

計畫編號：MOHW106-CDC-C-315-114710

衛生福利部疾病管制署 106 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

## 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：慕蓉蓉

協同主持人：

研究人員：林鈺棋

研究人員：葉弘毅

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	4	
計畫內容		
一、前言	5	
二、材料與方法	7	
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。

本計畫已完成建置以 NGS 檢驗不明原因死亡個案之技術，自本 (2017)年 1-10 月共 8 例不明原因死亡個案中，有 5 例檢出病原，病原種類包括病毒及細菌，陽性率達 63%；此外，也完成制訂未知感染原之檢驗流程圖，運用 NGS 作為釐清可能感染原之最後選擇。

本計畫期望運用高通量定序偵測平台建立病原體基因序列分析的流程，適當串連各種分子檢驗技術之優點，以期能即時發現不明與罕見的病原體，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, the NGS platform has been applied in the pathogen detection of sudden death and unknown etiology cases. There were 8 death cases with unknown causes reported from Jan – Oct this year. Five of them has been detected pathogens by using this NGS platform, including bacateria, viruses, the positive rate is 63%. The flow chart of detection of unknown pathogen has also been established in this project to clarify the possible casues for dieases or death by using NGS technology.

In this project, we expect to establish pathogen analysis plate form 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、檢體來源：依本署預防醫學辦公室遴選不明原因或法醫通報且其檢驗結果為陰性之死亡個案，並由各實驗室協助提供檢體或核酸進行 NGS 檢驗分析。
- 2、檢體核酸萃取：利用自動核酸萃取系統 MagNA Pure Compact Instrument (Roche Applied Science) 進行檢體核酸萃取，萃取完成的核酸置於-80°C 冷凍櫃保存。
- 3、基因庫建置及 NGS 定序：利用 QIAseq FX Single Cell RNA Library Kit 進行建庫，並以 Illumina MiSeq 或 MiniSeq 平台進行 NGS 定序。
- 4、資料分析：NGS 定序資料利用 MePIC (Metagenomic Pathogen Identification for Clinical Specimens) [13] 先濾除人類基因序列，再比對 Genbank 之病原資料庫，找出可能病原。
- 5、PCR 反應及定序比對：針對 NGS 檢出病原之個案，分別進行 PCR 反應及定序比對，使用引子如下：
  - (1) 個案編號 1、4 及 7：  
16S-8F：AGAGTTTGATCCTGGCTCAG [14]。  
16S-1492R(I)：GGTTACCTTGTTACGACTT [14]。
  - (2) 個案編號 6：  
HHV7-F：GTTACTTTCAAAAATGTTTGTCCC [15]。  
HHV7-R：GGAAATAGGATCTTTTCAAATTC [15]。



### 三、 結果

#### 1、 NGS 檢驗疑似腸病毒快速死亡個案

本次以通報疑似腸病毒快速死亡個案，經multiplex real-time PCR panel 46個病原檢驗陰性，隨即以NGS技術進行分析。傳染病檢測須掌握檢驗時效，因此本計畫將檢體核酸自行建庫，只需1天(委外廠商須有案件先後次序，則須1周左右)。此次實驗上述通報疑似腸病毒快速死亡個案外，加入一臨床個案經multiplex real-time PCR panel偵測出Toxoplasma gondii及CMV，此臨床個案檢體作為陽性對照組。檢驗結果分析快速死亡個案血清中有TTV(Torque teno virus)(表一)。經分析可得到約四百萬條定序片段數，其中疑似腸病毒感染併發重症個案血清檢體中偵測之Torque teno virus (以下簡稱TTV)序列有6640條，約佔總輸出量0.16%。由於數量夠多，收集定序深度達50倍以上之TTV片段進行組裝，組裝結果獲得單一長度達3,377 bps之TTV組裝片段，經與NCBI基因資料庫進行比對，與Torque teno virus 24 (accession number : AB060597)最為相近，序列相同度達86%。(圖一)

