저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

- 저작자표시. 귀하는 원저작자를 표시하여야 합니다.
- 비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.
- 변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.
A DISSERTATION FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Development of Pre-breeding Materials in
Capsicum spp. and Map-based Cloning of a Male
Sterile Gene, ms10^{35} in Tomato

고추에서 유용 유전 자원 개발 및 토마토 웅성 불임

ms10^{35}의 유전자 지도 이용 동정

February, 2014

HEE-JIN JEONG

MAJOR IN HORTICULTURAL SCIENCE
DEPARTMENT OF PLANT SCIENCE
THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY
Development of Pre-breeding Materials in *Capsicum* spp. and Map-based Cloning of a Male Sterile Gene, *ms10^35* in Tomato

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

BY
HEE-JIN JEONG

MAJOR IN HORTICULTURAL SCIENCE
DEPARTMENT OF PLANT SCIENCE

February, 2014

APPROVED AS A QUALIFIED DISSERTATION OF HEE-JIN JEONG FOR THE DEGREE OF DOCTOR OF PHILOSOPHY BY THE COMMITTEE MEMBERS

CHAIRMAN

Hee Jae Lee, Ph.D.

VICE-CHAIRMAN

Byoung-Cheorl Kang, Ph.D.

MEMBER

Jin Hoe Huh, Ph.D.

MEMBER

Nam Chon Baek Ph.D.

MEMBER

Chang Sik Oh, Ph.D.
Development of pre-breeding materials in *Capsicum* spp. and map based cloning of a male sterile gene, *ms10* in tomato

HEE-JIN JEONG

Department of Plant Science, Seoul National University

Abstract

Tomato and pepper belong to the Solanaceae family and economically important vegetable crops worldwide. The main objectives of tomato and pepper breeding are developing high yield, pest resistant, and quality cultivars. Narrowed genetic diversity in commercial lines due to under-utilization of germplasm is a serious problem in tomato and pepper breeding. Therefore, discovering of allelic diversity and characterization of useful traits are essential for breeding. In the first chapter, species identification markers for *Capsicum* spp. were developed. Identification of *Capsicum* species has always been controversial because of
ambiguous morphological behaviors. Moreover, most previous studies for species identification using molecular markers also had disadvantages that were low reproducibility and unsuitability for handling large quantities of samples. Here, SNP markers derived from conserved ortholog set II (COS II) and the Waxy gene were developed as candidate maker set for species identification. Furthermore, phylogenetic analysis supported that C. annuum, C. baccatum, and C. pubescens clades were clearly separated. Closely related species C. chinense and C. frutescens were divided into subclades, indicating that the two species are apparently different. C. chacoense, the most controversial species, was belonged to C. baccatum clade. These results provide the information of genetic relationship of Capsicum species and reliable species classification method. In the second chapter, allele mining was performed to find useful allelic variation of a virus resistance gene in natural germplasm pool and ethyl methanesulfonate (EMS) mutagenized population to enhance the genetic diversity of Capsicum species. A fast and accurate method was developed to mine the useful alleles in significant number of germplasm using high resolution melting (HRM) analysis. Of 248 pepper germplasm, 13 newly obtained allelic variations of eIF4E were identified and one accession of them, C03946 (pvrHRM13) from C. baccatum was shown to carry strong resistance to TEV-HAT virus. Furthermore, five more allelic variations from EMS-mutageneized M1 population were also identified. Therefore, this result provides the strength of genetic diversity in Capsicum spp.
and a useful tool for mining novel alleles. In the third chapter, a male sterile gene of tomato was characterized. Male sterility is the one of the most important characteristics for plant breeding. Here, the map-based cloning of the $ms10^{35}$ mutant of tomato was performed and a responsible gene encoding a bHLH transcription factor was discovered. With morphological defects and molecular evidence, $Ms10^{35}$ played a pivotal role for pollen development after meiotic stage. $ms10^{35}$ was firstly isolated gene in tomato discovered by map-based cloning among the pollen development related genes at early stage. This provides that the insight of molecular evidence for tomato pollen development and utilization of male sterility line to improve tomato breeding.

Keywords: allele mining, Capsicum, male sterility, map-based cloning, pollen development, potyvirus resistance, species identification, tomato

Student Number: 2008-21263
Contents

Abstract .......................................................................................................................... i
Contents ....................................................................................................................... iv
List of tables .................................................................................................................. ix
List of figures ................................................................................................................ xi
List of abbreviations ..................................................................................................... xiii

General introduction ................................................................................................... 1

Literature review ......................................................................................................... 6

Chapter I

Identification of Capsicum Species Using SNP Markers Based on High-Resolution Melting Analysis

Abstract ....................................................................................................................... 61
Introduction .................................................................................................................. 63

Materials and methods

Plant materials ............................................................................................................. 67
Morphological trait observation ................................................................................. 67
Chapter II

A Survey of Natural and EMS-induced Variations of eIF4E Using High-Resolution Melting Analysis in Capsicum Species
Plant materials ........................................................................................................99
Chemical mutagenesis ............................................................................................99
DNA extraction .......................................................................................................101
HRM analysis .........................................................................................................101
Sequencing and sequence alignment of eIF4E fragment ....................................102
Virus inoculation and disease resistant evaluation .............................................103

Results

Screening of pvrHRM variation in Capsicum germplasm using HRM ...........104
Nucleotide and amino acid changes in natural variants .........................105
Virus resistance test on new allelic variants ...........................................108
Generation of an EMS mutant population .............................................111
Application of HRM to mutation screening ........................................116

Discussion ..........................................................................................................119

References ..........................................................................................................124

Chapter III

Tomato Male Sterile Gene, ms10\textsuperscript{35} Regulates Tapetum Degeneration and Meiosis

Abstract .............................................................................................................131
Introduction........................................................................................................................................133

Materials and methods

Plant materials......................................................................................................................................137

Phenotypic observation and genetic analysis of *ms10^{35}*,.................................137

Scanning electron microscopic (SEM) analysis .................................................................138

Light microscopic analysis and transmission electron microscopic (TEM) observation................................................................................................................................138

DAPI staining analysis for meiotic process .......................................................................139

DNA extraction ........................................................................................................................140

Alignment of BAC clone and marker development using high-resolution melting (HRM) analysis ................................................................................................................................140

Fine mapping of *ms10^{35}*,........................................................................................................141

Total RNA isolation and RT-PCR ........................................................................................142

5’ RACE analysis for transcription start site survey .........................................................142

Phylogenetic analysis ..............................................................................................................145

Complementation of the *dyt1* mutant ...............................................................................145

Localization of expressed *Ms10^{35}* by in situ hybridization ........................................146

RNA transcriptome analysis by RNAseq ..........................................................................147

Results

*ms10^{35}* shows complete sterility with no ability to produce viable pollen ... 149

Pollen development of *ms10^{35}* is arrested at tetrad stages ................................. 149
Defects in degeneration of tapetal cells result in failure of pollen development in $msI0^{35}$ .................................................. 154

Meiosis of $msI0^{35}$ is arrested at anaphase I................................. 158

$msI0^{35}$ encodes a bHLH transcription factor............................. 160

Localization of the $MsI0^{35}$ gene by RNA in situ hybridization .......... 166

Complementation of male sterility with a full length $MsI0^{35}$ gene........ 169

Comparison of global gene expression between male sterile and fertile plants by RNAseq analysis.............................................. 173

Discussion.................................................................................. 186

References.................................................................................. 193

Abstract in Korean ..................................................................... 203
List of tables

Chapter I

Table 1. Capsicum species used in this study.................................................................68
Table 2. Morphological and agronomic traits and their description used in morphological characterization.................................................................69
Table 3. Species identification by morphological traits in Capsicum.................73

Chapter II

Table 1. Nucleotide and amino acid substitutions of eIF4E haplotypes in Capsicum germplasm, and frequencies of pvrHRM haplotypes.................100
Table 2. Effects of EMS on seed germination and fruit fertility of Yuwol-cho pepper........................................................................................................113
Table 3. Summary of EMS-induced mutation screening using HRM..............118

Chapter III

Table 1. Primers used for fine mapping.................................................................143
Table 2. Primer used for RT-PCR of ms10^{35} gene and its putative down-stream genes.................................................................144
Table 3. Genetic analysis of ms10^{35} gene using an F2 population derived from 2-517 and T-1082 ........................................................................162
Table 4. Summary of T1 transgenic plant analysis for complementation test.....172
Table 5. Transcription factors down-regulated in the ms10^{35} mutant anther......175

Table 6. Acyl lipid metabolism genes that are down-regulated in ms10^{35}........177

Table 7. Genes for cell wall modification/degradation down-regulated in 

\[ ms10^{35} \]

.................................................................178

Table 8. Transporter genes that are down-regulated in ms10^{35}.................181

Table 9. Genes encoding pollen wall or coat proteins down-regulated in 

\[ ms10^{35} \]

.................................................................182

Table 10. Known genes down regulated genes for pollen development in 

\[ ms10^{35} \]

.................................................................183
List of figures

Chapter I

Fig. 1. HRM profiles of markers developed from COSII........................................75
Fig. 2. Sequence alignment of C2_At5g04590 showing ten SNP positions ........78
Fig. 3. Whole Waxy gene sequence of each Capsicum species .......................79
Fig. 4. HRM analysis using the Waxy locus ......................................................80
Fig. 5. Phylogenetic tree showing genetic diversity of 31 Capsicum species ......82

Chapter II

Fig. 1. HRM analysis of allelic variation..........................................................106
Fig. 2. Examples of natural variations of pvrHRM analyzed with HRM analysis ..........................................................107
Fig. 3. Amino acid sequence alignment of pvrHRM alleles ..........................119
Fig. 4. Nucleotide sequence alignment of pvrHRM1 alleles in natural variation..... 110
Fig. 5. Virus accumulation in Capsicum germplasm containing various pvrHRM alleles ..................................................................................112
Fig. 6. EMS-induced nucleotide changes in pvrHRM1 .................................114
Fig. 7. Sensitivity of HRM analysis and screening of EMS-induced mutations in pvrHRM ............................................................................117
Chapter III

Fig. 1. Phenotypic analysis of ms10^{35} plant...............................................................150

Fig. 2. Anther development in male fertile flower and male sterile plant.........152

Fig. 3. Transmission electron microscopic analysis of tapetum development....156

Fig. 4. TEM analysis of nucleus development and DAPI staining for meiosis
observation..................................................................................................................159

Fig. 5. Map-based cloning of ms10^{35} gene ..........................................................163

Fig. 6. Neighbor-joining tree of Ms10^{35} gene and its homologues..............167

Fig. 7. Amino acid sequence alignment of Ms10^{35} and its homologues AtDYT1
and OsUDTI ................................................................................................................168

Fig. 8. Localization of ms10^{35} expression in the anther by RNA in situ
hybridization.................................................................................................................170

Fig. 9. Complementary test of ms10^{35} mutant by a full length Ms10^{35} gene.....171

Fig. 10. Molecular function of differentially down-regulated in ms10^{35} mutant.174

Fig. 11. Expression patterns of Ms10^{35} gene regulated by putative up- and down-
stream genes..............................................................................................................184

Fig. 12. Visualized expression patterns of up- and down-regulated genes at
different stages between male fertile and sterile plant.................................185
List of figures

Chapter I

Fig. 1. HRM profiles of markers developed from COSII.................................75
Fig. 2. Sequence alignment of C2_At5g04590 showing ten SNP positions .......78
Fig. 3. HRM analysis using Waxy locus ......................................................79
Fig. 4. Phylogenetic tree showing genetic diversity of 31 Capsicum species ......82

Chapter II

Fig. 1. HRM analysis of allelic variation......................................................106
Fig. 2. Examples of natural variations of pvrHRM analyzed with HRM analysis
......................................................................................................................107
Fig. 3. Amino acid sequence alignment of pvrHRM alleles .......................110
Fig. 4. Nucleotide sequence alignment of pvrHRM1 alleles in natural variation
......................................................................................................................111
Fig. 5. Virus accumulation in Capsicum germplasm containing various pvrHRM
alleles .............................................................................................................113
Fig. 6. EMS-induced nucleotide changes in pvrHRM1 ...............................116
Fig. 7. Sensitivity of HRM analysis and screening of EMS-induced mutations in
pvrHRM ........................................................................................................119

Chapter III
Fig. 1. Phenotypic analysis of ms10<sup>35</sup> plant ................................................................. 151

Fig. 2. Anther development in male fertile flower and male sterile plant ........ 153

Fig. 3. Transmission electron microscopic analysis of tapetum development .... 157

Fig. 4. TEM analysis of nucleus development and DAPI staining for meiosis observation ........................................................................................................... 161

Fig. 5. Map-based cloning of ms10<sup>35</sup> gene ................................................................. 164

Fig. 6. Neighbor-joining tree of Ms10<sup>35</sup> gene and its homologues .............. 168

Fig. 7. Amino acid sequence alignment of Ms10<sup>35</sup> and its homologues AtDYT1 and OsUDT1 ........................................................................................................... 169

Fig. 8. Localization of ms10<sup>35</sup> expression in the anther by RNA in situ hybridization ........................................................................................................... 171

Fig. 9. Complementary test of ms10<sup>35</sup> mutant by a full length Ms10<sup>35</sup> gene .... 172

Fig. 10. Molecular function of differentially down-regulated in ms10<sup>35</sup> mutant. 178

Fig. 11. Expression patterns of Ms10<sup>35</sup> gene regulated by putative up- and down-stream genes .............................................................................................. 186

Fig. 12. Visualized expression patterns of up- and down-regulated genes at different stages between male fertile and sterile plant ................................. 178
# List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFLP</td>
<td>Amplified fragment length polymorphism</td>
</tr>
<tr>
<td>BAC</td>
<td>Bacterial artificial chromosome</td>
</tr>
<tr>
<td>bHLH</td>
<td>Basic helix loop helix</td>
</tr>
<tr>
<td>CAPS</td>
<td>Cleaved amplified polymorphic sequences</td>
</tr>
<tr>
<td>CMS</td>
<td>Cytoplasmic male sterility</td>
</tr>
<tr>
<td>COS</td>
<td>Conserved ortholog set</td>
</tr>
<tr>
<td>CTAB</td>
<td>Cetyl trimethylammonium bromide</td>
</tr>
<tr>
<td>DAPI</td>
<td>4',6-Diamidino-2-phenylindole</td>
</tr>
<tr>
<td>eIF4E</td>
<td>Eukaryotic translation initiation factor 4E</td>
</tr>
<tr>
<td>EMS</td>
<td>Ethyl methane sulfonate</td>
</tr>
<tr>
<td>FDA</td>
<td>Fluorescent diacetate</td>
</tr>
<tr>
<td>GMS</td>
<td>Genic male sterility</td>
</tr>
<tr>
<td>HRM</td>
<td>High-resolution melting</td>
</tr>
<tr>
<td>NMU</td>
<td>N-Nitroso-N-methylurea</td>
</tr>
<tr>
<td>NJ</td>
<td>Neighbor joining</td>
</tr>
<tr>
<td>PCD</td>
<td>Programmed cell death</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PMC</td>
<td>Pollen mother cell</td>
</tr>
<tr>
<td>PMMoV</td>
<td><em>Pepper mild mottle virus</em></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>PPR</td>
<td>Pentatricopeptide repeat</td>
</tr>
<tr>
<td>RACE</td>
<td>Rapid amplification of complementary DNA ends</td>
</tr>
<tr>
<td>RAPD</td>
<td>Randomly amplified polymorphic DNA</td>
</tr>
<tr>
<td>RFLP</td>
<td>Restriction fragment length polymorphism</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcriptase polymerase chain reaction</td>
</tr>
<tr>
<td>SCAR</td>
<td>Sequence characterized amplified region</td>
</tr>
<tr>
<td>SEM</td>
<td>Scanning electron microscope</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SSR</td>
<td>Simple sequence repeat</td>
</tr>
<tr>
<td>TEM</td>
<td>Transmission electron microscope</td>
</tr>
<tr>
<td>TEV</td>
<td><em>Tobacco etch virus</em></td>
</tr>
<tr>
<td>TILLING</td>
<td>Targeting induced local lesions in genomes</td>
</tr>
<tr>
<td>TMV</td>
<td><em>Tobacco mosaic virus</em></td>
</tr>
<tr>
<td>TSS</td>
<td>Transcription start site</td>
</tr>
<tr>
<td>TSWV</td>
<td><em>Tomato spotted wild virus</em></td>
</tr>
<tr>
<td>TuMV</td>
<td><em>Turnip mosaic virus</em></td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated region</td>
</tr>
</tbody>
</table>
General introduction

Conservation and manipulation of genetic resources are important to find available accessions for crop breeding (Nass and Paterniani, 2000). However, most germplasms such as wild progenitor, landrace, and exotics cannot directly be used for crop breeding because of poor appearance, long time and huge economic support, lack of evaluation, and characterization database for breeders (Marshall, 1989). For this reason, breeders tend to develop a cultivar using the combination of limited number of elite lines which ensure to release new cultivar in short term without tedious and time consuming processes (Troyer and Rocheford, 2002). However, genetic diversity of elite inbred line has shown narrow genetic diversity as by-product (Goodman, 1999).

Pre-breeding is defined as a part of breeding activities for producing intermediate plant materials which can be utilized by breeders in near future through discovering new traits from exotic germplasms, developing useful breeding materials, and broadening the genetic diversity for crop breeding (Acosta-Gallegos et al., 2007). While current cultivated crops showed narrow genetic diversity, exotic germplasms have been reported to carry the tremendous amount of genetic diversity which has not been utilized ever for breeding. Recently, pre-breeding is believed to be a link between germplasm managements and this activity has being encouraged by the development of public-private partnership (Ferroni and Castle, 2011). For this reason, pre-breeding activities
have gradually been spotlighted because introduction of various useful traits from wild relatives enable to obtain success in plant breeding. For example, new and useful traits in wild type relatives such as yield, disease resistance, male sterility, and tolerance to biotic and abiotic stresses can be expected to improve crop quality (Sharma et al., 2013). The representative examples of pre-breeding activities are; 1) germplasm barcoding using molecular work (Jung et al., 2010), 2) discovering of useful allelic diversity (Rubio et al., 2009), and 3) characterization of useful traits for breeding (Martin et al., 1993).

Traditionally, species identification of germplasm with morphological characters has been problematic due to limitation of useful traits, and ambiguous scoring (Zewdie and Bosland, 2000). Various nuclear DNA sequences have also been used in species identification due to rapid evolution rate and abundant applicable sequence information (Liu et al., 2001; Senchina et al., 2003; Taberlet et al., 2007; Wolfe et al., 1987; Wray et al., 2003). According to modern classifications, there are five major domesticated species (*Capsicum annuum, C. frutescens, C. chinense, C. baccatum, and C. pubescens*) and over 22 wild species in genus *Capsicum* (Ballard et al, 1970). To identify *Capsicum* species, various DNA markers based on random polymorphic sequences have been used including restriction fragment length polymorphism (RFLP) (Lefebvre et al., 1993), randomly amplified polymorphic DNA (RAPD) (Votava et al., 2005), amplified fragment length polymorphism (AFLP) (Geleta et al., 2005), and simple sequence
repeat (SSR) (Kwon et al., 2005). In addition, specific polymorphic sequence, such as the nuclear Waxy gene has also been applied in Capsicum spp. (Jarret, 2008; Walsh and Hoot, 2001). However, most of previous research could not be answered correct identification of species in genus Capsicum. For example, closely related species, C. chinense and C. frutescens, were not easily distinguished, and involvement of C. chacoense has not been clearly established. Furthermore, these markers have not been utilized for handling large quantities of germplasm. However, with the development of advanced methodology, single nucleotide polymorphism (SNP) and indels containing the most abundant information could be applied for genotype discrimination, and identification of cultivars using SSRs and/or SNPs (Jung et al., 2010; Mackay et al., 2008; Wu et al., 2008). In chapter I, an effective method was developed to identify Capsicum species using nuclear sequence based markers.

Conservation and management of genetic diversity are urgently needed (Johnson, 2008), but exploring novel alleles in a large germplasm collection is like finding a needle in a haystack. Therefore, a high-throughput method is required to apply for identifying novel alleles so called allele mining (Kumar et al., 2010). Due to powerful ability to detect SNP and easiness to utilize them, high-resolution melting (HRM) analysis has been spotlighted as a high-throughput method to SNP (Wittwer et al., 2003). Furthermore, HRM analysis has been
proved to be a suitable method for allele mining in plants (Chagne et al., 2008; Croxford et al., 2008; Lehmensiek et al., 2008).

Construction of mutant population induced by chemical mutagen, such as ethyl methane sulfonate (EMS) is an alternative way to strength the genetic diversity. Mutant populations have been generated for mutation breeding and reverse genetics in various crops, including wheat (Dong et al., 2009), tomato (Menda et al., 2004; Minoia et al., 2010), and rapeseed (Stephenson et al., 2010; Wang et al., 2008c). Generating a unique allelic variation by artificial mutant from existing allele increases the probability to obtain success in crop breeding. In chapter II, allele mining for the eIF4E gene, which is responsible for potyvirus resistance in Capsicum, was performed using HRM analysis.

Genic male sterility (GMS) has widely been used for production of F1 hybrid in tomato (Georgiev 1991; Kaul et al., 1988). The ms1035 mutant of over 50 GMS lines was considered as a useful male sterile line because of lack of ability to produce pollens (Georgiev, 1991; Kumar and Singh, 2004). In previous studies, the ms1035 mutant showed morphological defect caused by breakdown in meiosis at the early stage of pollen development (Segui-Simarro and Nuez, 2007; Zamir et al., 1980). Furthermore, the genetic analysis demonstrated that the ms1035 was located on the long arm of chromosome 2 and highly linked to peroxisome-2,3 (Tanksley et al., 1992). However, precise cytological and molecular evidence how ms1035 mutant is generated has not been discovered yet.
During pollen development, large numbers of genes are involved in microsporegenesis and microgametogenesis of floral organs (Ma, 2005). Especially, cytological changes occur in sporophytic and gametophytic cells to support producing normal pollens. Moreover, complex coordinated interactions among responsible genes are triggered for operating the mechanism to support exact temporal and spatial gene expressions. Meiotic process and programmed cell death (PCD)-triggered tapetal cell degenerations are the representative examples. Pollen development pathway has been studied mostly in model plants. Since initiation of floral bud, several bHLH transcription factors are required for operating the cascade of pollen development. For example, *DYT1* in *Arabidopsis* (Zhang et al., 2006) and *UDT1* in rice (Jung et al., 2005) encoding bHLH transcription factors were considered as a key regulator for meiosis and tapetum development. In chapter III, map-based cloning of the *ms10* mutant was performed and the function of isolated gene was characterized.

In this dissertation, three different experiments were performed focusing on the development of molecular methods for efficient pre-breeding and its application to genetic resources. Furthermore, the functions of male sterility gene were characterized. The results from this dissertation would provide the basis of molecular breeding method for developing of pre-breeding materials in *Capsicum* species and better understanding of pathway for pollen development in tomato.
Literature review

Species identification of *Capsicum* species

The history of *Capsicum* species

The genus *Capsicum* is composed of around 30 species including five different cultivated species (Pickersgill, 1997). The origin of *Capsicum* species is presumed to be along the alpine region of South America, because genetic diversity of wild pepper species is the most complicated and concentrated in this region (Pickersgill, 1971). During domestication, *Capsicum* species were gradually migrated and differentiated into plain areas near Amazon basin, and subsequently, five species (*C. annuum*, *C. chinense*, *C. frutescens*, *C. baccatum*, and *C. pubescens*) were domesticated at least two separated sub-center of origin, the Caribbean coast and the Andean region, respectively (Pickersgill, 2007). Because of the geographical distinct of South America, domestication of *Capsicum* species was complex and dynamic evolutionary process. However, the most dramatic divergence of *Capsicum* species for breeding occurred after the voyage of Christopher Columbus (Djian-Caporalino et al., 2007; Zewdie and Bosland, 2000). A thousand of landraces and cultivars have been exploited by farmers and breeders to improve exotic germplasm. Since introduction to the Eurasia continent, *C. annuum* has been become the most domesticated species among five cultivated species.
The most curious questions about the species of *Capsicum* are: 1) where they were domesticated and 2) how they are closely related between wild type and its cultivated taxa. In a view of combination with morphological traits and molecular markers, genus *Capsicum* has been divided into three main lineages, *C. annuum*, *C. baccatum*, and *C. pubescens* complexes (Onus and Pickersgill, 2004). Especially, the three different lineages were evolved and domesticated independently from geographically isolated and different environmental conditions.

Considering intercross inability between other *Capsicum* clades (Onus and Pickersgill, 2004) and significant morphological differences such as seed and flower color, *C. pubescens* complex seemed to be formed as a distinct clade at first and separated from other *Capsicum* ancestral clade. Especially, *C. pubescens* has primarily been cultivated in South America and rarely spread to outside (Eshbaugh et al., 1983; McLeod et al., 1982), indicating that this species clade has strictly been isolated from other species. Recently, *C. eximium*, *C. cardenasii*, and *C. tovarii* were reported to be in *C. pubescens* complex as wild type relatives (Ince et al., 2010). However, which species is the exact wild progenitor of *C. pubescens* is still questionable.

*C. baccatum* clade was mainly cultivated in lowland of South America and formed another independent complex. For example, *C. baccatum* var. *pendulum* was cultivated species and *C. baccatum* var. *baccatum* and *C. praetermissum* were
constituted in *C. baccatum* complex as wild progenitor (Ince et al., 2010; Walsh and Hoot, 2001). On the other hands, involvement of *C. chacoense* in *C. baccatum* has been controversial so far (Tam et al., 2009; Walsh and Hoot, 2001), because the morphological characteristics of *C. chacoense*, such as corolla color and seed color, were similar with those of *C. annuum* clade. However, recent relationship results using molecular markers represented that *C. chacoense* appeared to be in *C. baccatum* clade (Ibiza et al., 2012; Nicolai et al., 2013; Walsh and Hoot, 2001).

