

行政院國家科學委員會專題研究計畫 成果報告

流感 A 型病毒 HA 蛋白的演化與變異 研究成果報告(精簡版)

計畫類別：個別型
計畫編號：NSC 98-2221-E-468-021-
執行期間：98 年 08 月 01 日至 99 年 07 月 31 日
執行單位：亞洲大學生物資訊學系

計畫主持人：張培均

計畫參與人員：碩士班研究生-兼任助理人員：鄭尹婷
碩士班研究生-兼任助理人員：陳憬輝
碩士班研究生-兼任助理人員：喬思敏
大專生-兼任助理人員：陳柏逢
博士班研究生-兼任助理人員：蕭雅莉

報告附件：出席國際會議研究心得報告及發表論文

公開資訊：本計畫可公開查詢

中華民國 99 年 10 月 31 日

流感 A 型病毒 HA 蛋白的演化與變異

計畫類別： 個別型計畫 整合型計畫

計畫編號：NSC 98-2221-E-468 -021 -

執行期間：98 年 08 月 01 日至 99 年 07 月 31 日

執行機構及系所：亞洲大學生物資訊學系

計畫主持人：張培均

共同主持人：

計畫參與人員：蕭雅莉、鄭尹婷、陳柏逢、陳憬輝、喬思敏

成果報告類型(依經費核定清單規定繳交)： 精簡報告 完整報告

本計畫除繳交成果報告外，另須繳交以下出國心得報告：

赴國外出差或研習心得報告

赴大陸地區出差或研習心得報告

出席國際學術會議心得報告

國際合作研究計畫國外研究報告

處理方式：除列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權， 一年 二年後可公開查

詢

中 華 民 國 99 年 10 月 30 日

英文摘要

Keywords : Influenza, Hemagglutinin, amino acid, molecular evolution, physiochemical property

Hemagglutinin (HA) protein is very important in the event of influenza (flu) infection, because it binds to the receptor on the cellular surface of human lung specifically, and bring about membrane fusion by the change of protein conformation. All of the processes are physiochemical interactions in the molecular level. There are two significant physiochemical features for the Influenza infection which broke out in 1918 may responsible to the high infection and mortality: (1) In general, the size of HA binding position is narrow, but becoming larger by one amino acid change on the binding position in the flu of 1918, that apt to promote the infection to human. (2) Two histidine-rich segments in the HA protein with basicity may promote the membrane fusion in the infection process. One of the two segments found only in avian flu.

According to these clues, we inferred that the infection host and the probability of cross-species infection could be determined by the physiochemical properties of residues in the HA protein sequence.

In this project, we will study the correlation regarding the properties of the HA protein of influenza A viruses that affecting the probability of cross-species infection, the infective host and the infection ability. Clustering analysis will be used for molecular evolution of the HA protein, at first. Secondly, we will define and quantify the HA protein sequences by physiochemical properties of residues, clustering analysis will apply. We also will utilize the method of data mining to discover the significant residues among the HA protein sequences that infecting different species. Finally, we will study the structural effects of residue change of the HA sequence that binding to the receptor about these significant residues. All the HA protein sequences was collected from public database.

中文摘要

關鍵詞：流行性感冒，血球凝集素，氨基酸，分子演化，物理化學性質

對於流行性感冒病毒來說，血球凝集素 (hemagglutinin, HA) 蛋白是非常重要的，因為它結合到人類的肺細胞表面受器，並改變形態使病毒的膜與宿主融合，整個過程在分子層次上是物理化學的交互作用。1918 年造成異常高的傳染性和死亡率的 HA 蛋白，在結構上有兩個特別突出的特徵異於一般：(1) 一般受器結合位置的範圍是狹小的，然而 1918 年流行性感冒病毒的 HA 蛋白卻以一個氨基酸突變，使結合位置稍微的變大，這可能會增加傾向於對人類細胞的感染。(2) 有兩個之前未被查覺的鹼性小片段(histidine rich)，當病毒融入細胞膜時，可能會提高 HA 的感染性，其中一個小片段被發現只有在感染鳥的病毒種類中才有。

根據這些線索，我們推論流行性感冒病毒的感染對象，以及禽流感感染人的可能性，應該跟它的 HA 蛋白質序列所具有的物理化學性質，有很密切的關係。

本計畫將探討 A 型流感病毒 HA 蛋白的氨基酸序列特徵與感染對象、感染力的強弱以及跨物種感染可能性的關連，首先，我們將從分子演化的角度出發，對歷年來的 A 型流感病毒 HA 蛋白做分群與分類，其次，再針對其氨基酸序列的物理化學性質給予適當的量化，並針對量化後的特徵，做分群與分類。接著比較分析兩種分類與分群，並探討 HA 蛋白在感染過程中的關鍵氨基酸，最後探討這些關鍵氨基酸的變異對於 HA 蛋白與受器結合的可能影響。

