

RESUMO - OMICS

PIPELINE PARA IDENTIFICAÇÃO DE GENES RELACIONADOS À PROMOÇÃO DE CRESCIMENTO EM PLANTAS: UM ESTUDO DE CASO DA RIZOBACTÉRIA BACILLUS VELEZENSIS

Beatriz Aparecida Leopoldino Da Silva (beatrizsilva@alunos.utfpr.edu.br)

Danilo Sipoli Sanches (danilosanches@professores.utfpr.edu.br)

Alisson Lucas De Souza (alissonsouza@alunos.utfpr.edu.br)

Admilton Gonçalves De Oliveira Junior (admilton@uel.br)

Gustavo Manoel Teixeira (gustavo.teixeira@uel.br)

João Paulo De Oliveira (jp.joaopaulo_oliveira@hotmail.com)

Jacques Uel Br (jacques@uel.br)

A utilização de rizobactérias promotoras de crescimento em plantas (PGPR) tornou-se uma solução sustentável para o controle de fitopatógenos e aumento da produtividade agrícola nas últimas décadas. Cada vez mais, essa área de estudo aumenta e novos resultados surgem. O objetivo desta pesquisa é analisar os genes relacionados à promoção de crescimento em plantas (PGP) descritos na literatura e, com isso, organizar um banco de dados com essas informações, para corroborar em estudos futuros desta área. Neste trabalho, iniciou-se o desenvolvimento de um pipeline com scripts na linguagem Python, utilizando a biblioteca Biopython, para buscar genes relacionados à PGP em genomas de referência encontrados na literatura, e realizar a comparação com os genomas de interesse, neste caso, as cepas *Bacillus velezensis* LABIM,

disponibilizadas pelo Laboratório de Biotecnologia Microbiana (LABIM), da Universidade Estadual de Londrina (UEL). Inicialmente, foi escolhido o genoma de referência *Bacillus velezensis* HNA3, estudado em Zaid et al. (2022), e os genomas de interesse LABIM22, LABIM40 e LABIM44, cujos códigos de acesso GenBank são, respectivamente, CP040881.1, CP045993.1, CP023748.1 e CP079719.1. A partir destes códigos, o primeiro script realiza o download das sequências em formato GenBank, via módulo Entrez. O segundo script, por sua vez, tem como objetivo extrair as principais informações dos genes contidos no arquivo gerado anteriormente, e recuperar informações faltantes, também via Entrez. Já o terceiro script utiliza o algoritmo BLASTp para encontrar os genes relacionados à PGP não identificados na etapa anterior. Por fim, o quarto script compara os genes encontrados pelos scripts anteriores e apresenta os resultados obtidos dessas comparações. Primeiramente, destaca-se que os arquivos Genbank das quatro cepas, disponibilizados no NCBI, não continham o nome de todos os genes, provavelmente devido à problemas de anotação, tornando-se necessário a busca via Entrez no segundo script, e o alinhamento BLASTp no terceiro script, aumentando expressivamente a quantidade de genes identificados. Após a comparação dos genomas, verificou-se que, dos 34 genes apresentados em Zaid et al. (2022), foram identificados os mesmos 29 genes em todas as cepas, indicando, possivelmente, que as cepas LABIM possam ser PGPR. Ademais, comparando todos os genes identificados, obteve-se um percentual aproximado de 84% de genes compartilhados entre as quatro cepas. Assim, nas etapas futuras desta pesquisa, serão acrescentados novos genomas de referência e de interesse para realizar novas comparações, além de explorar sequências do banco de dados Swissprot e incluir estratégias que utilizam mineração de texto e outros alinhamentos BLAST.