

國立臺灣大學生物資源暨農學院農藝學系



碩士論文

Department of Agronomy

College of Bioresources and Agriculture

National Taiwan University

Master Thesis

開發水稻之單一核苷酸多型性品種鑑別系統

Development of Rice Variety Identification System using Single
Nucleotide Polymorphic Markers

周佳霖

Chia-Lin Chou

指導教授：胡凱康 博士

Advisor: Kae-Kang Hwu, Ph.D.

中華民國 104 年 7 月

July, 2015

誌謝

從來沒想過自己會去念一個研究所，也從來沒想過這一念竟然念了五年，一路走來靠著許多人的幫助，要感謝的太多了，就感謝天吧。


能完成這本論文，由衷感謝胡凱康老師，總是看起來很輕鬆優雅地在學業和工作上給我極多的幫助，條理分明的思緒總是命中問題的核心，超級有耐心討論與解決各種疑難雜症，正向思維與不放棄每一個學生令我敬佩和感動。也感謝我的口試委員林順福老師、張孟基老師與楊佐琦場長，在論文完成初期給予許多寶貴的建議，親切認真地讓口試在輕鬆的氣氛下獲益良多，口試結束收到鉅細靡遺修改與建議的紙本，看得見的用心更是讓我充滿無限感激。

學校生活學到學術上的知識以及訓練我們有條理地思考和組織，成就這本論文最重要的骨架，無憂無慮的生活也是最令我懷念的，謝謝 Energy、笨魚和東東學長就像大鴨子一樣帶領我們衝往實驗室與田間的各種工作；芳華學姐和美玲的各種照顧和幫助；簡直以系館為家的林延諭和林昭京，能和你們組成 Lins group 一起生活一起討論一起大笑真的太讚啦；工友、宇珊、R 啟蒙老師高驥和跟我一起崩潰也要一起畢業的顏維萱，不論在學校或是離開學校都受到你們好多幫助，如果沒有你們我真的會被撕裂成兩半；系排的夥伴黑哥、晶晶、炮炮、糗哥、松鼠、佳穎、龍、心碎以及數不清的學妹們是學生生涯很大的心靈支柱，和你們一起打球感覺像可以浮出水面換一口新鮮的空氣，然後就有力量繼續潛入無盡念書與報告的大海；另外，感覺總像同學的研究所學長姐景雯張、聰明人、糗哥、大根、蒲蒲、康樂夫妻、拉炮、高高、科后、派大星、感覺總像學弟妹的研究所同學賣樞、婷玫、淨惠、癢癢、愛陵、欣歡、小朱和真的是同學的許栗子，有你們的陪伴超好的！

行政與研究工作大幅提昇我的文字表達能力與拓展視野，使我在寫作論文時幾乎沒有生不出來的煩惱。謝謝市府工作的夥伴陳柏宏課長、麗華姐、建威叔叔、芳小姐、雅君、盈如、祥洲大哥、家輝、詩軒、柏儒、永青、秀慧、汶宗……等等太多人了，除了帶領我從一個菜鳥變成一個還 OK 的人，也讓我知道面對困難最好的方法與態度；在論文寫作期間種苗場鍾文全課長、小毛、明燕姐、哲仁、惠如、秀滿姐的各種 cover，恩威並施地督促我完成這本論文，超級給力的助理智容和詩禴姐則讓我不用擔心實驗的部分，使這本論文可以如期完成；而病理室是我研究所生涯的開端，也是我在種苗場工作永遠的娘家和充電站，生技課的大小小營造出一個舒適的工作氣氛，因為有你們讓我覺得人生超美好。

也謝謝我的家人以及像家人一樣的馬仔和氣勢，總是給我無限的支持，讓我可以無後顧之憂地做我想做的事，也因為有你們我才能堅定地完成這本論文，謝謝你們。


中文摘要



本研究從水稻 44K SNP 晶片分析得到 35,746 筆包含 413 個廣泛來源的水稻基因型分析結果，經由篩選成功讀值率、SNP 類型、對偶基因位置、次要對偶基因型頻度等資訊，得到表現較佳的 130 個 SNP 基因座，再經實驗測試，先選擇可判讀且於次族群中遺傳距離較遠的品種間具多型性的 47 個基因座後，對本研究收集到的 49 個目標品種基因型分析，挑選出適用於我國優良水稻推薦品種組合的 10 個 SNP 組成水稻 SNP 品種鑑別套件。使用的 10 個 SNP 標誌在 49 個包含我國優良水稻推薦品種的水稻品種間之平均 PIC 值為 0.402，於 49 個品種間共 1,176 個兩兩組合之中，僅無法鑑別‘臺南 16 號’ / ‘越光’、‘臺東 30 號’ / ‘臺東 33 號’、‘臺中私糯 1 號’ / ‘臺中私糯 2 號’ 等背景相似度極高的三組品種。從篩選品種中選取遺傳距離最大且不具品種權保護的 3 個品種組成標準品種，其涵蓋 100% 的對偶基因型。

關鍵字：水稻、品種鑑別、單一核苷酸多型性

ABSTRACT



The goal of this study is to develop a set of single nucleotide polymorphism (SNP) markers for identification of Taiwan rice varieties. At first, 130 candidate SNP were selected based on call rates, SNP types, allele positions and minor allele frequency (MAF) from 35,746 SNP datasets which derived from 413 rice varieties analyzed by rice 44K SNP microarray chip. By laboratory analysis around 130 SNP and 8 genetically distinct japonica Taiwan rice varieties, we narrowed down the candidate SNP into 47 SNP loci. After genotyping 49 rice varieties with these 47 SNP markers, we found that a selected set of 10 SNP markers was capable to distinguish 1,173 pairs out of total 1,176 variety pairs except 3 pairs of closely related varieties, ‘TN16’/‘Koshihikari’, ‘TT30’/‘TT33’ and ‘TCSG1’/‘TCSG2’. The average of polymorphism information contents (PICs) is 0.402 for the selected set. We also selected 3 varieties as reference varieties that they have largest genetic distance within 49 Taiwan rice varieties and all without protect by Plant Variety Right. The reference varieties cover 100% of the alleles of 49 rice varieties used in this study.

Keywords: rice, variety identification, single nucleotide polymorphism (SNP)

目錄

口試委員會審定書	#
誌謝	i
中文摘要	ii
ABSTRACT	iii
目錄	iv
圖目錄	vi
表目錄	vii
重要名詞縮寫對照表	viii
第一章 前言	9
第二章 前人研究	12
一、 SNP 探勘與分析	12
二、 SNP 用於品種鑑別	15
三、 水稻 SNP 資料庫	15
四、 水稻的核心收集	16
五、 CoreHunter	18
第三章 材料與方法	20
一、 試驗材料	20
二、 水稻 DNA 萃取、定量與品質檢定	21
三、 聚合酶連鎖反應與電泳分析	21
四、 SNP 基因座挑選	22
五、 SNP 分子標誌設計	23
六、 評估 SNP 分子標誌與架構水稻 SNP 品種鑑別套件	23
七、 驗證套件標誌組合	24
八、 建立標準品種	24
第四章 結果與討論	29
一、 SNP 基因座挑選結果	29
二、 驗證基因座挑選組合	30
三、 架構臺灣水稻優良品種 SNP 鑑別套件	30



四、 標準品種	31
五、 開發階段對偶基因專一性引子之選擇	31
六、 可能的分析系統	33
七、 無法區別的臺灣品種	34
第五章 結論.....	45
參考文獻.....	48
附錄.....	55
附錄 1、以 47 個標誌對 49 個水稻品種基因型分析結果.....	55





圖目錄

圖 1、CoreHunter 輸入資料之格式	28
圖 2、Allele specific (AS)系統增幅結果	36
圖 3、兩兩品種間多型性量統計圖	37
圖 4、水稻 SNP 品種鑑別套件於染色體位置示意圖	40
圖 5、比較側翼條帶大小差異對 AS 系統之影響	42
圖 6、AS-primer 方向及序列造成之結果差異	43
圖 7、AS-primer 之序列配對錯誤有無於梯溫 PCR 之差異比較	44
圖 8、開發策略流程圖與研究結果概要	47



表目錄

表 1、水稻 SNP 資料庫.....	19
表 2、參試水稻品種.....	26
表 3、9 個候選 SNP 組合之兩兩品種間基因型差異數分佈.....	38
表 4、水稻 SNP 品種鑑別套件之 SNP 資訊.....	39
表 5、AS 系統與 Tetra primer ARMS 系統之比較表.....	41

重要名詞縮寫對照表



A-NE：最近區域收集系平均距離 (Average distance between each accession and the nearest entry)

ARM：香稻 (Aromatic)

AS：對偶基因專一性 (allele specific)

AS-marker：對偶基因專一性標誌 (allele specific marker)

AS-primer：對偶基因專一性引子 (allele specific primer)

CV：Coverage of alleles in the core subset 演算法 (計算所有基因型占全部基因型的覆蓋率，挑選最廣泛涵蓋所有基因座中所有可能基因型的核心收集)

E-E：平均遺傳距離 (Average genetic distances between entries)

E-NE：最近加權平均距離 (Average distance between each entry and the nearest neighbouring entry)

IND：秈稻 (Indica)

JPN：粳稻 (Japonica)

MAF：次要對偶基因頻度 (Minor Allele Frequency)

MR：Modified Rogers 演算法 (將每一個等位基因座視為獨立，計算兩兩收集系間的歐幾里德遺傳距離)

PCR：聚合酶鏈鎖反應 (Polymerase Chain Reaction)

PIC：多型性資訊含量值 (Polymorphism Information Contents)

RAD：限制酶位點標定 (Restriction-site Associated DNA)

SNP：單一核苷酸多型性 (Single Nucleotide Polymorphisms)

SSR：簡單重覆序列 (Simple Sequence Repeats)

TEJ：溫帶粳稻 (Temperate japonica)

TILLING：定向誘導基因组局部突變 (Targeting Induced Local Lesions in Genomes)

TRJ：熱帶粳稻 (Tropical japonica)

第一章 前言

水稻 (*Oryza sativa* L.) 是世界主要的糧食作物之一，依據聯合國糧食組織 () 的統計，2013 年全球水稻栽培面積為 1.6 億公頃，次於小麥 (2.1 億公頃) 和玉米 (1.8 億公頃)，是世界第三大栽培面積的作物；而其年產量 7.4 億噸，僅次於 10.1 億噸的玉米，為世界第二大產量的作物。亞洲是水稻最主要的生產地，全球超過 90% 的水稻於此區生產，世界前五大的水稻生產國亦皆位於亞洲，分別為中國、印度、印尼、孟加拉和越南。

我國的稻米種植記錄在新石器時代大坌坑文化考古物中即有大量碳化水稻的出土，這些水稻的米粒粒型較野生稻粒型更偏向私稻粒型，顯示最遲在距今約莫五千年前，已經有將野生稻馴化為栽培稻的紀錄，而隨著考古年代漸近，出土的碳化稻粒型更增加爪哇稻之熱帶粳型稻的遺跡 (王，2007)。水稻在臺灣的大面積種植與推廣始於明朝末年，漢人及荷蘭人在臺灣活動引進私稻，並因糧食自給之政策而大面積墾地種植，而後日據時期日本人將優良粳稻品種引入臺灣並成功馴化栽種，經過有系統的雜交育種與推廣，使我國人民食用米的習慣由私稻改變為粳稻。臺灣光復後各農業試驗場致力於水稻的育種改良與推廣，新品種不斷推陳出新，截至目前為止已育成 302 個私、粳稻品種，其中 197 個品種尚保存稻種可供利用 (呂與呂，2010)。

光復後為保持國內農作物優良品種遺傳特性與避免品種間之混雜造成生產品質下降，確保國內糧食作物優良種子之供應，我國於 1959 年建立良種繁殖檢查制度，列入檢查的作物有水稻、玉米、高粱、落花生等。自良種繁殖制度實施以來，成效最顯著的即為水稻，各育種機關選育出具有推廣價值的新品種後，經向農委會申請登記並命名後始得推廣繁殖；具有推廣價值之新品種，由育成機關供應基本種 (種原) 與繁殖原原種，並按原原種、原種及採種三級繁殖檢查制度生產優良種子，以供應農民種植。自三級繁殖檢查制度實施以來，已對我國糧食作物優良種子之供應，產生極大的助益 (林，2011)。

水稻良種繁殖之種子檢查規範主要依據「臺灣地區農作物種苗檢查須知」與國際種子檢查協會印行之「國際種子檢查規則」，辦理田間、室內及貯藏檢查，內容涵蓋水分測定、潔淨度分析、發芽試驗、種與品種的鑑別等。在品種鑑別部分，由於臺灣水稻育種常使用育成之優良品種做為雜交親本，導致品種間歧異度小，外觀形態相近，不易由外觀區別。使用分子標誌輔助品種鑑別，可有效避免人為誤判，提高品種鑑別可信度。此外，我國糧食管理法及其子法明文規定「包裝或容器標示稻米品種者，其內容物含標示之品種含量，應達百分之八十以上」，相關市場銷售糧食米抽查及檢驗辦法也於 103 年 12 月 18 日開始施行。專業檢查人員雖可由水稻植株或稻穀外觀形態鑑別品種，但脫殼糙米或是精白米粒提供的外觀形態資訊過少，已無法由外觀判別品種。分子標誌不受樣品外表形態的限制，並提供相對於外觀鑑別準確且客觀的數據，是市售米品種鑑別不可或缺的工具。

作物育種研究室目前已開發三版水稻 SSR 品種鑑別套件，第一版套件包含分散於水稻的 10 對染色體上的 12 個簡單重覆序列 (Simple Sequence Repeats, SSR) 標誌，其 SSR 為 3 或 4 個鹼基對重覆 11 至 20 次之完美重覆(perfect repeats)，於 192 個水稻品種中多型性資訊含量值(Polymorphism Information Contents, PIC) 皆大於 0.6，引子標定 4 色螢光後，可單次多重 (multiplex) 聚合酶鏈鎖反應 (Polymerase Chain Reaction, PCR) 完成增幅，並可於單次多重螢光毛細管電泳 (multiplex capillary electrophoresis) 完成基因型分析 (資料未發表)。第二版套件改善第一版套件因重覆次數過多或重覆單位過小導致痕跡條帶 (stutter band) 偏高，可能造成誤將同質結合基因座誤判為異質結合之情形，選取 4 鹼基對為單位、重覆次數 6 至 15 次的 SSR，組成包含 12 個分散於水稻 12 條染色體的 SSR 品種鑑別套件，也可於單次多重 PCR 增幅後，單次多重螢光毛細管電泳完成基因型分析 (柯，2014)，在第二版套件篩選 SSR 標誌時未考慮因重複單位提高與重複次數降低造成之 SSR 在品種間的多型性程度，使用的篩選族群未考量水稻次族群結果與臺灣梗稻品種間親緣關係較近，導致選得的 SSR 標誌之多型性多半來自私稻，造成在梗稻品種間之鑑別力不足。第三版套件則是綜合第一與第

二版套件的經驗，並針對臺灣粳稻品種篩選適用之標誌，結果以 12 個 SSR 標誌組成可於單次多重 PCR 增幅並於單次螢光毛細管電泳完成基因型分析之第三版水稻 SSR 品種鑑別套件（資料未發表），目前此套件已實際應用於我國水稻品種之分子檢測上。

SSR 標誌具有共顯性、高再現性、高多型性及具物種專一性等優勢，然 SSR 之特性亦常造成生物體內重覆次數改變或側翼序列點突變，降低分析穩定度。單一核苷酸多型性（Single Nucleotide Polymorphisms, SNP）相對較穩定，且於基因體中更廣泛存在（reference needed），並具分析系統多樣化、部分系統間可轉換使用及可自動化判讀提高分析效率等優勢，隨著 SNP 資料庫陸續公開，序列資訊取得更趨容易，使越來越多研究使用 SNP 做為分子的工具。本研究是作物育種研究室開發水稻 SNP 品種鑑別套件的起點，我們參考實驗室前人開發水稻 SSR 分子鑑別套件之經驗，逐步建立 SNP 篩選模式與優化的策略，以篩選具高多型性與高讀值機率（call rate）潛力且能鑑別「臺灣優良水稻推薦品種」的水稻 SNP 組合為本次研究的目標，期望本次開發的品種鑑別套件可實際運用於輔助水稻三級繁殖制度之種子檢查，也期許未來能拓展到開發可適用於世界水稻品種鑑別的平台，成為此系列研究的最終目標。

第二章 前人研究



一、SNP 探勘與分析

探勘 SNP 的方法主要有比對序列差異、利用 SNP 物性與化性上的差異加以區分及以酶切方式辨識 SNP 三個大方向。

由比對序列差異探勘 SNP 是最多人使用的方法，序列的來源可能來自現有資料庫或解序特定片段序列，除了模式植物阿拉伯芥 (Schmid *et al.*, 2003) 外，在大麥 (Kota *et al.*, 2001)、番茄 (Labate and Baldo, 2005) 等作物或是原核生物如分支桿菌 (*mycobacterium*) (Wang *et al.*, 2010) 等，都有成功探勘 SNP 例子。隨著科學發展演進，以次世代定序於全基因體探勘 SNP，已從大豆 (Lam *et al.*, 2010)、小鼠 (Frazer *et al.*, 2007)、西瓜 (Guo *et al.*, 2013) 等材料獲得數以萬計的 SNP；值得一提的是，Baird 等學者建立限制酶位點標定 (Restriction-site Associated DNA, RAD) 技術，以限制酶處理樣品並接上條碼序列 (barcode) 後，再以次世代定序集中定序限制酶切位相鄰之序列，能同時完成多個樣品的高密度 SNP 探勘與基因型鑑別，改善全基因體定序過於廣泛導致得到過多累贅序列的問題，並可於單次分析多個樣品 (Baird *et al.*, 2008)。

藉由偵測 SNP 造成 DNA 物性與化性不一致的表現探勘 SNP，如單股構型核苷酸多態性 (Single-strand Conformational Polymorphism(SSCP)-SNP) (Bertin *et al.*, 2005) 利用 SNP 造成單股 DNA 構型的差異探勘 SNP；溫度梯度膠體電泳 (Temperature Gradient Gel Electrophoresis, TGGE) (Rosenbaum and Riesner, 1987)、變性梯度膠體電泳 (Denaturing Gradient Gel Electrophoresis, DGGE) (Ge *et al.*, 1999) 及高效能液相層析 (Denaturing High Performance Liquid Chromatography, DHPLC) (Wolford *et al.*, 2000) 以溫度或化學藥劑使 DNA 變性 (denature)、復性 (renature) 後產生結構差異以探勘 SNP；Zn²⁺-cyclen-PAGE (Kinoshita-Kikuta *et al.*, 2002) 藉由 DNA 帶電荷差異影響電泳結果偵測 SNP；高解析熔點法 (High Resolution Melting, HRM) (Reed and Wittwer, 2004) 則利用

SNP 造成雙股 DNA 變性溫度差異偵測 SNP。

以酶切探勘 SNP 也是常用的方法，利用限制酶辨識並裁切 SNP 差異序列組合，如限制酶切片段長度多型性 (Restriction Fragment Length Polymorphism, RFLP)、酶切增幅片段多型性 (Cleavage Amplified Polymorphic Sequence, CAPS)；利用特異性核酸內切酶可辨識並水解單股 DNA 的特性，處理復性後產生的異源雜合雙鏈 (heteroduplex)，偵測 SNP，如裂解酶切片段長度多型性 (Cleavage Fragment Length Polymorphism, CFLP) (Lyamichev *et al.*, 1993)、定向誘導基因組局部突變 (Targeting Induced Local Lesions in Genomes, TILLING) (McCallum *et al.*, 2000)。

探勘 SNP 的系統同時也可做為 SNP 分析、檢測的工具，在已知 SNP 位置與序列的前提下，除了前文提到的方法，也可藉由專一性引子、探針等策略檢測 SNP 基因型。在專一性引子的應用方面，Soleimani 等學者利用引子在 3'端較具選擇性的特性，設計一組包含 SNP 側翼引子對 (flanking primers) 與 3'具 SNP 選擇性的對偶基因專一性引子 (allele specific primer, AS-primer) 組成之引子組合，單次可增幅一個 SNP 基因型 (Soleimani *et al.*, 2003)。Latorra 等學者則以甲基化修飾對偶基因專一性引子的構型，建立鎖核酸技術 (Locked Nucleic Acid, LNA)，優化對偶基因專一性 PCR (allele specific PCR, AS-PCR) 之效果 (Latorra *et al.*, 2003)；Ye 等學者於 SNP 的側翼序列設計成對的引子，並於 SNP 位置設計在 3'端具對偶基因專一性且兩個基因型互為方向相反的引子，共計 4 支引子，進行 tetra primer ARMS PCR，可於 1 個 PCR 反應同時偵測 2 個 SNP 基因型 (Ye *et al.*, 2001)；Orum 等學者利用黏合性較 DNA-DNA 佳的肽核酸 (Peptide nucleic acid, PNA)，與 DNA 序列互補為 PNA-NDA，阻礙 PCR 增幅，用以偵測 SNP (Orum *et al.*, 1993)；Nikiferov 等學者將增幅目標位置之其中一引子磷酸化，於增幅後高溫變性並以外切酶處理，保留磷酸化之單股，以標定不同顏色螢光的三磷酸雙脫氧核苷酸 (dideoxynucleotide triphosphates, ddNTP)，延展 (elongation) DNA 後，藉由比色法判讀 SNP 基因型 (Genetic Bit Analysis, GBA) (Nikiferov *et al.*,

1994)；而 Ahmadian 等學者則是以焦磷酸測序 (pyrosequencing) 原理為基礎，藉由偵測 DNA 延展時 dNTP 釋出磷酸根產生的能量激發光，成功判讀 SNP 基因型 (Ahmadian *et al.*, 2000)。以探針偵測 SNP 的方法更多且廣泛地被應用，Huang 等學者以專一性探針上標定不同類型抗原，與外加抗體結合後，偵測剩餘抗體量以判讀 SNP 基因型 (Huang *et al.*, 2002)；Nilsson 等學者以兩端分別為 SNP 基因型專一性序列與其鄰近序列之倒置序列設計 Padlock 探針，於雜合至目標位置時形成環狀構造，當 SNP 基因型與探針相符時，DNA 連接酶可使探針黏合成環狀，並以滾環式複製 (Rolling circle replication, RCR) 增幅後判讀 (Nilsson *et al.*, 1997)；Ullman 等學者利用氧分子受到光能激發後，經由載體間電子傳遞機制，建立 AlphaScreen 技術 (Ullman *et al.*, 1994)，Beaudet 等學者分別在 SNP 專一性引子或探針上標定電子傳遞載體，當成功增幅或雜合時，載體間距離足夠近，將進行電子傳遞並釋放不同波長的光波，成功開發可使用 AlphaScreen 偵測 SNP 的 12 個專一性引子與 7 個探針 (Beaudet *et al.*, 2001)；以螢光探針偵測 SNP，其好處是不需膠體電泳，可直接接收螢光強度進行結果判讀，也廣泛應用於科學研究。螢光系統偵測 SNP 主要的方法有 TaqMan 系統與分子信標 (Molecular beacon, MB)，TaqMan 探針包含二個螢光標定，5'端為發光子 (reporter)，3'端為消光子 (quencher)，探針與模板互補序列結合時，呈穩定狀態，當引子延伸時因為 *Taq* DNA 聚合酶的外切功能 (exonuclease) 將與模版穩定結合的探針水解，使得消光子脫離發光子而能接受外界能量激發螢光，藉由偵測螢光顏色，可判讀目標序列的 SNP 基因型；MB 探針則是在游離狀態時呈穩定狀態，當與目標序列結合時，激發螢光 (Mhlanga and Malmberg, 2001)。

此外，近年來解序技術進步，隨著高通量資訊的產生與生物資訊學的發展，有許多與農業相關物種的 SNP 晶片陸續且快速被開發，如水稻 (Ebana *et al.*, 2010; Tung *et al.*, 2010; Zhao *et al.*, 2010; Yu *et al.*, 2014)、玉米 (Ganal *et al.*, 2011)、櫻桃 (Peace *et al.*, 2012)、番茄 (Sim *et al.*, 2012)、蘋果 (Chagné *et al.*, 2012)、鮭魚 (*Salmo salar*) (Dominik *et al.*, 2010)、豬 (Ramos *et al.*, 2009)、鯰魚 (Liu

et al., 2011)、茶 (Fang *et al.*, 2014) 等，以此已知 SNP 序列之晶片分析未知樣品的 SNP 基因型，可快速且大量的獲得 SNP 資訊，許多這些資訊被放於公開的平台供學者們取用，使科學的研究更加快速有效率。



二、SNP 用於品種鑑別

SNP 於基因體中大量存在，其標誌分析系統彈性大，目前已有許多物種以 SNP 建立品種鑑別套件。除了前文已提到的晶片系統套件外，也有許多非晶片系統的套件被開發如橄欖 (Reale *et al.*, 2006; Consolandi *et al.*, 2007)、分支桿菌 (Wang *et al.*, 2010)、番椒 (Jung *et al.*, 2010)、葡萄 (Cabezas *et al.*, 2011)、肉牛 (Heaton *et al.*, 2002)、柚 (Wu *et al.*, 2014) 等。

這些品種鑑別套件在開發時，有各自選擇 SNP 的策略。在探勘 SNP 時，部分文獻為了可以直接做為育種時選拔外表性狀的對照，於基因序列編碼區 (coding regions) 內探勘能造成外表型差異的 SNP (Reale *et al.*, 2006; Consolandi *et al.*, 2007)，也有文獻以隨機探勘特定或非特定位置之 SNP 後，再由實驗材料挑選及驗證選取的 SNP 標誌 (Wang *et al.*, 2010; Jung *et al.*, 2010)；在篩選 SNP 時，考量標誌使用的效率與開發標誌的擴充性，評估基因型分析成功率、次要對偶基因頻度 (Minor Allele Frequency, MAF) 與 PIC 值等數值進行挑選 (Cabezas *et al.*, 2011; Wu *et al.*, 2014)，也有研究將側翼序列可與物理圖譜對照納入挑選 SNP 的考量條件 (Wu *et al.*, 2014)；另外考量標誌間的相關性，避免挑選到過份相似的標誌，挑選分散於染色體 (Cabezas *et al.*, 2011) 或兩兩 SNP 標誌間相關係數 (correlation coefficient) 低的標誌 (Heaton *et al.*, 2002)。

三、水稻 SNP 資料庫

隨著科技進步，SNP 資料庫公開取得資源增加，在水稻目前已有超過 10 個資料庫可做為研究的背景材料 (表 1)，這些資料庫有針對單一亞種內收集系探勘 SNP 之資料庫 (Thomson *et al.*, 2011; 陳等, 2013)、比較世界主要之水稻亞種秈稻 (Indica, IND) 與粳稻 (Japonica, JPN) 間之 SNP 資料庫 (Feltus *et al.*, 2004;

Kim *et al.*, 2009; Gao *et al.*, 2013; Yonemaru *et al.*, 2014) 、僅收集秈稻與粳稻之 SNP 資料庫 (Zhao *et al.*, 2010; Thomson *et al.*, 2011) 、廣泛收集來自世界各地水稻亞種之 SNP 資料庫 (McNally *et al.*, 2009; Zhao *et al.*, 2011, 2015; Thomson *et al.*, 2011; Yonemaru *et al.*, 2014; Alexandrov *et al.*, 2014; Zheng *et al.*, 2015) ，以及栽培稻與野生稻 (*O. rufipogon*) 間之 SNP 資料庫 (Thomson *et al.*, 2011) 等。

本研究於 2012 至 2013 年進行套件 SNP 的篩選，考量 SNP 資訊取得與需選取於品種間具較高多型性的 SNP 標誌，選擇 Zhao 等學者於 2011 年發表的 44K 水稻 SNP 晶片對包含秈稻、溫帶粳稻 (Temperate japonica, TEJ)、熱帶粳稻 (Tropical japonica, TRJ)、香稻 (Aromatic, ARM)、早中稻 (Aus) 及小部分其它 (Admix) 水稻亞種之 413 個水稻品種基因型分析結果，做為篩選標誌的起點，爰此為取得的資料中品種數較多且次族群的種類涵蓋較廣，可針對不同次族群進行分析與篩選 SNP。

四、水稻的核心收集

核心收集的概念起源於基因庫之種原收集，於保育的觀點下避免遺傳流失 (genetic erosion)，保持基因多樣性，或應用於育種，基因庫收集、保存、交換及評估各個收集到的種原。然而，隨著收集到的種原越來越多，數量越來越龐大，在有限的資源下難以維持，更遑論基因庫的交換與利用等收集種原的目的，科學家開始思考該如何從這些原始收集集中挑選一部分最具「代表性」的收集系，以適當的遺傳收集數量，盡可能最大限度代表整個遺傳資源的多樣性，提高基因庫之管理與利用效率。挑選與建立核心收集的策略，考量不同目的而可將核心收集組成的形式歸類為三種類型，(1) 組成核心收集的收集系廣泛涵蓋所有分佈；(2) 以表現極端的收集系彙集成核心收集；(3) 依收集的原始分佈等比例縮小組成核心收集，相對應的收集系挑選策略則有：最近區域收集系平均距離 (Average distance between each accession and the nearest entry, A-NE)、最近加權平均距離 (Average distance between each entry and the nearest neighbouring entry, E-NE)、

平均遺傳距離 (Average genetic distances between entries, E-E) 及以匯總統計 (summary statistics) 取得符合原始分佈之核心收集等方法。A-NE 先以最小平均遺傳距離分群，再從分群內挑選相對遺傳距離最大的收集系組合組成核心收集，E-NE 挑選收集系兩兩遺傳距離與最小遺傳距離加總最大的組合組成核心收集，E-E 則挑選兩兩收集系的遺傳距離總合最大者組成核心收集 (Odong *et al.*, 2012)。

第一類型的核心收集通常是代表一個地區範圍內的歧異之收集，例如日本 National Institute of Agrobiological Sciences (NIAS) 基因庫為有效管理基因庫，並解決重複收集日本水稻造成基因庫資源浪費問題，將取自日本各地收集到的 236 個水或陸稻收集系，以 32 個 SSR 標誌分析基因型，並調查 20 個外表性狀，依據基因型分析的結果，選取涵蓋原始收集 87.5% 基因型與大部分外表性狀的 50 個收集系組成核心收集，建立可代表日本稻米的核心收集 (Ebana *et al.*, 2008)；Kojima 等學者從 NIAS 基因庫選擇來自 23 個國家共計 332 個收集系，以 179 個 RFLP 標誌分析基因型，挑選涵蓋 91% 基因型的 67 個收集系與做為私稈稻序列背景參考用之‘日本晴’與‘Kasalath’共計 69 個收集系，組成適用於稻種歧異度分析的核心收集 (Kojima *et al.*, 2005)；第二類型的核心收集常見於育種材料之收集，Borba 等人從 EMBRAPA Rice Core Collection (ERiCC) 收集來自 25 個國家的 242 個稻收集系，以 82 個 SSR 標誌分析，選擇遺傳距離最大的 24 個收集系組合做為核心收集的子收集，目標是用於稻米育種研究 (Borba *et al.*, 2009)；Li 等學者為了提高育種材料的運用性與方便管理，有效將水稻收集系用於品種改良，以 15 個數量性狀與 34 個質量性狀從 2,262 個水稻收集系中挑出遺傳距離最大的 150 個收集系組成核心收集 (Li *et al.*, 2011)，Zhang 等學者以 274 個 SSR 標誌驗證此收集系與原始收集的關係，結果挑出的核心收集與原始收集中具有相似的遺傳歧異度與族群結構 (Zhang *et al.*, 2011)；第三類型則是目標為保持與原始收集系組成相類似的結構，如美國農部 (United States Department of Agriculture, USDA) 為了更有效的評估管理種原庫中稻收集系，從 18,412 個收集

系中以隨機方式選取來自 114 個國家之 1,790 個收集系做為核心收集，調查 14 個性狀評估以隨機方式選取的核心收集，此核心收集與原始收集具高度相關 ($r = 0.94, p < 0.0001$)，且可有效評估原始收集 88% 的訊息 (Yan *et al.*, 2007)。

五、CoreHunter

CoreHunter 於 2009 年發表，以改良之隨機區域搜尋法 (stochastic local search algorithm) 選取核心收集，於原始集中隨機選取核心收集的目標數量後，以每次替換 1 個收集系並比較替換前後之演算值，留下較佳的組合，組成核心收集。Tachuk 等學者依據核心收集的挑選策略，開發一些演算值的計算方法，Modified Rogers (MR) 將每一個等位基因座視為獨立，計算兩兩收集系間的歐幾里德遺傳距離；Cavalli-Sforza and Edwards (CE) 則模擬這些收集系是在有低突變率及選拔壓力下產生，計算兩兩收集系間的歐幾里德遺傳距離；Shannon's diversity index (SH) 會盡量保留所有的基因座，使遺傳訊息不會因為少量而被捨棄；expected proportion of heterozygous loci (HE) 著重基因座內多樣性的選擇，使選到核心收集的基因座帶有較多可能性；number of effective alleles (NE) 則是除了選擇基因座內多樣性，也考慮多樣性在挑選到的核心收集內的平均分散程度；Coverage of alleles in the core subset (CV) 計算所有基因型占全部基因型的覆蓋率，挑選最廣泛涵蓋所有基因座中所有可能基因型的核心收集，而 Proportion of non-informative alleles in the core subset (PN) 則相反於 CV，以計算未覆蓋全部基因型的程度，挑選最廣泛涵蓋所有基因座基因型的核心收集 (Thachuk *et al.*, 2009)。使用者可依據不同目標選擇不同的「挑選策略」，分配各種評估策略的權重計算演算值並組成核心收集。然由前人研究得知，歧異度的描述可由 E-NE、E-E 或 A-NE 表示，當單以 E-E 挑選核心收集時，可能造成選到的核心收集分群後群間差異極大但群內差異極小之情形，而使用 A-NE 或 E-NE 則可有效避免此情形發生。CoreHunter 於 2012 年的版本新增 minMR 與 minCE 參數，與 MR、CE 不同的地方是，MR、CE 是以 E-E 計算遺傳距離，而 minMR、minCE 則是以 A-NE 計算遺傳距離 (Beukelaer *et al.*, 2012)。

表 1、水稻 SNP 資料庫。

Database	Data	Material	Reference
Plant Genome Mapping Laboratory	384,431 SNPs	IND, JPN	Feltus <i>et al.</i> , 2004
OryzaSNP@MSU	162,478 SNPs × 20 varieties	IND, JPN, Aus	McNally <i>et al.</i> , 2009
RSNMDB	12,829 SNPs	IND, JPN	Kim <i>et al.</i> , 2009
RIS	4,723,468 SNPs × 2 variety	IND, JPN	Gao <i>et al.</i> , 2013
臺南區農業改良場 研究彙報第 61 號	123 SNPs	JPN-Taiwan	陳等, 2013
SNP-Seek system	20 M SNPs × 3 K accessions	IND, Aus, TEJ, TRJ, ARM, ADM	Alexandrov <i>et al.</i> , 2014
RiceVarMap	6,551,358 SNPs	IND, Aus, TEJ, TRJ, ARM, ADM	Zhao <i>et al.</i> , 2015
RFGB	>18,900,000 SNPs × 2,859 accessions	IND, Aus, TEJ, TRJ, ARM, ADM	Zheng <i>et al.</i> , 2015
Q-TARO	3,894 SNPs × 76 accessions	IND, Aus, TRJ, TEJ, ADM	Yonemaru <i>et al.</i> , 2014
	3,252 SNPs × 177 accessions	IND, JPN	Yonemaru <i>et al.</i> , 2014
Rice Diversity	44,100 SNPs × 413 accessions	IND, Aus, TEJ, TRJ, ARM, ADM	Zhao <i>et al.</i> , 2011
	1,536 SNPs × 395 accessions	IND, JPN	Zhao <i>et al.</i> , 2010
	96 SNPs (RiceOPA1.0)	IND, Aus, TEJ, TRJ, ARM, ADM	Thomson <i>et al.</i> , 2011
	384 SNPs (RiceOPA2.1)	IND	Thomson <i>et al.</i> , 2011
	384 SNPs (RiceOPA3.1)	IND, JPN	Thomson <i>et al.</i> , 2011
	384 SNPs (RiceOPA4.0)	JPN	Thomson <i>et al.</i> , 2011
	384 SNPs (RiceOPA5.0)	IND, <i>O. rufipogon</i>	Thomson <i>et al.</i> , 2011
	384 SNPs (RiceOPA6.0)	JPN, <i>O. rufipogon</i>	Thomson <i>et al.</i> , 2011
	384 SNPs (RiceOPA7.0)	IND, JPN	Thomson <i>et al.</i> , 2011

註 1：Oryza sativa 亞種之縮寫對照。

IND: Indica; JAP: Japonica; TEJ: Temperate japonica; ARM: Aromatic TRJ: Tropical japonica; ADM: Admixed

註 2：資料庫連結網址如下：

Rice Diversity: <http://www.ricediversity.org/data/>

Q-TARO (QTL Annotation Rice Online) database: <http://qtaro.abr.affrc.go.jp/index.html#>

RFGB (水稻功能基因組育種數據庫): <http://www.rmbreeding.cn/Index/>

SNP-Seek system: <http://www.oryzasnp.org/iric-portal/>

RiceVarMap: <http://ricevarmap.ncpgr.cn/>

RIS (Rice Information Syst): <http://rice.genomics.org.cn/rice/index2.jsp>

Plant Genome Mapping Laboratory: <http://www.plantgenome.uga.edu/snp>

OryzaSNP@MSU: <http://oryzasnp.plantbiology.msu.edu/index.html>

RSNMDB (RSN marker database): <http://nabic.niab.go.kr/SNP/>

第三章 材料與方法



一、試驗材料

先前開發水稻 SSR 品種鑑別套件的經驗，得知臺灣的水稻品種歧異度低，在粳稻中尤為明顯（柯，2014），為使開發之套件可實際運用在臺灣水稻品種鑑別上，本次研究 SNP 挑選的策略是先尋找在 8 個歧異度較大的粳稻品種間具有多型性的標誌，再擴大範圍篩選包含優良水稻推廣品種在內之 49 個常見的臺灣水稻品種，參試水稻品種清單詳如表 2。

第一階段目標為挑選能成功增幅且可用於鑑別粳稻品種間的標誌。從實驗室開發之第三版水稻 SSR 品種鑑別套件 RS3.4 的分析資料中，取其 53 個臺灣粳稻品種，以 CoreHunter 分析（設定參數為 -MR 0.5 -MRmin 0.5 -mixrep -sample_size 88），選出歧異度較大且基因型廣布於粳稻品種間的 8 個粳稻品種，用以作為臺灣粳稻的核心族群，計有‘花蓮 19 號’（‘HL19’）、‘花蓮 21 號’（‘HL21’）、‘高雄 139 號’（‘KH139’）、‘桃園 4 號’（‘TY4’）、‘桃園糯 2 號’（‘TYW2’）、‘臺梗 5 號’（‘TK5’）、‘臺梗糯 1 號’（‘TKGlu1’）、‘臺農 77 號’（‘TNG77’），其品種間最小遺傳距離為 0.791，平均的遺傳距離為 0.852。此外，由於已知第三版水稻 SSR 品種鑑別套件尚無法鑑別‘臺東 30 號’（‘TT30’）/‘臺東 33 號’（‘TT33’）與‘臺中秈糯 1 號’（‘TCSG1’）/‘臺中秈糯 2 號’（‘TCSG2’），本研究在第一階段挑選 SNP 多型性時，加入分析此 4 個水稻品種的基因型，期望能找到在此兩兩品種間具有多型性的標誌。

第二階段以 104 年臺灣優良水稻推廣品種與臺灣常見的水稻栽培品種為標誌篩選標的，共計收集到 49 個水稻品種，包含 40 個粳稻（‘花蓮 19、20、21 號’（‘HL19, 20, 21’）、‘高雄 139、144、145、146、147 號’（‘KH139, 144, 145, 146, 147’）、‘桃園 3、4 號’（‘TY3, 4’）、‘桃園糯 2 號’（‘TYW2’）、‘越光’、‘臺中 192、194 號’（‘TC192, 194’）、‘臺中在來 1 號’、‘臺東 30、32、33 號’（‘TT30, 32, 33’）、‘臺東糯 31 號’（‘TTG31’）、‘臺南 11、13、14、16

號’ (‘TN11, 13, 14, 16’)、‘臺稉 2、4、5、8、9、11、14、16 號’ (‘TK2, 4, 5, 8, 9, 11, 14, 16’)、‘臺稉糯 1、3、5 號’ (‘TKGlu1, 3, 5’)、‘臺農 71、74、77、79、84 號’ (‘TNG71, 74, 77, 79, 84’)、‘臺農糯 73 號’ (‘TNGW73’) 與 9 個私稻 (‘高雄私 7 號’ (‘KHS7’)、‘臺中私 10、17 號’ (‘TCS10, 17’)、‘臺中私糯 1、2 號’ (‘TCSG1, 2’)、‘臺私 2 號’、‘臺農私 14、22 號’ (‘TNGS14, 22’) 與 ‘臺農私糯 21 號’ (‘TNGSW21’))。

試驗材料來源以農委會種苗改良繁殖場種子檢查室進行水稻種子檢查合格之原原種、原種樣品為主，部分材料為各改良場提供之育種家種子。

二、水稻 DNA 萃取、定量與品質檢定

DNA 萃取方式修改自 Jobes 等學者開發的 SDS (sodium dodecyl sulfate) 法 (Jobes *et al.*, 1995)。單一品種取水稻單粒稻穀加入 400 μ L SDS 萃取液 (1.25% SDS、100 mM Tris-HCl (pH 8.0)、500 mM NaCl₂、50 mM EDTA (ethylenediaminetetra-acetic acid, pH 8.0))，以均質機 (SH-100, KURABO) 均質後置於 65 °C 水浴處理 30 分鐘。加入 125 μ L 5M KOAc (Potassium acetate)，並立即混勻，於 4 °C 靜置 20 分鐘後，以 15,000 rpm 在 4 °C 下離心 10 分鐘。於低溫下取上清液 400 μ l 加入 800 μ L 95 % 酒精，混合均勻後在 4 °C 下靜置 10 分鐘。以 15,000 rpm 在 4 °C 下離心 10 分鐘，倒去上清液，吸除殘滴，置於室溫下待 pellet 風乾，加入 120 μ l TE buffer 回溶 DNA。以分光光度計 (NanoPhotometer P360, IMPLEN) 定量，並將樣品稀釋為 10 ng/ μ L。

三、聚合酶連鎖反應與電泳分析

聚合酶連鎖反應 (Polymerase Chain Reaction, PCR) 以 *Taq* DNA Polymerase Master Mix Red --2x and 1.5 mM MgCl₂ (AMPLIQON) 進行，反應總體積為 13 μ L，包含 20 ng 樣品 DNA、1X *Taq* DNA Polymerase 2x Master Mix Red (Tris-HCl pH 8.5、(NH₄)₂SO₄、3 mM MgCl₂、0.2 % Tween 20、0.4 mM dNTP、0.2 U Ampliqon

Taq DNA polymerase)、0.3 μM AS-primer、0.3 μM 對應 AS-primer 之側翼引子、0.15 μM AS-primer 同向之側翼引子。

聚合酶連鎖反應使用熱循環反應器 (Veriti 96Well Thermal Cycler, Applied Biosystems) 進行, 反應條件為 94 $^{\circ}\text{C}$ 、2 分鐘; 94 $^{\circ}\text{C}$ 、1 分鐘, T_m $^{\circ}\text{C}$ 、40 秒, 72 $^{\circ}\text{C}$ 、40 秒, 共 2 個循環; 94 $^{\circ}\text{C}$ 、30 秒, T_m $^{\circ}\text{C}$ 、20 秒, 72 $^{\circ}\text{C}$ 、20 秒, 共 31 個循環; 72 $^{\circ}\text{C}$ 、10 分鐘; 反應完成以 25 $^{\circ}\text{C}$ 保存。 T_m 為引子黏合溫度, 當 AS-primer 為具有 3'端 -3 位置鹼基與參考序列不相符時, T_m 為 58 $^{\circ}\text{C}$; 當 AS-primer 序列與參考序列完全相符時, T_m 為 60 $^{\circ}\text{C}$ 。

電泳分析以 0.5x TAE Buffer (20 mM Tris、0.5 mM EDTA-Na、10 mM glacial acetic acid, pH 8.3, GeneMark) 加入 5 μl SafeView DNA Stain (GeneMark) 配置 200 mL 之 2% 瓊脂膠片, 將 PCR 擴增產物取 5 μl 載入瓊脂膠片, 以 0.5x TAE Buffer 作為緩衝液, 經 200 伏特分離 30 分鐘後以照相系統 (Alpha Innotech Imaging System, Alphamager 2200, Alpha Innotech) 記錄後, 以人工進行基因型判讀。

四、SNP 基因座挑選

從 Rice Diversity 網站 (<http://www.ricediversity.org/data/sets/44kgwas/>) 下載水稻 44K SNP 基因型資料與側翼資料, 取其交集 SNP 資訊後, 比對染色體編號與物理圖譜位置, 將結果完全一致的 35,746 個 SNP 進行初步篩選。考量標誌的可用性、可通用於不同分析系統、確保 SNP 在基因型間有充分流通性及避免選到存活基因 (survival gene) 或花粉競爭基因連鎖的相關分子標誌, 初步篩選階段選擇整體的讀值機率 > 0.9 、屬於非 A/T 或 C/G 組合型的 SNP、座落在非中節區且整體的異質結合率 < 0.05 的 SNP 基因座, 共獲得 22,213 個 SNP 標誌。我們從 22,213 個通過初步篩選的 SNP 中, 以 MAF 評估標誌的多型性, 另考量主要的水稻品種屬於秈稻與溫帶粳稻較多, 因此只採用這兩類水稻分群的 MAF 進行評估與篩選。由於溫帶粳稻內的歧異度較秈稻低很多, 且同時在這兩類之中皆具有高 MAF 的 SNP 數目相當少, 對於無法選到在這兩類都具有較大 MAF 的染

色體，將染色體分為兩臂（arms），先在一臂上盡可能選取對於其中一類中具有較高 MAF 後，在另一臂上選取對於另一類具有較高 MAF 的 SNP。此階段選到每一個染色體區段各 4-8 個 SNP，共獲得 130 個 SNP。



五、SNP 分子標誌設計

AS-PCR 與 tetra-primer ARMS PCR 不需訂購螢光引子且可以水平電泳判讀，操作方便且於探勘多型性階段最省成本，考量 tetra-primer ARMS PCR 單次反應需調合 4 支引子的相容性，設計成功率低且各引子組合 PCR 的最佳化條件較不易調整，故於 SNP 多型性探勘階段選擇以 AS-PCR 探勘目標族群多型性。

擷取 130 個 SNP 側翼各 500 bp 的序列，批次送入 BatchPrimer3 網站 (You *et al.*, 2008)，選擇 'Allele-specific primers and allele-flanking primers'，設定引子大小介於 15-30 mer，偏好值為 20 mer、黏合溫度 (primer T_m) 介於 57–63 °C，偏好值為 60°C、側翼產物大小介於 200–1,000 bp，偏好值為 300 bp，分別勾選或不勾選 3' 錯誤配對 (mismatch) 選項，進行引子設計。將設計結果挑選 2 個側翼序列與 2 個 SNP 基因型皆成功設計的引子組合，淘汰側翼條帶與對偶基因專一性條帶大小差異小於 50 bp 者，避免於水平電泳條帶無法分離，當 AS-primer 有超過 1 種組合可選擇時，以 Q-score 平均值較高者優先選擇。

六、評估 SNP 分子標誌與架構水稻 SNP 品種鑑別套件

130 個 SNP 以 BatchPrimer3 設計對偶基因專一性引子的結果，共成功設計 97 個在 SNP 專一性引子 3' 端 -3 位置鹼基與參考序列不相符的 SNP 引子組，與 88 個 SNP 專一性引子序列與參考序列完全相符的 SNP 引子組。

第一階段以設計之 185 個標誌對 8 個在歧異度較大的稉稻基因型分析，選取可正常增幅且於稉稻間具有多型性的標誌，當同一個 SNP 基因座具有 2 個可使用的標誌時，考量 AS-primer 同向者其系統通用性較高且引子黏合性較一致，較好判讀結果，以同向 AS-primer 之標誌為優先。篩選結果共計得到 47 個標誌，其中 40 個為 AS 引子方向相同的標誌，7 個為 AS 引子反向的標誌。

CoreHunter 輸入資料的形式為「分子標誌的所有可能基因型之頻度」對「收集系」的基因型分析兩向表，以 CoreHunter 軟體之 CV 參數可單純計算挑選到的收集系中，所有基因型占原始收集全部基因型的覆蓋率，進而挑選到涵蓋最多標誌基因型的最少收集系組合，以此概念將兩向表更改輸入資料的形式為「兩兩收集系間的差異」對「分子標誌」的基因型分析兩向表（圖 1），預期可利用 CV 參數找到涵蓋最廣泛收集系間差異的最少分子標誌組合，即為可辨識最多收集系的最少分子標誌組合。第二階段將 47 個標誌對 49 個臺灣水稻品種基因型分析，將資料整理為兩兩收集基因型系差異對分子標誌之兩向表，於 CoreHunter 設定參數 -CV 1.0，篩選可區別品種間差異的標誌組合，調整選取標誌數量至最小可涵蓋所有基因型讀值的數量，獲得可區分 49 個臺灣水稻品種的最小標誌組合。

七、驗證套件標誌組合

為確認以 CoreHunter 之 CV 參數挑選出來的標誌組合符合實際需求，以窮舉法排列出所有符合實際需求的可能性，驗證 CoreHunter 的挑選結果。由於無效等位基因（null allele）無法提供 SNP 多型性資訊，我們剔除無效等位基因過多的標誌，並依據 Patterson 等學者的方法（Patterson *et al.*, 2006），將二元的 SNP 基因型資料正規化後，兩兩比對 SNP 間有效等位基因的相關，將有效等位基因間完全一致的標誌剔除含無效等位基因較多的一個，若皆無無效等位基因，則隨機剔除一個。兩兩比對品種間的相似性，為確保建立的分子鑑別套件可達到辨識最多品種數且為縮小窮舉法排列組合的範圍，將基因型分析無差異的品種合併後，優先選擇只有 1 個 SNP 差異組合的 SNP 標誌，固定必須要存在於 SNP 組合的標誌後，以窮舉法尋找涵蓋水稻基因型最廣泛的最小標誌組合。

八、建立標準品種

為了利於套件的交換利用，確保套件交流時基因型分析的正確性，參考核心收集的概念，我們選擇建立第一類型的核心收集（組成核心收集的收集系廣泛涵蓋所有分佈），從水稻 SNP 品種鑑別套件候選標誌中，挑選能涵蓋全部基因型的

最小品種組合。將水稻 SNP 品種鑑別套件候選標誌對 49 個參試水稻品種基因型分析的結果，以 CoreHunter v2.0 提供之 R 套件介面 (Beukelaer *et al.*, 2012) 輸入參數 -MR 0.5 -MRmin 0.5 -mixrep，挑選以本套件分析基因型其品種間遺傳距離最大且涵蓋最廣泛套件 SNP 基因型的臺灣水稻品種最小組合，建立對本套件對偶基因判讀所需之標準品種 (reference variety)。

表 2、參試水稻品種。49 個水稻品種中包含 40 個粳稻與 9 個秈稻品種

No.	品種	英文縮寫	種子來源	優良水稻 推廣品種	第一 階段	第二 階段	備註
1	花蓮 19 號	HL19	作物育種研究室		○	○	
2	花蓮 20 號	HL20	作物育種研究室			○	#183
3	花蓮 21 號	HL21	種子檢查室	○	○	○	
4	高雄 139 號	KH139	種子檢查室	○	○	○	
5	高雄 144 號	KH144	作物育種研究室			○	#109
6	高雄 145 號	KH145	種子檢查室	○		○	原種
7	高雄 146 號	KH146	作物育種研究室	○		○	
8	高雄 147 號	KH147	種子檢查室	○		○	原種
9	高雄私 7 號	KHS7	作物育種研究室	○		○	
10	桃園 3 號	TY3	種子檢查室	○		○	原種
11	桃園 4 號	TY4	作物育種研究室		○	○	
12	桃園糯 2 號	TYW2	作物育種研究室		○	○	
13	越光		農業試驗所	○		○	
14	臺中 192 號	TC192	作物育種研究室	○		○	
15	臺中 194 號	TC194	作物育種研究室			○	#086
16	臺中在來 1 號		農業試驗所			○	
17	臺中私 10 號	TCS10	種子檢查室	○		○	
18	臺中私 17 號	TCS17	作物育種研究室	○		○	#046
19	臺中私糯 1 號	TCSG1	種子檢查室		○	○	
20	臺中私糯 2 號	TCSG2	種子檢查室	○	○	○	原種
21	臺私 2 號	TS2	作物育種研究室			○	
22	臺東 30 號	TT30	種子檢查室	○	○	○	原種
23	臺東 32 號	TT32	種子檢查室			○	原原種
24	臺東 33 號	TT33	臺東改良場	○	○	○	
25	臺東糯 31 號	TTG31	作物育種研究室	○		○	
26	臺南 11 號	TN11	種子檢查室	○		○	原種
27	臺南 13 號	TN13	臺南改良場			○	原原種
28	臺南 14 號	TN14	作物育種研究室			○	
29	臺南 16 號	TN16	臺南改良場			○	原始種子
30	臺粳 2 號	TK2	作物育種研究室	○		○	

31	臺稈 4 號	TK4	種子檢查室			○	原種
32	臺稈 5 號	TK5	作物育種研究室		○	○	
33	臺稈 8 號	TK8	種子檢查室	○		○	原種
34	臺稈 9 號	TK9	種子檢查室	○		○	原種
35	臺稈 11 號	TK11	作物育種研究室			○	
36	臺稈 14 號	TK14	種子檢查室	○		○	原種
37	臺稈 16 號	TK16	臺中改良場			○	展示用種子
38	臺稈糯 1 號	TKGlu1	種子檢查室	○	○	○	
39	臺稈糯 3 號	TKGlu3	種子檢查室	○		○	
40	臺稈糯 5 號	TKGlu5	作物育種研究室			○	
41	臺農 71 號	TNG71	種子檢查室	○		○	原種
42	臺農 74 號	TNG74	種子檢查室			○	原原種
43	臺農 77 號	TNG77	農業試驗所	○	○	○	原原種
44	臺農 79 號	TNG79	作物育種研究室			○	
45	臺農 84 號	TNG84	作物育種研究室			○	
46	臺農私 14 號	TNGS14	作物育種研究室	○		○	#251
47	臺農私 22 號	TNGS22	種子檢查室	○		○	原原種
48	臺農私糯 21 號	TNGSW21	作物育種研究室			○	#253
49	臺農糯 73 號	TNGW73	作物育種研究室	○		○	

	A	B	C	D	E	F	G	
1	marker\variety	TN11	TK14	TC192	TCS10	TCSG2	TK2	
2	dd1001675	A	A	A	G	G	G	
3	id11009466	A	A	A	G	G	A	
4	id12008090	C	C	C	C	C	C	
5	id2010155	A	A	A	G	A	G	
6	id4005236	A	A	A	G	G	G	
7	id4011022	T	T	T	T	T	T	

正規的
CoreHunter 輸入
資料格式

	A	B	C	D	E	F	G	H	
1	marker	variety	TN11	TK14	TC192	TCS10	TCSG2	TK2	
2	dd1001675	A	1	1	1	0	0	0	
3	dd1001675	G	0	0	0	1	1	1	
4	id11009466	A	1	1	1	0	0	1	
5	id11009466	G	0	0	0	1	1	0	
6	id12008090	C	1	1	1	1	1	1	
7	id12008090	G	0	0	0	0	0	0	
8	id2010155	A	1	1	1	0	1	0	
9	id2010155	G	0	0	0	1	0	1	
10	id4005236	A	1	1	1	0	0	0	
11	id4005236	G	0	0	0	1	1	1	
12	id4011022	A	0	0	0	0	0	0	
13	id4011022	G	1	1	1	1	1	1	

整理成以 CV 參
數挑選標誌的兩
兩品種比較表

	A	B	C	D	E	F	G	H	
1	variety1	variety2	dd1001675	id11009466	id12008090	id2010155	id4005236	id4011022	
2	TN11	TK14	0	0	0	0	0	0	
3	TN11	TC192	0	0	0	0	0	0	
4	TN11	TCS10	1	1	0	1	1	0	
5	TN11	TCSG2	1	1	0	0	1	0	
6	TN11	TK2	1	0	0	1	1	0	
7	TK14	TC192	0	0	0	0	0	0	
8	TK14	TCS10	1	1	0	1	1	0	
9	TK14	TCSG2	1	1	0	0	1	0	
10	TK14	TK2	1	0	0	1	1	0	
11	TC192	TCS10	1	1	0	1	1	0	
12	TC192	TCSG2	1	1	0	0	1	0	
13	TC192	TK2	1	0	0	1	1	0	
14	TCS10	TCSG2	0	0	0	1	0	0	
15	TCS10	TK2	0	1	0	0	0	0	
16	TCSG2	TK2	0	1	0	1	0	0	

圖 1、CoreHunter 輸入資料之格式。橫列為每個標誌之基因型頻度，直行為不同收集系（品種），填入的數值為各基因型的頻度。以 CV 參數挑選涵蓋最廣泛基因型的分子標誌時，需先將兩向表轉換為兩兩品種之分子標誌頻度差異比較表。

第四章 結果與討論



一、SNP 基因座挑選結果

對偶基因專一性標誌 (allele specific marker, AS-marker) 單次增幅一個 SNP 基因型，故偵測 2 個 SNP 對偶基因需由 2 次 PCR 完成。於增幅反應中加入與 AS-primer 同向之外側側翼引子，可避免錯誤增幅且可作為 PCR 正對照，故典型的增幅結果為具備側翼條帶與 SNP 專一性條帶 (圖 2 (A))；有時因為 SNP 專一性引子與同向之側翼引子互相競爭，造成能成功增幅 SNP 專一性條帶的產物其側翼條帶微弱甚至消失 (圖 2 (B))，在判讀時也可依據比較相同 SNP 基因座之 2 個 PCR 結果，側翼條帶與 SNP 專一性條帶強弱差異，判別 SNP 基因型 (圖 2 (C))。當(1) 任一 PCR 結果完全無增幅產物；(2) 增幅之條帶大小不符合預期；(3) 增幅之條帶混雜無法辨識 (圖 2 (D)、(E)) 時，將其視為不良分子標誌並淘汰。

第一階段以 130 個 SNP 基因座設計出的 185 組 AS-marker 分析 8 個歧異度較大的粳稻品種。結果有 61 個標誌能成功增幅與判讀，剔除重覆基因座位置的標誌，將散佈在水稻基因體不同 SNP 位置的 47 個標誌進行第二階段分析。

第二階段以 47 個 AS-marker 分析包含臺灣優良水稻推薦品種在內的 49 個臺灣水稻品種，結果 4,606 個資料點中，包含 8 個 SNP 位置為無效等位基因與 4,590 個同質 SNP 基因型，無異質結合基因型 (附錄 1)。比較標誌於品種間基因型讀值的結果，有 5 組標誌完全一致，分別為 id4011022 / id4011023、id10007165 / id10007166、id1000948 / id1000955、id1027576 / id1027513、id12008144 / id12008149；而 id9002505 與 id9002494 之間，除了 id9002505 有 2 個缺值之外，其餘完全一致；將高度相關的標誌刪去缺值較多的一個，完全相同者則隨機保留一個後，兩兩比較 41 個 AS-marker 在 49 個品種基因型間之差異，結果有 3 組品種在 47 個標誌之基因型讀值完全一樣，分別為‘TT30’/‘TT33’、‘TCSG1’/‘TCSG2’、‘TN16’/‘越光’，其餘品種兩兩間皆至少有 1 個 SNP 標誌以上的差異 (圖 3)。

輸入 41 個標誌基因型分析之讀值，以 CoreHunter 設定參數 $-CV\ 1.0$ ，篩選涵蓋最多品種間基因型差異的最少分子標誌組合，調整選取分子標誌數量至最小可涵蓋所有基因型讀值的數量，結果至少需 10 個 SNP 才能達到可區分最多兩兩品種差異的最小組合，將除了 3 組基因型完全相同的組合外，其餘水稻品種兩兩組合間皆至少具有 1 個基因型不同的 SNP (表 3)。

二、驗證基因座挑選組合

以窮舉法驗證 SNP 基因座挑選組合時，考量本次研究目的為組成能區分最多水稻品種的 SNP 組合，且為簡化排列組合的數量，優先選取兩兩品種間僅 1 個差異的必要 SNP。固定必須要存在的 4 個 SNP 之後，其餘的 37 個 SNP 取隨機的組合，合併先前選出的 4 個必要 SNP 之後，計算兩兩品種間的差異。結果顯示取隨機 4 與 5 (總數為 8 與 9) 個 SNP 都沒有辦法找到可以區分每一個品種的組合，取 6 (總數為 10) 時可以找到 9 組能區分每一個品種的 SNP 組合，其中的一組 SNP 組成與 CoreHunter 篩選出的結果相符。

CoreHunter 所得到的挑選結果與每次隨機抽換的 SNP 有關，當符合篩選標準的組合數大於 1 時，每次分析可能得到不同的結果。以相同的篩選標準執行 1,000 次，可以得到窮舉法獲得的全部 9 種組合 (資料未呈現)。此一結果顯示以兩兩品種差異涵蓋率取代對偶基因頻度，再以 $-CV\ 1.0$ 條件進行篩選，是一個尋找最小分子標誌組合的可行策略；雖然每次只能獲得一種組合，但是當需要鑑別的品種數與分子標誌數目較多，窮舉法所需要的運算量在一般電腦已無法負荷時，上述的策略不失為一個系統性、有效率的方法。

三、架構臺灣水稻優良品種 SNP 鑑別套件

觀察 9 個涵蓋所有基因讀讀的最小 SNP 基因座組合，其兩兩品種間差異數的分佈無明顯差異 (圖 3)，在挑選建構臺灣水稻優良 SNP 品種識別套件之 SNP 組合時，為使套件具有較高的可靠性，檢查兩兩品種僅 1 個 SNP 差異的組合之親緣關係，在確認其並非由少數幾個親緣關係非常近的品種組成後，採用差異只

有 1 個的兩兩品種組合數盡可能低的組合做為挑選標準，選擇組合 1 做為臺灣水稻優良 SNP 品種鑑別套件之 SNP 組合。

組合 1 包含 dd1001675、id1000948 (或 id1000955)、id6003492、id7000308、id8000978、id10000129、id10007137、id4005078、id6005779、id2010412 等 10 個 SNP 基因座，分散在水稻第 1、2、4、6、7、8、10 條染色體上，平均 PIC 值為 0.402。組合之 MAF 介於 0.302 至 0.497 之間，於私稻與溫帶粳稻有較佳的 MAF 表現，適用於臺灣水稻族群 (表 4、圖 4)。

四、標準品種

以 CoreHunter v2.0 挑選品種間遺傳距離最大的臺灣水稻組合做為標準品種，結果當選取品種數為 3 時始可達到最大的遺傳距離 (Modified Rogers' distance (MR) 為 0.81597; minimum Modified Rogers' distance (minMR) 為 0.77460) 與最高對偶基因型百分比 (100%)，以 CoreHunter 輸入相同的篩選條件執行 1,000 次，共得到 4 個不同的品種組合，分別為組合 1: 'TCS10'、'TK9' 及 'TNG71'; 組合 2: 'TCS10'、'TK9' 及 'KH146'; 組合 3: 'TCS10'、'KH145' 及 'TC194'; 組合 4: 'TNGW73'、'HL19' 及 'HL21'。這些候選品種組合的水稻品種中，'KH146'、'TC194' 及 'HL21' 具有植物品種權保護，考量標準品種取得與交流之實用性，優先選取不具有品種權保護的品種組合，結果僅 1 個品種組合未包含具有品種權保護的水稻品種，故以 'TCS10'、'TK9' 及 'TNG71' 做為本套件之標準品種。

五、開發階段對偶基因專一性引子之選擇

在已知 SNP 位置與側翼序列的前提下，以對偶基因專一性引子 PCR 增幅後用水平電泳分析基因型，不需於引子標定螢光或特殊處理，也不需以酵素酶切，是相對成本低、操作方便且分析結果容易判讀的方法，常見的對偶基因專一性引子分析系統有 Allele specific (AS) 和 tetra primer ARMS。(表 5)

AS 系統單一反應由一 AS-primer、一相對應引子及一側翼引子組成，SNP 位置位於 AS-primer 的 3' 端，在設計引子時也可將 AS-primer 3' 端 3 位置設計為

與參考序列不相符的鹼基，增加引子專一性之差異。比較側翼條帶長短對 PCR 增幅結果的影響，結果不論側翼條帶長或短皆能成功增幅目標條帶，故無明顯影響（圖 5）；比較同一 SNP 位置之 2 個 AS-primer 同向或互為反向之增幅結果，當 AS-primer 同向時，二 AS 條帶大小相同，於讀值時較易判別（圖 6 (B)、(D)、(E)），當 AS-primer 互為反向時，因引子黏合性不完全相同，當需要比較條帶強弱判別 SNP 基因型時，可能將黏合性較佳的條帶視為增幅之對偶基因，造成誤判（圖 6 (A)、(C)）。有關 AS-primer 3' 端錯誤配對有無之比較，AS-primer 3' 端無錯誤配對時，AS-primer 與側翼引子黏合性較接近，當 PCR 反應中 AS-primer 與側翼引子比例相當時，條帶強度相當（圖 6 (A)、(B)、(C)、(D)），而 AS-primer 3' 端錯誤配對時，AS-primer 黏合性明顯較差，故 AS 條帶強度較弱（圖 6 (E)）；錯誤配對產生的黏合性差異也可由梯溫 PCR 觀察到，當引子序列與樣品序列完全相配時，引子黏合性較不受溫度改變而影響，而具有錯誤配對的引子則隨溫度增高黏合性明顯降低（圖 7），此部分的結果與 Sipos 與 Ishii 等學者將吻合或具錯誤配對引子以梯溫 PCR 增幅的結果相符（Ishii and Fukui, 2001; Sipos *et al.*, 2007），推測錯誤配對雖造成引子黏合性減低，但其 PCR 黏合溫度可影響引子黏合性，隨著錯誤配對鹼基數增加，更甚地拉大引子黏合專一性差距，利用此特性可使設計之引子具有更高的敏感度。

Tetra primer ARMS 系統單一反應由一對側翼引子與一對分別偵測兩個不同 SNP 基因型且方向相反的基因型專一性引子組成，2 個 SNP 基因型增幅之條帶具大小差異，單次反應可增幅出一條側翼條帶與 1 至 2 條基因型專一性條帶，故可由 1 個 PCR 得到完整 SNP 基因型資訊。以 BatchPrimer 3 設計 tetra primer ARMS 系統之引子，成功率很低，130 個基因座僅有 35 個可成功設計，故引子設計階段以 3 個引子組成的 AS 系統做為優先考量。觀察 AS 系統之引子組成，當 AS-primer 互為相反方向時，單一基因座之引子組成恰為 tetra primer ARMS 系統之引子組合，我們將方向相反的 AS-marker 側翼引子與 AS-primer 共計 4 支引子混合後一起 PCR（配方：1X *Taq* DNA Polymerase 2x Master Mix Red、2 μ M flanking

primer F、2 μM flanking Primer R、4 μM AS-primer F、4 μM AS-primer R、20 ng DNA，反應總體積為 13.6 μL ；條件同材料與方法)，結果可成功增幅目標條帶，但因為引子間黏合性不同，造成部分異結合條帶過淡，需依實際情形個別調整 PCR 配方條件後方可使用。

開發階段期望引子組合能廣泛地適用同一實驗操作流程，在本研究初期考量 PCR 黏合適溫的廣泛性與 AS 系統和 tetra primer ARMS 系統互相轉換之可能性，選擇合成 AS-primer 3' 端與目標序列吻合且兩個基因型 AS-primer 方向相反的引子進行實驗，結果 AS-primer 黏合性差異大且 PCR 產物大小之差距造成結果不易判讀(圖 6 (A)、(C))。Tetra primer ARMS 系統雖然可於單一反應完成 2 個 SNP 基因型之判讀，但需將每個基因座之引子組合濃度最佳化，於實務上將造成實驗步驟複雜化，可能進而影響實驗或檢測結果之正確性。研究者認為 tetra primer ARMS 系統較適合應用於僅需偵測少數 SNP 基因座之情境，本次開發之水稻 SNP 品種鑑別套件預計使用 10 個 SNP 基因座，未來可能因為套件的擴充需納入更多的 SNP 標誌，tetra primer ARMS 系統於擴充性與相關操作等方面，並不適用於本研究之套件建立結果，故最終仍選擇以 AS 系統進行基因型分析。而研究過程中歸納設計 AS 標誌需注意的地方：(1) 側翼條帶的大小目前尚未發現對增幅表現的差異，唯需注意其與 AS 條帶之大小差異不應過小；(2) AS-primer 方向為同向時較好判讀；(3) AS-primer 具少數錯誤配對時，引子專一性較序列完全吻合之引子高，但對 PCR 的黏合溫度也較敏感；(4) 引子間或引子本身可能產生會自黏或互黏的次級結構造成 PCR 增幅效率低；(5) 同一反應中各引子的黏合溫度不可差距太大。

六、可能的分析系統

SNP 基因型分析系統非常多元，本研究成果期望未來能運用在大量樣品的定性檢定，故以此方向羅列幾個未來可能使用的分析系統。在已知能以 AS 系統成功判讀基因型的前提下，可將 AS 系統之兩個 AS-primer 標定不同色螢光，直接

用於 open array (Brenan and Morrison, 2005) 之分析；在品種純度或特定品種之判定時，可將 DNA 混合後以側翼引子增幅並進行 TILLING (McCallum *et al.*, 2000)，或重新設計符合 HRM 系統需求的短片段側翼引子進行 HRM (Reed and Wittwer, 2004) 分析。此外，Taqman 系統穩定性極高，是 SNP 基因型分析的“gold standard”，其他 SNP 分析技術的成效都是與 TaqMan 來比較分析的穩定性，縱其分析成本較高，於應用上恐有不切實際之疑慮，但可使用混合樣品或多重 PCR 方式，幫助節省品種檢查的成本。

七、無法區別的臺灣品種

SNP 標誌篩選結果，計有‘越光’/‘TN16’、‘TT30’/‘TT33’及‘TCSG1’/‘TCSG2’三組品種間無具有多型性的 SNP。‘越光’為日本品種，米質佳但因其短日照的特性，於臺灣種植時常因日照時間短照成提早抽穗，產量減低，‘TN16’以‘越光’與‘臺農 67 號’為親本，取‘越光’之優良米質背景與‘臺農 67 號’對日照鈍感之特性，將‘臺農 67 號’調控抽穗期相關基因 *hd1*、*Hd6* 和 *ehd1* 導入越光(陳等, 2010)，成功育種出與‘越光’優良的米質相當且生育期較長，產量較高的‘臺南 16 號’，其帶有高度與越光相同的基因背景(陳等, 2012)；‘TT33’以‘TT30’為母本，‘basmati 370’為父本，雜交後經過 14 代自交選拔，得到抗稻熱病、株型良好不易倒伏、米粒外觀優良、米飯食味佳且耐儲藏的‘TT33’(丁等, 2013)；‘TCSG2’以‘TCSG1’為母本，‘臺中私 17 號’與‘臺梗 16 號’的雜交後代為父本，選育出高產、強稈、不易倒伏、氮肥利用效率高、抗稻熱病、縐葉枯病、斑飛蝨與白背飛蝨，碾米品質佳但耐寒性較差的‘TCSG2’，其品種特性與‘TCSG1’的同質性高，但產量較‘TCSG1’較高，且碾米品質較佳(楊, 2006)為目前主要推廣的取代‘TCSG1’的品種。

品種鑑別主要的目的有二，一為保護育種者智慧財產權，一為監控農產品產銷和品質等(林, 2005)，本次套件無法區分的品種間，‘TT30’/‘TT33’、‘TCSG1’/‘TCSG2’屬相同育種者的智慧財產權，於智慧財產權保護上，可視為同時保護同

一育種者的多個品種；‘TN16’與‘越光’為近似同源系，其所導入的 3 個基因中，*Ehd1* 外顯子 (Exon) 上的 SNP 可以用來區分這兩個品種 (資料未顯示)，保護不同育種者的智慧財產權與必要時之監控農產品產銷及品質等使用；此外，‘TCSG2’高產且與‘TCSG1’同質性高，育成目的為取代‘TCSG1’成為新興推廣品種，此政策方向將慢慢使我國‘TCSG1’的種植面積與產出大幅減少，故其雖無法以現有套件區分，惟於監控農產品產銷部分將不會有太大的影響。

柯 (2014) 由篩選世界水稻品種建立可廣泛鑑別全球水稻品種的第二版水稻品種鑑別系統，發現平均分散於 12 條染色體的分子標誌在 64 個臺灣品種間，有 18 組無法區別的品種，實際相符率大於期望相符率，顯示臺灣水稻品種可能因次族群內親緣性等因素存在族群次結構，不利於以可廣泛鑑別全球水稻的品種鑑別套件使用在鑑別臺灣水稻品種 (柯，2014)。本研究目的之一為建立 SNP 品種鑑別系統之篩選模式與優化的策略，雖研究結果可能因為親緣關係造成基因背景相似度太高，使無法由建立的篩選模式找到能區分 3 組品種兩兩間的標誌，但由標誌之 MAF 與 PIC 值，可推測本次建立的套件應具有符合期望的辨識率。

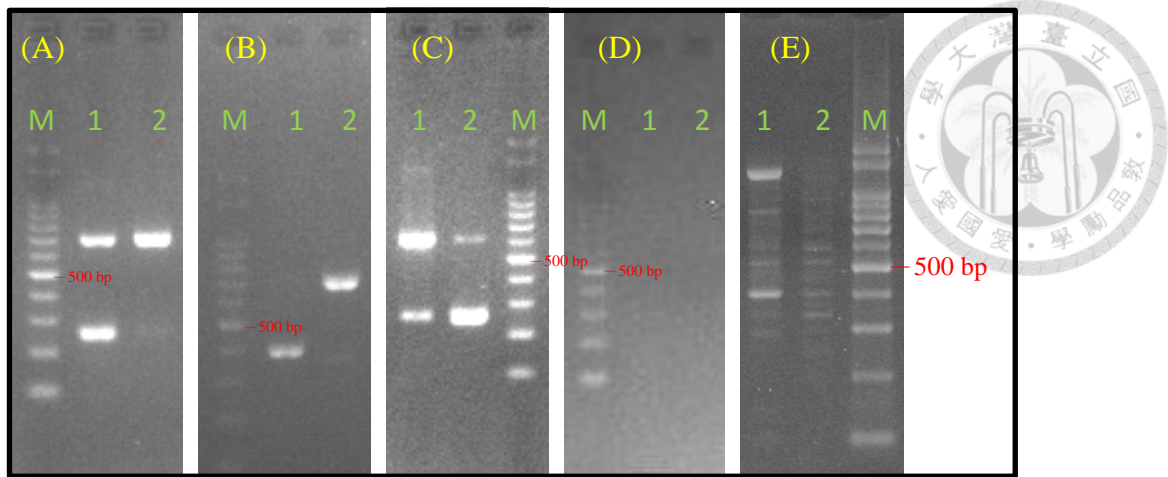


圖 2、Allele specific (AS)系統增幅結果。以 3 支引子組成的 Allele specific (AS) 系統，側翼條帶約 700 bp，對偶基因專一條帶約介於 200-300 bp 之間。M 為 100-3000 bp 之 DNA ladder，由小至大分別為 100、200、300、400、500、600、700、800、900、1000、2000、3000 bp；1 與 2 分別為 AS-primer 具有 SNP 基因座之 A 對偶基因或 B 對偶基因的專一性，分別偵測 SNP 對偶基因 A 與 SNP 對偶基因 B。(A)典型的增幅結果，具備側翼條帶與 SNP 專一性條帶，可藉由增幅出基因座 A 之 SNP 專一性條帶，判定基因型讀值為 A；(B)側翼引子與對偶基因專一性引子互相競爭的結果，有時對偶基因專一性引子黏合性較佳，可能造成側翼條帶因增幅效率較差而看不到條帶，圖之基因型讀值為 A；(C)PCR 增幅結果，雖側翼條帶與對偶基因專一條帶皆被增幅，但可藉由比較條帶強弱之差異，判讀基因型讀值。圖之基因型讀值為 B；(D)PCR 無增幅產物；(E)條帶混雜且無目標大小之條帶。

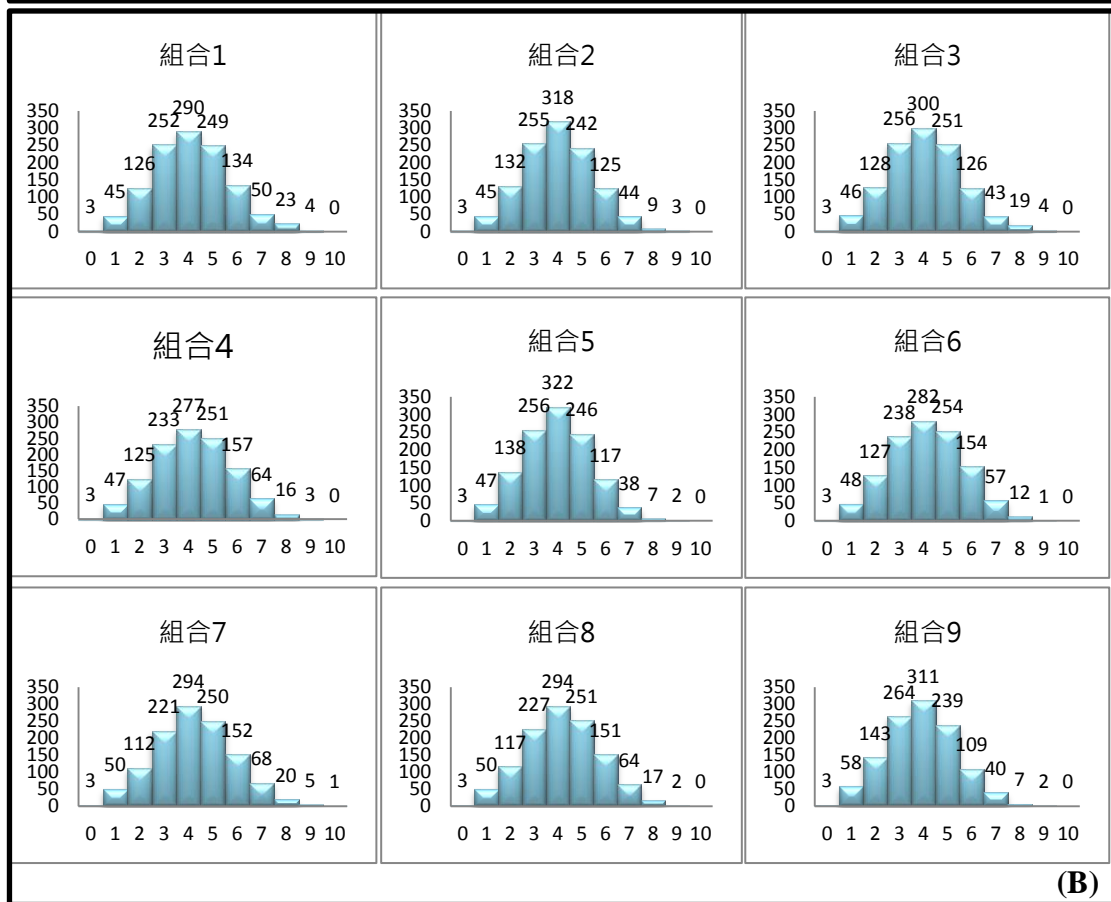
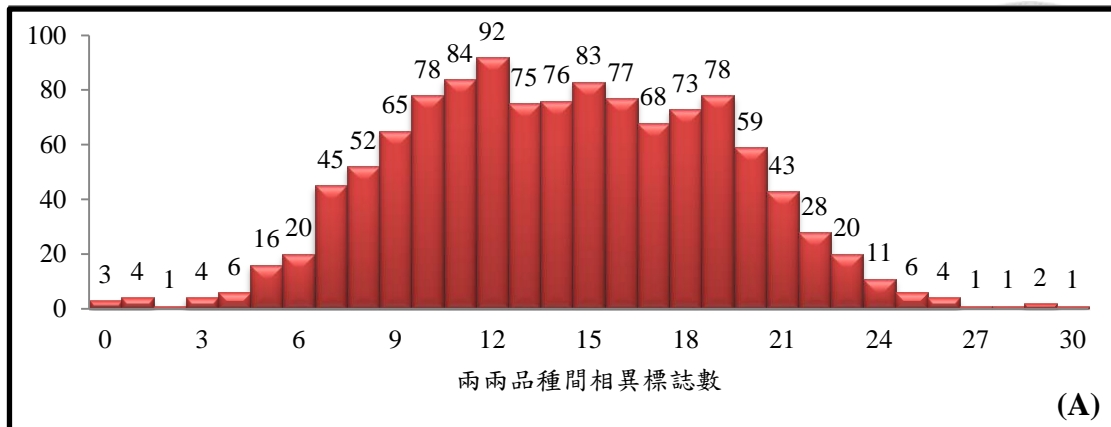


圖 3、兩兩品種間多型性量統計圖。(A) 以 41 個基因型分析結果不重覆的標誌分析 49 個品種，結果在 3 組兩兩品種間完全無多型性；4 組兩兩品種間僅 1 個標誌的差異，而兩兩品種間的差異數最多可達 30 個。(B) 以建立套件之候選組合中 10 個 SNP 標誌對 49 個品種基因型分析的結果分佈，有 3 組兩兩品種間完全無多型性；兩兩品種間差異數最多可達 10 個。

表 3、9 個候選 SNP 組合之兩兩品種間基因型差異數分佈。9 個候選 SNP 組合為可區分除了 TT30 / TT33、TCSG1 / TCSG2、TN16 / 越光品種間之 SNP 最小集合。

差異數	組合 1	組合 2	組合 3	組合 4	組合 5	組合 6	組合 7	組合 8	組合 9
0	3	3	3	3	3	3	3	3	3
1	45	45	46	47	47	48	50	50	58
2	126	132	128	125	138	127	112	117	143
3	252	255	256	233	256	238	221	227	264
4	290	318	300	277	322	282	294	294	311
5	249	242	251	251	246	254	250	251	239
6	134	125	126	157	117	154	152	151	109
7	50	44	43	64	38	57	68	64	40
8	23	9	19	16	7	12	20	17	7
9	4	3	4	3	2	1	5	2	2
10	0	0	0	0	0	0	1	0	0

註：各組合組成標誌如下

組合 1：dd1001675、id1000948 或 id1000955、id6003492、id7000308、id8000978、id10000129、id10007137、id4005078、id6005779、id2010412

組合 2：dd1001675、id1000948 或 id1000955、id6003492、id7000308、id8000978、id10000129、id6012242、id4005078、id6005779、id11009990

組合 3：dd1001675、id1000948 或 id1000955、id6003492、id7000308、id8000978、id10000129、id2010412、id10007137、id6005779、id4005236

組合 4：dd1001675、id1000948 或 id1000955、id6003492、id7000308、id8000978、id10000129、id6012242、id4005078、id6006625、id9002494

組合 5：dd1001675、id1000948 或 id1000955、id6003492、id7000308、id8000978、id10000129、id6012242、id11009990、id6005779、id4005236

組合 6：dd1001675、id1000948 或 id1000955、id6003492、id7000308、id8000978、id10000129、id6012242、id4005236、id6006625、id9002494

組合 7：dd1001675、id1000948 或 id1000955、id6003492、id7000308、id8000978、id10000129、id6012242、id4005078、id6006625、id8000987

組合 8：dd1001675、id1000948 或 id1000955、id6003492、id7000308、id8000978、id10000129、id6012242、id4005236、id6006625、id8000987

組合 9：dd1001675、id1000948 或 id1000955、id6003492、id7000308、id8000978、id10000170、id6012242、id4005078、id6005779、id11009990



表 4、水稻 SNP 品種鑑別套件之 SNP 資訊。SNP 分子標誌的遺傳圖譜位置資料來自於 MapDisto Genetics Software web site (<http://mapdisto.free.fr/cMconverter/>) 上的 cM Converter；除了 PIC 值依本次研究結果計算，本表其餘資訊來自 Zhao 等學者(2011)文獻提供之 44 K SNP 分子標誌分析 413 個水稻品種之結果，做為此 44K SNP 的基本資料。

SNP id	Chr ^{註 1}	Pos ^{註 2}	cM ^{註 3}	Alleles ^{註 4}	Call rate	Hetero ^{註 5}	MAFall ^{註 6}	MAFaus	MAFind	MAFtej	MAFtrj	PIC
id1000955	1 p	1043946	7.28	G/A	0.985472	0	0.425061	0.25	0.344828	0.451613	0.135417	0.389838
dd1001675	1 q	42335049	177.18	G/A	0.997579	0	0.395631	0.140351	0.471264	0.368421	0.43299	0.483132
id2010412	2 q	24560641	98.66	G/A	0.992736	0	0.397561	0.267857	0.413793	0.4	0.257732	0.499792
id4005078	4 q	17434802	38.41	T/C	0.937046	0	0.302326	0.227273	0.47619	0.46875	0.092784	0.324865
id6003492	6 p	5181055	25.39	G/A	0.992736	0	0.480488	0.22807	0.383721	0.389474	0.257732	0.499792
id6005779	6 p	8979465	54.92	T/C	0.968523	0	0.3025	0.017857	0.345679	0.419355	0	0.149938
id7000308	7 p	1627513	7.56	G/A	0.970944	0	0.433915	0	0.116279	0.122222	0	0.369846
id8000978	8 p	3088782	26.1	C/A	0.992736	0	0.407317	0	0.310345	0.442105	0.103093	0.474802
id10000129	10 p	631626	1.83	A/G	0.96368	0	0.497487	0.446429	0.357143	0.244681	0.042553	0.424823
id10007137	10 q	22497992	83.57	A/C	0.995157	0	0.313869	0.089286	0.022989	0.5	0.453608	0.408163

註 1：Chr 為 SNP 位於第幾條染色體的位置，p 為短臂，q 為長臂。

註 2：Pos 為 SNP 在該染色體上的第幾個鹼基位置。

註 3：Alleles 為 SNP 的對偶基因型組合。

註 4：Hetero 為 SNP 基因座在 413 個水稻品種中異質結合的機率。

註 5：MAFall、MAFaus、MAFind、MAFtej 為 SNP 基因座於所有水稻品種 (all)、早中稻 (aus)、私稻 (ind)、溫帶粳稻 (tej) 之次要對偶基因頻度 (MAF) 值。

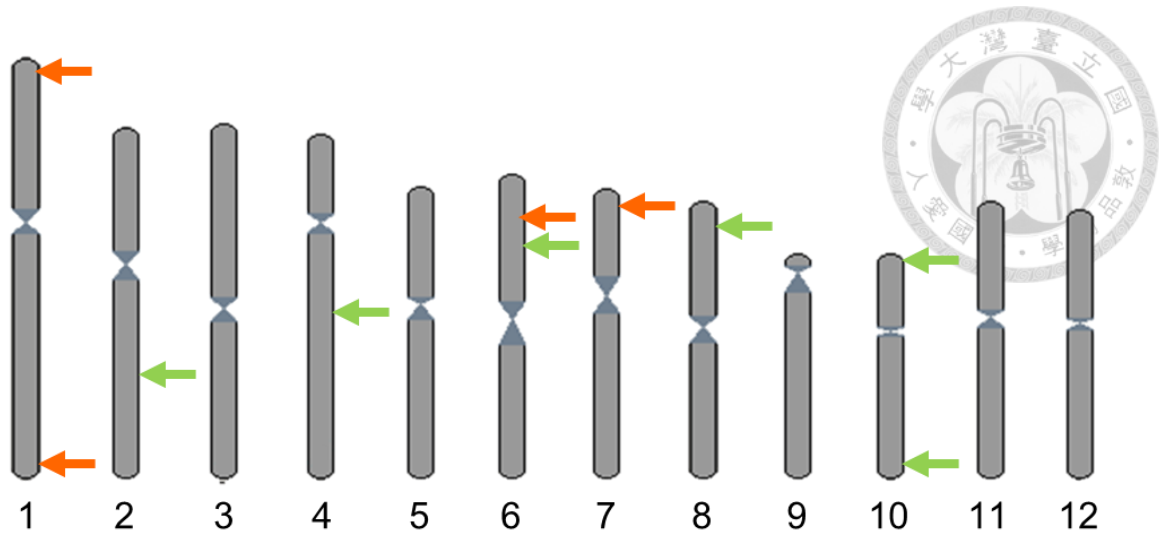
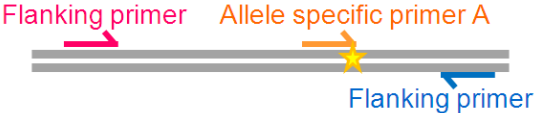
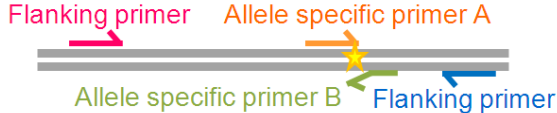

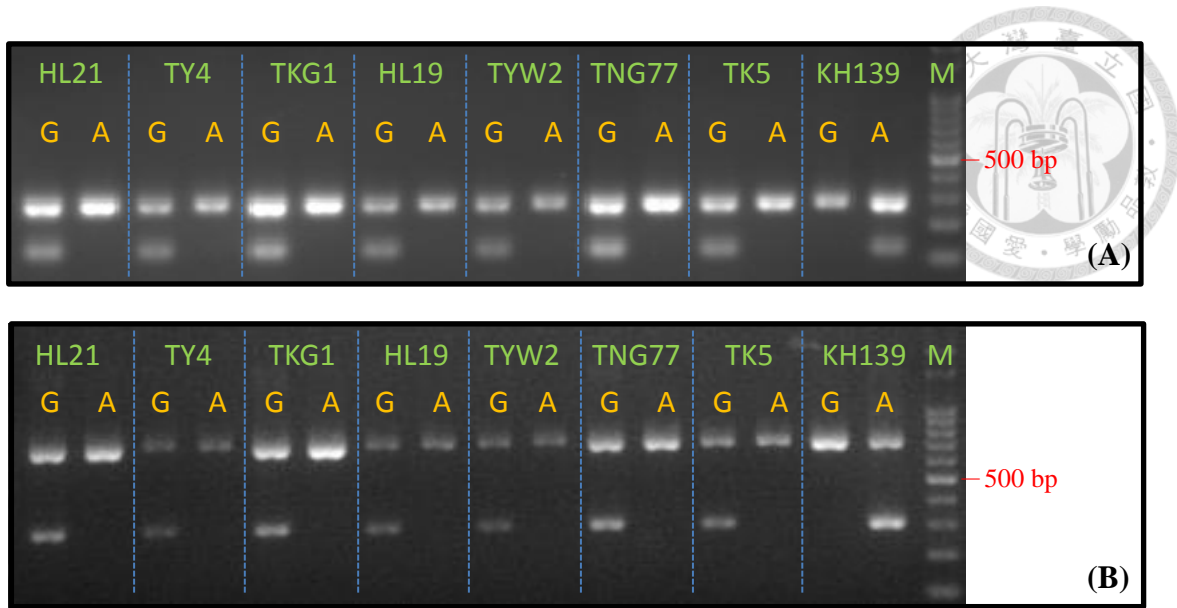


圖 4、水稻 SNP 品種鑑別套件於染色體位置示意圖。箭頭為 10 個 SNP 基因座之位置，橘色箭頭為兩兩品種間僅 1 個標誌差異的「必選」標誌，若不考慮必選之標誌，本次研究選到的 SNP 位置廣泛分佈於水稻的 12 條染色體。

表 5、AS 系統與 Tetra primer ARMS 系統之比較表。灰色為雙股 DNA，紅、藍、橘、綠色箭號分別代表側翼引子 (Flanking primer) 與對偶基因專一性引子 (Allele specific primer)，星號代表 SNP 位置，其序列與 AS-primer A 相符，與 AS-primer B 不符。

Allele specific system (AS-system)	Tetra primer ARMS system
(Reaction 1)	(Reaction 1)
	
(Reaction 2)	
	
3 primers / reaction	4 primers / reaction
1 genotype / reaction	2 genotype / reaction
設計成功率高 (130 個 SNP 位置可設計超過 90 組引子)	設計成功率低 (130 個 SNP 位置僅能設計 35 組引子)
PCR 適應性 ^{註1} 較高	PCR 適應性較低

註 1：PCR 適用之引子黏合溫度範圍越廣表示 PCR 適應性高；PCR 適用之引子黏合溫度範圍越窄表示 PCR 適應性低。



基因型讀值：

	HL21	TY4	TKG1	HL19	TYW2	TNG77	TK5	KH139
(A)	G	G	G	G	G	G	G	A
(B)	G	G	G	G	G	G	G	A

圖 5、比較側翼條帶大小差異對 AS 系統之影響。以 id6006625 SNP 基因座為例，設計不同側翼大小之引子，偵測 HL21、TY4、TKG1、HL19、TYW2、TNG77、TK5 和 KH139 等 8 個品種，每行分別偵測一個 SNP 基因型(A 或 G)；M 為 100-3,000 bp 之 DNA ladder。(A)側翼條帶大小為 286 bp，AS 條帶為 133 bp；(B)側翼條帶大小 706 bp，AS 條帶大小 305 bp。

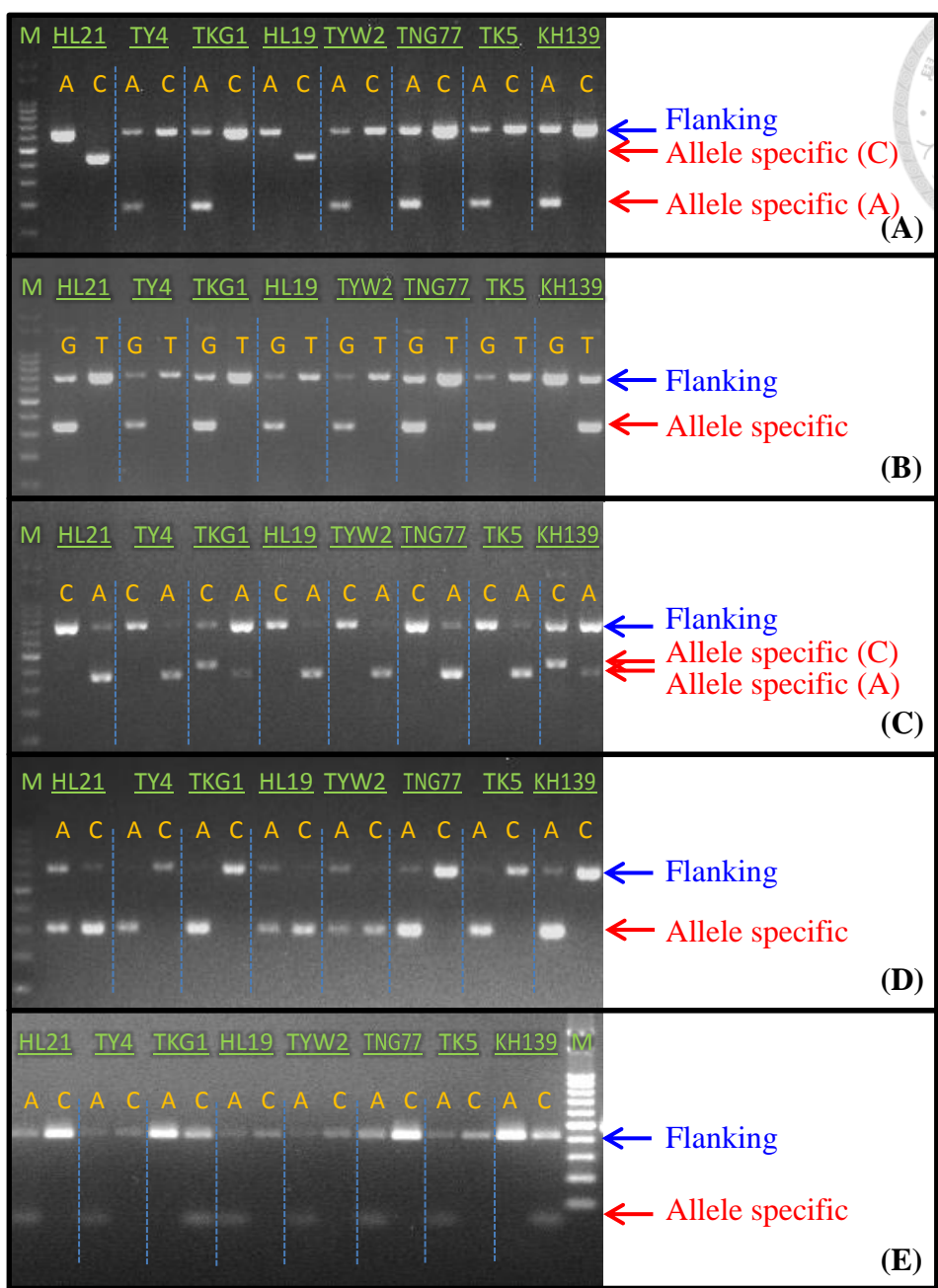
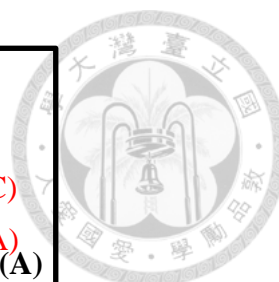


圖 6、AS-primer 方向及序列造成之結果差異。以 8 個水稻品種(HL21、TY4、TKG1、HL19、TYW2、TNG77、TK5 及 KH139)為例，M 為 100-3,000 bp 之 DNA ladder，每個行偵測 1 個 SNP 基因型(A、T、C 或 G)。(A)二 SNP 對偶基因之 AS-primer 互為反向且序列與參考序列吻合之典型結果；(B) 二 SNP 對偶基因之 AS-primer 同向且序列與參考序列吻合之典型結果；(C) 二 SNP 對偶基因之 AS-primer 互為反向且序列與參考序列吻合，AS-primer 與側翼引子互相競爭，造成結果非條帶有、無之差異，需從條帶強弱判斷基因型讀值；(D) 二 SNP 對偶基因之 AS-primer 同向且序列與參考序列吻合，AS-primer 與側翼引子互相競爭，造成結果非條帶有、無之差異；(E)二 SNP 對偶基因之 AS-primer 同向且在 AS-primer 3'端-3 位置設計一個配對錯誤。

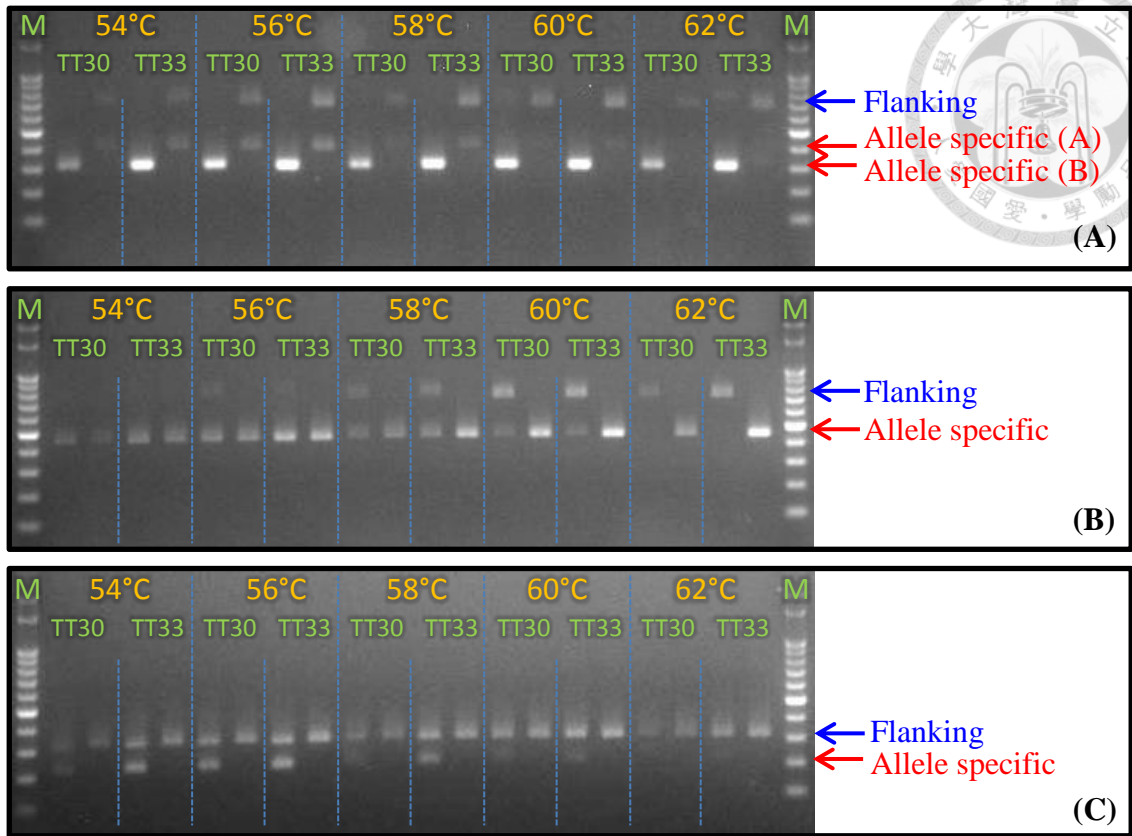



圖 7、AS-primer 之序列配對錯誤有無於梯溫 PCR 之差異比較。以 TT30 和 TT33 為例，M 為 100-3,000 bp 之 DNA ladder，每個行偵測 1 個 SNP 基因型。(A) AS-primer 序列與參考序列吻合時，通常增幅結果較不受溫度影響。(B) AS-primer 序列與參考序列吻合，判讀基因型可能需由條帶強弱分辨，當黏合溫度太低時可能誤判其基因型為異結合基因型。(C) AS-primer 3'端-3 位置具有 1 個錯誤配對鹼基，引子黏合溫度較低，但黏合性差異較大，條帶皆為有/無之差異。

第五章 結論

SNP 具突變率較低且於基因體中廣泛存在的優勢，使其適合作為品種鑑別的工具。偵測 SNP 方法多元，從單次偵測 1 個樣品中 1 個 SNP 基因型的低通量方法至以晶片或以次世代定序結合基因分型測序 (Genotyping-by-Sequencing, GBS) 之高通量方法皆有之。本研究目的為開發適用於鑑別臺灣優良水稻推薦品種的 SNP 品種鑑別系統，可實際應用在輔助水稻三級繁殖制度之水稻品種純度檢查及其他相關水稻品種或純度鑑別上，考量開發套件之使用彈性與檢測成本等因素，選擇以少量但優質的 SNP 組合組成本次開發的套件。

我們使用 Zhao 等學者 (2011) 以 44K SNP 晶片分析 413 個水稻得到的 35,746 筆 SNP 基因型分析資料為 SNP 篩選起點，先篩選 SNP 基因座的背景資訊，選擇成功讀值率高、SNP 類型為同質結合與異質結合變性 (denature) 溫度差異大的組合類型 (A/C、A/G、T/C 或 T/G)、位置廣泛分散於水稻 12 條染色體、於臺灣主要亞種秈稻與粳稻次族群間具有較佳 MAF 表現的標誌，得到共計 130 個候選 SNP 位置，再設計 AS-marker 進行實驗室測試，從這些 SNP 中先選擇可實際判讀且於臺灣粳稻次族群中遺傳距離較遠的品種間具多型性的 47 個標誌，對包含我國優良水稻推薦品種在內的 49 個目標品種基因型分析，最後以 CoreHunter 挑選涵蓋最廣泛品種間兩兩差異的最小 SNP 基因座組合，建立適用於我國優良水稻推薦品種組合的水稻 SNP 品種鑑別套件，此套件包含 10 個水稻 SNP 基因座，平均 PIC 值為 0.402，於本研究的 47 個水稻品種間的 1,176 個兩兩組合中，僅有 3 組遺傳背景極度相似的品種組合無法鑑別。而為方便套件建立與交流等，我們從篩選品種中選取遺傳距離最大且不具品種權保護的 3 個品種組成標準品種，其涵蓋 47 個品種中 100% 的對偶基因型。(圖 8)

藉由篩選 SNP 基因座背景資訊後再進行實驗測試，可有效提昇篩選到表現較優良的基因座，節省套件開發資源。本研究直接針對目標水稻品種挑選 SNP 基因座組成套件，選到的 SNP 組合僅適合本次非隨機取樣之子集合，雖選到的基因座為可達到該子集合最精簡、效率最高之組合，但其可能因取樣偏差造成建立的套件無法廣泛適用於其他水稻子集合，當目標品種改變時需進行調整以符合需求。理想中，以系統性的方式，依據每個 SNP 攜帶的資訊，挑選 SNP 基因座，組成水稻 SNP 品種鑑別套件，可廣泛適用於整體水稻，不需隨子集合改變調整套件內容，



然而，以臺灣水稻族群為例，其相對狹小的遺傳背景，將造成以系統性方式挑選出來的 SNP 組合，可能需要太多標誌才能達到完全區分目標水稻品種的效果，或在有限的標誌範圍中，無法完全辨別目標水稻品種，造成開發的套件無法符合實際兩兩品種鑑別的需求。兩相折衷的結果，倘若在分析平台或資源允許的前提下（如：open array 可單次分析 16 組 SNP 標誌），可結合本研究方法挑出最小適用標誌組合（10 個）後，依據候選 SNP 基因型分析結果，增加取用（6 個）攜帶資訊較佳的 SNP 至補足系統之單次分析數，使建立的套件能在「品種鑑別實務」及「帶有廣泛歧異度鑑別能力」間，取得一個平衡。


開發策略

研究結果

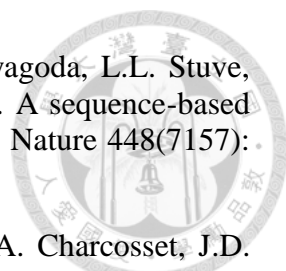


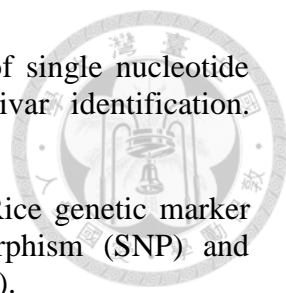
圖 8、開發策略流程圖與研究結果概要。本研究先篩選 SNP 背景資訊，留下廣佈於水稻基因體且表現較佳的 SNP 基因座，設計引子進行實驗篩選，將可實際使用的 SNP 對目標品種基因型分析後，挑選涵蓋最廣泛目標品種兩兩差異的 SNP 組合建立套件，再從建立的套件中挑選涵蓋最廣泛基因型且不具品種權保護的品種建立標準品種。

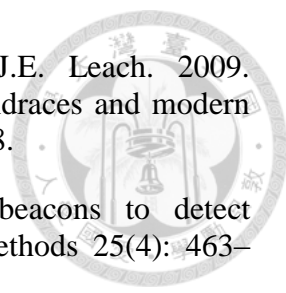
參考文獻

- 
- 丁文彥、林家玉、黃秋蘭、江瑞拱 (2013) 水稻新品種臺東 33 號之育成。臺東區農業改良場研究彙報 23: 1 - 16。
- 王映皓 (2007) 臺灣出土古稻米粒的初步研究。國立臺灣大學農藝所碩士論文。
- 呂椿棠、呂秀英 (2010) 台灣稻作資訊系統之研發及應用。台灣農業研究 59(1): 61-69。
- 林美瑄 (2011) 臺灣種子繁殖檢查制度。兩岸種苗科技研討會專刊 31 - 35。
- 林順福 (2005) 作物品種之 DNA 鑑定技術。21 世紀農業發展與新興科技應用研討會。
- 柯宇珊 (2014) 以簡單重覆序列開發水稻品種鑑別系統。國立臺灣大學農藝所碩士論文。
- 陳正昇、陳榮坤、金漢煊、林彥蓉 (2010) 以分子輔助選種導入 *hd1*、*Hd6* 和 *ehd1* 抽穗期基因至水稻越光品種。作物、環境與生物資訊 7: 1-20。
- 陳榮坤、林彥蓉、羅正宗 (2012) 水稻新品種臺南 16 號之育成。臺南區農業改良場研究彙報 60: 1 - 12。
- 陳榮坤、蔡孟勳、陳凱儀 (2013) 臺灣粳型水稻品種隨機型 SNP 分子標誌資料庫的建構。臺南區農業改良場研究彙報 61:15 - 28。
- 楊嘉凌 (2006) 豐產良質長糯米新品種台中私糯 2 號。台中區農業專訊 55: 1 - 1。
- Ahmadian, A., B. Gharizadeh, A.C. Gustafsson, F. Sterky, P. Nyren, M. Uhlén, and J. Lundeberg. 2000. Single-nucleotide polymorphism analysis by pyrosequencing. *Anal. Biochem.* 280(1): 103-110.
- Alexandrov, N., S. Tai, W. Wang, L. Mansueto, K. Palis, R.R. Fuentes, V.J. Ulat, D. Chebotarov, G. Zhang, Z. Li, R. Mauleon, R.S. Hamilton, and K.L. McNally. 2014. SNP-Seek database of SNPs derived from 3000 rice genomes. *Nucleic Acids Res.*
- Baird, N.A., P.D. Etter, T.S. Atwood, M.C. Currey, A.L. Shiver, Z.A. Lewis, E.U. Selker, W.A. Cresko, and E.A. Johnson. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3(10): e3376.
- Beaudet, L., J. Bédard, B. Breton, R.J. Mercuri, and M.L. Budarf. 2001. Homogeneous assays for single-nucleotide polymorphism typing using alphaScreen. *Genome Res.* 11(4): 600-608.
- Bertin, I., J.H. Zhu, and M.D. Gale. 2005. SSCP-SNP in pearl millet—a new marker

- system for comparative genetics. *Theor. Appl. Genet.* 110(8): 1467–1472.
- Beukelaer, H.D., P. Smýkal, G.F. Davenport, and V. Fack. 2012. Core Hunter II: fast core subset selection based on multiple genetic diversity measures using Mixed Replica search. *BMC Bioinformatics* 13(1): 312.
- Borba, T.C. de O., R.P.V. Brondani, P.H.N. Rangel, and C. Brondani. 2009. Microsatellite marker-mediated analysis of the EMBRAPA Rice Core Collection genetic diversity. *Genetica* 137(3): 293–304.
- Brenan, C., and T. Morrison. 2005. High throughput, nanoliter quantitative PCR. *Drug Discov. Today Technol.* 2(3): 247–253.
- Cabezas, J.A., J. Ibáñez, D. Lijavetzky, D. Vélez, G. Bravo, V. Rodríguez, I. Carreño, A.M. Jermakow, J. Carreño, L. Ruiz-García, M.R. Thomas, and J.M. Martínez-Zapater. 2011. A 48 SNP set for grapevine cultivar identification. *BMC Plant Biol.* 11(1): 153.
- Chagné, D., R.N. Crowhurst, M. Troglio, M.W. Davey, B. Gilmore, C. Lawley, S. Vanderzande, R.P. Hellens, S. Kumar, A. Cestaro, R. Velasco, D. Main, J.D. Rees, A. Iezzoni, T. Mockler, L. Wilhelm, E. Van de Weg, S.E. Gardiner, N. Bassil, and C. Peace. 2012. Genome-wide SNP detection, validation, and development of an 8K SNP array for apple. *PLoS ONE* 7(2): e31745.
- Consolandi, C., L. Palmieri, S. Doveri, E. Maestri, N. Marmiroli, S. Reale, D. Lee, L. Baldoni, N. Tosti, M. Severgnini, G. De Bellis, and B. Castiglioni. 2007. Olive variety identification by ligation detection reaction in a universal array format. *J. Biotechnol.* 129(3): 565–574.
- Dominik, S., J.M. Henshall, P.D. Kube, H. King, S. Lien, M.P. Kent, and N.G. Elliott. 2010. Evaluation of an Atlantic salmon SNP chip as a genomic tool for the application in a Tasmanian Atlantic salmon (*Salmo salar*) breeding population. *Aquaculture* 308, Supplement 1: S56–S61.
- Ebana, K., Y. Kojima, S. Fukuoka, T. Nagamine, and M. Kawase. 2008. Development of mini core collection of Japanese rice landrace. *Breed. Sci.* 58(3): 281–291.
- Ebana, K., J. Yonemaru, S. Fukuoka, H. Iwata, H. Kanamori, N. Namiki, H. Nagasaki, and M. Yano. 2010. Genetic structure revealed by a whole-genome single-nucleotide polymorphism survey of diverse accessions of cultivated Asian rice (*Oryza sativa* L.). *Breed. Sci.* 60(4): 390–397.
- Fang, W.P., L.W. Meinhardt, H.W. Tan, L. Zhou, S. Mischke, and D. Zhang. 2014. Varietal identification of tea (*Camellia sinensis*) using nanofluidic array of single nucleotide polymorphism (SNP) markers. *Hortic. Res.* 1: 14035.
- Feltus, F.A., J. Wan, S.R. Schulze, J.C. Estill, N. Jiang, and A.H. Paterson. 2004. An SNP resource for rice genetics and breeding based on subspecies Indica and Japonica genome alignments. *Genome Res.* 14(9): 1812–1819.
- Frazer, K.A., E. Eskin, H.M. Kang, M.A. Bogue, D.A. Hinds, E.J. Beilharz, R.V. Gupta,

- 
- J. Montgomery, M.M. Morenzoni, G.B. Nilsen, C.L. Pethiyagoda, L.L. Stuve, F.M. Johnson, M.J. Daly, C.M. Wade, and D.R. Cox. 2007. A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. *Nature* 448(7157): 1050–1053.
- Ganal, M.W., G. Durstewitz, A. Polley, A. Bérard, E.S. Buckler, A. Charcosset, J.D. Clarke, E.M. Graner, M. Hansen, J. Joets, M.C. Le Paslier, M.D. McMullen, P. Montalent, M. Rose, C.-C. Schön, Q. Sun, H. Walter, O.C. Martin, and M. Falque. 2011. A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. *PLoS ONE* 6(12): e28334.
- Gao, Z.-Y., S.C. Zhao, W.-M. He, L.B. Guo, Y.L. Peng, J.J. Wang, X.S. Guo, X.M. Zhang, Y.C. Rao, C. Zhang, G.J. Dong, F.Y. Zheng, C.X. Lu, J. Hu, Q. Zhou, H.J. Liu, H.Y. Wu, J. Xu, P.X. Ni, D.L. Zeng, D.H. Liu, P. Tian, L.H. Gong, C. Ye, G.H. Zhang, J. Wang, F.K. Tian, D.W. Xue, Y. Liao, L. Zhu, M.S. Chen, J.Y. Li, S.H. Cheng, G.Y. Zhang, J. Wang, and Q. Qian. 2013. Dissecting yield-associated loci in super hybrid rice by resequencing recombinant inbred lines and improving parental genome sequences. *Proc. Natl. Acad. Sci.* 110(35): 14492–14497.
- Ge, W., M.E. Davis, H.C. Hines, and K.M. Irvin. 1999. Polymorphism in exon 10 of the bovine GHR gene detected by PCR-DGGE. *Anim. Genet.* 30(2): 167–168.
- Guo, S., J. Zhang, H. Sun, J. Salse, W.J. Lucas, H. Zhang, Y. Zheng, L. Mao, Y. Ren, Z. Wang, J. Min, X. Guo, F. Murat, B.K. Ham, Z. Zhang, S. Gao, M. Huang, Y. Xu, S. Zhong, A. Bombarely, L.A. Mueller, H. Zhao, H. He, Y. Zhang, Z. Zhang, S. Huang, T. Tan, E. Pang, K. Lin, Q. Hu, H. Kuang, P. Ni, B. Wang, J. Liu, Q. Kou, W. Hou, X. Zou, J. Jiang, G. Gong, K. Klee, H. Schoof, Y. Huang, X. Hu, S. Dong, D. Liang, J. Wang, K. Wu, Y. Xia, X. Zhao, Z. Zheng, M. Xing, X. Liang, B. Huang, T. Lv, J. Wang, Y. Yin, H. Yi, R. Li, M. Wu, A. Levi, X. Zhang, J.J. Giovannoni, J. Wang, Y. Li, Z. Fei, and Y. Xu. 2013. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nat. Genet.* 45(1): 51–58.
- Heaton, M.P., G.P. Harhay, G.L. Bennett, R.T. Stone, W.M. Grosse, E. Casas, J.W. Keele, T.P.L. Smith, C.G. Chitko-McKown, and W.W. Laegreid. 2002. Selection and use of SNP markers for animal identification and paternity analysis in U.S. beef cattle. *Mamm. Genome* 13(5): 272–281.
- Huang, T.J., M. Liu, L.D. Knight, W.W. Grody, J.F. Miller, and C.M. Ho. 2002. An electrochemical detection scheme for identification of single nucleotide polymorphisms using hairpin-forming probes. *Nucleic Acids Res.* 30(12): e55.
- Ishii, K., and M. Fukui. 2001. Optimization of annealing temperature to reduce bias caused by a primer mismatch in multitemplate PCR. *Appl. Environ. Microbiol.* 67(8): 3753–3755.
- Jobes, D.V., D.L. Hurley, and L.B. Thien. 1995. Plant DNA isolation: A method to efficiently remove polyphenolics, polysaccharides, and RNA. *Taxon* 44(3): 379–386.

- 
- Jung, J., S.W. Park, W.Y. Liu, and B.C. Kang. 2010. Discovery of single nucleotide polymorphism in *Capsicum* and SNP markers for cultivar identification. *Euphytica* 175(1): 91–107.
- Kim, C., U. Yoon, G. Lee, H. Lee, Y. Kim, and J. Hahn. 2009. Rice genetic marker database: An identification of single nucleotide polymorphism (SNP) and quantitative trait loci (QTL) markers. *Afr. J. Biotechnol.* 8(13).
- Kinoshita-Kikuta, E., E. Kinoshita, and T. Koike. 2002. A novel procedure for simple and efficient genotyping of single nucleotide polymorphisms by using the Zn²⁺-cyclen complex. *Nucleic Acids Res.* 30(22): e126.
- Kojima, Y., K. Ebana, S. Fukuoka, T. Nagamine, and M. Kawase. 2005. Development of an RFLP-based rice diversity research set of germplasm. *Breed. Sci.* 55(4): 431–440.
- Kota, R., R.K. Varshney, T. Thiel, K.J. Dehmer, and A. Graner. 2001. Generation and comparison of EST-derived SSRs and SNPs in barley (*Hordeum Vulgare* L.). *Hereditas* 135(2-3): 145–151.
- Labate, J.A., and A.M. Baldo. 2005. Tomato SNP discovery by EST mining and resequencing. *Mol. Breed.* 16(4): 343–349.
- Lam, H.M., X. Xu, X. Liu, W. Chen, G. Yang, F.L. Wong, M.W. Li, W. He, N. Qin, B. Wang, J. Li, M. Jian, J. Wang, G. Shao, J. Wang, S.S.M. Sun, and G. Zhang. 2010. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nat. Genet.* 42(12): 1053–1059.
- Latorra, D., K. Campbell, A. Wolter, and J.M. Hurley. 2003. Enhanced allele-specific PCR discrimination in SNP genotyping using 3' locked nucleic acid (LNA) primers. *Hum. Mutat.* 22(1): 79–85.
- Li, X., Y. Lu, J. Li, H. Xu, and M.Q. Shahid. 2011. Strategies on sample size determination and qualitative and quantitative traits integration to construct core collection of rice (*Oryza sativa*). *Rice Sci.* 18(1): 46–55.
- Liu, S., Z. Zhou, J. Lu, F. Sun, S. Wang, H. Liu, Y. Jiang, H. Kucuktas, L. Kaltenboeck, E. Peatman, and Z. Liu. 2011. Generation of genome-scale gene-associated SNPs in catfish for the construction of a high-density SNP array. *BMC Genomics* 12(1): 53.
- Lyamichev, V., M.A. Brow, and J.E. Dahlberg. 1993. Structure-specific endonucleolytic cleavage of nucleic acids by eubacterial DNA polymerases. *Science* 260(5109): 778–783.
- McCallum, C.M., L. Comai, E.A. Greene, and S. Henikoff. 2000. Targeted screening for induced mutations. *Nat. Biotechnol.* 18(4): 455–457.
- McNally, K.L., K.L. Childs, R. Bohnert, R.M. Davidson, K. Zhao, V.J. Ulat, G. Zeller, R.M. Clark, D.R. Hoen, T.E. Bureau, R. Stokowski, D.G. Ballinger, K.A. Frazer, D.R. Cox, B. Padhukasahasram, C.D. Bustamante, D. Weigel, D.J. Mackill, R.M.

- 
- Bruskiewich, G. Rättsch, C.R. Buell, H. Leung, and J.E. Leach. 2009. Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. *Proc. Natl. Acad. Sci.* 106(30): 12273–12278.
- Mhlanga, M.M., and L. Malmberg. 2001. Using molecular beacons to detect single-nucleotide polymorphisms with Real-Time PCR. *Methods* 25(4): 463–471.
- Nikiforov, T.T., R.B. Rendie, P. Goelet, Y.H. Rogers, M.L. Kotewicz, S. Anderson, G.L. Trainor, and M.R. Knapp. 1994. Genetic Bit Analysis: a solid phase method for typing single nucleotide polymorphisms. *Nucleic Acids Res.* 22(20): 4167–4175.
- Nilsson, M., K. Krejci, J. Koch, M. Kwiatkowski, P. Gustavsson, and U. Landegren. 1997. Padlock probes reveal single-nucleotide differences, parent of origin and *in situ* distribution of centromeric sequences in human chromosomes 13 and 21. *Nat. Genet.* 16(3): 252–255.
- Odong, T.L., J. Jansen, F.A. van Eeuwijk, and T.J.L. van Hintum. 2012. Quality of core collections for effective utilisation of genetic resources review, discussion and interpretation. *Theor. Appl. Genet.* 126(2): 289–305.
- Orum, H., P.E. Nielsen, M. Egholm, R.H. Berg, O. Buchardt, and C. Stanley. 1993. Single base pair mutation analysis by PNA directed PCR clamping. *Nucleic Acids Res.* 21(23): 5332–5336.
- Patterson, N., A.L. Price, and D. Reich. 2006. Population structure and eigenanalysis. *PLoS Genet* 2(12): e190.
- Peace, C., N. Bassil, D. Main, S. Ficklin, U.R. Rosyara, T. Stegmeir, A. Sebolt, B. Gilmore, C. Lawley, T.C. Mockler, D.W. Bryant, L. Wilhelm, and A. Iezzoni. 2012. Development and evaluation of a genome-wide 6K SNP array for diploid sweet cherry and tetraploid sour cherry. *PLoS ONE* 7(12): e48305.
- Ramos, A.M., R.P.M.A. Crooijmans, N.A. Affara, A.J. Amaral, A.L. Archibald, J.E. Beever, C. Bendixen, C. Churcher, R. Clark, P. Dehais, M.S. Hansen, J. Hedegaard, Z.L. Hu, H.H. Kerstens, A.S. Law, H.J. Megens, D. Milan, D.J. Nonneman, G.A. Rohrer, M.F. Rothschild, T.P.L. Smith, R.D. Schnabel, C.P. Van Tassell, J.F. Taylor, R.T. Wiedmann, L.B. Schook, and M.A.M. Groenen. 2009. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. *PLoS ONE* 4(8): e6524.
- Reale, S., S. Doveri, A. Díaz, A. Angiolillo, L. Lucentini, F. Pilla, A. Martín, P. Donini, and D. Lee. 2006. SNP-based markers for discriminating olive (*Olea europaea* L.) cultivars. *Genome* 49(9): 1193–1205.
- Reed, G.H., and C.T. Wittwer. 2004. Sensitivity and specificity of single-nucleotide polymorphism scanning by high-resolution melting analysis. *Clin. Chem.* 50(10): 1748–1754.
- Rosenbaum, V., and D. Riesner. 1987. Temperature-gradient gel electrophoresis:

Thermodynamic analysis of nucleic acids and proteins in purified form and in cellular extracts. *Biophys. Chem.* 26(2–3): 235–246.

Schmid, K.J., T.R. Sørensen, R. Stracke, O. Törjék, T. Altmann, T. Mitchell-Olds, and B. Weisshaar. 2003. Large-scale identification and analysis of genome-wide single-nucleotide polymorphisms for mapping in *arabidopsis thaliana*. *Genome Res.* 13(6a): 1250–1257.

Sim, S.C., G. Durstewitz, J. Plieske, R. Wieseke, M.W. Ganal, A. Van Deynze, J.P. Hamilton, C.R. Buell, M. Causse, S. Wijeratne, and D.M. Francis. 2012. Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. *PLoS ONE* 7(7): e40563.

Sipos, R., A.J. Székely, M. Palatinszky, S. Révész, K. Márialigeti, and M. Nikolausz. 2007. Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targeting bacterial community analysis. *FEMS Microbiol. Ecol.* 60(2): 341–350.

Soleimani, V.D., B.R. Baum, and D.A. Johnson. 2003. Efficient validation of single nucleotide polymorphisms in plants by allele-specific PCR, with an example from barley. *Plant Mol. Biol. Report.* 21(3): 281–288.

Thachuk, C., J. Crossa, J. Franco, S. Dreisigacker, M. Warburton, and G.F. Davenport. 2009. Core Hunter: an algorithm for sampling genetic resources based on multiple genetic measures. *BMC Bioinformatics* 10: 243.

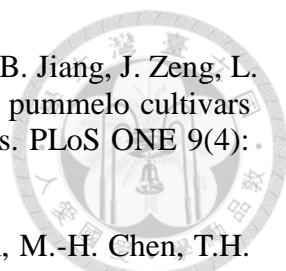
Thomson, M.J., K. Zhao, M. Wright, K.L. McNally, J. Rey, C.W. Tung, A. Reynolds, B. Scheffler, G. Eizenga, A. McClung, H. Kim, A.M. Ismail, M. de Ocampo, C. Mojica, M.Y. Reveche, C.J. Dilla-Ermita, R. Mauleon, H. Leung, C. Bustamante, and S.R. McCouch. 2011. High-throughput single nucleotide polymorphism genotyping for breeding applications in rice using the BeadXpress platform. *Mol. Breed.* 29(4): 875–886.

Tung, C.W., K. Zhao, M.H. Wright, M.L. Ali, J. Jung, J. Kimball, W. Tyagi, M.J. Thomson, K. McNally, H. Leung, H. Kim, S.N. Ahn, A. Reynolds, B. Scheffler, G. Eizenga, A. McClung, C. Bustamante, and S.R. McCouch. 2010. Development of a research platform for dissecting phenotype–genotype associations in rice (*Oryza* spp.). *Rice* 3(4): 205–217.

Ullman, E.F., H. Kirakossian, S. Singh, Z.P. Wu, B.R. Irvin, J.S. Pease, A.C. Switchenko, J.D. Irvine, A. Dafforn, and C.N. Skold. 1994. Luminescent oxygen channeling immunoassay: measurement of particle binding kinetics by chemiluminescence. *Proc. Natl. Acad. Sci.* 91(12): 5426–5430.

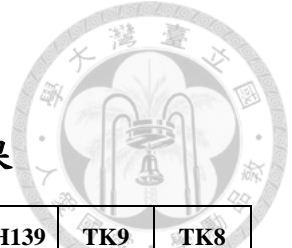
Wang, H., J. Yue, M. Han, J. Yang, and Y. Zhao. 2010. Rapid method for identification of six common species of mycobacteria based on multiplex SNP analysis. *J. Clin. Microbiol.* 48(1): 247–250.

Wolford, J., D. Blunt, C. Ballecer, and M. Prochazka. 2000. High-throughput SNP detection by using DNA pooling and denaturing high performance liquid chromatography (DHPLC). *Hum. Genet.* 107(5): 483–487.

- 
- Wu, B., G. Zhong, J. Yue, R. Yang, C. Li, Y. Li, Y. Zhong, X. Wang, B. Jiang, J. Zeng, L. Zhang, S. Yan, X. Bei, and D. Zhou. 2014. Identification of pummelo cultivars by using a panel of 25 selected SNPs and 12 DNA segments. *PLoS ONE* 9(4): e94506.
- Yan, W., J.N. Rutger, R.J. Bryant, H.E. Bockelman, R.G. Fjellstrom, M.-H. Chen, T.H. Tai, and A.M. McClung. 2007. Development and evaluation of a core subset of the USDA rice germplasm collection. *Crop Sci.* 47(2): 869–876.
- Ye, S., S. Dhillon, X. Ke, A.R. Collins, and I.N.M. Day. 2001. An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Res.* 29(17): e88.
- Yonemaru, J., K. Ebana, and M. Yano. 2014. HapRice, an SNP haplotype database and a web tool for rice. *Plant Cell Physiol.* 55(1): e9–e9.
- You, F., N. Huo, Y. Gu, M. Luo, Y. Ma, D. Hane, G. Lazo, J. Dvorak, and O. Anderson. 2008. BatchPrimer3: A high throughput web application for PCR and sequencing primer design. *BMC Bioinformatics* 9: 253.
- Yu, H., W. Xie, J. Li, F. Zhou, and Q. Zhang. 2014. A whole-genome SNP array (RICE6K) for genomic breeding in rice. *Plant Biotechnol. J.* 12(1): 28–37.
- Zhang, P., J. Li, X. Li, X. Liu, X. Zhao, and Y. Lu. 2011. Population structure and genetic diversity in a rice core collection (*Oryza sativa* L.) investigated with SSR markers. *PLoS ONE* 6(12): e27565.
- Zhao, K., C.-W. Tung, G.C. Eizenga, M.H. Wright, M.L. Ali, A.H. Price, G.J. Norton, M.R. Islam, A. Reynolds, J. Mezey, A.M. McClung, C.D. Bustamante, and S.R. McCouch. 2011. Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* 2: 467.
- Zhao, K., M. Wright, J. Kimball, G. Eizenga, A. McClung, M. Kovach, W. Tyagi, M.L. Ali, C.-W. Tung, A. Reynolds, C.D. Bustamante, and S.R. McCouch. 2010. Genomic diversity and introgression in *O. sativa* reveal the impact of domestication and breeding on the rice genome. *PLoS ONE* 5(5): e10780.
- Zhao, H., W. Yao, Y. Ouyang, W. Yang, G. Wang, X. Lian, Y. Xing, L. Chen, and W. Xie. 2015. RiceVarMap: a comprehensive database of rice genomic variations. *Nucleic Acids Res.* 43(D1): D1018–D1022.
- Zheng, T., H. Yu, H. Zhang, Z. Wu, W. Wang, S. Tai, L. Chi, J. Ruan, C. Wei, J. Shi, Y. Gao, B. Fu, Y. Zhou, X. Zhao, F. Zhang, McNALLY Kenneth L., Z. Li, G. Zhang, J. Li, D. Zhang, J. Xu, and Z. Li. 2015. Rice functional genomics and breeding database (RFGB)-3K-rice SNP and InDel sub-database. *Chin. Sci. Bull.* 60(4): 367–371.

附錄

附錄 1、以 47 個標誌對 49 個水稻品種基因型分析結果



Variety Marker	TN11	TK14	TC192	TCS10	TCSG2	TK2	TK16	KH139	TK9	TK8
dd1001675	A	A	A	G	G	G	A	G	A	A
id11009466	A	A	A	G	G	A	A	G	A	A
id12008090	C	C	C	C	C	C	C	C	C	C
id2010155	G	G	G	A	G	A	G	A	G	G
id4005236	A	A	A	G	G	G	A	G	G	G
id4011022	T	T	T	T	T	T	T	T	T	T
id4011023	G	G	G	G	G	G	G	G	G	G
id7000308	A	A	A	G	A	A	G	A	A	A
id7002851	T	G	G	N	N	G	G	T	G	G
id9002505	G	G	G	G	N	G	G	G	G	G
id10000129	G	A	A	A	G	G	G	A	G	G
id10000170	A	C	C	A	A	A	A	C	A	A
id10001501	G	G	G	G	A	G	G	A	G	G
id10007165	C	C	C	C	C	C	C	T	C	C
id10007166	C	C	C	C	C	C	C	T	C	C
id1000948	T	T	T	C	C	T	T	T	T	T
id1000955	G	G	G	A	A	G	G	G	G	G
id1027576	T	T	T	C	C	C	T	T	T	T
id11009990	T	C	C	T	T	T	T	C	T	T
id12003239	T	T	T	C	C	T	T	T	T	T
id12008144	C	C	C	C	C	C	C	C	C	C
id2010166	A	A	A	G	G	A	A	A	A	A
id2010412	A	A	A	A	G	G	A	G	A	A
id2010974	A	A	A	G	A	A	A	A	A	A
id2010992	C	C	C	C	C	C	C	C	C	C
id6005779	T	T	T	T	T	T	T	T	C	T
id6012242	G	G	G	G	G	G	G	T	G	G
id8000978	C	C	C	A	A	C	A	A	A	C
id8002106	G	G	G	G	T	G	G	T	G	G
id9002494	A	A	A	A	G	A	A	A	A	A
wd7002626	A	A	A	G	G	A	A	A	A	A
id6006625	G	A	A	A	A	A	A	A	G	G
id8000987	C	C	C	C	C	C	C	C	C	C
id11002184	G	A	A	G	G	G	G	A	G	G
id2010969	T	T	T	T	G	T	T	T	T	T
id4005078	T	T	T	T	C	C	T	C	C	C
id7000276	A	A	A	C	A	A	A	A	A	A
wd9002354	T	C	C	C	C	T	C	T	C	T
id1027513	A	A	A	C	C	C	A	A	A	A
id11001787	C	C	C	C	C	C	C	C	C	C
id11002213	G	A	A	N	N	G	G	A	G	G
id12008149	G	G	G	G	G	G	G	G	G	G
id3004807	A	A	A	G	G	A	A	A	A	A
id6003492	A	A	G	G	G	G	G	A	A	A
id7002784-4	A	A	A	G	G	A	A	A	A	A
id10007137	C	C	C	C	C	C	C	A	C	C
id4011112	T	T	T	C	C	T	T	T	T	T

Variety Marker	TT30	TNG84	TKGlu3	TK4	TCS17	TT33	KH146	TKGlu1	TNG71	TCSG1
dd1001675	A	G	A	G	A	A	G	A	G	G
id11009466	A	A	G	G	G	A	G	G	G	G
id12008090	C	C	C	C	T	C	C	C	C	C
id2010155	A	G	A	A	G	A	A	A	G	G
id4005236	G	G	G	G	G	G	G	G	G	G
id4011022	T	T	T	T	C	T	T	T	T	T
id4011023	G	G	G	G	A	G	G	G	G	G
id7000308	G	A	A	A	A	G	A	A	A	A
id7002851	T	G	G	G	T	T	T	G	T	N
id9002505	G	G	G	G	A	G	G	G	G	A
id10000129	G	G	G	A	G	G	G	G	A	G
id10000170	A	A	A	C	A	A	A	A	C	A
id10001501	G	G	G	G	A	G	G	G	G	A
id10007165	C	T	C	C	C	C	T	C	T	C
id10007166	C	T	C	C	C	C	T	C	T	C
id1000948	T	T	T	T	C	T	C	T	T	C
id1000955	G	G	G	G	A	G	A	G	G	A
id1027576	T	C	T	C	T	T	C	T	C	C
id11009990	T	C	T	T	C	T	C	T	C	T
id12003239	T	T	T	T	C	T	T	T	T	C
id12008144	C	C	C	C	T	C	C	C	C	C
id2010166	A	A	A	A	G	A	A	A	G	G
id2010412	G	A	G	A	G	G	G	G	G	G
id2010974	A	A	A	A	G	A	A	A	A	A
id2010992	C	C	C	C	T	C	C	C	C	C
id6005779	T	T	T	T	T	T	T	T	T	T
id6012242	G	G	T	G	G	G	G	G	G	G
id8000978	A	A	A	A	A	A	C	A	C	A
id8002106	G	G	G	G	G	G	G	G	G	T
id9002494	A	A	A	A	G	A	A	A	A	G
wd7002626	A	A	A	A	G	A	A	A	A	G
id6006625	G	A	G	G	A	G	G	G	G	A
id8000987	C	C	C	C	C	C	T	C	C	C
id11002184	G	G	G	G	G	G	G	A	G	G
id2010969	T	T	T	T	T	T	T	T	T	G
id4005078	C	C	C	C	C	C	C	C	C	C
id7000276	C	A	A	A	A	C	A	A	A	A
wd9002354	C	C	T	C	C	C	T	T	T	C
id1027513	A	C	A	C	A	A	C	A	C	C
id11001787	T	C	C	C	C	T	T	C	C	C
id11002213	G	G	G	G	N	G	G	A	G	N
id12008149	G	G	G	G	A	G	G	G	G	G
id3004807	G	A	A	A	G	G	A	A	A	G
id6003492	A	G	G	A	G	A	A	G	A	G
id7002784-4	A	A	A	A	G	A	A	A	A	G
id10007137	C	A	A	C	C	C	A	C	A	C
id4011112	C	T	T	T	T	C	T	T	T	C

Variety Marker	KH145	TY3	TN13	TK11	KH147	TNGW 73	TN14	TK5	TN16	KHS7
dd1001675	A	A	G	A	A	A	A	G	G	A
id11009466	A	G	G	G	G	G	G	G	G	A
id12008090	C	C	C	T	C	C	C	C	T	C
id2010155	A	A	G	A	A	A	G	A	G	G
id4005236	A	G	G	G	A	G	G	G	G	G
id4011022	T	T	T	T	T	T	C	T	C	T
id4011023	G	G	G	G	G	G	A	G	A	G
id7000308	A	A	A	A	A	A	A	A	A	G
id7002851	G	G	T	G	G	G	T	G	T	N
id9002505	G	G	G	G	G	G	G	G	G	A
id10000129	G	A	A	G	G	A	G	G	A	G
id10000170	A	C	C	A	A	A	A	A	C	A
id10001501	G	G	G	G	G	G	G	G	G	G
id10007165	T	C	C	T	C	C	C	T	T	C
id10007166	T	C	C	T	C	C	C	T	T	C
id1000948	T	T	C	T	T	T	T	T	C	T
id1000955	G	G	A	G	G	G	G	G	A	G
id1027576	T	T	C	T	T	T	T	T	C	T
id11009990	T	T	C	C	C	T	T	C	T	C
id12003239	T	T	T	T	T	T	T	T	T	C
id12008144	C	C	C	C	C	C	C	C	T	C
id2010166	A	A	A	A	A	A	A	A	G	G
id2010412	G	G	G	G	G	G	A	G	G	A
id2010974	A	A	A	A	A	A	A	A	A	G
id2010992	C	C	C	C	C	C	C	C	C	T
id6005779	T	T	T	T	T	T	T	T	T	T
id6012242	G	T	G	T	G	T	G	T	T	G
id8000978	A	A	C	C	A	C	C	A	A	A
id8002106	G	G	G	G	G	G	G	G	G	G
id9002494	A	A	A	A	A	A	A	A	A	G
wd7002626	A	A	A	A	A	A	A	A	A	A
id6006625	G	G	G	G	G	G	G	G	G	A
id8000987	C	C	C	T	C	T	T	C	C	C
id11002184	A	G	G	G	A	G	G	G	G	G
id2010969	T	T	T	T	T	T	T	T	T	T
id4005078	T	C	C	C	T	C	C	C	C	C
id7000276	A	A	A	A	A	A	A	A	A	C
wd9002354	C	C	C	T	C	T	C	T	T	C
id1027513	A	A	C	A	A	A	A	A	C	A
id11001787	C	C	C	T	C	C	C	C	C	C
id11002213	A	G	G	G	A	G	G	G	G	G
id12008149	G	G	G	G	G	G	G	G	A	G
id3004807	A	A	A	A	A	A	A	A	A	G
id6003492	A	A	A	A	G	A	G	A	G	A
id7002784-4	A	A	A	A	A	A	A	A	A	A
id10007137	A	C	C	C	C	C	C	A	A	C
id4011112	T	T	T	T	T	T	T	T	T	C

Variety Marker	HL21	TTG31	TNGS 14	TNG77	TC194	TKGlu5	TNGS 22	TNG79	TS2	台中 在來 1
dd1001675	A	A	A	G	G	A	A	A	G	A
id11009466	A	A	G	A	G	G	G	A	G	A
id12008090	C	C	C	C	C	C	C	C	C	C
id2010155	G	G	G	G	G	A	A	A	A	G
id4005236	A	G	G	G	G	G	G	G	G	G
id4011022	T	T	T	T	C	T	T	T	T	C
id4011023	G	G	G	G	A	G	G	G	G	A
id7000308	G	A	A	A	A	A	A	A	G	G
id7002851	N	G	N	G	G	G	N	G	N	T
id9002505	G	N	G	A	A	G	G	G	G	A
id10000129	G	G	A	G	G	G	G	G	A	A
id10000170	A	A	A	A	A	A	A	A	A	N
id10001501	G	G	G	G	G	G	A	G	G	A
id10007165	C	C	C	T	C	T	C	C	C	C
id10007166	C	C	C	T	C	T	C	C	C	C
id1000948	T	T	C	T	T	T	C	T	T	T
id1000955	G	G	A	G	G	G	A	G	G	G
id1027576	T	T	T	C	C	T	T	T	C	T
id11009990	T	T	T	C	T	C	C	T	T	T
id12003239	T	T	C	T	T	T	C	T	C	C
id12008144	C	C	C	C	C	C	C	C	C	C
id2010166	A	A	G	A	A	A	G	A	G	G
id2010412	A	A	A	A	A	G	A	G	A	A
id2010974	A	A	G	A	A	A	G	A	G	G
id2010992	C	C	T	C	C	C	C	C	C	T
id6005779	C	C	T	T	C	T	T	T	T	T
id6012242	G	G	G	G	G	T	G	T	G	G
id8000978	A	C	A	C	C	A	C	A	A	A
id8002106	G	G	G	G	G	G	T	G	G	T
id9002494	A	G	A	G	G	A	A	A	A	G
wd7002626	G	A	G	A	A	A	G	A	G	G
id6006625	G	G	A	G	G	G	A	A	A	A
id8000987	C	T	C	C	C	C	C	C	C	T
id11002184	G	G	G	G	G	G	G	A	G	A
id2010969	T	T	T	T	T	T	T	T	T	T
id4005078	T	C	C	C	C	C	C	C	T	C
id7000276	C	A	A	A	A	A	A	A	C	C
wd9002354	C	T	C	T	C	T	C	C	C	C
id1027513	A	A	A	C	C	A	A	A	C	A
id11001787	C	C	C	C	C	C	C	C	C	C
id11002213	N	G	N	G	G	G	G	A	N	A
id12008149	G	G	G	G	G	G	G	G	G	G
id3004807	A	A	A	A	A	A	G	A	G	G
id6003492	G	A	G	A	G	A	G	A	G	G
id7002784-4	G	A	G	A	A	A	G	A	G	G
id10007137	C	C	C	A	C	A	C	C	C	C
id4011112	T	T	C	T	C	T	C	T	C	T

Variety Marker	越光	TNG74	TT32	TYW2	HL19	TY4	TNGSW 21	HL20	KH144
dd1001675	G	A	A	G	G	G	G	A	A
id11009466	G	G	A	A	A	G	G	A	A
id12008090	T	C	C	C	C	T	C	C	C
id2010155	G	G	G	G	A	G	G	G	G
id4005236	G	G	A	G	G	G	G	G	G
id4011022	C	T	T	C	T	T	C	T	T
id4011023	A	G	G	A	G	G	A	G	G
id7000308	A	A	A	A	G	A	G	G	G
id7002851	T	G	G	G	T	T	N	G	N
id9002505	G	G	G	G	G	G	N	G	G
id10000129	A	A	G	G	G	G	G	G	G
id10000170	C	C	A	A	A	A	A	A	A
id10001501	G	G	G	G	G	G	G	G	A
id10007165	T	C	C	C	C	T	C	C	T
id10007166	T	C	C	C	C	T	C	C	T
id1000948	C	T	T	T	C	T	C	T	C
id1000955	A	G	G	G	A	G	A	G	A
id1027576	C	T	T	T	T	T	C	T	T
id11009990	T	C	T	T	T	N	T	T	C
id12003239	T	T	T	T	C	T	T	C	C
id12008144	T	C	C	C	C	T	C	C	C
id2010166	G	A	A	A	A	G	G	A	A
id2010412	G	A	A	A	G	G	G	A	A
id2010974	A	A	A	A	G	A	A	A	A
id2010992	C	C	C	C	T	C	C	C	C
id6005779	T	T	T	T	T	T	T	T	T
id6012242	T	G	G	G	T	T	G	G	G
id8000978	A	C	A	C	A	C	A	C	A
id8002106	G	G	G	G	G	G	N	G	G
id9002494	A	A	A	A	A	A	A	A	A
wd7002626	A	A	A	A	A	A	A	A	A
id6006625	G	G	A	G	G	G	A	G	A
id8000987	C	T	C	T	C	T	T	C	C
id11002184	G	G	G	G	G	G	G	G	G
id2010969	T	T	T	T	G	T	G	T	T
id4005078	C	C	T	C	C	C	C	C	C
id7000276	A	A	A	A	C	A	C	A	C
wd9002354	T	C	C	N	C	T	N	T	C
id1027513	C	A	A	A	A	A	C	A	A
id11001787	C	C	C	T	T	T	C	C	C
id11002213	G	N	G	G	G	G	N	G	G
id12008149	A	G	G	G	G	A	G	G	G
id3004807	A	A	A	A	G	G	G	A	A
id6003492	G	G	G	A	G	A	G	A	A
id7002784-4	A	A	A	A	A	A	G	A	G
id10007137	A	C	C	C	A	A	C	C	A
id4011112	T	T	T	C	T	T	T	T	T