

MINERAÇÃO DE GENOMAS DE CIANOBACTÉRIAS PARA A IDENTIFICAÇÃO DE AGRUPAMENTOS GÊNICOS ENVOLVIDOS NA BIOSSÍNTESE DE METABÓLITOS SECUNDÁRIOS

João Pedro Brandão Domingues¹

Laura Pavan Ióca,² Alessandra S. Eustáquio³

Camila Manoel Crnkovic¹

¹Faculdade de Ciências Farmacêuticas/Universidade de São Paulo,

²Universidade de Princeton, ³Universidade de Illinois em Chicago

joaopedro.brandao@usp.br

Objetivos

Metabólitos secundários, também conhecidos como produtos naturais, representam uma importante fonte de moléculas que inspiram a descoberta e o desenvolvimento de fármacos. No entanto, a descoberta de produtos naturais é um processo desafiador. Expandir a diversidade de fontes e aperfeiçoar as abordagens de bioprospecção são elementos essenciais para aprimorar a identificação de novas moléculas. Cianobactérias abrangem um filo diverso de microrganismos conhecidos por produzir metabólitos secundários com aplicações farmacêuticas e biotecnológicas. Metabólitos secundários de cianobactérias brasileiras permanecem vastamente inexplorados, representando uma fonte promissora de estruturas químicas. À medida que o sequenciamento de genomas se torna mais acessível e a bioinformática evolui, a abordagem da mineração genômica ganha um papel significativo na pesquisa de produtos naturais. Este projeto tem como objetivo investigar sequências de genomas de cianobactérias por mineração genômica para identificar e priorizar agrupamentos gênicos de biossíntese (BGCs, do inglês *biosynthetic gene clusters*) para a descoberta de produtos naturais.

Materiais e Métodos

Foi realizada uma busca no banco de dados do GenBank a procura de genomas de cianobactérias brasileiras. Entre os resultados, 55 sequências genômicas foram selecionadas e analisadas usando antiSMASH 6.1 (BLIN et al., 2021) para prever e anotar BGCs. Paralelamente, foi desenvolvida uma matriz de priorização de BGCs que codificam para a síntese de policetídeos do tipo I (T1PKS), para a síntese de peptídeo não ribossomais (NRPS) e para a síntese híbrida de policetídeos e peptídeos não ribossomais (NRPS-PKS). A matriz foi testada em genomas de cianobactérias previamente estudados que abrigam BGCs que codificam para diversos metabólitos secundários, e.g. *Moorea producens* JHB 22AUG96-1, conhecida por produzir jamaicamidas e hectoclorinas, duas classes de metabólitos secundários híbridos T1PKS-NRPS. (BOUDREAU et al., 2015; LEAO et al., 2017). A priorização levou em consideração a arquitetura e a composição dos módulos nos BGCs detectados, incluindo a integridade dos genes biossintéticos centrais, a presença de enzimas de adaptação, genes regulatórios e relacionados ao transporte e a semelhança com base em agrupamentos de genes conhecidos. A matriz será aplicada para analisar as 55 sequências genômicas para a priorização de BGCs de produtos naturais.

Resultados

Uma versão preliminar da matriz de priorização de BGCs foi construída utilizando os resultados

do antiSMASH para genomas de *Moorea producens* JHB e *Adonisia turfae* CCMR0081, uma cepa brasileira ainda não explorada para produção de produtos naturais. A matriz de priorização de BGCs para *Moorea producens* JHB foi composta por 23 BGCs elegíveis, sendo que o BGC das jamaicamidas (BGC 37) obteve 81 pontos, sendo a maior pontuação, e o BGC das hectoclorinas (BGC 14) obteve 45 pontos, estando entre as cinco maiores pontuações. O segundo, terceiro e quarto BGCs com as maiores pontuações, apesar de apresentarem estruturas biossintéticas interessantes, ainda não são associados a metabólitos secundários conhecidos.

