

Regulador transcricional LaeA em *Trichoderma* spp: identificação e arquitetura genômica

Fernandes, Joelma dos Santos^{1,4}; Queiroz, Claudia Afras^{1,4}; Sousa, Thiago Fernandes^{2,4}; Hanada, Rogério Eiji¹; Koolen, Hector Henrique Ferreira³; Silva, Gilvan Ferreira⁴

¹Instituto Nacional de Pesquisa da Amazônia, ²Universidade Federal do Amazonas, ³Universidade do Estado do Amazonas, ⁴Embrapa Amazônia Ocidental.
E-mails: jsan.fernandes@gmail.com, gilvan.silva@embrapa.br

Resumo

Ferramentas de bioinformática e sequenciamento de genomas são cada vez mais eficientes e com menor custo e tem contribuído com a identificação de um vasto número de grupos de genes relacionados à síntese metabólica. A identificação de genes relacionados a vias metabólicas por meio de mineração não é suficiente para obtenção de metabólitos secundários, visto que estes clusters gênicos podem estar silenciados ou crípticos. Neste sentido, o uso de fatores de transcrição capazes de ativar genes responsáveis pela biossíntese pode ser uma estratégia para desbloqueio de vias silenciadas. O fator de transcrição LaeA é considerado um regulador importante para desbloquear a expressão de metabólitos crípticos ou silenciados, já constatado em vários fungos. Este trabalho teve como objetivo localizar e realizar análise sintênica de LaeA em 16 isolados de *Trichoderma*. A análise ocorreu com a busca da sequência que codifica o regulador global LaeA e cerca de cinco genes a montante e a jusante ao LaeA nos genomas dos 16 isolados, utilizando o software Genious Prime. A arquitetura foi analisada com base na sintonia via pipeline Clinker & Clustermap.js. Os resultados indicam que dentre alguns dos genes que flanqueiam o *laeA*, estão enzimas relacionadas com a síntese de metabólitos primários e secundários. Grupos de genes específicos inseridos com alta similaridade também foram observados em quatro isolados, que podem ter se integrado na região por meio de eventos de recombinação e inserção de genes. Essas análises evidenciam a alta identidade do gene *laeA* nos dezesseis isolados, demonstrando ser um gene conservado em *Trichoderma* sp.

Palavras-chave: *Trichoderma*, Sintonia, Regulador LaeA.

Introdução

O sequenciamento de nova geração tem permitido a obtenção de genomas completos de microrganismos com menor custo. E a disponibilidade de ferramentas de bioinformática cada vez mais eficientes para estudos genômicos e metabolômicos tem contribuído na identificação de um vasto número de genes

relacionados à síntese metabólica, confirmando o potencial biotecnológico de vários microrganismos como fonte de novas moléculas (Gluck-Thaler et al. 2020; Martínez-Cárdenas et al. 2021; Song et al. 2022).

Os microrganismos são produtores de várias moléculas de interesse biotecnológico, e entre os fungos, o gênero *Trichoderma* destaca-se pela produção de diversos metabólitos de importância farmacêutica e agrônômica (Zingeilinger et al. 2016; Sun et al. 2022). Contudo, um arsenal de metabólitos secundários detectados via análise genômica tem revelado que a capacidade para produção de moléculas é bem maior do que o observado em condições de laboratório, que acabam não sendo sintetizados por falta de conhecimento das condições ambientais de síntese e/ou dos mecanismos de regulação genética que fazem com que em determinadas condições essas vias estejam silenciadas ou sejam críptica, apresentando uma expressão muito baixa do metabólito, não sendo possível detectar as moléculas nas análises químicas (Zingeilinger et al. 2016).

A identificação de compostos bioativos relacionados a vias silenciadas ou crípticas está sendo ampliada com estratégias de mineração de genoma, que auxiliam na busca por vias metabólicas, também chamados de *Clusters* Gênicos Biossintéticos (do inglês *Biosynthetic Gene Clusters* - BGCs) com base em genomas completos (Bok et al. 2006). Porém, conforme mencionado acima a identificação de BGCs no genoma não significa que em condições de laboratório os genes que compõem determinada via estejam sendo expressos, uma vez que, fatores ambientais podem influenciar na ativação desses BGCs. Neste contexto, a indução da expressão de BGCs silenciados ou crípticos pode ser uma alternativa ao desbloqueio dessas vias.

