计算机工程与应用 2009 45 (36): 44-46 ISSN: 1002-8331 CN: 11-2127/TP

研究、探讨

蛋白质二级结构预测方法研究

王艳春[1，2]

1.青岛农业大学 信息科学与工程学院，山东 青岛 266109
2.西北农林科技大学 机械与电子工程学院，陕西 杨陵 712100

收稿日期 2009-1-4 修回日期 2009-2-17 网络版发布日期 2009-12-30 接受日期

摘要 为提高蛋白质二级结构预测精度，提出一种新的网络模型和编码方法。首先利用基因表达式编程（GEP）的全局搜索能力同时进化设计神经网络的结构和连接权；其次，对神经网络输入层编码进行了改进，添加了氨基酸残基所处的疏水环境。用PDBSelect25中的36条蛋白质共6 122个残基进行测试，结果表明提出的网络模型和编码方法能有效提高蛋白质二级结构预测的精度。

关键词 蛋白质 二级结构预测 基因表达式编程 神经网络

分类号 TP183

Study of protein secondary structure prediction methods

WANG Yan-chun[1，2]

1.College of Information Science and Engineering，Qingdao Agricultural University，Qingdao，Shandong 266109，China
2.College of Mechanical and Electronic Engineering，Northwest A & F University，Yangling，Shaanxi 712100，China

Abstract

In order to improve the prediction accuracy of protein secondary structure，a new network model and its coding method are proposed.Firstly，the structure and connection weights of BP network are evolved simultaneously by using global research ability of GEP.Secondly，the coding method of neural network is improved by integrating the hydrophobic value around the residue.The model is employed to predict 36 nonhomologous protein sequences with 6，122 residues in PDBSelect25，the results show that the proposed model and coding method can efficiently improve the prediction accuracy.

Key words protein secondary structure prediction gene expression programming neural network

DOI: 10.3778/j.issn.1002-8331.2009.36.014

通讯作者 王艳春 wychun99@sina.com