研究目的

A 型流感病毒能夠感染多種物種，是最常出現的流行性感冒，1918 年感染人的流行性感冒，在全世界有二千萬到四千萬人死亡，造成如此高感染力的原因始終不明，一直到 James Stevens 和 Scripps 學院的 Ian Wilson 在 La Jolla 開始著手研究採自阿拉斯加受難者的屍體組織，他們找出病毒表面蛋白 hemagglutinin (HA) 的序列並分析其結構 (Woods *et al.*, 2004)，結果顯示，它跟感染鳥的流行性感冒病毒較接近較有關係。

對於人類流行性感冒病毒來說，HA 蛋白是最常被研究的部分，因為它結合到人類的特定細胞表面受器，並進行感染過程，HA 會改變形態幫助病毒的膜與宿主融合，整個過程在分子層次上是物理化學的交互作用。1918 年造成異常高的傳染性和死亡率的 HA 蛋白，在結構上有兩個特別突出的特徵異於一般：(1) 一般與受器結合位置的範圍是狹小的，然而 1918 年流行性感冒病毒的 HA 蛋白卻以一個氨基酸突變，使結合位置稍微的變大，這可能會增加傾向於對人類細胞的感染。(2) 有兩個之前未被查覺的鹼性小片段(histidine rich)，當病毒融入細胞膜時，可能會提高 HA 的感染性，其中一個小片段只有在感染鳥的病毒種類中才有 (Woods *et al.*, 2004, Gamblin *et al.*, 2004)。

根據這些線索，我們推論流行性感冒病毒的感染對象、感染力的強弱以及跨物種感染的可能性，應該與它的 HA 蛋白質序列片段所具有的物理化學性質，有很密切的關係。

本計畫為三年的研究計畫，探討 A 型流感病毒 HA 蛋白的氨基酸序列特徵與感染對象、感染力的強弱以及跨物種感染可能性的關連，首先，我們從分子演化的角度出發，對歷年來的 A 型流感病毒 HA 蛋白做分群與分類，其次，再針對其氨基酸序列的物理化學性質給予適當的量化，並針對量化後的特徵，做分群與分類。接著比較分析兩種分類與分群，以探討 HA 蛋白在感染過程中的關鍵氨基酸，再以結構生物資訊學的方法探討這些關鍵氨基酸的變異對於 HA 蛋白與受器結合的可能影響。最後，我們發展針對 HA 蛋白質序列的流感 A 行病毒疫苗設計。目前研究成果在於發展更精確的抗原決定位預測方法，將來將應用到流感病毒的疫苗設計上。

文獻探討

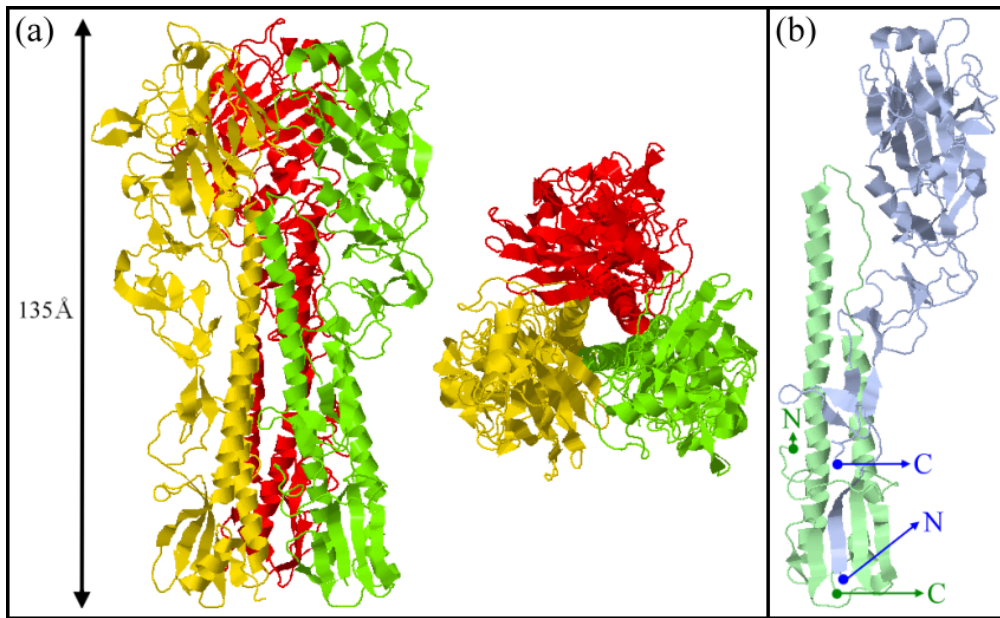
流行性感感冒病毒屬於正黏液病毒科(Family Orthomyxoviridae)，是一種具有外套膜的單股 RNA 病毒，由於它們都缺乏具有修正錯誤的 DNA 聚合酶，所以相較於 DNA 病毒，RNA 病毒具有很高的變異性，這種病毒的特徵是 RNA 可以分成很多段。流行性感感冒病毒依據其基質蛋白(matrix protein)及內部核糖核蛋白(ribonucleoprotein)的抗原特性分為 A、B、C 三種類型，其中 A 型流感病毒可傳染人類、豬、馬、海豹、鯨以及鳥類等，對豬而言可造成嚴重的壞死性肺炎。這個類型的流感病毒也是最常見最容易引起世界大流行的類型(Palese, 2004)，它總共有八段 RNA。

