A THESIS FOR THE DEGREE OF MASTER OF SCIENCE

Genetic Mapping of the Unstable *Restorer-of-fertility* Gene in Sweet Pepper (*Capsicum annuum* L.)

단고추 불안정 웅성불임회복 유전자의 유전자 지도 작성

FEBRUARY, 2019

MOO-CHAN KANG

MAJOR IN HORTICULTURAL SCIENCE AND BIOTECHNOLOGY
DEPARTMENT OF PLANT SCIENCE

THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

Genetic Mapping of the Unstable *Restorer-of-fertility* Gene in Sweet Pepper (*Capsicum annuum* L.)

UNDER THE DIRECTION OF DR. BYOUNG-CHEORL KANG
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF SEOUL NATIONAL UNIVERSITY

BY
MOO-CHAN KANG

MAJOR IN HORTICULTURAL SCIENCE AND BIOTECHNOLOGY
DEPARTMENT OF PLANT SCIENCE
THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

FEBRUARY, 2019

APPROVED AS A QUALIFIED DISSERTATION OF
MOO-CHAN KANG
FOR THE DEGREE OF MASTER OF SCIENCE
BY THE COMMITTEE MEMBERS

CHAIRMAN _____
Doil Choi, Ph.D.

VICE-CHAIRMAN _____
Byoung-Cheorl Kang, Ph.D.

MEMBER _____
Jin Hoe Huh, Ph.D.

Genetic Mapping of the Unstable *Restorer-of-fertility* Gene in Sweet Pepper (*Capsicum annuum* L.)

MOO-CHAN KANG

DEPARTMENT OF PLANT SCIENCE
THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

ABSTRACT

Cytoplasmic-genic male sterility (CGMS) has been used for F_1 hybrid seed production in several crop plants. The CGMS system is determined by interaction between cytoplasm male sterility (CMS) associated mitochondrial gene and the *Restorer-of-fertility (Rf)* gene in nucleus. In the previous studies, several *Rf*-linked markers and candidate genes were identified. However, the unstable *Rf* genes that are affected by environmental conditions has rarely been studied. In this study, we investigated on the male fertility phenotype of the unstable *Rf* gene and developed markers linked to the unstable *Rf* gene. The male fertility phenotype was inherited as a 3:1 ratio in a population derived from a cross between MSGR-A with stable sterility (*rf*/*rf*) and SPR03 with unstable fertility (Rf^u/Rf^u) indicating that the unstable *Rf* gene is controlled by a single dominant gene in the permissive

temperature. Genetic mapping revealed that 214.14MB-CAPS is co-segregation with the unstable *Rf* locus. Candidate genes were predicted using the sequences within the delimited region between 13T7-SCAR and G16-SCAR. A total of 37 and 80 predicted genes were detected in the target region from Zunla-1 and CM334 (UCD 10X), respectively. Predicted genes related to unstable *Rf* couldn't be found in Zunla-1. However, In CM334 (UCD 10X), pentaricopeptide (PPR) gene, protein gamete expressed 2 (GEX 2) and peroxidase were detected in the target region. 214.14MB-CAPS which is co-segregated with unstable *Rf* showed high accuracy of genotype (91.6%) in unstable fertile lines. Application of marker co-segregated with unstable *Rf* can be broadly use unstable male fertile lines. However, stable male sterile could not be distinguished. Phylogenetic tree based on the genotype of several markers showed that stable or unstable lines were prone to be cluster together.

Keywords: Cytoplasmic genetic male sterility (CGMS), Unstable restorer-of-fertility, Genetic mapping, Molecular marker

Student number: 2017-29073

CONTENTS

RESULTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BLAST	Basic local alignment search tool
CAPS	Cleaved amplified polymorphic sequence
CGMS	Cytoplasmic genetic male sterility
CMS	Cytoplasmic male sterility
cM	centi Morgan (the unit of genetic distance)
CTAB	Cetyl trimethylammonium bromide
LOD	Logarithm of the odds
MAS	Marker-assisted selection
PCR	Polymearase chain reaction
PPR	Pentatricopeptide repeat
Rf	Restorer of fertility
SCAR	Sequence characterized amplified region
SNP	Single nucleotide polymorphism

INTRODUCTION

F_1 hybrid cultivars have contributed significantly to crop production due to increased yield and plant vigor. For hybrid seed production, various methods are being used including physical and chemical, and genetic control. Although manual emasculation is the most common physical method, it is labor intensive and expensive. Chemical method can be efficient but F_1 seeds can be affected by toxicity of pollen killing substances and it is difficult to apply in the field (Colombo and Galmarini 2017). Cytoplasmic male sterility (CMS) is a maternally inherited trait conferred by the mitochondrial genes that cause failure to produce functional pollens (Chen et al., 2017). Mitochondrial genomes have drastically diverged in terms of molecular form, genome size, gene content and arrangement, and organization of intergenic regions (Kubo et al., 2011). Comparative analysis between the mitochondrial genome of male fertile and sterile line of hot pepper (*Capsicum annuum* L.) revealed that DNA region of *Orf507* and *Ψatp6-2* is highly rearranged by homologous recombinations (Jo et al., 2014). In CGMS (Cytoplasmic-Genic Male Sterility) system, the *Restorer-of-fertility* (*Rf*) gene suppresses the CMS gene function. The cloning of *Rf* genes have been tried in several crops. *Rf* genes have been identified in maize (Cui et al., 1996), rice (Wang et al., 2006; Fujii and Toriyama 2009; Itabash et al., 2011; Tang et al., 2014; Huang et al., 2015), sorghum (Jordan et al., 2010), radish (Koizuka et al., 2003), rapeseed

(Uyttewaal et al., 2008; Liu et al., 2016), sugar beet (Kitazaki et al., 2015). More than half of the cloned *Rf* genes encode pentatricopeptide repeat (PPR)-containing proteins. Besides the *Rf* gene encoding PPR protein, four *Rf* genes encode non-PPR proteins. *Rf2* in maize encodes aldehyde dehydrogenase (Cui et al., 1996), *Rf17* in rice encodes acyl-carrier protein synthase (Fujii and Toriyama, 2009) and *Rf2* in rice encodes a glycine-rich domain (Itabashi et al., 2011).

PPR proteins constitute one of the largest protein families in land plants, with more than 400 members in most species (Barkan and Small, 2014). PPR protein is mainly composed of a number of PPR motif consisted with a degenerate 35 amino acid motif repeated in tandem (Manna 2015). Combination of 5^{th} and 35^{th} amino acid of respective PPR motif can determine the specificity of one nucleotide from RNA molecule (Shen et al., 2016). Functions of PPR proteins include RNA editing, splicing, cleavage and degradation in plant. In CMS-BT rice, *Rf1a* and *Rf1b* which is the PPR type *Rf* gene cleaves and degrades *atp6-orf79* dicistronic transcripts by binding RNA molecule (Wang et al., 2006). The PPR type *Rf* gene is located closely each other in a cluster form, unlike other PPR genes. The *Rf4* locus is located in a typical PPR cluster region where *Rf1a* and *Rf1b* are also present (Tang et al., 2014).

In the CGMS system, the stability of *Rf* gene in various environments is critical. In CMS-S maize, *Rf9* shows consistent restorer ability below 28℃, however, when daily high maximum temperature is higher than 30℃, male fertility

changes to male sterility (Gabay-Laughnan et al., 2009). In rice CMS-BT, *Rf5* or *Rf6* exhibits normal seed setting, but the heterozygote of *Rf6* shows more stable seed setting rates than those of heterozygote of *Rf5* under the heat-stress conditions (Zhang et al., 2017). In pepper, the thermo-sensitive cytoplasmic male sterility (TCMS) is sterile at temperatures above 15℃, but becomes fertile when the night temperature drops below 13℃ (Kim et al., 2013). Lee (2000) suggested that beside the major restorer fertile gene, there is unstable dominant *Rf* (*St*$^{\mu}$) locus that affects male fertility. Lack of stable restorer lines in sweet pepper has been an obstacle for using the CGMS system. To overcome this, the stable *Rf* gene in hot pepper was introduced successfully to sweet pepper through marker-assisted backcrossing (Lin et al., 2015).

Although efforts to identify the *Rf* gene in pepper have led to the development of several molecular markers, the identity of *Rf* has not been confirmed. Recently, Jo et al., (2016) delimited genomic region of *Rf* and identified potential candidate genes (*CDS1*, *CDS2*, *CDS3* and *CaPPR6*). Among them, *CaPPR6* was suggested as a candidate gene (Jo et al., 2016). The Co1mod1-CAPS, 10 kb away from *CaPPR6*, showed high accuracy in predicting *Rf* phenotypes when applied to the breeding lines. *CaPPR6* was expressed only in Bukang C (*Rf*/*Rf*), however, this gene was expressed in MSGR-A (*rf*/*rf*) but not in SPR03 (*Rf*$^{\mu}$/ *Rf*$^{\mu}$) which has unstable male fertility (Ha 2017). In addition, SPR03 (*Rf*$^{\mu}$/ *Rf*$^{\mu}$) has the same genotype as A line (*rf*/*rf*) for *CaPPR6* and whereas male sterile MSGR-A has *Rf*/*Rf* (Ha 2017).

In this study, we investigated on the male fertility phenotype of the unstable *Rf* gene and mapped the unstable *Rf* gene. We applied previously and newly developed markers to the unstable *Rf* segregating population between MSGR-A and SPR03 to confirm and the position of the unstable *Rf* gene.

