

國立臺灣大學生物資源暨農學院農藝學系



博士論文

Department of Agronomy

College of Bioresources and Agriculture

National Taiwan University

Doctoral Dissertation

不結球白菜雄不稔性之遺傳分析、基因精細定位  
及基因體關聯性之研究

Genetic Analysis, Gene Fine-mapping, and Genome-wide  
Association Studies on Male Sterility  
in *Brassica rapa* ssp. *chinensis*

林子凱

Tzu-Kai Lin


指導教授：林順福 博士

Advisor: Shun-Fu Lin, Ph.D.

中華民國 109 年 7 月

July, 2020

## 誌 謝



進入農業試驗所埋首於作物育種已逾二十年，雖勉力精進，卻仍常感不足，承蒙所屬長官與師長的厚愛，讓子凱得以在職身份，重新回到校園繼續鑽研相關研究的學理和技術。從修習課業到完成博士論文的過程中，首要感謝恩師林順福老師的指導，追隨林老師門下多年，得老師對受業生的鞭策與鼓勵，除課業外，更有待人處世上的諸多提醒，令受業生銘誌於心，時時勿忘恩師的諄諄引導，並期勉自己能為台灣農業貢獻所學。修課過程中，感謝研究室智益、禎祐、銘錦、紹芬、文雅、亞平、鍾瑩、泰佑、尚誼、婷玫、淨惠等之相互切磋砥礪，讓我回學校不僅止於修課，更能重新體驗學生生活之青春氛圍。

修業期間感謝博士班資格考委員胡凱康老師強化育種理論、彭雲明老師建議試驗設計方法、王裕文老師啟發作物管理精義、林彥蓉老師闡釋基因體學研究方式、陳凱儀老師於基因分型技術之指導等，使得研究方向及論文內容架構可以更趨完備。此外，還要感謝中興大學農藝學系曾富生教授、台灣大學園藝暨景觀學系羅筱鳳教授、農委會科技處王任賢處長、農業試驗所林學詩所長、台南農改場義竹工作站謝明憲站長、台灣大學農藝學系黃永芬老師擔任口試委員，對論文初稿悉心斧正，並給予珍貴建議，讓研究論文得以順利完成。

研究過程感謝士佳、詠如、家郁、思好等人協助進行實驗，以及蔬菜研究室同仁的幫忙，投稿過程感謝林彥蓉老師對於寫作的建議及亞平協助圖表的修飾，以順利在 SCI 期刊刊登。另外感謝家人在我進修期間的支持，感謝愛妻俐文一路陪伴並分擔家中瑣事，小兒佑謙、新岳帶來的天倫歡樂，都是我身後無形的力量。最後，我要由衷感謝父母親，從小悉心栽培與無條件支持讓我順利完成學業，我謹將這份小小的榮譽獻給家人。

謹以此論文分享給大家，願大家共享我的喜悅。

## 中文摘要



白菜類作物是重要的蔬菜來源之一，為了應用雄不稔性生產具雜種優勢的商業  $F_1$  種子，急需豐富雄不稔基因的遺傳資源。本研究在小油菜 (*Brassica rapa* ssp. *chinensis* var. *oleifera* Makino) 地方品種中發現具雄不稔性的自發性突變體，利用後裔檢定進行系統性的遺傳分析，結果顯示該性狀符合單一基因座上三個複等位基因控制的遺傳模式，將該基因命名為 *BrMs*，而此三個複等位基因間的顯隱性關係為  $BrMs^a > BrMs^b > BrMs^c$ ，其中控制雄不稔性表現者為  $BrMs^b$ ，此遺傳模式為不結球白菜上首見，透過 genotyping-by-sequencing (GBS) 策略針對不同雄不稔等位基因建構之  $F_2$  及  $BC_1$  分離族群進行基因定位，不同雄不稔等位基因均被定位於 Chr07 染色體之相同區域而驗證遺傳分析的結論。

依據 GBS 粗定位之區間標誌 (bin markers) 設計為單一核苷酸多型性 (single nucleotide polymorphism, SNP) 分子標誌，並以包含 1,590 株的  $F_2$  分離族群進行驗證，其中有二個 SNP 分子標誌分別位於雄不稔基因的二側，與基因之遺傳距離各均為 0.3 cM 而將 *BrMs* 限縮於約 400 kb 範圍內，為比較不同雄不稔基因間之差異，結球白菜雄不稔基因 *Ms* 的序列特徵化增幅區域 (sequence characterized amplified region, SCAR) 分子標誌也成功的轉換為 SNP 分子標誌，並分析相同  $F_2$  族群，連鎖分析顯示 *Ms* 基因雖亦屬 Chr07 連鎖群但與本研究之 *BrMs* 基因位於不同區域，然而透過基因體比較分析發現大油菜雄不稔基因 *BnRf* 位於本研究 *BrMs* 基因粗定位範圍內，進一步依據 *BnRf* 之基因序列，分析本研究雄不稔系與恢復系其 *BnRf* 同源基因之差異而開發了一個 Indel 分子標誌，以該分子標誌偵測前述  $F_2$  族群之 923 個植株，結果顯示該分子標誌基因分型的結果與外表型完全一致。

為了進行 *BrMs* 基因的精細定位，將雄不稔系 NH80-A 與恢復系 TA95 進行全基因體定序，針對 *BrMs* 粗定位區間比較二品系 DNA 序列之差異，並開發了

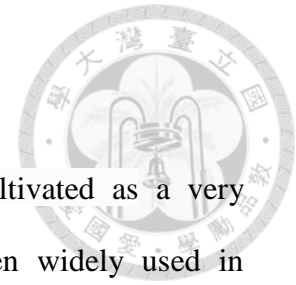
9 個 SNP 分子標誌用以分析 4,415 株 F<sub>2</sub> 族群，基因精細定位的結果顯示 BrA7\_6746K 與 BrA7\_6749K 等二個 SNP 分子標誌與雄不稔外表型共分離，然而依據 *BnRf* 而設計的 Indel 分子標誌進行基因分型時，發現 2 個體其外表型與分子標誌的基因型不一致，推測該個體可能為複合基因內重組的結果，透過基因精細定位將 *BrMs* 限縮於 24 kb 範圍內，該區域僅包含 *Bra014987*、*Bra014988* 及 *Bra014989* 等 3 個預測基因 (putative genes)，可做為據圖選殖 (map-based cloning) 之參考。

正常稔性個體其 *BrMs* 基因型大都不具 *BrMs<sup>b</sup>* 等位基因，為了開發可鑑別 *BrMs<sup>a</sup>* 與 *BrMs<sup>c</sup>* 等位基因之分子標誌，利用 140 個 *B. rapa* 自交系與 NH80-A 進行測交而分析其 *BrMs* 基因型，並針對整個族群進行全基因體關連性分析 (genome-wide association study, GWAS)，結果於 Chr07 染色體上探勘出 17 個 SNPs 與 *BrMs* 基因具關聯性，挑選 4 個設計為 SNP 分子標誌，其中 TqBrA7\_6750K 基因分型的結果顯示其可有效鑑別 *BrMs<sup>a</sup>* 與 *BrMs<sup>c</sup>* 等位基因。

本研究解析本土小油菜細胞核雄不稔性之遺傳機制，研究結果顯示其遺傳行為具商業應用潛力，並可擴大應用於白菜類作物，*BrMs* 基因定位之結果及開發的分子標誌，除可做為基因選殖之參考外，亦有助於透過分子育種加速選育白菜類作物之雄不稔品系。

**關鍵字：**不結球白菜、雄不稔性、遺傳分析、遺傳連鎖圖譜、基因定位、單一核苷酸多型性分子標誌、育種

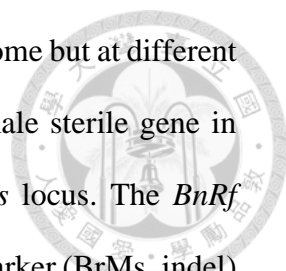
## Abstract



Consisting of many crop subspecies, *Brassica rapa* is cultivated as a very important source of vegetables. Despite male sterility has been widely used in commercial production of hybrid seeds in brassica crops, but not in *B. rapa* due to lack of useful genetic resources with male sterile genes. We discovered a spontaneous mutant from a local variety of rape (*B. rapa* ssp. *chinensis* var. *oleifera* Makino) in Taiwan.

Results from a series of progeny tests indicated that male sterility in *B. rapa* ssp. *chinensis* fits a one gene with multi-allelic model. The gene designated as *BrMs* is composed of three alleles with *BrMs<sup>a</sup>* dominant to *BrMs<sup>b</sup>*, and *BrMs<sup>b</sup>* dominant to *BrMs<sup>c</sup>* (*BrMs<sup>a</sup>* > *BrMs<sup>b</sup>* > *BrMs<sup>c</sup>*). Among three alleles, *BrMs<sup>b</sup>* conferring male sterility and the genotype with *BrMs<sup>b</sup>Ms<sup>b</sup>* and *BrMs<sup>b</sup>Ms<sup>c</sup>* would be sterile. This is the first report of such genetic model in non-heading Chinese cabbage. The model was confirmed according to the results of gene mapping under the genotyping-by-sequencing (GBS) strategy, which was that the male sterility loci were mapped at the same region of chromosome 07 using BC<sub>1</sub> and F<sub>2</sub> populations involved different *BrMs* alleles.

SNP markers were designed based on GBS bin markers co-segregating with *BrMs* trait to evaluate genetic distance between *BrMs* and markers loci based on 1,590 F<sub>2</sub> plants derived from male sterile (NH80-A) and restorer line (TA95). Two flanking SNP markers, BrA7\_6632K and BrA7\_7625K, were only 0.3 cM away from the upstream and downstream of *BrMs* locus respectively, narrowing the male sterility gene down to a 400 kb region. A SNP marker, syau\_scr04\_SNP, developed from sequence characterized amplified region (SCAR) marker closely linked to *Ms* locus in Chinese cabbage was used to investigate the same F<sub>2</sub> population. The result of linkage analysis



suggested that *BrMs* and *Ms* were located on the same A07 chromosome but at different loci. Interestingly, comparative genomic analysis showed that a male sterile gene in *Brassica napus*, *BnRf*, was within the 400 kb region of the *BrMs* locus. The *BnRf* orthologs of *B. rapa* ssp. *chinensis* were sequenced, and one indel marker (*BrMs\_indel*) was developed for genotyping the same F<sub>2</sub> population. Male sterility phenotypes of all 923 investigated plants were completely co-segregated with the marker genotypes.

To perform gene fine-mapping for the *BrMs*, the whole genome sequencing of the male sterile line NH80-A and restorer line TA95 was accomplished. Within the interval detected from primary *BrMs* gene mapping, the variations of DNA sequences between two lines were developed into 9 SNP markers. These SNP markers and *BrMs\_indel* marker were investigated in 4,415 plants of an F<sub>2</sub> population derived from NH80-A × TA95. The genotyping results showed that two SNP markers, *BrA7\_6746K* and *BrA7\_6749K*, were co-segregated with male sterility. Two out of 4,415 plants showed no co-segregation between male sterility phenotypes and *BrMs\_indel* genotypes, and it might be resulted from the intragenic recombination within the *BrMs* gene complex. In addition, the results of fine mapping had located *BrMs* to a 24 kb physical region. This area contains only 3 putative genes, *Bra014987*, *Bra014988* and *Bra014989*, providing solid information for map-based cloning.

Despite *BrMs<sup>b</sup>* conferring male sterility in *B. rapa*, most male fertile varieties retain *BrMs<sup>a</sup>* or *BrMs<sup>c</sup>* but not *BrMs<sup>b</sup>* allele in the locus. In order to develop molecular markers to distinguish *BrMs<sup>a</sup>* and *BrMs<sup>c</sup>* alleles, 140 *B. rapa* inbred lines were crossed with NH80-A to test their *BrMs* genotype, and followed by performing genome-wide association studies (GWAS). Seventeen SNPs on the Chr07 chromosome associated with *BrMs* were found, and four of them were selected to develop SNP markers. Genotyping result of examined *B. rapa* accessions indicated that *TqBrA7\_6750K*

marker could be used to discriminate *BrMs<sup>a</sup>*, and *BrMs<sup>c</sup>* alleles effectively.

This study explored the genetic mechanism of the new genic male sterility mutation in non-heading Chinese cabbage. The mutated gene has commercial application potential and can be extended to the other *B. rapa* crops. The results of gene fine mapping and marker development, not only have provided a solid foundation for *BrMs* gene cloning, but also speeded up the development of male sterile lines of *B. rapa* crops through molecular breeding.

**Keywords:** non-heading Chinese cabbage, male sterility, genetic analysis, genetic map, gene mapping, SNP marker, plant breeding



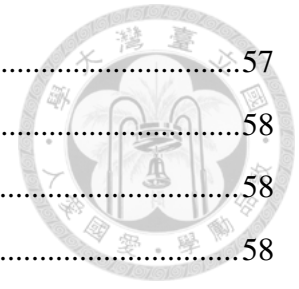




## 目錄

中文摘要 .....	i
Abstract .....	iii
<b>第 1 章 總論 .....</b>	<b>1</b>
1 前言 .....	2
2 前人研究 .....	3
2.1 小油菜之分類地位及其與芸苔屬物種間之關係.....	3
2.2 白菜類之起源.....	5
2.3 台灣小油菜之栽培史.....	6
2.4 雄不稔性的類型.....	6
2.5 十字花科作物細胞質雄不稔性的來源與利用 .....	7
2.6 細胞質雄不稔性應用上可能的限制.....	8
2.7 芸苔屬作物細胞核雄不稔性的來源與遺傳.....	10
2.7.1 隱性基因控制之細胞核雄不稔的遺傳與應用 .....	10
2.7.2 單一顯性基因控制之細胞核雄不稔的遺傳與應用 .....	12
2.7.3 細胞核內二對顯性基因所控制之雄不稔性的遺傳與應用 .....	14
2.7.4 細胞核內複等位基因所控制之雄不稔性的遺傳與應用 .....	15
2.7.5 細胞核內多個基因所控制之雄不稔性的遺傳與應用 .....	16
2.8 雄不稔性相關基因的分子標誌與基因定位.....	17
<b>第 2 章 不結球白菜雄不稔性的遺傳模式 .....</b>	<b>21</b>
1 前言 .....	22
2 材料與方法 .....	22
2.1 試驗材料.....	22
2.2 後裔檢定試驗之遺傳分析.....	24
2.2.1 雄不稔性遺傳模式之假設.....	24
2.2.2 試驗材料相互雜交建構後裔檢定分析族群 .....	25
2.2.3 後裔檢定試驗之規模及性狀調查 .....	26
2.2.4 利用後裔檢定試驗之結果推導雄不稔性之遺傳模式 .....	27

3 結果 .....	27
3.1 以 TA95 為恢復系之第一雜交群組後裔檢定試驗 .....	27
3.2 以 SG388 為恢復系之第二雜交群組後裔檢定試驗 .....	34
4 討論 .....	39
4.1 測交分析之方法 .....	39
4.2 大油菜與大白菜細胞核雄不稔遺傳之比較 .....	40
4.3 遺傳行為對爾後育種之差異 .....	40
<b>第 3 章 以雄不稔株 × 恢復系之 F<sub>2</sub> 族群建立遺傳連鎖圖定位 <i>BrMs</i> 基因 .....</b>	<b>41</b>
1 前言 .....	42
2 材料與方法 .....	43
2.1 建構基因定位分離族群 .....	43
2.2 GBS (genotyping-by-sequencing) 之進行 .....	43
2.2.1 建構 GBS library .....	43
2.2.2 序列分析 (raw sequence data processing) .....	44
2.3 連鎖分析 (linkage analysis) .....	44
3 結果 .....	44
3.1 GBS 的結果 .....	44
3.2 <i>BrMs</i> 之基因定位 (gene mapping) .....	45
4 討論 .....	53
4.1 透過 GBS 策略建構連鎖圖以進行基因定位 .....	53
4.2 本研究遺傳連鎖圖與其他不結球白菜連鎖圖之比較 .....	54
<b>第 4 章 以雄不稔株 × 維持系之 BC<sub>1</sub> 族群建立遺傳連鎖圖定位 <i>BrMs</i> 基因 .....</b>	<b>55</b>
1 前言 .....	56
2 材料與方法 .....	56
2.1 建構基因定位分離族群 .....	56
2.2 GBS 之進行 .....	56
2.2.1 建構 GBS library .....	56
2.2.2 序列分析 .....	57



2.2.3 連鎖分析 .....	57
3 結果 .....	58
3.1 GBS 的結果.....	58
3.2 <i>BrMs</i> 之基因定位.....	58
4 討論 .....	69
4.1 BC <sub>1</sub> 遺傳連鎖圖與 F <sub>2</sub> 連鎖圖之比較.....	69
4.2 基因定位結果驗證 <i>BrMs<sup>a</sup></i> 、 <i>BrMs<sup>b</sup></i> 及 <i>BrMs<sup>c</sup></i> 的等位性.....	69
<b>第 5 章 <i>BrMs</i> 基因之精細定位及開發其 SNP 分子標誌 .....</b>	<b>71</b>
1 前言 .....	72
2 材料與方法 .....	73
2.1 粗定位之 GBS bin markers 驗證.....	73
2.1.1 建構 <i>BrMs</i> 基因之分離族群 .....	73
2.1.2 將 GBS 及 SCARs markers 轉換為 SNP 分子標誌.....	73
2.1.3 參試植株基因分型及 <i>BrMs</i> 連鎖分析 .....	74
2.2 不結球白菜 <i>BrMs</i> 與大油菜 <i>BnRf</i> 之關係.....	74
2.2.1 不結球白菜 <i>BnRf</i> orthologs 間序列比較及引子設計 .....	74
2.2.2 <i>BnRf</i> orthologs 之 DNA 片段選殖與定序.....	75
2.2.3 <i>BrMs</i> 與 <i>BnRf</i> ortholog 之等位性分析 .....	76
2.3 <i>BrMs</i> 基因之精細定位.....	76
2.3.1 基因體定序及序列分析 .....	76
2.3.2 開發精細定位用 SNP 分子標誌.....	77
2.3.3 精細定位用分離族群之建構與基因分型 .....	77
3 結果與討論 .....	78
3.1 粗定位之 GBS markers 的驗證.....	78
3.1.1 將 GBS markers 轉換為 SNP 分子標誌 .....	78
3.1.2 結球白菜 <i>Ms</i> 基因之 SCAR 分子標誌轉換為 SNP 分子標誌 .....	78
3.1.3 SNP 分子標誌與 <i>BrMs</i> 基因之連鎖分析結果 .....	82
3.1.4 不結球白菜雄不稔 <i>BrMs</i> 基因與結球白菜 <i>Ms</i> 基因間可能之關係 .....	83
3.1.5 SNP 分子標誌開發策略及應用於 MAS 之潛力 .....	83

3.1.6 近緣種大油菜之 <i>BnRf</i> 基因位於不結球白菜 <i>BrMs</i> 之定位區間 .....	84
3.2 不結球白菜 <i>BrMs</i> 與大油菜 <i>BnRf</i> 之關係.....	84
3.2.1 不結球白菜 <i>BnRf</i> orthologs 間序列之選殖與比較.....	84
3.2.2 不結球白菜 <i>BrMs</i> 與油菜 <i>BnRf</i> 之等位性 .....	87
3.3 <i>BrMs</i> 基因之精細定位.....	87
3.3.1 不結球白菜基因體定序與組裝結果.....	87
3.3.2 開發精細定位用 SNP 分子標誌.....	89
3.3.3 精細定位 <i>BrMs</i> 基因 .....	91
<b>第 6 章 利用全基因體關連分析 (GWAS) 開發鑑別 <i>BrMs<sup>a</sup>/BrMs<sup>c</sup></i> 等位基因之</b>	
<b>SNP 分子標誌.....</b>	<b>95</b>
1 前言 .....	96
2 材料與方法 .....	97
2.1 試驗材料.....	97
2.2 GBS 之進行.....	98
2.2.1 建構 GBS library.....	98
2.2.2 序列分析 .....	99
2.3 全基因體關聯性分析 (GWAS) .....	100
2.4 SNP 分子標誌之設計與驗證 .....	100
3 結果與討論 .....	101
3.1 次世代定序所獲得之讀序.....	101
3.2 透過 GBS 探勘獲得之 SNPs.....	103
3.3 GWAS 分析之族群組成.....	107
3.4 GWAS 分析的結果.....	113
3.5 SNP 分子標誌的驗證.....	116
<b>第 7 章 總結 .....</b>	<b>125</b>
<b>參考文獻 .....</b>	<b>131</b>



附 錄 .....	141
附錄 1 兩種遺傳模式下不結球白菜參試品系可能的基因型 .....	141
附錄 2 Progeny test A 之各種雜交組合後裔雄不稔性狀分離情形.....	142
附錄 3 Progeny test B 之各種雜交組合後裔雄不稔性狀分離情形.....	143
附錄 4 Progeny test C 之各種雜交組合後裔雄不稔性狀分離情形.....	144
附錄 5 本研究部份內容所發表之報告 .....	145



## 圖目錄

圖 1-1	蕓苔屬物種間之關係.....	3
圖 1-2	利用單一隱性基因控制之雄不稔系生產雜交 F <sub>1</sub> 種子的模式.....	11
圖 1-3	利用單一顯性基因控制之雄不稔系生產雜交 F <sub>1</sub> 種子的模式.....	12
圖 1-4	利用對環境敏感的單一顯性基因控制之雄不稔系生產雜交 F <sub>1</sub> 種子的模式.....	13
圖 1-5	利用二對顯性基因控制之雄不稔系生產雜交 F <sub>1</sub> 種子的模式.....	15
圖 1-6	利用複等位基因控制之雄不稔系生產雜交 F <sub>1</sub> 種子的模式.....	16
圖 2-1	不結球白菜 NH80 花器形態及花粉形成之差異.....	23
圖 2-2	Progeny test A、Progeny test B 及 Progeny test C 等三種後裔檢定試驗.....	25
圖 2-3	第一雜交群組於三種後裔檢定之結果.....	33
圖 2-4	第二雜交群組於三種後裔檢定之結果.....	38
圖 3-1	以不結球白菜 NH80-A × TA95 之 F <sub>2</sub> 族群建立的高密度遺傳連鎖圖.....	47
圖 3-2	以不結球白菜 NH80-A × TA95 之 F <sub>2</sub> 族群建立的各染色體遺傳連鎖圖.....	52
圖 4-1	以不結球白菜 NH80-A × WH606 之 BC <sub>1</sub> 族群建立的高密度遺傳連鎖圖.....	61
圖 4-2	以不結球白菜 NH80-A × WH606 之 BC <sub>1</sub> 族群建立的各染色體遺傳連鎖圖.....	66
圖 4-3	不結球白菜雄不稔基因 <i>BrMs</i> 之定位區間.....	68
圖 5-1	將 GBS markers 轉換為 SNP 分子標誌後之基因分型情形.....	79
圖 5-2	syau_scr04 引子擴增產物於 NH80-A 與 TA95 間之多型性位點.....	80
圖 5-3	以 syau_scr04_SNP 進行基因分型之情形.....	80
圖 5-4	源自 GBS tags 及 SCARs 的 SNP 分子標誌所建構之遺傳連鎖圖.....	82
圖 5-5	針對不結球白菜 <i>BnRf</i> orthologs 不同區域進行 PCR 之結果.....	85
圖 5-6	不結球白菜 <i>BnRf</i> orthologs 與大油菜 <i>BnRf</i> 基因結構之比較.....	86
圖 5-7	<i>BnRf</i> 與 <i>BnRf</i> orthologs 於第 4 個 intron 之序列差異.....	86
圖 5-8	SNP 分子標誌序列比對至 <i>B. rapa</i> 及 <i>B. napus</i> 基因體的物理位置.....	90
圖 6-1	正反讀序重疊之 paired-end reads 合併成 single-end reads 的長度分布.....	101
圖 6-2	參試族群 SNP 位點之次等位基因頻度分布情形.....	104
圖 6-3	<i>B. rapa</i> 族群全基因體 SNPs 的 LD decay 情形.....	104

圖 6-4	<i>B. rapa</i> 族群各染色體 SNPs 的 LD decay 情形.....	106
圖 6-5	各 <i>B. rapa</i> 品系之異質結合率分布情形.....	112
圖 6-6	參試品系經主成分分析後之分群情形.....	112
圖 6-7	未控制族群結構並以廣義線性模型進行 GWAS 分析之結果.....	113
圖 6-8	採 Q+K 模型控制族群結構並以 CMLM 進行 GWAS 分析之結果.....	113
圖 6-9	經 GWAS 分析顯示 Chr 07 上具有與 <i>BrMs</i> 相關之 SNPs.....	114
圖 6-10	透過 GWAS 開發的 SNP 分子標誌基因分型效果.....	118
圖 7-1	利用雄不稔基因 <i>BrMs</i> 生產 F <sub>1</sub> 雜交種子的採種模式.....	127
圖 7-2	當自交系基因型為 <i>BrMs<sup>a</sup>Ms<sup>a</sup></i> 時的雄不稔系培育方法.....	129
圖 7-3	當自交系基因型為 <i>BrMs<sup>c</sup>Ms<sup>c</sup></i> 時的雄不稔系培育方法.....	130



## 表目錄

表 2-1	參試不結球白菜之品系來源與主要用途.....	24
表 2-2	參試不結球白菜品系之學名.....	24
表 2-3	第一雜交群組後裔之稔性表現.....	27
表 2-4	不結球白菜 NH80-A×TA95 之 F <sub>2</sub> 族群於 progeny test A 的稔性分離情形.....	29
表 2-5	不結球白菜 NH80-A×TA95 之雜交後裔於 progeny test B 的稔性表現.....	31
表 2-6	不結球白菜 WH606×TA95 之 F <sub>2</sub> 族群於 progeny test C 的稔性表現.....	32
表 2-7	第二雜交群組後裔之稔性表現.....	34
表 2-8	不結球白菜 NH80-A×SG388 之 F <sub>2</sub> 族群於 progeny test A 的稔性分離情形.....	35
表 2-9	不結球白菜 NH80-A×SG388 之雜交後裔於 progeny test B 的稔性表現.....	37
表 2-10	不結球白菜 WH606×SG388 之 F <sub>2</sub> 族群於 progeny test C 的稔性表現.....	37
表 3-1	以 F <sub>2</sub> 族群建立之連鎖群長度、SNP 數量及 marker 密度.....	46
表 3-2	與 <i>BrMs</i> 基因共分離的 GBS tags.....	53
表 4-1	以 BC <sub>1</sub> 族群建立之連鎖群長度、SNP 數量及 marker 密度.....	60
表 4-2	與 <i>BrMs</i> 基因共分離的 GBS tags.....	67
表 5-1	選殖不結球白菜 <i>BnRf</i> orthologs 時所使用之引子對序列.....	75
表 5-2	以 KASP 原理設計的分子標誌引子序列.....	81
表 5-3	不結球白菜 NH80-A 及 TA95 經次世代定序所獲得讀序摘要.....	88
表 5-4	不結球白菜 NH80-A 及 TA95 讀序比對至參考序列的結果.....	89
表 5-5	針對 <i>BrMs</i> 進行基因精細定位之結果.....	92
表 6-1	合成 Y-adapter 的 DNA 序列.....	98
表 6-2	參試族群 SNPs 分布及 LD decay 情形.....	103
表 6-3	進行 GWAS 分析的品系資訊.....	107
表 6-4	與 <i>BrMs</i> 相關 SNPs 的物理位置.....	115
表 6-5	本研究開發之分子標誌對 <i>B. rapa</i> 品系進行基因分型的結果.....	119





## 第1章 總論



## 1 前言

在蕓苔屬 (*Brassica*) 中，白菜類作物 (*Brassica rapa*) 之分化明顯，大致上可分為結球白菜 (*B. rapa* ssp. *pekinensis*) 與不結球白菜 (*B. rapa* ssp. *chinensis*) 二個亞種，該二者之間無生殖隔離，其中不結球白菜為因應不同生態環境、用途、食用文化與栽培習慣等，已分化出許多不同的生態型，台灣常見的不結球白菜包含小白菜、青梗白菜及油菜 (小油菜) 等。台灣結球白菜之年平均栽培面積約為 2,100 公頃，而不結球白菜則為 5,000 公頃，均為我國重要蔬菜作物品項，而屬不結球白菜之小油菜除了食用之外，少數地方品種則改做為綠肥作物，年平均栽培面積約為 7,500 公頃。

由於蕓苔屬作物大都為異花授粉，並且具有明顯的雜種優勢，在商業生產上利於開發一代雜交 ( $F_1$ ) 品種，亦為商業育種主流，目前生產  $F_1$  種子主要是利用自交不親和性及雄不稔性，利用自交不親和性採種具有雙親皆可留種之優點，然而存在下列兩個瓶頸，包括 (一)  $F_1$  種子雜交率不穩定：自交不親合性容易受環境因子影響，而減弱或喪失自交不親和特性，從而使雜交率降低，不但影響商業生產  $F_1$  之雜交純度，更重要的是非雜交種子的發生率增加，意謂著親本流失的風險也提高；(二) 親本用自交系繁殖成本高：進行  $F_1$  採種時的自交系最好具有強自交不親合性，但採種前又需打破自交不親合性以繁殖足量親本，親本繁殖普遍採用的方法為利用人工進行蕾期自交授粉，導致耗費大量勞力與採種成本。

植物雄不稔性之應用為  $F_1$  種子生產提供了另一個途徑，所謂雄不稔性係指無法產生有功能的花藥、花粉或雄配子，可商業化應用的雄不稔源，除了不易受環境因子影響外，母本繁殖時也不需人工蕾期自交授粉，可望突破前述自交不親合性採種之瓶頸，目前甘藍、花椰菜及青花菜等，均已普遍利用雄不稔性進行一代雜交 ( $F_1$ ) 品種之種子生產，然而白菜類蔬菜則較少，國內目前則尚無利用雄不稔性進行採種的白菜品種，主要原因之一為白菜類蔬菜中較缺乏具商業化利用價值的雄不稔源。

綠肥用的小油菜於每年水稻秋收後即大面積地栽培，除了提供肥份養地之外，盛花期形成之花海更成為賞心悅目的農村景觀，並帶動休閒觀光農業之發展，由於

種植面積大、栽培較為粗放，在台灣行之多年且鮮少進行人為選拔，因此保有相當之遺傳變異性與良好的環境適應性，是一個研究族群遺傳及發掘重要性狀的良好素材，而依據 Kaul (1988) 所述植物普遍存在雄不稔性之說法，暗示在足夠大的族群內應有機會發現雄不稔突變株，基此思維而於冬季的油菜花田間，逐株探索優良特性，期望加以開發且利用於育種工作，結果於 2004 年的調查過程中，在台中市霧峰區的油菜田區中發現了自發性的雄不稔植株，若能針對此小油菜雄不稔源深入研究，透過育種擴大其應用範圍，將成為白菜類作物利用雄不稔性採種的契機。

本研究過程首先從遺傳分析的面向切入，透過該雄不稔基因的遺傳行為評估其應用性，而運用基因定位技術驗證遺傳分析之結論外，並可透過相關基因體資訊以剖析此新雄不稔源與其他不稔源間之差異，進而開發分子標誌並將該性狀導入分子育種的範疇。

## 2 前人研究

### 2.1 小油菜之分類地位及其與蕓苔屬物種間之關係

蕓苔屬作物種類繁多，為了避免分類上的困擾，現今許多學者仍然延用 1935 年 Nagaharu 所提出之蕓苔屬作物間關係圖 (圖 1-1，部分文獻以 *B. campestris* 取代 *B. rapa*)。

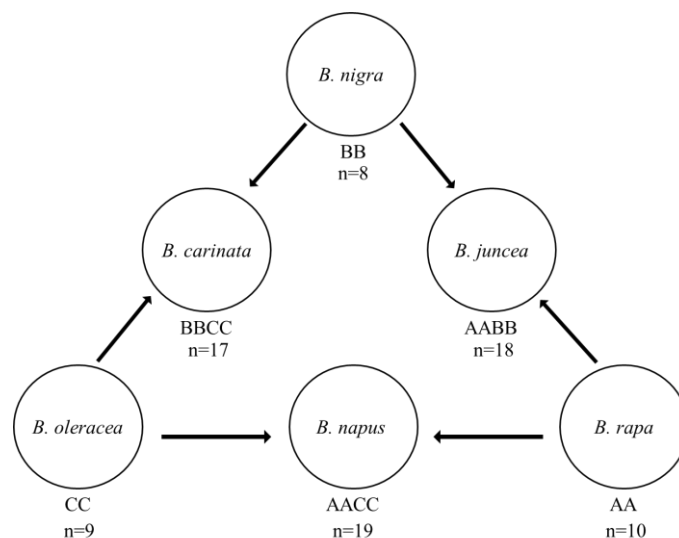




圖 1-1 蕓苔屬物種間之關係  
(修改自 Nagaharu, 1935)



其中白菜類 *B. rapa* (AA)、黑芥 *B. nigra* (BB) 及甘藍類 *B. oleracea* (CC) 被認為是其原始種，而其他的複二倍體則可能是由這些原始種天然雜交演化而來，其中大油菜 *B. napus* (AACC,  $2n = 38$ ) 主要為榨油用，芥菜 *B. juncea* (AABB,  $2n = 36$ ) 主要為蔬菜用或榨油用，至於阿比西尼亞芥 *B. carinata* (BBCC,  $2n = 34$ ) 則可為飼料用及油料用，雖然阿比西尼亞芥較不普遍，然近年在十字花科作物研究上常利用轉移位於 B 基因組的優良性狀，因而成為重要的近緣種種質資源 (Navabi et al., 2010)。二倍體物種中黑芥之種原較少，而甘藍類及白菜類則主要為蔬菜用。

在蕓苔屬作物中，白菜分化明顯，其下之物種學名較複雜，物種起源亦有不同的說法，1753 年瑞典植物學家林奈首先將二個白菜類物種蕓菁及蕓苔分別定名 *B. rapa* L. 及 *B. campestris* L.，其後於 1755 年亦將中國南方的不結球白菜定名為 *B. chinensis* L.，至 1833 年 Johann Metzger 將 *B. rapa* L. 及 *B. campestris* L. 合併為一個種，並採用 *B. rapa* 做為種名。由於白菜類作物種類繁多，因此如結球白菜（大白菜）、不結球白菜、菜薹、白菜型油菜及蕓菁等常被命名為不同之種名。結球白菜屬於蕓苔的亞種之一，分類上較無爭議，而目前較普遍使用的學名為 *B. rapa* ssp. *pekinensis* (Lour.)，然而不結球白菜在植株外觀上分化出許多不同的類型，本研究之材料小油菜即是屬於不結球白菜的一種類型。曹與李 (1982) 認為不結球白菜在分類上也屬於蕓苔的亞種之一，因此認定不結球白菜學名應為 *B. campestris* ssp. *chinensis*，並依據型態特徵之差異在該亞種下區分為普通白菜 (*B. campestris* ssp. *chinensis* var. *communis* Tsen et Lee)、烏塌菜 (*B. campestris* ssp. *chinensis* var. *rosularism* Tsen et Lee)、菜薹 (*B. campestris* ssp. *chinensis* var. *parachinensis* Bailey)、薹菜 (*B. campestris* ssp. *chinensis* var. *tai-tsai* Hort.)、多頭菜 (*B. campestris* ssp. *chinensis* var. *multiceps*) 及油菜 (*B. campestris* ssp. *chinensis* var. *oleifera* Makino) 等 6 個變種，台灣俗稱之小油菜即為 *B. campestris* ssp. *chinensis* var. *oleifera* Makino，然近年於白菜類作物的研究上，學者大多採用 *B. rapa* 取代 *B. campestris*，因此本研究中，小油菜以 *B. rapa* ssp. *chinensis* var. *oleifera* Makino 表示，而其他不結球白菜則以 *B. rapa* ssp. *chinensis* 表示。


## 2.2 白菜類之起源



關於小油菜於中國之起源，Li (1981)則認為小油菜是由小白菜(*B. campestris* L. ssp. *chinensis* Makino)發生變異而產生的變種之一，劉(1984)則認為中國的小油菜為因應不同地區環境種植而發展出不同生態型，原產於中國北部的北方小油菜應命名為*B. campestris* L.，北方小油菜發源於中國的中部或西北部，其中陝西可能為其起源地，由於*B. campestris* L.是世界上廣泛分佈的蕓苔屬植物，因此認為中國西北部僅為原產地之一，至於原產於中國南部的南方油白菜為*B. chinensis*，該類型為中國特有的且分化顯著，推測是由中國的蕓苔原始種(*B. campestris* L.)演化而得，南方油白菜發源於中國的中部或南部之長江流域，包括雲南、貴州、湖北及四川等可能為起源地，且依據RAPD (random amplified polymorphic DNA)分子標誌建構之系統發育樹顯示北方小油菜的起源早於南方油白菜(何等，2003)。

至於結球白菜起源則有二個假說，其一為雜交起源說(Li, 1981)，即結球白菜是由小白菜(ssp. *chinensis*)與蕓菁(ssp. *oleifera*)天然雜交後，經結球性的篩選與環境適應性的演化而得，另一則為分化起源說(譚，1979)，其認為結球白菜是由南方的小白菜向北方傳播時逐漸分化而形成的，二個假說均認為結球白菜中先有不結球的散葉類型，然後有半結球品種，最後出現結球充實的品種。曹與曹(1995)之研究指出，大白菜的發生可能是由小白菜分化至一定階段後，與蕓菁或是近緣種雜交後於中國北方之生態條件下演化而來的，郭等(2002)之研究結果亦說明大白菜可能起源於高度分化之小白菜，且暗示起源於不結球白菜中之臺菜(*B. campestris* ssp. *chinensis* var. *tai-tsai* Hort.)的可能性，雖然結球白菜的起源有不同的說法，但共同點均認為結球白菜之分化晚於不結球白菜，且由不結球白菜演化而得。至於另一個近緣物種大油菜，其原產於歐洲，一般認為它的起源可能是多源發生的，即在不同的地區由甘藍(*B. oleracea*)與蕓苔(*B. campestris*)於自然環境下因種間雜交並形成複二倍體後演化而得，而中國廣泛栽培的大油菜最早是於50年代由日本及英國引進栽培的(劉，1984)。

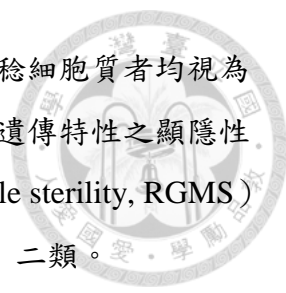
### 2.3 台灣小油菜之栽培史



前述萵苣屬的六個物種中，除了三個複二倍體物種可做為油料作物之外，白菜類中的小油菜 (*B. campestris* ssp. *chinensis* var. *oleifera* Makino; 或 *B. rapa* ssp. *chinensis* var. *oleifera* Makino) 是十字花科作物中唯一可做為油料用途的二倍體物種；在日據時代，台灣栽培在來種小油菜用以榨油，當時栽培面積約 1000 公頃(林，2006)，該類型小油菜具有早熟、抗病性佳及適應性廣等優點，但由於種子小致產量受限，且含油率低，因此於 1950 年代由日本引進之大油菜榨油用品種如‘近畿 32 號’及‘農林 20 號’等品種所取代(邱與彭，1995)，而 1960 年代台灣進行油菜育種工作後，分別由新竹及台中區農業改良場選育出數個大油菜品種，而在來種小油菜則改作為綠肥及蜜源作物(鄭等，1963)。經改良與推廣後，60 年代台灣大油菜栽培面積達 20,000 公頃為栽培生產之高峰，然而萃取的菜籽油因芥酸含量高不適用於人體或動物食用，致使 90 年代後台灣已鮮少栽培大油菜。在來種小油菜雖已超過半個世紀不以榨油為主要用途，但時至今日仍普遍被農戶當作冬裏作之綠肥用途，並為台灣冬季主要之蜜源作物，由於沒有經過太嚴格的人為篩選，族群內仍然保有一定之遺傳多樣性，且對台灣的生態環境早已演化具良好的適應性，因此極有可能在台灣本地小油菜中發掘重要性狀之遺傳資源，進而利用於白菜類作物或其近緣種作物改良。台灣雖已少栽培大油菜，但其仍為世界油菜栽培之主要類型，為了利用雜種優勢，國際上目前大都以培育雜交種為主，因此在進行大油菜育種時雄不稔性的利用非常廣泛。在台灣，白菜類作物主要做為蔬菜用，商業品種也大都改良為一代雜交種，而這些商業品種目前大都以自交不親合性生產 F<sub>1</sub> 種子，至於綠肥用小油菜則仍維持固定種，並以開放授粉(open pollination)的方式繁衍種子。

### 2.4 雄不稔性的類型

植物的雄不稔性依其遺傳行為可分為細胞質型雄不稔(cytoplasmic male sterility, CMS)及細胞核型雄不稔(genic male sterility, GMS)(Kaul, 1988)，其中細胞質型雄不稔，因常有細胞核基因與細胞質基因之相互作用，而又稱之為核質互作型雄不稔(gene-cytoplasmic male sterility)，有許多 CMS 例子因為是透過遠緣雜交或細胞融合方式，再輔以適當之回交育種程序而將一物種的雄不稔細胞質轉移至另一物種，致使接受親一方僅具雄不稔細胞質但不具有位於細胞核上之恢




復基因，然不論細胞核基因是否參與稔性表現，一般將具有雄不稔細胞質者均視為細胞質型雄不稔。至於在細胞核型雄不稔中，依據控制雄不稔遺傳特性之顯隱性關係，又可區分為受隱性基因控制的雄不稔（recessive genic male sterility, RGMS）及顯性基因控制的雄不稔（dominant genic male sterility, DGMS）二類。

## 2.5 十字花科作物細胞質雄不稔性的來源與利用

十字花科作物中，自發性的 CMS 最早於日本鹿兒島的野生蘿蔔（*Raphanus sativus* L.）中發現（Ogura, 1968），為 Ogu CMS 的起源，而另一個天然的‘Pol CMS’突變體則是 1972 年於大油菜品種的‘Polima’中發現（楊與傳, 1998），之後於大油菜的育種過程中亦陸續發現‘Nap CMS’（Thompson, 1972）及‘陝 2A CMS’（李, 1980），而近年則在芥菜（*B. juncea*）中發現‘126-1 CMS’（Sodhi et al., 2006）及‘hau CMS’（Wan et al., 2008）等 CMS 細胞質源。為了轉移這些重要的雄不稔細胞質以擴大應用於不同作物上，可利用種間雜交的策略，藉由細胞質為母性遺傳的特性，完整的將雄不稔性狀遺傳至不同物種上。事實上蕓苔屬種間雜交的研究相當多，歸納 326 篇有關蕓苔屬物種間天然雜交或人為控制授粉親合性的報告（FitzJohn et al., 2007），顯示數個蕓苔屬物種間並無嚴重的生殖隔離，而探討種間雜交後裔的細胞學行為也有助於不同物種間移轉優良性狀之效率，如 Leflon et al.（2006）探討種間雜交 *B. rapa* (AA, 2n=20) × *B. napus* (AACC, 2n=38) 之雜種 (AAC) 染色體的分離情形，發現所有雜種 (AAC) 植株中之雄可稔比率可達 60%~84%，進一步與雙親回交發現，若以 *B. napus* (AACC) 當母本且雜種 (AAC) 當父本時，其 BC<sub>1</sub> 的植株中約有 20%~26% 其配子染色體數達 n=18，已非常接近 *B. napus* (AACC) 之配子染色體數 n=19，若反交則其 BC<sub>1</sub> 的植株以 n=14 或 n=15 較多；這些研究結果可做為種間雜交導入性狀時，親本選擇之參考。

蕓苔屬作物具有易獲得種間雜種的特性，早在 1974 年 Bannerot 便將帶有 Ogu CMS 的蘿蔔（*Raphanus sativus*）與 *B. napus* 雜交，經染色體倍加後再回交三代而育成具有 Ogu CMS 的 *B. napus*，隨後在 1978 年 Heyn 將此 *B. napus* 與 *B. campestris* ssp. *trilocularis* 直接雜交後並回交四代而育成帶有 Ogu CMS 的 *B. campestris*（Kaul, 1988），Verma et al.（2000）則是將含 Pol CMS 的 *B. napus* 與 *B. campestris* 雜交，經回交五代而育成含有 Pol CMS 的 *B. campestris*，此外將具有 Pol CMS 之 *B.*




*napus* (AACC,  $2n = 38$ ) 與 *B. juncea* (AABB,  $2n = 36$ ) 進行種間雜交，發現其雜交  $F_1$  產生部份稔性，但經 5~7 世代回交後即可獲得 100% 雄不稔之 *B. juncea* (Sodhi et al., 2006)。而近年在 *B. juncea* 上發現的雄不稔細胞質‘*hau CMS*’，亦經由種間雜交導入 *B. napus* (Wan et al., 2008) 及 *B. rapa* 中 (Heng et al., 2015)，總括來說，十字花科的 CMS 來源中，目前以 Ogu CMS 與 Polima CMS 的利用上較為普遍，由於 Ogu CMS 具有不稔性穩定且不易受環境因子影響的特性，因此對許多以營養器官為主要收穫物之作物而言是理想的細胞質雄不稔源，目前許多甘藍類蔬菜如青花菜、花椰菜及甘藍等大多培育成具有 Ogu CMS 的母本以生產雜交種子。而大油菜的主要收穫物非營養器官，父本必須是恢復親才能使得  $F_1$  具有花粉而可生產榨油用的種子，由於 Polima CMS 源是在大油菜中發現的自發性突變體，比較容易在不同品系間搜尋出具稔性恢復基因的基因型，因此在應用上，目前大油菜仍以 Polima CMS 為主。

## 2.6 細胞質雄不稔性應用上可能的限制

CMS 於採種上的應用主要是進行三系雜交，依據這種模式，維持親 (B line) 需為正常細胞質且細胞核內為隱性同質的 *rf* 恢復基因，其與 CMS 品系 (A line) 雜交產生 100% 不稔後裔當作母本，在大油菜這種需產生種子的作物，其父本尚需具有核恢復基因 *Rf*；Bannerot et al. (1974) 將 Ogu CMS 轉入大油菜後即遭遇大部分品系不具 *Rf* 基因之問題，因此有學者進一步再將 *Rf* 基因由蘿蔔導入大油菜中 (Delourme et al., 1998)。除此之外，利用回交將‘Ogu CMS’細胞質導入大油菜時，同時出現了幼葉黃化及蜜腺不正常的問題 (Bannerot et al., 1974; 1977)，此種現象可能由於不同物種的細胞質與細胞核基因之間不親和所致，後來透過原生質體融合的方式將 CMS 貢獻親與接受親進行細胞融合，而獲得可利用之 Ogu CMS 種質 (Pelletier et al., 1983)。

Pol CMS 突變體於大油菜中發現，所以沒有缺少恢復基因的問題，但其不稔性卻易受環境影響而不穩定 (Fu et al., 1990)。根據統計不論是 CMS 或是 GMS，約 70% 以上品系會受環境因子影響而改變其稔性，其中 CMS 受影響的比例略高於 GMS，而環境因子中主要受溫度影響的例子較多，少部分會受光週期的影響 (Kaul, 1988)。於具有 Pol CMS 的大油菜研究指出，其雄不稔性不穩定之現象主要受溫





度影響，而可將其區分為高溫不稔型、低溫不稔型及穩定不稔型三種；其中高溫不稔型指不稔系在 15°C 以上不稔性表現完全，但低於該溫度會產生微量花粉，而低溫不稔型則在 12°C 以下不稔性穩定，反之亦出現微量花粉，品系間屬高溫不稔型的比例較高，而穩定不稔型比例較低 (Fu et al., 1990)。進一步探討 Pol CMS 大油菜‘溫度敏感因子’的遺傳性，顯示此表現型受對溫度敏感之核內一對隱性基因 (*ts*) 的影響，在高溫時 *ts* 基因會抑制恢復基因 *Rf* 的表現，而無法恢復稔性，且此 *ts* 基因與 *Rf* 不具等位性 (Fan et al., 2007)。Ogu CMS 大油菜的不稔系雖較不受溫度影響，但若與恢復系雜交而欲使其恢復稔性時，於日溫/夜溫為 32°C/26°C 的高溫環境條件下則稔性無法恢復 (Polowick and Sawhney, 1988)，Zhang and Hao (2001) 研究溫度對 CMS 結球白菜稔性恢復的影響，結果發現在人工控溫條件下，溫度 6~9°C 時處理 3 天以上其不稔性轉為可稔，而處理後至出現稔性約需 10~16 天。此種受溫度而影響雄不稔性之表現亦普遍出現在其他作物，如水稻雄不稔誘變株 ‘IR32364TGMS’ 在日/夜溫為 27°C/21°C 時可恢復稔性 (Borkakati and Virmani, 1996)，細香蔥 (chive) 則在 20°C/12°C 時表現不稔，但在相對高溫 24°C/24°C 時則稔性恢復 (Engelke et al., 2004)。除溫度外，番茄雄不稔突變體 ‘7B-1’ 的不稔性受光週期影響，當超過 12 小時的日長環境會表現出雄不稔，而在 8 小時以下的短日照則恢復稔性 (Sawhney, 2004)；水稻亦有受光週期影響雄不稔表現之例子 (He et al., 1999)，然而在蕓苔屬作物中則較少見受光週期影響的研究。

作物育種上除了考慮雄不稔的穩定性及恢復基因 *Rf* 是否普遍存在而能擴大雜交組合數之外，細胞質遺傳組成與農藝性狀之關係亦應一併考量，一些重要性狀如生長勢、種子產量、抗病性等可能與細胞質相關而產生副作用，Rajcan et al. (2002) 利用兩個 DH (Double haploid) 的油菜品系進行正反交，結果其後裔的次亞麻油酸 (linolenic) 含量、蛋白含量及開花期等性狀會受細胞質影響而偏向母本，因此若選用的 CMS 細胞質含有不良細胞質遺傳的性狀時也會限制其應用性，且有限的雄不稔細胞質來源，也可能提高因遺傳背景窄化而帶來遺傳脆弱性 (genetic vulnerability) 的風險；如玉米的 Texas CMS 品系其粒線體上具有 *T-urf13* 基因而易受玉米南方葉燒病 (southern corn leaf blight) 及玉米黃葉枯病 (yellow leaf blight) 的危害 (Levings, 1990) 便是著名的例子。由於現今可利用的 CMS 細胞質仍具前述限制因子，除了以各種手段克服這些問題之外，新 CMS 系統之開發以及 GMS

的使用便成為雄不稔性研究的新方向。



## 2.7 蕓苔屬作物細胞核雄不稔性的來源與遺傳

蕓苔屬作物當中已發現許多自發性 (spontaneous) 細胞核雄不稔源，如花椰菜 (Nieuwhof, 1961)、甘藍 (Nieuwhof, 1961; Fang et al., 1997)、不結球白菜 (Cao and Li, 1981; Ying et al., 2003)、大油菜 (Pan et al., 1988)、結球白菜 (Takahata et al., 1996; Feng et al., 2009) 及蕪菁 (Kenji et al., 2013) 等，由於雄不稔性狀之遺傳行為會影響後續的採種方式，因此通常會進一步分析這些細胞核雄不稔源其基因座之數量與對偶基因間之顯隱性關係。

### 2.7.1 隱性基因控制之細胞核雄不稔的遺傳與應用

十字花科作物中，大部份發現的自發性細胞核雄不稔源均受單一隱性核基因控制，如甘藍、花椰菜 (Nieuwhof, 1961)、不結球白菜 (Cao and Li, 1981; Ying et al., 2003; Li et al., 2016)、結球白菜 (Takahata et al., 1996; Zhang et al., 2010) 及蕪菁 (Kenji et al., 2013) 等，為了區分雄不稔性是否尚受到細胞質基因的影響，通常會於雄不稔株的雜交後代中，挑選具正常稔性的單株與其他品系進行正反交，進一步由正反交後裔的自交分離族群中檢測是否具有雄不稔株，以判別是否具有細胞質效應，經確認不受細胞質效應之影響後，便歸類於細胞核型雄不稔。一般認為受核基因控制的雄不稔源其稔性較細胞質型穩定而不易受環境因子影響，且在同物種或亞種之間透過雜交也可將穩定雄不稔特性遺傳至後代，例如日本在結球白菜 (*B. campestris* ssp. *pekinensis*) 品種 Hiratsuka 中所發現受一對隱性基因控制的雄不稔源，將其與結球白菜的其他品種 (cv. *Hiratsuka*) 或不結球白菜 (*B. campestris* ssp. *chinensis*) 的其他變種 var. *parachinensis*, var. *oleifera* 及 var. *chinensis* 等進行雜交，後代雄不稔性的遺傳行為及表現型均一致 (Takahata et al., 1996)。

核內單一隱性基因所控制的雄不稔遺傳機制中，假設雄不稔基因為 *ms*，當單株基因型為 *msms* 則呈現雄不稔性狀，則在雄不稔系內包含半數不稔株 *msms* 及半數可稔株 *Msms*，該雄不稔系繁殖時，需針對品系內不稔株與可稔株進行兄妹交 (sib mating) 後，再收穫不稔株上結實的種子，而利用該雄不稔系進行 F<sub>1</sub> 採種的模式如 (圖 1-2) 所示，由於該雄不稔系具有半數可稔株，因此於田間採種時雄不稔系需

密植，至開花初期需先拔除 50%可稔株，剩餘的不稔株才做為母本與花粉親雜交而生產雜交種子，但拔除雄不稔系中的可稔株極耗時費力，因此大規模商業化採種時難以擴大利用。

隱性基因控制的雄不稔性中，也曾發現涉及二對基因的例子。中國在大油菜的育種過程中，由低芥酸油菜品種‘Oro’中發現雄不稔突變株，其不稔性受兩對隱性核基因控制，當二基因均為隱性 *mslmsl ms2ms2* 時表現雄不稔外表特性，而當 *Ms1* 或 *Ms2* 任一存在則恢復稔性，此種材料若將雄不稔系其中一對隱性基因固定後，其  $F_1$  雜交種子生產模式相似於一對隱性基因所控制之雄不稔性；但欲利用雜交而轉移雄不稔性時，選拔上要同時針對二個基因座進行鑑定與篩選，實例如大油菜‘Oro’之後裔經過 25 世代兄妹交而純化出分別具有 *ms1* 與 *Ms1* 等位基因的雄不稔近似同源系（near isogenic lines, NILs）材料‘S45AB’（Pan et al., 1988）。

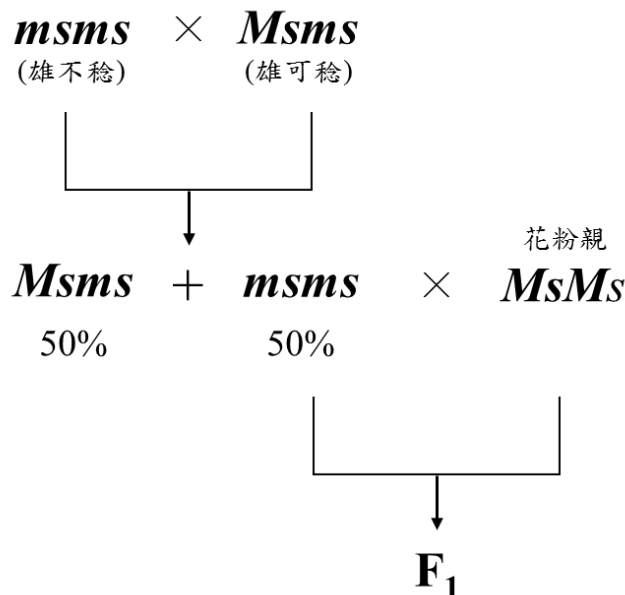


圖 1-2 利用單一隱性基因控制之雄不稔系生產雜交  $F_1$  種子的模式

## 2.7.2 單一顯性基因控制之細胞核雄不稔的遺傳與應用

十字花科作物中，極少數的例子顯示雄不稔性是受到核內單一顯性基因所控制，Van der Meer (1987) 在結球白菜品種‘Granaat’的後裔中發現雄不稔株，該不稔株於同品系內的兄妹交後裔均產生約 50% 的不稔株，以該不稔株與另一個結球白菜品種 ‘Monument’ 雜交，後裔也產生約 50% 雄不稔植株，持續將它們後裔的雄不稔株與 Monument 品系雜交連續三個世代結果均一致，而 Monument 品系各世代自交後裔均無產生不稔株，因此推測 Granaat 品系的雄不稔性是受到一對顯性基因控制。若以 *Ms* 表示雄不稔基因，理論上 *MsMs* 及 *Msms* 基因型將呈現雄不稔性狀，然而自然界中因不稔株無法天然自交，因而不存在顯性同質結合的 *MsMs* 基因型，故推測結球白菜品系 Granaat 中的雄不稔株與雄可稔株的基因型分別為 *Msms* 與 *msms*，至於 Monument 品系的基因型則為 *msms*，而利用該雄不稔系進行 F<sub>1</sub> 採種的模式如圖 1-3 所示，類似於核內單一隱性基因所控制的雄不稔系，由於雄不稔系與雄可稔系雜交後代中仍具有半數可稔株，因此於田間採種時也需拔除雄不稔系中的可稔株，而降低其實用價值。

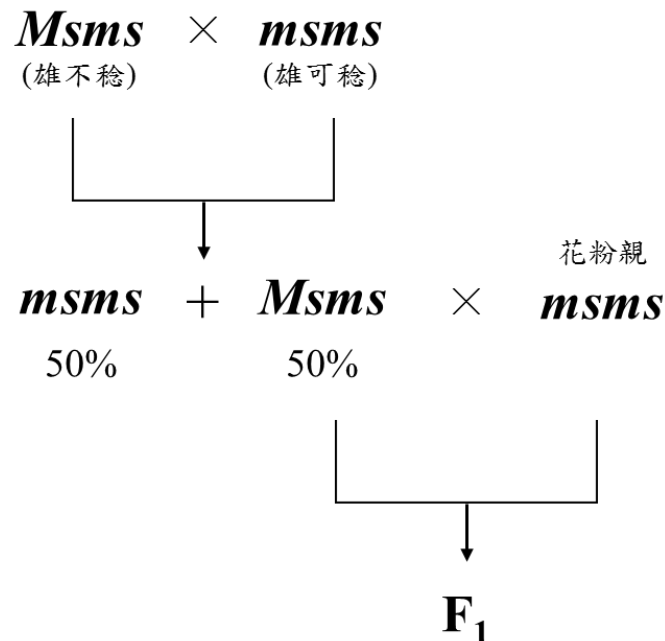


圖 1-3 利用單一顯性基因控制之雄不稔系生產雜交 F<sub>1</sub> 種子的模式

在花椰菜及甘藍也曾發現受核內單一顯性基因所控制的雄不稔株 (Ruffio-Chable et al., 1993; Fang et al., 1997)，然而這類雄不稔源有別於前述結球白菜的例子，研究顯示該雄不稔性狀會受環境因子影響 (特別是低溫環境) 而改變其稔性表現，如甘藍品系‘79-399-3’於溫度 < 15°C 時會產生微量花粉 (Wang et al., 2005) 可進行自交繁殖，因而有機會可獲得同質顯性 *MsMs* 基因型個體，經測交篩選出 *MsMs* 個體後，透過體細胞組織培養而量產繁殖成雄不稔系，由於雄不稔系所有個體均為 *MsMs*，與 *msms* 品系雜交可獲得 100% 雄不稔後裔，用於再次繁殖雜交種子 (圖 1-4)，此採種過程類似於 CMS 之三系採種，過程中不需拔除雄不稔系中的可稔株，具商業利用性。此外，於選育雄不稔品系時顯性雄不稔基因則具有性狀容易轉移的優勢，中國農科院培育之‘中甘 16 號’、‘中甘 17 號’、‘中甘 18 號’、‘中甘 19 號’及‘中甘 21 號’等即以此種方式生產。

