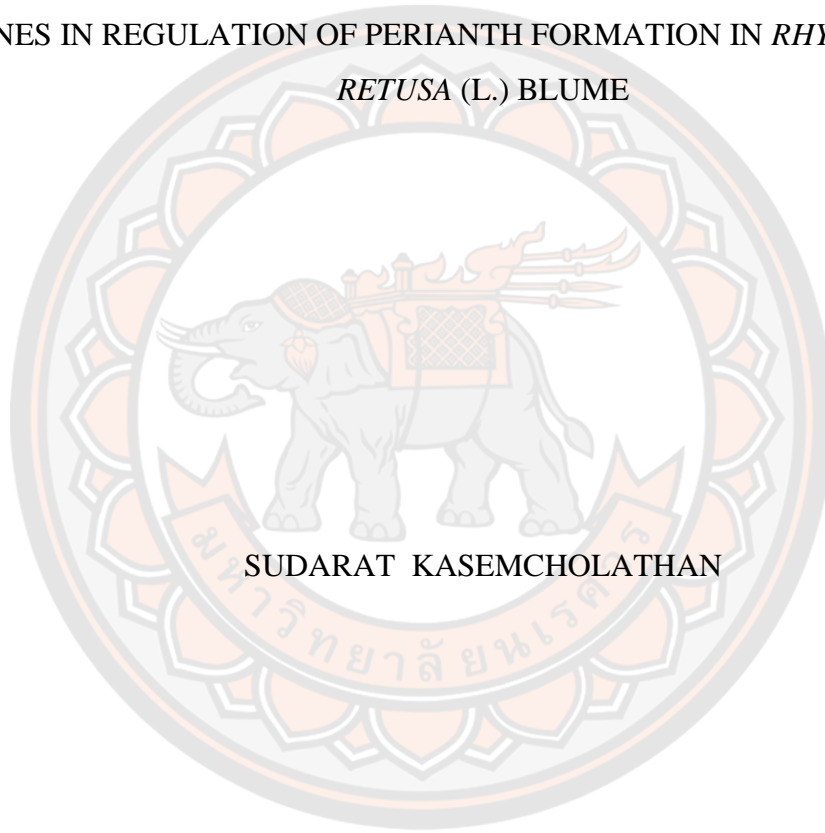




ISOLATION AND EXPRESSION ANALYSIS OF B- AND E-CLASS MADS BOX
GENES IN REGULATION OF PERIANTH FORMATION IN *RHYNCHOSTYLIS*
RETUSA (L.) BLUME



SUDARAT KASEMCHOLATHAN

A Thesis Submitted to the Graduate School of Naresuan University
in Partial Fulfillment of the Requirements
for the Master of Science in (Biotechnology)

2020

Copyright by Naresuan University

ISOLATION AND EXPRESSION ANALYSIS OF B- AND E-CLASS MADS BOX
GENES IN REGULATION OF PERIANTH FORMATION IN *RHYNCHOSTYLIS*
RETUSA (L.) BLUME



SUDARAT KASEMCHOLATHAN

A Thesis Submitted to the Graduate School of Naresuan University
in Partial Fulfillment of the Requirements
for the Master of Science in (Biotechnology)

2020

Copyright by Naresuan University

Thesis entitled "Isolation and expression analysis of B- and E-class MADS Box genes
in regulation of perianth formation in *Rhynchosyilis retusa* (L.) Blume"

By SUDARAT KASEMCHOLATHAN

has been approved by the Graduate School as partial fulfillment of the requirements
for the Master of Science in Biotechnology of Naresuan University

Oral Defense Committee

..... Chair
(Assistant Professor Kullanart Obsuwan, Ph.D.)

..... Advisor
(Assistant Professor Pattamon Sangin, Ph.D.)

..... Co Advisor
(Assistant Professor Anupan Kongbangkerd, Dr.rer.nat.)

..... Internal Examiner
(Assistant Professor Kittisak Buddhachat, Ph.D.)

Approved

.....
(Paisarn Muneesawang, Ph.D.)

Dean of the Graduate School

Title	ISOLATION AND EXPRESSION ANALYSIS OF B- AND E-CLASS MADS BOX GENES IN REGULATION OF PERIANTH FORMATION IN <i>RHYNCHOSTYLIS RETUSA</i> (L.) BLUME
Author	SUDARAT KASEMCHOLATHAN
Advisor	Assistant Professor Pattamon Sangin, Ph.D.
Co-Advisor	Assistant Professor Anupan Kongbangkerd, Dr.rer.nat.
Academic Paper	M.S. Thesis in Biotechnology, Naresuan University, 2020
Keywords	DEF-like genes, AGL6-like genes, Floral development, MADS- box genes, <i>Rhynchostylis retusa</i>

ABSTRACT

Rhynchostylis retusa (L.) Blume is a popular and scarce orchid species with attractive flowers arranged in a racemose cluster. To study this flower's complex floral perianth formation, MADS-box genes that play an important role in flower morphogenesis were investigated. Four B-class *DEF*-like MADS-box genes (*RrDEF1*, *RrDEF2*, *RrDEF3*, and *RrDEF4*) and two E-class *AGL6*-like MADS-box genes (*RrAGL6-1* and *RrAGL6-2*) were identified, and their expression was characterized by qRT-PCR. These genes were expressed in every developmental stage of reproductive organs but not in the vegetative leaves, except for *RrDEF2*. Most of these genes revealed similar expression patterns in stage 3 and mature flowers. Higher *RrDEF1* and *RrAGL6-1* expression was detected in the sepals and petals rather than the lips, while *RrDEF3*, *RrDEF4*, and *RrAGL6-2* were highly expressed in the lips but exhibited low expression in the sepals. *RrDEF3* was also strongly expressed in petals. These gene expression patterns supported current profiles involving sepal/petal/lip determination, with *RrDEF1* and *RrAGL6-1* promoting sepal/petal formation, while *RrDEF3*, *RrDEF4*, and *RrAGL6-2* promoted lip formation. Our findings support revised 'orchid code' and 'perianth code' hypotheses in which floral organs are regulated by combined levels of expression from each clade of *DEF*-like genes and *AGL6*-like genes. These four *RrDEF* and two *RrAGL6* genes promote and enrich our knowledge and understanding concerning the floral development of orchids.



ACKNOWLEDGEMENTS

This thesis will never be complete and successful without the valuable helps and kindness from these people whom I would like to express my sincerely thanks. First of all, I would like to thanks my advisor, Assistant Professor Dr. Pattamon Sangin for her precious suggestions from the first day of this project was started and for her kindly supports in every single thing to me. I would also express my sincere thanks to my co-advisor, Assistant Professor Anupan Kongbangkerd who always offer his expertise in orchid and support the samples using in this project. I also deliver my truly thanks to Associate Professor Dr. Akira Kanno for providing the guidance, taking care during my time and experiment has performed at Graduate School of Life Science, Tohoku University.

I would like to express my special thanks to the committee members, Assistant Professor Dr. Kullanart Obsuwan and Assistant Professor Dr. Kittisak Buddhachat for their suggestions and encouragement. I wish to acknowledge the Department of Biology, Faculty of Science, Naresuan University, Phitsanulok, Thailand providing laboratory facilities and all the technician for their merciful help. I would like to thanks Science Achievement Scholarship of Thailand (SAST) and Short-term Overseas Research Scholarship from Faculty of Science, Naresuan University and for their funding.

I am extremely grateful to my parents and sister for their unconditional love and sincerely understanding that started from the ground up. My completion of university life could not have been accomplished without the support of my close friends, Atcharapohn Jai-uean and Siriphat Muangpa for enjoying and overcoming all the obstacle together. Finally, I would like to thanks all the cups of coffee and cats especially Neptune and Pluto, to make my day happier and special.

Some part of this thesis was reproduced with permission from Springer Nature

SUDARAT KASEMCHOLATHAN

TABLE OF CONTENTS

	Page
ABSTRACT.....	C
ACKNOWLEDGEMENTS.....	E
TABLE OF CONTENTS.....	F
List of tables.....	H
List of figures.....	I
CHAPTER 1.....	1
INTRODUCTION.....	1
Background.....	1
Statements of the problem.....	3
Purposes of the study.....	3
Significance of the study.....	3
Scopes of the study.....	4
Hypothesis.....	4
CHAPTER II.....	5
LITERATURE REVIEWS.....	5
Orchid flowers.....	5
MADS-box genes.....	12
Floral organ identity.....	19
The important of <i>DEF</i> -like and <i>AGL6</i> -like in floral organ specification.....	30
The orchid MADS-box genes.....	37
CHAPTER III.....	47
RESEARCH METHODOLOGY.....	47
Plant materials.....	47
Chemicals and instruments.....	48
RNA extraction.....	50

RNA analysis	50
First strand cDNA synthesis	51
PCR amplification of partial sequence of <i>DEF</i> -like and <i>AGL6</i> -like genes	52
3'-Rapid amplification of cDNA ends (3'RACE)	53
Gene cloning	55
Transformation of plasmid DNA into <i>E. coli</i> using the heat shock method.....	57
Sequence data analysis	59
Expression analysis.....	60
CHAPTER IV	64
RESULTS AND DISCUSSION	64
RNA extraction	64
First partial sequences of <i>DEF</i> -like and <i>AGL6</i> -like genes from <i>R. retusa</i>	65
The 3'-Rapid amplification of cDNA ends of <i>DEF</i> -like and <i>AGL6</i> -like genes from <i>R. retusa</i>	68
Partial coding sequence of <i>DEF</i> -like and <i>AGL6</i> -like genes in <i>R. retusa</i>	74
Phylogenetic analysis.....	75
Expression analysis.....	79
CHAPTER V	88
CONCLUSION.....	88
REFERENCES	89
APPENDIX.....	92
BIOGRAPHY	112

List of tables

	Page
Table 1 Recent Class B and E MADS-box genes characterized in Orchidaceae	39
Table 2 List of oligonucleotide sequences.....	62
Table 3 Purity of total RNA from <i>R. retusa</i> using spectrophotometer	64
Table 4 The accession numbers of four DEF-like genes and two AGL6-like	73
Table 5 BLAST® results of first partial DEF-like clade 1 nucleotide sequence in <i>R. retusa</i> . The result was shown first 10 highest total scores.	100
Table 6 BLAST® results of first partial DEF-like clade 2 nucleotide sequence in <i>R. retusa</i> . The result was shown all sequences	101
Table 7 BLAST® results of first partial DEF-like clade 3 nucleotide sequence in <i>R. retusa</i> . The result was shown first 10 highest total scores.	102
Table 8 BLAST® results of first partial DEF-like clade 4 nucleotide sequence in <i>R. retusa</i> . The result was shown all sequences	103
Table 9 BLAST® results of first partial AGL6-like clade 1 nucleotide sequence in <i>R. retusa</i> . The result was shown first 10 highest total scores.	104
Table 10 BLAST® results of first partial AGL6-like clade 2 nucleotide sequence in <i>R. retusa</i> . The result was shown first 10 highest total scores.	105
Table 11 BLAST® results of final partial DEF-like clade 1 nucleotide sequence in <i>R. retusa</i> . The result was shown first 10 highest total scores.	106
Table 12 BLAST® results of final partial DEF-like clade 2 nucleotide sequence in <i>R. retusa</i> . The result was shown first 12 highest total scores.	107
Table 13 BLAST® results of final partial DEF-like clade 3 nucleotide sequence in <i>R. retusa</i> . The result was shown first 12 highest total scores.	108
Table 14 BLAST® results of final partial DEF-like clade 4 nucleotide sequence in <i>R. retusa</i> . The result was shown all sequences which.....	109
Table 15 BLAST® results of final partial AGL6-like clade 1 nucleotide sequence in <i>R. retusa</i> . The result was shown first 10 highest total scores.	110
Table 16 BLAST® results of final partial AGL6-like clade 2 nucleotide sequence in <i>R. retusa</i> . The result was shown first 10 highest total scores.	111

List of figures

	Page
Figure 1 Schematic diagrams of <i>Arabidopsis thaliana</i> and <i>Phalaenopsis</i>	6
Figure 2 Diagrams of orchid floral homeotic variants. Normal flower at	8
Figure 3 The gene duplication fate in evolution	10
Figure 4 The comparison of <i>Rhynchosstylis retusa</i> flower between wild type.....	11
Figure 5 (a) The origin of two major MADS-domain by gene duplication in.....	13
Figure 6 (a) the crystal structure of Human SRF MADS-domain binding to.....	15
Figure 7 (a) A schematic structure of type II MIKC-MADS protein	17
Figure 8 The phylogenetic tree of MADS box gene based on whole genome	19
Figure 9 The simplified phylogenetic tree of Angiosperm based on earliest	21
Figure 10 The classical ABC model and the floral homeotic mutant in Class A, B, C.....	22
Figure 11 The classical ABC model, the ABCDE model and the floral quartet	24
Figure 12 The autoregulation of MIKC-MADS domain.....	26
Figure 13 The flower induction and flower formation network	27
Figure 14 The regulation of floral quartet complex from DNA binding to.....	29
Figure 15 The nucleosome mimicry hypothesis for floral quartet complex.....	30
Figure 16 The consensus sequence in C-terminus indicating PI motif, PI.....	32
Figure 17 The B-class gene lineage duplication events along with land plants	34
Figure 18 The phylogenetic relationship of AGL6 subfamily.....	36
Figure 19 The phylogenetic tree of subfamily in Orchidaceae family. The Hypothesis	38
Figure 20 The orchid code and the hypothesis of the orchid perianth origin.	39
Figure 21 The comparison of previous orchid code model and the revised	43
Figure 22 Perianth code model	44
Figure 23 The summary of P code complexes regulation.....	46
Figure 24 Morphology of <i>R. retusa</i> wild type flower. (a) Developmental stages 1-4.....	47
Figure 25 Map of pJET1.2/blunt cloning vector (Thermo Scientific, USA).....	56
Figure 26 Assessment of RNA integrity by agarose gel electrophoresis stained	65

Figure 27	The PCR amplification of first partial sequences of DEF-like and.....	66
Figure 28	The consensus sequences from colonies sequencing using S-HrDEF	68
Figure 29	The consensus sequences from colonies sequencing using DEF2/FW	68
Figure 30	The consensus sequences from colony sequencing using	68
Figure 31	The physical map of primer using in DEF-like genes isolation in <i>R. retusa</i>	69
Figure 32	The physical map of primer using in AGL6-like genes isolation in <i>R. retusa</i>	69
Figure 33	The 3'RACE products of DEF-like from <i>R. retusa</i> were validated by	70
Figure 34	The 3'RACE products of AGL6-like from <i>R. retusa</i> were validated by	71
Figure 35	The partial sequences of four DEF-like genes in <i>R. retusa</i>	72
Figure 36	The partial sequences of two AGL6-like genes in <i>R. retusa</i>	73
Figure 37	Alignment of the consensus amino acid sequences of the C-terminal	74
Figure 38	Alignment of the consensus amino acid sequences of the C-terminal	75
Figure 39	Phylogenetic analysis of DEF-like gene nucleotide sequences.....	77
Figure 40	Phylogenetic analysis of AGL6-like gene nucleotide sequences	78
Figure 41	The expression pattern of RrDEF-like and RrAGL6-like genes in.....	81
Figure 42	The expression pattern of RrDEF-like and RrAGL6-like genes in.....	82
Figure 43	Expression patterns of four RrDEF and two RrAGL6 genes in different	83
Figure 44	Expression patterns of RrDEF1 genes in various organs of stage 3 and.....	84
Figure 45	Expression patterns of RrDEF2 genes in various organs of stage 3 and.....	84
Figure 46	Expression patterns of RrDEF3 genes in various organs of stage 3 and.....	85
Figure 47	Expression patterns of RrDEF4 genes in various organs of stage 3 and.....	85
Figure 48	Expression patterns of RrAGL6-1 genes in various organs of stage 3 and.....	86
Figure 49	Expression patterns of RrAGL6-2 genes in various organs of stage 3 and.....	87
Figure 50	Multiple sequence alignment of selected clones of amplified S-HrDEF ...	97
Figure 51	Multiple sequence alignment of selected clones of amplified S-HrDEF ...	98
Figure 52	Multiple sequence alignment of selected clones of amplified S-HrDEF ...	99

CHAPTER 1

INTRODUCTION

Background

Orchidaceae is one of the largest flowering plant families with extremely diverse and specialized floral morphology (Aceto & Gaudio, 2011). Orchid flowers consist of three sepals in the outermost whorl, while the petals in the second whorl include two similar lateral sepals and one modified shape which called lip or labellum, with the stamen and pistil fused together as the column in the innermost whorl. Orchids follow a basic morphology, but peloric mutants still occur in floral organs (Bateman & Rudall, 2006b; Mondragón-Palomino & Theissen, 2009; Rudall & Bateman, 2002).

Rhynchostylis retusa (L.), also called foxtail orchid, is a monopodial and epiphytic orchid species with beautiful inflorescent flowers arranged in a racemose cluster. Its unique fragrance makes this orchid species more attractive and valuable (Parab & Sellappan, 2008). Four species in the genus *Rhynchostylis* are endemic to Southeast Asian countries, and three are found in Thailand (Anuttato, Boonruangrod, Kongsamai, & Chanprame, 2017). The *R. retusa* species is in high demand and often featured in orchid competitions. Many studies have indicated that the flower structure has continuously diverged from wild types. However, studies of orchid genetic backgrounds are not common in Thailand.

Molecular genetic mechanisms relating to flower morphology have been extensively studied for more than 20 years in the two core eudicot model plants *Arabidopsis thaliana* and *Antirrhinum majus*, and the established ABCDE flower development model (Theissen et al., 2000; Theißen, Melzer, & Rümpler, 2016; Theißen & Saedler, 2001). In this model, each flower organ identity is determined by a combination of ABCDE class genes: sepal formation is specified by a combination of A- and E-class genes; petal formation is specified by A-, B-, and E-class genes; the stamen is specified by B-, C-, and E-class genes; the carpel is specified by C- and E-class genes; and the ovule is specified by D- and E-class genes. Most of these floral homeotic genes belong to the MIKC type in a superclade of MADS-box genes, which

encode the MADS-box transcription factors. The conserved MIKC domain structure includes a highly conserved MADS (M) domain, intervening (I) domain, keratin-like (K) domain, and the most varied C-terminal (C) domain. Functions of genes in distinct clades of the ABCDE model are different (Becker & Theißen, 2003). Many ABCDE class genes have been isolated and characterized in orchids. In particular, the class B MADS-box genes are essential for identifying the development of petals and stamens. There are two major lineages: *APETALA3 (AP3)/DEFICIENS (DEF)* and *PISTILLATA (PI)/GLOBOSA (GLO)* (different names from the loci of *A. thaliana* and *A. majus*, respectively) (Aceto & Gaudio, 2011). Several B-class genes have been isolated and characterized in orchids, including *Cymbidium* spp., *Dendrobium* spp., *Erycina pusilla*, *Oncidium* ‘Gower Ramsey’, *Phalaenopsis* spp., and *Habenaria radiata* (Hsu et al., 2015; Kim et al., 2007; Lin et al., 2016; Mitoma et al., 2019; Pan, Tsai, & Chen, 2017; Sirisawat, Ezura, Fukuda, Kounosu, & Handa, 2010; Sirisawat, Fukuda, Ezura, & Handa, 2009; Tsai, Kuoh, Chuang, Chen, & Chen, 2004; Xiang et al., 2018; Xu, Teo, Zhou, Kumar, & Yu, 2006).

To better explain the tepal development in orchids, a revised ‘orchid code’ hypothesis, as a combinatorial different expression profile of the four DEF-like genes, is proposed. This hypothesis suggests that the inner lateral tepals (petals) are specified by the high level of clade 1 and 2 and the low level of clade 3 and 4 DEF-like genes. By contrast, the lip formation requires a low level of clade 1 and 2 and a high level of clade 3 and 4 DEF-like genes (Mondragón-Palomino & Theissen, 2011). Accordingly, the formation of higher MADS-box protein complexes determines orchid tepal morphogenesis. Recently, the Perianth (P) code model has been proposed to validate perianth formation in orchids by manipulating the expression between different AP3(DEF)/AGL6 homolog complexes. The heteromeric sepal/petal (SP) complex (OAP3-1/OAGL6-1/OAGL6-1/OPI) determined sepal/petal formation, while the lip (L) complex (OAP3-2/OAGL6-2/OAGL6-2/OPI) is exclusively required for lip formation (Hsu et al., 2015).

Molecular studies on the orchid MADS-box genes have strongly enhanced understanding of flower development mechanisms over the past decade. However, questions remain regarding the evolution and diversification of flower morphology in orchids (Aceto & Gaudio, 2011). Various floral organ identity genes can explain many

orchid perianth formations; however, some have evolved independently. Further extensive study on the genes and analysis of their expressions in many orchids is still required. To better understand perianth development of *R. retusa* at the molecular level, the floral organ identity genes in the B-class and E-class of four *DEF*-like genes and two *AGL6*-like genes, respectively, were isolated, and their phylogenetic relationships and proposed expression profiles in the wild type were studied. This research aimed to increase the understanding of expression patterns and predict the cause of variations or speciation of *R. retusa*.

Statements of the problem

According to the background described above, the following problems has been stated as below.

1. Does *R. retusa* contain *DEF*-like and *AGL6*-like genes?
2. How is the relationship between *DEF*-like and *AGL6*-like genes from *R. retusa* and other homologs?
3. How are the expression levels of *DEF*-like and *AGL6*-like genes in different developmental stages of flower buds and in each floral organs?

Purposes of the study

The statement of problems described above lead to the purpose of this study:

1. To isolate *DEF*-like and *AGL6*-like genes from *R. retusa*.
2. To reconstruct the phylogenetic tree of *DEF*-like and *AGL6*-like genes of *R. retusa* between their homologs.
3. To analyze the expression level of *DEF*-like and *AGL6*-like genes in different developmental stages of flower buds and in each floral organs of *R. retusa*.

Significance of the study

1. This study provides the sequences and expression profiles of the floral organ identity genes in *R. retusa*.
2. This initial research data unlocks and enrich knowledge and understanding the floral development of an orchid.

Scopes of the study

This study focused on the fundamental profile of floral organ identity genes from *R. retusa* including B- and E-class MADS-box genes; *DEF*-like and *AGL6*-like respectively. The nucleotide sequence of *DEF*-like and *AGL6*-like genes from flower buds will be cloned and sequenced. After that, the 3' rapid amplification cDNA ends (3'RACE) were performed, cloned and sequenced. This results in the partial coding sequences. Finally, the expression level was analyzed using qPCR in different developmental stages and each floral organ; sepals, petals, lip and column.

Hypothesis

In last 40 years from the first MADS-box gene isolated, several orchid MADS-box genes have been isolated and characterized rapidly. According to the floral morphological transition in many orchids, all efforts lead to study about class B and E MADS-box genes especially from Epidendroideae subfamily which is the largest ones in Orchid family. Currently, several instances of B-and E-class MADS-box genes in orchids were summarized in “orchid code” and “perianth code” which is the complex pattern of genes involving the perianth formation. These evidences lead this study hypothesis as:

1. Partial sequence of *DEF*-like and *AGL6*-like genes are isolated from *R. retusa*
2. The *DEF*-like and *AGL6*-like genes sequence alignment and phylogenetic tree involve with other orchid homologs and represent their relationship.
3. The expression pattern of *DEF*-like and *AGL6*-like genes explain the perianth formation regarding orchid code and perianth code.

CHAPTER II

LITERATURE REVIEWS

Orchid flowers

Overview

Flower or the reproductive organ of the angiosperm is the most complicated and the most recent in plant lineage diversification which bring flowering plant separate from other land plant (Alvarez-Buylla et al., 2010). The basic morphological structure frequently consists of four types of organs including sepals, petals, stamen and carpel which are arranged in whorl, from whorl 1 to 4 respectively (Figure 1). Understanding how distinct organs have possessed their features has been investigated for decades (Causier, Schwarz-Sommer, & Davies, 2010; Coen & Meyerowitz, 1991; Weigel & Meyerowitz, 1994).

Orchidaceae is one of the largest families of angiosperm. There are approximately 30000 species distributing in five subfamilies; Apostasioideae, Cypridioideae, Vanilloideae, Orchidoideae and Epidendroideae (Aceto & Gaudio, 2011). The orchid flowers have greatly diversified and occupied complex and unique floral morphology which has fascinated the researcher for a long time (Mondragón-Palomino, 2013). Typically, the orchids showed zygomorphic or bilateral symmetry and in first two whorls of the orchid perianth have tepals which is almost identical petaloid organs surrounding the two inner whorl of reproductive organs. The outermost whorl consists of three outer tepals (also termed as 'sepals'). The second whorl consists of three inner tepals (also termed as 'petals'). One of the inner tepal, the median one distinctively differentiated from other two as lip or labellum. The innermost male and female in orchid flowers are fused into a single unit of reproductive organ called 'gynostemium' or 'column'. The abaxial orientation (the lowermost side) of lip is interesting because this organ homologous to the adaxial tepal (the uppermost side) in other monocots flower. The term 'resupination' was used to described this phenomenon which occurred in 180° rotations during orchid flowers development and turned the lip to the bottom position by torsion of the ovary/pedicele. These key innovations facilitated the co-evolution between orchids and pollinators leading to extend species diversity by

the adaptive radiation (Bateman & Rudall, 2006a; Dressler, 1993; Mondragón-Palomino & Theißen, 2008; Mondragón-Palomino & Theissen, 2009)

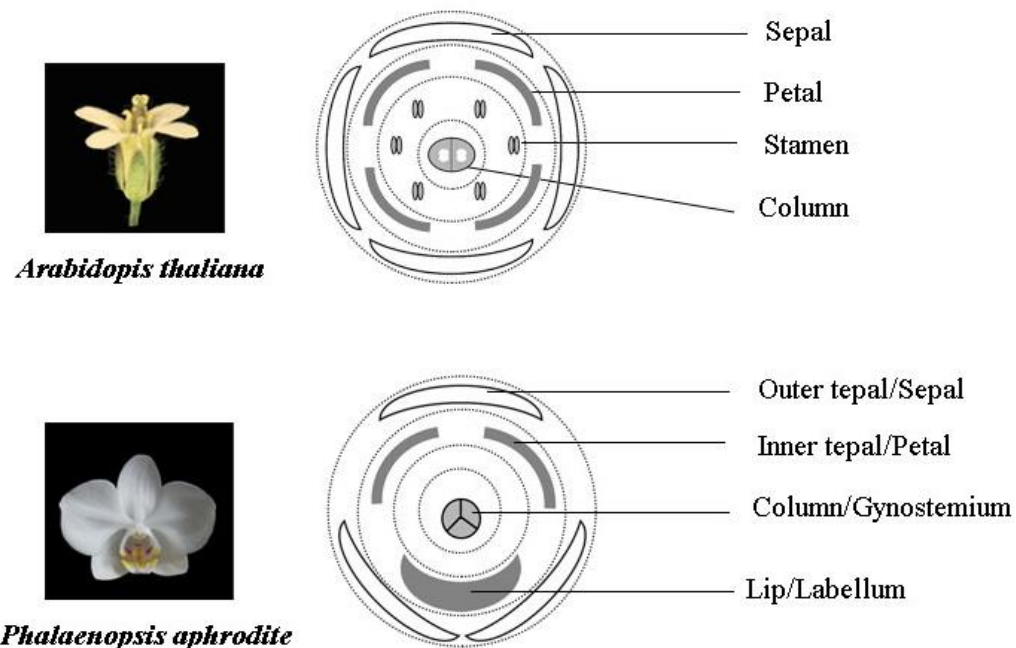


Figure 1 Schematic diagrams of *Arabidopsis thaliana* and *Phalaenopsis aphrodite* representing the morphology of basic of flower and orchid flower.

Source: modified from Alvarez-Buylla et al. (2010); Kanno (2016); Mondragón-Palomino (2013)

The orchid floral homeotic variants

Since late 19th century, Bateman and Rudall suggested the floral terata or phenotypic variant categorization in orchid flowers (Figure 2). There are six types of spontaneous and culture mutants which had been investigated for 25 years in field observation including three types of peloria; type A to C and three types of pseudopeloria; type A to C. The terms ‘peloria’ and ‘pseudopeloria’ depend on the wholly and partially transition of bilateral symmetry to radial symmetry respectively. Except to type A pseudopeloria mutant, all of the rest were homeosis or homeoheterotopy, the organs are either at least partial or complete replacement.

Firstly, type A peloria, both lateral petals are replaced by lip-like organs. These mutant occurred frequently from somaclonal variation in tissue culture. However, there were also 25% of British native orchids that were spontaneously occurred in nature; for example, *Ophrys insectifera* L., *Phalaenopsis equestris* and *Cymbidium goringii*. In type B peloria, the phenotype of lip was replaced by lateral petal. The frequency of this type was less common than type A peloria and there was reported in *Phragmipedium lindenii*, *Ophrys fuciflora* and *Ophrys araneola* Rchb. For type C peloria, all second whorl petals were replaced by sepal-like organs. The possible candidate of this type is in genus *Thelymitra* including *T. cucculata* and *T. formosa*. Consequently, the perianth of peloria type achieved a radial or actinomorphic symmetry rather than bilateral or zygomorphic symmetry. In type B pseudopeloria, the lip transformed to sepal but other organs were consistent. The *Platanthera chlorantha* was exemplified for this category. For type C pseudopeloria, both lateral petals were replaced by sepal. The relevant species is *Epidendrum pseudoepidendrum* (Mondragón-Palomino & Theissen, 2009).

Distinctively from other five categories of perianth transition, type A pseudopeloria is better involve in heterochrony than heterotopy. The temporal change in expression which is better or less expression affect the structure gets larger (termed as peramorphic heterochrony) or smaller (termed as paedomorphic heterochrony) respectively. The example candidates of this type are *Epipactis phyllanthes* var. *phyllanthes* and *Nigritella austriaca* which are categorized as paedomorphic heterochrony. In addition, the existence of transformation has two kind of gradient evidences. First, the event occurred in the same whorl, another one is shifted between two whorls. If the sepals from outer first whorl are shifted toward inner second whorl, will termed acropetal homeosis. In contrast, the second whorl organs are shifted to first whorl position or the pollinaria have existed in labellum or lateral petal, termed basipetal homeosis. Interestingly, all five fundamental types of homeosis are acropetal homeosis. Thus, Type D pseudopeloria is extra-categorized the transformation of outer tepal to lateral inner petal which is basipetal homeosis; for example, *Cattleya alvaroana* (Bateman & Rudall, 2006a; Mondragón-Palomino & Theissen, 2009; Rudall & Bateman, 2002).

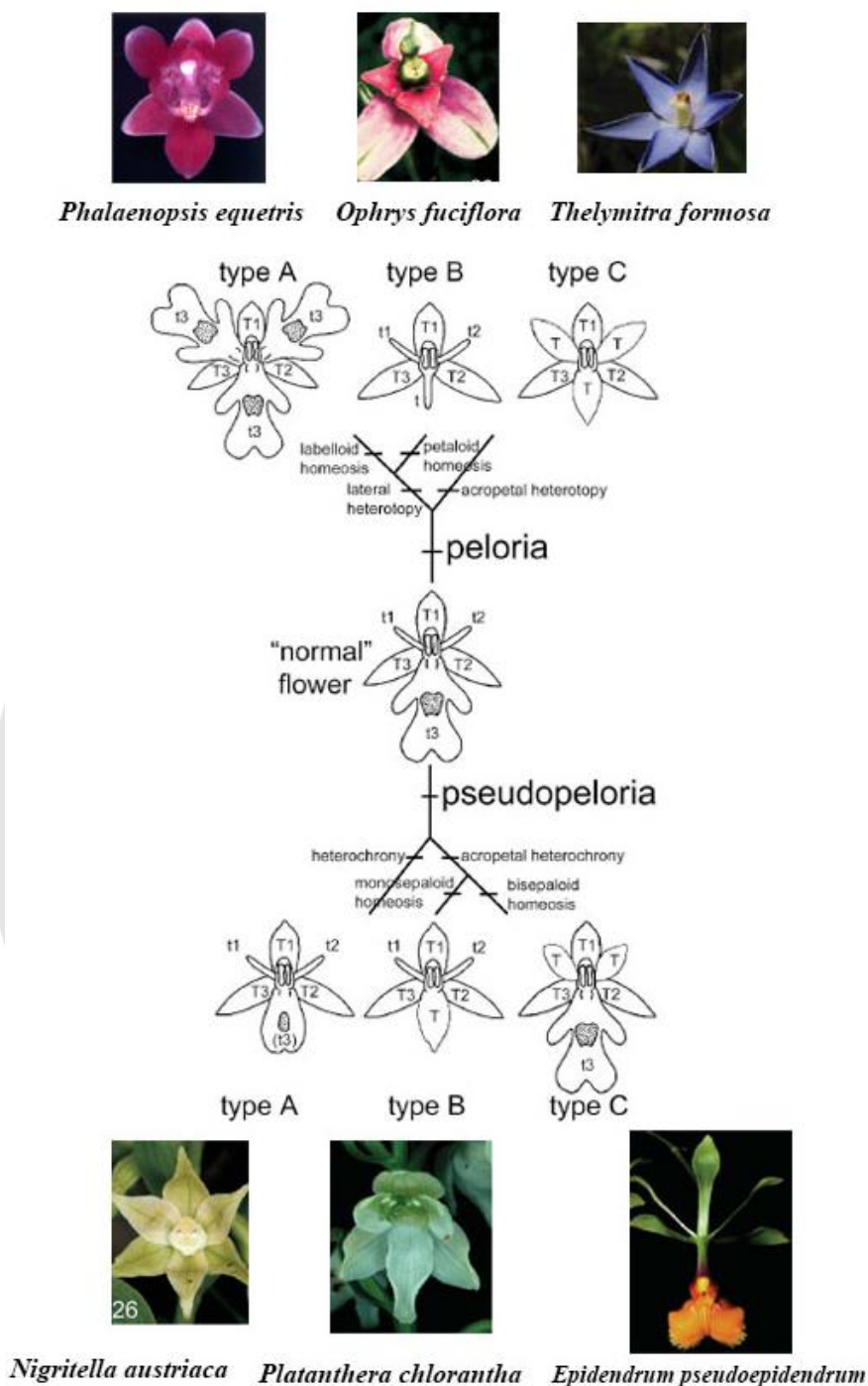


Figure 2 Diagrams of orchid floral homeotic variants. Normal flower at the center, Type A-C of peloria and pseudopeloria were indicated and the examples of each type were represent.

Source: modified from Bateman and Rudall (2006a); Mondragón-Palomino and Theissen (2009); Rudall and Bateman (2002)

The driving forces of orchid flowers diversification

According to the orchid flowers providing the most diversity and many key innovations, to approach what are the causation of evolutionary novelties is always one of the most challenge for researcher. A frequently explanation is ultimate causes and proximate causes (Mondragón-Palomino & Theissen, 2009). The ultimate causes in morphological change refer to the attraction of flowers to the specific pollinators that means focusing on the real reason or the evolutionary forces effect to the new traits. However, the proximate causes are more informative which explain the term of the mechanism or the biological function. In evolutionary development genetic field (EvoDev), the change in developmental process, especially the developmental control genes was assumed to be a major perspective of the change in floral morphology evolution (Mondragón-Palomino & Theissen, 2008; Theissen et al., 2000).

The developmental control genes are able to rise their varieties by generating genes paralogs from gene or genome duplication which resulted in the same function. The gene duplication is generally investigated in most species. According to lacking of selective pressure or evolutionary pressure to preserve both of duplicate genes, the frequent evolutionary fate will usually distinguish one of the copies to loss of function; non-functionalization. However, in neo-functionalization and sub-functionalization is essential matter of genetic novelty that innovate traits in evolution (Figure 3). The neo-functionalization is one of two copies of developmental control genes continuously functions correct and another one copy is free from selective pressure and generate the new and different function, this is an exciting event that lead to organs uncommon position and organ substitution; heterotopy and homeosis respectively. In addition, the sub-functionalization or duplication-degeneration-complementation model occurred when both copies are equally mutated in different cis-regulatory elements. Thus, the paralogous descendants possess the divided of original function; easily called job sharing fate. Neither of duplicated genes is able to lost or achieve novel function, this means it is neutral process without beneficial effect. However, in pleiotropic ancestral gene which can exhibit two functions, one function has changed without influencing to another. This situation permits the specialize adaptation of phenotypes and lets an adaptive benefit. Eventually, according to orchid peloria and pseudopeloria, these mutation was also the results from shift and loss of target genes such as the change in

recognition of DNA binding domain of transcription factors to their cis-regulatory elements of target genes (Mondragón-Palomino & Theißen, 2008).

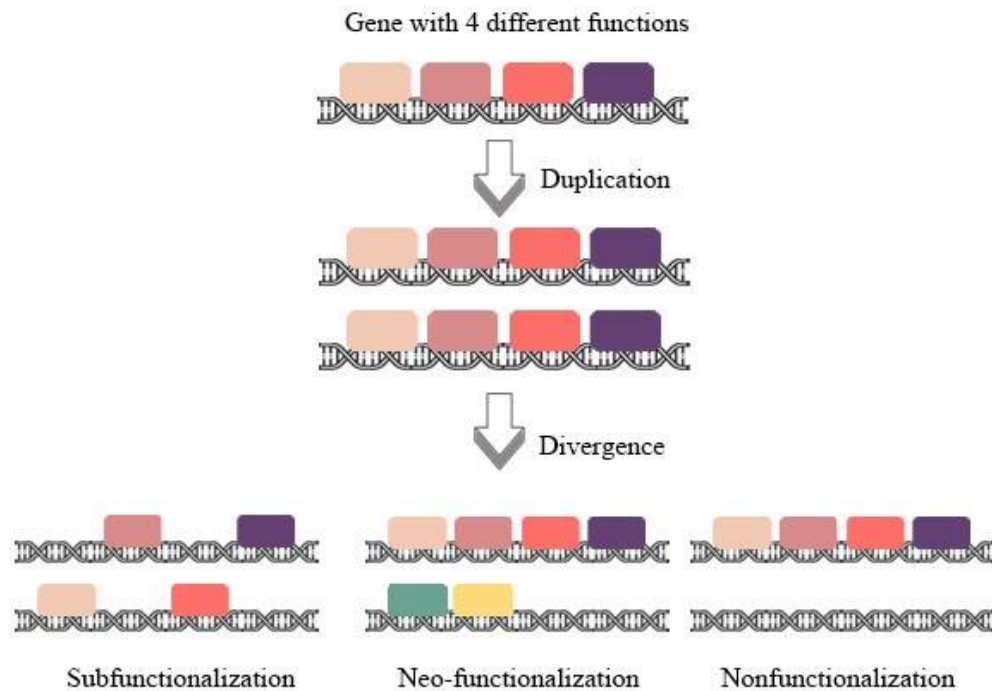


Figure 3 The gene duplication fate in evolution

Rhynchostylis retusa

Rhynchostylis retusa is in tribe Vandeeae, sub tribe Aeridinae and genus *Rhynchostylis* which is a small genus consisting of three species, *R. retusa*, *R. gigantea* and *R. coelestis* reported as endemic in Southeast Asian countries, such as Thailand, Laos, and Myanmar (Anuttato et al., 2017). *Rhynchostylis retusa* or commonly known as foxtail orchid. The name “*Rhynchostylis*” is derived from the Greek *rhykos* (“beak”) and *stylis* (“column”), it means the shape of beaked column exhibiting in this genus. *Rhynchostylis retusa* inflorescence is densely pendant raceme, consisting of more than 100 magenta-spotted white flowers and spicy smell. The plant has a short, stout, creeping stem, blunt leave apex and blooms on an axillary pendant to 60 cm in May to June. It is high value and demand in orchid market.