Em *Aspergillus* spp. foi identificado e analisado o gene *laeA* que codifica uma proteína nuclear (LaeA- S-adenosil-L-metionina dependente de metiltransferase) conhecida como regulador transcricional global do metabolismo secundário (Bok et al. 2004). A deleção e superexpressão deste gene levou a descoberta de novas moléculas e tem contribuído para o entendimento de vários fatores relacionados à regulação epigenética do metabolismo em *Aspergillus* spp. e outras espécies de fungos filamentosos (Bok et al. 2004; Karimi-Aghcheh et al. 2013; Jiang et al. 2016; Yu et al. 2019). Em *Penicillium dipodomyis* YJ-11 a superexpressão de LaeA induziu alterações

morfológicas evidentes e variações metabólicas com a ativação de BGCs silenciados (Yu et al. 2019).

Segundo Bok et al. (2006) apenas fungos filamentosos têm possíveis ortólogos de LaeA com regiões conservadas, como a S-adenosil-L-metionina dependente de metiltransferase identificada em isolados de *Trichoderma* neste estudo. Análise da sintenia de genes agrupados a família de genes ortólogos pode contribuir com novos *insights* sobre a evolução e a conservação genômica, e assim gerar informações críticas sobre o contexto da evolução genômica.

Tendo em vista o papel funcional do regulador transcricional LaeA para a expressão de BGCs silenciados ou crípticos, este trabalho teve como objetivo analisar a composição e arquitetura dos genes que flanqueiam o fator de transcrição LaeA em diferentes isolados de *Trichoderma*, com intuito de verificar a conservação gênica da região.

Material e Métodos

A realização do presente estudo ocorreu com a seleção de dezesseis isolados de *Trichoderma* (Tabela 1) obtidos na sua grande maioria do bioma amazônico. Os isolados selecionados tiveram seus genomas sequenciados e analisados. O acesso ao patrimônio genético é o SISGEN: N°AB6B14F.

O sequenciamento dos genomas foi realizado na plataforma Illumina com read length: 2 x 150 (Paired End) com cobertura de 100x. Os genes foram anotados por meio do plugin AUGUSTUS (Stanke e Morgenstern, 2005) do software Geneious Prime versão 2022.2.2 (<https://www.geneious.com>) usando *Fusarium graminearum* como referência e a predição funcional dos genes foi realizada pelas plataformas PFAM e NCBI (<https://www.ncbi.nlm.nih.gov/genome>). Após a anotação dos genomas foi realizada a localização das sequências homólogas ao laeA e cerca de cinco genes a montante e a jusante com maior identidade foram selecionados e analisados.

A análise de sintenia da região a montante e a jusante aos genes homólogos ao laeA nos dezesseis isolados foi realizada por meio do pipeline clinker & clustermap.js (Gilchrist et al. 2021). Para isso, arquivos genbank das sequências homólogas ao laeA foram obtidos via fungiSMASH (Blin et al. 2019).

Tabela 1: Acesso de culturas de *Trichoderma* spp. de diferentes Coleções Microbiológicas Amazônicas.