Finally, the most complicated clade, *C. annuum* clade contains three different cultivated species and numerous wild relatives (Onus and Pickersgill, 2004). Domestication of *C. annuum* clade seemed to occur in the Caribbean basin of South America (Pickersgill, 1971) and transferred to Europe and Asia through several routes. *C. annuum* has widely been cultivated all over the world, while *C. chinense* has been spread to lesser extent than *C. annuum*. In case of *C. frutescens* only Tabasco has been cultivated in Louisiana, USA. Besides cultivated species, *C. galapagoense*, *C. glabriusculum*, and *C. aviculare* have been reported as wild type progenitors of *C. annuum* clade. Especially, three cultivated species in *C. annuum* clade could not be easily identified due to primitive phenotyping such as flower color, fruit shape, and other morphological behaviors. In addition, some reports argued that *C. chinense* and *C. frutescens* were considered as same species or
should be integrated because of their cross-ability and morphological similarity (McLeod et al., 1982; Pickersgill, 1971; Walsh and Hoot, 2001).

**Identification of species using morphological traits**

*Capsicum* species has been identified based on morphological differences such as seed color, flower color, fruit shape, presence of calyx, and cross-compatibility (Ince et al., 2010). In the initial species identification, two distinct lineages were grouped as white and purple flower clades (Ballard et al., 1970), but this classification standard became useless soon because white-colored *Capsicum* species such as *C. annuum, C. chinense,* and *C. praetermissum* also showed purple flowered lines. To compensate the defect, corolla pattern was different and therefore, this has widely been applied as the key point for species identification instead of flower color (Jeong et al., 2010; Onus and Pickersgill, 2004; Pickersgill, 1988). While *C. annuum, C. chinense,* and *C. chacoense* showed white corolla color, *C. frutescens* exhibited greenish corolla and *C. baccatum* had yellow spot on the corolla. In addition, unique black seed color of *C. pubescens* could be used for identification of *C. pubescens* species (Pickersgill, 1997). Filament color, toothed calyx, and number of flower per node were also used for species identification.

However, species identification using morphological traits was frequently ambiguous, because most of characteristics could be affected by environmental
factor (Ince et al., 2010). In addition, due to cross-compatibility except C. pubescens, mixture of genetic background could make genetic variation complex to identify species (Onus and Pickersgill, 2004). Therefore, the trend of Capsicum species identification has been changed to determine species using molecular markers recently.

Application of molecular markers for Capsicum species identification

To precisely figure out the identification of Capsicum species, molecular markers have been applied to overcome the disadvantages of morphological traits based identification. Several types of molecular markers such as AFLP (Geleta et al., 2005; Ibiza et al., 2012), RAPD (Ince et al., 2010), and SSR markers (Kong et al., 2012; Nicolai et al., 2013; Shirasawa et al., 2013; Yi et al., 2006) have been applied for assessment of relationship and diversity within and between species in Capsicum. Recently, SNP markers have also been developed using conserved gene sequences such as conserved ortholog set (COS) II gene (Wu et al., 2006) and Waxy (Walsh and Hoot, 2001).

Since cytoplasmic genomes are inherited maternally and conserved well from generation to generation, chloroplast sequence-derived markers have been used for species identification (Ryzhova and Kochieva, 2004; Walsh and Hoot, 2001). Encouraged by accumulation of genome sequence and development of analysis method, nuclear DNA-derived markers have also been adopted for
species identification (Jeong et al., 2010; Walsh and Hoot, 2001; Wu et al., 2006). Especially, since nuclear DNA can be exchanged by cross pollination and rates of natural mutation is higher than cytoplasmic DNA, more detailed genetic analysis can be obtained.

Combined all the results performed with several types of molecular markers, genetic relationship of *Capsicum* species was divided into three groups, which almost corresponded with clustering result with morphological characteristics (McLeod, 1982; Ibiza et al., 2012; Ince et al., 2010; Jeong et al., 2010; Walsh and Hoot, 2001). Most unique data in genetic analysis were that 1) *C. chinense* and *C. frutescens* showing ambiguous morphological relationship because of frequent interspecific crosses actually closely related but clearly separated (Ibiza et al., 2012; Jeong et al., 2010; Walsh and Hoot, 2001), 2) although three wild type relatives showed distinct morphological behaviors in *C. pubescens* clade, *C. pubescens* and three different wild relatives, *C. eximum*, *C. cardenasii*, and *C. eximum* were grouped together (Ibiza et al., 2012; Pickersgill, 1991; Walsh and Hoot, 2001), and 3) involvement of *C. chacoense*, which was always controversial issue in *Capsicum*, was confirmed that it belongs to *C. baccatum* clade (Ibiza et al., 2012; Jeong et al., 2010; Walsh and Hoot, 2001).

*Species identification and pepper breeding*

Description of species identification is a major task to utilize germplasm to
conventional pepper breeding. Many Solanaceous crops have been improved by means of interspecific hybridization (Pickersgill, 1997). Especially, the introgression of disease resistance genes from wild relatives to *C. annuum* elite breeding line could attribute crop improvement. *Tobacco mosaic virus* (TMV) resistance in *C. chinense* (*L*3) and *C. chacoense* (*L*4) was successfully introduced into *C. annuum* (Berzal-Herranz et al., 1995; Boukema, 1980; De la Cruz et al., 1997). In addition, *Tomato spotted wild virus* (TSWV) resistance from *C. chinense* (Boiteux et al., 1993), resistance to bacterial leaf spot disease (*Bs*2) from *C. chacoense* (Hibberd et al., 1987), and potyvirus resistance from *C. chinense* to Tabasco breeding (Greenleaf 1956; Greenleaf et al., 1970) were also excellent examples. These have been successfully introduced into *C. annuum* commercial line worldwide. However, introgression of useful genes from wild progenitor to cultivated line is a quite difficult and tedious task because cross incompatibility barrier exists in *Capsicum* species. To solve this problem, interspecific cross-bridge can be accepted to efficiently introduce useful allele (Zhang et al., 2012). Therefore, to better understand relationship among species, cross-compatibility, utilization, and introgression of exotic germplasm, species identification should be processed.

**Utilization of useful allele for plant breeding**

*Reverse genetics using allele mining*
Reverse genetics is an alternative process of forward genetics to associate certain allelic difference of known gene and phenotypic variation (Hardy et al., 2010). In comparison with forward genetics, reverse genetic can be reduced endeavor and time to develop population for gene of interest and therefore, it has been widely used for characterization of gene function. Reverse genetics has been accompanied with discovery of allelic variation process so called allele mining (Barkley and Wang, 2008). Especially, since genetic reservoir of cultivated crops has been not enough in current plant breeding for yield, biotic, abiotic stresses, and other disease resistance, useful allele mining from wild type relatives, allied species, and artificially induced mutant population could exploit diversity to overcome genetic bottleneck (Glaszmann et al., 2010). Subsequently, the combination and accumulation of useful alleles for various characteristics enhanced the genetic potential and crop improvement for agronomic traits (Barton and Keightley, 2002; Morgante and Salamini, 2003; Slafer et al, 1994).

Recently, newly obtained allelic variations have been adopted for improvement of various crops and vegetables breeding, including rice, wheat, potato, and tomato. In rice, the mutant allele of prostrate growth habit (PROGI) enhanced higher yield induced by erect growth and number of grains (Tan et al., 2008). Furthermore, GW2 and SW5 which were responsible for seed size (Lu et al., 2013), rice grain-filling gene GIF1 (Wang et al., 2008a), and other yield related genes such as Gn1 (Ashikari et al., 2005), sd1 (Kovi et al., 2011), and qSH1
(Konishi et al., 2006) have been utilized wild type accessions to improve yield of cultivated rice crop.

Second example of allele mining is finding disease resistance genes. Especially, there are several examples for introduction of resistance alleles from wild type progenitors to overcome the occurrence of new virulent diseases induced by evolutionary driving force and environment change. Representative examples are; 1) bacterial leaf blast resistance gene of rice, Xa21, from Oryza longistaminata to O. sativa (Khush and Kinoshita, 1991), 2) brown stem rot resistance of soybean like Rbs1 and Rbs3 from soybean germplasm (Klos et al., 2000), 3) potato late blight disease resistance gene of potato, Rpi-blb2, from Solanum stoloniferum (Wang et al., 2008b), 4) Tomato yellow leaf curing virus resistance gene of tomato, Ty-1 to Ty-5, from S. chilense to S. lycopersicum (Hutton et al., 2012), and 5) potyvirus resistance gene of pepper, pvr1, from C. chinense to C. annuum (Rubio et al., 2009). In addition, all these useful genes have been introgressed into cultivated species through consecutive backcrossing process using marker assisted selection to facilitate plant breeding.

Third example for allele mining is the improvement of crop quality. For example, development of imidazolinone herbicides resistant crops in maize (Gerwick et al., 1993), wheat (Rauch and Thill, 2002), rice (Carlson et al., 2002), and rapeseed (Shaner et al., 1996) were developed from mutagenized population using allele mining.
Construction of mutagenized population to exploit new allelic diversity

Since naturally occurred plant germplasm has narrow genetic diversity and insufficient various allelic variations (Kumar et al., 2010), securement of wider sources in heritable variation is one of the most urgent missions for crop improvement (Mba, 2013). In addition, exploitations of genetic resources for crop breeding induced by artificial mutagens have been obtaining popularity in plant breeding. The common types of induced mutation for plant breeding are: 1) single base substitutions caused by chemical mutagen such as EMS and N-nitroso-N-methylurea (NMU), 2) alteration and/or rearrangement of chromosome structure induced by physical mutagens such as ultra violet light, gamma-ray and X-ray radiation (Caldwell et al., 2004; Close et al., 2009; Greene et al., 2003; Mba, 2013; Menda et al., 2004; Sato et al., 2006; Suzuki et al., 2008; Till et al., 2004). Since induced mutagenesis has been proved as a safe, powerful, and inexpensive strategy, mutagenized populations can be alternative to alleviate the extreme loather to genetically modified organism (GMO) (Parry et al., 2009). In transformation recalcitrant plants such as pepper, mutagenized population can be a great choice to study the function of gene of interest and to generate new genetic resources. With the accumulation of knowledge and advanced technologies for biology, genetic mutagenesis has been a general trend to identify genes, to study the function of genes, and to utilize mutants for crop improvement (Kumar et al., 2010; Mba, 2013)
Allele mining using molecular technique

One of the most important prerequisites for reverse genetics is that the sequence of interest or the homologues sequence in related species should be known. Genome sequence data, gene annotation, various omics, molecular marker platform, and bioinformatic tools have been developed to support allele mining (Kumar et al., 2010). Considering these, allele mining in reverse genetic has been spotlighted for crop breeding (Barkley and Wang, 2008). Especially, two major approaches have widely been utilized for identification of allelic variation of gene of interest.

The most popular method, targeting induced local lesions in genomes (TILLING), detects sequence variation in gene of interest by enzymatic cleavage of mismatched sequence between wild type and mutant individuals after polymerase chain reaction (PCR) hybridization (Till et al., 2004). Since firstly developed in Arabidopsis mutant population, TILLING method has widely been applied to detect mutations in various crops including rice (Suzuki et al., 2008; Till et al., 2007; Wu et al., 2005), maize (Till et al., 2004), sorghum (Xin et al., 2008), barley (Caldwell et al., 2004; Talame et al., 2008), soybean (Cooper et al., 2008), tomato (Gady et al., 2009; Minoia et al., 2010; Okabe et al., 2011), and rapeseed (Harloff et al., 2012; Wang et al., 2008c). Therefore, reverse genetics could obtain great success to identify diverse allelic variation using the combination between genetic mutagenesis and TILLING. Furthermore, modified
TILLING technique, Eco-TILLING also showed noticeable progress to detect spontaneous mutations in germplasm (Barkley and Wang, 2008; Till et al., 2007). However, PCR-based TILLING has challenges such as technical expertise, complexity, less efficiency, and time.

Development of analyzing method contributed the process of allele mining. For example, DNA sequencing analyzer such as Li-Cor 4300 using fluorescent labeling (Colbert et al., 2001), bioinformatics tools such as CODDLE analysis, and a high-throughput mutation/polymorphism discovery technique such as KeyPoint™ platform (Rigola et al., 2009) were established to discover mutations. Moreover, HRM analysis could identify SNP mutations depending on detection of the level of decreased fluorescent dye during thermal denaturation and applied to various crops for allele mining (Gady et al., 2009; Hofinger et al., 2009; Lochlainn et al., 2011)

**Meiocyte and tapetum development of pollen development**

*General pathway of pollen development at early stage in model plant*

The reproductive processes in plants such as flowering, pollination, and fruit set are the most important development in a whole plant life (Wilson and Zhang, 2009). Especially, phanerogamy produces pollen through meiosis process to maximize the genetic diversity by pollination (Deveshwar et al., 2011). During production and release of pollen, a highly organized pathway is required to control
the temporal and spatial expression of pollen development-related genes (Tiang et al., 2012; Zhou et al., 2011). For example, through transcriptomic analysis in Arabidopsis, expression of around 15,000 genes was presumed during anther and male gametophyte development (Honys and Twell, 2004). In addition, during pollen development, sporophytic cells have proven to be cooperated with gametophytic tissues to support pollen maturation by transferring nutrients and materials (Jia et al., 2008; Yang et al., 2003, 2005; Zhang et al., 2006). Although complex cellular and molecular processes are involved in pollen development, understanding how sporophytic and gametophytic tissues are interacted each other has still not been understood.

Since the initiation of stamen development in a floral bud, microgametogenesis in reproductive cells underwent with a series of developmental events such as primodium cell differentiation, anther cell differentiation, meiosis and mitosis, and subsequently pollen maturation (Brukhin et al., 2003; Koltunow et al. 1990; Smyth et al. 1990; Wilson and Zhang, 2009; Zhang et al., 2011). In Arabidopsis, pollen development has been identified and divided into 14 morphological stages (Ma, 2005; Sanders et al., 1999). During stage 1 to 5, archesporial cell differentiations, followed by the division of primary parietal and sporogenous cell from anther cell layer occurred (Chang et al., 2011; Zhang et al., 2006). Continuously, primary parietal cells are divided into secondary parietal cells and subsequently, generate the tapetum, the middle cell
layer, and the endothecium, respectively. In conclusion, archesporial cells are differentiated into five different cell layers, pollen mother cell, tapetum, middle cell layer, endothecium and epidermis from innermost to outermost (Chang et al., 2011; Wilson and Zhang, 2009).

After differentiation of anther cell layers, dramatic morphological changes are observed in the five different cell layers (Sorensen et al, 2003; Yang et al., 2007; Zhang et al., 2006). Pollen mother cells proceed into tetrads through meiosis (Nonomura et al., 2004, 2006, 2007; Yi et al., 2012; Zhang et al., 2006; Zhou et al., 2011). Meiocytes seem to contact with the tapetal cell during meiosis (Canales et al., 2002; Yang et al., 2003; Zhang et al., 2006; Zhao et al., 2002). Tapetal cells surround and protect meiocytes by secreting of callose during meiosis. Continuously, tapetal cells are condensed and degenerated by PCD to deliver tapetal cell fragments toward microspore for synthesizing of sporopollenin wall surrounding primexine of pollen. Abnormal function of meiocyte and tapetum causes loss of male fertility but little or no clue has been discovered (Li et al., 2006; Sorensen et al., 2003; Yi et al., 2012; Zhang et al., 2006).

At late stage of pollen development, mitotic division generating trinucleate, exine formation, and dehiscence are required to produce normal pollens (Hu et al., 2011; Li et al., 2011; Vizcay-barrena and Wilson, 2006). Microspores accumulate starch and lipid, and simultaneously divided into generative and sperm cells. Sporophytic cells surrounding microspore are completely degenerated. Exine
formation and patterning composing of nexine, baculae, and tectum are completely finished by sporopollenin accumulation onto primexine matrix. Finally, stomium degeneration followed by pollen dehiscence is initiated at the last stage of pollen development. At this time, additional PCD is progressed in stomium cell which is a connecting cell between two locules (Gorguet et al., 2009).

Significant progresses on pollen development have also been elucidated in rice. Even though the appearance of cell structure and subtle difference in pollen development are existed, basic mechanism of pollen production has a common denominator among crops such as initiation of cellular differentiation, tapetum degeneration, and pollen maturation (Ma, 2005; Zhang et al., 2011). In addition, analysis of significant numbers of genes has shown the high homologue relationship, indicating that pollen development pathway may be well conserved in higher plants. This information would broaden the knowledge to other crops like tomato to better understand the mechanism of pollen development.

Molecular evidence and expression of related genes during meiosis

Meiotic division is comprised of highly conserved dynamic processes in higher plants (Tiang et al., 2012). Several key genes have been identified to have essential roles for DNA replication, chromatid cohesion, chromosome condensation, pairing, and synapsis (Chang et al., 2011; Zhang et al., 2011). To elucidate the role of component genes, meiotic mutants have been identified at the
genetic and cytological levels in not only *Arabidopsis* (Armstrong et al., 2002; Hamant et al., 2006; McCormick, 1993; Yuan et al., 2009) but also other various crops including rice (Nonomura et al., 2004, 2006, 2007; Wang et al., 2010), maize (Pawlowski et al., 2009), and tomato (De Storme and Geelen, 2013). For example, *HOMOLOGOUS PAIRING ABERRATION IN RICE MEIOSIS1* series, *PAIR1/PAIR2/PAIR3* (Armstrong et al., 2002; Nonomura et al., 2004, 2006; Yuan et al., 2009), and their homologous genes of *Arabidopsis* are required for homologous chromosome paring and synapsis during prophase I in meiosis. In addition, *MEL1* is involved in chromosome condensation at leptotene stage and controlling the size and number of pollen mother cells (Nonomura et al., 2007). *ZEPI* and its homologue gene of *Arabidopsis*, *ZYP1*, played a role for normal formation of synapsis at early prophase I and chromosomal segregation at anaphase I (Higgins et al., 2005; Wang et al., 2010). *MPS1* (*Multipolar spindle 1*) was proposed to play essential role for organization of spindle and chromosomal segregation at metaphase I (Jiang et al., 2009). *PSSI* (*Pollen semi-sterility1*) encoding a kinesis-1 like protein is necessary for meiotic chromosomal pulling by two spindles at anaphase I and II, respectively (Zhou et al., 2011). Furthermore, several novel proteins were required for chromatid cohesion and bivalent formation during meiosis. *Arabidopsis SWII* and rice *REC8* were essential for the successful transmission of chromosome from interphase S to meiosis II (Mercier et al., 2003; Shao et al., 2011). In these mutants, meiocyte underwent abnormal
chromosome division and subsequently, showed meiocyte degradation. In addition, chromatid cohesion was also important to locate other essential proteins for meiosis such as *PAIR2*, *PAIR3*, and *ZEPI* indicating that homologues pairing and synapsis were established by the formation of chromatid cohesion (Shao et al., 2011).

Synthesis of sister cells appear to depend on the normal function of tapetum. For example, tapetum-specific expressed genes such as *EMS1/EXS*, *TDF1*, and *DYT1* are required for cytokinesis at anaphase I after chromosomal segregation (Canales et al., 2002; Yang et al., 2003; Zhang et al., 2006; Zhao et al., 2002). Therefore, tapetum tissues are also involved in normal meiotic division suggesting that sporophytic and gametophytic cells coordinate and support meiosis by cell to cell communications. However, precise mechanism and function of each cell have been remained unknown.

*Genetic approach to elucidate tapetum development and PCD*

The process of tapetum development from cell differentiation to PCD triggered degeneration has been reported in various crops as well as model plants (Chang et al., 2011; Parish and Li, 2010; Wilson and Zhang, 2009; Zhang et al., 2011). Same as meiotic process, cytological and genetic developments of tapetum are identical between monocot and dicot plants. Tapetum development-related genes also have similar structure and function. The critical molecular support for
tapetum development and function are: 1) cell to cell signaling between sporophytic and gametophytic cells mediated from EMS1/TPD1 and 2) transcriptional regulation to function tapetum identification, differentiation, and degeneration controlled by DYT1, AMS, and MS1.

Tapetum is generated through three cell divisions from second layer of differentiated archesporial cell controlled by SPL/NZZ (Schiefthaler et al., 1999; Yang et al., 1999). Continually, EMS1/EXS and their ligand TPD1 act for tapetal specification (Canales et al., 2002; Yang et al., 2003; Zhao et al., 2002). In addition, a number of leucine-rich repeat receptor like protein kinases such as BAMI/BAM2 (Hord et al., 2006), SERK1/SERK2 (Albrecht et al., 2005; Colcombet et al., 2005), and RPK2 (Mizuno et al., 2007) have functionally redundant for cell fate identification of tapetum. Moreover, glutaredoxins such as ROXY1/ROXY2 (Xing and Zachgo, 2008) were also necessary for biochemical reactions during tapetum development. Common phenotypic defects of these mutants were abnormal structure, or absence of tapetal and middle cell layer, indicating that most of these genes are involved in differentiation, maintenance, specification, and normal function of tapetum at the early stage.

Homologous genes of Arabidopsis have also been discovered in rice. MSP 1 (Multiple sporocyte 1) is the homologues of EMS1/EXS and also regulates the sporocyte cells division (Nonomura et al., 2003). The phenotypic defect of msp1 mutant resembled that of ems1/exs. Moreover, like TPD1, a ligand protein of
EMS1/EXS, OsTDL1A has been also demonstrated (Zhao et al., 2008). These genes showed similar function, gene structure, and phenotypic aberrant in each mutant, supporting that pollen development process is conserved in various crops.

After formation of tapetum surrounding sporogenous cells, tapetal cells are breakdown by apoptosis-like PCD (Parish and Li, 2010). Sacrificial process of sporophyte tissues contributes microspore development (Balk and Leaver, 2001; Papini et al., 1999; Wang et al., 1999). Numerous transcription factors and receptor genes have also been identified to activate and repress tapetum PCD (Parish and Li, 2010; Wilson and Zhang, 2009; Zhang et al., 2011). Specifically, a transcriptional cascade constituted by DYT1, AMS, and MSI genes is required for normal tapetum function and viable pollen production (Sorensen 2003; Vizcay-Barrena and Wilson, 2006; Xu et al., 2010; Zhang et al., 2006). Among them, DYT1 is the most upstream gene for tapetum development and regulates numbers of genes expression. In dyt1 mutant, tapetum showed hypertrophic and vacuolated morphology (Zhang et al., 2006). AMS and MSI genes were extremely down-regulated. Using a total RNA extracted from anther, transcriptome was analyzed to compare expression level between normal and dyt1 mutant lines (Feng et al., 2012). In dyt1 mutant, lipid binding transporter, hydrolase activity, exine wall synthesis, and phenylpropanoid pathway genes as well as several transcription factors were greatly down-regulated. Therefore, DYT1 might have function for key regulator for pollen development. Aborted microspores (AMS) are the results
of the downstream and counterpart of DYT1 (Feng et al., 2012). Furthermore, AMS interacted with ASHR3 in vivo, a SET domain protein, which acts an epigenetic regulator (Thorstensen et al., 2008). In ams mutant, callose deposition, and initial exine formation were normal, although tapetum became vacuolated and prematurely enlarged, tetrad formation (Sorensen et al., 2003). Microspores in ams mutant degenerated and never produced normal pollens. MS1 gene, which is expressed in tapetum from late tetrad stage to free microspore stage, is downstream of DYT1 (Ito and Shinozaki, 2002; Vizcay-Barrena and Wilson, 2006; Zhang et al., 2006). Especially, MS1 controlled the tapetal cell degradation and therefore, contributed tapetum function and pollen wall deposition as a transcriptional activator (Alves-Ferreira et al., 2007; Ito et al., 2007; Yang et al., 2007). In msl mutant, tapetal PCD were delayed and impaired (Vizcay-Barrena and Wilson, 2006). Tapetum was vacuolated and swollen as same as mutant phenotype of DYT1 and AMS. Three genes, DYT1, AMS, and MS1 were transcription factors and regulated the numerous genes having important role for tapetum development (Zhu et al., 2011). In their mutants, the degeneration were delayed or failed due to repression of PCD related gene. Moreover, considering the protein-protein interaction results among the three genes, not only independent roles but also controlled interactions among three genes are required for tapetal development. Beside these three genes, MYB33 and MYB65 were essential role for early stage of pollen development (Millar and Gubler, 2005). MYB33 and MYB65
were proved to be downstream of *SPL* and *EMS1/EXS*, but were on the parallel pathway of *DYT1*, because those expressions were not reduced in *dyt1* mutant. In *myb33* and *myb65* double mutant, tapetal cell were vacuolated and meiocyte were degenerated similarly to *dyt1* mutant. However, sterility caused by double mutantation could be restored by low temperature, indicating that *MYB33* and *MYB65* were not perfectly necessary for pollen development.

In rice, three genes for normal tapetum development have been characterized (Jung et al., 2005; Li et al., 2006; Li et al., 2011). *OsUDT1*, the homologues of *DYT1*, played a role for initiation of tapetum development as same as *DYT1* (Jung et al., 2005). *OsTDR1* was the homologue of *AMS* and showed similar function of *AMS*. In rice, this tapetal specific gene regulated PCD by inducing cysteine protease (Niu et al., 2013). *OsTDR1* controls the synthesis of aliphatic fatty acids for pollen wall deposition (Li et al., 2006). *PTC1*, *MS1* homologue in rice was also identified (Li et al., 2011). *PTC1* regulated the initiation of PCD and pollen exine formation. The *ptc1* mutant was recovered by complementation of *Arabidopsis MS1* gene, strongly supporting that the function of homologous genes was evolutionary similar between monocot and dicot plants.

Owing to large number of mutant isolation, development of genome sequencing method, and transcriptome analysis, tapetum degeneration by PCD has been studied in rice. For example, *OsC6* (Zhang et al., 2010) and *OsCP1* (Li et al., 2006) were regulated by *OsTDR1* and controlled orbicules development and
pollen exine formation. *API5* was identified as an important gene for post-meiotic development (Li et al., 2011). *API5* was also under the control of *OsTDR1* and interacted with two DEAD-box genes and cysteine protease genes. *EAT1*, a positive regulator of PCD was identified as one of downstream genes of *OsTDR1* (Niu et al., 2013). *EAT1* regulated two independent aspartic proteases triggering PCD in tapetum.

