

Yasmine Silmon
Genome Assignment

Chromosome Maps

1. What chromosome did you choose ?

I chose chromosome 8.

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

Chromosome 8 consists of over 1400 genes, with 140 million base pairs.

4. List one gene that is located on this chromosome.

One gene located on the chromosome is ZFPM2.

5. State the function of the gene listed.

The function of ZFPM2 is the development of the diaphragm and cardiovascular system.

Introduction to Nucleotide BLAST

*6. What is the second sequence description match for your query sequence? aattggaagc
aaatgacatc acagcaggtc agagaaaaag ggttgagcgg caggcaccca gagtagtagg tctttggcat taggagcttg
agcccgagcg gccctagcag ggacccagc.*

Homo sapiens CFTR (CFTR) gene, partial cds.

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

The encoded proteins' function is work work as a chloride channel, and control ion and water secretion and absorption in epithelial tissues.

8. For what disease is a mutated form of the gene responsible ?

The most commonly seen mutation causes cystic fibrosis.

9. On what chromosome is the gene located ?

This chromosome is located on the 7q31.2

10. Choose a species (state the scientific name) other than Homo Sapiens, that also has 100% identity for this sequence.

Another species that has 100% identity for this sequence is *Pongo Abellii*.

11. What is the common name for this species?

The common name for this species is an orangutan.

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

I am not surprised that pongo abellii also share this gene with homo sapiens due to the fact that genetically, and evolutionary they are very similar. They're one of the closest species to homo sapiens.

13.

a. What is the genus and species with this nucleotide sequence ?

The genus is *sapajus*, and the species is *sapajus apella*.

b. What is the common name ?

The common name for this species is *tufted capuchin*.

c. How many gaps occur between the two sequences?

The gaps that occur are 1/199 (0%).

14. What is a gap in sequence alignments ?

A gap in a sequence alignment simply means that one, or possibly more, amino acids residues have been deleted from the sequence.

<https://proteinstrutures.com/sequence/sequence-alignment/#:~:text=A%20gap%20in%20one%20of,insertion%20in%20the%20second%20sequence>.

15. NM_14556

Mus musculus TAR DNA binding protein (Tardbp), transcript variant 1, mRNA

16. NM_013444

Homo sapiens ubiquilin 2 (UBQLN2), mRNA

17. NM_001010850

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

18. KJ174530

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

19. What disease is associated with mutations of the genes referenced in #15-#18? What is the common name of this disease ?

Amyotrophic Lateral Sclerosis, ALS, is associated with mutation with these genes. It is commonly known as Lou Gehrig's Disease.

20. What is genbank ?

Genbank is the NIH genetic sequence database. It's a collection of all publicly available DNA sequences.

Introduction to Protein BLAST

21. What is cDNA ?

cDNA is synthetic DNA that has been transcribed from a specific mRNA through a reaction using the enzyme reverse transcriptase.

22. What is the sequence match ?

ACATTTGCTTCTGACACAATTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCATCTG
ACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTT
GGTGGTGAGGCCCTGGGCAG

Bet-a-globin (homo sapiens)

23. Do you see any difference between the two amino acid sequences ?

Yes, there is a difference between the two amino acid sequences.

24. If you say differences, what were they ?

There is a space on 361, where the query sequence is SYVGVG, but the subject sequence is SYRVG. In the amino acid comparison, we see the difference in the G and R represented by a gap.

25. Are there any gaps in the sequence alignment ?

There is a gap on 121, we see this represented by GAP on both amino acid sequences.

26. What gene encodes for the polypeptide you were analyzing ?

The encoded gene for the polypeptide is FDFR3.

27. What is the function of this protein ?

This protein binds acidic and basic fibroblast growth hormones and plays a role in bone development and maintenance.

28. What human disease is caused by a mutation in this gene ?

Mutations of this gene cause craniosynostosis and multiple types of skeletal dysplasia.

29. What is the connection among the following: NIH, NLM, NCBI, and HHS ?

They are all national databases for an array of resources relating to medicine, health, and biology, and biotechnology.

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

The biggest takeaway from this assignment that I had was how to use these databases to gather more information on chromosomes, genes, mutations, and amino acid sequences. The information can look intimidating but when you read through the instructions it is fairly simple to gain access and interpret the information given. I was genuinely impressed how quick and easy it was to access so much. In addition, with ALS, I did not know that the common name Lou Gehrig's Disease was named after the baseball player and to my surprise there is more than one baseball player who has the disease.