

計畫編號：MOHW108-CDC-C-315-134515

衛生福利部疾病管制署 108 年署內科技研究計畫

計畫名稱：高通量定序(NGS)平台開發與應用

108 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：慕蓉蓉

協同主持人：

研究人員：林鈺棋

研究人員：陳筱蓉

執行期間：108 年 1 月 1 日至 108 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	6
三、結果	8
四、討論	11
五、結論與建議	12
六、參考資料	13
六、圖、表	15

共 (21) 頁

計畫中文摘要：

高通量定序（high-throughput sequencing），又稱為下世代定序（Next Generation Sequencing, NGS）具有龐大數據之產出以及 de novo 定序分析的強大能力，有別於以往僅能對已知病原體基因序列進行偵測，近年來廣泛運用在傳染病原的偵測與追蹤，將分子診斷提升至全基因領域。目前 NGS 可應用於不明原因傳染病原偵測、新興微生物（病原）之鑑定分析、微生物全基因體定序、微生物分型、傳染病群聚調查、抗藥性基因分析比對以及微生物菌叢（microbiome）的變化追蹤等。

本計畫完成建置以第二代搭配第三代 NGS 全基因定序技術進行細菌全基因體組裝、基因分析以及比對，期望運用高通量定序偵測平台建立病原體基因序列組裝分析的流程，適當串連各種分子檢驗技術之優點，以期能於第一時間將罕見或變異之病原體全基因解密，強化防疫時效與避免不明傳染病對社會造成衝擊。加強病原體基因序列資料之彙整分析，瞭解其可能的感染源、疾病的流行趨勢及協助疫情調查外，還可獲取更多生物資訊以增進對該病原的瞭解及認識，作為未來防疫政策擬訂及相關疾病研究的重要參考。

關鍵詞：高通量，下世代定序

計畫英文摘要：

keywords : high-throughput, next generation sequencing

High-throughput sequencing, also referred to as next generation sequencing has been proved itself to be robust and expanded to the whole genome level. The NGS has dramatically changed the molecular diagnostic arena in clinical microbiology and applied to detection of unknown disease associated pathogens, genotypic resistance testing, outbreak investigation and investigation of microbial population diversity.

In this study, miseq and PacBio NGS platforms have been applied in the genome assembly and comparison. We expect to establish pathogen genome analysis platform by adequately integrating many kinds of molecular diagnostic techniques with NGS system. Through the analysis of pathogen sequences and related epidemiological information, we can assist the clarification of infectious pathogens and infection routes, provide a reference for disease prevention policy assessment and formulating monitor the drug resistance of pathogens and even provide a foundation for future developing diagnostic techniques.

本文

一、前言：

高通量定序又被稱為 deep sequencing, 下世代定序 (next generation sequencing, NGS), 其成功超越傳統 sanger sequencing, 不但可以一次產出 Gb 至 Tb 大量之序列資料, 且無需事先設計引子或探針, 直接針對未知基因進行序列分析, 再加上 de novo 定序組裝的強大能力, 可針對新興未知病原基因體完全解密[1]。

近年來高通量定序廣泛應用在傳染病原的偵測與追蹤, 例如不明原因傳染病原偵測、新興微生物 (病原) 之鑑定分析、微生物全基因體定序、微生物分型、傳染病群聚調查、抗藥性基因分析比對以及微生物菌叢 (microbiome) 的變化追蹤等[2-9]。

目前市面上有數種高通量定序平台針對不同目的, 數據資料的產出量, 每筆序列的長度及組裝功能的差異, 可供選擇使用。高通量定序在診斷應用上可分成兩種策略: whole exome sequencing (WES) 以及 whole genome sequencing (WGS) [10]。前者利用已知基因的序列, 大量篩選可能之變異 (例如抗藥變異) [11], 或是追蹤菌叢的變化 (例如以 16S 基因分析細菌種類) [12]。後者則利用 de novo sequencing 的方式, 可以組裝全基因體序列 (2011 年德國新型出血性大腸桿菌 E. Coli O104), 或者未知病原的偵測以及