**Pollen development pathway and key events in cultivated tomato**

Pollen development has been studied mainly in *Arabidopsis* and rice. In tomato, microgametogenesis from pollen mother cell differentiation to pollen maturation was described histologically and cytologically, and the developmental order was rearranged according to the correlation with bud length and morphological events (Brukhin et al., 2003). Tomato flower development was divided into 20 stages, and each sub-division was almost corresponded with histological event of *Arabidopsis* and rice. Furthermore, ultrastructural observation from meiosis to anther dehiscence was performed (Polowick and Sawhney, 1992, 1993, 1995).

In tomato, *ps-2* was isolated to be responsible for functional sterility which is important role for pollen dehiscence (Gorguet et al., 2006, 2009). Several microspore and tapetum specific-expressed gene or RNA fragments have also been studied in tomato. Four different genes, *TomA108*, *TGAS100*, *TGAS105*, and
LeGRP92 participating in early developmental stage have been identified. The TomA108 gene encoding a cysteine-rich protein with hydrophobic domain and a secretory signal was specifically expressed from early meiosis stage to free microspore in tapetum (Chen et al., 2006). Several cysteine-rich proteins have been reported to have important function for tapetum development in tomato (Aguirre and Smith, 1993; Chen and Smith, 1993). Even though precise gene characterization has not been identified, TomA108 appeared to have specific function in tapetum.

Two flower specific cDNAs, TGAS100 and TGAS105, were discovered in anther cDNA library (van den Huevel et al., 2002). TGAS105 encoded cysteine-rice extension-like protein, but TGAS100 does not have any homologue gene. However, they were activated by exogenous gibberellins, and expression of two genes was shown to be localized at tapetum cells. Therefore, TGAS100 and TGAS105 specifically expressed and regulated developmental process stimulated by gibberellins. LeGRP92, a glycine-rich protein was identified to be specifically expressed in microspore mother cell and tapetum (McNeil and Smith, 2010). Transgenic line having down-regulated LeGRP92 by antisense RNA showed alteration of pollen wall deposition and pollen viability, supporting that LeGRP92 may play a role for sporopollenin deposition and exine formation.

TA29 and its homologue gene were isolated in tobacco and tomato, respectively. Even though the precise function of TA29 has not been determined
yet, RNAi transgenic construct and TA29::barnase construct generated the male sterile plant showing aborted and agglomerated pollens (Mariani et al., 1990; Nawaz-ul-Rehman et al., 2007). Considering no destructive difference in growth habit and female reproductive tissues, TA29 was assumed to specifically involve in male gametophyte development. PCD-related genes have been demonstrated in PCD developing tissues of tomato (Senatore et al., 2009). By transcript analysis, a peptidase subfamily C1A, SGN-U321072 was presumed to be expressed in floral bud among multiple cysteine-rich proteinase genes. However, the function and expression pattern of SGN-U321072 have not been reported.

At late stage of pollen development, an endo-beta-mannannase, LeMAN5, was highly expressed only in anthers and pollens (Filichkin et al., 2004). By the localization and expression pattern, LeMAN5 was involved in pollen wall degradation. The pLAT52 gene of tomato encoding a heat-stable, glycosylated protein was also reported to be expressed in pollens (Muschietti et al., 1994). Using pLAT52-silenced line, abnormal hydrated pollen and inhibited pollen tube growth were observed, indicating that pLAT52 controlled pollen dehydration and germination after maturation. A tomato AGAMOUS-like 1 (TAGL1) mutant plant showed altered flower and fruit ripening pattern (Gimenez et al., 2010). Over-expression and RNAi construct showed altered flowering pattern and morphological change of fruit. In addition, TAGL1 was highly expressed at early development and gradually decreased, but accumulation of this gene was detected.
at the fruit maturation again. Considering the expression pattern and phenotypic aberrant of TAGLI, this gene participated in early and late pollen development.
References


Greenleaf, W. H., 1956. Inheritance of resistance to Tobacco etch virus in


(MPS1), a novel coiled-coil protein of *Arabidopsis thaliana*, is required for meiotic spindle organization. Plant J 59:1001-1010.


Nonomura, K., A. Morohoshi, M. Nakano, M. Eiguchi, A. Miyao, H. Hirochika, and N. Kurata. 2007. A germ cell specific gene of the ARGONAUTE family
is essential for the progression of premeiotic mitosis and meiosis during sporogenesis in rice. Plant Cell 19:2583-2594.


Rubio, M., M. Nicolai, C. Caranta, and A. Palloix. 2009. Allele mining in the pepper gene pool provided new complementation effects between pvr2-


Tam, S. M., V. Lefebvre, A. Palloix, A. M. Sage-Palloix, C. Mhiri, and M. A. Grandbastien. 2009. LTR-retrotransposons Tnt1 and T135 markers reveal


Xin, Z., M. L. Wang, N. A. Barkley, G. Burow, C. Franks, G. Pederson, and J. Burke. 2008. Applying genotyping (TILLING) and phenotyping analyses to
elucidate gene function in a chemically induced sorghum mutant population.

BMC Plant Biol 8:103.


Chapter I

Identification of *Capsicum* Species Using SNP Markers Based on High-resolution Melting Analysis

Abstract

Single nucleotide polymorphisms (SNPs) derived from both nuclear and cytoplasmic DNA sequences were developed to identify distinct species of *Capsicum*. Species identification was achieved by detecting allelic variations of these markers via high-resolution melting (HRM) analysis. COSII markers and the *Waxy* gene from the nuclear sequence as the SNP markers were used for detection of polymorphism by HRM analysis. A total of 31 accessions of *Capsicum*, representing six species, were analyzed using this method. While single markers were insufficient for identifying *Capsicum* species, combinations of all markers unambiguously identified all six of them. A phylogeny based on the SNP markers was consistent with the current taxonomy of *Capsicum* species. These observations demonstrate that the markers developed in this study are useful for rapid identification of new germplasm for management of *Capsicum* species.
Keywords: *Capsicum*, DNA barcoding, Germplasm, High-resolution melting (HRM) analysis, Single nucleotide polymorphism (SNP)
Introduction

Lack of genetic diversity in modern crops is a serious problem due to the concurrent vulnerabilities of crop plants to biotic and abiotic stresses (Hammer et al., 2003). The management and utilization of germplasms are important to increase and maintain genetic diversity in crop plants. Traditionally, germplasms have been classified based on morphological characters (Zewdie et al., 2004), but morphological identification can be problematic when useful traits are limited and the scoring of these traits is ambiguous. Instead, DNA markers have been used for rapid and accurate identification of plant species (Taberlet et al., 2007).

DNA sequences used for species classification should be conserved within each species, but variable enough to contain phylogenetic information ensuring proper assignment of species to taxonomic groups. The sequences should also retain priming sites conserved between species (Taberlet et al., 2007). DNA markers for species identification have been based on both organellar and nuclear genomes. Mitochondrial genes, such as coxI (CO1), have frequently been used for identification of animal species (Hebert et al., 2004). Mitochondrial genes of plants, however, are inadequate to classify species effectively because they evolve too slowly (Taberlet et al., 2007). Moreover, frequent rearrangements of mitochondrial DNA make it difficult to find universal intergenic sequences with this classification method (Palmer, 2000). Alternatively, highly conserved gene order and coding sequences, in addition to several highly variable non-coding
sequences of plastid DNA, enable efficient species identification and phylogenetic analysis in flowering plants (Taberlet et al., 1991). Chloroplast DNA sequences such as the trnH-psbA intergenic region (Kress et al., 2005), the trnL-trnF intergenic region (Taberlet et al., 1991), the trnL intron (Taberlet et al., 2007), and the matK gene have widely been used for this purpose. Various nuclear DNA sequences have also been used in species identification due to certain advantages. These advantages include the containment of predominant variations due to an evolution rate of 5 to 10 times higher than that of chloroplast and mitochondrial DNA, and abundant applicable loci scattered on different chromosomes (Liu et al., 2001; Senchina et al., 2003; Wolfe et al., 1987; Wray et al., 2003).

The genus Capsicum originated in the tropics and subtropics at least 8,000 years ago (Aguilar-Melendez et al. 2009). Capsicum is cultivated all over the world (Pickersgill, 1997). According to modern classifications, there are five major domesticated species (Capsicum annum, C. frutescens, C. chinense, C. baccatum, and C. pubescens) and over 22 wild species (Ballard, 1970). Although Capsicum species show enormous genetic and morphological diversity, most exotic germplasms are not being utilized efficiently for breeding due to preferential use of elite genetic resources and sexual barriers between species (Pickersgill, 1997).

Various DNA markers based on random polymorphic sequences have been
used to classify *Capsicum* species. These include restriction fragment length polymorphisms (RFLPs) (Lefebvre et al., 1993), randomly amplified polymorphic DNA (RAPD) (Votava et al., 2005), amplified fragment length polymorphisms (AFLPs) (Geleta et al., 2004), and simple sequence repeats (SSRs) (Kwon et al., 2005). Highly variable barcode sequences have recently been selected from the plastid genome and used for *Capsicum* species identification (Jarret, 2008; Walsh and Hoot, 2001). The nuclear *Waxy* gene was also found to contain polymorphic sequences between *Capsicum* species (Jarret, 2008; Walsh and Hoot, 2001). Most of these sequences, however, were unable to distinguish between the species included in the *C. annuum* clade: *C. annuum*, *C. frutescens*, and *C. chinense* (Bosland and Baral, 2007; Jarret, 2008). Moreover, polymorphisms in these markers can only be detected by direct sequencing, which is not suitable for handling large quantities of germplasm.

Recently, high-resolution melting (HRM) analysis was developed to detect SNPs and indels located on amplicons (Gundry et al., 2003; Ririe et al., 1997). In this method, amplified DNA is melted by a gradual temperature increase. Melting pattern differences are subsequently detected by subtle changes of fluorescence generated by double-stranded DNA binding dyes. Rapid and inexpensive detection of a broad range of SNPs and indels enables the application of HRM analysis for genotype discrimination (Lehmensiek et al., 2008) and genetic mapping (Chagne et al., 2008). Identification of cultivars using SSRs and/or SNPs
was also performed successfully via this technique (Mackay et al., 2008; Wu et al., 2008). In present study, an effective method was developed to identify Capsicum species using markers derived from nuclear. Furthermore, marker polymorphisms would be efficiently detected by HRM analysis.
Materials and methods

Plant materials

A total of 31 Capsicum accessions representing six species were obtained from the Asian Vegetable Research and Development Center in Tainan, Taiwan, and used to select species-specific markers (Table 1). Seeds were first treated with 2% sodium chlorate and 10% trisodium phosphate prior to being soaked in Petri dishes with water to induce germination. Germinated seeds were then grown in a growth chamber in trays containing sterilized soil. At the eight-leaf stage, the plantlets were transplanted to a greenhouse at Seoul National University, Suwon, Korea.

Morphological trait observation

Five plants of each 31 accession were evaluated for 25 morphological and agronomic traits (Table 2) based on germplasm characterization criteria provided by the Rural Development Administration, Suwon, Korea. Two independent observers evaluated all specimens twice to control for intra- and inter-observer variation in classification results.

Genomic DNA extraction

Two to three young leaves from plants of each accession were used for DNA extraction. The concentration and purity of DNA samples was measured using a
### Table 1. The *Capsicum* species used in this study.

<table>
<thead>
<tr>
<th>Species</th>
<th>Accession</th>
<th>Origin</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. annuum</em></td>
<td>Subicho</td>
<td>Korea</td>
<td>Pepper Experiment Station¹</td>
</tr>
<tr>
<td></td>
<td>Hanulcho</td>
<td>Korea</td>
<td>Pepper Experiment Station¹</td>
</tr>
<tr>
<td></td>
<td>Thai Hot</td>
<td>Thailand</td>
<td>Cornell Univ.²</td>
</tr>
<tr>
<td></td>
<td>Dempsey</td>
<td>USA</td>
<td>Cornell Univ.</td>
</tr>
<tr>
<td></td>
<td>Perennial</td>
<td>India</td>
<td>Cornell Univ.</td>
</tr>
<tr>
<td></td>
<td>CM334</td>
<td>Mexico</td>
<td>Cornell Univ.</td>
</tr>
<tr>
<td><em>C. chinense</em></td>
<td>Miscucho</td>
<td>Peru</td>
<td>Cornell Univ.</td>
</tr>
<tr>
<td></td>
<td>Jalapeno Rojo</td>
<td>Mexico</td>
<td>Cornell Univ.</td>
</tr>
<tr>
<td></td>
<td>Habanero</td>
<td>Mexico</td>
<td>Cornell Univ.</td>
</tr>
<tr>
<td></td>
<td>Peru</td>
<td>Peru</td>
<td>Cornell Univ.</td>
</tr>
<tr>
<td><em>C. frutescens</em></td>
<td>Tabasco</td>
<td>USA</td>
<td>Cornell Univ.</td>
</tr>
<tr>
<td></td>
<td>C00065</td>
<td>Costa Rica</td>
<td>AVRDC³</td>
</tr>
<tr>
<td></td>
<td>C00088</td>
<td>Thailand</td>
<td>AVRDC</td>
</tr>
<tr>
<td></td>
<td>C00309</td>
<td>Mexico</td>
<td>AVRDC</td>
</tr>
<tr>
<td></td>
<td>C00642</td>
<td>Peru</td>
<td>AVRDC</td>
</tr>
<tr>
<td><em>C. baccatum</em></td>
<td>C00045</td>
<td>Costa Rica</td>
<td>AVRDC</td>
</tr>
<tr>
<td></td>
<td>C00475</td>
<td>USA</td>
<td>AVRDC</td>
</tr>
<tr>
<td></td>
<td>C00773</td>
<td>Chile</td>
<td>AVRDC</td>
</tr>
<tr>
<td></td>
<td>C00774</td>
<td>USA</td>
<td>AVRDC</td>
</tr>
<tr>
<td></td>
<td>C00803</td>
<td>USA</td>
<td>AVRDC</td>
</tr>
<tr>
<td></td>
<td>C00948</td>
<td>UK</td>
<td>AVRDC</td>
</tr>
<tr>
<td><em>C. pubescens</em></td>
<td>Bangkok</td>
<td>Thailand</td>
<td>AVRDC</td>
</tr>
<tr>
<td></td>
<td>C01323</td>
<td>Ecuador</td>
<td>AVRDC</td>
</tr>
<tr>
<td></td>
<td>C01324</td>
<td>Guatemala</td>
<td>AVRDC</td>
</tr>
<tr>
<td></td>
<td>C01374</td>
<td>Guatemala</td>
<td>AVRDC</td>
</tr>
<tr>
<td></td>
<td>C01572</td>
<td>Peru</td>
<td>AVRDC</td>
</tr>
<tr>
<td></td>
<td>C04895</td>
<td>Guatemala</td>
<td>AVRDC</td>
</tr>
<tr>
<td><em>C. chacoense</em></td>
<td>C04389</td>
<td>Argentina</td>
<td>AVRDC</td>
</tr>
<tr>
<td></td>
<td>C04390</td>
<td>Bolivia</td>
<td>AVRDC</td>
</tr>
<tr>
<td></td>
<td>C04392</td>
<td>Bolivia</td>
<td>AVRDC</td>
</tr>
<tr>
<td></td>
<td>C04399</td>
<td>Argentina</td>
<td>AVRDC</td>
</tr>
</tbody>
</table>

¹Young Yang Pepper Experiment Station, Young Yang, Korea

²Cornell University, Ithaca, NY, USA

³Asian Vegetable Research and Development Center, Tainan, Taiwan
**Table 2.** Morphological and agronomic traits and their descriptions used in morphological characterization.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plant shape</td>
<td>1: Dwarf, 2: Intermediate, 3: Erect</td>
</tr>
<tr>
<td>Plant height</td>
<td>From soil to the highest point of individual plants. Average of 5 plants after growth is complete (cm)</td>
</tr>
<tr>
<td>Plant width</td>
<td>From left side to right side of individual plants. Average of 5 plants after growth is complete</td>
</tr>
<tr>
<td>Stem length</td>
<td>Stem length (cm)</td>
</tr>
<tr>
<td>Stem color</td>
<td>1: Green, 2: Purple, 3: Green with purple</td>
</tr>
<tr>
<td>Pubescence</td>
<td>1: Glabrous, 2: Intermediate, 3: Excessive</td>
</tr>
<tr>
<td>Leaf color</td>
<td>1: Green, 2: Dark green, 3: Green with purple, 4: Purple</td>
</tr>
<tr>
<td>Leaf length</td>
<td>Average of full leaf length after growth is complete</td>
</tr>
<tr>
<td>Leaf width</td>
<td>Average of full leaf width after growth is complete</td>
</tr>
<tr>
<td>Node length</td>
<td>Average length of internode between the 3rd and 4th node of main stem (cm)</td>
</tr>
<tr>
<td>Corolla color</td>
<td>1: White, 2: Light-green, 3: White with yellow spots, 4: Purple</td>
</tr>
<tr>
<td>Stamen number</td>
<td>Average number of stamens per flower</td>
</tr>
<tr>
<td>Stamen color</td>
<td>1: White, 2: Yellow, 3: Blue, 4: Purple</td>
</tr>
<tr>
<td>Flower size</td>
<td>1: Small, 2: Intermediate, 3: Big</td>
</tr>
<tr>
<td>Flower position</td>
<td>1: Pendant, 2: Medium, 3: Erect</td>
</tr>
<tr>
<td>Fruit position</td>
<td>1: Pendant, 2: Medium, 3: Erect</td>
</tr>
<tr>
<td>Fruit length</td>
<td>Average vertical length of fruit</td>
</tr>
<tr>
<td>Fruit width</td>
<td>Average horizontal length of fruit</td>
</tr>
<tr>
<td>Calyx margin</td>
<td>1: Smooth, 5: Intermediate, 9: Serrate</td>
</tr>
<tr>
<td>Fruit shape</td>
<td>1: Elongated, 2: Oblate, 3: Round, 4: Conical, 5: Campanulate, 6: Blocky</td>
</tr>
<tr>
<td>Immature fruit color</td>
<td>1: Dark green, 2: Green, 3: Green-yellow (light green), 4: Yellow, 5: White, 6: Purple, 7: Black</td>
</tr>
<tr>
<td>Mature fruit color</td>
<td>1: Dark green, 2: Green, 3: Green-yellow (light green), 4: Yellow, 5: White, 6: Purple, 7: Black</td>
</tr>
<tr>
<td>Fruit weight</td>
<td>Average weight of fruit (g)</td>
</tr>
<tr>
<td>Seed color</td>
<td>1: Yellow, 2: Brown, 3: Black</td>
</tr>
<tr>
<td>Seed number</td>
<td>Average number of seeds per fruit</td>
</tr>
<tr>
<td>Seed weight</td>
<td>Weight of 1,000 seeds per fruit</td>
</tr>
</tbody>
</table>
NanoDrop (NanoDrop Technologies, Wilmington, DE, USA) after DNA was extracted using the cetyl trimethylammonium bromide (CTAB) method (Park et al., 2009). DNA samples showing absorbance ratios at 260/280 nm above 1.8 were used for marker analysis.

**Marker screening**

In a previous study, 40 COSII markers were selected for F₁ cultivar identification in *Capsicum* (Jung et al., 2010). This marker set contained sequence variation not only within *C. annuum*, but also between other *Capsicum* species. The *Waxy* gene (Jarret, 2008; Walsh and Hoot, 2001) was also used for marker development. All of these markers were tested for species identification.

**HRM analysis conditions**

A Rotor-Gene 6000 real-time PCR thermocycler (Corbett Research, Sydney, Australia) was used to detect SNP markers via HRM analysis. The real-time PCR amplification conditions were 95°C for 10 min, then 50 cycles of 94°C for 20 s, 55°C for 20 s and 72°C for 40 s, followed by 95°C for 60 s, and finally 40°C for 60 sec. For HRM analysis, temperatures were increased 0.1°C every minute from 70 to 90°C as previously used for the development of COSII markers (Park et al., 2009).
DNA sequence analysis

Sequence analysis was performed using amplicons obtained from each species to confirm SNPs detected by HRM and to detect SNPs in the Waxy gene. Amplified PCR product from pooled DNA of the same species was sequenced directly. A Zymoclean PCR purification kit was used to purify PCR products according to the protocol supplied by the manufacturer (Zymo Research, Irvine, CA, USA), prior to sequencing with an ABI3730 sequencer (NICEM, Seoul, Korea). The sequences of PCR products amplified from accession in each species of Capsicum were analyzed. All obtained sequences were aligned using a MegAlign software.

Dendrogram tree and genetic distance analysis

A dendrogram of genetic distances was obtained using a Mega 4.0 statistical program (Kumar et al., 2008). Pairwise comparisons between species were performed through the measurement of genetic distances using the Tamura-Nei distance method. A genetic distance matrix was used for cluster analysis via the neighbor-joining (NJ) method.
Results

Species identification based on morphology

Twenty-five morphological traits were evaluated to confirm the species identity of Capsicum materials used in this study. Traits related to growth habit and plant features such as leaf, stem, flower, and fruit descriptions were included in the morphological key (Table 2). These characteristics included corolla color, anther color, seed color, and fruit shape (Table 3). Overall, the flower and fruit traits were the most distinct among species. All C. annuum and C. chinense accessions had white corolla, while anther color and fruit shape varied within species. All C. frutescens accessions had blue anthers and tapered fruit, while corolla color varied from white to greenish-white. In the case of C. baccatum, the corolla color was white with yellow spots, while anther color and fruit type differed greatly. In C. pubescens, both corolla and stamen color were purple. Additionally, the seeds of C. pubescens accessions was black-colored, while most Capsicum species seeds were tan (Table 2). C. chacoense accessions all had white corolla and small fruit, but showed either yellow or blue anther color. In summary, C. chacoense, C. baccatum, and C. pubescens could clearly be distinguished from other species of Capsicum based on morphological characteristics. However, none of the morphological traits could be used to identify species in the C. annuum complex due to their similarity.
Table 3. Species identification by morphological traits in *Capsicum*.

<table>
<thead>
<tr>
<th>Species</th>
<th>Accession</th>
<th>Corolla color</th>
<th>Anther color</th>
<th>Seed color</th>
<th>Fruit type</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. annuum</em></td>
<td>Thai Hot, Perennial, CM334</td>
<td>White</td>
<td>Blue-purple</td>
<td>Yellow</td>
<td>Elongated</td>
</tr>
<tr>
<td></td>
<td>Dempsey</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subicho</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hanulcho</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. frutescens</em></td>
<td>C00065, C00309, C00642</td>
<td>Greenish-white</td>
<td>Blue</td>
<td>Yellow</td>
<td>Tapered</td>
</tr>
<tr>
<td></td>
<td>C00088</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tabasco</td>
<td>White</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. chinense</em></td>
<td>Jalapeno Rojo, Habanero</td>
<td>White</td>
<td>White</td>
<td>Yellow</td>
<td>Blocky</td>
</tr>
<tr>
<td></td>
<td>Miscuchu</td>
<td></td>
<td>Blue</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Peru</td>
<td></td>
<td>Purple</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. baccatum</em></td>
<td>C00045</td>
<td>White with yellow spots</td>
<td>Yellow</td>
<td>Yellow</td>
<td>Oblate</td>
</tr>
<tr>
<td></td>
<td>C00475</td>
<td></td>
<td>White</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C00774, C00803, C00948</td>
<td>Yellow</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C00773</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. pubescens</em></td>
<td>Bangkok, C01323, C01324, C01374, C01572, C04895</td>
<td>Purple</td>
<td>Purple</td>
<td>Black</td>
<td>Round</td>
</tr>
<tr>
<td><em>C. chacoense</em></td>
<td>C04389, C04390, C04392</td>
<td>White</td>
<td>Yellow</td>
<td>Yellow</td>
<td>Round</td>
</tr>
<tr>
<td></td>
<td>C04399</td>
<td></td>
<td>Blue</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Development of species identification markers using nuclear DNA sequences

In a previous study, COSII markers were selected to use for cultivar identification in Capsicum (Jung et al., unpublished). Polymorphisms of 40 COSII markers in 31 accessions from six Capsicum species were surveyed by HRM analysis. Of the 40 COSII markers selected, six (C2_At1g60640, C2_At1g20575, C2_At1g50020, C2_At2g19560, C2_At4g03020, and C2_At5g04590) were chosen as candidate species identification markers showing highly species-specific polymorphism (Fig. 1). These six COSII markers showed at least three melting curve patterns representative of certain species. Using C2_At2g19560 and C2_At1g20575, C. annuum and C. pubescens showed distinct melting curve patterns while other species showed the same (Fig. 1A, B). C2_At1g50020 could distinguish five species. However, C. chinense accessions did not show a fixed melting curve pattern and the C. frutescens Tabasco accession followed the C. annuum pattern. The melting curve patterns for C2_At1g60640 and C2_At4g03020 of C. chinense accessions were unfixed (Fig. 1C). However, C2_At1g60640 could identify C. frutescens, C. baccatum, C. pubescens, and C. annuum (Fig. 1D), while C2_At4g03020 distinguished C. annuum, C. pubescens, and C. frutescens (Fig. 1E).

Among the six COSII markers, C2_At5g04590 was especially noteworthy for identification of cultivated Capsicum species (Fig. 1F). While other COSII markers showed less specific melting curves for each species, the melting curve
**Fig. 1.** HRM profiles of markers developed from COSII. Melting curves and difference plot graphs are in left and right columns, respectively. Red, *C. annuum*; pink curve, *C. chinense*; blue, *C. frutescens*; black, *C. baccatum*; purple, *C. chacoense*; and green, *C. pubescens*. (A, C) C2_At2g19560, (B, D) C2_At1g20575, (E, G) C2_At1g50020, (F, H) C2_At1g60640, (I, K) C2_At4g03020, (J, L) C2_At5g04590. (A, B, E, F, I, and J) Normalized fluorescence. (C, D, G, H, K, L) Difference plot.
pattern of C2_At5g04590 was specific to each species. The six different species showed five distinct melting curves, with \textit{C. baccatum} and \textit{C. chacoense} sharing the same pattern. In cultivated \textit{Capsicum} species, each melting curve pattern of C2_At5g04590 corresponded specifically to a certain species without exception. C2_At5g04590 sequences of the six \textit{Capsicum} species were analyzed due to the distinct nature of their melting curve patterns. The amplified product sizes from the six species were the same, without indels in the sequenced region. Compared to \textit{C. baccatum} and \textit{C. chacoense}, a total of 10 sequences specific SNPs were found (Fig. 2). Four SNPs corresponded to transition (3 A↔T, 1 C↔G), while the others were transversion (2 A↔G, 2 C↔T, 2 G↔T).

