

計畫編號：MOHW106-CDC-C-315-114719

衛生福利部疾病管制署 106 年署內科技研究計畫

計畫名稱：建立深度分析流感病毒基因序列與病毒特性關連之平台

106 年 度/全 程 研 究 報 告

執行單位：檢驗及疫苗研製中心

計畫主持人：劉銘燦

協同主持人：楊季融

研究人員：吳馥婷、邱之柔

執行期間：106 年 1 月 1 日至 106 年 12 月 31 日

目錄

	頁 碼
目錄	1
計畫中文摘要	2
計畫英文摘要	3
計畫內容	
一、前言	4
二、材料與方法	5
三、結果	8
四、討論	11
五、結論與建議	13
六、計畫重要研究成果及具體建議	13
七、參考文獻	14
八、圖、表	15

共 (23) 頁

計畫中文摘要：

流感病毒嚴重威脅人類健康，流感病毒變化快易產生新變異株，造成病毒抗原飄移、抗藥性與致病力的改變，而抗原飄移導致每年疫苗株組成需調整，抗藥性發生則須改變治療藥物，而致病力的改變造成死亡人數增加。所以瞭解流感病毒在演化和流行病學過程如何造成抗原飄移、抗藥性與致病力等的改變非常重要。病毒繁殖時突變造成病毒歧異度增加，而有夸種(quasispecies)的出現，quasispecies 的動態變化影響病毒抗原飄移、抗藥性與致病力，但低比例 quasispecies 無法以傳統的基因定序法檢測。近年來新的 DNA 定序技術，具高通量(high-throughput)特性，相對於傳統 Sanger 定序法又稱為次世代定序 (next-generation sequencing, NGS) 簡稱 NGS，NGS 可提供大量的 DNA 序列資訊，深度解析基因體，可發現微生物基因體序列低比例的變化，故 NGS 為研究病毒微量 quasispecies 動態變化的強力工具。本計畫擬建立流感病毒 NGS 平台，以 NGS 為工具，研究(1)流感病毒在原始臨床檢體與分離病毒株基因序列差異，(2) 流感病毒 quasispecies 基因序列動態變化，(3) 從個體層次分析流感病毒感染後，不同時期(早、後期)與不同部位(上、下呼吸道)基因序列之變化，瞭解流感病毒在單一宿主內的演變，(4)比較不同流感季之流感病毒 quasispecies 變化，研究當年的主流病毒株基因特徵是否在前季 quasispecies 變異株已存在，瞭解流感病毒變異株出現的機制，(5)分析抗流感藥物使用前後病毒 quasispecies 基因序列的變化，瞭解流感病毒抗藥性出現的機制，(6)比較不同亞型/次亞型流感 quasispecies 變化是否有差異。希望本計畫的結果對流感病毒的演化、適應、抗原飄移、抗藥性與致病力的變化更進一步瞭解，以更有效地防治流感。

關鍵詞：流感病毒、病毒演化、跨種、病毒歧異、次世代定序

計畫英文摘要：

Influenza viruses pose a serious threat to human health. The viruses with high mutations can generate new variants, resulting in the changes of viral antigenic drift, antiviral resistance and virulence. The antigenic drift makes the composition of the vaccine strains to be reformulated each year. When antiviral resistance occurs, the kinds of antivirals shall be changed and the increase of viral virulence causes increased deaths. High mutations in virus propagation make the increase of viral divergence, which contribute to formation of viral quasispecies. The dynamic changes of quasispecies influence the changes of antigen drift, resistance and virulence, however, but the low proportion of quasispecies is unable to be detected with conventional sequencing method. In recent years, new DNA sequencing technology with high throughput, compared to conventional Sanger sequencing method, is termed as next-generation sequencing (NGS). NGS can generate an extraordinary number of sequences to analyze genomics deeply and detect low frequency changes in microbial gene sequences. Therefore, NGS is a powerful tool to study the dynamics of virus quasispecies. In this project, we intend to establish NGS platform for influenza viruses, and address the issues including (1) sequence differences between viruses in the original clinical sample and isolates in cell culture, (2) dynamics of quasispecies, (3) sequence differences between viruses collected in early and late stages and in different location (upper and lower respiratory tract) in the individual to understand evolution of the influenza virus in an individual host, (4) comparison of quasispecies dynamics in different influenza seasons to determine whether predominant strain in this season has existed in the quasispecies of previous seasons, which uncovers the mechanism of variant emergence, (5) sequence differences between viruses collected before and after the treatments of antivirals to understand the development of resistant influenza viruses (6) comparison of quasispecies dynamics in the different types/subtypes of influenza viruses. The results obtained in this project will advance our understanding on influenza virus evolution, adaptation, antigenic drift, antiviral resistance and virulence, which make prevent and control influenza effectively.

keywords : Influenza; virus evolution; quasispecies; viral diversity; next-generation sequencing, NGS

本文

一、前言：

流感病毒嚴重威脅人類健康，每年流感全球侵襲率估計在兒童約20-30%、成年人約5-10%，主要導致嬰兒、老人或慢性病患等高危險群者住院和死亡，估計每年約造成3~5百萬例嚴重病例，約有25-50萬人死亡(WHO, Fact sheet on influenza, <http://www.who.int/mediacentre/factsheets/fs211/en/>)。流感病毒，屬於正黏液病毒科(*Orthomyxoviridae*)，病毒基因體為8個負單股RNA，病毒產生10-14個蛋白質[9, 木]。流感病毒突變率高易產生新變異株，造成病毒歧異度變大，形成quasispecies [6]，但流感病毒感染個體時因瓶頸現象造成選擇少數病毒繼續複製並釋出[4, 8, 13]，流感病毒基因的變化為經過多次的感染循環:感染時病毒歧異度窄化而釋出時歧異度增加，使得少數株變為主流株。故quasispecies中各族群的消長決定病毒抗原飄移、抗藥性與致病力的改變[6]。而抗原飄移導致每年疫苗株組成需調整，抗藥性發生則須改變治療藥物，而致病力的改變造成死亡人數增加。過去quasispecies中少數株無法以傳統的基因定序法檢測，但近年來新的DNA定序技術，具高通量(high-throughput)特性，相對於傳統Sanger定序法又稱為次世代定序(next-generation sequencing, NGS)簡稱NGS，NGS可提供大量的DNA序列資訊，深度解析基因體，可發現微生物基因體序列低比例的變化，且不用特定primers的特性，可用於決定未知微生物的基因序列[1]。NGS技術在2005年有Roche 454 FLXpyrosequencing定序平台(<http://www.454.com/>)，後來在2007年，Illumina公司SOLEXA Genome Analyzer (<http://www.illumina.com>)，Applied Biosystems SOLiD (<http://www.appliedbiosystems.com>)，Life Technologies 的 Ion Torrent PGM (<http://www.iontorrent.com/>)和Pacific Biosciences公司單分子即時定序(single molecule real-time sequencing, SMRT) (<http://www.pacificbiosciences.com/>)。這些NGS方法具有不同的特性如每次read序列的長短、序列的正確性等差異。在微生物學和病毒學NGS方法典型應用中，除了高通量全基因組定序，發現新微生物和病毒，NGS可以應用到

metagenomic 策略用於檢測未知疾病相關病毒和發現新的病毒等與過去利用微陣列的方法，NGS 提供更高的靈敏度檢測的新病毒[3, 5, 7, 11]。目前 NGS 被應用在研究流感病毒分子流行病學，病毒歧異度變化[12]傳播與致病性，如利用 NGS 研究流感病毒的演變與流感病毒在不同物種間的適應變化例，如原先少數 NA 缺失的變異株因在家禽適應強而成主流病毒[2]，在特定地理區域不同的人類流感病毒的流行和蔓延的模式[4]。

