

出國報告（出國類別：研習）

「食品生物基因辨識技術」出國報告

服務機關：衛生福利部 食品藥物管理署

姓名職稱：王鈺婷 薦任技士

派赴國家：日本

出國期間：105 年 7 月 24 日-105 年 8 月 6 日

報告日期：105 年 11 月 1 日

摘要

次世代定序自開發以來進展飛快，因其具有高通量、非目標性偵測之特性，數日內可獲得完整及準確之核酸序列資料，而廣用於各項研究領域，由近年文獻顯示其應用於食品生物檢測亦由未知微生物組成鑑定，至基因組成複雜之基因改造轉殖核酸片段分析與新興物種鑑別等，皆凸顯其優勢。本次研習拜訪日本國立感染症研究所 (National Institute of Infectious Disease, NIID) 位於戶山廳舍的 Pathogen Genomic Center 之黑田 誠教授 (Dr. Matoko Kuroda) 實驗室，主要目的為學習如何使用次世代定序技術(Next Generation Sequencing, NGS) 及其分析平台，包括核酸萃取純化、片段化(fragmentation)、全基因體基因庫(DNA library)製備、增殖 PCR)、illumina MiniSeq 上機、序列分析及註解(annotation)。研習期間筆者以過去承辦之檢驗案件與黑田教授討論，黑田教授針對次世代定序儀之優勢及盲點深入解析。此行自黑田教授之經驗分享及其實驗室同仁不吝指導次世代定序技術之分析策略，筆者收獲相當豐富且彌足珍貴，同時加強日方對於本署研究能力之良好印象，成功建立與日方友好之資訊溝通管道。

目 次

摘要.....	2
本文	
一、目的.....	4
二、過程.....	5
三、心得及建議.....	14

一、 目的

近年國際經貿交流頻繁，食品快速流通，近年國際間因不明生物性物質引發人體健康之爭議時有所聞。面對如此檢驗難度之提升，為保障國人食品安全，本署須建立一套全面性、解析度高且快速之分析平台，以提升未知物質之判明率。次世代分析技術因具高通量、高解析度及快速等特性，可在數日內完成全基因體定序，近年國際間知名研究機構及政府檢驗機構皆已具備該技術，本署已建置次世代定序儀一套(Life Ion Torrent Personal Genome Machine)，並運用於如高鐵炸彈客一案之突發案件或乳酸菌查驗登記檢驗案，但面對龐大之序列資訊，如何正確又快速解讀是一大考驗。有感於此，筆者此行參訪日本NIID，向熟捻於次世代定序技術、並建立數種次世代定序技術數據分析平台之黑田教授實驗室學習，實際操作現今市占率較高之 illumina 次世代定序分析平台及其後續生物資訊分析，除蒐集日本近年研發資訊及相關方法，將相關經驗應用於食品生物相關業務及研究工作，並藉此建立合作夥伴關係，建立國際聯繫管道與人脈，提升本署食品生物之分子檢驗的研究開發能力與國際同步，鞏固我國在國際食品檢驗技術之領先地位與能見度。

二、 過程

本次至日本NIID戶山廳舍內病原體基因組分析研究中心 (Pathogen Genomic Center) 位於新宿區的戶山廳舍，由黑田 誠博士 (Dr. Makoto Kuroda)所領導之實驗室研習2周，由加藤健吾博士指導筆者有關illumina MiniSeq平台之實驗技術學習及後續序列註解分析，藉此行機會，筆者亦將102年承辦之高鐵炸彈客未知物分析一案與黑田教授討論，交流雙方在NGS定序技術之經驗及分析策略，探討NGS定序技術於食品生物檢測上之定位。



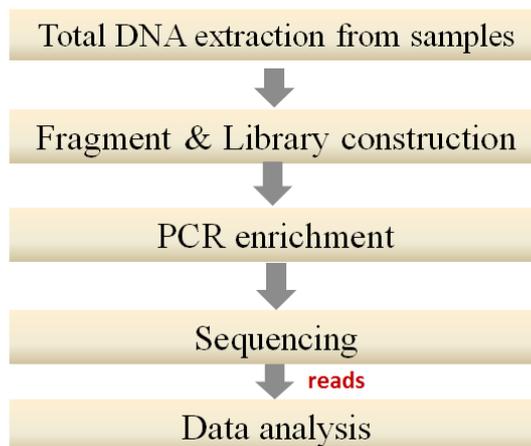
NIID 戶山庁舎外觀



筆者與黑田教授實驗室合影

1. NGS 定序技術之研習：

NGS experimental process



NGS 定序技術操作流程

檢體均質溶於 Tirs-EDTA 緩衝液後，經 achromopeptidase 分解外
壁並添加 RNase 去除 RNA，由 1% SDS 造成蛋白質變性後，續以

phenol/chloroform/isopropanol 粗略純化 DNA，續由 Promega Wizard SV gel & PCR clean-up system 再次純化，並以 Qubit® dsDNA BR Assay 精確測量萃取 DNA 之濃度。將 DNA 調整適當濃度，取出 50 ng 以 QIAseq FX DNA library kit 製備 NGS 基因資料庫，包含以限制酶將 DNA 分子均一片斷化(Fragmentation)，DNA 片段接上供組裝過程辨識之 adapter，續經 Beckman Ampure Bead 純化之，經由 Hi-Fi PCR 放大各 DNA 片段以完成 NGS 基因庫製備。