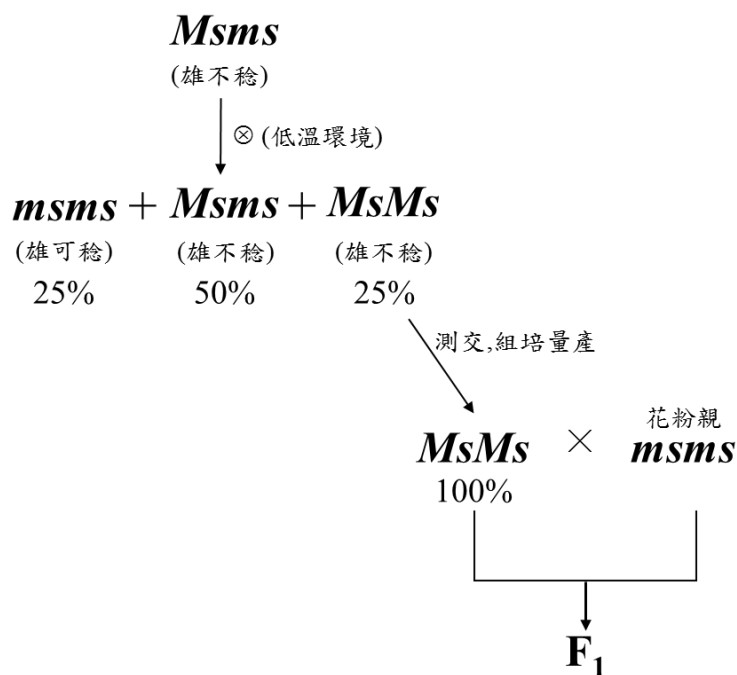


圖 1-4 利用對環境敏感的單一顯性基因控制之雄不稔系生產雜交 F<sub>1</sub> 種子的模式

### 2.7.3 細胞核內二對顯性基因所控制之雄不稔性的遺傳與應用

Li et al. (1985, 1988) 發現大油菜之‘Yi3A’品系材料受核內二對顯性基因所控制，所謂的二對顯性基因，其實一個是顯性的雄不稔基因 *Ms*，另一個則是顯性的恢復基因 *Rf*，顯性恢復基因 *Rf* 對顯性雄不稔基因 *Ms* 具上位性而可恢復稔性，因此只有當基因型為 *MsMs rfrf* 或 *Msms rfrf* 時可表現為雄不稔，由於雄不稔基因 *Ms* 為顯性因而被歸類為 DGMS，後續發現的大油菜‘Shaan GMS’（胡等，2004）、小油菜‘896AB’（董等，1999）以及由‘Yi3A’選育出的雄不稔系‘Rs1046AB’（Lu et al., 2004）均屬於此種類型。

在這種二對顯性基因交感作用的遺傳行為下，可培育出含有 *MsMs rfrf* 與 *MsMs Rfrf* 二種基因型組成的品系，即此雄不稔系中含 50% 的可稔株。繁殖雄不稔系時可透過雄不稔株 (*MsMs rfrf*) 與可稔株 (*MsMs Rfrf*) 進行兄妹交後，收穫不稔株結實的種子而達到增殖雄不稔系之目的，而進行雜交種採種則分為二階段進行，第一階段是將 *MsMs rfrf* (雄不稔) 與 *MsMs Rfrf* (雄可稔) 雜交後代 (含 50% *MsMs rfrf* 及 50% *MsMs Rfrf*) 與臨時維持系 (*msms rfrf*) 栽培田間後，當雄不稔系開花時拔除可稔株，保留雄不稔株與臨時維持系雜交，此時雄不稔系上所收穫的種子其基因型均為 *Msms rfrf*，因此後裔全數為雄不稔株，第二階段再利用此 100% 雄不稔後裔做為母本與父本花粉親雜交生產商業用雜交種子 (圖 1-5)，育種時若將臨時維持系改良為雄不稔系的近同源系 (即只有雄不稔基因型有差異)，則生產的雜交種子為單交種，若臨時維持系與雄不稔系並非近同源系且具不同遺傳組成則成為三交種，而此種三系雜交的採種模式也近似於 CMS 之三系雜交，主要差別在於第一階段母本繁殖時需針對雄不稔系中之可稔株進行淘汰，此法已應用於中國大油菜的雜交種生產上 (Li et al., 1985; 1993)。

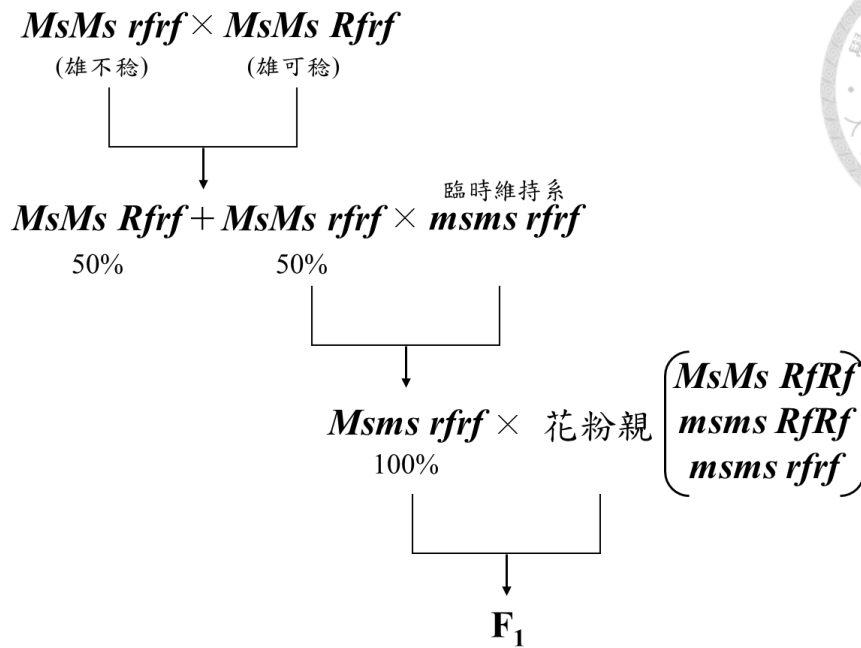


圖 1-5 利用二對顯性基因控制之雄不稔系生產雜交 F<sub>1</sub> 種子的模式

#### 2.7.4 細胞核內複等位基因所控制之雄不稔性的遺傳與應用

針對結球白菜上所發現的雄不稔源，馮等（1995）以該雄不稔系進行二種雜交組合，並利用後裔檢定方式分析其遺傳行為，而提出該結球白菜雄不稔性係受複等位基因控制的遺傳模式，其遺傳分析結果顯示雄不稔基因座具三個複等位基因  $Ms^f > Ms > ms$ ，其中只有  $Ms$  等位基因表現為雄不稔，但會受到另一個等位基因  $Ms^f$  作用而使稔性恢復，因此只有當基因型為  $MsMs$  或  $Msms$  時才產生雄不稔外表型（Feng et al., 1996）。宋等（2005）對雄不稔大油菜‘609AB’品系進行研究時，起初亦認為該不稔性受兩對非等位顯性基因  $Ms$  與  $Mf$  控制，但進行遺傳分析後也發現此二基因具有等位性，可套用與前述結球白菜相同之遺傳假說，而後透過分子標誌的連鎖關係分析進一步證實此為複等位基因遺傳模式（Song et al., 2006）。

複等位基因模式下進行雜交種採種的方式與二對顯性基因模式幾乎一樣，差別只在抑制雄不稔基因的 allele 是否位於相同基因座，以 Feng et al.（1996）所提之複等位基因  $Ms^f > Ms > ms$  為例，培育出的雄不稔系包含  $Ms^fMs$  與  $MsMs$  二種基因型雜交後代，因此同樣也含有 50% 的可稔株，繁殖雄不稔系時透過雄不稔株（ $MsMs$ ）與可稔株（ $Ms^fMs$ ）進行兄妹交後收穫不稔株結實的種子。雜交種子生產

時也是採二階段進行，第一階段是將雄不稔系含 50%  $M_sM_s$  及 50%  $M_s^fM_s$  植株與臨時維持系 ( $msms$ ) 栽培田間後，當雄不稔系開花時拔除可稔株 ( $M_s^fM_s$ )，保留雄不稔株 ( $M_sM_s$ ) 以與臨時維持系雜交，此時雄不稔系上所收穫的種子其基因型均為  $M_sms$  故後裔全數為雄不稔株，第二階段再利用此 100% 雄不稔後裔做為母本與父本花粉親雜交生產商業用雜交種子 (圖 1-6)，雖然複等位基因採種模式幾乎同於二對顯性基因模式，然而在雄不稔系的育種過程中，進行性狀轉移時複等位基因模式只需考慮單一基因座上可能產生之 6 種基因型，將比二對顯性基因模式中二基因座可能產生之 9 種基因型更為簡便。

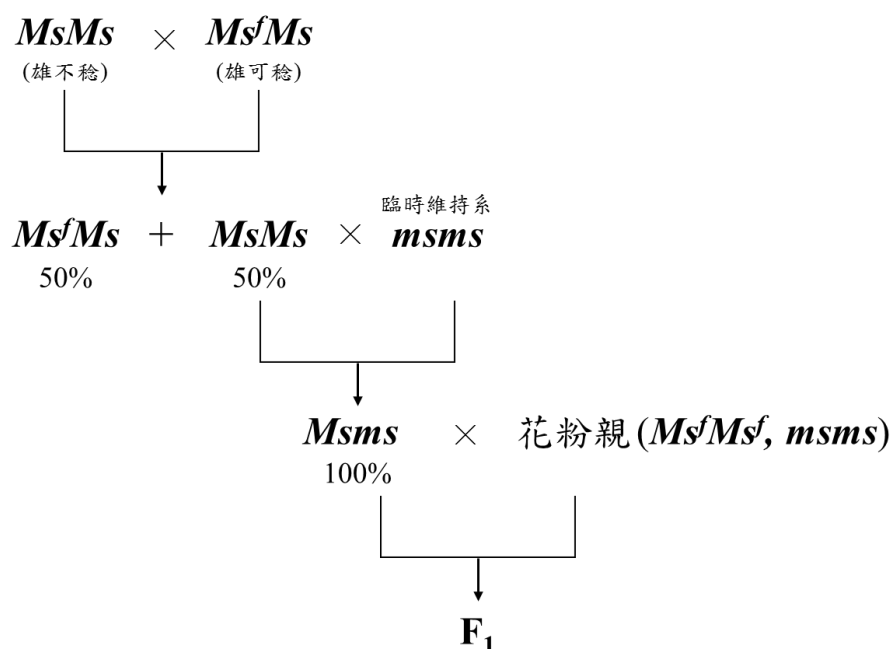


圖 1-6 利用複等位基因控制之雄不稔系生產雜交 F<sub>1</sub> 種子的模式

### 2.7.5 細胞核內多個基因所控制之雄不稔性的遺傳與應用

前述中國所培育的大油菜雄不稔系‘S45AB’主要受到兩對隱性雄不稔基因控制，另一個雄不稔大油菜品系‘9012AB’的遺傳雖類似，但它除了受兩對隱性雄不稔基因控制之外，還受另一個隱性恢復基因  $rf$  控制而影響稔性 (Chen et al., 1998)，由於有三個基因影響雄不稔性之表現，因此是目前十字花科蕓苔屬作物中雄不稔性遺傳機制最複雜的例子。經分析發現雄不稔系 9012A 與 S45A 之間的雄不稔基因不具等位性 (Tu et al., 1997; Chen et al., 1998)，為了與 S45A 的二個雄不稔基因




*Bnms1* 及 *Bnms2* 有所區別，而將 9012AB 的兩個隱性基因命名為 *Bnms3* 及 *Bnms4* (Huang et al., 2007; He et al., 2008)，在 9012AB 雄不稔性的遺傳模式中，由於隱性恢復基因 *rfrf* 具有上位性，因此當恢復基因為顯性 *Rf*，而二個雄不稔基因為同質隱性（即 *ms3ms3ms4ms4Rf\_*）時，則可表現為雄不稔；然而基因型若為 *ms3ms3ms4ms4 rfrf* 時，因隱性上位基因 *rfrf* 抑制雄不稔基因之表現而恢復稔性。

作物育種時常將隱性雄不稔基因之一（*ms3ms3* 或 *ms4ms4*）及顯性恢復基因 *RfRf* 固定之後，藉由異質結合的 *Ms3ms3*（或 *Ms4ms4*）基因培育出包含 50% 不稔株 *ms3ms3 ms4ms4 RfRf* 及 50% 可稔株 *Ms3ms3 ms4ms4 RfRf*（或 *ms3ms3 Ms4ms4 RfRf*），二個隱性雄不稔基因中的一個被固定之後，其遺傳行為就等於是前述二對顯性基因的遺傳模式，雜交種子生產時，第一階段中拔除其中的可稔株（*Ms3ms3 ms4ms4 RfRf* 或 *ms3ms3 Ms4ms4 RfRf*），由剩下的不稔株 *ms3ms3 ms4ms4 RfRf* 與基因型為 *ms3ms3 ms4ms4 rfrf* 之臨時保持系雜交，以繁殖出 100% 雄不稔後裔以做為雜交品種之母本，第二階段再與父本花粉親雜交生產商業用種子。

大油菜‘9012AB’雄不稔性的遺傳機制雖複雜，但後來透過分子層次的研究發現，*Bnms4* 與抑制基因 *BnRf* 具有等位性（Zu et al., 2010; Dong et al., 2012; Xie et al., 2012），而 *Rf* 實為 *Ms4* 基因座上的一個等位基因，因此在 *Ms4* 基因座上有三個複等位基因，若以原基因符號而言，其顯隱性應為  $Ms4 > Rf > ms4$ （之後修正為  $Rf^a > Rf^b > Rf^c$ ），其中 *Rf* 表現雄不稔性狀，而雄不稔株之基因型應為 *ms3ms3 RfRf*（修正後為 *ms3ms3 Rf<sup>b</sup>Rf<sup>b</sup>*）或 *ms3ms3 Rfrf*（修正後為 *ms3ms3 Rf<sup>b</sup>Rf<sup>c</sup>*），也就是說大油菜‘9012AB’的雄不稔性同時具有前述二對顯性基因模式與複等位基因模式，若將 *Ms3* 基因座固定成隱性同質 *ms3ms3* 基因型，則可視為複等位基因模式，以圖 1-6 的方式進行 F<sub>1</sub> 種子生產，反之，若將複等位基因座固定為 *Rf<sup>b</sup>Rf<sup>b</sup>* 基因型則可視為二對顯性基因模式，並以圖 1-5 的方式進行 F<sub>1</sub> 種子生產。

## 2.8 雄不稔性相關基因的分子標誌與基因定位

在雄不稔性育種的利用上，為避免鑑定不稔系、維持系或恢復系的基因型的繁瑣測交工作，分子標誌之應用則更顯重要，CMS 的雄不稔基因位於細胞質，係遺傳自母本，育種上不需篩選，因此 CMS 相關分子標誌之研究大都著重於細胞核恢復基因上。早年當作物基因體資訊有限時，欲搜尋特定性狀連鎖之分子標誌常利用



大量可產生多型性的分子標誌搭配混合 DNA 法 (bulked segregant analysis, BSA) 進行, 由於 BSA 法於分離族群中可抵消非目標基因遺傳背景之干擾, 為開發性狀連鎖分子標誌快速有效的策略 (Michelmore et al., 1991), 因此許多 CMS 恢復基因分子標誌即以此方式篩選而建立, 如蘿蔔 'Kosena CMS' 恢復基因分子標誌藉 BSA 的方式進行隨機增幅多型性 DNA (random amplified polymorphic DNA, RAPD) 及增殖片段長度多型性 (amplified fragment length polymorphism, AFLP) 分析而建立 (Imai et al., 2003); 大油菜利用 CMS 生產雜交種時需有恢復基因才能結實, 因此針對大油菜 Pol CMS 的恢復基因很早就已開發緊密連鎖之限制片段長度多型性 (restriction fragment length polymorphism, RFLP) 分子標誌 (Jean et al., 1997); 而轉移 Ogu CMS 至大油菜中的過程中, 配套的恢復基因亦已成功利用 RAPD 分子標誌標定 (Delourme et al., 1994; Hansen et al., 1997), 並進一步建構遺傳連鎖圖譜 (genetic map) 而定位出恢復基因的位置 (Delourme et al., 1998)。


細胞核雄不稔之利用逐漸受到重視, 利用細胞核雄不稔性進行雜交種子生產時, 儘管雄不稔性的遺傳行為多樣化, 然而在各採種模式下必須執行淘汰可稔株的程序, 此外, 當雄不稔性的遺傳機制愈複雜, 以傳統育種程序進行雄不稔系培育或性狀轉移的困難度越高, 更需藉助分子標誌輔助選種 (marker assisted selection, MAS) 來提高育種效率, 因此細胞核雄不稔基因相關分子標誌之開發以及基因定位往往都是研究熱點, 在十字花科作物中, 相關的研究以大油菜最為完整。

大油菜在二對隱性核基因控制的雄不稔系 'S45AB' 的研究中, 針對後裔族群將不稔與可稔的植株以 BSA 的方式進行 AFLP 分析, 而找到與 *Ms1* 基因共分離 (co-segregation) 的 AFLPs 分子標誌, 進一步轉換成共顯性標誌的 SCARs (sequence characterized amplified region) 分子標誌, 透過基因定位結果發現 *Ms1* 基因位於 *B. napus* 的 N7 連鎖群上 (Yi et al., 2006), Lei et al. (2007) 也用相同雄不稔源為材料進行測交及等位性分析, 分離出 *ms1ms1 Ms2ms2* 及 *ms1ms1 ms2ms2* 近同源系之同時, 再次確認 *Ms1* 及 *Ms2* 並非等位基因, 並找到與 *Ms2* 基因座連鎖的 AFLPs 分子標誌, 且轉換成 SCARs 分子標誌, 進一步透過比較基因體學之方法, 利用阿拉伯芥上的分子標誌將 *Ms2* 基因定位在 *B. napus* 的 N16 連鎖群。

大油菜‘Yi3A’雄不稔性受核內二對顯性基因所控制，Lu et al. (2004) 利用源自 Yi3A 之雄不稔基因之回交族群，以 BSA 的策略進行 AFLP 分析，而開發出與雄不稔 *Ms* 基因連鎖的分子標誌，Hong et al. (2006) 同樣以 Yi3A 之近同源系為材料，透過 BSA 法找到三個與恢復基因 *Rf* 共分離的 AFLPs 分子標誌，並進一步轉換為三個顯性的 SCARs 標誌，為了能快速簡便的進行基因分型，而有效率地用於育種篩選，Hong et al. (2008) 進一步把 Lu et al. (2004) 在 *Ms* 基因上所開發之 AFLPs 分子標誌也轉成 SCARs 分子標誌，並將前述與 *Rf* 基因有關的三個顯性 SCARs 分子標誌於其他不同的族群及品系進行驗證，結果顯示雖然只有一個 SCAR 與 *Rf* 基因緊密連鎖，但此原為顯性的 SCAR 卻能在 *Rf/rf* 基因型中擴增出兩個不同大小的條帶，而成為共顯性(codominance)的分子標誌(Hong et al., 2008)。

雄不稔油菜‘609AB’受單一基因座上之複等位基因控制，等位基因之顯隱性關係為  $Mf > Ms > ms$ ，利用 BSA 的方式分別建構不同等位基因組成之 *Ms/ms* 族群及 *Mf/Ms* 族群，結果於 *Ms/ms* 族群中找到與 *Ms* 連鎖的 AFLPs 分子標誌並轉換成 SCARs；其中一個 SCARs 序列經定序並再修飾後，可於另一個建構的 *Mf/Ms* 族群中與 *Mf* 基因共分離，因此，此研究除了開發出與雄不稔基因連鎖之分子標誌外，也證實該雄不稔性為複等位基因之遺傳模式 (Song et al., 2006)，然而，尚未有該雄不稔基因定位之相關發表。

大油菜‘9012AB’雄不稔遺傳機制最為複雜，由於‘9012AB’與後續發現的雄不稔系‘7-7365AB’其雄不稔基因間具有等位性(Huang et al., 2007; Xiao et al., 2008)，因此整合二個雄不稔系的分子研究可對該雄不稔源的所有基因有完整之描述。在 2010 年之前，‘9012AB’的雄不稔性認為是受雄不稔基因 *Bnms3* 及 *Bnms4* 以及隱性恢復基因 *Bnrf* 等三個基因控制，為了開發分子標誌以輔助雄不稔性狀之篩選，針對 9012AB 之不稔性，透過 BSA 之策略找到與 *ms3* 共分離之 AFLP 分子標誌 (Ke et al., 2005) 並進而轉為 SCAR 分子標誌 (He et al., 2008)；Huang et al. (2007) 以雄不稔系‘7-7365AB’為參試材料，同樣採取 BSA 之策略，利用基因型為 *Ms3ms3ms4ms4 RfRf* 之可稔株自交分離族群中找到與 *Ms3* 共分離之 SCAR 分子標誌，進一步利用二個公開之油菜 DH (doubled haploid) 族群分析該 SCAR 分子標誌，而將 *Bnms3* 定位於 *B. napus* 的 N19 連鎖群上，除了 *Bnms3* 之外，恢復基因 *Bnrf* 也透過 BSA 策略及基因精細定位方式，開發其 SCAR 分子標誌 (Xie et al., 2008; Xu et al.,



2009) , 此外 , Xiao et al. (2008) 則是利用 7-7365AB 雄不稔系為試驗材料建構恢復基因 *Bnrf* 的分離族群 , 並且將 *Bnrf* 基因定位於 *B. napus* 的 N7 連鎖群上。相對於 *Bnms3* 或 *Bnrf* 基因而言 , *Bnms4* 基因的研究則較少 , Zu et al. (2010) 針對 *BnMs4* 進行基因定位時 , 透過分子標誌間之連鎖失衡情形發現 *Bnms4* 與 *Bnrf* 均座落於 *B. napus* 的 N7 連鎖群相近區域 , 因而推測 *Bnms4* 與 *Bnrf* 可能具有等位性 , 透過系統性的遺傳分析及分子層次的基因精細定位 , 其結果支持 *Bnms4* 與 *Bnrf* 為複等位基因之論點 , 並將該基因座之物理圖譜區域限縮於 50 kb 的區域內 (Dong et al., 2012) , 而後藉由結球白菜及阿拉伯芥之基因體資訊進行比較基因體分析 , 進一步將 *BnRf* 基因座定位於 13.8 kb 之區間內 (Xie et al., 2012) , 整合前述結果及近年基因精細定位之資訊 , 針對含 *BnRf* 等位基因之 BAC library 進行篩選後終於選殖出 *BnRf* 基因 , 該基因可轉譯 mtHSP70-1-like 之蛋白質 , 進一步將其轉殖至阿拉伯芥後 , 由性狀表現證實 *BnRf* 基因之功能 (Deng et al., 2016) 。

事實上 , 受二對顯性核基因控制的大油菜雄不稔源 'Rs1046AB' (雄不稔基因源自 'Yi3A') , 後來證實其顯性雄不稔基因 *Ms* 與恢復基因 *Rf* 也具有等位性 (Liu et al., 2008) , 為了與大油菜其他雄不稔基因有所區別 , 將該基因名為 *Brms5* 且定位於 *B. napus* 的 A8 連鎖群上 (Lu et al., 2013)

相較於大油菜 , 白菜類作物細胞核雄不稔基因的研究較少 , 在結球白菜方面 , 針對單一隱性基因控制的雄不稔性 , 透過 BSA 之策略結合 SRAP (sequence-related amplified polymorphism) 分子標誌分析 , 篩選得到與雄不稔基因相距 2.98 cM 之分子標誌 , 並將該基因定位於 A8 連鎖群 (張等 , 2010) , 而針對大白菜的複等位基因遺傳之雄不稔性 , 則已開發出二個 SCARs 分子標誌 , 且將雄不稔基因定位於 R7 (即 A7) 連鎖群 (Feng et al., 2009) 。而在不結球白菜中 , Ying et al. (2003) 將單一隱性基因控制的雄不稔突變體與 'Huang-Bai-Qing' 品種回交 16 代後 , 利用其近同源系而開發 *Ms* 基因的 AFLPs 分子標誌。另一個不結球白菜雄不稔系 'WS24' 同樣是受單一隱性基因控制 , 利用 'WS24' 與可稔自交系的 F<sub>2</sub> 分離世代 , 以商業化的莖苔屬 SNP 晶片 (Illumina Brassica 60 K SNP array) 進行偵測 , 而找到與雄不稔性狀連鎖的 SNP , 並將該基因定位於 *B. rapa* 之 A2 連鎖群 (Li et al., 2016) 。



## 第2章 不結球白菜雄不稔性的遺傳模式



## 1 前言

雄不稔性在育種上主要應用於一代雜交品種種子之生產，為因應雄不稔基因之遺傳行為，就須建立配套的採種方式，因此有必要確認雄不稔性狀屬於細胞質遺傳或僅受細胞核基因影響，若屬於細胞核雄不稔遺傳，也須進一步釐清控制性狀的基因座數量與對偶基因間之顯隱性關係。

我們於台灣綠肥用小油菜中發現的雄不稔株，在培育雄不稔系的過程中，利用一些測交與正反交的結果，初步歸納及判定其非屬於細胞質遺傳，然而它與一些特定品系雜交後，卻可產生類似細胞質雄不稔之 100% 為雄不稔株的後裔，由於目前在不結球白菜上所發現之細胞核雄不稔基因，均受一對隱性基因所控制（Cao and Li, 1981; Ying et al., 2003; Li et al., 2016），尚未發現其他雄不稔遺傳模式，為開發該新雄不稔源之應用性而進行本試驗。

依據前人研究，屬於細胞核雄不稔又可產生 100% 雄不稔後裔的遺傳模式並不多，彙整本研究小油菜歷年測交之結果，顯示該雄不稔性可符合二對顯性遺傳及複等位遺傳等二種模式，實務上有必要進一步確認屬何種遺傳模式，而採用的分析方法也需考量族群之大小、統計上的效力及參試材料的遺傳背景等。在本章研究中，我們採取了由一系列後裔檢定所組成的分析方法，主要依據可能的遺傳模式先以理論值推導各種雜交後裔族群的外表型，並同時考量受測雄不稔個體為不同基因型時可能產生的結果，當實際完成各種測交後裔檢定時，依據各項檢定結果可歸納出該雄不稔性適配的遺傳模式與參試植株的基因型，此外，考量不結球白菜存在不同的生態型，參試材料中也加入不同類型的品系，且進行相同的遺傳分析試驗，期望能客觀的評估出該雄不稔性之實際遺傳模式。

## 2 材料與方法

### 2.1 試驗材料

具雄不稔性的自發性突變體，是 2004 年從台灣的綠肥用小油菜 (*Brassica rapa* ssp. *chinensis* var. *oleifera* Makino) 地方品種“NH80”中所發現，該突變體與同品種之雄可稔單株經兄妹交 8 代所培育出雄不稔系“NH80-AB”，若將該品系之雄不稔株 (NH80-A) 與可稔株 (NH80-B) 進行兄妹交，後裔可獲得 50% 的雄不稔植株，

雄不稔株花器型態上有花藥萎縮白化，內無花粉、花絲短縮的外觀，目視易於觀察（圖 2-1）。恢復系 TA95 與 SG388 分別為源自泰國及中國的蔬菜用油菜品種，經自交 7 代而培育之自交系，至於維持系 WH606 則為台灣葉用油菜地方品種自交系（表 2-1、2-2）。

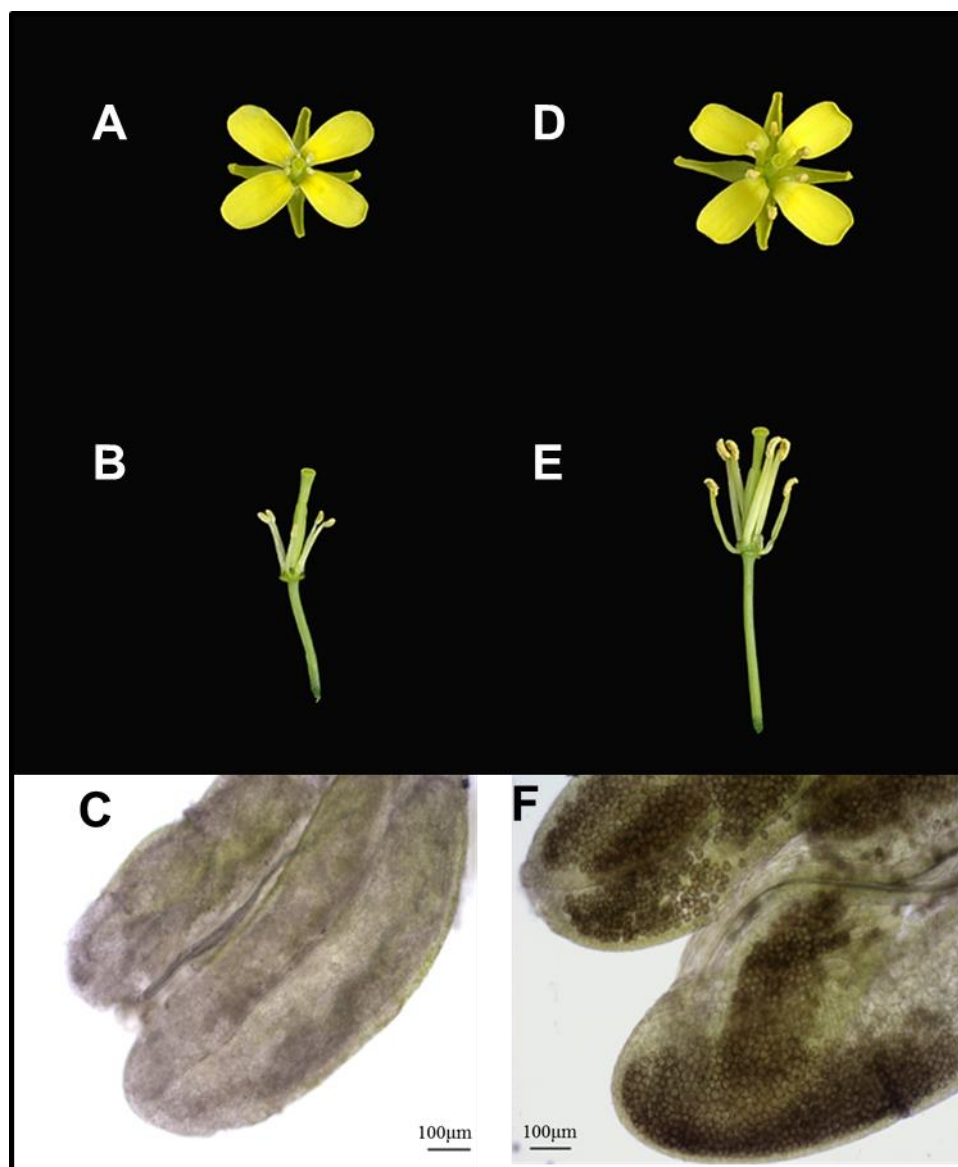


圖 2-1 不結球白菜 NH80 花器形態及花粉形成之差異  
A,B,C: 雄不稔株 (NH80-A)  
D,E, F: 雄可稔株 (NH80-B)



表 2-1 參試不結球白菜之品系來源與主要用途

品系	性質	世代	來源	用途
NH80-A	雄不稔系 (不稔株)	Sib <sub>8</sub>	台灣	綠肥、榨油
NH80-B	雄不稔系 (可稔株)	Sib <sub>8</sub>	台灣	綠肥、榨油
TA95	恢復系	S <sub>7</sub>	泰國	葉用蔬菜
SG388	恢復系	S <sub>4</sub>	中國	苔用蔬菜
WH606	維持系	S <sub>9</sub>	台灣	葉用蔬菜

表 2-2 參試不結球白菜品系之學名

品系	學名
NH80-A	<i>Brassica rapa</i> ssp. <i>chinensis</i> var. <i>oleifera</i> Makino
NH80-B	<i>Brassica rapa</i> ssp. <i>chinensis</i> var. <i>oleifera</i> Makino
TA95	<i>Brassica rapa</i> ssp. <i>chinensis</i> var. <i>communis</i> Tsen et Lee
SG388	<i>Brassica rapa</i> ssp. <i>chinensis</i> var. <i>parachinensis</i> Bailey
WH606	<i>Brassica rapa</i> ssp. <i>chinensis</i> var. <i>communis</i> Tsen et Lee

## 2.2 後裔檢定試驗之遺傳分析

### 2.2.1 雄不稔性遺傳模式之假設

雄不稔性狀遺傳分析主要參考劉 (1992) 設計之後裔檢定試驗略修改而進行，於歷年測交結果中，顯示該雄不稔性狀符合單基因座 (內含複等位基因) 的模式或二對顯性基因模式，因此先針對此二種模式分別定義基因 (或等位基因) 間之顯隐性關係。在單基因模式下，以 *BrMs* 表示該雄不稔基因，該基因座內有 *BrMs<sup>a</sup>*、*BrMs<sup>b</sup>* 及 *BrMs<sup>c</sup>* 三個等位基因，*BrMs<sup>b</sup>* 為表現雄不稔的等位基因且對 *BrMs<sup>c</sup>* 為顯性，因此當個體基因型為 *BrMs<sup>b</sup>Ms<sup>b</sup>* 及 *BrMs<sup>b</sup>Ms<sup>c</sup>* 時表現雄不稔，但 *BrMs<sup>a</sup>* 則會抑制 *BrMs<sup>b</sup>* 而使稔性恢復；在二對顯性基因模式下，以 *BrMs* 表示雄不稔基因，而另一個獨立的 *BrRf* 為恢復基因，*BrRf* 會抑制 *BrMs* 基因表現而恢復植株稔性，因此當個體基因型為 *BrMsMs rfrf* 及 *BrMsms rfrf* 時才表現出雄不稔性狀 (附錄 1)。這兩種遺





傳模式可透過一系列後裔檢定加以區別，主要為後裔檢定 A (progeny test A)、後裔檢定 B (progeny test B) 及後裔檢定 C (progeny test C) 等三種後裔檢定試驗 (圖 2-2、附錄 2~4)。

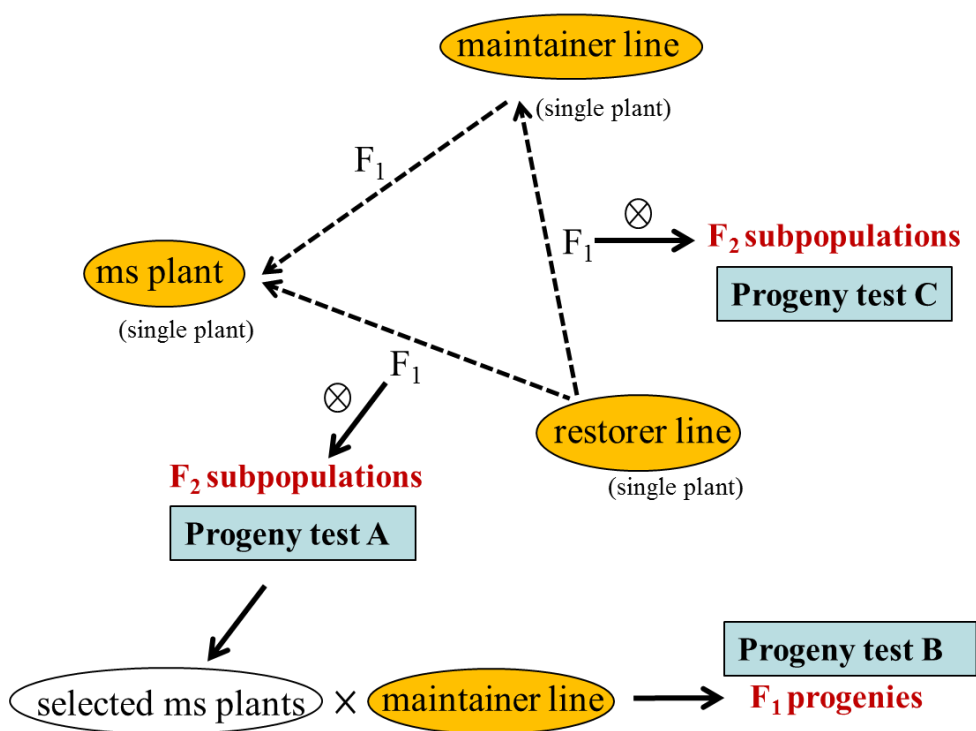


圖 2-2 Progeny test A、Progeny test B 及 Progeny test C 等三種後裔檢定試驗  
本圖修改自劉 (1992) 之後裔檢定試驗，虛線箭頭表示品系間雜交授粉之方向。ms plant: NH80-A; maintainer line: WH606; restorer lines: TA95, SG388

### 2.2.2 試驗材料相互雜交建構後裔檢定分析族群

本研究之後裔檢定試驗需以三個栽培季針對參試材料進行相互雜交授粉方可完成，因此於 2010 至 2012 年期間，每年 1 月份種植參試材料並雜交授粉，整體雜交過程如圖 2-2 所示，第一年將雄不稔株、恢復系及維持系分別取 1 單株相互授粉而獲得雄不稔株×恢復系、雄不稔株×維持系及恢復系×維持系之 F<sub>1</sub> 個體，此外恢復系與維持系單株也同時保留部分花枝進行自交以做為後續世代雜交之花粉親，第二年將雄不稔株×恢復系及恢復系×維持系之 F<sub>1</sub> 個體進行自交，而獲得進行

progeny test A 與 progeny test C 所需之分離次族群 (subpopulations) ，第三年則由 progeny test A 中的雄不稔株與維持系雜交而獲得可進行 progeny test B 的後裔種子，基於前述授粉方式，由於單株間相互雜交或自交時花粉來源均不同，為避免花粉污染，恢復系與維持系進行雜交時，以恢復系作為母本並於花蕾期先行除雄後再授予維持系之花粉。此外，在所有雜交或自交過程中，各花枝於授粉完成後均加以標記並套袋隔離，於果莢成熟後再分別採收。

由於本研究包含兩個恢復系，因此後裔檢定試驗可視為兩個雜交群組進行分析，各雜交群組第一年將產生五種後裔，因此以 TA95 為恢復系之第一雜交群組而言有 NH80-A×TA95、NH80-A×WH606、WH606×TA95、TA95 自交及 WH606 自交等五種後裔，而以 SG388 為恢復系之第二雜交群組而言則有 NH80-A×SG388、NH80-A×WH606、WH606×SG388、SG388 自交及 WH606 自交等後裔。

### 2.2.3 後裔檢定試驗之規模及性狀調查

配合品系間雜交或自交授粉進度，各種後裔檢定族群，分別於 2011 年至 2014 年冬裏作栽培於農業試驗所試驗田，個體的稔性調查於開花期以目視進行，各雜交群組獲得五種後裔後，再由其後裔進行 progeny test A、progeny test B 及 progeny test C 等三種後裔檢定試驗 (圖 2-3、2-4)，其中 progeny test A 主要檢測雄不稔株×恢復系雜交組合之所有 F<sub>2</sub> 次族群是否均產生稔性分離，以 TA95 為恢復親的第一雜交群而言，共包含 64 個 NH80-A×TA95 之 F<sub>2</sub> 次族群 (subpopulations)，每個次族群栽培至少 60 株，以調查各次族群內雄不稔性狀分離情形；progeny test B 則是將 progeny test A 之 F<sub>2</sub> 族群的雄不稔株與維持系進行雜交，並檢定雜交 F<sub>1</sub> 後裔各家系是否產生稔性分離情形，即由 NH80-A×TA95 的 F<sub>2</sub> 次族群中逢機選取 40 個雄不稔單株與 WH606 雜交，針對各 F<sub>1</sub> 家系調查至少 20 個體的雄蕊稔性；而 progeny test C 主要為檢定恢復系與維持系雜交後所有衍生之 F<sub>2</sub> 次族群是否產生雄不稔植株，以檢定雄不稔基因之等位性，即以 TA95×WH606 之 25 個 F<sub>2</sub> 次族群各栽培 60 株，以調查各次族群雄不稔性狀分離情形。

對於以 SG388 為恢復親的第二雜交群而言，progeny test A 包含 43 個 NH80-A×SG388 之 F<sub>2</sub> 次族群，progeny test B 包含 37 個雄不稔單株 (源自 NH80-A×SG388



F<sub>2</sub> 族群) 與 WH606 雜交衍生的 F<sub>1</sub> 家系, progeny test C 則有 25 個 SG388×WH606 之 F<sub>2</sub> 次族群。

#### 2.2.4 利用後裔檢定試驗之結果推導雄不稔性之遺傳模式

針對複等位基因模式與二對顯性基因模式之遺傳假說, 各品系所有可能的基因型如附錄 1, 其中若以雄不稔個體與恢復系進行雜交時共有 8 種可能的基因型組合, 在這 8 種組合下, 三種 progeny test 其後裔族群或家系雄不稔性狀分離情形的理論值如附錄 2~4, 試驗主要目的為檢定各後裔族群是否產生稔性分離, 而非性狀分離比, 整合 progeny test A、progeny test B 及 progeny test C 等三種後裔檢定試驗的結果, 可分析出參試品系的基因型, 進而推論本研究油菜突變體其雄不稔性的遺傳模式。

### 3 結果

#### 3.1 以 TA95 為恢復系之第一雜交群組後裔檢定試驗

為了確認 NH80-A 雄不稔性狀之遺傳模式, 而將雄不稔株 NH80-A、恢復系 TA95 及維持系 WH606 相互雜交, 結果顯示 TA95 及 WH606 的自交後裔、NH80-A × TA95 及 TA95 × WH606 的 F<sub>1</sub> 後裔均具正常稔性, 然而 NH80-A × WH606 之 F<sub>1</sub> 後裔所有植株均為雄不稔 (表 2-3)。

表 2-3 第一雜交群組後裔之稔性表現

Cross	No. of male sterile plants	No. of male fertile plants
NH80-A × TA95	0	70
NH80-A × WH606	40	0
WH606 × TA95	0	50
TA95⊗	0	72
WH606⊗	0	70

NH80-A × WH606 之所有 F<sub>1</sub> 後裔均為雄不稔，顯示參試的雄不稔 NH80-A 單株，其雄不稔基因對維持系而言為顯性同質結合體，而該雄不稔單株與恢復系 TA95 的 F<sub>1</sub> 後裔中，所有植株均表現為雄可稔，顯示恢復系的稔性恢復基因對參試的 NH80-A 單株而言也為顯性同質結合體。因此在二種雄不稔性遺傳模式的定義下，依據第一年品系間之雜交結果，參試雄不稔株於複等位基因模式下基因型應為 *BrMs<sup>b</sup>Ms<sup>b</sup>*，若在二對顯性上位基因模式下則應為 *BrMsMsrfjf*；而維持系 WH606 基因型則應分別為 *BrMs<sup>c</sup>Ms<sup>c</sup>* 與 *Brmsms rfjf*；恢復系 TA95 於複等位基因模式下，基因型應為 *BrMs<sup>a</sup>Ms<sup>a</sup>*，二對顯性上位基因模式下則應為 *BrMsMs RfRf*、*BrMsms RfRf*、*Brmsms RfRf*（附錄 1）。

在 progeny test A 中，雄不稔株與恢復系雜交後裔所有 F<sub>2</sub> 次族群均呈現稔性分離，在 NH80-A×TA95 的雜交組合中，共獲得 64 個 F<sub>2</sub> 次族群（表 2-4），其中同時符合 1：3 與 3：13 分離比者達 50 個，僅管有 8 個次族群雄不稔單株與可稔單株之比例經統計檢定推翻 1：3 分離比之假設，但這些次族群全都符合 3：13 之分離比，同樣地，雄不稔單株與可稔單株之比例經統計檢定推翻 3：13 分離比之次族群共有 6 個，但這些也都符合 1：3 之分離比，亦即所有 F<sub>2</sub> 次族群中並無產生既不符合 1：3 也不符合 3：13 分離比者。

進一步由 NH80-A × TA95 的 F<sub>2</sub> 次族群中逢機選取 40 個雄不稔單株與維持系 WH606 雜交，由衍生出 40 個 F<sub>1</sub> 品系進行 progeny test B，這些品系的所有單株均呈現雄不稔（表 2-5），即無任何一個雜交品系呈現稔性分離，綜整 progeny test A 及 progeny test B 的結果，顯示雄不稔株（NH80-A）與恢復系（TA95）雜交組合之基因型，可能為雜交組合編號 1 之 *BrMs<sup>b</sup>Ms<sup>b</sup> × BrMs<sup>a</sup>Ms<sup>a</sup>* 或雜交組合編號 2 之 *BrMsMs rfjf × BrMsMs RfRf*（附錄 2、3）。

progeny test C 主要是利用恢復系與維持系進行雜交，由其後裔 F<sub>2</sub> 次族群是否產生稔性分離情形，以區別雄不稔基因間之等位性。在 WH606×TA95 的組合中，共產生 25 個 F<sub>2</sub> 次族群，但所有後裔植株稔性均表現正常，亦即無產生稔性分離的 F<sub>2</sub> 次族群（表 2-6），顯示 *BrMs<sup>a</sup>/BrMs<sup>c</sup>*（或 *BrMs/BrRf*）間具有等位性（附錄 4），即此兩個 alleles 位於同一個基因座上。

表 2-4 不結球白菜 NH80-A $\times$ TA95 之 F<sub>2</sub> 族群於 progeny test A 的稔性分離情形

F <sub>2</sub> subpopulation	Total No.	ms plants	mf plants	$\chi^2 P_{(1:3)}$	$\chi^2 P_{(3:13)}$
13-B701	96	22	74	0.64	0.30
13-B702	93	19	74	0.31	0.68
13-B703	27	3	24	0.10	0.31
13-B704 <sup>x</sup>	88	14	74	0.05	0.49
13-B705 <sup>y</sup>	92	27	65	0.34	0.01
13-B706	72	12	60	0.10	0.65
13-B707	88	17	71	0.22	0.89
13-B708	74	15	59	0.35	0.74
13-B709	74	12	62	0.08	0.58
13-B710	102	21	81	0.30	0.63
13-B711	96	17	79	0.10	0.79
13-B712	79	16	63	0.33	0.73
13-B713	85	19	66	0.57	0.39
13-B714	83	19	64	0.66	0.33
13-B715	69	15	54	0.53	0.52
13-B716 <sup>y</sup>	94	26	68	0.55	0.03
13-B717	67	14	53	0.44	0.65
13-B718	94	19	75	0.28	0.72
13-B719 <sup>x</sup>	100	16	84	0.04	0.48
13-B720	83	16	67	0.23	0.90
13-B721	100	25	75	1.00	0.11
13-B722 <sup>x</sup>	106	16	90	0.02	0.33
13-B723 <sup>y</sup>	97	27	70	0.52	0.02
13-B724	74	17	57	0.69	0.35
13-B725 <sup>y</sup>	84	25	59	0.31	0.01
13-B726	77	12	65	0.06	0.48
13-B727 <sup>x</sup>	95	15	80	0.04	0.46
13-B728	75	13	62	0.13	0.75
13-B729	78	17	61	0.51	0.49
13-B730 <sup>x</sup>	61	8	53	0.03	0.26
13-B731	72	16	56	0.59	0.45
13-B732	80	19	61	0.80	0.25
13-B733 <sup>x</sup>	89	12	77	0.01	0.20
13-B734	96	22	74	0.64	0.30
13-B735	64	13	51	0.39	0.75
13-B736	77	18	59	0.74	0.30

表 2-4 (續)

F <sub>2</sub> subpopulation	Total No.	ms plants	mf plants	$\chi^2 P_{(1:3)}$	$\chi^2 P_{(3:13)}$
13-B737 <sup>x</sup>	85	10	75	0.005	0.10
13-B738	84	19	65	0.61	0.36
13-B739	75	16	59	0.46	0.57
13-B740	77	17	60	0.55	0.45
13-B741	55	13	42	0.82	0.35
13-B742	83	15	68	0.14	0.87
13-B743	85	17	68	0.29	0.77
13-B744	81	19	62	0.75	0.28
13-B745	95	19	76	0.26	0.75
13-B746	80	19	61	0.80	0.25
13-B748	79	21	58	0.75	0.07
13-B749	55	12	43	0.59	0.56
13-B750	46	6	40	0.06	0.32
13-B751 <sup>y</sup>	91	25	66	0.59	0.03
13-B752	78	19	59	0.90	0.20
13-B753	70	17	53	0.89	0.24
13-B754 <sup>x</sup>	78	9	69	0.01	0.10
13-B755 <sup>y</sup>	83	25	58	0.28	0.01
13-B756	96	23	73	0.81	0.19
13-B757	79	19	60	0.85	0.23
13-B758	66	13	53	0.32	0.84
13-B759	91	24	67	0.76	0.06
13-B760	95	19	76	0.26	0.75
13-B761	41	10	31	0.93	0.35
13-B762	66	11	55	0.12	0.66
13-B763	70	15	55	0.49	0.57
13-B764	80	18	62	0.61	0.39
13-B765	95	23	72	0.86	0.17

ms: male sterile ; mf: male fertile

<sup>x</sup> 表示該 F<sub>2</sub> 次族群之稔性表現不符合 1:3 分離比

<sup>y</sup> 表示該 F<sub>2</sub> 次族群之稔性表現不符合 3:13 分離比

表 2-5 不結球白菜 NH80-A×TA95 之雜交後裔於 progeny test B 的稔性表現

F <sub>1</sub> progeny	ms plants	mf plants	F <sub>1</sub> progeny	ms plants	mf plants
14-B301	31	0	14-B321	9	0
14-B302	13	0	14-B322	12	0
14-B303	27	0	14-B323	27	0
14-B304	28	0	14-B324	12	0
14-B305	26	0	14-B325	27	0
14-B306	24	0	14-B326	24	0
14-B307	28	0	14-B327	28	0
14-B308	19	0	14-B328	11	0
14-B309	24	0	14-B329	26	0
14-B310	28	0	14-B330	12	0
14-B311	28	0	14-B331	26	0
14-B312	20	0	14-B332	18	0
14-B313	21	0	14-B333	14	0
14-B314	27	0	14-B334	27	0
14-B315	26	0	14-B335	25	0
14-B316	29	0	14-B336	29	0
14-B317	20	0	14-B337	26	0
14-B318	22	0	14-B338	28	0
14-B319	28	0	14-B339	26	0
14-B320	16	0	14-B340	26	0

ms: male sterile ; mf: male fertile

表 2-6 不結球白菜 WH606×TA95 之 F<sub>2</sub> 族群於 progeny test C 的稔性表現

F <sub>2</sub> 品系	ms plants	mf plants	F <sub>2</sub> 品系	ms plants	mf plants
12-B326	0	74	12-B339	0	77
12-B327	0	76	12-B340	0	69
12-B328	0	79	12-B341	0	80
12-B329	0	76	12-B342	0	73
12-B330	0	79	12-B343	0	77
12-B331	0	73	12-B344	0	74
12-B332	0	76	12-B345	0	66
12-B333	0	78	12-B346	0	67
12-B334	0	64	12-B347	0	80
12-B335	0	79	12-B348	0	78
12-B336	0	75	12-B349	0	68
12-B337	0	76	12-B350	0	74
12-B338	0	79			

ms: male sterile ; mf: male fertile

整合三種 progeny test 的結果如圖 2-3 所示，progeny test A 所有後裔品系之稔性均呈現分離，progeny test B 所有品系均為雄不稔而無分離，而 progeny test C 所有後裔均為雄可稔亦無分離情形，對照至附錄 2~4 之遺傳模式外表型理論值，本試驗之雄不稔系、恢復系及維持系間，雄不稔性狀的遺傳符合單一基因座複等位基因模式，等位基因的顯隱性關係為  $BrMs^a > BrMs^b > BrMs^c$ ，可推測雄不稔系中，參試的雄不稔株 NH80-A 基因型為  $BrMs^bMs^b$ ，可稔株 NH80-B 基因型為  $BrMs^aMs^b$ ，恢復系 TA95 基因型為  $BrMs^aMs^a$ ，而維持系 WH606 基因型則為  $BrMs^cMs^c$ 。



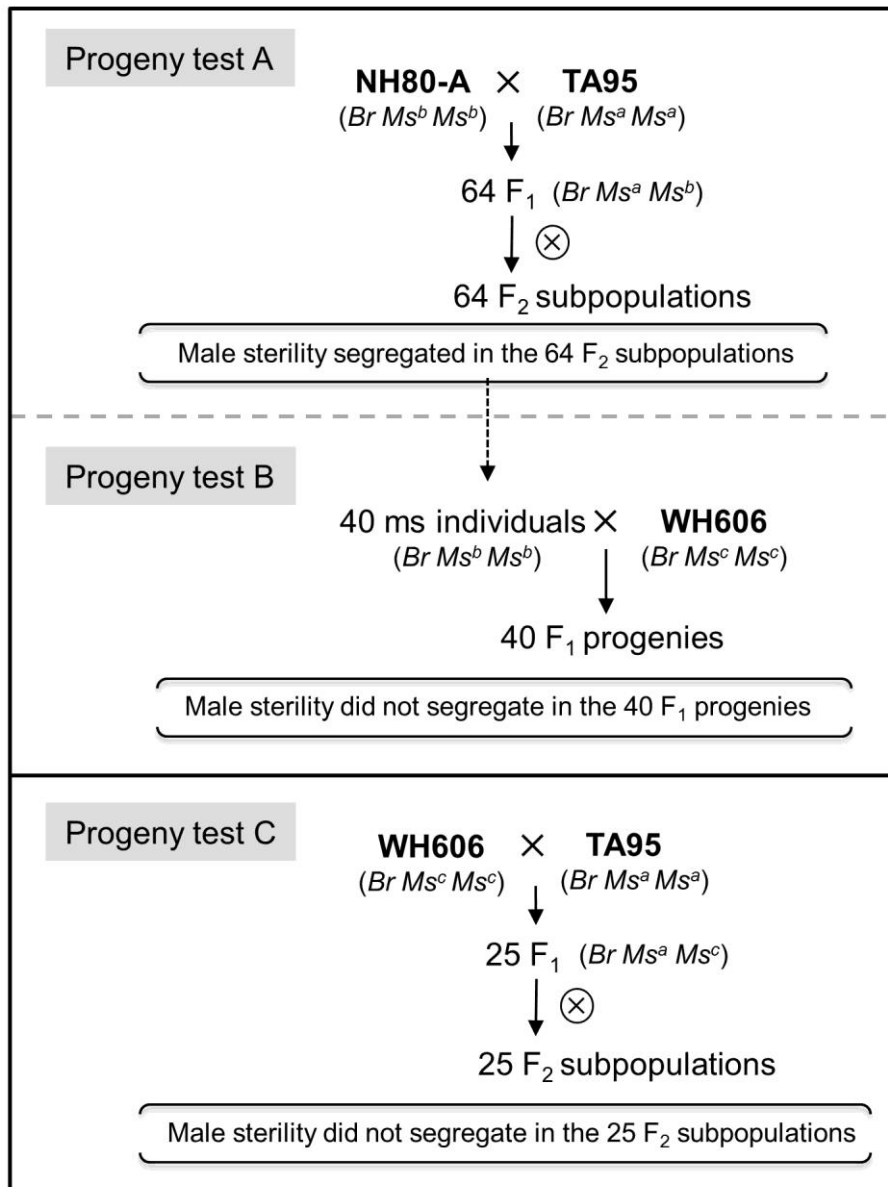


圖 2-3 第一雜交群組於三種後裔檢定之結果



### 3.2 以 SG388 為恢復系之第二雜交群組後裔檢定試驗

為了分析 NH80-A 雄不稔性狀的遺傳行為，以另一個恢復系 SG388 取代 TA95 進行相同之雜交組合及後裔檢定。SG388 與 TA95 屬不結球白菜之二個不同變種，SG388 源自中國南方食用花苔的蔬菜品種，而有別於源自泰國之葉用蔬菜 TA95。結果顯示 SG388 及 WH606 的自交後裔、NH80-A × SG388 及 WH606 × SG388 的 F<sub>1</sub> 後裔均具有正常稔性，而 NH80-A × WH606 之 F<sub>1</sub> 後裔均為雄不稔（表 2-7）。

表 2-7 第二雜交群組後裔之稔性表現

Cross	No. of male sterile plants	No. of male fertile plants
NH80-A × SG388	0	50
NH80-A × WH606	40	0
WH606 × SG388	0	42
SG388⊗	0	68
WH606⊗	0	70

NH80-A × WH606 之所有 F<sub>1</sub> 後裔均為雄不稔，顯示第二雜交群組中參試的雄不稔 NH80-A 單株，其雄不稔基因對維持系而言為顯性同質結合體，而該雄不稔單株與恢復系 SG388 的 F<sub>1</sub> 後裔中，所有植株均表現為雄可稔，顯示恢復系 SG388 的稔性恢復基因對參試的 NH80-A 單株而言也為顯性同質結合體，因此在二種雄不稔性遺傳模式定義下，依據第一年品系間之雜交結果，參試雄不稔株於複等位基因模式下基因型應為 *BrMs<sup>b</sup>Ms<sup>b</sup>*，若在二對顯性上位基因模式下則應為 *BrMsMs rfrf*；維持系 WH606 基因型在複等位基因模式為 *BrMs<sup>c</sup>Ms<sup>c</sup>*，在二對顯性上位基因模式為 *Brmsms rfrf*；而恢復系 SG388 基因型在複等位基因模式為 *BrMs<sup>a</sup>Ms<sup>a</sup>*，在二對顯性上位基因模式為 *BrMsMs RfRf*、*BrMsms RfRf*、*Brmsms RfRf*（附錄 1）。

在 progeny test A 中，雄不稔株與恢復系雜交後裔之所有 F<sub>2</sub> 次族群均呈現稔性分離情形，在 NH80-A × SG388 的雜交組合中，共獲得 43 個 F<sub>2</sub> 次族群（表 2-8），其中同時符合 1：3 與 3：13 分離比者達 34 個，有 5 個次族群雄不稔單株與可稔單株之比例經卡方檢定符合 3：13 之分離比但推翻 1：3 分離比，此外，也有 4 個次族群經檢定只符合 1：3 分離比，然而所有 F<sub>2</sub> 次族群中並無產生既不符合 1：3 也不符合 3：13 分離比者。

表 2-8 不結球白菜 NH80-A×SG388 之 F<sub>2</sub> 族群於 progeny test A 的稔性分離情形

F <sub>2</sub> subpopulation	Total No.	ms plants	mf plants	$\chi^2 P_{(1:3)}$	$\chi^2 P_{(3:13)}$
11-B31 <sup>x</sup>	63	8	55	0.02	0.22
11-B32	81	17	64	0.40	0.61
11-B33	21	5	16	0.90	0.55
11-B34	68	16	52	0.78	0.31
11-B35 <sup>x</sup>	71	9	62	0.02	0.19
11-B36	59	16	43	0.71	0.10
11-B37 <sup>x</sup>	82	11	71	0.02	0.22
11-B38	47	12	35	0.93	0.23
11-B39	58	11	47	0.29	0.97
11-B40	54	10	44	0.27	0.97
11-B41	63	13	50	0.42	0.70
11-B42	65	15	50	0.72	0.37
11-B43	52	11	41	0.52	0.66
11-B45 <sup>y</sup>	82	23	59	0.52	0.03
11-B46	13	5	8	0.26	0.07
11-B47 <sup>x</sup>	66	9	57	0.03	0.29
11-B48	60	13	47	0.55	0.56
11-B49	83	19	64	0.66	0.33
11-B50	61	16	45	0.82	0.13
11-B51	23	2	21	0.07	0.22
11-B52	64	13	51	0.39	0.75
11-B53	74	16	58	0.50	0.53
11-B54 <sup>x</sup>	78	12	66	0.05	0.45
11-B55	89	18	71	0.30	0.72
11-B56	60	16	44	0.77	0.12
11-B57	67	12	55	0.18	0.86
11-B58	61	9	52	0.06	0.42
11-B59 <sup>y</sup>	76	21	55	0.60	0.05
11-B60	63	11	52	0.17	0.79
11-B61 <sup>y</sup>	92	28	64	0.23	0.004
11-B62	71	16	55	0.63	0.41
11-B63	54	15	39	0.64	0.09
11-B64	82	14	68	0.10	0.70
11-B65	58	14	44	0.88	0.29
11-B66	19	5	14	0.89	0.40