Isolado	Código	Hospedeiro/substrato	Local	Coleção Microbiológica
T2475	INPA2475	<i>Scleronema micranthum</i>	Amazonas	Coleção de Microrganismos de Interesse Agrossilvicultural CMIAINPA
TM 8	CPAA-TM8	Sedimento	Rio Madeira/AM	CPAA-Embrapa Amazônia Ocidental
TM 12	MMSRG-12	<i>Mauritia flexuosa</i> (Fruto)	Amazonas	Central de coleções Microbiológicas - UEA
TM 14	CPAA-TM14	Sedimento	Rio Madeira/AM	CPAA-Embrapa Amazônia Ocidental
TM 15	MMSRG-15	<i>Mauritia flexuosa</i> (Fruto)	Amazonas	Central de coleções Microbiológicas - UEA
TM 26	CPAA-TM26	Sedimento	Rio Purus/AM	CPAA-Embrapa Amazônia Ocidental
TM 38A	MMSRG-38A	<i>Euterpe precatória</i> (Casca externa)	Amazonas	Central de coleções Microbiológica - UEA
TM 42	CPAA-TM42	Sedimento	Rio Purus/AM	CPAA-Embrapa Amazônia Ocidental
TM 44	CPAA-TM44	Sedimento	Rio Purus/AM	CPAA-Embrapa Amazonia Ocidental
TM 45	CPAA-TM45	Sedimento	Rio Purus/AM	CPAA-Embrapa Amazônia Ocidental
TM 58	CPAA-TM58	Sedimento	Rio Juruá/AM	CPAA-Embrapa Amazônia Ocidental
TM 63	CPAA-TM63	Sedimento	Rio Juruá/AM	CPAA-Embrapa Amazônia Ocidental
TM 64	CPAA-TM64	Sedimento	Rio Juruá/AM	CPAA-Embrapa Amazônia Ocidental
TM 67	CPAA-TM67	Sedimento	Rio Juruá/AM	CPAA-Embrapa Amazônia Ocidental
TM 85	MMSRG-85	<i>Botrylloides giganteus</i>	Mar/SP	Central de coleções Microbiológicas - UEA
TM 94	MMSRG-94	Sedimento	Rio Amazonas	Central de coleções Microbiológicas - UEA

A filogenia dos dezesseis isolados foi realizada com base na sequência parcial do fator de alongação 1- α (EF 1 α). Os alinhamentos múltiplos das sequências de nucleotídeos foram gerados utilizando-se a ferramenta CLUSTALW (Thompson et al. 1994), implementado pelo programa MEGA 7 (Kumar et al. 2016). As análises filogenéticas foram obtidas por meio da análise concatenada do locus e os ramos internos foram avaliados com 1000 repetições de bootstrap.

Resultados e Discussão

A análise do gene que codifica um ortólogo do regulador global LaeA e seus genes a montante e a jusante nos dezesseis isolados do gênero *Trichoderma*, revelou que LaeA apresenta identidade acima de 80% na maioria

dos isolados analisados (Figura 1), demonstrando tratar-se de um regulador transcricional conservado neste gênero. Os genes a montante e a jusante ao gene *laeA* apresentaram a mesma organização, com exceção de alguns isolados que apresentaram variações, sendo possível observar aglomerados específicos de genes introduzidos na região por eventos de recombinação ou inserção.

A montante do gene *laeA* foram observados aglomerados intergênicos nos isolados de *Trichoderma* TM 63 e TM 67, se tratando de um grupo de genes que não são observados nos demais (Figura 1). As sequências de aminoácidos destas proteínas apresentaram identidade com proteínas hipotéticas do gênero *Trichoderma*. Outro grupo intergênico a montante do gene *laeA* também foi observado nos isolados de *Trichoderma* TM 26 e TM 45 (Figura 1). Neste grupo são observados quatro genes inseridos na região, sendo que um deles codifica uma amidohidrolase e os outros genes são relacionados a proteínas hipotéticas.

Os demais genes a montante ao *laeA* estão presentes em quase todos os isolados, dentre eles alguns codificam proteínas relacionadas ao metabolismo primário e secundário, e algumas ainda não tem as proteínas analisadas em espécies de *Trichoderma*, porém possuem similaridade com proteínas de outros microrganismos.

Um dos genes, o primeiro a montante ao *laeA*, codifica uma proteína que possui homologia com um transportador de pantotenato Liz1 em *Fusarium sporotrichioides*. Em estudos realizados em *Schizosaccharomyces pombe*, o Liz1 pode ser o único transportador de pantotenato de membrana plasmática funcional (Stolz et al. 2004). Outro gene, o segundo a montante ao *LaeA*, está relacionado a uma proteína com domínio do tipo SnoaL em *Trichoderma simmonsii*. SnoaL pertence a uma família de pequenas ciclases de policetídeos, que catalisam etapas de fechamento de anel na biossíntese de antibióticos produzidos por policetídeos em *Streptomyces* spp. (Sultana et al. 2004; Ferrara et al. 2021). Já Ferrara et al. (2021), observaram que a ciclase SnoaL em *Aspergillus carbonarius* atuando na regulação positiva da biossíntese de ocratoxina A (OTA), que é uma micotoxina bem conhecida com ampla distribuição em alimentos e rações.