The \textit{Waxy} gene was used for HRM analysis. The size of the original \textit{Waxy} locus differs among \textit{Capsicum} species due to indels (Fig. 3A). PCR product sizes were 926 bp for \textit{C. annuum}, 927 bp for \textit{C. baccatum}, \textit{C. chinense}, and \textit{C. pubescens}, and 928 bp for \textit{C. chacoense} and \textit{C. frutescens}. Since DNA fragments of over 900 bp in length are not suitable for HRM analysis, a new set of primers was designed using inner sequences having distinct SNPs and indels (Fig. 3). All species showed species-specific melting curves with HRM analysis except for \textit{C. annuum} and \textit{C. frutescens}, which showed the same pattern (Fig. 4).

\textbf{Phylogenetic analysis of \textit{Capsicum} species}

The concatenated sequences of the six COSII markers and the \textit{Waxy} marker
Fig. 2. Sequence alignment of C2_At5g04590 showing ten SNP positions. Black boxes indicate putative positions of SNP substitution.
Fig. 3. Whole Waxy gene sequence of each Capsicum species. A red arrow indicates the position of newly developed HRM waxy primers.
Fig. 4. HRM analysis using the Waxy locus. Five types of normalized melting curves (up) and difference curves (down) detected by HRM profile of an amplicon from the Waxy gene.
were used to construct a phylogenetic tree. The resulting phylogenetic tree structure was consistent with the taxonomic classification of *Capsicum* species. *C. annuum*, *C. chinense*, and *C. frutescens* were shown to be closely related as they made up the *C. annuum* clade. *C. baccatum* and *C. chacoense* were presented in a separate clade. The *C. pubescens* clade, however, was separated from the other species (Fig. 5).
Fig. 5. Phylogenetic tree showing genetic diversity of 31 *Capsicum* species. The tree was produced using the neighbor-joining method based on the six COSII markers and the *Waxy* gene.
Discussion

Even among international organizations, marker-assisted management of germplasm has not widely been used in *Capsicum*. Rather, species identification of new introductions relies mainly on the investigation of morphological traits (Jarret, 2008). This classification method has several drawbacks including limited ranges of variation, environmental effects, and quantitative trait ambiguity (Small et al., 2004). In this study, six *Capsicum* species were characterized using morphological traits such as corolla color, anther color, and fruit type. These traits, however, were not useful for identifying species in the *C. annuum* clades, although they could be useful for identifying of other species. Additionally, a SNP marker set consisting of six COSII markers and the *Waxy* marker was developed and clearly distinguished six *Capsicum* species was developed.

Previously developed DNA markers, such as RFLP, AFLP, and SSR, have been limited in their ability to identify *Capsicum* species because the spectrum of DNA polymorphisms which can be detected by these markers are relatively narrow, thereby limiting the information that can be obtained from a single marker analysis. Sequence analysis of highly variable DNA has commonly been required for overcoming these limitations during the species identification and phylogenetic analysis. Although sequence analysis is precise, it is time-consuming and laborious especially when large numbers of samples need to be analyzed. In this study, HRM analysis was demonstrated to be used for detecting polymorphic
SNP markers and for identifying species. HRM analysis has been used for cultivar identification in grape (Mackay et al., 2008), apple (Chagne et al., 2008), barley (Lehmensiek et al., 2008), almond (Wu et al., 2008), and olive (Muleo et al., 2009). HRM analysis has an advantage over direct sequencing in both marker development and polymorphism detection, since it is sensitive to detect SNP, and sequence differences between species could clearly be detected without direct sequencing.

Six COSII markers were selected for Capsicum species identification. Among the COSII markers, C2_At5g04590 was found to contain ten SNPs classifying six Capsicum species. Except for C. baccatum and C. chacoense, which contain an identical sequence for C2_At5g04590, accessions included in the other four species could be distinguished without exception. Therefore, the C2_At5g04590 sequence could be used as a reliable marker to classify accessions of four of the Capsicum species used in this study. Similarly, COSII markers that amplified both the intronic and exonic regions were used to construct phylogenetic trees of both the Solanum species and Solanaceae family (Wu et al., 2006). Since intronic sequences across the Solanaceous species differed in both length and sequence, only the exonic sequences were suitable to generate a phylogenetic tree. Alternatively, both intronic and exonic sequences were useful for phylogenetic analysis across more closely related species such as the S. lycoperiscum clade. However, intronic sequences were more powerful than exonic
sequences in providing more detailed information about closely related species due to their longer branch length and higher bootstrap value (Wu et al., 2006). Therefore, intron sequences of COSII markers might be more useful for phylogenetic analysis of Capsicum species.

The Waxy gene, encoding an essential protein for granule-bound starch synthesis, was used for species identification in Capsicum (Walsh and Hoot, 2001). Together with the chloroplast atpB-rbcL region, the Waxy gene was suggested for a species identification marker in Capsicum, because it contains a considerable number of nucleotide substitutions and is universally distributed across plant species. It could also differentiate morphologically similar species C. frutescens and C. chinense. Therefore, this marker was applied for species identification in Capsicum. The Waxy gene primer was re-designed to contain eight nucleotide substitutions and one indel for the identification of Capsicum species. Eight nucleotide differences resulted in variations of melting curve patterns. However, there was one insertion on the amplicon of C. frutescens that was not clearly identified with HRM analysis. This might be due to the fact that the stability of the amplicon was not affected largely by the insertion. SantaLucia et al. (1996) reported that nearest-neighbor stabilities were smallest when the combinations of neighbor nucleotides were consisted by thymine or adenine. The inserted nucleotide of C. frutescens on Waxy sequence was thymine and surrounded by thymine and adenine. Therefore, the insertion of this nucleotide might not be
enough to induce the change in the stability of entire DNA duplex detectable by HRM analysis. In fact, the theoretical melting temperature calculated using the model suggested by SantaLucia et al. (1996) did not significantly differentiate the amplicons of *C. annuum* and *C. frutescens*.

A dendrogram was obtained from combination with six COS II markers, and the *Waxy* locus to determine the phylogenetic relationship among *Capsicum* species. The analysis results supported the evolutionary relationship of species reported in previous studies (Ballard, 1970; Jarret, 2008; Pickersgill, 1979; Walsh and Hoot, 2001). *C. chinense*, *C. frutescens*, and *C. annuum* accessions formed a clade separated from other species supporting the close evolutionary relationship among these three species. These species share similar morphological characteristics, thus have often been grouped together to compose the *C. annuum* complex, especially *C. chinense* and *C. frutescens*, in several studies (Heiser and Pickersgill, 1969; Pickersgill et al., 1979). The results in the present study also supported the close evolutionary relationship between *C. baccatum* and *C. chacoense* as reported previously (Walsh and Hoot, 2001), even though they are morphologically different. In addition, *C. pubescens* species was separated from other clades. Most of markers showed variations specific to *C. pubescens* in this study. Therefore, *C. pubescens* species seemed to be diverged early from other *Capsicum* species and evolved separately from *C. annuum* clade and *C. baccatum* clade.
An ideal DNA marker for species classification should be DNA region which is variable between species while conserved within species. However, a single sequence which satisfies this requirement is difficult to be identified and the determination of a universal DNA barcode sequence is controversial.

No single marker used in this study was sufficient for Capsicum species identification; C2_AT5g04590 did not show polymorphism between C. baccatum and C. chacoense, while the Waxy marker could not identify C. annuum and C. frutescens. However, a combination of all markers enabled reliable classification of the six Capsicum species.
References


Chapter II

A Survey of Natural and EMS-induced Variations of eIF4E Using High-resolution Melting Analysis in Capsicum species

Abstract

Allele mining is a method used to find undiscovered natural variations or induced mutations in a plant. Allele mining has become increasingly important as more genomic information is available in plants. A high-throughput method is required to facilitate the identification of novel alleles in a large number of samples. Herein the application of a high-resolution melting (HRM) method to detect natural variations was described and ethyl methane sulfonate (EMS)-induced mutations in Capsicum. Single polymorphic mutations in the first exon of the eIF4E gene were scanned, wherein the mutations confer resistance to potyviruses. Sixteen allelic variations out of 248 germplasm collections were identified through HRM analysis, and one accession carrying an allelic variation (pvrHRM113) was confirmed to be resistant to the TEV-HAT strain. In addition, five single polymorphic mutations in the eIF4E gene were identified in an EMS-
induced mutant population. These results demonstrate that HRM allows for the rapid identification of new allelic variants in both natural and artificial mutant populations.

Keywords: Allele mining, Capsicum, eIF4E, EMS mutagenized population, high-resolution melting (HRM) analysis
Introduction

Genetic diversity enables plant breeders to make a superior cultivar by combining useful alleles from different plants. Genetic diversity of agricultural traits, such as disease resistance, high yield, and tolerance to biotic and abiotic stresses has been exploited for crop improvement (Glaszmann et al., 2010). Insufficient genetic diversity is one of the major challenges in modern plant breeding. Landraces and traditional varieties have steadily been replaced by superior ones but genetically less diverse commercial cultivars during the last several decades (Tanksley, 1997). These changes caused the genetic vulnerability and erosion of genetic diversity. Therefore, the conservation and management of genetic diversity are urgently required (Johnson, 2008). Allele mining has been used to find unexplored alleles at known loci (Bhullar et al., 2009). However, exploring rare novel alleles in a large germplasm collection is a challenging task. Therefore, it is essential to develop a high-throughput method to identify novel alleles from a large collection (Kumar et al., 2010).

Mutant populations induced by chemical mutagens, such as ethyl methane sulfonate (EMS), have been generated for mutation breeding, and for reverse genetics in various crops, including wheat (Dong et al., 2009), tomato (Menda et al., 2004; Minoia et al., 2010), and rapeseed (Stephenson et al., 2010; Wang et al., 2008c). The combination of chemical mutagenesis and the mutation screening method, called targeted induced local lesions in genomes (TILLING), has been
used for the identification of mutations in a target gene. Even though the TILLING method has commonly been used in reverse genetics (Triques et al., 2008), it is labor-intensive and time-consuming processes, such as the isolation of mismatch-specific endonuclease, use of fluorescent-dye labeled primers, and polyacrylamide gel electrophoresis (Gady et al., 2009).

High-resolution melting (HRM) analysis is a recently developed method to detect single nucleotide polymorphism (SNP) (Wittwer et al., 2003). Several studies have shown that the HRM analysis is a sensitive and powerful method for the detection of sequence variations, including SNP, InDels, and even simple sequence repeats (SSRs) in plants (Chagne et al., 2008; Croxford et al., 2008; Lehmensiek et al., 2008). The strengths of this method include low cost, no need for restriction enzymes, and no electrophoresis after the polymerase chain reaction (PCR) reaction (Hofinger et al., 2009). Thus, HRM analysis is suitable for allele mining in plants.

Eukaryotic initiation factor 4E (eIF4E) is related to recessive resistance against several potyviruses including TEV, CVMV, TuMV, and PMMoV in Capsicum. Subtle nucleotide substitutions and amino acid changes in the eIF4E pocket site result in the inhibition of interactions with the VPg protein of the virus genome, which ultimately leads to resistance (Charron et al., 2008; Kang et al., 2005; Yeam et al., 2007). Four series of eIF4E variations (pvr1+, pvr1, pvr1l, and pvr1l) on the genome in Capsicum annuum have been discovered and confirmed
for specific resistance to different types of TEV strains (Kang et al., 2005). Seven natural polymorphic amino acid substitutions in the eIF4E gene were found to be related to virus resistance (Charron et al., 2008). Since these allelic variations have been discovered mostly in C. annuum, additional alleles are yet to be discovered in other Capsicum species, such as C. baccatum, the C. pubescens complex, and other Capsicum wild relatives.

The objective of the present study is to test the feasibility of using HRM analysis in allele mining for the eIF4E gene. The results show that HRM analysis is useful for mining novel alleles from both natural and EMS-induced variations in Capsicum.
Materials and methods

Plant materials

A total of 248 *Capsicum* accessions were obtained from several germplasm centers, including the Asian Vegetable Research and Development Center (AVRDC, Tainan, Taiwan), the Youngyang Pepper Experiment Station (Yeongyang, Korea), and Cornell University (Ithaca, NY, USA). These *Capsicum* accessions included 165 *C. annuum* accessions, 30 *C. frutescens* accessions, 25 *C. baccatum* accessions, 5 *C. pubescens* accessions, 6 *C. chinense* accessions, 16 *C. chacoense* accessions, and 1 *C. praetemissium* accession (Table 1). Ten seeds from each accession were pre-soaked in a petri-dish to induce germination after surface sterilization with 2% sodium chlorate and 10% trisodium phosphate. Germinated seeds were grown on trays in a growth chamber.

Chemical mutagenesis

Korean landrace Yuwol-cho was used to generate a mutant population using EMS. To set an optimal condition for the construction of a mutant population, 100 seeds were treated with EMS at different concentrations, ranging from 0% to 2% (v/v). Seeds were pre-soaked in distilled water at 24°C for 18 h. These seeds were soaked at five different concentrations (Sigma-Aldrich, St. Louis, MO, USA) solutions (0.1 M phosphate buffer, pH 7.0) with gentle agitation at 20°C for 12 h. EMS-treated seeds were washed with 0.5% (v/v) ethylacetate (Sigma- Aldrich) in
Table 1. Nucleotide and amino acid substitutions of *eIF4E* haplotypes in *Capsicum* germplasm, and frequencies of *pvrHRM* haplotypes.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Amino acid change</th>
<th>No. of accessions</th>
<th>C. annuum</th>
<th>C. frutescens</th>
<th>C. chinense</th>
<th>C. baccatum</th>
<th>C. pubescens</th>
<th>C. chacoense</th>
<th>C. praetemissum</th>
<th>Total</th>
<th>Species-specific haplotype</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>pvrHRM1</em></td>
<td></td>
<td>79</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>79</td>
<td>C. annuum</td>
</tr>
<tr>
<td><em>pvrHRM1</em></td>
<td>T51A, P66T, K71R</td>
<td>18</td>
<td>1</td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>23</td>
<td>Multiple</td>
</tr>
<tr>
<td><em>pvrHRM1</em></td>
<td>V67E, L79R</td>
<td>35</td>
<td>5</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>41</td>
<td>Multiple</td>
</tr>
<tr>
<td><em>pvrHRM1</em></td>
<td>K71R</td>
<td>18</td>
<td>-</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>28</td>
<td>Multiple</td>
</tr>
<tr>
<td><em>pvrHRM1</em></td>
<td>H53R, K71R</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>-</td>
<td>5</td>
<td>-</td>
<td>24</td>
<td>Multiple</td>
</tr>
<tr>
<td><em>pvrHRM1</em></td>
<td>G32D, K71R</td>
<td>2</td>
<td>12</td>
<td>-</td>
<td>4</td>
<td>-</td>
<td>9</td>
<td>-</td>
<td>-</td>
<td>28</td>
<td>Multiple</td>
</tr>
<tr>
<td><em>pvrHRM1</em></td>
<td>L20S, K71R</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>Multiple</td>
</tr>
<tr>
<td><em>pvrHRM1</em></td>
<td>E37K, N65D, K71R</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7</td>
<td>C. baccatum</td>
</tr>
<tr>
<td><em>pvrHRM1</em></td>
<td>N65D, K71R</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>C. baccatum</td>
</tr>
<tr>
<td><em>pvrHRM1</em></td>
<td>K71R</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>Multiple</td>
</tr>
<tr>
<td><em>pvrHRM1</em></td>
<td>K71R, A74D</td>
<td>2</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>C. frutescens</td>
</tr>
<tr>
<td><em>pvrHRM1</em></td>
<td>K71R, A73T</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>Multiple</td>
</tr>
<tr>
<td><em>pvrHRM1</em></td>
<td>V67E</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>Multiple</td>
</tr>
<tr>
<td><em>pvrHRM1</em></td>
<td>P66T, K71R, A73D</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>Multiple</td>
</tr>
<tr>
<td><em>pvrHRM1</em></td>
<td>N65D, K71R, A73D</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>C. frutescens</td>
</tr>
<tr>
<td><em>pvrHRM1</em></td>
<td>A68E, A73D</td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>C. annuum</td>
</tr>
<tr>
<td><em>pvrHRM1</em></td>
<td>A68E, A73D, A74D</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>C. annuum</td>
</tr>
</tbody>
</table>
0.1 M phosphate buffer (pH 7.0) for 45 min and then 0.5 mM sodium thiosulfate (pH 9.0) for 45 min. After these treatments, seeds were germinated in a growth chamber until the 8-leaf stage and transplanted in a greenhouse. Seed germination and fruit setting rates of M₁ plants were evaluated to determine an optimum EMS concentration. The mutation frequency for each gene was calculated as previously described by Wang et al. (2008c).

DNA extraction

Fresh leaves of pepper germplasm and M₁ mutants were used for DNA extraction. DNA was extracted using the cetyl trimethylammonium bromide (CTAB) method reported previously (Hwang et al., 2009). Concentrations of DNA samples were measured by Nano Drop (NanoDrop Technologies, Inc., Wilmington, DE, USA) and diluted with distilled water for PCR amplification and HRM analysis.

HRM analysis

Cleaved amplified polymorphic sequence (CAPS) markers were developed to distinguish four pvr1 alleles (Yeam et al., 2005). These CAPS markers were converted into a HRM marker. This marker, named pvrHRM, is 233bp long, corresponding to the 30th to 262nd nucleotides of the eIF4E full-length sequence. The pvrHRM1 forward primer was 5’-GACGGTTGATGAAGCTGAGAAGG-3’,
and the reverse primer was 5'-CAGTGGAGAAAGTGTAGACGTTGC-3'. A real-time PCR thermocycler (Corbett Research, Sydney, Australia) was used to detect nucleotide sequence variations using HRM analysis. PCR and HRM conditions were employed according to Park et al. (2009). Briefly, the real-time PCR amplification conditions consisted of a preheating at 95°C for 10 min; then 50 cycles of 94°C for 20 s, 55°C for 20 s, and 72°C for 40 s; followed by 95°C for 60 s; and finally 40°C for 60 s. For HRM analysis, the temperature was increased by 0.1°C every minute from 70 to 90°C.

**Sequencing and sequence alignment of eIF4E fragment**

To confirm SNPs detected by HRM, sequence analysis was performed using PCR products obtained from plants which were ultimately found to contain different alleles. PCR was performed in the same condition with HRM analysis using Ex-Taq polymerase (Takara, Shiga, Japan). The PCR products were purified by a Zymoclean PCR Purification Kit following the manufacturer’s protocol (Invitrogen Korea, Seoul, Korea), and cloned into a pGEM-T vector (pGEM®-T Easy Vector Systems, Promega, Seoul, Korea). Purified plasmids were sequenced by an ABI3730 sequencer using two universal primers, T7 and SP6 (NICEM, SNU, Seoul, Korea). Nucleotide sequences were aligned using the MegAlign program (DNA Star, DNASTAR, Inc., Madison, WI, USA).
Virus inoculation and disease resistant evaluation

The eIF4E gene variant plants were inoculated with TEV-HAT at the five to six leaf stages. Virus was mechanically inoculated on the third and fourth leaves after being dusted with carborundum. TEV-HAT inoculum was obtained from TEV-HAT infected tobacco tissues by grinding in 0.1 mM potassium phosphate buffer (pH 7.5), as described previously (Kang et al., 2005). Five individuals per accession were inoculated and monitored. At 21 days post-inoculation (dpi), leaf tissues were assayed for the presence or absence of the TEV-HAT virus using DAS-ELISA, as described Kang et al., (2005). ELISA substrate-enzyme reaction was read twice every 15 and 45 minutes. An uninoculated, healthy leaf was used as a negative control.
Results

Screening of pvrHRM variation in Capsicum germplasm using HRM

A 233-base pair fragment on the first exon of the eIF4E gene (the pvr1 locus) was screened by HRM analysis to identify a natural variation (Fig. 1A). First, a polymorphism test was performed for the eIF4E gene to confirm the sensitivity of HRM analysis of C. annuum cv. RNaky (pvr1+), 5502 (pvr1), and Dempsey (pvr12), which had been studied previously (Kang et al., 2005). The normalized melting curve and difference graph of the first exon of the eIF4E gene showed clearly distinguished genotypes of the reference samples (Fig. 1B, C). Thus, this HRM marker was designed as pvrHRM. For instance, pvrHRM of pvr1+, pvr1, and pvr12 are pvrHRM1+, pvrHRM1, and pvrHRM12, respectively.

The melting curve variation of pvrHRM was surveyed using HRM analysis to find allelic variation in the 248 pepper accessions. The differences in the HRM curves of pvrHRM were investigated for germplasm of the reference samples. The patterns of the normalized melting curves (Fig. 2A) and the corresponding difference curves (Fig. 2B) of the natural variants were completely different from those of the reference samples. Furthermore, new melting curves could be categorized into 13 different types from pvrHRM1 to pvrHRM15. From a total of 16 allelic variations, several haplotypes that followed were species-specific: pvrHRM1+, pvrHRM14, and pvrHRM15 were found only in C. annuum, and pvrHRM17, pvrHRM18, and pvrHRM13 were discovered only in C. baccatum.
Lastly, pvrHRM1\(^{10}\) was present only in *C. frutescens*. In contrast, the other haplotypes (*pvrHRM1, pvrHRM1^2, pvrHRM1^3, pvrHRM1^4, pvrHRM1^5, pvrHRM1^6, pvrHRM1^9, pvrHRM1^{11}, and pvrHRM1^{12}*) were detected in multiple species.

**Nucleotide and amino acid changes in natural variants**

Since HRM analysis could not identify the precise positions and the type of nucleotide substitutions, nucleotide sequences of each representative *pvrHRM* variant were determined in order to confirm the nucleotide and amino acid changes. Sequences of polymorphic alleles including three reference types (*pvrHRM1^+, pvrHRM1, and pvrHRM1^2*) were aligned and compared to each other. Sixteen allelic sequences of *pvrHRM* exhibited diverse variation in 33 nucleotide sites, but no InDel polymorphism was detected. Among twelve new haplotypes, nine haplotypes were homozygous and three haplotypes (*pvrHRM1^8, pvrHRM1^{11}, and pvrHRM1^{14}* ) were heterozygous combined with *pvrHRM1^+*. All 33 SNPs led to amino acid changes (Fig. 3A, Fig. 4). Nucleotide and amino substitutions were mainly discovered near the surface region of the *eIF4E* pocket site, as reported previously (Charron et al., 2008; Monzingo et al., 2007). Mutations of amino acid positions at 51, 66, 67, 68, 73, 74, and 79 were already identified in the natural variation from *C. annuum* accessions (Kang et al., 2005; Charron et al., 2008) (Table 1).
Fig. 1. HRM analysis of allelic variation. (A) The structure of the *pvrl* gene and nucleotide substitution within *pvrl* alleles. The red arrow on the first exon of the *pvrl* gene indicates the location of natural variation and an EMS-induced mutant. (B) Normalized fluorescence of three kinds of *pvrl* gene with HRM. (C) Difference plot graph corresponding to normalized curves of the *pvrl* gene. Red, *pvrl*+; Blue, *pvrl*; Black, *pvrl*2.
Fig. 2. Examples of natural variations of \textit{pvrHRM} analyzed with HRM analysis.

The representative individuals that were sequenced for validation of HRM analysis are shown. (A) Normalized fluorescence. (B) Difference plot.
In our survey, new polymorphic sites near the surface region of *eIF4E* were identified at position 65 in three haplotypes, and were changed from asparagine to asparatic acid (Fig. 3A). Moreover, the amino acid substitution from alanine to threonine, instead of aspartic acid, at position 73 in *pvrHRM1* was also a new amino acid change (Charron et al., 2008). Amino acid changes were also found at positions 53 (*pvrHRM1*), 32 (*pvrHRM*), and 20 (*pvrHRM*), which were not involved in the Cap-binding pocket of *eIF4E* (Table 1).

In comparison with the previous study (Charron et al., 2008), *pvrHRM1* sequence in the present study was the same as the first exon of *pvr2* or *pvr2*. In addition, *pvrHRM1* and *pvrHRM1* were the same sequences as *pvr2* and *pvr2*, respectively. However, new combinations of amino acid changes were also discovered in *pvrHRM1*, *pvrHRM1*, *pvrHRM1*, and *pvrHRM1* alleles.

**Virus resistance test on new allelic variants**

TEV-HAT virus was inoculated to confirm whether the amino acid changes confer resistance against potyvirus. Most of the accessions developed typical symptoms of TEV-HAT, such as the mosaic symptom and mild vein-clearing after 7 dpi. The symptoms were gradually spread to uninoculated upper leaves. At 21 dpi, the virus accumulation test was evaluated by DAS-ELISA. The *pvr1* and *pvr1* genotypes were shown to be resistant to TEV-HAT, as expected. In the case of 12 newly obtained allelic variants, the *pvrHRM1* haplotype from C03946 (C.
Fig. 3. Amino acid sequence alignment of *pvrHRM* alleles. (A) Amino acid sequence substitution in natural variation. (B) Amino acid sequence substitution in EMS-induced mutation. Black boxes indicate the variation of sequences in *pvrHRM*.
Fig. 4. Nucleotide sequence alignment of *pvrHRM1* alleles in natural variation.

Black boxes indicate the variation of sequences.
baccatum) was resistant to the TEV-HAT strain (Fig. 5). This allele contained three amino acid substitutions (N65D, K71R, and A73D) on \textit{pvrHRM1} (Table 1).

