



## **VALIDAÇÃO DA MATRIZ DE PESOS UTILIZADA COMO REFERÊNCIA EM UMA FERRAMENTA DE PREDIÇÃO DE PROMOTORES**

Julia Gabriele Dani (BIC-UCS), Scheila de Avila e Silva (Orientador(a))

Os promotores são regiões do DNA que iniciam a transcrição de um gene. Nas bactérias, o promotor é reconhecido pela RNA polimerase em conjunto com um fator sigma associado. Esse é uma proteína móvel que possui a função de auxiliar no reconhecimento do promotor. Desse modo, os fatores sigmas determinam quais genes serão expressos no organismo. Nesse sentido, o conhecimento obtido através das ferramentas de predição de sequências promotoras possibilita a compreensão sobre os seus genes e suas interações ecológicas. Um problema nas ferramentas de predição é a frequente classificação incorreta das sequências promotoras (falsos positivos e negativos). Isso ocorre devido a heterogeneidade dos promotores, juntamente com a grande quantidade de dados. Por isso, as ferramentas de predição precisam de ajustes para tornar os resultados cada vez mais precisos. A ferramenta computacional BACPP, desenvolvida na Universidade de Caxias do Sul, realiza a predição de sequências promotoras de bactérias Gram-negativas. Atualmente a ferramenta utiliza como referência uma matriz de pesos ponderados, que possui valores obtidos empiricamente após um estudo com redes neurais. Esses valores são constituídos por valores inteiros e são utilizados igualmente para todos os fatores sigmas existentes. O objetivo deste trabalho foi aprimorar a ferramenta BACPP por meio de alteração dos valores de peso da matriz ponderada utilizada para o cálculo de probabilidade de classificação de uma determinada sequência genômica. Para aumentar a assertividade no reconhecimento e identificação de promotores no BACPP, os valores da matriz de pesos foram substituídos por valores obtidos com uso de Algoritmos Genéticos. Esses valores estão sendo analisados em relação a sua capacidade de discriminar sequências promotoras de não-promotoras por meio de análise estatística a fim de definir pontos de corte. Em relação aos resultados, as análises para verificação de possível otimização da ferramenta BACPP estão em andamento.

Palavras-chave: bioinformática, sequências promotoras, genomas

Apoio: UCS