

## Abstract in Chinese

幾十年來已經有非常多的方法被發展來做為基因辨認及基因訊號預測，並提供分析基因及功能元件的主要工具。然而近年來大量不同生物基因定序的完成，比較基因、結構基因及基因訊號研究已成基因資訊的幾個重要主題。本總計畫主要目標主要是在能洞察整體生物基因機械結構及基因結構功能的細部辨認。

在總目標下，本子計畫的主要目的在於發展以模糊類神經網路演算法來有效完成基因辨認及基因結構預測，早期研究基因辨識的方法大都集中在用內容及訊號方法作個別功能元件的預測，最近基因預測則根據多種資料的分析及整合，如基因內容，訊號資訊，基因結構等方法包括類神經網路，語言（基因文法），隱藏式馬可夫法，及動態規劃等都被用來做基因辨識及預測。模糊類神經法是種基於專家意見的（無模型）估測及動態系統，其結合模糊理論及類神經演算法而有兩者長處。模糊理論可包含上述方法，如或然率，數學邏輯，測度理論。而類神經包含動態系統，適應方法及統計的概念在內，他們已成功地應用於系統控制及訊號工程領域

實際上，模糊類神經演算法特別有助於分類，函數逼近，映射等問題。加以訓練及加入專家的知識後，可應用於基因辨識、基因結構及基因訊號預測。本子計畫提出的模糊基因辨識法主要是以類神經網路結構來實現。它是種並行計算法，可以增進學習訓練過程的快速完成。本子計畫起出以標準基因庫的資料來訓練。然後再鎖定在三個基因（視覺基因及生長賀爾蒙及熱休克蛋白 Hsp70 基因）的辨識，希望能發展出有視覺效果的工具及方法。最後我們會與別種方法分別在 sensitivity, specificity 及 accuracy 方面作比較，並與各子計畫的方法相印證、修訂並改善，然後在曾賢晴及林彩雲兩位教授的實驗室做實驗來印證。

## Abstract in English

In the last decades, many methods have been developed for gene recognition, gene structure prediction and signal detection, and provide ways to predict genes and functional elements from DNA sequences. However, driven largely by the vast amounts of DNA sequence data recently, comparative genomics, structure genomics and signal detection become important topics in genome informatics. The goal of this program project is to gain insight into universal biological mechanisms and into the details of gene structure and functions.

The purpose of this component project is to develop a fuzzy neural network scheme to achieve efficient gene recognition, gene structural prediction and signal detection. Many early methods of gene recognition focused on predicting individual functional element using either gene search by content or signal methods. Recently, gene prediction programs have based on the analysis and integration of multiple types of content and signal information and have taken gene structure into account. Neural network, linguistic (i.e., gene grammar) method, hidden Markov model and dynamic programming etc. have been developed for gene recognition, gene structure prediction

and signal detection. Fuzzy logic theory overlaps with the above methods and with the probability, mathematical logic, and measure theory. Neural theory embeds in the mathematical fields of dynamical systems, adaptive (learning) scheme and statistics. Therefore, fuzzy neural network algorithms can integrate the above algorithms as expert-based numerical model-free estimators and dynamical systems.

In practice, fuzzy network algorithms are especially useful for classification and function approximation/mapping problems with training and a priori information (including expert's knowledge) and can tolerate some imprecision. Fuzzy neural network methods integrate expert knowledge (a priori information or rules of genes) and data, learning to improve the performance of gene recognition and gene structure prediction in genome informatics. The proposed fuzzy method is implemented with a neural network structure, which is added by parallel computations to achieve a fast convergence in the learning procedure.

**The main work of this component project is to develop and train a fuzzy neural network algorithm for gene recognition, gene structure prediction and signal detection from a benchmark data set at first. Then, we focus on the recognition of visual gene, growth hormone gene and Hsp70 gene and develop several visualization tools. The results may be improved by discussing with other component projects. Finally, the sensitivity, specificity and accuracy of the proposed method will be compared with other methods and confirmed in Tseng' and Lin' s labs.**