**Generation of an EMS mutant population**

Yuwol-cho (\textit{C. annuum} L.) is a pungent Korean landrace and has unique characteristics for the construction of an EMS mutant population. This landrace flowers and matures early, produces a large number of seeds per fruit and is susceptible to various viral diseases.

In order to develop an EMS mutant population, optimal EMS concentrations were investigated for mutagenesis. Yuwol-cho seeds were treated to five different concentrations of EMS ranging from 0\% to 2\% (v/v) at 0.5\% intervals. The germination rate and time were observed everyday until 7 days after treatment (Table 2). The germination rate gradually decreased with increasing EMS concentrations. Among the five EMS concentrations, the optimal germination rate (LD_{15}) was determined to be 1.5\% EMS to generate maximum number of mutant.

A total of 1,350 seeds were treated with 1.5\% EMS, and were sown in the green house to generate an \textit{M}_1 population. A total of 1,092 seeds (80.8\%) germinated and 735 set seeds (67.3\%). These \textit{M}_1 plants were evaluated to confirm the effect of EMS for distinctive phenotypes. Twenty-five morphological traits including morphology and growth habit of leaf, inflorescence, flowers, and seed fertility were evaluated (Jeong et al., 2010).
Fig 5. Virus accumulation in *Capsicum* germplasm containing various *pvrHRM* alleles. Virus accumulation was tested at 14 dpi. Each of three individuals was tested for virus accumulation. Bars are standard errors of the means. NC, un-inoculated leaves as a negative control.
Table 2. Effects of EMS on seed germination rate and fruit fertility of Yuwol-cho.

<table>
<thead>
<tr>
<th>EMS (%)</th>
<th>No. of treated seeds</th>
<th>No. of germinated seeds</th>
<th>Germination rate (%)</th>
<th>Fruit fertility (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>102</td>
<td>97</td>
<td>95.1</td>
<td>99.3</td>
</tr>
<tr>
<td>0.5</td>
<td>107</td>
<td>96</td>
<td>89.7</td>
<td>90.9</td>
</tr>
<tr>
<td>1.0</td>
<td>110</td>
<td>94</td>
<td>85.4</td>
<td>83.2</td>
</tr>
<tr>
<td>1.5</td>
<td>112</td>
<td>93</td>
<td>83.0</td>
<td>74.9</td>
</tr>
<tr>
<td>2.0</td>
<td>108</td>
<td>80</td>
<td>74.1</td>
<td>36.7</td>
</tr>
</tbody>
</table>
Fig. 6. EMS-induced nucleotide changes in \textit{pvrHRM1}. Black boxes indicate the variation of sequence.
Among 25 pooled DNA samples, HRM curves of 14 individuals were identified as individual mutants. PCR products from putative mutants were performed for the validation of HRM analysis. Finally, eleven mutants were found to contain EMS-induced nucleotide substitutions. Six of the mutant individuals did not produce seed. Therefore, five individuals were considered to be true mutants (Fig. 3B, Fig. 6). EMS-induced nucleotide substitutions occurred on the exon; three of them (60%) were silent mutations, and the others (40%) were mis-sense mutations. Of the five mutants, 60% of the total number of mutations represented the GC/AT transition, and the other mutations represented GC/TA transversion (Fig. 3B, Fig. 6). HRM analysis was repeated to confirm whether the mutations were inherited by their offsprings. Eight individuals obtained from five different M₁ mutants were segregating as homozygous wild-type, mutant type, and heterozygous form. Based on this result, the mutation frequency was found to be one mutation per every 101.77 Kb for each plant (Table 3).
Application of HRM to mutation screening

HRM analysis was used to confirm mutations induced by EMS in the M₁ population. The pvr gene was selected to screen EMS mutants. Before screening EMS mutants, the DNA pooling strategy was tested to expedite screening. Wild-type Yuwol-cho (pvrHRM₁³) and the pvrHRM₁⁴ haplotypes were used to test the sensitivity of HRM for pooled DNA screening. These two genotypes share identical sequences, except for a single nucleotide that is G/A on the 158th position. DNA from the two haplotypes was mixed to form an artificial mutant (pvrHRM₁⁴) and wild type (pvrHRM₁³) mixtures of 1:1, 1:3, 1:4, and 1:7, respectively. The mutant haplotype could clearly be distinguished in HRM analysis when the artificial mutant to wild type ratios were 1:1, 1:3, and 1:4, while the mutant haplotype could not be discriminated from wild type in the 1:7 mixture (Fig. 7A, B). Finally, 4-fold pooling DNA mixtures, containing a total of 1,092 mutants, were made and analyzed using HRM. The samples containing mutant DNA were easily differentiated, due to a change in the shape and shift of the melting curve (Fig. 7 C, D). Twenty-five pools containing putative mutants were identified in the first screening (Table 2). Pooled DNA was decollated to individuals, and HRM analysis was performed to find individual mutants.
Fig. 7. Sensitivity of HRM analysis and screening of EMS-induced mutations in *pvrHRM*. (A) Normalized fluorescence. (B) Difference plot. The fraction of mutant alleles in non-mutant pooled DNA being 0:1 (red), 1:1 (pink), 1:3 (green), 1:7 (purple) and 0:1 (Blue). The shape and shift of the melting curve of 4X pooled DNA containing mutant variants was differentiated from the group of non-mutated curves due to nucleotide changes. The thick red curve indicates putative mutants. (C) Normalized fluorescence. (D) Difference plot. EMS-induced mutant (red), Wild-type (grey).
**Table 3. Summary of EMS-induced mutation screening using HRM.**

<table>
<thead>
<tr>
<th>Screening</th>
<th>Mutation detected</th>
<th>False negative (%)</th>
<th>Mutation rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st screening</td>
<td>25</td>
<td>14 (48)</td>
<td></td>
</tr>
<tr>
<td>(4X pooled DNA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd screening</td>
<td>14</td>
<td>3 (21)</td>
<td></td>
</tr>
<tr>
<td>(Individuals DNA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3rd screening</td>
<td>11</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>(Sequencing)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4th screening</td>
<td>5</td>
<td>0</td>
<td>1 per 101.77kb</td>
</tr>
<tr>
<td>(M₂ segregation)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Discussion

Allele mining facilitates the discovery of new allelic variants among a huge number of genetic resources. Finding an unexplored allele has been achieved in agriculturally important crops, including rice (Wang et al., 2008a), wheat (Yu et al., 2008), barley (Hofinger et al., 2009), potatoes (Wang et al., 2008b), and peppers (Rubio et al., 2009). In addition, allele mining can be applied for not only the model plant for which genomic information is available, but also for crops whose whole genome sequence is not available (Hofinger et al., 2009). HRM analysis has been validated as an extremely convenient technology for the development of SNP markers in various crop plants, including barley (Hofinger et al., 2009), almonds (Wu et al., 2008), grapes (Mackay et al., 2008), and olives (Mackay et al., 2008; Muleo et al., 2009), and peppers (Park et al., 2009). Considering the accuracy, time, and cost of HRM analysis, it has also been suggested as a good method for linkage mapping using SNP (Chagne et al., 2008; Croxford et al., 2008; Lehmensiek et al., 2008; Wu et al., 2009). The nucleotide variations in the eIF4E gene could be detected in barley (Hofinger et al., 2009). The presence of mutations was also clearly identified in the phytochrome A (phy A) gene of olive (Muleo et al., 2009) using HRM analysis.

In the present study, the HRM method was applied to survey the nucleotide variation of the first exon of eIF4E in Capsicum. Potyvirus was achieved by key amino acid changes in the first exon of eIF4E (Gao et al., 2004; Kang et al., 2005).
Therefore, this region is proper to mine new allelic variants of the pvr1 gene in Capsicum germplasm. A CAPS marker, pvr1-R1S, which can identify the nucleotide substitutions among pvr1 allelic series (pvr1+, pvr1, and pvr12), was developed in the first exon region (Yeam et al., 2005). However, this CAPS marker could not be directly applicable for HRM analysis because of the size of the PCR product and intronic sequence variation. The resolution of melting curves of the PCR products have been found to be rapidly decreased when the PCR product size is over 500 bp (Hofinger et al., 2009; Park et al., 2009; Reed and Wittwer, 2004). Moreover, nucleotide substitution on the intronic region of pvr1-R1S could also affect the precise genotyping of allelic variants using HRM analysis. Therefore, a new HRM marker, containing four substitution in nucleotide sites was re-designed for optimizing the HRM analysis condition, and this HRM marker could successfully differentiate the pvr1 series (Fig. 1A, B).

In previous TEV-HAT infectivity studies in Capsicum, the pvr1 and pvr12 alleles could abolish virus accumulation, while pvr1+, pvr1, and pvr2 alleles from pvr21 to pvr29 allowed high accumulations of the virus (Charron et al., 2008). In the present study, the pvrHRM113 allele was discovered as a new resistant source to the TEV-HAT strain (Fig. 3, Fig 5.). Three amino acid changes of pvrHRM113 in eIF4E were distributed on the surface region of the cap-binding pocket (Charron et al., 2008; Monzingo et al., 2007). Among them, the amino acid at 66, 67, and 79 has been demonstrated to play an essential role for TEV and PVY,
while amino acid changes at positions 65 and 73 in the *pvr1* gene have not been reported (Charron et al., 2008; Kang et al., 2005). Therefore, TEV-HAT resistance in this accession might be due to amino acid changes of *pvrHRM1*<sup>13</sup>. Even though the precise correlation between amino acid change and virus resistance should be confirmed in further study, this allele could directly be used for virus-resistant breeding against TEV-HAT strain in pepper. Furthermore, since amino acid substitutions in the *eIF4E* gene could be involved in the broad-spectrum resistant to potyviruses, this allelic variant, *pvrHRM1*<sup>13</sup>, could also be effective to different potyviruses including *Pepper mottle virus*, and Potato virus Y (Charron et al., 2008; Kang et al., 2005; Yeam et al., 2005).

An EMS-induced mutant population was generated in *Capsicum*, and the mutations were confirmed through HRM analysis. Five mutants in the *pvrHRM* fragment were finally obtained from 1,092 M<sub>1</sub> individuals (Table 3). Therefore, the mutation frequency induced by 1.5% EMS was estimated at 1 mutation per 101.77 kb. This mutation frequency seemed to be relatively higher than that of two tomato EMS populations (1/322 kb and 1/737 kb in 1.0% EMS) (Gady et al., 2009; Minoia et al, 2010), *Arabidopsis* (1/170 kb) (Greene et al., 2003), barley (1/500 kb) (Caldwell et al., 2004), and rice (1/500 kb) (Wu et al., 2005). In most cases, almost 100% of the mutation has been identified as a transition mutation in plants (Dong et al., 2009; Greene et al., 2003; Till et al., 2004). In the present study, however, only 60% of the mutations induced by EMS were transition
mutations (GC/AC), while the others were transversion mutants. This result is similar to the tomato EMS-induced mutant population (Minoia et al., 2010). Only 70% and 30% of the mutations were discovered as transition mutations in the 0.7% and 1.0% EMS-treated populations, respectively.

In comparison with other species, this difference might result from the specific mechanism against EMS. Especially, the Solanaceae family including tomato and pepper might have anti-mutagenic response that would work differentially to combat EMS mutagenesis (Minoia et al., 2010). Furthermore, HRM could identify all kinds of SNP substitutions on the nucleotide which are not found in other crops, because of the relatively low specificity to recognize all mismatches using endonuclease.

HRM analysis has also proved to be a very precise method for detecting a point mutation in homozygous and heterozygous haplotypes (Hofinger et al., 2009; Muleo et al., 2009; Wittwer et al., 2003). Detection of the transition from G/C to A/T, induced by EMS, was shown to be highly sensitive in HRM analysis (Hofinger et al., 2009; Wu et al., 2008). Considering the advantages of HRM analysis, an EMS-induced mutation in a gene of interest could be identified in the M1 generation. A DNA pooling strategy was applied to reduce the number of samples to be analyzed. In tomato, DNA diluted 8 times and mixed with S. lycopersicon and S. pimpenellifolium were distinguishable (Gady et al., 2009). In the present study, 4X pooled DNA samples were easily identified due to a change
in the melting curve shape and shift, while 8X pooled samples were not
differentiated. Because the mutants generally exist as a heterozygous haplotype in
the M₁ generation, a mutant of 4X pooled DNA in M₁ mutants corresponds with
that of the 8X pooled DNA in M₂ mutants. Therefore, the precise sensitivity of
HRM analysis was able to detect the mutation in up to 12.5% diluted samples,
even in a heterozygous mutation. HRM analysis for allele mining was applied to
detect nucleotide substitution in a large number of germplasm and an EMS-
induced mutant population.

Considering the advantages of HRM analysis, this method was useful to find
the allelic variation in Capsicum. Therefore, HRM analysis could be used as an
alternative method for SNP genotyping, TILLING, and Eco-TILLING.
References


analysis: application to the verification of grapevine and olive cultivars. Plant Methods 4:8-17.


Chapter III

Tomato Male Sterile Gene, \textit{ms10}^{35}, Regulates Tapetum Degeneration and Meiosis

Abstract

Meiosis and tapetum development are required to produce pollens. Understanding of the meiotic process and tapetum development is essential, but a little has been discovered in tomato. Here, a tomato male sterile gene, \textit{ms10}^{35}, regulating for meiosis and tapetum development, was identified. A tomato mutant containing \textit{ms10}^{35} showed abnormal flower morphologies with a dysfunctional meiotic process and enlarged vacuolation of tapetum at the early stage of pollen development. Microspores of \textit{ms10}^{35} mutant were completely degenerated and never developed to functional pollens. A responsible gene encoding a basic helix-loop-helix (bHLH) transcription factor was cloned by map-based cloning. \textit{Ms10}^{35} was highly expressed in anther tissues at the early developmental stage and showed homology with \textit{Arabidopsis DYT1}. \textit{ms10}^{35} male sterility was completely recovered by transgenic complementation test with a full length of \textit{Ms10}^{35}, indicating that a bHLH transcription factor is a responsible gene for \textit{ms10}^{35}.
Transcriptome analysis revealed that 220 anther specific genes which are associated tapetal development and meiosis had changed expression in ms10^{35} mutant. Altogether, ms10^{35} is claimed to be essential for meiocyte and tapetum development and function in tomato.

Keywords: bHLH transcription factor, Male sterility, Map-based cloning, Pollen development, Tapetum
Introduction

Pollen development is one of the most fundamental processes in the plant life (Wilson and Zhang, 2009). Through pollen development, plants deliver the genetic materials from generation to generation and produce recombinant progenies in the next generation (Deveshwar et al., 2011). During this process, plant manipulates an exquisite pathway to produce pollen supported by regulation of enormous numbers of genes (Feng et al., 2012; Honys and Twell, 2004; Wilson and Zhang, 2009). Pollen development has been demonstrated to be controlled by the differential expression and coordinated complex interactions of anther-specific sporophytic and gametophytic genes (Ma, 2005; Jia et al., 2008; Yang et al., 2003; Wilson and Zhang, 2009; Zhao et al., 2008). After anther cell specification and differentiation, sporogenous cells undergo meiosis to generate microspores. Simultaneously, tapetal cells support the pollen development by transportation of nutrients and materials generated via programmed cell death (PCD). Defect of only a gene in the regulatory machinery often induces the failure of producing functional pollen, and result in male sterility (Goldberg et al., 1993; Sanders et al, 1999).

Several sporophytic genes required for meiosis in model plants have been reported (Wilson and Zhang, 2009). Homologous chromosome paring and synopsis-related gene such as homologous pairing aberrant in rice meiosis 1
PAIR1, PAIR2, and PAIR3 ensure to form bivalents and to proceed cytokinesis (Nonomura et al., 2004; Nonomura et al., 2006; Yuan et al., 2009). In pollen semi-sterility 1 (PSS1), spindle organization is failed and results in unusual chromosome division (Zhou et al., 2011). In addition, Arabidopsis SWII and rice REC8 are essential for chromatid cohesion and bivalent formation during meiosis (Mercier et al., 2003; Shao et al., 2011). In these mutants, meiocytes undergo abnormal chromosome division and result in cell death.

Several sporophytic genes required for tapetum development have also been reported. Transcription factors, Dysfunctional tapetum 1 (DYTI) in Arabidopsis, and Undeveloped tapetum 1 (UDTI) in rice play a role after anther differentiation and tapetum development at the early stage (Jung et al., 2005; Zhang et al., 2006). The dyt1 and udt1 mutants show phenotypic abnormalities such as vacuolated and enlarged tapetum, and absence of PCD-triggered degradations. Furthermore, DYTI and UDTI regulate the expression of tapetal-specific genes such as lipid binding transporter, proteases, and pollen wall-synthesis genes as well as several transcription factors (Feng et al., 2012; Jung et al., 2005). Therefore, DYTI and UDTI appeared to be a key regulator at early stage of pollen development. In dyt1 and udt1, tetrads were not observed, suggesting that sporophytic and gametophytic tissues coordinate to complete meiosis. It is largely unknown that how meiocyte and tapetum development are coordinated.
Genic male sterility (GMS) has been utilized for F₁ hybrid production in tomato since there is no available cytoplasmic male sterile line (Georgiev, 1991; Kaul, 1988). Over 50 male sterile mutants have been reported to date and grouped into functional, structural, and sporogenous sterility (Gorman et al., 1997). Most of mutants are classified as sporogenous sterility and these are divided into five different subclasses (premeiotic, meiotic, tetrad, microspore, and not determined class) according to stage at which pollen development arrests (Rick and Butler, 1956). Among the mutants aborted at meiosis stage, ms10 and its allelic mutants (ms10³⁵ and ms10³⁶) have been widely utilized for F₁ hybrid breeding because ms10 mutant shows completely stable sterility and has no defect in growth habit. In the previous studies, ms10³⁵ is known to have empty pollen mother cell (PMC) and to show abnormal formation of tetrad in the anther locules (Corral-Martínez et al., 2010; Rick, 1948; Segui-Simarro and Nuez, 2007; Zamir et al., 1980). The cell lethality in ms10³⁵ mutant was reported because of unbalanced chromosome rearrangement occurring after telophase II (Nikolova et al., 2009). Furthermore, tapetal development in ms10³⁵ was delayed and incomplete at anthesis (Rick, 1948). However, the precise mechanisms and a responsible gene for ms10³⁵ have not been identified.

To elucidate the molecular mechanism for regulation of meiosis and tapetum development, the ms10³⁵ was phenotypically characterized and the responsible gene was isolated by map-based cloning. In ms10³⁵, meiosis was arrested at
anaphase II stage and tapetal cells were abnormally vacuolated and enlarged, causing male sterility. The ms10^{35} encodes a bHLH transcription factor that is specifically expressed in anther tissues at early developmental stage. Transcriptome analysis revealed that Ms10^{35} regulates down-stream gene expressions involved in tapetum development and meiosis-related gene. These results, together with mutant phenotypes of ms10^{35}, demonstrated that Ms10^{35} plays an essential role in controlling meiosis and tapetum development in tomato.
Materials and methods

Plant materials

A tomato male sterile plant, 2-517 (ms1035) and a recurrent male fertile parent (T-1082) were kindly provided by Dr. Young Chae (National Institute of Horticultural and Herbal Science, Rural Development Administration, Suwon, Korea). Another ms1035 line, MLE353, was obtained from IPK, Germany. To perform a genetic analysis and fine mapping, F1, BC1F1, F2 populations were obtained by crossing 2-517 and T-1082. All these plants were grown in the farm of College of Agriculture and Life Science at Seoul National University, Suwon, Korea.

Phenotypic observation and genetic analysis of ms1035

Phenotypic abnormality and the presence/absence of pollens in mature flowers of parents and segregation populations were observed to evaluate male sterility. The floral phenotypes such as stigma protrusion, and flower size and color were checked at the mature flower stage. The presence/absence of pollen was visually evaluated with three different flowers in each plant on black cotton fabric. Fluorescein diacetate (FDA) was used to check pollen viability. After dust down pollens on a slide glass, FDA acetone stock solution (2 mg/ml) diluted in 0.5 M sucrose solution (final concentration of 0.5 μg FDA/0.5 M sucrose) was
dropped on the pollens and checked viability using Axiophot microscope (Zeiss, Germany).

**Scanning electron microscopic (SEM) analysis**

Wild-type and *ms10* pollen grains were collected from mature flowers and mounted on SEM stubs. Mounted samples were coated with palladium-gold in a sputter coater (BAL-TEC/SCD 005) and examined by field emission scanning electron microscope (SUPRA 55VP, Carl Zeiss, Germany) with an acceleration voltage of 15 kV. Photo- graphic images were obtained using high resolution digital image processing.

**Light microscopic analysis and transmission electron microscopic (TEM) observation**

Anther of various stages were fixed at room temperature in a solution of 1.5% paraformaldehyde and 1.5% glutaraldehyde in 0.03 M sodium cacodylate buffer, pH 7.2, and 0.03 M soren phosphate buffer at room temperature. Pre-fixed anthers were washed three times with 0.05 M sodium cacodylate buffer for 10 min and then, floral bud was post-fixed with 1% osmium tetroxide in 0.05 M sodium cacodylate buffer for 2 h at 4°C. Post-fixed anthers were stained with 0.5% uranyl acetate solution and dehydrated ethanol series in 50, 75, 90, and 100% for 10 min, and incubated in 100% ethanol at room temperature. After transferred to mixtures
of Spurr’s resin and propylenoxide (1:1) for 3 h under vacuum, dehydrated anther samples were infiltrated with pure resin for 2 h at room temperature with shaking. At last, samples were soaked into flat mold filling with resin, and polymerized for 24 h at 70°C.

For the light microscopy analysis, the resin-fixed anthers were cut (1,000 nm) by ultra-microtome (MTX, RMC, Tucson, AZ, USA) and stained with 0.05% toluidine blue O in 2.5% sodium carbonate solution. Semi-thin sections were observed and photographed using an Axiophot photomicroscope (Zeiss)

For TEM analysis, ultrathin floral bud fragments (80 nm) were obtained using a MTX ultramicrotome and were then double stained with 2% (w/v) uranyl acetate and 2.6% (w/v) lead citrate aqueous solution. TEM images were observed using JEM1010 transmission electron microscope (Jeol, Tokyo, Japan) at 80 kV. High resolution digital images were captured an 832 charge-coupled device camera (Gatan). All the microscopy observations were performed at the National Instrumentation Center for Environmental Management (NICEM), Seoul National University, Seoul, Korea.

**DAPI staining analysis for meiotic process**

For the observation of meiotic chromosome, a modified PMC spreading protocol was used (Ross et al., 1997). Floral buds around meiosis stage were fixed in Carnoy’s fixative solution (ethanol:acetic acid = 3:1) for 48 h. Fixed buds were
rinsed twice in distilled water and in 10 mM citrate buffer (pH 4.5), respectively. Samples were incubated in a 0.3% (w/v) digestion mix cellulase RS and pectolyase Y23 (Sigma, St. Louis, MO, USA) in 10 mM citrate buffer for 3 h at 37°C. After digestion, 10 μl of 60% acetic acid was added to the cell suspension. Cells were fixed on the heated slide with Carnoy’s fixative. After air drying, fixed cells were stained and mounted with a 2 μg/ml DAPI solution in Vectashield antifade mounting medium (Vector Laboratories, Burlingame, CA, USA).

**DNA extraction**

Genomic DNA was extracted from 2-3 young leaf tissues using hexadecyltrimethyl-ammonium bromide (CTAB) method (Jeong et al., 2010). Leaf tissues were fragmentized using TissueLyserII (Qiagen, Haan, Germany). DNA concentrations were measured by a Nanodrop spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE, USA) and diluted to a final concentration of 20 ng/μl with TE buffer (pH 7.0) for further experiment.

**Alignment of BAC clone and marker development using high-resolution melting (HRM) analysis**

Tomato bacterial artificial chromosome (BAC) clones corresponding to the sequence from 69 to 78 cM on Chromosome 2 provided by Korea Research Institute of Bioscience and Biotechnology (KRIBB) were aligned by Seqman.
software (DNA Star, DNASTAR, Inc., Madison, WI, USA) for developing a large BAC clone contig. COS II markers located on the target region were tested to determine whether the ms10^{35} gene is contained in these BAC clones. Marker sequences were derived from the cleaved amplified polymorphic sequence (CAPS) sequence in the SGN database (http://solgenomics.net/). Polymorphic COS II marker between male sterile and fertile parents were selected and used for fine mapping for further experiment.

In order to develop additional markers linked to ms10^{35}, primer sets were designed based on non-coding region sequences which were predicted using FGENESH program (http://linux1.softberry.com/). Polymorphism test was performed with parental (2-517 and T-1082) and F₁ DNA using HRM analysis (Rotor-Gene 6000 thermocycler machine, Corbett Research, Sydney, Australia). PCR amplification was initiated with a 4 min hold at 95°C as an initial denaturation step, followed by 50 cycles of 95°C for 15 s, annealing at 55°C for 15 s, 72°C for 30 s. Melting curve analysis was performed after the PCR termination and ramped temperature from 70 to 95°C, raised by 0.1°C per second.

**Fine mapping of ms10^{35}**

A total of 1,100 plants comprised of 236 F₂ individuals were used for genetic analysis and additional 864 F₂ individuals were used for fine mapping. Nine markers were used for genotyping of F₂ individuals (Table 1). For HRM analysis,
each genotype was called automatically by normalized curves. For analysis of sequence characterized amplified region (SCAR) marker, amplified PCR products were separated on an 1% agarose gel stained with ethidium bromide. Linkage analysis of molecular markers was conducted using the Carthagene 1.0 program (De Givry et al., 2005). A minimum LOD score of 3.0 and a maximum distance of 30 cM were used as the threshold values. Molecular linkage maps were obtained using MapChart program (Voorrips, 2002).

Total RNA isolation and reverse transcriptase PCR (RT-PCR)

Floral buds, leaves, stems, and fruits at different stages were collected from male fertile and sterile plants and quickly frozen in liquid nitrogen. Floral buds were classified into five different stages according to their size (stage 1, < 2 mm; stage 2, 2-4 mm; stage 3, 4-6 mm; stage 4, 6-8 mm (opening); stage 5, open flower). Total RNA was isolated by TRIZOL extraction buffer (Ambion, Carlsbad, CA, USA). cDNA was synthesized using 2 μg of total RNA according to the manufacturer’s instruction. RT-PCR was performed using primers shown in Table 2. About 200 ng of cDNA was used for RT-PCR. PCR cycle conditions were 30 cycles at 94°C, 1 min; annealing at 55°C for 15 s; 72°C, 2 min. Amplified PCR products were separated on an 1% agarose gel and stained with ethidium bromide.

5’ RACE PCR for transcription start site survey
Table 1. Primers used for fine mapping.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Forward sequence</th>
<th>Reverse sequence</th>
<th>Sequence position on chromosome 2</th>
<th>Marker</th>
<th>No. of recombinants</th>
</tr>
</thead>
<tbody>
<tr>
<td>324K</td>
<td>GGAAGTCGAATCAAGAAAGTAAC</td>
<td>GTTTCAGTATCGCTTGAGAAGC</td>
<td>38348930..38349305</td>
<td>SNP</td>
<td>7</td>
</tr>
<tr>
<td>410K</td>
<td>CACTTCAGTTCTCTTTTTTTTTA</td>
<td>TCTTTCATTGAACTCAAAA</td>
<td>38432869..38450769</td>
<td>SNP</td>
<td>4</td>
</tr>
<tr>
<td>556K</td>
<td>TGAATGTGTCAACAAGTGGG</td>
<td>TATCAGGAAGCTTTAGAAAGCTGG</td>
<td>38578887..38579140</td>
<td>SNP</td>
<td>3</td>
</tr>
<tr>
<td>743K</td>
<td>CGACCAATAAGAAAGAGACGTTGTC</td>
<td>GCGATTTTTCACATAGAATATACATTC</td>
<td>38768391..38768906</td>
<td>SNP</td>
<td>2</td>
</tr>
<tr>
<td>762K</td>
<td>TCTTACTCACCTCGACCTCCTT</td>
<td>CCAGAGCTCTCTAACGTTACTCC</td>
<td>38784095..38784419</td>
<td>SNP</td>
<td>1</td>
</tr>
<tr>
<td>796K</td>
<td>GCTAATGATTAATATGTTGAGTCA</td>
<td>CCCACTCTCTATTCTTTTCGTTTGG</td>
<td>38819385..38820019</td>
<td>SCAR</td>
<td>0</td>
</tr>
<tr>
<td>843K</td>
<td>CTACCTTATAGCAGATGCAAATAGGG</td>
<td>AAGGAGCAGGTAATAATGGGT</td>
<td>38866161..38866295</td>
<td>SNP</td>
<td>2</td>
</tr>
<tr>
<td>1142K</td>
<td>TATGATGGACGACTGAACTCTTC</td>
<td>ACAGCGCGTGTTAGTAAGTTGGAG</td>
<td>39168678..39169088</td>
<td>SNP</td>
<td>5</td>
</tr>
<tr>
<td>1289K</td>
<td>ATGCAGCGAAATAGGCCAG</td>
<td>ATGGCAATGTGTCATCATCTAC</td>
<td>39316042..39317225</td>
<td>SNP</td>
<td>7</td>
</tr>
</tbody>
</table>
Table 2. Primers used for RT-PCR of the ms10\textsuperscript{35} gene and its putative down-stream genes.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward sequence</th>
<th>Reverse sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ms10\textsuperscript{35}</td>
<td>AGATCTCTCTGATTGAGCTTCAG</td>
<td>TCTTGAATAGGAGCAACTCAGG</td>
</tr>
<tr>
<td>Tom108A</td>
<td>ATGCAATTAGGAGCCTTGATTC</td>
<td>CAGTTCCAGTTCCTGTGCAG</td>
</tr>
<tr>
<td>TGAS100</td>
<td>TATATAGACATGGCAATGAAATGC</td>
<td>AGTCAAGACAACGATCAAGAATGC</td>
</tr>
<tr>
<td>LeGPR92</td>
<td>ATGCAATTAGGAGCCTTGATTC</td>
<td>CAGTTCCAGTTCCTGTGCAG</td>
</tr>
<tr>
<td>Cys Protease</td>
<td>ATGGGTGTGGTGATTGGAGGAAG</td>
<td>CAAATGCACTTTCCAATAACCC</td>
</tr>
<tr>
<td>TA29</td>
<td>AAGATTATTAACCATGAACCTCCT</td>
<td>ACATTCTTCAGTGTCACATACATC</td>
</tr>
<tr>
<td>LeMAN5</td>
<td>GCTAGGGTTTTACTTGATGAA</td>
<td>CTTATTTTTCACATGTCCTC</td>
</tr>
<tr>
<td>pLAT52</td>
<td>AAGGTTGACATGATAAGATGGC</td>
<td>AACCCAACTCAAGAGCTTC</td>
</tr>
<tr>
<td>TAGL1</td>
<td>CTTGATGCCAGGGAGGTATCAT</td>
<td>ATCGAATGCTGAGGAGGAG</td>
</tr>
<tr>
<td>DEF1</td>
<td>ATGAGGATCTTTGTAAGGAAGGCTAATG</td>
<td>TTCGTAGATTTCTTTACATCACATTTAG</td>
</tr>
<tr>
<td>Actin</td>
<td>GAAATAGCATAAGATGGGACAGC</td>
<td>ATACCCACCACATCACACGATAT</td>
</tr>
</tbody>
</table>
To identify the transcription start site of the Ms10\textsuperscript{35} gene, 5’ rapid amplification of complementary DNA ends (RACE) PCR was performed using SMARTer\textsuperscript{TM} RACE cDNA Amplification Kit (Clontech, Takara, Shiga, Japan). RNA was extracted from male fertile anthers and cDNA was synthesized according to the manufacturer’s instruction. Sequencing analysis was performed in at NICEM.

**Phylogenetic Analysis**

Multiple alignments of the bHLH families were constructed using ClustalW program of MEGA 5.2 (http://www.megasoftware.net/index.html) and the Neighbor-Jointing (NJ) methods were used with p-distance model and Pairwise Deletion and Bootstrap (1,000 replicates; random seed) (Tamura et al., 2011). Max parsimony method of MEGA was also used to support the NJ tree using default parameter.

**Complementation of the ms10\textsuperscript{35} mutant**

For a functional complementation analysis using transgenic plant, a 2,992 bp genomic sequence of ms10\textsuperscript{35}, which contains the intact ms10\textsuperscript{35} coding region (1,002 bp) with a 1.4 kb upstream sequence and a 0.6 kb downstream sequence, was amplified. The genomic fragment was cloned into the binary vector pCAMBIA2300, which carries a kanamycin resistance as selection marker using
restriction endonucleases XbaI. *Agrobacterium tumefaciens* (LBA4404) was transformed with the pCAMBIA2300-DYT1 plasmid or the control plasmid pCAMBIA2300 and used to transform *ms10*<sup>35</sup> heterozygous plants (*Ms1035/ms1035*). For plant genetic transformation, 7-10 day old cotyledon and hypocotyl were pre-cultured on MS medium plus 10 mg/L Zeatin overnight, then explants were co-cultured with *Agrobacterium tumefaciens* strain LBA4404 harboring *ms10*<sup>35</sup> gene and NPTII or control plasmid (pCAMBIA2300) on the same medium for 2 days. Discs were blotted on the sterile filter paper and transferred to the selection medium including 2 mg/L Zeatin, 0.5 mg/L IAA, 250 mg/L cefotaxime and 50mg/L kanamycin. Subsequently, independent *T<sub>1</sub>* lines were selected by confirmation with PCR. The *T<sub>1</sub>* plants were tested for homozygosity at the *ms10*<sup>35</sup> locus using primers for *NPTII* and transgenes.

**Localization of expressed *Ms10*<sup>35</sup> by in situ hybridization**

The floral buds of male fertile plant from stages 1 to 5 were collected for in situ hybridization. Each floral bud was fixed by vacuum infiltration with FDA buffer (10% formaldehyde, 50% ethanol, and 5% acetic acid) and dehydrated by ethanol. After fixation, floral bud tissues were embedded in paraplast (Sigma). Samples were sectioned into 8 mm thickness using an HM 340E rotary microtome (Microm International, Walldorf, Germany) and attached on the poly *L*-lysine coated slide glass. Deparaffin was performed with xylene (Sigma) and incubated
with 10 mg/ml proteinase K (Sigma). \textit{Ms10} \textsuperscript{35} probe for hybridization were labelled with DIG-nick translation mix according to manufacturer’s instructions (Roche, Basel, Switzerland). Hybridization was performed with DIG-labelled probe in humid chamber for 30 h at 42°C. After the hybridization, slide glass was washed with SSC buffer and formamide buffer and incubated with anti-DIG-alkaline phosphatase (Roche) for 3 h at room temperature. Finally, the samples were stained and visualized with 4-nitro blue tetrazolium chloride (NBT)/5-bromo-4-chloro-3-indolyl-phosphate (BCIP) solutions (Roche). in situ hybridized samples were detected and photographed using a Dimis-M light microscopy (Siwon Optical Technology, Anyang, Korea).

**RNA transcriptome analysis by RNAseq**

Total RNA from floral bud at stages 1 to 3 were extracted for RNA transcriptome analysis. Three different biological replicates were used for each stage. RNA was extracted in the same manner as described above and resuspended to 250 ng/μl with DEPC water. For facilitation of RNA-seq, strand-specific RNA-seq library was constructed for the synthesis of cDNA as described by previously Zhong et al. (2011). With barcoded adapter, 101-bp paired-end sequencing was performed using Hiseq 2500 (Illumina/Solexa, San Diego, CA, USA) in NICEM. RNA-seq algorithm of CLC genomics workbench 6.0 was used for relative digital expression with 98% of identity threshold (CLC bio, Prismet, Denmark). CLC
program was run with default value settings. ITAG2.3_CDS was used as a reference genome for read mapping (http://solgenomics.net). Digital expression data was normalized and transformed by CLC genomics workbench 6.0 internal algorithms. The mean values of three replicates were calculated to identify the statistically significant fold changes using log2 values in expression compared with the wild type and a false discovery rate of <0.05. The DESeq tool of R package (http://www.bioconductor.org/) was used to identify differentially expressed genes and cutoff the non-significant value automatically (Anders et al., 2013). The differentially expressed genes were sorted in Excel and estimated putative function with gene annotations obtained from the Sol genomics network (http://solgenomics.net).
Results

*ms10*<sup>35</sup> shows complete sterility without producing viable pollens.

To compare difference in plant development and growth, male fertile (T-1082) and male sterile plants (2-517) were grown until flowering stage. There was no significant difference in vegetative growth between the two lines until floral bud stage. After flower anthesis, flower morphology became clearly different (Fig. 1A-D). In *ms10*<sup>35</sup> mutant, flowers were pale yellow and much smaller than those of male fertile plants (Fig. 1A, B). Flower size and length of all organs in male sterile plant were much smaller with dramatically shorter anther cone and stigma of sterile plants protruded over anther cone (Fig. 1C, D, E).

To check the pollen activity in male fertile and sterile flowers, FDA test was performed with mature pollens. While highly activated green signals were detected in pollens from male fertile lines, no pollen activity was observed for pollens form male sterile lines (Fig. 1F, I). Moreover, no single pollen was observed on the anthers of male sterile lines under SEM, whereas normal pollen was easily found in male fertile line (Fig. 1G, H, J, K).

**Pollen development of ms10<sup>35</sup> is arrested at tetrad stages**

Previous reports showed that *ms10*<sup>35</sup> had empty PMC in the anther locule and meiosis defect at the early stage of development (Segui-Simarro and Nuez, 2007). However, it remained to be unknown what causes those developmental defects.
**Fig. 1.** Phenotypic analysis of \textit{ms10}^{35} plant. Floral morphology of T-1082 (wild-type) (A, B; Bar = 0.5 cm) and 2-517 (\textit{ms10}^{35}) (C, D; bar = 0.5 cm). Length differences of organ in male fertile and sterile plants (E; unit = cm). AL; anther length, SL; style length. FDA test for activity of fertile (F) and sterile (I) pollens (bar = 100 μm). Observation of mature pollens by SEM in fertile (G, H) and sterile (J, K) anthers (bar = 100 μm for G and J, 10 μm for H, and 50 μm for K).
To investigate and characterize pollen development in the ms10<sup>15</sup> mutant, cellular events were carefully examined according to developmental stages using a light microscope.

The pollen development can be divided into 8 stages from cellular differentiation of pollen mother cell (PMC) and tapetum to anther dehiscence (Brukhin et al., 2003). At the premeiotic stage, five different cell layers were successfully developed from archesporial cells in both male fertile and sterile anthers (Fig. 2A, E). Hook-shaped sporogenous cells were formed in the innermost space in anther locules. Inner and outer tapetal cell layers surrounded sporogenous cells completely. Moreover, outside tapetal cells, middle cell layers were generated from second parietal cells that are also ancestral cell of tapetum. Endothecium cells were developed as multi cell layers and epidermal cells were generated on the outermost layer. At the meiosis stage, there was neither noticeable difference between male fertile and sterile flowers (Fig. 2B, F).

At tetrad stage, dramatic abnormalities in pollen development were observed in the male sterile plant (Fig. 2C, G). In the male fertile plant, PMC proceeded into meiotic cells and tetrad surrounded by callose. Strongly stained tapetal cell layer was observed due to condensation after cell differentiation. Degenerated middle cell layers were almost invisible at this stage. As a result, the boundary between tapetal cells and middle cell layers disappeared while endothecium and
**Fig. 2.** Anther development in male fertile (A-D, I-L) and male sterile (E-H, M-P) flowers. Premeiotic stage (A, E), meiosis stage (B, F), tetrad stage (C, G), free microspore stage (D, H), uni-nucleate stage (I, M), pollen mitosis stage (J, N), pollen maturation stage (K, O), and pollen dehiscence (L, P). dMs, degenerated microspore; En, endothecium; Ep, epidermis; ML, middle cell layer; Msp, microspore; SC, sporocyte; T, tapetum; Tds, tetrad.
epidermis maintained their intact structure. By contrast, PMC underwent abnormal meiotic division in the male sterile flower. Whereas tetrad cells occupied a large space in the locules of the male fertile anthers, microspores in the male sterile anther was greatly reduced in their size and became shrunken due to strong pressure exerted by tapetum expansion (Fig. 2C, G). Furthermore, no nucleus was visible under a microscopy. Especially, tapetal cells of the male sterile anther were excessively enlarged and vacuolated. Condensation and solidification of tapetal cells were rarely observed. Furthermore, middle cell layer still maintained their structure and looked the same as that of premeiotic stage. There were no differences in endothecium and epidermis structure.

After tetrad stage, callose wall was dissolved and tetrad was released to free microspores (Fig. 2D). At this stage, the microspores were vacuolated and stained very weekly in only boundary. Tapetal cells were more condensed and formed a very sharp layer stained very deeply. Tapetal cells appeared to undergo PCD-triggered degeneration. By contrast, degenerated microspores in ms1035 mutant were aggregated themselves and stained very deeply (Fig. 2H). Vacuolated tapetal cells were severely expanded and the microspores were shrunk simultaneously. The middle cell layer also became vacuolated and expanded.

After microspore release, collapsed fragments of tapetum were gradually deposited on the primexine of vacuolated microspore and formed pollen wall (Fig. 2I). Staining of tapetum and middle cell layer were greatly decreased. Vacuolated
microspores were stained black due to accumulated lipid bodies. By contrast, degenerated microspores were continuously dwindled due to tapetum expansion in \textit{ms}10^{35} mutant (Fig. 2M).

In the male fertile anthers, tapetal cell was completely degenerated and became invisible (Fig. 2J, K). Finally, anther dehisced and released pollens outside of the anther (Fig. 2L). In \textit{ms}10^{35} mutant, however, degenerated microspore was continuously shrunk and only residual microspores were remained (Fig. 2N, P). Tapetum remained swollen and vacuolated without degeneration. The middle cell layer was not degenerated either.

In conclusion, no pollens were produced in \textit{ms}10^{35} mutant because pollen development was defected at very early stage of pollen development for the following reasons: 1) meiotic cell division appeared to stop after initiation of meiosis and 2) tapetum and middle cell layer are not degenerated properly.

**Defects in degeneration of tapetal cells result in failure of pollen development in \textit{ms}10^{35}**

To better understand pollen development defect in male sterile mutant, TEM observation was performed. Since pollen development of the \textit{ms}10^{35} mutant appeared to be defected in meiocyte and tapetum degeneration simultaneously, TEM observation was focused on differentiation of tapetum from meiosis to microspore stage. In male fertile plants, tapetum and middle cell layer were
clearly identified and well-defined (Fig. 3A). Tapetal cell had nucleus, cellular organelles, and vacuoles (Fig. 3B). Middle cell layer was thinner than other surrounding cell, indicating that this layer was being disappeared at this stage. By contrast, noticeable differences were observed even at the early stage of tapetum development in male sterile plant (Fig. 3C). Vacuolated region of tapetal cell became much larger than that of previous stage (Fig. 3D). Furthermore, other cellular organelles as well as nucleus were not observed clearly. At the tetrad stage, dense tapetal cells and narrowed middle cell layer were observed in the male fertile plant (Fig. 3E). Nuclear and cellular organelles of tapetum disappeared (Fig. 3F). Middle cell layer was almost diminished and hardly observed. Considering condensed cytoplasm and destructed organelles, cytolytic reaction caused by PCD appeared to be started. By contrast, tapetum was greatly expanded in the male sterile plant and nucleus and cellular organelles were not observed in fully vacuolated tapetal cell (Fig. 3G). Especially, pressures exerted by tapetum expansion caused distortion of degenerating microspores (Fig. 3H). In the case of middle cell layer, their structure was still maintained but cellular organelles disappeared like tapetal cells. At the microspore release stage, tapetum was degenerated by PCD in the male fertile flower (Fig. 3I). Orbicules, which is a responsible organelle to export nutrient and sporopollenin materials, were generated along the degenerated tapetal cell. Middle cell layer completely disappeared and only very narrowed cell remnant was traced (Fig. 3J). However,
**Fig. 3.** Transmission electron microscopic analysis of tapetum development to compare between male fertile (A, B, E, F, I, and J) and male sterile (C, D, G, H, K, and L) flowers. The structure of anther locule at meiosis stage in male fertile (A) and sterile (C) flowers (Bar = 10 μm). Magnified image focusing on the appearance of tapetum at the same stage of A and C in male fertile (B) and sterile (D) anthers (Bar = 2 μm). Difference between male fertile (E) and sterile (G) anthers at tetrad stage (Bar = 10 μm). Magnified image of tapetal cell at the same stage of E and G
in male fertile (F) and sterile (H) flowers (Bar = 2 μm). Free microspore stage of male fertile (I) and sterile (K) plants (Bar = 10 μm) and its magnified image at same stage in male fertile (J) and sterile (L) plants (Bar = 2 μm). dMs, degenerated microspore; En, endothecium; Ep, epidermis; ML, middle cell layer; Msp, microspore; SC, sporocyte; T, tapetum; Tds, tetrad.
tapetum and middle cell layer of the male sterile plant were irregularly expanded (Fig. 3K). Furthermore, tapetum was continuously expanded and pressed degenerated microspores from both sides (Fig. 3L).

**Meiosis of ms10^{35} is arrested at anaphase I**

To figure out the defect of meiotic process in ms10^{35} mutant, TEM analysis and 4’, 6-diamidino-2-phenylindole (DAPI) staining was performed. In male fertile plants, nuclear division during meiosis (Fig. 4A-C) and microspore production (Fig. 4D-E) could be clearly observed. In male sterile flower, there was no noticeable difference until the premeiotic stage. Sporocytes with well-defined structure could be observed (Fig. 4F). Not before long, sporocyte started to be crushed and showed abnormal morphology (Fig. 4G, arrow). Subsequently, sporocytes were diminished steadily and only their traces remained (Fig. 4H, arrow). Finally, cell nuclear disappeared at the end of development (Fig. 4I).

In DAPI staining, similar defect was observed. In male fertile plant, meiotic division from premeiotic stage to tetrads was observed (Fig. 4J-S). Through prophase stage, leptotene/zygotene (Fig. 4J, K), pachytene (Fig. 4L and M), and diakinesis (Fig. 4N) were observed and 12 chromosomes were formed as bivalents at the diakinesis and metaphase I (Fig. 4N-O). Chromosomes were separated into two haploids at anaphase I (Fig. 4P). During meiosis II, diploid cells were formed and subsequently, tetrads were formed at the end of meiosis (Fig. 4Q-S). In male
Fig. 4. TEM analysis of nucleus development (A-I) and DAPI staining for meiosis observation (J-AB). Uni-nucleate of premeiotic stage in male fertile (A) and sterile (F). Nucleus dividing stage (black arrow) (B). Dyad stage in male fertile (C). Diminishing nucleus of male sterile flower at the same stage (G). Tetrad stage in male fertile (D) and its corresponding stage in male sterile flower (H). Released microspore (E) and nucleus remnant in male sterile flower (I). DAPI stagining of male fertile (J-S) and male sterile (T-AB).
sterile plant, meiosis was well proceeded until anaphase I stages (Fig. 4T-Y). However, at anaphase stage, diploids were never observed by meiosis disruption or cell death around anaphase I (Fig. 4Z) and telophase I (Fig. 4AA and AB). Chromosomes were abnormally separated, suggesting that genes responsible for chromatid separation, cytokinesis, or a transcription factor controlling the expression of these genes are dysfunctional.

In conclusion, there were two distinct features of pollen development in ms10<sup>35</sup> mutant, the defect of meiosis in sporogenous cell and excessive tapetal and middle cell layer expansion without cell disruption. All these defects resulted in the failure of pollen development in the ms10<sup>35</sup> mutant.

**ms10<sup>35</sup> encodes a bHLH transcription factor**

To reveal the inheritance of the ms10<sup>35</sup> gene, an individual was developed by crossing 2-517 and T-1082. The pollen fertility was determined by observing pollen production in anthers. All F<sub>1</sub> plants had normal flowers and pollens, suggesting that the function of female organ in ms10<sup>35</sup> plant was normal although flowers were smaller than wild type. In the F<sub>2</sub> population, fertile and sterile phenotypes were segregated in a ratio of approximately 3:1 indicating that male sterility of ms10<sup>35</sup> was controlled by a single recessive gene (Table 3).

In the previous studies, the ms10<sup>35</sup> locus was showed to be tightly linked to peroxidase genes, *per-2, 3* at near 41 cM of Chromosome 2 in Tomato EXPEN
Based on this information, BAC clones that are located around this region were collected and aligned to produce a single large contig (Fig. 5A). Finally, BAC sequence spanning a total of 1,851,216 bp was obtained and used for further experiment. Two hundred and seventy candidate genes were predicted in the BAC sequence and the peroxidase genes were also contained in this BAC contig sequence as expected. Before developing molecular markers for fine mapping of ms10\textsuperscript{35} mutant tested the linkage between the BAC contig and ms10\textsuperscript{35} was tested. COSII markers known to be contained in the BAC contig were converted into SNP marker for HRM analysis. Among them, C2_At4g05090 positioned at 69.8 cM showed polymorphism and was tightly linked to ms10\textsuperscript{35}. Therefore, it was confirmed that the BAC contig contained the ms10\textsuperscript{35} gene.

To perform a fine mapping of the ms10\textsuperscript{35} gene, SNP markers showing polymorphism between male fertile and sterile plants were designed based on the non-coding region of BAC sequence (Fig. 5A). A total of 236 F\textsubscript{2} individuals were screened for recombinant event with SNP markers. Furthermore, additional 864 individuals were used to narrow down the region containing the ms10\textsuperscript{35} gene. Based on the recombinant patterns of each SNP marker, the candidate region of the ms10\textsuperscript{35} gene was determined to be a region spanning 82.2 kb between 762K 843K marker (Fig. 5A). In order to identify the responsible gene for ms10\textsuperscript{35} in this region, DNA sequence was analyzed and ten genes were predicted in the region.
Table 3. Genetic analysis of the *ms10* gene using an F₂ population derived from 2-517 and T-1082.

<table>
<thead>
<tr>
<th>Line</th>
<th>Male sterile</th>
<th>Male fertile</th>
<th>Ratio</th>
<th>Total</th>
<th>$X^2$ (P-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1 (2-517)</td>
<td>9</td>
<td>-</td>
<td>1:0</td>
<td>9</td>
<td>-</td>
</tr>
<tr>
<td>P2 (T-1082)</td>
<td>-</td>
<td>10</td>
<td>0:1</td>
<td>10</td>
<td>-</td>
</tr>
<tr>
<td>F₁</td>
<td>-</td>
<td>8</td>
<td>0:1</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>BC₁F₁</td>
<td>9</td>
<td>8</td>
<td>1:1</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>F₂</td>
<td>62</td>
<td>174</td>
<td>1:3</td>
<td>236</td>
<td>0.1968 (0.6573)</td>
</tr>
</tbody>
</table>
Fig. 5. Map-based cloning of ms10^{35} gene. (A) Physical map around ms10^{35}. Marker position on BAC contig is indicated on the right-down side. The number of recombinant of each marker are represented on the left-up side. (B) Gene prediction of the target region. Putative ORFs are indicated as black boxes. (C) The difference of expression pattern of the ms10^{35} gene. S1-S5 indicated different stage of anther, St: stem, Lf: leaf, Fr: fruit. (D) The
structure of the $Ms10^{35}$ gene encoding a bHLH transcription factor. A part of a transposon element was inserted on the promoter region of male sterile gene. (E) The spatial expression of the $Ms10^{35}$ gene in various flower organs. $Ms10^{35}$ was expressed only in stamen (F) The expression of the $Ms10^{35}$ gene in different $ms10^{35}$ male sterile line.
including two bHLH transcription factor genes, glycine-rich protein, pentatricopeptide repeat-like genes, phosphoinositide binding protein, translation initiation factor and the others (Fig. 5B). To check difference in expression between male fertile and sterile plants, RT-PCR was performed for all the candidate genes using total RNA extracted from different plant tissues including floral bud, stem, leaf and fruits (Fig. 5C). Among them, a bHLH transcription factor showed noticeable differences in its expression. A bHLH transcription factor was highly expressed at stage 1 in male fertile plant (Fig. 5C). The expression was gradually reduced and completely disappeared at stage 3 in male fertile plant. However, expression of this gene could not be detected in anthers of male sterile plant. When the expression of bHLH was checked in four different organ including stamen, pistil, petal and sepal, this gene was expressed only in the stamen tissue suggesting that this gene was specifically activated in male gametogenesis at early stage (Fig. 5E). Another line containing ms10^{35}, MLS353, also showed the same pattern of bHLH gene expression (Fig. 5F). Therefore, the bHLH was chosen as a strong candidate gene for ms10^{35} and performed further analysis. The bHLH gene encoded 209 amino acids composing of 4 exons and 3 introns with 154bp 5’ UTR and 419 bp 3’ UTR and contained bHLH domain. Transcription start site (TSS) was located 151 bp ahead of ATG site as confirmed by 5’ RACE PCR analysis. When nucleotide sequences of this gene was compared, sequence variation was found between male fertile and sterile plant. There was a
single nucleotide difference in third intron, however, this difference do not seem to affect gene function. By contrast, a large insertion was detected 135bp upstream of ATG site in male sterile plants. To reveal the precise structure of insertion in male sterile plant, genome walking was performed. A 398 bp partial region of kielia retrotransposon was found in ms10\textsuperscript{35}. Furthermore, a SCAR marker was developed and confirmed to be co-segregated with ms10\textsuperscript{35} (Fig. 5A).

To better understand the genetic relationship of the tomato bHLH gene, phylogenetic analysis was performed with other bHLH genes from tomato, rice, Arabidopsis, and other crops. Among these genes, DYT1, AMS from Arabidopsis and UDT1, TDR1 from rice were clustered together as reported. Analysis of full length of protein showed that Ms10\textsuperscript{35} gene is highly similar to AtDYT1 and OsUDT1 and formed a separated clade with AtAMS and OsTDR1 (Fig. 6). Amino acid sequence alignment of Ms10\textsuperscript{35} gene with AtDYT1 and OsUDT1, demonstrated that a bHLH domain is highly conserved between three species (Fig. 7).

**Localization of the Ms10\textsuperscript{35} gene by RNA in situ hybridization**

To examine the spatial and temporal expression of the ms10\textsuperscript{35}, RNA in situ hybridization was performed using tissues at various stages from premeiotic stage to pollen dehiscence. A ms10\textsuperscript{35}-specific DNA fragment was used as a probe. Transcript signal was weakly observed on the sporogenous and tapetal tissues at
Fig. 6. Neighbor-joining tree of the $Ms10^{35}$ gene and its homologues. $Ms10^{35}$ was clustered together with $UDT1$ and $DYT1$ genes in the red box.
**Fig. 7.** Amino acid sequence alignment of *Ms10* and its homologues *AtDYT1* and *OsUDT1*. Putative bHLH domain was indicated by red box.
meiotic stage, and abundantly on the meiocyte at tetrad stage (Fig. 8). The signal was continuously detected until free microspore stage. No signal was detected at vacuolated stage or dehiscence stage. By contrast, signal was not observed during all stages in negative control. Therefore, ms10<sup>35</sup> expression appeared to be localized on the meiocyte from early development to free microspore stage.

**Complementation of male sterility with a full length Ms10<sup>35</sup> gene**

To confirm the responsible gene for ms10<sup>35</sup>, the full length Ms10<sup>35</sup> gene with 1.2kb upstream and 0.8 kb downstream sequence from transcription start site was transformed into plants having the heterozygous genotype for ms10<sup>35</sup> (Ms10<sup>35</sup>/ms10<sup>35</sup>) plant. To select plants having homozogous ms10<sup>35</sup> genotype and carrying the complementary construct, a systemic PCR analysis were performed using three ms10<sup>35</sup> linked markers, NPTII gene, and a SCAR marker specific for male sterile gene (Table 4). When three ms10<sup>35</sup>-linked markers were analyzed, 46 transgenic plants of 185 T<sub>1</sub> plants had ms10<sup>35</sup>/ms10<sup>35</sup> homozygous background (Fig. 9, Table 4). Among 46 plants, 34 plants turned out to have complementary construct. When male sterility phenotype was observed, 8 of 34 plants were male fertile phenotypes indicating that male sterility was successfully complemented by the functional Ms10<sup>35</sup> allele. Stigma of all these plants was covered by anther cone. No complementary phenotype was observed transgenic plant with empty vector. These results demonstrated that the 3 kb genomic fragment of Ms10<sup>35</sup> could be
**Fig. 8.** Localization of $ms10^{35}$ expression in the anther by RNA in situ hybridization. RNA in situ hybridization with $ms10^{35}$ probe (A-E). Negative control without $ms10^{35}$ probe (F-J). Meiotic stage (A, F). Tetrad stage (B, G). Free microspore (C, H). Mitosis stage (D, I). Dehiscence stage (E, J).
Fig. 9. Complementary test of $ms10^{35}$ mutant by a full length $Ms10^{35}$ gene. (A) Position of $ms10^{35}$-linked markers and HRM analysis for selection of $ms10^{35}/ms10^{35}$. (B) Primer position for transgenic plant analysis. (C) PCR analysis of selected transgenic line. M: Size marker, (+): T-1082, (-): 2-517, PC: Transformation construct carrying $Ms10^{35}$ gene with up- and downstream sequence, NC: Transgenic plant carrying empty vector (pCAMBIA 2300). (D) Flower morphology and pollen activity test of transgenic plants.
Table 4. Summary of T1 transgenic plant analysis for complementation test.

<table>
<thead>
<tr>
<th>Name of plants</th>
<th>T0 No. of transgenic plants</th>
<th>T1 Genetic background</th>
<th>Km* (+)</th>
<th>TSP* (+)</th>
<th>Phenotype</th>
<th>Male sterile</th>
<th>Male fertile</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>54</td>
<td>39</td>
<td>15</td>
<td>13</td>
<td>11</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>40</td>
<td>31</td>
<td>9</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>11</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>J</td>
<td>67</td>
<td>53</td>
<td>14</td>
<td>11</td>
<td>9</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Empty</td>
<td>9</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

*TSP: Complementary gene specific gene SCAR marker
successfully complemented male sterility of $ms10^{35}$.

**Comparison of global gene expression between male sterile and fertile plants by RNAseq analysis**

Transcriptomes were analyzed in wild type and $ms10^{35}$ mutant by RNAseq using floral bud at three different developmental stages (Stages 1, 2, and 3). A total of 19,170 genes were expressed in 3 different stages of floral bud. Among them, 246 genes including 220 down-regulated and 26 up-regulated genes was discovered by statistical analysis with a false discovery cutoff of $< 0.05$ and identified >log2 fold changed gene ($P<0.05$, Fig. 10). Differentially expressed genes were classified using GO annotation categories.

Transcriptional regulatory pathway is quite important for controlling the expression pattern to produce normal pollen. Expressions of transcription factors such as bHLH, MYB, NAC, and Zinc finger domain were decreased (Table 5).

Overall 21 transcription factors including $Ms10^{35}$ (Solyc02g079810.1.1) were apparently reduced (Tables 5, 10). AMS-like gene (Solyc08g062780.1.1), $TDF1$-like gene (Solyc03g059200.1.1), $MYB35$-like gene (Solyc03g113530.2.1), $MS1$-like gene (Solyc04g008420.1.1) which are the most well known transcription factor in *Arabidopsis*, were dramatically reduced, respectively (Table 10). These three genes were important for pollen development as an activator of downstream gene. Besides, $bHLH89$-like gene (Solyc01g081100.1.1), $MYB21$-like gene
**Fig. 10.** Molecular function of differentially down-regulated in *ms10^{35}* mutant (*P*<0.05). Percentage of each category is shown.
Table 5. Transcription factors down-regulated in the ms10<sup>35</sup> mutant anther.

<table>
<thead>
<tr>
<th>ID</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solyc05g009230.1.1</td>
<td>-5.90</td>
<td>-</td>
<td>-</td>
<td>MYB</td>
</tr>
<tr>
<td>Solyc08g006470.2.1</td>
<td>-3.23</td>
<td>-</td>
<td>-</td>
<td>C2H2</td>
</tr>
<tr>
<td>Solyc01g081100.1.1</td>
<td>-3.20</td>
<td>-</td>
<td>-</td>
<td>bHLH</td>
</tr>
<tr>
<td>Solyc01g107170.2.1</td>
<td>-2.46</td>
<td>-</td>
<td>-</td>
<td>C2H2</td>
</tr>
<tr>
<td>Solyc08g079700.1.1</td>
<td>-2.01</td>
<td>-</td>
<td>-</td>
<td>Zinc finger</td>
</tr>
<tr>
<td>Solyc02g087860.2.1</td>
<td>-1.01</td>
<td>-</td>
<td>-</td>
<td>bHLH</td>
</tr>
<tr>
<td>Solyc02g079810.1.1</td>
<td>Down</td>
<td>-</td>
<td>-</td>
<td>bHLH</td>
</tr>
<tr>
<td>Solyc03g059200.1.1</td>
<td>Down</td>
<td>-</td>
<td>-</td>
<td>MYB</td>
</tr>
<tr>
<td>Solyc03g113530.2.1</td>
<td>Down</td>
<td>-</td>
<td>-</td>
<td>MYB</td>
</tr>
<tr>
<td>Solyc08g062780.1.1</td>
<td>Down</td>
<td>-9.47</td>
<td>-</td>
<td>bHLH</td>
</tr>
<tr>
<td>Solyc08g067960.2.1</td>
<td>-</td>
<td>-2.05</td>
<td>-</td>
<td>Zinc finger</td>
</tr>
<tr>
<td>Solyc12g009240.1.1</td>
<td>-</td>
<td>-3.60</td>
<td>-</td>
<td>ERF</td>
</tr>
<tr>
<td>Solyc05g006650.2.1</td>
<td>-</td>
<td>-1.52</td>
<td>-</td>
<td>Zinc finger</td>
</tr>
<tr>
<td>Solyc05g052570.2.1</td>
<td>-</td>
<td>-1.65</td>
<td>-</td>
<td>Zinc finger</td>
</tr>
<tr>
<td>Solyc06g069710.2.1</td>
<td>-</td>
<td>-4.88</td>
<td>-</td>
<td>NAC</td>
</tr>
<tr>
<td>Solyc04g008420.1.1</td>
<td>-</td>
<td>Down</td>
<td>-</td>
<td>PHD finger</td>
</tr>
<tr>
<td>Solyc04g077980.1.1</td>
<td>-</td>
<td>-2.39</td>
<td>-1.82</td>
<td>C2H2</td>
</tr>
<tr>
<td>Solyc07g063420.2.1</td>
<td>-</td>
<td>-4.01</td>
<td>-3.36</td>
<td>NAC</td>
</tr>
<tr>
<td>Solyc10g006880.2.1</td>
<td>-</td>
<td>Down</td>
<td>-4.07</td>
<td>NAC</td>
</tr>
<tr>
<td>Solyc03g093550.1.1</td>
<td>-</td>
<td>-</td>
<td>-1.80</td>
<td>ERF</td>
</tr>
<tr>
<td>Solyc12g005890.1.1</td>
<td>-</td>
<td>-</td>
<td>Down</td>
<td>MYB</td>
</tr>
</tbody>
</table>
(Solyc05g009230.1.1), and *MYB101*-like gene (Solyc03g059200.1.1) were also decreased in their expression. Interestingly, a bHLH transcription factor, *se2.1* which is a responsible major QTL for the development of stamen length was also decreased significantly. Acyl lipid metabolic genes were also down-regulated in *ms10* mutant (Tables 6, 10). Twenty genes including fatty acid CoA reductases, lipases, MtN3 family gene, and lipid transferase were down-regulated in *ms10*. In these genes, a fatty acid CoA reductase (Solyc03g051960.2.1) showed the homologue with *MS2* gene in *Arabidopsis* which is an important for pollen wall development. In lipid transferase, *TomA108* gene (Solyc01g009590.2.1) was also down-regulated. Fifty-one cell wall modification and degradation-related genes such as glycosyl transferase, pectin lyase, cysteine protease, and other peptidase were dramatically reduced in *ms10* mutant (Table 7). PCD-inducing proteases, C1A cysteine proteinase (Solyc07g053460.2.1) and aspartic proteinase (Solyc06g069220.1.1), were decreased their expression in *ms10* mutant. A total of 21 transporter genes for transport of pollen developmental materials also inhibited their expression in *ms10* mutant (Table 8). Besides, pollen wall synthesis-related genes were down-regulated in the *ms10* mutant (Table 9).

Twelve pollen wall-synthesized genes including glycoprotein, pectin lyase, oleosin, glycosyl transferase, and glucosyl transferase gene were also affected. Mannan endo-1, 4-beta-mannosidase (Solyc06g064520.2.1) was reported to be involved in pollen maturation. Down regulated in meiosis-related gene and meio-
Table 6. Acyl lipid metabolism genes down-regulated in ms10Δ.

<table>
<thead>
<tr>
<th>ID</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solyc03g051960.2.1</td>
<td>-11.11</td>
<td>Down</td>
<td>Down</td>
<td>Fatty acyl coA reductase</td>
</tr>
<tr>
<td>Solyc08g076790.2.1</td>
<td>-6.58</td>
<td>-</td>
<td>-</td>
<td>Cinnamoyl-CoA reductase-like protein</td>
</tr>
<tr>
<td>Solyc08g082770.2.1</td>
<td>-6.25</td>
<td>-</td>
<td>-</td>
<td>Nodulin MtN3 family protein</td>
</tr>
<tr>
<td>Solyc03g119620.1.1</td>
<td>-5.04</td>
<td>-</td>
<td>-</td>
<td>Lipid transfer protein</td>
</tr>
<tr>
<td>Solyc03g097170.2.1</td>
<td>-3.69</td>
<td>-</td>
<td>-</td>
<td>Cinnamoyl-CoA reductase-like protein</td>
</tr>
<tr>
<td>Solyc03g117800.2.1</td>
<td>-3.13</td>
<td>-</td>
<td>-</td>
<td>Fatty acid hydroxylase</td>
</tr>
<tr>
<td>Solyc09g090350.2.1</td>
<td>-3.31</td>
<td>-</td>
<td>-4.27</td>
<td>Long-chain fatty alcohol dehydrogenase</td>
</tr>
<tr>
<td>Solyc01g086830.2.1</td>
<td>Down</td>
<td>-</td>
<td>-</td>
<td>Plant lipid transfer</td>
</tr>
<tr>
<td>Solyc03g026090.2.1</td>
<td>Down</td>
<td>-</td>
<td>-</td>
<td>Plant lipid transfer</td>
</tr>
<tr>
<td>Solyc06g084190.2.1</td>
<td>Down</td>
<td>-</td>
<td>-</td>
<td>Lipid transfer protein</td>
</tr>
<tr>
<td>Solyc01g009590.2.1</td>
<td>Down</td>
<td>Down</td>
<td>-</td>
<td>Plant lipid transfer</td>
</tr>
<tr>
<td>Solyc01g090350.2.1</td>
<td>-</td>
<td>Down</td>
<td>-</td>
<td>Lipid transfer protein</td>
</tr>
<tr>
<td>Solyc11g065530.1.1</td>
<td>-</td>
<td>-2.15</td>
<td>-2.09</td>
<td>Lipase</td>
</tr>
<tr>
<td>Solyc01g095780.2.1</td>
<td>-</td>
<td>-9.87</td>
<td>Down</td>
<td>Lipid transfer protein</td>
</tr>
<tr>
<td>Solyc07g053780.2.1</td>
<td>-</td>
<td>Down</td>
<td>Down</td>
<td>Lipid transfer protein</td>
</tr>
<tr>
<td>Solyc06g059790.2.1</td>
<td>-</td>
<td>-</td>
<td>-5.99</td>
<td>Lipid transfer protein</td>
</tr>
<tr>
<td>Solyc01g066620.2.1</td>
<td>-</td>
<td>-</td>
<td>-5.27</td>
<td>Fatty acid oxidation complex</td>
</tr>
<tr>
<td>Solyc10g009430.2.1</td>
<td>-</td>
<td>-</td>
<td>-4.53</td>
<td>Diacylglycerol O-acyltransferase</td>
</tr>
<tr>
<td>Solyc09g075770.1.1</td>
<td>-</td>
<td>-</td>
<td>-2.49</td>
<td>Long-chain fatty acid CoA ligase</td>
</tr>
<tr>
<td>Solyc04g081780.2.1</td>
<td>-</td>
<td>-</td>
<td>Down</td>
<td>Lipase</td>
</tr>
</tbody>
</table>
Table 7. Genes for cell wall modification/degradation down-regulated in *ms10*.

<table>
<thead>
<tr>
<th>ID</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solyc01g080010.2.1</td>
<td>-5.16</td>
<td>-</td>
<td>-</td>
<td>Xylanase inhibitor</td>
</tr>
<tr>
<td>Solyc07g064180.2.1</td>
<td>-5.10</td>
<td>-</td>
<td>-</td>
<td>Pectinesterase</td>
</tr>
<tr>
<td>Solyc02g093580.2.1</td>
<td>-4.53</td>
<td>-</td>
<td>-</td>
<td>Pectate lyase</td>
</tr>
<tr>
<td>Solyc01g088790.2.1</td>
<td>-3.81</td>
<td>-</td>
<td>-</td>
<td>Sugar transferase</td>
</tr>
<tr>
<td>Solyc12g006380.1.1</td>
<td>-3.68</td>
<td>-</td>
<td>-</td>
<td>Oxoglutarate and iron-dependent oxygenase</td>
</tr>
<tr>
<td>Solyc01g088430.2.1</td>
<td>-3.59</td>
<td>-</td>
<td>-</td>
<td>Fatty acid hydroxylase</td>
</tr>
<tr>
<td>Solyc01g066470.2.1</td>
<td>-3.55</td>
<td>-</td>
<td>-</td>
<td>Epoxide hydrolase</td>
</tr>
<tr>
<td>Solyc03g044880.2.1</td>
<td>-2.72</td>
<td>-</td>
<td>-</td>
<td>Exostin family protein</td>
</tr>
<tr>
<td>Solyc04g082200.2.1</td>
<td>-2.63</td>
<td>-</td>
<td>-</td>
<td>Dehydrid</td>
</tr>
<tr>
<td>Solyc02g084410.2.1</td>
<td>-2.39</td>
<td>-</td>
<td>-</td>
<td>Glyoxalase</td>
</tr>
<tr>
<td>Solyc07g056570.1.1</td>
<td>-2.15</td>
<td>-</td>
<td>-</td>
<td>9-cis-epoxycarotenoid dioxygenase</td>
</tr>
<tr>
<td>Solyc04g099960.2.1</td>
<td>-1.71</td>
<td>-</td>
<td>-</td>
<td>L-allo-threonine aldolase</td>
</tr>
<tr>
<td>Solyc07g006660.1.1</td>
<td>Down</td>
<td>-</td>
<td>-</td>
<td>Glyoxal oxidase</td>
</tr>
<tr>
<td>Solyc07g064190.1.1</td>
<td>Down</td>
<td>-</td>
<td>-</td>
<td>Pectinesterase</td>
</tr>
<tr>
<td>Solyc08g077860.2.1</td>
<td>Down</td>
<td>-</td>
<td>-</td>
<td>Serine protease</td>
</tr>
<tr>
<td>Solyc11g018600.1.1</td>
<td>Down</td>
<td>-</td>
<td>-</td>
<td>Glucose/ribitol dehydrogenase</td>
</tr>
<tr>
<td>Solyc06g069220.1.1</td>
<td>-8.99</td>
<td>-8.64</td>
<td>-</td>
<td>Aspartic proteinase</td>
</tr>
<tr>
<td>Solyc06g062540.2.1</td>
<td>-6.68</td>
<td>-5.02</td>
<td>-</td>
<td>Pyridoxal phosphate phosphatase</td>
</tr>
<tr>
<td>Solyc10g079060.1.1</td>
<td>-6.59</td>
<td>-5.14</td>
<td>-</td>
<td>Squalene synthase</td>
</tr>
<tr>
<td>Solyc01g090710.2.1</td>
<td>-6.12</td>
<td>-5.16</td>
<td>-</td>
<td>Malate dehydrogenase</td>
</tr>
<tr>
<td>Solyc01g009240.2.1</td>
<td>-4.89</td>
<td>-3.92</td>
<td>-</td>
<td>Beta-D-glucosidase</td>
</tr>
<tr>
<td>Solyc05g005700.2.1</td>
<td>-2.59</td>
<td>-2.34</td>
<td>-</td>
<td>Aldehyde dehydrogenase 1</td>
</tr>
<tr>
<td>Solyc04g073990.2.1</td>
<td>-2.47</td>
<td>-1.71</td>
<td>-</td>
<td>Annexin</td>
</tr>
<tr>
<td>Solyc03g046200.1.1</td>
<td>-2.18</td>
<td>-</td>
<td>-5.70</td>
<td>Endo-1 3-beta-glucanase</td>
</tr>
<tr>
<td>Solyc01g081270.2.1</td>
<td>-</td>
<td>-6.33</td>
<td>-</td>
<td>Glutathione S-transferase</td>
</tr>
<tr>
<td>Solyc11g022530.1.1</td>
<td>-</td>
<td>-4.86</td>
<td>-</td>
<td>Glutathione S-transferase</td>
</tr>
<tr>
<td>Solyc02g092580.2.1</td>
<td>-</td>
<td>-4.21</td>
<td>-</td>
<td>Pyridoxal phosphate phosphatase</td>
</tr>
<tr>
<td>Solyc08g082440.2.1</td>
<td>-</td>
<td>-3.02</td>
<td>-</td>
<td>UDP-glucose 4-epimerase</td>
</tr>
<tr>
<td>Solyc12g007030.1.1</td>
<td>-</td>
<td>-2.98</td>
<td>-</td>
<td>Aldehyde dehydrogenase</td>
</tr>
</tbody>
</table>
Table 7. Continued.

<table>
<thead>
<tr>
<th>ID</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solyc12g007030.1.1</td>
<td>-</td>
<td>-2.98</td>
<td>-</td>
<td>Aldehyde dehydrogenase</td>
</tr>
<tr>
<td>Solyc11g008530.1.1</td>
<td>-</td>
<td>-2.65</td>
<td>-</td>
<td>Ribonuclease</td>
</tr>
<tr>
<td>Solyc06g059840.2.1</td>
<td>-</td>
<td>-2.49</td>
<td>-</td>
<td>3-methyl-2-oxobutanoate dehydrogenase</td>
</tr>
<tr>
<td>Solyc05g024260.2.1</td>
<td>-</td>
<td>-2.49</td>
<td>-</td>
<td>RAG1-activating protein 1</td>
</tr>
<tr>
<td>Solyc01g106820.2.1</td>
<td>-</td>
<td>-2.44</td>
<td>-</td>
<td>Peptidase</td>
</tr>
<tr>
<td>Solyc08g080570.2.1</td>
<td>-</td>
<td>-2.24</td>
<td>-</td>
<td>UDP-glucose 4-epimerase</td>
</tr>
<tr>
<td>Solyc03g065250.2.1</td>
<td>-</td>
<td>-1.58</td>
<td>-</td>
<td>Annexin</td>
</tr>
<tr>
<td>Solyc11g006570.1.1</td>
<td>Down</td>
<td>-</td>
<td>-</td>
<td>Inositol oxygenase</td>
</tr>
<tr>
<td>Solyc07g053460.2.1</td>
<td>-</td>
<td>-13.42</td>
<td>-11.00</td>
<td>C1A cysteine proteinase</td>
</tr>
<tr>
<td>Solyc06g074030.1.1</td>
<td>-</td>
<td>-1.82</td>
<td>-1.61</td>
<td>Ribonuclease</td>
</tr>
<tr>
<td>Solyc01g081250.2.1</td>
<td>Down</td>
<td>Down</td>
<td></td>
<td>Glutathione S-transferase</td>
</tr>
<tr>
<td>Solyc03g098680.2.1</td>
<td>-</td>
<td>-</td>
<td>-10.62</td>
<td>Kunitz trypsin inhibitor</td>
</tr>
<tr>
<td>Solyc11g044910.1.1</td>
<td>-</td>
<td>-</td>
<td>-6.84</td>
<td>Beta-xylosidase 1</td>
</tr>
<tr>
<td>Solyc04g081810.2.1</td>
<td>-</td>
<td>-</td>
<td>-6.47</td>
<td>Subtilisin-like protease</td>
</tr>
<tr>
<td>Solyc03g031510.2.1</td>
<td>-</td>
<td>-</td>
<td>-4.70</td>
<td>metallocarboxypeptidase inhibitor</td>
</tr>
<tr>
<td>Solyc07g045090.2.1</td>
<td>-</td>
<td>-</td>
<td>-4.49</td>
<td>Alcohol dehydrogenase</td>
</tr>
<tr>
<td>Solyc10g047030.1.1</td>
<td>-</td>
<td>-</td>
<td>-2.06</td>
<td>Glycoside hydrolase</td>
</tr>
<tr>
<td>Solyc04g054710.2.1</td>
<td>-</td>
<td>-</td>
<td>-1.80</td>
<td>Aspartate aminotransferase</td>
</tr>
<tr>
<td>Solyc10g078550.1.1</td>
<td>-</td>
<td>-</td>
<td>-1.72</td>
<td>Glutamate dehydrogenase</td>
</tr>
<tr>
<td>Solyc11g011930.1.1</td>
<td>-</td>
<td>-</td>
<td>-1.71</td>
<td>Isocitrate dehydrogenase</td>
</tr>
<tr>
<td>Solyc01g094910.2.1</td>
<td>-</td>
<td>-</td>
<td>-1.64</td>
<td>Ferric reductase oxidase</td>
</tr>
<tr>
<td>Solyc04g076650.1.1</td>
<td>-</td>
<td>-</td>
<td>Down</td>
<td>Rhamnogalacturonate lyase</td>
</tr>
<tr>
<td>Solyc03g098310.1.1</td>
<td>-</td>
<td>-</td>
<td>Down</td>
<td>Ornithine decarboxylase</td>
</tr>
</tbody>
</table>
cyte specific gene was also found (Table 10). Sister chromatid cohesion (Solyc03g116930.2.1) was completely down regulated in ms10^{35} gene. In addition, meiocyte-specific expressed gene, LeGRP92 encoding a glycine-rich protein (Solyc02g032910.1.1), was also down regulated.