A 型流感病毒是變異最為頻繁的一個類型，每隔十幾年就會發生一個抗原性大變異，產生一個新的病毒株，A 型流感病毒的高變異性增加了人們應對流行性感冒的難度(Webby and Webster, 2003)，人們如果無法準確預測即將流行的病毒亞型，便不能有針對性地進行預防性疫苗接種。

流行性感感冒病毒的外層蛋白包括血球凝集素(hemagglutinin, HA)、神經胺酸酶(neuraminidase, NA)與蛋白 M2 三種，也是組成流行性感感冒抗原的成份。流行性感感冒病毒依據其血球凝集素(hemagglutinin)與神經胺酸酶(neuraminidase)的抗原性差異可以再分成數種亞型，現在已知的血球凝集素(hemagglutinin)一共有十五種(H1 到 H15)，神經胺酸酶(neuraminidase)則有九種(N1 到 N9)，在人體上發現的病毒中只有找到五種 HA(H1, H2, H3, H5, H9)和兩種 NA(N1, N2)，其中 H5 和 H9 為偶發個案。根據世界衛生組織在 1980 年通過的流行性感感冒病毒株命名法修正案，流感病毒株的命名一共包括六個要素：型別/宿主/分離地區/毒株序號/分離年份 (HnNn)，其中感染對象為人類的病毒則省略宿主訊息，對於 B 型與 C 型流感病毒則省略亞型訊息。例如 A/swine/Lowa/15/30 (H1N1)表示其核蛋白為 A 型，1930 年時在 Lowa 分離出來的，是以豬為宿主的 H1N1 亞型流感病毒毒株，其病毒株的序號是 15。

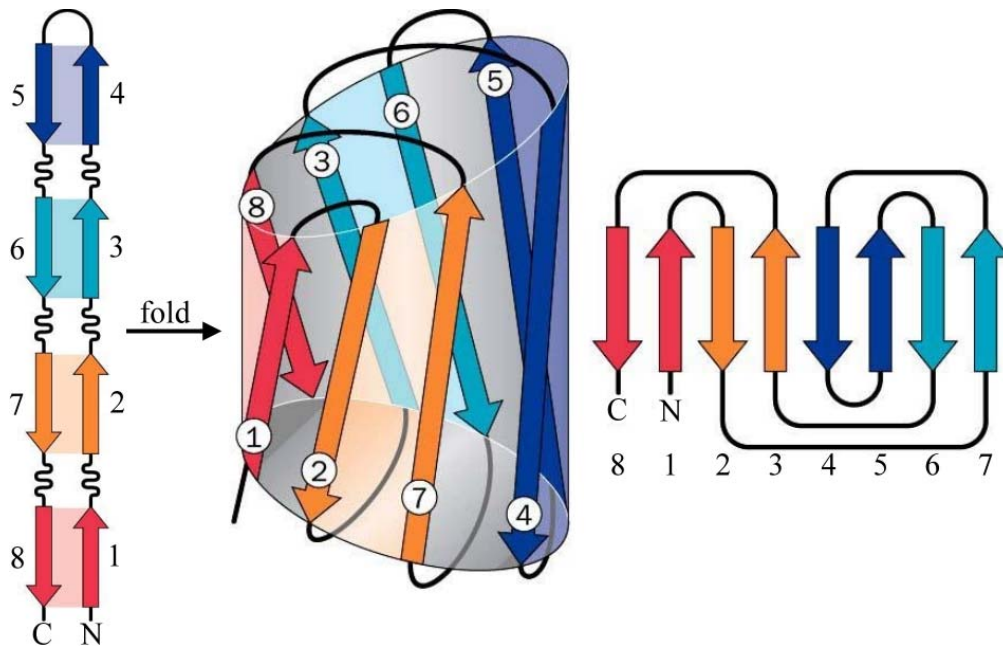
血球凝集素呈柱狀，能與人、鳥、豬、豚鼠等動物的紅血球表面受體相互結合引起凝血，所以被稱為血球凝集素。血球凝集素蛋白水解後分為輕鏈和重鏈兩個部分，重鏈可以與宿主細胞膜上的唾液酸(sialic acid)受體相互結合，唾液酸(sialic acid)受體的微酸性環境使得血球凝集素的結構轉為可產生膜融合的狀態，使病毒的內含遺傳物質得以進入細胞質內，輕鏈則可以協助病毒包膜與宿主細胞膜相互融合。血球凝集素在病毒導入宿主細胞的過程中扮演了重要角色，血球凝集素具有免疫原性，抗血球凝集素抗體可以中和流感病毒。隨著分子生物學的發展以及相關資料庫的建立，HA 蛋白序列的數目已經可以用來建立反應其演化過程的分子演化樹(Bush *et al.*, 1999, Fitch *et al.*, 1997, Holmes *et al.*, 2005)。