群聚事件的調查[4-6]。

高通量序列分析平台彌補傳統的病原體檢測技術僅能對已知病原體基因序列進行檢測，且單次的試驗裡只能檢測一個或數種已知的病原體的不足，且具有大規模 de novo 定序分析的強大能力，提高病原全基因解析的能力與速度，隨著基因定序方法的進步，能迅速累積病原體基因序列，得以全面性分析病原體之基因體結構。因此，本計畫將選取不同之高通量定序平台分別運用於未知病原之偵測、全基因體定序組裝以及群聚事件的調查等。未來面對未知的新興傳染病時，可以此系統進行即時偵測及基因解密；且針對發生群聚時，進行分析比對尋找可能知感染來源，以期能即時發現不明與罕見的病原體，快速找出感染源，強化防疫時效與避免不明傳染病對社會造成衝擊。

二、材料與方法

1、實驗菌株

使用疾病管制署收集之 CRE 送驗菌株。將挑選送至昆陽實驗室的 CRE 菌株經實驗室確認為 carbapenemase 基因陽性的菌株做進一步抗藥基因及質體分析。

2、脈衝膠電泳分析 (Pulsed field gel electrophoresis, PFGE)

使用限制酶進行菌株分型實驗，以脈衝電泳儀 CHEF-MAPPER (BIO-RAD, USA)，在 0.5x TBE buffer 將切斷的片段以電泳槽 CHEF-Mapper 跑膠質；以 H9812 菌株 (XbaI 限制酶切割) 當作片段大小指標。使用限制酶之脈衝圖譜，利用電腦將圖片掃描儲存成圖片檔，接著以套裝軟體 Phoretix 1D gel analysis advanced version 5.01 (Nonlinear Dynamics, UK) 對菌株進行親緣性樹狀圖分析，其原理是利用不同 DNA 片段電泳圖譜進行分析，以 UPGMA (unweighted pair group method using arithmetic averages) 的方式畫出樹狀圖 (dendrogram)，由樹狀圖對應出相似指數，分析菌株間分子關聯性。

3、抗藥質體的確認：S1-PFGE/Southern hybridization

使用 S1 nuclease 限制酶進行菌株分型實驗，以脈衝電泳儀 CHEF-MAPPER (BIO-RAD, USA)，變換時間：0.5-30 秒，電場值

為 6V/cm²，200V 電壓值，20 小時電泳時間，使用 1 % SeaKem Gold agarose (BMA, Rockland, ME, USA) 及 0.5×TBE 電泳液，以 H9812 菌株 (XbaI 限制酶切割) 當作片段大小指標。電泳膠須以 HCl depurination、NaOH denature 以及 Tris-HCl (pH 7.5) neutralization 處理後，轉漬至 NC paper 上。與 DIG 標誌之探針 (Roche Applied Science) 經 hybridization 20 小時，清洗後，以化學冷光偵測儀 (VersaDoc, BioRad) 偵測反應訊息

4、NGS 定序

利用自動核酸萃取系統 MagNA Pure Compact Instrument (Roche Applied Science) 進行核酸萃取，將核酸以 Nextera DNA Flex Library Prep Kit (illumina, 美國) 進行建庫，並以 Illumina iSeq 平台進行定序。

5、Bioinformatics 分析

NGS 定序資料以 BaseSpace Sequence Hub (illumina) 之 SPAdes Genome Assembler 進行 de novo assembly 組裝，組裝完成之片段再以 MLST 2.0 及 ResFinder 3.2 (Center for Genomic Epidemiology, DTU, Denmark, <https://cge.cbs.dtu.dk/services/>) 進行基因型別及抗藥基因分析。

三、 結果

1、 收集全國分區帶 NDM-5 菌株

統計103年至108（本）年5月期間，自19名CRE通報個案共分離21株帶NDM-5菌株，分別來自5家醫院，包括北部2家（HH及TH）、南部2家（CHA及CHB）與東部1家（TzHH）。本計畫分析103-107年共17名個案之帶NDM-5菌株，包括16株*E. coli*及1株*Enterobacter cloacae*（圖一）。

