

# 基于氨基酸序列预测蛋白质功能性点突变位点

宋卓<sup>1</sup>、张宁<sup>1</sup>、阮吉寿<sup>2</sup>、杨卓<sup>3</sup>、张涛<sup>\*1</sup>

1 南开大学生命科学学院

2 2) 南开大学数学科学学院与LPMC

3 南开大学医学院

突变是研究蛋白质结构和功能的重要方法。点突变实验中，突变位点的选择随机性大，若能对突变后蛋白质功能是否发生变化做出预测将大大减少实验的盲目性。为此，设计了一个基于信号处理的单点替换突变预测模型，对序列上每个位点所有可能的氨基酸替换的效果进行估计。使用蛋白质突变数据库(Protein Mutant Database, PMD)里的12个蛋白共2700多个点突变的数据集，对以上模型进行了验证。结果表明正确率高达81%，并且推荐出的替换选择位点仅占所有可能位点的3.5%。在体外定点突变实验中，使用本模型推荐的高可能性功能突变位点将有助于提高实验的成功率。该模型仅使用蛋白质的氨基酸序列信息，特别是对未知结构的蛋白质同样适用。然而，由于缺乏足够的突变实验数据，该模型仍有待于进一步完善和验证。

## Prediction of Functional Point Mutants of Proteins Based on The Amino Acid Sequences

Mutation is a powerful tool to understand protein structure and function. It will be great helpful if one can predict with certainty whether a protein mutant is likely to be functional. A novel approach based on signal processing is developed to estimate the effect of every possible substitution on each site along protein sequence and then to predict functional point mutants of the protein. This method was applied to a testing dataset of more than 2,700 point mutants belong to 12 proteins collected from Protein Mutant Database (PMD). Among the recommended predictions in PMD, the accuracy of performance is up to 81%. This new approach, only based on the information of amino acid sequence, will be very helpful to choose functional point mutants with high possibility for biologists, especially when the structure of the protein is still not clear.

### 关键词

点突变(Point mutant); 替换突变(Substitution); 蛋白质功能预测(Protein function prediction); 信号处理(Signal processing)