為了瞭解流感病毒演化和流行病學過程中抗原漂移(antigenic drift)、抗藥性、主流病毒株在個體與族群中出現的變化，本計畫擬建立流感病毒次世代(NGS)基因定序平台，並使用 NGS 分析比較流感病毒在原始臨床檢體與分離病毒株基因序列差異，流感病毒 quasispecies 基因序列動態變化，分析從個體層次感染流感病毒後，不同時期(早、後期)與不同部位(上、下呼吸道)基因序列之變化。使用 NGS 比較不同年之流感病毒 quasispecies 變化，研究當年的主流病毒株基因特徵是否在以前 quasispecies 變異株已存在，瞭解流感病毒變異株出現的機制。使用 NGS 分析使用抗流感藥物前後病毒 quasispecies 變異株基因序列的變化。比較不同亞型/次亞型流感 quasispecies 變化。希望本計畫的進行對流感病毒的演化、適應、抗原飄移、抗藥性與致病力的變化更進一步瞭解，以更有效地防治流感。

二、 材料與方法

(一) 檢體與病毒株收集

本計畫檢體來源為通報流感併發重症、上呼吸道群聚、新型 A 型流感及不明原因肺炎等個案檢體。流感核酸檢測 real-time RT-PCR 之步驟與引子、探針序列，以及病毒培養方法，詳見先前文獻[15]。

藉由上述檢驗陽性之臨床檢體，本研究從中挑選作為建置 NGS 分析平台之樣本，挑選的條件如下：

1. 該原始檢體可藉由 Sanger sequencing 完成病毒基因定序，以確保檢體中的病毒

量不至於過低。

2. 該檢體具有經細胞培養後所分離的病毒株，如此可直接比較原始檢體以及病毒分離株的 quasispecies 趨勢。
3. 挑選檢體時將參考原始檢體的 Sanger 定序結果，於各病毒 genetic clade 挑選，以利分析的病毒標的多元化。

(二) 流感病毒基因序列的分析流程

為能即時分析大量流感資料與易於監測流感病毒的變化，發展分析流程如下：

1. 將流感病毒氨基酸基因序列，使用 Clustal W 軟體作 alignment。
2. 使用 Bio-Eidt 將 alignment 之序列，調整 gaps 與切齊各序列的長度，以利後續之分析。
3. 使用 Amino acid variation 軟體，分析各位點變化的情形，選擇變化值大之位點，進行 proteotyping 之分析。
4. 使用 Excel 之巨集功能，將各種氨基酸以不同顏色代表，使各位點氨基酸一目了然，完成 proteotyping 之分析。此分析易於觀察其氨基酸位點變化情形，並依採檢日或發病日觀察這些氨基酸位點變化情形。
5. 將流感病毒基因的序列以 MEGA5 軟體 [10]分析其親緣關係。

(三) 流感病毒株培養

將臨床檢體接種於 MDCK (Madin-Darby canine kidney cell) 細胞，分離流感病毒。MDCK 細胞以 EMEM 培養基(內含 10% 胎牛血清)於 34°C，5% CO₂ 下繼代培養。

(四) 病毒與抗血清價位測定

本研究利用紅血球凝集試驗測定病毒量；以紅血球凝集抑制試驗測定抗血清力價，評估病毒抗原性。

1. 紅血球凝集試驗

- (1) 取 U 形底的 96 孔盤，於第二列至第八列加入 50 uL 的 PBS 溶液，於第一列加入 100 ul 的病毒抗原原液，negative control 行則以 100 uL PBS 取代抗

原。

- (2) 取第一列的抗原 50 uL 加入第二列，以微量吸管充份混合後，再取 50 uL 加入第三列，如此序列稀釋至第八列，抗原呈現 2 倍~128 倍稀釋。
- (3) 每孔分別加入 50 ul 的天竺鼠紅血球 (0.75%)，以手輕微搖晃孔盤後，之後以膠膜封住孔盤，置於室溫或 4 °C 下靜置 30—60 分鐘，之後觀察血球凝集，記錄病毒價位。

2. 紅血球凝集抑制試驗

- (1) 進行血球凝集抑制試驗前，須先以 PBS 溶液稀釋抗原原液至每 50 uL 稀釋液中含有 8 HA unit 的抗原。
- (2) 取 U 形底的 96 孔盤，於第二列至第八列加入 25 uL 的 PBS 溶液。於第一列加入 50 uL 的各標準病毒株的標準抗血清，negative control 行則以 25uL PBS 取代抗血清。取第一列的抗體 25 uL 加入第二列，以微量吸管充份混合後，再取 25 uL 加入第三列，如此序列稀釋至第八列。抗血清呈現 2 倍~128 倍稀釋。抗血清須經 RDE 處理以去除非專一性凝集。
- (3) 分別加入 25 uL (8 HA unit/50 uL) 的待測抗原及標準抗原，以手輕微搖晃孔盤後，置於室溫下反應 10—15 分鐘。
- (4) 加入以 PBS 稀釋的 0.75% 的天竺鼠紅血球 50 uL/well，之後以膠膜封住孔盤，至於室溫或 4°C 下靜置 30—60 分鐘，之後記錄抗血清價位結果。

(五) 病毒 RNA 萃取及反轉錄

病毒 RNA 萃取利用自動核酸萃取系統 MagNA Pure Compact Instrument (Roche Applied Science) 進行樣品核酸萃取，取 5 µL 萃取之核酸，利用隨機核苷酸(random octamer) 進行反轉錄反應，合成第一股 cDNA (first strand cDNA)：核酸與引子於 65°C 作用 5 分鐘後，置於冰上，再利用 PrimeScript RTase reverse transcriptase 進行反轉錄反應，反應條件為先 30°C 作用 10 分鐘後，再次 50°C 作用 60 分鐘，最後 95°C 作用 5 分鐘。

(六) NGS 平台

本研究使用 Illumina MiSeq sequencer 或 Pacific Biosciences 公司 SMRT 平台進行流感病毒序列分析，並比較與傳統方法序列差異。不同平台詳細步驟依廠商提供步驟進行。