2、 帶 NDM-5 *E. coli* 菌株之親緣關係圖

將16株帶NDM-5 *E. coli* 菌株進行PFGE分析（圖二），結果顯示來自北部HH醫院的8株菌株（HH1-8）與另一北部醫院TH之TH3菌株間之親緣關係相同度高於80%，屬於同源菌株。TH醫院的4株菌株（TH1-4）中，TH2及TH4之相同度達80%以上，屬於同源菌株。東部TzHH醫院的2株菌株（TzHH1-2），自成同一來源。此外，TH1、來自南部CHA醫院（CHA1）及CHB醫院（CHB1）的菌株與其他醫院的菌株相同度偏低，屬於散發事件。

3、 NDM-5 之質體分析

為進一步分析NDM-5質體，將17株帶NDM-5菌株先利用S1-PFGE分離質體，接續使用NDM-5基因之部分片段作為探針

進行 Southern Hybridization，以標示 NDM-5 質體，分析結果如圖三。依據 NDM-5 質體大小可區分為 4 類：(1) TH2 的 NDM-5 質體大小最大，約 220kb；(2) HH1-8 及 TH3 之 NDM-5 質體大小介於 150-180kb；(3) TH1、TH4 及 TzHH1-2 之 NDM-5 質體大小介於 80-110kb；(4) CHA1 及 CHB1-2 之 NDM 質體最小，約 40kb。

4、 利用 NGS 分析 7 株帶 NDM-5 菌株

為瞭解帶 NDM-5 菌株間之差異性，遴選 7 株 (HH7、TH1-2、TH4、CHB1-2 及 TzHH1) 帶 NDM-5 菌株進行 NGS。經 de novo assembly 後分別以 MLST 2.0 及 ResFinder 3.2 進行分析。依 MLST 分析結果，HH7 型別為 ST617，TH1 及 TzHH1 均為 ST410，TH2 與 TH4 則皆為 ST405，CHB1 型別為 ST3997 (表一)。此外，依 ResFinder 分析結果，7 株帶 NDM-5 菌株除具有 carbapenemase 的 *bla*_{NDM-5} 基因外，亦包括 extended-spectrum β -lactamase 基因 (*bla*_{CTX-M-15}、*bla*_{TEM-1B}) 及 aminoglycoside adenytransferase 基因 (*aadA2*、*aadA5*) 等多種抗藥基因。

為比較分析 NDM-5 質體間之差異性，將含有 NDM-5 基因之組裝片段以 Easyfig (<https://mjsull.github.io/Easyfig/>) 軟體進行雙序列

並列分析 (Pairwise sequence alignment)，依序列差異性可分為3類 (圖四)：

- (1) HH7、TH4及TzHH1組裝出序列相同之NDM-5組裝片段，包含2組抗藥基因結構，分別為 $bla_{\text{NDM-5}}-ble_{\text{MBL}}$ (bleomycin resistance gene) 及class 1 integron。class 1 integron包含可抵抗trimethoprim (*dfrA17*)、streptomycin (*aadA5*) 及sulfonamide (*sul1*) 等抗生素的抗藥基因，與具有多重抗藥功能的efflux pump基因 (*qacE Δ 1*)。此外，組裝片段兩端為插入序列IS26的部分片段，並形成正向重複序列 (direct repeat)。
- (2) TH1及TH2組裝出序列相同之NDM-5組裝片段，同樣包含 $bla_{\text{NDM-5}}-ble_{\text{MBL}}$ 及class 1 integron共2組抗藥基因結構，其中*aadA2*及*dfrA12*基因相異於HH7組裝片段之*aadA5*及*dfrA17*基因，其他序列均類似於HH7組裝片段。
- (3) CHB1及CHB2亦則組裝出序列相同之NDM-5組裝片段，其中僅包含1組抗藥基因結構 $bla_{\text{NDM-5}}-ble_{\text{MBL}}$ ，其中 $bla_{\text{NDM-5}}$ 上游分別為IS5及ISAb125序列，下游為IS26序列，明顯不同於上述2類的基因組成。

