

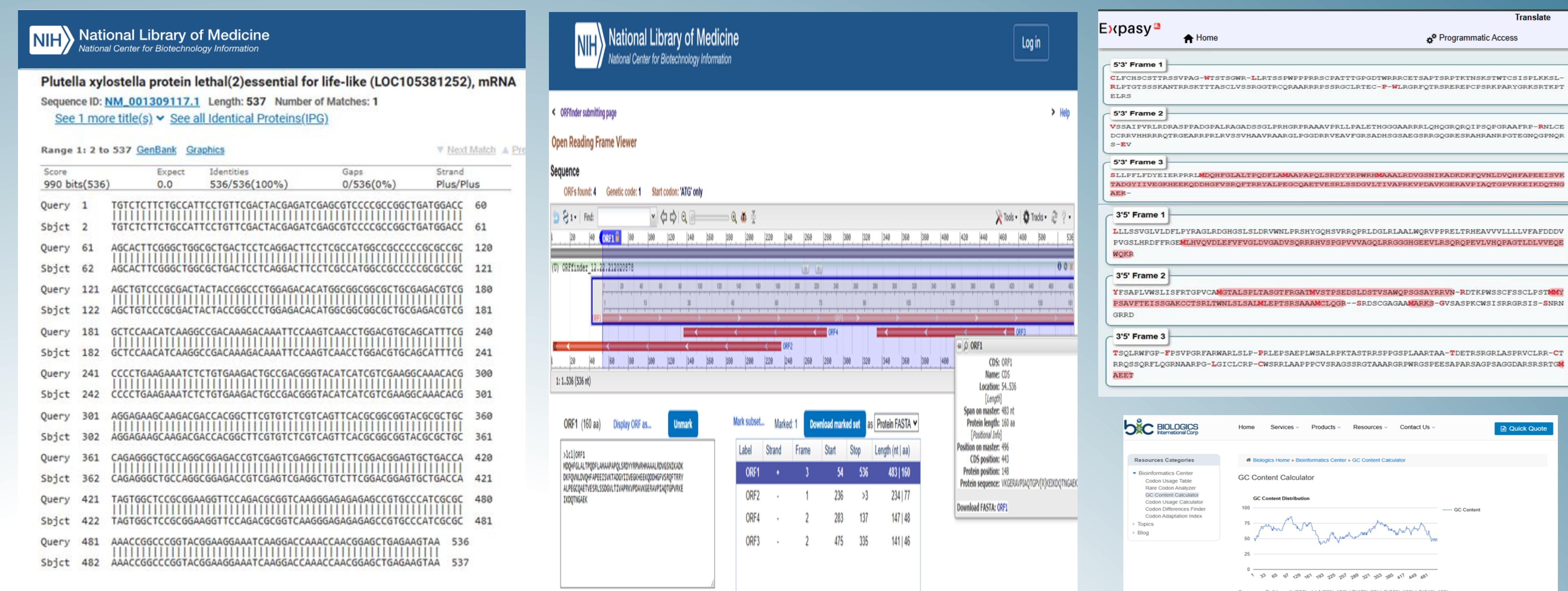
# Computational Analysis of Unknown DNA Sequence Using Integrated Bioinformatics Tools

## Introduction

**Organism:** *Plutella xylostella* (Diamondback moth), a major agricultural pest.  
**Gene:** Lethal(2)essential for life-like (l(2)efl).  
**Function:** Encodes a small Heat Shock Protein (sHSP) that acts as a molecular chaperone to protect cells from stress.