(七) NGS 資料分析

本研究使用 CLC Genomics Workbench version 7.0.3 (CLC Bio, Qiagen)軟體進行資料分析比對，本年除規劃將待測病毒的全基因體完成組裝外，亦後續進行病毒 quasispecies 的分析，比較流感病毒在原始臨床檢體與分離病毒株基因序列差異，分析流感病毒 quasispecies 基因序列動態變化。

三、 結果

本年度為四年計畫執行的第一年，截至 106 年 11 月止，研究成果敘述如下：

(一) 檢體收集、檢驗與挑選

106 年 1 月至 11 月間，共自通報流感併發重症、新型 A 型流感及呼吸道群聚等檢體中，檢出 1871 件流感陽性檢體，其中 71 件為 A(H1N1)pdm09 陽性、1657 件為 A(H3N2) 陽性、5 件為 A(H7N9)陽性(同一個案)、138 件為 FluB 陽性。後續自這些陽性檢體經培養後的病毒中，第一階段先挑選 12 株作為後續建立流感病毒次世代全基因體定序平台之測試材料，其中包含 6 株 A(H1N1)pdm09、5 株 A(H3N2)以及 1 株 A(H7N9)病毒，並以定序後之序列片段進行組裝及後續分析，確認該平台之可運作性，並同時評估病毒不同基因片段是否有定序效率的差異，優化整體組裝及分析程序。

依實驗設計，分析程序完成建立後，第二階段將再挑選 18 對樣本，包含 18 支 A(H3N2)病毒陽性臨床檢體，以及該 18 支檢體經由細胞培養後所增殖的病毒株，利用第一階段建立好的 NGS 定序流程，分析其全基因體序列，並比較病毒於細胞培養前後 quasispecies 的變異情形。

(二) 建立並優化流感病毒次世代基因定序平台

本研究本年已完成建立流感病毒次世代基因定序分析流程，包括三大部分：(1)以

RT-PCR 進行樣本流感病毒全基因片段放大；(2)以 RT-PCR 放大後的基因片段產物製備次世代定序序列庫(library)；(3)將次世代定序序列庫進行定序以及後續組裝分析。各部分的成果如下：

1. 以 RT-PCR 進行樣本流感病毒全基因片段放大

經參考文獻及自行比對流感病毒基因序列結果，本研究已設計單一引子對，規劃以該引子對於單一、單管 RT-PCR 反應，將原始臨床檢體或經細胞培養過後的流感病毒全基因體，包括 PB2、PB1、PA、HA、NP、NA、MP 以及 NS 等 8 個片段增殖放大。依據第一階段所選取 6 株 A(H1N1)pdm09、5 株 A(H3N2) 以及 1 株 A(H7N9)等 12 株病毒所放大的基因體產物結果(圖一)，顯示該引子對可成功於單次 RT-PCR，同步增殖 A 型流感病毒之全基因體基因片段，且可偵測對象涵蓋多種病毒亞型(H1, H3, H7)，這些產物將作為後續製備次世代定序序列庫的起始材料。

2. 以 RT-PCR 放大後的基因片段產物製備次世代定序序列庫(library)

本階段規劃以製備 DNA 序列庫的方式進行。方法建立所選用的試劑包括 (1) NEBNext dsDNA Fragmentase，作為起始 DNA 片斷化(fragmentation)之用、(2) NEBNext Ultra II DNA Library Prep Kit，將次世代定序儀分析所需之 adaptor 接上片斷化後的 DNA，並進行初步放大、以及(3) NEBNext Multiplex Oligos，將不同序列的 index 置入各序列庫，使各序列庫可於同一次世代定序反應中進行，並於後續分析時作進一步區分。

利用各試劑原廠所建議方法並作些許修改後，已成功完成第一階段所選 12 株 A 型流感病毒各別全基因體次世代定序序列庫的製備(圖二)，電泳膠的結果顯示，各序列庫的產物大小主要介於 250~300 bp 之間。將該 12 個各別序列庫以等量方式混合後，經 Bioanalyzer 儀器進行品質分析結果如圖三，結果顯示製備完成且混和後的序列庫 DNA 片段大小約在 200~500 bp 之間，且分布趨勢的波峰集中於 305 bp，此結果與 NEBNext Ultra II DNA Library Prep Kit 所提供

的預期結果相符，亦達後續以 Illumina Miseq 定序儀進行次世代定序分析的上機標準。

3. 將次世代定序序列庫進行定序以及後續組裝分析。

本年次世代定序序列庫的定序係以 Illumina Miseq 定序儀進行。定序完成後的 contig 組裝及後續分析則使用 CLC Genomics Workbench version 7.0.3 (CLC Bio, Qiagen)軟體進行。經測試組裝 12 個序列庫所得序列，組裝時同步採用 De novo assembly 以及 resequencing 的方式，可有效將各序列庫的 contig 完成全基因體組裝，各病毒全基因體組裝後的序列經與 NCBI 資料庫比對，顯示過程並無錯誤黏合，故以 CLC 軟體進行組裝分析程序已完成建立，後續針對其他樣本的 NGS 定序將以此流程進行。

(三) 使用次世代定序分析流感病毒在原始臨床檢體與分離病毒株基因序列差異

為進一步比較流感病毒在未經培養的原始臨床檢體與經細胞培養增殖後的分離株基因序列差異，瞭解流感病毒 quasispecies 基因序列動態變化，本研究接著利用建置完成的單管 RT-PCR 反應系統，先以第二階段所選取的 18 對 A(H3N2)病毒陽性樣本進行 RT-PCR 放大，結果如圖四 A(原始檢體)及 B(病毒株)。電泳膠圖顯示本年所設定的 RT-PCR 反應除可同步放大經培養後的病毒株全基因體基因外，對於存在於未經培養臨床檢體中較為微量的病毒，具相似的放大效率。後續針對 36 個各別樣本次世代定序序列庫的製備電泳膠結果如圖五 A(原始檢體)及 B(病毒株)，各序列庫的產物大小主要介於 250~300 bp 之間，與第一階段建立分析平台時所得趨勢相同。將該 36 個各別序列庫以等量方式混合後，經 Bioanalyzer 儀器進行分析結果如圖六，結果顯示製備完成且混和後的序列庫 DNA 片段大小約在 200~400 bp 之間，且分布趨勢的波峰集中於 282 bp。

將各 36 個次世代定序序列庫以 Illumina Miseq 進行定序，再以 CLC Genomics Workbench version 7.0.3 (CLC Bio, Qiagen)軟體進行個基因片段 contig 組裝後，針對各組樣本病毒 quasispecies 之全基因序列比較(原始臨床檢體病毒 vs. 病毒培養株)結果綜