There are variations in floral organ phenotype of *R. retusa* but they have not been reported or studied before; lip-loss and tri-lip. However, these two variants were found by authors and it was not available to study. This was very unfortunately because the tri-lip variants interestingly possessed lip-like structure in both ventral sepals that is basipetal homeosis which is the shifted of inner whorl to outer whorl (Figure 4). Thus, to better understand perianth development of the orchids at the molecular level, this study will identify the floral organ identity genes, B-class (*DEF*-like) and E-class (*AGL6*-like) in wild type, characterize and construct the phylogenetic tree to investigate the extensive evolution coupled with previous model. Furthermore, Thailand is a natural habitat for several diverse species of orchids which is high economic value floral species and one of the most leading exporters of tropical orchids. Although many interesting resource of specimen, the research about the floral homeotic genes have been scarcely studied. Thus, this would be the initial research to unlock the further study in many orchids in Thailand. Our results will be of interest to a broader context of researchers worldwide.

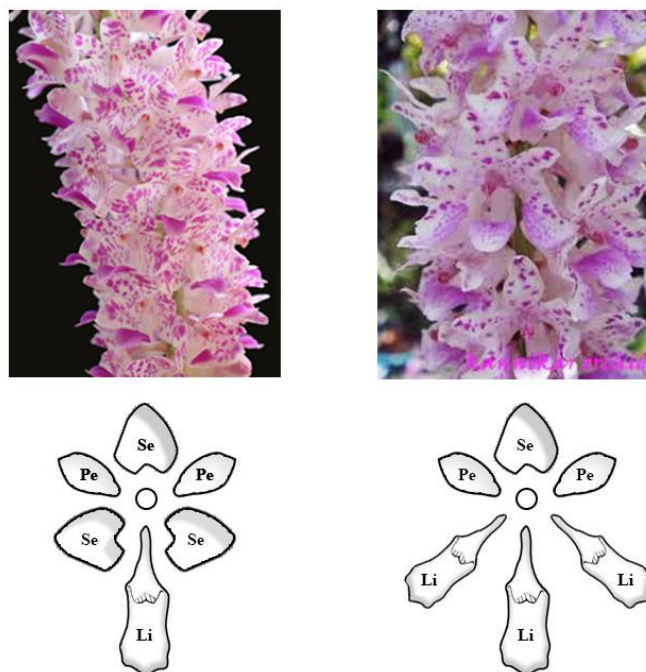


Figure 4 The comparison of *Rhynchostylis retusa* flower between wild type (a) and tri-lip (b)

MADS-box genes

Introduction

The MADS-box genes encode the MADS-domain family of transcription factors. The term “MADS” was derived from first four founding family members; *MINICHROMOSOME MAINTENANCE 1 (MCM1)* from *Saccharomyces cerevisiae*, *AGAMOUS (AG)* from *A. thaliana*, *DEFICIENS (DEF)* from *Antirrhinum majus* and *SERUM RESPONSE FACTOR (SRF)* from *Homo sapiens*. The first MADS-box gene was isolated from *S. cerevisiae*; ARG80 but that moment it did not fascinate and represent as the large or essential transcription factor class. In addition, ARG80 was assumed to correlate with the *MCM1* gene, so it was neglected to be part of the acronym (Gramzow & Theissen, 2010; Theißen & Gramzow, 2016).

The most recent data indicated that MADS-box genes originated from one region of topoisomerases IIA subunit A sequence (TOPOIIA-A) (Figure 5a). Topoisomerases typically function in the process of replication, transcription and recombination of DNA level and also the compaction and segregation of chromosome level. Forterre, Gribaldo, Gadelle, and Serre (2007) declared the phylogeny of prokaryotic and eukaryotic TOPOIIA-A which was only one gene existing in all eukaryotic species. Thus, this indicated that the duplication of one ancestral TOPOIIA-A gene in lineage and generated the most recent common ancestor (MRCA) of extant eukaryotes. Then, the another evolution of the ancestral TOPOIIA-A led for the ancestral MADS domain and the duplication of ancestral MADS domain led two major type including the SRF-like and MYOCYTE ENHANCER FACTOR 2 (MEF2)-like gene (termed as Type I and type II MADS-box gene respectively) in approximately 1.5 billion years ago (BYA) (Forterre et al., 2007; Gramzow, Ritz, & Theissen, 2010; Gramzow & Theissen, 2010). In contrast to the wide range distribution of MADS-box gene in eukaryotic species and biological process, the number of genes in protists, animals and fungi quite low but remarkably high in plant lineage. The MADS-domain transcription factor involved in all major aspect of algae (charophyte) and land plants (Embryophytes) including bryophytes (liverworts, hornworts, mosses), lycophytes, lycophytes, monilophytes (fern and its allies) gymnosperm and angiosperm. Thus, some clades of MADS-box genes have the origin and diversification point closely

correlated to the origin of evolutionary innovation point of seeds, flowers and fruits (Becker & Theißen, 2003).

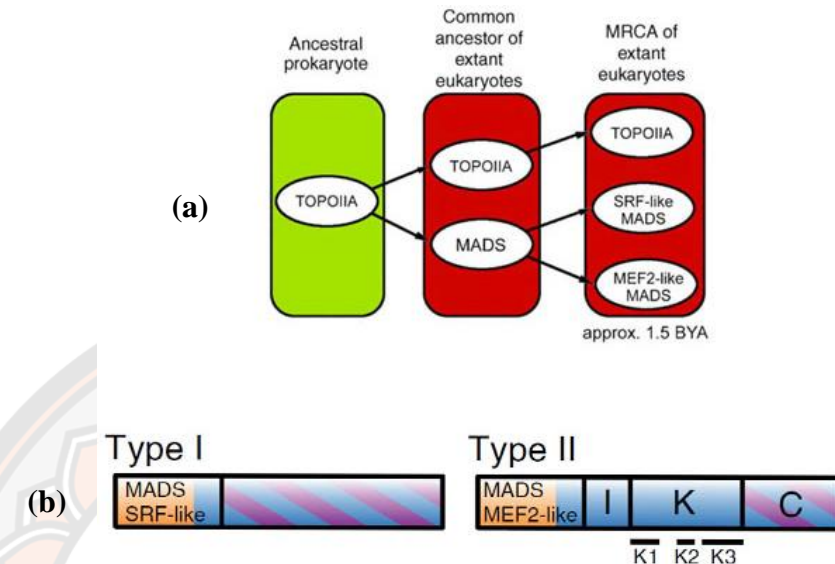


Figure 5 (a) The origin of two major MADS-domain by gene duplication in ancestral prokaryote TOPOIIA-A. (b) Two type of MADS-domain structure. Orange represents in DNA binding region, blue indicates a role in protein-protein interaction and purple represents transactivation region.

Source: Gramzow et al. (2010); Gramzow and Theissen (2010)

Structure and functions

There are 3-10% of all genes in genome of eukaryote encoding transcription factors. Transcription factors protein normally influence the process of gene transcription by binding to the regions of genome. The transcription factors contain DNA-binding domains that involve in specific sequence recognition of the regulated promoter regions (Gonzalez, 2016; Gramzow & Theissen, 2010). Then, the activation or repression will achieve depending on the protein-protein interaction with other transcription factor components or the interaction with chromatin-modifying enzyme which regulate directly through the genome accessibility. The different of DNA-binding domain structure and interaction is caused by alpha helix and beta sheet adoption. The same and closer of family or type of DNA-binding domains which have same structure tend to have more specificity of DNA sequence dimerization. So, the

diversification in DNA-binding specificity of gene in same family become interesting because the change in DNA-binding domain sequence can gain the novel character of DNA-binding domain and generate the evolution (Gonzalez, 2016).

The structure of MADS-box gene encoding MADS-domain protein consists of approximately 174 nucleotides long or 58 amino acids. According to the three-dimension structure of MADS domain in plant still not be available, the structure of MADS domain binding to DNA from SRF, MCM1 and MEF2 were isolated using X-ray crystallization (Figure 6a). According to the greatly conserved sequence, this could assume that plant MADS-box domain has similar structure. The X-ray crystallography reveal that the dimerization of MADS-domain recognizing to consensus DNA sequence CC(A/T)₆GG termed CARG box (CC-A-rich-GGG-box). The CARG motif are short, variable and ubiquitous in genome so it is very tough to predict the target genes using only this motif and it still unclear about how MADS-domain reach specifically to target gene (Gramzow & Theissen, 2010). The folding of MADS-domain beginning from N-terminus at 14th amino acid followed by a long amphipathic of α -helix and two β -strands (Figure 6b). The long antiparallel of α -helices structure is encoded from the central part of the MADS-domain and interestingly N-terminal stretch extensively interact with minor groove of the DNA. The narrow minor groove is recognized by the conserved arginine (R) position 2 side chain. This involves in specific sequence DNA binding mechanism through shape recognition. Above the coil of α -helices, there is a four-stranded antiparallel of β -sheet which is encoded by the C-terminal of MADS-domain. Thus, the MADS domain is categorized in beta-Scaffold factors with minor groove contacts superclass of transcription factors (Gramzow & Theissen, 2010; Theißen & Gramzow, 2016).

The half of N-terminus of MADS-domain is essential for DNA binding, but the half of C-terminus is required for dimerization to build the homo- or heterodimers or multimers not just single polypeptides. The localization of MADS-domain proteins is assumed to restricted in nucleus because it contains the nuclear localization signal (NLS) motif (KR[K/R]X₄KK) at position 22-30 (Figure 6b) which facilitate in transportation of transcription factor from cytoplasm (where the translation occurred) to nucleus (where the main function occurred) (Theißen & Gramzow, 2016).

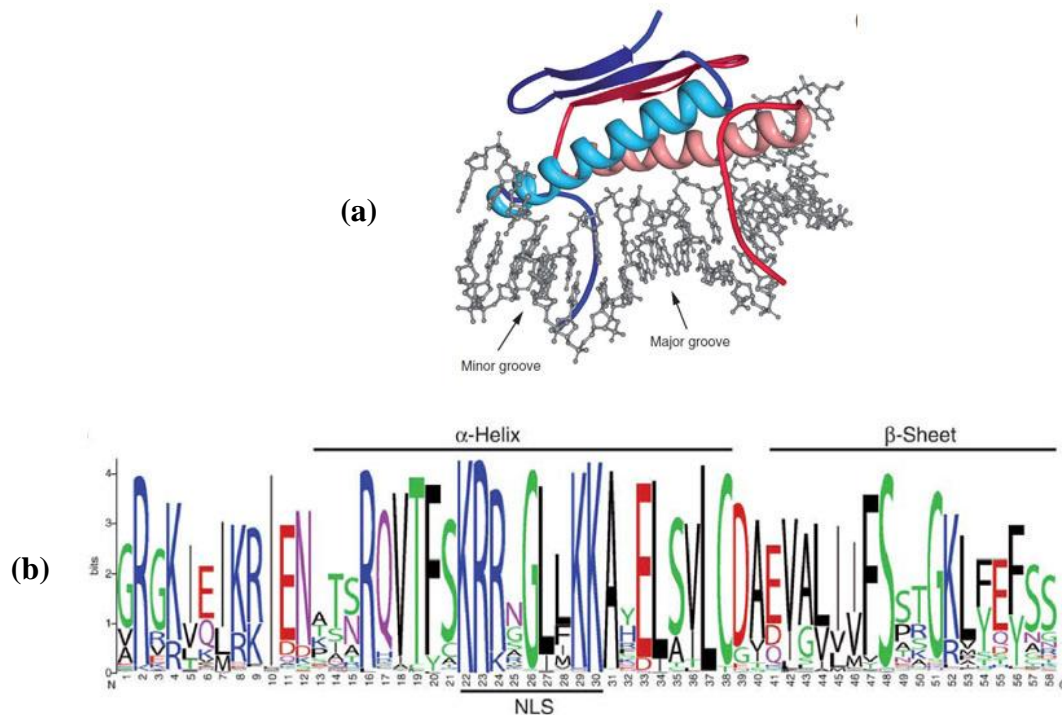


Figure 6 (a) the crystal structure of Human SRF MADS-domain binding to DNA. DNA is shown in grey ball-and-stick structure while the dimers of two MADS domain units were represented by red and blue color. The spring-like and arrow ribbons represented the α -helix and two β -strands. (b) The sequence logo showed summarization of amino acids based on 867 plant MADS domain sequences (Type I and II); the stack of letters indicated the conservation and the height of letters referred the frequency of amino acid in that position. The NLS underlined as the nuclear localization signal conserved sequence.

Source: Gramzow and Theissen (2010); Theißen and Gramzow (2016)

Types of plant MADS-box genes

From two major type of MADS-box genes that have been mentioned earlier; type I and type II of plant were orthologous to SRF-like and MEF2-like genes respectively (Figure 5b). Type I MADS-box genes contain one or two exons, but type II MADS-box genes have approximately seven exons. Type I and II MADS-box genes also could be distinguished by their evolutionary rate which type I MADS-box genes

have a faster rate of gene birth and death and experienced more duplication fate than type II MADS-box genes (Bemer, Gordon, Weterings, & Angenent, 2010; Nam et al., 2004). Consequently, type I genes gave rise into three groups: $M\alpha$; $M\beta$ and $M\gamma$ depending on their conserved region and variable of C-terminus (Gramzow & Theissen, 2010). Type I gene is like the ‘dark matter of the MADS universe’. It has been studied for long time but only genes from *A. thaliana* were found and characterized. The briefly results revealed that plant type I control the plant ‘female side’ including female gametophyte, embryo and seed development.

In this study, the type II MADS-box genes are notably focused because it widely known as the floral organ identity genes which encoded the highly conserved domain structure called ‘MIKC-type’ (Figure 7a). The structure of MIKC-type MADS-box genes consist of the most highly conserved MADS (M) domain which is the main factor for DNA-binding and also dimerization, intervening (I) domain which constitute in determining the specificity of DNA dimerization, Keratin-like (K) domain which mainly form protein-protein interaction and the most variable C-terminal (C) domain which involve in transcriptional activation and the multimeric formation respectively.

The different of Type I and II MADS-box genes in plant from animal and fungi, SRF-like and MEF2-like genes respectively, is the represent of K domain. Moreover, it is also the criteria to separate type I and type II. K-domain protein has approximately 70 amino acids length and is encoded from three exons; K1, K2 and K3 subdomain which characterized by a heptad repeat ($[abcdefg]_n$). The position ‘a’ and ‘d’ are usually hydrophobic amino acid; especially leucine, form amphipathic α -helices and lead to coiled coils structure that is look alike keratin so that why its name keratin-like but not even homolog (Kaufmann, Melzer, & Theißen, 2005). Thus, it is required in protein-protein interaction or multimeric complex formation. In case of class B protein (extensively described later), DNA binding heterodimer require mostly K1 and some K2. In addition, K2 and K3 involve in higher complex multimerization (Kaufmann et al., 2005). The essential of K-domain can describe through the absence of C-domain of SEP3 in *A. thaliana* but present of K3 subdomain could still perform mutimerization. The X-ray crystal structure of the K-domain (Figure 7b) was first and recently performed in SEP3 which showed the coiled coil of α -helices and revealed that K-domain involving in dimerization and tetramer formation. The tetramers of MIKC-

MADS domain proteins are able to bind to two different CArG boxes by looping two DNA binding sites (Puranik et al., 2014; Theißen & Gramzow, 2016).

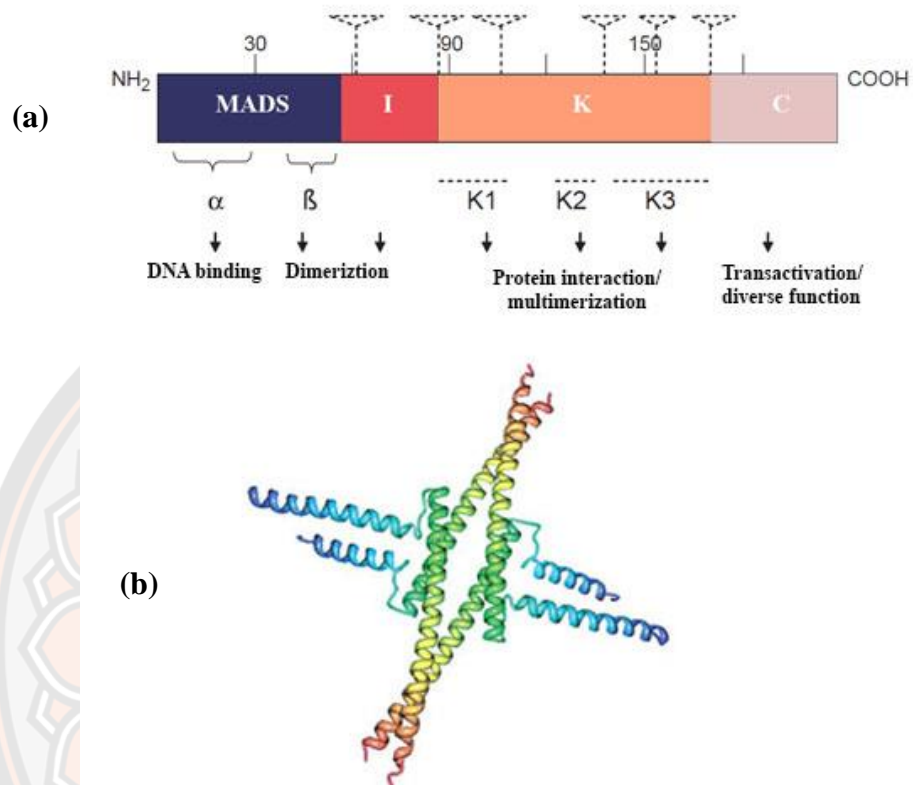


Figure 7 (a) A schematic structure of type II MIKC-MADS protein (APETALA3); the number above indicated the amino acid position and the dot triangles showed the intron position. (b) the X-ray crystal structure of K domain tetramer from SEP3 (*A. thaliana*).

Source: modified from Kaufmann et al. (2005); Theißen et al. (2016)

The type II MIKC-MADS domain are further distinguished into two type including MIKC^c (classic) and MIKC^{*} which different in the length and the exon numbers encoding I-domain. The MIKC^c have shorter and only one or two exons of I-domain protein, whereas MIKC^{*} proteins possess longer sequence and encoded by four or five exons. The MIKC^{*} genes are much like type I MADS groups because of their

function and absence of the K-domain. MIKC*-group was found in various of land plants. Especially only in ferns and seed plants, two different clades of MIKC* (P and S clade) which excepted in bryophyte and lycophyte were identified and had interacted to form heterodimers. Then, these dimers mostly bound to different type of CArG box that behave conversely to MIKC^c protein dimers. In addition, MIKC* transcription factors are required for the male gametophyte, proper pollen maturation and germination.

Previously, the chlorophyte lineage including green algae; *Ostreococcus tauri* and *Ostreococcus lucimarinus* contain just one MADS-box gene lacking of K-domain and probably have type I MADS-box domain (Figure 8). However, the MIKC-MADS domain was discovered in charophytes which is the most basal of streptophytes (Group of charophytes an embryphytes), this referred that the K-domain has occurred since more than 700 million years ago (MYA) in the extant streptophytes lineage. Then, this ancestral MIKC-MADS domain duplicated and led MIKC* and MIKC^c in the lineage of extant land plants more than 450 MYA which was predicted from the of gene characterization in moss *Physcomitrella patens*, the most basal species of land plants which is currently available. This suggested that at least one of each types of MIKC already occurred in the last common ancestor of mosses and vascular plants (Becker & Theißen, 2003; Gramzow & Theissen, 2010; Theißen & Gramzow, 2016).

The MIKC^c MADS-box genes are rather focused because they contribute to wide range of flowering developmental processes and become clearer and clearer (Theissen et al., 2000). The MIKC^c-type genes represent a monophyletic clade which can be further subdivided into ancient clade or 'gene subfamilies'. The *Arabidopsis* MADS-box genes early provided the member of well-characterized gene subfamilies including; *AG*-like, *AGL2*-like, *AGL6*-like, *AGL15*-like, *AGL17*-like, *DEF*-like, *GLO*-like, *FLC*-like, *GGM13*-like, *SQUA*-like, *STMADS11*-like and *TM3*-like genes. The eight of *AG*-like, *AGL6*-like, *AGL12*-like, *DEF*-like, *GLO*-like, *GGM13*-like, *STMADS11*-like and *TM3*-like genes existed in the most recent common ancestor of angiosperms and gymnosperms (seed plants) about 300 MYA but *AGL2*-like, *AGL17*-like, and *SQUA*-like genes occurred already in the most recent ancestor of monocots and eudicots about 200 MYA. In addition, only *AGL15*-like and *FLC*-like genes were reported solely in Brassicaceae (Becker & Theißen, 2003).

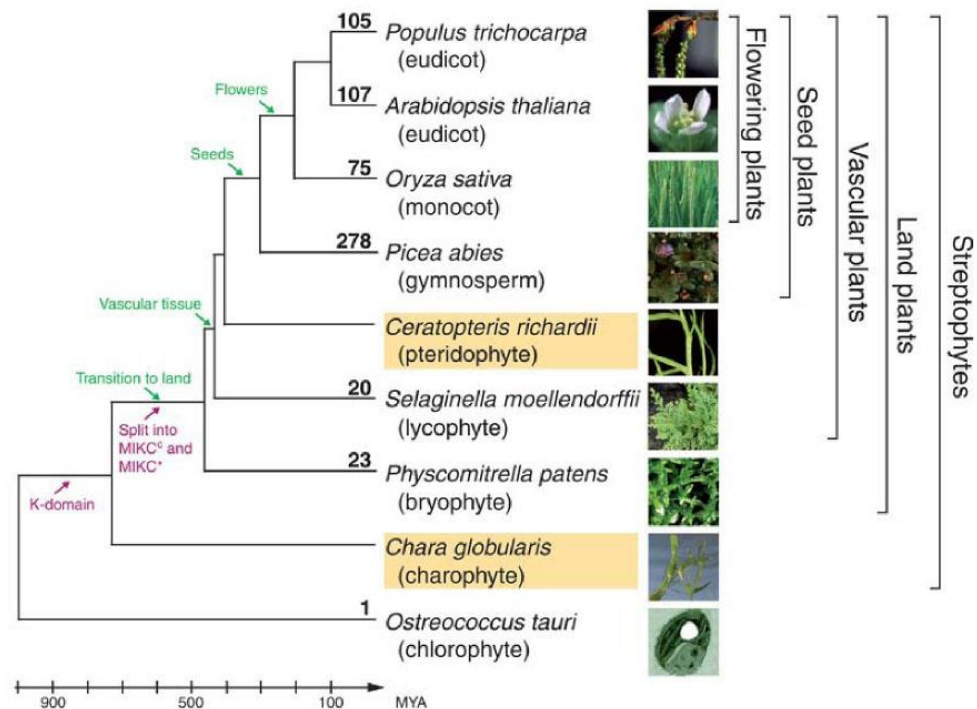


Figure 8 The phylogenetic tree of MADS box gene based on whole genome sequence data. The numbers on branch indicated the numbers of MADS box genes in genome. The purple and green arrow showed the major genetic changes and morphological innovation respectively.

Source: Gramzow and Theissen (2010)

Floral organ identity

Introduction

The developmental control genes are genes which significantly involve in developmental processes. The change in developmental control genes will affect the organ identity and formation. The developmental control genes were usually in multigene families encoding transcription factor (Theissen et al., 2000). One of the best paradigm is the homeobox genes, which marks the beginning of new era of evo-devo. The first homeotic mutation was found in *Drosophila* in early 1915 including *bithorax* mutants and *Antennapedia*

mutants. Thus, the striking parallels between the Homeobox in animal and the floral homeotic genes which almost encode MADS-box protein (Gehring, 1992; Ng & Yanofsky, 2001; Theissen et al., 2000).

In angiosperm, there are “ANITA” clades which is the earliest branch and sister to all angiosperm species, including Amborellaceae and Nymphaeaceae (Water lilies). Except to the ANITA group, the angiosperms are termed “euangiosperm/core angiosperm” that consist of the magnoliid complex, the monocots, the Chloranthaceae, and the eudicots (almost 75% of all angiosperms). In eudicots group further achieved Ranunculales clade (lower eudicots) that sister to all the rest eudicots; termed “core eudicots” (higher eudicots). Two great important clades in core eudicots are the rosids and asterids that distinguished in 70 MYA. One family within the rosids is Brassicaceae or the cabbage family, including model plant; *A. thaliana*. For asterids, another model plant *A. majus* is represented for this group in Lamiales clades and followed by the Solanales clade that comprise of *Petunia hybrid*, *Lycopersicon esculentum* and *Solanum tuberosum*; petunia, tomato and potato respectively. According to the earliest APG IV system, the phylogenetic tree of angiosperms is represented below (Figure 9).

ABC model

The floral homeotic changes were studied in early 1990s, based on two model plants, *Arabidopsis thaliana* and *Antirrhinum majus*. Both of these species received the floral mutants that affected the organ identity in particular whorls (Figure 10). The wild type flower organs structure is sepal/petal/stamen/carpel; arranged from whorl 1 to whorl 4. However, three classes of floral homeotic mutant gained the organ defects. The mutant in *apetala1* (*ap1*) *apetala2* (*ap2*) gene from *Arabidopsis* and *ovulata* (*ov*) gene from *Antirrhinum* altered the floral organ identity into leafy or carpel/stamen/stamen/carpel; A-function mutants. The mutant in *Arabidopsis apetala3* (*apetala3*) and *pistillata* (*pi*) genes from *Antirrhinum deficiens* (*def*) and *globosa* (*glo*) caused sepal/sepals/carpel/carpel; B-function mutants. Finally, the mutant in *agamous* (*ag*) and *plena* (*ple*) from *Arabidopsis* and *Antirrhinum* respectively defected the floral whorl into sepal/petal/petal/sepals or petals; C-function mutants. In addition, the mutant of all three functional classes (triple mutant) caused the transition of flower to leaves. These phenotypic mutations led the elegant model that classify the floral identity genes into to 3 functional classes, called classical “ABC model”. This model had contained

in many modern biology textbooks. Except to *apetala2*, all genes in ABC model shared the conserved homolog to type II MIKC type MADS-box proteins. Then, lots of ABC model genes were isolated from various species in the middle of 1990s (Causier et al., 2010; Coen & Meyerowitz, 1991; Weigel & Meyerowitz, 1994). This indicated that MADS-box transcription factors play as a key regulator in plant developmental processes.

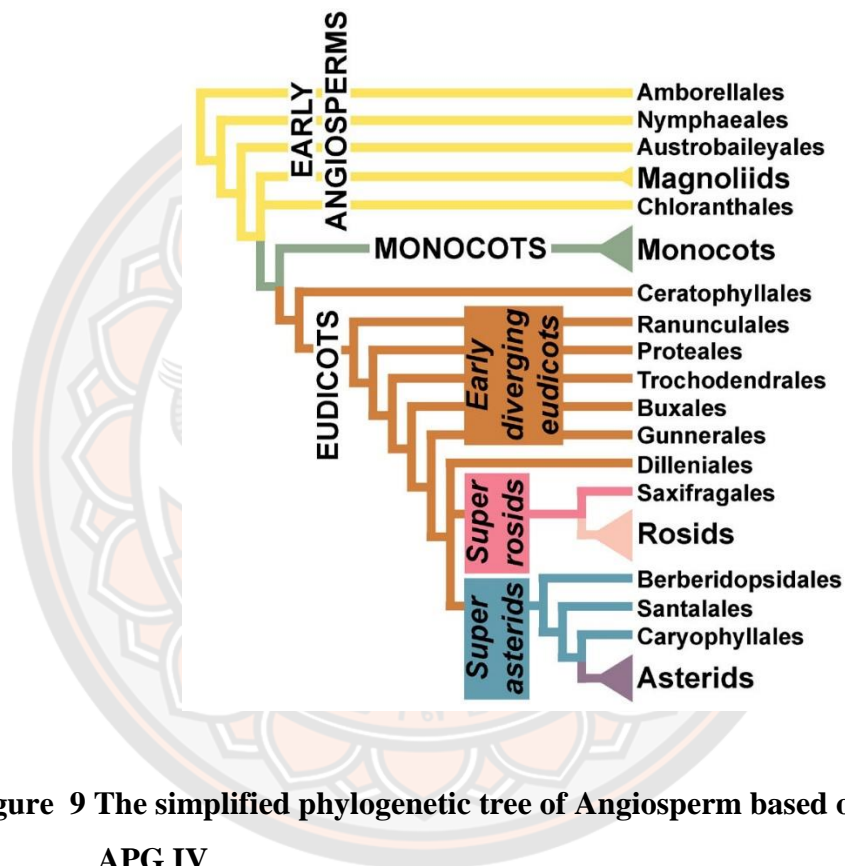


Figure 9 The simplified phylogenetic tree of Angiosperm based on earliest APG IV

Source: Byng et al. (2018)

ABCDE model

According to Johann Wolfgang Goethe had proposed the theory in 1790, he suggested that floral organs are modified from vegetative leaves referred that the expression of floral homeotic genes in vegetative tissue are able to converse into flower-like structure. However, the only co-expression of B-class or C-class expression in ABC model do not sufficiently change the vegetative organs to floral organs. So, this means the ABC model required the function that determine the organ identity. From the

MIKC MADS-box proteins that were mentioned before, the MADS-box protein-protein interaction, DNA dimerization and hetero-dimerization, involved in promoting floral organ identity (Causier et al., 2010). The DEF/GLO dimer and SQUAMOSA (SQUA) are the first MIKC MADS-box proteins interaction using C-domain to gain the multimeric complex (Causier et al., 2010; Theißen et al., 2016; Theißen & Saedler, 2001).

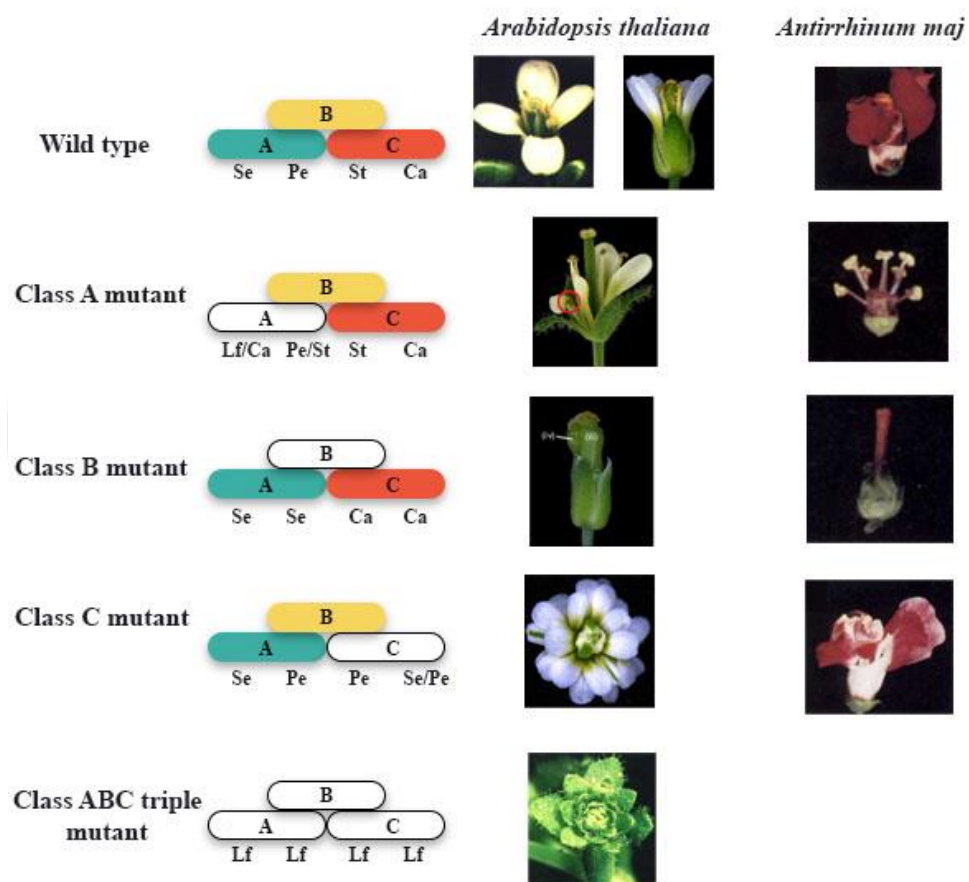


Figure 10 The classical ABC model and the floral homeotic mutant in Class A, B, C and triple ABC functions affecting in *A. thaliana* and *A. majus* phenotypes. **Source:** modified from Causier et al. (2010); Coen and Meyerowitz (1991); Weigel and Meyerowitz (1994)

Consequently, other novel floral mutants were studied and expanded the ABC model to ABCDE model (Figure 10). D-class function was suggested to combine with

C-class and specify ovule identity. However, the E-class was remarkably interesting for ABC model. The E-class function first found in tomato and petunia, *TM5* and *FBP2* genes (Angenent, Franken, Busscher, Weiss, & van Tunen, 1994; Ferrario, Immink, Shchennikova, Busscher-Lange, & Angenent, 2003; Pnueli, Hareven, Broday, Hurwitz, & Lifschitz, 1994). Then, these two genes were isolated from *A. thaliana* including *SEPALLATA 1 (SEP1)*, *SEP2* and *SEP3*. These three genes mutant (*sep1sep2sep3*) shared characters to silenced *tm5* and *fbp2*. The triple mutant of *SEP* phenotype is the replacement of sepal in every whorl, same as B- and C-function double mutant. The ectopic expression of *SEP* genes had not efficiency to change vegetative organs but the co-expression of A-, B-functional class and *SEP* sufficiently altered the rosette leaves to petaloid structure. These results indicated that *SEP* expression is sufficient to provide the floral identity. The ability of conversion from vegetative leaves to floral organs confirmed the theory of Johann Wolfgang Goethe which had predicted more than 200 years ago. With the function of *SEPs* proteins mediating the protein-protein interaction, *SEPs* form and stabilize the higher-order complexes between MADS-box proteins, this will further suggest that it plays an important role in other developmental processes regulation (Causier et al., 2010; Weigel & Meyerowitz, 1994).

The floral quartet model

The ABCDE model is the gene-based data. The further studies in protein-protein interaction of MIKC-MADS box proteins that form higher complex (tetrameric complex) required the protein-based experiment. The integration of MIKC MADS-box gene-based and encoded proteins information provided the new model for floral organ identity determination termed “Floral quartet model” (Theißen & Saedler, 2001) (Figure 11). The Floral quartet model was first introduced in 1999 when Egea-Cortines and team characterized DEF, GLO and SQUA from *A. majus* forming in tetramer (DEF-GLO-SQUA-SQUA) including DEF-GLO heterodimer and SQUA-SQUA homodimer (Egea-Cortines, Saedler, & Sommer, 1999). The Floral quartet complex formation also comprises *SEP* that was mentioned earlier to complete ABC model. The other complex had been investigated, AP3-PI-AG-*SEP* complex and AP3-PI-AP1-AP1 complex were important in stamen-like and petaloid structure respectively. For class D, the SHATTERPROF (SHP) and SEEDSTICK (STK) proteins had the mutant phenotype (*shp1shp2stk* triple mutant) same as partial loss of *SEP* gene also form multimeric

complex with SEP3 using yeast three-hybrid assay. These indicated that class D also included in Floral quartet model involving in ovule development.

This process of protein interaction impact on specificity of target genes. The important of tetramers or quartets is when only one protein of complex loss, the rest of protein will be disabled that brought evolutionary achievement to keep many paralogous MIKC-type genes proposed to control floral organ during seed plant evolution (Theissen et al., 2000; Theißen et al., 2016; Theißen & Saedler, 2001).

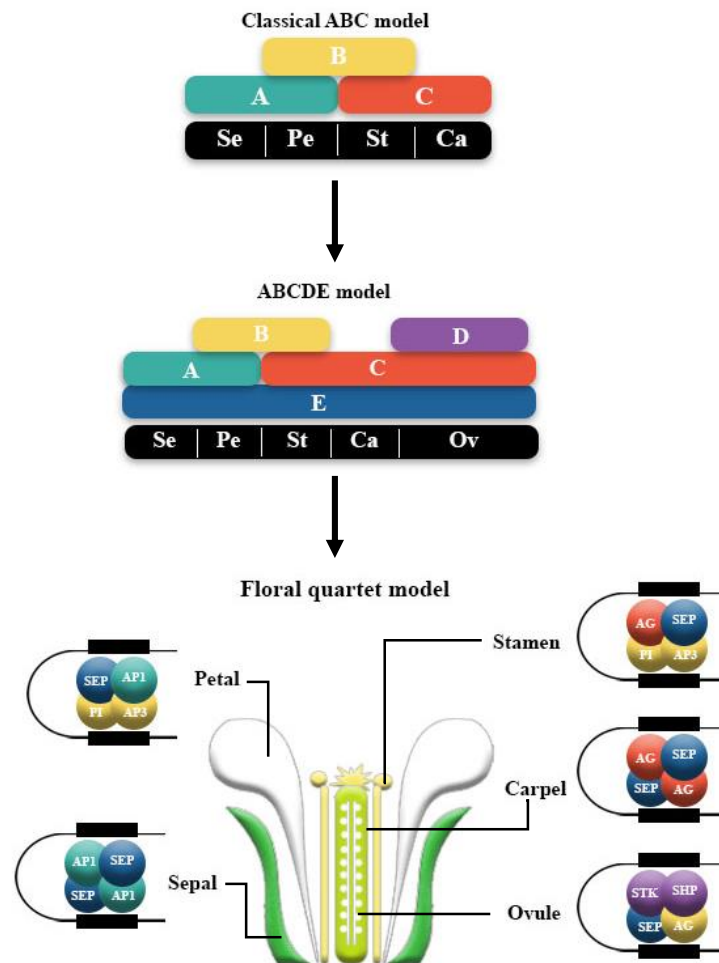


Figure 11 The classical ABC model, the ABCDE model and the floral quartet model was represented.