LITERATURE REVIEW

Cytoplasmic male sterility

CMS refer to inability to produce functional anthers and pollens. CMS circumvents the need for removal of anthers, thereby encouraging hybrid technology to generate dramatically superior F_1 progenies that exhibit significant advantage over its parents and existing popular cultivars in terms of yield, stress tolerance, adaptability (Bohra et al., 2016). CMS genes regulate the phenotype of male sterility by diverse mechanisms. Emerging research on CMS systems of diverse crops supports four models for the mechanisms that cause CMS : the cytotoxicity model, the energy deficiency model, the aberrant programmed cell death (PCD) model, and the retrograde regulation model (Chen and Liu 2014). CMS protein can directly cause influence by the cytotoxicity model. ORF138 protein responsible for the Ogura CMS strongly inhibit bacterial growth, but not by inhibition respiration (Duroc et al., 2005). In *Brassica juncea, orf288* for CMS-Hau is transcribed with *atp6* gene. The growth of *Escherichia coli* was significantly repressed in the presence of ORF288, which indicated that this protein is toxic to the *E. coli* host cells (Jing et al., 2012). An electron transport chain (ETC) in mitochondria creates a proton that drives the producing of adenosine triphosphate (ATP). CMS genes induce mitochondria fail to synthesis ATP for development of anther. OrfH79 was identified as CMS-HL (Hong Lian). Through bacterial two-

hybrid library screening, P61, a subunit of the ETC complex III, was selected as a candidate that interacts with ORFH79 (Wang et al., 2013). The spatiotemporal-specific accumulation of cytotoxic CMS proteins within the tapetum, the innermost somatic layer of anther wall or microspores during microspore development, frequently causes altered mitochondrial function such as ATP synthesis and abnormal PCD in the sporophytic or gametophytic cells, leading to male sterility (Touzet and Meyer 2014). In rice of CMS-WA lines, WA352 accumulates preferentially in the anther tapetum, thereby inhibiting COX11 function in peroxide metabolism and triggering premature tapetal PCD and consequent pollen abortion (Luo et al., 2013). *CW-orf307* was identified as the candidate most likely responsible for the CW-CMS event (Fujii et al., 2010). The CMS-CW cytoplasm upregulates, through one or more unknown retrograde signals, the expression of *rf17*, whereas *Rf17* is not upregulated by the CMS cytoplasm, owing to the presence of a mutation in the *Rf17* promoter regulatory region (Chen and Liu., 2014)

Restorer-of-fertility

The plant trait CMS is associated with a pollen sterility phenotype that can be suppressed or counteracted by nuclear genes known as *Rf* genes (Chase CD., 2007). Until now 14 *Rf* genes have been isolated from nine crop plants : *Rf2* (Maize) (Cui et al., 1996), *Rf1a, Rf1b (Rf5)* (Rice) (Wang et al., 2006), *Rf4* (Rice) (Tang et

al., 2014), *Rf2* (Rice) (Itabash et al., 2011), *Rf17* (Fujii and Toriyama et al., 2009), *Rf1, Rf2* (Sorghum) (Jordan et al., 2010), *Rfk1* (Radish) (Koizka et al., 2003), *Rf0* (Rapeseed) (Uyttewaal et al., 2008), *Rfp* (Rapeseed) (Liu et al., 2016), *BrRfp1* (Chinese cabbage) (Zhang et al., 2017). About half of the identified *Rf* genes encode pentatricopeptide repeat (PPR)-containing proteins (Kim and Zhang, 2018). *Rf1a, Rf1b* (*Rf5*), *Rf6, Rf4, Rf1, Rf2, Rfk1, Rf0, Rfp, BrRfp1* encode the PPR protein. Pentatricopeptide repeat (PPR) proteins are a large family of modular RNA-binding proteins which mediate several aspects of gene expression primarily in organelles but also in the nucleus (Manna 2015). Restoring PPR (Rf-PPR) proteins involve various mechanisms, including RNA cleavage, RNA destabilization, or translation inhibition (Dahan and Mireau, 2013). RF1A and RF1B are both targeted to mitochondria and can restore male fertility by blocking ORF79 production via endonucleolytic cleavage (RF1A) or degradation (RF1B) of dicistronic B-atp6/orf79 mRNA (Wang et al., 2006). *Rf0* (also called *Rfk1*) encodes PPR-B (or ORF687) that binds to and suppresses the translation of orf138 mRNA in CMS-Ogu of Brassica and CMS-Kos of radish (Kim and Zhang, 2018). Rf-PPR proteins interact with non-PPR protein to restore male sterility. GRP162 was found to physically interact with RF5 and to bind to *atp6-orfH79* via an RNA recognition motif (Hu et al., 2012). RF6 with hexokinase 6 function together in mitochondria to promote the processing of the aberrant CMS-associated transcript *atp6-orfH79*, thereby restoring fertility of HL-CMS rice (Huang et al., 2015). Although the majority of cloned *Rf* gene encode PPR proteins, four *Rf* genes encode non-PPR

protein. *Rf2* which restores CMS-T in maize encodes an aldehyde dehydrogenase (Cui et al., 1996). *Rf2* which restores CMS-Lead rice (LD) in rice encodes Non-PPR protein containing a glycine-rich domain (Itabash et al., 2011). *Rf17* which restores CMS-Chinese wild rice type (CW) in rice encodes Acyl-carrier protein synthase (Fujii and Toriyama, 2009). *Rf1* (bvORF20) which restores CMS-Owen in sugar beet encodes peptidase-like protein (Kitazaki et al., 2015).

Markers linked *Restorer-of-fertility* in pepper

Markers linked *Rf* gene was developed by genetic mapping in cultivars. CMS in pepper was discovered by Peterson (1958) in an Indian line, PI164835. In pepper CMS line, Two candidate CMS gene were identified as *Orf507* and *Ψatp6-2*. On the other hand, markers linked *Rf* gene has been developed in pepper. $OP13_{1400}$ of two random amplified polymorphic DNA (RAPD) markers is a tightly linked marker with a genetic distance of 0.37 cM (Zhang et al., 2000). CRF-SCAR (5cM) was developed from a Sequence tagged site (STS) marker (Gulyas et al., 2006). (Kim et al., 2006) detected eight AFLP markers by using bulked segregant analysis (BSA). Only AFRF8-CAPS (1.8cM) was successfully converted to a cleaved amplified polymorphic sequence (CAPS) marker (Kim et al., 2006). To identify partial restoration (*pr*) locus, PR-CAPS (1.6cM) marker was developed by using BSA and AFLP method (Lee et al., 2008b). For useful marker assisted selection in commercial hybrid pepper breeding, OPP13-CAPS (1,180bp) and PR-CAPS (640

bp) marker in 91 Korean inbred lines were clearly divided into three haplotypes by analysis of marker internal sequence (Lee et al., 2008a). To overcome lack of agreement between marker genotype and phenotype, BAC13T7-SCAR and BAC17T7-HRM was developed from the end sequence of a tomato BAC clone containing three genes which are homologous to petunia *Rf* gene (Jo et al., 2010). *Rf* locus was delimited to be within 821kb of sequence and then *CaPPR6* was identified as strong candidate pepper *Rf* gene (Jo et al., 2016).

MATERIALS AND METHODS

Plant materials and evaluation of the male fertility phenotype

Two hot pepper breeding and two sweet pepper breeding lines were used to identify molecular marker polymorphism (Table 1). A population F_4BC_5 plants derived from a cross between SPR03 (*C. annuum*) and MSGR-A (*C. annuum*) by marker assisted back crossing (MAB) (Figure. 1). The population was produced for introducing stable male sterility and useful trait in CGMS. MSGR-A is stable male sterility line. Stable male sterility from MSGR-A was introduced to the population as donor parent. SPR03 is unstable fertile line. SPR03 was use for recurrent line to introduce the traits. PR-CAPS was used for MAB. From BC_4F_4 to BC_4F_5, we used PPR6_3'UTR-SCAR to select heterozygous (Rf^u/rf). F_4BC_5 individuals were grown in the greenhouse of Seoul National University farm (Suwon, Korea) (Figure. 1a). The male fertility phenotype was observed at least four times by observing pollen production on anthers. Optical microscopy was also used to confirm presence of viable pollens in the plants. Pollen was collected from dehiscent flowers of plants and stained with FDA (fluorescein diacetate). After a 2 min incubation at room temperature, pollen viability was examined under a fluorescence microscope. Pollen from at least five flowers was stained for each line, and this was replicated three times. Leaf samples of 12 stable male sterile lines and 11

unstable male fertile lines given by Eco-Seed (Gimje, South Korea) were used to test a broad marker application.

Genomic DNA extraction

Young leaves from each of plant were used for DNA extraction. The cetyl trimethylammonium bromide (CTAB) method was used for DNA extraction (Park et al., 2009). DNA samples were measured for concentration and purity using Takes3 (BioTek, Winooski, USA). Extracted DNA samples were diluted to 50ng/μL for molecular marker analysis.

PCR conditions for molecular marker analysis

PCR reaction was conducted in a total volume of 20ul with 50ng DNA template, 2.5 mM dNTP mix, 2 ul 10X Hi-fi DNA polymerase buffer, 10 pmol of each primer, and 1 unit DNA *Taq* DNA polymerase. PCR analysis was carried out using an initial cycle of 94℃ for 5 min; 35 cycles of 94℃ for 30 s, annealing temperature of each marker (Table 2) for 30 s and 72℃ for 20 s; followed by a final extension of 72℃ for 5 min. For Cleaved Amplified Polymorphic Sequence (CAPS) marker, PCR product was digested with 5 unit appropriate restriction enzymes (Table 2). Digested PCR products were separated on 2% agarose gels.