血球凝集素的 X-ray 結構已經由 Wiley 等人精確地定出來(Weis *et al.*, 1990, Wilson *et al.*, 1981)。HA 是一個長約 135Å 的圓柱狀三聚體，在 35Å 和 70Å 之間順著放射狀的方向變化。它由三個相似度很高的單體所組成，聚集在中央的 α 螺旋圍繞成捲狀(central α -helical coiled coil)，形成一個類似樹幹的區域，三個球狀的頭包含唾液酸(sialic acid)的結合位置，如下圖(Weis *et al.*, 1990, Wilson *et al.*, 1981)。



圖中，(a)由三個單體所組成的 HA 示意圖，其中三個單體分別由黃、綠、紅三個顏色所表示，(a)右邊的圖是 HA 頂部鳥瞰圖。(b)為 HA 其中一個單體 HA0，其中淡藍色為 HA1，淡綠色為 HA2。HA1 的 N 端與 C 端由藍色的英文字母與圓點所標示，HA2 的 N 端與 C 端由綠色的英文字母與圓點所標示。

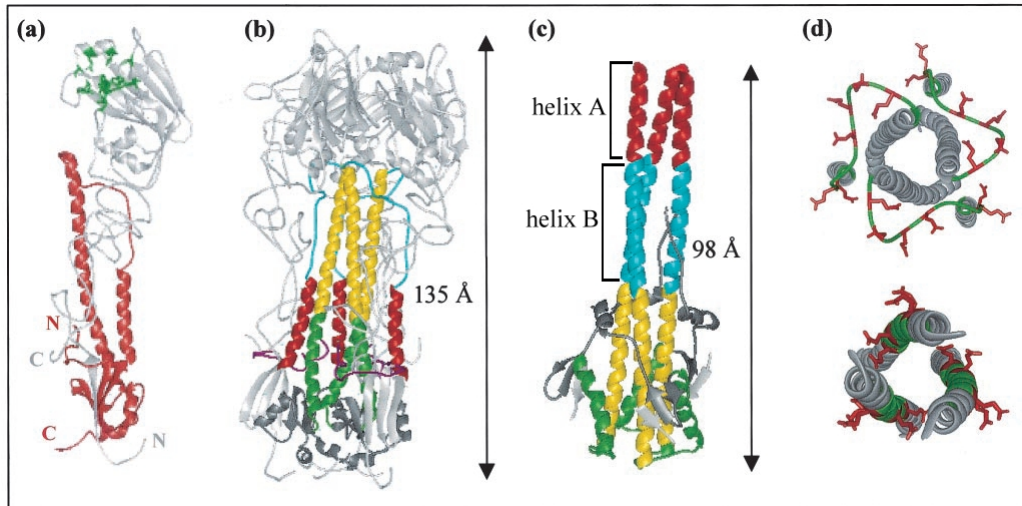
HA 的單體是由分裂成兩個 polypeptides(HA1 和 HA2)的無活性融合前驅蛋白 (fusion-inactive precursor)(HA0)所構成的。HA1 和 HA2 這兩個 polypeptide 之間藉由兩個單體內的雙硫鍵連接，這可能是 HA0 在內質網裡摺疊時形成的。無活性融合前驅蛋白分裂的兩端形成 HA1 的 C 端和 HA2 的 N 端，間隔 20Å。每一個單體在靠近各自 HA2 C 端的地方藉由 27 個氨基酸組成的穿膜螺旋 peptide 被固定在病毒的膜裡，這個區域在 x-ray 結構裡尚未顯示出來。HA 球狀的頭部是由 HA1 的第 116 到 261 個殘基摺疊成八個反平行 β -鏈的瑞士捲(jelly-roll)模樣，如下圖(Weis *et al.*, 1988)。



球狀的頭部包含高度保守性的桶狀袋形區域，這是屬於受器結合的位置，病毒利用此部分與宿主細胞受器蛋白的醣類側鏈唾液酸(sialic acid)殘基結合而引起感染(Bizebard *et al.*, 1995)。

病毒的感染需要 HA 產生胞膜融合作用，使病毒與宿主細胞膜融合。當病毒結合到細胞膜會使膜由外向內凹陷形成囊泡(vesicles)。內吞囊泡(endocytic vesicles)有質子泵(Proton pumps)會導致質子堆積和囊泡內 pH 值的降低，酸性的 pH 值(pH 值在 5 和 6 之間)會引起 HA 分子結構的外在形態改變，這會促進 HA 的融合功能(White *et al.*, 1987)。

當 HA 處於低 pH 值，其形態會產生大範圍的改變，這個改變需要球狀頭部優先開啟以避免空間上的干擾，如下圖。如此大規模的分子形態，應該是一個相當緩慢的過程，需要比融合過程還要長的時間(Shangguan *et al.*, 1998)。



圖中(a)為正常 pH 值狀態下的單體 HA1(灰色)、HA2(紅色)，以綠色表示的區域為受器結合位置。圖(b)同樣是正常 pH 值狀態下的 HA 三聚體結構，以紅色表示的區域為融合 peptide。圖(c)為在酸性 pH 值環境下，由 HA2 的第 38 到 175 個殘基以雙硫鍵連接到 HA1 的第 1 到 27 個殘基所組成的三聚體片段。圖(d)中 Glu 殘基(上圖中以紅色表示)在低 pH 值環境下，分子結構改變外在形態時會促使上圖中綠色部分纏繞 helix B 區域形成下圖的樣子。