Source: modified from Coen and Meyerowitz (1991); Theißen et al. (2016)

The deep regulatory mechanism of MADS-box transcription factor

According to MADS domain protein in animals and fungi efficiently bind CArG-box, the plant MIKC-MADS domain proteins binding with CArG motif in target genes were also investigated using various methods including Random binding site selection, Chromatin immunoprecipitation (ChIP), Transcriptional induction system, Electrophoretic mobility shift assay and *in silico* binding search. This found that the MIKC-MADS transcription factors bind to thousands of sites in the *Arabidopsis* genome and those sites strongly conserved as CArG-box sequence (S. de Folter & Angenent, 2006).

There were two type of target gene of MIKC-MADS transcription factors, ABC genes and non-ABC genes. The ABC target genes defined as the floral organ identity genes. For example, the *SHATTERPROOF* genes (*SHP*), *SHP1* and *SHP2*, these D-class MADS-box genes were identified as the target genes for class C MADS transcription factor, AG protein and specify for carpels.

The another type is non-ABC genes. These genes involved generally in cellular maintenance. For example, *NAC-LIKE ACTIVATED BY AP3/PI (NAP)* was identified as the target gene for DEF/GLO complex and have roles in cell division and cell expansion during petal and stamen development (B-class function) (Alvarez-Buylla et al., 2010; S. de Folter & Angenent, 2006).

These example shown that plant MADS domain have “autoregulation” mechanism (Figure 12). The auto-regulatory loop consists of negative feedback loop and positive feedback loop. The positive feedback loop found in B- and C-class MADS domain and described in two examples above. In addition, both DEF/GLO and AP3/PI from *A. majus* and *A. thaliana* respectively was also identified as direct target that bind to their own promoter controlling themselves. For the negative feedback loop, this mechanism switches off the gene expression that especially depends on the interaction of proteins between floral induction (Flowering) and floral organ formation. The floral induction protein; SUPPRESSOR OF CONSTANS 1 (SOC1), AGAMOUS-LIKE 24 (AGL24) and SHORT VEGETATIVE PHASE (SVP) form the heterodimer with floral organ identity; AG, SEP1/2/3, SHP1/2. Therefore, the gene expression involving in flowering was negatively suppressed by these heterodimers. Furthermore, both floral

induction and floral organ formation form the regulatory network with flowering time regulation (Figure 13); APETALA 1 (AP1) and FRUITFULL (FUL) to determine early and late flowering function (S. de Folter & Angenent, 2006; Stefan de Folter et al., 2005).

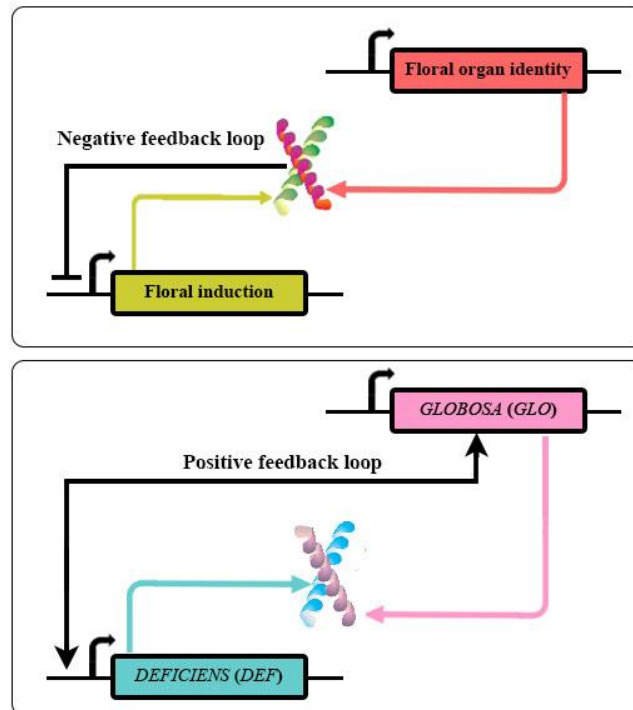


Figure 12 The autoregulation of MIKC-MADS domain.

Source: modified from S. de Folter and Angenent (2006)

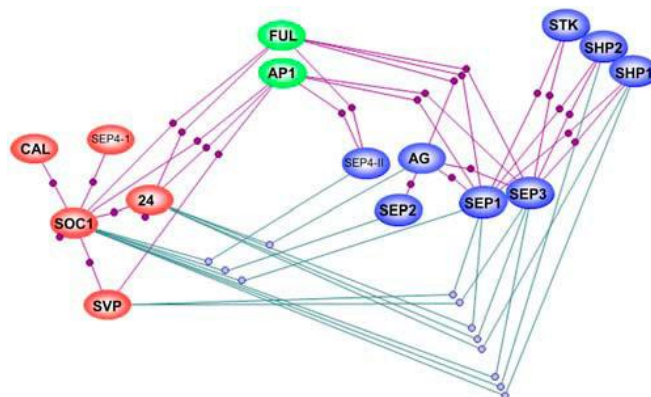


Figure 13 The flower induction and flower formation network

Source: Stefan de Folter et al. (2005)

The story of MADS-transcription factors story has not finished yet. According to the nature of transcription factors always requires the specific DNA sequences of target genes. So, the questions that have been obscure are “How MADS-domain protein acquire target gene specificity” and “What is the mechanism behind the activation and repression of target gene expression?”

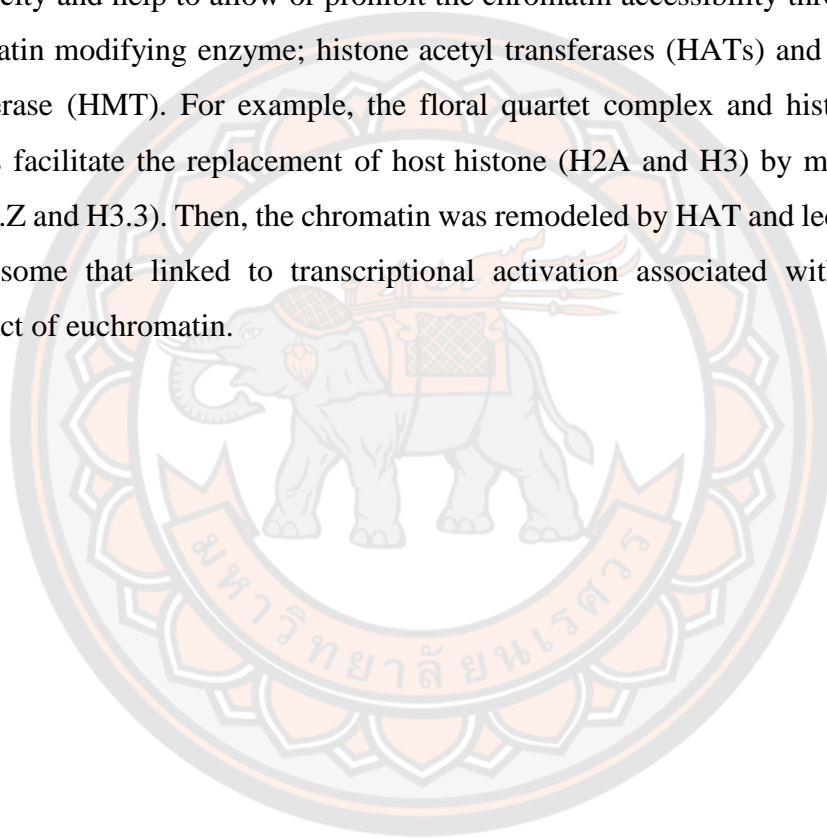
The limitation to achieved the specificity including two major causes. Firstly, The CARG-box alone is not enough reason to gain the specificity of MIKC-MADS transcription factor because this binding site was shared ubiquitously over thousand times in *A. thaliana* genome. Second is in approximately 45 types of MIKC-MADS protein possess the most highly conserved MADS-domain (M) from I-, K- and C-domain. This indicated that the DNA binding efficiency of MIKC-MADS protein will relatively have same specificity.

Consequently, in the most recent study took much effort in chromatin structure in the context of chromatin remodeling and chromatin modifying factors (Figure 14). These processes interacted with MIKC-MADS proteins. For example, the H3K27me3 was removed the methyl group in RELATIVE OF EARLY FLOWERING 6 (REF6) and SEP3 promoter by demethylase from AP1, the SEP3 subsequently activated. This resulted in AP1 and SEP3 bind to the enhancer site during the early of flower development and modify the chromatin accessibility. This defined the role of AP1 and SEP3 complex as the “Pioneer transcription factor”, means the ability of transcription factors to bind the inaccessible of nucleosome-associated sites by eliminating host nucleosome and open the chromatin and allow the “Non-pioneer transcription factor” to access.

This brought the interesting to clarify floral quartet complexes function. According to AP1 and SEP protein act as pioneer transcription factor, they facilitate the invasion of nucleosomal DNA during floral quartet complex formation by ejecting host nucleosome and replacing by floral quartet tetramer. The most efficiency of pioneer transcription factor could eject the nucleosome is the gap between pioneer transcription

factor and the binding site up to 74 bp. This supported by the length of CArG-box region which has known as the binding site before.

The nucleosome structure is octamer which consist of 2 copies of H2A, H2B, H3 and H4 histones. Thus, the half-nucleosome (H2A-H2B-H3-H4) is one copy and similar to floral quartet complex/tetramers. This hypothesized the “nucleosome mimicry” model of floral quartet function (Figure 15). The similarity of floral quartet tetramer at CArG box and half-nucleosome structure and sequence would represent the specificity and help to allow or prohibit the chromatin accessibility through triggering chromatin modifying enzyme; histone acetyl transferases (HATs) and histone methyl transferase (HMT). For example, the floral quartet complex and histone-modifying factors facilitate the replacement of host histone (H2A and H3) by modified histone ((H2A.Z and H3.3). Then, the chromatin was remodeled by HAT and led the acetylated nucleosome that linked to transcriptional activation associated with less densely compact of euchromatin.



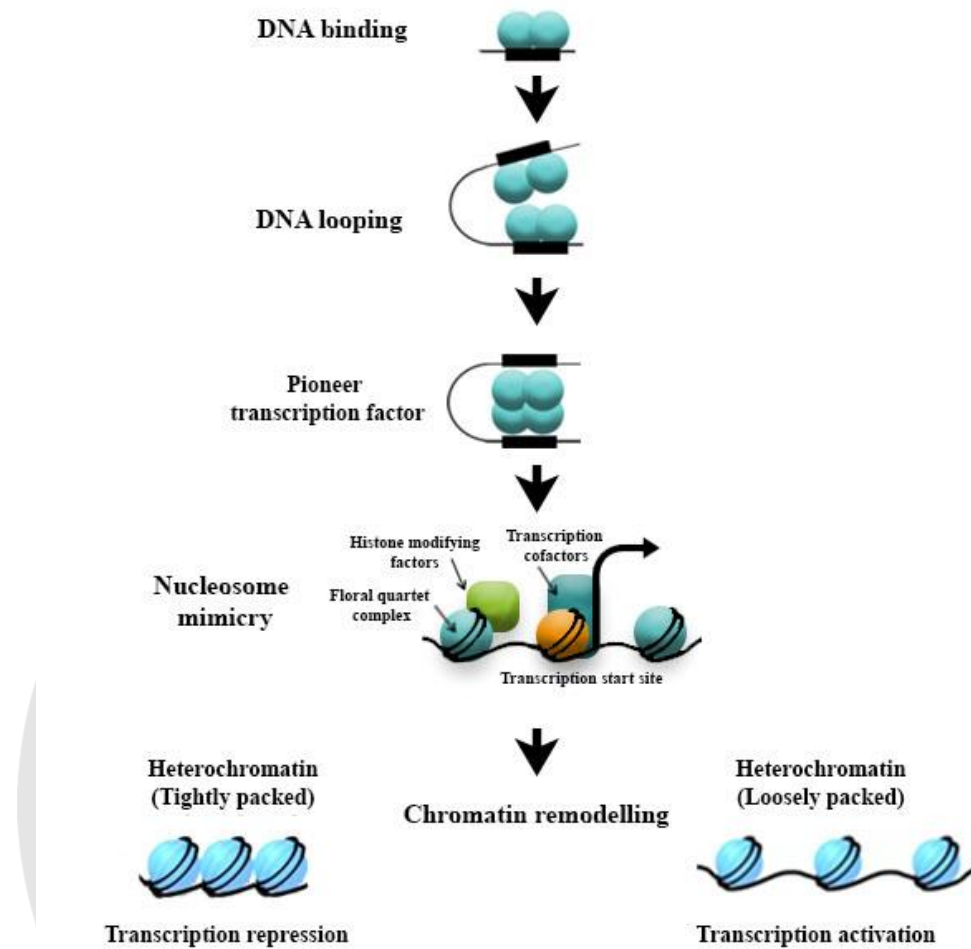


Figure 14 The regulation of floral quartet complex from DNA binding to chromatin remodeling

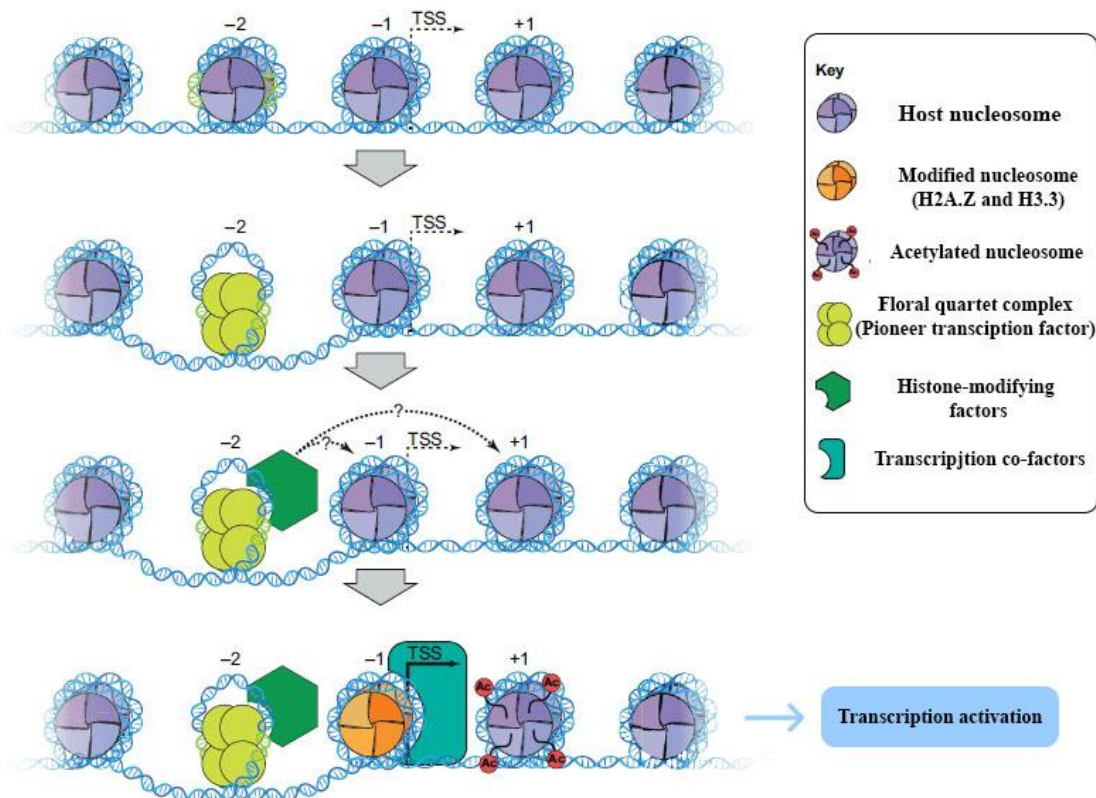


Figure 15 The nucleosome mimicry hypothesis for floral quartet complex functions

Source: modified from Theißen et al. (2016)

The important of *DEF*-like and *AGL6*-like in floral organ specification

DEF-like subfamily

According to ABC model, two genes, *DEF*-like and *GLO*-like genes, are specify for petals and stamen in *A. majus*. Their orthologs in *A. thanlina*, *AP3* and *PI*, have similar functions. Both of these genes has been found in angiosperm and gymnosperm, indicating that B-class genes have existed more than 300 MYA. Thus, the border ancestral functions may include separating male reproductive organ from female reproductive organ in gymnosperms followed by petal identity in angiosperms (Becker & Theißen, 2003). This B-class genes lineage was further investigated how petal independently achieved their unique morphology, *DEF*-like and *GLO*-like genes were characterized in lower eudicots, core eudicots and magnolid dicots. The results shown

the conserved sequences which could distinguish *DEF*- and *GLO*-like gene from other MADS-box genes. In K box, the *DEF* homologs shared the conserved sequence (H/Q)YExM in the position of residues 85 to 89. For *GLO* homologs possessed the conserved KHExL in residues 88 to 92 of the K box.

In addition, in the C-terminus also identified the notably conserved motif (Figure 16). In *GLO* homologs, the consensus sequence of MPF_xFRVQP_xQPNLQE was diagnosed and termed “PI motif”. Furthermore, for *DEF* lineage, two core dicots *DEF* orthologs, TM6 and PD2 from *Lycopersicon esculentum* and *Solanum tuberosum* (tomato and potato respectively) were previously found. However, there were another genes orthologous to *DEF* lineage, *STDEF* (*S. tuberosum*) and *LeAP3* (*L. esculentum*). Then, the lower eudicot and magnolid species group was compared to the core eudicots group. The distinctively difference between two groups was considered. In core eudicots, they shared the consensus sequence D(L/I)TTFALLE, termed as “euAP3 motif” which located in C-terminus. For the lower eudicots and magnolid group, the conserved sequence in C-terminal region was absolutely different as YG_xHDLRLA, termed as “paleoAP3 motif”.

Moreover, paleoAP3 motif perfectly aligned in *SILKY-1* (*Zea mays*), *PtAP3-1/-2* (*Pachysandra terminalis*), *TM6* (*L. esculentum*), *PD2* (*S. tuberosum*) even slightly diverged but still recognized. Interestingly, these genes represented the monocots and core eudicots. However, the paleoAP3 lower conserved in Pteridophytes (ferns), CRM genes. These suggested that Pteridophytes had diverged from the embryophytes (Land plants) approximately 400 MYA.

		PI Motif	
GLOBOSA	-----HHHH QNIA--D-Y EA-----	Q-MP--FAFR VQPMQPNLQE	RF-----
SvPI	-----H QQGVG--D-Y EY-----	Q-MP--FAFR VQPMQPNLQD	RF-----
FBP1	-----QQREN-HD-Y QN-----	H-MP--FAFR VQPMQPNLQE	RL-----
pMADS2	-----H Q-RDR--D-Y EY-----	Q-MP--FALR VQPMQPNLHE	RM-----
NTGLO	-----H Q-REN--EYQT-----	Q-MP--FAFR VQPMQPNLQE	RF-----
PI	-----DHD G-----	Q-----FGYR VQPIQPNLQE	KIMSLVID--
SLM2	NPS-----DRD-Y HY-----	QNPIPPYGFRR VQPMQPNLQD	RM-----
DaPI	-----H QNGR--D-Y-----	PS H-MP--FTFR AQPMQPNLQE	NQ-----
CpPI	-----H QNGR--D-Y-----	PS Q-MP--PFTFQ LHPSQPNLQE	IK-----
RbPI-1	-----DHD G-----	Q-MPMPFTFR VQPAQPNLQD	N-----
RbPI-2	-----DH- GY---PPPS Q-----	Q-MP--FTFL VQPIHPNFQD	N-----
DePI	QHSHHHHHH QKGRG--DHY QAAAAHASS	Q-MP--PFAFR VQPIQPNLHN	NNNNTNNTNN K
PnPI-1	NNN-----QKADGTRD-Y PAHNDNHH--	Q-VF--FGFQ VPPMQPNLTT	VTTTTTNNK
PnPI-2	-----I-----	-P--IAFH VQPLHPNLQE	MK-----
PmPI-1	DL-----	-P--FAFR VQPIQPNLQE	QK-----
PmPI-2	LNNFAPK-----	-P--IAFH VQPLHPNLQE	MK-----
PhPI	-----N Q--RE-RE-Y-----	H Q-LP--FTFR LQPIQPNLHQ	NQ-----
LPI	-----H P---DRD-F--AA-----	-QMP--ITFR VQPSHPNLQE	NN-----
OsMADS2	-----HH D---DRD-F--AA-----	-S-MP--FTFR VQPSHPNLQE	EK-----
OsMADS4	-----D-Y-----	Q-MP--F.FR VQP.QPNLQE	-----
Consensus	-----D-Y-----	Q-MP--F.FR VQP.QPNLQE	-----

PI motif sequence "MPFxFRVQPxQPNLQE"

	PI Motif-Derived	EuAP3 Motif	
DEFICIENS	--GPRILALR LPTNH--P T--LH--S GGGS--	D LITTFALLE	227
SvAP3	--GPRILALR LPSNH--P N--LH--S GGGS--	D LITTFALLE	212
PMADS1	--GHRILALR LQPNH--QP MHHHLH--S GGGS--	D IITTFALLE	231
NTDEF	--GPRILALR LQPNH--QP MHH--LH--S GGGS--	D IITTFALG	227
LeAP3	--GPRILDLR LQPNH--YH MH--LH--S GGGS--	D IITTFALG	212
STDef	--GHRILALG LQPNH--LH--S GGGS--	D IITTFALG	228
AP3	--GSRALALR FQNH--MHA--LHAPS AS--	D IITFHILLE	232
BobAP3	--LR FQNH--MHA--LHAPS AS--	D IITFHILLE	224
Boi1AP3	--GSRALALR YQNH--MHA--LHAPS AS--	D IITFHILLE	232
Boi2AP3	--LR FQNH--MHA--LHAPS AS--	D IITFHILLE	224
RAD1	--SHLVGLH F-PREAH-IP-----S AGGS--	C LITTYTILE	220
RAD2	-----GAD -P-----T-----AAGS--	Y LITTYTILE	195
SLM3	--SRVGLALR LQPC--QP N--LHAGA GSGS--	C VITTYALL	227
NMH7	--LGPRMVALS LQPTH--P N--PHN--GGASAAS--	D LITTYLLESH FSLRIRTTNT TITFQQ	245

euAP3 motif sequence "D(L/I)ITTFALLE"

PD2	--VHNLVAFR LQPLH--P N--LQNE- GG--	F GSRDLRLS	222
TM6	--VHNLVAFR LQPLH--P N--LQNE- GG--	F GSRDLRLS	222
CMB2	AAA-NLVALS RHPIT-----	F GSRDLRLS	214
PtAP3-1a	--SHLFAFR LHPN--QP N--LHIN- GGG--	Y GFHNLHLA	211
PtAP3-1b	--SHLFAFR LHPN--QP N--LHIN- GGG--	Y GFHNLHLA	211
PcAP3	DCENSQITFQ LQPS--QP N--LHAA- GGG--	Y LYNQHYV	234
PnAP3-1	DCENSQITFQ LQPS--QP N--LHAA- GGG--	Y FYSQHYA	231
PnAP3-2	--PNI FAFR LQPS--QP N--LHN- GGG--	Y NCHDLRLA	228
DeAP3	--QNI FAFR LQPS--QP N--LND- GGG--	Y GSHDLRLA	209
CpAP3	--VFSFR LQPS--QP N--LND- EE--	Y EIHDLRLA	202
RbAP3	APQ--VFSFR LQPS--QP N--LND- EE--	Y EIHDLRLV	205
PhAP3	--PHFLQYN MQGN--P YHES-ARSDVTTANISSAY	Y GIYDLRLA	216
MfAP3	--AHI-----LHPT- G-----	F GIHDLRLA	200
CRM3	-----MT-S---ERS D-----	-SFLDLRLA	220
Consensus	-----F.FR LQP.---QP N---LH---		
PI Motif	-----F.FR VQP.---QP N---LQE--		
Core Consensus	-----F.FR VQP.---QP N---LQE--		

PaleoAP3 Motif

paleo motif sequence "YGxHDLRLA"

PI Motif-derived sequence "FxRFLQPSQPNL"

Figure 16 The consensus sequence in C-terminus indicating PI motif, PI motif-derived and euAP3 motif

Source: Kramer, Dorit, and Irish (1998)

Consequently, the other conserved region FxRFLQPSQPNL was found almost in DEF lineage proteins which greatly similar to PI motif, so termed as "PI motif-derived". The degree of PI motif-derived conserve the PI motif sequence is different in

species, TM6 and PD2 possess the most conserved level of PI motif among core eudicot DEF genes and the most divergent of PI motif is belong to RAD1 and RAD2 from *Rumex acetosa*. Thus, the TM6 lineage was first suggested, the member of this clades include CMB2 from *Dianthus caryophyllus* (same subclass to *R. acetosa*) and AsAP3 from *Argyroxiphium sanwicense* (Asteridae subclass) which shared the synapomorphic characters over PI-derived motif including amino acid Met, Gly Lys and Val in position 69, 118, 147 and 211 respectively. The euAP3 lineage also shared the synapomorphic sequence include Leu, His Leu, Asn and Ile in position 54, 55, 72, 148 and 150 respectively. The lower eudicots and magnolids represented truly paleoAP3 motif, as paleoAP3 lineage. In addition, the evolutionary rate of DEF/GLO lineage is faster than other MADS-box genes around 20-40% (Purugganan, Rounsley, Schmidt, & Yanofsky, 1995). This data supported the important of B-class genes that the changes in B functions will directly affect the independent diversification of floral morphology in core eudicots, lower eudicots, magnolid dicots and monocots.

The phylogenetic relationship between B-class genes was analyzed and suggested that there were two duplication events occurred (Figure 17). Because of all of flowering plants have both *DEF*- and *GLO*-like genes, that means the first duplication generated two lineage of DEF and GLO lineage before the diversification of angiosperms. With the highly conserved of PI motif and lost paleoAP3 motif of GLO lineage (PI lineage), this may achieve their lineage through single clade. However, the DEF lineage (AP3 lineage) had another duplication event occurred before the diversification of core eudicots. The absence of euAP3 motif in lower eudicots and magnolids and only presence in core eudicots, the production of second duplication event subsequently have distinguished euAP3 lineage and TM6 lineage from paleoAP3 lineage after the diversification of Buxaceae or before the major subclasses of core eudicots emerged (Kramer et al., 1998).

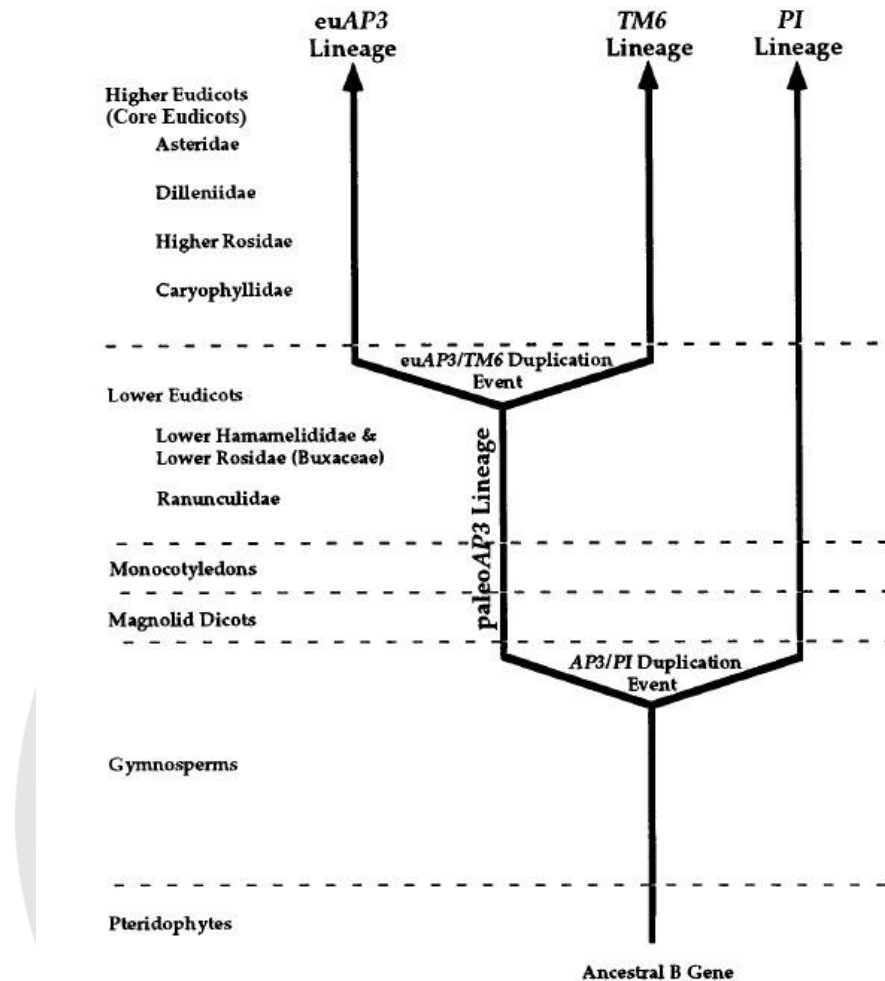


Figure 17 The B-class gene lineage duplication events along with land plants diversification

Source: Kramer et al. (1998)

***AGL6*-like subfamily**

According to floral quartet model, the E function became interesting because it interacts with ABC proteins and specify floral organ identity. The *AGL6*-like subfamily is one of the ancient group and interestingly sister clade to *AGL2* (*SEP*) and *SQUA*; *AGL2/AGL6/SQUA* superclade. In addition, the *FLOWERING LOCUS C* (*FLC*) is the key regulator gene in vernalization in grasses and eudicots also belong to this group.

Thus, *AGL6* protein might form multimeric complex and gain the E-functions. However, only *AGL6*-like was found in gymnosperm. This indicated that not only *AGL2*-like and

SQUA-like but also *AGL6*-like existed in common ancestor of extant seed plants. Moreover, the absence of *AGL2*-like and *SQUA*-like hypothesized that *AGL6*-like gene in gymnosperm is the basal lineage of angiosperms.

In gymnosperm, *DAL1* and *DAL14* from *Picea abies* are in *AGL6* subfamily. The *PrMADS2* and *PrMADS3* from *Pinus radiata* orthologous to *DAL14* and *DAL1* respectively. The *PrMADS2* and *DAL14* orthologs are presented in only male and female reproductive organs but the *PrMADS3* and *DAL1* are found in both reproductive organs and vegetative organs. In contrast, the *GGM9* and *GGM11* from *Gnetum gnemon* that were *DAL1* and *DAL14* orthologs respectively were expressed only in male and female reproductive organs. In core eudicots, the *AGL6*-like genes were divided into two groups, *AGL6*-like group and eu*AGL6* group. The *AGL6*-like group highly expressed only vegetative organ but eu*AGL6* group expressed in both floral organ and sometimes in vegetative tissue.

In *Oryza sativa* (rice) *AGL6*-like gene, *OsMADS6* or *MOSAIC FLORAL ORGANS1* (*MFO1*) is expressed in floral meristem especially in first, second and fourth whorl and the *OsMADS6* form tetrameric complex with *OsMADS2*-*OsmADS16* and *OsMADS4*-*OsMADS16* heterodimer to determine the second whorl identity. According to investigate grasses *AGL6*, two highly conserved in C-terminal region of *AGL6*-like genes had found, *AGL6*-I and *AGL6*-II motif. The *AGL6*-I and *AGL6*-II sequence is DCEPTLQIGY and ENNFMLGWVL respectively. This similar to SEP I and SEP II motif in *AGL2* lineage. Thus, the functions could be predicted as the transcriptional activation in floral quartet complex.

The phylogenetic relationship of *AGL6*-like shown that there were 4 clades of *AGL6*-like genes in monocots, *AGL6*-I to *AGL6*-IV (Figure 18). The orchids *AGL6* genes were found in *AGL6*-III and *AGL6*-IV. The genome sequence of *Phalaenopsis equestris* found three *Phalaenopsis AGL6* genes and suggested the expansion of diversified class B, C, D MADS-box genes including *AGL2*, *AG*, *DEF* and *AGL6* might involve in the origin of orchid evolutionary novelties or the unique morphology achievement of orchid flower. This supported with the expression profiles that distinctively represented spatial pattern; *PaAGL6-1* (flower specific), *PaAGL6-1* (lip specific), *PaAG-2* and *PaAG-3* (column specific) and *PaAG-4* (pedicel specific).

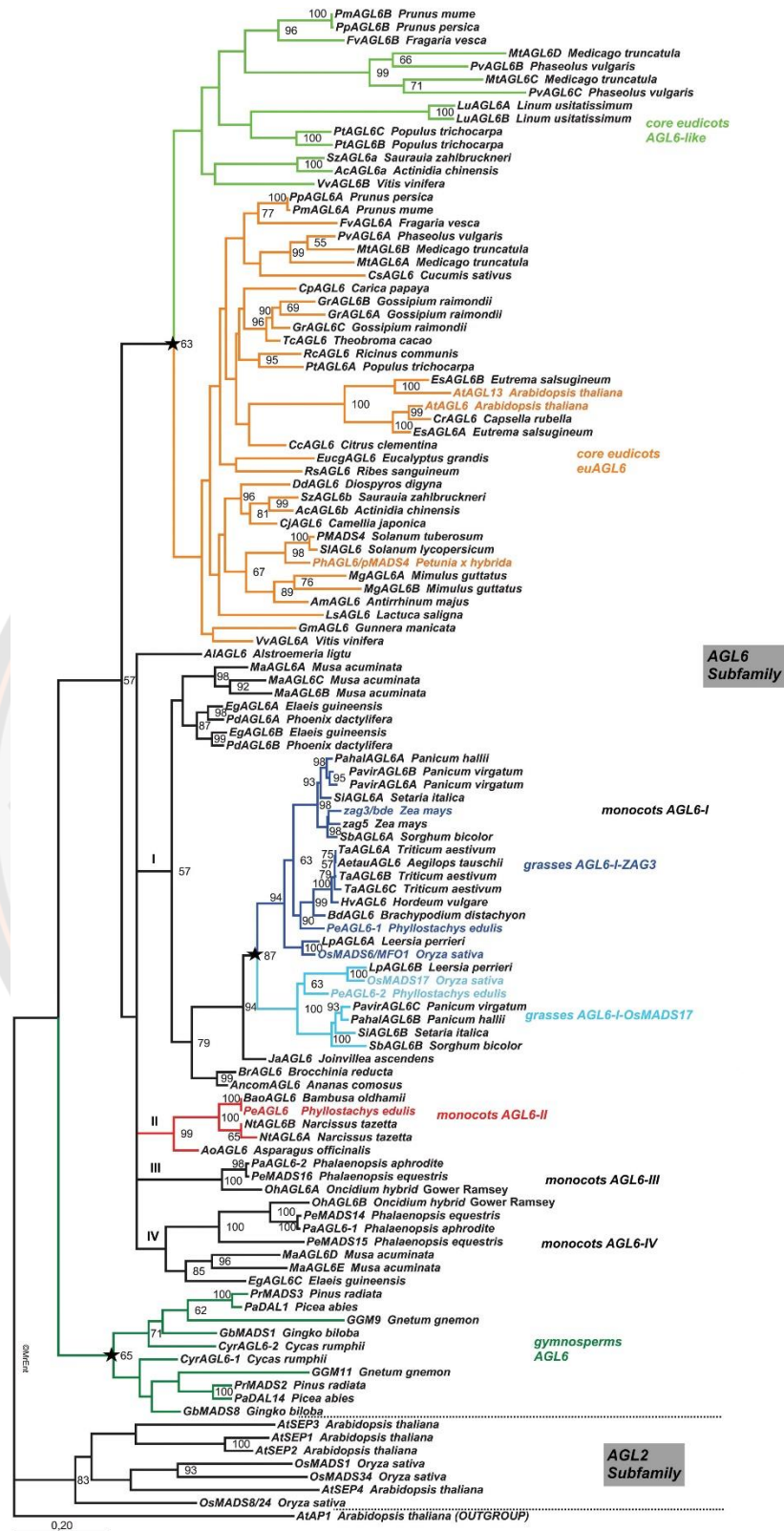


Figure 18 The phylogenetic relationship of AGL6 subfamily.

Source: Dreni and Zhang (2016)

The orchid MADS-box genes

Along the diversification of Orchid family led the questions concerning the evolution since the 19th century. The origin of labellum and gynostemium are great interesting. More than 20 years ago after the finding that floral organ identity specified by the interaction between MADS-domain transcription factors. The first MADS-box gene was isolated from orchid is *om1* from *Aranda* “Deborah” which was first considered as FBP2 (*Petunia hybrida*) homologs from sequence but the expression patterns were different. After more MADS-box genes were characterized and reconstructed the phylogenetic tree, the *om1* was verified as *SEPALLATA3*-like gene and orthologs to *DOMADS* and *DcOSEP1* (Zahn et al., 2005).

According peloric tetara have been reported in spontaneous and cultured orchids, it is essential to identify the genetically organization of orchids flower through MADS-box genes isolation, expression pattern analysis and developmental pathway to specify each organ. The first generation research of expression comparison between wild type and peloric mutant was in *Oncidium* Gower Ramsey, *Phalaenopsis equetris* *Dendrobium crumenatum* and *Habenaria radiata* using only B-class MADS-box genes (Hsu & Yang, 2002; Tsai et al., 2004; Xu et al., 2006). In the last two decades, isolation and characterization of in class B (*DEF*-like) and class E (*AGL6*-like) MADS-box genes from various orchids occurred at faster pace. Except to Apostasioideae which was the most initially diverging group, did not possess accurate lip organs and species poor, there were investigated widespread in all subfamilies of orchid family (Figure 19) including Vanilloideae, Cypripedioideae, Orchidoideae and Epidendroideae which indicated in Table 1. Most of them were focused on perianth speciation and led the genetic model for conserved identities of different orchid flower organs.

Orchid code model

Orchid code originated from the question “How the orchids got their lip?”. This question has provoked the scientists to find the answer. Based on the developmental genetic basic behind the morphological diversity, the most recent expression pattern data of *OMADS3*, *PeMADS2*, *PeMADS5*, *DcOAP3A* and *HrDEF*, the orchid code suggested the combination of B-class MADS-box genes determining specific floral

organ identity (Figure 20). The expression of *DEF*-like clade 1 and 2 expressed in outer tepals (sepals). The expression of *DEF*-like clade 1, 2 and 3 expressed in inner tepals (petals). The expression of *DEF*-like clade 3 and 4 expressed in lip. This hypothesized that clade 1 and 2 expressed in all tepals, the distinction between outer tepals and inner tepals depend on the expression of clade 3 and the identity of lip structure is specified by the expression of clade 4. This study additionally suggested that the orchid might diverged from lily-like flower and differentiated during two round of gene/genome duplication that occurred before the separation of Vanilloideae from higher orchids. The emergence of clade 3 and 4 was in first duplication event to distinguish inner from outer tepals. Then, the second duplication provided clade 4 to distinguish lip from inner tepals. These supported by the Apostasioideae representing the intermediate state which have not achieved lip structure yet.

