



INTELIGÊNCIA ARTIFICIAL E ANÁLISE DE DADOS BIOLÓGICOS NO DESENVOLVIMENTO DE FERRAMENTAS COMPUTACIONAIS APLICADAS À GENÔMICA.

Eduarda Salvador Terhorst (PROBIC-FAPERGS), Scheila de Avila e Silva (Orientador(a))

As sequências promotoras desempenham um papel central na expressão gênica, sendo reconhecidas por fatores sigma específicos da RNA polimerase bacteriana. A identificação precisa desses promotores é essencial para entender a regulação gênica e os mecanismos de transcrição. No entanto, devido às suas características intrínsecas, como sequências curtas e baixa conservação, a predição computacional de promotores enfrenta desafios, incluindo uma taxa significativa de falsos positivos. O BacPP (Bacterial Promoter Prediction) foi o primeiro método computacional a classificar sequências de acordo com o fator sigma que as reconhece. Utiliza redes neurais treinadas com sequências promotoras específicas para cada fator sigma, permitindo a classificação precisa de promotores em *Escherichia coli* e outras enterobactérias. Considerando os avanços e desafios na área, a presente pesquisa tem como objetivo propor o modelo BACPP 2.0, que introduz uma abordagem de classificação em duas etapas para melhorar a precisão na predição de promotores dependentes de fatores sigma. A pesquisa iniciou com a extração de sequências promotoras anotadas com seus respectivos fatores sigma (?24, ?28, ?32, ?38, ?54, ?70) a partir do RegulonDB. Após a coleta, as sequências passaram por um tratamento de padronização e extração de dados relevantes. Posteriormente, os promotores foram definidos pelos seus valores de estabilidade. Essa conversão foi essencial para a fase subsequente, na qual as sequências foram comparadas com regras previamente estabelecidas por inteligência artificial e pesquisas anteriores do grupo. Por fim, realizou-se a identificação e análise de padrões para classificação de promotores por meio de métodos estatísticos. Com o presente estudo será possível ajudar a preencher a lacuna entre a geração e a análise de dados na era pós-genômica. Os possíveis resultados serão aplicados para o aprimoramento do BacPP como classificador de segunda etapa aumentando o grau de assertividade da predição dessa ferramenta.

Palavras-chave: BacPP, Predição de Promotores, Fatores Sigmas

Apoio: UCS, FAPERGS