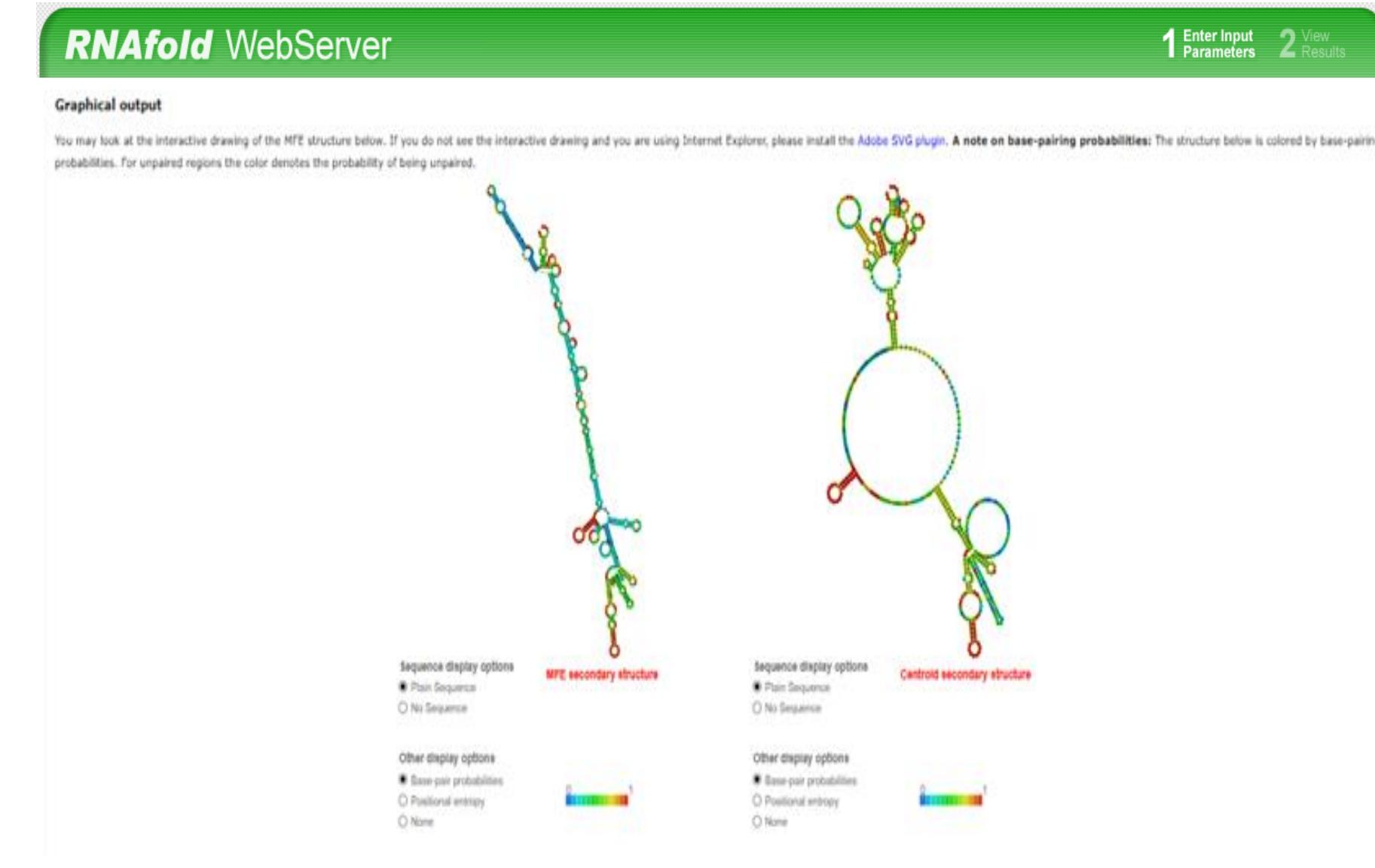
## DNA Analysis

**Tools:** NCBI BLASTn, ORF Finder and ExPASy Translate.  
**Identity:** 100% match to *P. xylostella* locus LOC105381252 with Query Cover 100% and E-value 0.0.  
**GC Content:** 61.75% (Indicates high structural stability)  
**ORF:** Frame 3 (483 bp) selected as the functional coding region because it represents the most biologically plausible coding sequence. It is the only frame that provides a complete, uninterrupted translation starting with a Start Codon (ATG) and ending with a Stop Codon (TAA), resulting in a protein of 160 amino acids. While other frames are interrupted by random stop codons



## RNA Analysis

**Tools:** NCBI Gene and RNAfold (ViennaRNA Webserver)  
**RNAfold Interpretation:** The RNAfold analysis provides two models to represent the folding potential of the mRNA. The MFE (Minimum Free Energy) structure represents the **most stable** and dominant conformation with a highly favorable energy of -194.40 kcal/mol. The Centroid structure represents the average conformation across the entire population of possible folds. The high degree of similarity between these two diagrams indicates that the mRNA transcript has a very well-defined, robust, and consistent structure, which is a characteristic of highly stable "essential for life" genes.



