Sequence Identification Details

(PART – I)

Sequence Type: Protein

BLAST Program Used: BLASTp (Protein BLAST)

Protein Name: Envelope glycoprotein gp160

Organism: Human immunodeficiency virus 1 (HIV-1)

Database Used: ClusteredNR (nr)

Query Length: 100 amino acids

Percent Identity: 80.43%

Query Coverage: 92%

E-value: 2e-44

Interpretation of BLAST Results:

The E-value obtained is extremely close to zero, indicating a highly significant match.

The percent identity is greater than 80%, which confirms that the sequence closely matches the identified protein and supports reliable identification.

Alignment Verification

The alignment of the query sequence with the top BLAST hit was examined to confirm sequence similarity. The alignment showed a high number of identical amino acids with very few gaps or mismatches. The conserved regions were clearly visible throughout the alignment.Based on the alignment results, the query sequence corresponds to the identified protein, indicating that both sequences represent the same protein.

Biological Function

The identified sequence corresponds to the envelope glycoprotein gp160 from Human immunodeficiency virus 1. This protein is a viral structural protein that plays a crucial role in viral infection. It is involved in the attachment of the virus to host cell receptors and facilitates viral entry into host cells. The protein is later cleaved into gp120 and gp41, which are essential for membrane fusion and viral infectivity.

Classification and Interpretation

The unknown sequence is identified as the envelope glycoprotein gp160 from Human immunodeficiency virus 1. BLAST analysis showed a percent identity of ~80.43% with an E-value of 2e-44, indicating a strong, statistically significant match. Alignment and functional analysis confirm that it corresponds to a viral structural protein involved in host cell attachment and viral entry.

Final Conclusion

The given unknown sequence is identified as the envelope glycoprotein gp160 protein from Human immunodeficiency virus 1. BLAST analysis revealed a percent identity of approximately 80.43% and an E-value of 2e-44, indicating a strong and statistically significant match. The alignment and functional analysis further support that the sequence corresponds to a viral structural protein involved in host cell attachment and viral entry.

(PART – II)

Sequence Type: DNA

BLAST Program Used: BLASTn (Nucleotide BLAST)

Identified Sequence: Zaire ebolavirus genomic sequence

Organism: Zaire ebolavirus

Database Used: core_nt

Query Length: 220 nucleotides

Percent Identity: 100%

Query Coverage: 100%

E-value: 4e-109

Interpretation of BLAST Results

The BLASTn analysis of the nucleotide sequence revealed a highly significant match with Zaire ebolavirus. The E-value obtained is extremely close to zero, indicating a statistically significant alignment. The percent identity and query coverage are both 100%, confirming that the query sequence is an exact match to the identified viral sequence. These results strongly support reliable identification of the sequence.

Alignment Verification

The alignment between the query sequence and the top BLAST hit was examined and showed complete nucleotide similarity across the entire sequence length. No gaps or mismatches were observed in the alignment. The full-length alignment with 100% identity confirms that the query sequence corresponds exactly to the identified viral sequence.

Biological Function

The identified nucleotide sequence belongs to Zaire ebolavirus, a virus known to cause Ebola virus disease. The sequence corresponds to a viral genomic region involved in viral replication and infection. Viral genomic sequences play an essential role in encoding proteins required for virus structure, replication, and host infection.

Classification and Interpretation

The analyzed sequence is a DNA sequence of viral origin, identified as belonging to Zaire ebolavirus. As a viral genome, it does not fall under prokaryotic or eukaryotic classification. The 100% sequence identity observed across known viral sequences indicates that this region is highly conserved among related viral strains.

Final Conclusion

The given unknown nucleotide sequence is identified as a genomic sequence from Zaire ebolavirus. BLASTn analysis revealed a percent identity of 100% with full query coverage and an E-value of 4e-109, indicating an extremely strong and reliable match. The alignment and classification confirm that the sequence is of viral origin and corresponds to a conserved region of the Ebola virus genome.