#### 2、 NGS 檢驗不明原因之死亡個案

依本署預防醫學辦公室遴選不明原因死亡個案共7例進行NGS檢驗，個案通報疾病、檢體種類及NGS檢驗結果如表二所示。依NGS檢驗結果，

於7例不明原因死亡個案之檢體中共有4例檢出病原，摘述如下：

- (1)個案編號1：庫賈氏病及日本腦炎之通報個案，於全血、血清及腦脊髓液之混樣檢體中比對出*Propionibacterium acnes*，其序列佔混樣檢體比率約0.1%。
- (2)個案編號4：未知感染源肺炎心肌炎之通報個案，分別於心、肝、脾及肺臟之混樣檢體中發現有*Pseudomonas spp.*序列（佔混樣檢體比率約0.02%）；於右肋膜腔、左/右肺及氣管病毒拭子之混樣檢體中比對出*Streptococcus spp.*（佔混樣檢體比率約6%）及*Stenotrophomonas spp.*（佔混樣檢體比率小於0.01%）；另於脾、腦幹、腦基底、腦脊髓液及血清之混樣檢體中檢出*Staphylococcus aureus*，佔混樣檢體比率約0.01%。
- (3)個案編號6：不明原因腦炎通報個案，於腦脊髓液、鼻咽病毒拭子及血清之混樣檢體中比對到Human herpesvirus的序列，佔混樣檢體比率小於0.01%。
- (4)個案編號7：不明原因肺炎通報個案，於尿液、鼻咽病毒拭子、痰及血清之之混樣檢體中發現有*Serratia marcescens*序列，佔混樣檢體比率約0.1%。

### 3、PCR 定序比對結果

將 NGS 檢驗結果檢出病原之個案編號 1、4、6 及 7 進行 PCR 反應及定序比對，結果如圖二所示，除個案編號 1 未檢出外，其他個案之 PCR 定序比對結果說明如下：

- (1) 個案編號 4：心、肝及脾臟檢體均比對出 *P. putida*；肺臟檢出 *Weissella confusa*；右肋膜腔檢出 *Stenotrophomonas sp.*；左/右肺及氣管之病毒拭子比對出 *S. salivarius*；另於腦脊髓液中檢出 *S. aureus*。
- (2) 個案編號 6：鼻咽病毒拭子中檢出 human herpesvirus 7。
- (3) 個案編號 7：自痰中比對出 *S. marcescens*。

#### 四、討論

- 1、 本計畫收集樣本以人類臨床檢體為主，並採檢體混樣上機。NGS 檢驗結果大部分為人類序列，造成病原佔檢體比率偏低，後續將評估使用更大輸出量的 NGS 試劑，以提升病原檢出率。
- 2、 於疑似腸病毒快速死亡個案之血清（圖一）中發現 TTV 序列，目前 TTV 被認為與肝臟疾病、呼吸道疾病、血液惡性腫瘤及自體免疫疾病有關，但其致病能力仍未清楚。
- 3、 由於 TTV 並非絕對致病原，因此並未納入 Multiplex Real-time PCR 之檢測項目，本次試驗證明 NGS 有不受限於檢測標的之優點，具檢測未知感染源之潛力，但由於自取得樣本至分析完成需時近 2 週，考量 NGS 所需時間、技術、成本（表三及表四）及資料分析能力均遠較 Multiplex Real-time PCR 檢驗方法為高，不建議作為每日例行性檢驗方法，但對於不明原因快速死亡個案之檢驗結果為陰性者，NGS 可作為釐清可能感染源之最後選擇。

## 五、結論與建議

- 1、目前傳染病檢驗方法以血清學及 PCR 核酸檢測為主，惟該等技術受限於檢測標的數量，對於通報多項傳染病之檢驗結果均為陰性者，例如不明原因重症或快速死亡個案，仍無法釐清其可能感染源。次世代定序技術(Next Generation Sequencing，以下簡稱 NGS)相較於 PCR 核酸檢測方法，具有產量大且無檢測標的數量限制之優勢，目前應用在臨床醫學相關領域上已日趨成熟，包括癌症研究及非侵入式胎兒染色體篩檢等。爰此，為提升本中心傳染病檢驗能力，以本計畫建置 NGS 平台以應用於傳染病之檢驗，並制訂未知感染源之檢驗流程(圖二)，以利遵循。
- 2、以陽性對照組之臨床檢體做為 NGS 檢測正確性之指標，可得到與例行性檢驗方法 Multiplex Real-time PCR 之檢測結果相同(表一)，足可證實 NGS 檢測之正確性。
- 3、本計畫自本署預防醫學辦公室遴選之不明原因死亡個案中共檢出 4 例帶有病原，後續已安排於本年 12 月初與各防疫醫師進行病例討論會，以利瞭解病例與檢出病原間之可能關聯性。

## 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] Takeuchi F, Sekizuka T, Yamashita A, Ogasawara Y, Mizuta K, Kuroda M. MePIC, metagenomic pathogen identification for clinical specimens. *Japanese journal of infectious diseases*. 2014;67:62-5.

[14] Turner S, Pryer KM, Miao VP, Palmer JD. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *The Journal of eukaryotic microbiology*. 1999;46:327-38.