To validate the transcriptome analysis, RT-PCR using anther-specific genes which were already reported to be expressed in meiocyte and tapetum during pollen development in tomato was performed. Among these genes, TomA108, TGAS100, LeGPR92, SGN-U321072, TA29, LeMAN5 and pLAT52 gene were shown to be significant expression difference between male fertile and sterile plants same as the transcriptome analysis (Figs. 11, 12).

All these results suggested that Ms10^{35} gene controlled tapetum development through supporting the expression of responsible genes for the regulation of transcription factor, biosynthesis, transport, metabolic pathway, lipid transfer gene, establishment of cell wall, and several protease genes. Simultaneously, Ms10^{35} gene contributed the sister chromatid cohesion during meiosis through unknown cell to cell signal transduction.
Table 8. Transporter genes that are down-regulated in ms10^{15}.

<table>
<thead>
<tr>
<th>ID</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solyc03g082700.2.1</td>
<td>-6.12</td>
<td>-</td>
<td>-</td>
<td>Oligopeptide transporter</td>
</tr>
<tr>
<td>Solyc09g005260.2.1</td>
<td>-1.56</td>
<td>-</td>
<td>-</td>
<td>Calcium/proton exchanger</td>
</tr>
<tr>
<td>Solyc06g064480.2.1</td>
<td>Down</td>
<td>-</td>
<td>-</td>
<td>Protein TAPI</td>
</tr>
<tr>
<td>Solyc12g013630.1.1</td>
<td>Down</td>
<td>Down</td>
<td>-</td>
<td>ABC-2 type transporter</td>
</tr>
<tr>
<td>Solyc11g009100.1.1</td>
<td>Down</td>
<td>-</td>
<td>-5.86</td>
<td>ABC-2 type transporter</td>
</tr>
<tr>
<td>Solyc09g014220.2.1</td>
<td>-3.88</td>
<td>-3.77</td>
<td>-2.33</td>
<td>Peptide transporter</td>
</tr>
<tr>
<td>Solyc01g079890.2.1</td>
<td>-</td>
<td>Down</td>
<td>-</td>
<td>Aquaporin</td>
</tr>
<tr>
<td>Solyc06g005820.2.1</td>
<td>-</td>
<td>-</td>
<td>-5.40</td>
<td>Copper transporter</td>
</tr>
<tr>
<td>Solyc02g086920.1.1</td>
<td>-</td>
<td>-</td>
<td>-4.19</td>
<td>Nodulin like protein</td>
</tr>
<tr>
<td>Solyc02g083510.2.1</td>
<td>-</td>
<td>-</td>
<td>-4.17</td>
<td>Aquaporin</td>
</tr>
<tr>
<td>Solyc01g104780.2.1</td>
<td>-</td>
<td>-</td>
<td>-3.91</td>
<td>Nodulin like protein</td>
</tr>
<tr>
<td>Solyc03g095900.2.1</td>
<td>-</td>
<td>-</td>
<td>-3.50</td>
<td>Oxoglutarate and iron-dependent oxygenase</td>
</tr>
<tr>
<td>Solyc03g080070.2.1</td>
<td>-</td>
<td>-</td>
<td>-3.11</td>
<td>Nodulin like protein</td>
</tr>
<tr>
<td>Solyc06g076930.1.1</td>
<td>-</td>
<td>-</td>
<td>-3.01</td>
<td>ABC-2 type transporter</td>
</tr>
<tr>
<td>Solyc04g072760.2.1</td>
<td>-</td>
<td>-</td>
<td>-2.13</td>
<td>High affinity sulfate transporter</td>
</tr>
<tr>
<td>Solyc01g103140.2.1</td>
<td>-</td>
<td>-</td>
<td>-1.74</td>
<td>Voltage-gated chloride channel</td>
</tr>
<tr>
<td>Solyc02g093650.2.1</td>
<td>-</td>
<td>-</td>
<td>Down</td>
<td>Early nodulin like protein</td>
</tr>
<tr>
<td>Solyc03g119880.2.1</td>
<td>-</td>
<td>-</td>
<td>Down</td>
<td>Early nodulin like protein</td>
</tr>
<tr>
<td>Solyc04g011340.2.1</td>
<td>-</td>
<td>-</td>
<td>Down</td>
<td>Early nodulin like protein</td>
</tr>
<tr>
<td>Solyc09g005970.1.1</td>
<td>-</td>
<td>-</td>
<td>Down</td>
<td>ABC transporter</td>
</tr>
</tbody>
</table>
Table 9. Genes encoding pollen wall or coat proteins down-regulated in \textit{ms10}^{35}.

<table>
<thead>
<tr>
<th>ID</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solyc02g032910.1.1</td>
<td>-8.81</td>
<td>-</td>
<td>-</td>
<td>Glycine rich protein</td>
</tr>
<tr>
<td>Solyc11g042810.1.1</td>
<td>-8.54</td>
<td>-</td>
<td>-</td>
<td>Glycosyl transferase</td>
</tr>
<tr>
<td>Solyc02g065670.1.1</td>
<td>-5.67</td>
<td>-</td>
<td>-</td>
<td>UDP-glucosyltransferase</td>
</tr>
<tr>
<td>Solyc11g072680.1.1</td>
<td>-5.27</td>
<td>-</td>
<td>-</td>
<td>Fasciclin-like arabinogalactan protein</td>
</tr>
<tr>
<td>Solyc02g084980.2.1</td>
<td>-3.68</td>
<td>-4.32</td>
<td>-</td>
<td>Glycosyl transferase</td>
</tr>
<tr>
<td>Solyc01g008100.2.1</td>
<td>-1.91</td>
<td>-1.65</td>
<td>-</td>
<td>Glycosyl transferase</td>
</tr>
<tr>
<td>Solyc09g008510.1.1</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>UDP-glucosyltransferase</td>
</tr>
<tr>
<td>Solyc02g068400.2.1</td>
<td>-</td>
<td>-6.21</td>
<td>-</td>
<td>Pectin lyase</td>
</tr>
<tr>
<td>Solyc12g010920.1.1</td>
<td>- Down</td>
<td>Down</td>
<td>Down</td>
<td>Oleosin</td>
</tr>
<tr>
<td>Solyc06g064520.2.1</td>
<td>-</td>
<td>-</td>
<td>-6.40</td>
<td>Mannan endo-1 4-beta-mannosidase</td>
</tr>
<tr>
<td>Solyc02g091700.2.1</td>
<td>-</td>
<td>-</td>
<td>-2.51</td>
<td>Hydroxyproline-rich glycoprotein</td>
</tr>
<tr>
<td>Solyc09g097770.2.1</td>
<td>-</td>
<td>-</td>
<td>-2.17</td>
<td>Glycine rich cell wall protein</td>
</tr>
<tr>
<td>Solyc07g065070.1.1</td>
<td>-</td>
<td>-</td>
<td>Down</td>
<td>Wax synthase isoform</td>
</tr>
</tbody>
</table>
Table 10. Known genes down regulated genes for pollen development in *ms10*.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
<th>Putative function</th>
<th>Regulation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solyc03g116930.2.1</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>Sister chromatid cohesion</td>
<td>Meiosis</td>
<td>SWI1 in Arabidopsis (Mercier et al., 2003)</td>
</tr>
<tr>
<td>Solyc02g032910.1.1</td>
<td>-6.81</td>
<td>-</td>
<td>-</td>
<td>Glycine rich protein</td>
<td>LeGRP92 in tomato (McNeil and Smith, 2010)</td>
<td></td>
</tr>
<tr>
<td>Solyc11g072680.1.1</td>
<td>-5.27</td>
<td>-</td>
<td>-</td>
<td>Fasciclin-like arabinogalactan</td>
<td>Tapetum</td>
<td>MTR1 in rice (Tan et al., 2012)</td>
</tr>
<tr>
<td>Solyc06g064480.2.1</td>
<td>Down</td>
<td>-</td>
<td>-</td>
<td>Protein TAP1</td>
<td>TGA5100 in tomato (van der Huebel et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>Solyc02g068400.2.1</td>
<td>-</td>
<td>-6.21</td>
<td>-</td>
<td>Pectin lyase</td>
<td>QRT3 in Arabidopsis (Rhee et al., 2003)</td>
<td></td>
</tr>
<tr>
<td>Solyc06g069220.1.1</td>
<td>-8.99</td>
<td>-8.64</td>
<td>-</td>
<td>Aspartic proteinase</td>
<td>EAT1 in rice (Niu et al., 2013; Huang et al., 2013)</td>
<td></td>
</tr>
<tr>
<td>Solyc02g078370.1.1</td>
<td>-</td>
<td>-12.93</td>
<td>-11.73</td>
<td>Anther-specific protein</td>
<td>TA29 in tomato (Nawaz-ul-Rehman et al., 2007)</td>
<td></td>
</tr>
<tr>
<td>Solyc07g053460.2.1</td>
<td>-</td>
<td>-13.42</td>
<td>-11.00</td>
<td>C1A Peptidase</td>
<td>Cystein protease in tomato (Santorete et al., 2009)</td>
<td></td>
</tr>
<tr>
<td>Solyc05g009230.1.1</td>
<td>-5.90</td>
<td>-</td>
<td>-</td>
<td>MYB transcription factor</td>
<td>Transcription factor</td>
<td>AtMYB21 in Arabidopsis (Shin et al., 2002)</td>
</tr>
<tr>
<td>Solyc02g079810.1.1</td>
<td>Down</td>
<td>-</td>
<td>-</td>
<td>bHLH transcription factor</td>
<td>DY1, UDT1 (Zhang et al., 2006; Jung et al., 2006)</td>
<td></td>
</tr>
<tr>
<td>Solyc03g059200.1.1</td>
<td>Down</td>
<td>-</td>
<td>-</td>
<td>MYB transcription factor</td>
<td>TDF1 in Arabidopsis (Zhu et al., 2008)</td>
<td></td>
</tr>
<tr>
<td>Solyc08g062780.1.1</td>
<td>Down</td>
<td>-9.47</td>
<td>-</td>
<td>bHLH transcription factor</td>
<td>AMS in Arabidopsis (Sorensen et al., 2003)</td>
<td></td>
</tr>
<tr>
<td>Solyc04g008420.1.1</td>
<td>Down</td>
<td>-</td>
<td>-</td>
<td>PHD finger protein</td>
<td>MS1 in Arabidopsis (Vizay-Barrena et al., 2006)</td>
<td></td>
</tr>
<tr>
<td>Solyc01g009590.2.1</td>
<td>Down</td>
<td>Down</td>
<td>-</td>
<td>Plant lipid transfer protein</td>
<td>Pollen wall</td>
<td>TomA108 in tomato (Chen et al., 2006)</td>
</tr>
<tr>
<td>Solyc01g008100.2.1</td>
<td>-1.91</td>
<td>-1.65</td>
<td>-</td>
<td>Glycosyl transferase</td>
<td>Osg1 (Moon et al., 2013)</td>
<td></td>
</tr>
<tr>
<td>Solyc01g095780.2.1</td>
<td>-</td>
<td>-9.87</td>
<td>Down</td>
<td>Plant lipid transfer protein</td>
<td>TomA108 in tomato (Chen et al., 2006)</td>
<td></td>
</tr>
<tr>
<td>Solyc03g051960.2.1</td>
<td>-11.11</td>
<td>Down</td>
<td>Down</td>
<td>Fatty acyl coA reductase</td>
<td>MS2 in Arabidopsis (Chen et al., 2011)</td>
<td></td>
</tr>
<tr>
<td>Solyc05g056170.2.1</td>
<td>-</td>
<td>-</td>
<td>-1.56</td>
<td>Phenylalanine ammonia-lyase</td>
<td>Phe in broccoli (Kishitani et al., 1993)</td>
<td></td>
</tr>
<tr>
<td>Solyc06g064520.2.1</td>
<td>-</td>
<td>-</td>
<td>-6.40</td>
<td>Endo-1,4-β-mannosidase</td>
<td>LeMAN5 in tomato (Filichkin et al., 2004)</td>
<td></td>
</tr>
</tbody>
</table>
**Fig. 11.** Expression patterns of *Ms10^{ss}\) gene regulated by putative up- and down-stream genes.

<table>
<thead>
<tr>
<th>Male sterile</th>
<th>Male fertile</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>S2</td>
</tr>
<tr>
<td>DYT1</td>
<td></td>
</tr>
<tr>
<td>TomA108</td>
<td></td>
</tr>
<tr>
<td>TGAS100</td>
<td></td>
</tr>
<tr>
<td>LeGRP92</td>
<td></td>
</tr>
<tr>
<td>SGN-U321072</td>
<td></td>
</tr>
<tr>
<td>TA29</td>
<td></td>
</tr>
<tr>
<td>LeMAN5</td>
<td></td>
</tr>
<tr>
<td>pLAT52</td>
<td></td>
</tr>
<tr>
<td>TAGL1</td>
<td></td>
</tr>
<tr>
<td>DEF1</td>
<td></td>
</tr>
<tr>
<td>Actin</td>
<td></td>
</tr>
</tbody>
</table>
Fig. 12. Visualized expression patterns of up- and down-regulated genes at different stages between male fertile and sterile plant.
Discussion

Male sterility in \textit{ms10}^{35} was demonstrated to be caused by impaired pollen development process of both meiocyte and tapetum at the early stage. \textit{Ms10}^{35} was identified to encode a bHLH transcription factor and showed homologues relationship with \textit{AtDYT1} and \textit{OsUDT1}. By comparative analysis of transcriptome, meiotic process and tapetum differentiation were reduced in \textit{ms10}^{35} mutant indicating that \textit{Ms10}^{35} plays an important role for tapetum development and completion of meiosis.

Male sterility of the \textit{ms10}^{35} gene fulfills requirements as a maternal line for producing F\textsubscript{1} hybrid because of its stability of male sterility and morphological characteristics (Gorman et al., 1997). Among the tomato sterile mutants, the sterility of \textit{ms10}^{35} is the most stable. For example, \textit{ms-5}, \textit{ms-7}, \textit{ms-16} appeared to produce viable pollens even though the pollen development in these mutants were also aborted during meiosis. Moreover, \textit{ms10}^{35} can be selected with a linked morphological marker, absence of anthocyanin (aa) on cotyledonary hypocotyl (Atanassova and Georgiev, 1986). This characteristic is useful to discard unwanted male fertile lines during maintenance of male sterile line because male fertile individuals could be easily recognized as early as the cotyledon stage. However, absence of anthocyanin was not observed in male sterile line used in this study. Furthermore, due to protruded stigma, a typical characteristic of \textit{ms10}^{35}
mutant, cross-pollination can be directly performed without anther cone removal (Levin et al., 1994).

The Se2.1 locus located near ms10$^{35}$ on chromosome 2, which is one of the QTLs consisting 5 different genes responsible for exerted stigma was isolated (Chen and Tanksley, 2004). In the Se2.1 locus, Style 2.1 was identified as a bHLH transcription factor containing 92 amino acids and involved in cell elongation (Chen et al., 2007). To test whether the decrease of anther length in ms10$^{35}$ was related with Style 2.1 in a F2 population, however, phenotypic variation was not correlated with Style 2.1, indicating that stigma protrusion was an pleiotrophic effect of ms10$^{35}$ (Data not shown). In trascriptome analysis, Style2.1 (Solyc02g087860.2.1) was reduced in the significant area at stage 1 (Table 6). However, in stage 2 and stage 3, there was not significant difference between male fertile and sterile plant, indicating that the elongation of stamen length appeared to be affected by Ms10$^{35}$ at early stage. However, direct evidence how this pleiotropic effect occurred in ms10$^{35}$ was not determined. Considering complementary transgenic line showed normal flower shape and anther length, Ms10$^{35}$ might regulate elongation of stamen cell by itself or by regulation of down-stream gene which is responsible for stamen development.

Proper meiosis and differentiation of sporophytic cell layers including tapetum are essential for successful pollen development in plant. Success of meiosis depends on tapetum differentiation in Arabidopsis. Tapetum-specific
expressed genes such as *EMS1/EXS*, *TDF1*, and *DYT1* are required for cytokinesis at anaphase I after chromosomal segregation (Canales et al., 2002; Yang et al., 2003; Zhang et al., 2006; Zhao et al., 2002). In *ems1/exs* and *tdf1* mutant, meiosis could be initiated meiosis but could not be completed due to failure of cytokinesis (Canales et al., 2002; Zhao et al., 2002). In *dyt1* mutant, meiocyte was not able to complete cytokinesis process and tetrad cells were not developed normally (Zhang et al., 2006). In addition, meiocyte did not produce tetrads in *udt1* mutant, suggesting that meiocytes could not complete meiosis II stage (Jung et al., 2005). Transcript of the meiosis-related gene, *PAIR1* was significantly down-regulated, suggesting that homologues chromosome pairing might be affected in *udt1*. However, precise mechanism has not discovered.

Considering close relationship with *DYT1* and *UDT1*, *Ms10³⁵* is also necessary to support completion of meiosis in tomato. The establishment of sister chromatid cohesion and the segregation by organization of spindle network was important for normal haploid production (Mercier et al., 2003). *Swi1* and *OsREC8* were a responsible gene for chromatid cohesion, respectively. In each mutant, unequal homolous paring and chromosome segregation at the metaphase I were observed and resulted in male sterility. In *ms10³⁵* mutant, chromosome rearrangement along the division plate was abnormally stuck together (Fig. 5N). Abnormal chromosome division at the anaphase II in *ms10³⁵* might result in remanent trace of meiocytes in the mutant locules due to impaired sister
chromatid cohesion. In RNAseq result, a chromatid cohesion protein (Solyc03g116940.1.1) was significantly reduced from stage 1 to stage 3 in significant level ($P < 0.05$).

Another significant defect of pollen development in $ms10^{35}$ mutant is abnormal tapetum enlargement and vacuolation at the time of meiosis. Vacuolated tapetal cell expansion was commonly observed in the $dytl$ and $udtl$ mutant (Jung et al., 2005; Zhang et al., 2006). Late maturation or delayed degradation of tapetal cells caused abnormal pollen development in plants (Li et al., 2006; Li et al., 2011; Niu et al., 2013; Yi et al., 2012). In $ms10^{35}$ mutant, nucleus and organelles in tapetum were disappeared rapidly during meiosis stage. Continuously, microspore abortion, inhibition of tapetum PCD, defect of middle cell layer degeneration was observed. In transcriptome analysis, several proteases such as a cysteine protease (Solyc06g069220.1.1) and an asparatic protease (Solyc07g053460.2.1) were highly down-regulated, indicating that $Ms10^{35}$ was presumed to activate and/or inhibit PCD in tapetum by regulation of several proteases. Among down-regulated genes in $ms10^{35}$ mutaant, genes for cell wall modification and degradation were most abundant, suggesting that $Ms10^{35}$ might contribute pollen development through triggering PCD operation system since the early stage of pollen development.

The other major role of $Ms10^{35}$ is supporting for pollen wall deposition by delivering materials and nutrient toward developing microspore. In wild type,
orbicule was successfully generated along the degenerated tapetum (Fig. 3J). Even thought presice function has not been identified, sporopollenin materials and nutrients are delivered toward microspore throught this molecule. However, in \textit{ms10}^{35} mutant, aberrant meiocyte and abnormal tapetum never allowed to produce orbicule. In addition, no normal sporopollenin structure was observed, indicating that pollen exine development presumed to be completely blocked. Lipid transferases, phenylalanine lyase, long chain fatty acid reductases, fatty acid hydroxylase and other several catalytic activity genes which are involved in sporopollenin biosynthesis, were down-regulated in \textit{ms10}^{35} mutant. In addition, oleosin (Solyc12g010920.1.1) and lipases (Solyc04g081780.2.1 and Solyc11g065530.1.1) were also inhibited their expression. Besides, a glycine-rich protein, LeGPR92 (Solyc02g032910.1.1), which are important for sporopollenin deposition and exine formation from tetrad and microspore stage was down-regulated. These results demonstrated that \textit{Ms10}^{35} also required for accumulation of lipids and coat protein onto primexine during pollen maturation. Together with morphological defect and transcriptome analysis, exine deposition was also highly inhibited in \textit{ms10}^{35} mutant. Transcription factors which showed high similarity with \textit{TDF1}, \textit{AMS}, \textit{MS1}, and \textit{MYB35} in \textit{Arabidopsis} were dramatically down-regulated. In \textit{Arabidopsis}, \textit{DYT1} regulate directly or indirectly a number of genes through forming a transcription cascade with those transcription factors which were confirmed by in protein-protein interaction assay and in vitro binding assay.
This regulation system is proved to exist in rice (Jung et al., 2006; Li et al., 2006; Niu et al., 2013). Therefore, a transcription factor cascade of DYT1 was presumed to be conserved in higher plant. In this study, even though direct evidence for regulation and controlling of these transcription factors has not been performed, the Ms10^{35} gene may involved in the stamen development through interaction with these transcription factor.

In summary, the most dramatic cytological changes observed in the ms10^{35} mutant can be summarized as follows: 1) breakdown meiocyte during meiosis process in sporogenous cells, 2) excessive vacuolation and expansion of tapetal cells, and 3) inhibition of nutrient delivery and sporopollenin material to developing microspore. These phenotypic defects were commonly observed in mutation of homologous genes of Arabidopsis and rice (Jung et al., 2005; Zhang et al., 2006). The Ms10^{35} gene encoded a bHLH transcription factor, and specifically expressed in anther. Histological and transcriptome analysis revealed that Ms10^{35} may function as a core factor to control meiosis at the sister chromatid cohesion and multiple progress of tapetum condensation, PCD, and cell wall development. Ms10^{35} plays a key role in regulatory of meiotic process and tapetum development at the early developmental stage. These results confirm that meiotic division and tapetal development is tightly related (Ma, 2005; Zhu et al., 2008). These findings suggested that a single transcription factor Ms10^{35} controls
both meiotic and sporogenous cell development. This research will provide a molecular clue of male sterility not only in tomato but also in other crop plants.
References


in tomato flowers. Heredity 73:72-77.


초 록

가지과 작물에 속하는 토마토와 고추는 전세계적으로 가장 중요한 채소 작물이다. 토마토와 고추 육종에서 가장 중요한 목표는 수량이 높은 품종, 병해충으로부터 저항성을 보이는 품종, 그리고 고품질의 품종을 육성하는 것이다. 하지만, 토마토와 고추의 상용 품종은 유용한 유전 자원이 제대로 활용되지 못해 유전적 다양성이 낮은 문제점이 있다. 그러므로, 다양한 유전 자원을 육종에 활용할 수 있도록 유전적 다양성에 대한 연구와 유용 유전자에 대한 활용하는 육종 전 과정(prebreeding)에 대한 중요성이 더욱 높아지고 있다.

제1장에서는, 고추 종을 구별하는 분자 표지를 개발하였다. 표현형을 기반으로 한 고추의 종구별과 초기 분자 표지를 활용한 종 구별에 대한 연구가 많이 있었다. 하지만, 이들은 분석이 모호하거나, 재현성이 부족하고 다량의 분석에는 적합하지 않은 문제점이 있었다. 따라서, COSII 분자 표지와 Waxy 유전자에서 SNP 분자 표지를 개발하여 종을 구별하고자 하였다. 이를 통하여, 5개의 재배종과 1개의 야생종을 구별할 수 있는 분석 방법을 통해 고추 종에 대한 기원을 연구하였다. 이 장에서 연구한 내용은 고추 종에 대한 분석 방법을 제시하는 동시에 유전적 근연 관계를 밝히는 연구 자료가 될 것이다.

제2장에서는, 고추의 유전 자원과 EMS를 처리하여 유기한 돌연변이 집단에서 고추 바이러스에 대해 저항성을 가지는 유용 유전자를 찾는 allele mining에 대한 연구를 하였다. 특히, 이러한 과정을 위하여 신속하고 높은 효율로 유전자를 감별할 수 있는 HRM 분석 방법을 이용하였다. 전체 248개의 고추 유전 자원을 분석하여 TEV-HAT 바이러스에 저항성에 연관된 eIF4E의 유전적 다양성을
연구하여 13개의 대립 유전자로 분류하였으며, 이 중 C. baccatum의 계통인 C03946(pvrHRM112)가 바이러스에 강한 저항성을 보였다. 또한, EMS에 유기된 M1 집단에서 동일한 유전자를 통해 다양한 대립 유전자를 찾아냈다. 이러한 결과를 토대로 고추의 유전 자원의 다양성을 확보하고 육종에 활용이 가능한 유용 유전자를 동정하는 방법에 대한 초석이 될 것으로 기대한다.

제3장에서는, 토마토 웅성 불임 ms1035의 유전자 동정을 수행하였다. 웅성 불임은 육종에 있어서 가장 중요한 형질 중에 하나이지만, 토마토에서 웅성 불임에 대한 연구는 많이 이루어지지 않았다. 지도 기초 유전자 동정 방법을 이용하여 토마토 ms1035는 bHLH 전사 인자 중 하나를 암호화하고 있는 것으로 밝혀졌다. 특히, 표현형적인 변이와 분자 유전학적인 방법을 통해 밝혀진 바에 따르면, ms1035는 화분 초기 형성 과정 중에서 감수 분열과 융단 조직 발달에 특히 관여하는 것으로 보인다. ms1035는 토마토 화분 형성에 관련된 유전자 중 처음으로 동정된 것으로 앞으로 화분 발달에 대한 연구와 웅성 불임성을 이용한 육종에 중요한 소재가 될 것으로 기대된다.

주요어: 고추, 대립 유전자 감별, 웅성 불임, 종 구별, 지도 기초 유전자 동정법, 토마토, 화분 발달상