Table 1. List of hot pepper and sweet pepper lines used for marker development

Name	Pungency	Fruit shape	CGMS line phenotype	Genotype
Chungyang A	Pungent	Elongated	Stable CMS line	*rfrf*
Chungyang C	Pungent	Elongated	Stable restorer line	*RfRf*
MSGR-A	Sweet	Small blocky	Stable CMS line	*rfrf*
SPR03	Sweet	Large blocky	Unstable restorer line	$Rf^{u}Rf^{u}$

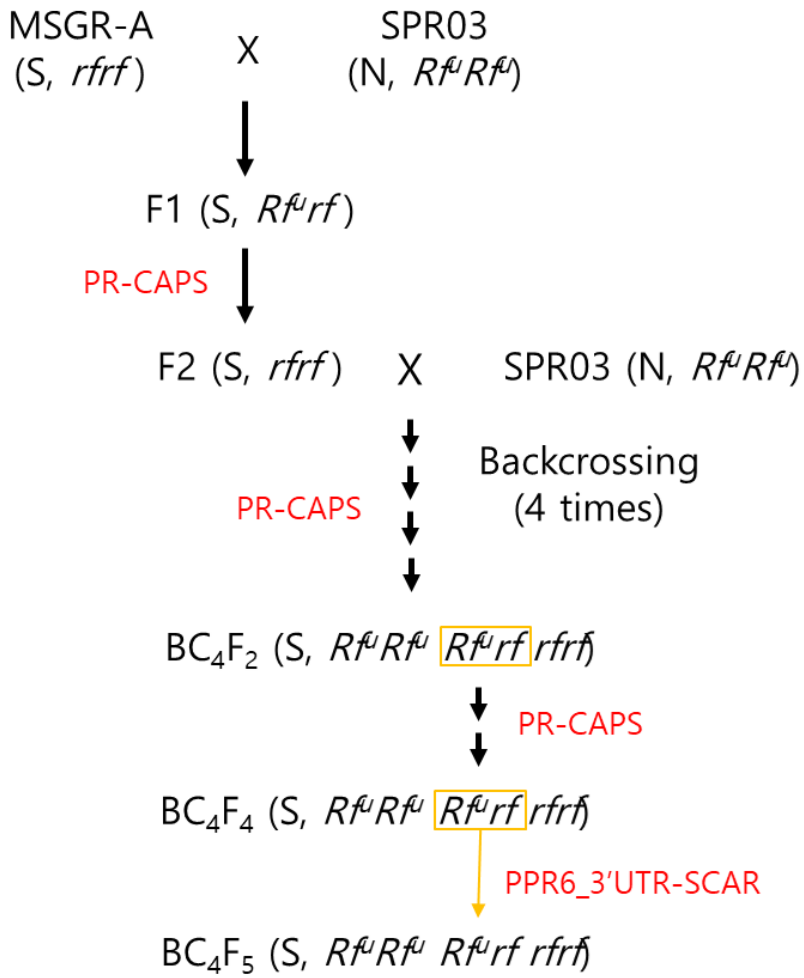

Figure 1. Pedigree of BC_4F_5 population. MSGR-A is a donor parent and stable male sterile line; SPR03 is a recurrent parent and unstable male fertile line. PR-CAPS is marker linked partial restoration locus, PPR6_3'UTR-SCAR is marker tightly linked to *CaPPR6*.

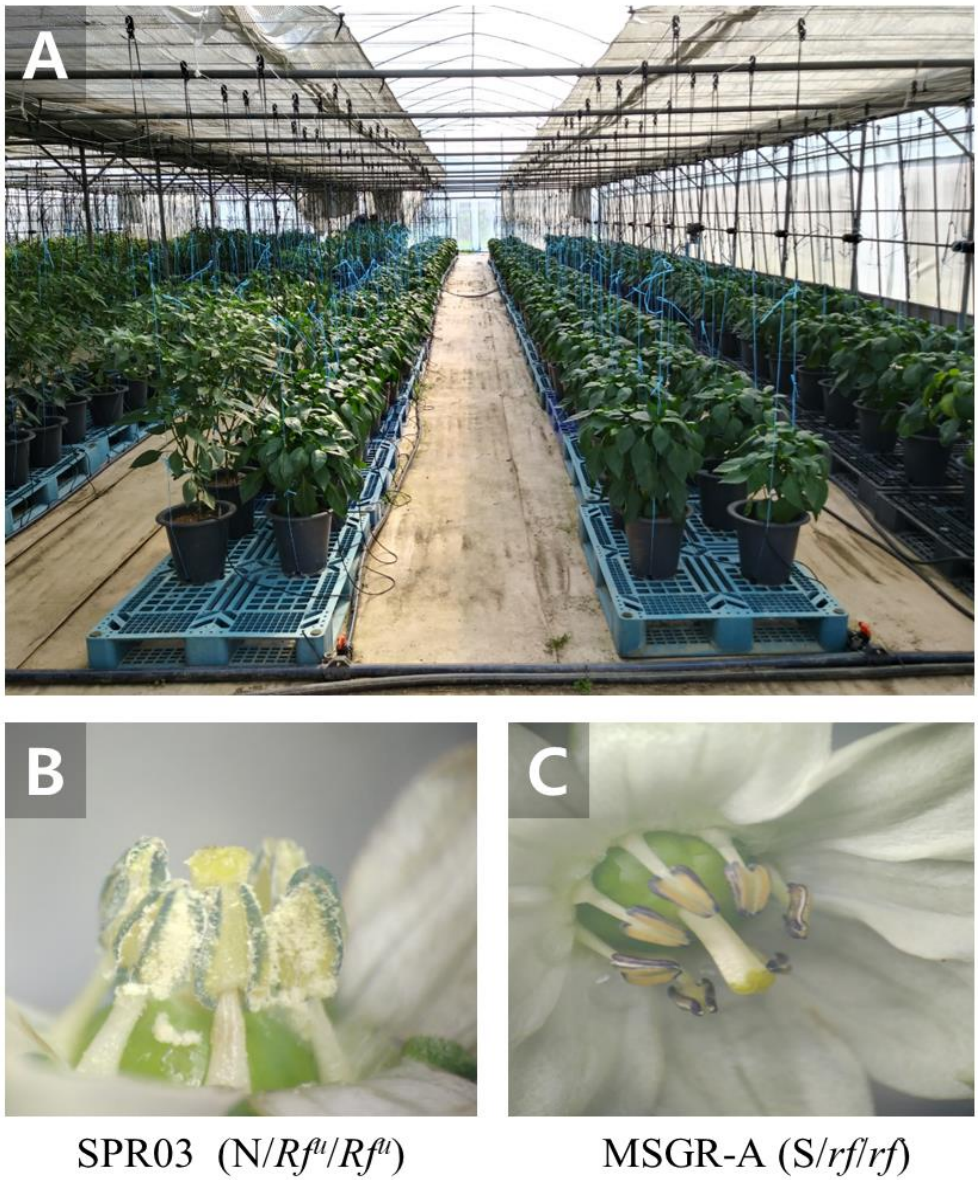

Figure 2. Phenotype scoring on unstable *Rf* population and sweet pepper parent line. A) BC_4F_5 300 population between SPR03 and MSGR-A in Seoul National University farm Suwon, B) unstable C line phenotype, C) stable A line phenotype

Table 2. Molecular marker developed by this study

Marker name	Primer sequence (5'-3'')	Annealing temperature (℃)	Marker type / Restriction enzyme	Remark
214.14MB CAPS	F : AAAGCTAAAACAGGAGCCTGAG R : ATGTTTCTTTGGGCTGTTGAGA	60℃	CAPS / *NaeI*	Developed in this study
213.85MB CAPS	F : GGAAACCAATCACATGTTGAAGA R : ACCCGCTGGATATGTCTTGTC	62℃	CAPS / *FokI*	Developed in this study
213.3MB CAPS	F : GAGCGGTTTTGATCCAAGTAACA R : ACTCTGATATACGCCACAAGGAC	60℃	CAPS / *BsrI*	Developed in this study
212.9MB CAPS	F : GGAAACCAATCACATGTTGAAGA R : ACCCGCTGGATATGTCTTGTC	60℃	CAPS / *MseI*	Developed in this study
13T7 SCAR	F : CACTAAGCCCGATGTATGAATC R : GAAGTAGGCCAAAACTTATCACG	53℃	SCAR	(Jo et al., 2010)
G16 SCAR	F : GATTTTACGATGCTCAACCC R : AAGTTGAACAATCTTCGCTG	60℃	SCAR	(Jo et al., 2016)
4162 CAPS	F : GCAGTTCGGTTTTACGGAGTTAC R : CCATTGGACAAAAGGGGATC	56℃	CAPS / *EcoRI*	(Jo et al., 2016)
PR CAPS	F : ATGTCACCCCCACACACTCCTTCACC R : TCCCATCTAGCCTCTGCCTTCTCAAATG	56℃	CAPS / *MseI*	(Lee et al., 2008)
PPR6_3'UTR SCAR	F : CTTGATGTTATAGCTTCTGA R : GTACAACTACATGTCTTGAG	56℃	SCAR	Unpublished

Development and introduction of *Rf*-linked markers

To identify polymorphism between MSGR-A and SPR03, direct sequencing was carried out by macrogen (Seoul, Korea). CAPS marker loci were identified by Basic local alignment search tool (BLAST) on 'L_Zunla-1' (http://peppersequence.genomics.cn) and 'CM334' (Hulse-kemp et al., 2018). Additional *Rf*-linked marker information was retrieved from literatures: Co1Mod1-CAPS, 4162-SCAR and G16-CAPS from Jo et al., 2016 and PR-CAPS and 13T7-SCAR from Lee et al., 2008 and Jo et al., 2010, respectively.

Linkage analysis

Linkage analysis was performed using molecular markers and 300 BC_4F_5 population from an introgression between MSGR-A and SPR03. Linkage map was constructed using Carthagene software with a LOD threshold of 3.0 and distance threshold 30cM (Givry et al., 2005). Genetic distance in linkage map was determined by using Kosambi mapping function. Based on genetic distance between markers, we create linkage map using Map Chart 2.3 software.

Gene prediction analysis

Annotation of target region was done by FGENESH program

(https://www.softberry.com). BLASTX (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to characterize predicted genes. We search database non-redundant protein sequences (nr) using Blastp (protein-protein BLAST). Optional organism is selected by *Capsicum annuum* (taxid : 4072).

