

Lista 3

Análise entre Cytoglobina e Biomphalaria glabrata

Vilmar Dorneles Aprato Neto¹

¹Departamento de Informática
UFRGS

UFRGS, 2018

1 Pseudo-código

- Algoritmo de Smith-Watermann

2 Resultados encontrados

- Conclusões

1 Pseudo-código

- Algoritmo de Smith-Watermann

2 Resultados encontrados

- Conclusões

Algoritmo Smith-Watermann

- Sendo A e B , duas sequências de tamanho n e m
- 1. Determine uma matrix de substituição e a penalidade do gap:
 - $s(a, b)$: pontuação de similiaridade de elementos que constituem as sequências;
 - W_k : penalidade de um gap de tamanho k ;
- 2. Construa uma matrix de pontuação H , inicializando sua primeira coluna e sua primeira linha. O tamanho dessa matrix será de $(m + 1) * (n - 1)$:
 - $H_{k0} = H_{0l} = 0$ para $0 \leq k \leq n$ e $0 \leq l \leq m$;
- 3. Preencha os valores da matrix com os seguintes valores:

$$H_{ij} = \max \begin{cases} 0, \\ H_{i-1,j-1} + s(a_i + b_j), \\ H_{i-1,j} + W_k(A), \\ H_{i,j-1} + W_k(B) \end{cases} \quad (1)$$

- 4. Faça o traceback começando no valor mais alto na matrix de score H e terminando na matrix célula que tem score 0, o traceback é baseado no resultado de cada score recursivamente para geração do melhor resultado.

1 Pseudo-código

- Algoritmo de Smith-Watermann

2 Resultados encontrados

- Conclusões

- Os resultados tiveram semelhança de 2,6455% entre as duas sequências de DNA;
- É possível afirmar que é pouco provável que exista uma semelhança entre as duas sequências;
- Pode se afirmar que exista uma pequena semelhança entre as duas sequências;

Imagens do Resultados

```
vilmar@THE-V:~/BioComp/BioComp2018/Trabalho3$ python3 e3-12345.py
Identity = 2.645503 percent
Score = -181
E--KALGVSSNSVESKSWTNLLRAFNKVLKEHSVKKIGLSEEDRKAIVSSWRKLVGRAGGRDNAGTNLVLMFENVPNMR
DRFTKFNAYQPDSALRQDREFGAQVDRITSGLESLVNNVENPGQFQAALERLSTLHKNKTPSVGLQYFGPLQRYIHLYIE
QNLNVASDSVESRAWTNLFAFNEVLKKA
-
-
-
-
-
EKVPGEMEIERERSEELSEAERKAVQAMWARLYANCEDVGVAAILVRFFVNFPSAKQYFSQFKHMEDPLEMERSPQLRKH
ACRVMGALNTVVENLHDPDKVSSVLALVGKAHALKHKVEPVYFKILSGVILEVVAEEFASDFPPETQRAWAKLRGLIYSH
VTAAYKEVGWVQQVPNATTPPATLPSSGP
```

Figure: Análise das duas sequências de DNA



Jonathan Pevsner

Bioinformatics and Functional Genomics.

Wiley Blackwell, 2015.



Wikipedia

https://en.wikipedia.org/wiki/Smith%E2%80%93Waterman_algorithm

Wikipedia, 2018.



T. F. Smith ,M. S. Watermann

Identification of Common Molecular Subsequence

Journal of Molecular Biology,1981.