Old Dominion University/BIOL294-Genetics/Rinehart-Kim/Module 9/Genome Assignment/SP24

- You are to address the statements/questions that are in red. You do not need to use complete sentences in your replies. You may either answer directly on this worksheet, or just your answers on another word document.
- The following is an excellent reference: Genetics Home Reference: http://ghr.nlm.nih.gov
- **As of now (3/31), all links are working and you should not have to pay to use them. However, remember that you might need to use different browsers. If you absolutely cannot get a link to work, please state that. Otherwise, an unanswered question will be considered as "incorrect".**
- If you use any source other than websites I have listed to answer your questions, you need to cite the source(s) that you used.

1. WHAT CHROMOSOME DID YOU CHOOSE?

Chromosome 13

2 & 3. STATE THE NUMBER OF GENES **AND** BASE PAIRS ON THE CHROMOSOME YOU CHOSE.

800 genes and over 110 million base pairs

4. LIST ONE GENE THAT IS LOCATED ON THIS CHROMOSOME.

BRCA2

5. STATE THE FUNCTION OF THE GENE YOU LISTED IN #4.

Protection of replication forks, determines the risk of ovarian cancer(higher risk if mutated), and provides cellular resistance to DNA damage

6. WHAT IS THE **SECOND** SEQUENCE DESCRIPTION MATCH FOR YOUR QUERY SEQUENCE?

Homo sapiens CFTR gene (cystic fibrosis transmembrane conductance regulator)

7. WHAT DOES THE ENCODED PROTEIN DO IN THE BODY? The protein encodes a member of the **ATP-binding cassette (ABC) transporter superfamily. The encoded protein functions as a chloride channel, making it unique among members of this protein family, and controls ion and water secretion and absorption in epithelial tissues.**

8. FOR WHAT DISEASE IS A MUTATED FORM OF THIS GENE RESPONSIBLE?

Cystic Fibrosis

9. ON WHAT CHROMOSOME IS THE GENE LOCATED?

Chromosome 7

10. Return to the original nucleotide sequence alignment descriptions. CHOOSE A SPECIES (STATE THE SCIENTIFIC NAME) OTHER THAN *HOMO SAPIENS* THAT ALSO HAS A 100% IDENTITY (Per. Ident) FOR THIS SEQUENCE?

Pongo Abelii

11. WHAT IS THE COMMON NAME FOR THIS SPECIES?

Sumatran Orangutan

12. DOES IT SURPRISE YOU THAT THIS SPECIES ALSO HAS A 100% SIMILARITY IN IDENTITY? WHY OR WHY NOT?

No, it doesn't surprise me that the similarity in identity is 100%. Due to the two species having evolved from a common ancestor, it is logical to assume that as true.

13. Return to the original nucleotide sequence alignment **DESCRIPTION**. Find the first match that has less than 100% similarity identity. Click on the description to answer this question.

a. **Sapajus Apella; *Genus: Sapajus***
b. **Tufted Capuchin**
c. **Gaps:1/119**

14. WHAT IS A GAP IN SEQUENCE ALIGNMENTS? (This is something you'll have to search for on the Internet.)

At least one amino acid, there can be more, was was removed from the sequence

FOR EACH, STATE WHAT THE GENE IS (#15-18). (Again, give the description of the gene or gene product, not the nucleotide sequence.)

15. NM_145556

TAR DNA Binding Protein; Disease: Amyotrophic Lateral Sclerosis

16. NM_013444

Ubiquilin 2, mRNA; Disease: Frontotemporal Dementia or Amyotrophic Lateral Sclerosis

17. NM_001010850

FUS RNA Binding Protein; Disease: Amyotrophic Lateral Sclerosis or Frontotemporal Lobar Degeneration

18. KJ174530

Superoxide dismutase 1; Disease: Amyotrophic Lateral Sclerosis or Parkinson's Disease

19. Search Google to answer the following: WHAT DISEASE IS ASSOCIATED WITH MUTATIONS OF THE GENES REFERENCED IN #15-#18? WHAT IS A "COMMON NAME" OF THE DISEASE? (The name of a person; Hint, hint…Baseball season is just beginning…)

Lou Gehrig's Disease

20. BLAST is possible because of the submission of DNA sequences to GenBank. WHAT IS GENBANK? (You can do an Internet search to find this information.)

