Penicillium brasilianum* secondary metabolism and its biotechnological potential. **Molecules**, v. 22, n. 6, 23 p.
- Blin, K; Medema, M H; Kazempour, D; *et al.* (2013) antiSMASH 2.0 — a versatile platform for genome mining of secondary metabolite producers. **Nucleic Acids Research**. v. 39, p. W339–W346.
- Blin, K; Medema, M H; Kottmann, R; *et al.* (2017) The antiSMASH database, a comprehensive database of microbial secondary metabolite biosynthetic gene clusters. **Nucleic Acids Research**, v. 45, n. D1, p. D555–D559.
- Blin, K; Shaw, S; Steinke, K; *et al.* (2019) antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. **Nucleic Acids Research**, v. 47, n. W1, p. W81–

W87.

- Bloudoff, K; Schmeing, T. M. (2017) Structural and functional aspects of the nonribosomal peptide synthetase condensation domain superfamily: discovery, dissection and diversity. **Biochimica et Biophysica Acta - Proteins and Proteomics**, v. 1865, n. 11, p. 1587–1604.
- Breitling, R; Cenicerós, A; Jankevics, A; *et al.* (2013) Metabolomics for Secondary Metabolite Research. **Metabolites**, v. 3, n. 4, p. 1076–1083.
- Camassola, M.; Dillon, A. J.P. (2007) Effect of methylxanthines on production of cellulases by *Penicillium echinulatum*. **Journal of Applied Microbiology**, v. 102, n. 2, p. 478–485.
- Carrau, J L *et al* (1981) Produção de enzimas celulolíticas por microrganismos. In: Simpósio Internacional de Engenharia Genética. **Anais**. p. 39. Piracicaba, SP, 1981.
- De Vries, R P; Tsang, A; Grigoriev, I V. (2018) **Fungal Genomics Methods and Protocols**. 2 ed: Methods in Molecular Biology. 1775 p.
- Demain, A. L.; Fang, A. (2000) The natural functions of secondary metabolites. **Advances in Biochemical Engineering/Biotechnology**, v. 69, p. 1–39.
- Dillon, A J P.; Zorgi, C; Camassola, M; *et al.* (2006) Use of 2-deoxyglucose in liquid media for the selection of mutant strains of *Penicillium echinulatum* producing increased cellulase and β -glucosidase activities. **Applied Microbiology and Biotechnology**, v. 70, n. 6, p. 740–746.
- Dillon, A J P; Bettio, M; Pozzan, F G; *et al.* (2011) A new *Penicillium echinulatum* strain with faster cellulase secretion obtained using hydrogen peroxide mutagenesis and screening with 2-deoxyglucose. **Journal of Applied Microbiology**. v.111 p. 48-53.

- Fang, X; Shen, Y; Zhao, J; *et al.* (2010) Status and prospect of lignocellulosic bioethanol production in China. **Bioresource Technology**, v. 101, n. 13, p. 4814–4819.
- Fill, T P; da Silva, B F; Rodrigues-Fo, E. (2010) Biosynthesis of phenylpropanoid amides by an endophytic *Penicillium brasilianum* found in root bark of melia azedarach. **Journal of Microbiology and Biotechnology**, v. 20, n. 3, p. 622–629.
- Fleming, A. (1929) On the Antibacterial Action of Cultures of a *Penicillium*, with Special Reference to their Use in the Isolation of *B. influenzae*. **British Journal of Experimental Pathology**. p. 226-236.
- Ghanbari, M A T; Mohammadkhani, H S; Babaeizad, V; *et al.* (2014) Identification of some secondary metabolites produced by four *Penicillium* species. **Mycologia Irania**, v. 1, n. October 2013, p. 107–113.
- Guzmán-Chávez, F; Zwahlen, R D.; Bovenberg, R A.L.; *et al.* (2018) Engineering of the filamentous fungus *Penicillium chrysogenum* as cell factory for natural products. **Frontiers in Microbiology**, v. 9, n. NOV, p. 1–25.
- Hautbergue, T; Puel, O; Tadrist, S; *et al.* (2017) Evidencing 98 secondary metabolites of *Penicillium verrucosum* using substrate isotopic labeling and high-resolution mass spectrometry. **Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences**, v. 1071, p. 29–43.
- Houbraken, J.; Samson, R. (2011) A. Phylogeny of *Penicillium* and the segregation of Trichocomaceae into three families. **Studies in Mycology**, v. 70, p. 1–51.
- Keller, N P. (2019) Fungal secondary metabolism: regulation, function and drug discovery. **Nature Reviews Microbiology**, v. 17, n. 3, p. 167–180.
- Keller, N P.; Turner, G; Bennett, J W. (2005) Fungal secondary metabolism - From biochemistry to genomics. **Nature Reviews Microbiology**, v. 3, n. 12, p. 937–947.

