Expt. No.	Bioinformatics: Genomics & Proteomics	Date:	Roll No.

Aim: To retrieve information about gene and protein from molecular biology databases and websites

Referral Websites
1. NCBI Entrez: https://www.ncbi.nlm.nih.gov/
2. PubMed: https://pubmed.ncbi.nlm.nih.gov/
3. Genome Database: https://www.ncbi.nlm.nih.gov/genome/
4. Protein Data Bank: https://www.rcsb.org/
5. Enzyme Database: https://enzyme.expasy.org/
6. Pathway Database: https://www.genome.jp/kegg/pathway.html
7. OMIM: https://www.omim.org/

Gene to be investigated:

Queries to be answered using the websites.

1) What scientific papers are associated with the protein/enzyme of humans?
2) What is the size of the mRNA? How many amino acids are present?
3) How do we find the amino acid sequence?
4) What are the protein polymorphisms observed?
5) Where is the protein found in the cell?
6) Which Tissue expresses the protein/enzyme the most?
7) How is the gene organized with respect to exons/introns?
8) How many catalytic subunits of the enzyme/protein?
9) Where is the gene positioned in the human genome? On what chromosome is the gene?
10) What genes are adjacent to our gene of interest in humans?
11) Is the 3D structure of the protein known? How do we find it?
12) Which enzyme subclass does the protein/enzyme belong to?
13) Are there genetic diseases related to the protein? What are the clinical symptoms?
14) Which proteins are associated with the given protein/enzyme?
15) Can you write down the pathway in which the protein participates?

A: DNA Sequence

Go to the NCBI Entrez: https://www.ncbi.nlm.nih.gov/ website. Under search, type name of enzyme "_____". In the dropdown adjacent to the search bar, choose "Gene". Select the results with respect to Homo Sapiens. Based on the search results, answer the following.

Q.1: How many genes of the enzyme family are present?
Q.2: Are the genes on the same chromosome?
Q.3: Write the chromosomal locations of the genes.
Q.4: Write the Gene Ids of all the genes.

Answers:

Find out more information about the gene. Select the gene for (Homo sapiens). A new window will open. Based on information in the window, answer the following.

Chromosomal Location:

Is the chromosomal location on the short arm or the long arm:

Number of exons:

Click on the "Genome Data Viewer" seen on the right section of the information in the webpage. A new window will open. Find the genes upstream, downstream of your given gene (adjacent genes) in the chromosome.

Gene Upstream of the given gene:

Gene Downstream of the given gene:

Length of Gene

Exact Location of gene on chromosome:

Go back to the gene page in NCBI. From the gene page find out the following.

Tissues in which the gene is expressed:

The tissue in which the gene is expressed the most:

The number of isoforms of gene i.e mRNA:

Protein (UniProtKB/SwissProt) ID of gene:

Click on the Protein ID of gene and go the UniProt site of the gene.

B. Protein Sequence

From the UniProtKB website. Write the following information related to the protein.

Function of the protein/enzyme:

Write the EC number of the enzyme:

Subcellular Location of the protein/enzyme:

Number of amino acids in various isoforms (if Present):

Types of post-translational modifications in the protein.

Write down the variants which cause the protein to lose its function.

Write down the disorder caused by the protein/enzyme dysfunction.

Using "Blast", blast the longest isoform against the UniProt Database.

Which species shows greater than 98% similarity with the given protein/enzyme in Homo sapiens:

Go to "Tools" and select "compute pI/MW" of the longest isoform of the protein.

Write down the pI and the molecular weight of the protein/enzyme.

Is the 3D structure of the protein or enzyme determined experimentally or is it a predicted structure?

If structure has been experimentally deduced write down the PDB ID (s) of the protein/enzyme:

Go to the Enzyme database https://enzyme.expasy.org/. ***Write the EC number in the search tool box and click enter.***

Find out the Uniprot Protein IDs of the enzyme in webpage belonging to mouse (mus musculus) and brown rat (rattus norvegicus).

From this page click on KEGG Pathway.

Enlist the pathways where the enzyme participates.

Go to the OMIM (Online Mendelian Inheritance in Man) database
OMIM: https://www.omim.org/

Are there genetic diseases related to the enzyme? What are the clinical symptoms?

Attendance (01)	Performance (03)	Discipline (01)	Total (05)	Teachers Signature