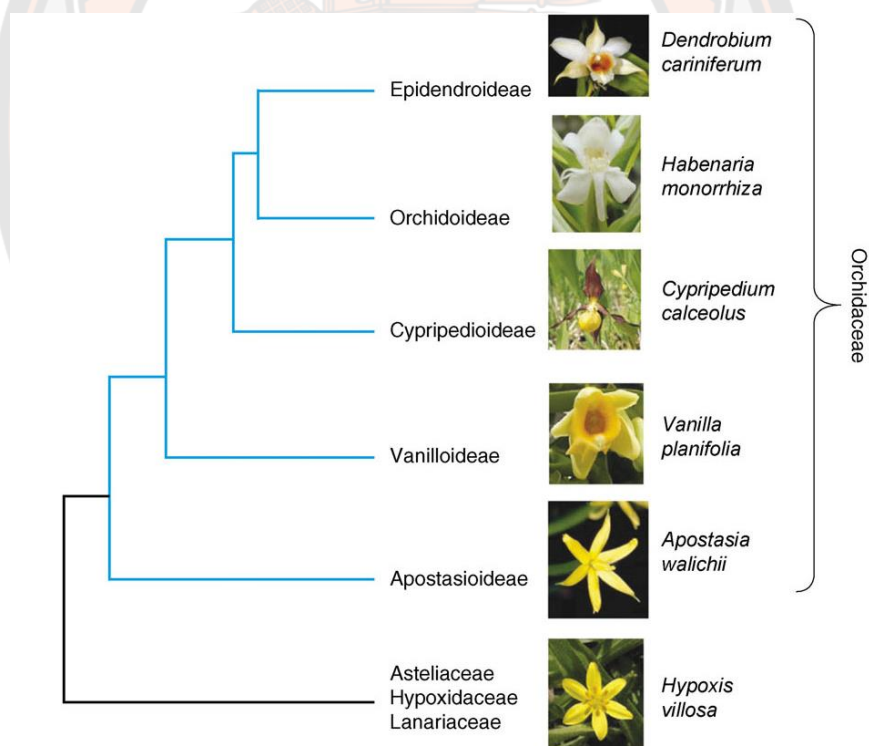


Figure 19 The phylogenetic tree of subfamily in Orchidaceae family. The *Hypoxis villosa* from Hypoxidaceae represented as the closest clade of the orchids.

Source: Mondragón-Palomino and Theißen (2008)

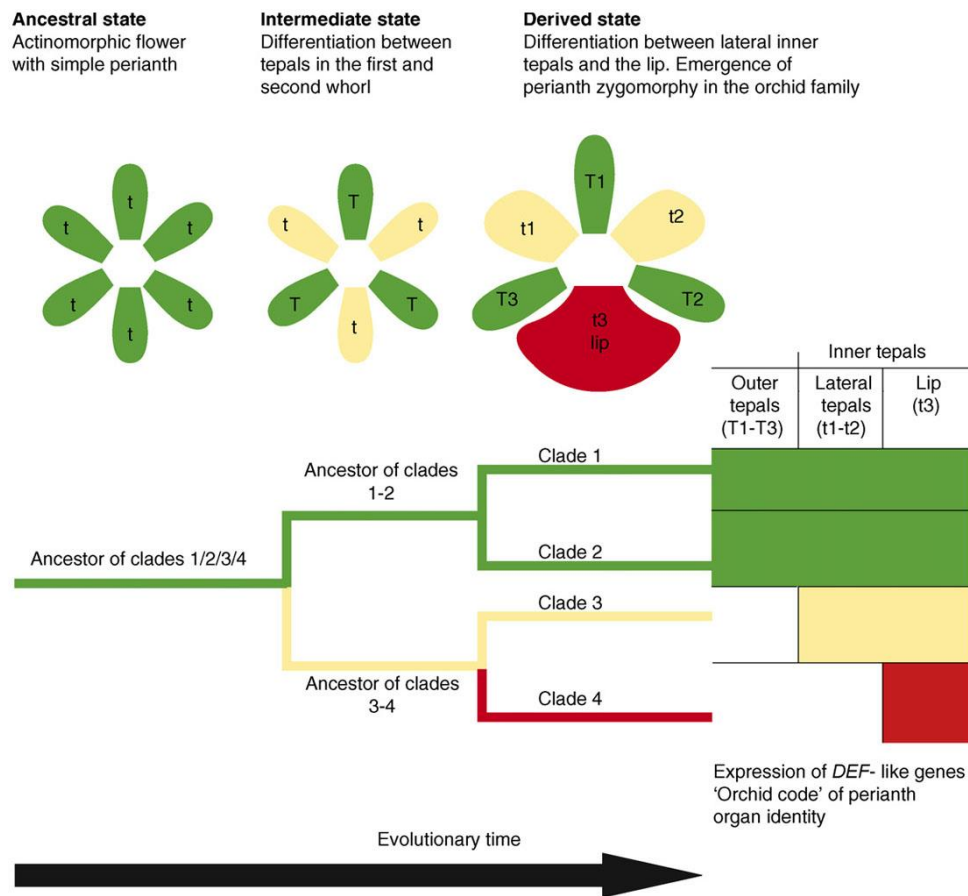


Figure 20 The orchid code and the hypothesis of the orchid perianth origin.
Source: Mondragón-Palomino and Theißen (2008)

Table 1 Recent Class B and E MADS-box genes characterized in Orchidaceae

Species	Genes	Class/Clade	References
<i>Oncidium</i> "Gower Ramsey" (Epidendroideae)	<i>OMADS3</i>	Class B/DEF/Clade 2	Hsu and Yang (2002)
	<i>OMADS5</i>	Class B/DEF/Clade 1	Chang et al. (2010)
	<i>OMADS9</i>	Class B/ DEF/Clade 3	
	<i>OMADS1</i>	Class E/AGL6/Clade 2	Chang, Chiu, Wu, and Yang (2009); (Hsu, Huang, Chou, & Yang, 2003)
	<i>OMADS7</i>	Class E/AGL6/Clade 1	Chang et al. (2009)
	<i>OncAP3-1</i>	Class B/DEF/Clade 1	
	<i>OncAP3-2</i>	Class B/DEF/Clade 3	
	<i>OncAP3-3</i>	Class B/ DEF/Clade 4	

Species	Genes	Class/Clade	References
	<i>OncAP3-4</i>	Class B/DEF/Clade 2	(Pan et al., 2011) Hsu et al. (2015)
	<i>OAP3-1</i>	Renamed from OMADS3	
	<i>OAP3-2</i>	Renamed from OMADS3	
	<i>OAGL6-1</i>	Renamed from OMADS3	
	<i>OAGL6-2</i>	Renamed from OMADS3	
<i>Vanilla ptilifera</i> (Vanilloideae)	<i>VaAP3-1</i>	Class B/DEF/Clade 1	
	<i>GalAP3-1</i>	Class B/DEF/Clade 1	
	<i>GalAP3-2</i>	Class B/DEF/Clade 3	
	<i>GalAP3-3</i>	Class B/ DEF/Clade 4	
<i>Paphiopedilum</i> Macabre (Cypripedioideae)	<i>PaphPe</i>	Class B/DEF/Clade 1	
	<i>PaphAP3-1</i>	Class B/DEF/Clade 3	
	<i>PaphAP3-2</i>	Class B/ DEF/Clade 4	
<i>Anoectochilus formosanus</i> (Orchidoideae)	<i>AfAP3-1</i>	Class B/DEF/Clade 1	
<i>Habenaria petelotii</i> (Orchidoideae)	<i>HpAP3-1</i>	Class B/DEF/Clade 1	
<i>Liparis distans</i> (Epidendroideae)	<i>LdAP3-1</i>	Class B/DEF/Clade 1	
	<i>LdAP3-2</i>	Class B/DEF/Clade 3	
<i>Phaius tankervilleae</i> (Epidendroideae)	<i>PtAP3-1</i>	Class B/DEF/Clade 1	
	<i>PtAP3-2</i>	Class B/DEF/Clade 3	
	<i>PtAP3-3</i>	Class B/ DEF/Clade 4	
<i>Brassavola nodosa</i> (Epidendroideae)	<i>BnAP3-1</i>	Class B/DEF/Clade 4	
	<i>BnAP3-2</i>	Class B/DEF/Clade 1	
	<i>BnAP3-3</i>	Class B/ DEF/Clade 3	
<i>Dendrobium</i> Spring Jewel (Epidendroideae)	<i>DenAP3-1</i>	Class B/DEF/Clade 1	
	<i>DenAP3-2</i>	Class B/DEF/Clade 3	
	<i>DenAP3-3</i>	Class B/ DEF/Clade 4	
<i>Phalaenopsis equestris</i> (Epidendroideae)	<i>PeMADS2</i>	Class B/DEF/Clade 1	Tsai et al. (2004)
	<i>PeMADS3</i>	Class B/DEF/Clade 3	
	<i>PeMADS4</i>	Class B/ DEF/Clade 4	
	<i>PeMADS5</i>	Class B/DEF/Clade 2	
<i>Dendrobium crumenatum</i> (Epidendroideae)	<i>DcOAP3A</i>	Class B/DEF/Clade 1	Xu et al. (2006)
	<i>DcOAP3B</i>	Class B/DEF/Clade 3	
<i>Dendrobium moniliforme</i> (Epidendroideae)	<i>DMAP3A</i>	Class B/DEF/Clade 1	Sirisawat et al. (2010)
	<i>DMAP3B</i>	Class B/DEF/Clade 3	Sirisawat et al. (2009)
	<i>DMMADS4</i>	Class B/ DEF/Clade 4	
<i>Vanilla planifolia</i> (Vanilloideae)	<i>VaplaDEF1</i>	Class B/DEF/Clade 1	Mondragón-Palomino and Theissen (2011)
	<i>VaplaDEF2</i>	Class B/DEF/Clade 4	
	<i>VaplaDEF3</i>	Class B/DEF/Clade 3	
<i>Phragmipedium longifolium</i> (Cypripedioideae)	<i>PhlonDEF1</i>	Class B/DEF/Clade 2	
	<i>PhlonDEF2</i>	Class B/DEF/Clade 1	
	<i>PhlonDEF3</i>	Class B/DEF/Clade 3	
	<i>PhlonDEF4</i>	Class B/DEF/Clade 4	
<i>Phalaenopsis</i> hyb. "Athens" (Epidendroideae)	<i>PeMADS2</i>	Class B/DEF/Clade 1	
	<i>PeMADS3</i>	Class B/DEF/Clade 3	
	<i>PeMADS4</i>	Class B/ DEF/Clade 4	
	<i>PeMADS5</i>	Class B/DEF/Clade 2	

Species	Genes	Class/Clade	References
<i>Phalaenopsis Aphrodite</i> (Epidendroideae)	<i>PaAP3-1</i>	Class B/DEF/Clade 1	
	<i>PaAP3-2</i>	Class B/DEF/Clade 2	
	<i>PaAP3-3</i>	Class B/DEF/Clade 4	
	<i>PaAP3-4</i>	Class B/DEF/Clade 3	
	<i>PaAGL6-1</i>	Class E/AGL6/Clade 2	
	<i>PaAGL6-2</i>	Class E/AGL6/Clade 1	
<i>Habenari radiata</i> (Orchidoideae)	<i>HrDEF</i>	Class B/DEF/Clade 3	Kim et al. (2007) Mitoma et al. (2019)
	<i>HrDEF-C1</i>	Class B/DEF/Clade 1	
	<i>HrDEF-C2</i>	Class B/DEF/Clade 2	
	<i>HrDEF-C3</i>	Class B/DEF/Clade 3	
	<i>HrDEF-C4</i>	Class B/DEF/Clade 4	
	<i>HrAGL6-1</i>	Class E/AGL6/Clade 1	
<i>Erycina pusilla</i> (Epidendroideae)	<i>EpMADS13</i>	Class B/DEF/Clade 1	Lin et al. (2016)
	<i>EpMADS14</i>	Class B/DEF/Clade 2	
	<i>EpMADS15</i>	Class B/DEF/Clade 3	
	<i>EpMADS3</i>	Class E/AGL6/Clade 1	
	<i>EpMADS4</i>	Class E/AGL6/Clade 3	
	<i>EpMADS5</i>	Class E/AGL6/Clade 2	
<i>Cymbidium goeringii</i> (Epidendroideae)	<i>CgDEF1</i>	Class B/DEF/Clade 1	Xiang et al. (2018)
	<i>CgDEF2</i>	Class B/DEF/Clade 2	
	<i>CgDEF3</i>	Class B/DEF/Clade 3	
	<i>CgDEF4</i>	Class B/DEF/Clade 4	
	<i>CgAGL6-1</i>	Class E/AGL6/Clade1	
	<i>CgAGL6-2</i>	Class E/AGL6/Clade3	
	<i>CgAGL6-3</i>	Class E/AGL6/Clade2	

The revised “orchid code” model

The expression data to generated orchid code limited only in Orchidoideae and Epidendroideae which is the most recently evolved subfamily. However, the other two subfamily remain vague. Thus, to understand precisely about orchid flower identity, Vanilloideae and Cypridioideae were extensively studied. The *DEF*-like genes were characterized from *Vanilla planifolia* and *Phragmidipedium longifolium*, represented Vanilloidea and Cypridioideae subfamily. In addition, the comparison of expression pattern of *DEF*- and *GLO*-like genes between wild type and peloric mutant from *Phalaenopsis* hybrid ‘Athens’ (Epidendroideae) had also been performed. The results indicated that the expression of *DEF*-like genes clade 1 and 2 with the absent expression of clade 3 and 4 specify for outer tepals. For lateral inner tepals, they required the higher level expression of clade 3 and 4 with the lower level expression of clade 1 and 2. In addition, both of the lower or absent of *DEF*-like genes clade 1 and 2 and the notably high level of clade 3 and 4 will determine the lip identity (Figure 21).

Upon the perianth determination, all four clades of *DEF*-like genes also represented in reproductive organs, clade 3 and 4 were observed in column which are higher in gynostemium than ovary. However, the expression pattern of clade 1 and 2 of *DEF*-like genes in reproductive organs depended on species and developmental stages. Moreover, the clade 1 and 2 genes are relative lower than clade 3 and 4 genes in reproductive organs.

Additionally, these three subfamilies diverged around 70 MYA but the expression pattern relatively conserved and four *DEF* paralogs functions are correlated, this shown that these genes existed for long time and these genes still keep the regular mechanism of upstream factors after the duplication not depending on the downstream that will affect in B-class target genes and have no change in floral organ identity (Mondragón-Palomino & Theißen, 2008; Mondragón-Palomino & Theissen, 2011). This refining model, the revised “orchid code” model is proposed the combination of different levels of four *DEF*-like genes that distinctively differ from previous orchid code model that simply refer only the gene expression through “on” or “off” activities.

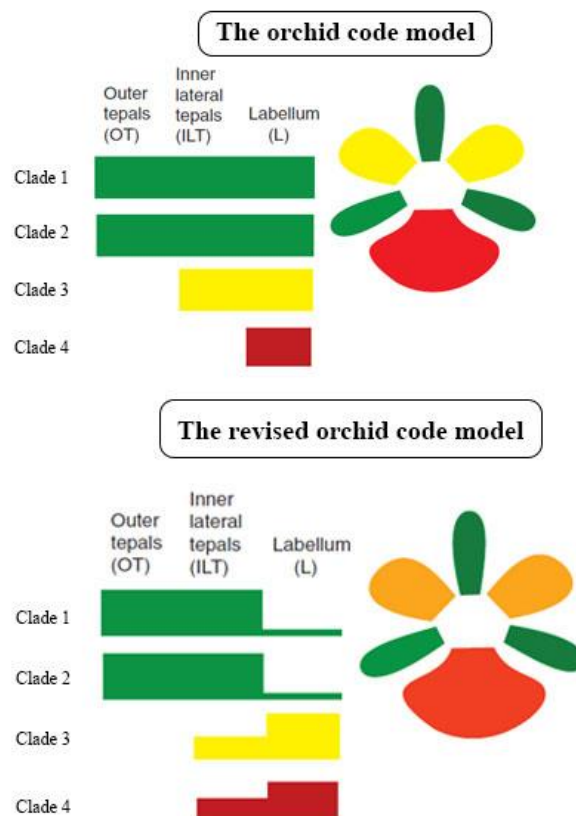


Figure 21 The comparison of previous orchid code model and the revised orchid code model

Source: Mondragón-Palomino and Theissen (2011)

The perianth code model

The perianth code model (P-code model) is the most recent and complete model that describe the orchid perianth formation at the moment (Figure 22). The origin of this model was from the expression examination in *Oncidium* orchids. Particularly in *Oncidium* Gower Ramsey, the *OPI* (GLO-like gene) expressed generally in all perianth organs. The *OAP3-1* and *OAGL6-1* genes, *DEF*-like gene clade 1 and *AGL6*-like gene clade 1 respectively both highly expressed in sepals and petals but loss in lip. However, the *OAP3-2*, or *DEF*-like gene clade 3 are notably detected in petals and lip. Additionally, the *OAGL6-2*, or *AGL6*-like gene clade 2 only expressed in lip and some in lateral sepals. Consequently, the perianth code was proposed the heteromeric complex of *DEF* and *AGL6* homologs determine the perianth formation including sepal/petal (SP) complex and lip (L) complex. The SP complex consist of *OAP3-1/OAGL6-1/OAGL6-1/OPI* specify for sepal/petal organs. The L complex consist of *OAP3-2/OAGL6-2/OAGL6-2/OPI*. In perianth code, *OPI* form ubiquitously in all perianth and only specifically apply for perianth not in stamen or carpel formation. This model could generally refer in other orchids that the SP complex is *DEF-1/AGL6-1/AGL6-1/GLO* and the L complex is *DEF-3/AGL6-2/AGL6-2/GLO*.

The competition between sepal/petal (SP) complex and lip (L) complex is similar to the balance. The presence of SP complex and loss L complex will promote sepal/petal development. Conversely, the presence of L complex and loss of SP complex will promote lip development (F). The coexistence and co-absence of SP and L complex will turn the balance into middle and generate the intermediate structures. For example, the GRtrip mutant of *O. Gower Ramsey* has only expression of L complex (*OAP3-2/OAGL6-2*) and loss of SP complex (*OAP3-1*) resulted in the full transformation of petals to lip-like structure. The x Beallara Eurosta (Bllra) shown the mutant type (Bllra-Trip) that received the sepal/petal-lip intermediate structure effecting from the coexistence of SP and L complex. The co-absence also causes the

transformation of petal to sepal/petal like lip intermediate structure in mutant of *Psychopsis papilio*.

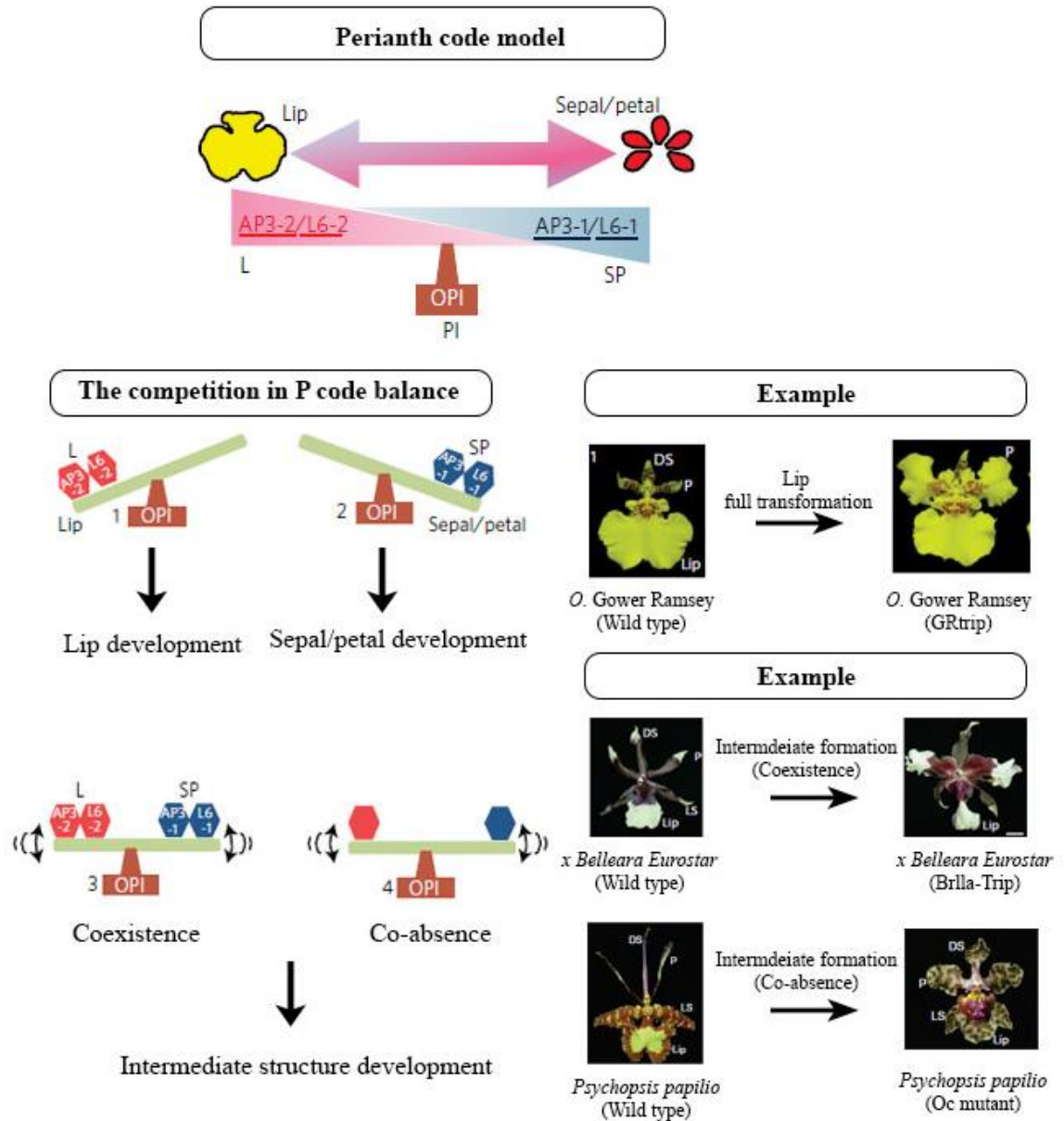


Figure 22 Perianth code model

Source: modified from Hsu et al. (2015)

The regulation of P code complexes was suggested from performing fluorescence resonance energy transfer (FRET). The physical protein interaction has different efficiency to localize to nucleus and switch on the “Lip program” (Figure 23). The L complex had evolved to promote the lip program but the SP complex conversely suppress Lip program and promote sepal/petal formation. This hypothesis supported by virus-induced gene silencing in *OAGL6-2*; *OAGL6-2* VIG. The results shown the suppression in L function, the reduced size of lip and the existence of green sepal/petal sector on lip organ were occurred. The homodimer of each *OAGL6-1* and *OAGL6-2* were formed without OPI protein and localize to nucleus and the heterodimers of *OAP3-1/OPI* and *OAP3-2/OPI* were formed and localized to nucleus. This suggested the higher heterotetrameric complexes, following floral quartet complex consist of one AP3, one PI and two E class protein in controlling petal development, SP tetramer (*OAP3-1/OAGL6-1/OAGL6-1/OPI*) switching off the Lip program and L complex (*OAP3-2/OAGL6-2/OAGL6-2/OPI*) switching on the Lip program.

The P code model clarify the perspective of orchid perianth diversity. The SP complex (PI/euAP3/E-class) initially maintain in petal formation of eudicots. Then, SP complex (PI/paleoAP3/AGL6-like) extend to all perianth and replace sepals with petaloid structure in monocots. In orchid family, except to Apostasioideae, the P code hypothesis also investigated in four subfamilies. The L complex is the novel genes from second gene/genome duplication. This event occurred after the first duplication of SP complex. Thus, the variation of lip in most orchid species is distinct lip organs and the intermediate structure, resulting from loss of SP complex and the coexistence of SP and L complex respectively (Hsu et al., 2015).

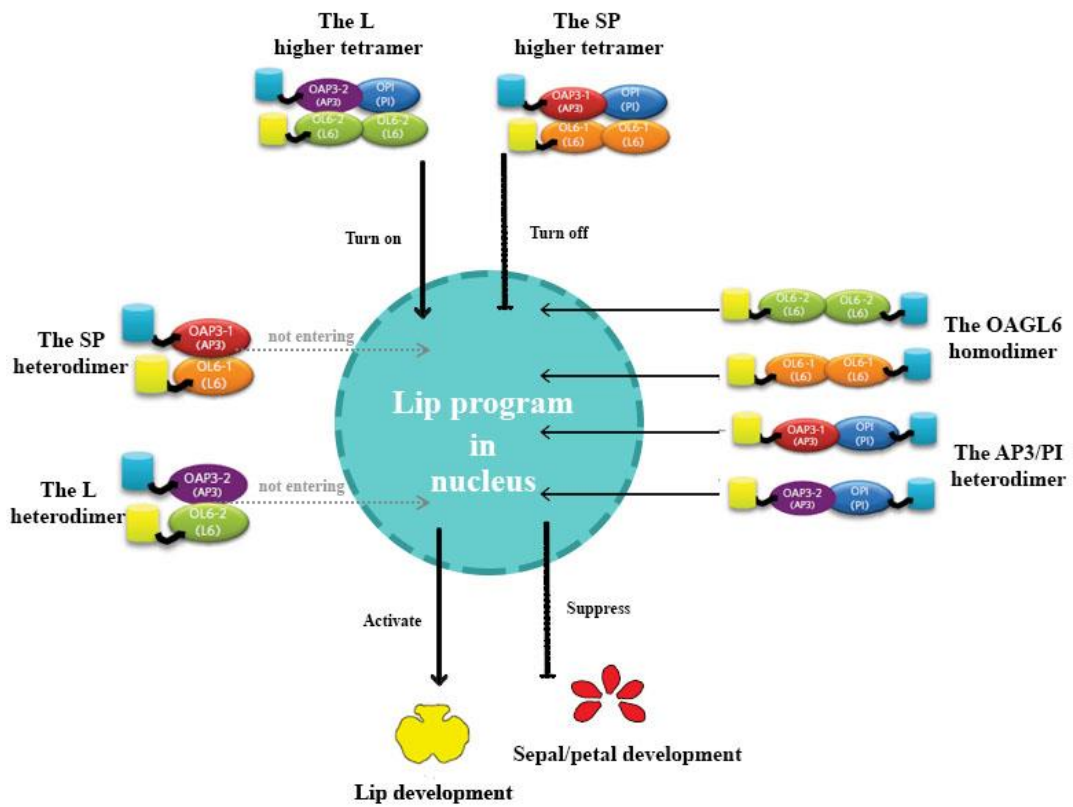


Figure 23 The summary of P code complexes regulation

Source: modified from Hsu et al. (2015)

CHAPTER III

RESEARCH METHODOLOGY

Plant materials

Rhynchostylis retusa plants used in this study were grown in an experimental nursery at the Department of Biology, Faculty of Science, Naresuan University, Phitsanulok, Thailand. Different developmental stages of *R. retusa* flowers from 1-4 (stage 1: 0.5 cm; stage 2: 1.0 cm; stage 3: 1.5 cm; stage 4: mature) and various organs of the flowers at stage 3 and 4 (mature) including sepals, petals, lips, columns. From stage 1 to stage 4 approximately took 15 days. In addition, leaves were also collected and represented as vegetative organs (Figure 24).

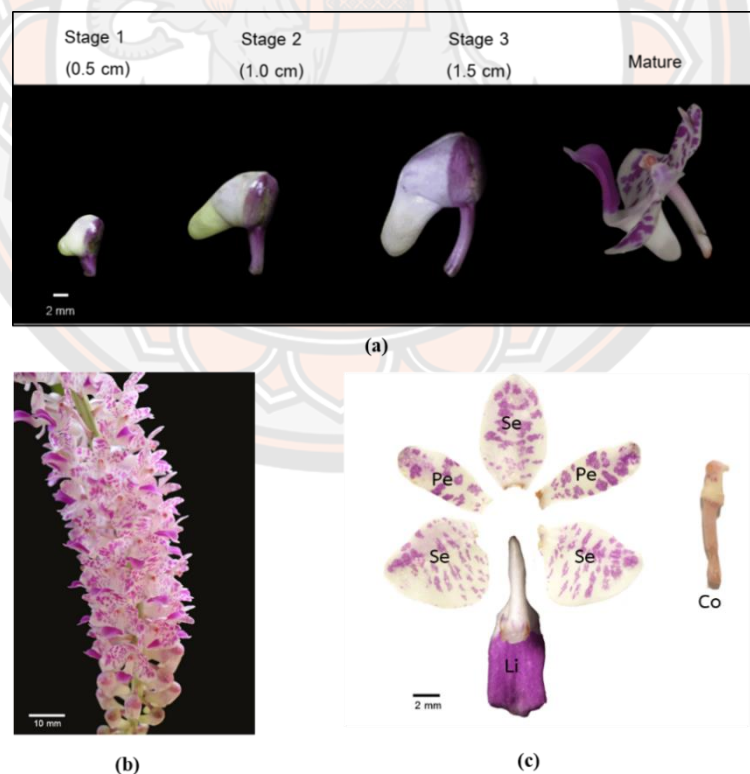


Figure 24 Morphology of *R. retusa* wild type flower. (a) Developmental stages 1-4 (stage 1: 0.5 cm; stage 2: 1.0 cm; stage 3: 1.5 cm; stage 4: mature),

(b) Racemose cluster of *R. retusa* inflorescence, and (c) Dissected flower organs including sepals, petals, lip and column.

Chemicals and instruments

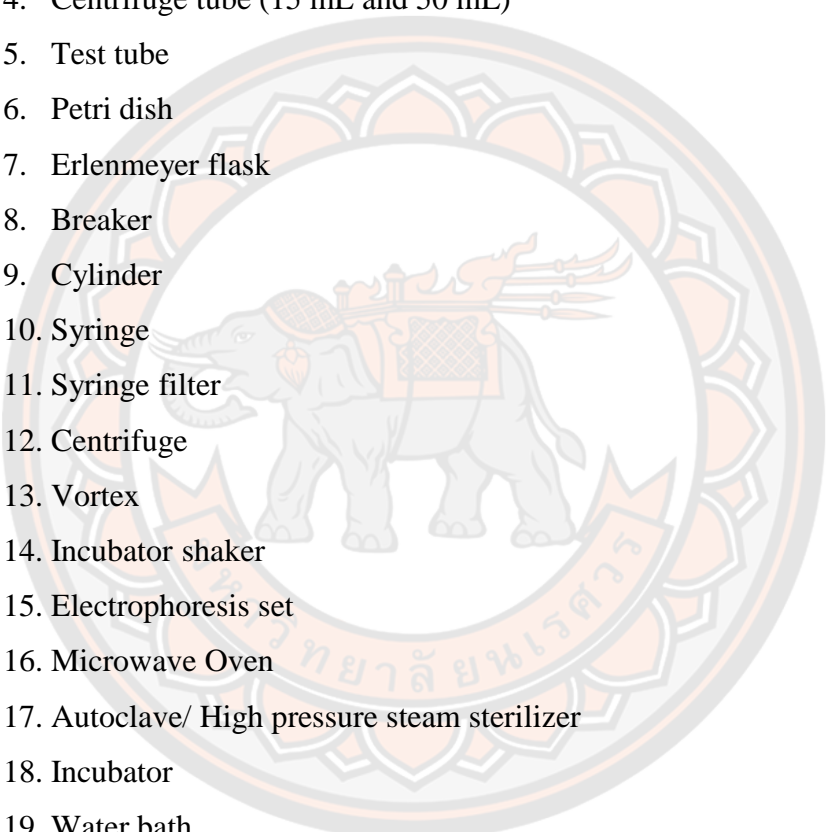
Chemicals

1. TRIzol™ reagent (Thermo Fisher Scientific, USA)
2. DNase I (Thermo Scientific, USA)
3. RevertAid First Strand cDNA synthesis kit (Thermo Scientific, USA)
4. Phusion™ High-Fidelity DNA polymerase (Thermo Scientific, USA)
5. iTaq Plus DNA polymerase (iNtRON Biotechnology, USA)
6. Thermo Scientific GeneJET gel extraction kit (Thermo Scientific, USA)
7. CloneJET PCR cloning kit (Thermo Scientific, USA)
8. *Eschericia coli* strain DH5 α
9. Plasmid miniPREP kit (PureDirex, Taiwan)
10. SensiFAST SYBR No-ROX kit (Bioline, USA)
11. SOB medium
12. SOC medium
13. 2X YT medium
14. 100 mg/ml Ampicillin
15. 2M D(+)-Glucose
16. 1M MgSO₄
17. 1M MgCl₂
18. 1M CaCl₂
19. Nuclease-free water
20. Autoclaved distilled water
21. 1X TAE buffer
22. 75% ethanol
23. Liquid nitrogen
24. Chloroform
25. Agarose
26. Ethidium bromide (EtBr)
27. 6X DNA Loading Dye

28. 100 bp DNA ladder

29. 1 Kb DNA ladder

Instruments

1. Mortar and pestle
 2. Pipette and Tip
 3. Microcentrifuge tube (0.2 mL and 1.5 mL)
 4. Centrifuge tube (15 mL and 50 mL)
 5. Test tube
 6. Petri dish
 7. Erlenmeyer flask
 8. Breaker
 9. Cylinder
 10. Syringe
 11. Syringe filter
 12. Centrifuge
 13. Vortex
 14. Incubator shaker
 15. Electrophoresis set
 16. Microwave Oven
 17. Autoclave/ High pressure steam sterilizer
 18. Incubator
 19. Water bath
 20. Heat box
 21. UV box
 22. Biometra TAdvance Thermal Cycler (AnalytikJena, Germany)
 23. Nabi-UV/Vis Nano Spectrophotometer (Laborimpex, Belgium)
 24. CFX connect real-time PCR detection system (BIO-RAD, USA)
- 

RNA extraction

Total RNA from all four developmental stages flower buds, flower organs and leave was extracted by TRIzol™ reagent (Thermo Fisher Scientific, USA) according to manufacturer' protocol.

1. 100 mg of tissue were cleaned and wiped before grinding with liquid nitrogen in mortar
2. Add 1000 µL of TRIzol™ Reagent to the fine sample
3. Incubated for 5 minutes to permit complete dissociation of the nucleoproteins complex
4. Added 200 µL of chloroform per 1000 µL of TRIzol™ Reagent used for lysis
5. Incubated for 5 minutes
6. Centrifuge the sample for 15 minutes at 12000 rpm at 4°C
7. Transferred the colorless upper aqueous phase to new tube (Avoid transferring any of the interphase or organic layer into the pipette when removing the aqueous phase.)
8. Added 500 µL of isopropanol to the aqueous phase and mixed by gently inverting
9. Incubated overnight at -20°C
10. Centrifuged for 15 minutes at 12,000 rpm at 4°C
11. Total RNA precipitate forms a white gel-like pellet at the bottom of the tube. Discarded the supernatant with a micropipettor
12. Resuspended the pellet in 1000 µL of 75% ethanol
13. Vortexed the sample briefly, then centrifuged for 5 minutes at 7500 rpm at 4°C.
14. Discarded the supernatant with a micropipettor
15. Air dried the RNA pellet
16. Incubated in a water bath or heat block set at 55°C for 10 minutes
17. Stored the RNA at -80°C until proceeding to downstream applications
18. Determined for its amount and purity by a Nabi-UV/Vis Nano Spectrophotometer (Laborimpex, Belgium) using absorbance at 260 and 280 nm
19. The integrity of the RNA samples was evaluated on 1% (w/v) agarose TAE gel electrophoresis stained with ethidium bromide (EtBr).

RNA analysis

The total extracted RNA was determined the amount and purity by Nabi-

UV/Vis Nano Spectrophotometer (Laborimpex, Belgium) using the absorbance at 260 and 280 nm. In addition, the integrity of RNA samples was evaluated on 1% (w/v) agarose TAE gel electrophoresis stained with Ethidium bromide (EtBr). The total RNA was migrated and represented the 28S and 18S rRNA of eukaryote samples. The ratio of 28S:18S should be 2:1 for good indication of the intact RNA.

First strand cDNA synthesis

The total RNA samples from *R. retusa* were used as template to generate first strand cDNA in a reverse transcription reaction. Firstly, DNA in RNA sample was removed. Total RNA (3000 ng) was treated with DNase I (Thermo Scientific, USA) following the manufacturer's protocol. Then, eleven microliters of DNase-treated RNA were immediately reverse transcribed with RevertAid First Strand cDNA synthesis kit (Thermo Scientific, USA) using P019HA adapter primer (5'-GACTCGTGACGACATCGATTTTTTTTTTTTTTTTTT -3') according to the manufacturer's protocol.

1. Genomic DNA from RNA was removed by adding the following reagents into a sterile, nuclease free tube on ice in the indicated order

Total RNA	1 µg
10X Reaction Buffer with MgCl ₂	1 µL
DNase I, RNase-free	1 µL (1 U)
Water, nuclease-free	to 10 µL

2. Incubated at 37°C for 30 minutes

3. Added 1 µL 50 mM EDTA and incubated at 65°C for 10 minutes

4. The cDNA synthesis was performed by Thermo Scientific RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific, USA)

5. Added the following reagents into a sterile, nuclease-free tube on ice in the indicated order

Total RNA	0.1-5 µg
Oligo (dT) ₁₈ Primer/ P019HA adapter primer	1 µL
Water, nuclease-free	To 12 µL

6. Incubated at 65 °C for 5 minutes

7. Chilled on ice, spun down and placed tube back on ice

8. Added the following components in the indicated order

5X Reaction Buffer	4 μ L
RiboLock RNase Inhibitor (20U/ μ L)	1 μ L
10 mM dNTP Mix	2 μ L
RevertAid M-MuLV RT (200 U/ μ L)	1 μ L
Total volume	20 μ L

9. Mixed gently and centrifuged briefly
10. Incubated at 42°C for 60 minutes
11. Terminated the reaction by heating at 70°C for 10 minutes
12. Directly used in PCR applications or stored at -20°C for less than one week. For longer storage, -70°C is recommended.

PCR amplification of partial sequence of *DEF*-like and *AGL6*-like genes

The degenerate primers were designed base on the nucleotide sequence of *AGL6*–like and *DEF*-like genes in orchids specifically targeting the MADS and K domain. The first strand cDNA for partial *AGL6*-like and *DEF*-like genes isolation were synthesized from stage 3 flower buds. The PCR amplification was performed using iTaq Plus DNA polymerase (iNtRON Biotechnology, USA) with specific primers. Gene specific primer for *RrDEF-1*, *RrDEF-3* and *RrDEF-4* was S-Hr*DEF* primer pair (5'- AACTGCGYGGTCTTGAGCAA-3' and 5'- AYYADGCRAGRCKDAGATCCTG-3'). A *RrDEF-2* was amplified with *DEF-2* primer pair (5'-ATGGGGAGAGGGAAGGTAGAGATAA-3'and 5'-GAACTACTTTCTGCACAATTGGC-3'). A *RrAGL6-1* was amplified with Orchid*AGL6-1* primer pair (5'- CTGAAGAGGATTGAGAAC-3' and 5'- GCATCCACCCAAGCATAA-3'). Finally, a *RrAGL6-2* was amplified with *AGL6-2* primer pair (5'-AGGCAAAGAGGACGCAGATA-3' and 5'-GTTCTGTGTCCATGTTACTTGAA-3'). PCR reactions were prepared in nuclease-free tube and set up on ice and followed the steps below.