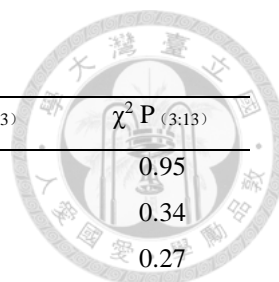


表 2-8 (續)

F <sub>2</sub> subpopulation	Total No.	ms plants	mf plants	$\chi^2 P_{(1:3)}$	$\chi^2 P_{(3:13)}$
11-B67	65	12	53	0.22	0.95
11-B68	27	7	20	0.91	0.34
11-B69	35	4	31	0.06	0.27
11-B70	54	10	44	0.27	0.97
11-B71	54	9	45	0.16	0.69
11-B72	43	10	33	0.79	0.45
11-B73 <sup>y</sup>	68	22	46	0.16	0.004
11-B74	60	14	46	0.77	0.36

ms: male sterile ; mf: male fertile

<sup>x</sup> 表示該 F<sub>2</sub> 次族群之稔性表現不符合 1:3 分離比

<sup>y</sup> 表示該 F<sub>2</sub> 次族群之稔性表現不符合 3:13 分離比

在 progeny test B 之試驗中，主要由 NH80-A × SG388 的 F<sub>2</sub> 次族群中隨機選取 37 個雄不稔單株與維持系 WH606 雜交，並衍生出 37 個 F<sub>1</sub> 品系，這些品系的所有單株均呈現雄不稔（表 2-9），即無任何一個雜交品系呈現稔性分離，綜整 progeny test A 及 progeny test B 的結果，顯示雄不稔株（NH80-A）與恢復系（SG388）雜交組合之基因型，可能為雜交組合編號 1 之  $BrMs^bMs^b \times BrMs^aMs^a$  或雜交組合編號 2 之  $BrMsMs\ rfrf \times BrMsMs\ RfRf$ （附錄 2、3）。

progeny test C 主要是區別恢復系與維持系的雄不稔基因間之等位性，在 WH606 × SG388 的組合中，共產生 25 個 F<sub>2</sub> 次族群，但這些次族群的所有後裔稔性均表現正常，亦即無產生稔性分離的 F<sub>2</sub> 次族群（表 2-10），顯示  $BrMs^a/BrMs^c$ （或  $BrMs/BrRf$ ）間具有等位性（附錄 4）。

整合三種 progeny test 的結果如圖 2-4 所示，progeny test A 所有後裔品系稔性均分離、progeny test B 所有品系均為雄不稔而無分離，而 progeny test C 所有後裔均為雄可稔亦無分離情形，對照附錄 2~4 遺傳模式之外表型理論值，本試驗第二雜交群組參試材料間，雄不稔性狀的遺傳亦符合單一基因座複等位基因模式，等位基因的顯隱性關係為  $BrMs^a > BrMs^b > BrMs^c$ ，可推測雄不稔系中雄不稔株 NH80-A 基因型為  $BrMs^bMs^b$ ，可稔株 NH80-B 基因型為  $BrMs^aMs^b$ ，恢復系 SG388 基因型為  $BrMs^aMs^a$ ，而維持系 WH606 基因型則為  $BrMs^cMs^c$ 。

表 2-9 不結球白菜 NH80-A×SG388 之雜交後裔於 progeny test B 的稔性表現

F <sub>1</sub> progeny	ms plants	mf plants	F <sub>1</sub> progeny	ms plants	mf plants
12-B201	13	0	12-B220	17	0
12-B202	19	0	12-B221	12	0
12-B203	19	0	12-B222	21	0
12-B204	21	0	12-B223	18	0
12-B205	22	0	12-B224	5	0
12-B206	12	0	12-B225	17	0
12-B207	15	0	12-B226	22	0
12-B208	10	0	12-B227	11	0
12-B209	11	0	12-B228	19	0
12-B210	16	0	12-B229	20	0
12-B211	24	0	12-B230	23	0
12-B212	12	0	12-B231	24	0
12-B213	21	0	12-B232	18	0
12-B214	15	0	12-B233	22	0
12-B215	20	0	12-B234	13	0
12-B216	24	0	12-B235	23	0
12-B217	23	0	12-B236	24	0
12-B218	10	0	12-B237	21	0
12-B219	13	0			

ms: male sterile ; mf: male fertile

表 2-10 不結球白菜 WH606×SG388 之 F<sub>2</sub> 族群於 progeny test C 的稔性表現

F <sub>2</sub> 品系	ms plants	mf plants	F <sub>2</sub> 品系	ms plants	mf plants
12-B301	0	77	12-B314	0	61
12-B302	0	69	12-B315	0	63
12-B303	0	66	12-B316	0	36
12-B304	0	59	12-B317	0	68
12-B305	0	76	12-B318	0	61
12-B306	0	63	12-B319	0	71
12-B307	0	77	12-B320	0	61
12-B308	0	60	12-B321	0	66
12-B309	0	72	12-B322	0	77
12-B310	0	60	12-B323	0	72
12-B311	0	52	12-B324	0	78
12-B312	0	66	12-B325	0	64
12-B313	0	67			

ms: male sterile ; mf: male fertile

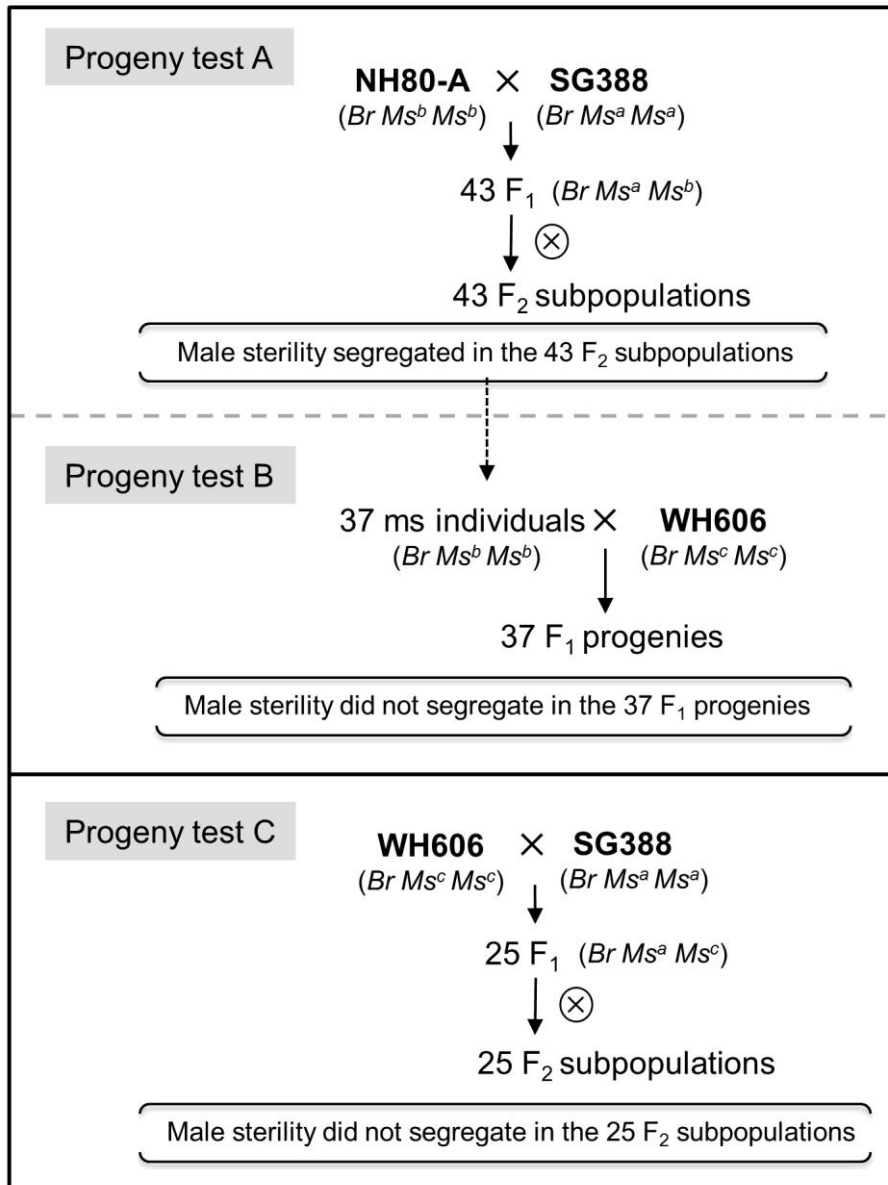


圖 2-4 第二雜交群組於三種後裔檢定之結果



## 4 討論

### 4.1 測交分析之方法

本試驗利用雄不稔單株、恢復系及維持系等基因型植株，以 progeny test A、progeny test B 及 progeny test C 等三種後裔檢定設計（劉，1992）分析其後裔稔性之表現，結果顯示本研究小油菜之雄不稔性狀，極可能為同一基因座上之不同等位基因控制性狀之表現，等位基因顯隱性關係為  $BrMs^a > BrMs^b > BrMs^c$ ，當基因型為  $BrMs^bMs^b$  或  $BrMs^bMs^c$  時表現雄不稔。

本試驗參試的不結球白菜分屬相同亞種下的三個變種，其中雄不稔系為 *B. rapa* ssp. *chinensis* var. *oleifera* Makino，恢復系 TA95 及維持系 WH606 為 *B. rapa* ssp. *chinensis* var. *communis* Tsen et Lee，另一個恢復系 SG388 則為 *B. rapa* ssp. *chinensis* var. *parachinensis* Bailey，此三個變種因分化而有形態上及用途上之差異，然而雄不稔等位基因型並未因分化而有所不同，如 TA95 與 SG388 雖為不同變種，但基因型一致，因此推測在不結球白菜的共同祖先已存在不同雄不稔等位基因，由於不結球白菜還有許多其他類型之變種，若能納入更多不同類型之不結球白菜進行雄不稔等位基因之遺傳分析，將能證實此一推論，此外，屬另一個亞種 ssp. *pekinensis* 之結球白菜，其是否具有相同雄不稔基因值得進一步探討。

一般進行遺傳分析時常以雄不稔株與恢復系雜交後，利用卡方檢定分析其  $F_2$  族群中雄不稔株與可稔株之分離比，當該性狀受單基因座控制時分離比為 1:3（不稔：可稔），而受二基因座控制時為 3:13（不稔：可稔），然而此二種外表型分離比相近，不易以卡方測驗驗證屬於何種模式，如本研究的 progeny test A 試驗，不論是以 TA95 為恢復系的第一雜交群組或是 SG388 為恢復系的第二雜交群組，約 80% 次族群（第一雜交群組 50/64，第二雜交群組 34/43）同時符合 1:3 與 3:13 分離比，其餘 20% 次族群有些檢定為符合 1:3，有些則符合 3:13，因此，若只參考  $F_2$  分離比的結果難以確認遺傳模式，故本試驗搭配不同基因型間之後裔檢定試驗，且 progeny test A、progeny test B 及 progeny test C 等三種後裔檢定主要都是檢測次族群或雜交後裔是否產生稔性分離，而非檢測稔性分離的比率，因此能更容易明確地判別各後裔檢定之結果，儘管需進行較多次雜交工作，然綜合三種測交後裔結果更能有效分析出實際遺傳模式。



## 4.2 大油菜與大白菜細胞核雄不稔遺傳之比較

在蕓苔屬白菜類作物中，首先於結球白菜中發現受複等位基因控制之雄不稔性 (Feng et al., 1996; Feng et al., 2009)，而在其近緣物種大油菜中，也在多個不同的遺傳資源上發現同基因座上之超過二個等位基因控制之雄不稔性狀 (Song et al., 2006; Liu et al., 2008; Xie et al., 2012)，本研究則是首次於不結球白菜上證實此種遺傳模式，顯示細胞核雄不稔性受複等位基因遺傳之模式是普遍存在於蕓苔屬作物中的，此與 Lu et al., (2013) 之觀點一致，由於等位基因間具有顯隱性關係，當以部分等位基因進行遺傳分析時，就可能產生不同遺傳模式之結論，例如本研究若僅以雄不稔株 NH-80A 與恢復系 TA95 進行遺傳分析時，結果將會推論該性狀是受一對隱性基因控制，反之若僅以雄不稔株 NH-80A 與維持系 WH606 進行分析則結果則呈現單一顯性雄不稔基因遺傳，至今十字花科作物已發現許多受單基因控制的細胞核雄不稔性 (Nieuwhof, 1961; Cao and Li, 1981; Van der Meer, 1987; Ruffio-Chable et al., 1993; Takahata et al., 1996; Fang et al., 1997; Ying et al., 2003; Zhang et al., 2010; Kenji et al., 2013; Li et al., 2016) 等，然而以往這些研究大都以少數品系進行雜交，而分析其遺傳模式，若能將發現之雄不稔系與廣泛地與眾多種原進行測交，將很可能發現同基因座上更多之其他等位基因參與雄不稔性狀之控制。

## 4.3 遺傳行為對爾後育種之差異

在作物育種上，利用細胞質雄不稔性進行雜交種子之生產已相當普遍，然而部份實例指出，CMS 細胞質因與不良性狀基因連鎖導致利用上受到限制，由於細胞核雄不稔性具有不稔性完全、較無異源細胞質產生之副作用 (Ke et al., 2005)，因此為雜交種子生產提供另一個良好途徑，由於單隱性或雙隱性核基因控制之雄不稔性均需於不稔系中拔除約 50% 可稔株，而限制其發展，然而在複等位基因模式或雙顯性基因模式下，利用維持系與雄不稔系中之同質結合的雄不稔株雜交即可獲得 100% 雄不稔後裔，故以其作為雜交種之母本極具應用潛力，此二種遺傳模式均適合作為商業雜交種子生產用，但於育種過程進行性狀轉移時，複等位基因模式只需考慮單基因座之 6 種基因型，而比雙顯性基因兩個基因座之 9 種基因型更為簡便。





第3章 以雄不稔株 × 恢復系之 F<sub>2</sub> 族群建立遺傳連鎖圖定位  
*BrMs* 基因

## 1 前言

本研究第二章針對小油菜細胞核雄不稔源進行遺傳分析的結果，顯示該雄不稔性符合單基因座上，受複等位基因控制的遺傳模式，這種遺傳模式具實務應用之潛力。然而，關於該基因座實際位於何染色體上、是否與其他性狀連鎖、利用周邊序列開發分子標誌輔助選拔等，則是育種上迫切需求的資訊，因此次階段研究上若能更進一步定位出該基因之所在範圍，不僅能將遺傳分析的結論轉為視覺化之描述，相關資訊更有助於剖析該雄不稔源與前人研究結果之異同，由於蕓苔屬之大油菜其複二倍體基因組 (AACC) 中的 A 基因組與白菜基因組 (AA) 同源，因此定位出 A 基因組上雄不稔基因之所在深具意義。

為了輔助育種或進行基因選殖 (positional cloning) 等目的，在蕓苔屬作物細胞核雄不稔基因的研究上，已開發許多與 GMS 基因連鎖的分子標誌，透過這些分子標誌及遺傳連鎖圖之建立，目前大油菜有五個雄不稔基因被定位於 A7, N16, N19 及 A8 等連鎖群 (Yi et al., 2006; Huang et al., 2007; Lei et al., 2007; Xiao et al., 2008; Lu et al., 2013)，而結球白菜的單一隱性基因與複等位基因座分別定位於 A8 及 A7 連鎖群 (Feng et al., 2009; Zhang et al., 2010)，不結球白菜的單一隱性基因則定位於 A2 連鎖群 (Li et al., 2016)。儘管有許多雄不稔基因被定位於染色體的特定區域，但只有少數雄不稔基因恰巧透過相同分子標誌之連鎖關係而證實其等位性 (Zu et al., 2010; Xie et al., 2012)，仍有許多雄不稔基因其連鎖圖上的分子標誌密度不夠高、於分離族群間不具共同多型性標誌，或缺乏實際物理位置可供比對等因素，導致於不同遺傳背景下所發現的雄不稔基因難以比較其異同。

隨著次世代定序技術之應用，為了簡化基因組的複雜性而僅針對基因組特定區域進行定序之 reduced-representation sequencing (RRS) 是獲得大量分子標誌的有效方法，而 genotyping-by-sequencing (GBS) 則是基於相同策略的基因分型技術之一 (Elshire et al., 2011)，當有作物的參考序列 (reference genome) 時，該技術透過對不同個體之 DNA 條碼可將不同個體同時進行基因分型，使得 GBS 可用快速且低成本之方式偵測大量包含物理位置資訊的 SNPs，以建構高密度遺傳連鎖圖，進而應用於重要性狀之 QTL mapping，目前已普遍應用於許多作物如玉米、小麥、大麥、甘藍及甜瓜等 (Elshire et al., 2011; Poland et al., 2012; Lee et al., 2015;

Chang et al., 2017) , 藉此技術結合已建構完成之白菜基因體資訊應可有效率的定位基因，而基因定位所獲得之物理位置資訊，亦將有助於系統性地研究芸苔屬作物細胞核雄不稔基因間之關係。



## 2 材料與方法

### 2.1 建構基因定位分離族群

對於  $BrMs^a/BrMs^b$  之定位，2013 年 2 月將雄不稔株 NH80-A 與恢復系 TA95 進行雜交，同年 10 月將其 F<sub>1</sub> 植株自交以獲得 F<sub>2</sub> 種子，2014 年 2 月進一步培育 88 株 F<sub>2</sub> 植株做為基因定位用族群，所有植株以盆栽種植 30 天後取葉片萃取 DNA 以建構 GBS 用之 DNA library，各單株透過 GBS 程序分析基因型，而外表型則於開花期調查各單株之雄不稔性狀後而獲得。

### 2.2 GBS (genotyping-by-sequencing) 之進行

#### 2.2.1 建構 GBS library

GBS libraries 之建構主要參考 Poland et al. (2012) 之方法而略加修飾，每一植株取約 200 ng DNA 後同時以 *Pst*I-HF 及 *Msp*I (New England Biolabs (NEB), USA) 二種內切核酸酶於 37 °C 下反應 4 小時後，再於 65 °C 維持 20 分鐘，酶切完成之 DNA 片段以 barcoded adapters 及 reverse Y-adaptor (Poland et al., 2012) 進行連接 (ligation)，各樣本連接反應體積為 40 μL，包含酶切 DNA (200 ng)、barcoded adapter (20 nM)、reverse Y-adaptor (300 nM)、T4 DNA ligase (200 U)、1X NEB Buffer 4 (NEB, USA) 及 ribo ATP (1 mM)，於 22 °C 下進行 2 小時後再於 65 °C 維持 20 分鐘，完成 adapters 連接之 DNA 各取 15 μL 混合成 DNA library，隨後用 AMPure XP 磁珠 (Beckman Coulter, USA) 以 DNA:磁珠為 1:1.8 的比例進行純化，最後利用經修飾之 Illumina 引子對 (Poland et al., 2012) 於 95 °C (30 sec)，62 °C (25 sec)，68 °C (40 sec) 條件下進行 PCR 反應 18 個 cycles，建庫完成的 GBS library 則由國立陽明大學基因體研究中心以 Illumina HiSeq 2000 定序平台 (Illumina Inc., SA) 進行次世代定序 (Next Generation Sequencing, NGS)，定序方式為進行 100 bp 的單端讀序 (single-end reads)。



## 2.2.2 序列分析 (raw sequence data processing)

本研究以下載自 Brassica database (BRAD) (<http://brassicadb.org/brad/index.php>) 的 *B. rapa* var. chiifu40 (v1.5) 為參考序列 (Cheng et al., 2011)，經次世代定序獲得之讀序 (reads) 以軟體 CLC Genomics Workbench v8.0.2 (CLC Bio, <http://www.clcbio.com>) 篩選序列品質分數 (quality scores) 大於 Q20 者，依據各樣品 DNA 條碼進行分類，去除條碼序列的讀序，隨之比對 (alignment) 至參考序列並以 sam 的檔案格式輸出，而基因座暨基因座目錄之建立、SNP calling、分離後裔之基因分型等則以軟體 Stacks v1.37 (<http://creskolab.uoregon.edu/stacks/>) (Catchen et al., 2011) 之“ref\_map.pl”指令進行，各讀序中的專一性序列 (unique sequence) 被定義為 tag，只有讀序深度 (depth)  $\geq 5$ ，且於族群內之缺失率 (missing rate) 少於 20% 的 tags，才會輸出成為基因型資料。

## 2.3 連鎖分析 (linkage analysis)

族群的基因型資料以 R/qlt 程式建立遺傳連鎖圖 (Broman et al., 2003; Broman, 2010)。簡言之，將個體或 tags 缺值大於 20% 之資料捨棄後，篩選出未偏離常態分布的 tags，以重組率 (recombination fraction)  $\leq 0.3$  且 LOD 值  $\geq 7$  之條件分類為相同連鎖群，並以 Kosambi 函數 (Kosambi, 1943) 估計遺傳距離。在建立初級連鎖圖之後，將族群內具有相同分離情形，或遺傳距離小於 0.5 cM 的 tags 合併為相同的 bin，由各 bin 中隨機選取一個 tag 做為 bin marker，再次計算 bins 間之遺傳距離以建構遺傳連鎖圖，該連鎖圖相鄰 bin 之間至少包含 1 次重組。此外，族群內所有個體之雄不稔性外表型資料，連同 bins 一併進行連鎖分析，並估計雄不稔基因與相鄰 bins 間之遺傳距離。

## 3 結果

### 3.1 GBS 的結果

NH80-A  $\times$  TA95 之 F<sub>2</sub> 族群被用來定位含有 *BrMs<sup>a</sup>* 及 *BrMs<sup>b</sup>* 等位基因之 *BrMs* 基因座，該 DNA library 經 Illumina 定序平台解序後產生約 25 Gb 序列資料，經定序品質篩選 (Quality Trim) 後共獲得 245 millions 短序列，平均長度為 93.6 bp，將這些序列與參考序列比對後，雄不稔母本 (NH-80A) 與恢復系父本 (TA95) 之成

功比對率 (mapping rate) 分別為 81.7% 與 41.9%，F<sub>2</sub> 族群整體比對率為 80.2%。

F<sub>2</sub> 族群中共有 33,632 個 tags 可比對至參考序列上，雙親間具多型性的 tags 有 5,284 個，經排除於個體間缺失率大於 20% 或不符合分離比率者，共獲得 2,243 個 tags 用於建構遺傳圖譜 (表 3-1)，將基因型相同的 tags 合併後，所獲得之遺傳圖譜係由 569 個 bins 組成之 10 個連鎖群 (圖 3-1)，其中包含 4,347 個 SNPs，連鎖圖總長 1,011 cM，其中以 A09 連鎖群最長達 135.8 cM，最短則為 A10 連鎖群僅 75.1 cM，各連鎖群 bins 的平均間隔的介於 1.6 cM 至 2.1 cM 之間，整體而言平均為 1.8 cM，每個 bin 有 3.5 至 5 個具多型性之 tags，整體平均有 3.9 個多型性 tags。進一步將各連鎖群之 bin markers 比對其位於參考序列上的物理位置，結果顯示在 569 個 bin markers 中，大部份 bin markers 順序與物理位置一致，僅有 18 個 bin markers 其順序與相鄰的 bins 之間有倒置 (inversion) 現象 (圖 3-2)。

### 3.2 *BrMs* 之基因定位 (gene mapping)

在 NH80-A × TA95 之 F<sub>2</sub> 族群所建構的連鎖圖中，*BrMs* 基因被定位於 Chr07 連鎖群 (圖 3-2)，Chr07 連鎖群全長約 82 cM，共含 51 個 bins，其中 bin “F2\_21293” 與 *BrMs* 基因共分離，兩側相鄰的二個 bins “F2\_21158” 與 “F2\_21316” 分別與該基因相距 1.8 及 0.6 cM，其物理圖 (physical map) 上的對應位置分別位於 Chr07 染色體之 6.0 Mb 及 7.6 Mb 附近 (圖 3-2、表 3-2)，顯示 *BrMs* 基因座落於此 1.6 Mb 的範圍內，在 bin “F2\_21293” 之位置尚有 “F2\_21230”、“F2\_21231”、“F2\_21236” 及 “F2\_21271” 等 4 個之 tags 與 *BrMs* 基因共分離，基因兩側的 bins “F2\_21158” 與 “F2\_21316” 則分別有 3 個及 2 個 tags 與之相鄰，這 10 個 tags 當中共含 15 個 SNP 位點 (表 3-2)，依據這些 tags 之物理位置可將 *BrMs* 座落位置再限縮於標定位置 6.6 Mb ~ 7.6 Mb 區間之 1.0 Mb 範圍內。



表 3-1 以 F<sub>2</sub> 族群建立之連鎖群長度、SNP 數量及 marker 密度

Characteristics	Linkage Groups										Total
	A01	A02	A03	A04	A05	A06	A07	A08	A09	A10	
Length (cM)	107.8	98.2	132.6	93.4	107	102.8	82	76.5	135.8	75.1	1,011
No. of SNP	384	468	527	306	400	572	495	335	584	276	4,347
No. of polymorphic Tag	206	233	263	176	209	279	256	185	288	148	2,243
No. of Bin	58	60	68	44	54	62	51	49	82	41	569
average SNPs (SNPs/Tag)	1.9	2.0	2.0	1.7	1.9	2.1	1.9	1.8	2.0	1.9	1.9
average Tags (Tags/bin)	3.6	3.9	3.9	4.0	3.9	4.5	5.0	3.8	3.5	3.6	3.9
average interval (cM/bin)	1.9	1.6	2.0	2.1	2.0	1.7	1.6	1.6	1.7	1.8	1.8

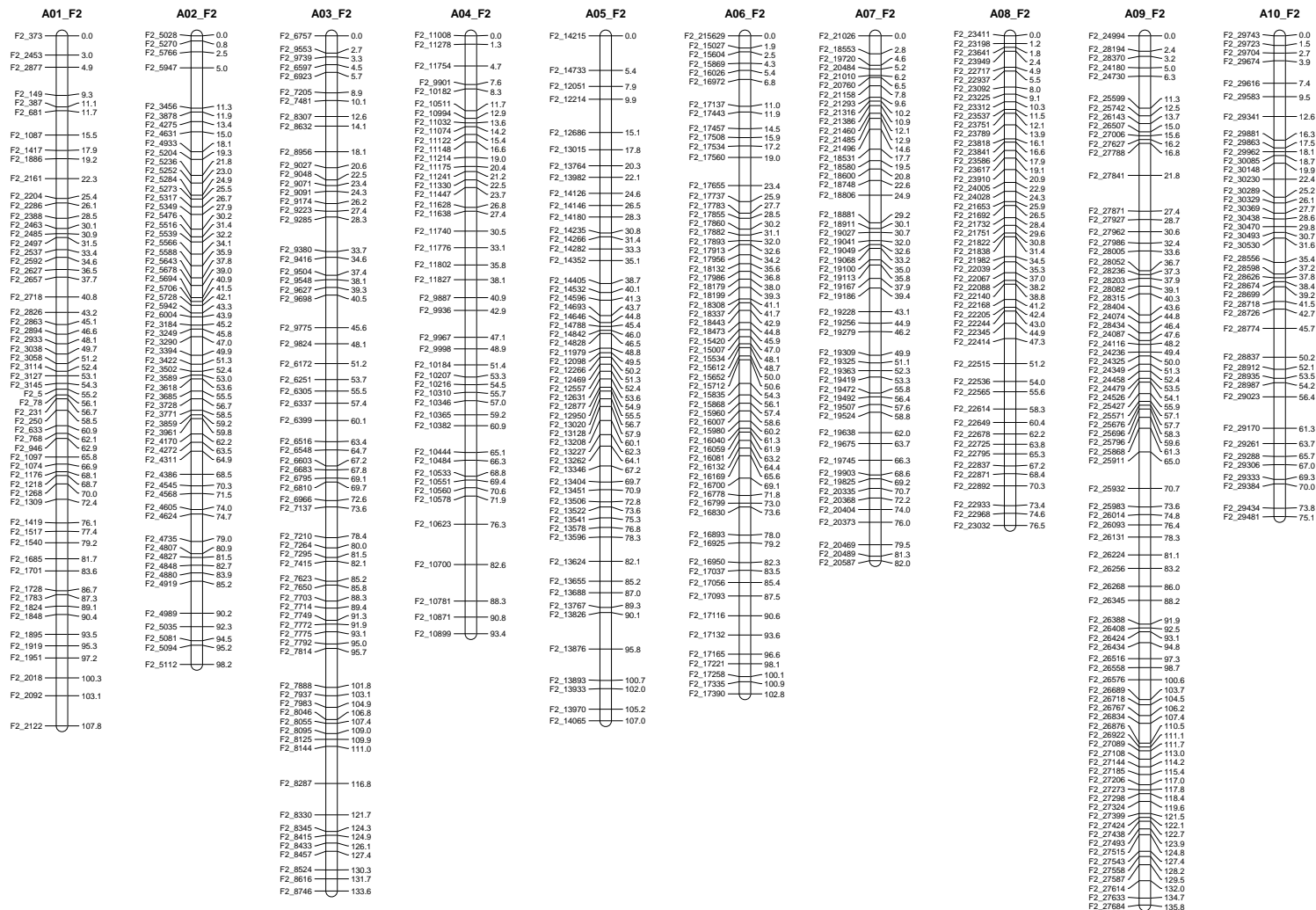
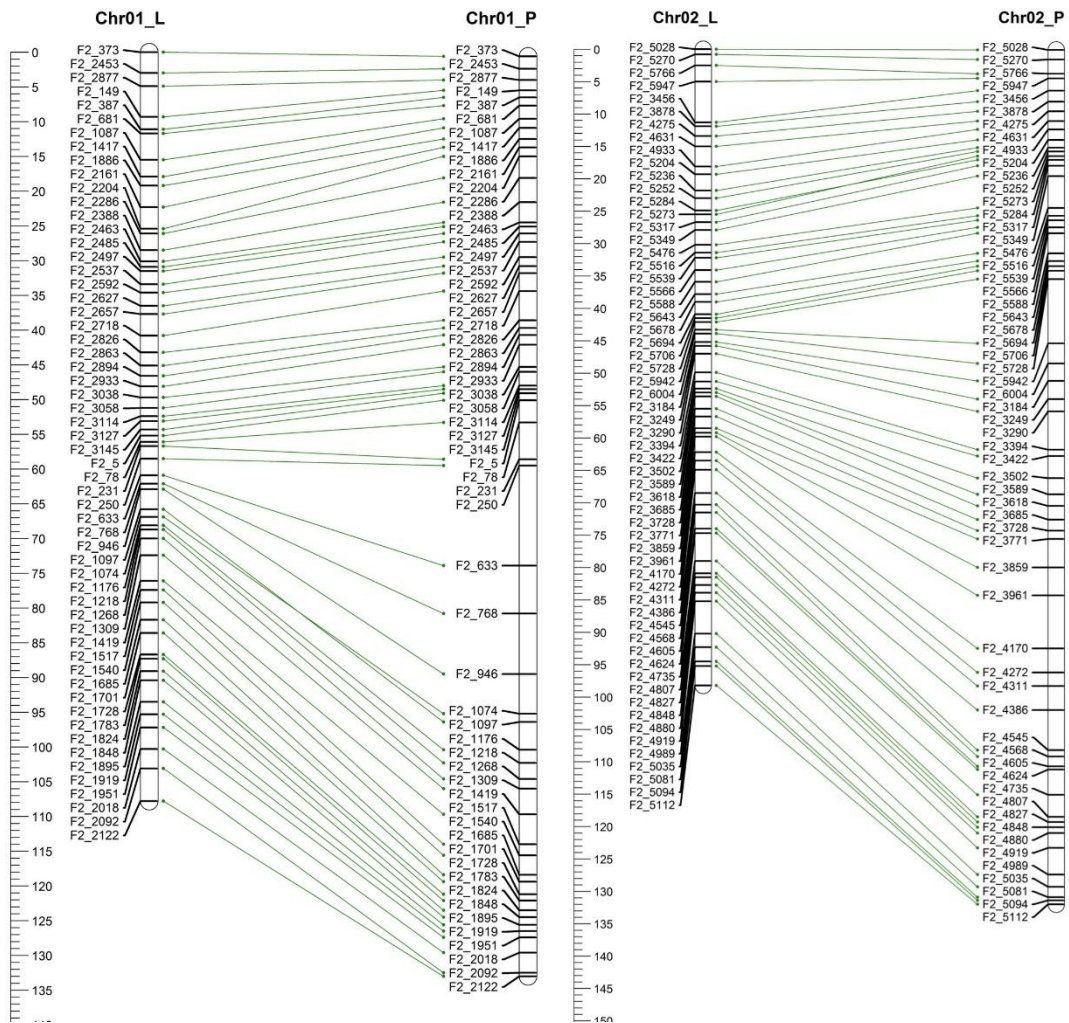
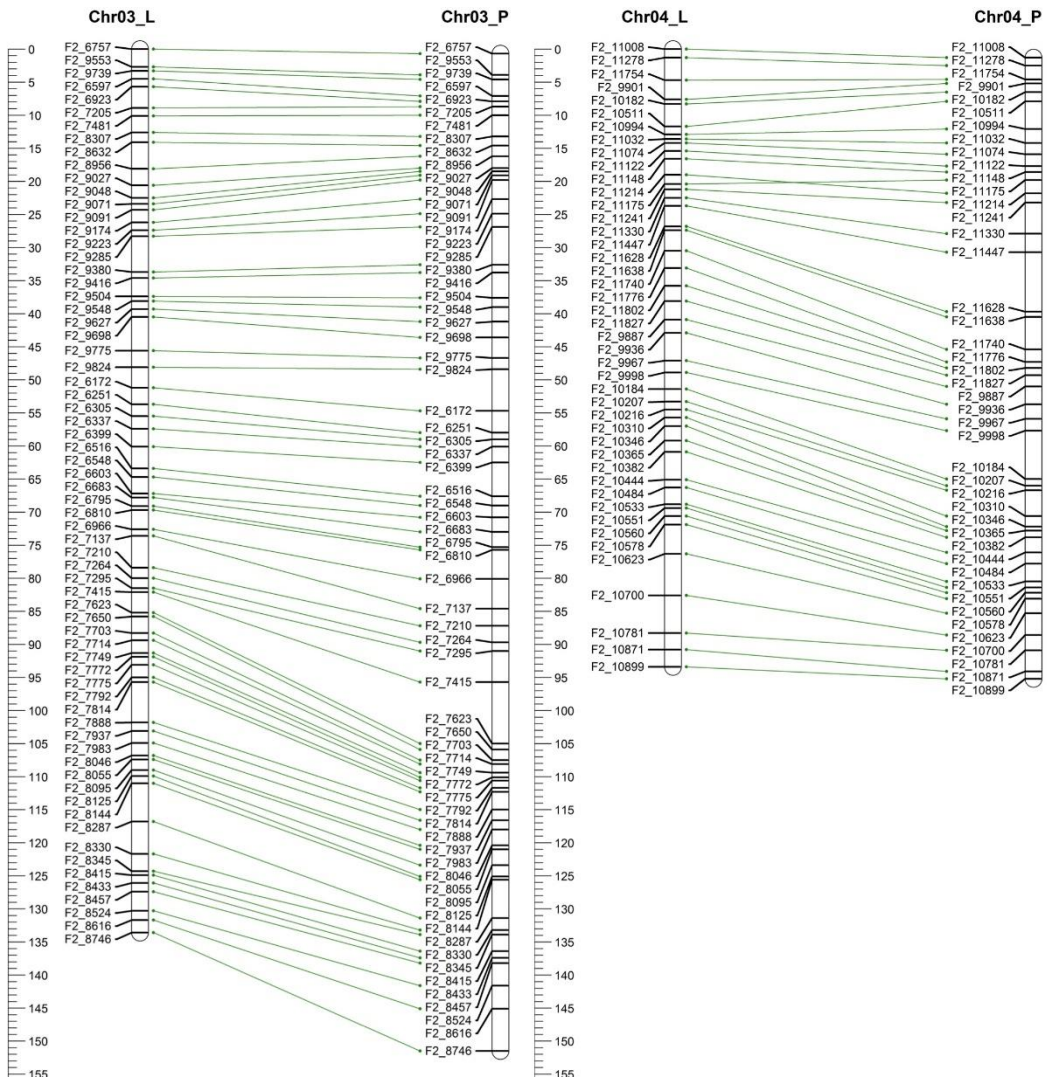
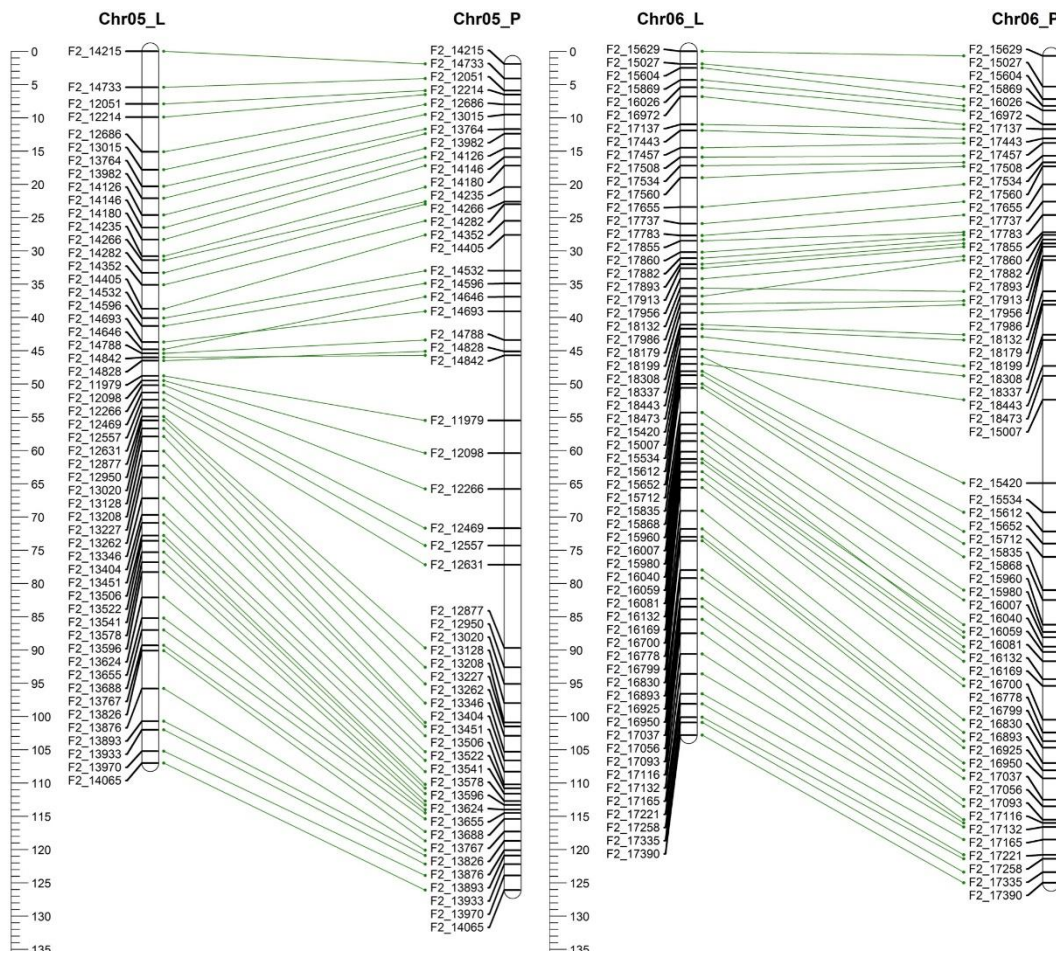


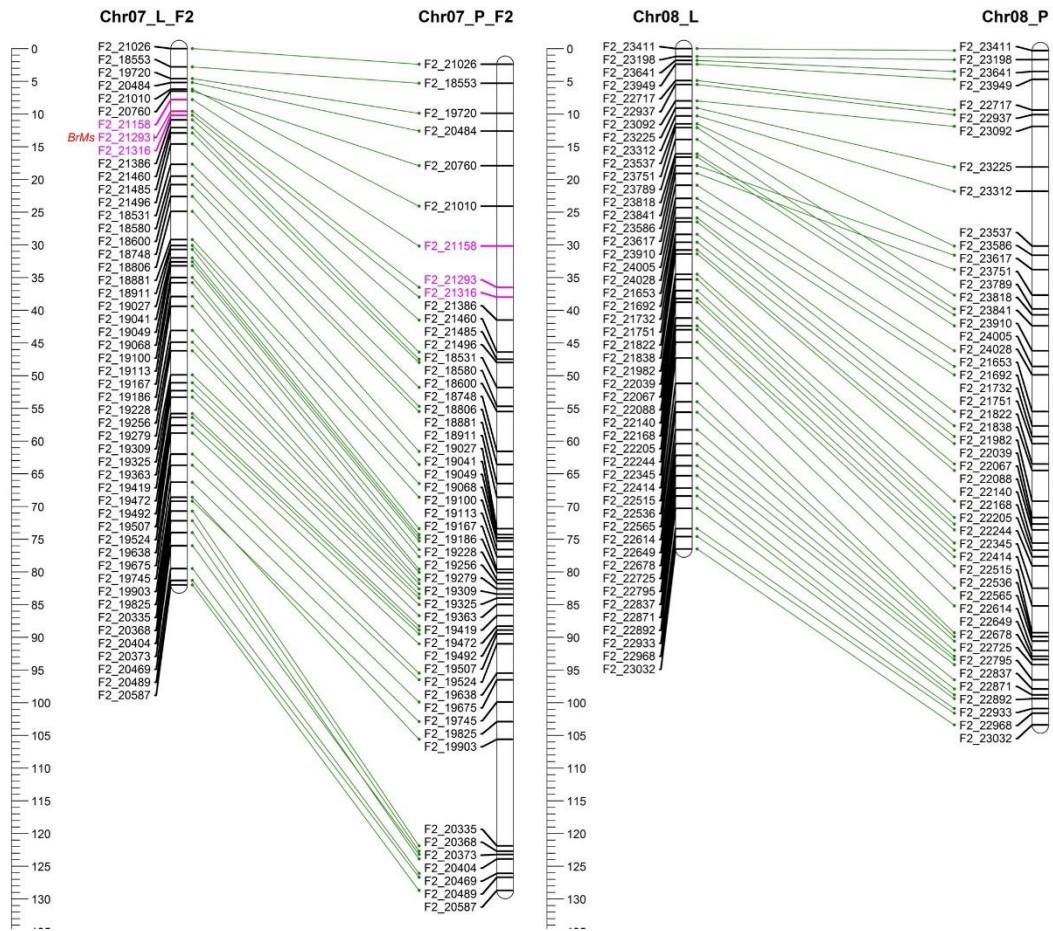
圖 3-1 以不結球白菜 NH80-A × TA95 之 F<sub>2</sub> 族群建立的高密度遺傳連鎖圖











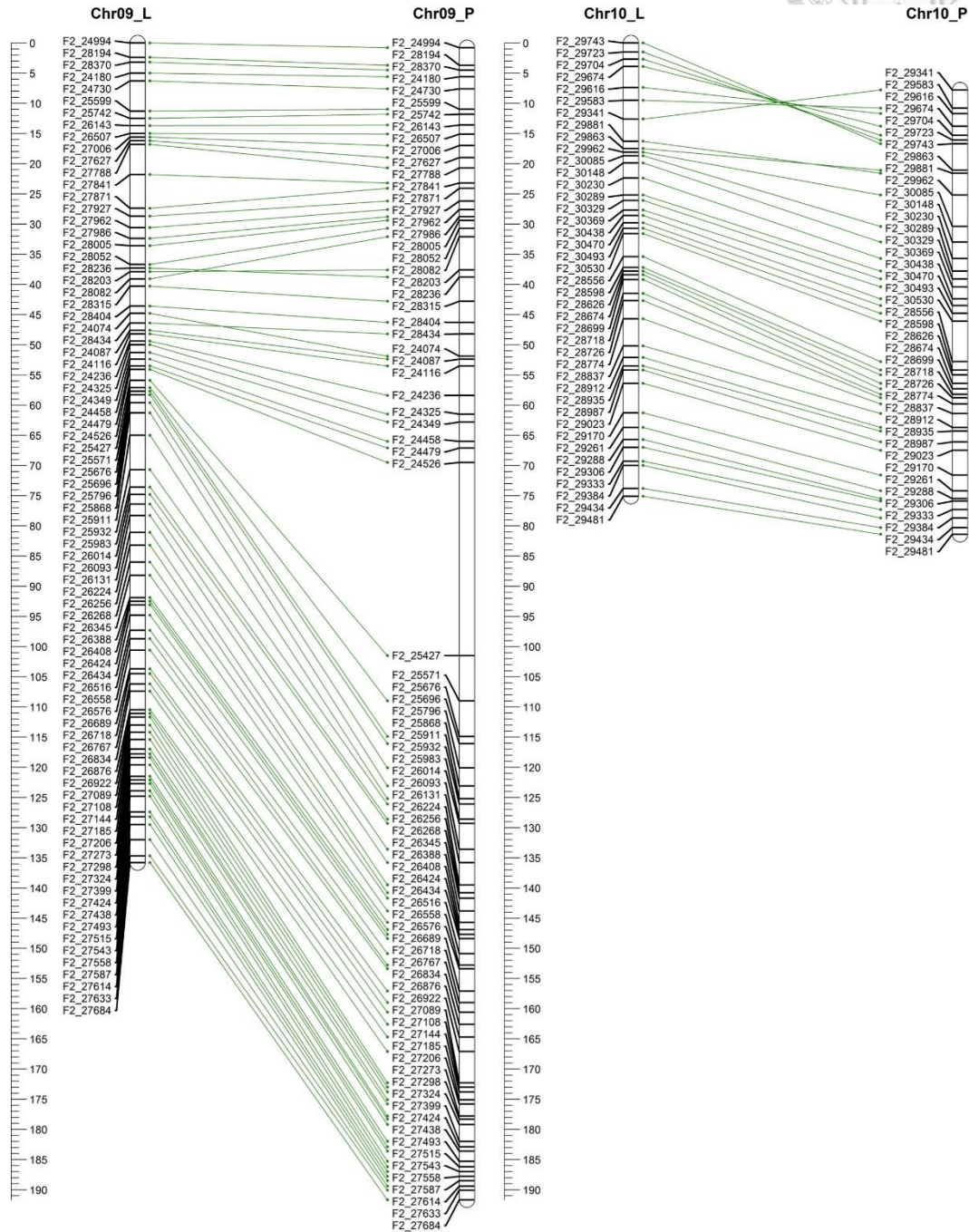


圖 3-2 以不結球白菜 NH80-A × TA95 之 F<sub>2</sub> 族群建立的各染色體遺傳連鎖圖  
 圖左側\_L: 表示遺傳連鎖圖  
 圖右側\_P: 表示連鎖圖上 bin markers 所對應之物理位置

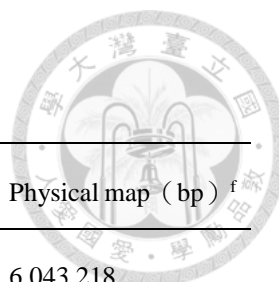


表 3-2 與 *BrMs* 基因共分離的 GBS tags

Bin <sup>d</sup>	Tags <sup>e</sup>	SNP numbers	Genetic map (cM)	Physical map (bp) <sup>f</sup>
F2_21158	F2_21158	1	7.78	6,043,218
	F2_21208	1		6,498,810
	F2_21222	1		6,632,406
F2_21293*	F2_21230	1	9.63	6,700,179
	F2_21231	2		6,700,608
	F2_21236	3		6,710,545
	F2_21271	3		7,038,629
	F2_21293	1		7,297,575
F2_21316	F2_21316	1	10.24	7,597,365
	F2_21322	1		7,625,231

<sup>d</sup>: \*表示與 *BrMs* 共分離的 bin marker

<sup>e</sup>: 在相同 bin 之下的各個 tags 其於族群內所有單株間的基因型皆相同

<sup>f</sup>: 的物理位置參照 *B. rapa* var. Chiifu40 (v1.5)，所列數字為 tags 第 1 個核苷酸的位置

## 4 討論

### 4.1 透過 GBS 策略建構連鎖圖以進行基因定位

次世代定序技術具有短時間內獲得大量 DNA 序列資料的優勢，因此對於基因體之研究有莫大之幫助，而 GBS 則是利用 NGS 之特色針對多個樣本進行 reduced-representation sequencing，可在特定育種族群內偵測大量 SNPs 並完成基因分型，成為近年普遍運用的基因定位策略之一。本研究中利用 GBS 方法建立不結球白菜之 F<sub>2</sub> 分離族群之高密度遺傳連鎖圖譜（圖 3-1），結果顯示於 NH80-A × TA95 之 F<sub>2</sub> 族群所發掘的多型性 tags 超過 2,000 個，其中包含之 SNPs 數量達 4,000 個以上（表 3-1），利用如此高密度之標誌建構連鎖圖而成功將 *BrMs* 基因定位於 A7 連鎖群之 2.4 cM 範圍內。此外，輔以結球白菜之基因體資訊，透過共分離的 SNPs 物理位置資訊，也進一步將目標基因定位於 A07 染色體所標示位置之 6.6 Mb~7.6 Mb 區間範圍內，相較於其他常用之分子標誌如 AFLP (Yi et al., 2006; Huang et al., 2007; Lei et al., 2007; Xiao et al., 2008) 或簡單重複序列 (simple sequence repeats, SSR) (Feng et al., 2009) 等，建立連鎖圖時，利用 sequencing-based 的 GBS 策略，著實具有低成本且可快速完成基因定位之優點。

## 4.2 本研究遺傳連鎖圖與其他不結球白菜連鎖圖之比較

本研究 F<sub>2</sub> 族群所構建的遺傳連鎖圖總長度為 1,011 cM，與其他白菜連鎖圖之總長相近（970 cM~1,230 cM）（Suwabe et al., 2006; Choi et al., 2007; Kim et al., 2009; Wang et al., 2011; Chung et al., 2014; Huang et al., 2017）。在以 GBS 策略建構高密度連鎖圖時，由於分子標誌數量大幅增加，資料運算效率的限制，及基因分型錯誤率提高是二個主要之挑戰（Van Os et al., 2005），其中基因分型錯誤率更是造成遺傳連鎖圖遺傳距離膨脹的主因（Hackett and Broadfoot, 2003），利用“bin”將相同分離情形的眾多標誌整合成單一標誌，可有效增加資料運算效能及降低基因分型錯誤的影響，已普遍應用於構建作物高密度連鎖圖（Sun et al., 2007; Ganai et al., 2011），故本研究將數以千計之 tags 整合成“bin markers”，有效地抑制連鎖圖膨脹，而使得連鎖圖總長與 PCR-based markers 所建構之連鎖圖接近，而這些 bin marker 的順序大多與物理圖上之順序一致（圖 3-2），顯示本研究所建立遺傳連鎖圖之可靠性。Huang et al.（2017）也以相似策略利用 restriction-site-associated DNA sequencing（RAD-seq）建立不結球白菜的高密度遺傳圖譜，與本研究 F<sub>2</sub> 連鎖圖相較時，二者 bins 間的平均間隔是相似的，本研究獲得較多 SNPs，可能是由於對多型性 tags 篩選之閾值不同，以及使用不同內切酵素建構 GBS library 所致，由於該連鎖圖以 *EcoRI* 切位建構 RAD-seq library，而本研究則採用 *PstI*，彼此依據不同酵素切位的周邊序列進行分析，因而影響探勘多型性 SNPs 的結果，亦造成二個遺傳圖譜之組成有些差異，但是均可做為白菜基因體研究之參考，若能整合這兩個連鎖圖之 SNPs 成為更高密度之遺傳圖譜，將助益於不結球白菜基因體分析、基因定位或分子育種等研究及應用。





第4章 以雄不稔株 × 維持系之 BC<sub>1</sub> 族群建立遺傳連鎖圖定位  
*BrMs* 基因



## 1 前言

本研究第二章之遺傳分析結果，顯示小油菜雄不稔性性狀符合單一基因座之複等位基因模式，即  $BrMs^a$ 、 $BrMs^b$ 、 $BrMs^c$  三個等位基因位於同一基因座上，且顯隱性關係為  $BrMs^a > BrMs^b > BrMs^c$ ，然而遺傳分析結果雖適配於該模式，但卻不必然為唯一模式，例如當二個控制雄不稔性的基因座緊密連鎖時，亦有可能產生類似前述遺傳分析之結果，或者還有其他目前未知的遺傳模式存在。在第三章中，我們利用雄不稔株×恢復系之  $F_2$  世代，建構  $BrMs^a/BrMs^b$  之分離族群，成功將  $BrMs$  基因座定位於第 A07 號染色體，在本章中我們仍設定雄不稔株 NH80-A 的基因型為  $BrMs^bMs^b$ ，維持系 WH606 基因型為  $BrMs^cMs^c$ ，利用該二者之回交  $BC_1$  世代，可以建構  $BrMs^b/BrMs^c$  之分離族群，假若  $BrMs^b/BrMs^c$  之定位區域明顯偏離第三章  $BrMs^a/BrMs^b$  之區間時，表示  $BrMs^c$  與第三章的  $BrMs^a/BrMs^b$  不具等位性，換言之，本研究希望利用具有已知物理位置的分子標誌，定位不同的等位基因，透過彼此間物理位置的比較更精確地說明這些等位基因（或基因座）間之關係。

## 2 材料與方法

### 2.1 建構基因定位分離族群

對於  $BrMs^b/BrMs^c$  之定位，2013 年 2 月將雄不稔株 NH80-A 與維持系 WH606 進行雜交，於同年 10 月將其  $F_1$  植株與 WH606 回交，以獲得  $BC_1$  種子，2015 年 3 月培育 186 株  $BC_1$  植株做為基因定位用族群，所有植株以盆栽種植 30 天後，取葉片萃取 total DNA 以建構 GBS 用之 DNA library，各單株透過 GBS 程序以分析基因型，而外表型資料則於開花期調查各單株之雄不稔性狀後獲得。

### 2.2 GBS 之進行

#### 2.2.1 建構 GBS library

GBS library 之建構主要參考 Poland et al. (2012) 之方法而略加修飾，每一植株取約 200 ng DNA 後同時以 *Pst*I-HF 及 *Msp*I (New England Biolabs, USA) 二種內核核酸酶於 37 °C 下反應 4 小時後，再於 65 °C 維持 20 分鐘，酶切完成之 DNA 片段以 barcoded adapters 及 reverse Y-adaptor (Poland et al., 2012) 進行連接，各樣本連接反應體積為 40  $\mu$ L，包含酶切 DNA (200 ng)、barcoded adapter (20 nM)、reverse



Y-adapter (300 nM)、T4 DNA ligase (200 U)、1X NEB Buffer 4 (NEB, USA) 及 ribo ATP (1 mM)，於 22 °C 下進行 2 小時後再於 65 °C 維持 20 分鐘，完成 adapters 連接之 DNA 各取 15  $\mu$ L 混合成 DNA library，隨後用 AMPure XP 磁珠 (Beckman Coulter, USA) 以 DNA:磁珠為 1:1.8 的比例進行純化，最後利用經修飾之 Illumina 引子對 (Poland et al., 2012) 於 95 °C (30 sec)，62 °C (25 sec)，68 °C (40 sec) 條件下進行 PCR 反應 18 個 cycles，建庫完成的 GBS library 同樣由國立陽明大學基因體研究中心以 Illumina Hiseq2000 定序平台 (Illumina Inc., USA) 進行次世代定序，定序方式為 100 bp 的單端讀序。

### 2.2.2 序列分析

本研究以下載自 Brassica database (BRAD) (<http://brassicadb.org/brad/index.php>) 的 *B. rapa* var. chiifu40 (v1.5) 為參考序列 (Cheng et al., 2011)，經次世代定序獲得之讀序以軟體 CLC Genomics Workbench v8.0.2 (CLC Bio, <http://www.clcbio.com>) 篩選序列品質分數大於 Q20 者，依據各樣品 DNA 條碼進行分類，去除條碼序列後隨之與參考序列比對，並以 sam 的檔案格式輸出，而基因座暨基因座目錄之建立、SNP calling、分離後裔之基因分型等則以軟體 Stacks v1.37 (<http://creskolab.uoregon.edu/stacks/>) (Catchen et al., 2011) 之“ref\_map.pl”指令進行，各讀序中的專一性序列被定義為 tag，只有讀序深度  $\geq 5$ ，且於族群內缺失率少於 20% 的 tags 才會輸出成為基因型資料。

### 2.2.3 連鎖分析

族群的基因型資料以 R/qtl 程式建立遺傳連鎖圖 (Broman et al., 2003; Broman, 2010)，簡言之，將個體或 tags 缺值大於 >20% 之資料捨棄後，選取未偏離常態分布的 tags，以重組率  $\leq 0.3$  且 LOD 值  $\geq 7$  之條件分類為相同連鎖群，並以 Kosambi 函數 (Kosambi, 1943) 估計遺傳距離，建立初級連鎖圖之後，將具有相同分離情形或遺傳距離小於 0.5 cM 的 tags，合併為相同的 bin，由各 bin 中隨機選取一個 tag 做為 bin marker，再次計算 bins 間之遺傳距離而建構出遺傳連鎖圖，相鄰 bin 間至少包含 1 次重組，此外，族群內所有個體之雄不稔性外表型資料，連同 bins 一併進行連鎖分析，以估計雄不稔基因與相鄰 bins 間之遺傳距離。



### 3 結果

#### 3.1 GBS 的結果

NH80-A × WH606 之 BC<sub>1</sub> 族群被用來定位含有 *BrMs<sup>b</sup>* 及 *BrMs<sup>c</sup>* 等位基因之 *BrMs* 基因座，進行 GBS 程序後，BC<sub>1</sub> 族群產生約 25 Gb 序列資料，經序列品質篩選後獲得 250 millions 個平均長度為 92.6 bp 的序列，讀序平均長度小於第三章 F<sub>2</sub> 族群之 93.6 bp，主要原因是 BC<sub>1</sub> 之個體數較多，建構其 DNA library 時 DNA 條碼由 5~8 bp 構成，而 F<sub>2</sub> DNA library 其條碼則僅由 5~6 bp 組成，經刪除條碼序列後，讀序平均長度如預期約減少 1 bp，將 BC<sub>1</sub> 之所有讀序比對至參考序列後，維持系父本 (WH606) 之成功比對率為 85.4%，族群整體比對率則為 84.4%。

BC<sub>1</sub> 族群共產生 22,224 個 tags，少於第三章 F<sub>2</sub> 族群之 GBS 結果 (33,632 個)，經排除個體間缺失率大於 20% 或不符合分離比率者，篩選獲得 1,016 個具多型性的 tags，其中含 1,791 個 SNPs，合併為 493 個 bins 後所建立之連鎖圖全長為 1,051 cM，同樣由 10 個連鎖群構成，仍以 A09 連鎖群最長達 142.1cM，最短之 A10 連鎖群則為 79.8 cM (表 4-1、圖 4-1)，整個連鎖圖中 bins 的平均間隔為 2.1cM，然而有 4 區域其相鄰 bins 之間隔大於 10 cM，分別為 Chr03 之 BC\_4152~BC\_4588、Chr05 之 BC\_8085~BC\_8541、Chr07 之 BC\_12149~BC\_12305 與 BC\_13338~BC\_13433 等 (圖 4-2)，每個 bin 有 1.3 至 2.4 個具多型性之 tags，整體平均含 2.1 個多型性 tags (表 4-1)，密度明顯小於 F<sub>2</sub> 族群之連鎖圖 (3.9 tags/bin)。進一步將各連鎖群之 bin markers 比對其位於參考序列上的物理位置，結果顯示在 493 個 bin markers 中，大部份 bin markers 順序與物理位置一致，僅有 11 個 bin markers 其順序與相鄰的 bins 之間有倒置 (inversion) 現象 (圖 4-2)。

