CALIFORNIA STATE SCIENCE FAIR
2014 PROJECT SUMMARY

Name(s)

Eric Chen; Nathan Ng

Project Number

34657

Project Title

Novel B-cell Epitope Prediction from Intrinsically Disordered Protein Region Positioning

Abstract

Objectives/Goals

The goal of this project is to utilize the location of intrinsically disordered protein regions on invading pathogens to aid in the identification of B-Cell epitopes, the short protein sequences recognized by the immune system. B-Cell epitope predictors can expedite the creation of peptide vaccines, allowing them to become a feasible alternative to traditional vaccines with lengthy development times.

Methods/Materials

Data culled from the Immune Epitope Database (IEDB) were parsed through various means, including the NCBI Basic Local Alignment Search Tool (BLAST), BioPython, and Python, in order to map epitope sequences onto their proteomes. The sequences were passed through IUPred, a protein disorderness predictor in order to generate an individual disorderness tendency score for each residue. The residue scores and positioning of epitopes were then statistically analyzed using R in order to generate a logistical regression model. This model was optimized by changing its threshold level.

Results

Compared to benchmark B-cell epitope predictors that used single amino acid propensity scales, our model outperformed all of them. For a certain level of specificity, both our models had higher levels of sensitivity, which demonstrates the importance of disorderness as an epitope prediction variable. Through development and testing of the logistic regression model, the area under the ROC curve was determined to be 0.594 for bacteria and 0.636 for viruses, using 200 bootstrap resamplings. The virus disorderness threshold score was 0.454 with 24.08% specificity and 89.20% sensitivity and the bacteria disorderness threshold score was 0.531 with 84.78% specificity and 22.66% sensitivity.

Conclusions/Discussion

We have identified intrinsic protein disorderness as a viable and efficient epitope prediction criterion. By integrating it into a metapredictor, the process of epitope prediction can be made more accurate and efficient. The prediction can be used to expedite the process of peptide vaccine development, making it an economically viable and structurally sound method of vaccination. Compared to live attentuated vaccines, peptide vaccines contain less risk for mutation, are more stable, and can immunize more effectively.

Summary Statement

By utilizing computational biology, our project uncovers a novel method of B-Cell epitope prediction that uses the disordered protein regions of an invading pathogen and can greatly improve the prospects of peptide vaccine development.

Help Received

Dr. Ponomarenko helped with graphics generation for the poster.