四、討論

本計畫利用 NGS 分析國內帶有 NDM-5 質體之 CRE。依據

NDM-5 組裝結果，推論 3 種可能的 NDM-5 傳播方式如下：

- (1) 來自相同HH醫院的HH1-8，其PFGE親緣高度相似，亦具有類似的質體圖譜及大小相近的NDM-5質體，因此可能為帶NDM-5質體同源菌株之傳播方式。
- (2) 來自相同TH醫院的TH1及TH2，其PFGE親緣低度相似，且質體圖譜及NDM-5質體大小均有差異，顯示兩菌株之來源不同，但兩菌株卻具有相同的NDM-5組裝片段，此結果可能與組裝片段兩端的IS26有關。IS26常見於各種抗藥質體，且多位於抗藥基因附近。IS26由一個Tnp26轉位酶及兩端各14bp反向重複序列(indirect repeat)組成，具有稱為translocatable units的新型易位機制，可攜帶一種或多種抗藥基因，插入至其他IS26存在的地方，已被認為是傳播抗藥基因之重要角色[14]。在本計畫可從TH1與TH2兩者在不同質體上卻發現相同的NDM-5組裝片段獲得證實，故推論此為IS序列攜帶NDM-5基因易位之傳播方式。
- (3) 來自相同醫院不同菌種的CHB1及CHB2，兩者具有相同大小的NDM-5質體，且具有相同的NDM-5組裝片段，因此可能為NDM-5質體之傳播方式。

五、 結論與建議

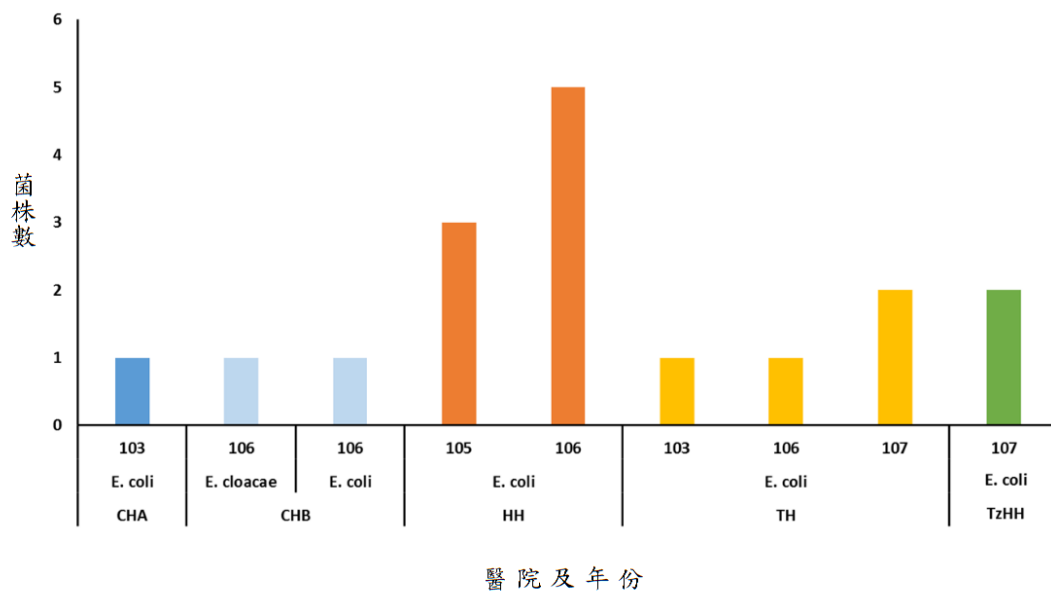
- 1、 本計畫以建置 NGS 平台以應用於傳染病為主軸，第一（106）年已完成於不明原因感染原之檢驗應用，並制訂不明原因知感染源之檢驗流程。第二（107）年，運用兩種 NGS 平台（短片段搭配長片段組合校正），完成全基因組裝定序，並進行解密分析比對。
- 2、 本第三（108）年，利用 NGS 產出之大量序列，進行 NDM-5 群聚事件追蹤分析。過去針對細菌抗藥性相關研究，多僅聚焦於某些已知抗藥基因，但對於其他與抗藥基因相關部分，如 IS 序列及未知的新型抗藥基因等，通常難以同步進行分析。此外，由於細菌可透過質體交換或 IS 序列的易位機制傳播抗藥基因，因此若單以傳統 PFGE 進行菌株分型，僅能掌握同源菌株之傳播方式，難以瞭解其他如抗藥質體或抗藥基因易位等傳播方式。本計畫利用 NGS 解序抗藥基因，可獲得更細緻的分子生物資訊以釐清菌株及抗藥基因之來源，並可掌握抗藥菌株之流行演化趨勢。
- 3、 本計畫建置之流程，在未來面對新型變異之病原，提供全基因序列解密，更可與國際序列進行比對，與國際接軌。