Também foi observado um gene, o quarto a montante ao *laeA*, conservado em todos os 16 isolados de *Trichoderma*, e de acordo com análises realizadas no banco de dados do NCBI (Blast) codifica uma proteína chamada

Switch 2 em *Trichoderma asperellum*, ortólogo de RAD26 da superfamília SNF2 em *Fusarium solani*. Trata-se de subunidade catalítica do complexo SWI/SNF que atua para alterar a estrutura da cromatina (Durr et al. 2006; Ceballos e Heyer 2011).

Na região a jusante ao gene *laeA* foram observados dois genes conservados entre os dezesseis isolados. O primeiro a jusante, apresenta homologia com citocromo P450 em *Trichoderma* spp., um importante grupo de enzimas relacionadas ao metabolismo primário e secundário, e metabolismo de xenobióticos (Lah et al. 2011; Chadha et al. 2018). Bansal e Mukherjee (2016), ao analisar o genoma de três espécies de *Trichoderma* (*T. reesei*, *T. atroviride* e *T. virens*) observou que o citocromo P450 está sempre nas proximidades de clusters de genes relacionados ao metabolismo secundário, como NRPS, PKS e aglomerados de ciclase terpênic.

O segundo gene a jusante, compartilhado em todas as espécies analisadas, codifica uma proteína homóloga a proteína de ativação chamada E1-like em *Trichoderma reesei*, que faz parte de uma superfamília central para a conjugação de ubiquitina (Ub), biossíntese de cisteína, tiamina e vários metabólitos secundários em diversos organismos (Burroughs et al. 2009). Segundo Burroughs et al. (2009), as proteínas semelhantes a E1 estão mais próximas das desidrogenases dependentes de NAD(P)/FAD e das metiltransferases dependentes de S-AdoMet, com atividade de fosfotransferência por “invenção” catalítica. Os demais genes a jusantes ao *laeA* ainda não tem sua caracterização bem definida em fungos.

Além da análise de sintenia dos genes que flanqueiam o regulador *LaeA*, as correlações filogenéticas entre os 16 isolados de *Trichoderma* foram avaliadas com base no gene *tef1-α* (Figura 2) com intuito de observar a arquitetura dos genes encontrados na análise sintênica e filogenia.