整如表一~八。分析結果顯示，經細胞培養後的 A(H3N2)病毒株，在其八個基因體片段都會出現具特定核苷酸位點突變的 quasispecies。這些帶有突變的 quasispecies 所占比例大多小於 20%，於單一樣本並非主流族群，亦即這些微量病毒群體的序列資訊並無法以傳統的 Sanger 定序偵測。在 18 對樣本所觀察到的病毒 quasispecies 中，胺基酸因核苷酸同步改變(non-synonymous mutation)且出現次數最多的基因體為 HA(18 次)，其次依序為 PB2(17 次)、PA(16 次)、NA(12 次)、PB1(8 次)、NS(5 次)及 NP(1 次)；MP 並未觀察到具胺基酸突變的病毒 quasispecies。此外，部分病毒 quasispecies 所帶有的核苷酸突變並不會改變其蛋白質的胺基酸序列(synonymous mutation)，此類同義突變出現次數最多的基因體為 PB1(11 次)，其次依序為 PA(7 次)、PB2(5 次)及 NP(5 次)、HA(4 次)、NS(2 次)與 NA(1 次)及 MP(1 次)。

此外，分析經細胞培養後各病毒 quasispecies 基因體蛋白突變種類的位點分布，結果顯示同一位點出現突變次數最多的前三名熱點(hot spot)為 HA 蛋白的位點 176，共出現 8 次(5 種突變)，以 T176K 頻率最高；其次為 NA 蛋白的位點 151，出現 7 次(2 種突變)，以 D151N 頻率最高；第三為 HA 蛋白的位點 174，共出現 5 次(1 種突變)，皆為 N174K。綜觀病毒 quasispecies 全基因體的突變趨勢(表九)，在所有觀察到的胺基酸位點中，出現 1 次的突變佔最多，比例為 83.8%，其對於病毒演化或病原特徵所扮演的角色需進一步評估。

四、 討論

流感病毒對於公共衛生防治所帶來的威脅，即在於其高度變異性，導致具特殊突變的變異株無預警的出現，進而影響疫苗的保護效力。以往利用傳統 Sanger 定序分析流感病毒變異情形時，礙於科技本身限制，僅能得到病毒 quasispecies 中的主流病毒序列，容易忽略較為 minor 的變異情形，無法看出病毒演變全貌。利用 NGS 的定序技術將有助於克服此一限制，針對流感病毒的演變進行更為全面的監控，尤其針對 minor 變異株

消長趨勢的掌握，可協助我國及早掌握新流感變異株出現的時機，對疾病防治實務有實質助益。本年為四年期計畫的第一年，藉由以 NGS 技術同步分析 18 對未經培養且存在於原始臨床檢體中的 A(H3N2)病毒，以及同一檢體經由細胞培養後所分離的病毒株，已可有效偵測病毒 quasispecies。這些 quasispecies 在各基因體片段皆有核苷酸的突變，而這些變異會伴隨蛋白質胺基酸產生非同義或同義突變，此外，偵測結果也觀察到流感病毒的 HA 及 NA 表面蛋白均出有非同義突變的熱點，包含 HA 蛋白的 174 與 176 兩位點，以及 NA 蛋白的 151 位點等。這些因細胞培養而產生的突變對於病毒演化或病原毒力特徵的意義將於未來的研究期程進行完整評估。此外，這些病毒 quasispecies 的發現也同時代表新一代的 NGS 確實較傳統 Sanger 定序更有能力進行微量病毒變異偵測，本分析流程將對於我國流感疫情預警具實質助益。

本計畫於上半年陸續完成 NGS 分析流程後，下半年已依原規劃，分析過去的檢體以及今年新收案的檢體，探討流感病毒於細胞培養複製後產生的 quasispecies 趨勢，其策略即著重原始檢體與培養後病毒基因序列的差異分析，以利更為全面的掌握病毒基因序列動態變化。依現行流程，目前已成功將單次上機可分析全基因體的病毒樣本數提高至 36 件，並於同一實驗比較各時期流感病毒的樣態。未來進行更進一步的跨物種分析時，將規劃與本署防疫醫師或我國各大型醫院之醫師合作，協助取得不同時期與不同部位之檢體，再以 NGS 定序比較各個時期與部位的病毒序列差異。

若評估以 NGS 進行流感病毒全基因體定序的成本效益(表十)，依本計畫現階段流感病毒 NGS 的分析量能，每次同時進行 36 個樣本全基因體定序所需經費估計約 109,160 元，平均每個樣本約 3,033 元。若與採用傳統 Sanger 定序法相比，36 個病毒樣本進行全基因體定序則需 468,000 元，平均每個樣本約 13,000 元。這個趨勢顯示，NGS 不但可有效作為分析微量流感病毒 quasispecies 的變化情形，應用於流感病毒序列監測，亦可大幅節省經費，惟若考量兩種分析方法的即時性，以 NGS 進行病毒全基因體序列分析，由 RT-PCR 起算，經過序列庫製備以及 Miseq 上機，約需 5~10 天可進行序列組裝分析；反之，傳統 Sanger 定序耗時較短，約 2~3 天即可得到定序結果進行組裝。因此，NGS

現階段仍無法完全取代 Sanger 定序，未來兩者方法的選用時機，可依照不同時效需求或經費預費等因素進行評估。

五、結論與建議

- (一) 本計畫已於第一年完成建立流感病毒NGS全基因定序流程，包括(1)最佳化RT-PCR進行流感病毒全基因片段放大；(2)以RT-PCR放大後的基因片段產物製備次世代定序序列庫(library)；(3)將次世代定序序列庫進行定序以及後續組裝分析。此分析流程經由兩批定序之實驗結果證實具穩定性，現階段一批次病毒全基因體最大分析量能為36個樣本，除可有效偵測微量病毒quasispecies的變異情形，亦可大幅節省經費。未來將持續增加單次NGS上機的可分析樣本數，精進分析品質及成本效益。
- (二) 整體考量經費及時效等因素，NGS現階段仍無法完全取代Sanger定序，未來針對流感病毒的監測與演化分析，如何將兩種方法互相搭配，獲得最佳效益，將是提升我國流感防治量能的重要工作。本計畫執行成果將可提供佐證，俾使實驗室權責單位依照不同時效或經費需求等因素評估NGS的可應用性。

六、計畫重要研究成果及具體建議

- (一) 完成建立流感病毒NGS全基因定序流程。
- (二) 比較流感病毒quasispecies動態變化，找出位於表面蛋白HA與NA的變異熱點(Hot spot)供後續病毒演化及毒力特徵等分析。
- (三) NGS對於流感病毒基因體監測分析具實際應用價值，經有效評估後，將可協助實驗室進行新病毒風險評估等相關研究。