六、參考資料

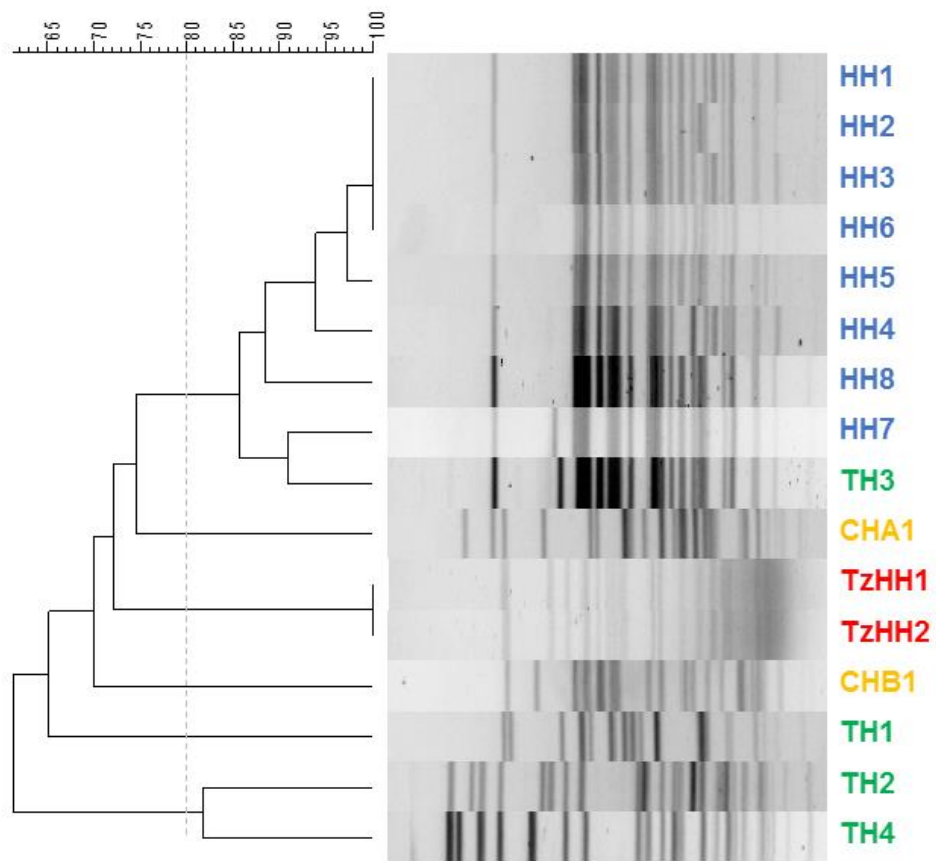
- [1] Lefterova MI, Suarez CJ, Banaei N, Pinsky BA. Next-Generation Sequencing for Infectious Disease Diagnosis and Management: A Report of the Association for Molecular Pathology. *The Journal of molecular diagnostics : JMD*. 2015;17:623-34.
- [2] Monira S, Nakamura S, Gotoh K, Izutsu K, Watanabe H, Alam NH, et al. Metagenomic profile of gut microbiota in children during cholera and recovery. *Gut pathogens*. 2013;5:1.
- [3] Koser CU, Ellington MJ, Cartwright EJ, Gillespie SH, Brown NM, Farrington M, et al. Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. *PLoS pathogens*. 2012;8:e1002824.
- [4] Deng X, den Bakker HC, Hendriksen RS. Genomic Epidemiology: Whole-Genome-Sequencing-Powered Surveillance and Outbreak Investigation of Foodborne Bacterial Pathogens. *Annual review of food science and technology*. 2016;7:353-74.
- [5] Brzuszkiewicz E, Thurmer A, Schuldes J, Leimbach A, Liesegang H, Meyer FD, et al. Genome sequence analyses of two isolates from the recent *Escherichia coli* outbreak in Germany reveal the emergence of a new pathotype: Entero-Aggregative-Haemorrhagic *Escherichia coli* (EAHEC). *Archives of microbiology*. 2011;193:883-91.
- [6] Nakamura S, Yang CS, Sakon N, Ueda M, Tougan T, Yamashita A, et al. Direct metagenomic detection of viral pathogens in nasal and fecal specimens using an unbiased high-throughput sequencing approach. *PloS one*. 2009;4:e4219.