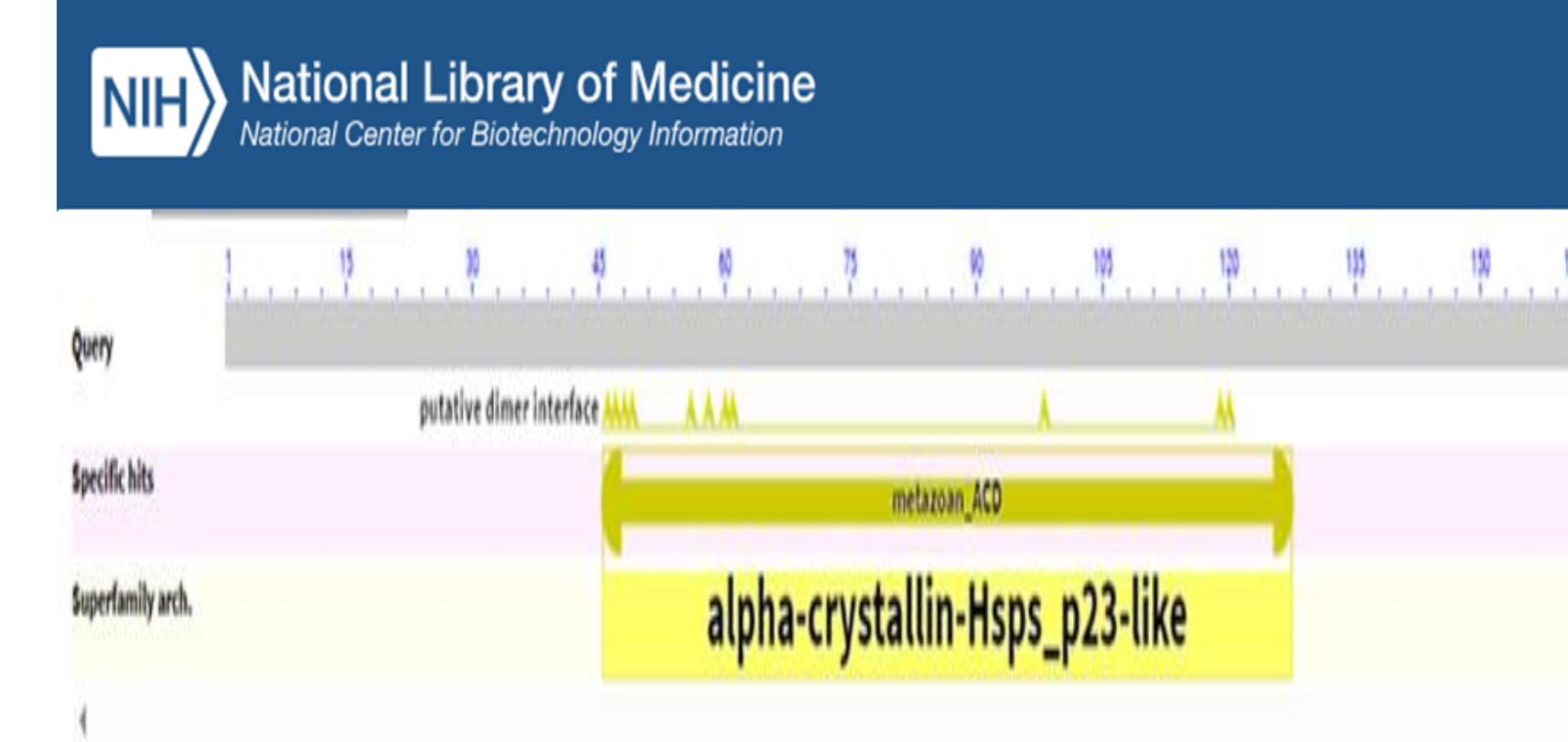
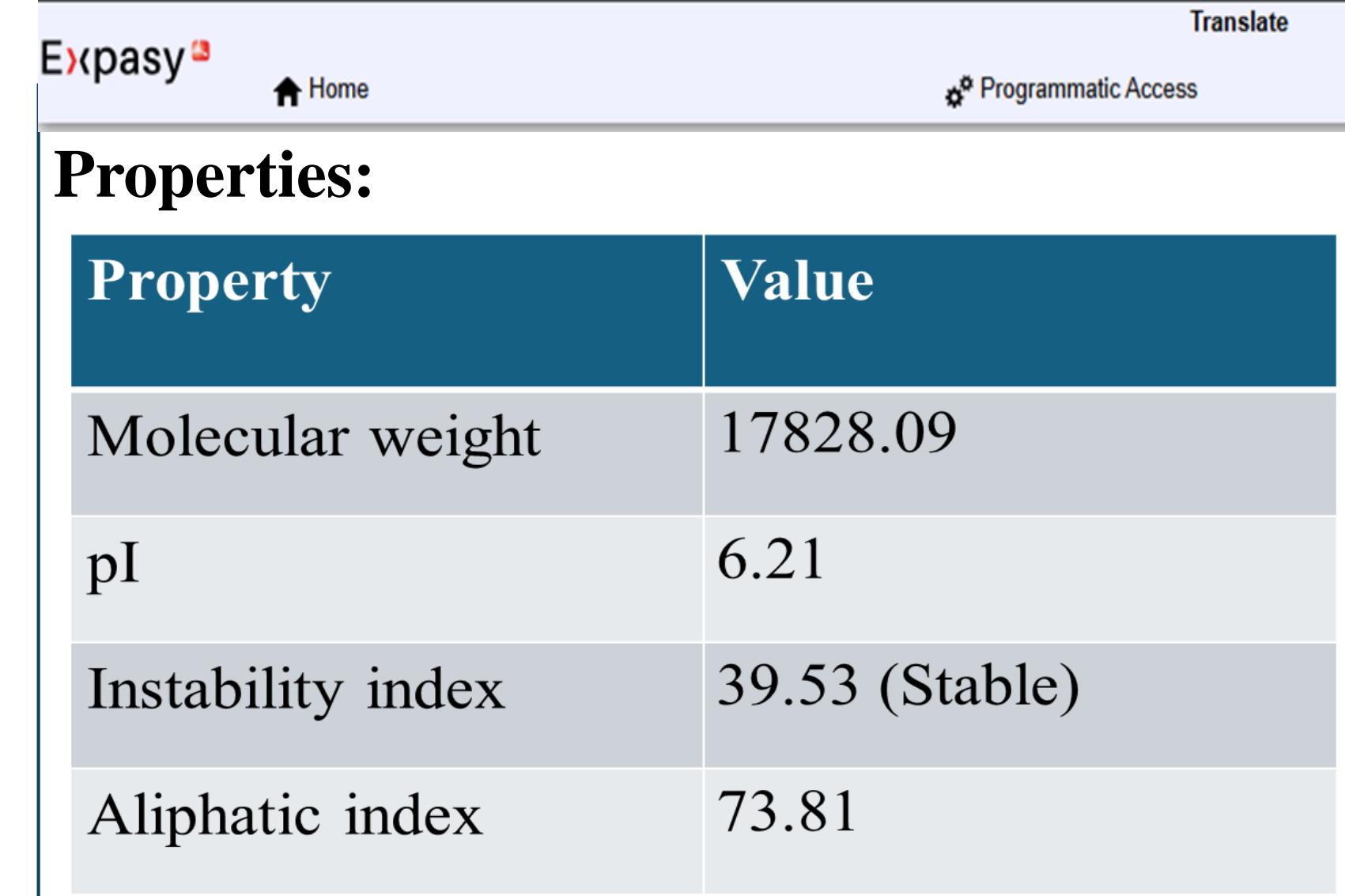
## mRNA Structural Features (Annotations):