七、 参考文献：

1. Barzon L, Lavezzo E, Militello V, Toppo S, and Palu G: Applications of next-generation sequencing technologies to diagnostic virology. *Int J Mol Sci* 2011; 12:7861-7884.
2. Croville G, Soubies SM, Barbieri J, Klopp C, Mariette J, Bouchez O, Camus-Bouclainville C, and Guerin JL: Field monitoring of avian influenza viruses: whole-genome sequencing and tracking of neuraminidase evolution using 454 pyrosequencing. *J Clin Microbiol* 2012; 50:2881-2887.
3. Dunne WM, Jr., Westblade LF, and Ford B: Next-generation and whole-genome sequencing in the diagnostic clinical microbiology laboratory. *Eur J Clin Microbiol Infect Dis* 2012; 31:1719-1726.
4. Fordyce SL, Bragstad K, Pedersen SS, Jensen TG, Gahrn-Hansen B, Daniels R, Hay A, Kampmann ML, Bruhn CA, Moreno-Mayar JV, Avila-Arcos MC, Gilbert MT, and Nielsen LP: Genetic diversity among pandemic 2009 influenza viruses isolated from a transmission chain. *Virology* 2013; 45:116.
5. Kuroda M, Katano H, Nakajima N, Tobiume M, Aina A, Sekizuka T, Hasegawa H, Tashiro M, Sasaki Y, Arakawa Y, Hata S, Watanabe M, and Sata T: Characterization of quasispecies of pandemic 2009 influenza A virus (A/H1N1/2009) by de novo sequencing using a next-generation DNA sequencer. *PLoS One* 2010; 5:e10256.
6. Lauring AS, and Andino R: Quasispecies theory and the behavior of RNA viruses. *PLoS Pathog* 2010; 6:e1001005.
7. Lin Z, Farooqui A, Li G, Wong GK, Mason AL, Banner D, Kelvin AA, Kelvin DJ, and Leon AJ: Next-generation sequencing and bioinformatic approaches to detect and analyze influenza virus in ferrets. *J Infect Dev Ctries* 2014; 8:498-509.
8. Moncla LH, Zhong G, Nelson CW, Dinis JM, Mutschler J, Hughes AL, Watanabe T, Kawaoka Y, and Friedrich TC: Selective Bottlenecks Shape Evolutionary Pathways Taken during Mammalian Adaptation of a 1918-like Avian Influenza Virus. *Cell Host Microbe* 2016; 19:169-180.
9. Nelson MI, and Holmes EC: The evolution of epidemic influenza. *Nat Rev Genet* 2007; 8:196-205.
10. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S: MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011; 28:2731-2739.
11. Tang P, and Chiu C: Metagenomics for the discovery of novel human viruses. *Future Microbiol* 2010; 5:177-189.
12. Van den Hoek S, Verhelst J, Vuylsteke M, and Saelens X: Analysis of the genetic diversity of influenza A viruses using next-generation DNA sequencing. *BMC Genomics* 2015;

16:79.

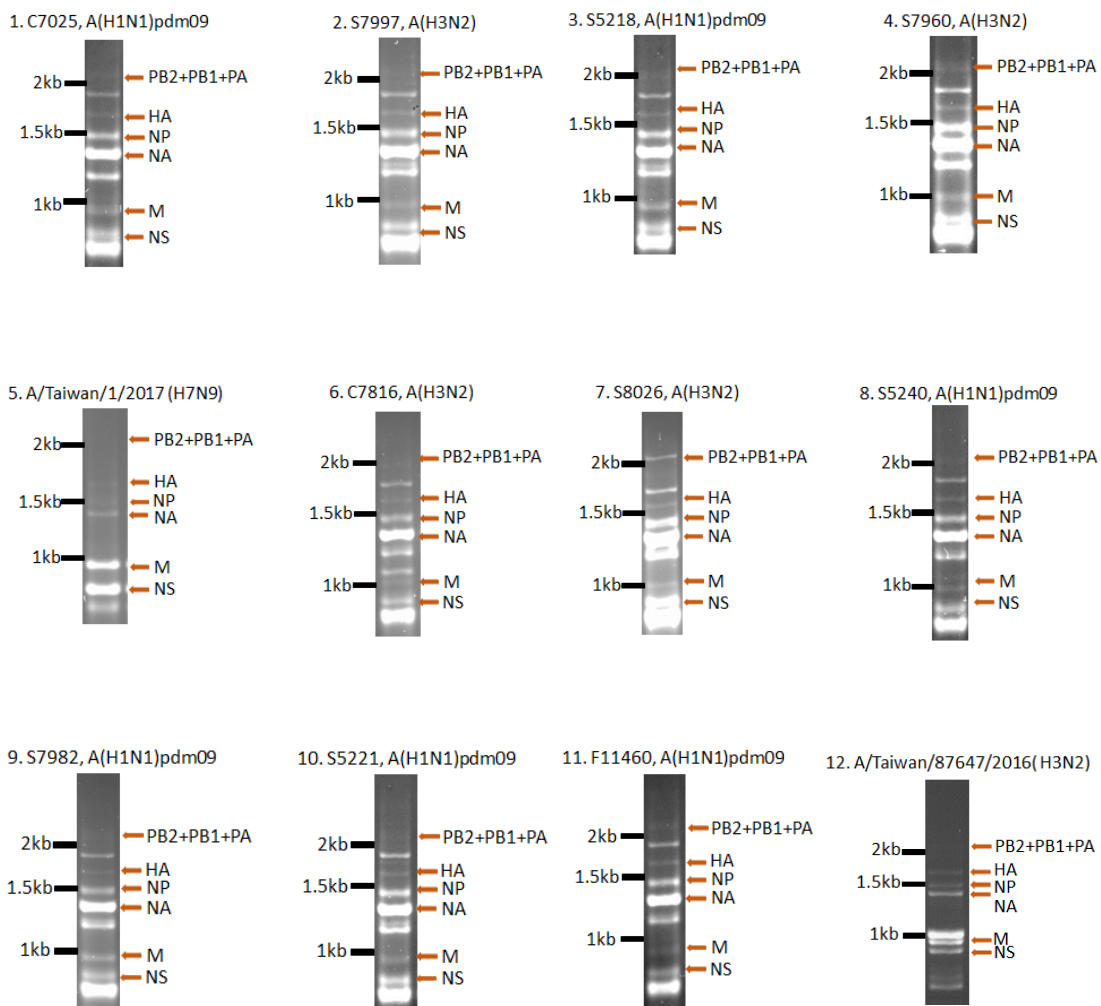
13. Varble A, Albrecht RA, Backes S, Crumiller M, Bouvier NM, Sachs D, Garcia-Sastre A, and tenOever BR: Influenza A virus transmission bottlenecks are defined by infection route and recipient host. *Cell Host Microbe* 2014; 16:691-700.

14. Wu Y, Wu Y, Tefsen B, Shi Y, and Gao GF: Bat-derived influenza-like viruses H17N10 and H18N11. *Trends Microbiol* 2014; 22:183-191.