- [7] Palacios G, Druce J, Du L, Tran T, Birch C, Briese T, et al. A new arenavirus in a cluster of fatal transplant-associated diseases. *The New England journal of medicine*. 2008;358:991-8.
- [8] Clausen PT, Zankari E, Aarestrup FM, Lund O. Benchmarking of methods for identification of antimicrobial resistance genes in bacterial whole genome data. *The Journal of antimicrobial chemotherapy*. 2016.
- [9] Cannas A, Mazzarelli A, Di Caro A, Delogu G, Girardi E. Molecular Typing of *Mycobacterium Tuberculosis* Strains: A Fundamental Tool for Tuberculosis Control and Elimination. *Infectious disease reports*. 2016;8:6567.
- [10] Lapin V, Mighion LC, da Silva CP, Cuperus Y, Bean LJ, Hegde MR. Regulating whole exome sequencing as a diagnostic test. *Human genetics*. 2016;135:655-73.
- [11] Fisher RG, Smith DM, Murrell B, Slabbert R, Kirby BM, Edson C, et al. Next generation sequencing improves detection of drug resistance mutations in infants after PMTCT failure. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology*. 2015;62:48-53.
- [12] Allen HK, Bayles DO, Looft T, Trachsel J, Bass BE, Alt DP, et al. Pipeline for amplifying and analyzing amplicons of the V1-V3 region of the 16S rRNA gene. *BMC research notes*. 2016;9:380.

