Name:_____

BIO/CMPSC 300
UniProt and String Online Tools
Activity 5
Fall 2019

Please save this file to act5/ in your repository. Due at the end of class.
GitHub link: https://classroom.github.com/a/DuBJW7yi

1. Go to UniProt online tool http://www.uniprot.org/ to choose a protein sample which has domains. What is the ID (i.e., Entry number) that you chose?

2. What is the general function (in a brief summary) of this protein?

3. What kinds of information does the "Names and Taxonomy" tab provide? You may need to spend some time researching this.

4. What organism(s) does this protein come from? Name up to three.

5. How many domains does this protein have according UniProt?

6. What kinds of information does the "PTM and Processing" tab provide? You may need to spend some time researching this.

7. What is the information which is relayed in the "Feature Key" column of the table in the "PTM and Processing" tab?

8. How many domains does this protein have?

9. Name up to three of the domains for this protein.

10. Run a Blast search on one of the domains. What is the result of using Blast

11. Describe what you see in terms of where else this domain turns up across biology.

12. Submit a screen shot of your Blast output page.

13. What types of functions to you find are being fulfilled by the domains?

14. Are the functions of these other domains similar to the definition that you summarized above? Briefly explain.

15. Go to the Strings online tool at https://string-db.org/ and search for your UniProt ID. Change the settings to include a network screenshot of the protein's interaction with other proteins according to both Text mining and Experiments. As two separate graphics, take screen shots of the text mining and then the experiment network for your protein.

16. Are you able to reach a conclusion about how the protein has been mentioned in the text? What contexts did you find?

17. Also at the Strings online tool, spend some time to try-out and test the other forms of analysis. What types of analysis did you choose and what have you /did you learn about your protein? What other interesting artifacts did you notice?

18. Compare and contrast the above networks (the *text mining*, *experimental* and any others that you would like to implement as well). Which nodes do you

see are still the same and which are different? What makes the difference between a node being included and not included in a network.

19. For each network, perform a k-means clustering. What is a k-mean clustering? What do you learn from this analysis?

20. Describe the outcome of the clustering tasks between your Text mining and Experiments networks and attach a screenshots.

21. What evidence, according to the Analysis tab, can you find concerning domain activity? Describe what you observe.