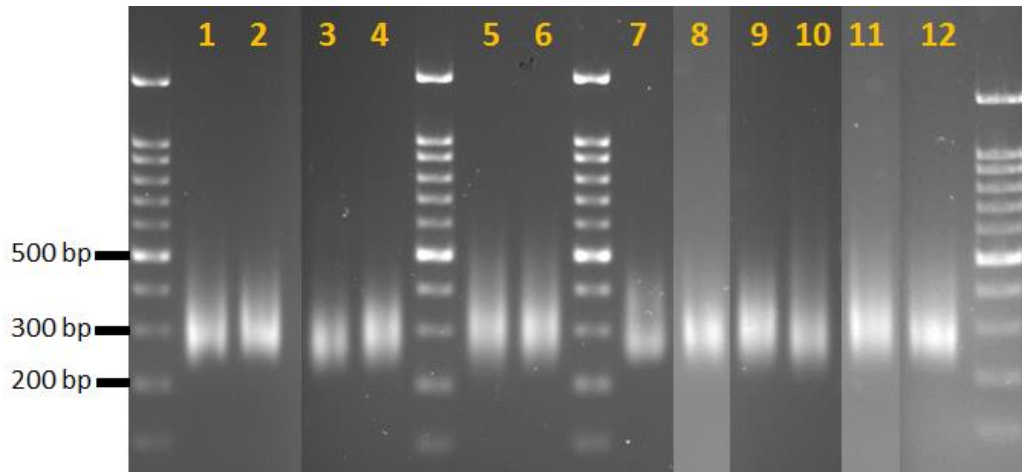
15. Yang JR, Kuo CY, Huang HY, Wu FT, Huang YL, Cheng CY, Su YT, Chang FY, Wu HS, and Liu MT: Newly emerging mutations in the matrix genes of the human influenza A(H1N1)pdm09 and A(H3N2) viruses reduce the detection sensitivity of the real-time RT-PCR assay. *J Clin Microbiol* 2014; 52:76-82.

八、圖、表

圖一、以 RT-PCR 放大各亞型 A 型流感病毒株全基因體片段結果

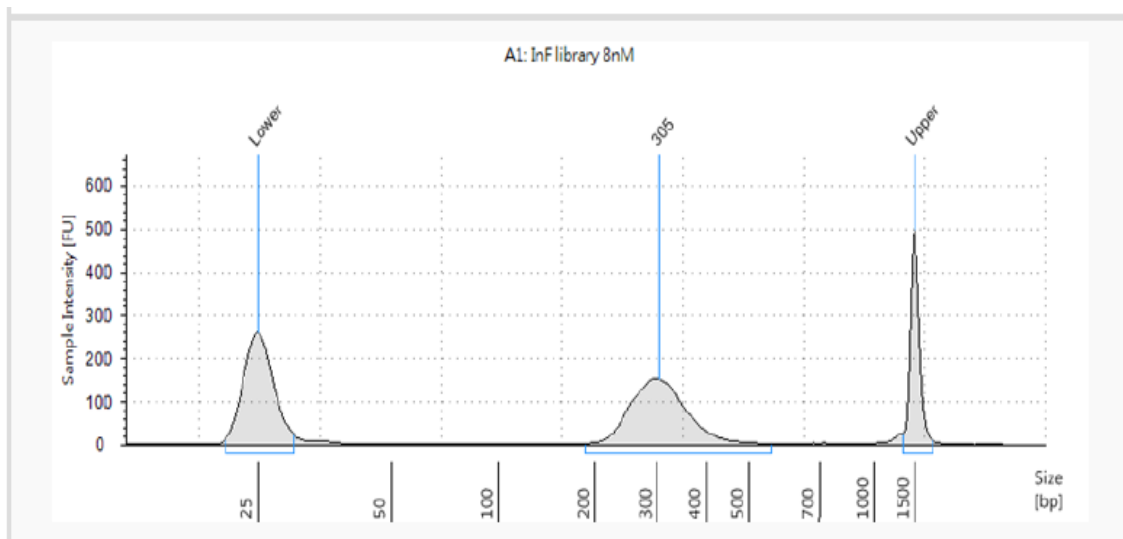


圖二、以傳統電泳分析 12 個流感病毒次世代定序序列庫之結果

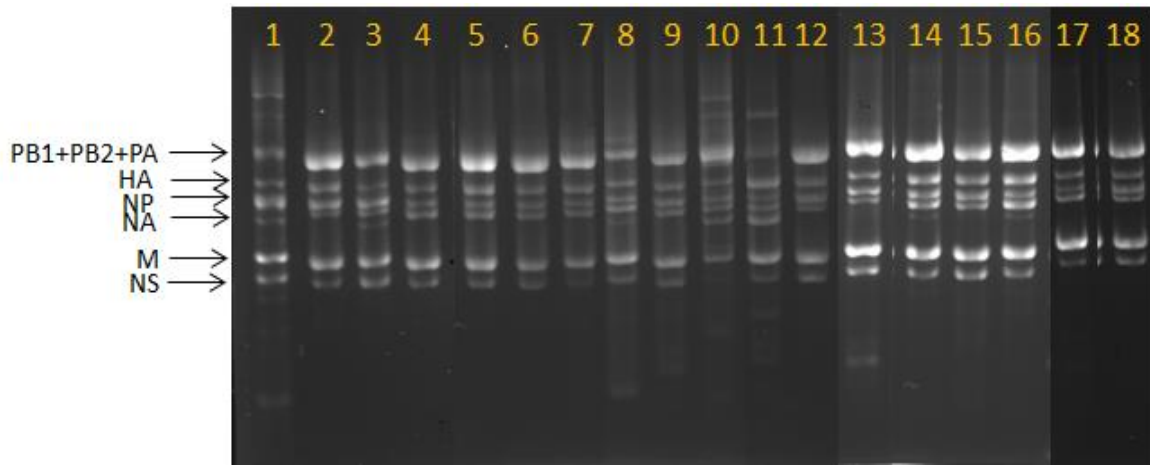


- 1: C7025, A(H1N1)pdm09; 2: S7997, H3N2; 3: S5218, A(H1N1)pdm09;
4: S7960, H3N2; 5: A/Taiwan/1/2017(H7N9); 6: C7816, H3N2; 7: S8026, H3N2
8.S5240, A(H1N1)pdm09; 9: S7982, A(H1N1)pdm09; 10: S5221, A(H1N1)pdm09
11: F11460, A(H1N1)pdm09; 12: A/Taiwan/876472016(H3N2)