以往筆者在 NGS 基因庫製備過程，會以 e-Gel (下圖左) 選取特定片段 DNA 進行後續純化，然而加藤博士卻以傳統電泳(下圖右)取代 e-Gel，操作上較為不便且需時較長，此令筆者不解。經黑田博士解釋過去實驗室亦大量使用 e-Gel 以選取 DNA 片段，後來發現因 e-Gel 以電泳時間決定選取 DNA 片段大小，故若因電流不穩或是電泳時間稍有誤差，容易錯失最佳 DNA 片段大小，反而成為整個實驗流程最大誤差來源，因此之後改用傳統電泳，經 Ethidium bromide 染色後精確選取所需之理想 DNA 片段，以免除後續分析上過短片段 DNA 序列所造成之判讀干擾。



e-Gel



研習當日之電泳情形

此次試驗以近 2 年推出之 illumina MiniSeq 次世代定序儀為基因庫解序平台，相較於同廠牌、全球普及率較高之 MiSeq，同樣以 Paired-end、單邊讀長 150 bp 為分析模式，MiniSeq 價格較為親民，上機時間亦較短，適用於數目較少、基因體不大且須時效性之檢體分析。下圖分別為 illumina MiniSeq 外觀、試劑套裝卡匣、上樣晶片及依照電腦圖示指令之上機流程：



Illumine MiniSeq 外觀



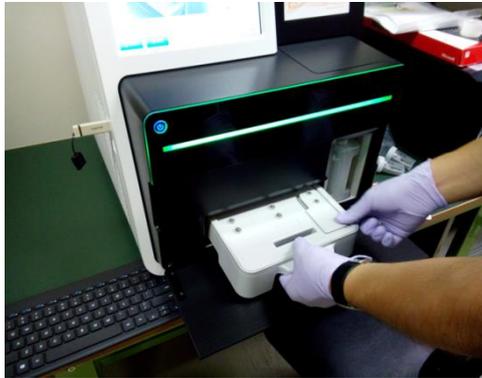
MiniSeq 上機用試劑套組



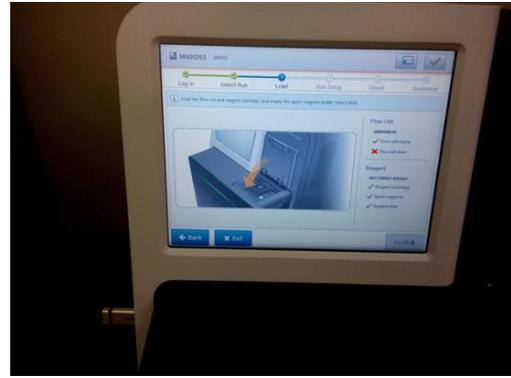
上機晶片



進行上機前儀器前清洗



放入試劑套裝卡匣



放入上樣晶片

因該平台藉 CCD camera 偵測晶片內螢光訊號，故晶片表面不能有毛屑或其他雜質殘留，因此加藤博士在上機前以拭鏡紙將晶片之訊號偵測窗反覆擦拭並面對日光燈以確認晶片潔淨度(如下圖)。



對光確認晶片潔淨度

定序結果判讀分析

經由約莫 20 小時完成定序，所產生之 raw reads 以該平台相對應之 illumina 內建軟體確認該次定序品質，確認符合分析所需品質後續執行序列拼裝 (contig assembly) ，藉由 Tablet NGS 確認 contig assembly 情況，以及找出因定序覆蓋率過低而需以 PCR 確認之區域，以確認完整序列資訊。由於該次試驗之檢體為臨床細菌 DNA 檢體，加藤博士續以黑田博士實驗室研發之生物資訊軟體 Annotation of Microbial Genomic Analysis (AMIGA) 進行分析，同時利用 Metagenomic Pathogen Identification for Clinical Specimens (MePIC) 平台，利用元基因組分析 (Megagenomics) 原理確認該檢體所含之菌種及純度，確認該樣品是否在培養過程中遭受污染。

