

## Protein Remote Homology Detection by Using Profile-Based Matrix Transformation Approaches

**Authors :** Bin Liu

**Abstract :** As one of the most important tasks in protein sequence analysis, protein remote homology detection has been studied for decades. Currently, the profile-based methods show state-of-the-art performance. Position-Specific Frequency Matrix (PSFM) is widely used profile. However, there exists noise information in the profiles introduced by the amino acids with low frequencies. In this study, we propose a method to remove the noise information in the PSFM by removing the amino acids with low frequencies called Top frequency profile (TFP). Three new matrix transformation methods, including Autocross covariance (ACC) transformation, Tri-gram, and K-separated bigram (KSB), are performed on these profiles to convert them into fixed length feature vectors. Combined with Support Vector Machines (SVMs), the predictors are constructed. Evaluated on two benchmark datasets, and experimental results show that these proposed methods outperform other state-of-the-art predictors.

**Keywords :** protein remote homology detection, protein fold recognition, top frequency profile, support vector machines

**Conference Title :** ICBCBI 2019 : International Conference on Biological Computing and Biomedical Informatics

**Conference Location :** Paris, France

**Conference Dates :** February 21-22, 2019