圖三、以 Bioanalyzer 分析流感病毒次世代定序序列庫(12 個序列庫混合)之結果

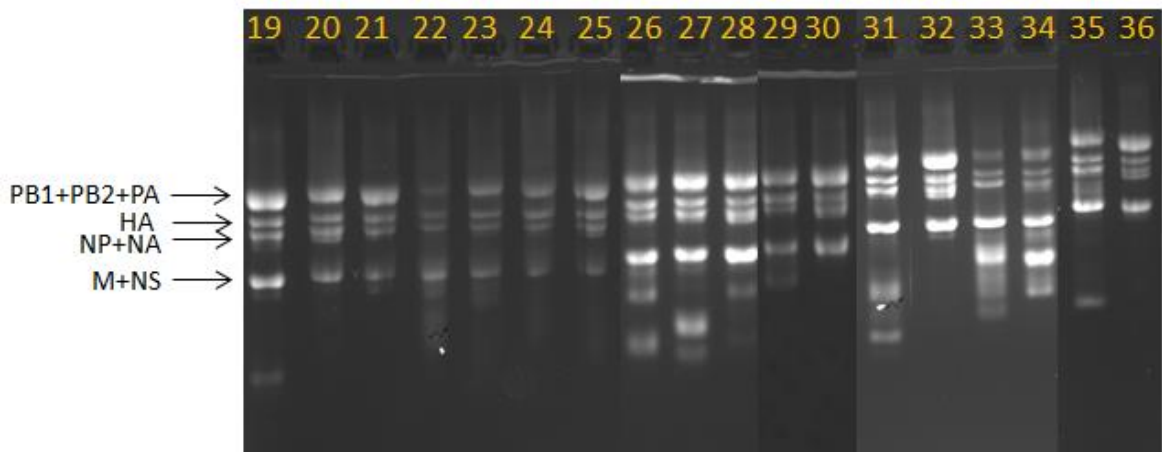


圖四 A、以 RT-PCR 放大 18 支 A(H3N2) 流感陽性原始檢體病毒之全基因體片段結果



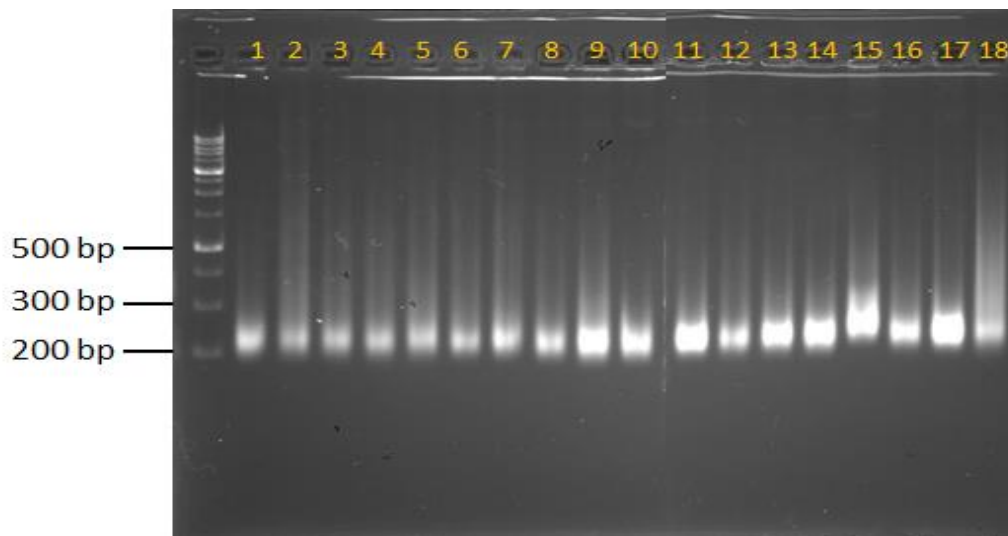
1:C8197; 2:C8203; 3:C8215; 4:C8217; 5:C8238; 6:C8243; 7:C8244; 8:C8269; 9:C8282;
10:C8284; 11:C8311; 12:C8312; 13:S8408; 14:S8425; 15:S8430; 16:S8434; 17:S8488;
18:S8517

圖四 B、以 RT-PCR 放大 18 支 A(H3N2) 流感病毒細胞培養株之全基因體片段結果



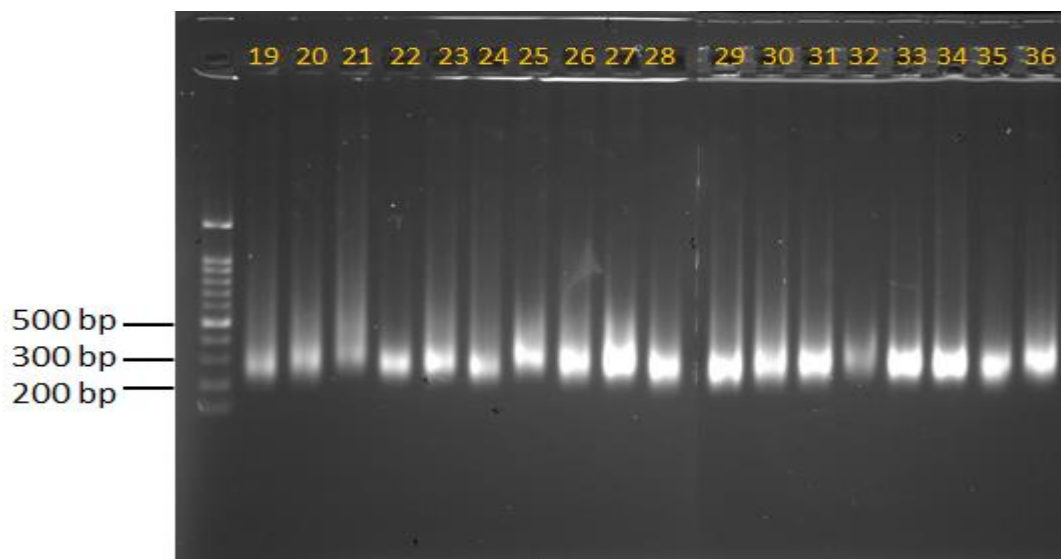
19:C8197; 20:C8203; 21:C8215; 22:C8217; 23:C8238; 24:C8243; 25:C8244; 26:C8269;
27:C8282; 28:C8284; 29:C8311; 30:C8312; 31:S8408; 32:S8425; 33:S8430; 34:S8434;
35:S8488; 36:S8517

圖五 A、以傳統電泳分析 18 個 A(H3N2)流感病毒陽性原始檢體次世代定序序列庫之結果



1:C8197; 2:C8203; 3:C8215; 4:C8217; 5:C8238; 6:C8243; 7:C8244; 8:C8269; 9:C8282;
10:C8284; 11:C8311; 12:C8312; 13:S8408; 14:S8425; 15:S8430; 16:S8434; 17:S8488;
18:S8517

圖五 B、以傳統電泳分析 18 個 A(H3N2)流感病毒細胞培養株次世代定序序列庫之結果



19:C8197; 20:C8203; 21:C8215; 22:C8217; 23:C8238; 24:C8243; 25:C8244; 26:C8269;
 27:C8282; 28:C8284; 29:C8311; 30:C8312; 31:S8408; 32:S8425; 33:S8430; 34:S8434;
 35:S8488; 36:S8517

