

## QBQ204 – 2016s1 – Odontologia

### Aula de bioinformática na sala multimídia – 31 de março de 2016

#### 1. Tarefa 1

- a. Nesta tarefa você vai realizar o alinhamento de duas sequências de DNA usando um programa disponível na internet.
- b. Acesse <http://blast.ncbi.nlm.nih.gov/Blast.cgi>. Esta é a página central do programa BLAST, para comparação e busca de sequências.
- c. Clique na opção **Align two (or more) sequences using BLAST (bl2seq)**. Ela consta da seção **Specialized BLAST**.
- d. Em cada caixa de entrada na tela que aparece, cole **as sequências A e B** abaixo. A seguir aperte o botão BLAST.
- e. Cole no relatório o alinhamento obtido (que aparece na seção **alignments**).

#### 2. Tarefa 2

- a. Nesta tarefa você vai comparar uma das sequências da tarefa anterior com todas as sequências disponíveis num banco de sequências do centro NCBI dos Estados Unidos. Note que a sequência de entrada é de DNA mas o banco de consulta é de proteínas. Isto significa que o programa comparador irá traduzir a sequência de entrada nos 6 quadros de leitura para poder achar resultados.
- b. Volte a <http://blast.ncbi.nlm.nih.gov/Blast.cgi>
- c. Clique na opção **blastx**
- d. Na caixa de entrada na tela que aparece, cole **a sequência A** abaixo. A seguir aperte o botão BLAST.
- e. Cole no relatório o alinhamento do primeiro hit dessa busca.

#### 3. Tarefa 3

- a. Esta tarefa é semelhante à anterior, mas a sequência de entrada é de uma proteína.
- b. Volte a <http://blast.ncbi.nlm.nih.gov/Blast.cgi>
- c. Clique na opção **protein blast**
- d. Na caixa de entrada na tela que aparece, cole **a sequência C** abaixo. A seguir aperte o botão BLAST.
- e. Cole no relatório o alinhamento do primeiro hit dessa busca.

4. **Tarefa 4:** comente os resultados obtidos, relacionando-os aos conceitos aprendidos em aula. Inclua detalhes do tipo: a que organismos pertencem essas sequências? que proteínas são codificadas por cada uma das sequências?

5. **Tarefa 5:** use os programas abaixo para rever os conceitos aprendidos em aula:

- a. síntese proteica
- b. replicação de DNA
- c. transcrição de DNA
- d. *não é preciso entregar nada para esta tarefa.* Utilize esta oportunidade para fixar conceitos e aprender mais sobre esses três processos.

**Formato do relatório:**

- a) nome e números dos participantes do grupo
- b) resultados das tarefas de 1 a 4

Envie o relatório até o dia **7 de abril de 2016** para os dois monitores nos endereços [alexandre.redondo@hotmail.com](mailto:alexandre.redondo@hotmail.com) e [alexandre.kuniyoshi@butantan.gov.br](mailto:alexandre.kuniyoshi@butantan.gov.br)

## Dados de entrada

>seqA

```
ATGCATTTTACCATTCAACGCGAAGCCCTGCTGAAACCGCTGCAACTGGTTCGCTGGCGTCGTCGAGCGCC
GACAGACCCTGCCGGTGCCTTTCCAACGTGCTGCTGGTTCGTCGAGAAGCAGCAGTTGTGCTCACCAGGTAC
CGATCTGGAGGTGCAATTGGTTCGGTTCGCGTGGCGCTGGAGGAGAACGCCGAGCCCGGCGAGATCACCCTG
CCGGCGCGCAAGCTGATGGACATCTGCAAGAGCCTGCCGAACGACACGCTGATCGATATCCGTCTGGACG
AGCAGAAGCTGCTGATCAAGGCCGGGCGTAGCCGTTTCAGCCTGTCCACCTTGCCGGCCAGCGACTTTCC
CACCGCCGAGGAAGGTTTGGGCTCGCTGACCTTCTCCCTGGGGCAGAGCAAGCTGCGCCGGCTGATCGAA
CGCACAGTTTTCGCCATGGCCCAGCAGGACGTGCGCTACTACCTGAACGGCATGCTGCTGGAGATGAACG
GCGGGCTGCTGCGCGCGGTGCCACCGACGGGCACCGCCTGGCGCTGTGCTCCATGCAATCCGGCATCGA
GCATGCCGACCGCCATCAGGTCATAGTGCCGCGCAAGGGTATTCTGGAGCTGGCGCGCCTGCTCACCAGC
CAAGACGGCGAAGTCAGCATCGTCCTCGGCCAGTACCATATCCGTGCCACCACCGGCGAATTCACCTTCA
CCTCCAAGCTGGTCGACGGCAAGTTCCCTGACTACGAACGGGTACTCCCGCGCGGTGGCGACAAGAAGGT
ACTCGGCGATCGACAGTTGCTGCGTGAAGCCTTCAGCCGCACGGCGATCCTCTCCAACGAGAAGTACCGC
GGTATCCGCTGCAATTGGCCAGTGGCCTGCTGAAGATCCAGGCCAACAACCCGGAGCAGGAAGAGGCCG
AAGAGGAAGTGGCGGTGACTACAGCGGCGATGCTCTGGAGATCGGCTTCAACGTGAGCTATCTGCTCGA
CGTGCTGGGTGTGATGAGTGCCGAGCAGGTGTGCCTGACCCTGTCCGACTCCAACAGCAGCGCCTTGCTC
CAGGAAGCCGATAATGACGATTCCGCCTATGTCGTCATGCCGATGCGGCTGTAA
```

