

LISTA DE EXERCÍCIOS III

Instruções:

- A resolução do exercício deve ser feita **individualmente**. Cópias evidentes entre trabalhos não serão aceitas.
- A entrega deve ser online via Moodle (exclusivamente), somente até a data especificada.
- Para cada uma das tarefas deve-se entregar o com código fonte. O nome do arquivo deve identificar a tarefa, exemplo "e3-1a.py" referente ao item "1a" da tarefa. Arquivos corrompidos serão desconsiderados.
- Além do código fonte deve-se entregar um único arquivo PDF com o nome "e3-il.pdf" apresentando o pseudocódigo do algoritmo desenvolvido e os resultados encontrados.
- Não serão aceitos trabalhos atrasados.
- Data de entrega: 28.08.2018 (terça-feira) até as 13:00 via Moodle (<https://moodle.ufrgs.br/login/index.php>).

NOME: CARTÃO:

Considere as seguintes seqüências:

```
>gi|110319967|emb|CAJ44466.1| haemoglobin type 1 [Biomphalaria glabrata]
MFVLKGSVVQAFVLLSIVCLEITIADDGVRYVNAEWKRPEQSQEGRHSCTARRLEDNSEEVACSTEVKFR
QRAPAEYANKIKKAKDKLRLESQFDDCQQENDRKDRLIQLQANLTDTIHRLVTDSDIQLRSSWATLTA
GADGRNNFNGNVLWLLNTIPNIRERFEKFNAHQSDALKNDNEFVKQVKLVGGGLQSFIDNLENPGQLQ
ATIERLASVHLKMRPTIGLEYRPLQENIAQYVASALGVGADDAAPKAWERLLNAFNEVLNSFANYNIGL
SDTDKVALQSSWSRLTAGADGKRNAVRLVLWMFNNVPMNRERFTKFNAHQSDALKTDAEFLKQVDAII
GGFETLINLNDADLLNRLSLADEHLEKKAISSNYFGPLQKNIHLFIEGTLNFGSDSDEARAWTHLV
GALNKVIKDHAIHNLGLSDIDRDALVSSWNQLTGRAGGSRNAGTNLVLMLENVPMNRDQFSKFNAHQSD
DNLRKDAEFVRQVDLITGGLESVDNVNPIFLQEALVRLADAHLNLKPRVGLEYFGPLQRYIHAYIEKA
LGVSADSAAPRAWTDLLTAFNNVLKDRFTLRIVSDDRRALQSSWSRLQSQAGNKQEAGIKLVTLWFDNV
PNMRDRFTKFNAHSSDEALRANNEFLRQVDVIVGGDLINNVDNSDNFQAAIERLVDVHLHMSPSVGL
YFGPLQQNIRSYIQNALGVAADSAEARSWTNLFTAFNEFLADHTIQKIGLSPTDRKVLDRTWKQLTSGGK
QEAGVKLVLMFQVPMNRDQFSKFDARKSEAEALRNDAEFINQVNNIVGGDLINLNDKPGQLQAALER
LADYHLDHKPRIGLEFFGPLQKYIHLYIESALNVAVGSEESRAWTDLLTALNKVIRDHAIDRLGLSDNDR
EAIDSSWKKLRSGAGGRRNAGIKLVQWMLRTIPNMRDRFNKFDKQSDAALQRDPEFLAQVDRILGGVES
LVNNVDDPVALKAAIDRLADAHLSFDPRVGLDYFGPLQTYIHDIIEALGVGADSDEAKGWTDLFAAFNK
VLKERTVLKIVSDNERAALRSSWDSLKSAAGGTQEAGVRLVLWMLQNVPMNRERFNKFNALQGDDALRAD
AEFVKQVERIAGGLESINNVDNQGLQAAIDRLVNAHLNFRPSVGLYFQPLQENIYKYLEALIVAAD
SNEAKAWTHLLSAFNTVLRHSLEKIGLSVDVRKALESSWKKLTDAAAGGSENAGTNLVFWLLDNVPMNRD
RFTKFNAHQSNAAALQDDDEFNRQVRAITRGIESFVNNVNNPAALQSSIENLVDHLNFPQSIGLSYFGSV
QQYIHLYIAKALGVASNSDEAKSWTNLFAAFNKVLKEHSLEKIGISDSKRALVSSWKKLTAGGRQNFV
DLVLWMFNNVPMNREQFTKFDKQSDADLRDPNFKQVNNIINGVDLVDSVNDPGQLQANLERLTDVH
LHFVPSVGPFEFFGPLQKNIHTFIEQALGVGADSDEPKAWTDLIGAFNKVLNDHAIQHIGLSETDRRALDS
SWKRLTAGENGVQKAGVNLVLWFFNNIPNMRERFTKFDANQADDALRADPEFQKQVNVIVGGLKSFLDSV
NDPIALQANMDRVAEHLSDMPVVGVPYFSALSQNIHRFIEISLGTADSDSQAWTDLLAGFTRVVRNR
AVLRKVSDSDKSAFVSSWNLIRKAASRRNAGVNLVLWLFNNVPMNRNHFTEKFGNQPDAAALRNDQEFLN
QVDRIAGGLESVKVNNPARFLDALERLSSAHLNMKPSIGLEYFGPLQKNIHTYIESALGVAAGSDEAN
AWTDVFGAFNEILKYSSVEKIGLSDSKQALTSSWSSLIAEGKDTAGEKLVLMFNDVPMNRRFTKFDA
```