- [13] Dortet L, Nordmann P, Poirel L. Association of the emerging carbapenemase NDM-1 with a bleomycin resistance protein in Enterobacteriaceae and *Acinetobacter baumannii*. *Antimicrobial agents and chemotherapy*. 2012;56:1693-7.
- [14] Harmer CJ, Moran RA, Hall RM. Movement of IS26-associated antibiotic resistance genes occurs via a translocatable unit that includes a single IS26 and preferentially inserts adjacent to another IS26. *mBio*. 2014;5:e01801-14.

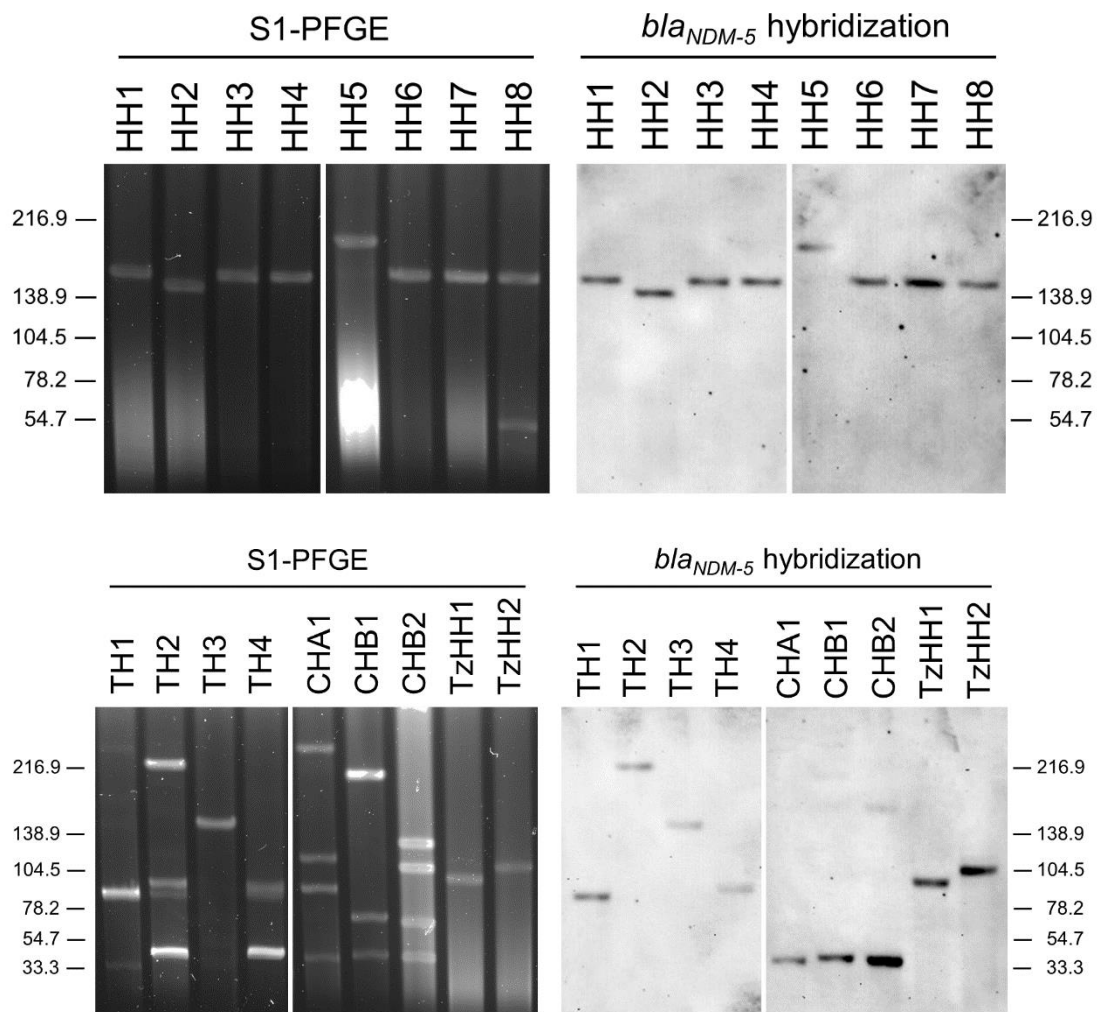
七、圖、表



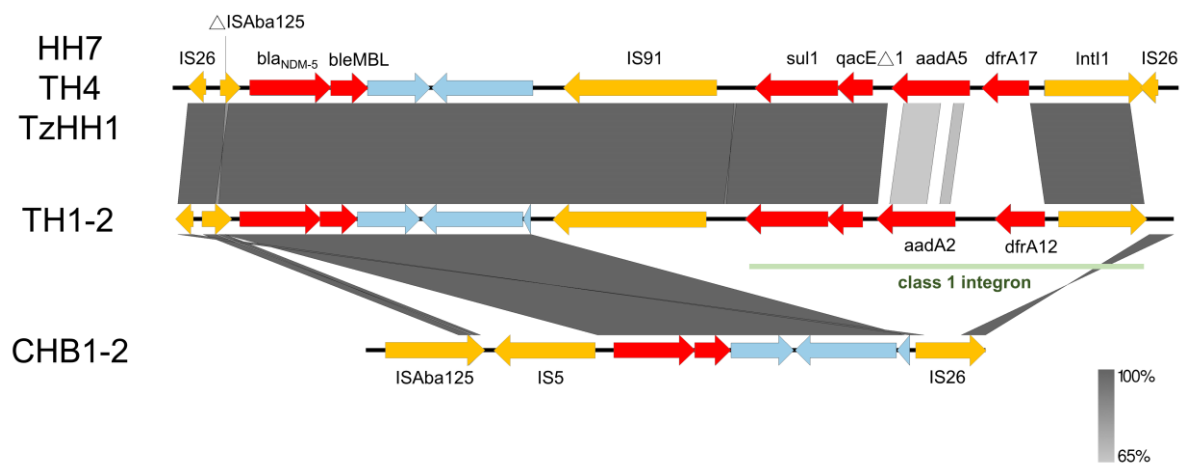
圖一、17株帶NDM-5菌株之醫院及年份分布



圖二、16株帶NDM-5的*E. coli*菌株之PFGE親緣關係圖



圖三、NDM-5 質體分析



圖四、並列比較分析 NDM-5 質體間之差異性

表一、7 株帶 NDM-5 菌株之基因型別及抗藥基因

Strain ^o	ST type ^o	Antimicrobial resistant genes ^o
HH7 ^o	617 ^o	aac(6')-Ib-cr, aadA5, aph(3'')-Ib, aph(6)-Id, blaCTX-M-15, blaNDM-5, blaOXA-1, catA1, dfrA17, mdf(A), sul1, sul2, tet(B) ^o
TH1 ^o	410 ^o	aac(6')-Ib-cr, aadA2, blaCMY-146, blaCTX-M-15, blaNDM-5, blaOXA-1, blaTEM-1B, dfrA12, mdf(A), mph(A), rmtB, sul1, tet(A) ^o
