



Colorado Science and Engineering Fair

2024 Individual Project Abstract Form

Please print 2 copies of the completed form. Sign both copies, keep 1 for your notebook and submit 1 copy to your Regional Fair Director with your other paperwork.

Title of Project: ProteinFlow - A Novel Biocomputational Approach to Discovering Bacteria with Applications to Global Issues Using Protein Sequence

Finalist's Name: Ayush Vispute

School and City: Rampart High School, Colorado Springs

Sponsor's Name: Pankaj Vispute

Category: Micro & Molecular Biology (MCRO)

Division: Senior (9th - 12th grades)

Abstract (250 words or less):

Purpose

This project aims to develop a novel biocomputational pipeline, ProteinFlow, to identify similarities between bacteria based on protein sequence to discover bacteria applicable to global issues.

Procedure

ProteinFlow first identifies bacteria with specific functions through literature and performs searches against a protein database to uncover similar protein sequences. Then, phylogenetic analyses are conducted, employing Neighbor-Joining and Parsimony trees to elucidate evolutionary relationships. Protein structures are predicted and compared using structural alignment values. The methodology encompasses a novel computational workflow, integrating sequence alignment, phylogenetics, and structural analysis.

Results

Two case studies focusing on *Synechococcus Elongatus*, a bacteria providing a novel option for carbon capture technologies, and Tuberculosis, the world's leading infectious killer, demonstrates the efficacy of ProteinFlow and applicability.

ProteinFlow revealed that Cyanobacteria *Okeania*, *Hydrocoleum*, and *Pseudanabaena Salubria* share multiple similarities with *Synechococcus Elongatus*, probably sharing the same functional capabilities to capture carbon at significant rates.

Analyzing *Mycobacterium Tuberculosis* revealed that *Mycobacterium Asiaticum*, *Kiyosense*, and *Tuberculosis Caprae*, also share many similarities under the metrics of ProteinFlow, identifying them as crucial bacteria to research in an effort to find alternate treatments for TB.

Both of these result sets were heavily supported by E-Values (sequence similarity), RMSD Scores, and Evolutionary Analysis, further validating ProteinFlow's capabilities and its results.

Conclusions

1. Developed ProteinFlow, a novel pipeline to analyze proteins with immense application to all fields
2. Identified 6 bacteria that may provide insights into global issues (Tuberculosis and Climate Change)
3. Future Research
 - a. Experimental investigation of key bacteria.
 - b. Integration of 3d modeling software directly into ProteinFlow.

I hereby certify that the above statements are correct and the information provided in the Abstract is the result of one year's research. I also attest that the above properly reflects my own work.

Finalist's Signature: *Ayush Vispute*

Date: 3/4/24

In addition, all students must complete the ISEF Student Checklist (1A), Research Plan, Approval Form (1B), and Checklist for Adult Sponsor (1), and any other ISEF forms required for this type of project. See the International Rules and Guidelines for form requirements. Return COPIES of all of these forms to your Regional Fair Director with your Finalist Verification/Permission Form. A signed copy of this form must be included in your notebook.