圖六、以 Bioanalyzer 分析流感病毒次世代定序序列庫(36 個序列庫混合)之結果



表一、經培養後病毒 quasispecies 於 PB2 蛋白產生之突變分析

(A) 非同義突變

	PB2																
a.a position	27	49	50	105	108	113	114	194	196	598	649	657	670	723	740	755	
specimen	H	W	I	V	A	K	V	R	C	T	V	K	N	K	N	D	R
cultured virus	N	*	M	M	T	*	L	Q	Y	S	G	E*	K*	E	K	N	Q
percentage (%)	5.09	6.74	6.44	5.98	22.65	9.03	8.10	10.78	12.57	5.43	9.55	44.05	55.48	5.21	8.39	12.60	16.38

(B) 同義突變

	PB2				
a.a position	31(93)	84(252)	144(430)	200(600)	461(1383)
specimen (nt)	T	T	C	T	C
cultured virus (nt)	C	C	A	C	T
percentage (%)	7.84	6.07	10.38	5.56	16.66

表二、經培養後病毒 quasispecies 於 PB1 蛋白產生之突變分析

(A) 非同義突變

PB1								
a.a position	37	47	84		111	163	192	676
specimen	W	H	Q		M	L	G	S
cultured virus	G	L	K	H	I	I	R*	N
percentage (%)	6.11	9.45	30.12	30.24	13.49	-	84.53	12.44

(B) 同義突變

PB1											
a.a position	8(22)	14(42)	74(222)	136(408)	162(486)	206(618)	222(666)	428(1284)	686(2058)	687(2061)	728(2184)
specimen (nt)	C	A	T	C	G	A	G	A	A	A	T
cultured virus (nt)	T	G	A	T	A	G	A	G	G	G	C
percentage (%)	15.46	9.52	7.55	12.68	7.97	4.37	11.69	11.52	20.47	20.06	7.67

表三、經培養後病毒 quasispecies 於 PB1 蛋白產生之突變分析

(A) 非同義突變

PA																
a.a position	31	33	35	92	113	139	140	155	156	211	391	557	605	608	652	653
specimen	E	N	F	N	K	K	S	I	A	M	E	I	K	T	S	P
cultured virus	A	T	S	K	*	E	A	M	T	K	K	V	R	I	A	I
percentage (%)	6.34	6.58	7.41	9.11	28.71	22.52	22.88	7.91	5.42	5.31	20.01	7.33	17.44	9.26	9.87	9.57

(B) 同義突變

PA							
a.a position	87(261)	114(342)	140(420)	171(513)	342(1026)	544(1632)	545(1635)
specimen (nt)	C	A	T	T	G	G	A
cultured virus (nt)	A	G	C	C	A	A	C
percentage (%)	4.92	18.85	22.01	55.13	6.78	30.98	5.73

表四、經培養後病毒 quasispecies 於 HA 蛋白產生之突變分析

(A) 非同義突變

HA																			
a.a position	120	174	176						237	257	264	351			399	401	408	452	465
specimen	D	N	T						P	D	T	K	I	R	I	F	T	E	
cultured virus	G	K	I	R*	A*	E	K	L	H	E	I	I	K	Q*	V	L	I	G*	
percentage (%)	9.68	10.43 ~ 46.86	28.79	63.14	-	-	10.04 ~ 23.31	6.37 ~ 8.02	32.93	5.65	5.21	6.41	24.48	15.99	5.39	11.05	33.2	68.25	

(B) 同義突變

HA				
a.a position	236(708)	269(807)	550(1650)	
specimen (nt)	G	A	T	C
cultured virus (nt)	A	T	C	T
percentage (%)	43.15	8.24	8.15	52.39

表五、經培養後病毒 quasispecies 於 NP 蛋白產生之突變分析

(A) 非同義突變

	NP	
a.a position	190	385
specimen	A	Y
cultured virus	V	H
percentage (%)	7.63	21.7

(B) 同義突變

	NP				
a.a position	27(81)	56(168)	108(324)	150(450)	417(1251)
specimen (nt)	A	A	C	A	C
cultured virus (nt)	T	G	A	G	T
percentage (%)	17.39	5.19	13.54	8.15	8.53

表六、經培養後病毒 quasispecies 於 NA 蛋白產生之突變分析

(A) 非同義突變

	NA											
a.a position	21	21	27	99	148	151	362	392	393	397	423	
specimen	C		A	P	T	D		M	I	N	I	Y
cultured virus	G	F	D	S	I	N	G	V	K	H	K	D
percentage (%)	11.83	28.57	12.57 ~ 18.83	7.13	8.94	11.15 ~ 44.53		29.11	6.98	6.85	14.8	7.75

(B) 同義突變

	NA
a.a position	414(1242)
specimen (nt)	T
cultured virus (nt)	C
percentage (%)	20.25

表七、經培養後病毒 quasispecies 於 MP 蛋白產生之突變分析

(A) 非同義突變

	MP
a.a position	-
specimen	-
cultured virus	-
percentage (%)	-

(B) 同義突變

	MP
a.a position	243(729)
specimen (nt)	G
cultured virus (nt)	A
percentage (%)	6.2

表八、經培養後病毒 quasispecies 於 NS1 及 NS2 蛋白產生之突變分析

(A) 非同義突變

a.a position	NS1			NS2	
	64	71	229	72	112
specimen	I	G	K	K	E
cultured virus	T	E	E	E	G
percentage (%)	6.72	63.29	5.37 ~ 23.95	5.37 ~ 23.95	10.02

(B) 同義突變

a.a position	NS1		NS2
	64(192)	269(807)	-
specimen (nt)	C	A	-
cultured virus (nt)	T	G	-
percentage (%)	8.28	10.02	-

表九、病毒 quasispecies 全基因體突變次數分析

No. of mutation	Viral proteins								total (%)
	PB2	PB1	PA	HA	NP	NA	MP	NS1/NS2	
1	15	6	16	8	2	7		2/1	57 (83.8)
2	1	1		1		2			5 (7.3)
3				1				1/1	3 (4.4)
4									
5				1					1 (1.5)
6									
7						1			1 (1.5)
8				1					1 (1.5)
Total	16	7	16	12	2	10	0	3/2	68

表十、流感病毒 NGS 與 Sanger 定序成本效益比較分析

To perform full-genome sequencing of 36 influenza A isolates	Sanger		Miseq	
	Reactions	Cost	Reactions	Cost
RT-PCR reaction (NT 110/ rxn)	20 x 36	\$79,200	1 x 36	\$3,960
Library preparatrion (NT 1,200/ rxn)	-	-	1 x 36	\$43,200
Sanger sequencing (NT 135/ rxn)	20 x 4 x 36	\$388,800	-	-
MiSeq sequencing (15M reads, NT 62,000/ rxn)	-	-	1	\$62,000
Total cost		\$468,000		\$109,160