

Multiple Sequence Alignment of Hemoglobin Beta Chain

Task 2 – Bioinformatics Internship (CodeAlpha)

Abstract

Multiple Sequence Alignment (MSA) is a fundamental bioinformatics technique used to identify conserved regions, functional motifs, and evolutionary relationships among biological sequences. In this study, Multiple Sequence Alignment was performed on Hemoglobin Beta Chain protein sequences obtained from different vertebrate species using Clustal Omega. The results demonstrate a high degree of conservation across species, confirming the essential functional role of Hemoglobin in oxygen transport.

Introduction

Hemoglobin is a vital oxygen-transport protein present in red blood cells of vertebrates. The beta subunit of hemoglobin plays a critical role in oxygen binding and release. Studying conserved regions of this protein across species helps in understanding evolutionary conservation and functional significance. Multiple Sequence Alignment enables the comparison of homologous sequences to identify conserved residues and evolutionary divergence.

Objective

The main objectives of this task are:

- To perform Multiple Sequence Alignment of Hemoglobin Beta Chain proteins.
- To identify conserved amino acid residues.
- To analyze evolutionary relationships among different species.

Materials and Methods

Protein Sequences

Hemoglobin Beta Chain protein sequences from the following organisms were retrieved from UniProt in FASTA format:

- Homo sapiens (Human)
- Pan troglodytes (Chimpanzee)
- Mus musculus (Mouse)
- Rattus norvegicus (Rat)
- Bos taurus (Cow)

Alignment Tool

- **Clustal Omega** (European Bioinformatics Institute)

Methodology

1. Protein sequences were downloaded from the UniProt database.
2. Sequences were uploaded in FASTA format to Clustal Omega.
3. Multiple Sequence Alignment was performed using default parameters.
4. The alignment output was analyzed to identify conserved regions.
5. Screenshots and alignment files were saved for documentation.

Results

The Multiple Sequence Alignment revealed:

- A high level of conservation across all five species.
- Strong similarity between human and chimpanzee sequences.
- Minor variations in mouse and rat sequences.
- Conserved histidine residues associated with heme binding.
- Minimal insertions or deletions, indicating structural stability.

Discussion

The conserved regions observed in the alignment highlight the biological importance of Hemoglobin Beta Chain. The high conservation across species suggests strong evolutionary pressure to maintain protein structure and function. Small amino acid substitutions in rodent species reflect evolutionary divergence without affecting the core function of the protein.

Conclusion

Multiple Sequence Alignment of Hemoglobin Beta Chain proteins successfully identified conserved regions critical for oxygen transport. The results confirm the evolutionary stability and functional importance of Hemoglobin across vertebrates. This study demonstrates the effectiveness of MSA in comparative protein analysis.

References

1. UniProt Consortium. UniProt: the universal protein knowledgebase.
2. Sievers F, Higgins DG. Clustal Omega for making accurate alignments of many protein sequences.
3. European Bioinformatics Institute (EBI) – Clustal Omega Tool.

Appendix

- FASTA sequences file
- Multiple Sequence Alignment output (.aln)
- Alignment screenshots