1. Added the following components in the indicated order

Nuclease-free water	to 20 μ L
10X iTaq MgCl ₂ free buffer	2 μ L
25 mM MgCl ₂	2 μ L
10 mM dNTPs	0.4 μ L

10 μ M forward primer	0.4 μ L
10 μ M reverse primer	0.4 μ L
Template cDNA	2 μ L
iTaq TM DNA polymerase (5 U/ μ L)	0.15L

2. The components were mixed and spun down. The thermal cycling was performed in Biometra TAdvance Thermal Cycler (AnalytikJena, Germany) and programmed as follow

Initial denaturation	94°C	for 3 minutes	} 35 cycles
Denaturation	94°C	for 30 seconds	
Annealing	54°C	for 30 seconds	
Extension	72°C	for 30 seconds	
Final extension	72°C	for 5 minutes	
Hold	4°C	∞	

3. Evaluated the amplified products on 1.5% (w/v) agarose TAE gel electrophoresis stained with ethidium bromide (EtBr)

3'-Rapid amplification of cDNA ends (3'RACE)

The 3'-rapid amplification of cDNA ends (3'-RACE) is a procedure for amplification of nucleic acid sequences from cDNA which synthesized from adapter primer and utilized the poly (A) region by anchored PCR using gene specific primers that anneal to known sequence with a primer that attaches to an adapter sequence, P018HA (5'- GACTCGTGACGACATCG-3').

The gene specific primers for DEF-like and AGL6-like was obtained from previous amplified sequence and designed in the conjunction. Thus, the forward primers were provided the specific sequence to get 3' ends of each genes. For RrDEF-1, RrDEF-2, RrDEF-3, RrDEF-4, RrAGL6-1 and RrAGL6-2, the primers were qRrDEF-1 (5'-AGCACAAGGGAACTTACCGC-3'), qRrDEF-2(5'-AGGAAGGG GGAGAATCTGGA-3'), qRrDEF-3 (5'CTCTCAAGAAACCCACAGGAAC-3'), qRrDEF-4 (5'- CTCTCAAGAAACACACCGAAAC-3'), qRrAGL6-1 (5'-CGTCAACTTGGAGAGATCAATAAG-3') and qRrAGL6-2 (5'-GTTGGACCAG

ATGGAAGAGC-3') respectively. The PCR amplification were performed using Phusion™ High-Fidelity DNA polymerase (Thermo Scientific, USA). The reaction component was performed following the steps below.

1. PCR reactions were prepared in nuclease-free tube and set up on ice. Added the following components in the indicated order

Nuclease-free water	to 20	μL
5X Phusion™ HF buffer	4	μL
10 mM dNTPs	0.4	μL
10 μM forward primer	0.4	μL
10 μM adapter primer (P018HA)	0.4	μL
Template cDNA	2	μL
Phusion™ High-Fidelity DNA polymerase	0.2	μL

2. The components were mixed and spun down. The thermal cycling was performed in Biometra TADvance Thermal Cycler (AnalytikJena, Germany) and programed as follow

Initial denaturation	98°C	for 30 seconds	} 35 cycles
Denaturation	98°C	for 10 seconds	
Annealing	54°C	for 30 seconds	
Extension	72°C	for 30 seconds	
Final extension	72°C	for 5 minutes	
Hold	4°C	∞	

3. Evaluated the amplified products on 1.5% (w/v) agarose TAE gel electrophoresis stained with ethidium bromide (EtBr)

Gel purification

The gel purification was performed using Thermo Scientific GeneJET gel extraction kit (Thermo Scientific, USA) following the manufacturer' protocol.

1. The target PCR product fragment was excised from 1.5% agarose gel TAE electrophoresis stained with EtBr

2. Excised gel slice containing the desire fragment using a clean scalpel or razor blade.

3. Pre-weighed 1.5 mL tube, placed the gel slice into the tube and weighed.

Then, record the weight of the gel slice.

4. Added 1:1 volume of Binding Buffer to the gel slice (volume: weight)
5. Incubated at 60 °C for 10 min in heat box until the gel slice is completely dissolved
6. Mixed by inversion and vortexed briefly
7. Transferred up to 800 µL of the mixture to column (with collection tube)
8. Centrifuged at 12000 rpm for 1 minute
9. Discarded the flow-through and placed the column back into the collection tube
10. Added 700 µL of Wash Buffer into the column
11. Centrifuged at 12000 rpm for 1 minute
12. Discarded the flow-through and placed the column back into the collection tube
13. Centrifuged the empty column at 12000 rpm for 1 minute to completely remove residual wash buffer
14. Transferred the column into new nuclease-free 1.5 mL tube
15. Added nuclease-free water to the center of the column membrane
16. Waited for 15 minutes and centrifuged at 12000 rpm for 1 minute
17. Discarded the column, collected the flow-through and evaluated the purified products on 1.5% (w/v) agarose TAE gel electrophoresis stained with ethidium bromide (EtBr)
18. Stored at -20°C

Gene cloning

Construction of a recombinant DNA

The purified PCR products were first ligated into the pJET1.2/blunt cloning vector using CloneJET PCR cloning kit (Thermo Scientific, USA). This vector system contains a lethal gene that is interrupted by ligation. Therefore, only cells with recombinant plasmids are able to propagate (Figure 25). The reaction component was added following the steps below.

1. Ligation reactions were prepared in nuclease-free tube and set up on ice.

Added the following components in the indicated order

2X Reaction Buffer	10 µL
purified PCR product	1 µL
Water, nuclease-free	Up to 17 µL
DNA Blunting Enzyme	1 µL

- | | |
|-------|------------|
| Total | 18 μ L |
|-------|------------|
- Vortexed briefly and spun down
 - Incubated the mixture at 70°C for 5 min and chilled on ice
 - Added the following components

pJET1.2/blunt Cloning Vector (50 ng/ μ L)	1 μ L
T4 DNA Ligase	1 μ L
Total volume	20 μ L
 - Vortexed briefly and spun down
 - Incubated at room temperature (22°C) for 5 minutes
 - Used the ligation mixture directly for transformation or stored at -20°C

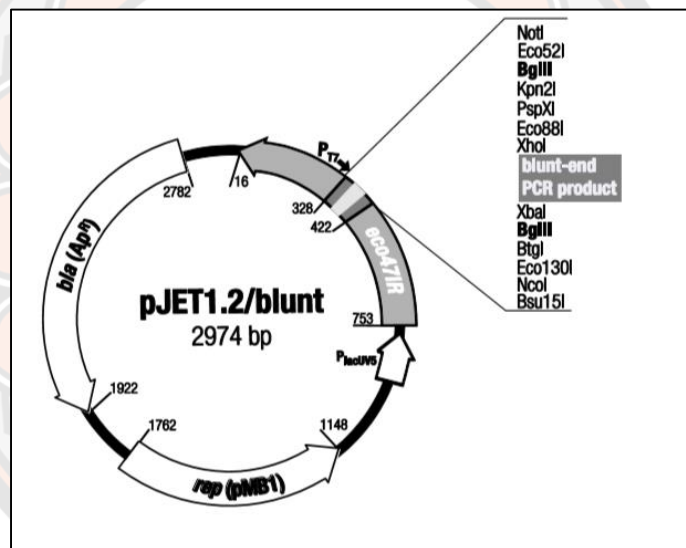


Figure 25 Map of pJET1.2/blunt cloning vector (Thermo Scientific, USA)

Source: CloneJET PCR cloning kit' manufacture protocol (Thermo Scientific, USA)

Preparation of competent cells

- Streaked the *Escherichia coli* strain DH5 α on the 2X YT plate and incubated at 37 °C for 16-20 hours
- Aliquoted 2 mL of 50 mL SOB medium (added MgSO₄ and MgCl₂) into 15 mL centrifuge tube
- Picked a single colony from plate and transferred it into 3 mL of SOB medium
- Incubated cells at 37 °C with vigorous shaking at 200 rpm for 16-18 hours

on a incubator shaker

5. Transferred 0.5 mL of the grown cells to residual 47 mL of SOB medium and Incubated cells at 37 °C with vigorous shaking at 200 rpm for 2-3 hours
6. Transferred cells aseptically to 50 mL centrifuge tube and leaved on ice for 30 minutes
7. Centrifuged cells at 3000 rpm at 4 °C for 10 minutes and discarded the supernatant
8. Resuspended the cell pellet in 10 mL of 100 mM CaCl₂ and chilled on ice
9. Centrifuged cells at 3000 rpm at 4 °C for 10 minutes and discarded the supernatant
10. Added 1 mL of 100 mM CaCl₂ to dissolve the cell pellet and aliquoted 0.2 mL of competent cells into 1.5 mL microcentrifuge tube
11. Stored in ice for immediately transformation

Transformation of plasmid DNA into *E. coli* using the heat shock method

1. Added 7 µL of ligation mixture to competent cell tube and gently mixed by pipetting
2. Chilled on ice for 30 minutes
3. Heated shock at 42 °C for 1 minutes and immediately placed on ice for 5 minutes
4. Added 800 mL of SOC broth and incubated at 37 °C with vigorous shaking at 200 rpm for 1 hour
5. Centrifuged at 3000 rpm for 10 minutes and discarded 700 µL of supernatant
6. Mixed the cell by pipetting and spread 50-100 µL of mixture on 2X YT agar plate (added Ampicillin)
7. Incubated at 37 °C overnight and selected the colonies to check insertion

Colony PCR

1. For screening the colonies with the presence of an insert, the white colonies were picked by a sterile toothpick or pipette tip into 20 µL of 2X YT broth and waited for 1 minutes
2. The PCR amplification were performed using iTaq Plus DNA polymerase

(iNtRON Biotechnology, USA).

3. PCR reactions were prepared in nuclease-free tube and set up on ice. Added the following components in the indicated order

Nuclease-free water	to 20 μ L
10X iTaq MgCl ₂ free buffer	2 μ L
25 mM MgCl ₂	2 μ L
10 mM dNTPs	0.4 μ L
10 μ M pJET1.2/Fw	0.4 μ L
10 μ M pJET1.2/Rv	0.4 μ L
Colony mixture	2 μ L
iTaq TM DNA polymerase (5 U/ μ L)	0.15 μ L

4. The components were mixed and spun down. The thermal cycling was performed in Biometra TAdvance Thermal Cycler (AnalytikJena, Germany) and programed as follow

Initial denaturation	94°C	for 3 minutes	} 30 cycles
Denaturation	94°C	for 30 seconds	
Annealing	60°C	for 30 seconds	
Extension	72°C	for 30 seconds	
Final extension	72°C	for 5 minutes	
Hold	4°C	∞	

5. Evaluated the amplified products on 1.5% (w/v) agarose TAE gel electrophoresis stained with ethidium bromide (EtBr)

6. The correct size products were selected to inoculate by pipetting 2 μ L of colony mixture to 2 mL of 2X YT broth

7. Incubated at 37°C with vigorous shaking at 200 rpm for 16-18 hours

Plasmid extraction

1. The inserted *E. coli* cells were harvested by transferred the bacterial culture to microcentrifuge tube

2. The pelleted bacterial cells were extracted by Plasmid miniPREP kit (PureDirex, Taiwan) following the manufacturer' protocol.

3. Transferred bacterial culture to a 1.5 mL microcentrifuge tube

4. Centrifuged at 12000 rpm for 1 minutes and discarded the supernatant
5. Added 200 μ L of the Buffer S1 (added RNase A) to resuspend the bacterial cell pellet
6. Added 200 μ L of the Buffer S2 and mix thoroughly by inverting the tube 10 times (Do not vortex)
7. Chilled at the room temperature for 2 minutes or until the lysate is homologous
8. Added 300 L of the Buffer S3 and mixed immediately and inverted thoroughly 10 times (Do not vortex)
9. Centrifuged at 12000 rpm for 3 minutes
10. Pipetted the supernatant to column (with collection tube)
11. Centrifuged at 12000 rpm for 30 seconds, discarded the flow-through and placed the column back to the collection tube
12. Added 400 μ L of the Buffered W1 into the column
13. Centrifuged at 12000 rpm for 30 seconds, discarded the flow-through and placed the column back to the collection tube
14. Added 600 μ L of the Buffered W2 (added Ethanol) into the column
15. Centrifuged at 12000 rpm for 30 seconds, discarded the flow-through and placed the column back to the collection tube
16. Centrifuged at 12000 rpm for 2 minutes to remove the residual Buffer W2
17. To elute DNA, placed the PM column in a clean 1.5 ml microcentrifuge tube
18. Added 50-200 μ L of the nuclease-free water to the center of each column, let it stand for 15 minutes, and centrifuged at 12000 rpm for 2 minutes
19. Discarded the column, collected the flow-through and evaluated the purified products on 1.5% (w/v) agarose TAE gel electrophoresis stained with ethidium bromide (EtBr)

Sequence data analysis

The extracted plasmid was sent to Macrogen Co., LTD (Korea) to sequence. All sequencing reactions contained the standard pJET1.2 sequencing primer. Raw DNA sequence data were edited with GeneStudioTM software (GeneStudio, Inc, USA) to remove vector and poly A sequences as well as poor quality data. The multiple sequence alignment was aligned by the ClustalW (EMBL-EBI, UK). The amino acid sequences were generated using EMBOSS Transeq translates nucleic acid sequences (EMBL-EBI,

UK). Each sequences were compared with National Center for Biotechnology Information (NCBI) database. The phylogenetic analysis of *DEF*- and *AGL6*-like gene nucleotide sequences were constructed by maximum-likelihood tree with 1000 bootstrap replicates in MEGA 7.0.26 (Kumar et al., 2015)

Expression analysis

RNA extraction

Total RNA was isolated from different developmental stages of *R. retusa* flower buds from stages 1-4 and sepals, petals, lip and column from stage 3 and 4 by TRIzol™ reagent (Thermo Fisher Scientific, USA) according to manufacturer' protocol followed previous section.

cDNA synthesis

1. The cDNA synthesis was performed by ReverTra Ace™ qPCR RT master mix with gDNA remover (TOYOBO, Japan)
2. Prepared the 4xDN master mix and gDNA remover mixture by adding a 1 in 50 volumes of gDNA remover to 4xDN master mix (This mixture can be stored at -20°C for at least 3 months)
3. Genomic DNA from RNA was removed by adding the following reagents into a sterile, nuclease free tube on ice in the indicated order

4x DN master mix	2 µL
RNA template	0.5 pg – 0.5µg
Nuclease-free Water	To 8 µL

4. Incubated at 37°C for 5 minutes
5. Prepared the reverse transcription solution by adding the following reagents

Reaction mixture from 11.2.4.	8 µL
5x RT Master Mix II	2 µL
Total	10 µL
6. Incubated at 37°C for 15 minutes
7. Incubated at 98°C for 5 minutes
8. Stored at 4°C or -20°C until use in RT-PCR and Real-time PCR.

Reverse transcription PCR (RT-PCR)

1. to examine the expression patterns of *RrDEF-1*, *RrDEF-2*, *RrDEF-3*, *RrDEF-4*, *RrAGL6-1* and *RrAGL6-2*. The primers of *eukaryotic translation elongation factor 1A (eEF1A)* was used as reference gene for internal control.

2. PCR reactions were performed using iTaq Plus DNA polymerase (iNtRON Biotechnology, USA) and prepared in nuclease-free tube and set up on ice. Added the following components in the indicated order

Nuclease-free water	to 20 μ L
10X iTaq MgCl ₂ free buffer	2 μ L
25 mM MgCl ₂	2 μ L
10 mM dNTPs	0.4 μ L
10 μ M forward primer	0.4 μ L
10 μ M reverse primer	0.4 μ L
Template cDNA	2 μ L
iTaq™ DNA polymerase (5 U/ μ L)	0.15L

3. The components were mixed and spun down. The thermal cycling was performed in Biometra TAdvance Thermal Cycler (AnalytikJena, Germany) and programed as follow

Initial denaturation	94°C	for 3 minutes	} 35 cycles
Denaturation	94°C	for 30 seconds	
Annealing	54°C	for 30 seconds	
Extension	72°C	for 30 seconds	
Final extension	72°C	for 5 minutes	
Hold	4°C	∞	

4. Evaluated the amplified products on 1.5% (w/v) agarose TAE gel electrophoresis stained with ethidium bromide (EtBr)

Quantitative real-time PCR

1. The quantitative real-time PCR (qPCR) of cDNA were performed using 2X SensiFAST SYBR No-ROX kit (Bioline, USA) and prepared in nuclease-free tube and set up on ice. Added the following components in the indicated orderly

Nuclease-free water	to 20 μ L
---------------------	---------------

2x SensiFAST SYBR [®] No-ROX Mix	10 μ L
10 μ M forward primer	0.8 μ L
10 μ M reverse primer	0.8 μ L
Template cDNA	Up to 8.4 μ L

2. The components were mixed and spun down. The thermal cycling was performed in the CFX connect real-time PCR detection system (BIO-RAD, USA) and programmed as follow

Initial denaturation	94°C	for 5 minutes	} 35 cycles
Denaturation	94°C	for 30 seconds	
Annealing	54°C	for 30 seconds	
Extension	72°C	for 30 seconds	
Melting curve	60-95°C		

3. All qPCR experiments were performed in triplicate and the results were analyzed using the comparative Ct method ($2^{-\Delta\Delta C_t}$ method)

Table 2 List of oligonucleotide sequences

Primer name	Sequence (5' → 3')	Application
S-HrDEF/Fw	AACTGCGYGGTCTTGAGCAAA	DEF-specific primer
S-HrDEF/Rv	AYYADGCRAGRCKDAGATCCTG	DEF-specific primer
DEF-2/Fw	ATGGGGAGAGGGAAGGTAGAGATAA	DEF-2-specific primer
DEF-2/Rv	GAACTACTTTCTGCACAATTGGC	DEF-2-specific primer
OrchidAGL6-1/Fw	CTGAAGAGGATTGAGAAC	AGL6-1-specific primer
OrchidAGL6-1/Rv	GCATCCACCCAAGCATAA	AGL6-1-specific primer
AGL6-2/Fw	AGGCAAAAGAGGACGCAGATA	AGL6-2-specific primer
AGL6-2/Rv	GTTCTGTGTCCATGTTACTTGAA	AGL6-2-specific primer
P018HA	GACTCGTGACGACATCG	Adapter primer
P019HA	GACTCGTGACGACATCGATTTTT TTTTTTTTTTTT	Adapter primer
eEF1A/Fw	TAAGTCTGTTGAGATGCACC	Reference gene primer
eEF1A/Rv	CTGGCCAGGGTGGTTCATGAT	Reference gene primer
qRrAGL6-1/Fw	CGTCAACTTGGAGAGATCAATAAG	Real time PCR/3'RACE for <i>RrAGL6-1</i> gene
qRrAGL6-1/Rv	TGAATTCGAGTGGTAAGGGTGC	Real time PCR for <i>RrAGL6-1</i> gene
qRrAGL6-2/Fw	GTTGGACCAGATGGAAGAGC	Real time PCR/3'RACE for <i>RrAGL6-2</i> gene

Primer name	Sequence (5' → 3')	Application
qRrAGL6-2/Rv	GCTTCTGGGGCCAATATTGATA	Real time PCR for <i>RrAGL6-2</i> gene
qRrDEF-C1/Fw	AGCACAAGGGAACTTACCGC	Real time PCR/3'RACE for <i>RrDEF-C1</i> gene
qRrDEF-C1/Rv	CAACCCTAAAGGAAAACATCTGAG	Real time PCR for <i>RrDEF-C1</i> gene
qRrDEF-C2/Fw	AGGAAGGGGGAGAATCTGGA	Real time PCR/3'RACE for <i>RrDEF-C2</i> gene
qRrDEF-C2/Rv	CAGAGAAAGTATCATGTGATCGC	Real time PCR for <i>RrDEF-C2</i> gene
qRrDEF-C3/Fw	CTCTCAAGAAACCCACAGGAAC	Real time PCR/3'RACE for <i>RrDEF-C3</i> gene
qRrDEF-C3/Rv	GCTTGGTTGGGACGAAATGAAT	Real time PCR for <i>RrDEF-C3</i> gene
qRrDEF-C4/Fw	CTCTCAAGAAACACACCGAAAC	Real time PCR/3'RACE for <i>RrDEF-C4</i> gene
qRrDEF-C4/Rv	CGGAAGGCATACATGTGAGAC	Real time PCR for <i>RrDEF-C4</i> gene
pJET1.2/Fw	CRACTACTATAGGGAGAGCGGC	Colony PCR/sequencing primer
pJET1.2/Rv	AAGAACATCGATTTTCCATGGCAG	Colony PCR/sequencing primer



CHAPTER IV

RESULTS AND DISCUSSION

RNA extraction

Total RNA was extracted from whole flower bud from stage 1-3 and mature flower of *R. retusa*, dissected flower organs at stage 3 and mature flower including sepals, petals, lips, columns and leaves. The purity of extracted total RNA was evaluated by Nabi-UV/Vis Nano Spectrophotometer (Table 3). The ratios of $A_{260/280}$ were approximately 2.0. These indicated that there were significant low contamination of proteins or polysaccharide. The ratios of $A_{260/230}$ were constantly <2.0 , these were lower than expected which may had been contaminated of contaminant which absorbed at 230 nm including TRIzol™ reagent which is phenol based solution. The integrity of total RNA samples was analyzed on 1.5% (w/v) agarose TAE gel electrophoresis stained with EtBr (Figure 26). The results showed that most of them had clear 28S and 18S rRNA bands with ratio 2:1 of 28S:18S rRNA and low degradation.

Table 3 Purity of total RNA from *R. retusa* using spectrophotometer

Samples	OD			Ratio		Concentration (ng/μl)
	A ₂₆₀	A ₂₈₀	A ₂₃₀	A _{260/280}	A _{260/230}	
Flower buds stage 1	30.60	15.41	18.66	1.99	1.64	1224
Flower buds stage 2	18.03	9.09	10.61	1.98	1.70	721
Flower buds stage 3	7.37	3.96	8.99	1.86	0.82	295
Mature flowers	16.53	8.41	10.33	1.97	1.60	661
Flower bud stage 3 organs						
Sepal	11.83	6.18	6.88	1.91	1.72	473.2
Petal	3.43	1.84	8.17	1.87	0.42	137.3
Lip	18.17	8.89	18.93	2.04	0.96	726.9
Column	22.27	12.14	20.62	1.83	1.08	890.9
Mature flower organs						
Sepal	19.23	9.46	11.25	2.03	1.71	769.3
Petal	7.53	3.94	4.25	1.91	1.77	301.2
Lip	14.24	7.43	8.79	1.92	1.62	569.7
Column	22.81	11.86	19.50	1.92	1.17	912.2
Vegetative organ						

Leave (L)	3.534	2.019	5.556	1.75	0.64	141.4
-----------	-------	-------	-------	------	------	-------

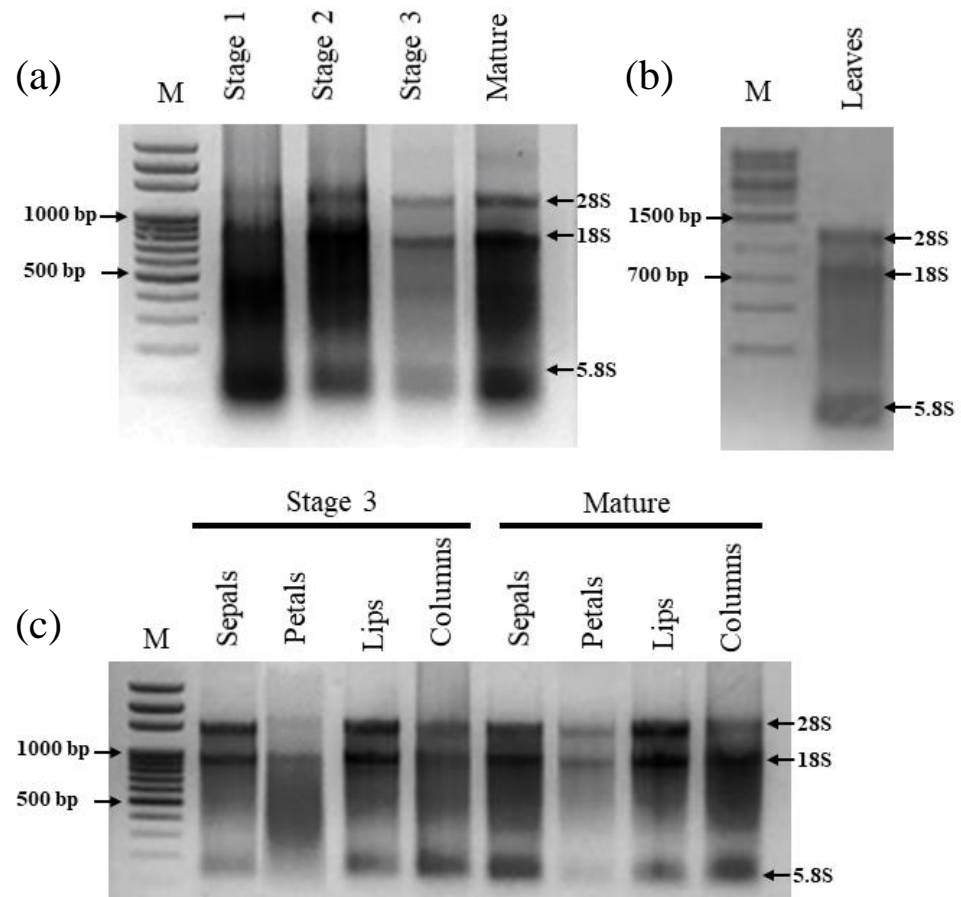


Figure 26 Assessment of RNA integrity by agarose gel electrophoresis stained with EtBr. Total RNA from stage1-3 whole bud and mature flower. (b) Total RNA from leaves. (c) Total RNA from dissected flower organs; sepals, petals, lips and columns from stage3 and mature flowers.

First partial sequences of *DEF*-like and *AGL6*-like genes from *R. retusa*

1. PCR amplification

To isolate *DEF*-like and *AGL6*-like MADS-box genes from *R. retusa*, the degenerated primers were used to obtain partial sequences of cDNA. The whole floral bud cDNA was amplified using primers specific to *DEF*-like and *AGL6*-like genes. S-

HrDEF forward primer and S-HrDEF reverse primer is specific for *DEF*-like genes. The expected amplification size was approximately 350 bp (Figure 27a). OrchidAGL6-1 forward and reverse primer was specific for *AGL6*-like genes. The expected amplification size was approximately 700 bp (Figure 27b). The specific product sizes were obtained. In addition, the DEF2 and AGL6-2 forward and reverse primers were separately designed for *DEF*-like clade 2 and *AGL6*-like clade 2 genes respectively. The DEF2/FW and DEF2/RV primer was designed from *DEF*-like clade 2 orthologs in *O.* ‘Gower’ Ramsey (*OMADS3*; accession number AY196350.1), *P. equetris* (*PeMADS5*; accession number Y378148.1) and *C. goeringii* (*CgDEF2*; accession number KX347446.1) which generated the approximately band size as 600 bp (Figure 27c). In addition, for *AGL6*-like clade2, the specific primers were designed from orthologous *AGL6*-like clade 2 in *C. goeringii* (*CgAGL6-3*; accession number KU058679) and *E. pusilla* (*EpMADS5*; accession number KJ002730) name as AGL6-2/FW and AGL6-2/RV primers corresponding with the size as 200 bp (Figure 27d). The purified PCR products were ligated into the pJET1.2/blunt cloning vector using CloneJET PCR cloning kit (Thermo Scientific, USA) and randomly selected clones to sequence and compare to the NCBI sequence database.

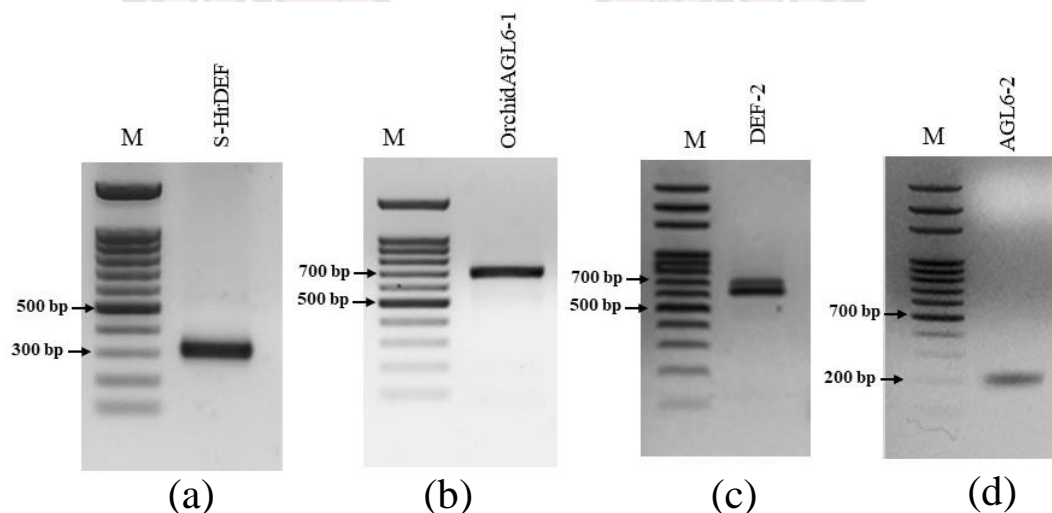


Figure 27 The PCR amplification of first partial sequences of *DEF*-like and

AGL6-like from *R. retusa* by agarose gel electrophoresis stained with EtBr; (a) primer S-HrDEF, (b) Primer OrchidAGL6-1, (c) Primer DEF-2 and (d) Primer AGL6-2

2. Sequence analysis

The raw sequence data (abi) were received from Macrogen Co., LTD (Korea). The DNA sequences from each colony were edited and aligned using GeneStudio™ software (GeneStudio, Inc, USA) and the ClustalW (EMBL-EBI, UK) respectively. The colony sequences from S-HrDEF primer were arranged into 3 clades of *DEF*-like genes (clade 1, 3 and 4) representing in APPENDIX B Figure 50-52. The consensus sequence alignment indicated that S-HrDEF degenerated primers producing *DEF*-like clade 1, clade 3 and clade 4 in *R. retusa* and the products had size 330, 312 and 318 bp respectively (Figure 28). The DEF2/FW and DEF2/RV primers generated *DEF*-like clade 2 which had size 721 bp (Figure 29). The OrchidAGL6-1 also produced the expected PCR product size as 706 bp and for *AGL6*-like clade 2 genes, AGL6-2/FW and AGL6-2/RV primer produced the target sequence as size 205 bp (Figure 30). All the sequences were confirmed with BLAST® results which compared to NCBI database (APPENDIX B Table 5-10).

>Consensus_S-HrDEF_DEF1

```
A A C T G C G C G G T C T T G A G C A A A C T T T G G A A G A G T C T C T G A G A A T T G T T A G G C A T A G A A A G T A T C A T G T G A T
C G C C A C A C A A A C T G A C A C T T A C A G A A A A A G C T T A A A A G C A C A A G G G A A A C T T A C C G C G C T C T A A T A C A T
G A A C T G G A T A T G A A A G A G G A G A A T C C G A A C T A C G G T T T T A A T G T G G A A A G C C A T A G T A G A A T T T A T G A A A
A T T C A A T T C C A A T G G T G A A T G A G T G C C C T C A G A T G T T T T C C T T T A G G G T T G T T C A T C C A A A T C A G C C C A A
T C T G C T T G G C T T A G G T T A T G A A T C A C A G G A T C T T A G C C T C G C C T A A T A A T
```

>Consensus_S-HrDEF_DEF3

```
A A C T G C G C G G T C T T G A G C A A A A T A T G G A C G A G G C C C T A A A G C T T G T A A G G A A T C G A A A G T A T C A C G T C A T
C A G C A C G C A G A C A G A T A C C T T C A A A A A A A G T T G A A A A A C T C T C A A G A A A C C C A A G G A A C T T A C T A C G G
G A G C T G G A A A C T G A G C A T G C C G T C T A C T A T G T G G A T G A T G A T C C A A C A A C T A T G A T G G C G C G C T T G C A C
T T G G A A A T G G G G C T T C C T A C T T G T A T T C A T T T C G T A C C C A A C C A A G C C A G C C A A A C C T T C A A G G A T G G G
A T A T G T C C C T C A G G A T C T T C G T C T C G C C T A A T
```

>Consensus_S-HrDEF_DEF4

```
A A C T G C G C G G T C T T G A G C A A A A C A T C G A C G A G G C A T T G A A G C T A G T A C G A A A T A G A A A T A T C A T G T A A T
C A G T A C T C A A A C G G A C A C C T A C A G A A G A A G C T G A A G A A C T C T C A A G A A A C A C A C C G A A A C T T A A T G C A C
G A A C T G G A A A T C G T T G A G G A C C A C C C A G T C T T T G G G T A C C A C G A G G A T T C A A G C A A T T A T G A G G G C G T T C
T T G C T C T C G C A A A T G A T G G G T C T C A C A T G T A T G C C T T C C G A G T G C A C C C A A C C A A C C A A A T C T T C A T G G
A A T G G G A T A T G G C T C C C A G G A T C T C C G C C T C G C C T G A T
```

Figure 28 The consensus sequences from colonies sequencing using *S-HrDEF* degenerated primer

```
>Consensus_DEF-2FwRv
AGAAGATAGAGAATCCAACAAGCAGGCAAGTAACGTATTCAAAGAGGCGACTTGGGATCAT
GAAGAAGGCCGAGGAACTCACAGTGCTCTGCGACGCTCAACTCTCACTCATCATCTTCTCC
GGCTCCGGCAAGTTAGCTGATTTCTGCAGCCCTTCCACAGAGTAAACTCTGTTCTCATCCC
CTTTTCTTTTTCTGCTCAAACACCAATTACTTCATGCAATCTTTTGGCAGCGTTAAAGATA
TATTTGAGAGGTATCAAAAAGTTACCGGAATTGATATATGGGATGCGCAACATCAGAGAAT
GCAGAACACTCTGAAGAATCTCAGGGAGACTAATGGTAATCTTCAGAAGGAGATAAGGTGG
GTTTGAATTTGGGGGATTGGGCTTGATTGTCTGATGAGGTTAAGAGAAGGTGTTGTGGTTa
TGCAGGCAGAGGAAGGGGGAGAATCTGGAAGGGTTGAGCTTTAAAGAGCTGCGCGTCTTG
AGCAAAAATTGGAGGAGTCCATGAAGATTGTTCCGCAGAGAAAGTATCATGTGATCGCTAC
GCAAACAGATACTTACAGGAAAAAGCTCAGAAGCAGCAGACAAATATACACTGCCCTAACG
CATGAACTGAAGCTCGAAAAAGAGAGTCAACTGTGCAGTTTGGTCCGAGAAGATCTTAGCG
GCATCTACAGCAGCTTGAATCTCAATGGCAAATCAGCAGCACCAGAGTGG
```

Figure 29 The consensus sequences from colonies sequencing using DEF2/FW and DEF2/RV primer

```
>Consensus_OrchidAGL6-1
CTCAAGAGGATTGAGAACAAGATCAATCGCCAGGTGACCTTCTCCAAGCGCAGGAATGGCCTCCTCAAAAAGGC
TTATGAGCTTTCTGTTCTGTGATGCCGAGGTGCCCCATCATCTTCTCAAGCCGAGGCAAGCTCTATGAAT
TCGGCAGTGCTGGCACTTGCAAAACACTGGAACGATATCAACGTACCTGCTACAGTTCTCAAGCTGCCAATCCC
GTAGATCGTGAAACACAGAGCTGGTATCAAGAAGTATCCAAATGAAGGCAAAGTTCGATTTCATTACAACGCTC
CCACAGGAATTTACTTGGAGAGGATCTTGGACCTTGAACGTGAAGGAATTACAGCAGTTGGAGCGGCAACTTG
AATCTGCTTTATCGCAGGCCAGGCAAAGAAAGACACAAATAATGCTGGATCAAATGGAGGAGCTACGTAAAAAG
GAACGTCAACTTGGAGAGATCAATAAGCAGCTAAAAATGAAGCTTGGAGGCTGGTGGTGGCTCTCTTAGGCTTAT
CCAAGGCTCATGGGATTCTGATGCGGCGGTGGTTGAAGGCAATGCGTTCCAAATGCACCCTTACCACTCGAATT
CATTGGAAATGCGAGCCAACCTTACATATAGGGTATCACCAGTTTGGTTCCTCCAGAACTGTAATTCCAGAAC
CCTGGTGTAGAGAATAATAATTTTATGCTTGGATGGATGC
```

```
>Consensus_AGL6-2
AGGCAAAAGAAGGACGCAAAATATGTTGGACCAGATGGAAGAGCTAAAGAAAAAGGAACGCCACCTCGGTGATAT
TAACAAGCAGCTTAAACATAAGCTTGGGGCAGATGGTGGATCGATGAGAGCTCTCCAAAGTTCTGGCGGCCTG
CTTCTGGGGCCAATATTGATACTTTTCGTAATCATTCAAGTAACATGGACACAGAAC
```

Figure 30 The consensus sequences from colony sequencing using OrchidAGL6-1 and AGL6-2 primer.

The 3'-Rapid amplification of cDNA ends of *DEF*-like and *AGL6*-like genes from *R. retusa*

1. Physical map of primers

The synthesized cDNA from total RNA of *R. retusa* flower buds was used as a template to amplify the fragment of *DEF*-like and *AGL6*-like genes. The primer maps were shown in figure 31 and figure 32 for *DEF*-like and *AGL6*-like genes respectively. For *DEF*-like gene, two pairs of primer including S-HrDEF/Fw and S-HrDEF/Rv primer and DEF2/Fw and DEF2/Rv primer were used to isolate middle sequence of *DEF*-like clade 1, 2, 3 and 4 genes. Then, qRrDEF-C1, qRrDEF-C2, qRrDEF-C3, and qRrDEF-C4 forward primer were designed to amplified 3' end of mRNA with P018HA primer (Figure 31). For *AGL6*-like gene, two pairs of primer including OAGL6-1/Fw and OAGL6-1/Rv primer and OAGL6-2/Fw and OAGL6-2/Rv primer were amplify middle sequence of *AGL6*-like and 3' RACE primers, qRrAGL6-1/Fw and qRrAGL6-2/Fw were further designed for 3' end of *AGL6*-like clade 1 and 2 gene respectively (Figure 32).