在圖(b)與(c)中的 HA2 片段，紅色區域(helix A，第 38 到 55 個殘基)、藍色區域(helix B，第 56 到 75 個殘基)、綠色區域(helix D，第 106 到 129 個殘基)以及深灰色(helix E 到 H 之間，第 130 到 175 個殘基)，在酸性 pH 值環境下會有大量結構改變，而黃色區域(helix C，第 76 到 105 個殘基)則會維持它們原本的形態(Isin *et al.*, 2002)。

在生物學與分子生物學上，已經確定 HA2 上第 1 到 20 個位置的氨基酸，這是流感病毒 HA 的融合 peptide。這個融合 peptide 較不具極性的，含有較多的 glycines，而且形態上具有高彈性。在正常 pH 值之下的形狀，三個融合 peptide 的 N 端，隱藏於病毒外套雙層脂質膜下約 30Å。在酸性 pH 值的環境時，這個區域因 HA 分子上的形態改變而顯露出來，其疏水性促進了病毒外膜與宿主細胞膜的融合，這是整個感染過程的關鍵之一。

研究方法

本研究方法為本計畫第三年部分之研究方法。

蛋白質序列的抗原決定位是指會引起免疫反應的蛋白質片段(peptide)，抗原決定位可分為屬於 T 細胞與屬於 B 細胞兩大類，其中又可分為屬於 MHC I 與 MHC II 兩類，此部分我們研究的對象，主要是 T 細胞抗原決定位。T 細胞的抗原決定位相關資料庫頗多，例如 MHCPEP (Brusic *et al.*, 1998)、MHCBN (Bhasin *et al.*, 2003)、EPIMHC (Reche *et al.*, 2005)、AntiJen (Toseland *et al.*, 2005)、SFYPEITHI (Rammensee *et al.*, 1999)以及 IGMGT (Giudicelli *et al.*, 1997)，其預測的方法目前有多種被發展出來，其中以支援向量機為基礎的方法 (Donnes *et al.*, 2002) 有較高之準確性。

(1) MHCPEP 的抗原決定位資料

MHCPEP (Brusic *et al.*, 1997) 是一個 MHC class I 與 MHC class II 蛋白質片段的資料庫，創立於 1994 年。資料庫中的資訊經由已發表或直接透過實驗的數據取得，其抗原活性的分級如下表：

PD50 Value	
>10 μ M	None
10 μ M - 99 nM	yes, little
100 nM - 1 nM	yes, moderate
<1nM	yes, high

我們經由 POPI 伺服器 (Tung and Ho, 2007) 取得 MHC class I 與 MHC class II 的抗原片段分別為 428 與 226 條胜肽序列，四個類別的各別序列數如下表所示：

Immunogenicity class of peptides	MHC class I	MHC class II
None	144	45
Little	83	60
Moderate	100	64
High	101	57

(2) 抗原決定位序列的特徵擷取(feature extraction)

(i) 多序列排比與量化

以公用軟體 ClustalW 3.14 (Thompson *et al.*, 1997) 做多序列排比，每個氨基酸以 AAindex 的物理化學性質 (Kawashima *et al.*, 2007) 加以量化，並以 Pearson's coefficient 估計與免疫強度最相關的三個物理化學性質，並產生三個量化後之矩陣。

(ii) 數量之正規化

對於每一個矩陣，針對每一個排比後之序列位置(column)做數量的正規化。令 $A_i^{l,m}$ 代表物理化學性質 m 的矩陣中第 l 條序列的第 i 個位置的元素，則如下式做數量的正規化：

$$X_i^{l,m} = \frac{A_i^{l,m} - \min_l \{A_i^{l,m}\}}{\max_l \{A_i^{l,m}\} - \min_l \{A_i^{l,m}\}}$$

(iii) 模糊測度與 Choquet 積分 (Chang *et al.*, 2010)

對每一個排比後之序列位置(column)計算每個物理化學性質的模糊測度。對於物理化學性質 m 的矩陣中第 i 個位置(column)，我們以下式來定義其模糊測度：

$$\nu(\{X_i^m\}) = \frac{1 + \text{corr}(X_i^m, Y)}{2}$$

其中 $\text{corr}(X_i^m, Y)$ 為向量 X_i^m 與 Y 的 Pearson 相關係數， $Y = (Y^1, Y^2, \dots, Y^k)'$ 為屬性向量， k 為資料集中之蛋白質片段序列條數。若第 l 條之蛋白質片段序列會引起免疫反應，則令 $Y^l = 1$ ，反之則令 $Y^l = 0$ 。定出每個位置每個物理化學性質以及物理化學性質之間共同的模糊測度後，便能以下式計算每一條序列中每個氨基酸本身三種物理化學性質之 Choquet 積分值。

$$C_\nu(x) = \sum_{i=1}^n x_{(i)} [\nu(A_{(i)}) - \nu(A_{(i+1)})]$$

其中 $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ ； $n=4$ 為正規化後三種物理化學性質值的大小排序， ν