RESULT

Evaluation of male fertility in a populating segregating for unstable *Rf*

The male fertility phenotype was evaluated for 300 individual plants in the BC_4F_5 population segregating for the unstable *Rf* gene. SPR03 with normal cytoplasm produced pollens regardless of temperature. However, the male fertile phenotype in BC_4F_5 population was not consistent throughout growing season. Pollens were observed in some plants in the segregating population during the spring season in clear days, but in small amounts (Figure 3A and 3B). When viability of pollens in plants producing pollens, most of pollens were viable (Figure 3C and 3D). And most of fertile plants produced fruits with seeds even though they produced small amounts of pollens indicating these plants are male fertile. To avoid confusion of phenotyping male fertility, we tried to observe pollens below 30℃ in clear days. Segregating analysis of the BC_4F_5 show 228 fertility and 72 sterility. This segregating ratio fit to a 3:1 ratio ($\chi^2 = 0.016$; $P = 0.899$) (Table 3) suggesting that the unstable restorer of fertility is controlled by single dominant gene. Unstable *Rf* is appropriate for genetic mapping of single gene.

Polymorphism test of previously developed marker in sweet pepper

To test previously developed markers in the unstable *Rf* segregating population developed in this study, we first tested polymorphisms of markers

between SPR03 (Rf^{u}/Rf^{u}) and MSGR-A (*rf*/*rf*). These markers (4162-SCAR, 120kb-end-SCAR, PPR12-SCAR, 3336-last-SCAR, Co1mod1-CAPS, 4940-CAPS) showed polymorphisms in Chung-yang C (*Rf*/*Rf*), however, could not distinguish MSGR-A and SPR03 (Figure. 4A). PPR12-SCAR and 4940-CAPS showed polymorphism in between SPR03 and MSGR-A, but PCR product was not consistent. We tried to convert previously developed *Rf* markers for application to parents of the segregating population of this study. 4162-SCAR was converted to CAPS marker (4162-CAPS) and G16-CAPS was converted to G16- SCAR (Figure. 4B). We identified marker polymorphism (13T7-SCAR, G16-SCAR, PR-CAPS, PPR6_3'UTR-SCAR and 4162-CAPS) in sweet pepper. Based on Zunla-1 genome sequence information, we identified marker loci. 13T7-SCAR, G16-SCAR, PR-CAPS, and 4162-CAPS were located at the physical positions of 214,034,516, 214,279,034, 214,746,352 and 217,546,562bp on chromosome 6. PPR6_3'UTR-SCAR is located at the physical position 213,995,089bp in Zunla-1 genome sequence information. PPR6_3'UTR-SCAR is closely located to 13T7-SCAR but the marker was developed from 3'untranslated region of *CaPPR6* and Physical distance is 800bp between *CaPPR6* and PPR6_3'UTR-SCAR. We consider that PPR6_3'UTR-SCAR is near the *CaPPR6* which is at the physical positions of 217,273,851bp. When 31 BC_4F_4 individuals were tested, 4162-CAPS showed two recombinants, G16-SCAR, PR-CAPS and PPR6_3'UTR-SCAR showed one recombinant and 13T7-SCAR doesn't show recombinant (Figure. 5). This indicates that unstable *Rf* gene is near the 13T7-SCAR.

Table 3. Segregation of fertile and sterile phenotype in the BC_4F_5 population

Number of individual			Expected ratio	χ^2	P-value
Total	Fertile	Sterile			
300	228	72	3:1	0.016	0.899

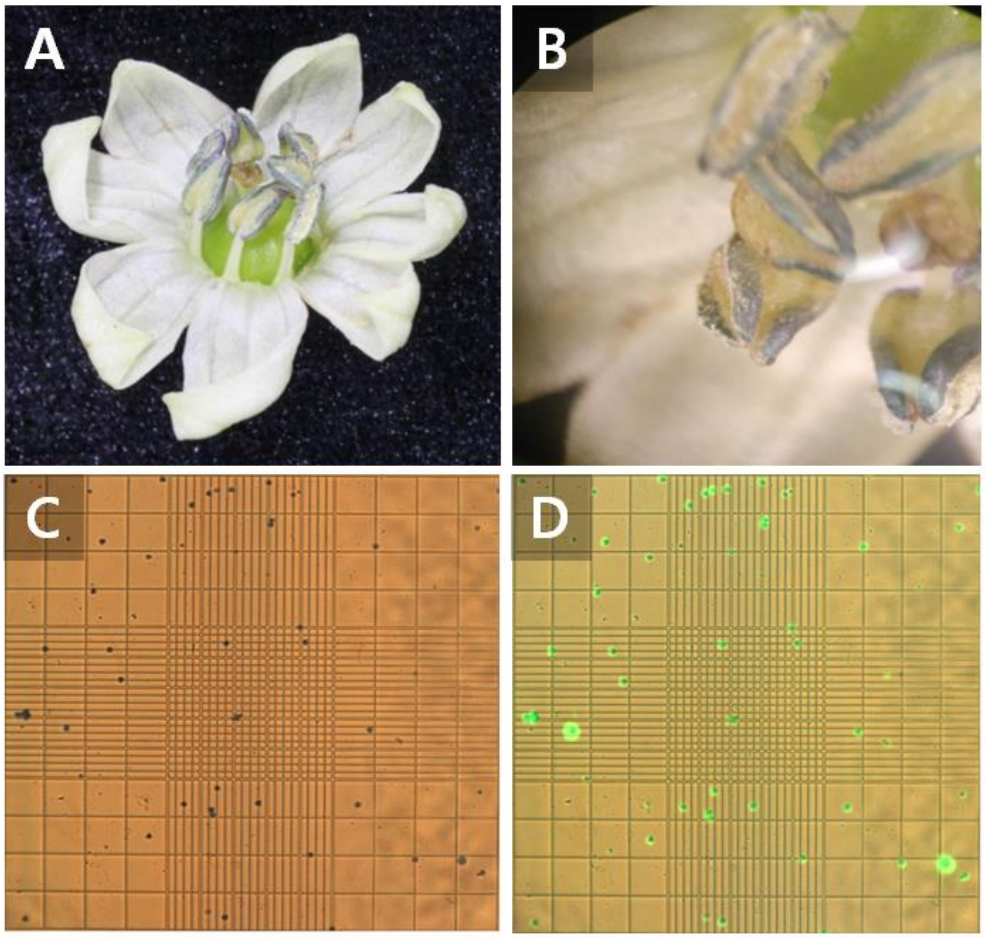

Figure 3. Unstable *Rf* phenotype and pollens viability Unstable *Rf* phenotype observed by camera (A), unstable *Rf* phenotype observed close by optical microscope (B), before FDA staining of pollen from unstable *Rf* (C), FDA staining of pollen from unstable *Rf* (D)

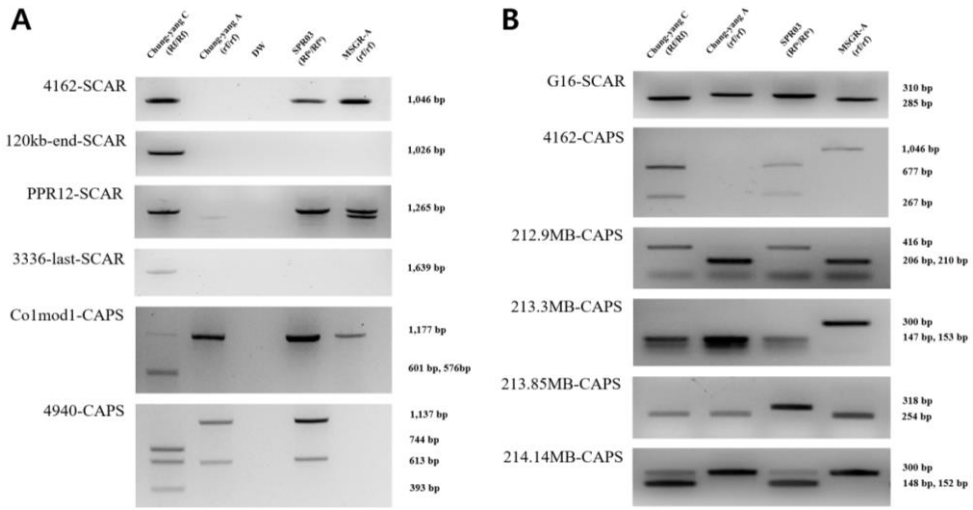

Figure 4. Application of molecular markers to hot pepper parent and sweet pepper parent line Previously developed marker test to hot pepper and sweet pepper (A), converting previously developed *Rf* markers to use in sweet pepper parent line and Newly developed markers around 13T7-SCAR (B)

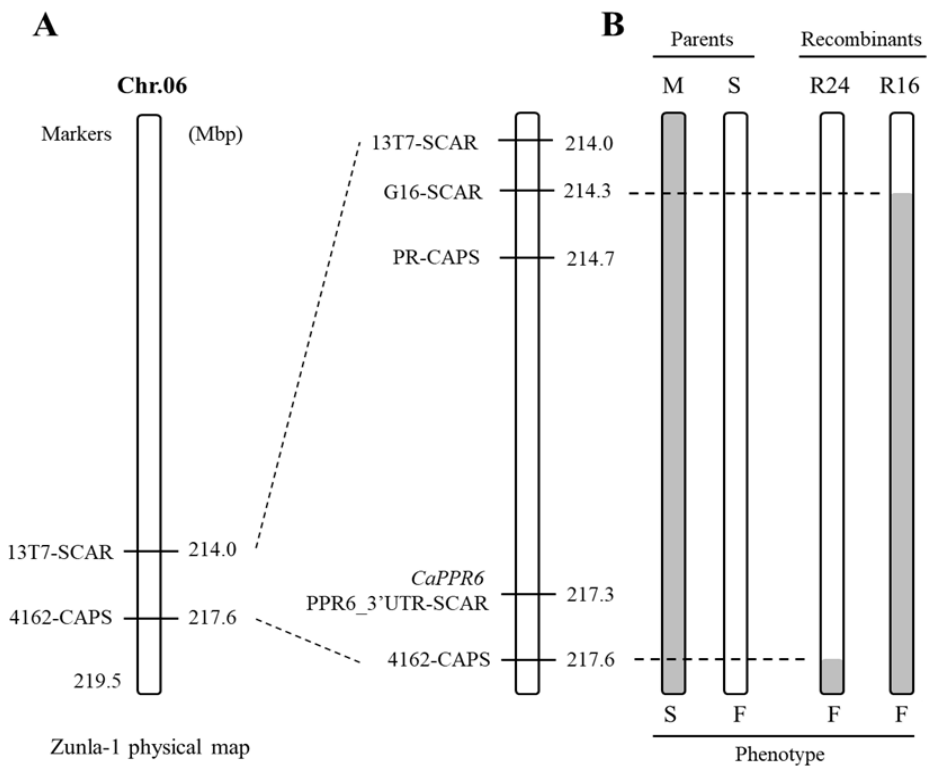

Figure 5. Genotype result of previously developed *Rf* markers to BC_4F_4

Location of previously developed marker on chromosome 6 in Zunla-1 physical map (A), result of recombinants by application of previously developed marker (B)