同時，黑田實驗室中的山下明史博士介紹其自行架設之 UHmin tool 線上網站，網站內為其自行開發之生物資訊分析路徑，其中亦包含可供 Sanger 定序較為方便使用之分析軟體，山下博士亦針對網站內容細心講解，令筆者收穫不少。山下博士於過去曾作為交換學者至台灣清華大學交流，對於台灣的印象相當好，亦表示很開心能夠在工作實驗室中再次遇到來自台灣的研究同好，筆者亦在研習期間受到山下博士諸多照顧，相當感恩！另一位生物資訊專才之關塚剛史博士亦

撥空指導筆者使用如 GViewer 及 FigEasy 等序列比對用繪圖軟體，關塚博士亦致力於黑田教授實驗室內序列資料庫之建置及繪圖軟體之開發。與山下博士及關塚博士之談話，2 位皆表示現行 NGS 分析軟體皆有其限制，唯熟悉 Linux 系統才能打破藩籬，靈活運用各項生物資訊分析技巧，更能符合實際研究需求。

2. 以 102 年高鐵炸彈客一案與黑田教授討食品中未知物檢驗策略

筆者於 102 年承辦高鐵炸彈客一案，當時有 3 件自嫌犯住處搜出標有「肉毒」或「肉桿」之未知物送至本署檢驗，分析其中是否含有嫌犯聲稱之肉毒桿菌或其他生物性有害物質：該檢體呈現深灰色~棕色、黏稠且具惡臭，針對該 3 件未知物檢體，本署運用已建置之 Life Ion Torrent Personal Genome Machine 次世代定序儀進行定序分析，並同時以傳統微生物培養、病原菌專一性 real-time PCR 及小鼠試驗，確認該檢體是否含有肉毒桿菌。在確認該次檢體定序數據品質符合分析需求後，筆者將原始定序數據(raw reads)及其預測之 Open reading frame，直接與美國 National Center for Biotechnology Information (NCBI) 基因體資料庫配對，尋找是否含有肉毒桿菌或其他有害物質相關之核酸片段，續利用軟體將組裝序列(assembly)後以 BLAST 比對，收集上述比對到病原微生物之 raw reads 及組裝序列(contig)，確認其

序列專一性。由配對(mapping)或比對(BLAST)方式皆未找出肉毒桿菌神經毒素基因，但卻發現部分短片段序列的確比對到肉毒桿菌基因體，後來筆者針對該序列設計 primer，以實驗室現有之肉毒桿菌及親緣性相近之 *Clostridium sporogenes* 進行 PCR 分析，發現僅 *Clostridium sporogenes* 呈現陽性訊號，後發現 NCBI 資料庫內並無 *Clostridium sporogenes* 全基因體定序資料，因此僅以 NGS 定序片段直接與 NCBI 資料庫比對僅能比對出肉毒桿菌基因體序列，無法忠實呈現其為 *Clostridium sporogenes* 之結果。由 MG-RAST 進行 Megagenomics 分析，結果顯示該檢體中基質為米，超過 9 成比例組成為細菌，其中更以 *Lactobacillus* spp. 為大宗。由該 3 件檢體之傳統微生物培養、real-time PCR 及小鼠試驗，亦皆為肉毒桿菌陰性。

筆者將上述案情及分析策略向黑田教授說明，黑田教授亦同意以其實驗室建立之 AMIGA 平台及 MePIC-megablast 再次分析，結果顯示該 3 件檢體確實主要由一群發酵細菌及酵母菌組成，未偵測到病原微生物專一性核酸片段。藉此機會筆者與黑田教授請教若以 NGS 定序技術該如何正確判讀數據，黑田教授表示未知物中有害物質成分可能極為微量，因此以 raw reads 會比經組裝所得之 contig 更能忠實呈現檢體原始樣貌，並建議以 nucleotide 資料庫比對會比 genome 資料

庫、protein 資料庫更具全面性，較能避免因 genome 資料庫資料有限而無法順利比對。且 NGS 定序技術因以核酸為偵測目標，若未知物之核酸回收率過低，將大幅影響 NGS 定序結果之代表性，且對於未知物中是否含有病原微生物，並非經資料庫比對就能定論，要先確認比對到的序列是否具該病原微生物專一性 (pathogen specific)，很多時候比對到的都是供微生物生理運作所需之一般體基因 (somatic gene)，並非致病基因，且即使比對到致病基因，仍需由 PCR 及 Sanger 定序確認之，以免除因序列組裝錯誤而造成判讀錯誤；倘若傳統微生物培養結果與 NGS 定序結果不符，須各自檢視該次傳統培養及 NGS 定序之試驗確效情況才能決定判讀方向，且 NGS 定序僅能顯示核酸序列數值，而無法反映出檢體內否仍有病原微生物存活。