The Nucleotide Sequence Database

Introduction to Protein BLAST

Assignment Goal: To use the Internet-based site BLAST, Basic Local Alignment Search Tool (NCBI), to 1) translate cDNA and 2) to search for similarities between amino acid sequences.

Assignment: Access the BLAST site at http://blast.ncbi.nlm.nih.gov/Blast.cgi

Click on "blastx". (This is a smaller button next to "Nucleotide BLAST".

21. First, answer this question: WHAT IS cDNA? (Again, you can do an Internet search to find this information.)

DNA that is synthesized/ reverse transcribed from an RNA template strand

Enter the following cDNA sequence on "blastx":
ACATTTGCTTCTGACACAATTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCATCTGACTCC
TGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAG
GCCCTGGGCAG

22. WHAT IS THE SEQUENCE MATCH?

Hemoglobin Beta

Assignment: While still on BLAST, choose "Protein BLAST". Check the box which says "Align two or more sequences".

Copy and paste the following sequence into the "Enter Query Sequence" box:

Person 1/Sequence 1:

MGAPACALALCVAVAIVAGASSESLGTEQRVVGRAAEVPGPEPGQQEQLVFGSGDAVELSCPPP
GGGPMGPTVWVKDGTGLVPSERVLVGPQRLQVLNASHEDSGAYSCRQRLTQRVLCHFSVRVTD
APSSGDDEDGEDEAEDTGVDTGAPYWTRPERMDKKLLAVPAANTVRFRCPAAGNPTPSISWLK
NGREFRGEHRIGGIKLRHQQWSLVMESVVPSDRGNYTCVVENKFGSIRQTYTLDVLERSPHRPI
LQAGLPANQTAVLGSDVEFHCKVYSDAQPHIQWLKHVEVNGSKVGPDGTPYVTVLKTAGANT
TDKELEVLSLHNVTFEDAGEYTCLAGNSIGFSHHSAWLVVLPAEEELVEADEAGSVYAGILSYG
VGFFLFILVVAAVTLCRLRSPPKKGLGSPTVHKISRFPLKRQVSLESNASMSSNTPLVRIARLSSGE
GPTLANVSELELPADPKWELSRARLTLGKPLGEGCFGQVVMAEAIGIDKDRAAKPVTVAVKML
KDDATDKDLSDLVSEMEMMKMIGKHKNIINLLGACTQGGPLYVLVEYAAKGNLREFLRARRPP
GLDYSFDTCKPPEEQLTFKDLVSCAYQVARGMEYLASQKCIHRDLAARNVLVTEDNVMKIADF
GLARDVHNLDYYKKTTNGRLPVKWMAPEALFDRVYTHQSDVWSFGVLLWEIFTLGGSPYPGIP
VEELFKLLKEGHRMDKPANCTHDLYMIMRECWHAAPSQRPTFKQLVEDLDRVLTVTSTDEYLD
LSAPFEQYSPGGQDTPSSSSGDDSVFAHDLLPPAPPSSGGSRT

Copy and paste the following sequence into the "Enter Subject Sequence" box:
Person 2/Sequence 2:

MGAPACALALCVAVAIVAGASSESLGTEQRVVGRAAEVPGPEPGQQEQLVFGSGDAVELSCPPP
GGGPMGPTVWVKDGTGLVPSERVLVGPQRLQVLNASHEDSGAYSCRQRLTQRVLCHFSVRVTD
APSSGDDEDGEDEAEDTGVDTGAPYWTRPERMDKKLLAVPAANTVRFRCPAAGNPTPSISWLK
NGREFRGEHRIGGIKLRHQQWSLVMESVVPSDRGNYTCVVENKFGSIRQTYTLDVLERSPHRPI
LQAGLPANQTAVLGSDVEFHCKVYSDAQPHIQWLKHVEVNGSKVGPDGTPYVTVLKTAGANT
TDKELEVLSLHNVTFEDAGEYTCLAGNSIGFSHHSAWLVVLPAEEELVEADEAGSVYAGILSYR
VGFFLFILVVAAVTLCRLRSPPKKGLGSPTVHKISRFPLKRQVSLESNASMSSNTPLVRIARLSSGE
GPTLANVSELELPADPKWELSRARLTLGKPLGEGCFGQVVMAEAIGIDKDRAAKPVTVAVKML
KDDATDKDLSDLVSEMEMMKMIGKHKNIINLL
GACTQGGPLYVLVEYAAKGNLREFLRARRPPGLDYSFDTCKPPEEQLTFKDLVSCAYQVARGME
YLASQKCIHRDLAARNVLVTEDNVMKIADFGLARDVHNLDYYKKTTNGRLPVKWMAPEALF
DRVYTHQSDVWSFGVLLWEIFTLGGSPYPGIPVEELFKLLKEGHRMDKPANCTHDLYMIMREC
WHAAPSQRPTFKQLVEDLDRVLTVTSTDEYLDLSAPFEQYSPGGQDTPSSSSSGDDSVFAHDLLP
PAPPSSGGSRT

Click on the "BLAST" box.

After you receive results, choose the "Description" box for more information about the alignment.

23. DO YOU SEE ANY DIFFERENCES BETWEEN THE TWO AMINO ACID SEQUENCES? (Look for a space between same amino acid comparison for both sequences.)

Yes there was a difference in the sequences

24. IF YOU SAW DIFFERENCES, WHAT WERE THEY?

```
Query  361  EELVEADEAGSVYAGILSYGVGFFLFILVVAAVTLCRLRSPPKKGLGSPTVHKISRFPLK  420
            EELVEADEAGSVYAGILSY VGFFLFILVVAAVTLCRLRSPPKKGLGSPTVHKISRFPLK

Sbjct  361  EELVEADEAGSVYAGILSYRVGFFLFILVVAAVTLCRLRSPPKKGLGSPTVHKISRFPLK  420
```

25. ARE THERE ANY GAPS IN THE SEQUENCE ALIGNMENT?

Gaps:1/806

Return to the BLAST home page (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Return to the Protein BLAST search to identify the polypeptide which you have been analyzing. (You may use either sequence.). Copy and paste the sequence into the "Enter Query Sequence box. Be sure that the box "Align two or more sequences" is **NOT** checked. Choose "Quick BLASTP" under the Program Selection. Click the "Blast" button at the bottom.

26. WHAT GENE ENCODES FOR THE POLYPEPTIDE YOU WERE ANALYZING?

Click on "Description", then choose "Gene" under "Related Information" to answer the following two questions.

Fibroblast Growth Factor Receptor 3

27. WHAT IS THE FUNCTION OF THIS PROTEIN? Click on “Description” and then “Gene” under “Related Information” to the right to get this information.
The gene encodes a member of the Fibroblast Growth Factor Receptor family

28. WHAT HUMAN DISEASE IS CAUSED BY A MUTATION IN THIS GENE? Again, this information can be gained by clicking on “Description” and then “Gene” under “Related Information” to the right to get this information
Achondroplasia, also known as Dwarfism

29. WHAT IS THE CONNECTION AMONG THE FOLLOWING: NIH, NLM, NCBI, and HHS? (What do the abbreviations stand for? Who oversees what?)

NIH: National Institutes of Health (primary federal agency for conducting and supporting medical and behavioral research for the Nation

NLM:National Library of Medicine (conducts and supports research in methods of recording/storing/retrieving/preserving/communicating health information)

NCBI: National Center for Biotechnology Information (creating automated systems for storing and analyzing knowledge about molecular biology/biochemistry/genetics)

HHS: Health and Human Services (promoting and enhancing the health of citizens of the United States).

30. WHAT WAS ONE POSITIVE THING AND ONE NEGATIVE THING YOU ENCOUNTERED WHILE DOING THIS ASSIGNMENT?

The only negative thing I encountered while doing this assignment was when I ran into trouble trying to do question 23. A positive was that I learned how/where to find more specific gene information.