- Lin, S; Shi, T; Chen, K Y *et al.* (2011) Cyclopenicillone, a unique cyclopentenone from the cultures of *Penicillium decumbens*. **Chemical Communications**, v. 47, n. 37, p. 10413–10415.
- Mamma, D; Hatzinikolaou, D G.; Christakopoulos, P. (2004) Biochemical and catalytic properties of two intracellular β -glucosidases from the fungus *Penicillium decumbens* active on flavonoid glucosides. **Journal of Molecular Catalysis B: Enzymatic**, v. 27, n. 4–6, p. 183–190.
- Medema, M H.; Blin, K; Cimermancic, P; *et al.* (2011) AntiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. **Nucleic Acids Research**, v. 39, n. SUPPL. 2, p. 339–346.
- Medema, M H.; Kottmann, R; Yilmaz, P; *et al.* (2015) Minimum Information about a Biosynthetic Gene cluster. **Nature Chemical Biology**, v. 11, n. 9, p. 625–631.
- Nielsen, J C. (2018) Systems Biology of the Secondary Metabolism in Filamentous Fungi. **Tese de doutoramento**. Department of Biology and Biological Engineering, Chalmers University of Technology. Göteborg, Sweden.
- Nielsen, J C; Grijseels, S; Prigent, S; *et al.* (2017) Global analysis of biosynthetic gene clusters reveals vast potential of secondary metabolite production in *Penicillium* species. **Nature Microbiology**, v. 2. n. 17044, 9 p.
- Peberdy, J F. (1987) ***Penicillium and Acremonium***. v.1. Nottingham, England: University of Nottingham. 305 p.
- Roze, L V.; Chanda, A; Linz, J E. (2011) Compartmentalization and molecular traffic in secondary metabolism: A new understanding of established cellular processes. **Fungal Genetics and Biology**, v. 48, n. 1, p. 35–48.

- Rubini, M R; Dillon, A J P; Kyaw, C M; *et al.* (2009) Cloning, characterization and heterologous expression of the first *Penicillium echinulatum* cellulase gene. **Journal of Applied Microbiology**, v. 108, p. 1187–1198.
- Schneider, W D H; Reis, L; Camassola, M; Dillon, A J P. (2014) Morphogenesis and Production of Enzymes by *Penicillium echinulatum* in Response to Different Carbon Sources. **BioMed Research International**. 10 p.
- Schneider, W D H; Gonçalves, T A; Uchima, C A; *et al.* (2016) *Penicillium echinulatum* secretome analysis reveals the fungi potential for degradation of lignocellulosic biomass. **Biotechnology for Biofuels**. v. 9:66.
- Schürmann, B T.M.; Sallum, W S.T.; Takahashi, J A. (2010) Austin, dehydroaustin and other metabolites from *Penicillium brasilianum*. **Quimica Nova**, v. 33, n. 5, p. 1044–1046.
- Shen, B. (2003) Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. **Current Opinion in Chemical Biology**, v. 7, n. 2, p. 285–295.
- Shen, S; Li, W; Wang, J. (2013) A novel and other bioactive secondary metabolites from a marine fungus *Penicillium oxalicum* 0312F₁. **Natural Product Research**, v. 27, n. 24, p. 2286–2291.
- Soltani, J. (2016) **Secondary Metabolite Diversity of the Genus *Aspergillus*: Recent Advances**. In: Gupta. V K. ***Aspergillus* System Properties and Applications**. Elsevier B.V.; p. 275-292.
- Tannous, J; Keller, N P.; Atoui, A; *et al.* (2018) Secondary metabolism in *Penicillium expansum*: Emphasis on recent advances in patulin research. **Critical Reviews in Food Science and Nutrition**, v. 58, n. 12, p. 2082–2098.
- Visagie, C. M.; Houbraken, J.; Frisvad, J. C.; *et al.* (2014) Identification and nomenclature of the genus *Penicillium*. **Studies in Mycology**, v. 78, n. 1, p. 343–371.

Weber, T; Blin, K; Duddela, S; *et al.* (2015) AntiSMASH 3.0 - A comprehensive resource for the genome mining of biosynthetic gene clusters. **Nucleic Acids Research**, v. 43, n. W1, p. W237–W243.

Zampieri, D. (2015) Expressão gênica e atividades de celulases, β -glicosidases, xilanases e swoleninas de *Penicillium echinulatum* S1M29. **Tese de Doutorado**. Instituto de Biotecnologia. Universidade de Caxias do Sul. Caxias do Sul, Brasil.