#### 3.2 *BrMs* 之基因定位

在以 NH80-A × WH606 之 BC<sub>1</sub> 族群而建立之連鎖圖中，*BrMs<sup>b</sup>/BrMs<sup>c</sup>* 基因也被定位於 A07 連鎖群 (圖 4-2)，該連鎖圖中，A07 連鎖群全長約 99.8 cM，*BrMs<sup>c</sup>* 等位基因與 bin“BC\_13924”共分離並與兩側相鄰 bins “BC\_13879”及“BC\_13971”分別相距 0.9 cM 及 1.1 cM，由於 tags “BC\_13914”、“BC\_13917”及“BC\_13923”與 bin “BC\_13924”基因型相同，因此共有 4 個 tags 與 *BrMs<sup>c</sup>* 共分離，這 6 個 tags 當中共計

有 7 個 SNP 位點（表 4-2），整合其物理位置資訊顯示 *BrMs<sup>b</sup>/BrMs<sup>c</sup>* 座落於 A07 染色體標示為 6.2 Mb 與 7.6 Mb 位置間之 1.37 Mb 範圍內（表 4-2）。相較於第三章 *BrMs<sup>a</sup>/BrMs<sup>b</sup>* 被定位於 A07 染色體 6.6–7.6 Mb 之區域，二個基因定位結果區間重疊於位置 6.6–7.6 Mb 之間（圖 4-3），此外，BC<sub>1</sub> 族群之 tag “BC\_13914”與源自 F<sub>2</sub> 族群之 tag “F2\_21230”實際上座落於相同物理位置而為相同的 tag（表 3-2、4-2）。



表 4-1 以 BC<sub>1</sub> 族群建立之連鎖群長度、SNP 數量及 marker 密度

Characteristics	Linkage Groups										Total
	A01	A02	A03	A04	A05	A06	A07	A08	A09	A10	
Length (cM)	108.4	116.5	135.3	81.2	99.6	108.8	99.7	79.9	142.1	79.8	1,051
No. of SNP	192	229	189	87	167	217	156	155	283	116	1,791
No. of polymorphic Tag	117	117	100	42	90	132	87	98	175	58	1,016
No. of Bin	55	59	55	32	38	57	43	44	77	33	493
average SNPs (SNPs/Tag)	1.6	2.0	1.9	2.1	1.9	1.6	1.8	1.6	1.6	2.0	1.8
average Tags (Tags/bin)	2.1	2.0	1.8	1.3	2.4	2.3	2.0	2.2	2.3	1.8	2.1
average interval (cM/bin)	2.0	2.0	2.5	2.5	2.6	1.9	2.3	1.8	1.8	2.4	2.1

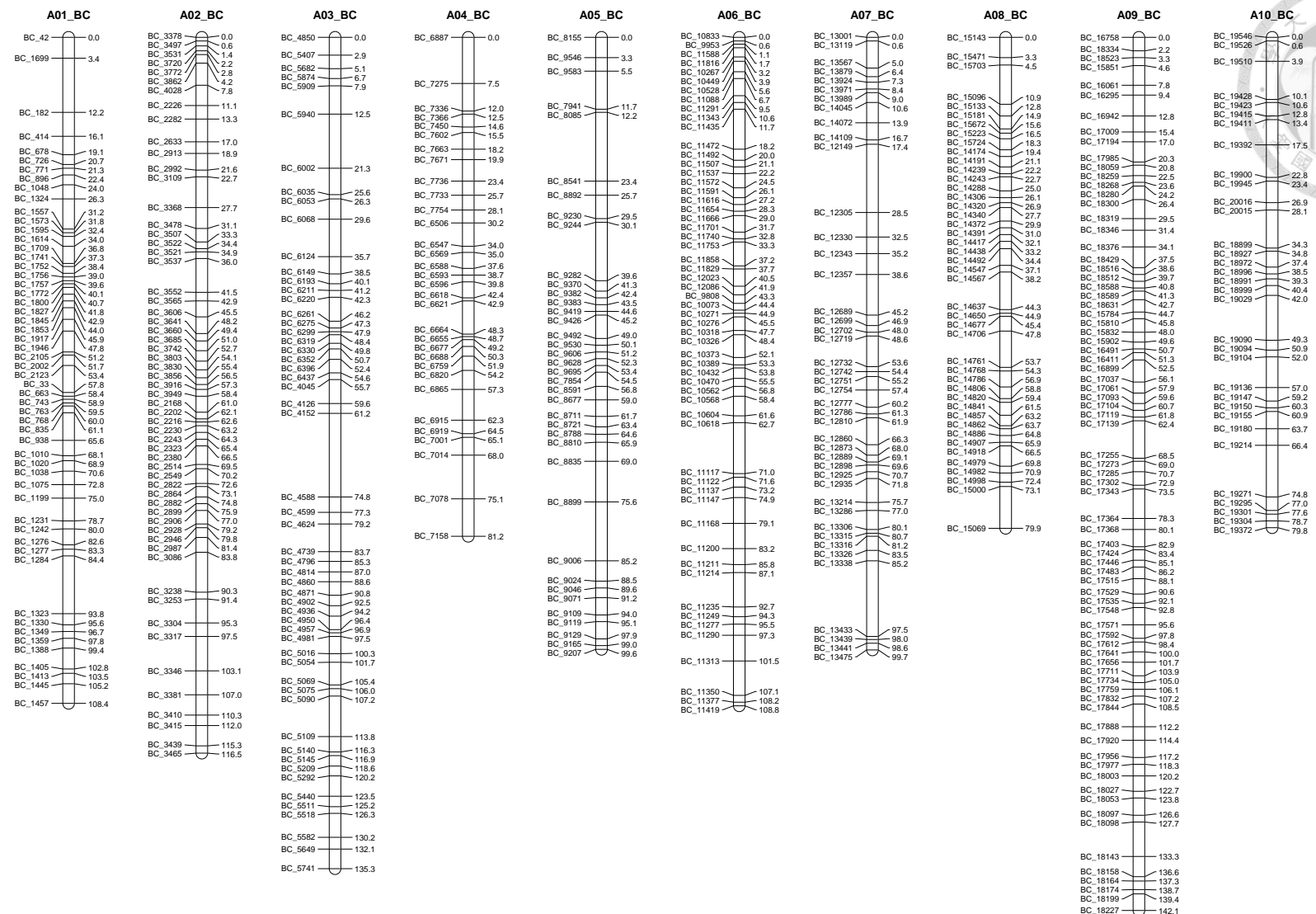
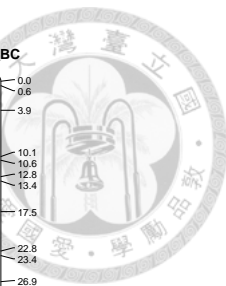
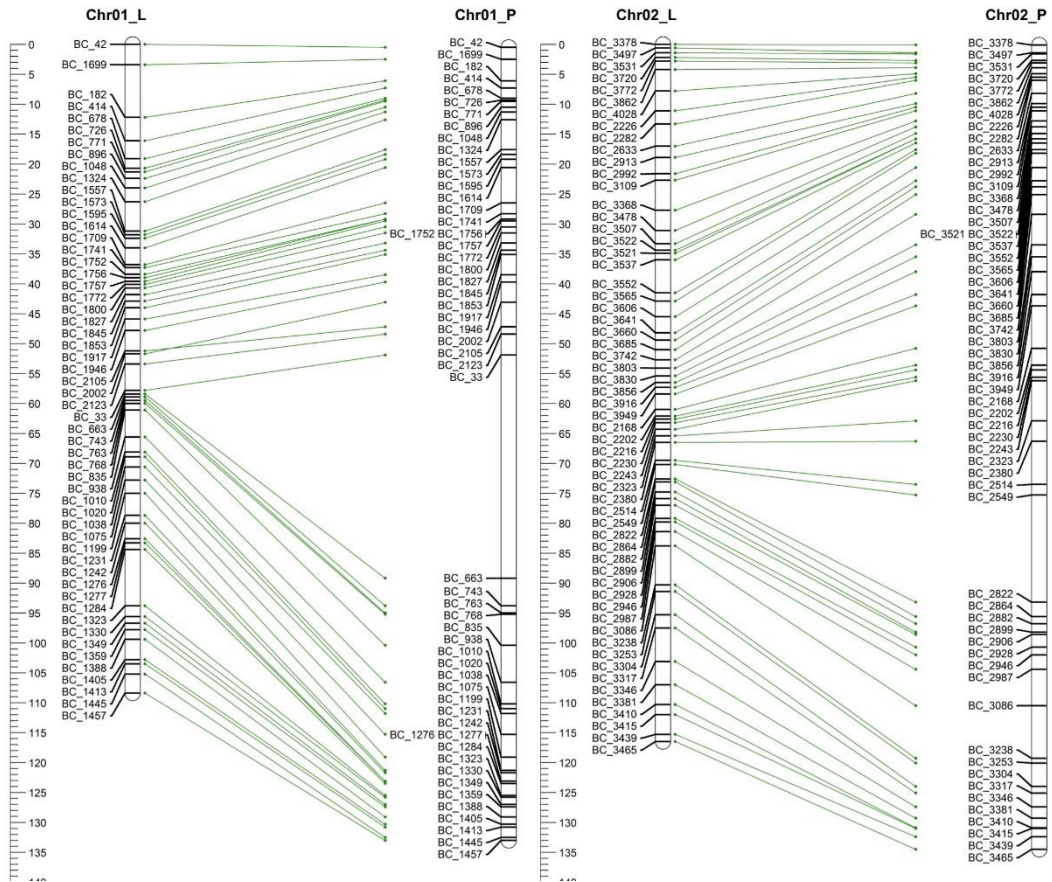
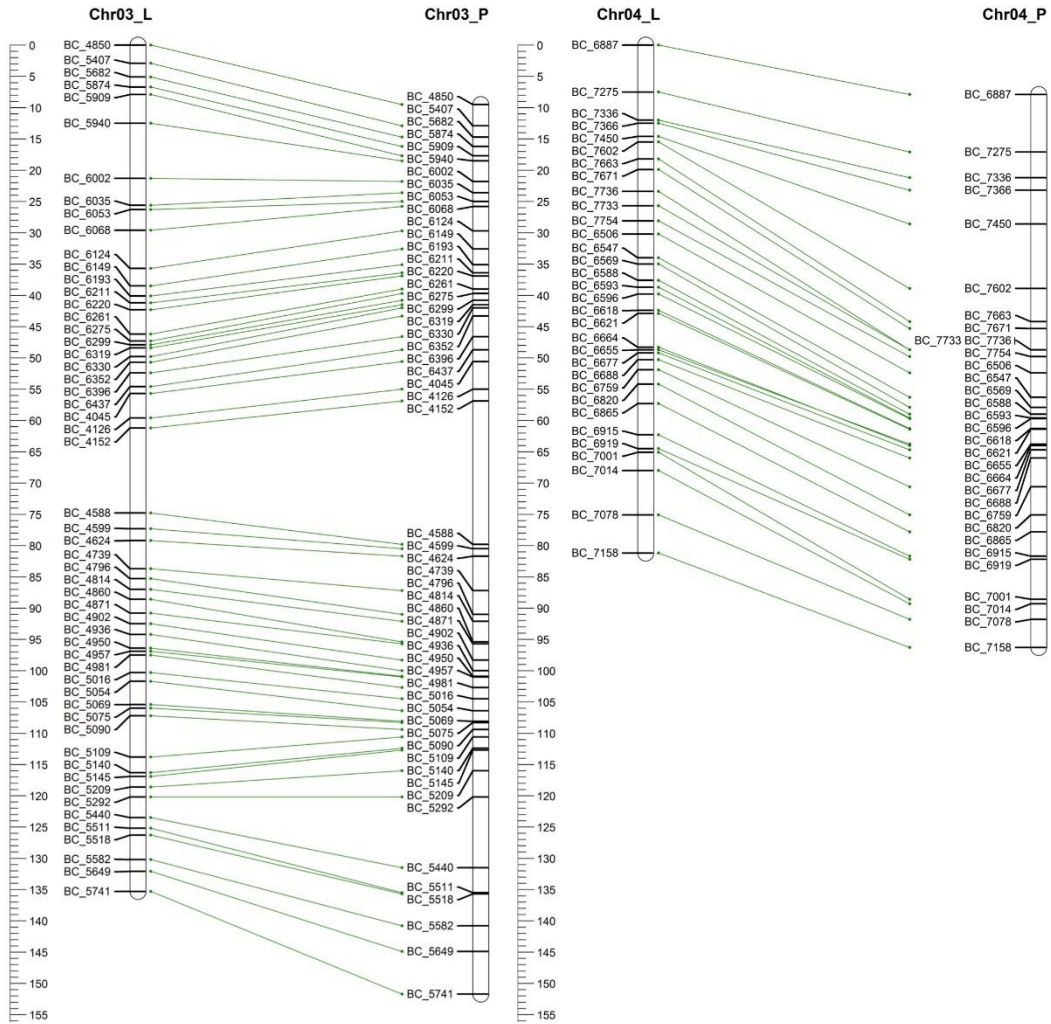
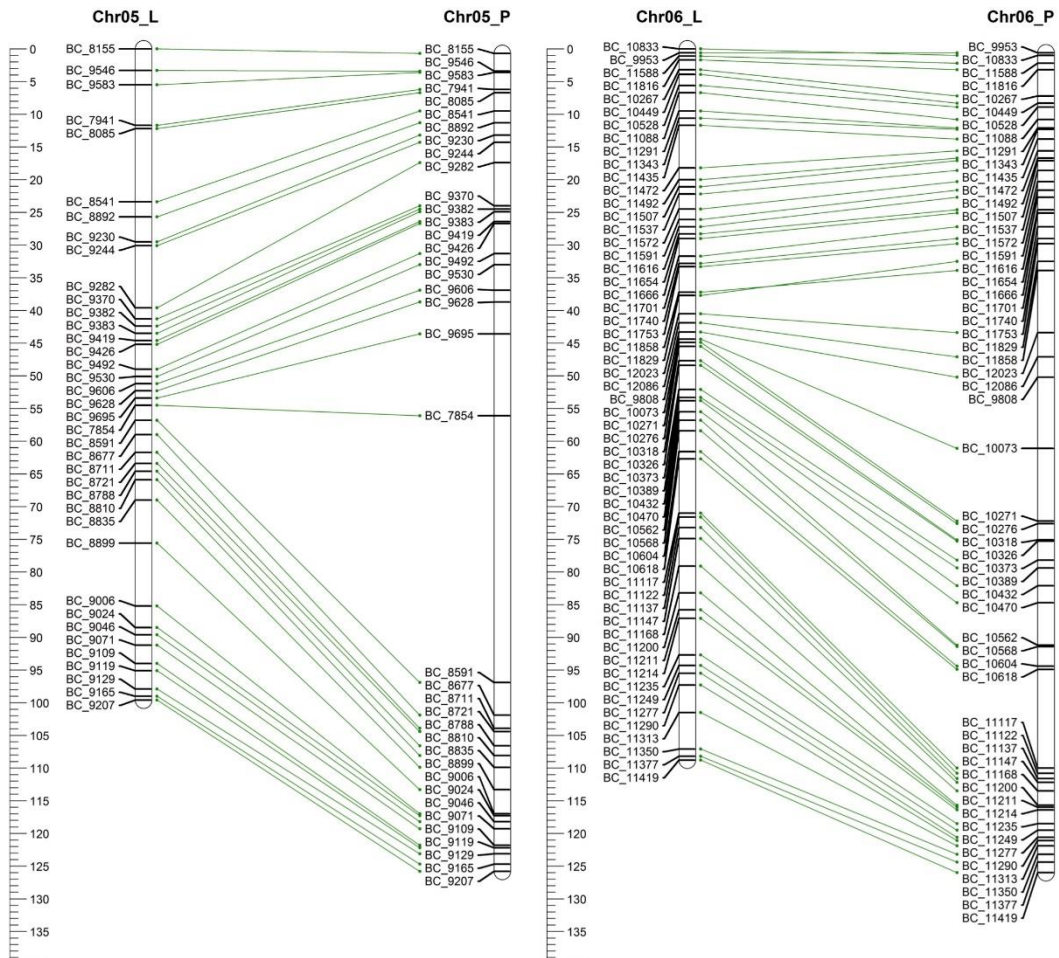


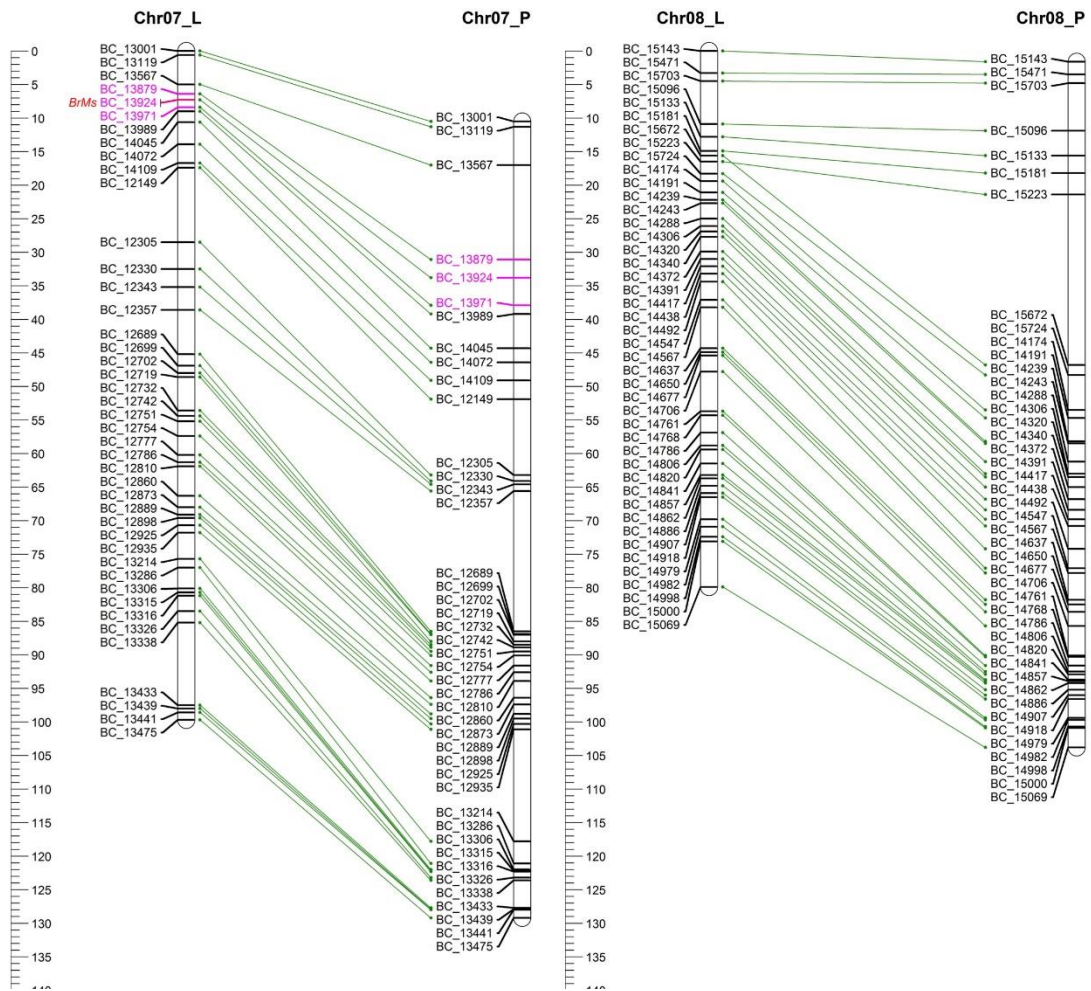
圖 4-1 以不結球白菜 NH80-A × WH606 之 BC<sub>1</sub> 族群建立的高密度遺傳連鎖圖











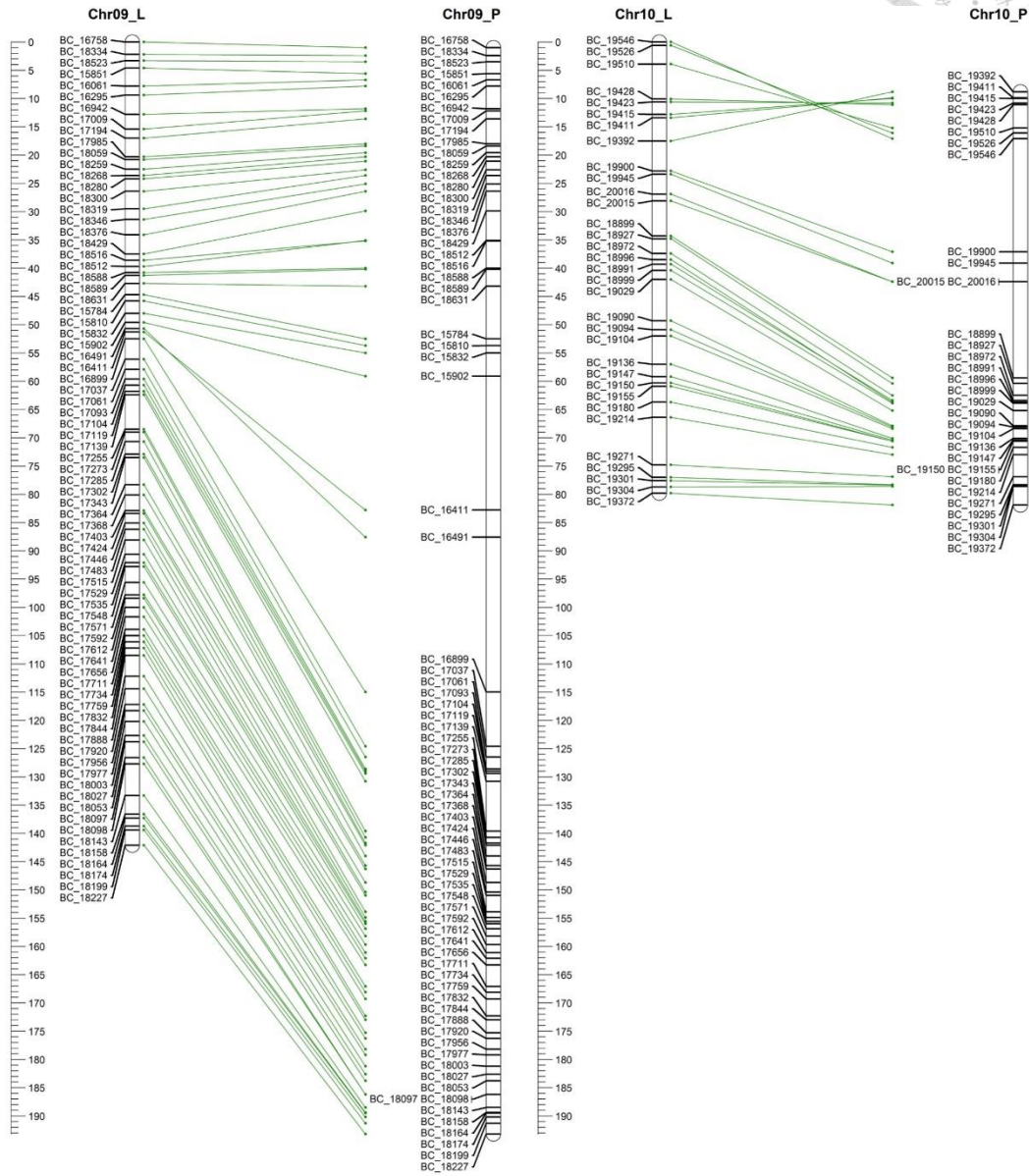


圖 4-2 以不結球白菜 NH80-A × WH606 之 BC<sub>1</sub> 族群建立的各染色體遺傳連鎖圖  
 圖左側\_L: 表示遺傳連鎖圖  
 圖右側\_P: 表示連鎖圖上 bin markers 所對應之物理位置



表 4-2 與 *BrMs* 基因共分離的 GBS tags

Bin <sup>d</sup>	Tags <sup>e</sup>	SNP numbers	Genetic map (cM)	Physical map (bp) <sup>f</sup>
BC_13879	BC_13879	2	6.44	6,211,318
BC_13924*	BC_13914	1	7.30	6,700,179
	BC_13917	1		6,700,978
	BC_13923	1		6,736,962
	BC_13924	1		6,750,243
BC_13971	BC_13971	1	8.41	7,583,322

<sup>d</sup>: \*表示與 *BrMs* 共分離的 bin marker

<sup>e</sup>: 在相同 bin 之下的各個 tags 其於族群內所有單株間的基因型皆相同

<sup>f</sup>: 物理位置參照 *B. rapa* var. Chiifu40 (v1.5)，所列數字為 tags 第 1 個核苷酸的位置

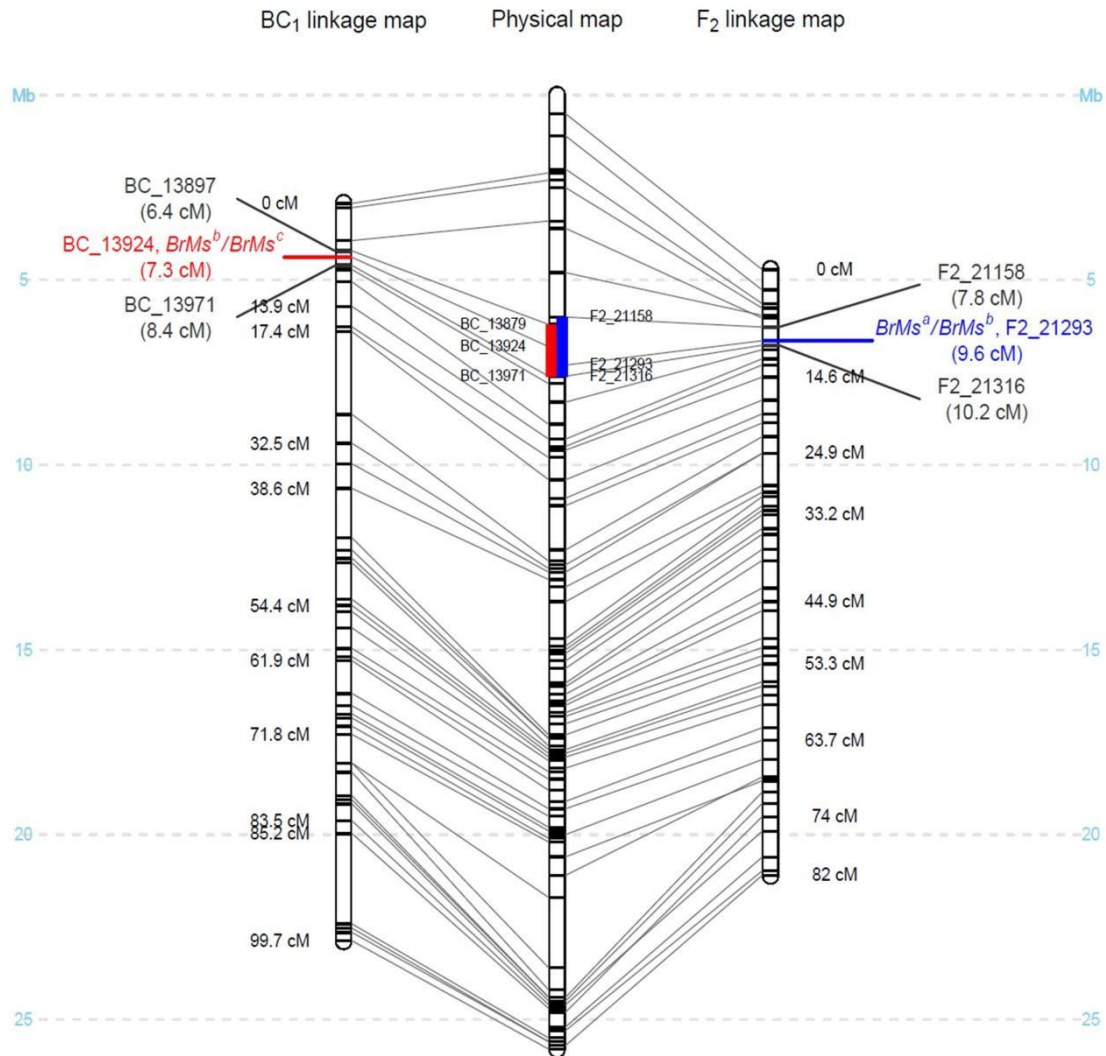


圖 4-3 不結球白菜雄不稔基因 *BrMs* 之定位區間

圖左側為針對 *BrMs<sup>b</sup>/BrMs<sup>c</sup>* 基因座建構之遺傳連鎖圖，其基因定位區間為物理圖譜上紅色區域；圖右側為針對 *BrMs<sup>a</sup>/BrMs<sup>b</sup>* 基因座建構之遺傳連鎖圖，其基因定位區間為物理圖譜上藍色區域



## 4 討論

### 4.1 BC<sub>1</sub> 遺傳連鎖圖與 F<sub>2</sub> 連鎖圖之比較

本研究中利用 GBS 方法分別建立不結球白菜之 F<sub>2</sub> 與 BC<sub>1</sub> 分離族群之高密度遺傳連鎖圖，第三章中於 NH80-A×TA95 之 F<sub>2</sub> 族群連鎖圖上所發掘的 SNPs 數量超過 NH80-A × WH606 之 BC<sub>1</sub> 族群連鎖圖二倍以上，其主要原因為雜交親本間遺傳歧異度不同所致，源自台灣的雄不稔系 NH80-A 與維持系 WH606 其與參考序列間之比對率分別為 81.7% 與 85.4%，而 TA95 源自泰國其比對率僅為 41.9%，顯示 TA95 與二個台灣品系間之遺傳距離較遠，致使 F<sub>2</sub> 族群的親本間可產生較多多型性 tags 用以進行連鎖分析 (表 3-1)。此外，F<sub>2</sub> 及 BC<sub>1</sub> 族群之定序資源一樣，皆是以 Illumina HiSeq2000 定序平台進行二個 lanes 的定序量，然而 F<sub>2</sub> 族群個體數小於 BC<sub>1</sub> 族群，也就是說 F<sub>2</sub> 每單一個體所獲得之讀序數量大於 BC<sub>1</sub> 的個體，因此以獲得之序列資料與參考序列比對時，F<sub>2</sub> 個體各位點的定序深度較大且缺值減少，因而容易達到篩選的條件，而獲得較多之 tags，進而發掘較多之 SNPs 位點。整體而言，F<sub>2</sub> 連鎖圖上 bins 的分佈較為均勻，BC<sub>1</sub> 連鎖圖上 bins 的平均區間大於 F<sub>2</sub> 連鎖圖，且有 4 個 bins 區間大於 10 cM，可能是由於該區域的染色體同源性高缺乏多型性所致，僅管二個連鎖圖的 bins 密度略有差異，若將這二個不結球白菜連鎖圖上的 bin markers，對應至結球白菜參考序列之物理位置，不結球白菜各連鎖群上 bins 的順序，與結球白菜參考序列的物理圖間，均具有良好的共線性與一致性，間接證實二個連鎖圖的可靠性。二個連鎖圖上有少部份區域與物理圖間呈現順序上之差異，以 A10 連鎖群為例，當與結球白菜物理圖比較時，可發現 F<sub>2</sub> 及 BC<sub>1</sub> 連鎖圖在該連鎖群上臂相似位置處均存在染色體倒置的現象 (圖 3-2、4-2)，這可能是本試驗材料均屬不結球白菜而與參考序列之結球白菜在基因體演化上的差異。

### 4.2 基因定位結果驗證 *BrMs<sup>a</sup>*、*BrMs<sup>b</sup>* 及 *BrMs<sup>c</sup>* 的等位性

本研究以雄不稔株 (NH80-A) 與維持系 (WH606) 之 BC<sub>1</sub> 族群進行連鎖分析以定位雄不稔相關基因，結果顯示維持系之 *BrMs<sup>c</sup>* 被定位於 A7 染色體約 6.2~7.6 Mb 區間 (表 4-2)，相較於第三章以雄不稔株 (NH80-A) 與恢復系 (TA95) 之 F<sub>2</sub> 族群試驗結果，恢復系之 *BrMs<sup>a</sup>* 被定位於 A07 染色體約 6.6~7.6 Mb 之區域 (表 3-2)，二個基因定位區間重疊於 6.6~7.6 Mb 位置間 (圖 4-3)，並於該區間內發現

有相同之 tags (“BC\_13914”與“F2\_21230”)與這些等位基因共分離，由於恢復系 TA95 與維持系 WH606 二者與相同雄不稔單株雜交後裔的稔性表現不同，表示 TA95 與 WH606 在雄不稔性狀上必為不同之基因型，但二者之雄不稔基因又定位於相同區間，顯示 *BrMs<sup>a</sup>*、*BrMs<sup>b</sup>* 及 *BrMs<sup>c</sup>* 間為等位性之可能性極高，此結果亦支持 *BrMs* 為複等位基因之遺傳模式，此與第二章雜交後裔檢定分析結果一致。

前人研究在結球白菜及大油菜也曾發現受複等位基因控制的雄不稔性 (Song et al., 2006; Feng et al., 2009; Dong et al., 2012)，其中結球白菜主要以二種雜交後裔檢定進行遺傳分析而證實其遺傳模式 (Feng et al., 1996)，而大油菜則是先以 BSA 的策略找到與雄不稔等位基因連鎖之分子標誌後，再利用相同分子標誌與不同等位基因間存在之連鎖關係，而證實複等位基因之遺傳模式 (Song et al., 2006; Liu et al., 2008; Xie et al., 2012)，本研究除了發現相同分子標誌而證實等位基因間之連鎖關係外，尚解析出各等位基因之實際物理區間，加上第二章之測交遺傳分析的結果，對於此不結球白菜雄不稔性遺傳模式之解析更為直接與清晰，儘管在分子的層次上 *BrMs<sup>a</sup>* 與 *BrMs<sup>c</sup>* 仍無法排除為此區間內不同基因座之可能性，即使如此，則 *BrMs<sup>a</sup>* 與 *BrMs<sup>c</sup>* 必然也存在緊密連鎖關係，廣義遺傳上也可視其為相同基因座上之複等位基因。本試驗雄不稔基因座被定位於 A07 連鎖群，與結球白菜之雄不稔複等位基因座位於相同連鎖群 (Feng et al., 2009)，二者間互為同物種下之不同變種，且雄不稔性狀之遺傳模式相同，彼此間是否為相同基因是值得進一步探討。綜合結球白菜、大油菜及本研究於不結球白菜雄不稔性之結果，推測蕓苔屬作物中普遍存在複等位基因控制的雄不稔性遺傳模式，此與 Lu et al. (2013) 之推論一致。此外，本研究之基因定位結果，將有助於做為親源相近作物之細胞核雄不稔基因精細定位 (fine mapping) 及據圖選殖 (map-based cloning) 的參考。



## 第5章 *BrMs* 基因之精細定位及開發其 SNP 分子標誌



## 1 前言

本研究第三及第四章，利用小油菜 *BrMs* 不同等位基因分離族群進行基因定位之結果，不僅驗證由傳統遺傳分析所推論之複等位基因控制遺傳模式，也將 *BrMs* 基因初步定位於 A07 染色體上約 1 Mb 之範圍內，此定位結果是基於 GBS 所獲得之 bin markers 的生物資訊，由於目前並沒有專屬小油菜（不結球白菜）之全基因體資訊，本研究採用結球白菜基因體做為分析之參考序列，二者間是否存在差異不得而知，若能利用 GBS bin markers 之資訊實際開發為 SNP 分子標誌，並於較大的分離族群中進行驗證，有助於確認研究結果之可信度。

由於本研究 *BrMs* 基因定位區間具有實際的染色體物理位置（以結球白菜基因體為標準），這些資訊提供了 *BrMs* 與其他雄不稔基因間進行比對的可能性，如同物種結球白菜 *Ms* 基因位於 A07 染色體上（Feng et al., 2009），而大油菜也有二個雄不稔基因被定位於 A07 染色體上（Yi et al., 2006; Xiao et al., 2008），若能透過相同分子標誌之連鎖關係，或以比較基因體之方式比對雄不稔基因物理位置，則可釐清 *BrMs* 是否為小油菜特有之雄不稔基因。

此外，目前 *BrMs* 基因定位範圍內被預測的候選基因數量多，若能進一步進行精細定位而縮小 *BrMs* 的座落區間，不僅可有效減少基因的數量而助於選殖出 *BrMs*，若能在更小範圍的區間內設計分子標誌也利於提升選拔的準確度。由於本研究第三及第四章採用之 GBS 技術，相當於針對基因體部份取樣而經濟地分析參試樣品間之基因型（只分析 *Pst*I 切位周邊約 100 bp 序列），然因部份取樣的策略，導致區間內有限的 GBS bin markers 數量不易進行基因精細定位，為增加該區間內分子標誌的密度，本研究擬針對雄不稔株 NH80-A 與恢復系 TA95，進一步利用全基因體定序（whole genome resequencing）的方式獲得前述定位區間內之序列資訊，再發掘彼此間之多型性 SNP 位點，據以開發為分子標誌，並用之分析更大之分離族群，除完成基因精細定位外，*BrMs* 周邊眾多的緊密連鎖 SNP 分子標誌將有助於育種應用。





## 2 材料與方法

### 2.1 粗定位之 GBS bin markers 驗證

#### 2.1.1 建構 *BrMs* 基因之分離族群

為了能準確估計 *BrMs* 與周邊 GBS markers 間之遺傳距離，2013 年 10 月栽培雄不稔株 NH80-A 與恢復系 TA95 並進行雜交，次年度 3 月將其 F<sub>1</sub> 植株自交以獲得 F<sub>2</sub> 種子，2016 年 1 月種植 1,590 株 F<sub>2</sub> 個體，以做為 GBS markers 驗證用之分離族群。將參試 F<sub>2</sub> 種子播種於 104 孔穴盤後，於農業試驗所溫室培育至 5 葉苗齡時移入生長箱，以 8°C 進行春化處理 30 天後，再次移植至溫室栽培，取葉片萃取 DNA 後，利用本章所開發之 SNP 分子標誌進行基因分型，並於開花期調查各單株之雄不稔性狀及記錄其外表型。

#### 2.1.2 將 GBS 及 SCARs markers 轉換為 SNP 分子標誌

選取本研究第三章 GBS 試驗中與雄不稔基因 *BrMs* 共分離或相鄰之 tags，依據 NH80-A 與 TA95 之序列差異設計為 SNP 分子標誌，進行時擷取這些 tags 上 SNP 位點兩側的 DNA 序列資訊至少 50 bp 以“KASP by Design”設計成 Kompetitive Allele Specific PCR (KASP) assays (LGC Genomics, UK)。此外，亦將結球白菜上所開發之二個與雄不稔基因 *Ms* 相關之 SCARs 分子標誌 “syau\_SCR01” 及 “syau\_SCR04” (Feng et al., 2009) 轉換為 SNP 分子標誌，轉換時將該 SCARs 引子組對 NH80-A 及 TA95 進行 PCR，將所獲得之擴增產物 (amplicons) 選殖至 pGEM-T Easy Vector (Promega, USA) 並轉型 (transform) 至 *E. coli* DH5 $\alpha$ ，於含有 Amp<sup>r</sup>/X-gal/IPTG 之 LB 固體培養基中各挑選三個轉型成功的菌落，接著繼代培養於 LB 液體培養基後萃取 DNA，並以 ABI Prism 3730 DNA 定序儀 (Applied Biosystems, USA) 進行定序，源自 NH80-A 與 TA95 之擴增產物序列以 BioEdit 7.2.0 軟體 (Hall, 1999) 進行比對，再依據序列間之 SNP 位點設計為 KASP assay。



### 2.1.3 參試植株基因分型及 *BrMs* 連鎖分析

為了驗證 GBS markers 的正確性，而針對 1,590 株 NH80-A × TA95 之 F<sub>2</sub> 個體以 SNP 分子標誌進行基因分型。進行時各單株以 QuickExtract™ DNA extraction solution (Epicentre, WI) 萃取 DNA，隨之將經稀釋之 DNA 萃取液、KASP assay 及 KASP 2X master mix，以 5 µl 總反應體積於 Applied Biosystems Viia 7 Real-Time PCR System (Applied Biosystems) 進行反應及螢光偵測。PCR 反應條件依據 LGC 提供之操作指南進行 (94°C 15 分鐘，隨之以 94°C 20 秒及 61-55°C 60 秒 (每 cycle 降低 0.6°C) 進行 10 個 cycles，再以 94°C 20 秒，55°C 1 分鐘進行 26 個 cycles，最後維持反應於 40°C 1 分鐘，並進行等位基因專一性之螢光偵測)，等位基因專一性之螢光偵測完成後，各單株基因型由 Real-Time PCR 隨附之 Viia 7 software (v1.2) 進行基因型判斷。進行連鎖分析時，將分離族群中雄不稔性外表型資料視為單一標誌，併同 SNP 分子標誌 (由 GBS 及 SCARs markers 轉換) 的基因型資料，利用 R/qlt 程式進行連鎖分析 (Broman et al., 2003; Broman, 2010)，以重組率 ≤ 0.3 且 LOD 值 ≥ 7 之條件分類為相同連鎖群，並以 Kosambi 函數 (Kosambi, 1943) 估計遺傳距離。

## 2.2 不結球白菜 *BrMs* 與大油菜 *BnRf* 之關係

### 2.2.1 不結球白菜 *BnRf* orthologs 間序列比較及引子設計

大油菜之 *BnRf* 基因透過比較基因體學之方式，發現其相對物理位置座落於本研究之 *BrMs* 基因定位範圍內，為研究 *BrMs* 與 *BnRf* 之關係而分析不結球白菜上同源基因之 DNA 序列差異及等位性，由於 *BnRf* 屬於大油菜的基因，因此本研究於不結球白菜所分析者以 *BnRf* 之同源基因 (orthologs) 稱之。分析時參考大油菜 *BnRf* 基因的 DNA 序列 (NCBI accession No. KT818624 及 KT818625) 設計引子，並對本研究之雄不稔株 NH80-A 與恢復系 TA95 進行 PCR 反應進而解序，KT818624 全長 8,572 bp，共設計可涵蓋基因全長之四組引子對，此外，由於本研究進行全基因體解序之研究中，TA95 之組裝序列顯示其 *BnRf* ortholog 基因上游有明顯序列差異，因此也依據 TA95 之 *BnRf* ortholog 基因上游約 2,000 bp 設計一組引子對參與分析。總計設計 5 組引子如表 5-1 所示。

表 5-1 選殖不結球白菜 *BnRf* orthologs 時所使用之引子對序列

Primers	Sequence	Reference
F_1	5'-GTCCAGTCAAACATTTACACAGG-3'	KT818624
R_2568	5'-TCTAGTGGCGGCCTTAGGAT-3'	KT818624
F_2376	5'-GCCAACTAAAGCCCTCGACT-3'	KT818624
R_4477	5'-ACTGCGAAACTTGCCAGGTA-3'	KT818624
F_4150	5'-TAGCTGAGATTACCCACCCCT-3'	KT818624
R_6207	5'-CCGGGTAAACACCTTTGCAC-3'	KT818624
F_6051	5'-ACAAGAAGCCCGTGCCAAAA-3'	KT818624
R_8572	5'-GATCGACAAATGCATCAACAAC-3'	KT818624
B60214F_1	5'-TCCTTTGCCCATTTTCTCTACTCA-3'	TA95 genome
B60214R_2020	5'-CCATGTCTTGTCGGAGGAGTT-3'	TA95 genome

## 2.2.2 *BnRf* orthologs 之 DNA 片段選殖與定序

由 NH80-A 及 TA95 各取 1  $\mu$ l 之 DNA ( $\sim$ 150 ng/ $\mu$ l)，以前述源自 *BnRf* 基因 DNA 序列設計之引子各 1  $\mu$ l (10  $\mu$ M) 及 25  $\mu$ l 2X Fast-Run Advanced Taq Master Mix (PROTECH) 以總體積 50  $\mu$ l 進行 PCR 反應，獲得之擴增產物以膠體電泳確認長度後，以載體:擴增產物 = 1:3 莫耳比例條件下，選殖至 pGEM-T Easy Vector (Promega, USA) 並轉型至 *E. coli* DH5 $\alpha$ ，隨後將菌液均勻塗在含 Ampicillin (100  $\mu$ g/ml) + IPTG (0.5 mM) + X-Gal (80  $\mu$ g/ml) 的 LB Agar 固態培養基中，置於 37  $^{\circ}$ C 繼夜培養後，針對不同擴增產物各挑選 4 顆經固態培養長出之白色菌落，以含 0.2  $\mu$ l T7 promoter (10  $\mu$ M)、0.2  $\mu$ l SP6 promoter (10  $\mu$ M)、1  $\mu$ l reverse primer (10  $\mu$ M)、4  $\mu$ l 2X Fast-Run Advanced Taq Master Mix (PROTECH) 於總體積 50  $\mu$ l 下進行 PCR 反應，並以膠體電泳確認篩選菌落所轉入之 DNA 片段長度，經確認後之菌落以液體培養基(LB Broth, Miller)於 37  $^{\circ}$ C 繼代培養，隨後萃取質體 DNA(GeneSpin Miniprep Purification Kit, PROTECH)，並以 ABI Prism 3730 DNA 定序儀 (Applied Biosystems, USA) 進行定序。



### 2.2.3 *BrMs* 與 *BnRf* ortholog 之等位性分析

將 NH80-A 與 TA95 之 *BnRf* orthologs 基因序列，以 BioEdit 7.2.0 軟體進行比對，分析二品系間之多型性位點，再擷取具 SNP 或 Indel 之周邊序列設計成 KASP 分子標誌，並以該分子標誌對本研究前述之 NH80-A × TA95 之 F<sub>2</sub> 分離族群裔進行基因分型，基因分型時 Real-Time PCR 之設定與基因型判斷方式同前述 SNP 分子標誌之分析方式，所獲得之各單株之基因型資料再與雄不稔性之外表型進行基因連鎖分析，以判別 *BrMs* 與 *BnRf* ortholog 是否具等位性。

## 2.3 *BrMs* 基因之精細定位

### 2.3.1 基因體定序及序列分析

為了探勘更多位於 *BrMs* 基因周邊之分子標誌，以利用於精細定位該基因，本研究擬透過全基因體定序，比較參試品系間基因座落區域之序列資訊，進而設計一系列分子標誌並於更大的分離族群中進行分析，以有效限縮該基因座落之範圍，以建立據圖選殖之基礎。進行時以雄不稔株 NH80-A 與恢復系 TA95 之葉片萃取 DNA 後，送至華大基因 (BGI) 進行 DNA 建庫與全基因體定序，構建 DNA library 時，DNA 片段之 Insert size 設定為 500 bp，建庫完成後以 Illumina Hiseq 2000 定序平台 (Illumina Inc., USA) 進行基因體定序，進行時採片段雙端 (paired-end) 定序，各端讀序長度為 100 bp。

經基因體定序獲得之初級讀序以軟體 CLC Genomics Workbench v8.0.2 (CLC Bio, <http://www.clcbio.com>) 進行品質篩選 (quality trim)，保留序列品質分數大於 Q20 之讀序，以下載自 Brassica databas (BRAD) (<http://brassicadb.org/brad/index.php>) 的 *B. rapa* var. *chiifu40* (v1.5) 為參考序列 (Cheng et al., 2011)，以進行序列比對組裝。序列比對時，各鹼基之錯誤配對 (mismatch)、插入 (insertion) 及缺失 (deletion) 之加扣分參數採 CLC Genomics Workbench v8.0.2 之預設值，進行序列比對時的長度比率 (length fraction) 設定為 0.5，序列相似比率 (similarity fraction) 設定為 0.8，當讀序與參考序列間發生非專一性配對 (non-specific match) 時，則採逢機方式比對至其中一個位置。

除了將 NH80-A 與 TA95 之基因體序列資料與參考序列比對組裝之外，也將該二品系之序列資料進行無參考序列組裝（de novo assembly），進行時各項參數均採 CLC Genomics Workbench v8.0.2 之預設值，完成 de novo assembly 後，將該二品系組裝後之所有 contigs 分別建構為該品系之序列資料庫。

### 2.3.2 開發精細定位用 SNP 分子標誌

依據本研究第三章及第四章之基因定位結果，*BrMs* 基因座落於第 7 號染色體的 6.6-7.6 Mb 區域之間，而本章將 GBS markers 轉成 KASP SNP 分子標誌，並以 1,590 株之 F<sub>2</sub> 族群進行驗證之結果，進一步將 *BrMs* 限縮於該區域之 400 kb 範圍內，因此針對該區域內之 DNA 序列進行分析。將二參試品系的全基因體序列資料比對參考序列而完成組裝後，針對前述特定之 400 kb 區域內，依序各擷取約 10 kb 序列以 BioEdit 7.2.0 軟體進行品系間相互比對，當品系間探勘出 SNP 位點時，進一步將該 10 kb 序列，於前述經 de novo assembly 而建立的 NH80-A 與 TA95 序列資料庫中，分別進行 local blast 分析，當搜尋出該序列於二品系對應之特定 contig 後，再將該二個 contigs 以 BioEdit 7.2.0 軟體再次進行比對，以確認 SNP 位點序列資訊無誤，再以該 contig 上 SNP 兩側的 DNA 序列資訊設計為 KASP SNP 分子標誌。

### 2.3.3 精細定位用分離族群之建構與基因分型

為了提高能 *BrMs* 基因精細定位之解析度，2016 年 9 月擴大培育雄不稔株 NH80-A 與恢復系 TA95 雜交後裔之 F<sub>2</sub> 族群共 4,415 株，做為基因精細定位用族群，其開花期之誘導與雄不稔性狀外表型之調查方法與前述 1,590 株組成之 F<sub>2</sub> 族群相同，基因分型時，先以 GBS 試驗中與 *BrMs* 相鄰的 BrA7\_6632K 與 BrA7\_7038K 對整個族群之 4,415 株進行分析，篩選出位於該二個 SNP 分子標誌間產生重組的個體（recombinants），再利用源自基因體定序而開發的 SNP 分子標誌分析重組個體的基因型。



### 3 結果與討論

#### 3.1 粗定位之 GBS markers 的驗證

##### 3.1.1 將 GBS markers 轉換為 SNP 分子標誌

為了應用於 GMS 之分子輔助選拔，參考第三章以 GBS bin markers 進行粗定位之結果，將與 *BrMs* 相鄰或共分離之 tags 上的 SNP 位點設計為 SNP 分子標誌，並以 1,590 株 NH80-A × TA95 之 F<sub>2</sub> 族群評估這些 SNP 分子標誌進行篩選的精確性，在所有設計的 SNP 分子標誌中，由於“BrA7\_7597K”對於基因型間之區別效果不佳，而改以“BrA7\_7625K”取代之，最終將 GBS markers 轉換為 3 個 SNP 分子標誌，其螢光訊號清晰，為可明顯分群之共顯性分子標誌（圖 5-1）。

##### 3.1.2 結球白菜 *Ms* 基因之 SCAR 分子標誌轉換為 SNP 分子標誌

為了比較不結球白菜與結球白菜二個亞種間 GMS 基因之異同，二個與結球白菜 *Ms* 基因相關之 SCAR 分子標誌 “syau\_scr01”與“syau\_scr04”（Feng et al., 2009）也用於分析相同之 1,590 株 F<sub>2</sub> 族群，二個 SCAR 分子標誌於族群親本 NH80-A 與 TA95 間均可擴增出預期之 PCR 產物，然彼此間卻無長度上的多型性，經定序分析後顯示“syau\_scr01”之擴增產物於二品系間完全一致，而“syau\_scr04”之擴增產物於 NH80-A 與 TA95 間具有 10 個多型性位點（圖 5-2）。利用其中 1 個多型性位點設計成 SNP 分子標誌，並名為“syau\_scr04\_SNP”（表 5-2），以該 SNP 分子標誌對 F<sub>2</sub> 族群進行基因分型時，其螢光訊號清晰，為可明顯分群之共顯性標誌（圖 5-3）。

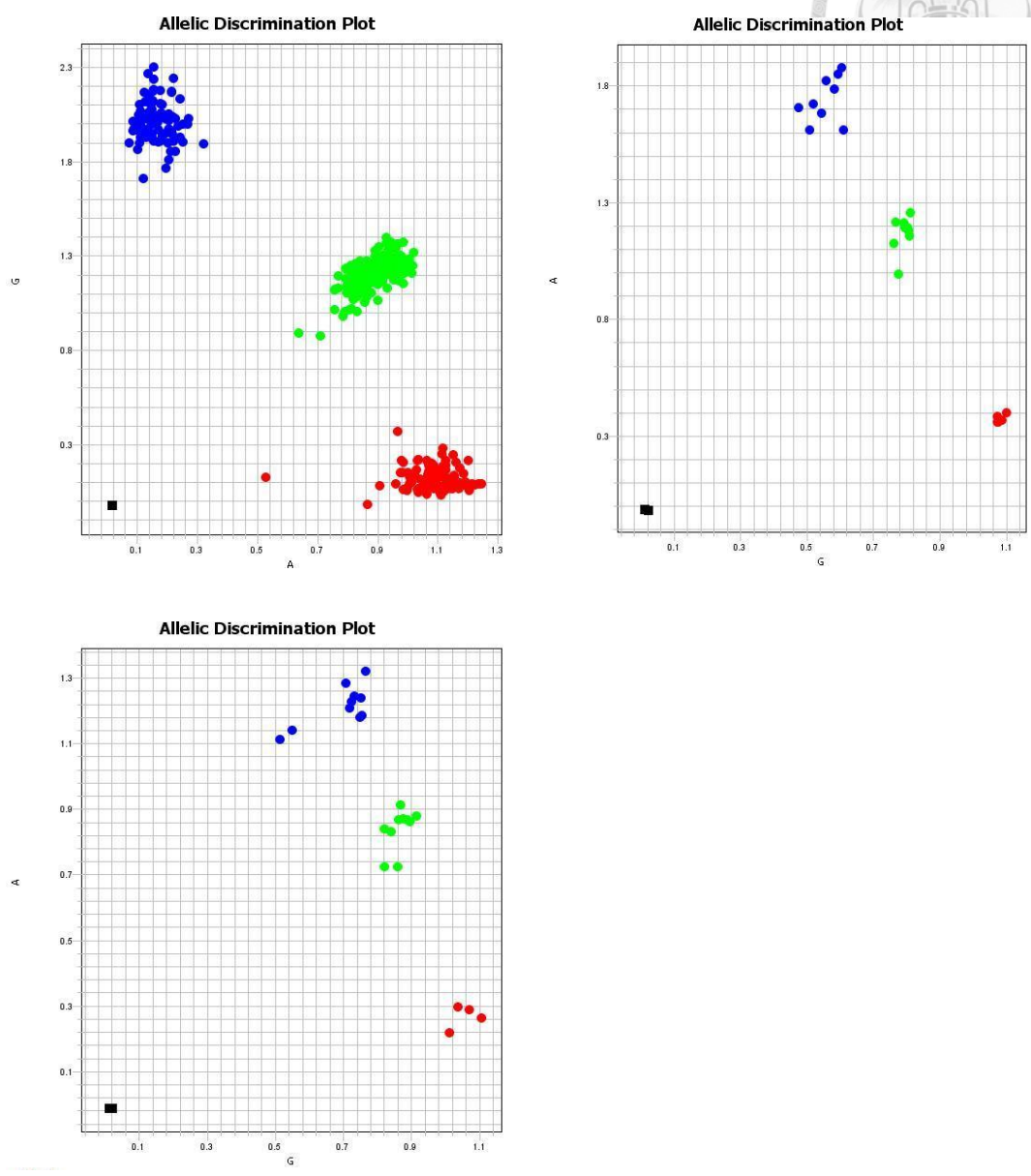
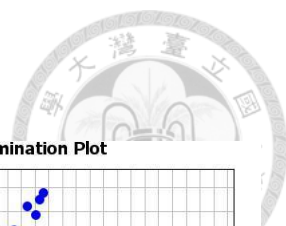


圖 5-1 將 GBS markers 轉換為 SNP 分子標誌後之基因分型情形  
BrA7\_6632K (左上), BrA7\_7038K (左下), BrA7\_7625K (右上)

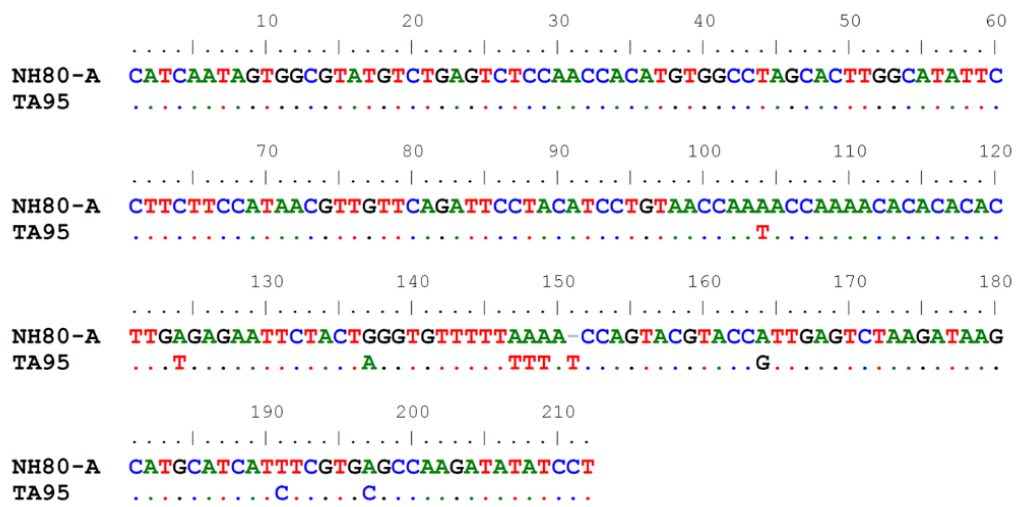


圖 5-2 syau\_scr04 引子擴增產物於 NH80-A 與 TA95 間之多型性位點

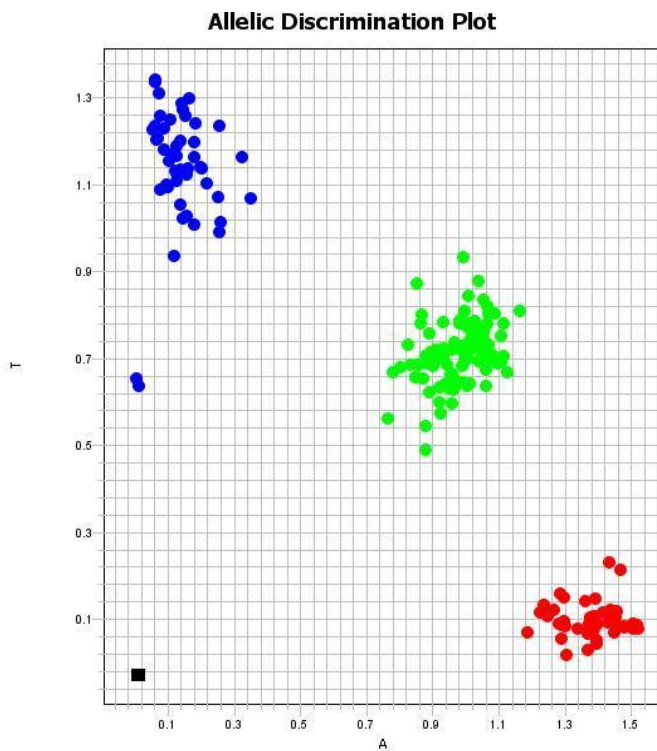


圖 5-3 以 syau\_scr04\_SNP 進行基因分型之情形



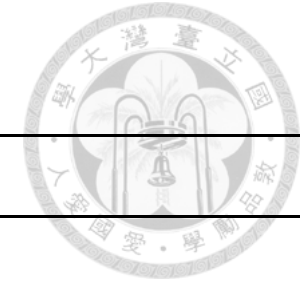


表 5-2 以 KASP 原理設計的分標子引子序列

SNP marker	Tag	KASP Assay	Sequence
BrA7_6632K	F2_21222	Allele 1 Primer	5'-CACCAACAACGCATCATCGTCC-3'
		Allele 2 Primer	5'-CACCAACAACGCATCATCGTCT-3'
		Common Primer	5'-GTTTGACCTAGCTCGACGAAACCAA-3'
BrA7_7038K	F2_21271	Allele 1 Primer	5'-CTGAAGCTTTGTCGATTTTGTAAGTAATA-3'
		Allele 2 Primer	5'-CTGAAGCTTTGTCGATTTTGTAAGTAATG-3'
		Common Primer	5'-CCGGCAACGCTTCCTTTACCTCTA-3'
BrA7_7625K	F2_21322	Allele 1 Primer	5'-AAGAAGTTTGATGAAGAGGTTTCTTCAATA-3'
		Allele 2 Primer	5'-GAAGTTTGATGAAGAGGTTTCTTCAATG-3'
		Common Primer	5'-CACCAAATTACCCAACCTGTTCATTTCTT-3'
syau_scr04_SNP	-	Allele 1 Primer	5'-TGAGTCTAAGATAAGCATGCATCATC-3'
		Allele 2 Primer	5'-GAGTCTAAGATAAGCATGCATCATT-3'
		Common Primer	5'-TTAACACAGGTTGCTAACAGGATATATCTT-3'
BrMs_indel	-	Allele 1 Primer	5'- CGATAAGATCCCTCTTAAGTCTTAAG -3'
		Allele 2 Primer	5'- CGATAAGATCCCTCTTAAGTCTTAAC -3'
		Common Primer	5'- CCAGTACAACAATTGTAGATACAGAGACAA -3'



