**Bioinformatics Worksheet #2: BLAST Practice - Assignment**

**Group Number:\_\_\_\_\_3\_\_\_\_\_\_\_\_**

**Names of individuals in group:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Mystery bacterial operon #: Operon 3**

**Table 1: Table of results for 8 largest proteins in operon**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Protein #** | **Length (base pairs)** | **Length (amino acids)** | **E-value** | **Query Cover** | **% Identity** | **Protein name – brief description** |
| 1 | 585 | 194 | 4e-143 | 100 | 100 | [AraC family transcriptional regulator [Xenorhabdus nematophila]](http://www.ncbi.nlm.nih.gov/blast/Blast.cgi#alnHdr_498912027) |
| 2 | 351 | 116 | 8e-77 | 100 | 100 | [transcriptional regulator [Xenorhabdus nematophila]](http://www.ncbi.nlm.nih.gov/blast/Blast.cgi#alnHdr_498912028) |
| 3 | 165 | 54 | 5.6 | 70 | 42 | [hypothetical protein [Enterococcus faecalis]](http://www.ncbi.nlm.nih.gov/blast/Blast.cgi#alnHdr_552936822) |
| 4 | 147 | 48 |  |  |  | No protein found |
| 5 | 147 | 48 |  |  |  | No protein found |
| 6 | 120 | 39 |  |  |  | No protein found |
| 7 | 114 | 37 |  |  |  | No protein found |
| 8 |  |  |  |  |  |  |

1. What organism did this operon come from? What evidence did you use to come to this conclusion?

**Xenorhabdus nematophilus – (**[**http://www.ncbi.nlm.nih.gov/nuccore/AJ012828.1**](http://www.ncbi.nlm.nih.gov/nuccore/AJ012828.1)**)**

**Evidence from the fact that the two best matching proteins were from this organism.**

1. Are the functions of the genes found in this operon related? What cellular system do you think this operon and the genes contained are involved in?

**By visiting the actual site describing the operon itself (**[**http://www.ncbi.nlm.nih.gov/nuccore/AJ012828.1**](http://www.ncbi.nlm.nih.gov/nuccore/AJ012828.1)**) and scrolling through the genes in the operon, they are labeled as “AraC family transcriptional regulator” and “transcriptional regulator”, indicating that this operon is required for regulating transcription. Further investigation indicated that the operon, flhDC is the flagellar master operon of Xenorhabdus nematophilus, required for motility.**

1. Do some research using NCBI (<http://www.ncbi.nlm.nih.gov>) and other internet resources to learn about the system. What role does this system have in this organism? Create a hypothesis about what might happen to the organism if these genes were deactivated.

**These genes code for transcriptional regulators for the flagellar proteins. This would mean that if removed, the bacteria would not be able to activate the genes for their flagella proteins, making this motile bacteria unable to move.**

1. Share what you learned with the other groups and write down what you learned about the other systems.

**Various answers**

**Bonus**: Go to RCSB Protein Data Bank (<http://www.rcsb.org/pdb/home/home.do>) and search for one of the proteins you found by name. You may not be able to find it for your organism, or at all, but here you can see the structure of the protein determined by protein crystallography. First, view the **Summary** tab. Count the number of beta pleatings and alpha helixes (can also be found under the **Sequence tab**). Are there any ligands associated with this protein? Are there any other units for this protein? If so, name them. Be sure to look at the protein in the **3-D view**.

Name of protein you investigated:\_\_ **AraC family transcriptional regulator** \_

Number of Beta-pleated sheets:\_\_\_\_ 23% beta sheet (13 strands; 48 residues)) \_\_\_\_\_\_\_\_\_\_\_\_\_\_

Number of alpha helix sections:\_\_\_ 37% helical (8 helices; 76 residues) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Associated ligands?\_\_Sulfate Ion, Magnesium Ion, 1,2-ETHANEDIOL \_\_\_\_\_

Additional subunits for protein?\_\_\_none\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Provide a rough 3-D sketch of your protein:

