Task 2: Multiple Sequence Alignment (MSA) Using Clustal Omega

Aim :

This task aims to perform multiple sequence alignment of hemoglobin beta protein sequences from different species and identify conserved regions using the Clustal Omega tool.

Introduction :

Multiple Sequence Alignment (MSA) is an advanced tool used for aligning three or more biological sequences including proteins and nucleotides. The procedure aims to determine the conserved parts of biological sequences belonging to the same family.

Proteins that carry out similar functions to living organisms often possess similar amino acid sites. Analysis of the conserved sites helps understand the function and conservation of the protein. In this assignment, the hemoglobin beta protein from various species was aligned using Clustal Omega.

Materials and Tools Used :

Database: UniProt

Tool: Clustal Omega (EMBL-EBI)

Type of sequences: Protein sequences

Protein family: Hemoglobin beta

Number of sequences used: 5

Sequences Used :

Hemoglobin subunit beta – Homo sapiens

Hemoglobin subunit beta 1 – Mus musculus

Hemoglobin subunit beta 1 – Rattus norvegicus

Hemoglobin subunit beta – Bos taurus

Hemoglobin subunit beta 1– Danio rerio

Methodology :

The protein sequences for the Hemoglobin beta chain based on various species were obtained from the UniProt database, which provides the FASTA format for proteins. Five sequences obtained were merged and processed using the Clustal Omega tool for multiple sequence alignment. Analyzing the alignment results revealed the conserved regions as well as the variations within the sequences.

Result :

A multiple sequence alignment of the hemoglobin beta protein from five different species (Homo sapiens, Mus musculus, Rattus norvegicus, Bos taurus, and Danio rerio) was successfully carried out using the Clustal Omega system. The results of the experiment reveal a high degree of similarity among the five mammalian species.

A few conserved regions that appear throughout all the sequences are identified, marked with asterisks (*) in the alignment result. These conserved residues are related to the important regions of the hemoglobin beta chain that function during the process of oxygen binding and structural stability.

Minor discrepancies were found in certain areas of the sequences, particularly in the hemoglobin beta gene from Danio rerio, or the zebra fish. These discrepancies illustrate the evolutionary differences that have occurred between fish and other animals that live on land.

The presence of conserved residues in different species confirms the crucial biological function of the hemoglobin beta protein, and the utility of the multiple sequence alignments in the identification of the conserved regions is evident.

Discussion :

The conserved regions that were identified using the multiple sequence alignment indicate that hemoglobin beta is a conserved protein. This implies that these amino acids play a vital role in protein function. Non-conserved regions could account for differences between species that are adaptive but that do not affect protein function.

Conclusion :

The multiple sequence alignment technique employing the tool “Clustal Omega” has successfully identified the conserved and varying parts of the hemoglobin beta proteins of different species. The significance of the multiple sequence alignment technique has been brought out through the present study, which demonstrates the role of proteins in characterizing the functionality and evolution of species. The study clearly illustrates the conservation and significance of the hemoglobin beta protein to vertebrates as an oxygen-carrying molecule.

References :

UniProt Consortium. UniProt Protein Database. https://www.uniprot.org

Clustal Omega – Multiple Sequence Alignment Tool. https://www.ebi.ac.uk/Tools/msa/clustalo/

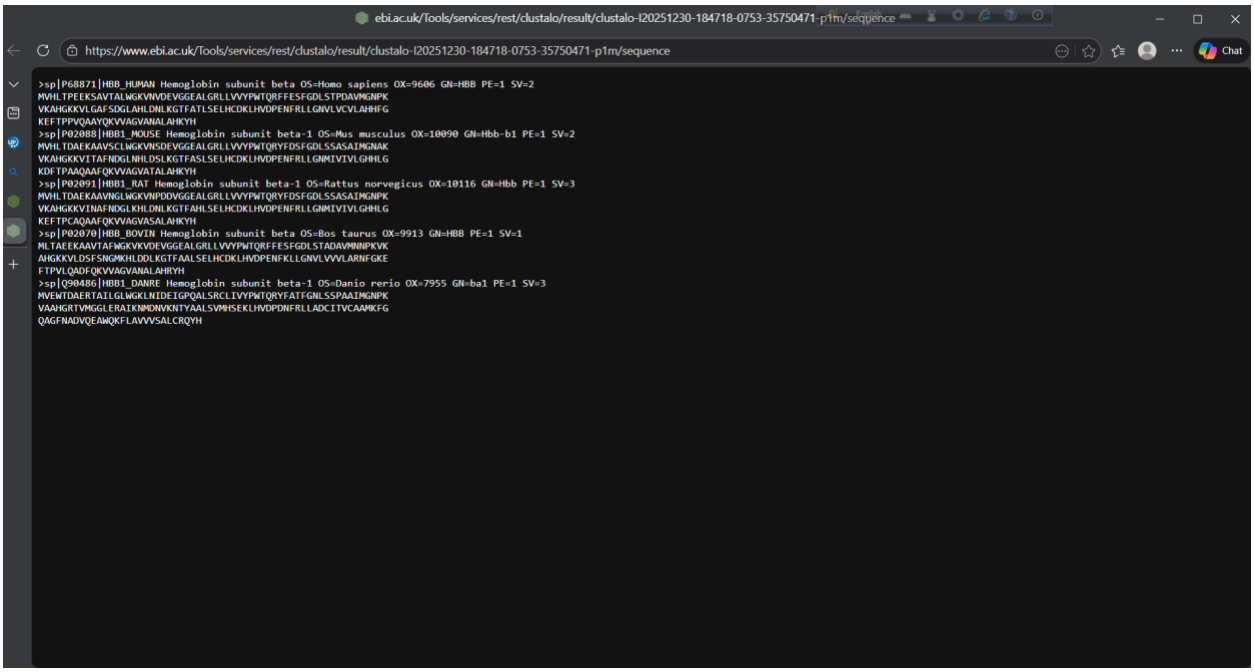

Figure 1: Input protein sequences used for MSA

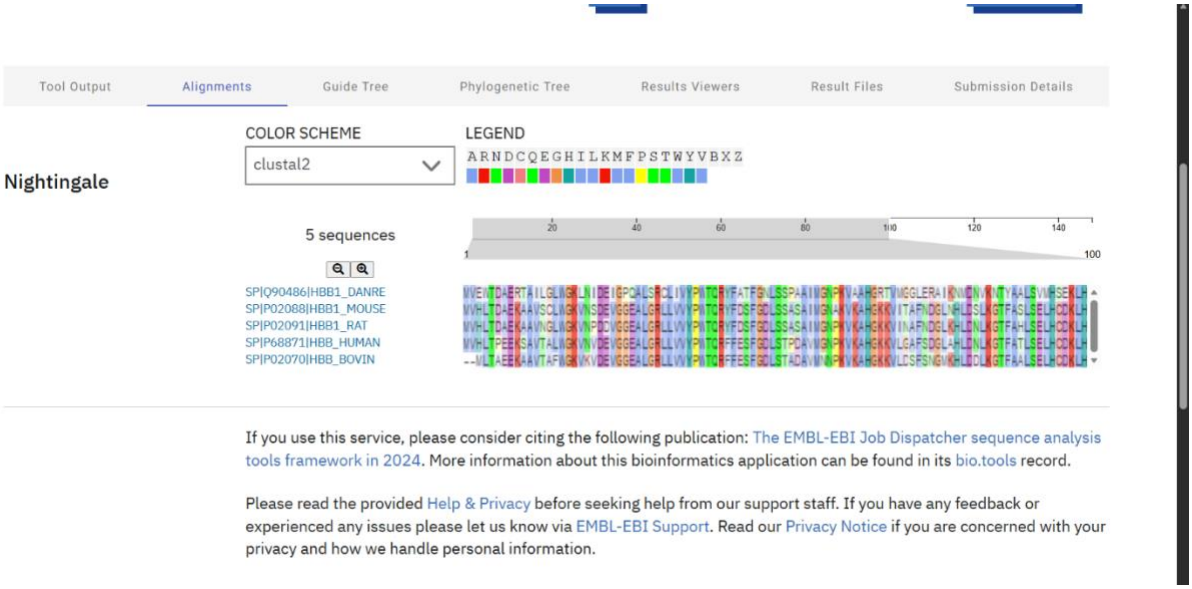

Figure 2: Colored multiple sequence alignment output

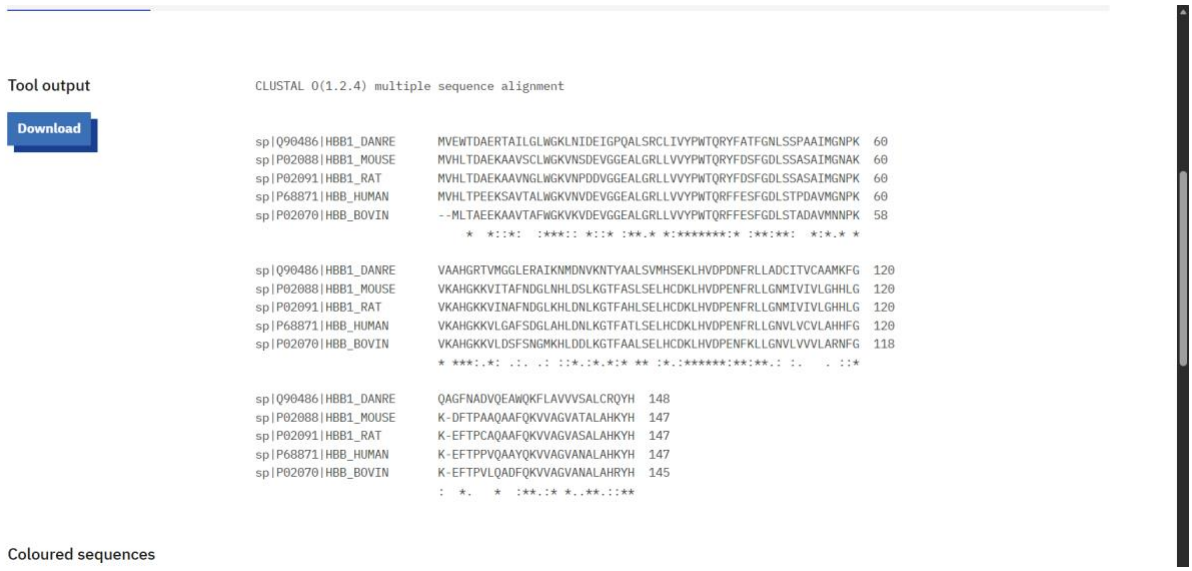

Figure 3: Conserved regions indicated by asterisks (*)