### 3.1.3 SNP 分子標誌與 *BrMs* 基因之連鎖分析結果

以 1,590 株 F<sub>2</sub> 分離族群進行連鎖分析，其中“BrA7\_6632K”及“BrA7\_7625K”為源自與 *BrMs* 相鄰 tags “F2\_21222”及“F2\_21322”開發而得之 SNP 分子標誌，分析結果顯示該二者位於 *BrMs* 基因兩側而分別相距 0.3 及 1.3 cM（圖 5-4），此 1.6 cM 區間小於第 3 章 GBS markers 定位結果之 2.4 cM（表 3-2），此外，“BrA7\_7038K”是依據與 *BrMs* 共分離 tag “F2\_21271”上 SNP 位點而設計之分子標誌，連鎖分析結果顯示其與“BrA7\_7625K”位於同一側，但最接近 *BrMs* 距離僅 0.3 cM，因此 *BrMs* 基因被限縮於此 0.6 cM 範圍內，而此二個 SNPs 間的物理距離僅約 400 kb。

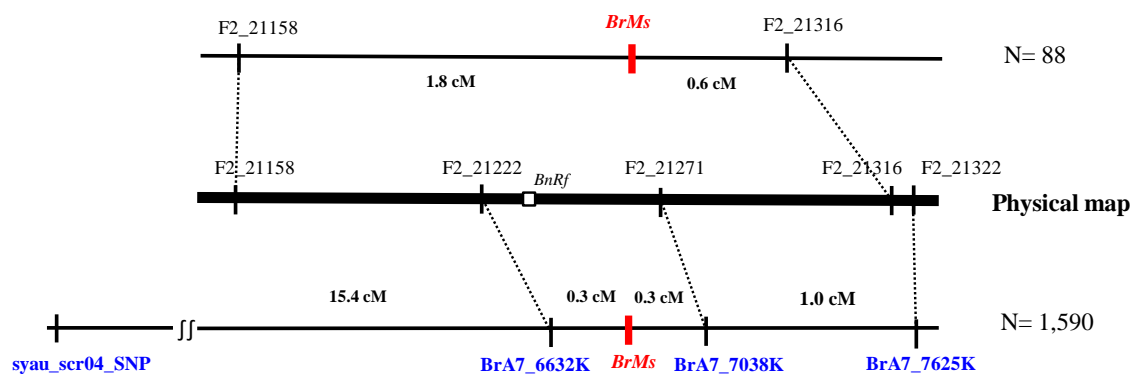


圖 5-4 源自 GBS tags 及 SCARs 的 SNP 分子標誌所建構之遺傳連鎖圖

不結球白菜與結球白菜間屬於同一物種 (species) 下的不同亞種 (subspecies)，由於本研究不結球白菜之 *BrMs* 基因與結球白菜之 *Ms* 基因均位於相同之 A07 染色體 (Feng et al., 2009)，探討其是否為相同雄不稔基因及相關分子標誌的適用性將有助於分子選種之進行，我們成功的將結球白菜 GMS 的 SCAR marker “syau\_scr04” 轉換為 SNP 分子標誌並用於分析不結球白菜 F<sub>2</sub> 族群，結果顯示 syau\_scr04\_SNP 也位於 A07 連鎖群，但與不結球白菜 *BrMs* 間相距 15.4 cM（圖 5-4），此遺傳距離與 Feng et al. (2009) 的研究中描述 syau\_scr04 與結球白菜 *Ms* 相距 2.5 cM 之結果有明顯差別。

### 3.1.4 不結球白菜雄不稔 *BrMs* 基因與結球白菜 *Ms* 基因間可能之關係

儘管目前結球白菜 *Ms* 基因之實際物理位置並無發表，無法確認其與不結球白菜 *BrMs* 基因間之關係，且於不同族群間所估算之遺傳重組距離可能有差異，然依據本研究結果顯示，*syau\_scr04\_SNP* 與 *BrMs* 基因間之遺傳距離明顯大於 *syau\_scr04* 與 *Ms* 之遺傳距離，初步推測 *BrMs* 與 *Ms* 可能為同連鎖群上之不同基因座。事實上，此種同一連鎖群具有二個不同雄不稔基因的現象在大油菜上已有先例，即大油菜品系“9012A”之複等位基因 *BnRf* 定位於 N7 連鎖群 (Xiao et al., 2008; Xie et al., 2012)，而源自品系“S45A”之 GMS 基因 *Bnms1* 也被定位於相同連鎖群 (Yi et al., 2006)，但二個 GMS 基因不具等位性 (Chen et al., 1998)，說明大油菜之 N7 連鎖群上至少存在二個不等位之 GMS 基因座，對照於本研究結果，*BrMs<sup>a</sup>/BrMs<sup>b</sup>* 及 *BrMs<sup>b</sup>/BrMs<sup>c</sup>* 都被定位於 A7 連鎖群上的相同狹小範圍內，顯示這些等位基因均位於不結球白菜的 *BrMs* 基因座上，至於遺傳距離較遠的結球白菜 *Ms* 基因座上可能存在其他雄不稔基因。

### 3.1.5 SNP 分子標誌開發策略及應用於 MAS 之潛力

本研究為了育種目的而設計的雄不稔性 SNP 分子標誌可有效輔助選拔，其中“BrA7\_6632K”及“BrA7\_7038K”是依據 GBS 試驗所產生的資訊進一步開發而得，策略上為有效利用基因體定序資源，因此先以小族群 (88 株) 進行粗定位 (coarse mapping)，其定位區間為 2.4 cM，然而受益於 GBS 技術可獲得大量含有物理位置之 SNP 位點資訊，因此可進一步於定位區間內設計少量 SNP 分子標誌，再於大族群中 (1,590 株) 中驗證其鑑別效率並準確估計遺傳距離，依此方式開發的二個 SNP 分子標誌與 *BrMs* 基因間均只相距 0.3 cM (圖 5-4)，實際以該二者檢測 1,590 株 F<sub>2</sub> 族群進行驗證時，其基因分型結果與雄不稔外表型間各只出現 3 個重組株 (3/1590 = 0.19%)，顯示應用於育種篩選時，其中任何一個標誌即可達良好的分子輔助選拔效果，若同時使用與這二個 SNP 分子標誌進行分析則可更加提高篩選的精確度，此外，本研究也成功的將結球白菜 GMS 之 SCAR 分子標誌 “*syau\_scr04*”轉換為 SNP 分子標誌，雖然結球白菜之雄不稔基因 *Ms* 與本研究之 *BrMs* 是否為相同基因尚有待確認，假若該二個基因確實為白菜物種中之不同雄不稔基因，在育種實務上，本研究開發之二個不結球白菜 *BrMs* 基因之 SNP 分子標誌加上結球白菜 *Ms* 基因之分

子標誌 syau\_scr04\_SNP，將助益於 *B. rapa* 之各亞種如結球白菜、不結球白菜或蕪菁等作物之雄不稔系選育。



### 3.1.6 近緣種大油菜之 *BnRf* 基因位於不結球白菜 *BrMs* 之定位區間

由於大油菜之 N7 連鎖群與白菜 A7 連鎖群演化自共同祖先之相同染色體 (<http://www.brassica.info/resource/maps/lg-assignments.php>)，基因體比較之結果顯示 *BnRf* 基因區間在大油菜 N7 連鎖群與白菜 A7 連鎖群間具有高度共線性，*BnRf* 的 13.8 kb 基因區間可對應至結球白菜 A7 連鎖群之 17.8 kb 範圍內，這區間預測包含三個基因 “*Bra014989*”、“*Bra014990*”及“*Bra014991*” (Xie et al., 2012)，以我們所使用的參考序列進行比對，發現 *BnRf* 的定位區間也座落於本研究 *BrMs* 之 400 kb 定位區間內 (圖 5-4)，推測大油菜 *BnRf* 基因很可能與本研究之 *BrMs* 基因具同源性，若此，將可證明大油菜 (AACC 基因體) 的 *BnRf* 基因源自不結球白菜 (AA 基因體)。隨後在大油菜的研究中，透過據圖選殖而確認了 *BnRf* 基因 (Deng et al., 2016)，研究結果顯示大油菜雄不稔等位基因 *BnRf<sup>b</sup>* 雖位於該區域，但並非前述之 “*Bra014989*”、“*Bra014990*”或 *Bra014991*”這 3 個基因，*BnRf<sup>b</sup>* 有可能為白菜基因組以外的片段，因此，蕓苔屬中各種不同 GMS 的基因結構仍需進一步解析，而本研究對於 *BrMs* 之基因定位結果，將有助於該基因進行據圖選殖之參考。

## 3.2 不結球白菜 *BrMs* 與大油菜 *BnRf* 之關係

### 3.2.1 不結球白菜 *BnRf* orthologs 間序列之選殖與比較

由於大油菜 *BnRf* 與本研究不結球白菜 *BrMs* 之定位區間重疊，為了釐清二者之關係，擬依據 *BnRf* 序列資訊分析不結球白菜的同源基因結構，依據大油菜 *BnRf* 基因序列 (及 TA95 的基因體資訊) 所設計的引子組，對雄不稔株 NH80-A 與恢復系 TA95 進行 PCR 結果如圖 5-5 所示，所有引子組均可擴增出專一性片段，其中 lane 1-4 為 NH80-A 利用 *BnRf* 基因的引子進行 PCR 之產物，而 lane 6-9 則為 TA95 以相同引子進行 PCR 之結果，其中除 lane 1 及 lane 6 之擴增產物長度小於預期外，其餘擴增片段大小均與 *BnRf* 基因預期一致，也就是說不結球白菜之 *BnRf* ortholog，其基因總長可能小於 *BnRf*，而 lane 6 長度小於 lane 1，顯示 TA95 之 *BnRf* ortholog 基因總長也可能小於 NH80-A。此外，lane 5 及 lane 10 是依據 TA95 基因體組裝序列中，



*BnRf* ortholog 基因上游設計的引子對所進行 PCR 之結果，顯示 TA95 之擴增片段也小於 NH80-A。

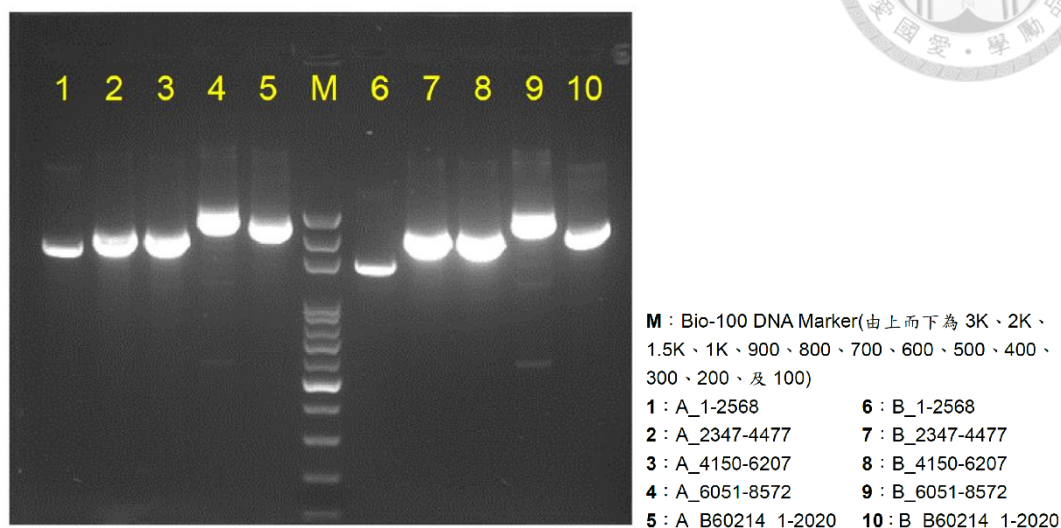


圖 5-5 針對不結球白菜 *BnRf* orthologs 不同區域進行 PCR 之結果  
A : NH80-A; B : TA95, lane 1-4 及 lane 6-9 為依據 *BnRf* 基因序列設計引子之 PCR 結果, lane 5 及 lane 10 為參考 TA95 基因體中 *BnRf* ortholog 基因上游序列設計引子之 PCR 結果

進一步將參試品系的擴增片段進行解序並組裝後，將大油菜 *BnRf* 二個等位基因 (KT818624 及 KT818625) 及不結球白菜之二個 *BnRf* orthologs 進行比對的結果如圖 5-6 所示。大油菜中 KT818624 及 KT818625 基因全長分別為 8,572 及 7,726 bp，然而雄不稔株 NH80-A 與恢復系 TA95 的 *BnRf* orthologs 基因全長僅為 7,562 及 7,411 bp，儘管基因長度均不同，但四個 *BnRf* orthologs 間所有外顯子 (exon) 的序列完全一致，主要的差異在啟動子 (promoter) 的區域，以大油菜 KT818624 的基因序列為基礎，KT818625 在轉譯起始點上游 1,410 bp 處有一 850 bp 缺失 (deletion)，而不結球白菜 NH80-A 與 TA95 之 *BnRf* orthologs 在相同區域則分別有 1,015 bp 及 1,180 bp 缺失，這同時也說明了 PCR 結果中 lane 6 的擴增片段長度小於 lane 1 的原因，此外，*BnRf* orthologs 間之 non-coding regions 則發現 4 個插入/缺失 (indel) 及 2 個 SNP 位點 (圖 5-6)。比對基因定序結果顯示 NH80-A 與 TA95 的 *BnRf* orthologs 之間，在第 4 個內含子 (intron) 存在一個 14 bp 的 indel 變異 (圖 5-7)。

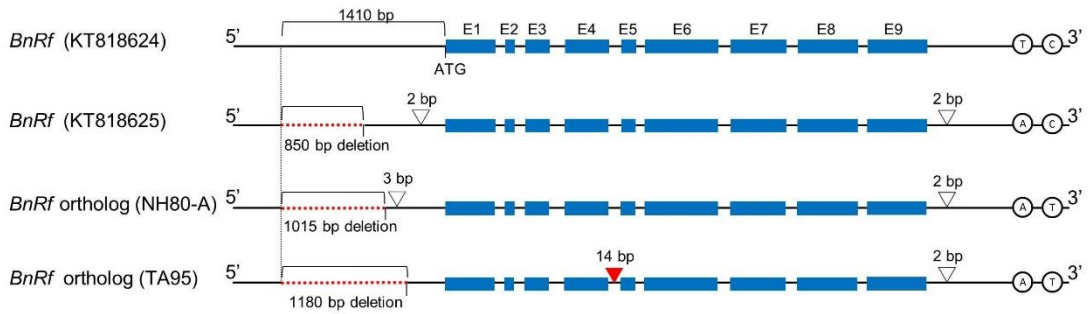


圖 5-6 不結球白菜 *BnRf* orthologs 與大油菜 *BnRf* 基因結構之比較

由上至下分別為大油菜雄不稔系 *BnRf* 等位基因 (KT818624)、大油菜可稔系 *BnRf* 等位基因 (KT818625)、不結球白菜雄不稔系 (NH80-A) *BnRf* ortholog 及不結球白菜恢復系 (TA-95) *BnRf* ortholog，倒三角形表示發生 indel 的位置

```

      10      20      30      40      50      60
      |.....|.....|.....|.....|.....|.....|
BnRf (KT818624)  GTAAATTCCTCTGCCACTTTTCATTTATTTATTTTCTTCTTATACATCTTATTTCGATAAG
BnRf (KT818625)  GTAAATTCCTCTGCCACTTTTCATTTATTTATTTTCTTCTTATACATCTTATTTCGATAAG
BnRf_ortholog (NH80-A) GTAAATTCCTCTGCCACTTTTCATTTATTTATTTTCTTCTTATACATCTTATTTCGATAAG
BnRf_ortholog (TA95)  GTAAATTCCTCTGCCACTTTTCATTTATTTATTTTCTTCTTATACATCTTATTTCGATAAG

      70      80      90      100     110     120
      |.....|.....|.....|.....|.....|.....|
BnRf (KT818624)  ATCCCTCTTAAGTCTTAA-----CCTTTGCTCTGTATCTACAATTGTTGT
BnRf (KT818625)  ATCCCTCTTAAGTCTTAA-----CCTTTGCTCTGTATCTACAATTGTTGT
BnRf_ortholog (NH80-A) ATCCCTCTTAAGTCTTAA-----CCTTTGCTCTGTATCTACAATTGTTGT
BnRf_ortholog (TA95)  ATCCCTCTTAAGTCTTAAAGTTAAGACTCTTAACCTTTGCTCTGTATCTACAATTGTTGT

      130     140     150     160
      |.....|.....|.....|.....|
BnRf (KT818624)  ACTGGTCTAGAGTCATTCTTCTTCCCTGATTGTTTAATTCAG
BnRf (KT818625)  ACTGGTCTAGAGTCATTCTTCTTCCCTGATTGTTTAATTCAG
BnRf_ortholog (NH80-A) ACTGGTCTAGAGTCATTCTTCTTCCCTGATTGTTTAATTCAG
BnRf_ortholog (TA95)  ACTGGTCTAGAGTCATTCTTCTTCCCTGATTGTTTAATTCAG

```

圖 5-7 *BnRf* 與 *BnRf* orthologs 於第 4 個 intron 之序列差異



### 3.2.2 不結球白菜 *BrMs* 與油菜 *BnRf* 之等位性

為了研究 *BrMs* 與 *BnRf* 的等位性，將 *BnRf* orthologs 第 4 個內含子的 14 bp 多型性區域設計為 KASP assay，因而成為一個可藉由螢光反應區分二品系間 *BnRf* ortholog 基因型的 Indel 分子標誌，將此標誌名為 *BrMs\_indel*，以之對 NH80-A × TA95 之 F<sub>2</sub> 後裔進行基因分型，並與雄不稔外表型進行連鎖分析，結果參試的 923 個 F<sub>2</sub> 植株中，*BrMs\_indel* 分析的基因型與外表型完全一致，即此分子標誌與雄不稔性狀完全共分離，暗示 *BrMs* 與 *BnRf* 可能為同源基因，且於演化過程基因序列上產生了一些變異。

在 *BnRf* orthologs 之間，於啟動子區域所發現之長片段的插入或缺失可能為影響雄不稔基因表現的重要因子，在油菜雄不稔品系 RG206A 中，其雄不稔等位基因 (KT818624) 於啟動子區域較雄可稔等位基因 (KT818625) 多了一段 850 bp 序列，可能導致該基因轉錄效率大幅減低，而為造成雄不稔性之主因 (Deng et al., 2016)，而本研究的雄不稔系 NH80-A，其 *BnRf* ortholog 之啟動子相同區域也比恢復系 TA95 多了 165 bp 的插入片段，因此推測此啟動子區域 DNA 序列之變異性，有可能影響白菜類或其近緣物種雄不稔基因之表現，進而影響稔性表現之差異，因此，若能透過基因選殖或其他研究策略，分析此關鍵區域的變異情形，及其與雄不稔性狀表現之相關性，則可說明蕓苔屬作物間，細胞核雄不稔基因結構與功能性之關係。

## 3.3 *BrMs* 基因之精細定位

### 3.3.1 不結球白菜基因體定序與組裝結果

NH80-A 進行次世代定序所獲得之基因體序列資料如表 5-3 所示，本次定序的長度設定為 100 bp，原始讀序經過序列品質篩選後保留 100,865,760 個讀序，進一步以參考序列為模版進行組裝，組裝完成總計可比對至參考序列者，約佔所有讀序的 91.5%，這些讀序的平均長度約 99.94 bp，而無法比對至參考序列者約佔 8.5%，由於這類讀序的平均長度為 99.95 bp，它與可比對至參考序列的讀序長度一致，因此推測無法正確比對至參考序列者，主要為參試品系基因體與採用的參考序列間，存在遺傳變異所致。

表 5-3 不結球白菜 NH80-A 及 TA95 經次世代定序所獲得讀序摘要

	NH80-A	TA95
Mapped reads	92,299,418 (91.51%)	83,991,652 (90.95%)
Average length of mapped reads	99.94	99.94
Not mapped reads	8,566,342 (8.49%)	8,360,830 (9.05%)
Average length of not mapped reads	99.95	99.95
Total reads	100,865,760	92,352,482

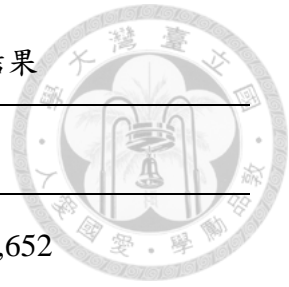
進一步針對已完成比對之讀序進行分析，與參考序列比對過程中，屬於非專一性配對 (non-specific matches) 的約佔 18.44% (表 5-3)，所謂“non-specific matches”指的是特定讀序在整個參考序列中不只比對到一個位置，在參數設定上可選擇排除這些序列，或將其逢機對應到其中的一個位置，但由於植物的種化過程中本來就存在許多序列重複 (duplicate) 的現象，基因體上不同位置自然就有許多相同或相似的序列，致使序列比對時產生一定比例的“non-specific matches”，若參數設定上將這些讀序刪除，則組裝後的基因體勢必會產生更多缺口 (gap)，因此試驗中仍選擇將這些讀序逢機對應到其中一個序列相同之位置，由於參考序列的總長度為 284,855,3731 bp，NH80-A 完成組裝序列 (consensus sequence) 之總長度為 251,861,735 bp，約涵蓋參考序列的 88%，而定序深度則達 31.5 倍 (表 5-4)。

TA95 進行次世代定序所獲得之基因體序列資料中，經過品質篩選之後共獲得 92,352,482 個讀序，其中可比對至參考序列上者約佔 90.95%，平均長度約 99.94 bp。而無法比對至參考序列者約佔 9.05%，相較於 NH80-A 而言，有較多讀序無法比對至參考序列上，可能它與參考序列基因體之遺傳距離較 NH80-A 遠，可與參考序列完成比對的讀序當中，屬 non-specific matches 者約佔 17.01%，TA95 完成組裝的序列長度為 249,304,092 bp，相似於 NH80-A 之結果，約佔參考序列的 88% (表 5-4)，由於所採用的參考序列屬結球白菜，因此這 88% 覆蓋率主要應是不結球白菜與結球白菜演化過程產生序列變異之結果。



表 5-4 不結球白菜 NH80-A 及 TA95 讀序比對至參考序列的結果

	NH80-A	TA95
Mapped reads	92,299,418	83,991,652
Non-specific matches	17,023,598 (18.44%)	14,283,498 (17.01%)
Mapped read length (bp)	9,224,861,154	8,394,542,513
Total consensus length (bp)	251,861,735	249,304,092
Total reference length (bp)	284,855,373	284,855,373
Reference coverage (%)	88	88
Depth (X)	31.52	28.62



### 3.3.2 開發精細定位用 SNP 分子標誌

本研究將 GBS markers 轉成 KASP SNP 分子標誌後，以 1,590 株 F<sub>2</sub> 族群進行驗證之結果將 *BrMs* 限縮於該區域之 400 kb 範圍內，因此利用 NH80-A 及 TA95 進行全基因體定序所獲得之 consensus sequences，針對該 400 kb 區域之 DNA 序列相互比對，以探勘彼此之間的多型性位點，並於該區域內共設計了 8 個 SNP 分子標誌，連同 GBS 試驗開發之 BrA7\_6632K 與 BrA7\_7038K 分子標誌，進一步取這些 SNP 周邊約 60-100 bp 序列於 Brassica database (<http://brassicadb.org/brad/>) 依據結球白菜基因體 *B. rapa* (chromosome V1.5) 及大油菜基因體 *B. napus* (chromosome V1.0) 為參考序列進行比對 (blastn)，10 個 SNP 位點於二個物種中之比對結果如圖 5-8 所示，10 個 SNP 中有 8 個可同時比對至 *B. napus* 之 A07 染色體上及 *B. napus* 之 C07 染色體上；而其餘 2 個 SNP 中，BrA7\_6795K 僅位於結球白菜及油菜之 A07 染色體，在油菜之 C07 染色體未發現該位點，而 BrA7\_6862K 僅在結球白菜基因體發現。

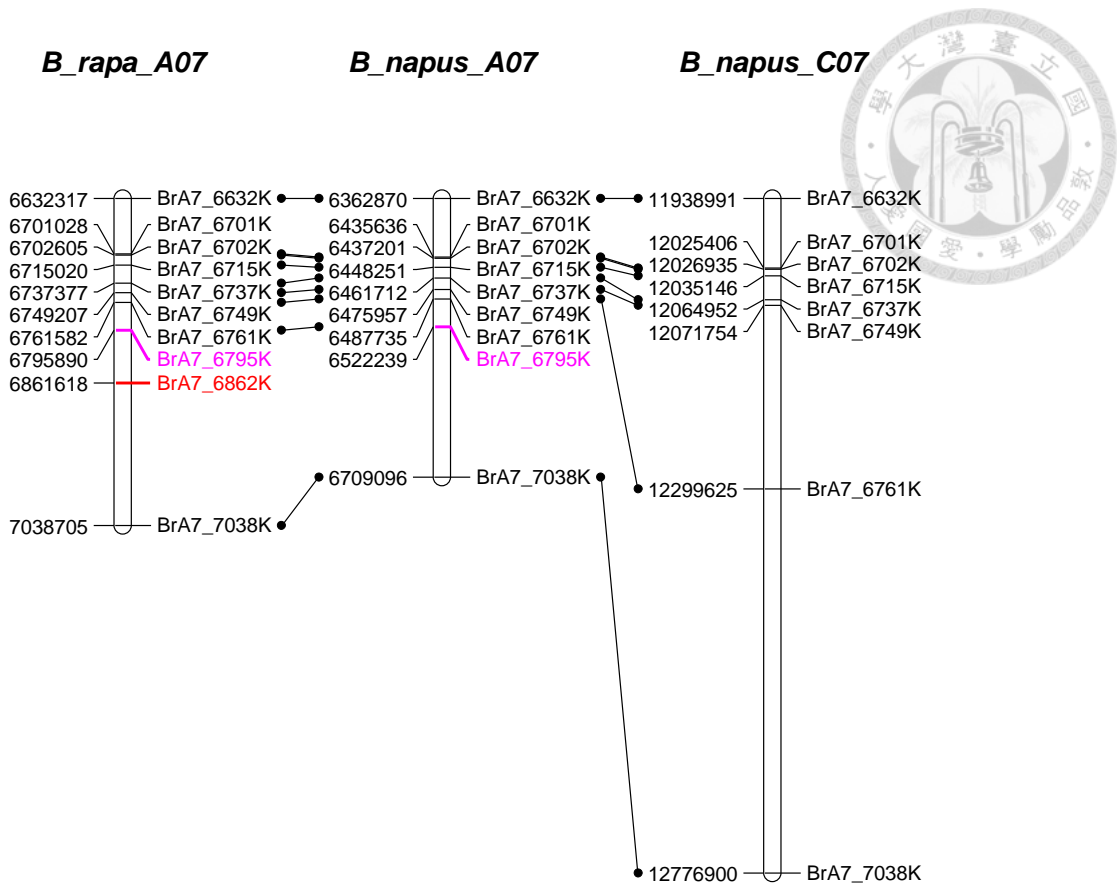


圖 5-8 SNP 分子標誌序列比對至 *B. rapa* 及 *B. napus* 基因體的物理位置

依據 SNP 周邊序列之比對結果，顯示這些 SNPs 不論在結球白菜之 A07 染色體或大油菜之 A07、C07 染色體，均具共線性且順序一致，依據各 SNPs 間之相對位置而言，結球白菜之 A07 染色體與大油菜之 A07 染色體最為相似，此符合大油菜之 N7 連鎖群（即本研究之大油菜 A07 連鎖群）與白菜 A7 連鎖群演化自共同祖先相同染色體（<http://www.brassica.info/resource/maps/lg-assignments.php>）的前人研究結果。在此 10 個 SNPs 範圍內，結球白菜之染色體區間較大油菜之染色體區間長，且大油菜 A07 染色體已失去 BrA7\_6862K 位點，暗示演化過程中大油菜 A07 染色體在 BrA7\_6795 至 BrA7\_7038K 間可能發生部分缺失，若以 BrA7\_6633K 至 BrA7\_6795K 間的 162 kb 範圍而言，二個染色體間一致性相當高，長度也只差異約 4 kb，顯示這區域於演化過程中仍維持保守性，適合做為比較基因體之定位（comparative mapping）用，此與 Xie et al.（2012）發現，*BnRf* 基因區間在大油菜 N7 連鎖群與結球白菜 A07 連鎖群間具有高度一致性的結果相似。然而，本研究另發現不結球白

菜 *BrMs* 基因座落之 400 kb 區域，除了與結球白菜、大油菜之 A07 染色體具高度一致性外，也與大油菜之 C07 染色體具有共線性關係，大油菜 C07 染色體上之相對區域也可能源自相同祖先之同源染色體，但在該 C07 染色體上，BrA7\_6749K 與 BrA7\_6761K 間的序列距離明顯增長，而 BrA7\_6761k 與 BrA7\_7038K 間不僅序列距離增長，且 BrA7\_6795K 及 BrA7\_6862K 等序列均已遺失（圖 5-8），顯示 C07 染色體 BrA7\_6749k 至 BrA7\_7038K 間已產生較大之序列變異性。

除了利用全基因體定序開發出 8 個新的 SNP 分子標誌之外，本研究也針對 NH80-A 與 TA95 進行無參考序列之基因體組裝，並建立成二個序列資料庫，依據前述 8 個 SNP 周邊序列對 NH80-A 與 TA95 的序列資料庫進行比對（local blast）以確認位點時，發現一組 contigs 可探勘出 NH80-A 與 TA95 間之 SNP 位點，但該區域不存在參考序列上，在本研究也將該位點開發為 SNP 分子標誌。整體而言，透過基因體定序共開發 9 個 SNP 分子標誌，連同 GBS 試驗開發之 BrA7\_6632K 及 BrA7\_7038K，共 11 個 SNP 分子標誌用於 *BrMs* 之精細定位。

### 3.3.3 精細定位 *BrMs* 基因

進一步利用前述 11 個 SNP 分子標誌及 1 個源自不結球白菜 *BnRf* ortholog 之 *BrMs*\_indel，以 4,415 株 NH80-A × TA95 之 F<sub>2</sub> 族群對 *BrMs* 進行基因精細定位。首先以 BrA7\_6632K 與 BrA7\_7038K 分子標誌對整個族群進行分析後，在 4,415 株中發現 61 個體於該 2 個分子標誌間產生重組（recombination），即此重組個體於該二個分子標誌的基因分型結果中，其中一個分子標誌呈現為異質結合（H）且另一個則呈現與 NH80-A 相同之同質結合基因型（A），或是其中一個分子標誌呈現為異質結合（H）且另一個則呈現與 TA95 相同之同質結合基因型（B）。由於該分離族群是由 *BrMs*<sup>a</sup>/*BrMs*<sup>b</sup> 等位基因所組成，唯有隱性同質基因型 *BrMs*<sup>b</sup>/*BrMs*<sup>b</sup> 的個體（與 NH80-A 相同）才能表現雄不稔性狀，因此在所有 61 株重組體（recombinants）中，僅有 30 株基因分型為 H/A 的重組體，其外表型性狀可能產生稔性分離之差異，可用以進行基因型與外表型之比對，其餘 31 株基因分型為 H/B 的重組體外表型均為雄可稔而無法進行比對，進一步將 30 株基因分型為 H/A 的重組體，以其他 9 個 SNP 分子標誌進行分析，這些重組體之 11 個分子標誌的基因型及在染色體上的斷點如表 5-5 所示。

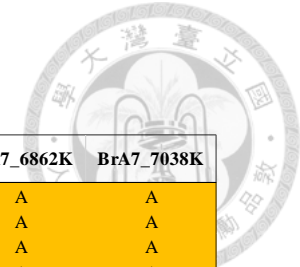



表 5-5 針對 *BrMs* 進行基因精細定位之結果

Sample	Phenotype	BrA7_6632K	BrA7_6701K	BrA7_6702K	BrA7_6715K	BrA7_6737K	Br_Ms_indel	BrA7_6746K	BrA7_6749K	BrA7_6761K	BrA7_6795K	BrA7_6862K	BrA7_7038K
16_796	Ms	H	A	A	A	A	A	A	A	A	A	A	A
16_3080	Ms	H	A	A	A	A	A	A	A	A	A	A	A
16_3484	Ms	H	A	A	A	A	A	A	A	A	A	A	A
16_4792	Ms	H	A	A	A	A	A	A	A	A	A	A	A
16_4293	Ms	H	H	A	A	A	A	A	A	A	A	A	A
16_3055	Mf	A	A	H	H	H	H	H	H	H	H	H	H
16_3927	Mf	A	A	H	H	H	H	H	H	H	H	H	H
16_4710	Mf	A	A	H	H	H	H	H	H	H	H	H	H
16_580	Ms	H	H	H	H	A	A	A	A	A	A	A	A
16_3789	Ms	H	H	H	H	A	A	A	A	A	A	A	A
16_4253	Ms	H	H	H	H	A	A	A	A	A	A	A	A
16_4769	Ms	H	H	H	H	A	A	A	A	A	A	A	A
16_4100	Mf	A	A	A	A	H	H	H	H	H	H	H	H
16_905	Ms	H	H	H	H	H	A	A	A	A	A	A	A
16_265	Mf	A	A	A	A	A	H	H	H	H	H	H	H
16_5220	Mf	A	A	A	A	A	H	H	H	H	H	H	H
16_3157	Ms	H	H	H	H	H	H	A	A	A	A	A	A
16_3236	Mf	A	A	A	A	A	A	H	H	H	H	H	H
16_13	Mf	H	H	H	H	H	H	H	H	A	A	A	A
16_4124	Mf	H	H	H	H	H	H	H	H	A	A	A	A
16_62	Ms	A	Undetermined	A	A	A	A	A	A	A	A	H	H
16_471	Mf	H	H	H	H	H	H	H	H	H	H	H	A
16_525	Mf	H	H	H	H	H	H	H	H	H	H	H	A
16_713	Mf	H	H	H	H	H	H	H	H	H	H	H	A
16_87	Mf	H	H	H	H	H	H	H	H	H	H	H	A
16_282	Mf	H	H	H	H	H	H	H	H	H	H	H	A
16_245	Mf	H	H	H	H	H	H	H	H	H	H	H	A
16_231	Mf	H	H	H	H	H	H	H	H	H	H	H	A
16_4969	Mf	H	H	H	H	H	H	H	H	H	H	H	A
16_704	Ms	A	A	A	A	A	A	A	A	A	A	A	H

表列樣本為 BrA7\_6632K 與 BrA7\_7038K 間之重組體 (recombinants) 以所有分子標誌進行基因分型的結果，其中 A：與 NH80-A 相同之基因型；H：異質結合體；Ms：雄不稔外表型；Mf：雄可稔外表型；橘色邊緣粗體直線表示該重組各體其染色體之斷點



利用全基因體定序而新開發之 9 個 SNP 分子標誌中，有 8 個是以參考序列為模板對 NH80-A 與 TA95 進行序列組裝後，再依二者間序列之差異探勘而得，然而 BrA7\_6746K 並非以參考序列為模板，而是由本身序列進行無參考序列之自行組裝後，再於 NH80-A 與 TA95 序列資料庫中，篩選出對應之 contigs 相互比對而獲得，由於在參考序列上並無此位點，其物理位置是依據 contig 上可比對至參考序列的位置推估而得。所有 SNP 分子標誌依照其物理位置及基因型，可排出各重組體在此區域染色體互換之斷點，但假若分子標誌間的順序並非與物理位置一致，則各重組體在此區域呈現的染色體斷點可能不只一處，然而本研究係以 SNP 分子標誌預期之物理位置為基準，將 30 個重組體的染色體斷點依序排列後，發現各重組體均只有一個斷點，顯示 SNP 分子標誌的順序無誤。由於目前不結球白菜尚未有全基因體解序之資料庫可供參考，本研究採用分屬不同亞種的結球白菜基因體為參考序列以開發分子標誌，儘管大部份 SNPs 均可由參考序列比對後獲得且正確無誤，然而利用近緣物種的基因體為參考序列，難以探勘因演化過程產生之變異，本研究搭配針對染色體焦點區域之 de novo assembly 策略而開發出 BrA7\_6746K 分子標誌，可做為彌補因演化差異造成生物資訊缺失的有效方法。

進一步比對重組體外表型與基因型的資料，結果顯示 BrA7\_6746K 與 BrA7\_6749K 分子標誌基因分型的結果與外表型完全一致，因此 *BrMs* 基因被定位於 BrA7\_6737K 與 BrA7\_6761K 間之 24 kb 範圍內 (表 5-5)。本研究先前推測 *BrMs* 與 *BnRf* 為同源基因之證據，是依據 *BnRf* 而設計之分子標誌 (*BrMs\_indel*) 分析 923 個 NH80-A × TA95 之 F<sub>2</sub> 植株時，該分子標誌基因型與雄不稔性狀完全共分離之結果，然而為進行精細定位而將族群放大至 4,000 株以上時，隨著定位解析度提高，而發現了此分子標誌基因型與外表型不一致的 2 株個體 (16-3157 及 16-3236)，推測 *BrMs* 基因可能為存在一組複合基因 (gene complex)，而該 2 株基因型與外表型不同的個體，可能是該複合基因內發生重組的結果。事實上於驗證大油菜 *BnRf* 基因功能的過程中，將雄不稔株之 *BnRf* 等位基因 (*BnRf<sup>ph</sup>*) 轉殖至阿拉伯芥 (*A. thaliana*) 時，可誘導雄不稔性之表現因而確認其為雄不稔基因，然該研究也意外發現將可稔系之另一個 *BnRf* 等位基因 (*BnRf<sup>ru</sup>*)，轉殖至阿拉伯芥時也會產生雄不稔，當時即推測 *BnRf* 並非一般基因單純組成，可能為一組複合基因座 (Deng et al. 2016)。此種複合基因的實例曾在

水稻抗稻熱病基因 (*Pikm1-TS*及*Pikm2-TS*) 發現 (Ashikawa et al. 2008)。

呈現該 2 株個體表現型的另一個可能原因則是同源基因重複 (gene duplication) 的結果，相同功能基因叢聚 (clusters) 在鄰近基因體已有許多實例，如水稻白葉枯病 (bacterial blight) 的抗病 R 基因即叢聚於第 11 號染色體之特定區域，其可能係種化過程中基因不對稱分離而造成的基因重複 (gene duplication) 現象 (Mizuno et al. 2020)，若不結球白菜的 *BrMs* (或大油菜的 *BnRf*) 在基因座落區域曾發生基因重複現象，使得該基因附近不只有一個 copy 而有二個以上之 paralogs，而在演化過程中該二物種各只有一個 paralog 維持基因的雄不稔性功能，但因 paralogs 間，位置具些微的差異亦有可能產生類似本研究之結果，未來若 *BrMs* 被選殖確認後，進一步針對不結球白菜與大油菜之相對染色體區域，進行結構上之定序與解析，則能進一步解析基因間功能表達與分子演化之關係。

依據 Brassica database (BRAD) (<http://brassicadb.org/brad/index.php>) 對 *B. rapa* var. *chiifu40* (v1.5) 之基因註解 (annotation) 結果，本研究精細定位之 24 kb 範圍內預測含 *Bra014987*、*Bra014988* 及 *Bra014989* 等 3 個基因，以該 3 個基因可能轉譯的蛋白對阿拉伯芥進行比對 (BlastX)，分別比對出 *AT3G23820*、*AT3G23840* 及 *AT3G23990* 等最近似的阿拉伯芥基因，其中 *AT3G23840* 轉譯之 transferase family protein 為維持阿拉伯芥花粉稔性所必需，該基因突變體會影響花粉受精能力而造成雄不稔 (Haslam et al., 2015)，雖然於白菜基因體所預測的三個基因中，以 *Bra014988* 的功能與本研究雄不稔性狀相似，較有可能為本研究之候選基因，但 *AT3G23840* 影響稔性之機制為花粉受精能力之喪失，有別於本研究之無花粉造成之雄不稔，此有賴進一步選殖出 *Bra014987*、*Bra014988* 及 *Bra014989* 並加以驗證其基因功能。此外，Xie et al. (2012) 利用比較基因體的方式選殖大油菜 *BnRf* 時，該基因的 13.8 kb 定位區間恰巧對應至結球白菜 A7 連鎖群之 17.8 kb 範圍，並包含 *Bra014989*、*Bra014990* 及 *Bra014991* 等三個預測基因，其中 *Bra014989* 與我們的結果有交集，儘管在大油菜的研究中最終證實 *BnRf* 並非這些基因 (Deng et al., 2016)，但我們也證實 *BnRf* ortholog 與 *BrMs* 緊鄰或重疊，因此 *BrMs* 為 *Bra014989* 的可能性仍存在。



第6章 利用全基因體關連分析 (GWAS) 開發鑑別  $BrMs^a/BrMs^c$   
等位基因之 SNP 分子標誌




## 1 前言

本研究透過基因精細定位將 *BrMs* 限縮於約 24 kb 範圍內，此區間所開發之分子標誌如第五章之 BrA7\_6746k 與 BrA7\_6749K 等，於 NH80-A 與 TA95 雜交分離後裔中區分 *BrMs<sup>a</sup>* 與 *BrMs<sup>b</sup>* 等位基因之基因型時準確度相當高，雖然於該區間內預測只有 3 個基因，然而本研究尚未將該 3 個推測的基因序列轉殖至其他稔性正常的個體以驗證基因功能，因此在無法確認何者為雄不稔基因的情況下，只能保守推論前述這些 SNP 分子標誌可能與 *BrMs* 緊密連鎖，或恰位於基因上而非功能性分子標誌 (functional marker)。當這些標誌不是功能性分子標誌時，儘管它與 *BrMs* 基因有緊密的連鎖關係，通常仍難以直接用來鑑別其他小油菜品系或其他白菜類物種之 *BrMs* 基因型。

事實上本研究小油菜之雄不稔株是被發掘獲得的，自然情況下該 NH80 族群大都稔性正常，而大部份白菜類作物品種也都是稔性正常的個體，也就是說除了 NH80-A 個體具 *BrMs<sup>b</sup>* 等位基因之外，大部份品系其 *BrMs* 基因型主要由 *BrMs<sup>a</sup>* 與 *BrMs<sup>c</sup>* 等位基因所構成，由於這二種等位基因組成之基因型，均呈現雄可稔外表型，育種時若欲以本研究之雄不稔性進行採種工作，則必需先區分育種材料之基因型為何，作法上可能需先將育種材料與雄不稔個體進行測交，由後裔檢定再判斷原測交親之基因型，此舉勢必耗時費力。倘若第五章開發之 SNP 分子標誌均非功能性分子標誌，則有必要開發出可區別 *BrMs<sup>a</sup>* 與 *BrMs<sup>c</sup>* 等位基因的分子標誌，才符合育種選拔或品系基因型鑑別之需求。

異交作物之種原大都為經過多年多世代之雜交授粉及染色體重組分離等過程演化 (或馴化) 而成，相較於以雙親交配而產生之分離族群而言，利用多樣化種原所構建之族群，其共線性的染色體區間內因經過頻繁的重組，若利用這類族群進行基因定位將有更高之解析度。因此以種原族群配合大量基因體資訊進行所謂的全基因體關連性分析 (genome-wide association study; GWAS) 為近年盛行的基因研究策略之一，其主要原理乃是建立高密度分子標誌與性狀間的相關，進而定位或找到控制該性狀之基因座。





連鎖失衡 (linkage disequilibrium; LD) 指的是二基因座間存在非逢機的關聯性，透過連鎖失衡可連結分子標誌與目標基因間之關係，所以連鎖失衡一直是 GWAS 分析時很重要的議題。當分子標誌與目標基因距離越遠時，連鎖失衡越不顯著，即所謂連鎖失衡衰退 (LD decay)，此連鎖失衡衰退的平均物理距離與作物繁殖特性有關，通常自交作物連鎖失衡衰退的平均距離遠大於異交作物，連鎖失衡衰退的平均距離短時，一旦分子標誌與目標基因建立強烈關聯性，則表示該分子標誌有可能極靠近候選基因，甚至直接命中基因位置，但這也意味著需要較高密度的分子標誌才能涵蓋整個基因體，有效地偵測到目標基因。反之，當 LD decay 的平均距離大時，分子標誌密度不需太高即可與目標基因間建立關聯性，該分子標誌也可做為選拔用。總結來說，進行 GWAS 分析時，配合種原族群的遺傳背景，所需的分子標誌最好具有高密度且遍及整個基因體之特性，現今受惠於次世代定序之普遍應用，獲得基因體資料並發掘大量 SNP 位點已非難事，因此已有許多作物如高粱 (Morris et al., 2013)、玉米 (Farfan et al., 2015)、大油菜 (Hatzig et al., 2015; 魏等, 2017)、水稻 (Spindel et al., 2016)、小麥 (Zhou et al., 2017) 及花生 (Otyama et al., 2019) 等透過 GWAS 以發掘重要性狀之基因。

本研究小油菜屬異交作物，推測種原族群中的 *BrMs* 基因座以 *BrMs<sup>a</sup>* 與 *BrMs<sup>c</sup>* 等位基因為主，為建立鑑別 *BrMs<sup>a</sup>* 與 *BrMs<sup>c</sup>* 基因型之分子工具以做為育種用途，本章將針對 *BrMs* 基因進行 GWAS 分析，此外，試驗族群中也將納入結球白菜及不同類型的不結球白菜做為分析材料，除增加族群的遺傳變異外，也可一併評估該 *BrMs<sup>b</sup>* 等位基因衍生之雄不稔性狀是否亦可作用於不同種類之 *B. rapa* 作物，進而將該基因的應用範圍擴大至其他白菜類作物。

## 2 材料與方法

### 2.1 試驗材料

為了區分 *BrMs<sup>a</sup>* 與 *BrMs<sup>c</sup>* 基因型，以 140 個 *B. rapa* 自交系進行 GWAS，這些品系原始來源分布於台灣、中國、日本、韓國及泰國等地 (表 6-3)。2018 年 10 月將所有參試品系各栽培 2 株於 7 吋盆，當發育至四葉苗齡時全數移至生長箱，以日溫 8°C，夜溫 6°C 進行人工春化處理 45 天，處理完成後移至溫室持續栽培，於開花期將各品系分別與雄不稔系 NH80-A 進行測交授粉，並於採收期收穫測交後裔 F<sub>1</sub> 種



子，此外各品系之二植株均採取葉片萃取 DNA，其中一株之 DNA 部份用於建構 DNA library 以進行 GWAS 分析，其餘 DNA 保存備用，待 GWAS 分析完成並開發出候選 SNP 分子標誌時，可供進行基因分型驗證用。

2019 年 4 月進行後裔檢定，將參試品系與雄不稔系 NH80-A 之測交 F<sub>1</sub> 種子各播種 20 株以上，並種植於 104 孔穴盤，當發育至三葉苗齡即移至生長箱，以前述相同低溫條件進行人工春化處理 50 天，春化處理完成後先移植至 35 孔穴盤，並持續於溫室栽培至開花以調查各單株之雄蕊稔性，各品系測交後裔全數為雄不稔者，判定該品系（或品系內之單株）之基因型為 *BrMs<sup>c</sup>Ms<sup>c</sup>*，當測交後裔全數具正常花粉則判定其花粉親基因型為 *BrMs<sup>a</sup>Ms<sup>a</sup>*，若測交後裔產生分離時則判定其花粉親為異質結合體，將所有參試品系測交結果的基因型資料，再與參試族群經定序而獲得之 SNPs 資訊整合以進行 GWAS 分析。

## 2.2 GBS 之進行

### 2.2.1 建構 GBS library

DNA libraries 之建構主要參考第三章 GBS library 之建構方法修改，針對參試植株取約 200 ng DNA 後同時以 8U 之 *Pst*I-HF 及 *Bsa*HI（New England Biolabs, USA）二個內核酸酶於 37 °C 下進行酶切 4 小時後，再於 65 °C 維持 20 分鐘，酶切完成之 DNA 片段以 barcoded adapters 及 reverse Y-adapter（Poland et al., 2012）進行連接，其中 Y-adapter 由表 6-1 之二種單股 DNA（各 10 μM）於 Adapter Buffer（50 mM NaCl, 10 mM Tris-HCl）環境下置於 95 °C 1 分鐘，隨後每分鐘下降 1 °C 至最終溫度達 30 °C，而完成雙股之黏合。

表 6-1 合成 Y-adapter 的 DNA 序列

Primers	Sequence
<i>Bsa</i> HI_top	5'- CGAGATCGGAAGAGCGGGGACTTTAAGC -3'
<i>Bsa</i> HI_bot	5'- GATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT -3'



各樣本進行連接反應時反應體積為 40  $\mu$ L，包含酶切 DNA (200 ng)，barcoded adapter (0.02 pmol)，reverse Y-adapter (3 pmol)，T4 DNA ligase (200U)，1X NEB CutSmart® Buffer (NEB, USA) 及 ribo ATP (1 mM)，於 22  $^{\circ}$ C 下進行 2 小時，再於 65 $^{\circ}$ C 維持 20 分鐘，完成 adapters 連接之 DNA 各取 10  $\mu$ L 混合成 DNA library，隨後用 AMPure XP 磁珠 (Beckman Coulter, USA) 以 DNA:磁珠為 1:1 的比例進行純化，最後利用經修飾之 Illumina 引子對 (Poland et al., 2012) 於 95 $^{\circ}$ C (30 sec)，62 $^{\circ}$ C (25 sec)，68 $^{\circ}$ C (40 sec) 條件下進行 PCR 反應 18 個 cycles，建庫完成的 GBS library 以 Illumina Hiseq X Ten 定序平台 (Illumina Inc., USA) 進行次世代定序，進行時採片段雙端 (paired-end) 定序，各端讀序長度為 150 bp。

### 2.2.2 序列分析

經次世代定序獲得之序列以軟體 CLC Genomics Workbench v8.0.2 (CLC Bio, <http://www.clcbio.com>) 進行處理，首先刪除讀序上 Illumina 定序平台之接頭(adapter) 序列，其次以 “merge overlapping pairs” 指令將各片段上正反讀序可能重疊者合併成一個 single-end read，再篩選出序列品質分數大於 Q20 者成為後續分析之 “clean reads” 的一部分。此 clean reads 包含 paired-end reads 及合併後較長片段之 single-end reads 二種，這些序列分別依據各參試樣品之 DNA 條碼進行分類，隨之以 “map reads to reference” 指令與參考序列進行比對歸位(mapping)，參考序列採用 *B. rapa* var. *chiifu40* (v1.5) (Cheng et al., 2011)，比對過程中各鹼基之錯誤配對、插入及缺失之加扣分參數採 CLC Genomics Workbench v8.0.2 之預設值，而當讀序與參考序列間發生非專一性配對時則將該序列排除，各參試樣品之序列比對完成後，其結果以 bam 的檔案格式輸出。

各樣品完成與參考序列的比對歸位的結果進一步以軟體 Stacks v1.37 (<http://creskolab.uoregon.edu/stacks/>) (Catchen et al., 2011) 整合成整個參試族群的資訊，主要以 “ref\_map.pl” 指令建立基因座暨基因座目錄、進行 SNP calling 及各品系之基因分型等，在所有讀序中，專一性序列被定義為 tag，只有讀序深度  $\geq 5$ ，且於族群內之缺失率少於 20% 的 tags，才會輸出成為基因型資料，再以 “populations” 指令將整個參試族群的 SNP 資訊以 vcf 的檔案格式輸出。



### 2.3 全基因體關聯性分析 (GWAS)

GWAS 之分析主要利用 TASSEL v5.2.52 (Bradbury et al. 2007) 軟體進行，族群中篩選各 SNP 位點之條件設定為需同時符合下列要件：1.) 至少出現於 105 個參試品系以上。2.) 其次等位基因頻度 (minor allele frequency; MAF)  $\geq 0.05$ 。3.) 淘汰異質結合  $\geq 0.75$  者。篩選的 SNPs 進一步依據其物理位置分析各染色體及全基因體的 LD decay 情形，其中平均 LD 距離以最高  $r^2$  值的 1/2 時之距離估計。

篩選後獲得 25,514 個 SNPs，先以廣義線性模型 (generalized linear model, GLM) 及壓縮混合線性模型 (compressed mixed linear model, CMLM) 進行 GWAS 分析，參考該二種分析方法之 Q-Q plot 圖形，擇定以壓縮混合線性模型對雄不稔性狀與 SNPs 進行關聯性之檢測，分析時以主成分分析 (principal component analysis, PCA) (Q=3 個 PCA 向量) 及相對親緣矩陣 (kinship matrix, K) 控制族群結構 (population structure)，其中相對親緣矩陣主要利用前述通過條件之所有 SNPs 以 Centered-Identity by State (Centered IBS) 法產生，而變方成分估計則採 P3D 法 (Zhang et al. 2010) 估計一次，基於  $P = G + Q + K + E$  模型計算各 SNP 位點與雄不稔性狀之相關，且以  $-\text{Log}_{10}(P)$  呈現於 Manhattan Plot，並以  $-\text{Log}_{10}(0.05/\text{SNPs 總數})$  做為篩選具關聯性 SNPs 之閾值 (threshold)。

### 2.4 SNP 分子標誌之設計與驗證

在 GWAS 結果中，首先評估機率值 (P value) 超過閾值的 SNPs，其所涵蓋之物理區間，於該區間內挑選部份 SNPs 開發為 SNP 分子標誌，以驗證其對於不同 *BrMs* 基因型之鑑別能力。挑選時以最靠近區間邊緣的 SNP 做為第一個候選位點，其次選擇間距不小於 10 kb 者做為第二個候選位點，循此間距依序於區間挑出所有候選位點，進而設計成 KASP SNP assays 或 Taqman assays。

為了驗證獲選 SNPs 位點對 *BrMs* 基因型之鑑別能力，以這些 SNP 分子標誌對前述參試族群進行基因分型，進行時各單株以 QuickExtract™ DNA extraction solution (Lucigen Co. Ltd, USA) 萃取 DNA，隨之將經稀釋之 DNA 萃取液、KASP assay 及 KASP master mix (或 Taqman assay 及 Taqman GTXpress master mix)，於 5  $\mu\text{l}$  反應

體積下依據原廠建議操作程序以 Applied Biosystems Viia 7 Real-Time PCR System (Applied Biosystems) 進行反應及螢光偵測。KASP assays 的 PCR 反應條件係依據 LGC 所提供之操作指南進行 (為: 94°C 15 分鐘, 隨之以 94°C 20 秒及 61-55°C 60 秒 (每 cycle 降低 0.6°C) 進行 10 個 cycles, 再以 94°C 20 秒, 55°C 1 分鐘進行 26 個 cycles, 最後維持反應於 40°C 1 分鐘並進行等位基因專一性之螢光偵測), 至於 Taqman assay 的反應條件則依據 Viia 7 Real-Time PCR 對於 Taqman assay 的原廠預設程式進行, 當反應結束並完成螢光偵測後, 各單株基因型由 Real-Time PCR 隨付之 Viia 7 software (v1.2) 進行基因型判別。

### 3 結果與討論

#### 3.1 次世代定序所獲得之讀序

經次世代定序後共產生 463 million paired-end reads, 序列總長累計約 138.9 Gb 資料量, 經品質篩選後獲得 447 million paired-end reads, 表示可進行後續分析之序列約佔 96.5%, 嘗試將 paired-end reads 中重疊的序列進行合併, 結果共有 179.6 million paired-end reads 可合併成為較長的 single-end reads 序列, 其片段長分布於 75-292 bp 之間 (圖 6-1), 這些序列的佔比約為 40%, 其餘 60% 仍維持 paired-end, 將這些讀序依樣品 DNA 條碼進行分類後, 平均每個參試樣本可獲得 3.7 million 個讀序。

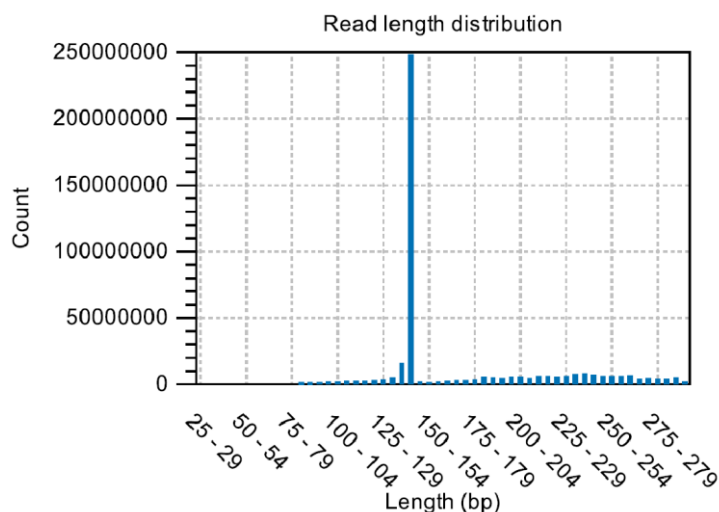
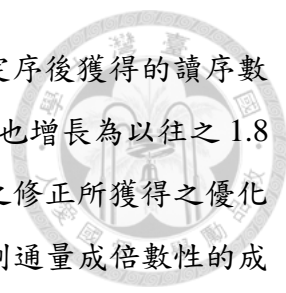


圖 6-1 正反讀序重疊之 paired-end reads 合併成 single-end reads 的長度分布



本研究為了進行 GWAS 而建構之 DNA library，經次世代定序後獲得的讀序數量，經換算約為第 3、4 章 GBS 試驗之 3.6 倍，而平均序列長度也增長為以往之 1.8 倍，主要是因為次世代定序平台之更新，以及 DNA 建庫方法之修正所獲得之優化結果。隨著次世代定序平台之更新，每單一個 lane 所獲得之序列通量成倍數性的成長，有助於在單次定序過程中，獲得更多的基因型資料或增加參試的樣品數，然而為配合不同定序平台，本研究嘗試創新的 DNA 建庫方式加以因應，由於目前定序公司大都以 paired-end 定序為主要服務項目，經建庫之 DNA 片段以 paired-end 進行雙端定序 150 bp 時，即所建構之 DNA library 平均序列至少需大於 300 bp 以上，本研究第 3、4 章 GBS 之 DNA 建庫主要參考 Poland et al. (2012) 之方法以 *Pst*I-HF 及 *Msp*I 內切酵素酶切 DNA 成小片段，其中 *Msp*I 識別序列與切位為 C/CGG，當不考慮物種基因體之 GC 含量時，其機率約為  $(1/4)^4$  即遇可辨識的序列而酶切，產生之小片段平均約 256 bp，然此小片段還需被 *Pst*I-HF 酶切才可被定序，因此這種片段大小不適用於前述 150 bp paired-end 定序之設定，然而若改以識別序列為 6 個鹼基的內切酵素，機率上約  $(1/4)^6$  才有可辨識之序列，酶切後之片段又太大，且因酶切位點將大幅減少而影響可獲得之序列資訊，本研究改採 *Pst*I-HF 及 *Bsa*HI 酵素之組合改進 DNA 建庫方式，雖然此二者皆為辨識 6 個鹼基的內切酵素，然 *Bsa*HI 識別序列及切位為 GR/CGYC，因此除了 *Pst*I-HF 之外，相當於另外投入 4 種辨識 6 個鹼基的內切酵素進行反應，除了可獲得適當大小之片段外，酶切後之端點 (sticky end) 序列一致，可用相同的 Y-adapter 進行連接而省卻操作時需黏合不同 adapter 之麻煩。

完成次世代定序及序列整理後，約 60 % 仍為 paired-end reads，顯示其原 DNA 片段大於 300 bp，利用此 paired-end reads 比對至參考序列時，由於相同 reads 二端各有 150 bp 同時進行序列比對，且成對的序列若無法比對至相近的位置時（例如兩端序列被比對至不同染色體），可透過生物資訊軟體以適當參數加以過濾而提高序列比對之正確性，定序結果中有 40% 的 paired-end reads 因其進行定序的 DNA 片段小於 300 bp 而被合併成 single-end reads，合併後大多數的 reads 長度大於 150 bp，較長的序列除了能增進比對的準確度之外，日後欲開發分子標誌時，由於候選 SNP 位點的周邊序列資訊增加，有助於提高開發的成功率。



### 3.2 透過 GBS 探勘獲得之 SNPs

針對整個參試族群進行 GBS 的結果，共探勘出 53,611 個 SNPs 位點，進一步針對各 SNPs 於整個族群中之次等位基因頻度 (minor allele frequency, MAF)、參試品系間之缺值、異質結合佔比等進行篩選，最後共獲得 25,514 個 SNPs 位點，這些 SNPs 於 *B. rapa* 各染色體之分布及密度如表 6-2 所示，其中染色體 Chr03 及 Chr09 相對較長，因此所獲之 SNPs 也較多，整個族群中異質結合的 SNPs 約佔 5%，這是由於參試種原大都為自交系所致，整個族群的次等位基因頻度 (MAF) 平均約 0.24，其分布如圖 6-2，MAF 介於 0.1~0.5 的 SNPs 數量均相似。由於所使用之參考序列基因體總長約 257 Mb，因此以全基因體 SNP 密度分布而言，平均每 10 kb 有 1 個 SNP，進一步依據連鎖失衡衰退的情形 (圖 6-3、6-4; 表 6-2)，估計 *B. rapa* 全基因體的平均 LD decay 為 12.5 kb，顯示本研究之 SNP 密度應足以進行 GWAS 分析。

表 6-2 參試族群 SNPs 分布及 LD decay 情形

	SNPs	MAF	Heterozygous	Reference length	Density (bp)	LD decay
Chr01	2,728	0.24	0.06	26,791,028	9,821	10,500
Chr02	2,438	0.24	0.05	26,939,826	11,050	12,000
Chr03	3,629	0.21	0.04	31,765,688	8,753	12,000
Chr04	1,498	0.21	0.04	19,269,589	12,864	8,500
Chr05	2,478	0.20	0.05	25,303,532	10,211	9,000
Chr06	2,960	0.24	0.09	25,210,368	8,517	10,500
Chr07	2,310	0.24	0.08	25,876,096	11,202	9,000
Chr08	2,048	0.24	0.04	20,826,945	10,169	7,500
Chr09	3,526	0.24	0.06	38,884,800	11,028	13,500
Chr10	1,899	0.24	0.05	16,405,180	8,639	12,500
Total	25,514	0.24	0.05	257,273,052	10,084	12,500

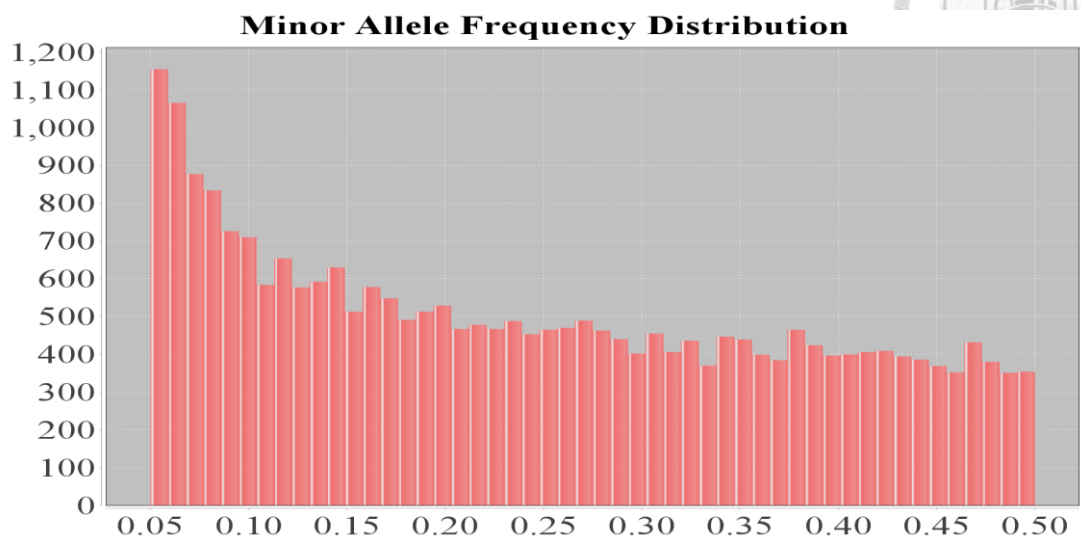
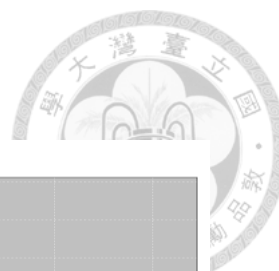


圖 6-2 參試族群 SNP 位點之次等位基因頻度分布情形

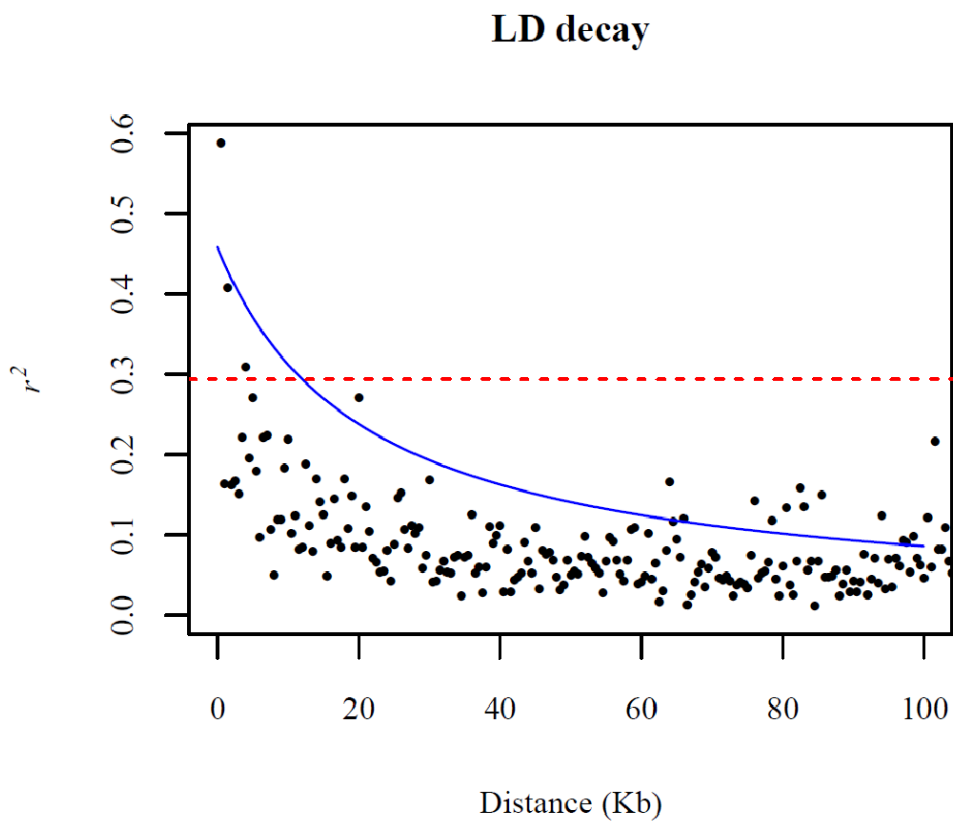
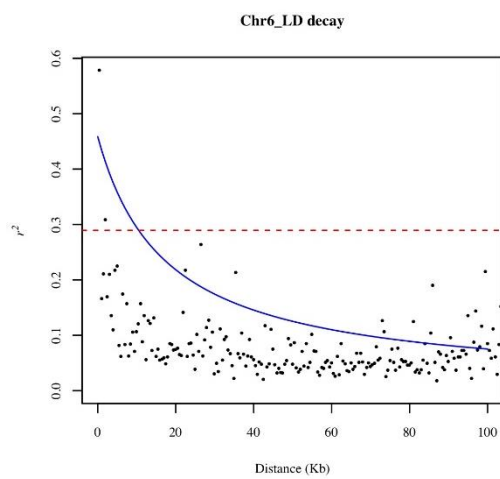
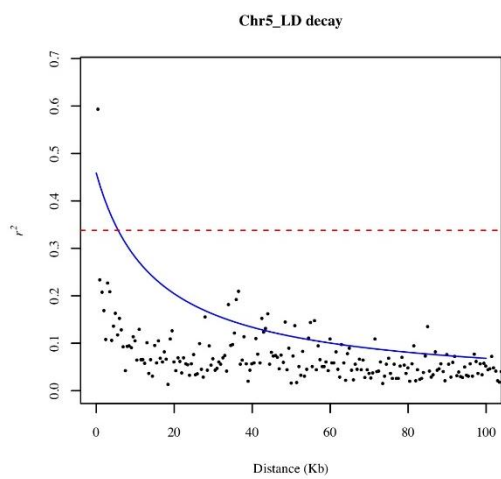
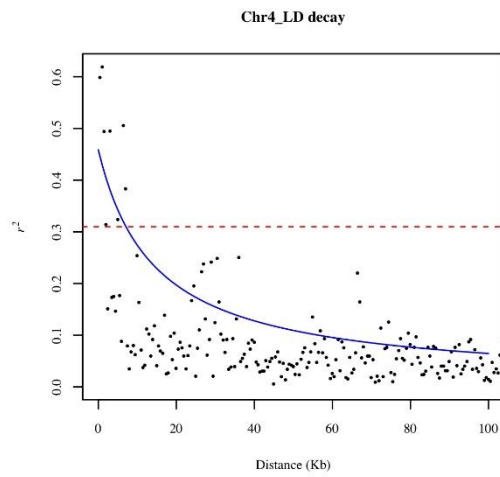
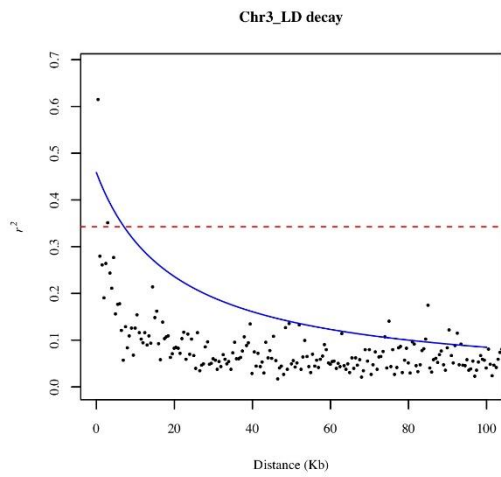
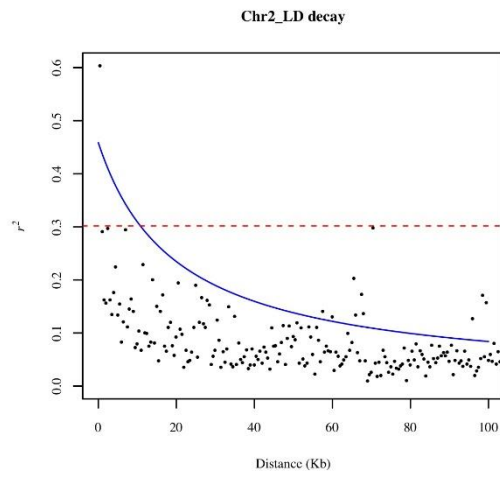
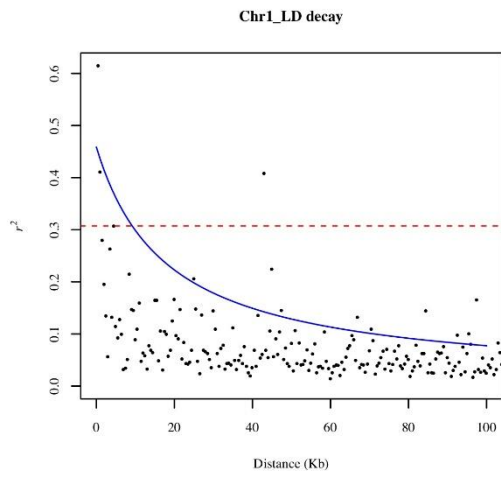
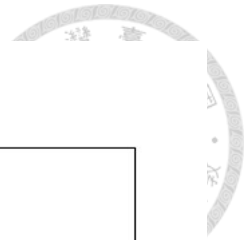


圖 6-3 *B. rapa* 族群全基因體 SNPs 的 LD decay 情形

紅色虛線為最高  $r^2$  的 1/2





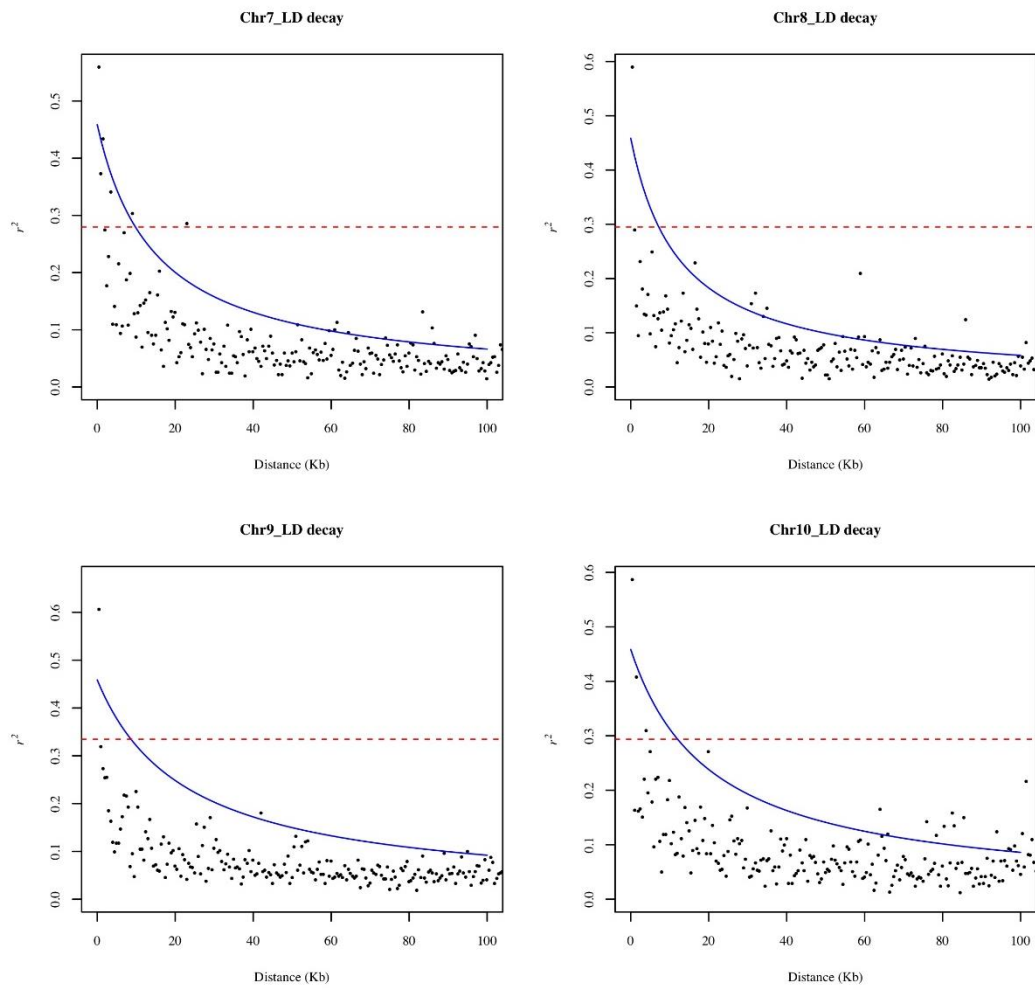


圖 6-4 *B. rapa* 族群各染色體 SNPs 的 LD decay 情形  
 紅色虛線為最高  $r^2$  的 1/2，表示各染色體之平均 LD decay 長度

### 3.3 GWAS 分析之族群組成

參試的 140 個 *B. rapa* 品系，其中包含 30 個不結球白菜及 110 個結球白菜，種原來源分布於台灣、中國、日本、韓國及泰國等地（表 6-3）。



表 6-3 進行 GWAS 分析的品系資訊

編號	品系原始親源	來源	類型	異質結合率	<i>BrMs<sup>2</sup></i>
GBS01	早熟 8 號-A	中國	結球白菜	0.038	a
GBS02	早熟 8 號-B	中國	結球白菜	0.075	-
GBS03	早熟 8 號-C	中國	結球白菜	0.055	a
GBS04	夏季王	中國	結球白菜	0.036	c
GBS05	太空 6 號	中國	結球白菜	0.041	a
GBS06	太空 9 號-A	中國	結球白菜	0.041	a
GBS07	太空 9 號-B	中國	結球白菜	0.038	c
GBS08	先風橘紅心-A	中國	結球白菜	0.047	a
GBS09	先風橘紅心-B	中國	結球白菜	0.040	a
GBS10	航天八號	中國	結球白菜	0.040	a
GBS11	耐儲狀元-A	中國	結球白菜	0.034	a
GBS12	耐儲狀元-B	中國	結球白菜	0.052	c
GBS13	新星鐵將軍	中國	結球白菜	0.039	c
GBS14	真好吃-A	中國	結球白菜	0.059	c
GBS15	真好吃-B	中國	結球白菜	0.042	a
GBS16	真好吃-C	中國	不結球白菜	0.062	a/c
GBS17	真好吃-D	中國	結球白菜	0.049	c
GBS18	膠蔬夏季王	中國	結球白菜	0.040	c
GBS19	膠蔬夏霸王	中國	結球白菜	0.058	c
GBS20	靚冠熱白菜-A	中國	結球白菜	0.043	c
GBS21	靚冠熱白菜-B	中國	結球白菜	0.040	-
GBS22	夏秋 55-A	中國	結球白菜	0.052	a/c
GBS23	夏秋 55-B	中國	結球白菜	0.078	c
GBS24	小義和秋	中國	結球白菜	0.043	c
GBS25	義和夏陽白	中國	結球白菜	0.093	c
GBS26	津寶 2 號	中國	結球白菜	0.093	a
GBS27	天白 50	中國	結球白菜	0.038	c

表 6-3 (續)

編號	品系原始親源	來源	類型	異質結合率	BrMs <sup>2</sup>
GBS28	巨龍 58	中國	結球白菜	0.038	a
GBS29	鴻盛三優 523	中國	結球白菜	0.047	a
GBS30	厚葉 5 號	中國	結球白菜	0.045	a
GBS31	紅寶娃-A	中國	結球白菜	0.051	a/c
GBS32	紅寶娃-B	中國	結球白菜	0.039	-
GBS33	快菜先鋒	中國	結球白菜	0.042	c
GBS34	夏綠 5 號-A	中國	結球白菜	0.044	c
GBS35	夏綠 5 號-B	中國	結球白菜	0.039	a/c
GBS36	傳祺大白菜-A	中國	結球白菜	0.043	c
GBS37	傳祺大白菜-B	中國	結球白菜	0.041	a
GBS38	日本抗病夏陽白-A	中國	結球白菜	0.040	c
GBS39	日本抗病夏陽白-B	中國	結球白菜	0.064	a/c
GBS40	瑞春白菜	中國	結球白菜	0.039	c
GBS41	勝春	中國	結球白菜	0.041	a
GBS42	東京春-A	中國	結球白菜	0.038	a
GBS43	東京春-B	中國	結球白菜	0.118	-
GBS44	夏將軍	中國	結球白菜	0.034	a/c
GBS46	CC-200-A	台灣	結球白菜	0.040	a
GBS47	CC-200-B	台灣	結球白菜	0.042	-
GBS48	CC-801	台灣	結球白菜	0.043	c
GBS49	合歡三和 059 球白菜	台灣	結球白菜	0.034	a/c
GBS50	曉曉橘芯	韓國	結球白菜	0.043	a/c
GBS51	一品菜	韓國	結球白菜	0.039	a
GBS52	珍珠娃娃菜	韓國	結球白菜	0.146	-
GBS53	金福娃	韓國	結球白菜	0.034	a
GBS54	高麗貝貝-A	韓國	結球白菜	0.046	a
GBS55	高麗貝貝-B	韓國	結球白菜	0.081	a
GBS56	首爾娃娃-A	韓國	結球白菜	0.051	a
GBS57	首爾娃娃-B	韓國	結球白菜	0.050	a
GBS58	高山娃娃菜	韓國	結球白菜	0.035	a
GBS59	極品娃娃菜	韓國	結球白菜	0.036	a
GBS60	春黃美	日本	結球白菜	0.033	a

表 6-3 (續)

編號	品系原始親源	來源	類型	異質結合率	BrMs <sup>2</sup>
GBS61	春大將	日本	結球白菜	0.032	c
GBS62	大福 234	日本	結球白菜	0.039	a
GBS63	美琪黃芯	日本	結球白菜	0.035	a
GBS64	黃優白菜	日本	結球白菜	0.032	c
GBS65	黃金包心白菜	日本	結球白菜	0.031	a
GBS66	勝利結球白菜	日本	結球白菜	0.025	c
GBS67	黃ごころ 65	日本	結球白菜	0.078	-
GBS68	オレソヅクイソ-A	日本	結球白菜	0.033	-
GBS69	オレソヅクイソ-B	日本	結球白菜	0.037	-
GBS70	無雙	日本	結球白菜	0.034	a
GBS71	金將二號	日本	結球白菜	0.034	c
GBS73	冬峠-A	日本	結球白菜	0.048	c
GBS74	冬峠-B	日本	結球白菜	0.094	-
GBS75	晚輝	日本	結球白菜	0.068	-
GBS76	サラダ	日本	結球白菜	0.044	-
GBS77	チヒリ 70	日本	結球白菜	0.037	c
GBS78	プチヒリ	日本	結球白菜	0.038	c
GBS80	晚抽プチヒリ	日本	結球白菜	0.037	-
GBS82	おいしい白菜	日本	結球白菜	0.039	-
GBS83	甘みがあっておいしい	日本	結球白菜	0.046	c
GBS84	耐病早生 60 日型白菜	日本	結球白菜	0.032	c
GBS85	泰國油菜	泰國	油菜	0.041	-
GBS86	天綠 711 特油青菜心	中國	油菜	0.043	a
GBS87	夏綠 388 油青早菜心	中國	油菜	0.044	a
GBS88	澳選短葉油綠菜心王	中國	油菜	0.040	-
GBS89	中南短葉油綠菜心王	中國	油菜	0.046	a
GBS90	廣東高級 50 天菜心	中國	油菜	0.047	a
GBS91	京研 3 號-A	中國	青梗白菜	0.086	-
GBS92	京研 3 號-B	中國	青梗白菜	0.053	a
GBS93	早生金品 21-A	中國	青梗白菜	0.037	a
GBS94	早生金品 21-B	中國	青梗白菜	0.074	a
GBS95	新綠洲 G88-A	中國	青梗白菜	0.041	a

表 6-3 (續)

編號	品系原始親源	來源	類型	異質結合率	BrMs <sup>2</sup>
GBS96	新綠洲 G88-B	中國	青梗白菜	0.057	-
GBS101	葵扇矮腳黑葉白菜	中國	白梗白菜	0.044	a
GBS102	黃葉白菜種	中國	青梗白菜	0.075	a
GBS103	夏寶黑葉白梗白菜	中國	白梗白菜	0.066	a
GBS105	鴻盛金牌六號白菜	中國	結球白菜	0.071	a
GBS106	早熟五號小王子	中國	結球白菜	0.097	a
GBS107	鴻盛金牌五號白菜	中國	結球白菜	0.066	a
GBS108	四季皇小白菜	中國	結球白菜	0.053	c
GBS109	油菜	台灣	油菜	0.045	a
GBS111	AG90	台灣	油菜	0.093	a
GBS113	秀珍	台灣	油菜	0.057	a
GBS115	青龍	台灣	油菜	0.045	a
GBS117	白油 1	台灣	油菜	0.044	-
GBS122	阿麗	台灣	油菜	0.071	-
GBS124	日本油美一號	日本	油菜	0.047	-
GBS125	福祿甜油菜	日本	油菜	0.130	a/c
GBS127	中原奶白菜	日本	白梗白菜	0.121	a
GBS129	夏三味	日本	青梗白菜	0.085	a
GBS132	ハルピン夏	日本	青梗白菜	0.057	-
GBS133	CC-200	台灣	結球白菜	0.044	-
GBS135	Hercules 60	泰國	結球白菜	0.041	-
GBS137	早熟 5 號 (優選)	中國	結球白菜	0.046	-
GBS140	德陽 01 -A	中國	結球白菜	0.041	-
GBS141	德陽 01 -B	中國	結球白菜	0.039	-
GBS143	536 小白菜	中國	結球白菜	0.039	-
GBS145	德高 87-114	中國	結球白菜	0.034	-
GBS148	津育 60	中國	結球白菜	0.090	-
GBS149	TIEN JIN	泰國	結球白菜	0.069	-
GBS151	GIANT	泰國	結球白菜	0.039	-
GBS153	GIANT 029	泰國	結球白菜	0.186	-
GBS156	優選快菜 369	中國	結球白菜	0.079	-
GBS158	無敵快菜	中國	結球白菜	0.038	-



表 6-3 (續)

編號	品系原始親源	來源	類型	異質結合率	<i>BrMs</i> <sup>z</sup>
GBS160	豐抗 78	中國	結球白菜	0.039	-
GBS161	西白 45 (金童子) -A	中國	結球白菜	0.043	-
GBS163	西白 45 (金童子) -B	中國	結球白菜	0.040	-
GBS165	揚夏早 38 天 (626)	中國	結球白菜	0.212	-
GBS167	夏暉早 38 天 (632) -A	中國	結球白菜	0.088	-
GBS169	夏暉早 38 天 (632) -B	中國	結球白菜	0.143	-
GBS171	和冠 (403) -A	中國	結球白菜	0.045	-
GBS174	和冠 (403) -B	中國	結球白菜	0.039	-
GBS176	春冠 -A	中國	結球白菜	0.039	-
GBS177	春冠 -B	中國	結球白菜	0.051	-
GBS180	正板夏秋 3 號	中國	結球白菜	0.047	-
GBS182	興滇高抗 4	中國	結球白菜	0.138	-
GBS183	板扎黃帝	中國	結球白菜	0.030	-
GBS185	吉祥	日本	結球白菜	0.154	-
GBS187	NH80-A	台灣	油菜	0.073	-
GBS188	NH80-B	台灣	油菜	0.075	-
GBS191	NH80-A×泰國油菜	台灣×泰國	油菜	0.186	-

<sup>z</sup>*BrMs* 基因型：a 表示 *BrMs<sup>a</sup>Ms<sup>a</sup>*；c 表示 *BrMs<sup>c</sup>Ms<sup>c</sup>*；a/c 表示 *BrMs<sup>a</sup>Ms<sup>c</sup>*；- 表示無測交結果

利用 25,514 個 SNP 分析所有參試品系的異質結合率 (heterozygosity)，其中最低為編號 GBS66 之 0.025，最高為 GBS165 之 0.212，GBS191 為 NH80-A× 泰國油菜之 F<sub>2</sub> 單株，其異質結合率為 0.186，整個族群異質結合率的分布情形如圖 6-5 所示，大部份品系均集中於 0.03-0.055 之間，族群平均為 0.057，顯示族群大都由自交系構成，而其中大於 0.08 者約有 19 個，這些品系可能因自交代數不足導致異質結合率偏高。140 個品系中，有 88 個完成與 NH80-A 雄不稔株測交之後裔檢定，其中 *BrMs* 基因型為 *BrMs<sup>a</sup>Ms<sup>a</sup>* 者有 50 個，29 個為 *BrMs<sup>c</sup>Ms<sup>c</sup>*，有 9 個則為異質結合體 *BrMs<sup>a</sup>Ms<sup>c</sup>*。

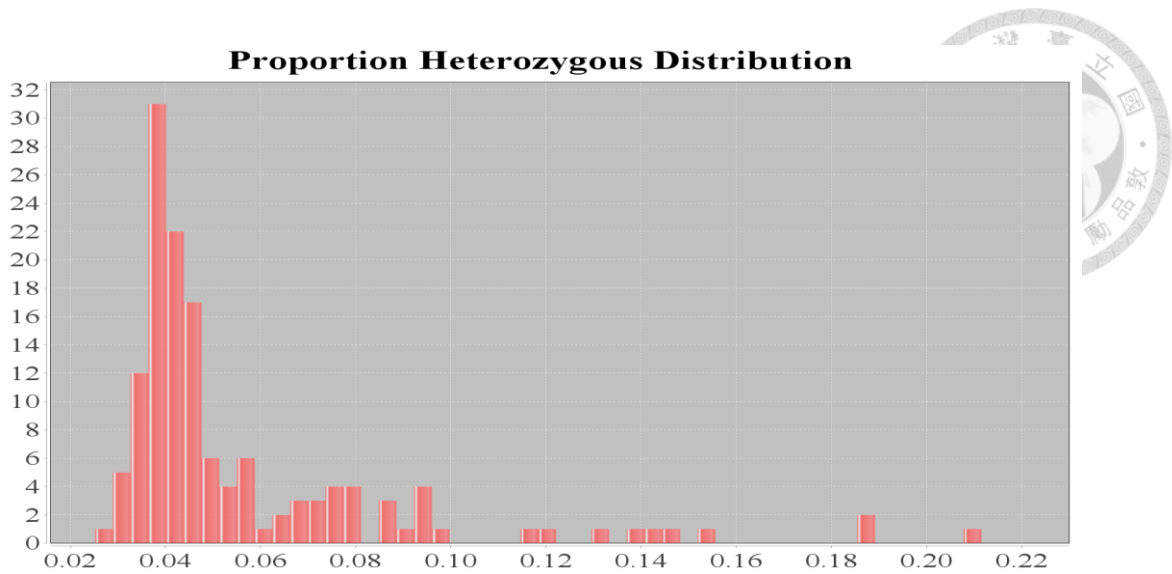


圖 6-5 各 *B. rapa* 品系之異質結合率分布情形

透過主成分分析顯示族群可明顯區分為結球白菜與不結球白菜二大群(圖 6-6), 其中 GBS165 屬結球白菜, 然而被歸類至不結球分群中, 由於該品系之異質結合率為族群中最高者(表 6-3), 推測可能為留種時受不結球白菜之花粉污染而雜交之結果。

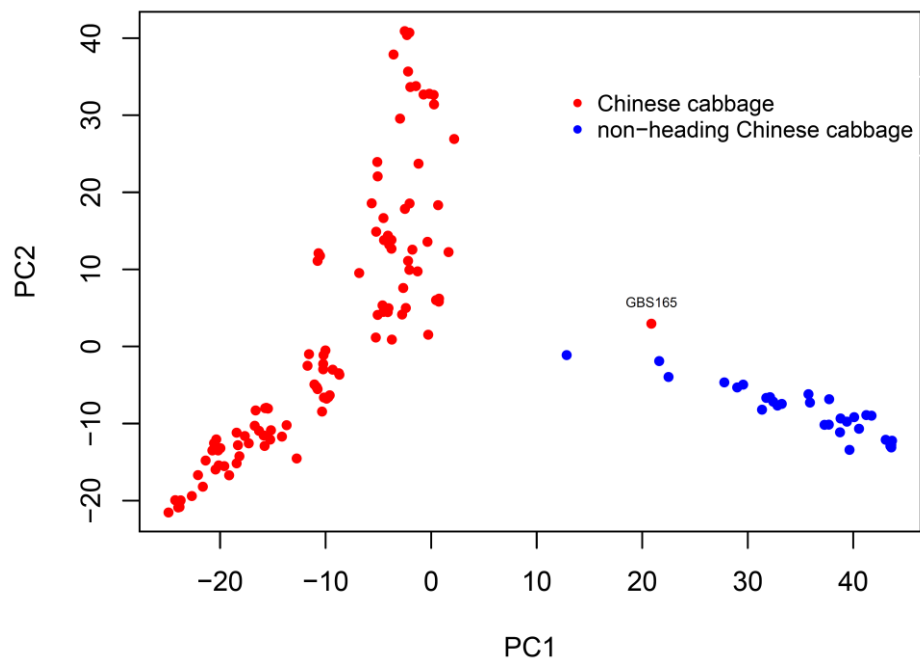


圖 6-6 參試品系經主成分分析後之分群情形





### 3.4 GWAS 分析的結果

整合 88 個品系的測交結果及 25,514 個 SNPs 基因分型資訊進行 GWAS 分析，當不考慮族群結構直接以廣義線性模型（general linear model, GLM）進行分析時，其 $-\text{Log}_{10}(P)$  實際觀測值與期望值如圖 6-7，即使 P-value 很大，例如當 $-\text{Log}_{10}(P) = 0.2$  開始，觀測值與期望值即無法吻合，篩選獲得與 *BrMs* 相關的 SNPs 容易產生偽陽性，顯示此模型不適合，改採  $P = G + Q + K + E$  模型控制族群結構，以壓縮混合線性模型（compressed mixed linear model, CMLM）進行分析後，顯示該模型較適合（圖 6-8），於該模型下進行 GWAS 分析後，發現位於 Chr 07 上的 SNPs 有一顯著波峰（圖 6-9），表示這些 SNP 位點與 *BrMs* 具相關性。

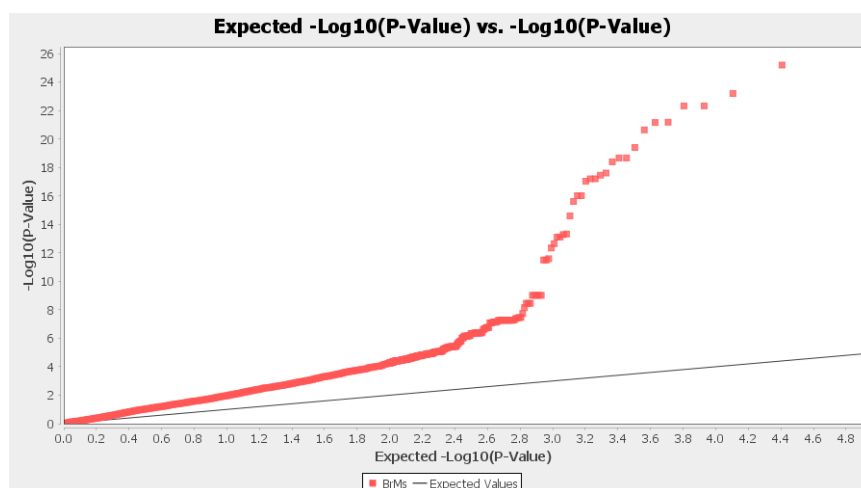


圖 6-7 未控制族群結構並以廣義線性模型進行 GWAS 分析之結果

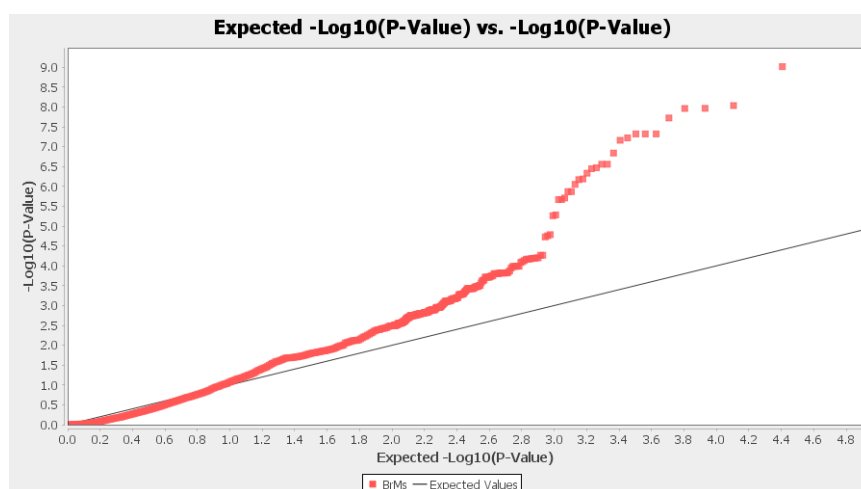


圖 6-8 採 Q+K 模型控制族群結構並以 CMLM 進行 GWAS 分析之結果

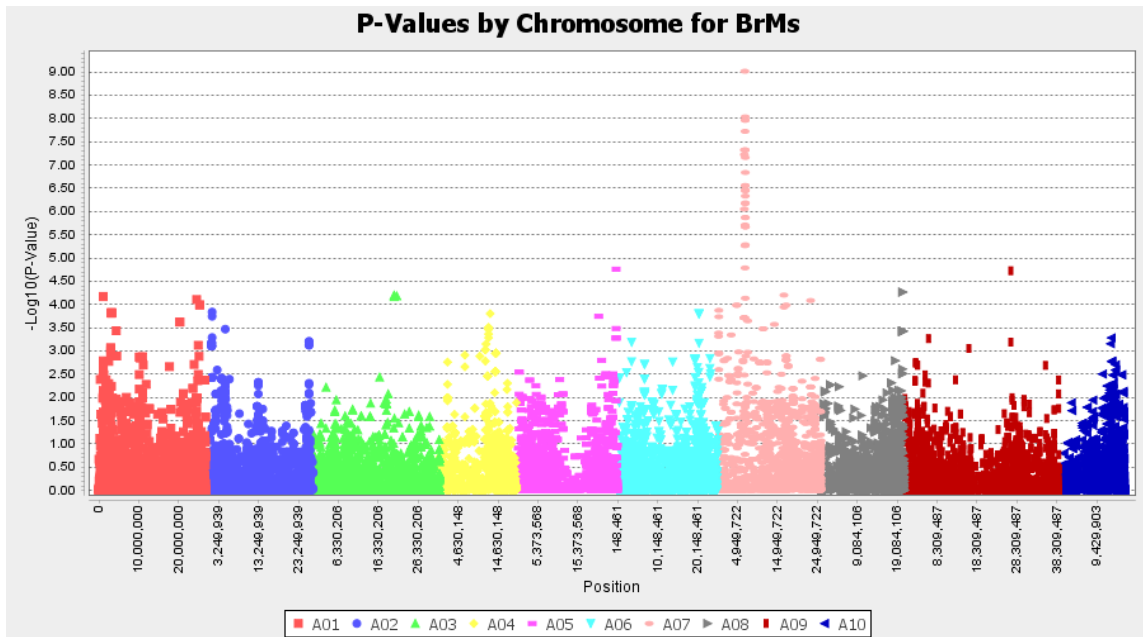


圖 6-9 經 GWAS 分析顯示 Chr 07 上具有與 *BrMs* 相關之 SNPs

與 *BrMs* 相關並通過閾值的 SNPs 共 17 個，其物理位置分佈於 Chr07 之 6,633,783 至 6,869,553 之間 (表 6-4)，其中以 A07 之 6,750,217 位點的  $-\text{Log}(p)$  最高，該 SNP 位點之  $R^2$  為 0.89，這些 SNPs 中，位於 6,633,783、6,750,217、6,824,432 及 6,863,231 等 4 個位點被用來設計為 KASP SNP 分子標誌 (或 Taqman SNP 分子標誌)。

表 6-4 與 *BrMs* 相關 SNPs 的物理位置

Tag No.	Chromosome	Position	$-\text{Log}(p)$	$R^2$
17945	A07	6,633,783*	6.1	0.38
17945	A07	6,633,876	7.3	0.49
17945	A07	6,633,878	7.2	0.49
17959	A07	6,750,217*	9.0	0.89
17965	A07	6,823,967	7.7	0.76
17966	A07	6,824,432*	5.7	0.36
17970	A07	6,863,231*	6.6	0.51
17970	A07	6,863,267	6.6	0.51
17970	A07	6,863,305	7.3	0.58
17970	A07	6,863,308	7.3	0.58
17970	A07	6,863,309	6.2	0.53
17972	A07	6,869,429	5.9	0.41
17972	A07	6,869,430	6.3	0.44
17972	A07	6,869,433	5.9	0.41
17972	A07	6,869,535	6.8	0.44
17972	A07	6,869,547	6.4	0.41
17972	A07	6,869,553	7.2	0.47
Threshold			5.7	


\*用以設計為 SNP 分子標誌的位點



### 3.5 SNP 分子標誌的驗證

依據 GWAS 結果中與 *BrMs* 相關連的位點所設計的 SNP 分子標誌分別名為 BrA7\_6633K、TqBrA7\_6750K、BrA7\_6824K 及 BrA7\_6863K，這些 SNP 分子標誌均可有效區別該位點 allele 間之差異（圖 6-10）。將具有測交結果的 70 個品系共 88 個單株，以上述這 4 個 SNP 分子標誌加上第五章中與 *BrMs<sup>a</sup>/BrMs<sup>b</sup>* 共分離的 BrA7\_6746K、BrA7\_6749K，及代表 *BnRf* ortholog 之 *BrMs\_indel* 等進行基因分型，彙整結果如表 6-5 所示。在本章開發的 4 個 SNP 分子標誌中，以 TqBrA7\_6750K 進行基因分型的結果與測交結果相當吻合，在 88 個測交單株中，TqBrA7\_6750K 鑑別為 *BrMs<sup>a</sup>Ms<sup>a</sup>* 及 *BrMs<sup>a</sup>Ms<sup>c</sup>* 基因型者共有 57 個，這些與測交結果完全一致，而 31 個被鑑別為 *BrMs<sup>c</sup>Ms<sup>c</sup>* 基因型的個體中，24 個也與測交結果一致，但 B022-1、B022-2、B032-1、B033-2、B117-1、B122-1、B133-1 等 7 個的測交後裔卻產生分離現象（表 6-5），而與 TqBrA7\_6750K 鑑別有出入，整體而言該標誌鑑別的正確率達 92%（81/88），至於 BrA7\_6633K、BrA7\_6824K 及 BrA7\_6963K 鑑別的正確率分別為 75.0%（66/88）、80.6%（71/88）及 72.4%（64/88），這 3 個分子標誌距離 TqBrA7\_6750K 越近者正確率越高。

參試的 88 個植株以 *BrMs\_indel* 及 BrA7\_6746K 進行檢測，發現有許多測交檢測屬 *BrMs<sup>c</sup>Ms<sup>c</sup>* 基因型的植株未產生螢光反應（表 6-5），顯示這些植株並無該 SNP 位點，也就是說這些植株在該位點的區域呈現片段缺失的狀態。回顧這 2 個分子標誌開發的過程，*BrMs\_indel* 是由 NH80-A 與 TA95 的 *BnRf* orthologs 間進行序列比對後而獲得，然而該片段於結球白菜參考序列（*B. rapa* var. *chiifu40* v1.5）上並未比對出一致的序列，而 BrA7\_6746K 則是由本研究第五章對 NH80-A 與 TA95 進行 *de novo* assembly 時，由所建構的序列資料庫中篩選出之 contigs 比對而獲得，其在參考序列上也無此位點，因此推測提供參考序列之 *B. rapa* var. *chiifu40* 也屬 *BrMs<sup>c</sup>Ms<sup>c</sup>* 基因型，由於帶有 *BrMs<sup>a</sup>*（或 *BrMs<sup>b</sup>*）等位基因之的個體大都具含 *BrMs\_indel* 及 BrA7\_6746K 之片段，但許多帶有 *BrMs<sup>c</sup>* 等位基因的個體卻又缺乏此片段，暗示演化上 *BrMs<sup>a</sup>*（或 *BrMs<sup>b</sup>*）與 *BrMs<sup>c</sup>* 周邊序列已產生序列結構上的差異，這與 Deng et al. (2016) 於 *B. napus* 上的研究結果相似。此外，BrA7\_6749K 鑑別的正確率僅為 72.7%（64/88），該位點與 TqBrA7\_6750K 只相距約 1 kb，而本研究對於 *B. rapa* 全



基因體所估算之平均連鎖失衡衰退為 12.5 kb，理論上這二個相距僅 1 kb 的分子標誌對於檢測結果應相去不遠，然而 BrA7\_6749K 檢測的正確率甚或低於相距 117 kb 的 BrA7\_6633K，推測 *BrMs<sup>a</sup>* 與 *BrMs<sup>c</sup>* 周邊結構的變異，除包含 BrMs\_indel 及 BrA7\_6746K 之區域外，亦延伸至包含 BrA7\_6749K 之區域，唯 6,749 kb 附近之序列尚未完全缺失。

TqBrA7\_6750K 檢測結果中，被該分子標誌鑑別為 *BrMs<sup>c</sup>Ms<sup>c</sup>* 基因型的 7 個體因測交後裔產生雄可稔植株，致正確率無法達 100%，然而這些個體的測交後裔均為分離狀態而非全為雄可稔植株，由於並沒有同質結合的個體被錯誤鑑別，推測該可稔植株花粉的產生，可能另由其他遺傳機制控制，事實上 TqBrA7\_6750K 恰位於預測基因 “*Bra014989*” 上，其與相鄰之 “*Bra014988*” 均為本研究第六章基因精細定位結果中的候選基因，由於 “*Bra014988*” 近似於阿拉伯芥基因 “*AT3G23990*”，而 “*AT3G23990*” 突變體之花粉具不稔性且受環境因子影響 (Haslam et al., 2015)，由於該二基因相距甚近，而本研究之 7 個外表型與 TqBrA7\_6750K 鑑別有出入的個體中，其於前述二個候選基因是否產生變異導致稔性表現之差異，有待進一步探究，然本研究開發之 TqBrA7\_6750K 可用以檢測 *B. rapa* 之 *BrMs* 基因型，於雄不稔種原之篩選及雜交後代選拔之育種實作過程上仍具有顯著之效益。

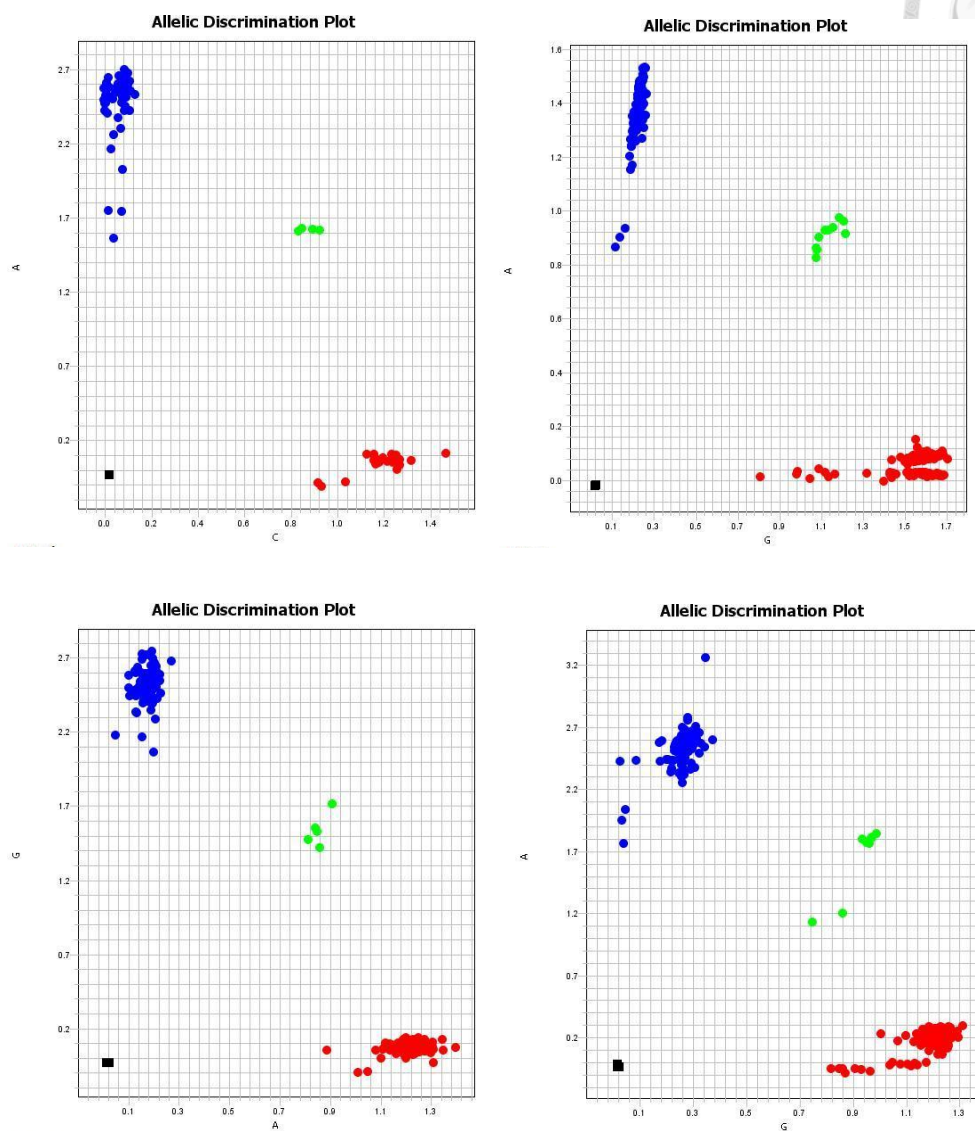
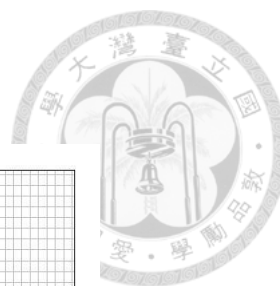
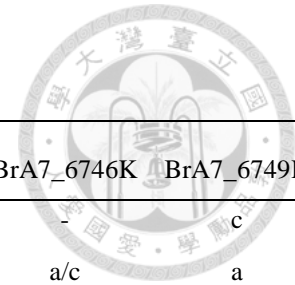


圖 6-10 透過 GWAS 開發的 SNP 分子標誌基因分型效果  
左上:BrA7\_6633K; 右上:TqBrA7\_6750K;  
左下:BrA7\_6824K; 右下:BrA7\_6963K

表 6-5 本研究開發之分子標誌對 *B. rapa* 品系進行基因分型的結果



No.	Origin	Genotype	TqBrA7_6750K	BrA7_6633K	BrA7_6824K	BrA7_6863K	BrMs_indel	BrA7_6746K	BrA7_6749K
B004-1	夏季王	c	c	c	c	c	-	-	c
B009-1	先風橋紅心-B	a	a	a	a	c	a	a/c	a
B009-2	先風橋紅心-B	a	a	a	a	c	a	a/c	a
B012-2	耐儲狀元-B	c	c	c	c	c	a	-	a
B016-1	真好吃-C	a/c	a/c	a	c	c	-	-	c
B016-2	真好吃-C	c	c	a	a/c	a/c	a	a	a/c
B017-2	真好吃-D	c	c	c	c	c	-	-	a
B018-1	膠蔬夏季王	c	c	c	c	c	-	-	a
B018-2	膠蔬夏季王	c	c	c	c	c	-	-	a
B022-1	夏秋 55-A	a/c	c	c	c	c	-	-	c
B022-2	夏秋 55-A	a/c	c	c	c	c	-	-	c
B023-1	夏秋 55-B	c	c	a	c	c	a	-	c
B023-2	夏秋 55-B	c	c	a/c	c	c	-	-	c
B025-2	義和夏陽白	c	c	c	c	c	-	-	c
B026-1	津寶 2 號	a	a	a/c	a/c	a/c	a	c	a/c
B026-2	津寶 2 號	a/c	a/c	a	a	a	a	c	a
B028-1	巨龍 58	a	a	c	a	a	a	a	a
B028-2	巨龍 58	a	a	c	a	a	a	a	a

表 6-5 (續)

No.	Origin	Genotype	TqBrA7_6750K	BrA7_6633K	BrA7_6824K	BrA7_6863K	BrMs_indel	BrA7_6746K	BrA7_6749K
B030-2	鴻盛三優 523	a	a	a	a	a	a	a	a
B032-1	紅寶娃-A	a/c	c	a	c	c	a	-	c
B033-2	紅寶娃-B	a/c	c	c	c	c	-	-	c
B043-2	東京春-A	a	a	a	a	c	a	a/c	a
B046-1	CC-200-A	a	a	a	a	a	a	c	a
B049-2	合歡三和 059 球白菜	c	c	c	c	c	-	-	c
B050-1	曉曉橘芯	a/c	a/c	a	c	c	-	-	c
B050-2	曉曉橘芯	c	c	a	a/c	a/c	a	c	a/c
B051-1	一品菜	a	a	a	a	a	a	c	a
B054-1	高麗貝貝-A	a	a	a	a	a	a	c	a
B055-1	高麗貝貝-B	a	a	a	a	a	a	c	a
B055-2	高麗貝貝-B	a	a	a	a	a	a	c	a
B057-1	首爾娃娃-B	a	a	a	a	a	a	c	a
B057-2	首爾娃娃-B	a	a	a	a	a	a	c	a
B060-1	春黃美	a	a	a	a	a	a	c	a
B061-2	春大將	c	c	c	c	c	-	-	a
B062-1	大福 234	a	a	a	a	a	a	c	a
B066-1	勝利結球白菜	c	c	c	c	c	a	-	a
B066-2	勝利結球白菜	c	c	c	c	c	-	-	a



表 6-5 (續)

No.	Origin	Genotype	TqBrA7_6750K	BrA7_6633K	BrA7_6824K	BrA7_6863K	BrMs_indel	BrA7_6746K	BrA7_6749K
B068-2	オレンゾクイソ-A	a	a	a	a	c	a	a/c	a
B071-1	金將二號	c	c	c	c	c	-	-	a
B072-1	冬峠-A	c	c	c	c	c	-	-	a
B077-2	プチヒリ	c	c	c	c	c	-	-	c
B078-2	晩抽プチヒリ	c	c	c	c	c	-	-	c
B079-2	おいしい白菜	c	c	c	c	c	a	-	a
B083-1	天綠 711 特油青菜心	a	a	a	a	a	c	c	a
B084-1	夏綠 388 油青早菜心	a	a	a	a	a	c	a/c	a
B086-1	中南短葉油綠菜心王	a	a	a	a	a	c	a/c	a
B086-2	中南短葉油綠菜心王	a	a	a	a	a	c	a/c	a
B090-2	早生金品 21-A	a	a	a	a	a	a	a/c	a
B094-1	葵扇矮腳黑葉白菜	a	a	a	a	a	c	a/c	a
B095-2	黃葉白菜種	a	a	a	a	a	a	a	a
B096-1	夏寶黑葉白梗白菜	a	a	a	c	c	a	a/c	a
B096-2	夏寶黑葉白梗白菜	a	a	a	c	c	a	a/c	a
B097-1	鴻盛金牌六號白菜	a	a	a	a	a	a	a	a
B097-2	鴻盛金牌六號白菜	a	a	a	a	a	a	a	a
B098-2	早熟五號小王子	a	a	a	a	a	a	a/c	a
B099-1	鴻盛金牌五號白菜	a	a	a	a	a	a	c	a

表 6-5 (續)

No.	Origin	Genotype	TqBrA7_6750K	BrA7_6633K	BrA7_6824K	BrA7_6863K	BrMs_indel	BrA7_6746K	BrA7_6749K
B099-2	鴻盛金牌五號白菜	a	a	a	a	a	a	c	a
B102-1	AG90	a	a	a	a	a	c	a/c	a
B104-1	青龍	a	a	a	a	a	-	a/c	a
B109-1	福祿甜油菜	a/c	a/c	c	c	c	-	-	c
B109-2	福祿甜油菜	c	c	a/c	a/c	a/c	a	a/c	a/c
B110-1	中原奶白菜	a	a	a	a	a	c	a/c	a
B110-2	中原奶白菜	a	a	a	a	a	c	a/c	a
B111-1	夏三味	a	a	a	a	c	a	c	a
B111-2	夏三味	a	a	a	a	c	a	c	a
B115-1	早熟5號(優選)	a	a	a	a	a	a	c	a
B116-1	德陽01-A	c	c	c	c	c	-	-	c
B117-1	德陽01-B	a/c	c	c	c	c	a	-	c
B118-1	536小白菜	a	a	a	a	a	a	c	a
B120-2	津育60	c	c	a	c	c	-	-	c
B121-1	TIEN JIN	c	c	c	c	c	-	-	c
B122-1	GIANT	a/c	c	c	c	c	-	-	c
B123-1	GIANT 029	a/c	a/c	a/c	a/c	a/c	a	a/c	a/c
B124-2	優選快菜369	a	a	a	a	a	a	c	a
B125-1	無敵快菜	a	a	a	a	a	a	a	a

表 6-5 (續)

No.	Origin	Genotype	TqBrA7_6750K	BrA7_6633K	BrA7_6824K	BrA7_6863K	BrMs_indel	BrA7_6746K	BrA7_6749K
B126-2	豐抗 78	c	c	c	c	c	-	-	a
B127-1	西白 45 (金童子) -A	c	c	c	c	c	-	-	c
B128-1	西白 45 (金童子) -B	c	c	a	c	c	-	-	c
B130-1	夏暉早 38 天 (632) - A	a	a	a	a	c	a	a/c	a
B131-1	夏暉早 38 天 (632) - B	a	a	a	a	a	a	a	a
B133-1	和冠 (403) -A	a/c	c	c	c	c	-	-	c
B134-2	和冠 (403) -B	c	c	a	c	c	-	-	c
B135-2	春冠 -A	a	a	a	a	c	a	a/c	a
B136-1	春冠 -B	c	c	c	c	c	-	-	c
B137-1	正板夏秋 3 號	c	c	c	c	c	a	-	c
B138-2	興滇高抗 4	a	a	a	a	a	a	c	a
B139-1	板扎黃帝	c	c	c	c	c	-	-	a
B140-1	吉祥	c	c	c	c	c	-	-	c

a:鑑別為  $BrMs^aMs^a$ ; c:鑑別為  $BrMs^cMs^c$ ; a/c:鑑別  $BrMs^aMs^c$ ; -:無螢光反應





## 第7章 總結

白菜是重要的蔬菜來源之一，台灣每年栽培結球白菜與不結球白菜合計超過 7,000 公頃，雖然大都採用一代雜交品種，但因缺乏適當雄不稔源，致使目前大多數品種主要仍以自交不親合性採種，本研究針對於台灣小油菜田區中發現的自發性雄不稔株進行研究，期望能將該特性利用於白菜類作物之採種用途。

本研究第二章整合三種後裔檢定方式進行該雄不稔性之遺傳分析，結果顯示其遺傳符合單基因座之複等位基因模式，並將該基因命名為 *BrMs*，而三個複等位基因則分別為 *BrMs<sup>a</sup>*、*BrMs<sup>b</sup>* 與 *BrMs<sup>c</sup>*，其中控制雄不稔性表現者為 *BrMs<sup>b</sup>*，彼此的顯隱性關係為 *BrMs<sup>a</sup>* > *BrMs<sup>b</sup>* > *BrMs<sup>c</sup>*；而本研究第三章及第四章分別以不同等位基因建構之分離族群進行基因定位，由於 *BrMs<sup>a</sup>/BrMs<sup>b</sup>* 與 *BrMs<sup>b</sup>/BrMs<sup>c</sup>* 之基因定位區間重疊，因此進一步驗證了遺傳模式的正確性。

為了釐清小油菜 *BrMs* 與近緣物種之細胞核雄不稔基因之異同，並開發其緊密連鎖之分子標誌，而於第五章中進行了基因的精細定位，透過遺傳距離的差異，我們推測同物種結球白菜上所發現的雄不稔基因 *Ms*，與本研究 *BrMs* 應非相同基因。此外，藉由比較基因體之方法發現大油菜 *BnRf* 基因座落於本研究的定位區間內，而利用小油菜之 *BnRf* ortholog 開發的 Indel 分子標誌 *BrMs\_indel* 於 923 個 F<sub>2</sub> 植株中進行驗證時，由於外表型與基因型完全一致，初步認定 *BrMs* 與 *BnRf* 間為同源基因，然而進一步於更大的分離族群 (4,415 株) 中進行精細定位時，發現有 2 株分離後裔其 *BnRf* ortholog 上的 Indel 分子標誌之基因型與雄不稔外表型不一致，由於 *BrMs* 與 *BnRf* 遺傳行為相似，且不結球白菜基因體上 *BrMs* 與 *BnRf* ortholog 位置相鄰，推測其為複合基因 (gene complex) 內重組之結果，而本章開發之分子標誌亦為目前白菜類物種中最接近雄不稔基因者，其中不僅開發出與 *BrMs* 共分離的 *BrA7\_6746K* 與 *BrA7\_6749K* 分子標誌，基因精細定位過程中於基因周邊的 10 個 SNP 分子標誌亦可應用於育種時之輔助篩選，除此之外，為了進一步區分正常稔性個體之 *BrMs* 基因型，本研究第六章透過全基因體關連性分析開發出可鑑別 *BrMs<sup>a</sup>* 與 *BrMs<sup>c</sup>* 等位基因之 SNP 分子標誌 *TqBrA7\_6750K*，整合前述這些分子標誌，已建構完成雄不稔性育種時完整的分子輔助工具。

在全基因體關聯性分析過程中，小油菜雄不稔系與不同白菜類作物之測交結果，證實該雄不稔基因亦可作用於不同白菜類作物，顯示可擴大 *BrMs* 之應用範圍。而

基因精細定位之結果將 *BrMs* 基因之物理區間限縮於 24 kb 範圍內，該區間預測只有 3 個基因，分別為 *Bra014987*、*Bra014988* 及 *Bra014989* 等，事實上全基因體關聯性分析所獲得之 TqBrA7\_6750K 恰位於 *Bra014989* 上，並與 BrA7\_6749K 僅相距 56 bp，很有可能即是 *BrMs* 基因，此結果可供未來進行選殖該基因時之參考。

實務應用上，本研究 *BrMs* 的遺傳模式經評估具有商業應用價值，其雜交 F<sub>1</sub> 種子之採種模式如圖 7-1 所示。

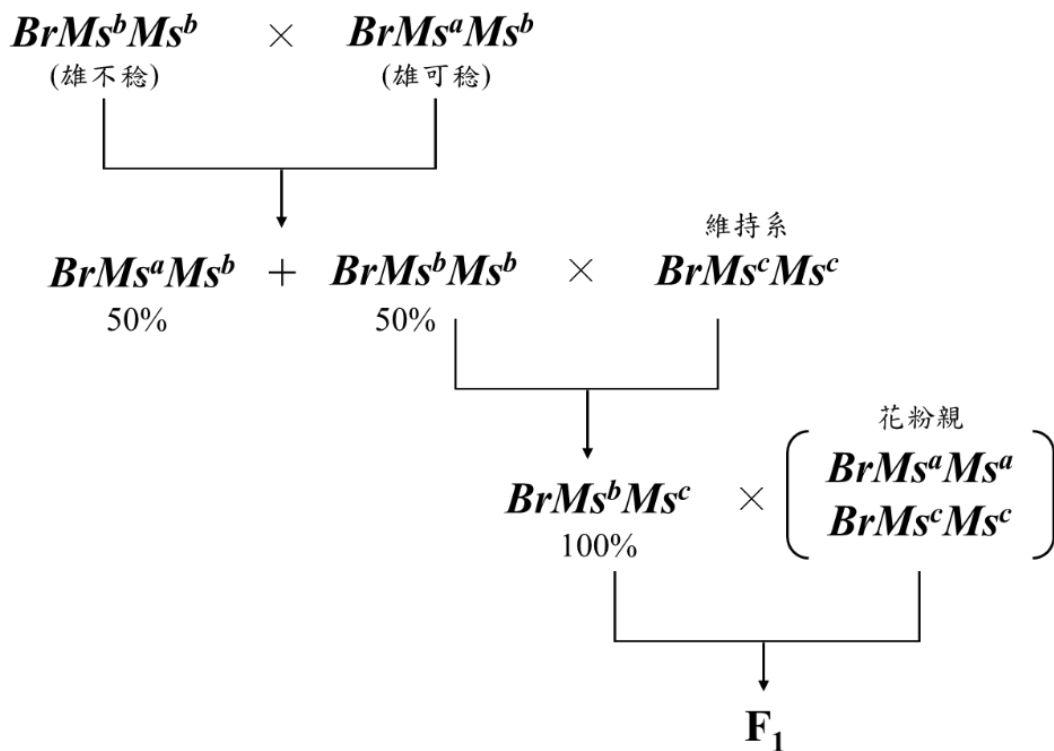



圖 7-1 利用雄不稔基因 *BrMs* 生產 F<sub>1</sub> 雜交種子的採種模式



進行採種時需包含三個品系，即雄不稔系、維持系與雜交種父本（花粉親）等，其中雄不稔系中包含二種基因型，品系內會產生 50% 雄不稔株，雜交  $F_1$  種子生產時採二階段進行，第一階段是將雄不稔系與維持系栽培田間後，於雄不稔系開花時拔除可稔株，保留雄不稔株與臨時維持系雜交，事實上淘汰可稔株可透過本研究開發之緊密連鎖分子標誌於育苗期進行，可省卻田間拔除之困擾，第一階段所收穫的種子將全數為雄不稔株，第二階段再利用此 100% 雄不稔後裔做為母本與父本花粉親雜交生產商業用  $F_1$  種子，整個採種模式中，如果維持系園藝性狀異於雄不稔系，則本法採收之雜交後裔屬三交種，若為求  $F_1$  之整齊度，可利用回交將維持系培育成與雄不稔系相似之近同源系，其與雄不稔系之差異僅在雄不稔基因型，如此採得之雜交種子則為單交種。

這種二階段採種模式之所以具可行性，與作物的繁殖力有關。一般白菜類作物單株可以產種超過 5,000 粒，換言之其繁殖倍率可達 5,000 倍，透過二階段採種相當於繁殖了  $(5,000)^2 = 25,000,000$  倍，以台灣不結球白菜每年 5,000 公頃栽培面積之種子需求總量進行估算，第一階段只需於雄不稔系中針對 400 株進行分子篩選，所獲得之雄不稔株數量即可滿足需求。此外，利用此法採種也具有近似於透過 CMS 採種之品種保護優點，即第一階段之母本種子生產於台灣進行，由於該母本種子全數均具雄不稔性，第二階段採種時可將其連同花粉親送至採種成本低廉之國家生產  $F_1$  種子，而免除雙親本同時流失的風險。



除了採種模式外，欲進行 *BrMs* 基因轉移時建議可採下列方式，首先，利用 TqBrA7\_6750K 先鑑別試驗材料之 *BrMs* 基因型，由於整個採種系統中，雄不稔系與維持系合計需包含三種等位基因（圖 7-1），因此當欲轉移的自交系對象經判定為 *BrMs<sup>a</sup>Ms<sup>a</sup>* 時，它需與基因型為 *BrMs<sup>b</sup>Ms<sup>c</sup>* 之個體雜交（圖 7-2），並由後代分離的 *BrMs<sup>a</sup>Ms<sup>b</sup>* 及 *BrMs<sup>a</sup>Ms<sup>c</sup>* 分別持續與該自交系回交，回交過程中以本研究開發之 SNP 分子標誌進行前景選拔（foreground selection），若能搭配該品系之遺傳背景進行選拔（background selection），更能提高回交效率，回交完成後再以前景選拔分子標誌確認，進而培育出雄不稔基因型分別為 *BrMs<sup>a</sup>Ms<sup>b</sup>*、*BrMs<sup>b</sup>Ms<sup>b</sup>* 及 *BrMs<sup>c</sup>Ms<sup>c</sup>* 之近同源系，再利用於採種工作。

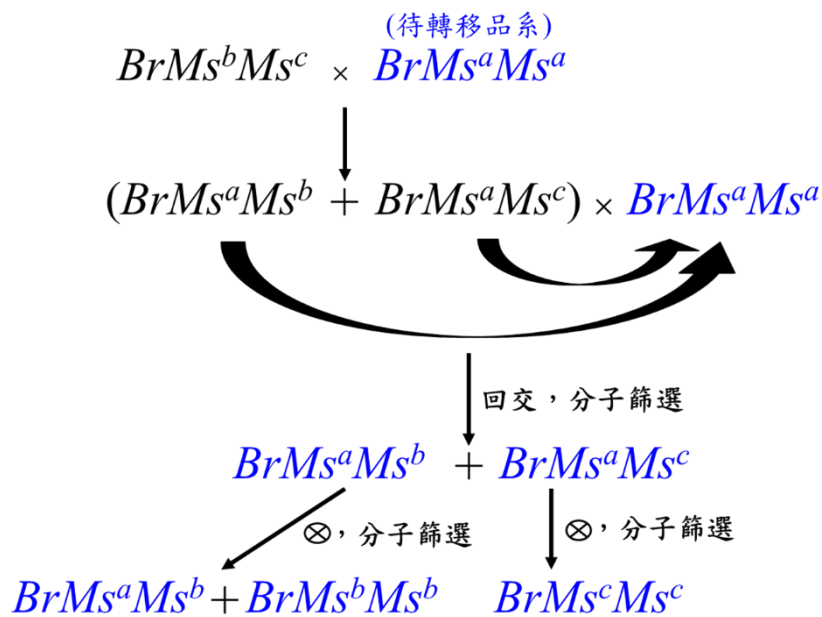


圖 7-2 當自交系基因型為 *BrMs<sup>a</sup>Ms<sup>a</sup>* 時的雄不稔系培育方法

經判定自交系基因型為  $BrMs^cMs^c$  時，它需與基因型為  $BrMs^aMs^b$  之個體雜交(圖 7-3)，並由後代分離的  $BrMs^aMs^c$  及  $BrMs^bMs^c$  基因型分別持續與該自交系回交，回交過程同樣進行分子篩選，完成後再以前景選拔分子標誌篩選出  $BrMs^aMs^b$  個體，自交一次再篩選出  $BrMs^aMs^b$  與  $BrMs^bMs^b$  基因型以組成雄不稔系。雄不稔性轉移時，伴隨待轉移品系  $BrMs$  基因型之不同，往往需要以  $BrMs^bMs^c$  或  $BrMs^aMs^b$  基因型之個體當作貢獻親進行雜交，因此進行育種前，可先培育出含  $BrMs^bMs^c$  與  $BrMs^aMs^b$  基因型之近同源系，再視待轉移品系基因型，選擇適當個體進行雜交，由於近同源系遺傳背景一致，欲於回交過程進行遺傳背景選拔時，可減輕開發背景選拔用分子標誌的工作量。

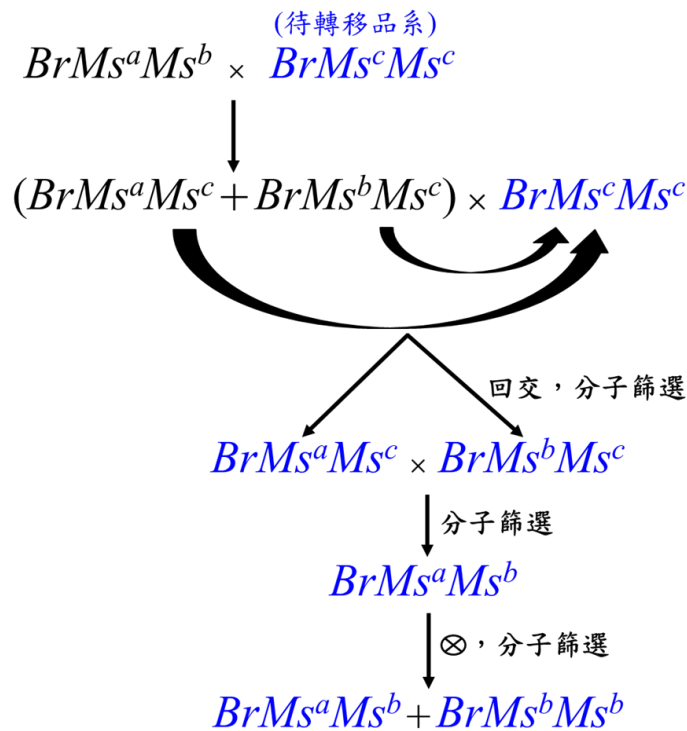




圖 7-3 當自交系基因型為  $BrMs^cMs^c$  時的雄不稔系培育方法

關於後續研究方向，可著重於  $BrMs$  基因之確認與選殖，甚或進行基因編輯加以調控，此外，部份品系經 TqBrA7\_6750K 鑑別為  $BrMs^cMs^c$  者，其與雄不稔系雜交後裔會出現部份雄可稔個體，將持續探討該現象是否存在其他遺傳機制，相信這些研究結果，將更能強化本研究新發現雄不稔源的開發與應用。

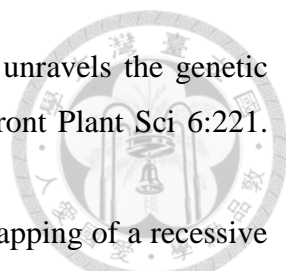
## 參考文獻

- 
- 何余堂、陳寶元、傅廷棟、李殿榮、涂金星 (2003) 白菜型油菜在中國的起源與進化。遺傳學報 30 :1003-1012。
- 宋來強、傅廷棟、楊光聖、涂金星、馬朝芝 (2005) 1 對複等位基因控制的油菜 (*Brassica napus* L.) 顯性核不育系 609AB 的遺傳驗證。作物學報 31:869-875。
- 李殿榮 (1980) 甘藍型油菜三系選育初報。陝西農業科學 (1):26-29。
- 林俊義 (2006) 能源作物之國外推動經驗與國內發展展望(下)。技術服務 65:5-10。
- 邱祥發、彭武男 (1995) 低芥酸油菜品種之育種研究。桃園區農業改良場研究彙報第 22 號。
- 胡勝武、于澄宇、趙惠賢、路明、張春紅、俞延軍 (2004) 甘藍型油菜核不育材料 Shaan-GMS 恢復基因的篩選及其遺傳分析。西北農林科技大學學報(自然科學版) 32:9-18。
- 張慧、張淑江、武劍、李菲、章時蕃、王曉武、張新梅、孫日飛 (2010) 大白菜細胞核隱性雄性不育系恢復基因 BrMf2 的標記及定位。中國農業科學 43 :993-999。
- 曹家樹、曹壽椿 (1995) 大白菜起源的雜交驗證初報。園藝學報 22 :93-94。
- 曹壽椿、李式軍 (1982) 白菜地方品種的初步研究—Ⅲ.不結球白菜品種的園藝學分類。南京農業大學學報 (2):30-37。
- 郭晶心、周乃元、馬榮才、曹鳴慶 (2002) 白菜類蔬菜遺傳多樣性的 AFLP 分子標記研究。農業生物技術學報 10:138-143。
- 馮輝、魏毓棠、許明 (1995) 大白菜核基因雄性不育系遺傳假說及其驗證。中國科協第二屆青年學術年會園藝論文集 pp 458-466。
- 楊光堃、傅廷棟 (1998) 甘藍型油菜細胞質雄性不育的遺傳分類研究。中國農業科學 31:27-31。
- 董振生、劉創社、景軍勝、冉龍貴、張修森、董軍剛、劉詢俠 (1999) 白菜型油菜雙顯性核不育 896AB 恢復系基因型的鑑定。作物學報 25 :193-198。
- 劉后利 (1984) 幾種芸薹屬油菜的起源和進化。作物學報 10:9-18。
- 劉定富 (1992) 植物顯性核不育恢復性遺傳的理論探討。遺傳 14:31-36。
- 鄭仲孚、趙致康、吳國璋、林興 (1963) 台灣雜糧作物品種圖說。台灣省政府農林廳 pp 251-267。

- 
- 魏大勇、譚傳東、崔藝馨、吳道明、李加納、梅家琴、錢偉 (2017) 甘藍型油菜 *polCMS* 育性恢復位點的全基因組關聯分析。中國農業科學 50:802-819。
- 譚其猛 (1979) 試論大白菜品種的起源、分布和演化。中國農業科學 12 :68-75。
- Ashikawa I, Hayashi N, Yamane H, Kanamori H, Wu J, Matsumoto T, Ono K, Yano M (2008) Two adjacent nucleotide-binding site-leucine-rich repeat class genes are required to confer *Pikm*-specific rice blast resistance. *Genetics* 180 :2267-2276.
- Bannerot H, Boulidard L, Cauderon Y, Tempe J (1974) Transfer of cytoplasmic male sterility from *Raphanus sativus* to *Brassica oleracea*. In: Wills AB, North C (Eds) *Proc Eucarpia Meet Cruciferae*. Scottish Horticult Res Inst, Invergowrie, UK, pp 52–54.
- Bannerot H, Boulidard L, Chupeau Y (1977) Unexpected difficulties met with the radish cytoplasm in *Brassica oleracea*. *Eucarpia Cruciferae Newsletter* 2: 16.
- Borkakati RR, Virmani SS (1996) Genetics of thermosensitive genic male sterility in rice. *Euphytica* 88:1-7.
- Broman KW (2010) Genetic Map Construction with R/qtl, pp. Technical Report # 214. University of Wisconsin-Madison, Department of Biostatistics & Medical Informatics.
- Broman KW, Wu H, Sen S, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19:889-890.
- Cao S, Li S (1981) Breeding of summer Chinese cabbage "Dwarf Hybrid No. 1" and male sterile AB lines. *Acta Horticulturae Sinica* 8:35-42
- Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) *Stacks*: building and genotyping loci *De Novo* from short-read sequences. *G3: Genes Genomes Genetics* 1:171-182.
- Chang C-W, Wang Y-H, Tung C-W (2017) Genome-wide single nucleotide polymorphism discovery and the construction of a high-density genetic map for melon (*Cucumis melo* L.) using genotyping-by-sequencing. *Front Plant Sci* 8:125.
- Chen F, Hu B, Li C, Li Q, Chen W, Zhang M (1998) Genetic studies on GMS in *Brassica napus* L: I. Inheritance of recessive GMS line 9012A. *Acta Agronomica Sinica* 24:431-438.
- Cheng F, Liu S, Wu J, Fang L, Sun S, Liu B, Li P, Hua W, Wang X (2011) BRAD, the genetics and genomics database for Brassica plants. *BMC Plant Biol* 11 :136.
- Choi S, Teakle G, Plaha P, Kim J, Allender C, Beynon E, Piao Z, Soengas P, Han T, King

- G, Barker G, Hand P, Lydiate D, Batley J, Edwards D, Koo D, Bang J, Park B-S, Lim Y (2007) The reference genetic linkage map for the multinational *Brassica rapa* genome sequencing project. *Theor Appl Genet* 115:777-792.
- Chung H, Jeong Y-M, Mun J-H, Lee S-S, Chung W-H, Yu H-J (2014) Construction of a genetic map based on high-throughput SNP genotyping and genetic mapping of a TuMV resistance locus in *Brassica rapa*. *Mol Genet Genomics* 289 :149-160.
- Delourme R, Bouchereau A, Hubert N, Renard M, Landry BS (1994) Identification of RAPD markers linked to a fertility restorer gene for the *Ogura* radish cytoplasmic male sterility of rapeseed (*Brassica napus* L.). *Theor Appl Genet* 88:741-748.
- Delourme R, Foisset N, Horvais R, Barret P, Champagne G, Cheung WY, Landry BS, Renard M (1998) Characterisation of the radish introgression carrying the *Rfo* restorer gene for the *Ogu*-INRA cytoplasmic male sterility in rapeseed (*Brassica napus* L.). *Theor Appl Genet* 97:129-134.
- Deng Z, Li X, Wang Z, Jiang Y, Wan L, Dong F, Chen F, Hong D, Yang G (2016) Map-based cloning reveals the complex organization of the *BnRf* locus and leads to the identification of *BnRf<sup>b</sup>*, a male sterility gene, in *Brassica napus*. *Theor Appl Genet* 129:53-64.
- Dong F, Hong D, Xie Y, Wen Y, Dong L, Liu P, He Q, Yang G (2012) Molecular validation of a multiple-allele recessive genic male sterility locus (*BnRf*) in *Brassica napus* L. *Mol Breed* 30:1193-1205.
- Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6 (5):e19379. doi:10.1371/journal.pone.0019379
- Engelke T, Gera D, Tatlioglu T (2004) Determination of the frequencies of restorer- and maintainer-alleles involved in CMS<sub>1</sub> and CMS<sub>2</sub> in German chive varieties. *Plant Breed* 123:51-59.
- Fan ZX, Lei WX, Hong DF, He JP, Wan LL, Xu ZH, Liu PW, Yang GS (2007) Development and primary genetic analysis of a fertility temperature-sensitive polima cytoplasmic male sterility restorer in *Brassica napus*. *Plant Breed* 126:297-301.
- Fang Z, Sun P, Liu Y, Yang L, Wang X, Hou A, Bian C (1997) A male sterile line with dominant gene (Ms) in cabbage (*Brassica oleracea* var. *capitata*) and its utilization for hybrid seed production. *Euphytica* 97 :265-268.
- Farfan IDB, De La Fuente GN, Murray SC, Isakeit T, Huang P-C, Warburton M, Williams

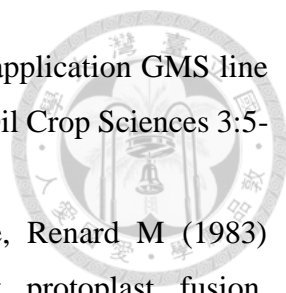
- P, Windham GL, Kolomiets M (2015) Genome Wide Association Study for Drought, Aflatoxin Resistance, and Important Agronomic Traits of Maize Hybrids in the Sub-Tropics. PLoS ONE 10:e0117737. doi:10.1371/journal.pone.0117737
- Feng H, Wei P, Piao Z-Y, Liu Z-Y, Li C-Y, Wang Y-G, Ji R-Q, Ji S-J, Zou T, Choi S-R, Lim Y-P (2009) SSR and SCAR mapping of a multiple-allele male-sterile gene in Chinese cabbage (*Brassica rapa* L.). Theor Appl Genet 119:333-339.
- Feng H, Wei Y, Ji S, Jin G, Jin J, Dong W Multiple allele model for genic male sterility in Chinese cabbage. In: III International Symposium Diversification of Vegetable Crops 467, 1996. pp 133-138.
- FitzJohn RG, Armstrong TT, Newstrom-Lloyd LE, Wilton AD, Cochrane M (2007) Hybridisation within *Brassica* and allied genera: evaluation of potential for transgene escape. Euphytica 158:209-230.
- Fu T, Yang G, Yang X (1990) Studies on “Three Line” Polima Cytoplasmic Male Sterility Developed in *Brassica napus* L. Plant Breed 104:115-120.
- Ganal MW, Durstewitz G, Polley A, Bérard A, Buckler ES, Charcosset A, Clarke JD, Graner E-M, Hansen M, Joets J, Le Paslier M-C, McMullen MD, Montalent P, Rose M, Schön C-C, Sun Q, Walter H, Martin OC, Falque M (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:e28334. doi:10.1371/journal.pone.0028334
- Hackett C, Broadfoot L (2003) Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. Heredity 90:33.
- Hall TA BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In: Nucleic Acids Symposium Series, 1999. vol 41. Information Retrieval Ltd., [London] c1979-c2000., pp 95-98.
- Hansen M, Halldén C, Nilsson N-O, Säll T (1997) Marker-assisted selection of restored male-fertile *Brassica napus* plants using a set of dominant RAPD markers. Mol Breed 3:449-456.
- Haslam TM, Haslam R, Thoraval D, Pascal S, Delude C, Domergue F, Fernández AM, Beaudoin F, Napier JA, Kunst L, Joubès J (2015) ECERIFERUM2-LIKE proteins have unique biochemical and physiological functions in very-long-chain fatty acid elongation. Plant Physiol 167:682-692.
- Hatzig SV, Frisch M, Breuer F, Nesi N, Ducournau S, Wagner MH, Leckband G, Abbadi

- 
- A, Snowdon RJ (2015) Genome-wide association mapping unravels the genetic control of seed germination and vigor in *Brassica napus*. *Front Plant Sci* 6:221. doi:10.3389/fpls.2015.00221
- He J, Ke L, Hong D, Xie Y, Wang G, Liu P, Yang G (2008) Fine mapping of a recessive genic male sterility gene ( *Bnms3* ) in rapeseed ( *Brassica napus* ) with AFLP- and *Arabidopsis* -derived PCR markers. *Theor Appl Genet* 117:11-18
- He YQ, Yang J, Xu CG, Zhang ZG, Zhang Q (1999) Genetic bases of instability of male sterility and fertility reversibility in photoperiod-sensitive genic male-sterile rice. *Theor Appl Genet* 99:683-693.
- Heng S, Shi D, Hu Z, Huang T, Li J, Liu L, Xia C, Yuan Z, Xu Y, Fu T, Wan Z (2015) Characterization and classification of one new cytoplasmic male sterility (CMS) line based on morphological, cytological and molecular markers in non-heading Chinese cabbage (*Brassica rapa* L.). *Plant Cell Rep* 34:1529-1537.
- Hong D, Wan L, Liu P, Yang G, He Q (2006) AFLP and SCAR markers linked to the suppressor gene (*Rf*) of a dominant genetic male sterility in rapeseed (*Brassica napus* L.). *Euphytica* 151:401-409.
- Hong DF, Liu J, Yang GS, He QB (2008) Development and characterization of SCAR markers associated with a dominant genic male sterility in rapeseed. *Plant Breed* 127:69-73.
- Huang L, Yang Y, Zhang F, Cao J (2017) A genome-wide SNP-based genetic map and QTL mapping for agronomic traits in Chinese cabbage. *Sci Rep* 7:46305. doi:10.1038/srep46305
- Huang Z, Chen Y, Yi B, Xiao L, Ma C, Tu J, Fu T (2007) Fine mapping of the recessive genic male sterility gene ( *Bnms3* ) in *Brassica napus* L. *Theor Appl Genet* 115:113-118.
- Imai R, Koizuka N, Fujimoto H, Hayakawa T, Sakai T, Imamura J (2003) Delimitation of the fertility restorer locus *Rfk1* to a 43-kb contig in Kosena radish (*Raphanus sativus* L.). *Mol Genet Genomics* 269:388-394.
- Jean M, Brown GG, Landry BS (1997) Genetic mapping of nuclear fertility restorer genes for the 'Polima' cytoplasmic male sterility in canola (*Brassica napus* L.) using DNA markers. *Theor Appl Genet* 95:321-328.
- Kaul MLH (1988) Male sterility in higher plant. Monographs on Theoretical and Applied Genetics, vol 10. Springer-Verlag, Germany
- Ke LP, Sun YQ, Hong DF, Liu PW, Yang GS (2005) Identification of AFLP markers

- linked to one recessive genic male sterility gene in oilseed rape, *Brassica napus*. *Plant Breed* 124:367-370.
- Kenji W, Takashi S, Eimi F, Kenji K, Daizo I, Junzo F (2013) Genic male sterility in *Brassica rapa* ssp. *rapa* cv. 77B. *J of Agri Sci* 57 :287-292.
- Kim H, Choi SR, Bae J, Hong CP, Lee SY, Hossain MJ, Van Nguyen D, Jin M, Park B-S, Bang J-W, Bancroft I, Lim YP (2009) Sequenced BAC anchored reference genetic map that reconciles the ten individual chromosomes of *Brassica rapa*. *BMC Genomics* 10 (1):432. doi:10.1186/1471-2164-10-432
- Kosambi DD (1943) The estimation of map distances from recombination values. *Annals of Eugenics* 12:172-175.
- Lee J, Izzah NK, Choi B-S, Joh HJ, Lee S-C, Perumal S, Seo J, Ahn K, Jo EJ, Choi GJ, Nou I-S, Yu Y, Yang T-J (2015) Genotyping-by-sequencing map permits identification of clubroot resistance QTLs and revision of the reference genome assembly in cabbage (*Brassica oleracea* L.). *DNA Res* 23:29-41.
- Leflon M, Eber F, Letanneur J, Chelysheva L, Coriton O, Huteau V, Ryder C, Barker G, Jenczewski E, Chèvre A (2006) Pairing and recombination at meiosis of *Brassica rapa* (AA) × *Brassica napus* (AACC) hybrids. *Theor Appl Genet* 113:1467-1480.
- Lei S, Yao X, Yi B, Chen W, Ma C, Tu J, Fu T (2007) Towards map-based cloning: fine mapping of a recessive genic male-sterile gene (*BnMs2*) in *Brassica napus* L. and syntenic region identification based on the *Arabidopsis thaliana* genome sequences. *Theor Appl Genet* 115:643-651.
- Levings CS, III (1990) The Texas cytoplasm of maize: cytoplasmic male sterility and disease susceptibility. *Science* 250:942-947. doi:10.1126/science.250.4983.942
- Li CW (1981) The origin, evolution, taxonomy and hybridization of Chinese cabbage. Chinese cabbage, *Proceedings of the First International Symposium*:3-11.
- Li SL, Qian YX, Wu ZH (1985) Inheritance of genic male sterility in *Brassica napus* and its application to commercial production. *Acta Agriculturae Shanghai* 1:1-12.
- Li SL, Qian YX, Wu ZH, Stefansson BR (1988) Genetic male sterility in rape (*Brassica napus* L.) conditioned by interaction of genes at two loci. *Canadian Journal of Plant Science* 68:1115-1118.
- Li SL, Zhou ZJ, Zhou XR (1993) Inheritance of recessive genic male sterile line S45AB of rape (*Brassica napus* L.). *Acta Agriculturae Shanghai* 9:1-7
- Li X, Wang A, Zu F, Hu Z, Lin J, Zhou G, Tu J (2016) Identification of a nuclear-recessive gene locus for male sterility on A2 chromosome using the Brassica 60 K SNP array



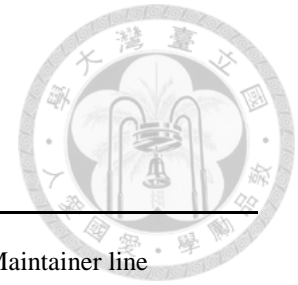
- in non-heading Chinese cabbage. *Genes & Genomics* 38:1151-1157.
- Liu J, Hong D, Lu W, Liu P, He Q, Yang G (2008) Genetic analysis and molecular mapping of gene associated with dominant genic male sterility in rapeseed (*Brassica napus* L.). *Genes & Genomics* 30:523-532.
- Lu GY, Yang GS, Fu TD (2004) Molecular mapping of a dominant genic male sterility gene *Ms* in rapeseed (*Brassica napus*). *Plant Breed* 123:262-265.
- Lu W, Liu J, Xin Q, Wan L, Hong D, Yang G (2013) A triallelic genetic male sterility locus in *Brassica napus*: an integrative strategy for its physical mapping and possible local chromosome evolution around it. *Ann Bot* 111:305-315.
- Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *PNAS (USA)* 88:9828-9832.
- Mizuno H, Katagiri S, Kanamori H, Mukai Y, Sasaki T, Matsumoto T, Wu J (2020) Evolutionary dynamics and impacts of chromosome regions carrying R-gene clusters in rice. *Sci Rep* 10:872. doi:10.1038/s41598-020-57729-w
- Morris GP, Ramu P, Deshpande SP, Hash CT, Shah T, Upadhyaya HD, Riera-Lizarazu O, Brown PJ, Acharya CB, Mitchell SE, Harriman J, Glaubitz JC, Buckler ES, Kresovich S (2013) Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *PNAS (USA)* 110:453-458.
- Nagaharu U (1935) Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization *Jpn J Bot* 7:389-452.
- Navabi ZK, Parkin IAP, Pires JC, Xiong Z, Thiagarajah MR, Good AG, Rahman MH (2010) Introgression of B-genome chromosomes in a doubled haploid population of *Brassica napus* × *B. carinata*. *Genome* 53:619-629. doi:10.1139/G10-039
- Nieuwhof M (1961) Male sterility in some cole crops. *Euphytica* 10 (3):351-356.
- Ogura H (1968) Studies on the new male sterility in Japanese radish, with special reference to the utilization of this sterility towards the practical raising of hybrid seeds. *Memoirs of the Faculty of Agriculture* 6:39-78.
- Otyama PI, Wilkey A, Kulkarni R, Assefa T, Chu Y, Clevenger J, O'Connor DJ, Wright GC, Dezern SW, MacDonald GE, Anglin NL, Cannon EKS, Ozias-Akins P, Cannon SB (2019) Evaluation of linkage disequilibrium, population structure, and genetic diversity in the U.S. peanut mini core collection. *BMC Genomics* 20 (1):481. doi:10.1186/s12864-019-5824-9

- 
- Pan T, Zeng F, Wu S, Zhao Y (1988) A study on the breeding and application GMS line of low eruci acid in rapeseed (*B. napus*). Chinese Journal of Oil Crop Sciences 3:5-8.
- Pelletier G, Primard C, Vedel F, Chetrit P, Remy R, Rousselle, Renard M (1983) Intergeneric cytoplasmic hybridization in cruciferae by protoplast fusion. Molecular and General Genetics 191:244-250.
- Poland J, Brown P, Sorrells M, Jannink J-L (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS ONE 7 (2):e32253. doi: 10.1371/journal.pone.0032253
- Polowick PL, Sawhney VK (1988) High Temperature Induced Male and Female Sterility in Canola (*Brassica napus* L.). Ann Bot 62:83-86.
- Rajcan I, Kasha KJ, Kott LS, Beversdorf WD (2002) Evaluation of cytoplasmic effects on agronomic and seed quality traits in two doubled haploid populations of *Brassica napus* L. Euphytica 123:401-409.
- Ruffio-Chable V, Bellis H, Herve Y (1993) A dominant gene for male sterility in cauliflower (*Brassica oleracea* var. *botrytis*): phenotype expression, inheritance, and use in F1 hybrid production. Euphytica 97: 9-17.
- Sawhney VK (2004) Photoperiod-sensitive male-sterile mutant in tomato and its potential use in hybrid seed production. The Journal of Horticultural Science & Biotechnology 79:138-141.
- Sodhi YS, Chandra A, Verma JK, Arumugam N, Mukhopadhyay A, Gupta V, Pental D, Pradhan AK (2006) A new cytoplasmic male sterility system for hybrid seed production in Indian oilseed mustard *Brassica juncea*. Theor Appl Genet 114:93-99.
- Song L-Q, Fu T-D, Tu J-X, Ma C-Z, Yang G-S (2006) Molecular validation of multiple allele inheritance for dominant genic male sterility gene in *Brassica napus* L. Theor Appl Genet 113:55-62.
- Spindel JE, Begum H, Akdemir D, Collard B, Redoña E, Jannink JL, McCouch S (2016) Genome-wide prediction models that incorporate de novo GWAS are a powerful new tool for tropical rice improvement. Heredity 116:395-408.
- Sun Z, Wang Z, Tu J, Zhang J, Yu F, McVetty PBE, Li G (2007) An ultradense genetic recombination map for *Brassica napus*, consisting of 13551 SRAP markers. Theor Appl Genet 114:1305-1317.
- Suwabe K, Tsukazaki H, Iketani H, Hatakeyama K, Kondo M, Fujimura M, Nunome T,

- 
- Fukuoka H, Hirai M, Matsumoto S (2006) Simple Sequence Repeat-Based Comparative Genomics Between *Brassica rapa* and *Arabidopsis thaliana*: The Genetic Origin of Clubroot Resistance. *Genetics* 173:309-319.
- Takahata Y, Nagasaka M, Kondo H, Kaizuma N (1996) Genic male sterility in *Brassica campestris* L. *Acta Hort* 407:147-150.
- Thompson KF (1972) Cytoplasmic male sterility in oilseed rape. *Heredity* 29:253-257
- Tu J, Fu T, Zheng Y (1997) Analysis on inheritance and isolocus of the rapeseed GMS 90-2441A (*B. napus* L.). *Journal of Huazhong Agricultural University* 16:255-258.
- Van der Meer QP (1987) Chromosomal monogenic dominant male sterility in chinese cabbage (*Brassica rapa* subsp. *pekinensis* (Lour.) hanelt). *Euphytica* 36:927-931.
- Van Os H, Stam P, Visser RGF, van Eck HJ (2005) SMOOTH: a statistical method for successful removal of genotyping errors from high-density genetic linkage data. *Theor Appl Genet* 112:187-194.
- Verma JK, Sodhi YS, Mukhopadhyay A, Arumugam N, Gupta V, Pental D, Pradhan AK (2000) Identification of stable maintainer and fertility restorer lines for 'polima' CMS in *Brassica campestris*. *Plant Breed* 119:90-92.
- Wan Z, Jing B, Tu J, Ma C, Shen J, Yi B, Wen J, Huang T, Wang X, Fu T (2008) Genetic characterization of a new cytoplasmic male sterility system (*hau*) in *Brassica juncea* and its transfer to *B. napus*. *Theor Appl Genet* 116 :355-362.
- Wang X, Lou P, Bonnema G, Yang B, He H, Zhang Y, Fang Z (2005) Linkage mapping of a dominant male sterility gene *Ms-cd1* in *Brassica oleracea*. *Genome* 48:848-854.
- Wang Y, Sun S, Liu B, Wang H, Deng J, Liao Y, Wang Q, Cheng F, Wang X, Wu J (2011) A sequence-based genetic linkage map as a reference for *Brassica rapa* pseudochromosome assembly. *BMC Genomics* 12 (1):239. doi:10.1186/1471-2164-12-239
- Xiao L, Yi B, Chen Y, Huang Z, Chen W, Ma C, Tu J, Fu T (2008) Molecular markers linked to *Bn;rf*: a recessive epistatic inhibitor gene of recessive genic male sterility in *Brassica napus* L. *Euphytica* 164:377-384.
- Xie Y, Dong F, Hong D, Wan L, Liu P, Yang G (2012) Exploiting comparative mapping among *Brassica* species to accelerate the physical delimitation of a genic male-sterile locus (*BnRf*) in *Brassica napus*. *Theor Appl Genet* 125:1-12.
- Xie YZ, Hong DF, Xu ZH, Liu PW, Yang GS (2008) Identification of AFLP markers linked to the epistatic suppressor gene of a recessive genic male sterility in rapeseed

- and conversion to SCAR markers. *Plant Breed* 127:145-149.
- Xu Z, Xie Y, Hong D (2009) Fine mapping of the epistatic suppressor gene (*esp*) of a recessive genic male sterility in rapeseed (*Brassica napus* L.). *Genome* 52:755-760.
- Yi B, Chen Y, Lei S, Tu J, Fu T (2006) Fine mapping of the recessive genic male-sterile gene (*Bnms1*) in *Brassica napus* L. *Theor Appl Genet* 113:643-650.
- Ying M, Dreyer F, Cai D, Jung C (2003) Molecular markers for genic male sterility in Chinese cabbage. *Euphytica* 132:227-234.
- Zhang H, Zhang S-j, Wu J, Li F, Zhang S-f, Wang X-w, Zhang X-m, Sun R-f (2010) Mapping of recessive genic male sterile restoring gene (*BrMf2*) in *Brassica rapa* L. ssp. *pekinensis*. *Scientia Agricultura Sinica* 43:993-999.
- Zhang L-G, Hao D-F (2001) Investigation on the sterility changeover of male sterility Line CMS7311 in heading Chinese cabbage. *Acta Botanica Sinica* 43:1123-1128.
- Zhou Y, Tang H, Cheng MP, Dankwa KO, Chen ZX, Li ZY, Gao S, Liu YX, Jiang QT, Lan XJ, Pu ZE, Wei YM, Zheng YL, Hickey LT, Wang JR (2017) Genome-wide association study for pre-harvest sprouting resistance in a large germplasm collection of Chinese wheat landraces. *Front Plant Sci* 8:401. doi:10.3389/fpls.2017.00401
- Zu F, Xia S, Dun X, Zhou Z, Zeng F, Yi B, Wen J, Ma C, Shen J, Tu J, Fu T (2010) Analysis of genetic model of a recessive genic male sterile line 7-7365AB in *Brassica napus* L. based on molecular markers. *Scientia Agricultura Sinica* 43:3067-3075.

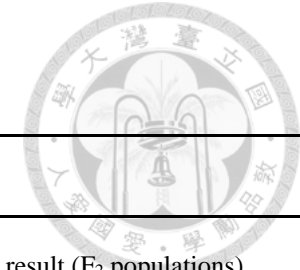
## 附 錄



附錄 1 兩種遺傳模式下不結球白菜參試品系可能的基因型

Inheritance model	Male sterile plant (NH80-A)	Restorer line (TA95, SG388)	Maintainer line (WH606)
1 gene with multiple allele	$BrMs^bMs^b$ $BrMs^bMs^c$	$BrMs^aMs^a$	$BrMs^cMs^c$
2 dominant gene	$BrMsMsrfrf$ $BrMsmsrfrf$	$BrMsMsRfRf$ $BrMsmsRfRf$ $BrmsmsRfRf$	$Brmsmsrfrf$

附錄 2 Progeny test A 之各種雜交組合後裔雄不稔性狀分離情形



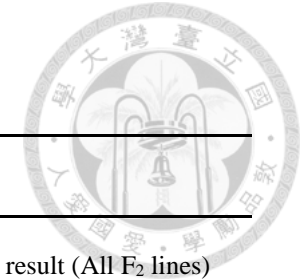
Inheritance model	No. of combination	Genotype of ms plant and restorer line	Progeny test A		
			F <sub>1</sub> genotype	F <sub>2</sub> phenotype ( <i>mf:ms</i> ratio)	Test result (F <sub>2</sub> populations)
1 gene with multiple alleles	1	<i>BrMs<sup>b</sup>Ms<sup>b</sup> × BrMs<sup>a</sup>Ms<sup>a</sup></i>	<i>BrMs<sup>a</sup>Ms<sup>b</sup></i>	segregate (3:1)	All subpopulations segregated
2 dominant genes	2	<i>BrMsMs rfrf × BrMsMs RfRf</i>	<i>BrMsMs RfRf</i>	segregate (3:1)	All subpopulations segregated
2 dominant genes	3	<i>BrMsMs rfrf × BrMsms RfRf</i>	<i>BrMsMs RfRf</i>	segregate (3:1)	All subpopulations segregated
			<i>BrMsms RfRf</i>	segregate (13:3)	
2 dominant genes	4	<i>BrMsMs rfrf × Brmsms RfRf</i>	<i>BrMsms RfRf</i>	segregate (13:3)	All subpopulations segregated
1 gene with multiple alleles	5	<i>BrMs<sup>b</sup>Ms<sup>c</sup> × BrMs<sup>a</sup>Ms<sup>a</sup></i>	<i>BrMs<sup>a</sup>Ms<sup>b</sup></i>	segregate (3:1)	Partial subpopulations segregated
			<i>BrMs<sup>a</sup>Ms<sup>c</sup></i>	no segregate (1:0)	
2 dominant genes	6	<i>BrMsms rfrf × BrMsMs RfRf</i>	<i>BrMsMs RfRf</i>	segregate (3:1)	All subpopulations segregated
			<i>BrMsms RfRf</i>	segregate (13:3)	
2 dominant genes	7	<i>BrMsms rfrf × BrMsms RfRf</i>	<i>BrMsMs RfRf</i>	segregate (3:1)	Partial subpopulations segregated
			<i>BrMsms RfRf</i>	segregate (13:3)	
			<i>Brmsms RfRf</i>	no segregate (1:0)	
2 dominant genes	8	<i>BrMsms rfrf × Brmsms RfRf</i>	<i>BrMsms RfRf</i>	segregate (13:3)	Partial subpopulations segregated
			<i>Brmsms RfRf</i>	no segregate (1:0)	

附錄 3 Progeny test B 之各種雜交組合後裔雄不稔性狀分離情形



Inheritance model	No. of combination	Genotype of ms plant and restorer line	Progeny test B			
			ms plant genotype from Test cross A	maintainer line genotype	progeny phenotype (mf:ms ratio)	Test result (All progeny lines)
1 gene with multiple alleles	1	$BrMs^bMs^b \times BrMs^aMs^a$	$BrMs^bMs^b$	$BrMs^cMs^c$	no segregate (0:1)	All lines ms
2 dominant genes	2	$BrMsMs rfrf \times BrMsMs RfRf$	$BrMsMs rfrf$	$Brmsms rfrf$	no segregate (0:1)	All lines ms
2 dominant genes	3	$BrMsMs rfrf \times BrMsms RfRf$	$BrMsMs rfrf$	$Brmsms rfrf$	no segregate (0:1)	Partial lines segregated
			$BrMsms rfrf$	$Brmsms rfrf$	segregate (1:1)	
2 dominant genes	4	$BrMsMs rfrf \times Brmsms RfRf$	$BrMsMs rfrf$	$Brmsms rfrf$	no segregate (0:1)	Partial lines segregated
			$BrMsms rfrf$	$Brmsms rfrf$	segregate (1:1)	
1 gene with multiple alleles	5	$BrMs^bMs^c \times BrMs^aMs^a$	$BrMs^bMs^b$	$BrMs^cMs^c$	no segregate (0:1)	All lines ms
2 dominant genes	6	$BrMsms rfrf \times BrMsMs RfRf$	$BrMsMs rfrf$	$Brmsms rfrf$	no segregate (0:1)	Partial lines segregated
			$BrMsms rfrf$	$Brmsms rfrf$	segregate (1:1)	
2 dominant genes	7	$BrMsms rfrf \times BrMsms RfRf$	$BrMsMs rfrf$	$Brmsms rfrf$	no segregate (0:1)	Partial lines segregated
			$BrMsms rfrf$	$Brmsms rfrf$	segregate (1:1)	
2 dominant genes	8	$BrMsms rfrf \times Brmsms RfRf$	$BrMsMs rfrf$	$Brmsms rfrf$	no segregate (0:1)	Partial lines segregated
			$BrMsms rfrf$	$Brmsms rfrf$	segregate (1:1)	

附錄 4 Progeny test C 之各種雜交組合後裔雄不稔性狀分離情形



Inheritance model	No. of combination	Genotype of maintainer line and restorer line	Progeny test C		
			F <sub>1</sub> genotype	F <sub>2</sub> phenotype ( <i>mf:ms</i> ratio)	Test result (All F <sub>2</sub> lines)
1 gene with multiple alleles	1	<i>BrMs<sup>c</sup>Ms<sup>c</sup> × BrMs<sup>a</sup>Ms<sup>a</sup></i>	<i>BrMs<sup>a</sup>Ms<sup>c</sup></i>	no segregate (1:0)	All lines mf
2 dominant genes	2	<i>Brmsms rfrf × BrMsMs RfRf</i>	<i>BrMsms Rfrf</i>	segregate (13:3)	All lines segregated
2 dominant genes	3	<i>Brmsms rfrf × BrMsms RfRf</i>	<i>BrMsms Rfrf</i> <i>Brmsms rfrf</i>	segregate (13:3) no segregate (1:0)	Partial lines segregated
2 dominant genes	4	<i>Brmsms rfrf × Brmsms RfRf</i>	<i>Brmsms Rfrf</i>	no segregate (1:0)	All lines mf
1 gene with multiple alleles	5	<i>BrMs<sup>c</sup>Ms<sup>c</sup> × BrMs<sup>a</sup>Ms<sup>a</sup></i>	<i>BrMs<sup>a</sup>Ms<sup>c</sup></i>	no segregate (1:0)	All lines mf
2 dominant genes	6	<i>Brmsms rfrf × BrMsMs RfRf</i>	<i>BrMsms Rfrf</i>	segregate (13:3)	All lines segregated
2 dominant genes	7	<i>Brmsms rfrf × BrMsms RfRf</i>	<i>BrMsms Rfrf</i> <i>Brmsms rfrf</i>	segregate (13:3) no segregate (1:0)	Partial lines segregated
2 dominant genes	8	<i>Brmsms rfrf × Brmsms RfRf</i>	<i>Brmsms Rfrf</i>	no segregate (1:0)	All lines mf





## 附錄 5 本研究部份內容所發表之報告

(<https://www.g3journal.org/content/10/4/1309>)



# Genetic Analysis and Fine Mapping of a Spontaneously Mutated Male Sterility Gene in *Brassica rapa* ssp. *chinensis*

Tzu-Kai Lin,<sup>\*†</sup> Ya-Ping Lin,<sup>‡</sup> and Shun-Fu Lin<sup>\*1</sup>

<sup>\*</sup>Department of Agronomy and <sup>‡</sup>Institute of Ecology and Evolutionary Biology, National Taiwan University, Taipei City, Taiwan, R.O.C. 10617, and <sup>†</sup>Crop Science Division, Taiwan Agricultural Research Institute, Taichung, Taiwan, R.O.C. 41362

**ABSTRACT** Male sterility has been widely used in hybrid seed production in Brassica, but not in *B. rapa* ssp. *chinensis*, and genetic models of male sterility for this subspecies are unclear. We discovered a spontaneous mutant in *B. rapa* ssp. *chinensis*. A series of progeny tests indicated that male sterility in *B. rapa* ssp. *chinensis* follows a three-allele model with *BrMs*<sup>a</sup>, *BrMs*<sup>b</sup>, and *BrMs*<sup>c</sup>. The male sterility locus has been mapped to chromosome A07 in BC<sub>1</sub> and F<sub>2</sub> populations through genotyping by sequencing. Fine mapping in a total of 1,590 F<sub>2</sub> plants narrowed the male sterility gene *BrMs* to a 400 kb region, with two SNP markers only 0.3 cM from the gene. Comparative gene mapping shows that the *Ms* gene in *B. rapa* ssp. *pekinensis* is different from the *BrMs* gene of *B. rapa* ssp. *chinensis*, despite that both genes are located on chromosome A07. Interestingly, the DNA sequence orthologous to a male sterile gene in *Brassica napus*, *BnRf*, is within 400 kb of the *BrMs* locus. The *BnRf* orthologs of *B. rapa* ssp. *chinensis* were sequenced, and one KASP marker (*BrMs\_indel*) was developed for genotyping based on a 14 bp indel at intron 4. Cosegregation of male sterility and *BrMs\_indel* genotypes in the F<sub>2</sub> population indicated that *BnRf* from *B. napus* and *BrMs* from *B. rapa* are likely to be orthologs. The *BrMs\_indel* marker developed in this study will be useful in marker-assisted selection for the male sterility trait.

**KEYWORDS**

Non-heading Chinese cabbage  
Genic male sterility  
Mutant  
Genotyping by sequencing (GBS)  
SNP  
Plant breeding

Male sterility is often used in plant breeding to facilitate the production of hybrid seeds. It can be classified into cytoplasmic male sterility (CMS), which relies on cytoplasmic organelle genes; and genic male sterility (GMS), which is determined in the nucleus. Although CMS was widely applied in plant breeding, it was associated with undesirable agricultural traits such as poor growth, decreased seed production and reduced disease resistance (Levings 1990). Thus, GMS has become an alternative approach to overcome the disadvantages of CMS. Many spontaneous GMS mutants have been discovered in Brassica, including in cauliflower (Nieuwhof 1961), cabbage (Fang *et al.* 1997),

non-heading Chinese cabbage (Cao and Li 1981; Ying *et al.* 2003), rapeseed (Pan *et al.* 1988), Chinese cabbage (Takahata *et al.* 1996; Feng *et al.* 2009) and turnip (Kenji *et al.* 2013).

The need to produce large quantities of hybrid seed limits the commercial application of GMS for two reasons. First, when a gene controlling GMS is recessive, *msms* for male sterility for example, half of the progenies in a male sterile line will restore fertility when male sterile plants are crossed to male fertile plants (*Msms*) during seed propagation of the male sterile line. Those male fertile plants in the male sterile line must be eliminated during hybrid seed production, increasing labor costs. Second, when GMS is controlled by a dominant gene (van der Meer 1987), male sterile plants have genotype *Msms*. The male sterile plant has no pollen so it cannot self-pollinate to obtain *Msms* inbred lines. Dominant male sterility is commercially feasible when genes controlling GMS are either multi-allelic (Song *et al.* 2006; Feng *et al.* 2009) or involve interaction of multiple genes (Li *et al.* 1985; Chen *et al.* 1998; Lu *et al.* 2004). Under these genetic models, dominant male sterile homozygotes are obtained through sib-mating within male sterile lines. Crossing dominant GMS plants with recessive temporary maintainer lines can produce all male sterile progenies as female parents for hybrid seed production.

Copyright © 2020 Lin *et al.*

doi: <https://doi.org/10.1534/g3.120.401091>

Manuscript received June 11, 2019; accepted for publication February 7, 2020; published Early Online February 11, 2020.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplemental material available at figshare: <https://doi.org/10.25387/g3.11663970>.

<sup>1</sup>Corresponding author: No.1, Sec. 4, Roosevelt Rd., Da'an Dist., Department of Agronomy, National Taiwan University, Taipei City, Taiwan, R.O.C. 10617. E-mail: shunfu@ntu.edu.tw

Numerous DNA markers closely linked with GMS genes have been developed for marker-assisted breeding and have contributed to positional cloning (Ying *et al.* 2003; Ke *et al.* 2004; Lu *et al.* 2004; Ke *et al.* 2005; Hong *et al.* 2006; Song *et al.* 2006; He *et al.* 2008; Hong *et al.* 2008; Feng *et al.* 2009; Xu *et al.* 2009). A total of five GMS genes were mapped on linkage groups N7, N16, N19 and A8 in *B. napus* (Yi *et al.* 2006; Huang *et al.* 2007; Lei *et al.* 2007; Xiao *et al.* 2008; Lu *et al.* 2013). Two GMS genes were mapped on linkage groups A8 and R7, respectively, in Chinese cabbage (*B. rapa* ssp. *pekinensis*), (Feng *et al.* 2009; Zhang *et al.* 2010), while only one recessive GMS gene was mapped on linkage group A2 in non-heading Chinese cabbage (*B. rapa* ssp. *chinensis*) (Li *et al.* 2016). Although many gene mapping studies have located genetic regions conferring GMS, few proved allelism of the GMS genes (Zu *et al.* 2010; Xie *et al.* 2012). Moreover, comparative mapping of GMS genes from various genetic materials remains difficult due to insufficient marker density and lack of common polymorphic markers with known physical positions.

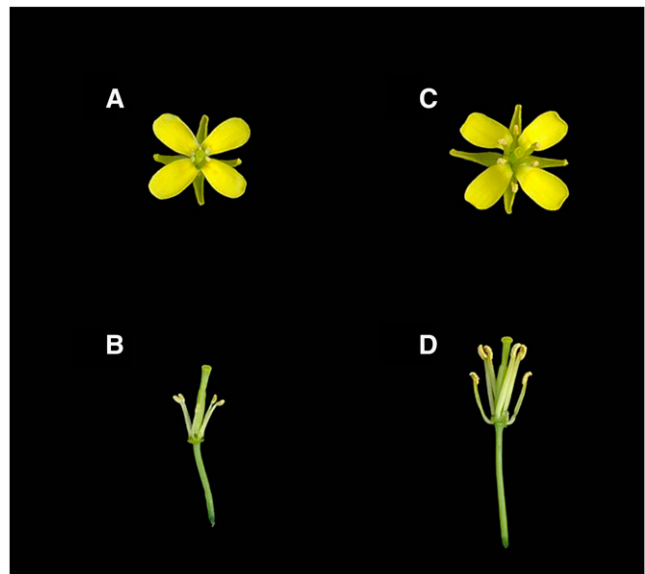
Construction of reduced-representation libraries (RRL) followed by next-generation sequencing is a convenient method to obtain large numbers of DNA markers to characterize genetic variation. Genotyping by sequencing (GBS) allows multiplexing large numbers of individuals at low cost in an efficient manner (Elshire *et al.* 2011). GBS is particularly useful for generating high density single nucleotide polymorphism (SNP)-based genetic maps with physical positions revealed by alignment to a reference genome. GBS-based QTL mapping has been applied to many crop species including but not limited to corn, wheat, barley, cabbage and melon (Elshire *et al.* 2011; Poland *et al.* 2012; Lee *et al.* 2015; Chang *et al.* 2017).

We found a male sterile individual originating from a spontaneous mutation in landraces of non-heading Chinese cabbage. The male sterile line derived from this mutant has a commercial advantage because it is able to produce crossbred progenies entirely with male sterility for hybrid seed production. This new GMS mutant provides an important genetic resource since few studies have reported the location of GMS and in fact, only one recessive GMS gene was discovered previously (Cao and Li 1981; Ying *et al.* 2003; Li *et al.* 2016). Moreover, comparing this new GMS mutation in *B. rapa* ssp. *chinensis* with close relatives is interesting, since *B. rapa* ssp. *pekinensis* might be derived from natural crosses between *B. rapa* ssp. *chinensis* and *B. rapa* ssp. *rapifera* (Li 1981); and *B. rapa* ( $2n = 20$ , genome AA) is one of the progenitors of *B. napus* ( $2n = 38$ , genome AACC). This study focuses on determining the inheritance of GMS, constructing genetic maps to locate the GMS locus using GBS, developing SNP markers closely linked to GMS for marker-assisted selection, and clarifying the evolution of GMS genes in the genus *Brassica*.

## MATERIALS AND METHODS

### Plant materials

A spontaneous mutant of GMS was discovered from landrace NH80 of non-heading Chinese cabbage (*B. rapa* ssp. *chinensis* var. *oleifera* Makino), which is used as green manure in Taiwan. NH80-AB was a GMS line developed by sib mating of male sterile and male fertile NH80 individuals for eight generations. A cross of male sterile (NH80-A) with male fertile (NH80-B) plants would result in half of the progenies being sterile. Morphologically, the male sterile plants display shrunken and pale anthers with no fertile pollen and shortened filaments, clearly distinguishable from the male fertile plants (Figure 1). The restorer line, TA95, and maintainer line, WH606, were *B. rapa* ssp. *chinensis* inbred lines respectively derived from germplasm collected in Thailand and Taiwan, and developed by at least eight generations of selfing.



**Figure 1** Floral organ morphology of male sterility: A and B are NH80-A (male sterile); C and D are NH80-B (male fertile).

### Genetic analysis of male sterility

Genetic analysis of spontaneous male sterility in this study used Liu's (1992) progeny test method with minor modifications. There are two genetic male sterility models considered: one is a 'multi-allelic' model determined by one gene with multiple alleles, the other is a 'two-gene' model determined by two dominant genes. Under the first model, the male sterility gene, *BrMs*, has a total of three alleles: *BrMs<sup>a</sup>*, *BrMs<sup>b</sup>* and *BrMs<sup>c</sup>*, with *BrMs<sup>b</sup>* conferring GMS and dominant to *BrMs<sup>c</sup>*. If this model is true, then *BrMs<sup>b</sup>Ms<sup>b</sup>* and *BrMs<sup>b</sup>Ms<sup>c</sup>* would be sterile while *BrMs<sup>a</sup>* could restore fertility by suppressing *BrMs<sup>b</sup>*. Under the second model, *BrMs* represents the GMS gene while *BrRf* is a different gene recovering male fertility. Individuals with the genotypes *BrMsMs rrf* and *BrMsms rrf* would be male sterile because *BrRf* will suppress *BrMs* to restore fertility.

Three single cross combinations, NH80-A × TA95, NH80-A × WH606, and TA95 × WH606, were completed by intercrossing single plants from male sterile line NH80-A, restorer TA95 and maintainer WH606 (Table 1). Progeny tests A, B and C were conducted to examine male sterility at flowering time. Progeny test A, including 64 F<sub>2</sub> sub-populations of NH80-A × TA95, was used to test segregation of GMS in at least 60 individuals for each F<sub>2</sub> subpopulation. Progeny test B was used to test the segregation of GMS in hybrids from crosses between 40 randomly chosen male sterile F<sub>2</sub> individuals from progeny test A and WH606, the maintainer line. At least 20 progeny per cross were investigated for male sterility. Progeny test C was used to test allelism of GMS by crossing the restorer (TA95) with the maintainer line (WH606), examining sixty individual plants from each of 25 F<sub>2</sub> sub-populations for segregation of male sterility. The genotypes of each line under the two proposed genetic models are listed in Table S1. A total of eight possible genotypic combinations by crossing male sterile individuals with restorer lines, and their theoretical segregation ratios, are shown in Tables S2-S4.

### Mapping populations

A total of 88 F<sub>2</sub> plants derived from NH80-A × TA95 and 186 BC<sub>1</sub> plants derived from NH80-A × WH606 were used for mapping of *BrMs<sup>a</sup>/BrMs<sup>b</sup>* and *BrMs<sup>b</sup>/BrMs<sup>c</sup>*, respectively. Plants were grown in

**Table 1 Sterility of progenies in different crosses between male sterile, restorer, and maintainer lines**

Cross	No. of male sterile plants	No. of male fertile plants
NH80-A × TA95	0	70
NH80-A × WH606	40	0
TA95 × WH606	0	50
TA95 ⊗ <sup>a</sup>	0	72
WH606 ⊗ <sup>a</sup>	0	70

<sup>a</sup> ⊗ indicates self-pollination.

individual pots. DNA was extracted from leaves of each 30-day old plant for constructing GBS libraries. Phenotypic data were collected at flowering for each plant.

### Genotyping by sequencing

GBS libraries were constructed based on Poland *et al.* (2012) with some modifications. Approximately 200 ng DNA of each individual was digested with restriction enzymes *Pst*I-HF and *Msp*I (New England Biolabs [NEB], Ipswich MA, US) at 37° for 4 h and then kept at 65° for 20 min. The digested DNA was ligated with the barcoded and reverse Y-adapters (Poland *et al.* 2012), in a total volume of 40 μL containing 200 ng of digested DNA, 20 nM barcoded adapter, 300 nM reverse Y-adapter, 200 U of T4 DNA ligase, 1 × NEB Buffer 4 (NEB, Ipswich, MA, USA), and 1 mM ribo-ATP at 22° for 2 h and then at 65° for 20 min. Two DNA libraries were established by mixing 15 μL of the above samples from each individual of F<sub>2</sub> and BC<sub>1</sub> populations, respectively. The DNA libraries were purified using AMPure XP beads at the ratio of 1:1.8 for DNA:beads (Beckman Coulter, Brea CA, USA); modified with Illumina primers (Poland *et al.* 2012); and amplified for 18 cycles at 95° for 30 s, at 62° for 25 s, and at 68° for 40 s. These two GBS libraries were sequenced with Illumina HiSeq2000 (100-bp, single-end; Illumina Inc., San Diego, CA, USA) in the Genome Research Center, National Yang-Ming University, Taiwan.

The reference genome *B. rapa* var. Chiifu40 version 1.5 (<http://brassicadb.org/brad/index.php>) was used for data analysis (Cheng *et al.* 2011). The raw reads of 100-bp single-end sequences were selected for high-quality reads (Q > 20), then sorted according to barcodes using CLC Genomics Workbench (version 8.0.2; CLC Bio, <http://www.clcbio.com>). Clean reads without barcode sequences were mapped onto the reference genome and saved in the .sam file format. Locus indexing, SNP calling, and genotyping of segregants were performed using the command 'ref\_map.pl' in Stacks v. 1.37 (<http://creskolab.uoregon.edu/stacks/>) (Catchen *et al.* 2011). Unique sequences of clean reads were defined as tags. Only tags with depth of more than five reads in each individual and missing in less than 20% of a population were used as DNA markers for genotyping. Files S1 and S2 contain physical positions of polymorphic tags and genotypes for each individual of F<sub>2</sub> and BC<sub>1</sub> populations, respectively.

### Linkage analysis of BrMs

Genetic maps of the F<sub>2</sub> and BC<sub>1</sub> populations were constructed with the R/qtl program (Broman *et al.* 2003; Broman 2010). Linkage groups were formed with pair-wise recombination of tag markers smaller than 0.3 and LOD scores greater than 7. The Kosambi function (Kosambi 1943) was used to calculate genetic distance. Closely linked tag markers (genetic distance smaller than 0.5 cM) were merged into the same bin, with the threshold being at least one recombination event between adjacent bins. A random tag marker within each bin was selected as

a bin marker. The genetic distance was recalculated with respect to the randomly selected bin markers. Male sterility was considered a morphological marker, and was added to the linkage analysis.

### Development of SNP and Indel markers linked to BrMs

Tags co-segregating with the *BrMs* gene were developed into SNP markers with Kompetitive Allele Specific PCR (KASP) assays (LGC Genomics, Teddington, UK). In addition, two SCAR markers, syau\_SCR01 and syau\_SCR04 (Feng *et al.* 2009), were converted into SNP markers. The two primer pairs of the SCAR markers were used to amplify DNA segments of NH80-A and TA95, and the amplicons were sequenced using an ABI Prism 3730 DNA sequencer (Applied Biosystems, Foster City, USA). BioEdit (version 7.2.0) (Hall 1999) was used for DNA sequence alignment. In addition, the DNA sequence of the *BnRf* gene (GenBank accessions No. KT818624 and No. KT818625) was referenced to design primer pairs for comparative sequencing of NH80-A and TA95 lines, and variation regions in the DNA sequences within the candidate gene were used to design Indel marker with Kompetitive Allele Specific PCR (KASP) assays.

A total of 1,590 F<sub>2</sub> plants derived from NH80-A × TA95 were genotyped with SNP markers for fine-mapping. The developed Indel KASP marker was used for allelism analysis related to *BnRf* and *BrMs* genes. DNA samples of each individual plant were isolated with Quick-Extract DNA extraction solution (Epicentre, Madison, WI, USA). Applied Biosystems Viia 7 Real-Time PCR System (Applied Biosystems, Foster City, USA), was used for fluorescence detection according to the manufacturer's protocol from KASP, and Viia 7 software (v. 1.2) was used for genotype calling.

### Data availability

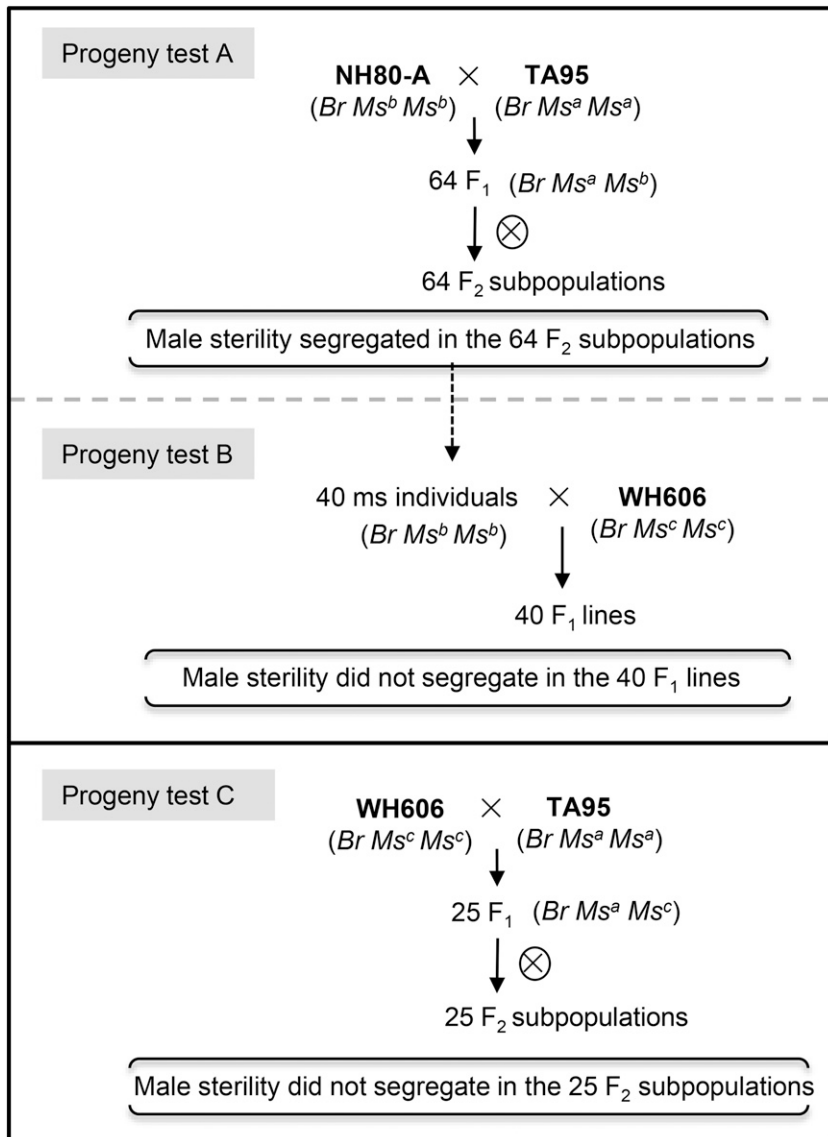
All polymorphic tags and genotypic data are available in Supplementary Files S1 and S2. DNA sequences of *BnRf* orthologs in NH80-A and TA95 are available in Supplementary Files S3. Linkage maps of two populations are available in Supplementary Figures S1 and S2. The predicted genotypes of two genetic models of GMS are listed in Supplementary Table S1, as well as segregation ratios of male sterility in three sets of progeny tests are available in Supplementary Tables S2, S3, and S4. Supplemental material available at figshare: <https://doi.org/10.25387/g3.11663970>.

## RESULTS

### Genetic model of GMS in *Brassica rapa* ssp. *chinensis*

Selfed progenies of TA95 and WH606, and F<sub>1</sub> progenies of NH80-A × TA95, TA95 × WH606 show male fertility, while F<sub>1</sub> progenies of NH80-A × WH606 are male sterile (Table 1). The GMS allele in NH80-A is homozygous and dominant to WH606. The GMS genotype of NH80-A should be *BrMs<sup>b</sup>Ms<sup>b</sup>* under the multi-allelic model described in the M&M, or *BrMsMs rfrf* under the two-gene model. The corresponding maintainer genotypes should be *BrMs<sup>c</sup>Ms<sup>c</sup>* under the multi-allelic model or *Brmsms rfrf* under the two-gene model, and the corresponding restorer genotypes should be *BrMs<sup>a</sup>Ms<sup>a</sup>* or *BrMsMs RfRf*, *BrMsms RfRf*, *Brmsms RfRf*, as shown in Table S1.

In progeny test A, segregation of male sterility was observed in a total of 64 F<sub>2</sub> subpopulations derived from NH80-A × TA95. In progeny test B, all plants in each of 40 F<sub>1</sub> lines were male sterile (Figure 2). These two progeny tests suggest that the possible genotypes of NH80-A and TA95 are *BrMs<sup>b</sup>Ms<sup>b</sup>* and *BrMs<sup>a</sup>Ms<sup>a</sup>* respectively, under the multi-allelic model or *BrMsMs rfrf* and *BrMsMs RfRf* under the two-gene model (Table S2, Table S3). In progeny test C, all F<sub>2</sub> plants derived from TA95 × WH606 were fertile (Figure 2), suggesting *BrMs<sup>a</sup>* and *BrMs<sup>c</sup>*



**Figure 2** Inheritance of genetic male sterility in three different progeny tests. The genotypes of male sterile line NH80-A, restorer line TA95, maintainer line WH606, and progenies are indicated in brackets under the multiple-allele model.

are allelic (Table S4). All three progeny tests suggest that the inheritance of GMS follows a multi-allelic model, with *BrMs<sup>a</sup>* dominant to *BrMs<sup>b</sup>*, and *BrMs<sup>b</sup>* dominant to *BrMs<sup>c</sup>* (*BrMs<sup>a</sup>* > *BrMs<sup>b</sup>* > *BrMs<sup>c</sup>*). Therefore, the best fit of the model to the observed data are that the genotype of NH80-A is *BrMs<sup>b</sup>Ms<sup>b</sup>*; the genotype of NH80-B is *BrMs<sup>a</sup>Ms<sup>b</sup>*; the genotype of restorer line TA95 is *BrMs<sup>a</sup>Ms<sup>a</sup>*, and the genotype of WH606 is *BrMs<sup>c</sup>Ms<sup>c</sup>*.

### Linkage analysis of BrMs

An F<sub>2</sub> population derived from NH80-A × TA95 was used to map the *BrMs* locus containing *BrMs<sup>a</sup>* and *BrMs<sup>b</sup>* alleles. A total of 25 Gb of raw data were acquired from the Illumina sequencing platform, and 245 million cleaned reads with an average length of 93.6 bp were obtained after quality trimming and mapped to the reference genome. Read mapping rates of the male sterile line NH80-A, restorer line TA95, and the F<sub>2</sub> populations are 81.7%, 41.9%, and 80.2%, respectively. A total of 33,632 tags were aligned to the reference genome with 5,284 being polymorphic between the parents. After eliminating markers with more than 20% missing data, we obtained a total of 2,243 tag

markers to construct the linkage map (Table 2). After assigning tags with the same genotypes into bins, the final genetic map comprises 569 recombination bins with 4,347 SNPs on ten linkage groups. The length of the genetic map is 1,011 cM, with the longest linkage group spanning 135.8 cM (chromosome A09), and the shortest spanning 75.1 cM (chromosome A10). The average distance between consecutive bins is 1.8 cM, ranging from 1.6 to 2.1 cM. The number of polymorphic tags per bin ranges from 3.5 to 5.0, with an average of 3.9.

The *BrMs* locus was mapped on chromosome A07 in the F<sub>2</sub> population of NH80-A × TA95 (Figure 3; Figure S1). The length of chromosome A07 is 82 cM, containing 51 bins. The bin “F2\_21293” is cosegregating with *BrMs<sup>a</sup>*, while the two flanking bins, “F2\_21158” and F2\_21316”, are 1.8 and 0.6 cM from *BrMs<sup>a</sup>*, respectively. The corresponding physical position of *BrMs<sup>a</sup>* is approximately 6.0 to 7.6 Mb (Table 3, Figure 3). Within this 1.6 Mb, there are a total of four tags (F2\_21230, F2\_21231, F2\_21236, and F2\_21271) cosegregating with the *BrMs<sup>a</sup>* gene in addition to F2\_21293, and the flanking bins, F2\_21158 and F2\_21316 contain 3 and 2 tags, respectively. The regions of the ten tags contain a total of 15 SNPs (Table 3). The *BrMs<sup>a</sup>* locus can

■ **Table 2** Characteristics of the linkage groups established in the F<sub>2</sub> and BC<sub>1</sub> populations for mapping *BrMs*

Characteristics	Chromosomes										Total
	A01	A02	A03	A04	A05	A06	A07	A08	A09	A10	
<b>F<sub>2</sub> population (H<sub>0</sub>: <i>BrMs<sup>a</sup>/BrMs<sup>b</sup></i>)</b>											
Length (cM)	107.8	98.2	132.6	93.4	107	102.8	82	76.5	135.8	75.1	<b>1,011</b>
No. of SNP	384	468	527	306	400	572	495	335	584	276	<b>4,347</b>
No. of polymorphic Tag	206	233	263	176	209	279	256	185	288	148	<b>2,243</b>
No. of Bin	58	60	68	44	54	62	51	49	82	41	<b>569</b>
average SNPs (SNPs/Tag)	1.9	2.0	2.0	1.7	1.9	2.1	1.9	1.8	2.0	1.9	<b>1.9</b>
average Tags (Tags/bin)	3.6	3.9	3.9	4.0	3.9	4.5	5.0	3.8	3.5	3.6	<b>3.9</b>
average interval (cM/bin)	1.9	1.6	2.0	2.1	2.0	1.7	1.6	1.6	1.7	1.8	<b>1.8</b>
<b>BC<sub>1</sub> population (H<sub>0</sub>: <i>BrMs<sup>b</sup>/BrMs<sup>c</sup></i>)</b>											
Length (cM)	108.4	116.5	135.3	81.2	99.6	108.8	99.7	79.9	142.1	79.8	<b>1,051</b>
No. of SNP	192	229	189	87	167	217	156	155	283	116	<b>1,791</b>
No. of polymorphic Tag	117	117	100	42	90	132	87	98	175	58	<b>1,016</b>
No. of Bin	55	59	55	32	38	57	43	44	77	33	<b>493</b>
average SNPs (SNPs/Tag)	1.6	2.0	1.9	2.1	1.9	1.6	1.8	1.6	1.6	2.0	<b>1.8</b>
average Tags (Tags/bin)	2.1	2.0	1.8	1.3	2.4	2.3	2.0	2.2	2.3	1.8	<b>2.1</b>
average interval (cM/bin)	2.0	2.0	2.5	2.5	2.6	1.9	2.3	1.8	1.8	2.4	<b>2.1</b>

be further narrowed to a 1.0 Mb interval within the 1.6 Mb region based on the physical locations of the 10 tags.

The BC<sub>1</sub> population derived from NH80-A × WH606 was also used to map the *BrMs* locus, containing *BrMs<sup>b</sup>* and *BrMs<sup>c</sup>* alleles. A total of 25 Gb of raw data were also acquired from the Illumina sequencing platform, and 250 million cleaned reads with an average length of 92.6 bp were obtained after quality trimming. After aligning the clean reads to the reference genome, we found that the mapping rate for the maintainer (WH606) and the BC<sub>1</sub> populations were 85.4%, and 84.4%, respectively. A total of 22,224 tags were obtained in the BC<sub>1</sub> population, fewer than those from the F<sub>2</sub> population (33,632). After filtering, 1,016 polymorphic tags were obtained including 1,791 SNPs which were combined into 493 recombination bins. A linkage map for the BC<sub>1</sub> populations comprises the 493 recombination bins spanning a total of 1,051 cM on ten chromosomes. The longest linkage group is chromosome A09, spanning 142.1 cM and the shortest is chromosome A10, spanning 79.8 cM (Table 2). The average distance between consecutive bins was 2.1 cM, while four pairs of bins have gaps larger than 10 cM (Figure S2). Bins contain an average of 2.1 polymorphic tags, ranging from 1.3 to 2.4 tags (Table 2).

In this NH80-A × WH606 BC<sub>1</sub> population, the *BrMs* locus was also mapped to chromosome A07 (Figure 3; Figure S2). *BrMs<sup>c</sup>* was cosegregating with a total of four tags, bin BC\_13924, BC\_13914, BC\_13917 and BC\_13923 and flanked by bins BC\_13879 and BC\_13971 at distances of 0.9 and 1.1 cM from the *BrMs* locus. The physical position of *BrMs<sup>c</sup>* is in the range from 6.2 to 7.6 Mb (Table 3, Figure 3). The fact that the physical positions of *BrMs<sup>a</sup>* and *BrMs<sup>c</sup>* overlap in the range of 6.6–7.6 Mb (Figure 3) and tags BC\_13914 and F2\_21230 were located at the same position, essentially the same tag, indicates that *BrMs<sup>a</sup>*, *BrMs<sup>b</sup>*, and *BrMs<sup>c</sup>* are allelic, in concordance with the results from the three progeny tests (Figure 2).

### Fine-mapping of *BrMs*

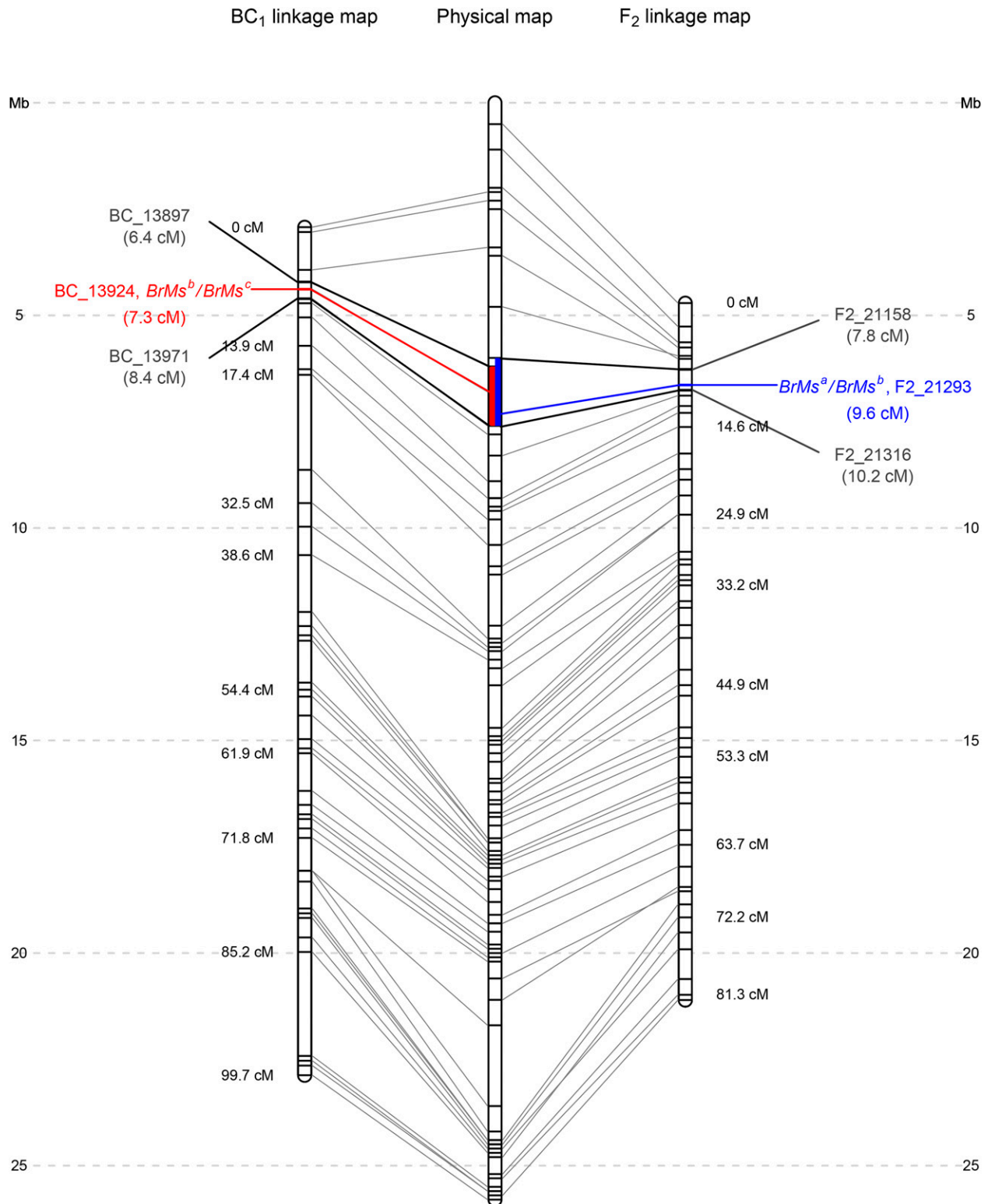
SNPs cosegregating with or flanking *BrMs* within the tags were designed into SNP markers for marker-assisted selection (Table 4). A total of 1,590 individuals in the NH80-A × TA95 F<sub>2</sub> population were used to fine map the *BrMs* locus. The flanking SNP markers, BrA7\_6632K and BrA7\_7625K, are 0.3 and 1.3 cM from the *BrMs* locus, respectively (Figure 4). Another SNP marker, BrA7\_7038K, obtained from the cosegregating tags, was on the same side as BrA7\_7625K, but only 0.3 cM from the *BrMs* locus. Therefore, the

*BrMs* locus was finely mapped to a 0.6 cM region, corresponding to a 400 kb physical region. Only three (0.19%) of 1,590 plants were recombinants, suggesting that these two SNP markers are appropriate for marker-assisted selection.

Two SCAR markers, syau\_SCR01 and syau\_SCR04, developed from Chinese cabbage (Feng *et al.* 2009) was used to compare GMS between *B. rapa* ssp. *chinensis* and *B. rapa* ssp. *pekinensis* in the same F<sub>2</sub> population. These two SCAR markers produced the same amplicons between Chinese cabbage and non-heading Chinese cabbage, with no length differences. Sequencing showed that the amplicon of the primer “syau\_scr04” contained a SNP between NH80-A and TA95. This SNP was designed into a SNP marker named “syau\_scr04\_SNP” (Table 4). Linkage analysis of the F<sub>2</sub> population suggests that syau\_scr04\_SNP is located on chromosome A07, 15.4 cM from the *BrMs* locus (Figure 4). This result is different from a previous study (Feng *et al.* 2009) that reported a distance of 2.5 cM between syau\_SCR04 and *Ms* in *B. rapa* ssp. *pekinensis*.

The GMS gene of *B. napus*, *BnRf*, was mapped to a 13.8 kb interval of LG N7, corresponding to a 17.8 kb interval of LG A7 which contains a total of three putative genes, *Bra014989*, *Bra014990*, and *Bra014991* (Xie *et al.* 2012). These three genes are located in an interval of 6,733,932–6,752,184 bp on chromosome A07, which is within the range containing the 400 kb *BrMs* locus from this study (Figure 4).

To understand the relationship between *BnRf* and *BrMs* genes, the orthologs of *BnRf* in lines NH80-A and TA95 were sequenced. The sequence length of *BnRf* orthologs in *B. napus* was longer than those of *B. rapa*. The orthologs in male sterile line 9012A and male fertile line RG206H of *B. napus* were 8,572 bp (KT818624) and 7,726 bp (KT818625), respectively (Deng *et al.* 2016). The length of orthologs in male sterile line NH80-A and in male fertile line TA95 of *B. rapa chinensis* were 7,562 bp and 7,411 bp, separately (Supplementary Files S3). Despite variation in sequence lengths, the four *BnRf* orthologs had the same DNA sequences in all 9 exons. The notable difference among the four orthologs was located at the promoter region. Compared to KT818624, the other orthologs had different lengths of deletion starting from position -1410 upstream of the translation start site (ATG). In addition to a deletion of 850 bp in *B. napus* male fertile line RG206H (KT818625) (Deng *et al.* 2016), 1,015 bp and 1,180 bp deletions were found in the orthologs of lines NH80-A and TA95, respectively (Figure 5). Besides, 4 insertions (2 bp and 3 bp at 5' UTRS; 14 bp at intron 4; 2 bp at 3' UTRS)



**Figure 3** Linkage maps of the male sterility gene in non-heading Chinese cabbage. The *BrMs<sup>b</sup>/BrMs<sup>c</sup>* locus mapped on chromosome A07 in the BC<sub>1</sub> population (left), and the *BrMs<sup>a</sup>/BrMs<sup>b</sup>* locus mapped on chromosome A07 in the F<sub>2</sub> population (right). The overlapping physical maps of *BrMs<sup>b</sup>/BrMs<sup>c</sup>* and *BrMs<sup>a</sup>/BrMs<sup>b</sup>* are indicated by red and blue bars.

and 2 SNP (at 3' UTRS) were observed at non-coding regions. Based on the 14 bp insertion at intron 4, a KASP assay was conducted to investigate the indel marker (called *BrMs\_indel*) for understanding

the relationship between *BnRf* and *BrMs* (Table 4). A total of 923 F<sub>2</sub> plants derived from NH80-A × TA95 were genotyped with *BrMs\_indel*. The marker completely cosegregated with male sterility,

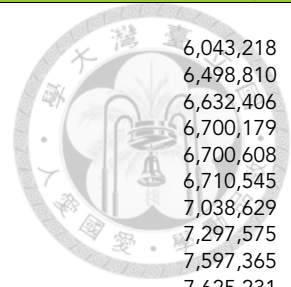
■ Table 3 GBS tags cosegregating with the *BrMs* gene

Bin <sup>a</sup>	Tags <sup>b</sup>	SNP numbers	Genetic map (cM)	Physical map (bp) <sup>c</sup>
F <sub>2</sub> population (H <sub>0</sub> : <i>BrMs</i> <sup>a</sup> / <i>BrMs</i> <sup>b</sup> )				
<b>F2_21158</b>	F2_21158	1	7.78	6,043,218
	F2_21208	1		6,498,810
	F2_21222	1		6,632,406
<b>F2_21293*</b>	F2_21230	1	9.63	6,700,179
	F2_21231	2		6,700,608
	F2_21236	3		6,710,545
	F2_21271	3		7,038,629
	F2_21293	1		7,297,575
<b>F2_21316</b>	F2_21316	1	10.24	7,597,365
	F2_21322	1		7,625,231
BC <sub>1</sub> population (H <sub>0</sub> : <i>BrMs</i> <sup>b</sup> / <i>BrMs</i> <sup>c</sup> )				
<b>BC_13879</b>	BC_13879	2	6.44	6,211,318
<b>BC_13924*</b>	BC_13914	1	7.30	6,700,179
	BC_13917	1		6,700,978
	BC_13923	1		6,736,962
	BC_13924	1		6,750,243
<b>BC_13971</b>	BC_13971	1	8.41	7,583,322

<sup>a</sup>Bold markers indicate bin markers. Bins segregated with *BrMs*<sup>a</sup> or *BrMs*<sup>c</sup> are indicated by \*.

<sup>b</sup>Tags in the same bins displayed the same segregation pattern.

<sup>c</sup>The physical map of the first nucleotide of each tag is indicated. *B. rapa* var. Chiifu40 (v1.5) was used for reference sequences, downloaded from the Brassica database (<http://brassicadb.org/brad/index.php>).



suggesting that *BrMs* and *BnRf* genes were orthologs differentiated by mutations.

## DISCUSSION

### A multiple allelic model of male sterility in *Brassica rapa* ssp. *chinensis*

This study first proved that the GMS in *B. rapa* ssp. *chinensis* fits a one gene multi-allelic model. A conventional method uses segregation ratios to test genetic models for GMS by conducting progeny tests in F<sub>2</sub> populations derived from male sterile lines crossed with restorer lines. The theoretical ratio of male sterile and male fertile individuals in a segregating population for the multi-allele model is 1:3 while for the two-gene model is 3:13. However, it is quite difficult to separate these two ratios and draw a firm conclusion. Our study, using three progeny tests, is advantageous because we only need to determine whether there is (or is not) segregation of male sterility in the progeny populations and not distinguish subtle differences in segregation ratios. In three progeny tests, the 'two dominant genes' model proposed by

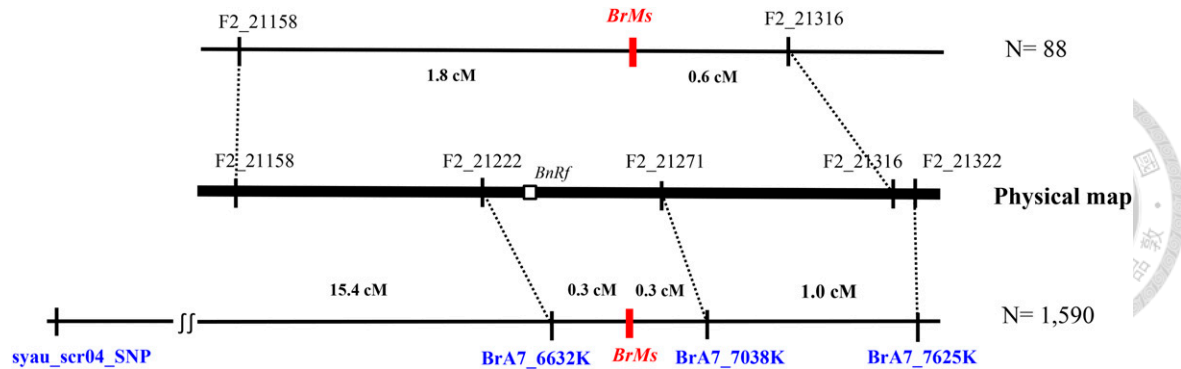
Liu (1992) assumed that the two dominant genes were located on different chromosomes and segregated independently. However, if the two genes were linked on the same chromosome, the results of the progeny test might be biased and thus lead to a different conclusion. In our study, we used different populations to map the GMS locus in each population conferring a different allele. The result suggests that *BrMs*<sup>a</sup>/*BrMs*<sup>b</sup> and *BrMs*<sup>b</sup>/*BrMs*<sup>c</sup> map to an interval from 6.0 Mb to 7.6 Mb on chromosome A07 (Figure 3) in two different populations (F<sub>2</sub> and BC<sub>1</sub>). Because the same tags are cosegregating with the GMS locus in both populations, it suggests that *BrMs*<sup>a</sup>, *BrMs*<sup>b</sup> and *BrMs*<sup>c</sup> alleles are located at the same locus.

Multiple allelism of GMS has also been validated in *Brassica rapa* ssp. *pekinensis* using two types of progeny tests and in *B. napus* using closely linked markers (Feng *et al.* 1996; Song *et al.* 2006; Liu *et al.* 2008; Xie *et al.* 2012). Our study, integrating the results of progeny tests, the same cosegregated tag in two linkage maps and the physical positions of *BrMs*, supports the multi-allelic genetic model for GMS in *B. rapa* ssp. *chinensis*. Collectively, results from *B. rapa* ssp. *pekinensis*, *B. napus* and our study, suggest that multi-allelic control of GMS in Brassica is a widespread phenomenon, consistent with the hypothesis proposed by

■ Table 4 Primer sequences of SNP and Indel markers for KASP assays

SNP marker	Tag	KASP Assay	Sequence
BrA7_6632K	F2_21222	Allele 1 Primer	5'-CACCAACAACGCATCATCGTCC-3'
		Allele 2 Primer	5'-CACCAACAACGCATCATCGTCT-3'
		Common Primer	5'-GTTTGACCTAGCTCGACGAAACCAA-3'
BrA7_7038K	F2_21271	Allele 1 Primer	5'-CTGAAGCTTTGTGCGATTTTGTAAAGTAATA-3'
		Allele 2 Primer	5'-CTGAAGCTTTGTGCGATTTTGTAAAGTAATG-3'
		Common Primer	5'-CCGGCAACGCTTACCTCTA-3'
BrA7_7625K	F2_21322	Allele 1 Primer	5'-AAGAAGTTTGATGAAGAGGTTTCTTCAATA-3'
		Allele 2 Primer	5'-GAAGTTTGATGAAGAGGTTTCTTCAATG-3'
		Common Primer	5'-CACCAAATTACCAACCTGTTTCATTTCTT-3'
syau_scr04_SNP	—	Allele 1 Primer	5'-TGAGTCTAAGATAAGCATGCATCATC-3'
		Allele 2 Primer	5'-GAGTCTAAGATAAGCATGCATCATT-3'
		Common Primer	5'-TTAACACAGGTTGCTAACAGGATATATCTT-3'
BrMs_indel	—	Allele 1 Primer	5'-CGATAAGATCCCTCTTAAGTCTTAAG-3'
		Allele 2 Primer	5'-CGATAAGATCCCTCTTAAGTCTTAAC-3'
		Common Primer	5'-CCAGTACAACAATTGTAGATACAGAGACAA-3'





**Figure 4** Fine mapping of the GMS gene. *BrMs* was fine mapped using 1,590 individuals in the F<sub>2</sub> population of NH80-A × TA95 and SNP markers. Loci on the linkage and physical maps are illustrated. The corresponding physical positions of *BnRf* on *B. napus* are indicated by an open box.

Lu *et al.* (2013). In practice, GMS by either model can produce 100% male sterile progenies for commercialization by crossing the male sterile lines to the maintainer lines. In addition, applying GMS from a multi-allelic model with 6 possible genotypes is easier than the two dominant gene model with 9 possible genotypes.

### Construction of high-density linkage maps with GBS

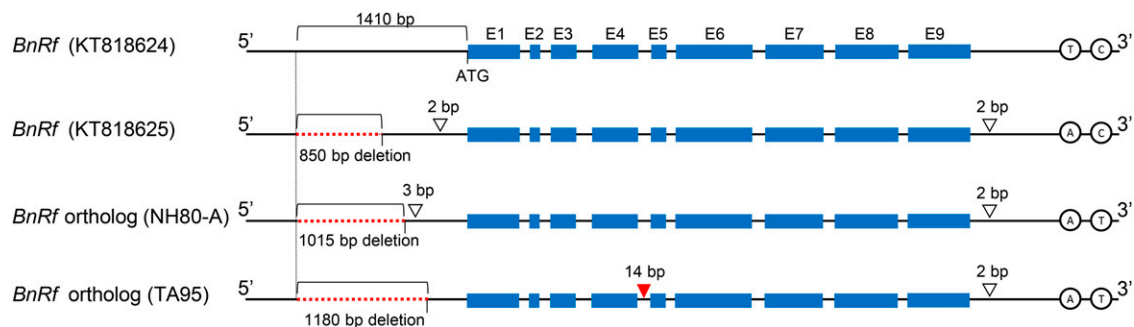
We constructed high-density linkage maps of the NH80-A × TA95 F<sub>2</sub> and NH80-A × WH606 BC<sub>1</sub> populations, respectively, using GBS (Table 2, Figures S1–S2). Approximately twice as many SNPs were discovered in the F<sub>2</sub> population as the BC<sub>1</sub> population, mainly due to higher genetic diversities between the parents. A total of 81.7% and 85.4% of the Illumina reads from the parents NH80-A and WH606, originating from Taiwan, were mapped to the reference genome, while only 41.5% of the reads from TA95, originating from Thailand, were mapped. This difference in mapping rate suggests that TA95 is more genetically distant from NH80-A and WH606, corresponding to the higher frequency of polymorphic tags in the F<sub>2</sub> population derived from NH80-A × TA95 (Table 2). In general, the distribution of bins in the F<sub>2</sub> population is more even than that in the BC<sub>1</sub> population. The two linkage maps display extensive colinearity and consistency (Figure 3).

Genetic maps of the F<sub>2</sub> and BC<sub>1</sub> populations are 1,101 and 1,051 cM in length, respectively, within the range of previously published linkage maps of *B. rapa* ranging from 970 cM to 1,230 cM (Suwabe *et al.* 2006; Choi *et al.* 2007; Kim *et al.* 2009; Wang *et al.* 2011; Chung *et al.* 2014; Huang *et al.* 2017). Many SNP markers from GBS are prone to genotypic errors; increased marker numbers challenge computational speed

and efficiency (van Os *et al.* 2005), which inflates total genetic distances (Hackett and Broadfoot 2003). Using the ‘bin’ method to combine closely linked markers can substantially increase computational efficiency and reduce genetic distance inflation caused by genotypic errors. This method of constructing linkage groups has been widely applied to many other crops (Sun *et al.* 2007; Ganai *et al.* 2011). Benefitting from the bin approach, the total genetic distance in our study is similar to those of maps constructed using PCR-based markers (Suwabe *et al.*, 2006; Choi *et al.*, 2007; Wang *et al.*, 2011). A similar approach was applied to construct high-density linkage maps in non-heading Chinese cabbage with restriction-site-associated DNA sequencing (RAD-seq) (Huang *et al.* 2017). The resulting genetic map contains a similar number of intervals as our F<sub>2</sub> map. However, our study contains more polymorphic tags, possibly due to using different restriction enzymes or different thresholds to define tags. The previous study used the restriction enzyme *EcoRI* while our study used *PstI* to construct GBS libraries. Flanking sequences of different restriction enzyme recognition sites may generate different sets of SNPs. Both linkage maps provide useful information for the *B. rapa* genome. Integrating these two linkage maps to obtain a relatively high-density map would facilitate investigation and applications of genomic analysis, gene mapping, and molecular breeding in *B. rapa* ssp. *chinensis*.

### Comparing the GMS mutation in *Brassica rapa* ssp. *chinensis* with those in other close relatives

The *BrMs* locus of *B. rapa* ssp. *chinensis* discovered in this study and the *Ms* locus of *B. rapa* ssp. *pekinensis* belong to the same chromosome, but



**Figure 5** The sequence alignment of *BnRf* orthologs. The accession numbers KT818624 and KT818625 in GenBank represent the *BnRf* genes in the male sterile line 9012A and male fertile line RG206H, respectively. E1 to E9 indicate the exons of *BnRf* while the black lines indicate introns or untranslated regions. Red dotted lines indicate the deletion regions in the promoter. The colorless inverted triangles show the insertions, and inverted triangle in red displays the region of *BrMs*\_indel marker. The circles at 3'UTR show the location of SNPs. All the labeled positions are in comparison to ATG on KT818624.

the two loci appear to be different. We developed a SNP marker (syau\_scr04\_SNP), equivalent to the SCAR marker syau\_scr04, closely linked to GMS in *B. rapa* ssp. *pekinensis* and found that it is 15.4 cM from the *BrMs* locus, while a previous study (Feng *et al.* 2009) reported a genetic distance of 2.5 cM between syau\_scr04 and the *Ms* locus. Although genetic distances can vary in different populations, the long genetic distance revealed by gene fine mapping in our study suggests that *BrMs* and *Ms* are likely two different loci. Similar results have been published in previous studies. Two independent GMS loci mapped to the same LG were also reported in *B. napus*. The *BnRf* locus was mapped to LG N7 in a *B. napus* population developed from accession 9012A (Xiao *et al.* 2008; Xie *et al.* 2012), and *Bnms1* was mapped to the same LG in a different population developed from accession S45A (Yi *et al.* 2006). Nevertheless, the GMS of 9012A and S45A were not allelic (Chen *et al.* 1998), indicating that at least 2 different GMS loci were on the same LG.

The *BnRf* gene from *B. napus* and *BrMs* gene from *B. rapa* are likely to be orthologs. In fact, chromosome N7 of *B. napus* and chromosome A7 of *B. rapa* are evolved from a common ancestral chromosome (<http://www.brassica.info/resource/maps/lg-assignments.php>); and the regions of *BnRf* on LG N7 of *B. napus* and LG A7 of *B. rapa* show high colinearity (Xie *et al.* 2012). In this study, the finding of the *BrMs\_indel* marker cosegregating with the *BrMs* locus has demonstrated that *BnRf* and *BrMs* are allelic. Since the *BnRf* gene and the *BrMs* gene are orthologs, it is very conceivable that the *BnRf* gene of the AACG genome of *B. napus* was derived from the AA genome of *B. rapa* ssp. *chinensis*.

Long sequence insertion or deletion in promoters of the four *BnRf* orthologs is a potentially important factor affecting gene function. In *B. napus*, reduction in transcription efficiency may result from an 850 bp insertion in the promoter region of *BnRf* gene in male sterile line RG206A (Deng *et al.* 2016). In this study, a 165 bp insertion or deletion was also observed in the corresponding regions of NH80-A and TA95 *BnRf* orthologs. We speculated that sequence variation of this region may change male sterile gene expression in Brassica families. However, to understand the key sequences conferring variation among the four *BnRf* orthologs that affect male sterility, a further study such as genetic transformation is needed.

### Development of BrMs markers for marker-assisted selection

Three SNPs and one KASP (*BrMs\_indel*) marker associated with male sterility of *B. rapa* ssp. *chinensis* were developed in this study. The *BrMs\_indel* marker designed on the gene is predominantly suggested for marker-assisted selection. However, if the *BrMs\_indel* is not available due to lack of polymorphism between breeding lines, SNP markers, BrA7\_6632K and BrA7\_7038K, respectively only 0.3 cM away from the *BrMs* locus are recommended (Figure 4). Either marker can assist in selecting male sterile plants in a segregating population. Using both SNP markers simultaneously will improve the selection accuracy by detecting single recombinants, only missing very rare double recombinants. These markers together with syau\_scr04\_SNP may facilitate selection for male sterile lines or individuals among subspecies of *B. rapa*.

### ACKNOWLEDGMENTS

We thank Dr. Tze-Tze Liu in the Genome Research Center of National Yang-Ming University, Taiwan, R.O.C. for providing advice on the massive parallel DNA sequencing work. This project was supported by the Council of Agriculture Executive Yuan, Taiwan, R.O.C. (Project Nos. 106AS-8.4.1-CI-C2).

### LITERATURE CITED

- Broman, K. W., 2010 Genetic map construction with R/qtl, pp. Technical Report # 214. University of Wisconsin-Madison, Department of Biostatistics & Medical Informatics.
- Broman, K. W., H. Wu, S. Sen, and G. A. Churchill, 2003 R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19: 889–890. <https://doi.org/10.1093/bioinformatics/btg112>
- Cao, S., and S. Li, 1981 Breeding of summer Chinese cabbage “Dwarf Hybrid No. 1” and male sterile AB lines. *Yuan Yi Xue Bao* 8: 35–42.
- Catchen, J. M., A. Amores, P. Hohenlohe, W. Cresko, and J. H. Postlethwait, 2011 *Stacks*: building and genotyping loci *De Novo* from short-read sequences. *G3 (Bethesda)* 1: 171–182.
- Chang, C.-W., Y.-H. Wang, and C.-W. Tung, 2017 Genome-wide single nucleotide polymorphism discovery and the construction of a high-density genetic map for melon (*Cucumis melo* L.) using genotyping-by-sequencing. *Front. Plant Sci.* 8: 125. <https://doi.org/10.3389/fpls.2017.00125>
- Chen, F., B. Hu, C. Li, Q. Li, W. Chen *et al.*, 1998 Genetic studies on GMS in *Brassica napus* L: I. Inheritance of recessive GMS line 9012A. *Zuo Wu Xue Bao* 24: 431–438.
- Cheng, F., S. Liu, J. Wu, L. Fang, S. Sun *et al.*, 2011 BRAD, the genetics and genomics database for Brassica plants. *BMC Plant Biol.* 11: 136. <https://doi.org/10.1186/1471-2229-11-136>
- Choi, S. R., G. R. Teakle, P. Plaha, J. H. Kim, C. J. Allender *et al.*, 2007 The reference genetic linkage map for the multinational *Brassica rapa* genome sequencing project. *Theor. Appl. Genet.* 115: 777–792. <https://doi.org/10.1007/s00122-007-0608-z>
- Chung, H., Y.-M. Jeong, J.-H. Mun, S.-S. Lee, W.-H. Chung *et al.*, 2014 Construction of a genetic map based on high-throughput SNP genotyping and genetic mapping of a TuMV resistance locus in *Brassica rapa*. *Mol. Genet. Genomics* 289: 149–160. <https://doi.org/10.1007/s00438-013-0798-9>
- Deng, Z., X. Li, Z. Wang, Y. Jiang, L. Wan *et al.*, 2016 Map-based cloning reveals the complex organization of the *BnRf* locus and leads to the identification of *BnRf(b)*, a male sterility gene, in *Brassica napus*. *Theor. Appl. Genet.* 129: 53–64. <https://doi.org/10.1007/s00122-015-2608-8>
- Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto *et al.*, 2011 A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6: e19379. <https://doi.org/10.1371/journal.pone.0019379>
- Fang, Z., P. Sun, Y. Liu, L. Yang, X. Wang *et al.*, 1997 A male sterile line with dominant gene (*Ms*) in cabbage (*Brassica oleracea* var. *capitata*) and its utilization for hybrid seed production. *Euphytica* 97: 265–268. <https://doi.org/10.1023/A:1003026523150>
- Feng, H., P. Wei, Z.-Y. Piao, Z.-Y. Liu, C.-Y. Li *et al.*, 2009 SSR and SCAR mapping of a multiple-allele male-sterile gene in Chinese cabbage (*Brassica rapa* L.). *Theor. Appl. Genet.* 119: 333–339. <https://doi.org/10.1007/s00122-009-1042-1>
- Feng, H., Y. Wei, S. Ji, G. Jin, J. Jin *et al.*, 1996 Multiple allele model for genic male sterility in Chinese cabbage, pp. 133–138 in *III International Symposium Diversification of Vegetable Crops* 467.
- Ganal, M. W., G. Durstewitz, A. Polley, A. Bérard, E. S. Buckler *et al.*, 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: e28334. <https://doi.org/10.1371/journal.pone.0028334>
- Hackett, C., and L. Broadfoot, 2003 Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. *Heredity* 90: 33–38. <https://doi.org/10.1038/sj.hdy.6800173>
- Hall, T. A., 1999 BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, pp. 95–98 in *Nucleic Acids Symposium Series*. [London]: Information Retrieval Ltd., c1979-c2000.
- He, J., L. Ke, D. Hong, Y. Xie, G. Wang *et al.*, 2008 Fine mapping of a recessive genic male sterility gene (*Bnms3*) in rapeseed (*Brassica napus*) with AFLP- and *Arabidopsis*-derived PCR markers. *Theor. Appl. Genet.* 117: 11–18. <https://doi.org/10.1007/s00122-008-0747-x>

- Hong, D., L. Wan, P. Liu, G. Yang, and Q. He, 2006 AFLP and SCAR markers linked to the suppressor gene (*Rf*) of a dominant genetic male sterility in rapeseed (*Brassica napus* L.). *Euphytica* 151: 401–409. <https://doi.org/10.1007/s10681-006-9162-z>
- Hong, D. F., J. Liu, G. S. Yang, and Q. B. He, 2008 Development and characterization of SCAR markers associated with a dominant genetic male sterility in rapeseed. *Plant Breed.* 127: 69–73.
- Huang, L., Y. Yang, F. Zhang, and J. Cao, 2017 A genome-wide SNP-based genetic map and QTL mapping for agronomic traits in Chinese cabbage. *Sci. Rep.* 7: 46305. <https://doi.org/10.1038/srep46305>
- Huang, Z., Y. Chen, B. Yi, L. Xiao, C. Ma *et al.*, 2007 Fine mapping of the recessive genetic male sterility gene (*Bnms3*) in *Brassica napus* L. *Theor. Appl. Genet.* 115: 113–118. <https://doi.org/10.1007/s00122-007-0547-8>
- Ke, L., Y. Sun, P. Liu, and G. Yang, 2004 Identification of AFLP fragments linked to one recessive genetic male sterility (RGMS) in rapeseed (*Brassica napus* L.) and conversion to SCAR markers for marker-aided selection. *Euphytica* 138: 163–168. <https://doi.org/10.1023/B:EUPH.0000046800.29308.b0>
- Ke, L. P., Y. Q. Sun, D. F. Hong, P. W. Liu, and G. S. Yang, 2005 Identification of AFLP markers linked to one recessive genetic male sterility gene in oilseed rape, *Brassica napus*. *Plant Breed.* 124: 367–370. <https://doi.org/10.1111/j.1439-0523.2005.01115.x>
- Kenji, W., S. Takashi, F. Eimi, K. Kenji, I. Daizo *et al.*, 2013 Genic male sterility in *Brassica rapa* ssp. *rapa* cv. 77B. *Journal of Agriculture Science*. Tokyo University of Agriculture 57: 287–292.
- Kim, H., S. R. Choi, J. Bae, C. P. Hong, S. Y. Lee *et al.*, 2009 Sequenced BAC anchored reference genetic map that reconciles the ten individual chromosomes of *Brassica rapa*. *BMC Genomics* 10: 432. <https://doi.org/10.1186/1471-2164-10-432>
- Kosambi, D. D., 1943 The estimation of map distances from recombination values. *Ann. Eugen.* 12: 172–175. <https://doi.org/10.1111/j.1469-1809.1943.tb02321.x>
- Lee, J., N. K. Izzah, B.-S. Choi, H. J. Joh, S.-C. Lee *et al.*, 2015 Genotyping-by-sequencing map permits identification of clubroot resistance QTLs and revision of the reference genome assembly in cabbage (*Brassica oleracea* L.). *DNA Res.* 23: 29–41.
- Lei, S., X. Yao, B. Yi, W. Chen, C. Ma *et al.*, 2007 Towards map-based cloning: fine mapping of a recessive genetic male-sterile gene (*BnMs2*) in *Brassica napus* L. and syntenic region identification based on the *Arabidopsis thaliana* genome sequences. *Theor. Appl. Genet.* 115: 643–651. <https://doi.org/10.1007/s00122-007-0594-1>
- Levings, III, C. S., 1990 The Texas cytoplasm of maize: cytoplasmic male sterility and disease susceptibility. *Science* 250: 942–947. <https://doi.org/10.1126/science.250.4983.942>
- Li, C. W., 1981 The origin, evolution, taxonomy and hybridization of Chinese cabbage. *Chinese Cabbage, Proceedings of the First International Symposium*: 3–11.
- Li, S. L., Y. X. Qian, and Z. H. Wu, 1985 Inheritance of genetic male sterility in *Brassica napus* and its application to commercial production. *Acta Agriculturae Shanghai* 1: 1–12.
- Li, X., A. Wang, F. Zu, Z. Hu, J. Lin *et al.*, 2016 Identification of a nuclear-recessive gene locus for male sterility on A2 chromosome using the Brassica 60 K SNP array in non-heading Chinese cabbage. *Genes Genomics* 38: 1151–1157. <https://doi.org/10.1007/s13258-016-0459-6>
- Liu, D., 1992 Discussions on the inheritance of fertility restoring for the dominant genetic male sterility in plants. *Hereditas* 14: 31–36.
- Liu, J., D. Hong, W. Lu, P. Liu, Q. He *et al.*, 2008 Genetic analysis and molecular mapping of gene associated with dominant genetic male sterility in rapeseed (*Brassica napus* L.). *Genes Genomics* 30: 523–532.
- Lu, G. Y., G. S. Yang, and T. D. Fu, 2004 Molecular mapping of a dominant genetic male sterility gene *Ms* in rapeseed (*Brassica napus*). *Plant Breed.* 123: 262–265. <https://doi.org/10.1111/j.1439-0523.2004.00957.x>
- Lu, W., J. Liu, Q. Xin, L. Wan, D. Hong *et al.*, 2013 A triallelic genetic male sterility locus in *Brassica napus*: an integrative strategy for its physical mapping and possible local chromosome evolution around it. *Ann. Bot.* 111: 305–315. <https://doi.org/10.1093/aob/mcs260>
- Nieuwhof, M., 1961 Male sterility in some cole crops. *Euphytica* 10: 351–356. <https://doi.org/10.1007/BF00039105>
- Pan, T., F. Zeng, S. Wu, and Y. Zhao, 1988 A study on the breeding and application GMS line of low eruci acid in rapeseed (*B. napus*). *Zhongguo Youliao Zuowu Xuebao* 3: 5–8.
- Poland, J., P. Brown, M. Sorrells, and J.-L. Jannink, 2012 Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS One* 7: e32253. <https://doi.org/10.1371/journal.pone.0032253>
- Song, L.-Q., T.-D. Fu, J.-X. Tu, C.-Z. Ma, and G.-S. Yang, 2006 Molecular validation of multiple allele inheritance for dominant genetic male sterility gene in *Brassica napus* L. *Theor. Appl. Genet.* 113: 55–62. <https://doi.org/10.1007/s00122-006-0271-9>
- Sun, Z., Z. Wang, J. Tu, J. Zhang, F. Yu *et al.*, 2007 An ultradense genetic recombination map for *Brassica napus*, consisting of 13551 SRAP markers. *Theor. Appl. Genet.* 114: 1305–1317. <https://doi.org/10.1007/s00122-006-0483-z>
- Suwabe, K., H. Tsukazaki, H. Iketani, K. Hatakeyama, M. Kondo *et al.*, 2006 Simple sequence repeat-based comparative genomics between *Brassica rapa* and *Arabidopsis thaliana*: the genetic origin of clubroot resistance. *Genetics* 173: 309–319. <https://doi.org/10.1534/genetics.104.038968>
- Takahata, Y., M. Nagasaka, H. Kondo, and N. Kaizuma, 1996 Genic male sterility in *Brassica campestris* L. *Acta Hort.* (407): 147–150 (ISHS). <https://doi.org/10.17660/ActaHortic.1996.407.17>
- van der Meer, Q. P., 1987 Chromosomal monogenic dominant male sterility in chinese cabbage (*Brassica rapa* subsp. *pekinensis* (Lour.) hanelt). *Euphytica* 36: 927–931. <https://doi.org/10.1007/BF00051877>
- Van Os, H., P. Stam, R. G. F. Visser, and H. J. van Eck, 2005 SMOOTH: a statistical method for successful removal of genotyping errors from high-density genetic linkage data. *Theor. Appl. Genet.* 112: 187–194. <https://doi.org/10.1007/s00122-005-0124-y>
- Wang, Y., S. Sun, B. Liu, H. Wang, J. Deng *et al.*, 2011 A sequence-based genetic linkage map as a reference for *Brassica rapa* pseudochromosome assembly. *BMC Genomics* 12: 239. <https://doi.org/10.1186/1471-2164-12-239>
- Xiao, L., B. Yi, Y. Chen, Z. Huang, W. Chen *et al.*, 2008 Molecular markers linked to *Bnyf1*: a recessive epistatic inhibitor gene of recessive genetic male sterility in *Brassica napus* L. *Euphytica* 164: 377–384. <https://doi.org/10.1007/s10681-008-9679-4>
- Xie, Y., F. Dong, D. Hong, L. Wan, P. Liu *et al.*, 2012 Exploiting comparative mapping among Brassica species to accelerate the physical delimitation of a genetic male-sterile locus (BnRf) in *Brassica napus*. *Theor. Appl. Genet.* 125: 211–222. <https://doi.org/10.1007/s00122-012-1826-6>
- Xu, Z., Y. Xie, and D. Hong, 2009 Fine mapping of the epistatic suppressor gene (*esp*) of a recessive genetic male sterility in rapeseed (*Brassica napus* L.). *Genome* 52: 755–760. <https://doi.org/10.1139/G09-049>
- Yi, B., Y. Chen, S. Lei, J. Tu, and T. Fu, 2006 Fine mapping of the recessive genetic male-sterile gene (*Bnms1*) in *Brassica napus* L. *Theor. Appl. Genet.* 113: 643–650. <https://doi.org/10.1007/s00122-006-0328-9>
- Ying, M., F. Dreyer, D. Cai, and C. Jung, 2003 Molecular markers for genetic male sterility in Chinese cabbage. *Euphytica* 132: 227–234. <https://doi.org/10.1023/A:1024642717416>
- Zhang, H., S.-J. Zhang, J. Wu, F. Li, S.-F. Zhang *et al.*, 2010 Mapping of recessive genetic male sterile restoring gene (BrMf2) in *Brassica rapa* L. ssp. *pekinensis*. *Zhongguo Nong Ye Ke Xue* 43: 993–999.
- Zu, F., S. Xia, X. Dun, Z. Zhou, F. Zeng *et al.*, 2010 Analysis of genetic model of a recessive genetic male sterile line 7-7365AB in *Brassica napus* L. based on molecular markers. *Zhongguo Nong Ye Ke Xue* 43: 3067–3075.

Communicating editor: S. Pearce