Development of additional markers linked to unstable *Rf*

13T7-SCAR was most closely linked to the unstable *Rf* gene among the previously developed *Rf* markers. To map the locus, we developed additional markers around 13T7-SCAR. We identified four SNPs between SPR03 and MSGR-A. Four SNPs can be converted to CAPS markers (Table.2). Newly developed markers were applied to Chung-yang and sweet pepper parent lines. As a result, 212.9MB-CAPS and 214.14MB-CAPS showed polymorphism in both parental lines, whereas 213.3MB-CAPS and 213.85MB-CAPS polymorphism was observed only in sweet pepper parent lines (Figure. 4B). Based on Zunla-1 genome sequence information, we identified marker loci. 212.9MB-CAPS, 213.3MB-CAPS, 213.85MB-CAPS, 214.4MB-CAPS were located at the physical positions of 212,936,173, 213,336,263, 213,854,539 and 214,148,114 bp on chromosome 6.

Genetic mapping of the unstable *Rf* gene

Four newly and three previously developed markers were applied to 300 individual plants in BC_4F_5 population and a local linkage map was constructed (Figure 5). PPR6_3'UTR-SCAR is closest to *CaPPR6* which is a candidate the *Rf* gene. The genetic distance of PPR6_3'UTR-SCAR from unstable the *Rf* locus was estimated to be 2.4cM (Figure. 6). The genetic distance of G16-SCAR from the unstable *Rf* locus was about 0.9cM. This suggests that the unstable *Rf* gene is located at a different position from the stable *Rf* gene. 13T7-SCAR was reported to be linked to the *Rf* locus at distances of 1.4cM (Jo et al., 2010). The genetic

distance of 13T7-SCAR from the unstable *Rf* locus was 0.4cM. Newly developed markers (213.3MB-CAPS, 212.9MB-CAPS, 213.85MB-CAPS) were linked to be the unstable *Rf* locus at genetic distance of 2.2, 4.3, 8.9cM, respectively. 214.14MB-CAPS was co-segregating with unstable *Rf*. 214.14MB-CAPS and was flanked by 13T7-SCAR and G16-SCAR. By comparing the physical position of the markers using reference genome of CM334 and Zunla-1, we delimited the target region less than 0.5 Mbp.

Analysis of gene prediction on target region

The target region of unstable *Rf* in zunla-1 and CM334 were 214,034,516 bp -214,279,034 bp (244,518 bp) and 4,574,641 bp - 5,052,969 bp (478,328 bp), respectively. A total of 37 and 80 genes were predicted in Zunla-1 and CM334, respectively (Table 4 and Table 5). In Zunla-1, protein blast analysis showed that only 13 predicted genes were characterized (Table. 4). Although one PPR protein was detected, this region was overlapped with 13T7-SCAR. Whereas, in CM334, 30 predicted proteins were characterized (Table. 5). One PPR protein was detected and 25 kb away from the co-segregated marker. Protein GAMETE EXPRESSED 2 (GEX2) was also detected. This gene was 43.5kb away from the co-segregated marker.

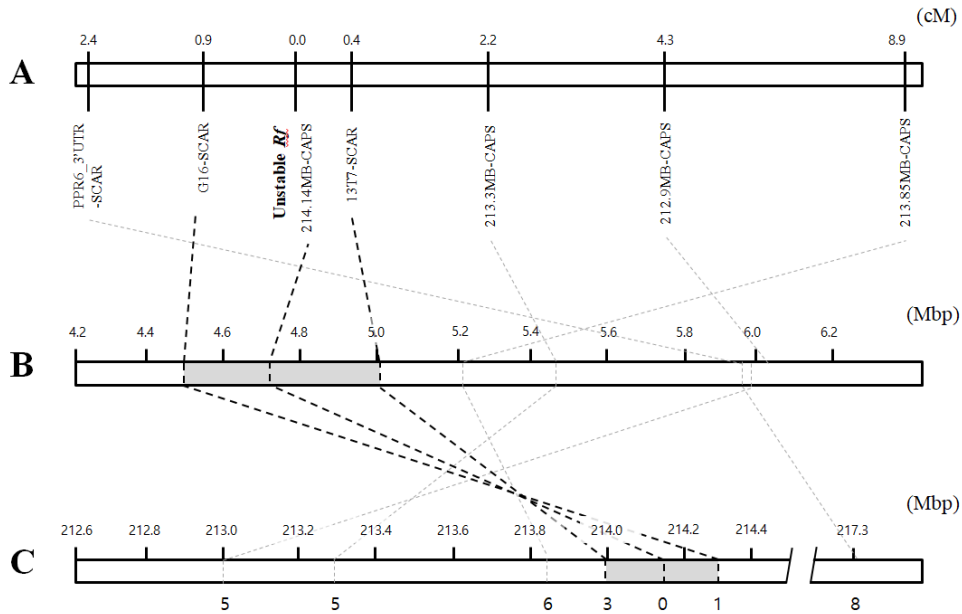

Figure 6. Comparison between physical map and linkage map from markers

Genetic location of unstable *Rf* in 300 BC_4F_5 population (A), and physical map of the markers in 'CM334' (B), and 'Zunla-1' and the number of recombinant (C)

Table 4. Gene prediction of target region in Zunla-1

Location on Zunla-1 physical map (bp)	Length of CDS	Sequence ID	Predicted characteristic
214,058,286 - 214,058,627	342 bp	PHT80469.1	Protein transport protein Sec24-like CEF
214,076,189 – 214,077,827	807 bp	XP_016576773.1	Adenylate isopentenyltransferase 5, chloroplastic-like
214,091,427 – 214,092,472	783 bp	PHT78430.1	Delta-1-pyrroline-5-carboxylate synthase
214,117,942 – 214,118,352	411 bp	XP_016546408.1	Keratin, type I cytoskeletal 12-like
214,120,445 - 214,121,074	630 bp	PHT68913.1	Aquaporin TIP1-2
214,136,519 – 214,139,789	816 bp	XP_016576011.1	Putative ER lumen protein-retaining receptor C28H8.4
214,150,353 – 214,165,195	993 bp	XP_016576775.1	Ycf54-like protein
214,171,792 – 214,172,277	486 bp	PHT66284.1	S-norcoclaurine synthase 2-like
214,173,124 – 214,197,827	1,296 bp	XP_016576015.1	E3 ubiquitin-protein ligase RNF4-like
214,206,261 – 214,206,986	654 bp	PHT94129.1	Beta-glucosidase-like SFR2, chloroplastic
214,212,376 – 214,234,076	1,356 bp	XP_016577264.1	Tripartite motif-containing protein 44-like
214,234,210 – 214,239,565	492 bp	XP_016565630.1	Calcium-dependent protein kinase 4

Table 5. Gene prediction of target region in CM334

Location on CM334 physical map (bp)	Length of CDS	Sequence ID	Predicted characteristic
4,614,934 - 4,623,229	666 bp	XP_016577264.1	PREDICTED: tripartite motif-containing protein 44-like sequence
4,623481 - 4,627,451	375 bp	XP_016576019.1	PREDICTED: xyloglucan endotransglucosylase/hydrolase protein 2-like isoform X3
4,629,673 - 4,643,871	828 bp	XP_016576017.1	PREDICTED: xyloglucan endotransglucosylase/hydrolase protein 2-like isoform X1
4,649,255 - 4,654,855	654 bp	PHT94129.1	Beta-glucosidase-like SFR2, chloroplastic Sequence
4,658,846 – 4,674,450	1,296 bp	XP_016576015.1	PREDICTED: E3 ubiquitin-protein ligase RNF4-like sequence
4,675,297 – 4,675,782	486 bp	XP_016577263.1	PREDICTED: S-norcoclaurine synthase 2-like sequence
4,698,285 – 4,702,396	747 bp	XP_016575979.1	PREDICTED: pentatricopeptide repeat-containing protein At1g62670, mitochondrial-like isoform X2
4,703,515 – 4,704,000	486 bp	PHT84892.1	Peroxidase 72 Sequence
4,705,463 – 4,709,462	963 bp	PHT87666.1	Cytochrome 89A2 Sequence
4,714,812 – 4,721,599	732 bp	XP_016576775.1	PREDICTED: ycf54-like protein
4,732,119 – 4,735,388	816 bp	XP_016576011.1	PREDICTED: putative ER lumen protein-retaining receptor C28H8.4
4,750,659 – 4,751,288	630 bp	PHT68913.1	Aquaporin TIP1-2 Sequence
4,752,264 – 4,753,790	825 bp	XP_016546408.1	PREDICTED: keratin, type I cytoskeletal 12-like Sequence
4,766,938 – 4,774,670	927 bp	PHT94469.1	Protein GAMETE EXPRESSED 2 Sequence
4,775,333 – 4,785,632	954 bp	PHT78430.1	Delta-1-pyrroline-5-carboxylate synthase Sequence
4,790,633 – 4,792,271	807 bp	XP_016576773.1	PREDICTED: adenylate isopentenyltransferase 5, chloroplastic-like Sequence
4,811,092 – 4,826,611	1,407 bp	PHT78430.1	Delta-1-pyrroline-5-carboxylate synthase Sequence
4,827,133 – 4,832,097	912 bp	XP_016576773.1	PREDICTED: adenylate isopentenyltransferase 5, chloroplastic-like Sequence

