

Desenvolvimento de uma Ferramenta Computacional para Clusterização Molecular e Identificação do Máximo Subgrafo Comum

Wenderson Rodrigues Lopes¹; Gabriel Teixeira Guerra²; Christiane Mariotini Moura Vasconcellos³; Juliana Lopes Rangel Fietto²; Raphael De Souza Vasconcellos²

1 - Departamento de Informática; 2 - Departamento de Bioquímica e Biologia Molecular; 3 - Departamento de Medicina e Enfermagem

ODS3 – SAÚDE E BEM-ESTAR

Introdução

O avanço das técnicas de bioinformática tem revolucionado o processo de descoberta e desenvolvimento de novos fármacos, tornando possível analisar grandes volumes de dados e prever interações moleculares de forma mais rápida e econômica. Ferramentas computacionais como o RDKit oferecem recursos essenciais para simulações *in silico*, permitindo a análise de similaridade estrutural, clusterização de compostos e a identificação de padrões químicos relevantes. Nesse contexto, a clusterização molecular e a identificação do Máximo Subgrafo Comum (MCS) destacam-se como estratégias poderosas para revelar características estruturais compartilhadas entre moléculas, auxiliando no desenho racional de fármacos.

Objetivos

Desenvolver um Jupyter Notebook interativo, utilizando a biblioteca RDKit em Python e o Google Colab, para:

- Realizar a clusterização molecular de compostos químicos com base em *fingerprints* e no algoritmo de Butina.
- Identificar o Máximo Subgrafo Comum (MCS) entre moléculas agrupadas.
- Fornecer uma interface clara e de fácil utilização, que permita a análise de similaridade estrutural e a seleção de moléculas promissoras para estudos posteriores.

Metodologia

Com o banco de dados de moléculas da PkiDB, foi aplicado o algoritmo de Butina juntamente com o algoritmo de tanimoto para obter as matrizes de similaridade, gerando os *clusters* das moléculas:

