

中華大學生物資訊學系系統開發專題報告

利用基因組序列比對建立北蕉及其抗黃葉病品系台五與台七之鑑別序列 Genome-wide study (GWS) of Asian Northern banana cultivar and its wilt-resistant breeding varieties Ty5 and Ty7.

專題組員: B10120001 郭岱伶、B10120018 林佩君、B10120039 鄭謹慧、
B10120003 黃元蓉、B10120004 文聖閔
專題編號: PROJ2015-BIOINFO-10108
指導老師: 劉志俊老師、張慧玫老師

1. 摘要

政府加入 WTO 後，農產品貿易競爭激烈，使各國對植物種源和優良品種加強智慧財產權觀念。香蕉黃葉病在全球肆虐，不但造成經濟損失，也導致東南亞以香蕉為主食的國家糧食短缺的問題。台灣香蕉研究所從北蕉培養出具有中高度抗病性的台蕉 5 號、與台蕉 7 號等優良品系，為申請「種苗權」品種專利以保護優良蕉種，我們利用基因組序列比較的技术，尋找北蕉與台蕉 5 號、與台蕉 7 號香蕉品系差異序列做分子標誌，以成為作物品種專利的鑑定工具。

一般我們分辨物種，最普遍方式是依外型性狀(phenotype)來分辨。由於台蕉 5 號與台蕉 7 號為同源於北蕉之近繼代營養品系(clone variants)變化而來，使得這三種品種間外表的相當酷似不易分辨時，就必須以利用序列差異，即基因型(genotype)的差異性，之分子鑑定較能區分以便記錄、保存下來，使物種得到保障。

2. 簡介

(一) 黃葉病

黃葉病(*Fusarium wilt*)是由尖孢镰刀菌古巴專化 (*Fusarium oxysporum f.sp. cubense*) 真菌所引起出現在香蕉根部

的病症。此真菌會透過分泌毒素，引發香蕉細胞程序性的死亡。

(二) 香蕉品種

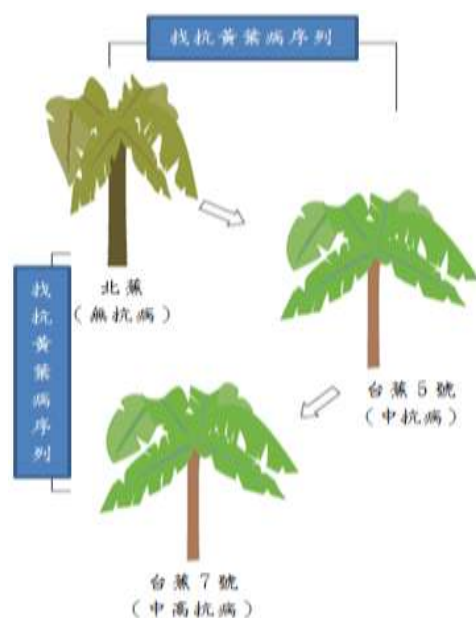
北蕉、V34 為一般常見香蕉品系做不抗性對照。台蕉 5 號與台蕉 7 號是台灣香蕉研究所先後培養出抗病性的優良品系。本實驗以上品系之蕉苗、蕉葉樣品皆由台灣香蕉研究所提供。

(三) 研究動機

香蕉黃葉病，固然會造成經濟損失與糧食短缺的問題，有別於一般植物病蟲害的襲擊，由於香蕉是從親代植株以無性生殖出芽生長而來，缺乏遺傳變異性，(1)一旦染病，同一品系會幾乎全面性地同時滅亡。(2)潛在病株分芽再傳下一代的隔離技術，會有漏網之魚，防疫很難。因此像台蕉 5 號、與台蕉 7 號這樣經過長時間篩選並拓寬遺傳變異性，培養達到穩定並具有高度抗病性的優良品系就很珍貴。

普遍且廣泛使用的分子鑑定方式有 RFLP、RAPD、ADLP 及 SSLP...等，常出現的親子鑑定、DNA 指紋。但以上所提及的分子鑑定方式，工作量大、技術步驟繁瑣以及對同品系先後變種之辨別度不佳，不適合本計畫使用，我們利用基因組地毯式比對，設計差異性

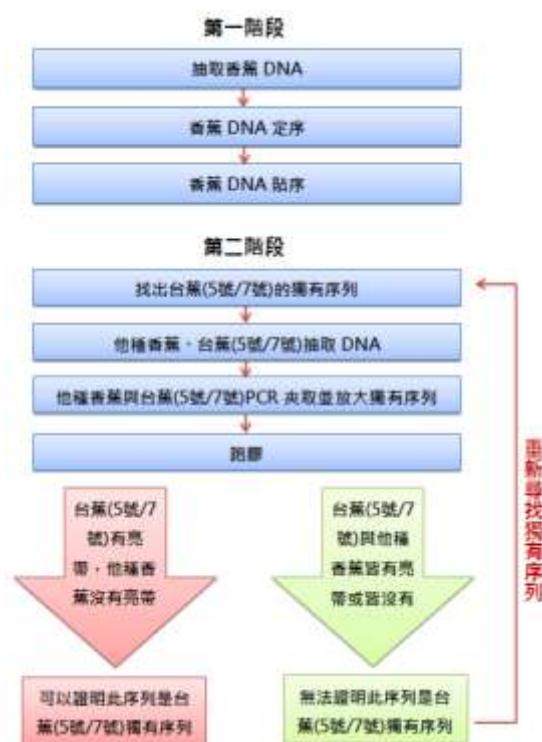
專用引子且選用聚合酶連鎖反應，準確性、專一性皆非常優良。本計畫就是研發如何協助台灣香蕉研究所分子鑑定序列可作為維持苗種的純度、異品種 (off-type)、混雜率、均一度等品種純度的依據。進一步可保護優良蕉種與智慧財產，以申請「種苗權」品種專利。我們以「V34」做為常見香蕉品系做一般不抗性香蕉之背景值對照，「北蕉」做不抗性之種源對照，與抗病性台 5 與台 7 的優良品系的為分析目標樣本，我們利用基因組序列比較的技術，發展出品系差異序列，以有效的鑑別不抗病性與抗病性香蕉品系[2]。



3. 專題進行流程

我們將來自香蕉研究所提供白化「V34」與「北蕉」品系之蕉苗 (為了排除葉綠體 DNA 的干擾) 抽取數個樣本，做全基因組 DNA 的萃取，再去除香蕉特殊膠質的問題，成功取得高品質的基因組 DNA，送基隆米克斯廠商做全基因組讀序，之後利用軟體作序列組合和貼序成全基因組，進行「北蕉」與台 5、台 7 全基因組比對有差異

性，再以「V34」序列作背景扣除，找出台 5、台 7 專一 Indel 序列。這些差異性 (Indel) 序列經過挑選設計成引子 (primer)，再做聚合酶連鎖反應 (Polymerase Chain Reaction, PCR) 對「北蕉」與台 5、台 7 樣品做鑑別測試 (如圖一)。



圖一 實驗流程

4. 實驗材料與方法

(一) 實驗藥品

CTAB、 $(\text{NH}_4)_2\text{SO}_4$ 、Tris、EDTA、NaCl、 β -mercaptoethanol 等都是從 SIGMA-ALDRICH CO. 購得，所使用的藥品均為試藥級或以上等級藥品。

(三) CTAB 植物基因體 DNA 萃取法

白化的「V34」與「北蕉」品系之蕉苗取 12g 種苗，將 45 ml 的 CTAB(cetyl trimethyl ammonium bromide)[1] [0.1M Tris HCl (pH8.0), 1.4M NaCl, 20mM EDTA, 2% (w/v) CTAB]加入

1% (v/v) 的 β -mercaptoethanol，以液態氮磨碎。放入水浴槽 55 °C/15 min，低溫離心 8500 rpm 10 min 後，抽取上清液。加入 1/10 體積的醋酸銨並且加 95% 酒精至滿管，放入 -20 °C 沉澱 1 hr。沉澱離心將酒精倒掉，以 70~75% 酒精清洗、離心再將酒精倒掉。自然烘乾後回溶在 2 mL TE 溶液。為去除樣本的蛋白質，加入等體積的苯氯仿混合後離心 13000 rpm 1 min，重複至中間層變透明為止。離心後以 95% 酒精 -20°C 靜置沉澱 1 hr。之後離心 13000 rpm 1 min 將酒精倒掉，以 70~75% 酒精洗淨離心後再倒掉酒精。自然烘乾後回溶 200 μ L TE。加入 2 μ L 10mg/mL RNase A 放入水浴槽 50 °C 20 min，之後加入 2 μ L 100 mg/mL Proteinase K 放入水浴槽 50 °C 20 min。樣本純化後加入苯氯仿清洗 2 次在做沉澱的步驟。自然烘乾後再加入 1000 μ L TE 回溶。其中 CTAB 溶液活性只能維持 2-3 天，為使 CTAB 能更有效率保護 DNA，故須於實驗當天泡製溶液，泡製比例如【表 1】，其中所含 β -mercaptoethanol(β -ME) 是用來去除香蕉特殊膠質。

表 1 CTAB 泡製比例

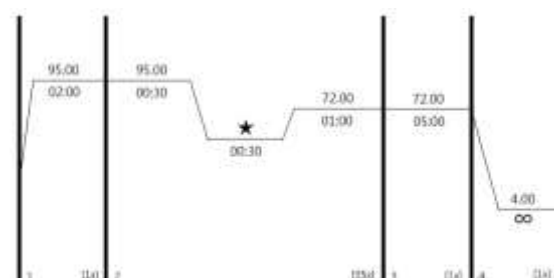
CTAB	PVP	β -ME
0.5 mL	0.02 g	2.5 μ L
5 mL	0.2 g	25 μ L

(四) 設計 Primer

由序列比對結果中找尋可能區域，在從中找出適合的序列片段做為設計 Primer 及 Probe，在此我們將使用的程式為 Vector NTI、PerlPrimer。

(五) 聚合酶連鎖反應 (PCR)

取出 1 μ L Genomic DNA (0.5 μ g/ μ L) 被當作模板，在 90°C 去二級結構 5 分鐘。PCR 連鎖反應將以 30 μ L 在 PCR 反應，20 mM 引子對各 1 μ L、3 μ L 的 10 \times PCR buffer、1 μ L VioTaq DNA 聚合酶 (5 units/ μ L)、15 μ L DEPC 處理過的水。PCR 連鎖反應程式設定如下：以 94°C 作用 5 分鐘模板單股化，再以 55°C 作用 30 秒、72°C 作用 1 分鐘、94°C 作用 30 秒，這三相溫度進行 30 個循環，經最後延長步驟才反應完全。所得擴大大子 (Amplicon)，以 2.0% 洋菜膠作電泳分析，經 Ethidium Bromide 染色後切出 DNA 擴大大子，再以 DNA 純化試劑 (Viogene)，得到純化 DNA 擴大大子。片段經過克隆再讀序後，將以 NCBI BLAST program 作進一步序列分析。



★：依各 primer 特性，合適溫度會有所不同。

圖 2 PCR 溫度程序設定

星號部分的溫度，依各 primer 的特性，合適溫度會有所不同。

4. 主要成果

(一) genomic DNA 品質檢測結果

由於香蕉的全基因體是 300 Mb，用 0.8 的洋菜膠和 1kb 的 ladder 跑膠，結果 DNA 的 band 帶會在 10kb 上。依跑膠結果分析 DNA 是否純淨



無降解斷裂【圖 3】。

Qubit：染料試劑，DNA 樣本濃度高時 (0.05–5 $\mu\text{g}/\text{mL}$) 用 dsDNA BR 試劑；DNA 樣本濃度低時 (1–500 ng/mL) 用 dsDNA HS 試劑。

調製標準 DNA Assay【圖 4a, b】及調製待測樣本 DNA Assay【表 2】，依步驟放入機器後，即可測定 DNA 濃度。在測定過後機器會產生一份詳細資訊以便記錄

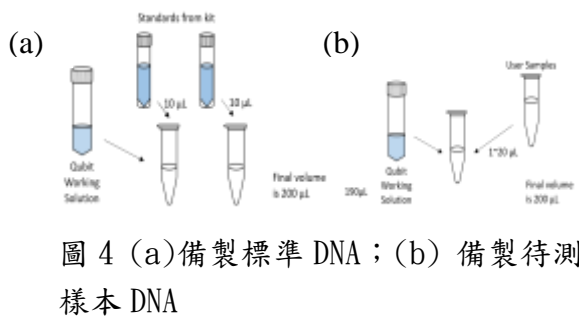


圖 4 (a) 備製標準 DNA；(b) 備製待測樣本 DNA

表 2 Qubit 定量後資料

Name	Date/Time	Assay Conc.	Units
#####	2014/10/4 23:05		225 ng/ml
Stock Conc	Units	Assay Type	
4.5 ng/µl		dsDNA HS	
Sample Vol	Dilution Factor	Std 1 RFU	Std 2 RFU
1	200	16.14	3793.76

(二) 基因體定序結果

北蕉、台五與台七定序結果，有相似大小的預估讀序資料與接近的 GC%。

表 3 北蕉、台五與台七初始定序結果

Sample Name	Clean Reads	Clean bases	Read length (bp)	Q20(%)	GC(%)
sample1 北蕉	433,591,508	54,198,938,500	125	94.37%	39.97%
sample2 台五	412,678,494	51,584,811,750	125	94.65%	41.29%
sample3 台七	442,413,666	55,301,708,250	125	94.50%	40.42%

(三) 依照 NCBI 香蕉公布版貼序

基因體貼序主要有三種方法：BWA, SOAP2, Bowtie2。(1) 建立 Bowtie2 所需之貼序索引。(2) 進行 Bowtie2 基因體貼序。(3) SAM 格式貼序結果轉為 BAM 格式。(4) BAM 格式貼序結果排序。(5) 建立 BAM 格式貼序結果索引。如【圖 5】。我們將香蕉的基因體組序後，會到 BLAST 進行核對，我們所使用的範本是 Musaceae[2]，核對的目的在於檢查基因體的組序是否有誤。

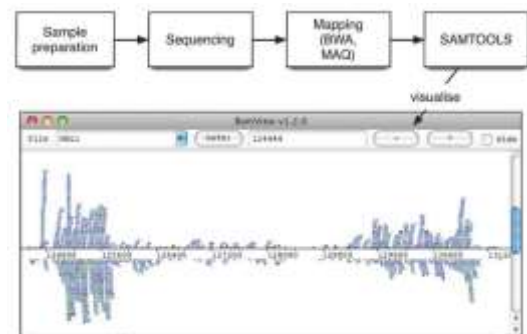


圖 5 貼序示意圖暨貼序過程

(四) 基因體貼序結果

北蕉、台五與台七定序結果，有相似大小的貼序資料。

表 4 北蕉、台五與台七貼序結果

品種	V34	北蕉	台五	台七
染色體	Size (bp), GC%	Size (bp), GC%	Size (bp), GC%	Size (bp), GC%
chr1	27,583,757 38.83%	27,585,734 38.85%	27,587,716 38.88%	27,587,028 38.88%
chr2	22,063,848 39.40%	22,065,050 39.43%	22,066,613 39.46%	22,066,447 39.46%
chr3	30,485,541 39.37%	30,487,100 39.39%	30,489,420 39.41%	30,488,727 39.41%
chr4	30,063,158 38.77%	30,065,707 38.78%	30,067,048 38.81%	30,066,472 38.81%
chr5	29,391,890 38.92%	29,393,881 38.93%	29,395,240 38.96%	29,393,969 38.95%
chr6	34,915,272 39.14%	34,916,436 39.16%	34,918,681 39.19%	34,917,832 39.19%
chr7	28,630,624 39.23%	28,632,070 39.25%	28,633,953 39.28%	28,633,526 39.27%
chr8	35,452,171 39.01%	35,454,482 39.04%	35,456,432 39.08%	35,456,170 39.08%
chr9	34,162,990 38.89%	34,165,729 38.92%	34,283,660 39.43%	34,166,920 38.94%
chr10	33,678,821 39.22%	33,681,640 39.25%	33,683,031 39.28%	33,682,250 39.28%
chr11	25,527,507 39.11%	25,528,558 39.13%	25,530,601 39.15%	25,529,975 39.16%
全基因體	331,955,579	331,976,387	332,112,395	331,989,316

(五) 基因體序列差異分析結果

將北蕉與台蕉 5 號比對差異後共發現 51,967 處 Indel 基因變異，如（表 5）；將北蕉與台蕉 7 號共發現 47,798 處 Indel 基因變異，如（表 6）。

表 5 台五的 Indel

品種	Chromosome	Indel
Taiwan 7	1	2923
Taiwan 7	2	3014
Taiwan 7	3	4906
Taiwan 7	4	3824
Taiwan 7	5	6377
Taiwan 7	6	4027
Taiwan 7	7	4175
Taiwan 7	8	3467
Taiwan 7	9	5885
Taiwan 7	10	4588
Taiwan 7	11	4612
		47798

表 6 台七的 Indel

品種	Chromosome	Indel
Taiwan 5	1	3209
Taiwan 5	2	3221
Taiwan 5	3	5344
Taiwan 5	4	4237
Taiwan 5	5	6664
Taiwan 5	6	4418
Taiwan 5	7	4603
Taiwan 5	8	3829
Taiwan 5	9	6407
Taiwan 5	10	5102
Taiwan 5	11	4933
		51967

(六) 候選引子之鑑定測試



圖 6 抽驗香蕉流程。

在實驗中我們找出適合分辨蕉種的引子後，便利用 PCR 實驗來測驗未知香蕉品種【圖 6】。如【圖 7】經由跑膠確認 DNA 片段的長度為原先設計的長度但不唯一，確實分辨北蕉及台蕉 5 號。而圖片上也很明顯得看出，引子編號 INS01 是有效擴充台蕉 5 號 DNA 但北蕉沒有這段序列。



圖 7 確認 INS01 引子之 PCR 跑膠結果

M: 100 bp marker ; 1: 台五 A ; 2: 台五 B ; 3: 北蕉 A ; 4: 北蕉 B 。

在設計引子的過程，為使引子達到最佳化的夾取動作，我們測試引子的最佳化溫度， Tw5_INS01 引子在 51 ° C 時，可擴大台蕉 5 號呈現 200-300bp 的兩條條帶，是北蕉所沒有的序列【圖 8】。

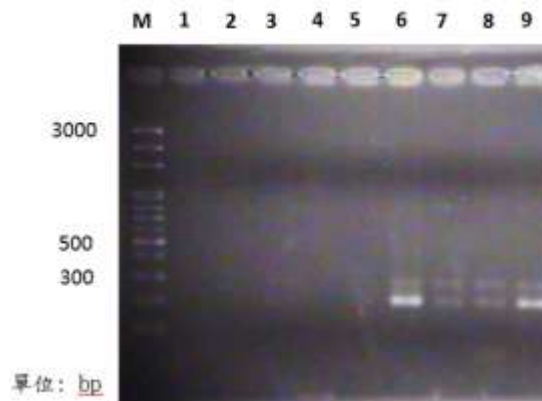


圖 8 Tw5_INS01 引子之 51 ° C，PCR 作用後台蕉 5 號夾取到兩條條帶

M: 100 bp marker ; 1: 北蕉 A ; 2: 北蕉 B ; 3: 北蕉 C ; 4: 北蕉 D ; 5: 無 ; 6: 台五 A ; 7: 台五 B ; 8: 台五 C ; 9: 台五 D 。

當溫度修正為 53° C 如【圖 9】，夾取到兩條條帶，只剩下唯一專一性的條帶且正確的位置上。

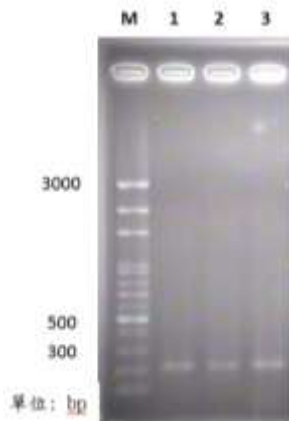


圖 9 Tw5_INS01 引子之 53° C，PCR 跑膠結果

M: 100bp marker ; 1: 台五 A ; 2: 台五 B ; 3: 台五 C。

5. 評估與展望

香蕉不僅對台灣來說是相當重要的一項農產品與經濟來源，對國外一些亞熱帶國家亦同。專家在香蕉還是種苗時就可依據香蕉之外型的判斷台五與台七的外表相異之處，但是一般人卻無法在種苗時判斷。但若完整外型來判斷品種的話，需要花約兩年的時間。所以若要申請種苗權的話，進行基因體的鑑別序列來鑑定品中相對較快。

我們針對不同品種香蕉各自專有的 marker 序列去做辨別，以便日後能簡單利用成本低廉且準確度高的 PCR 技術產生條帶來鑑定香蕉品種與確定品質的依據，進一步保護台灣所研發的栽培香蕉種與維護研發者的權力及利益等。

研究實驗已成功將四種品種之相交的基因體 DNA 抽出來。並以達成了定序標準，也同時成功的將四種香蕉定

序完成，再定序之後順利的組成香蕉之基因體的草圖。在使生物資訊的相關軟體進行比對後，目前得到了台五與北蕉比對出來的 Indel，共有 51967 個位點以及台七與北蕉比對出來的 Indel，共有 47798 個位點。再從中找尋差異序列來設計 primer 後，我們先將設計出來的 primer 與 Musaceae (範本香蕉) 進行 BLAST 比對，比對的結果如果一對 primer 是在相同之基因片段上的話即可進行 PCR 的步驟來做進一步的確認了。

然而我們和一般植物的品種鑑定有何不同？普遍實驗都是做種、亞種間的鑑定，然而我們的難度在於要辨別演化而來的兩代。如同辨別是否同父母簡單，但辨別誰是大哥、誰是小弟則有一定的困難度。不過我們不單單只是要分辨北蕉、台蕉 5 號，更希望我們找到的 primer 中，或許未來可以從這些 primer 之中找到專抗黃葉病的 DNA 基因片段，進而研發出黃葉病疫苗造福農民。

擁有這項技術，我們能推展更多台灣的農產品使其發揚國際，擁有科學的資料能夠更確保台灣的未能站在世界上屹立不搖。不僅為台灣帶來經濟的效益，更向全球宣示台灣不僅擁有品種，更擁有保護品種的科學能力，這將為台灣利下一個全新的里程碑，讓我們的未來擁有無限可能。

6. 結語

由於一開始需要大量濃度的 DNA 做定序，因此採取 CTAB 法抽取 DNA，得到的 DNA 量值大、品質穩定，但缺點是費時長。進行 PCR 的時候需要抽取許多不同種類的香蕉 DNA，CTAB 法

過於耗時，因此改利用 Kit 抽取 DNA，此方法速度快且適用於 PCR，能在一樣的時間內獲取更多種不同的 DNA，同時也因實驗時間短，降低了 PCR 受到雜質干擾或污染的可能性。

台蕉五號比起北蕉，更有對抗黃葉病的特性，香蕉為世界重要糧食，全球香蕉都有染上黃葉病的隱憂，因此保護種苗權就變得極為重要。

利用引子夾取特定序列的特性，來正確判別北蕉與台蕉五號。由於北蕉與台蕉五號序列太相似，跑 PCR 時只要溫度有偏差，就可能導致引子誤夾，我們便要找出最適當的溫度，其溫度在多次實驗後發現確認不會誤夾序列，便利用此引子配合 PCR 設定溫度，可順利地分辨未知香蕉為北蕉或台蕉五號，依此分辨機制可將台灣農作物放心的推向世界的舞台。

Françoise Carreel, Olivier Garsmeur, Benjamin Noel, Stéphanie Bocs, Gaëtan Droc, Mathieu Rouard, Corinne Da Silva, Kamel Jabbari, Céline Cardi, Julie Poulain, Marlène Souquet, Karine Labadie, Cyril Jourda, Juliette Lengellé, Marguerite Rodier-Goud, Adriana Alberti, Maria Bernard, Margot Correa, Saravanaraj Ayyampalayam, Michael R. Mckain, Jim Leebens-Mack, et al. **“The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants,”** 2012 Nature 488, 213 – 217.

7. 銘謝

過程中很感謝兩位指導老師，在實驗、資訊領域上，都給我們很大的資源。也帶我們到校外的實驗室學習很多必要的實驗基本功力，讓我們後續在實驗室中的觀念變得更加清晰且得心應手。

8. 參考文獻

- [1] Sue Porebski, L. Grant Bailey, Bernard R. Baum **“Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components,”** March 1997, Volume 15, Issue 1, pp 8-15
- [2] Angélique D’Hont, France Denoeud, Jean-Marc Aury, Franc-Christophe Baurens,