4,832,995 – 4,845,817	1,665 bp	PHT78430.1	Delta-1-pyrroline-5-carboxylate synthase
4,861,621 – 4,862,307	687 bp	XP_016546408.1	PREDICTED: keratin, type I cytoskeletal 12-like Sequence
4,918,842 – 4,927,910	636 bp	PHT68913.1	Aquaporin TIP1-2 Sequence
4,928,320 – 4,930,253	519 bp	XP_016546408.1	PREDICTED: keratin, type I cytoskeletal 12-like Sequence
4,952,133 – 4,962-629	339 bp	PHT68430.1	ADP-ribosylation factor GTPase-activating protein AGD1 Sequence
4,965,070 – 4,977,134	588 bp	PHT68266.1	Cell differentiation protein RCD1 -like protein Sequence
4,979,066 – 4,985,748	4,050 bp	XP_016556020.1	PREDICTED: putative late blight resistance protein homolog R1A-3 isoform X1 Sequence
5,004,914 – 5,012,194	3,753 bp	XP_016556023.1	PREDICTED: LOW QUALITY PROTEIN: putative late blight resistance protein homolog R1A-3 Sequence
5,012,570 – 5,013,235	318 bp	XP_016558148.1	PREDICTED: uncharacterized PE-PGRS family protein PE_PGRS54-like Sequence
5,023,335 – 5,036,703	4,278 bp	XP_016556034.1	PREDICTED: LOW QUALITY PROTEIN: putative late blight resistance protein homolog R1A-3 Sequence
5,039,329 – 5,039,930	468 bp	XP_016544454.1	PREDICTED: cytochrome P450 76A2-like Sequence
5,043,072 – 5,047,639	378 bp	PHT77060.1	Cytochrome Sequence

Application of *Rf*-linked markers to stable and unstable breeding lines

Co-segregated marker and four previously developed markers were used to genotype stable and unstable breeding lines provided by Eco-seed (Table 6). When 13T7-SCAR was applied to stable male sterile lines, nine lines show showed *rf* genotype and five lines showed *Rf* genotype. The marker showed low accuracy (64.2%). Co1Mod1-CAPS is located to the region close to the *CaPPR6* gene (11kb away). All of stable male sterile lines showed *rf* genotype and thus the marker genotype correspond to stable *rf* genotype. However, all of unstable male fertile lines showed *rf* genotype by application of 13T7-SCAR and Co1Mod1-CAPS. Unlikely other markers, 4162-CAPS showed two genotype classes of sterility (*rfrf*, *rfrf2*) (Table. 6). We expected that two genotype classes of sterility can distinguish stable male sterile and unstable male fertile lines. However, unstable male fertile and stable male sterile lines showed second genotype (*rfrf2*). In stable male sterile lines, two, seven and five lines showed first genotype (*rfrf*), second genotype (*rfrf2*) and *Rf* genotype. In unstable male fertile lines, two lines and ten lines showed *Rf* genotype and second genotype (*rfrf2*). Genotype accuracy of 4162-CAPS is low (16%) in unstable male fertile lines. PPR6_3'UTR-SCAR was developed in the 3'untranslated region of *CaPPR6*. In stable male sterile lines, six and eight lines showed *rf* genotype and *Rf* genotype respectively and this marker showed low accuracy (42.8%). In unstable male fertile lines, two and ten lines showed *rf* genotype and *Rf* genotype. The marker is more accurate than other previously developed markers in unstable male fertile lines (76.9%). The 214.14MB-CAPS

which is co-segregated with unstable *Rf* showed high accuracy (91.6%) in unstable male fertile lines. Twelve of thirteen lines showed *Rf* genotype in unstable fertile lines. On the other hand, eight and six lines showed *Rf* and *rf* genotype respectively and thus the marker showed low accuracy (42.8%) in stable male sterile lines. This suggests that 214.14MB-CAPS can use for detecting unstable male fertile lines. These markers can broadly use marker application to stable male sterile lines or unstable male fertile lines but no marker can detect stable male sterile lines or unstable male fertile lines. We generated phylogenetic treed based on marker genotypes from stable or unstable breeding lines. Most unstable male fertile or stable male sterile lines gather cluster respectively (Figure. 7). This suggested that application of several markers rather than single marker can distinguish unstable male fertile lines from stable male sterile lines. However, a few unstable male fertile or stable male sterile lines belong to different cluster (Figure. 7). We considered that these lines are necessary to confirm phenotype.

Table 6. Application of *Rf*-linked markers to stable and unstable breeding lines

Breeding line	Male fertility	Marker genotype							Remark
		212.9MB	13T7	214.14 MB	G16	Co1Mod1	PPR6_ 3'UTR	4162	
		212.94 Mbp*	214.04 Mbp*	214.14 Mbp*	214.3 Mbp*	217.3 Mbp*	217.3 Mbp*	217.6 Mbp*	
Chungyang A	Stable rf	rfrf	rfrf	rfrf	rfrf	rfrf	rfrf	rfrf	Stable sterility
A-178	Stable rf	rfrf	rfrf	RfRf	rfrf	rfrf	RfRf	rfrf	
MSGR-A	Stable rf	rfrf	rfrf	rfrf	RfRf	rfrf	RfRf	rf2rf2	Stable sterility
UA-2	Stable rf	rfrf	rfrf	rfrf	RfRf	rfrf	RfRf	rf2rf2	Cluster and erect type
UA-3	Stable rf	rfrf	rfrf	rfrf	RfRf	rfrf	RfRf	rf2rf2	Cluster and erect type
UA-4	Stable rf	rfrf	rfrf	rfrf	RfRf	rfrf	RfRf	rf2rf2	Cluster and erect type
A-136	Stable rf	rfrf	rfrf	RfRf	RfRf	rfrf	RfRf	rf2rf2	Southern type
A-138	Stable rf	rfrf	rfrf	rfrf	RfRf	rfrf	RfRf	rf2rf2	Southern type
A-85	Stable rf	rfrf	rfrf	RfRf	rfrf	rfrf	RfRf	rf2rf2	GMS--> CGMS
A-76	Stable rf	rfrf	RfRf	RfRf	RfRf	rfrf	rfrf	RfRf	GMS--> CGMS
A-81	Stable rf	rfrf	RfRf	RfRf	RfRf	rfrf	rfrf	RfRf	GMS--> CGMS
A-82	Stable rf	rfrf	RfRf	RfRf	RfRf	rfrf	rfrf	RfRf	GMS--> CGMS
A-83	Stable rf	rfrf	RfRf	RfRf	RfRf	rfrf	rfrf	RfRf	GMS--> CGMS
A-84	Stable rf	rfrf	RfRf	RfRf	RfRf	rfrf	rfrf	RfRf	GMS--> CGMS
SPR03	Unstable Rf	RfRf	RfRf	RfRf	rfrf	rfrf	rfrf	RfRf	Unstable Rf
A-148	Unstable Rf	rfrf	rfrf	RfRf	RfRf	rfrf	RfRf	rf2rf2	Chinese origine (Ligeumjo type)
A-176	Unstable Rf	rfrf	rfrf	rfrf	RfRf	rfrf	RfRf	rf2rf2	Southern MS line
A-1	Unstable Rf	rfrf	rfrf	RfRf	rfrf	rfrf	RfRf	rf2rf2	AC2258origin (Daedeulbo)
A-2	Unstable Rf	rfrf	rfrf	RfRf	rfrf	rfrf	Rfrf	rf2rf2	AC2258 origin (Daedeulbo)
A-4	Unstable Rf	rfrf	rfrf	RfRf	rfrf	rfrf	RfRf	rf2rf2	AC2258 origin (Daedeulbo)
A-5	Unstable Rf	rfrf	rfrf	RfRf	rfrf	rfrf	RfRf	rf2rf2	AC2258 origin (Daedeulbo)
A-72	Unstable Rf	rfrf	rfrf	RfRf	rfrf	rfrf	RfRf	rf2rf2	AC2259 origin PR line MS line
A-73	Unstable Rf	rfrf	rfrf	RfRf	rfrf	rfrf	RfRf	rf2rf2	(GMS--> CGMS)
A-74	Unstable Rf	rfrf	rfrf	RfRf	rfrf	rfrf	RfRf	rf2rf2	(GMS--> CGMS)
A-75	Unstable Rf	rfrf	rfrf	RfRf	rfrf	rfrf	RfRf	rf2rf2	(GMS--> CGMS)
A-79	Unstable Rf	rfrf	rfrf	RfRf	rfrf	rfrf	rfrf	RfRf	Geumtap
Chungyang C	Stable Rf	RfRf	RfRf	RfRf	RfRf	RfRf	RfRf	RfRf	Stable fertility

***** Marker physical location based on Zunla-1 genome sequence information

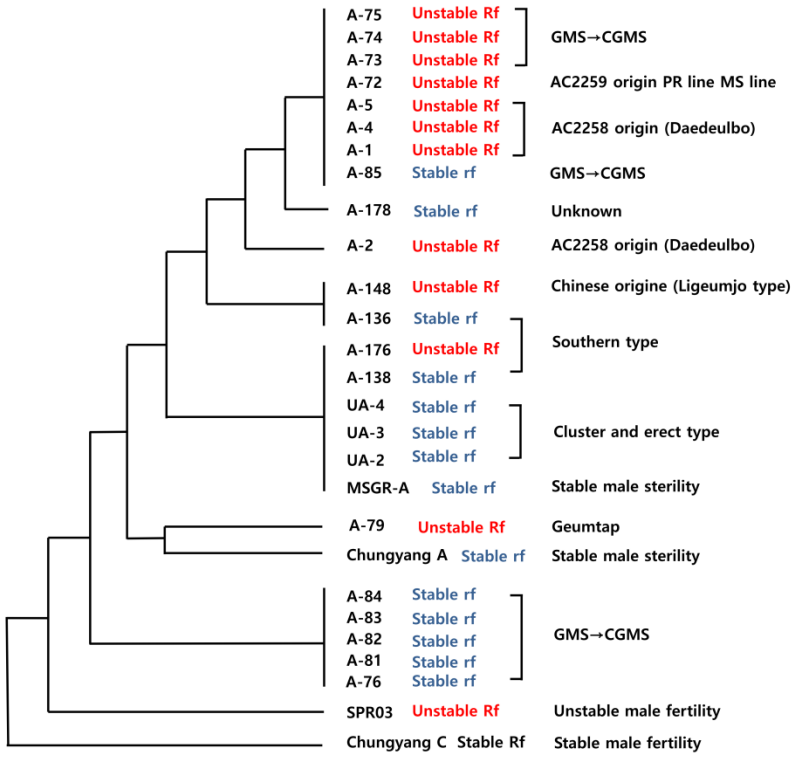

Figure 7. Phylogenetic tree based on marker genotypes from stable or unstable breeding lines