[15] Zaravinos A, Bizakis J, Spandidos DA. Prevalence of human papilloma virus and human herpes virus types 1-7 in human nasal polyposis. *Journal of medical virology*. 2009;81:1613-9.

## 七、圖、表

樣本來源 (通報編號)	發病及 死亡日期	檢體種類	例行性 檢驗結果*	NGS 檢驗結果	說明
疑似腸病毒感染併 發重症個案 (1060700002430)	發病：106/4/27 死亡：106/5/2	肛門拭子	陰性	陰性	106/5/10 晨會指 示本中心對本個案 進行NGS檢驗
		鼻咽拭子	陰性	陰性	
		血清	陰性	Torque teno virus (TTV)	
不明原因腦炎個案 (1050500004240)	發病：105/11/15 死亡：無	血清	CMV	CMV	臨床檢體陽性對照
		腦脊髓液	Toxo	Toxo	

\*註：「例行性檢驗結果」為Multiplex Real-time PCR之結果：CMV：Cytomegalovirus，Toxo：Toxoplasma gondii。

表一、疑似腸病毒感染併發重症個案 NGS 檢測結果



比對結果	定序片段數 (百分比)	
總數	4,041,234	(100%)
人類	3,511,523	(87%)
微生物類	18,567	(0.46%)
<i>Toxoplasma gondii</i> (Toxo)	86	對照組
Human herpesvirus (CMV)	58	對照組
<b>Torque teno virus (TTV)</b>	<b>6,440</b>	<b>(0.16%)</b>

### De novo assembly

單一長度達3,377 bps之片段，經比對與Torque teno virus 24-SAa-01序列相同度達86%。



圖一、NGS序列分析及病原序列組裝

個案編號	通報疾病	檢體種類	NGS檢驗結果	PCR定序比對結果
			(佔檢體比率)	
1	庫賈氏病	全血	<i>Propionibacterium acnes</i> (0.1%)	未檢出
		腦脊髓液		未檢出
	日本腦炎	血清		未檢出
		腦脊髓液		未檢出
		血清		未檢出
2	弓形蟲感染症	全血	未檢出	
		血清		
3	恙蟲病	全血	未檢出	
		血清		
	鈎端螺旋體病	全血		
		血清		
4	未知感染源肺炎心肌炎	心	<i>Pseudomonas spp.</i> (0.02%)	<i>Pseudomonas putida</i>
		肝		<i>Pseudomonas putida</i>
		脾		<i>Pseudomonas putida</i>
		肺臟		<i>Weissella confusa</i>
		腎	未檢出	
		腦		
		心包囊病毒拭子		
		左肋膜腔		
		右肋膜腔	<i>Streptococcus spp.</i> (6%) <i>Stenotrophomonas spp.</i> (>0.01%)	<i>Stenotrophomonas sp.</i>
		左肺病毒拭子		<i>Streptococcus salivarius</i>
		右肺病毒拭子		<i>Streptococcus salivarius</i>
		氣管病毒拭子		<i>Streptococcus salivarius</i>
		脾	<i>Staphylococcus aureus</i> (0.01%)	未檢出
		腦幹		未檢出
腦基底	未檢出			
腦脊髓液	<i>Staphylococcus aureus</i>			
血清	未檢出			
5	Q熱	全血	未檢出	
		血清		
6	不明原因腦炎	腦脊髓液	<i>Human herpesvirus</i> (>0.01%)	未檢出
		鼻咽病毒拭子		<i>Human herpesvirus 7</i>
		血清		未檢出
7	不明原因肺炎	尿液	<i>Serratia marcescens</i> (1%)	未檢出
		鼻咽病毒拭子		未檢出
		痰		<i>Serratia marcescens</i>
		血清		未檢出

表二、不明原因死亡個案之NGS檢驗結果

所需時間	現行操作流程
檢體處置及 <b>自建庫</b> 建庫樣本寄送 上機前QC 上機 資料分析及確認	1-2 日 1-2 日 1-2 日 1-2 日 2-4 日
*總 計	6-12 日

\* 試劑及儀器設備均能及時配合之預估日數

表三、NGS 檢驗時效分析

檢驗項目	每件檢體平均檢測費用(元)*
Multiplex Real-time PCR	1,500
<b>NGS</b>	
本中心自行上機	
批次檢測10件檢體(中輸出量)	6,100
批次檢測20件檢體(高輸出量)	4,700
委外上機	
批次檢測10件檢體(中輸出量)	7,500
批次檢測20件檢體(高輸出量)	6,050

\*註：檢測費用包含萃取核酸及檢驗項目所需試劑。

表四、成本經費分析