DNA sequence is 536 bp and chosen ORF (Frame 3) is 483 bp, the mRNA is divided as follows:

- 5' UTR (Untranslated Region):** The first 50 bp (approx). This region contains the ribosome binding signals.
- Coding Sequence (CDS):** Bases 51 to 533. This is the functional part that translates into the 160 aa protein.
- 3' UTR:** The final 3 bp (the stop codon area).
- Motifs:** The sequence ends with TAA, the essential Stop Codon motif that signals the end of translation.

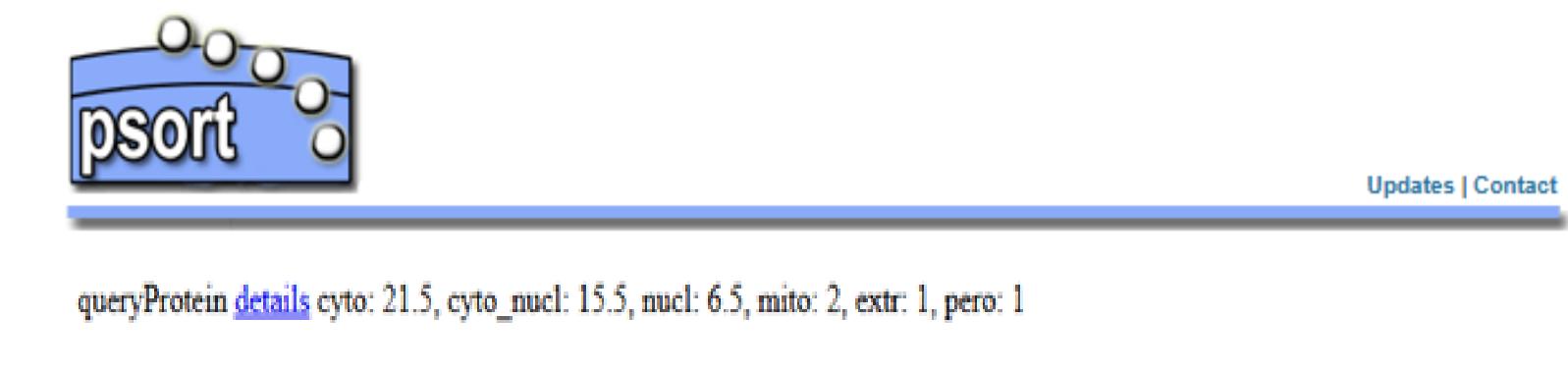
## Protein Analysis

**Tools:** NCBI and ProtParam.  
The analyzed protein is identified as a small Heat Shock Protein (sHSP) belonging to the "lethal (2) essential for life-like" family. Its primary function is to act as a molecular chaperone, which prevents the irreversible aggregation of other proteins during cellular stress (such as heat or chemical changes).  
**Key Domain:** Alpha-crystallin (shsp family)  
The presence of the highly conserved **alpha-crystallin domain** confirms its role in maintaining protein folding and cellular stability. Because it is an "essential" gene, it is critical for the survival, development, and stress resistance of the *Plutella xylostella* organism.



