



CURVATURA E MALEABILIDADE COMO PARÂMETROS DE CLASSIFICAÇÃO DE REGIÕES PROMOTORAS EM BACTÉRIAS GRAM-NEGATIVAS

Ivaine Taís Sauthier Sartor (BIC/UCS), Tahila Andrighetti, Günther J. L. Gerhardt, Sergio Echeverrigaray, Scheila de Avila e Silva (Orientador(a))

O controle da expressão gênica é um processo fundamental nas atividades celulares, realizado através da interação de múltiplos mecanismos de regulação. Os elementos promotores são reconhecidos como as principais regiões regulatórias da transcrição, uma vez que recrutam a maquinaria transcripcional através da ligação de proteínas reguladoras em suas sequências de DNA. A caracterização *in silico* de regiões promotoras apresenta algumas dificuldades, uma vez que estes elementos são curtos e degenerados, gerando uma alta probabilidade de encontrar sequências semelhantes em outras partes do genoma, fora de regiões regulatórias. Portanto, a incorporação de características estruturais, peculiares às regiões promotoras, poderá aumentar a precisão dos métodos de caracterização. Este estudo objetiva prever, reconhecer e caracterizar regiões promotoras reconhecidas pelo fator sigma 28 (σ^{28}) empregando uma abordagem de redes neurais artificiais utilizando como parâmetro de entrada dados de curvatura e maleabilidade da sequência. As regiões promotoras, obtidas do banco de dados *RegulonDB*, foram os exemplos positivo do aprendizado. Os valores para curvatura e maleabilidade destas sequências foram gerados a partir do algoritmo *Bend* (aplicado com diferentes tamanhos de janela deslizante) e suavizados pelo software *Tisean*, em diferentes graus. As sequências intergênicas foram os exemplos negativos. O número de neurônios na camada de entrada foi variável, conforme o tamanho da janela aplicada. As simulações foram realizadas no ambiente R através da utilização do algoritmo *back-propagation* e a metodologia empregada foi *2-fold cross validation*. Foram realizadas 6400 simulações utilizando os valores de curvatura e maleabilidade para o reconhecimento e predição de promotores. As análises destes resultados demonstraram que a curvatura destas sequências não se consolidou como uma característica ótima para a classificação, visto que, a melhor rede (janelamento 2) exibiu os seguintes valores: 55% de exatidão, 56% de sensibilidade e 54% de especificidade. Já para a maleabilidade, os valores observados para a melhor rede (janelamento 3) foram: 79% de exatidão, 78,66% de sensibilidade e 79,33% de especificidade. Estes resultados evidenciam que, para o fator σ^{28} , a maleabilidade é uma característica importante para utilização como parâmetro de classificação de regiões promotoras.

Palavras-chave: redes neurais, DNA, regiões promotoras.

Apoio: UCS.