Discussion

CGMS is a useful system for producing F1 hybrid seeds in pepper and other crops. However, the unstable fertility in restorer lines showing inconsistent male fertility depending on environmental conditions is a hindrance to use the CGMS system in sweet pepper. Many prior studies have attempted to identify the stable *Rf* gene, and as a result various markers and candidate gene have been identified. However, only a few researches have been done on the unstable *Rf* gene in pepper. Even though sweet pepper is a high value crop, CGMS system is not being used efficiently. The reason may be because sweet peppers do not have an *Rf* allele or restorer lines in sweet peppers may have weak fertility restoration abilities (Lin et al., 2015).

In this study, we investigated genetic factors of unstable fertility of a sweet pepper line SPR03 by phenotype analysis and genetic mapping. SPR03 is a blocky type pepper with a normal cytoplasm but has unstable fertility. We used MSGR-A as a maternal parent which is a stable sterile line. The unstable male fertility is affected by environmental conditions, especially high temperature. The interaction between genotype and environment is a common feature of quantitative traits (QTL). Wang et al. (2004) found one major QTL on chromosome 6 and additional minor QTLs by QTL analysis of restorer fertility in pepper. The unstable *Rf* was also considered to be related to QTL, but it was confirmed to be a single

dominant gene in this study.

Before genetic mapping of unstable male fertility, previously developed markers were applied to 31 BC_4F_4 individuals to test whether the unstable *Rf* gene is also associated with the *CaPPR6* gene. 4162-CAPS showed two recombinants and PPR6_3'UTR-SCAR, PR-CAPS and G16-SCAR showed one recombinant in this population. This suggested that the genetic factor of unstable fertility in SPR03 is linked to *Rf* but controlled by a different gene. Of the previously developed markers, only 13T7-SCAR showed no recombinant indicating that the unstable fertility gene is located near 13T7-SCAR. We then applied three previously developed markers and four newly developed markers to 300 individual plants from the BC_4F_5 population and created a genetic map. Genetic mapping showed that unstable fertility locus was between 13T7-SCAR and G16-SCAR and co-segregated with 214.14MB-CAPS. Although the physical distance based on Zunla-1 genome between 13T7-SCAR and G16-SCAR is about 200 kb, genetic distance between two markers is relatively far (1.3cM). It appears that this region may be a recombination hot spot or the genome assembly of this region is incorrect due to because the repeat sequences.

Gene prediction analysis from DNA region between 13T7-SCAR and G16-SCAR showed possible candidate genes for the unstable fertility. In Zunla-1 and CM334, putative ER lumen protein-retaining receptor C28H8.4 (XP_016576011.1) and Ycf54-like protein (XP_016576775.1) were located close

to 214.14MB-CAPS. ER lumen protein-retaining receptor C28H8.4 is required for the retention of luminal endoplasmic reticulum proteins. This protein is required for normal vesicular traffic through the golgi. The function of Ycf54-like protein has been unknown, but it is located in the chloroplast genomes. Even though these genes are closed linked to the unstable fertility, the functions of the genes are not directly related to the phenotype. By contrast, one gene encoding PPR protein was detected in CM334 genome. This gene is consisted of 3 exon and 5 PPR repeats. Besides PPR protein, Protein GAMETE EXPRESSED 2 (GEX2) was also detected. Although this gene has not been identified as an *Rf*, it is associated with pollen development. Peroxidase was also detected on the target region. Transcriptome level analysis between CMS line and its near isogenic restorer line revealed differentially expressed uni-gene including peroxidase (Liu et al., 2013). The function of peroxidase is removal of H_2O_2, oxidation of toxic reductants, biosynthesis and degradation of lignin. These three candidate gene are needed to be analyzed sequence further.

Until now, genetic mapping has been done only the stable *Rf* gene and broadly applicable markers have not been developed. In this study we developed markers for the unstable male fertility 214.14MB-CAPS and found that this marker is tightly linked the phenotype. Although application of marker co-segregated with unstable *Rf* can be broadly use unstable male fertile lines, stable male sterile or unstable male fertile could not be distinguished. This may be because breeders may have misjudged male fertility phenotype in some lines. However, phylogenetic tree

based on genotype from several markers showed that stable or unstable lines are prone to be clustered together. In further study, we need to confirm candidate genes identified in this study and male fertility phenotype of these lines.

References

Barkan A, Small I (2014) Pentatricopeptide Repeat Proteins in Plants. *Annu Rev Plant Biol* 65:415-442

Bohra A, Jha UC, Adhimoolam P, Bisht D, Singh NP (2016) Cytoplasmic male sterility (CMS) in hybrid breeding in field crops. *Plant Cell Rep* 35:967–993

Chase CD (2007) Cytoplasmic male sterility: window to the world of plant mitochondrial-nuclear interactions. *Trends Genet* 23(2):81-90

Chen L and Liu YG (2014) Male sterility and fertility restoration in crops. *Annu Rev Plant Biol* 65:579–606

Chen Z, Zhao N, Li S, Grover C, Nie H, Wendel J, Hua J (2017) Plant mitochondrial genome evolution and cytoplasmic male sterility. *CRC Crit Rev Plant Sci* 36:55-69

Colombo N, Galmarini C (2017) The use of genetic, manual and chemical methods to control pollination in vegetable hybrid seed production: a review. *Plant Breeding* 136:287-299.

Cui X, Wise R, Schnable P (1996) The *rf2* nuclear restorer gene of male-sterile T-cytoplasm maize. *Science* 272:1334-1336.

Dahan J, Mireau H (2013) The *Rf* and *Rf*-like PPR in higher plants, a fast-evolving

subclass of PPR genes. *RNA Biol* 10(9):1469-1476

Duroc Y, Gaillard C, Hiard S, Defrance MC, Pelletier G, Budar F (2005) Biochemical and functional characterization of ORF138, a mitochondrial protein responsible for Ogura cytoplasmic male sterility in *Brassiceae*. *Biochimie* 87(12):1089-1100

Fujii S, Kazama T, Yamada M, Toriyama K (2010) Discovery of global genomic re-organization based on comparison of two newly sequenced rice mitochondrial genomes with cytoplasmic male sterility-related genes. *BMC Genomics* 11:209

Fujii S, Toriyama K (2009) Suppressed expression of *RETROGRADE-REGULATED MALE STERILITY* restores pollen fertility in cytoplasmic male sterile rice plants. *Proc. Natl Acad Sci U S A* 106:9513-9518.

Gabay-Laughnan S, Kuzmin E, Monroe J, Roark L, Newton K (2009) Characterization of a novel thermosensitive restorer of fertility for cytoplasmic male sterility in maize. *Genetics* 182(1):91-103

Givry SD, Bouchez M, Chabrier P, Milan D, Schiex T (2005) CARTHAGENE: multi-population integrated genetic and radiation hybrid mapping. *Bioinformatics* 21:1703-1704

Gulyas G, Pakozdi K, Lee J, Hirata Y (2006) Analysis of fertility restoration by using cytoplasmic male sterile red pepper (*Capsicum annuum* L.) lines.

Breed Sci 56:331-334

Ha Y (2017) Candidate gene analysis of the *Restorer-of-fertility* gene in pepper and transgenic expression of the pepper CMS associated gene *orf507* in *Nicotiana benthamiana*. M S. Thesis. Seoul National University, Seoul., Korea. 71pp.

Hulse-Kemp AM, Maheshwari S, Hill TA, Jaffe D, Williams SR, Weisenfeld N, Ramakrishnan S, Kumar V, Shah P, Schatz MC, Church DM, Van Deynze A (2018) Reference quality assembly of the 3.5-Gb genome of *Capsicum annuum* from a single linked-read library. *Hortic Res* 5:4

Hu J, Wang K, Huang W, Liu G, Gao Y, Wang J, Huang Q, Ji Y, Qin X, Wan L, Zhu R, Li S, Yang D, Zhu Y (2012) The rice pentatricopeptide repeat protein RF5 restores fertility in Hong-Lian cytoplasmic male-sterile lines via a complex with the glycine-rich protein GRP162. *Plant Cell* 24(1): 109-122

Huang W, Yu C, Hu J, Wang L Dan Z, Zhou W, He C, Zeng Y, Yao G, Qi J, Zhang Z, Zhu R, Chen X, Zhu Y (2015) Pentatricopeptide-repeat family protein RF6 functions with hexokinase 6 to rescue rice cytoplasmic male sterility. *Proc. Natl. Acad. Sci. USA* 112:14984–14989

Itabashi E, Iwata N, Fujii S, Kazama T, Toriyama K (2011) The fertility restorer gene, *Rf2*, for Lead Rice-type cytoplasmic male sterility of rice encodes a mitochondrial glycine-rich protein. *Plant J* 65:359-367.