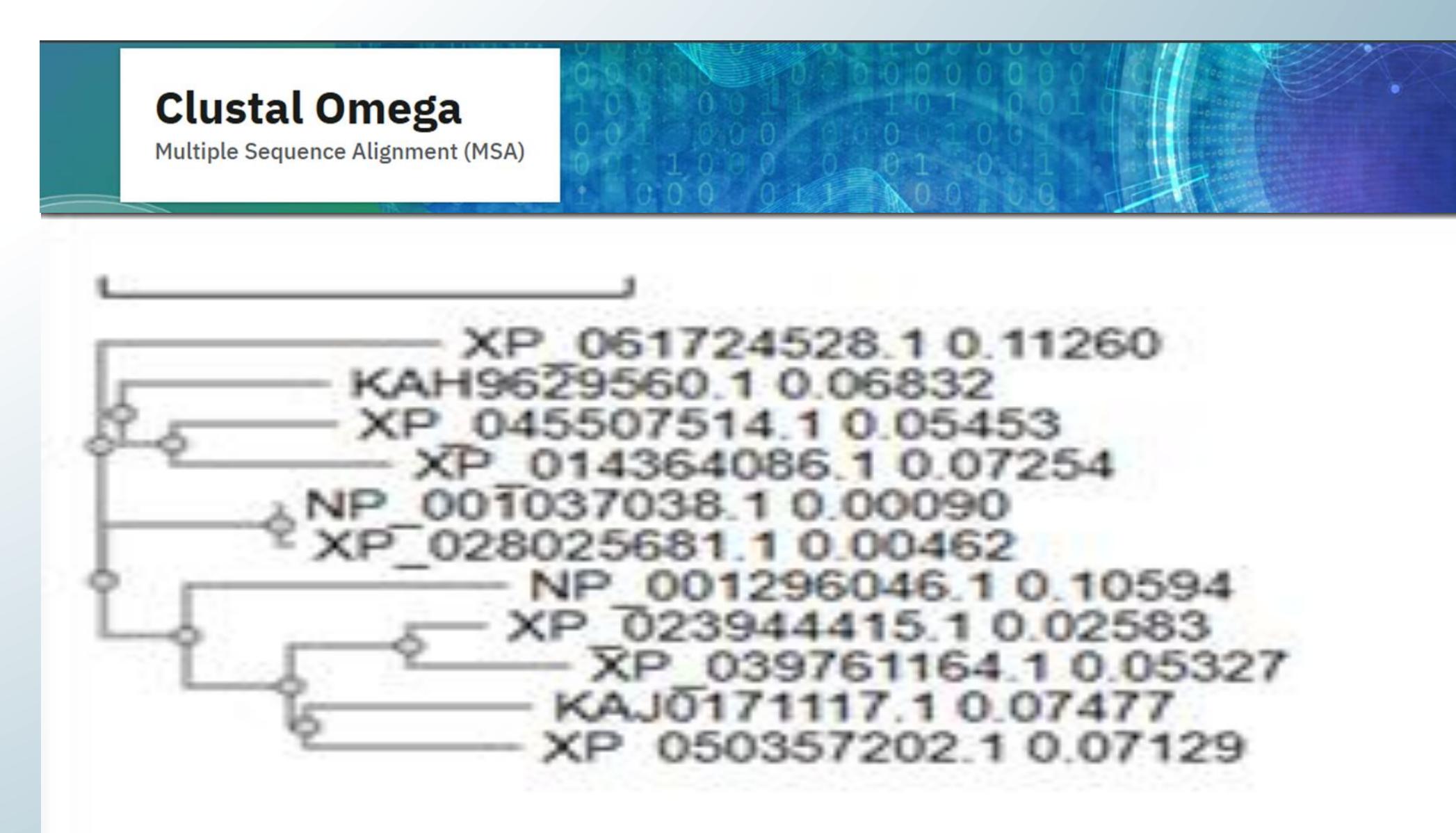
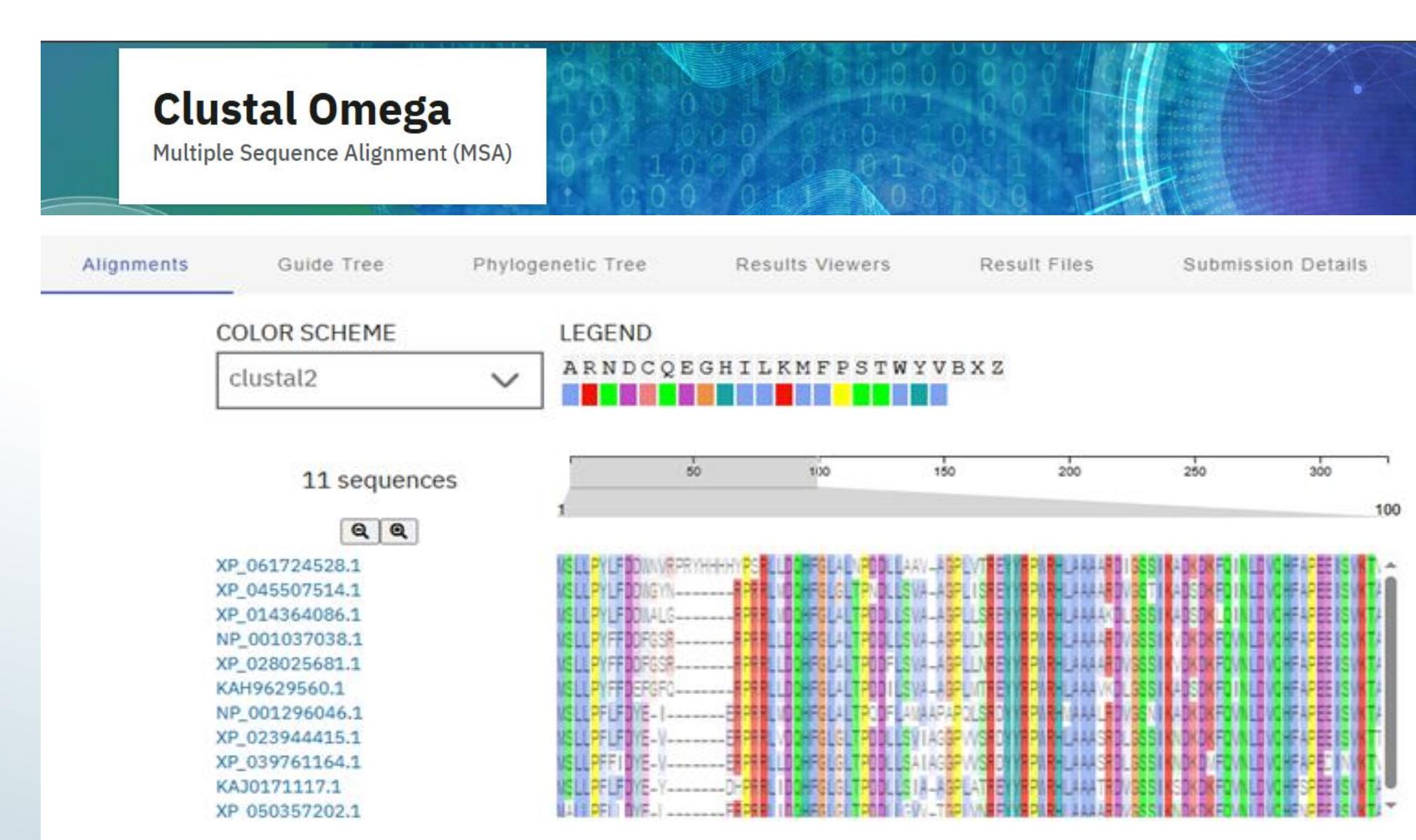
## Subcellular Localization & Interactions

