



LMQC COREDB: PLATAFORMA BIOQUIMIOINFORMÁTICA PARA CURADORIA E TRIAGEM VIRTUAL DE METABÓLITOS DO GÊNERO *Trichoderma* E SEUS ANÁLOGOS ESTRUTURAIS

Silva M. A. L.¹; Bitencourt J. K. P. S.¹; Nascimento T. P.¹; Santos C. B. R.¹

¹Universidade Federal do Amapá, Macapá, AP, Brasil

marcoaleal@unifap.br

Palavras-Chave: similaridade molecular, espaço químico, bioprospecção computacional

Introdução

A bioprospecção computacional de metabólitos secundários de fungos do gênero *Trichoderma* representa uma estratégia promissora para a descoberta de candidatos bioativos com atividades antimelanogênica e fotoprotetora. Contudo, a ausência de plataformas digitais integradas que centralizem dados estruturais, propriedades físico-químicas e resultados de triagem virtual desse gênero constitui uma lacuna relevante na literatura, dificultando o aproveitamento sistemático do potencial biotecnológico dessas espécies.

Abordagens bioquimioinformáticas integradas, que combinam curadoria de dados químicos, triagem virtual baseada em ligante LBVS (*Ligand-Based Virtual Screening*) e predição de propriedades farmacocinéticas, têm demonstrado eficiência na identificação de candidatos bioativos a partir de scaffolds naturais². No contexto do LBVS, o índice de Tanimoto sobre impressões digitais moleculares é amplamente empregado na expansão do espaço químico a partir de compostos de referência¹. Entretanto, a aplicação sistemática dessas ferramentas a metabólitos de *Trichoderma* permanece incipiente, tanto pela dispersão dos dados em múltiplas bases públicas quanto pela inexistência de pipelines automatizados voltados especificamente para esse gênero fúngico.

Nesse contexto, o Laboratório de Modelagem e Química Computacional (LMQC/UNIFAP), em parceria com o Escritório Modelo do Curso de Ciência da Computação da Universidade Federal do Amapá, desenvolveu a LMQC CoreDB Chemical Library, uma plataforma bioquimioinformática registrada junto ao Instituto Nacional da Propriedade Industrial (INPI)⁴, que integra curadoria de estruturas químicas, expansão do espaço químico por LBVS e interface de consulta interativa, com foco nos metabólitos do gênero *Trichoderma* e seus análogos estruturais. O objetivo deste trabalho é apresentar a arquitetura da plataforma, o pipeline computacional e os resultados preliminares de triagem, demonstrando sua aplicabilidade na identificação de candidatos bioativos para aplicações cosméticas e farmacêuticas.

Material e Métodos

A LMQC CoreDB⁴ foi concebida com arquitetura em três camadas integradas. A primeira camada compreende o banco de dados molecular, construído a partir de coleta automatizada de metabólitos de *Trichoderma* spp. nas bases PubChem e NPASS via API RESTful, com curadoria de estruturas em formato SMILES e metadados taxonômicos e



biológicos, armazenados em banco de dados relacional PostgreSQL. O desenvolvimento foi realizado em parceria com o Escritório Modelo do Curso de Ciência da Computação da UNIFAP, responsável pela implementação da infraestrutura de software.

A segunda camada compreende o pipeline LBVS, implementado em Python. Quatro metabólitos isolados de *Trichoderma harzianum* (Th1–Th4)³ foram empregados como moléculas-consulta. A similaridade molecular foi calculada pelo índice de Tanimoto ($\geq 0,60$) sobre impressões digitais 2D do PubChem¹, via PubChemPy. O conjunto resultante passou por filtragem pelas regras de Lipinski e critérios ADMET preditivos via SwissADME.

A terceira camada corresponde à interface web interativa, desenvolvida em Flask/Python pelo Escritório Modelo do Curso de Ciência da Computação, com funcionalidades de busca estrutural, exportação em SDF e SMILES, e visualização bidimensional das estruturas. O código-fonte está disponível em repositório público de acesso aberto

Resultados e Discussão

O banco de dados da LMQC CoreDB Chemical Library⁴ contém 312 estruturas químicas curadas de metabólitos de *Trichoderma* spp., com cobertura de 47 espécies distintas e metadados incluindo atividade biológica reportada, fonte de isolamento e propriedades físico-químicas calculadas. A integração com PubChem via API RESTful permitiu a recuperação automatizada de identificadores, SMILES canônicos e descritores moleculares para todas as estruturas catalogadas.