為模糊測度， $A_{(i)} = \{x_{(i)}, \dots, x_{(n)}\}$ 且 $A_{(n+1)} = \phi$ 。

(iv) 序列特徵向量的定義

根據每一條蛋白質片段序列的氨基酸計算出的 Choquet 積分值數列，即是此蛋白質片段序列之特徵向量，其維度即為多序列排比後之序列長度，可視為在高維空間中的一個資料點。

(3) 支援向量機 (SVM)

對於已知屬性資料點 $(\underline{x}_i, y_i), i=1, 2, \dots, N$ ，其中 $\underline{x}_i \in R^n, y_i \in \{1, -1\}, i=1, 2, \dots, N$ ， \underline{x}_i 即為第 i 條蛋白質片段序列之特徵向量， y_i 則為此蛋白質片段序列之屬性， N 為蛋白質片段資料集總條數，以支援向量機針對屬性 y_i 來對 \underline{x}_i 分類必須滿足下列要件 (Boser *et al.*, 1992; Cortes *et al.*, 1995; Vapnik *et al.*, 1995; Chang *et al.*, 2004;)：

$$\begin{aligned} \min_{\underline{w}, b, \xi} \quad & \frac{1}{2} \underline{w}' \underline{w} + c \sum_{i=1}^N \xi_i \\ \text{subject to} \quad & y_i (\underline{w}' \phi(\underline{x}_i) + b) \geq 1 - \xi_i, \\ & \xi_i \geq 0, \\ \text{where} \quad & b, c \in R, \underline{w}, \phi(\underline{x}_i) \in R^m \\ & \phi: R^n \rightarrow R^m \end{aligned}$$

則對於未知屬性的資料點 $\underline{x}_i \in R^n$ ，可計算下式來預測其屬性 y_i ：

$$\begin{aligned} d(\underline{x}_i) &= [\underline{w}' \phi(\underline{x}_i) + b - (1 - \xi_i)] \\ y_i &= \begin{cases} +1, & \text{if } d(\underline{x}_i) \geq 0 \\ -1, & \text{if } d(\underline{x}_i) < 0 \end{cases} \end{aligned}$$

我們以 LOOCV 來評估分類的效果。

結果與討論

本研究結果為本計畫第三年部分之研究結果。

我們在結果評估上，應用了精確度百分比(percentage accuracy;ACC)、馬修斯相關係數(Matthew's correlation coefficient; MCC)、總體精確度(overall accuracy;OA)以及平均精確度(averaged accuracies;AA)，並與 POPI 等其他方法作比較：

$$ACC_i = \frac{TP_i}{TP_i + FN_i} * 100\%$$

$$MCC_i = \frac{TP_i * TN_i - FP_i * FN_i}{\sqrt{(TP_i * FN_i) * (TP_i * FP_i) * (TN_i * FP_i) * (TN_i * FN_i)}}$$

$$OA = \sum \frac{TP_i}{N}$$

$$AA = \sum \frac{ACC_i}{h}$$

MHC class I 的結果如下表，顯示我們的方法仍有改進的空間，而有無模糊測度與積分的結果相當，這表示每個氨基酸的物理化學性質之間，沒有明顯的交互作用，或者使用的測度無法描述其交互作用。

class	ALIGN		PSI-BLAST		POPI		未經模糊測度與 Choquet 積分		經模糊測度與 Choquet 積分(P)	
	ACC(%)	MCC	ACC(%)	MCC	ACC(%)	MCC	ACC(%)	MCC	ACC(%)	MCC
0	69.44	0.61	82.14	0.59	83.33	0.63	81.94	0.47	75.69	0.5
1	39.76	0.32	45.59	0.40	50.60	0.44	31.33	0.36	39.76	0.32
2	39.00	0.22	34.67	0.12	55.00	0.47	32	0.18	32	0.18
3	62.38	0.37	46.99	0.37	59.41	0.49	56.44	0.45	60.4	0.48
OA	54.91		53.23		64.72		54.44		54.91	
AA	52.64		52.35		62.09		50.43		51.96	

MHC class II 的結果如下表，顯示有無模糊測度與積分的結果有明顯差異，這表示每個氨基酸的物理化學性質之間，有明顯的交互作用。

Immunogenicity class	未經模糊測度與 Choquet 積分		經模糊測度與 Choquet 積分 (Lambda)		經模糊測度與 Choquet 積分 (P)		經模糊測度與 Choquet 積分 (L(0,1))	
	ACC(%)	MCC	ACC(%)	MCC	ACC(%)	MCC	ACC(%)	MCC
None	64.44	0.62	64.44	0.57	66.67	0.68	66.67	0.69
Little	48.33	0.35	38.33	0.23	51.67	0.29	48.33	0.32
Moderate	56.25	0.31	54.69	0.32	51.56	0.32	51.56	0.32
Hight	64.91	0.51	68.42	0.55	64.91	0.53	71.93	0.53
OA		57.96		55.75		57.96		58.85
AA		58.49		56.47		56.47		59.62

