

1. What chromosome did you choose?

21

2/3. State the number of genes and base pairs on the chromosome you chose.

This chromosome contains over 400 genes, and 40 million base pairs

4. List one disease (due to a mutated gene) associated with this chromosome.

SOD1- Amyotrophic Lateral Sclerosis

5. State the normal function of the gene (not the disease) you listed in #4.

The normal function is that it removes dangerous superoxide radicals by converting them into non-harmful substances.

6. What is the top sequence description match for your query sequence?

Homo sapiens CFTR promoter region (LOC111674463) on chromosome 7

7. What is the function of this sequence?

This sequences function is to participate in long range interactions with enhancers locating within a topologically associating domain which extends upstream and downstream of the locus. These interactions show complex regulation of expression in different types of cells.

8. What gene (Protein) is the sequence associated with?

CFTR gene-cystic fibrosis transmembrane conductance regulator.

9. What does the encoded protein do in the body?

Plays a pivotal role in chloride transport and overall tissue homeostasis.

Sala, A. et al. Role of Protein Kinase A-Mediated Phosphorylation in CFTR Channel Activity Regulation. *Frontiers in Physiology* **12**. (2021)

10. For what disease is a mutated form of this gene responsible?

Cystic fibrosis.

11. On what chromosome is the gene located?

Chromosome 7

12. What species (stated the scientific name) other than Homo sapiens also has a 100% similarity for this sequence?

Pan troglodytes

13. Does it surprise you that this species also has 100% similarity in identity? Why or why not?

No it does not surprise me because humans share a majority of our DNA with chimpanzees, and Pan troglodytes are a type of chimpanzees.

14. NM_145556

Mus musculus TAR DNA binding protein, transcript variant 1, mRNA

15. NM_013444

Homo sapiens ubiquilin 2, mRNA

16. NM_001010850

Homo sapiens fusion (involved in t(12;16) in malignant liposarcoma) (FUS), transcript variant 2, mRNA

17. KJ174530

Homo sapiens superoxide dismutase-1 gene, exon 1 and partial cds

18. What disease is associated with mutations of the genes referenced in 14-17? What is a common name of the disease?

Amyotrophic Lateral Sclerosis

19. What is GenBank?

It is a public database that shows nucleotide sequences and the protein translations of these sequences.

20. What is the hierarchy among the National institutes of health, the national library of medicine, and the national center for biotechnology information? Where do GenBank and Blast belong?

The hierarchy goes biggest from top to bottom

National institutes of health

National library of medicine

national center of biotechnology information

Gen bank and Blast both belong below the national center of biotechnology information, specifically gen bank is bigger than blast because blast is a subsection of GenBank.

21. Reflect on one thing that you learned from doing this assignment.

I learned the gene mutations that cause ALS and I also learned that ALS is named after a baseball player.