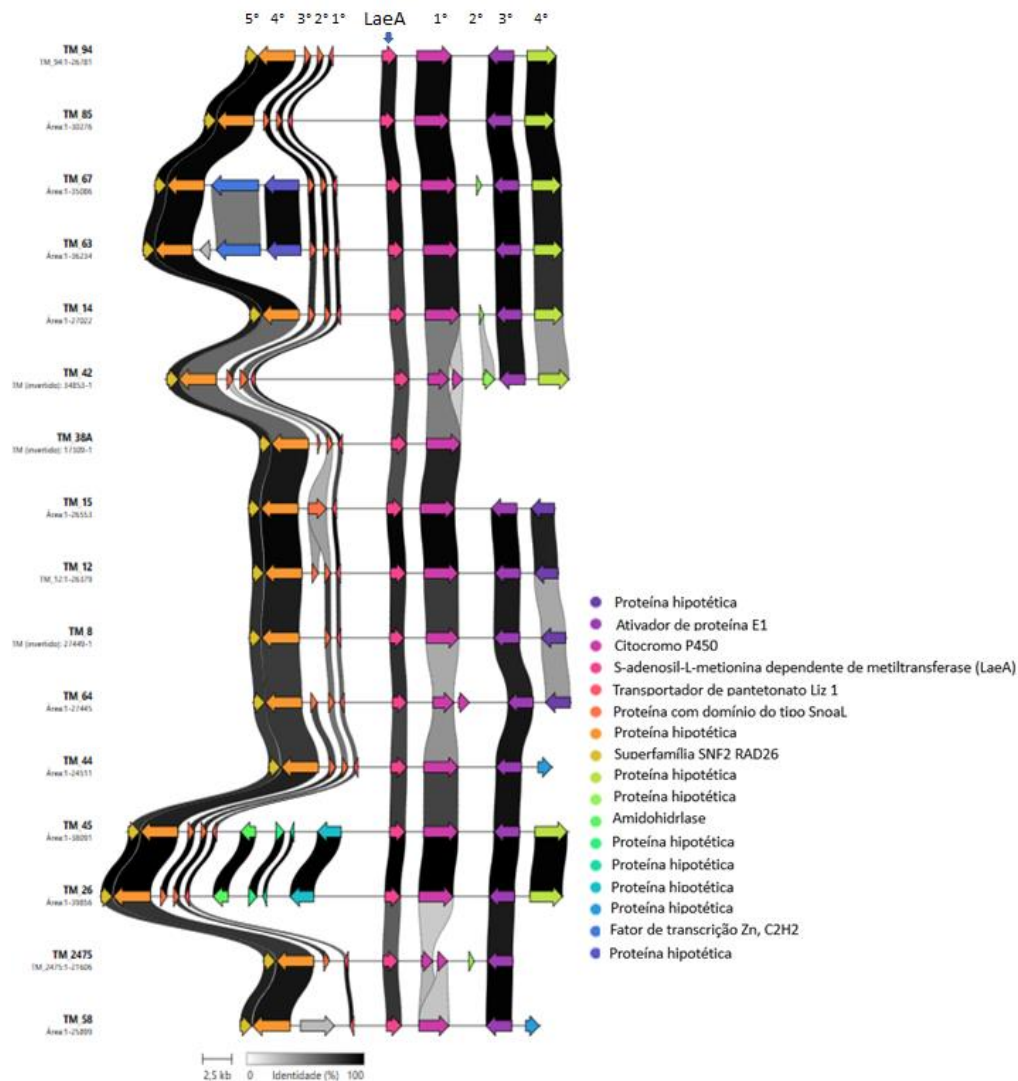


Figura 1. Localização e comparação da sequência do gene ortólogo ao *laeA* (S-adenosil-L-metionina dependente de metiltransferase) e dos genes a montante e a jusante ao regulador transcricional em dezesseis isolados de *Trichoderma*.

Na análise filogenética, o isolado *Trichoderma* TM2475 identificado como *Trichoderma agriamazonicum* sp. nova (Sousa et al. 2023) e o isolado *Trichoderma* TM58 ficaram agrupados com alto suporte de bootstrap (Figura 2) e apresentam arquitetura de genes semelhantes (Figura 1). O mesmo foi observado com os isolados de *Trichoderma* TM 12 e TM 15 que ficaram agrupados e apresentam genes com quase 100% de identidade (Figura 1 e 2).

Os isolados de *Trichoderma* (TM 26 e TM 45) estão agrupados no mesmo clado e próximos também na análise sintênica. Observa-se um grupo de quatro genes inseridos na região e todos estes apresentam 100% de identidade (Figura 1). Já os isolados de *Trichoderma* TM 63 e TM 67 que também apresentam similaridade na arquitetura de genes não estão filogeneticamente próximos.