未來我們將進一步探討氨基酸之間的耦合效應，也將試著尋找其他更有效的測度。

参考文献

- Bizebard, T., Gigant, B., Rigolet, P., Rasmussen, B., Diat, O., Boesecke, P., Wharton, S.A., Skehel, J.J and Knossow, M. (1995) Structure of influenza virus haemagglutinin complexed with a neutralizing antibody. *Nature*, **376**, 92–94.
- Brusic, V., Rudy, G., Kyne, A.P., and Harrison, L.C. (1997) MHCPEP, a database of MHC-binding peptides:update 1996. *Nucleic Acids Research*, **25**.
- Bush, R.M., Fitch, W.M., Bender, C.A. and Cox, N.J. (1999) Positive selection on the H3 hemagglutinin gene of human influenza virus A. *Mol. Biol. Evol.*, **16**, 1457–1465.
- Chang, H.J., Chang, P.C. Lee, K.J, Wang, J.D., and Liu. Y.L. (2010) Choquet Integral Algorithm For Thermostable Proteins Based On Huret Exponent And Generalized L-measure. *ICMLC*.
- Fitch, W.M., Bush, R.M., Bender, C.A. and Cox, N.J. (1997) Long term trends in the evolution of H(3) HA1 human influenza type A. *Proc. Nat. Acad. Sci.*, **94**, 7712–7718.
- Gamblin, S.J., Haire, L.F., Russell, R.J., Stevens, D.J., Xiao, B., Ha, Y., Vasisht, N., Steinhauer, D.A., Daniels, R.S., Elliot, A., Wiley, D.C. and Skehel, J.J. (2004) The Structure and Receptor Binding Properties of the 1918 Influenza Hemagglutinin. *Science*, **303**, 1838-1842.
- Holmes, E.C., Ghedin, E., Miller, N., Taylor, J., Bao, Y., George, K.S., Grenfell, B.T., Salzberg, S.L., Fraser, C.M., Lipman, D.J. and Taubenberger, J.K. (2005) Whole-Genome Analysis of Human Influenza A Virus Reveals Multiple Persistent Lineages and Reassortment among Recent H3N2 Viruses. *PLoS BIOLOGY*, **3**, 1579-1589.
- Isin, B., Doruker, P. and Bahar, I. (2002) Functional Motions of Influenza Virus Hemagglutinin: A Structure-Based Analytical Approach. *Biophysical Journal*, **82**, 569–581.
- Kawashima, S., Pokarowski, P., Pokarowska, M., Kolinski, A., Katayama, T., and Kanehisa, M. (2008) AAindex:amino acid index database, progress report 2008. *Nucleic Acids Research*, **36**.
- Palese, P. (2004) Influenza: old and new threats. *Nat. Med.*, **10**, s82–s87.
- Shangguan, T., Siegel, D.P., Lear, J.D., Axelsen, P.H., Alford, D. and Bentz, J. (1998) Morphological changes and fusogenic activity of influenza virus hemagglutinin. *Biophys. J.*, **74**, 54–62.
- Tung, C.W. and Ho, S.Y. (2007) POPI:predicting immunogenicity of MHC class I binding peptides by mining informative physicochemical properties. *Bioinformatics*, **23**.
- Webby, R.J. and Webster, R.G. (2003) Are we ready for pandemic influenza? *Science*, **302**, 1519–1522.

Weis, W.I., Brown, J.H., Cusack, S.C., Paulson, J.C., Skehel, J.J. and Wiley, D.C. (1988) Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. *Nature*, **333**, 426–431.

Weis, W.I., Brünger, A.T., Skehel, J.J. and Wiley, D.C. (1990). Refinement of the influenza virus hemagglutinin by simulated annealing. *J. Mol. Biol.*, **212**, 737–761.

White, J.M. and Wilson, I.A., (1987) Anti-peptide antibodies detect steps in a protein conformational change: low-pH activation of the influenza virus hemagglutinin. *J. Cell Biol.*, **105**, 2887–2895.

Wilson, I.A, Skehel, J.J. and Wiley, D.C. (1981) Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3Å resolution. *Nature*, **289**, 366–373.

Woods, H.R., Stevens, J. and Wilson, I.A. (2004) Resurrecting the Dead and the Deadly. http://www-ssrl.slac.stanford.edu/research/1918flu_summary.html

計畫成果自評

本計畫已經發展出一個抗原決定位的預測方法，並將應用於流感病毒疫苗的預測。

本計畫並未如預期完成全部三年的研究，主要原因在於 MCMC 的計算量十分龐大，研究生之訓練與相關知識的學習進度亦不如預期。

國科會補助專題研究計畫項下出席國際學術會議心得報告

日期：99年10月30日

計畫編號	NSC 98-2221-E-468 -021 -		
計畫名稱	流感 A 型病毒 HA 蛋白的演化與變異		
出國人員姓名	張培均	服務機構及職稱	亞洲大學生物資訊學系
會議時間	98年9月7日至 98年9月11日	會議地點	新加坡
會議名稱	(中文) 2009 生物資訊學國際研討會 (英文) International Conference on Bioinformatics (InCoB 2009)		
發表論文題目	(中文) 利用波性特徵與非線性指標偵測來癲癇 (英文) Detection of Epilepsy Events in The EEG by Mimetic Signatures and Nonlinear Dynamic Indications		