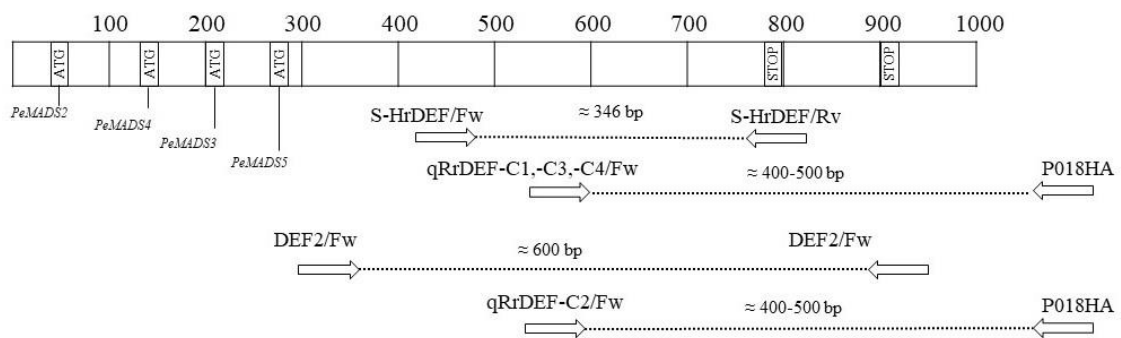


Figure 31 The physical map of primer using in *DEF*-like genes isolation in *R. retusa*

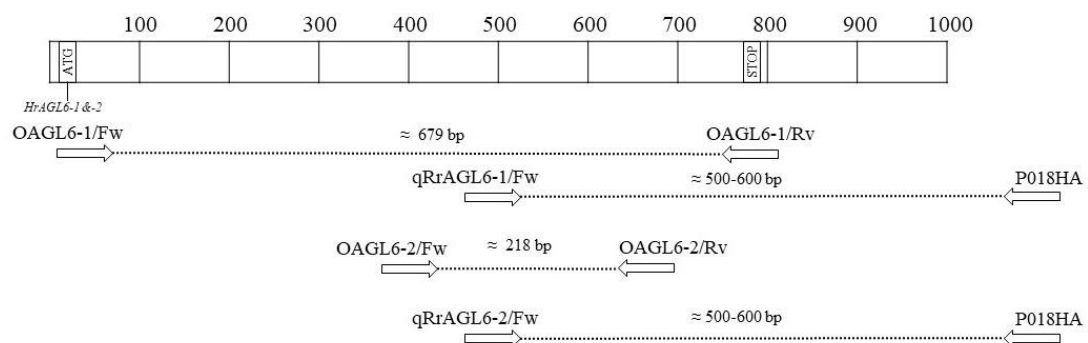


Figure 32 The physical map of primer using in *AGL6*-like genes isolation in *R. retusa*

2. PCR amplification

To isolate 3' ends of *DEF*-like and *AGL6*-like MADS-box genes from *R. retusa*, the forward primers were design from the first partial sequences. The whole floral bud cDNA which are synthesized by P019HA adapter primer was amplified using with P018HA primers. The expected amplification size for *DEF*-like and *AGL6*-like were approximately 500 and 600 respectively (Figure 31 and 32). The amplified products were validated using 1.5% agarose gel electrophoresis stained with EtBr (Figure 33 and 34). The results showed that all of 3'RACE products were able to amplify except to *AGL6-like* clade 1.

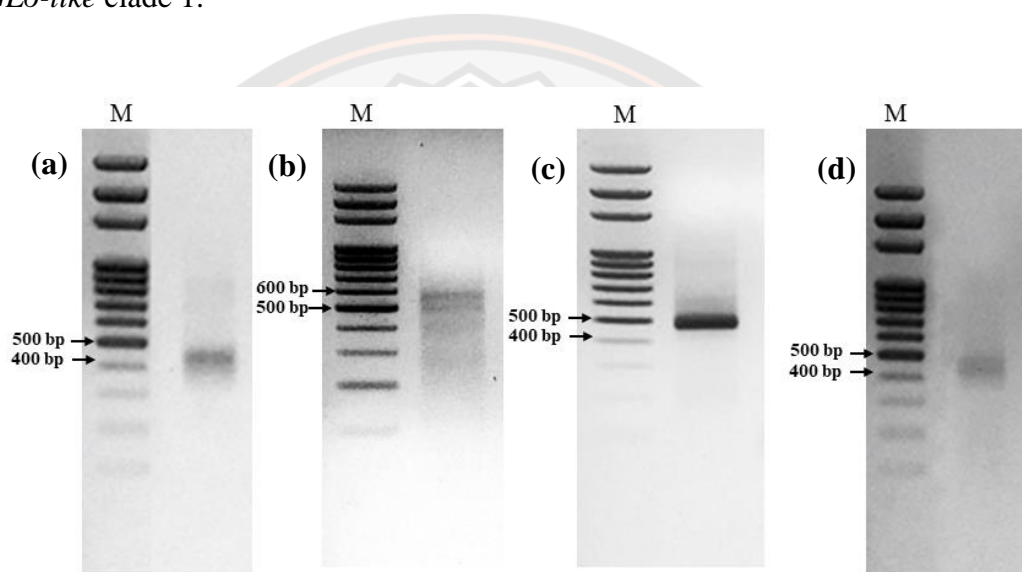


Figure 33 The 3'RACE products of *DEF*-like from *R. retusa* were validated by agarose gel electrophoresis stained with EtBr; (a) *DEF* clade 1, (b) *DEF* clade 2 (c) *DEF* clade 3 and (d) *DEF* clade 4

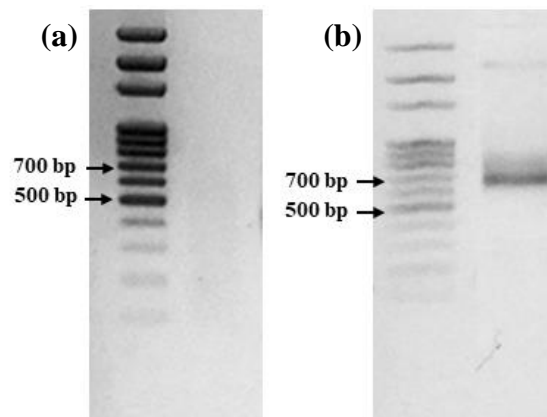


Figure 34 The 3'RACE products of AGL6-like from *R. retusa* were validated by agarose gel electrophoresis stained with EtBr; (a) AGL6-like clade 1, (b) AGL6-like clade 2

3. Sequence analysis

The raw sequence data (abi) were received from Macrogen Co., LTD (Korea). The DNA sequences from each colony were edited and aligned using GeneStudio™ software (GeneStudio, Inc, USA) and the ClustalW (EMBL-EBI, UK) respectively. The consensus sequence alignments from each gene led the partial coding sequences of DEF-like and AGL6-like genes in *R. retusa* (Figure 35 and 36) and were confirmed with BLAST® results which compared to NCBI database (APPENDIX B Table 11-16). The length of *RrDEF1*, *RrDEF2*, *RrDEF3*, *RrDEF4*, *RrAGL6-1* and *RrAGL6-2* are length 466, 829, 429, 538, 670 and 630 respectively.

>RrDEF1

CTTTGGAAGAGTCTCTGAGAATTGTTAGGCATAGAAAAGTATCATGTGATCGCCACACAAACTGACACTTA
CAAGAAAAAGCTTAAAAGCACAAAGGAAACTTACCGCGCTCTAATACATGAACTGGATATGAAAGAGGAG
AATCCGAAC TACGGTTTTAATGTGAAAGCCATAGTAGAATTTATGAAAATTC AATTCCAATGGTGAATG
AGTGCCTCAGATGTTTTCTTTAGGGTTGTTTCATCCAAATCAGCCCAATCTGCTTGGCTTAGGTTATGA
ATCACAGGATCTTAGCCTCGCCTAATAATCAGTAATATTATAAAAAGTTTGGTTTTATTGTATTTTTATTA
TATGTTTTGAACTTTAGAGTTATGAGATGGGTGATCTATTCAGAGAAAAC TGTCTTTATTTAGATTTTT
CAGTGTTCCTCTTCAAGTTCGGTGAAATTTGTTTGATGTTTTTTTCG

>RrDEF2

AGAAGATAGAGAATCCAACAAGCAGGCAAGTAACGTATTTCAAAGAGGCGACTTGGGATCATGAAGAAGGC
CGAGGAACTCACAGTGTCTGCGACGCTCAACTCTCACTCATCATCTTCTCCGGCTCCGGCAAGTTAGCT
GATTTCTGCAGCCCTTCCACAGACGTTAAAGATATATTTGAGAGGTATCAAAAAGTTACCGGAATTGATA
TATGGGATGCGCAACATCAGAGAAATGCAGAACACTCTGAAGAATCTCAGGGAGACTAATGGTAATCTTCA
GAAGGAGATAAGGCAGAGGAAGGGGAGAATCTGGAAGGGTTGAGCTTTAAAGAGCTGCGCGGTCTTGAG
CAAAAATTGGAGGAGTCCATGAAGATGTTTCGGCAGAGAAAGTATCATGTGATCGCTACGCAACAGATA
CTTACAGGAAAAAGCTCAGAAGCAGCAGACAAAATATACACTGCCCTAACGCATGAAGCTCGAAAA
AGAGAGTCAACTGTGCAGTTTGGTCGCAGAAGATCTTAGCGGCATCTACAGCAGCTTGAATCTCAATGGC
AAATCAGCAGCACCAGAGTGGGCCAAATGTGCAGAAAGTAGTTCATGAGTGCATGAGTTGGCTTTGAT
TGACCTGCAATTTCTATTACTTGTGTTATAATGTGGATTTGGGTTTCATGGCTTAACATCATAGCCTTGTT
TAAACTATTTTTTTGTGCAATGGTTAAGTTTTGGTCTTAATAGTATTGCAATAAGTTTTTGAGATATATA
AGGGGCAGTTGTAATCACACGTGCAATGTTCAAATATTTTGAATATTTAACAAAACGGC

>RrDEF3

ATATGGACGAGGCCCTAAAGCTTGTAAGGAATCGAAAGTATCACGTCATCAGCACGCAGACAGATACCTT
CAAAAAAAGTTGAAAAACTCTCAAGAAACCCAAAGGAACTTACTACGGGAGCTGGAAACTGAGCATGCC
GTCTACTATGTGGATGATGATCCAAACAACATGATGGCGCGCTTGCACCTGGAAATGGGGCTTCCTACT
TGTATTCAATTCGTACCCAACCAAGCCAGCCAAACCTTCAAGGAGTGGGATATGTCCCTCAGGATCTTCG
TCTCGCCTAATCTTTTATTATCTGCATGCCAAGTCTTAATTTATTTATGTATGGAGCTTCACTTTGACT
TGTTCTGATGTTCTTACGCTTACAAGTAGGGTCTAAGCACGGCAATTTAGATACTGGTATTTGTGCTCTA
CTTGATTTT

>RrDEF4

ACATCGACGAGGCATTGAAGCTAGTACGAAATAGAAAATATCATGTAATCAGTACTCAAACGGACACCTA
CAAGAAGAAGCTGAAGAACTCTCAAGAAACACACCCGAACTTAATGCACGAACTGGAAATCGTTGAGGAC
CACCCAGTCTTTGGGTACCACGAGGATTC AAGCAATTATGAGGGCGTTCTTGCTCTCGAAATGATGGGT
CTCACATGTATGCCTTCCGAGTGC AACCAACCAACCAATCTTCATGGAATGGGATATGGCTCCAGGA
TCTCCGCCTCGCCTGATATAATCGCGTAAGTACTACAATCACATATGCTATTTTCGTTTTATGGTTTCGCA
AATTATGCGCTTTGTAGCTGATATCTAATGTAGAATAACTACTGCAACTTGTTCCTTACTATGTG
TGATTCTGTGGTTATCTGGACTTAAAAGTATTTGTTGCATTGTGTTTTTTGCATAATAATACCACTCATC
CCTATGTCAAATCTGTTATTTATTTATTTATACAACTCTTCCGTC

Figure 35 The partial sequences of four *DEF*-like genes in *R. retusa*

```

>RrAGL6-1
AAGATCAATCGCCAGGTGACCTTCTCCAAGCGCAGGAATGGCCTCCTCAAAAAGGCTTATGAGCTTTCTG
TTCTCTGTGATGCCGAGGTGCGCCTCATCATCTTCTCAAGCCGAGGCAAGCTCTATGAATTCGGCAGTGC
TGGCACTTGCAAACACTGGAACGATATCAACGTACCTGCTACAGTTCTCAAGCTGCCAATCCCCTAGAT
CGTGAAACACAGAGCTGGTATCAAGAAGTATCCAAATTGAAGGCAAAGTTCGATTCATTACAACGCTCCC
ACAGGAATTTACTTGGAGAGGATCTTGGACCCCTTGAACGTGAAGGAATTACAGCAGTTGGAGCGGCAACT
TGAATCTGCTTTATCGCAGGCCAGGCAAAGAAAGACACAAATAATGCTGGATCAAATGGAGGAGCTACGT
AAAAAGGAACGTCAACTTGGAGAGATCAATAAGCAGCTAAAAATGAAGCTTGGAGCTGGTGGTGGCTCTC
TTAGGCTTATCCAAGGCTCATGGGATTCTGATGGCGGGTGGTTGAAGGCAATGCGTTCCAAATGCACCC
TTACCACTCGAATTCAATTGGAATGCGAGCCAACCTTACATATAGGGTATCACCAGTTTGTTCCTCCAGAA
ACTGTAATTCCCAGAACCCCTGGTGTAGAGAATAATAATT

>RrAGL6-2
TGTGGACCAGATGGAAGAGCTAAAGAAAAAGGAACGCCACCTCGGTGATATTAACAAGCAGCTTAAACA
TAAGCTTGGGGCAGATGGTGGATCGATGAGAGCTCTCCAAAGTTCCTGGCGCCTGCTTCTGGGGCCAAT
ATPGATACTTTTCGTAATCATCAAGTAACATGGACACCGAACCCACTCTTCAAATGGGAGGTACAATC
AGTATGTTCCCTTCTGAAGCAACAATTCTAGAAACGGTGGAGCTGGAACAGTTTCATGCCTGGATGGGG
CGCAGTTTGAGAGAGTTTGACTGAAAACCTTCCATAAATGTAATTTTAGGTGTTCCGCTTCTGTTTAAATAA
CGTACCTGTCTGTTGGAGGCTTTTTTTTCTCAGCTCTTACACTATGACTGTTTTGGGTATCAAACATATG
TTGATATCTGGTCACTAAACTTGAATATGGTGGTTGAGGAACTAAATGGCAATTTTAGTTTAGCAGGGCT
TCACAGGCAATGCTGTTGGTTGGGTTGCTCTTTAGATTAGTGACGACATGCACATATATTTTCTCACATT
AGTTTGCTTCCGTAAAAATCAAGTTTTTCATGTTTAAATAAAAA

```

Figure 36 The partial sequences of two *AGL6*-like genes in *R. retusa*

4. The nucleotide sequence deposited to gene bank database

The partial sequences of four *DEF*-like genes and two *AGL6*-like genes in *R. retusa* were deposited into the GeneBank database. The accession numbers were obtained and shown in Table 4.

Table 4 The accession numbers of four *DEF*-like genes and two *AGL6*-like genes in *R. retusa*

Genes	Accession number
<i>RrDEF1</i>	MW033595
<i>RrDEF2</i>	MW033596
<i>RrDEF3</i>	MW033597
<i>RrDEF4</i>	MW033598
<i>RrAGL6-1</i>	MW033599
<i>RrAGL6-2</i>	MW033600

Partial coding sequence of *DEF*-like and *AGL6*-like genes in *R. retusa*

Two well-conserved sequences of *DEF*-like and *AGL6*-like proteins were found in the C-terminal region. The C-terminal region of B-functional genes showed most *DEF*-like genes as two consensus sequences of a PI-derived motif (FXFRLOPSQPNLH) and a paleoAP3 motif (YGXHDLRLA) (Kramer et al., 1998). Both of these motifs were found in *RrDEF1*, *RrDEF3*, and *RrDEF4*, but were absent in the C-terminal region of *RrDEF2* (Figure 37). This evidence strongly indicated that all four *RrDEF* genes were members of the *DEF*-like family and the presence of paleoAP3 motif supported the duplication events of *DEF* genes family that paleoAP3 limited in low eudicots, magnolid dicots and monocots. The *AGL6*-I motif (DCEPTLQIGY) and *AGL6*-II motif (ENNFMLGWVL) (Ohmori et al., 2009) were also located in the middle part and at the end of the C-terminal regions of the *RrAGL6-1* and *RrAGL6-2* genes (Figure 38).

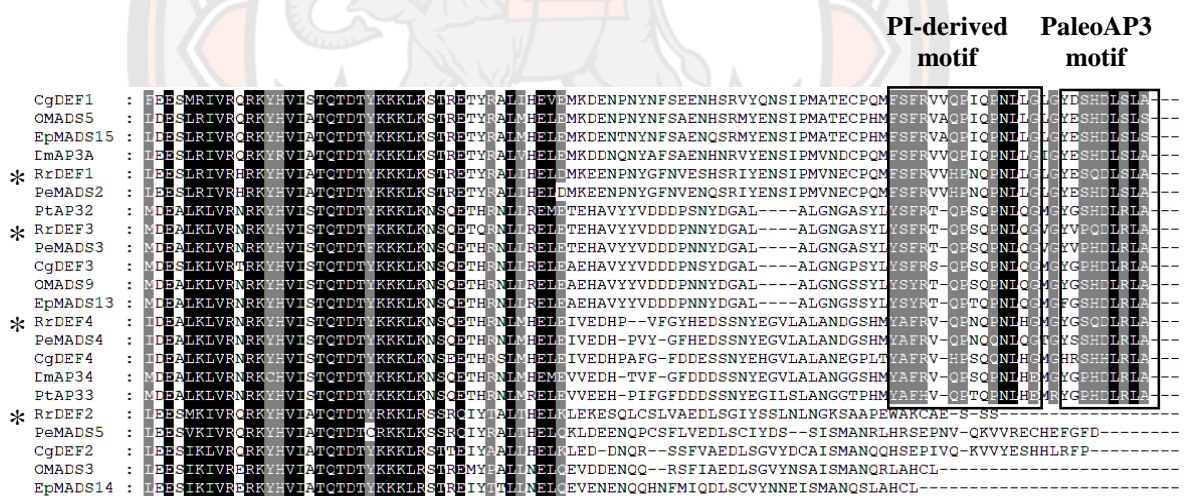


Figure 37 Alignment of the consensus amino acid sequences of the C-terminal

region for *RrDEF1*, *RrDEF2*, *RrDEF3*, and *RrDEF4* compared to other orchid species *DEF*-like MADS-box proteins. The two highly conserved PI-derived motif and PaleoAP3 motif are indicated above the columns. Except for *RrDEF2*, all *RrDEF* genes were detected as two conserved motifs in the C-terminal region. Identical and similar amino acids are shaded black and gray, respectively. The multiple sequence alignment was generated by ClustalW.

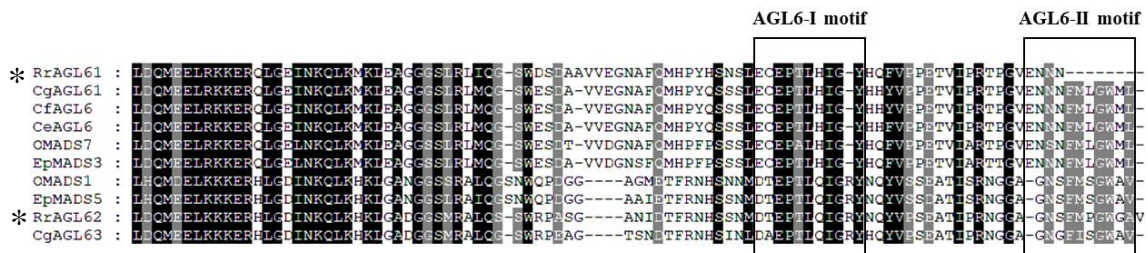


Figure 38 Alignment of the consensus amino acid sequences of the C-terminal region for *RrAGL6-1* and *RrAGL6-2* compared to other orchid species *AGL6*-like MADS-box proteins. The two conserved *AGL6-I* and *AGL6-II* motifs at the C-terminal region were detected and indicated above the columns. Identical and similar amino acids are shaded black and gray, respectively. The multiple sequence alignment was generated by ClustalW.

Phylogenetic analysis

To analyze the phylogenetic relationships of orchid B- and E-class genes, a phylogenetic tree of DEF and AGL6 lineages from Orchidaceae was reconstructed using the most published orchid B-class genes (Figure 39) and E-class genes (Figure 40). The maximum-likelihood analysis indicated that the orchid B-class homologs consisted of two lineages, DEF and GLO. Orchid DEF homologs were subdivided into two groups, AP3A and AP3B, and each consisted of two monophyletic clades. Within clade 1, *RrDEF1* was in the same subclade as *PeMADS2* and *DmAP3A* from *Phalaenopsis equestris* and *Dendrobium moniliforme*. *RrDEF2*, *PeMADS5*, *CgDEF2* (*Cymbidium goeringii*), *GgDEF2* (*Gongora galeata*), *OMADS3* (*Oncidium Gower Ramsey*), and *EpMADS14* (*Erycina pusilla*) were grouped in the same subclade 2, clade 2. Clade 3 contained *RrDEF3* with other clade 3 DEF-like genes from various orchid genera, such as *Phalaenopsis*, *Cymbidium*, *Oncidium*, and *Erycina*. Clade 4 comprised *RrDEF4*, *PeMADS4*, *CgDEF4*, *DmAP3-4*, and *PtAP3-3* (*Phaius tancarvilleae*). All four *RrDEF* genes were closely related to four DEF-like genes in *P. equestris*. The phylogenetic tree showed that the four *RrDEF* genes were separated into two monophyletic clades corresponding to previous data in *P. equestris*, *O. 'Gower*

Ramsey', *E. pusilla*, *D. moniliforme*, *C. goeringii*, and *Habenaria radiata* (Hsu & Yang, 2002; Kim et al., 2007; Mitoma et al., 2019; Pan et al., 2011; Tsai et al., 2004; Xiang et al., 2018). These results showed that two major duplication events from ancestral genes occurred, confirming the previous B-class gene evolution data. The first duplication from the ancestral genes produced both paleoAP3 and PI lineages of orchids, while the second duplication only occurred in the lineage of paleoAP3 (Kramer et al., 1998; Mondragón-Palomino & Theißen, 2008; Mondragón-Palomino & Theissen, 2009).

Phylogenetic analysis of the *AGL6*-like genes showed that monocot *AGL6*, eudicot *AGL6*, and gymnosperm *AGL6* were separated into three clades. Two genes, *RrAGL6-1* and *RrAGL6-2*, were assigned to two distinct clades in monocot *AGL6*, orchid *AGL6-1* and orchid *AGL6-2*. *RrAGL6-1* was closely related to *CeAGL6* (*C. ensifolium*), *CfAGL6* (*C. faberi*), and *CgAGL6-1* (*C. goeringii*) with high bootstrap support. *RrAGL6-2* was closely related to *PaAGL6* (*Phalaenopsis aphrodite*), *DAGL6* (*Dendrobium* hybrid cultivar), *OMADS1* (*O.* 'Gower Ramsey'), *EpMADS5* (*E. pusilla*), and *CgAGL6-3* (*C. goeringii*). These results were similar to previous reports of *AGL6*-like genes from *O.* 'Gower Ramsey', *P. equestris*, *E. pusilla*, *C. goeringii*, and *H. radiata* (Dreni & Zhang, 2016; Hsu et al., 2015; Lin et al., 2016; Mitoma et al., 2019; Xiang et al., 2018). Phylogenetic trees with two distinct subdivided clades of *DEF*- and *AGL6*-like genes from *R. retusa* suggested that the common ancestor of Orchidaceae possibly consisted of two *DEF*-like and *AGL6*-like genes.

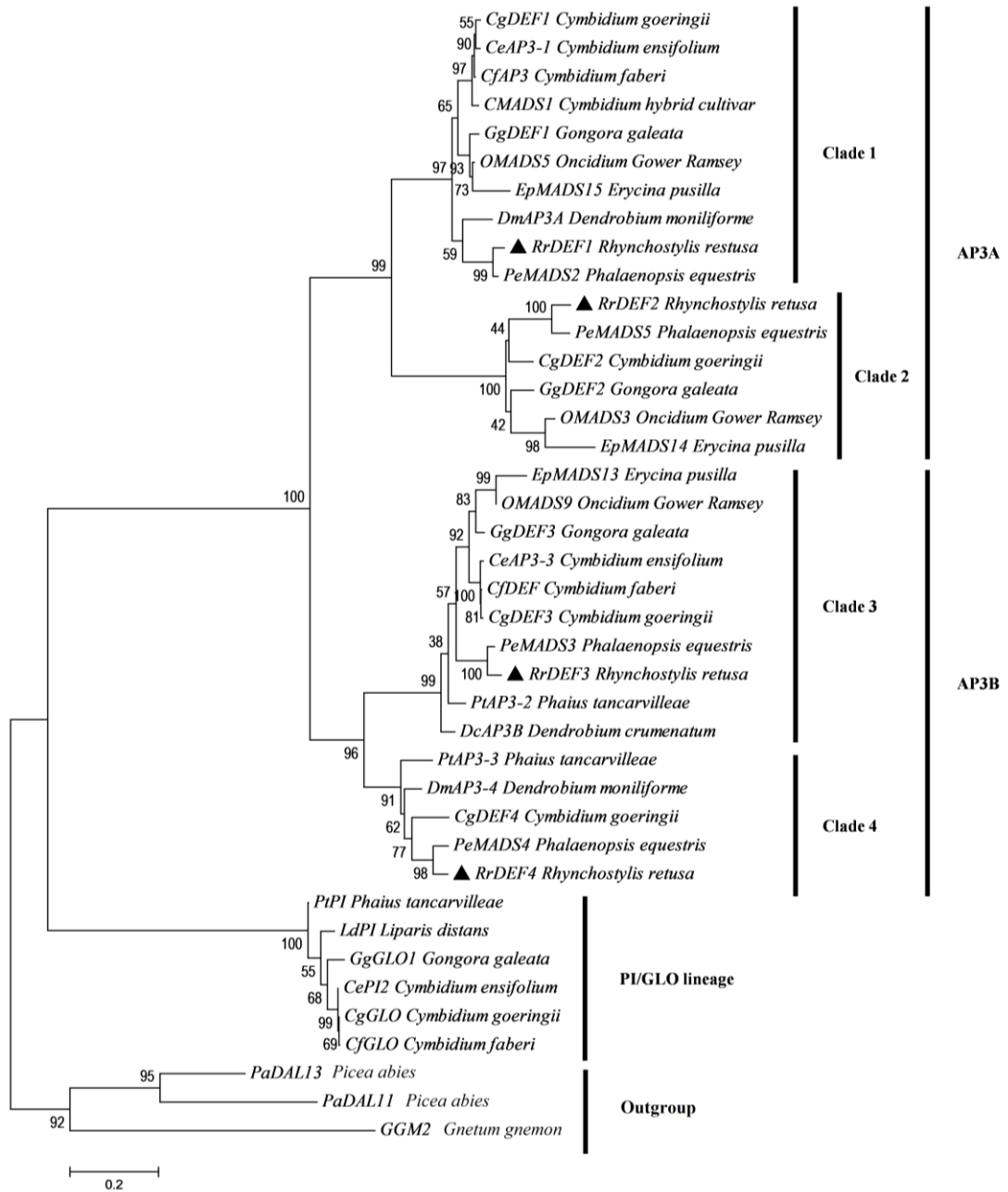


Figure 39 Phylogenetic analysis of DEF-like gene nucleotide sequences constructed by a maximum-likelihood tree with 1000 bootstrap replicates in MEGA version 7.0.26.

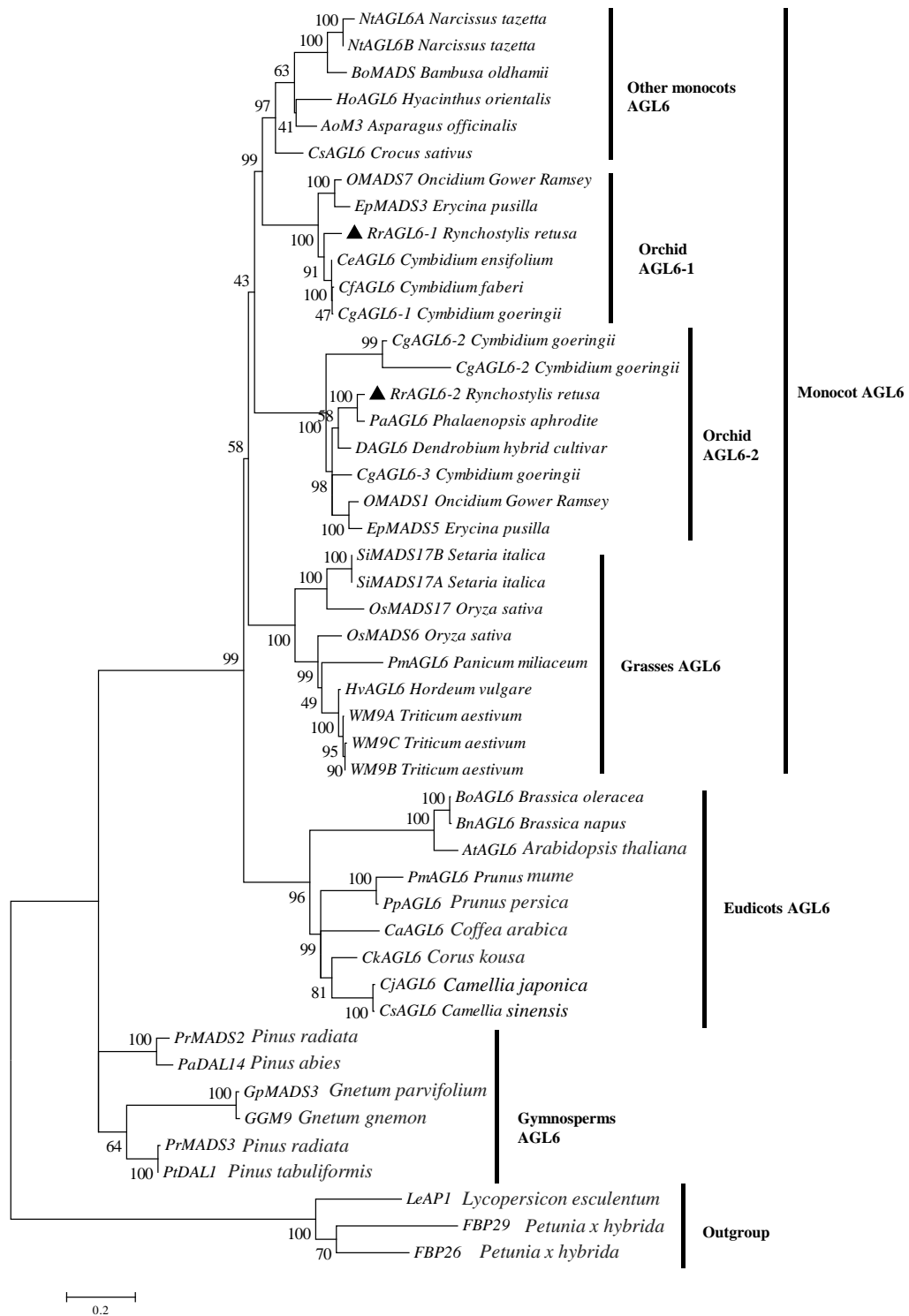


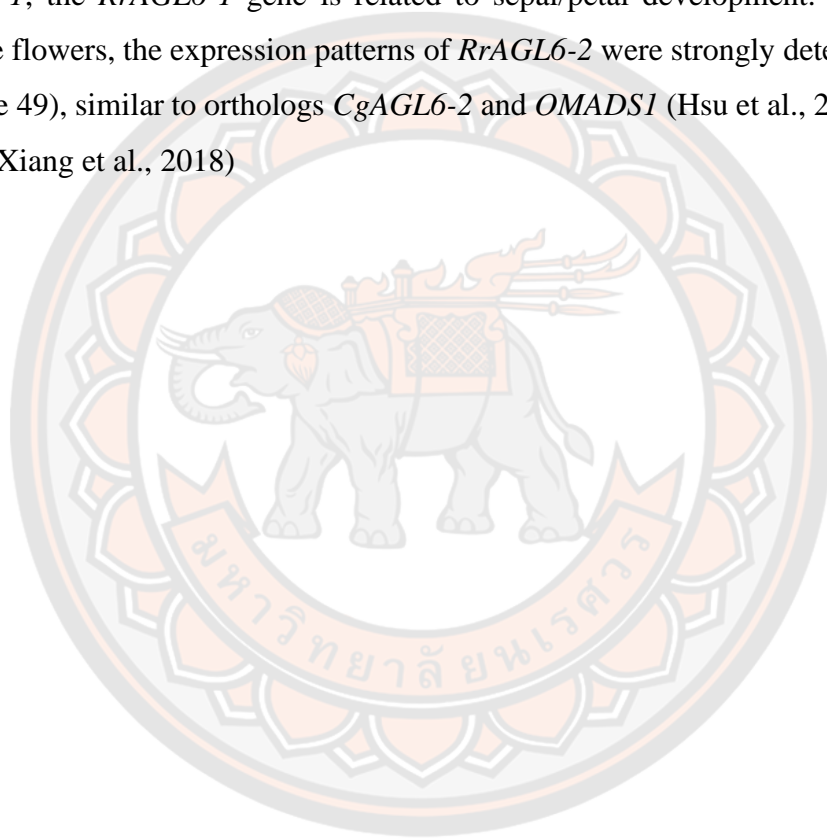
Figure 40 Phylogenetic analysis of *AGL6*-like gene nucleotide sequences constructed by a maximum-likelihood tree with 1000 bootstrap replicates in MEGA version 7.0.26.

Expression analysis

The expression of *DEF*-like and *AGL6*-like genes was further studied in various floral development stages and dissected flowers using reverse transcription PCR (RT-PCR) (Figure 41 and 42) and quantitative real-time polymerase chain reaction (qRT-PCR) (Figure 43-49). The RT-PCR and qRT-PCR relatively have the same direction. All genes were detectable in all developmental stages of the flower buds and mature flowers but not in leaves, apart from the *RrDEF2* gene that was slightly expressed in vegetative leaves (Figure 43). When the floral organs from floral bud stage 3 and mature flower were examined in the B-class group, the expression of *RrDEF1* was high in the sepals and petals but low in the lips and columns of both stages (Figure 44). This result supported previous studies that investigated sepal/petal development (Chang et al., 2010; Hsu et al., 2015; Lin et al., 2016; Mitoma et al., 2019; Mondragón-Palomino & Theissen, 2011; Pan et al., 2011; Xiang et al., 2018). In stage 3 and mature flowers, *RrDEF2* was more highly expressed in the petals than in sepals, lips, and columns. The *RrDEF2* expression was notably expressed lower than other genes in every developmental stages (Figure 45). *RrDEF2* was detected in vegetative and all floral organs and highly expressed in sepals/petals, similar to *OMADS3* from *O. 'Gower Ramsey'* (Chang et al., 2010) and *OMADS3* was presented the low copy number in *O. Gower Ramsey* genome (Hsu & Yang, 2002). *RrDEF3* was expressed in the petals, lips, and columns but not in the sepals (Figure 46). This expression pattern of *RrDEF3* was only evident in inner tepals similar to its orthologs *PeMADS3*, *CgDEF3*, *EpMADS13*, *OMADS9*, *HrDEF*, and *DMAP3B* (Chang et al., 2010; Kim et al., 2007; Lin et al., 2016; Pan et al., 2011; Sirisawat et al., 2010; Xiang et al., 2018), while the *RrDEF3* and *RrDEF4* genes were notably expressed in the lips. Thus, *RrDEF3* has a possible role in regulating lip formation, as reported in previous studies (Hsu et al., 2015; Mondragón-Palomino & Theissen, 2011). The expression of *RrDEF4* was detected in all floral organs. However, most orchid *DEF*-like clade 4 genes have identical expression patterns to *DEF*-like clade 3 genes, with *RrDEF4* distinctly expressed either in sepals or petals (Figure 47). Orchid *DEF*-like clade 4 genes showed various expression profiles supported by *CgDEF4*, *PaphAP3-2*, *OAP3-4*, and *BnAP3-1* from *C. goeringii*, *Paphiopedilum Macabre*, *O. 'Gower Ramsey'*, and *Brassavola nodosa*, respectively (Hsu et al., 2015; Pan et al., 2011; Xiang et al., 2018). Our results

further indicated that *DEF*-like clade 4 genes may play different roles in the development of floral organs in various orchids.

