

大肠杆菌与酵母菌基因特定序列信息参量的研究

陈颖丽、李前忠、马克健
内蒙古大学理论生物物理研究室

提出核酸序列的矩阵表示形式,按位点定义了有生物学意义的信息参数 $M_1(1)$ 、 $M_2(1)$ 和 $M_3(1)$ 。着重研究了不同表达水平的大肠杆菌(*Escherichia coli*, E.coli)的SD序列(Shine-Dalgarno region, SD)以及大肠杆菌(*E. coli*)和酵母菌(*Yeast*)基因起始、终止密码子邻近区域核酸序列的碱基关联性与保守性。并求出相应矩阵的本征值,给出了信息参量与基因表达水平的关系。发现信息参量体现了原核生物和真核生物翻译起始区域的显著差异,而且真核生物碱基起始区域的单碱基保守性程度及碱基关联性程度要比原核生物强。

A STUDY OF THE INFORMATIONAL PARAMETERS FOR THE SPECIFIC SEQUENCES OF E. COLI GENES AND YEAST GENES

The matrix representation of nucleic acid sequences and the definition by $M_1(1)$ 、 $M_2(1)$ and $M_3(1)$ are presented. The definition is biological of significance. The *E. coli* in different expression level were investigated for SD region. The conservative and correlative properties of the bases of the specific sequences between *E. coli* and *Yeast* genes were studied comparatively, and the eigenvalues of the corresponding matrix were calculated. The relation between informational parameters and gene expression level was given. The results show that the informational parameters indicate remarkable difference in the translated starting regions, and the conservativity of a single base and the correlation between base pair in the starting regions of eukaryote are stronger than those of prokaryote.

关键词

SD区域(Shine-Dalgarno region); 矩阵表示(Matrix representation); 信息参数(Informational parameter); 基因表达水平(Gene expression level); 保守位点(Conserved site)