三、心得與建議

首先，感謝本署長官給予此次研習機會，容筆者利用 2 周時間赴日學習次世代定序分析技術，並熟悉運用各項生物資訊分析平台進行後續核酸註解及其他比對分析。

NGS 定序技術雖自開發至今不到 10 年時間，卻已運用於臨床、農業等生物學研究，近年使用該技術進行分析研究之相關文獻更絡繹不絕，其提供一個高通量之快速定序平台，對於全基因體分析而言，價格與時間成本上皆遠低於傳統 Sanger 定序，且可免除傳統定序上仰賴 PCR 產物之分析而有資訊片斷化之缺點，NGS 定序技術於短於 1 周內提供檢體全基因體序列資訊，再與資料庫之參考序列 (reference sequence) 比較，可直接由數解註解分析進一步發現非預期之檢體特性，此令研究人員趨之若鶩，預計未來將有一定之普及度及分析市場之重要性。

NGS 定序技術之新穎性深具魅力，但其在檢測領域上之定位卻與其他分子生物技術如 PCR、real-time PCR 及 digital PCR 大大不同：筆者於過去承辦高鐵炸彈客之經驗，發現對於檢測人員，如已知檢測目標，real-time PCR 及近年發展出的 digital PCR 皆能於數小時內由儀器直接輸出可供解讀之數據，專一性強且敏感度甚至可達單一核酸

分子，適用於常規性檢測或是快速篩檢試驗，以滿足具時效性之業務需求；然而 NGS 定序提供原始之樣品完整序列資料，龐大且繁雜，分析人員須對該次任務已有相當清楚之邏輯分析策略，且鑒於當今各大核酸資料庫內序列資料並非涵蓋所有物種，其編寫概念亦受限於具某些特性之序列資料，如僅針對單一資料庫進行比對，則容易造成解讀錯誤，因此對於序列解讀，需要同時比對多種資料庫加以確認，並佐以其他分析平台驗證其結果，方能確認最終檢測結果。

NGS 定序技術自樣品製備至上機，雖然過程繁雜，需注意之細節不少，但對於一位熟悉分子生物技術之操作人員，幾次練習後自然不成問題；然而後續面臨龐大之序列資料，如同密碼般令人茫然無所解，對於非生物資訊專才人員，不啻是惡夢一場。此行出發前，筆者於幾次研討會上碰到幾位傑出的研究人員皆無奈於空有數據難解其義之窘境，在赴日參訪黑田教授研究室期間，幾位自分子生物學跨足生物資訊分析領域之博士後研究員亦反映學習過程相當辛苦，相較於 bench work 單線式實驗流程上手容易，生物資訊分析在邏輯上屬多面向，除非相當熟捻於各項序列註解分析平台，更甚者須運用 Linux 語言，才能隨心所欲解讀 NGS 定序所得序列資料，建議本署未來可編列經費，招募序列分析相關之生物資訊分析專才人員，為本署充實足

夠之分析量能，方能將次世代定序之優點體現並活用於食品生物相關研究業務。

筆者與黑田教授討論過程中，黑田教授表示以短片段定序原理為主之 NGS 定序平台，雖以高通量特性加強序列數據之連續性及正確度，但如何將短片段定序結果正確拼裝至關重要，且受限於偵測原理，目前市面上各項次世代定序平台皆無法完全克服重複序列 (repeat sequence) 或同元序列 (homopolymer) 上的定序錯誤或無法定序，為此需後續進行耗時的 PCR 及傳統定序分析，大幅降低利用 NGS 定序預期之數據產出效能。有鑑於此，NIID 將於今年冬天引進第三代定序技術平台—以單一核酸分子進行即時長片段定序，藉此可直接克服上述序列特性所造成的定序錯誤。本署食品生物業務範圍極廣，其中基因改造作物 (動物、植物) 更因外來基因插入過程中亦帶入相當量之重複序列，筆者於 101 年參訪國外基因改造生技公司時已目睹其將第三代定序技術運用於外來轉殖基因座之組成、插入點分析及其於基因體內拷貝數分析，本署作為基因改造食品查驗登記之主管單位，未來亦希望能以該技術全面性分析基因改造食品之特性，以強化本署為民眾食品安全把關之量能。

此行赴日，折服黑田教授及其實驗室同仁在工作上的專注、負責及執著，平日黑田教授主動指導筆者有關核酸註解分析，並關切筆者於日本生活所需，對於筆者的提問，黑田教授總願意撥空與筆者詳加解說並討論，由於黑田教授及實驗室同仁不吝指導，筆者無論在 NGS 定序技術上的學習、經驗分享、台日文化交流皆有長足進步，此行筆者亦為黑田教授及實驗室同仁介紹本署業務及在食品檢測上的專業表現，黑田教授亦對本署印象良好，並期待往後雙方能有更多互動或合作機會。