

# 不同结构的蛋白编码基因的密码子偏性研究

顾万君、马建民、周童、孙啸、陆祖宏  
东南大学吴健雄实验室

利用聚类分析方法,对两类具有不同三级结构的75个蛋白的编码基因的密码子使用偏性进行了分析。75个基因样本序列按照对应蛋白的三级结构被很清晰的分成了两类,从而发现密码子的使用与蛋白质的三级结构有很大的相关性。这一重要结果证实了DNA的一维信息中蕴含着蛋白质的三级结构信息。

## CODON USAGE IN GENES CODING FOR PROTEINS WITH DIFFERENT TERTIARY STRUCTURES

Based on gene codon usage distance 75 genes coding for proteins with two different fingerprints were clustered, to analyze the relationship between gene codon usage and protein tertiary structure. It was found that 75 genes coding for proteins with different tertiary structure can be clearly classified into two groups according to the difference of codon usage bias, which suggests that codon usage is highly correlated to protein tertiary structure. The relationship between codon usage and protein tertiary structure is firstly found in the research of codon usage, which proves that gene sequence contains protein tertiary structure information.

### 关键词

密码子使用(Codon usage); 蛋白质三级结构(Protein tertiary structure); 聚类分析(Cluster analysis)