一、參加會議經過

本次會議共有五天，討論的議題有 structural bioinformatics、genome analysis、systems biology、sequence analysis、genetic and population analysis、transcriptome analysis、industry track、biomedical informatics、text mining、sequence and structural bioinformatics、emerging areas 以及 interactome analysis，我發表的論文屬於 biomedical informatics 領域，被排在第三天，由於國內尚有工作待完成，因此，我是在會議的第二天才前往註冊，聽了最後幾場演講後，因飛行途中身體略感不適，因此離開會場回到住宿的飯店，第二天一早，我便前往會場做準備，會場在捷運站附近，是一棟新大樓，生活機能完整，因此十分方便，會議討論的議題十分豐富，很有前瞻性。由於國內尚有工作待完成，因此，我在會議的最後一天中午便前往機場搭機返國，

傍晚抵達國門。

二、與會心得

新加坡的學生學習態度積極，英文聽講能力非常好，國際化程度很高，我想這是他們所以很有競爭力的原因。新加坡對於生物資訊這個領域十分重視，研究人員亦十分積極，我在參與會議期間，經常看到研究人員熱烈參與討論的情景，很值得我們深思與效法。

三、考察參觀活動(無是項活動者略)

無.

四、建議

發展生物資訊的成本遠比發展生物科技來得少，台灣目前的資源有限，生物資訊這一富有創意的領域，政府更應先行投資並扶助其發展，建議仿效新加坡模式，禮聘國際上此一領域的專家學者，來參與我國生物資訊的發展方向規劃與人才培育，整合可用的國內外資源，並國際接軌。

五、攜回資料名稱及內容

大會手冊.

六、其他

無.

國科會補助專題研究計畫項下出席國際學術會議心得報告

日期：99年10月30日

計畫編號	NSC 98-2221-E-468 -021 -		
計畫名稱	流感 A 型病毒 HA 蛋白的演化與變異		
出國人員姓名	張培均	服務機構及職稱	亞洲大學生物資訊學系
會議時間	99年7月11日至 99年7月14日	會議地點	中國青島
會議名稱	(中文) 2010 人工智慧與機器學習國際研討會 (英文) International Conference of Machine Learning and Cybernetics (ICMLC 2010)		
發表論文題目	(中文) 基於賀氏指數與泛型 L 測度的熱穩定性蛋白積分演算法 (英文) Choquet Integral Algorithm for Thermostable Proteins Based on Hurst Exponent and Generalized L-measure		

一、參加會議經過

本次會議共有四天，討論的議題集中在人工智慧與機器學習，與生物資訊相關的議題很少，討論議題主要集中在非生物領域，我參與發表的論文被排在第二天。此行在研究工作上的主要收穫，則在於大會中許多資料探勘的方法，未來很有潛力可應用在生物資訊上。我在會議的最後一天便前往機場搭機返國，由於轉機香港，翌日凌晨才抵達國門。

二、與會心得

中國大陸的進步令人驚訝，其城市發展，學生素質，甚至連舉辦國際會議的品

質，感覺上都已超越台灣，我們若不提升，將來勢必失去競爭力，國人應虛心面對，積極奮鬥。

三、 考察參觀活動(無是項活動者略)

無.

四、 建議

建議政府增加大陸來台學生，使我國學生能對中國大陸有更充分的瞭解，如此，未來可提升我國學生的競爭力。

五、 攜回資料名稱及內容

大會論文集.

七、 其他

無

無衍生研發成果推廣資料

98 年度專題研究計畫研究成果彙整表

計畫主持人：張培均		計畫編號：98-2221-E-468-021-					
計畫名稱：流感 A 型病毒 HA 蛋白的演化與變異							
成果項目		量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）	
		實際已達成數（被接受或已發表）	預期總達成數（含實際已達成數）	本計畫實際貢獻百分比			
國內	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	1	0	100%		
		研討會論文	1	0	50%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	3	0	100%	人次	
		博士生	1	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	0	1	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	2	1	50%		
		專書	0	0	100%		章/本
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

<p>其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)</p>	<p>無</p>
--	----------

	成果項目	量化	名稱或內容性質簡述
科 教 處 計 畫 加 填 項 目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與(閱聽)人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

達成目標

未達成目標（請說明，以 100 字為限）

實驗失敗

因故實驗中斷

其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文： 已發表 未發表之文稿 撰寫中 無

專利： 已獲得 申請中 無

技轉： 已技轉 洽談中 無

其他：（以 100 字為限）

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

本計畫之研究成果為原計畫提案中之第三年部分，前二年之部分由於 MCMC 的計算量龐大，目前逐步進行中，目前的研究結果，未來可利用於流感病毒疫苗的設計。