

101 學年度生醫系專題研究競賽報名表

■ 碩士班

□ 大學部

年級：二 學號：0088002 姓名：陳勁含 指導教授：王怡鈞

題目 (中文)：演化分析的新方法：使用皮爾森相關係數評估蛋白質序列間相對演化關係

(英文)：A Novel Approach to Evaluate the Relative Evolutionary Distance of Proteins via Pearson correlation coefficient.

摘要：

Introduction

Evolutionary relationship of all living organisms can be viewed by the phylogenetic tree. So far there are many methods have been developed to construct the tree, e.g. distance based neighbor-joining method, character based maximum parsimony method, etc. However, multiple sequence alignment (MSA) should be performed before those methods. Here we propose a novel approach to replace the MSA process. We combine pairwise sequence alignment (BLAST) and Pearson correlation coefficient (PCC) to simulate the interactive relationship of compared protein sequences.

Rationale

Up to now, several studies have shown that the order in which sequences are added to a multiple sequence alignment could significantly affect the end result. Therefore we want to find if there is another method that makes more reliable results. Our goal is to construct a unique and reasonable phylogenetic tree-building method better than the others.

Material and methods

There were three main processes: 1. Computation of score matrix that stores the total scores of sequence alignments between any two proteins for comparison. 2. Calculate Pearson correlation coefficient from step 1 and create the protein relative distance matrix of one protein to the others. 3. Processes of proteins clustering to derive the relationship of selected proteins from protein relative distance matrix.

Results and Discussion

In order to compare the present developed method with our method, we preliminary fetch protein sequences (using protein arginine methyltransferase 1 as example) from several different species. The sequences are input to currently developed methods and our method respectively, and then compare their output trees. The currently developed methods may cause some unreasonable situations, e.g. mammals seem to be closer to the ancestor than yeasts. Without this problem, our method has made a more reasonable tree.

It has been reported that Neighbor-Joining method has an undesirable feature that it may produce negative lengths to some of the branches. However, after our improvement, this situation has not been seen in our method. Therefore, we adapted the advantage of Neighbor-Joining method that the tree is the only tree of all taxa and avoid the drawbacks of traditional phylogenetic analysis methods that the tree may affect by MSA and outgroup selected.

指導教授_____ (請簽名)