

Predição de estabilidade em pontos de mutação em proteína através de ferramentas computacionais

**CAMARGO, Alex D.; WERHLI, Adriano V.
MACHADO, Karina dos S.
alexcamargo@furg.br**

**Evento: Encontro de Pós-Graduação
Área do conhecimento: Ciência da Computação**

Palavras-chave: Bioinformática; Aprendizado de máquina; Pontos de mutação

1 INTRODUÇÃO

Em Bioinformática, o estudo sobre ferramentas computacionais, desde a interpretação de saídas geradas até as limitações de cada programa, contribui para a compreensão de inúmeros problemas atualmente enfrentados como, por exemplo, a função das mutações em doenças de tratamento complexo. Recentes trabalhos (MORT *et al*, 2014; GIOLLO *et al*, 2014) apresentam significativos avanços na predição de estabilidade em pontos de mutação através de técnicas de aprendizado de máquina. A presente pesquisa aborda uma revisão de ferramentas biológicas *in-silico*, ou seja, com execução através do computador. Para isso, foram comparados os programas: MAESTRO (*Multi AgEnt STability pRedictiOn*) e mCSM (*mutation Cutoff Scanning Matrix*). A escolha se deu pelo fato de que, conforme a literatura, tais ferramentas são comumente utilizadas em análises desse tipo.

2 REFERENCIAL TEÓRICO

A busca de variantes causadoras de doenças geralmente é focada em substituições de bases únicas que trazem uma mudança direta na sequência primária de uma proteína (MORT *et al*, 2014). As mutações desempenham um papel fundamental no organismo por apresentar vantagens ou desvantagens pelo fato de afetar a estabilidade da proteína (PIRES *et al*, 2014). Para tarefas de análise em pontos de mutação, geralmente são utilizadas técnicas de aprendizado de máquina através de um determinado conjunto de dados (GIOLLO *et al*, 2014).

3 MATERIAIS E MÉTODOS

MAESTRO é uma plataforma multi-agente de predição de estabilidade em pontos de mutação com base em funções estatísticas de pontuação e diferentes abordagens de aprendizado de máquina (LAIMER *et al*, 2015). Outra ferramenta estudada nesta pesquisa é o mCSM que, além do aprendizado de máquina, usa um conceito de assinaturas estruturais baseadas em grafos a fim de estudar e prever o impacto das mutações (PIRES *et al*, 2014). O método de comparação adotado visa a interpretação dos resultados gerados pelas ferramentas supra citadas elencando aspectos importantes de sua execução. Para estudo de caso foi utilizada a estrutura tridimensional da proteína *Beta-galactosidase* (PDB ID: 3THC) considerando a mutação na posição Y64F. Problemas nessa proteína, codificada pelo gene GLB1, causam as doenças humanas: Gangliosidose GM1 e Morquio B.

4 RESULTADOS e DISCUSSÃO

Tanto o MAESTRO quanto o mCSM requerem uma estrutura tridimensional da proteína do tipo selvagem a fim de realizar a predição da (des)estabilidade em um dado ponto de mutação. O resultado gerado pelo MAESTRO é demonstrado na Figura 1, posteriormente, a mesma tarefa foi submetida ao mCSM, tendo a sua saída ilustrada na Figura 2.

Figura 1 – MAESTRO: sumário dos resultados

```
alex@alex-pc ~/MAESTRO_linux_x64 $ ./maestro config.xml 3THC.pdb,A --evalmut="Y64.A{F}"
#structure      seqlength      mutation      score      delta_score      ddg      ddg_confidence
3THC.pdb,A,-    605      wildtype      -10.054000      0.000000      0.000000      1.000000
3THC.pdb,A,-    605      Y64.A{F}      -10.045675      0.008325      0.180983      0.955627
```

Fonte: Os autores

Figura 2 – mCSM: sumário dos resultados

PDB File	Chain	Wild Residue	Residue Position	Mutant Residue	RSA (%)	Predicted $\Delta\Delta G$	Outcome
3THC.pdb	A	Y	64	F	0.2	-1.301	Destabilizing

Fonte: Os autores

É importante destacar que o resultado é dado pelas diferenças de energia do tipo selvagem e o valor $\Delta\Delta G$ (kCal/mol) predito. Para a proteína estudada, os programas apresentaram resultados semelhantes, sendo que para o MAESTRO, valores $\Delta\Delta G$ positivos (coluna 6 da Figura 1) indicam desestabilização, o que ocorre de maneira inversa no mCSM, onde valores negativos (coluna 7 da Figura 2) indicam a desestabilização em um determinado ponto de mutação.

5 CONSIDERAÇÕES FINAIS

A presente pesquisa apresentou uma comparação entre ferramentas para predição de estabilidade em pontos de mutação em proteína. Um item interessante no MESTRO é que, além da predição de estabilidade, a ferramenta tem a possibilidade de busca por novos possíveis pontos de mutação. O mCSM tem o seu funcionamento diretamente na Web, o que facilita bastante a sua execução. Por fim, o uso da ciência da computação se mostra cada vez mais viável diante de problemas multidisciplinares, evidenciando a sua aplicabilidade na biologia.

REFERÊNCIAS

GIOLLO, M.; MARTIN, A. J.; WALSH, I.; FERRARI, C.; TOSATTO, S. C. **NeEMO: a method using residue interaction networks to improve prediction of protein stability upon mutation**. BMC genomics, v. 15, n. Suppl 4, p. S7, 2014.

LAIMER, J.; HOFER, H.; FRITZ, M.; WEGENKITTL, S.; LACKNER, P. **MAESTRO-multi agent stability prediction upon point mutations**. BMC bioinformatics, v. 16, n. 1, p. 116, 2015.

MORT, M.; STERNE-WEILER, T.; LI, B.; BALL, E. V.; COOPER, D. N.; RADIVOJAC, P.; MOONEY, S. D. **MutPred Splice: machine learning-based prediction of exonic variants that disrupt splicing**. Genome Biol, v. 15, n. 1, p. R19, 2014.

PIRES, D. E.; ASCHER, D. B.; BLUNDELL, T. L. **mCSM: predicting the effects of mutations in proteins using graph-based signatures**. Bioinformatics, v. 30, n. 3, p. 335-342, 2014.