TQSDEALRNDKEFRNQVNVIVGGLELLINSVNEPGQLQANLEKLVDHLMVPSVGLEYFGPLQKYIHLF
IEKALGVSSNSVESKSWTNLLRAFNVKLVKEHSVKKIGLSEEDRKAIVSSWRKLVGRAGGRDNAGTNLVW
MFENVPMNRDRFTKFNAYQPDSALRQDREFGAQVDRITSGLESLVNNVENPGQFQAALERLSTLHKNKTP
SVGLQYFGPLQRYIHLIEQNLNVASDSVESRAWTNLFASFNEVLKKA

>gi|19549331|ref|NP599030.1| cytoglobin [Homo sapiens]
MEKVPGEIERRERSEELSEAERKAVQAMWARLYANCEDVGVAAILVRFFVNFPSAKQYFSQFKHEDPL
EMERSPQLRKHACRVMGALNTVVENLHDPDKVSSVLALVGKAHALKHKVEPVYFKILSGVILEVVAEEFA
SDFPPETQRAWAKLRGLIYSHVTAAYKEVGWVQVNPATTPPATLPSSGP



Figure 1: Biomphalaria Glabrata (esquerda) e Homo Sapiens (direita).

Implemente um conjunto de rotinas computacionais para realizar as tarefas que seguem:

1. Desenvolva um algoritmo que leia duas seqüências n e m .
2. Implemente e utilize o algoritmo de Smith-Waterman para alinhar estas seqüências.
3. O programa deve imprimir na tela a tabela final de alinhamento, as duas cadeias alinhadas e identidade do alinhamento.
4. Regras:
 - Valor para GAP: -2
 - Valor para Match: +1
 - Valor para Mismatch: -1
 - $S_{i,j} = \text{MAX} \begin{bmatrix} 0 \\ S_{i-1,j-1} + s(a_i, b_j) \\ S_{i,j-1} + \text{gap}(\text{seq1}) \\ S_{i-1,j} + \text{gap}(\text{seq2}) \end{bmatrix}$
5. O que podemos concluir a partir do alinhamento?

⇒ Biomphalaria glabrata é uma espécie de caramujo da família Planorbidae, hospedeiro intermediário do Schistosoma mansoni causador da doença chamada esquistossomose. B.glabrata é uma espécie aquática e habita ambientes de água doce, principalmente em lagoas, canais e valas de irrigação e em pequenos cursos de água.

⇒ Cytoglobin is a globin molecule ubiquitously expressed in all tissues and most notably utilized in marine mammals. It was discovered in 2001 and named cytoglobin in 2002. It is thought to protect against hypoxia. The predicted function of cytoglobin is the transfer of oxygen from arterial blood to the brain.