TH2 ^o	405 ^o	aadA2, blaCTX-M-15, blaNDM-5, blaTEM-1B, dfrA1, dfrA12, erm(B), mdf(A), mph(A), qnrS1, rmtB, sul1, sul2 ^o
TH4 ^o	405 ^o	aac(3)-IIId, aac(6')-Ib-cr, aadA2, aadA5, blaCTX-M-15, blaNDM-5, blaOXA-1, blaTEM-1B, dfrA12, dfrA17, mdf(A), sul1, tet(B) ^o
CHB1 ^o	3997 ^o	aac(3)-IIId, ant(3'')-Ia, aph(3'')-Ib, aph(6)-Id, blaCMY-2, blaNDM-5, blaTEM-1B, dfrA14, mdf(A), sul2, tet(A) ^o
CHB2 ^o (<i>E. cloacae</i>) ^o	93 ^o	ARR-3, aac(6')-Ib-cr, aadA16, aadA2, aph(3')-Ia, blaACT-7, blaNDM-5, cmlA1, dfrA12, dfrA27, fosA, qnrB2, qnrS1, sul1, sul3, tet(A) ^o
TzHH1 ^o	410 ^o	aac(3)-IIId, aac(6')-Ib-cr, aadA2, aadA5, aph(3'')-Ib, aph(6)-Id, blaCMY-2, blaCTX-M-15, blaNDM-5, blaOXA-1, blaTEM-1B, dfrA12, dfrA17, mdf(A), mph(A), sul1, sul2, tet(B) ^o

衛生福利部疾病管制署 108 年科技研究計畫

期末審查意見回復

計畫編號：MOHW108-CDC-C-315-134515

計畫名稱：高通量定序(NGS)平台開發與應用

計畫主持人：慕蓉蓉

*修正處在報告中加底線標示

序號	審查意見	主持人回復說明	修正處頁碼
1	如果能萃取 plasmid 可以不必 WGS。	目前朝向只針對 plasmid 做分析，以及使用 WGS 互補各個方式之不足。	
2	結論有參考價值。臺灣基本資料建立也有部份特色。	謝謝委員鼓勵。	
3	Carbapenemase(一)的 CRE 是否有規劃研究?	謝謝委員建議，此計畫為技術建立，未來建立之技術會移至 Carbapenemase(一)的 CRE 分析。	

備註:請將此表單附在期末報告後方，如有修正期末報告內容請註明頁碼，並務必至 GRB 系統完成資料抽換。