A execução do pipeline LBVS com Th1–Th4 como moléculas-consulta resultou em 18.421 compostos análogos recuperados do PubChem com Tanimoto $\geq 0,60$. Após filtragem pelas regras de Lipinski ($PM \leq 500$ Da, $\log P \leq 5$, doadores de ligação de hidrogênio ≤ 5 , aceptores ≤ 10) e exclusão de compostos com alertas de toxicidade predita via SwissADME, foram retidos 4.387 candidatos com perfil farmacocinético favorável para aplicações dérmicas e cosméticas.

A interface web desenvolvida em parceria com o Escritório Modelo do Curso de Ciência da Computação permite consulta, filtragem e exportação dos resultados em tempo real, com visualização das estruturas 2D e tabela de descritores moleculares. A modularidade do pipeline permite sua extensão a outros gêneros de fungos bioativos sem modificações estruturais no código-base, ampliando o escopo de aplicação em estudos de bioprospecção computacional.

A abordagem bioquimioinformática integrada², composta por banco de dados curado, pipeline de triagem LBVS e interface web, diferencia a LMQC CoreDB de repositórios generalistas como PubChem e ChEMBL por seu foco taxonômico específico no gênero *Trichoderma*, pela rastreabilidade completa do pipeline e pela proteção intelectual formal junto ao INPI⁴. Esta especificidade é fundamental para estudos de relação estrutura-atividade direcionados à inibição de tirosinase e regulação da melanogênese⁵.



Conclusões

A LMQC CoreDB Chemical Library, desenvolvida pelo LMQC/UNIFAP em parceria com o Escritório Modelo do Curso de Ciência da Computação e registrada junto ao INPI (processo nº BR512026005830-9)⁴, constitui uma plataforma bioquimioinformática funcional e de acesso aberto para curadoria e triagem virtual de metabólitos do gênero *Trichoderma* e seus análogos estruturais. O pipeline integrado resultou em 4.387 candidatos bioativos com perfil farmacocinético favorável a partir de 18.421 análogos identificados. A interface web facilita o acesso aos dados por pesquisadores sem expertise em programação, ampliando o potencial de uso da plataforma em estudos de bioprospecção de fungos amazônicos com aplicação farmacêutica e cosmética.

Agradecimentos

Ao LMQC/UNIFAP, ao Escritório Modelo do Curso de Ciência da Computação da UNIFAP e ao Programa Bionorte/UFPA pelo suporte ao desenvolvimento da plataforma.

Referências

BAJUSZ, D.; RÁCZ, A.; HÉBERGER, K. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? **J. Cheminform.**, 7, 20, 2015. DOI: 10.1186/s13321-015-0069-3.

SANCHES, P. H. G. *et al.*. From Computational Screening to Antifungal Agents: Biochemoinformatics Applied to the Discovery of New Drug Candidates against *Candida* spp. **Pharmaceuticals**, 17(11), 1491, 2024. DOI: 10.3390/ph17111491.

GUO, Q. *et al.*. Structures and Biological Activities of Secondary Metabolites from the *Trichoderma* genus (Covering 2018–2022). **J. Agric. Food Chem.**, 71, 13612–13632, 2023. DOI: 10.1021/acs.jafc.3c04540.

SILVA, M. A. L. *et al.*. LMQC CoreDB Chemical Library: plataforma bioquimioinformática para curadoria e triagem virtual de metabólitos de *Trichoderma*. Programa de Computador. **INPI** - Instituto Nacional da Propriedade Industrial. Processo nº BR512026005830-9. Data de depósito: 01 jun. 2026.

FERREIRA, A. M. *et al.*. *Trichoderma asperellum* Extract Isolated from Brazil Nuts: In Vivo and In Silico Studies on Melanogenesis in Zebrafish. **Microorganisms**, 11(4), 1089, 2023. DOI: 10.3390/microorganisms11041089.