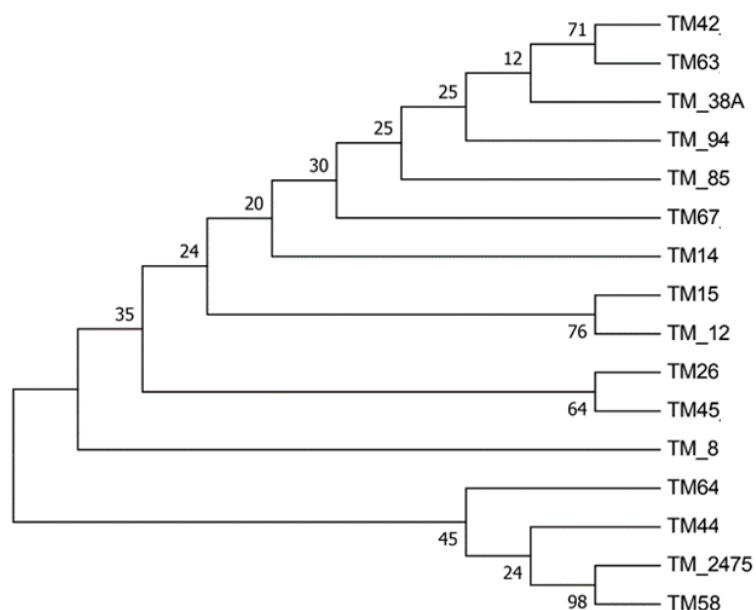


Figura 2. Árvore filogenética baseada em Neighbor joining usando o *barcode tef1- α* de dezesseis isolados de *Trichoderma* amazônicos.

Conclusão

A análise da arquitetura da região onde está localizado o *laeA* juntamente com a filogenia, evidenciam que a região a jusante e a montante são conservadas em *Trichoderma* spp., embora em alguns isolados foram observadas inserção gênica e inserção de DNA intergênico a montante ao *laeA*. Além disso, em todos os isolados a montante foi identificado o gene Switch 2, subunidade catalítica do complexo SWI/SNF que atua para alterar a estrutura da cromatina, a jusante foi localizada o gene P450 que codifica grupos de enzimas relacionadas ao metabolismo primário e secundário, e o E1-like gene que codifica a proteína de ativação relacionadas ao metabolismo secundário.

Agradecimentos

Os autores agradecem ao Conselho Nacional de Aperfeiçoamento de Pessoal de Nível Superior CAPES-Procad AmazonMicro e CAPES-Amazônia Legal pelo apoio financeiro e concessão de Bolsas de estudo. À Fundação de Amparo à Pesquisa do Estado do Amazonas por meio do programa Biodiversa (Edital N°007/2021).

Referências

- Bansal, R.; Mukherjee, P.K. 2016. Identification of novel gene clusters for secondary metabolism in *Trichoderma* genomes. *Microbiology* 85(2): 185-190.
- Blin, K.; Andreu, V.P.; Santos, E.L.C.L.; Carratore, F.; Lee, S.Y.; Medema, M.H.; Weber, T. 2019. The antiSMASH database version 2: A comprehensive resource on secondary metabolite biosynthetic gene clusters. *Nucleic Acids Research* 47(1): 625-630.
- Bok, J.W.; Keller, N.P. 2004. LaeA, a regulator of secondary metabolism in *Aspergillus* spp. *Eukaryotic Cell* 3(2): 527–535.
- Bok, J.W.; Hoffmeister, D.; Maggio-Hall, L.A.; Murillo, R.; Glasner, J.D.; Keller, N.P. 2006. Genomic mining for *Aspergillus* natural products. *Chemistry & Biology* 13(1): 31-37.
- Burroughs, A.M.; Iyer, L.M.; Aravind, L. 2009. Natural history of the E1-like superfamily: implication for adenylation, sulfur transfer and ubiquitin conjugation. *Proteins* 75(4): 895–910.
- Ceballos, S.J.; Heyer, W. 2011. Functions of the Snf2/Swi2 family Rad54 motor protein in homologous recombination. *Biochimica et Biophysica Acta* 1809(9): 509-523.
- Chadha, S.; Mehetre, S.T.; Bansal, R.; Kuo, A.; Aerts, A.; Grigoriev, I.V.; Druzhinina, I.S.; Mukherjee, P.K. 2018. Genome-wide analysis of cytochrome P450s of *Trichoderma* spp.: annotation and evolutionary relationships. *Fungal Biology and Biotechnology* 5(12): 1-15.
- Durr, H.; Flaus, A.; Owen-Hughes, T.; Hopfner, K. 2006. Snf2 family ATPases and DExx box helicases: differences and unifying concepts from high-resolution crystal structures. *Nucleic Acids Research* 34(15): 4160–4167.
- Ferrara, M.; Gallo, A.; Cervini, C.; Gambacorta, L.; Solfrizzo, M.; Padeiro, S.E.; Perrone, G. 2021. Evidence of the involvement of a cyclase gene in the biosynthesis of Ochratoxin A in *Aspergillus carbonarius*. *Toxins* 13(12): 892-904.
- Gilchrist, C.L.; Chooi, Y.H. 2021. Clinker & clustermap. js: Automatic generation of gene cluster comparison figures. *Bioinformatics* 37(16): 2473-2475.
- Gluck-Thaler, E.; Haridas, S.; Binder, M.; Grigoriev, I.V.; Crous, P.W.; Spatafora, J.W.; Bushley, K.; Slot, J.C. 2020. The Architecture of Metabolism Maximizes Biosynthetic Diversity in the Largest Class of Fungi. *Molecular Biology and Evolution* 37(10): 2838-2856.
- Karimi-Aghcheh, R.; Bok, J.W.; Phatale, P.A.; Smith, K.M.; Baker, S.E.; Lichius, A.; et al. 2013. Functional analyses of *Trichoderma reesei* LAE1 reveal conserved and contrasting roles of this regulator. *G3 (Bethesda, Md.)* 3(2): 369-378.

