

HOMEWORK -- Please don't try to write on this sheet! Put your answers on another sheet of paper and turn that in. I encourage you to explore the NCBI site in groups, but I expect each of you to find the answers to these questions by yourself, working alone, after you feel comfortable with the site. **This is due Tuesday, Oct. 2.**

Resources: my own handout (www.life.umd.edu/labs/mount/Handout-Internet.pdf)

NCBI tutorial: www.ncbi.nlm.nih.gov/Education/

1. (2 points) Use Entrez to find the protein associated with accession number NP_002007. What is the name of this protein and in what species is this protein found?
2. (Questions 2-6 are related; 1 point each) What **disease** is associated with the *FMR1* gene?
3. What **gene** is associated with the disease known as SPCH1? Provide the human gene name (in the form *ABCI*).
4. What name is given to the protein product of this gene?
5. Give a **refseq** accession number for the **protein**.
6. What is the degree of amino acid **identity** between that human protein (your answer to number 5) and the corresponding mouse protein? (give the number of identical amino acids / length of match for the entire alignment).
7. (1 point) What is the map position for the **gene** encoding the protein in question 1 (NP_002007) in humans (e.g. 10q11 would be a map position)
8. (1 point) **Use tblastn -- protein query and nucleotide database** -- to search the *Homo sapiens* DNA sequence with the protein sequence NP_001025253. What result do you get (describe your top hits and report their E values)?
9. (1 point) Now **use blastn -- nucleotide query and nucleotide database** -- to search the *Homo sapiens* DNA sequence with the nucleotide accession NM_001030082 (the nucleotide sequence from which the protein accession used in 8 was generated). What result do you get?

Questions 10 and 11 are 5 points each.

Please write between 100 and 400 words (no more) and provide between 2 and 5 citations to primary sources.

10. Where does this sequence come from? (Which species? What gene? If it encodes a protein, what do you think that protein does? Why do you think that? What caveats are there?)

```
AGCTTCCCAAAAATTCATTGTTGGGATTTCTGAACACATGTAGAAAAGTTGCAGTGTFTTCCCTCTTGGGCAT
```

11. What fraction of the human genome is "junk DNA?" I am asking you to use the definition that deletion of junk DNA would be selectively neutral. In this case, your answer will be graded not according to your actual numerical answer, but by your ability to find and cite relevant primary literature. You do need to provide a numerical answer (which can be a range), but you must also make a reasonable case for that answer citing relevant primary sources.