生物物理学报(2003) 19(2), 161-166

PDF文档

利用序列比较寻找胰岛素基因表达启动区与DNA结合蛋白的结合位点

周士新、孙啸、陆祖宏、谢建明、董献军、徐伟、王崎
东南大学分子与生物分子电子学国家重点实验室

NDF1、IPF1和HNF4是与胰岛素基因表达有关的DNA结合蛋白，通过比较SWISSPROT蛋白质数据库中人类、小鼠、大鼠这三种核蛋白氨基酸一级序列、模体和结构域，发现其结构十分相似，根据蛋白质结构和功能的关系，推测这些DNA结合蛋白与胰岛素基因结合的核苷酸序列相似；从GenBank核酸数据库中获得人类、小鼠、大鼠胰岛素DNA序列，用ClustalW比较三者Promoter区的核苷酸序列，显示有一段核苷酸序列较为相似，同时搜索TRANSFAC基因转录数据库中NDF1、IPF1和HNF4蛋白核苷酸结合位点，发现核酸比对保守的部分序列与TRANSFAC数据库中这三个转录因子的DNA结合位点一致，另外一些核酸保守序列可能为其他未知DNA结合蛋白的结合位点。这种核酸序列比对设计为分子生物学实验寻找和验证胰岛素DNA结合蛋白与核苷酸的结合位点提供了简单而实用的方法。

TO FIND NUCLEOTIDE BINDING SITES OF DBPS BY COMPARING THE PROMOTER SEQUENCES OF INSULIN GENE

There were several DNA binding proteins (DBPs) related to insulin gene expression, such as NDF1, IPF1 and HNF4. The structures of three DBPs was homologous by comparing the amino acid sequences, motifs and domains of human, mouse and rat in SWISSPROT protein database. According to the relationships between structure and function of protein, it was supposed that the nucleotide binding sites of DBPs were similar in the promoters of insulin gene. The DNA sequences of human, mouse and rat were obtained in GenBank nuclear database. The ClustalW program of multiple alignment was used to compare nucleotide sequences of the three mammals. It showed that there were some similar sequences in the promoters of insulin gene. In the meanwhile, the nucleotide binding sites of NDF1, IPF1 and HNF4 were searched in the TRANSFAC gene regulation database. The conservative nucleotide sequences were found between the promoters alignment of insulin gene and the nucleotide binding sites of DBPs in the TRANSFAC database. Some other conservative nucleotide sequences in the alignment were perhaps the binding sites of unknown proteins. This device of sequences alignment offered a simple and applicable method to find and confirm the nucleotide binding sites of DBPs in the molecular biology experiments.

关键词

DNA结合蛋白(DBP)(DNA binding protein(DBP)); 胰岛素基因启动区(Promoters of insulin gene); 模体和结构域(Motif and domain); 结合位点(Binding site)