Jing B, Heng S, Tong D, Wan Z, Fu T, Tu J, Ma C, Yi B, Wen J, Shen J (2012) A male sterility-associated cytotoxic protein ORF288 in *Brassica Juncea* causes aborted pollen development. *J Exp Bot* 63(3):1285-1295

Jordan D, Mace E, Henzell R, Klein P, Klein R (2010) Molecular mapping and candidate gene identification of the *Rf2* gene for pollen fertility restoration in sorghum [*Sorghum bicolor* (L.) Moench]. *Theor Appl Genet* 120:1279–1287

Jo YD, Choi Y, Kim DH, Kim BD, Kang BC (2014) Extensive structural variations between mitochondrial genomes of CMS and normal peppers (*Capsicum annuum* L.) revealed by complete nucleotide sequencing. *BMC Genomics* 15(1): 561

Jo YD, Kim YM, Park MN, Yoo JH, Park MK, Kim BD, Kang BC (2010) Development and evaluation of broadly applicable markers for *Restorer-of-fertility* in pepper. *Mol Breeding* 25:187-201

Jo YD, Ha Y, Lee JH, Park M, Bergsma AC, Choi HI, Goritschnig S, Kloosterman B, van Dijk PJ, Choi D, Kang BC (2016) Fine mapping of Restorer-of-fertility in pepper (*Capsicum annuum L.*) identified a candidate gene encoding a pentatricopeptide repeat (PPR)-containing protein. *Theor Appl Genet* 129(10):2003-2017

Koizuka N, Imai R, Fujimoto H, Hayakawa T, Kimura Y, Kohno-Murase J, Sakai T,

Kawasaki S, Imamura J (2003) Genetic characterization of a pentatricopeptide repeat protein gene, *orf687*, that restores fertility in the cytoplasmic male-sterile Kosena radish. *Plant J* 34:407–415

Kim DS, Kim DH, Yoo JH, Kim BD (2006) Cleaved amplified polymorphic sequence and amplified fragment length polymorphism markers linked to the fertility restorer gene in chili pepper. *Mol Cells* 21:135-140

Kim YJ, Zhang D (2018) Molecular Control of Male Fertility for Crop Hybrid Breeding. *Trends Plant Sci* 23(1):53-65

Kim YM, Jo YD, Kwon JK, Kim IT, Kang BC (2013) Characterization and inheritance of a novel thermo-sensitive restoration of cytoplasmic male sterility in *Capsicum annuum. Sci Hortic* 164:512-520

Kitazaki K, Arakawa T, Matsunaga M, Yui-Kurino R, Matsuhira H, Mikami T, Kubo T (2015) Post-translational mechanisms are associated with fertility restoration of cytoplasmic male sterility in sugar beet (*Beta vulgaris*). *Plant J* 83:290–299

Kubo T, Kitazaki K, Matsunaga M, Kagami H, Mikami T (2011) Male sterility-inducing mitochondrial genomes: How do they differ? *CRC Crit Rev Plant Sci* 30:378-400

Lee DH (2001) Studies on unstable fertility of CGMS (cytoplasmic-genic male sterility) in *Capsicum annuum* L. Ph D. Diss., Seoul Natl Univ,

Seoul., Korea. 94pp.

Lee J, Yoon JB, PARK HG (2008a) Likage analysis between the partial restoration (*pr*) and the restorer of fertility (*Rf*) loci in pepper cytoplasm male sterility. *Theor Appl Genet* 117:383-389

Lee J, Yoon JB, Park HG (2008b) A CAPS marker associated with the partial restoration of cytoplasmic male sterility in chilli pepper (*Capsicum annuum* L.). *Mol breed* 21:95-104

Lin S, Shieh H, Wang Y, Tan C, Schafleitner R, Yang W, Kumar S (2015) Restorer breeding in sweet pepper: Introgressing *Rf* allele from hot pepper through marker-assisted backcrossing. *Sci Hortic* 197:170-175

Liu C, Ma N, Wang PY, Fu N, Shen HL (2013) Transcriptome sequencing and de novo analysis of a cytoplasmic male sterile line and its near-isogenic restorer line in chili pepper (*Capsicum annuum L.*). *PLoS One* 8(6):e65209

Liu Z, Yang Z, Wang X, Li K, An H, Liu J, Yang G, Fu T, Yi B, Hong D (2016) A mitochondria-targeted PPR protein restores pol cytoplasmic male sterility by reducing orf224 transcript levels in oilseed rape. *Mol. Plant* 9:1082–1084

Luo D, Xu H, Guo J, Li H, Chen L, Fang C, Zhang Q, Bai M, Yao N, Wu H, Ji C, Zheng H, Chen Y, Ye S, Li X, Zhao X, Li R, Liu YG (2013) A

detrimental mitochondrial-nuclear interaction causes cytoplasmic male sterility in rice. *Nat Genet* 45(5):573-577

Manna S (2015) An overview of pentatricopeptide repeat proteins and their applications. *Biochimie* 113:93-99

Park SW, An SJ, Yang HB, Kwon JK, Kang BC (2009) Optimization of high resolution melting analysis and discovery of single nucleotide polymorphism in *Capsicum. Hort Environ Biotechnol* 50:31-39

Peterson PA (1958) Cytoplasmically inherited male sterility in *Capsicum. The Amcerican Naturalist* 92:111-119

Shen C, Zhang D, Guan Z, Liu Y, Yang Z, Yang Y, Wang X, Wang Q, Zhang Q, Fan S, Zou T, Yin P (2016) Structural basis for specific single-stranded RNA recognition by designer pentatricopeptide repeat proteins. *Nat Com* 7:11285

Tang H, Luo D, Zhou D, Zhang Q, Tian D, Zheng X, Chen L, Liu Y (2014) The rice restorer Rf4 for wild-abortive cytoplasmic male sterility encodes a mitochondrial localized PPR protein that functions in reduction of WA352 transcripts. *Mol. Plant* 7:1497–1500.

Touzet P, Meyer EH (2014) Cytoplasm male sterility and mitochondrial metabolism in plants. *Mitochondrion* Pt B:166-171

Uyttewaal M, Arnal N, Quadrado M, Martin-Canadell A, Vrielynck N, Hiard S,

Gherbi H, Bendahmane A, Budar F, Mireau H (2008) Characterization of Raphanus sativus pentatricopeptide repeat proteins encoded by the fertility restorer locus for Ogura cytoplasmic male sterility. *Plant Cell* 20:3331– 3345

Wang K, Gao F, Ji Y, Liu Y, Dan Z, Yang P, Zhu Y, Li S (2013) ORFH79 impairs mitochondrial function via interaction with a subunit of electron transport chain complex III in Honglian cytoplasmic male sterile rice. *New Phytol* 198(2):408-18

Wang Z, Zou Y, Li X, Zhang Q, Chen L, Wu H, Su D, Chen Y, Guo J, Luo D, Long Y, Zhong Y, Liu Y (2006) Cytoplasmic Male Sterility of Rice with Boro II Cytoplasm Is Caused by a Cytotoxic Peptide and Is Restored by Two Related PPR Motif Genes via Distinct Modes of mRNA Silencing. *The Plant Cell* 18:676-687.

Zhang B, Sanwen H, Guimei Y, Jiazhen G (2000) Two RAPD markers linked to a major fertility restorer gene in pepper. *Euphytica* 113(2):155-161

Zhang H, Che J, Ge Y, Pei Y, Zhang L, Liu Q, Gu M, Tang S (2017) Ability of Rf5 and *Rf6* to restore fertility of chinsurah boro II-type cytoplasmic male sterile *Oryza sativa* (spp. Japonica) lines. *Rice* 10(1):2

Zhang H, Wu J, Dai Z, Qin M, Hao L, Ren Y, Li Q, Zhang L (2017) Allelism analysis of *BrRfp* locus in different restorer lines and map-based cloning

of a fertility restorer gene, *BrRfp1*, for pol CMS in Chinese cabbage (*Brassica rapa* L.). *Theor Appl Genet* 130(3):539-547

ABSTRACT IN KOREAN

세포질유전자적웅성불임(CGMS)은 F_1 잡종종자를 생산하기 위해 여러 작물에서 활용된다. 식물에서 CGMS 시스템은 미토콘드리아 유전체 있는 불임 유전자와 핵내 회복유전자의 상호작용에 의해 결정된다. 지금까지 안정한 회복친에서 유전분석이 시도되어 왔고 회복유전자 연관 마커와 후보유전자 들이 밝혀졌다. 그러나 불안정한 회복친 관련 연구는 거의 없었다. 따라서 본 연구에서는 불안정 불임 회복 유전자의 표현형과 연관된 마커를 관찰 및 개발하고자 하였다. 먼저 표현형 분리비가 3:1로 분석 되었기 때문에 불안정 웅성 회복 표현형은 단일 우성 유전자에 의해 결정된다고 결론을 내렸다. 그리고 유전자 맵핑을 통하여 불안정 불임회복 유전자좌와 가까이 연관되어 있는 마커 214.14MB-CAPS를 개발 하였다. 새로 개발된 마커와 양옆에 위치한 13T7-SCAR와 G16-SCAR 사이 annotation을 하였고 기능을 분석 하였다. CM334 (UCD10X) 유전체서열 정보에서 PPR 유전자와 GEX2 유전자 그리고 Peroxidase를 발견하였다. 우리는 맵핑에 사용된 마커들이 다양한 육종계통에 쓸 수 있는지 확인하기 위해 12종류의 안정한 불임계통과 11종류의 불안정한 가임계통 을 얻어 마커를 적용 하였다. 불안정 회복 유전자와 아주 가까운 마커인 214.14MB-CAPS를 적용 하였을 때 불안정 회복 계통에서 높은 정확도를 보였다 (91.6%). 이러한 결과를 통해 214.14MB-CAPS는 불안정 회복계통에서 넓게 사용할 수 있다고 판단하였다. 그러나 안정 불임계통에서 낮은 정확도를 보여 불안정 회복과 안정 불임계통을 구별하수 없었다. 그래서 여러 불임 회복 유전자와 연관된 마커를 적용한 유전자형을 기반으로 하여 계통도를 작성하였다. 그 결과 불임 계통과 불안정 계통들이 분류가

되는 것을 확인하였다.

학번: 2017-29073