

Predicting Protein Stability Change upon Double Mutation from Partial Sequence Information Using Data Mining Approach

Lai, Lien-Fu; Wu, Chao-Chin; Huang, L. T.

Abstract

Prediction of protein stability upon amino acid substitution is an important problem in designing stable proteins. We have developed a classification rule generator for integrating the knowledge of amino acid sequence and experimental stability change upon single mutation. These rules are human readable and hence the method enhances the synergy between expert knowledge and computational system. Utilizing the information about wild type, mutant, three neighboring residues and experimentally observed stability data, we have developed a method based on decision tree for discriminating the stabilizing and destabilizing mutants and predicting the protein stability changes upon single point mutations, which showed an accuracy of 82% and a correlation of 0.70, respectively. In addition, we have developed a fuzzy query method to predict protein stability with partial information. We have developed a web server for predicting the protein stability changes upon single mutations by using fuzzy query mechanism and it is available at <http://bioinformatics.myweb.hinet.net/fqstab.htm>.

Key words: Data mining; Double mutation; Fuzzy query; Protein stability; Sequence information