>seqB

```
ATGCGTTTTTACACTGCAGCGCGAAGCCTTCCCTCAAACCATTGGCACAAGTGGTCAATGTGGTTCGAACGCC
GTCAAACCCTGCCCGTCTCGGCCAACTTGTGGTACAGGTCAAGGATGGACAGGTCTCGTTGACCGGCAC
CGACCTGGAAGTGGAAATGATCTCGCGCACGCTGGTGGAGGACGCGCAGGACGGCGAAACCACGATTCCG
GCGCGCAAGCTGTTTCGACATCCTGCGCGCCCTGCCGACGGCAGCCGCGTACCATTTTCGAGACCGGCG
ACAAGGTCACTGTGCAAGCCGGGCGGAGCCGCTTACCCTGGCGACGTTGCCTTCCAACGATTTTCCGTC
TGTGGACGAAGTCGAAGCGACCGAACCGCTCGTGGTGGCCGAAGCCGGGCTCAAGGAGCTGATCGAGCGC
ACCGCGTTTCGCCATGGCCCAGCAGGACGTGCGCTATTACCTCAATGGCCTGTTGTTTCGACCTGCGCGACG
GCTTGTGCGCTGCGTGGCTACCGACGGCCATCGTCTGGCGCTATGTGAGATGGAGCTGGAGAAAGCCGG
TGGCGCCAAGCGCCAGATCATCGTGCCGCGCAAGGGCGTGACCGAGCTGCAGCGTTTTGCTGGAAGGCGCC
GATCGCGAGGTTCGAGCTGGAAGTGGGCCCGCAGCCATATCCGCGTCAAGCGCGGTGATGTGACCTTTACCT
CGAAGCTGATCGACGGGCGGTTCCCGGACTACGAGGCAGTGATTCCGATCGGCGCCGATCGCGAGGTCAA
GGTGGATCGTGAAGCGCTGCGCGCGTTCGCTACAGCGTGGCGCCATCCTGTCCAACGAGAAGTACCGCGGT
GTCGCTGTGGAAGTGTGCGCGGGTCAGCTGAAGATCAGCGCGCATAACCCCGAGCAGGAGGAGGCGCAGG
AAGAGATCGAGGCCGATAACCAAGTTCGACGATCTGGCCATTGGCTTCAACGTCAACTATTTGCTCGACGC
GCTGTGCGCATTCGCGGATGAGCACGTGGTGTGATCCAGCTGCGTGTGATGCAATTCGTCTGCATTGGTGCCT
GAGGCCAGCAGCGAGAAGTTCGCGTCATGTGGTGTGATGCCGTTGCGTCTCTGA
```

>seqC

```
GGGTGETLGEKWKARLNQMSALEFYSSYKSGITEVCREEARRALKDGVATGGHAVSRGSAKIRWLEERGY
LQPYGKVVLDLGCGRGWSYAAATIRKVQEVRYTKGGPGHEEPMLVQSYGWNIVRLKSGVDVFMMAEPC
DTLLCDIGESSSSPEVEETRTRLRVLSMVGDWLEKRPGAFCIKVLCPYTSTMMETMERLQRRHGGGLVRVP
LCRNSTHEMYWVSGAKSNI IKSVSSTTSQLLLGRMDGPRRPVKYEEDVNLGSGTRAVASCAEAPNMKI IGR
RIERIRNEHAETWFLDENHPYRTWAYHGSYEAPTQGSASSLVNGVVRLLSKPWDVVTGVTGIAMTDTTPY
GQQRVFKEKVDTRVPDPQEGTRQVMNIVSSWLWELGKRKRPRVCTKEEFINKVRSNAALGAI FEEEEKEW
KTAVEAVNDPRFWALVDREHHLRGECHSCVYNMMGKREKKQGEFGKAKGSRAI WYMWLGARFLEFEAL
GFLNEDHWMGRENSGGGVEGLGLQRLGY ILEEMNRPAGGKMYADDTAGWDTRISKFDLENEALITNQMEE
GHRTLALAVIKYTYQNKVVKVLRPAEGGKTVMDI I SRDQDRGSGQVVTYALNTFTNLVVQLIRNMEAEEV
LEMQDLWLLRKPEKVTRWLQSNWDRLKRMVSGDDCVVKPIDDRFAHALRFLNDMGKVRKDTQEWKPS
GWSNWEEVFPFCSHHFNKLYLKDGRSIVVPCRHQDELIGRARVSPGAGWSIRETACLAKSYAQMWQLLYFH
RRDLRLMANAICSAVPVDWVPTGRTTWSIHGKGEWMTTEDMLMVWNRVWIEENDHMEDKTPVTKWTDIPY
L GKREDLWCGSLIGHRPRTTWAENIKDVTVMVRRII GDEEKYMDYLSQVRYLGEEGSTPGVL
```