

CHROMOSOME MAPS

1. What chromosome did you choose?
 - a. Chromosome 10
2. (&3) State the number of genes and base pairs on the chromosome you chose.
 - a. 1400 genes and 130 million base pairs
4. List one gene that is on the chromosome.
 - a. PAHX
5. State the normal function of the gene you listed in #4.
 - a. The PAHX gene codes for the PAHX (phytanic acid hydrolase) protein which metabolizes phytanic acid (a type of lipid).
6. State the possible diseases related to this gene.
 - a. Refsum disease is a recessive condition in which the PAHX protein does not function properly to metabolize phytanic acid.

GENBANK

7. What is GenBank?
 - a. According to the National Institutes of Health, GenBank is a sequence database that contains all publicly available nucleotide sequences and their protein translations for over 300,000 organisms.

INTRODUCTION TO BLAST

8. What is the top sequence description match for your query sentence?
 - a. Homo sapiens CF transmembrane conductance regulator (CFTR), RefSeqGene (LRG_663) on chromosome 7
9. What does the encoded protein do in the body?
 - a. The CF transmembrane conductance regulator (CFTR) protein is a chloride channel that facilitates the movement of chloride anions and water molecules in epithelial tissues (Meng et al.).
10. For what disease is a mutated form of this gene responsible?

- a. A mutated CFTR gene causes cystic fibrosis (Meng et al.).
11. On what chromosome is the gene located?
- a. Chromosome 7
12. What organism is the source of this DNA (look for the first sequence that doesn't have a 100% query cover)?
- a. Olive baboon
13. HOW MANY GAPS OCCUR BETWEEN THE TWO SEQUENCES (THE ONE YOU SUBMITTED AND THE FIRST ONE THAT HAS < 100% QUERY COVER)?
- a. The first submitted sequence has 0 gaps while the sequence with the <100% query cover has 2 gaps.
14. WHAT IS A GAP IN SEQUENCE ALIGNMENTS?
- a. A gap represents an insertion or deletion mutation in a sequence and is denoted with a (-).
15. State what the following gene is: NC_045512.2
- a. Severe acute respiratory syndrome coronavirus, 2 isolate, Wuhan-Hu-1
16. SCROLL DOWN THE LIST BELOW THIS SEQUENCE. THERE ARE MANY SEQUENCES THAT LOOK SIMILAR. CAN YOU DETECT WHAT IS DIFFERENT ABOUT THE OTHER SEQUENCES?
- a. Compared to the first sequence which doesn't identify the organism source, all the other sequences came from humans.

INTRODUCTION TO SWIS-PROT TO STUDY PROTEIN SEQUENCES

17. What is cDNA?
- a. Complementary DNA is DNA synthesized by the enzyme reverse transcriptase from an RNA template.
18. Enter the following sequence and state the sequence match using the same BLAST tool.
- a. Homo sapiens partial HBB gene for hemoglobin beta chain, exon 1, isolate 04593664
 - i. (hemoglobin beta protein)
19. Using the ExPASy translate tool, translate the same sequence and state how many 5' to 3' frames you obtained.
- a. I obtained three 5' to 3' frames.

20. Based upon the length of the polypeptide, which frame(s) is (are) more likely to be the correct one(s)?

- a. It is most likely that the third 5' to 3' frame is the most correct frame because the length of the polypeptide sequence is the longest. When selecting this frame and searching the sequence on the internet, the results came back as the homo sapiens hemoglobin beta protein.

AMINO ACID SEQUENCE COMPARISONS

21. Do you see any absences between the two amino acid sequences?

- a. Comparing the two amino acid sequences, there was one different nucleotide base between base pairs 361-421, and there was a gap in the first sequence at base pair 784 while the second sequence did not have a gap.

22. Why would differences be important and possibly clinically useful?

- a. Slight differences in amino acid sequences could completely change the polypeptide sequence. If an insertion or deletion mutation were to enter, the reading frame shifts which could change the amino acid sequence that is being coded. Therefore, identifying differences in amino acid sequences could identify mutations and their causes and effects.

23. What is the protein that you were analyzing?

- a. fibroblast growth factor receptor 3 isoform 1 precursor [Homo sapiens]

24. Reflect on one thing you already knew before doing this assignment.

- a. When I was a freshman, I did a similar assignment where I was given nucleotide sequences and had to identify them using BLAST. BLAST is an effective tool for identifying protein and nucleotide sequences, no matter how small or big they are.

25. Reflect on one thing you learned from doing this assignment.

- a. I learned that identifying gaps in genetic sequence alignments are important because gaps indicate mutations. Knowing where mutations occur in genetic sequences is important so that researchers can isolate and identify the characteristics and effects of these mutations.

Works Cited:

Meng, X., Clews, J., Kargas, V., Wang, X., & Ford, R. C. (2017). The cystic fibrosis transmembrane conductance regulator (CFTR) and its stability. *Cellular and molecular life sciences : CMLS*, 74(1), 23–38. <https://doi.org/10.1007/s00018-016-2386-8>