Kumar, S.; Stecher, G.; Tamura, K. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33(7): 1870-1874.

Lah, L.; Podobnik, B.; Novak, M.; Korošec, B.; Berna, S.; Vogelsang, M.; Kraševac, N.; Zupanec, N.; Stojan, J.; Bohlmann, J.; Komel, R. 2011. The versatility of the fungal cytochrome P450 monooxygenase system is instrumental in xenobiotic detoxification. *Molecular Microbiology* 81(5): 1374-1389.

Martínez-Cárdenas, A.; Cruz-Zamora, Y.; Fajardo-Hernández, C.A.; Villanueva-Silva, R.; Cruz-García, F.; Raja, H.A.; Figueroa, M. 2021. Genome mining and molecular networking-based metabolomics of the marine facultative *Aspergillus* sp. MEXU 27854. *Molecules* 26(17): 5362-5374.

Song, H.; Kim, K.; Park, S.; Lee, G.; Choi, J.; Jeon, J.; Cheong, K.; Choi, G.; Hur, J.; Lee, Y.A. 2022. A comparative genomic analysis of lichen-forming fungi reveals new insights into fungal lifestyles. *Scientific Reports* 12(1): 10724-10738.

Sousa, T.F.; Reça, B.N.P.V.; Castro, G.S.; Silva, I.J.S.; Caniato, F.F.; Júnior, M.B.A.; et al. 2023. *Trichoderma agriamazonicum* sp. nov. (Hypocreaceae), a new ally in the control of phytopathogen. *Microbiological Research* 275: 127469-127481.

Stanke, M.; Morgenstern, B. 2005. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Research* 33: W465-W467.

Stolz, J.; Caspari, T.; Carr, A.M.; Sauer, N. 2004. Cell division defects of *Schizosaccharomyces pombe* liz 1 mutants are caused by defects in Pantothenate uptake. *Eukaryotic Cell* 3(2): 406-412.

Sultana, A.; Kallio, P.; Jansson, A.; Wang, J.; Niemi, J.; Mantsala, P.; Schneid, G. 2004. Structure of the polyketide cyclase SnoaL reveals a novel mechanism for enzymatic aldol condensation. *The Embo Journal* 23(9): 1911-1921.

Sun, X.; Cai, Y.; Yuan, Y.; Bian, G.K.; Ye, Z.; Deng, Z.; Liu, T. 2019. Genome mining in *Trichoderma viride* J1-030: discovery and identification of novel sesquiterpene synthase and its products. *Beilstein Journal of Organic Chemistry* 15(1): 2052-2058.

Thompson, J.D.; Higgins, D.G.; Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22(22): 4673-4680.

Zingeilinger, S.; Gruberr, S.; Bansal, R.; Mukherjee, P.K. 2016. Secondary metabolism in *Trichoderma* – Chemistry meets genomics. *Fungal Biology Review* 30(2): 74-90.