In the E-class group, except for the high level of *RrAGL6-1* in stage 3 column, *RrAGL6-1* was highly expressed in sepals and petals but rarely in lips in both stage 3 and mature flowers (Figure 48). The expression patterns of *RrAGL6-1* were similar to its orthologs *CgAGL6-1*, *OMADS7*, and *EpMADS3* in *C. goeringii*, *O. 'Gower Ramsey'* and *E. pusilla* (Hsu et al., 2015; Lin et al., 2016; Xiang et al., 2018). Thus, similar to *RrDEF1*, the *RrAGL6-1* gene is related to sepal/petal development. In stage 3 and mature flowers, the expression patterns of *RrAGL6-2* were strongly detected in the lips (Figure 49), similar to orthologs *CgAGL6-2* and *OMADS1* (Hsu et al., 2015; Hsu et al., 2003; Xiang et al., 2018)



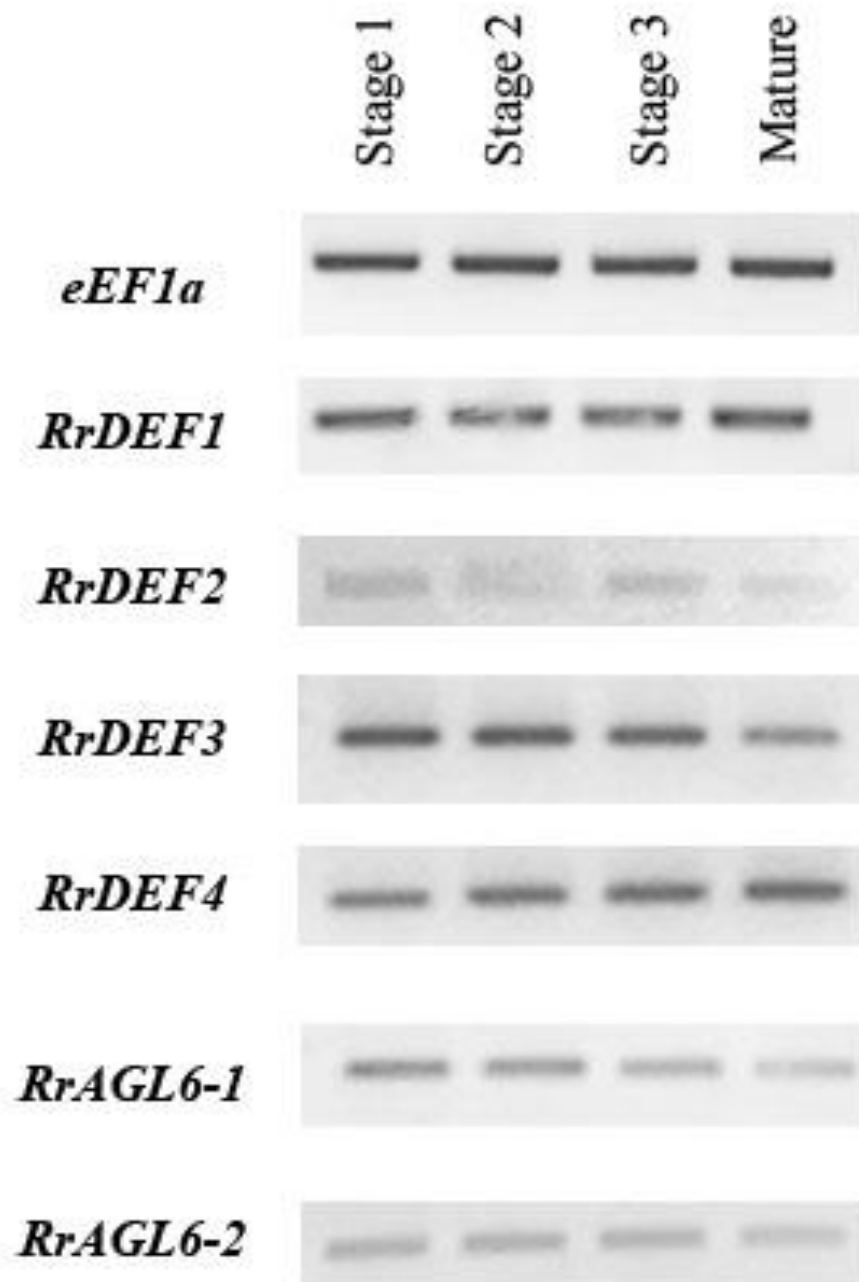


Figure 41 The expression pattern of *RrDEF*-like and *RrAGL6*-like genes in various floral development stages using RT-PCR and evaluated on 1.5% (w/v) agarose TAE gel electrophoresis stained with Ethidium bromide (EtBr)

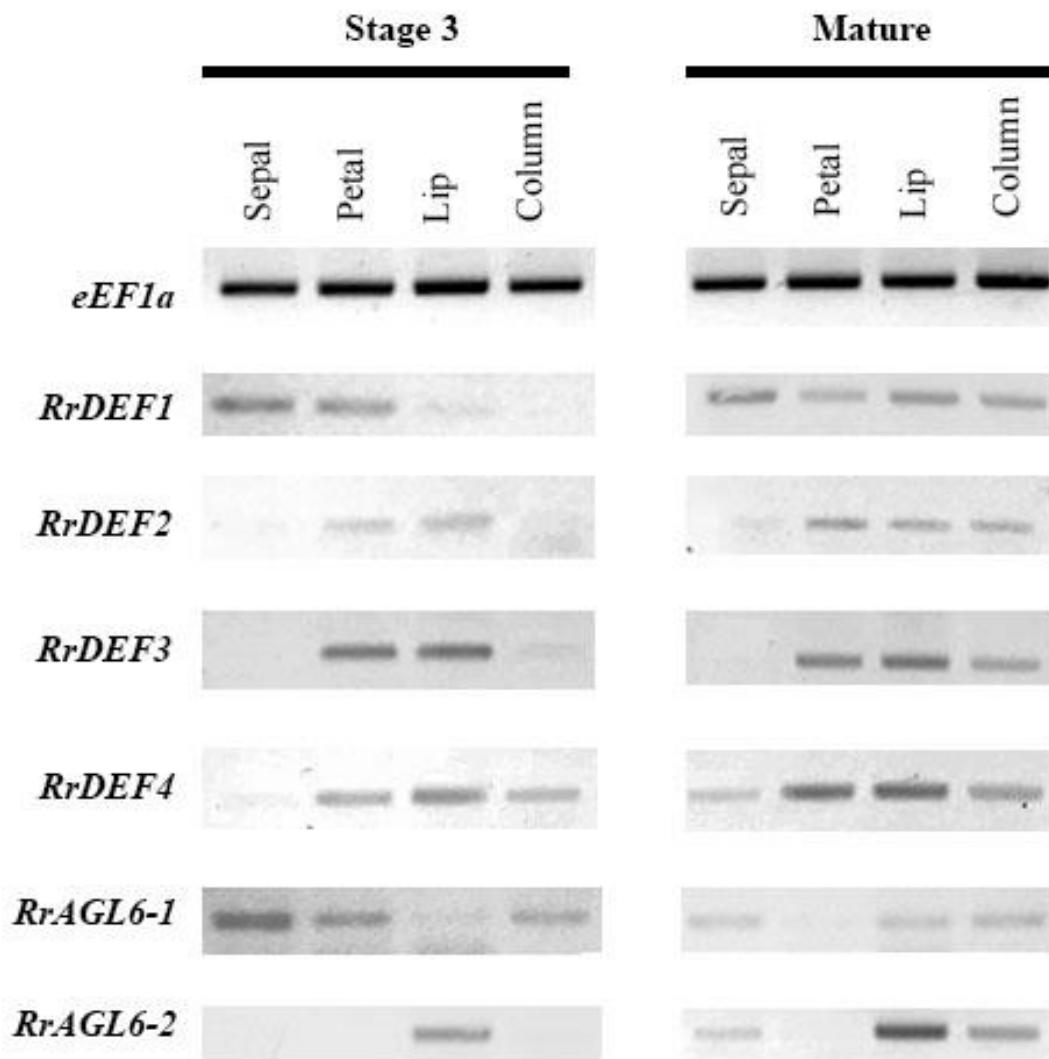


Figure 42 The expression pattern of *RrDEF*-like and *RrAGL6*-like genes in various floral development stages using RT-PCR and evaluated on 1.5% (w/v) agarose TAE gel electrophoresis stained with Ethidium bromide (EtBr)

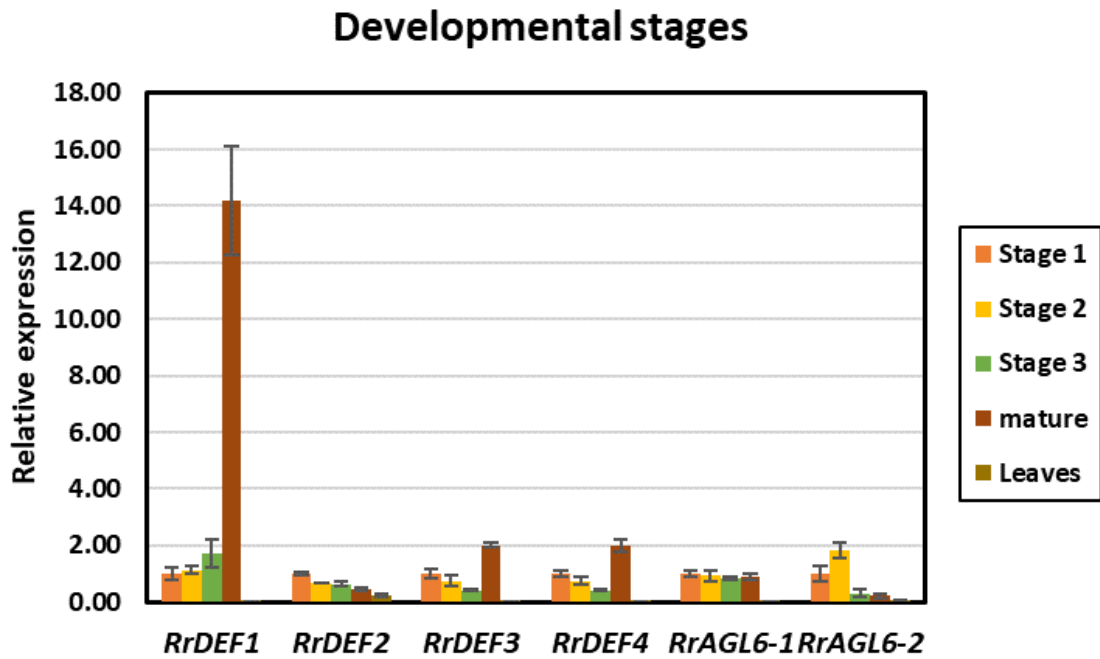


Figure 43 Expression patterns of four *RrDEF* and two *RrAGL6* genes in different developmental stages from 1–4 (stage 1: 0.5 cm; stage 2: 1.0 cm; stage 3: 1.5 cm; stage 4: mature). All experiments were performed in triplicate with data presented as means \pm SD.

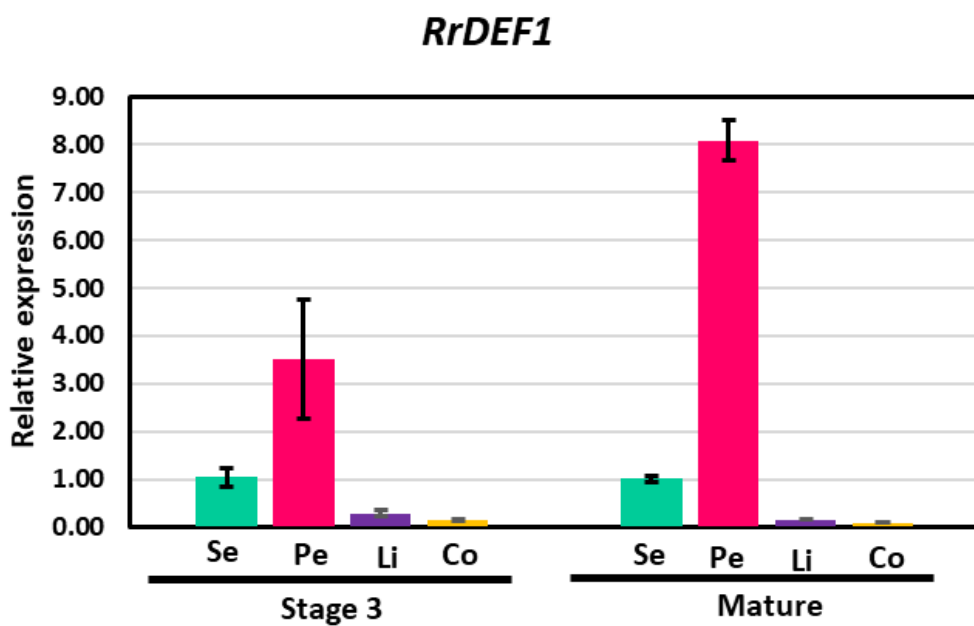


Figure 44 Expression patterns of *RrDEF1* genes in various organs of stage 3 and mature flowers (stage 4), including sepals, petals, lips, and columns. All experiments were performed in triplicate with data presented as means \pm SD

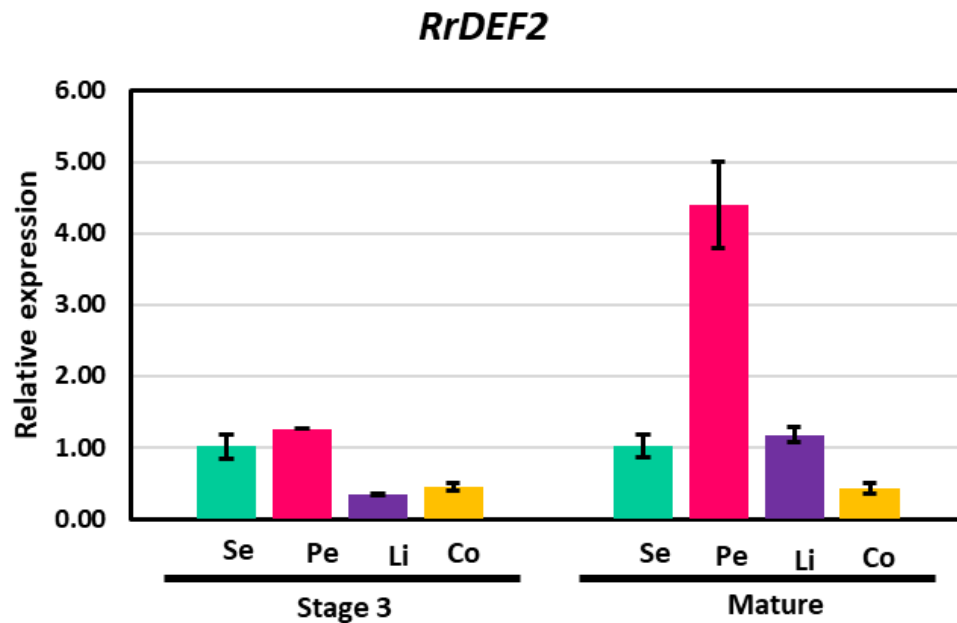


Figure 45 Expression patterns of *RrDEF2* genes in various organs of stage 3 and mature flowers (stage 4), including sepals, petals, lips, and columns. All experiments were performed in triplicate with data presented as means \pm SD

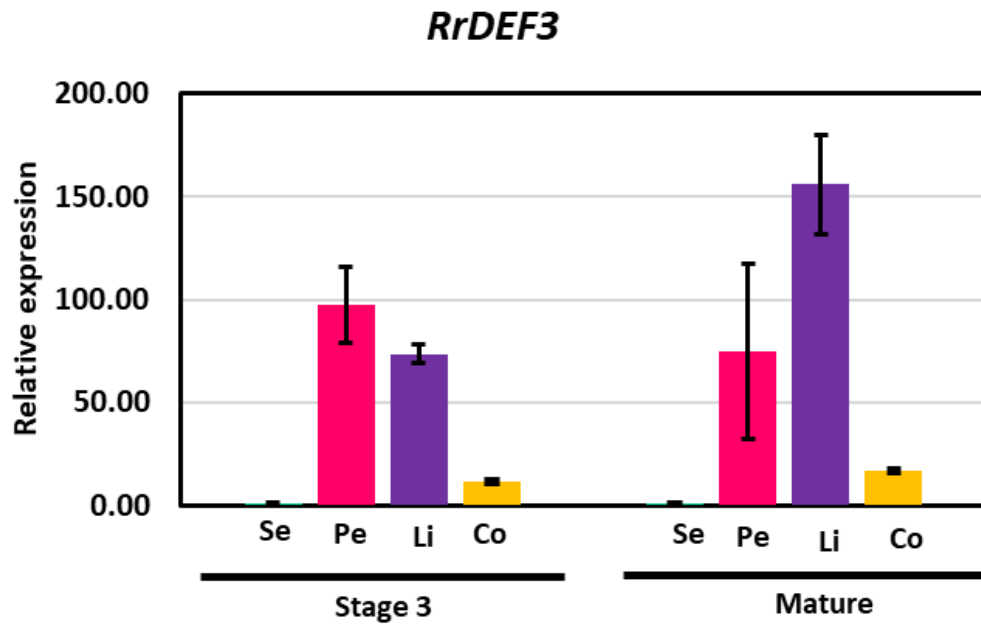


Figure 46 Expression patterns of *RrDEF3* genes in various organs of stage 3 and mature flowers (stage 4), including sepals, petals, lips, and columns. All experiments were performed in triplicate with data presented as means \pm SD

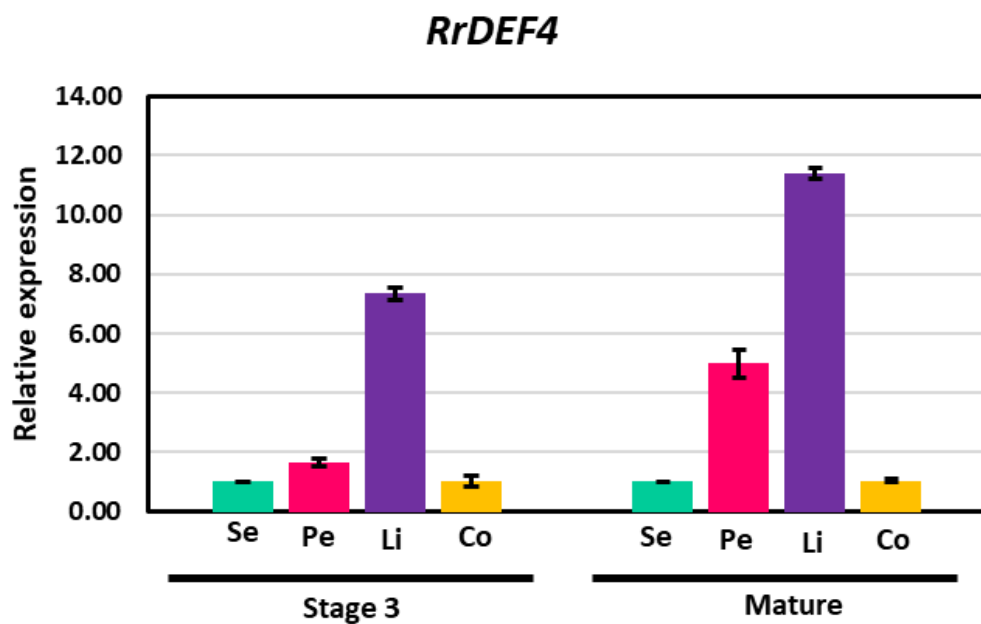


Figure 47 Expression patterns of *RrDEF4* genes in various organs of stage 3 and

mature flowers (stage 4), including sepals, petals, lips, and columns. All experiments were performed in triplicate with data presented as means \pm SD

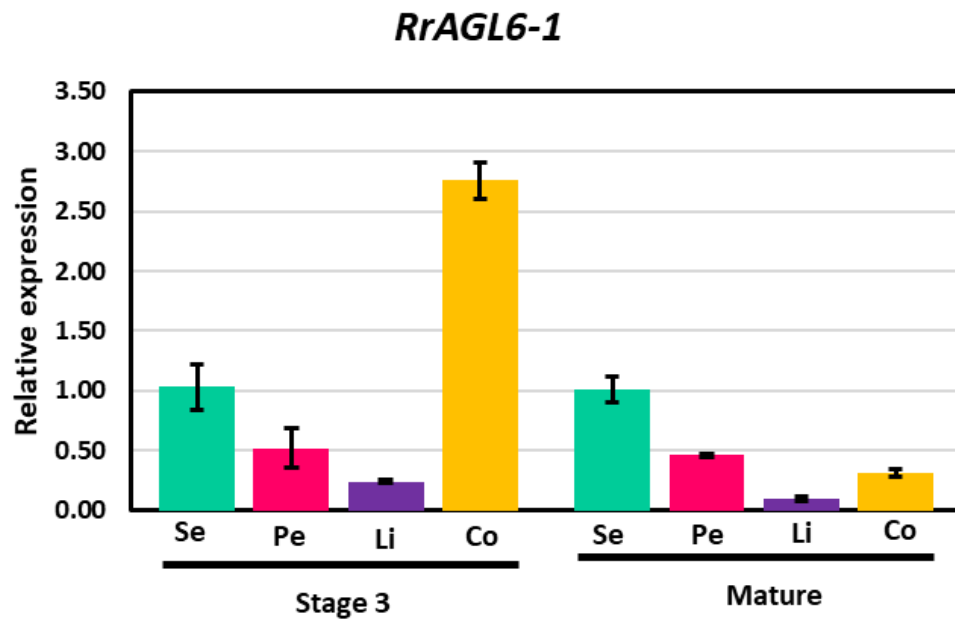


Figure 48 Expression patterns of *RrAGL6-1* genes in various organs of stage 3 and mature flowers (stage 4), including sepals, petals, lips, and columns. All experiments were performed in triplicate with data presented as means \pm SD

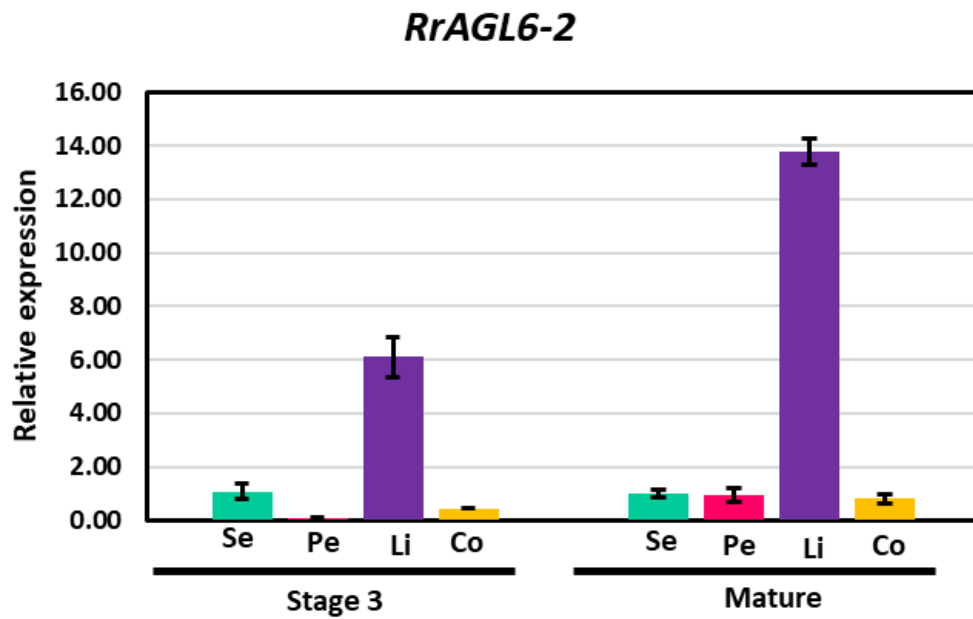


Figure 49 Expression patterns of *RrAGL6-2* genes in various organs of stage 3 and mature flowers (stage 4), including sepals, petals, lips, and columns. All experiments were performed in triplicate with data presented as means \pm SD

CHAPTER V

CONCLUSION

Four *DEF*-like genes and two *AGL6*-like genes found in *R. retusa* were investigated. Interestingly, four types of *DEF* genes were isolated from many orchid species, but rarely were all four *DEF*-like genes identified in a single orchid species (*Phalaenopsis* and *H. radiata*); *R. retusa* was one species in which they are all expressed. Expression patterns of *RrDEF1* and *RrAGL6-1* suggested that these genes may be involved in sepal/petal formation, while expression of the *RrDEF3*, *RrDEF4*, and *RrAGL6-2* genes related to lip development. These results confirmed the revised ‘orchid code’ and ‘P-code’ hypotheses that the expression levels of *DEF*-like and *AGL6*-like genes play a key role in orchid tepal speciation. The revised ‘orchid code’ hypothesized that different clades of *DEF*-like genes determine perianth development, with a high level of clades 1 and 2 and a low level of clades 3 and 4 specified for petal (inner lateral tepal). By contrast, a high level of clades 3 and 4 related to lip formation. However, the ‘P-code’ hypothesized that there was competition among the sepal/petal (SP) complex (OAP3-1/OAGL6-1/OAGL6-1/OPI) and the Lip (L) complex (OAP3-2/OAGL6-2/OAGL6-2/OPI). The SP complex specified the formation of sepals and petals, while the L complex determined lip formation. Moreover, the *RrDEF2* gene was expressed in all floral organs and in leaves. This indicated the possibility of functional diversification from other *RrDEF* paralogs during the second gene duplication in orchid evolution. However, the determination of accurate mechanisms of *RrDEF* and *RrAGL6* genes in floral developmental control requires more data from downstream experiments such as protein-protein interaction.

REFERENCES

- Aceto, S., & Gaudio, L. (2011). The MADS and the Beauty: Genes Involved in the Development of Orchid Flowers. *Curr Genomics*, 12(5), 342-356. doi:10.2174/138920211796429754
- Alvarez-Buylla, E. R., Benítez, M., Corvera-Poiré, A., Chaos Cador, A., de Folter, S., Gamboa de Buen, A., . . . Sánchez-Corrales, Y. E. (2010). Flower development. *The arabidopsis book*, 8, e0127-e0127. doi:10.1199/tab.0127
- Angenent, G. C., Franken, J., Busscher, M., Weiss, D., & van Tunen, A. J. (1994). Co-suppression of the petunia homeotic gene *fbp2* affects the identity of the generative meristem. *Plant J*, 5(1), 33-44. doi:10.1046/j.1365-313x.1994.5010033.x
- Anuttato, S., Boonruangrod, R., Kongsamai, B., & Chanprame, S. (2017). Morphological characterization of wild *Rhynchostylis gigantea* in Thailand. *J ISSAAS*, 23, 20-32.
- Bateman, R., & Rudall, P. (2006a). The good, the bad and the ugly: using spontaneous terata to distinguish the possible from the impossible in orchid floral evolution. *Aliso*, 22, 481-496.
- Bateman, R., & Rudall, P. (2006b). The Good, The Bad, The Ugly: Using Naturally Occurring Terata to Distinguish the Possible from the Impossible in Orchid Floral Evolution. *Aliso*, 22, 481-496. doi:10.5642/aliso.20062201.38
- Becker, A., & Theißen, G. (2003). The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Mol Phylogenet Evol*, 29(3), 464-489. doi:[https://doi.org/10.1016/S1055-7903\(03\)00207-0](https://doi.org/10.1016/S1055-7903(03)00207-0)
- Bemer, M., Gordon, J., Weterings, K., & Angenent, G. C. (2010). Divergence of Recently Duplicated My-Type MADS-Box Genes in Petunia. *Mol Biol Evol*, 27(2), 481-495. doi:10.1093/molbev/msp279
- Byng, J., Smets, E., van Vugt, R., Bidault, E., Davidson, C., Kenicer, G., . . . Christenhusz, M. (2018). The phylogeny of angiosperms poster: a visual summary of APG IV family relationships and floral diversity. *The Global Flora*, 1, 4-35.
- Causier, B., Schwarz-Sommer, Z., & Davies, B. (2010). Floral organ identity: 20 years of ABCs. *Seminars in Cell & Developmental Biology*, 21(1), 73-79. doi:<https://doi.org/10.1016/j.semcdb.2009.10.005>
- Chang, Y. Y., Chiu, Y. F., Wu, J. W., & Yang, C. H. (2009). Four orchid (*Oncidium Gower Ramsey*) AP1/AGL9-like MADS box genes show novel expression patterns and cause different effects on floral transition and formation in *Arabidopsis thaliana*. *Plant Cell Physiol*, 50(8), 1425-1438. doi:10.1093/pcp/pcp087
- Chang, Y. Y., Kao, N. H., Li, J. Y., Hsu, W. H., Liang, Y. L., Wu, J. W., & Yang, C. H. (2010). Characterization of the possible roles for B class MADS box genes in regulation of perianth formation in orchid. *Plant Physiol*, 152(2), 837. doi:<https://doi.org/10.1104/pp.109.147116>
- Coen, E. S., & Meyerowitz, E. M. (1991). The war of the whorls: genetic interactions controlling flower development. *Nature*, 353(6339), 31-37. doi:10.1038/353031a0

- de Folter, S., & Angenent, G. C. (2006). trans meets cis in MADS science. *Trends Plant Sci*, 11(5), 224-231. doi:10.1016/j.tplants.2006.03.008
- de Folter, S., Immink, R. G. H., Kieffer, M., Pařenicova, L., Henz, S. R., Weigel, D., . . . Angenent, G. C. (2005). Comprehensive Interaction Map of the Arabidopsis MADS Box Transcription Factors. *Plant Cell*, 17(5), 1424. doi:10.1105/tpc.105.031831
- Dreni, L., & Zhang, D. (2016). Flower development: the evolutionary history and functions of the *AGL6* subfamily MADS-box genes. *J Exp Bot*, 67(6), 1625-1638. doi:<https://doi.org/10.1093/jxb/erw046>
- Dressler, R. (1993). *Phylogeny and Classification of Orchid Family*.
- Egea-Cortines, M., Saedler, H., & Sommer, H. (1999). Ternary complex formation between the MADS-box proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in the control of floral architecture in *Antirrhinum majus*. *The EMBO Journal*, 18(19), 5370-5379. doi:<https://doi.org/10.1093/emboj/18.19.5370>
- Ferrario, S., Immink, R. G., Shchennikova, A., Busscher-Lange, J., & Angenent, G. C. (2003). The MADS box gene FBP2 is required for SEPALLATA function in petunia. *Plant Cell*, 15(4), 914-925. doi:10.1105/tpc.010280
- Forterre, P., Gribaldo, S., Gabelle, D., & Serre, M.-C. (2007). Origin and evolution of DNA topoisomerases. *Biochimie*, 89(4), 427-446. doi:<https://doi.org/10.1016/j.biochi.2006.12.009>
- Gehring, W. J. (1992). The homeobox in perspective. *Trends in Biochemical Sciences*, 17(8), 277-280. doi:[https://doi.org/10.1016/0968-0004\(92\)90434-B](https://doi.org/10.1016/0968-0004(92)90434-B)
- Gonzalez, D. H. (2016). Chapter 1 - Introduction to Transcription Factor Structure and Function. In D. H. Gonzalez (Ed.), *Plant Transcription Factors* (pp. 3-11). Boston: Academic Press.
- Gramzow, L., Ritz, M. S., & Theissen, G. (2010). On the origin of MADS-domain transcription factors. *Trends Genet*, 26(4), 149-153. doi:10.1016/j.tig.2010.01.004
- Gramzow, L., & Theissen, G. (2010). A hitchhiker's guide to the MADS world of plants. *Genome Biol*, 11(6), 214. doi:10.1186/gb-2010-11-6-214
- Hsu, H. F., Hsu, W. H., Lee, Y. I., Mao, W. T., Yang, J. Y., Li, J. Y., & Yang, C. H. (2015). Model for perianth formation in orchids. *Nat Plants*, 1(5), 15046. doi:<https://doi.org/10.1038/nplants.2015.46>
- Hsu, H. F., Huang, C. H., Chou, L. T., & Yang, C. H. (2003). Ectopic expression of an orchid (*Oncidium Gower Ramsey*) *AGL6*-like gene promotes flowering by activating flowering time genes in *Arabidopsis thaliana*. *Plant Cell Physiol*, 44(8), 783-794. doi:<https://doi.org/10.1093/pcp/pcg099>
- Hsu, H. F., & Yang, C. H. (2002). An orchid (*Oncidium Gower Ramsey*) *AP3*-like MADS gene regulates floral formation and initiation. *Plant Cell Physiol*, 43(10), 1198-1209. doi:<https://doi.org/10.1093/pcp/pcf143>
- Kanno, A. (2016). Molecular Mechanism Regulating Floral Architecture in Monocotyledonous Ornamental Plants. *The Horticulture Journal*, 85(1), 8-22. doi:10.2503/hortj.MI-IR05
- Kaufmann, K., Melzer, R., & Theißen, G. (2005). MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. *Gene*, 347(2), 183-198. doi:<https://doi.org/10.1016/j.gene.2004.12.014>

- Kim, S. Y., Yun, P. Y., Fukuda, T., Ochiai, T., Yokoyama, J., Kameya, T., & Kanno, A. (2007). Expression of a *DEFICIENS*-like gene correlates with the differentiation between sepal and petal in the orchid, *Habenaria radiata* (Orchidaceae). *Plant Sci*, 172(2), 319-326. doi:<https://doi.org/10.1016/j.plantsci.2006.09.009>
- Kramer, E. M., Dorit, R. L., & Irish, V. F. (1998). Molecular evolution of genes controlling petal and stamen development: duplication and divergence within the *APETALA3* and *PISTILLATA* MADS-box gene lineages. *Genetics*, 149(2), 765-783.
- Lin, C. S., Hsu, C. T., Liao, D. C., Chang, W. J., Chou, M. L., Huang, Y. T., . . . Shih, M. C. (2016). Transcriptome-wide analysis of the MADS-box gene family in the orchid *Erycina pusilla*. *Plant Biotechnol J*, 14(1), 284-298. doi:<https://doi.org/10.1111/pbi.12383>
- Mitoma, M., Kajino, Y., Hayashi, R., Endo, M., Kubota, S., & Kanno, A. (2019). Molecular mechanism underlying pseudopetalia in *Habenaria radiata* (Orchidaceae). *Plant J*, 99(3), 439-451. doi:<https://doi.org/10.1111/tpj.14334>
- Mondragón-Palomino, M. (2013). Perspectives on MADS-box expression during orchid flower evolution and development. *Front Plant Sci*, 4, 377. doi:10.3389/fpls.2013.00377
- Mondragón-Palomino, M., & Theissen, G. (2008). MADS about the evolution of orchid flowers. *Trends Plant Sci*, 13(2), 51-59. doi:<https://doi.org/10.1016/j.tplants.2007.11.007>
- Mondragón-Palomino, M., & Theissen, G. (2009). Why are orchid flowers so diverse? Reduction of evolutionary constraints by paralogues of class B floral homeotic genes. *Ann Bot*, 104(3), 583-594. doi:<https://doi.org/10.1093/aob/mcn258>
- Mondragón-Palomino, M., & Theissen, G. (2011). Conserved differential expression of paralogous *DEFICIENS*- and *GLOBOSA*-like MADS-box genes in the flowers of Orchidaceae: refining the 'orchid code'. *Plant J*, 66(6), 1008-1019. doi:<https://doi.org/10.1111/j.1365-3113X.2011.04560.x>
- Nam, J., Kim, J., Lee, S., An, G., Ma, H., & Nei, M. (2004). Type I MADS-box genes have experienced faster birth-and-death evolution than type II MADS-box genes in angiosperms. *Proc Natl Acad Sci U S A*, 101(7), 1910-1915. doi:10.1073/pnas.0308430100
- Ng, M., & Yanofsky, M. F. (2001). Function and evolution of the plant MADS-box gene family. *Nat Rev Genet*, 2(3), 186-195. doi:10.1038/35056041
- Ohmori, S., Kimizu, M., Sugita, M., Miyao, A., Hirochika, H., Uchida, E., . . . Yoshida, H. (2009). *MOSAIC FLORAL ORGANS1*, an *AGL6*-like MADS box gene, regulates floral organ identity and meristem fate in rice. *Plant Cell*, 21(10), 3008-3025. doi:<https://doi.org/10.1105/tpc.109.068742>
- Pan, Z. J., Cheng, C. C., Tsai, W. C., Chung, M. C., Chen, W. H., Hu, J. M., & Chen, H. H. (2011). The duplicated B-class MADS-box genes display dualistic characters in orchid floral organ identity and growth. *Plant Cell Physiol*, 52(9), 1515-1531. doi:<https://doi.org/10.1093/pcp/pcr092>
- Pan, Z. J., Tsai, W. C., & Chen, H. H. (2017). Flower Development of *Phalaenopsis* Orchids Involves Functionally Divergent B-Class and E-Class MADS-Box Genes *Orchid Biotechnology III* (pp. 251-288).