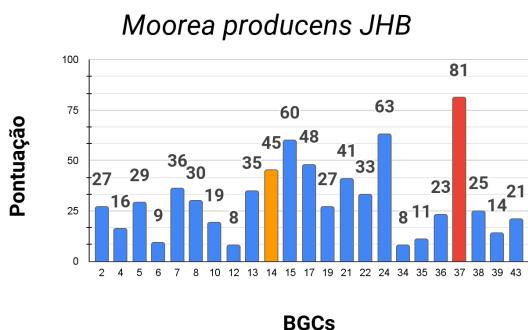


Figura 1: Pontuação total calculada para BGCs da *Moorea producens* JHB utilizando a matriz de priorização de BGCs. A coluna vermelha representa o BGC para a biossíntese de jamaicamidas e a coluna laranja representa o BGC para biossíntese de hectoclorinas.

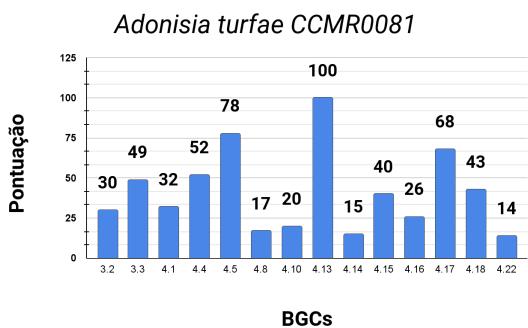


Figura: Pontuação total obtida por cada BGC da *Adonisia turfae* CCMR0081 utilizando a matriz de priorização de BGCs.

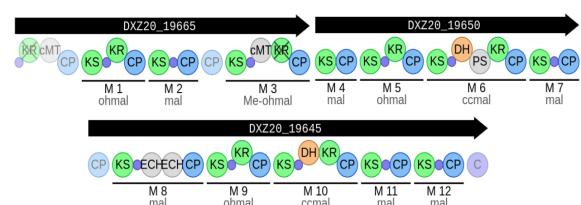


Figura 3: Representação visual do BGCs 4.13 priorizado pela matriz de priorização de BGCs para *Adonisia turfae* CCMR0081.

A matriz de priorização de BGCs para *Adonisia turfae* CCMR0081 foi composta por 14 BGCs. A partir da pontuação da matriz de priorização de BGCs para *Adonisia turfae* CCMR0081 seria possível priorizar três BGCs. O primeiro BGC priorizado da matriz foi o BGC 4,13 com pontuação de 100 pontos, seguido pelo BGC 4,17 com pontuação igual a 68 pontos e pelo BGC 4,5 com pontuação de 78 pontos.

Conclusões

Resultados preliminares da matriz de priorização de BGCs para *Moorea producens* JHB indicam que nossa matriz permitiu a priorização de BGCs, como o BGC das jamaicamidas. Além disso, revelou três BGCs promissores de *Adonisia turfae* CCMR0081, demonstrando que nossa matriz tem o potencial de priorizar BGCs de metabólitos secundários a partir de resultados antiSMASH. A matriz de priorização de BGCs será aplicada a todos os 55 genomas de cianobactérias brasileiras selecionados. Essa matriz de priorização de BGCs pode ser integrada a outras técnicas de prospecção para melhor direcionar os esforços de pesquisa de bioprospecção de produtos naturais, destacando BGCs com maior potencial para produzir metabólitos secundários promissores.

Referências

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

BOUDREAU, P. D. et al. Expanding the Described Metabolome of the Marine Cyanobacterium *Moorea producens* JHB through Orthogonal Natural Products Workflows. *PLOS ONE*, 10(7), p. e0133297, 29 jul. 2015.

LEAO, T. et al. Comparative genomics uncovers the prolific and distinctive metabolic potential of the cyanobacterial genus *Moorea*. *Proceedings of the National Academy of Sciences*, 114(12), p. 3198–3203, 21 mar. 2017.