**Tools:** DeepLoc 2.0 and PSORT.  
**Location:** The subcellular localization of the protein was cross-validated using DeepLoc 2.0 and WOLF PSORT, both of which identify the **cytoplasm** as the primary location. These findings are biologically consistent with the protein's role as a small heat shock protein (sHSP), which acts as a molecular chaperone in the cytosol to maintain proteostasis. While both tools also suggest a secondary potential for nuclear localization, this likely reflects the "shuttling" nature of sHSPs, which can move into the nucleus during extreme stress to protect nuclear integrity.



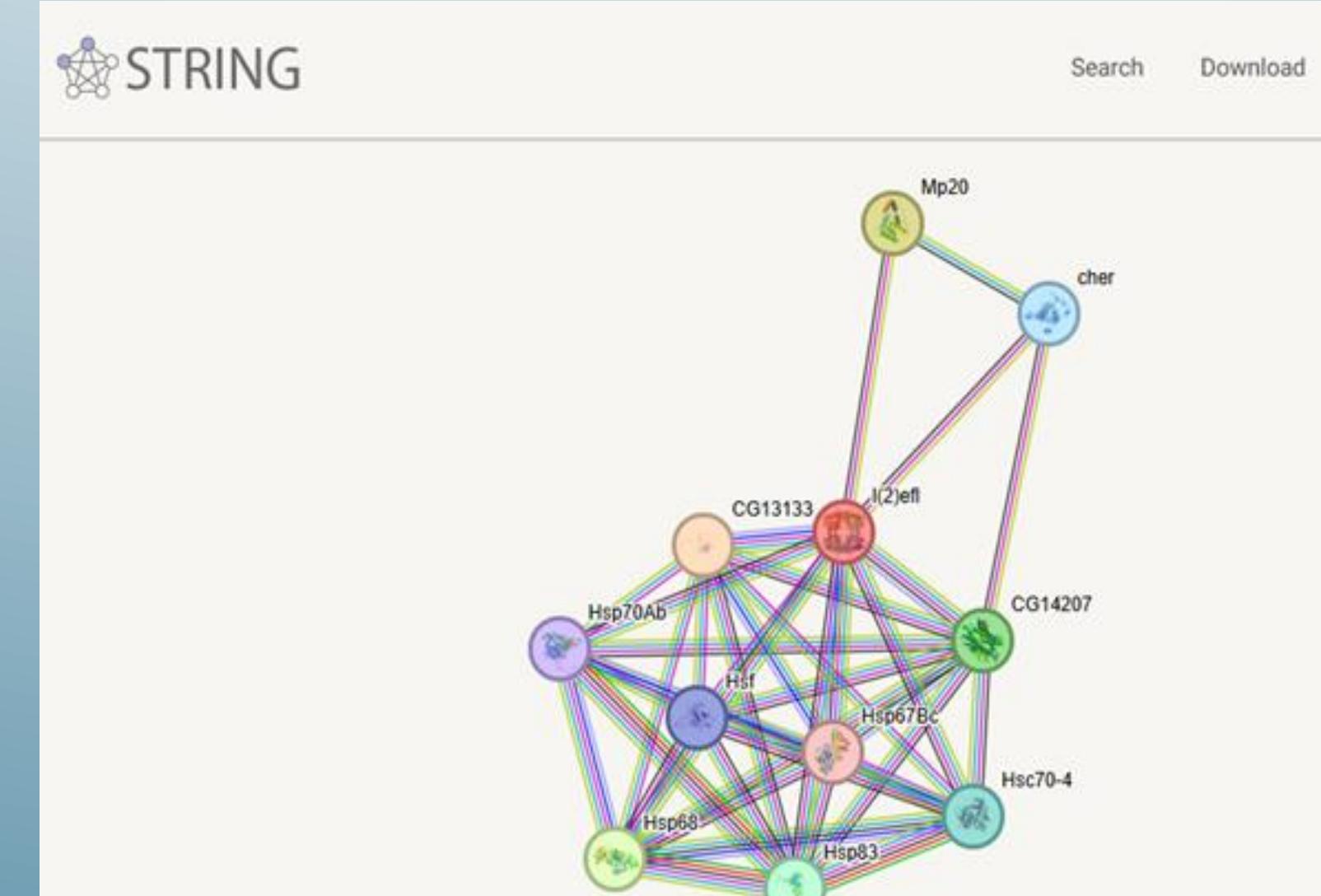
## Phylogenetic Analysis

**Tool:** Clustal Omega.  
**Interpretation:** The phylogenetic tree was constructed to determine the evolutionary relationship between the *Plutella xylostella* protein and other insect homologs. The analysis shows that the 160 aa protein clusters closely with the "lethal(2)essential for life-like" (l(2)efl) protein family, particularly within the order **Lepidoptera**. The strong conservation of the alpha-crystallin domain confirms its identity as an essential small heat shock protein and highlights its conserved role as a molecular chaperone.  
**Conclusion:** The gene is highly conserved across moths and butterflies, proving its vital role in survival and heat tolerance.



## Protein-Protein Interaction Analysis

**Tool:** STRING Database.  
**Network:** Forms a high-confidence cluster with **HSP70** and **HSP90** families.  
**Summary:** The STRING interaction network shows that the 160 aa lethal(2)essential for life-like protein functions as a molecular chaperone. It interacts with other heat-shock proteins to maintain protein quality by stabilizing and refolding damaged proteins during stress. This network confirms its central role in cellular stress response and highlights its importance for the survival and development of *Plutella xylostella*.



## Final Conclusion

The l(2)efl gene encodes a stable, cytoplasmic small Heat Shock Protein (l(2)efl) in the Diamondback moth. It is essential for protecting cellular proteins during environmental stress.

## References:

- <https://www.ncbi.nlm.nih.gov/>
- <https://web.expasy.org/translate/>
- <http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>
- <https://services.healthtech.dtu.dk/services/DeepLoc-2.0/>
- <https://www.genscript.com/psort.html>
- <https://www.ebi.ac.uk/jDispatcher/msa/clustalo>
- <https://string-db.org/>