- Parab, G., & Sellappan, K. (2008). Assessment of genetic variation among populations of *Rhynchostylis retusa*, an epiphytic orchid from Goa, India using ISSR and RAPD markers. *Indian J Biotechnol*, 7.
- Pnueli, L., Hareven, D., Broday, L., Hurwitz, C., & Lifschitz, E. (1994). The TM5 MADS Box Gene Mediates Organ Differentiation in the Three Inner Whorls of Tomato Flowers. *Plant Cell*, 6(2), 175-186. doi:10.1105/tpc.6.2.175
- Puranik, S., Acajjaoui, S., Conn, S., Costa, L., Conn, V., Vial, A., . . . Zubieta, C. (2014). Structural Basis for the Oligomerization of the MADS Domain Transcription Factor SEPALLATA3 in *Arabidopsis*. *Plant Cell*, 26(9), 3603. doi:10.1105/tpc.114.127910
- Purugganan, M. D., Rounsley, S. D., Schmidt, R. J., & Yanofsky, M. F. (1995). Molecular evolution of flower development: diversification of the plant MADS-box regulatory gene family. *Genetics*, 140(1), 345-356.
- Rudall, P. J., & Bateman, R. M. (2002). Roles of synorganisation, zygomorphy and heterotopy in floral evolution: the gynostemium and labellum of orchids and other lilioid monocots. *Biol Rev Camb Philos Soc*, 77(3), 403-441. doi:<https://doi.org/10.1017/s1464793102005936>
- Sirisawat, S., Ezura, H., Fukuda, N., Kounosu, T., & Handa, T. (2010). Ectopic expression of an AP3-like and a PI-like genes from 'Sekkoku' orchid (*Dendrobium moniliforme*) causes the homeotic conversion of sepals to petals in whorl 1 and the suppression of carpel development in whorl 4 in *Arabidopsis* flowers. *Plant Biotechnol*, 27(2), 183-192. doi:<https://doi.org/10.5511/plantbiotechnology.27.183>
- Sirisawat, S., Fukuda, N., Ezura, H., & Handa, T. (2009). *DMMADS4*, a *DEF*-like gene from *Dendrobium* is required for floral organ identity and flower longevity of orchid. *Acta Hort*, 836, 259-264. doi:<https://doi.org/10.17660/ActaHortic.2009.836.37>
- Theissen, G., Becker, A., Di Rosa, A., Kanno, A., Kim, J. T., Münster, T., . . . Saedler, H. (2000). A short history of MADS-box genes in plants. *Plant Mol Biol*, 42(1), 115-149.
- Theißen, G., & Gramzow, L. (2016). Chapter 8 - Structure and Evolution of Plant MADS Domain Transcription Factors. In D. H. Gonzalez (Ed.), *Plant Transcription Factors* (pp. 127-138). Boston: Academic Press.
- Theißen, G., Melzer, R., & Rümpler, F. (2016). MADS-domain transcription factors and the floral quartet model of flower development: linking plant development and evolution. *Development*, 143(18), 3259-3271. doi:<https://doi.org/10.1242/dev.134080>
- Theißen, G., & Saedler, H. (2001). Floral quartets. *Nature*, 409(6819), 469-471. doi:<https://doi.org/10.1038/35054172>
- Tsai, W. C., Kuoh, C. S., Chuang, M. H., Chen, W. H., & Chen, H. H. (2004). Four *DEF*-like MADS box genes displayed distinct floral morphogenetic roles in *Phalaenopsis* orchid. *Plant Cell Physiol*, 45(7), 831-844. doi:<https://doi.org/10.1093/pcp/pch095>
- Weigel, D., & Meyerowitz, E. M. (1994). The ABCs of floral homeotic genes. *Cell*, 78(2), 203-209. doi:[https://doi.org/10.1016/0092-8674\(94\)90291-7](https://doi.org/10.1016/0092-8674(94)90291-7)

- Xiang, L., Chen, Y., Chen, L., Fu, X., Zhao, K., Zhang, J., & Sun, C. (2018). B and E MADS-box genes determine the perianth formation in *Cymbidium goeringii* Rchb.f. *Physiol Plant*, 162(3), 353-369. doi:<https://doi.org/10.1111/ppl.12647>
- Xu, Y., Teo, L. L., Zhou, J., Kumar, P. P., & Yu, H. (2006). Floral organ identity genes in the orchid *Dendrobium crumenatum*. *Plant J*, 46(1), 54-68. doi:<https://doi.org/10.1111/j.1365-313X.2006.02669.x>
- Zahn, L. M., Kong, H., Leebens-Mack, J. H., Kim, S., Soltis, P. S., Landherr, L. L., . . . Ma, H. (2005). The evolution of the SEPALLATA subfamily of MADS-box genes: a preangiosperm origin with multiple duplications throughout angiosperm history. *Genetics*, 169(4), 2209-2223. doi:10.1534/genetics.104.037770





APPENDIX

APPENDIX A CHEMICALS AND SOLUTIONS

75% ethanol (50 mL)

Add following components

Absolute ethanol	37.5 mL
Distilled water	63.5 mL

Chloroform (50 mL)

Separated 50 mL of Chloroform to bottle with aluminium foil wrap

0.5M EDTA, pH

1. Add following components

EDTA	9.306
------	-------

2. Dissolved completely in 40 mL of distilled water and adjusted the pH to 8.0 with HCl
3. Adjusted the volume to 300 mL
4. Autoclaved at 121°C for 30 minutes and stored at room temperature

50X TAE buffer

1. Add following components

Tris-base	242 g
Glacial acetic acid	57.1 mL
0.5M EDTA (pH 8.0)	100

2. Dissolved completely and adjusted the volume to 1000 mL
3. Stored at room temperature
4. Prepared 1X TAE buffer before using

100 mg/ml Ampicillin (5 mL)

1. Add following components

Ampicillin	0.5 g
------------	-------

Distilled water 5 mL

2. Dissolved completely and filtered with sterile syringe filter
3. Separated 1 mL to 1.5 mL microcentrifuge tube with aluminium foil wrap
4. Stored at 4°C

2M D (+) Glucose (10 mL)

1. Add following components

D(+) Glucose 0.5 g

2. Dissolved completely with distilled water and adjusted the volume to 10 mL
3. Filtered with sterile syringe filter and separated 1 mL to 1.5 mL microcentrifuge tube
4. Stored at -20°C

1M MgSO₄.7H₂O (10 mL)

1. Add following components

MgSO₄.7H₂O 2.64 g

2. Dissolved completely with distilled water and adjusted the volume to 10 mL
3. Filtered with sterile syringe filter and separated 1 mL to 1.5 mL microcentrifuge tube
4. Stored at -20°C

1M MgCl₂.6H₂O (10 mL)

1. Add following components

MgCl₂.6H₂O 2.03 g

2. Dissolved completely with distilled water and adjusted the volume to 10 mL
3. Filtered with sterile syringe filter and separated 1 mL to 1.5 mL microcentrifuge tube, then, stored at -20°C

1M CaCl₂.2H₂O (100 mL)

1. Add following components

CaCl₂.2H₂O 14.7g

2. Dissolved completely in 80 mL of distilled water and adjusted the volume to 100 mL
3. Filtered with sterile syringe filter and stored at 4°C
4. Prepared 100mM CaCl₂ every time before using

1.5% Agarose gel (40 mL)

1. Add following components

Agarose	0.6 g
1X TAE buffer	40 mL

2. Heated and dissolved completely using microwave oven
3. Cast agarose gel in tray for 30 minutes and put into the electrophoresis tank to run
4. The gel was visualized by staining with ethidium bromide under UV light

SOB medium (300 mL)

1. Add following components

Tryptone	6 g
Yeast extract	4.5 g
NaCl	0.175 g
KCl	0.55

2. Dissolved completely in 150 mL of distilled water and adjusted the volume to 300 mL
3. Separated 50 mL of SOB medium to Erlenmeyer flask and capped with cotton plug
4. Autoclaved at 121°C for 30 minutes, cool down and stored at 4°C

SOC medium (60 mL)

1. Add following components

Tryptone	1.2 g
Yeast extract	0.3 g
NaCl	0.0351 g
KCl	0.111

2. Dissolved completely in 50 mL of distilled water, adjusted the volume to 60 mL
3. Transferred to Erlenmeyer flask and capped with cotton plug
4. Autoclaved at 121°C for 30 minutes, cool down and stored at 4°C

5. Added (Before using)

1M MgSO ₄ .7H ₂ O	600 μL
1M MgCl ₂ .6H ₂ O	600 μL
2M D(+)-Glucose	600

2X YT broth (200 mL)

1. Add following components

Tryptone	3.2 g
Yeast extract	2 g
NaCl	1 g

2. Dissolved completely in 150 mL of distilled water and adjusted the volume to 200 mL
3. Transferred to Erlenmeyer flask and capped with cotton plug
4. Autoclaved at 121°C for 30 minutes, cool down and stored at 4°C

2X YT plate (800 mL)

1. Add following components

Tryptone	12.8 g
Yeast extract	8 g
NaCl	2 g

2. Dissolved completely in 500 mL of distilled water and adjusted the volume to 800 mL
3. Transferred to Erlenmeyer flask which has 12 g of Agar and capped with cotton plug
4. Autoclaved at 121°C for 30 minutes, cool down and added Ampicillin
5. Cast the 2X YT medium in plate aseptically (approximately 20 mL per plate)

APPENDIX B SUPPLEMENTARY DATA

```

                *           20           *           40           *           60
S-HrDEF-1 : AACTGCGcGGTCTTGAGCAAACCTTGGAAAGAGTCTCTGAGAATTGTTAGGCATAGAAAAGT : 60
S-HrDEF-2 : AACTGCGcGGTCTTGAGCAAACCTTGGAAAGAGTCTCTGAGAATTGTTAGGCATAGAAAAGT : 60
S-HrDEF-3 : AACTGCGcGGTCTTGAGCAAACCTTGGAAAGAGTCTCTGAGAATTGTTAGGCATAGAAAAGT : 60
S-HrDEF-5 : AACTGCGcGGTCTTGAGCAAACCTTGGAAAGAGTCTCTGAGAATTGTTAGGCATAGAAAAGT : 60
S-HrDEF-7 : AACTGCGcGGTCTTGAGCAAACCTTGGAAAGAGTCTCTGAGAATTGTTAGGCATAGAAAAGT : 60
S-HrDEF-17 : AACTGCGcGGTCTTGAGCAAACCTTGGAAAGAGTCTCTGAGAATTGTTAGGCATAGAAAAGT : 60

                *           80           *           100          *           120
S-HrDEF-1 : ATCATGTGATCGCCACACAAAACCTGACACTTACAAGAAAAAGCTTAAAAGCACAAGGGAAA : 120
S-HrDEF-2 : ATCATGTGATCGCCACACAAAACCTGACACTTACAAGAAAAAGCTTAAAAGCACAAGGGAAA : 120
S-HrDEF-3 : ATCATGTGATCGCCACACAAAACCTGACACTTACAAGAAAAAGCTTAAAAGCACAAGGGAAA : 120
S-HrDEF-5 : ATCATGTGATCGCCACACAAAACCTGACACTTACAAGAAAAAGCTTAAAAGCACAAGGGAAA : 120
S-HrDEF-7 : ATCATGTGATCGCCACACAAAACCTGACACTTACAAGAAAAAGCTTAAAAGCACAAGGGAAA : 120
S-HrDEF-17 : ATCATGTGATCGCCACACAAAACCTGACACTTACAAGAAAAAGCTTAAAAGCACAAGGGAAA : 120

                *           140          *           160          *           180
S-HrDEF-1 : CTTACCGCGCTCTAATACATGAACCTGGATATGAAAGAGGAGAATCCGAACTACGGTTTTTA : 180
S-HrDEF-2 : CTTACCGCGCTCTAATACATGAACCTGGATATGAAAGAGGAGAATCCGAACTACGGTTTTTA : 180
S-HrDEF-3 : CTTACCGCGCTCTAATACATGAACCTGGATATGAAAGAGGAGAATCCGAACTACGGTTTTTA : 180
S-HrDEF-5 : CTTACCGCGCTCTAATACATGAACCTGGATATGAAAGAGGAGAATCCGAACTACGGTTTTTA : 180
S-HrDEF-7 : CTTACCGCGCTCTAATACATGAACCTGGATATGAAAGAGGAGAATCCGAACTACGGTTTTTA : 180
S-HrDEF-17 : CTTACCGCGCTCTAATACATGAACCTGGATATGAAAGAGGAGAATCCGAACTACGGTTTTTA : 180

                *           200          *           220          *           240
S-HrDEF-1 : ATGTGAAAGCCATAGTAGAATTTATGAAAATTC AATTC CAATGGTGAATGAGTGCCCTC : 240
S-HrDEF-2 : ATGTGAAAGCCATAGTAGAATTTATGAAAATTC AATTC CAATGGTGAATGAGTGCCCTC : 240
S-HrDEF-3 : ATGTGAAAGCCATAGTAGAATTTATGAAAATTC AATTC CAATGGTGAATGAGTGCCCTC : 240
S-HrDEF-5 : ATGTGAAAGCCATAGTAGAATTTATGAAAATTC AATTC CAATGGTGAATGAGTGCCCTC : 240
S-HrDEF-7 : ATGTGAAAGCCATAGTAGAATTTATGAAAATTC AATTC CAATGGTGAATGAGTGCCCTC : 240
S-HrDEF-17 : ATGTGAAAGCCATAGTAGAATTTATGAAAATTC AATTC CAATGGTGAATGAGTGCCCTC : 240

                *           260          *           280          *           300
S-HrDEF-1 : AGATGTTTTCCCTTTAGGGTTGTTTCATCCAAATCAGCCCAATCTGCTTGGCTTAGGTTATG : 300
S-HrDEF-2 : AGATGTTTTCCCTTTAGGGTTGTTTCATCCAAATCAGCCCAATCTGCTTGGCTTAGGTTATG : 300
S-HrDEF-3 : AGATGTTTTCCCTTTAGGGTTGTTTCATCCAAATCAGCCCAATCTGCTTGGCTTAGGTTATG : 300
S-HrDEF-5 : AGATGTTTTCCCTTTAGGGTTGTTTCATCCAAATCAGCCCAATCTGCTTGGCTTAGGTTATG : 300
S-HrDEF-7 : AGATGTTTTCCCTTTAGGGTTGTTTCATCCAAATCAGCCCAATCTGCTTGGCTTAGGTTATG : 300
S-HrDEF-17 : AGATGTTTTCCCTTTAGGGTTGTTTCATCCAAATCAGCCCAATCTGCTTGGCTTAGGTTATG : 300

                *           320          *
S-HrDEF-1 : AATCACAGGATCTCAGCCTCGCCTGATAAT : 330
S-HrDEF-2 : AATCACAGGATCTCAGCCTCGCCTGATAAT : 330
S-HrDEF-3 : AATCACAGGATCTCAGCCTCGCCTGATAAT : 330
S-HrDEF-5 : AATCACAGGATCTCAGCCTCGCCTGATAAT : 330
S-HrDEF-7 : AATCACAGGATCTCAGCCTCGCCTGATAAT : 330
S-HrDEF-17 : AATCACAGGATCTCAGCCTCGCCTGATAAT : 330

```

Figure 50 Multiple sequence alignment of selected clones of amplified S-HrDEF fragment in *R. retusa*. Colony 1, 2, 3, 5, 7 and 17 were represented DEF-like clade 1.

```

          *           20           *           40           *           60
S-HrDEF-4 : AACTGCGCGGTCTTGAGCAAAATATGGACGAGGCCCTAAAGCTTGTAAAGGAATcGAAAGT : 60
S-HrDEF-6 : AACTGCGCGGTCTTGAGCAAAATATGGACGAGGCCCTAAAGCTTGTAAAGGAATcGAAAGT : 60
S-HrDEF-8 : AACTGCGCGGTCTTGAGCAAAATATGGACGAGGCCCTAAAGCTTGTAAAGGAATcGAAAGT : 60
S-HrDEF-12 : AACTGCGCGGTCTTGAGCAAAATATGGACGAGGCCCTAAAGCTTGTAAAGGAATcGAAAGT : 60

          *           80           *           100          *           120
S-HrDEF-4 : ATCACGtCATCAGCACGCaGACAGATACCTTCAAAAAAAGTTGAAAAACTCTCAAGAAA : 120
S-HrDEF-6 : ATCACGtCATCAGCACGCaGACAGATACCTTCAAAAAAAGTTGAAAAACTCTCAAGAAA : 120
S-HrDEF-8 : ATCACGtCATCAGCACGCaGACAGATACCTTCAAAAAAAGTTGAAAAACTCTCAAGAAA : 120
S-HrDEF-12 : ATCACGtCATCAGCACGCaGACAGATACCTTCAAAAAAAGTTGAAAAACTCTCAAGAAA : 120

          *           140          *           160          *           180
S-HrDEF-4 : CCCAaAGGAACTTACTACGGGAGCTGGAAActGAGCATGCCGTCTACTATGTGGATGATG : 180
S-HrDEF-6 : CCCAaAGGAACTTACTACGGGAGCTGGAAActGAGCATGCCGTCTACTATGTGGATGATG : 180
S-HrDEF-8 : CCCAaAGGAACTTACTACGGGAGCTGGAAActGAGCATGCCGTCTACTATGTGGATGATG : 180
S-HrDEF-12 : CCCAaAGGAACTTACTACGGGAGCTGGAAActGAGCATGCCGTCTACTATGTGGATGATG : 180

          *           200          *           220          *           240
S-HrDEF-4 : ATCCAAACAACtATGATGGCGCGCTTGCACTTGGAATGGGGCTTCCTACTTGTATTCAT : 240
S-HrDEF-6 : ATCCAAACAACtATGATGGCGCGCTTGCACTTGGAATGGGGCTTCCTACTTGTATTCAT : 240
S-HrDEF-8 : ATCCAAACAACtATGATGGCGCGCTTGCACTTGGAATGGGGCTTCCTACTTGTATTCAT : 240
S-HrDEF-12 : ATCCAAACAACtATGATGGCGCGCTTGCACTTGGAATGGGGCTTCCTACTTGTATTCAT : 240

          *           260          *           280          *           300
S-HrDEF-4 : TTCGTACCCAACCAAGCCAGCCAAACCTTCAAGGAGTGGGATATGTCCCTCAGGATCTtC : 300
S-HrDEF-6 : TTCGTACCCAACCAAGCCAGCCAAACCTTCAAGGAGTGGGATATGTCCCTCAGGATCTtC : 300
S-HrDEF-8 : TTCGTACCCAACCAAGCCAGCCAAACCTTCAAGGAGTGGGATATGTCCCTCAGGATCTtC : 300
S-HrDEF-12 : TTCGTACCCAACCAAGCCAGCCAAACCTTCAAGGAGTGGGATATGTCCCTCAGGATCTtC : 300

          *
S-HrDEF-4 : GtCTcGCcTaAT : 312
S-HrDEF-6 : GtCTcGCcTaAT : 312
S-HrDEF-8 : GtCTcGCcTaAT- : 311
S-HrDEF-12 : GtCTcGCcTaAT : 312

```

Figure 51 Multiple sequence alignment of selected clones of amplified S-HrDEF fragment in *R. retusa*. Colony 4, 6, 8 and 12 were represented *DEF*-like clade 3.

```

*           20           *           40           *           60
S-HrDEF-9 : AACTGCGcGGTCTTGAGCAAAAACATCGACGAGGCATTGAAGCTaGTACGAAATAGAAAA : 60
S-HrDEF-10 : AACTGCGcGGTCTTGAGCAAAAACATCGACGAGGCATTGAAGCTaGTACGAAATAGAAAA : 60
S-HrDEF-11 : AACTGCGcGGTCTTGAGCAAAAACATCGACGAGGCATTGAAGCTaGTACGAAATAGAAAA : 60
S-HrDEF-13 : AACTGCGcGGTCTTGAGCAAAAACATCGACGAGGCATTGAAGCTaGTACGAAATAGAAAA : 60
S-HrDEF-14 : AACTGCGcGGTCTTGAGCAAAAACATCGACGAGGCATTGAAGCTaGTACGAAATAGAAAA : 60
S-HrDEF-16 : AACTGCGcGGTCTTGAGCAAAAACATCGACGAGGCATTGAAGCTaGTACGAAATAGAAAA : 60

*           80           *           100          *           120
S-HrDEF-9 : ATCATGTAATCAGTACTCAAACGGGACCTACAAGAAGAAGCTGAAGAACTCTCAAGAAA : 120
S-HrDEF-10 : ATCATGTAATCAGTACTCAAACGGGACCTACAAGAAGAAGCTGAAGAACTCTCAAGAAA : 120
S-HrDEF-11 : ATCATGTAATCAGTACTCAAACGGGACCTACAAGAAGAAGCTGAAGAACTCTCAAGAAA : 120
S-HrDEF-13 : ATCATGTAATCAGTACTCAAACGGGACCTACAAGAAGAAGCTGAAGAACTCTCAAGAAA : 120
S-HrDEF-14 : ATCATGTAATCAGTACTCAAACGGGACCTACAAGAAGAAGCTGAAGAACTCTCAAGAAA : 120
S-HrDEF-16 : ATCATGTAATCAGTACTCAAACGGGACCTACAAGAAGAAGCTGAAGAACTCTCAAGAAA : 120

*           140          *           160          *           180
S-HrDEF-9 : CACACCGAAACTTAATGcACGAACTGGAAATCGTTGAGGACCACCCAGTCTtTGGGTACC : 180
S-HrDEF-10 : CACACCGAAACTTAATGcACGAACTGGAAATCGTTGAGGACCACCCAGTCTtTGGGTACC : 180
S-HrDEF-11 : CACACCGAAACTTAATGcACGAACTGGAAATCGTTGAGGACCACCCAGTCTtTGGGTACC : 180
S-HrDEF-13 : CACACCGAAACTTAATGcACGAACTGGAAATCGTTGAGGACCACCCAGTCTtTGGGTACC : 180
S-HrDEF-14 : CACACCGAAACTTAATGcACGAACTGGAAATCGTTGAGGACCACCCAGTCTtTGGGTACC : 180
S-HrDEF-16 : CACACCGAAACTTAATGcACGAACTGGAAATCGTTGAGGACCACCCAGTCTtTGGGTACC : 180

*           200          *           220          *           240
S-HrDEF-9 : aCGAGGATTCAAGCAATTATGAGGGCGTCTTGCTCTCGCAAAATGATGGGTCTCACATGT : 240
S-HrDEF-10 : aCGAGGATTCAAGCAATTATGAGGGCGTCTTGCTCTCGCAAAATGATGGGTCTCACATGT : 240
S-HrDEF-11 : aCGAGGATTCAAGCAATTATGAGGGCGTCTTGCTCTCGCAAAATGATGGGTCTCACATGT : 240
S-HrDEF-13 : aCGAGGATTCAAGCAATTATGAGGGCGTCTTGCTCTCGCAAAATGATGGGTCTCACATGT : 240
S-HrDEF-14 : aCGAGGATTCAAGCAATTATGAGGGCGTCTTGCTCTCGCAAAATGATGGGTCTCACATGT : 240
S-HrDEF-16 : aCGAGGATTCAAGCAATTATGAGGGCGTCTTGCTCTCGCAAAATGATGGGTCTCACATGT : 240

*           260          *           280          *           300
S-HrDEF-9 : ATGCCTTCCGAGTGCAACCCAACCAACCAAAATCTTCATGGAATGGGATATGGCTCCCAGG : 300
S-HrDEF-10 : ATGCCTTCCGAGTGCAACCCAACCAACCAAAATCTTCATGGAATGGGATATGGCTCCCAGG : 300
S-HrDEF-11 : ATGCCTTCCGAGTGCAACCCAACCAACCAAAATCTTCATGGAATGGGATATGGCTCCCAGG : 300
S-HrDEF-13 : ATGCCTTCCGAGTGCAACCCAACCAACCAAAATCTTCATGGAATGGGATATGGCTCCCAGG : 300
S-HrDEF-14 : ATGCCTTCCGAGTGCAACCCAACCAACCAAAATCTTCATGGAATGGGATATGGCTCCCAGG : 300
S-HrDEF-16 : ATGCCTTCCGAGTGCAACCCAACCAACCAAAATCTTCATGGAATGGGATATGGCTCCCAGG : 300

*
S-HrDEF-9 : ATCTccGcCTCGCcTgaT : 318
S-HrDEF-10 : ATCTccGcCTCGCcTgaT : 318
S-HrDEF-11 : ATCTccGcCTCGCcTgaT : 318
S-HrDEF-13 : ATCTccGcCTCGCcTgaT : 318
S-HrDEF-14 : ATCTccGcCTCGCcTgaT : 318
S-HrDEF-16 : ATCTccGcCTCGCcTgaT : 318

```

Figure 52 Multiple sequence alignment of selected clones of amplified S-HrDEF fragment in *R. retusa*. Colony 9, 10, 11, 13, 14 and 16 were represented DEF-like clade 4.

Table 5 BLAST® results of first partial *DEF*-like clade 1 nucleotide sequence in *R. retusa*. The result was shown first 10 highest total scores.

Description	Scientific Name	Max Score	Total Score	Query Cover	Per. ident	Accession
PREDICTED: Phalaenopsis equestris MADS-box transcription factor 16-like (LOC110028792), mRNA	<i>Phalaenopsis equestris</i>	544	544	99%	96.64	XM_020730780.1
Phalaenopsis equestris MADS box transcription factor (MADS2) mRNA, complete cds	<i>Phalaenopsis equestris</i>	544	544	99%	96.64	AY378149.1
Phalaenopsis hybrid cultivar MADS box AP3-like protein 17 mRNA, complete cds	<i>Phalaenopsis</i> hybrid cultivar	532	532	99%	96.02	AY771993.1
Phaius tancarvilleae B-class MADS-box protein AP3-1 mRNA, partial cds	<i>Phaius tancarvilleae</i>	424	424	96%	90.6	EU4444050.1
Cymbidium ensifolium APETALA3-like MADS-box protein mRNA, complete cds	<i>Cymbidium ensifolium</i>	385	385	99%	87.8	JN613151.1
Dendrobium hybrid cultivar B-class MADS-box protein AP3-1 mRNA, partial cds	<i>Dendrobium</i> hybrid cultivar	385	385	97%	88.2	EU4444025.1
Cymbidium faberi AP3 mRNA, complete cds	<i>Cymbidium faberi</i>	385	385	99%	87.8	HM208536.1
Cymbidium goeringii clone ep123.comp36586_c0_seq2 MADS-box protein-like protein mRNA, complete cds	<i>Cymbidium goeringii</i>	379	379	99%	87.5	MF474248.1
Cymbidium goeringii AP3 mRNA, complete cds	<i>Cymbidium goeringii</i>	379	379	99%	87.5	HM106983.1
Cymbidium hybrid cultivar AP3-like MADS box protein (MADS1) mRNA, complete cds	<i>Cymbidium</i> hybrid cultivar	379	379	96%	88.09	DQ683575.1

Table 6 BLAST® results of first partial DEF-like clade 2 nucleotide sequence in *R. retusa*. The result was shown all sequences which had similarity from NCBI database.

Description	Scientific Name	Max Score	Total Score	Query Cover	Per. ident	Accession
Phalaenopsis equestris MADS box transcription factor (MADS5) gene, complete cds	<i>Phalaenopsis equestris</i>	497	906	79%	96.67	AY378151.1
Cymbidium goeringii clone ep456.comp52699_c0_seq3 MADS-box protein-like protein mRNA, complete cds	<i>Cymbidium goeringii</i>	379	379	46%	87.02	MF462080.1
PREDICTED: Phalaenopsis equestris MADS-box transcription factor 16-like (LOC110024739), transcript variant X2, mRNA	<i>Phalaenopsis equestris</i>	374	854	81%	90	XM_020724886.1
PREDICTED: Phalaenopsis equestris MADS-box transcription factor 16-like (LOC110024739), transcript variant X1, mRNA	<i>Phalaenopsis equestris</i>	361	841	81%	89.08	XM_020724880.1
Phalaenopsis equestris MADS5 transcription factor (MADS5) mRNA, complete cds	<i>Phalaenopsis equestris</i>	361	841	81%	89.08	AY378148.1
PREDICTED: Dendrobium catenatum MADS-box transcription factor 16 (LOC110103364), mRNA	<i>Dendrobium catenatum</i>	303	522	61%	86.01	XM_020832037.2
Cymbidium goeringii DEF2 mRNA, complete cds	<i>Cymbidium goeringii</i>	294	294	39%	85.66	KX347446.1
Gongora galeata DEFICIENS-like MADS-box transcription factor (DEF2) mRNA, complete cds	<i>Gongora galeata</i>	292	292	39%	85.51	FJ804098.1
Cymbidium goeringii clone ep123.comp41679_c0_seq4 MADS-box protein-like protein mRNA, complete cds	<i>Cymbidium goeringii</i>	281	281	39%	84.72	MF474240.1
Oncidium cv. 'Gower Ramsey' MADS box protein (MADS3) mRNA, complete cds	<i>Oncidium</i> hybrid cultivar	265	265	38%	83.99	AY196350.1
Oncidium hybrid cultivar Ap3-like MADS box protein mRNA, complete cds	<i>Oncidium</i> hybrid cultivar	261	261	38%	83.81	GU644447.1
Oncidium hybrid cultivar B-class MADS-box protein AP3-4 mRNA, partial cds	<i>Oncidium</i> hybrid cultivar	255	255	38%	83.64	EU444044.1

Table 7 BLAST® results of first partial *DEF*-like clade 3 nucleotide sequence in *R. retusa*. The result was shown first 10 highest total scores.

Description	Scientific Name	Max Score	Total Score	Query Cover	Per. ident	Accession
PREDICTED: Phalaenopsis equestris MADS-box transcription factor 16-like (LOC110018630), mRNA	<i>Phalaenopsis equestris</i>	505	505	99%	96.12	XM_020715991.1
Phalaenopsis equestris MADS box transcription factor (MADS3) mRNA, complete cds	<i>Phalaenopsis equestris</i>	505	505	99%	96.12	AY378150.1
PREDICTED: Dendrobium catenatum MADS-box transcription factor 16 (LOC110093040), mRNA	<i>Dendrobium catenatum</i>	418	418	98%	91.21	XM_020817789.2
Dendrobium hybrid cultivar B-class MADS-box protein AP3-2 mRNA, partial cds	<i>Dendrobium hybrid cultivar</i>	418	418	98%	91.21	EU4444026.1
Dendrobium hybrid cultivar MADS box AP3-like protein 1 mRNA, complete cds	<i>Dendrobium hybrid cultivar</i>	407	407	98%	90.55	EU927372.1
Dendrobium crumenatum MADS box AP3-like protein B (AP3B) mRNA, complete cds	<i>Dendrobium crumenatum</i>	407	407	98%	90.55	DQ119839.1
Phaius tancarvilleae B-class MADS-box protein AP3-2 mRNA, partial cds	<i>Phaius tancarvilleae</i>	396	396	98%	89.9	EU4444051.1
Brassavola nodosa B-class MADS-box protein AP3-3 mRNA, partial cds	<i>Brassavola nodosa</i>	396	396	98%	89.9	EU4444023.1
Dendrobium moniliforme AP3-related protein B mRNA, complete cds	<i>Dendrobium moniliforme</i>	396	396	98%	89.9	EU056328.1
Cymbidium sinense AP3-1 mRNA, complete cds	<i>Cymbidium sinense</i>	390	390	98%	89.58	KY797298.1

Table 8 BLAST® results of first partial *DEF*-like clade 4 nucleotide sequence in *R. retusa*. The result was shown all sequences which had similarity from NCBI database.

Description	Scientific Name	Max Score	Total Score	Query Cover	Per. ident	Accession
PREDICTED: Phalaenopsis equestris MADS-box transcription factor 16-like (LOC110030201), mRNA	<i>Phalaenopsis equestris</i>	472	472	100%	93.4	XM_020732801.1
Phalaenopsis equestris MADS box transcription factor (MADS4) mRNA, complete cds	<i>Phalaenopsis equestris</i>	472	472	100%	93.4	AY378147.1
Oncidium hybrid cultivar B-class MADS-box protein AP3-3 mRNA, partial cds	<i>Oncidium hybrid cultivar</i>	396	396	98%	89.46	EU4444043.1
PREDICTED: Dendrobium catenatum MADS-box transcription factor 16 (LOC110096227), mRNA	<i>Dendrobium catenatum</i>	394	394	100%	88.99	XM_020822071.2
Dendrobium moniliforme AP3-related protein 4 (MADS4) mRNA, complete cds	<i>Dendrobium moniliforme</i>	387	387	99%	88.64	GU132995.1
Dendrobium hybrid cultivar B-class MADS-box protein AP3-3 mRNA, partial cds	<i>Dendrobium hybrid cultivar</i>	385	385	98%	88.82	EU4444027.1
Cymbidium goeringii DEF4 mRNA, complete cds	<i>Cymbidium goeringii</i>	340	340	99%	86.16	KU058678.1
Cymbidium goeringii clone ep456.comp52927_c0_seq1 MADS-box protein-like protein mRNA, partial cds	<i>Cymbidium goeringii</i>	281	281	99%	82.78	MF462092.1
Phalaenopsis equestris MADS box transcription factor 4 (MADS4) gene, promoter region and complete cds	<i>Phalaenopsis equestris</i>	99	177	31%	96.61	KJ127931.1

Table 9 BLAST® results of first partial *AGL6*-like clade 1 nucleotide sequence in *R. retusa*. The result was shown first 10 highest total scores.

Description	Scientific Name	Max Score	Total Score	Query Cover	Per. ident	Accession
PREDICTED: Phalaenopsis equestris MADS-box transcription factor 6-like (LOC110035879), partial mRNA	<i>Phalaenopsis equestris</i>	1116	1116	100%	96.72	XM_020740192.1
<i>Cymbidium ensifolium</i> SEP-like MADS-box protein mRNA, complete cds	<i>Cymbidium ensifolium</i>	896	896	100%	90.9	JN613148.1
<i>Cymbidium goeringii</i> AGL6 mRNA, complete cds	<i>Cymbidium goeringii</i>	896	896	100%	90.9	HM208533.1
<i>Cymbidium goeringii</i> AGL6-like protein 1 mRNA, complete cds	<i>Cymbidium goeringii</i>	896	896	100%	90.9	GQ265900.1
<i>Cymbidium goeringii</i> clone ep123.comp42499_c0_seq1 MADS-box protein-like protein mRNA, complete cds	<i>Cymbidium goeringii</i>	891	891	100%	90.75	MF474236.1
<i>Cymbidium faberi</i> AGL6 mRNA, complete cds	<i>Cymbidium faberi</i>	891	891	100%	90.75	HM208534.1
PREDICTED: <i>Dendrobium catenatum</i> MADS-box transcription factor 6 (LOC110111596), mRNA	<i>Dendrobium catenatum</i>	880	880	100%	90.45	XM_020843535.2
<i>Oncidium Gower Ramsey</i> MADS box transcription factor 7 (MADS7) mRNA, complete cds	<i>Oncidium hybrid cultivar</i>	774	774	100%	87.61	HMI140845.1
<i>Erycina pusilla</i> MADS-box protein 3 mRNA, complete cds	<i>Erycina pusilla</i>	652	652	100%	84.33	KJ002728.1
<i>Phyllostachys edulis</i> MADS-box protein mRNA, complete cds	<i>Phyllostachys edulis</i>	316	316	99%	75.52	EU327784.1

Table 10 BLAST® results of first partial *AGL6*-like clade 2 nucleotide sequence in *R. retusa*. The result was shown first 10 highest total scores.

Description	Scientific Name	Max Score	Total Score	Query Cover	Per. ident	Accession
Dendrobium hybrid cultivar <i>AGL6</i> (<i>Agl6</i>) mRNA, complete cds	Dendrobium hybrid cultivar	228	228	100%	89.07	KF550139.1
PREDICTED: Dendrobium catenatum MADS-box transcription factor 6-like (LOC110095336), mRNA	Dendrobium catenatum	195	195	100%	85.79	XM_020820844.2
Erycina pusilla MADS-box protein 5 mRNA, complete cds	Erycina pusilla	174	174	100%	83.87	KJ002730.1
Oncidium Gower Ramsey MADS box transcription factor 1 (MADS1) mRNA, complete cds	Oncidium hybrid cultivar	174	174	100%	83.96	HM140843.1
Erycina pusilla MADS5 (MADS5) gene, complete cds	Erycina pusilla	71.3	71.3	24%	95.45	KJ715213.1

Table 11 BLAST® results of final partial *DEF*-like clade 1 nucleotide sequence in *R. retusa*. The result was shown first 10 highest total scores.

Description	Scientific Name	Max Score	Total Score	Query Cover	Per. ident	Accession
Phalaenopsis equestris MADS box transcription factor (MADS2) mRNA, complete cds	<i>Phalaenopsis equestris</i>	641	641	100%	91.7	<u>AY378149.1</u>
Phalaenopsis hybrid cultivar MADS box AP3-like protein 17 mRNA, complete cds	<i>Phalaenopsis</i> hybrid cultivar	612	612	96%	91.56	<u>AY771993.1</u>
PREDICTED: Phalaenopsis equestris MADS-box transcription factor 16-like (LOC110028792), mRNA	<i>Phalaenopsis equestris</i>	545	545	78%	93.75	<u>XM_020730780.1</u>
Phaius tancarvilleae B-class MADS-box protein AP3-1 mRNA, partial cds	<i>Phaius tancarvilleae</i>	385	385	63%	89.93	<u>EU444050.1</u>
<i>Cymbidium faberi</i> AP3 mRNA, complete cds	<i>Cymbidium.faberi</i>	364	364	90%	82.41	<u>HM208536.1</u>
<i>Cymbidium goeringii</i> clone ep123.comp36586_c0_seq2 MADS-box protein-like protein mRNA, complete cds	<i>Cymbidium goeringii</i>	363	363	94%	81.78	<u>MF474248.1</u>
<i>Oncidium</i> Gower Ramsey MADS box transcription factor 5 (MADS5) mRNA, complete cds	<i>Oncidium</i> hybrid cultivar	359	359	70%	86.4	<u>HM140840.2</u>
<i>Cymbidium ensifolium</i> APETALA3-like MADS-box protein mRNA, complete cds	<i>Cymbidium ensifolium</i>	353	353	70%	86.1	<u>JN613151.1</u>
<i>Cymbidium goeringii</i> AP3 mRNA, complete cds	<i>Cymbidium goeringii</i>	348	348	70%	85.8	<u>HM106983.1</u>
<i>Cymbidium</i> hybrid cultivar AP3-like MADS box protein (MADS1) mRNA, complete cds	<i>Cymbidium</i> hybrid cultivar	348	348	70%	85.8	<u>DQ683575.1</u>

Table 12 BLAST® results of final partial *DEF*-like clade 2 nucleotide sequence in *R. retusa*. The result was shown first 12 highest total scores.

Description	Scientific Name	Max Score	Total Score	Query Cover	Per. ident	Accession
PREDICTED: Phalaenopsis equestris MADS-box transcription factor 16-like (LOC110024739), transcript variant X2, mRNA	<i>Phalaenopsis equestris</i>	1062	1193	99%	92.49	<u>XM_020724886.1</u>
PREDICTED: Phalaenopsis equestris MADS-box transcription factor 16-like (LOC110024739), transcript variant X1, mRNA	<i>Phalaenopsis equestris</i>	1050	1180	99%	92.12	<u>XM_020724880.1</u>
Phalaenopsis equestris MADS5 transcription factor (MADS5) mRNA, complete cds	<i>Phalaenopsis equestris</i>	1050	1050	89%	92.12	<u>AY378148.1</u>
PREDICTED: Dendrobium catenatum MADS-box transcription factor 16 (LOC110103364), mRNA	<i>Dendrobium catenatum</i>	686	686	80%	85.27	<u>XM_020832037.2</u>
Cymbidium goeringii DEF2 mRNA, complete cds	<i>Cymbidium goeringii</i>	665	665	79%	84.96	<u>KX347446.1</u>
Cymbidium goeringii clone ep123.comp41679_c0_seq4 MADS-box protein-like protein mRNA, complete cds	<i>Cymbidium goeringii</i>	651	651	79%	84.56	<u>MF474240.1</u>
Gongora galeata DEFICIENS-like MADS-box transcription factor (DEF2) mRNA, complete cds	<i>Gongora galeata</i>	588	588	72%	84.39	<u>FJ804098.1</u>
Oncidium hybrid cultivar B-class MADS-box protein AP3-4 mRNA, partial cds	<i>Oncidium</i> hybrid cultivar	536	536	68%	83.89	<u>EU4444044.1</u>
Oncidium cv. 'Gower Ramsey' MADS box protein (MADS3) mRNA, complete cds	<i>Oncidium</i> hybrid cultivar	529	529	68%	83.54	<u>AY196350.1</u>
Oncidium hybrid cultivar Ap3-like MADS box protein mRNA, complete cds	<i>Oncidium</i> hybrid cultivar	525	525	68%	83.45	<u>GU644447.1</u>
Erycina pusilla MADS-box protein 14 mRNA, complete cds	<i>Erycina pusilla</i>	388	388	57%	81.33	<u>KJ002739.1</u>
Phalaenopsis equestris MADS box transcription factor (MADS5) gene, complete cds	<i>Phalaenopsis equestris</i>	291	901	75%	87.5	<u>AY378151.1</u>

Table 13 BLAST® results of final partial DEF-like clade 3 nucleotide sequence in *R. retusa*. The result was shown first 12 highest total scores.

Description	Scientific Name	Max Score	Total Score	Query Cover	Per. ident	Accession
PREDICTED: Phalaenopsis equestris MADS-box transcription factor 16-like (LOC110018630), mRNA	<i>Phalaenopsis equestris</i>	547	547	100%	90.47	XM_020715991.1
Phalaenopsis equestris MADS box transcription factor (MADS3) mRNA, complete cds	<i>Phalaenopsis equestris</i>	538	606	87%	95.81	AY378150.1
Dendrobium crumenatum MADS box AP3-like protein B (AP3B) mRNA, complete cds	<i>Dendrobium crumenatum</i>	427	427	89%	86.82	DQ119839.1
PREDICTED: Dendrobium catenatum MADS-box transcription factor 16 (LOC110093040), mRNA	<i>Dendrobium catenatum</i>	425	425	94%	85.57	XM_020817789.2
Dendrobium hybrid cultivar MADS box AP3-like protein 1 mRNA, complete cds	<i>Dendrobium</i> hybrid cultivar	422	422	89%	86.56	EU927372.1
Dendrobium hybrid cultivar B-class MADS-box protein AP3-2 mRNA, partial cds	<i>Dendrobium</i> hybrid cultivar	385	385	66%	90.91	EU444026.1
Cymbidium faberi DEFICIENS mRNA, complete cds	<i>Cymbidium faberi</i>	366	366	84%	85.21	HM208535.1
Phaius tancarvilleae B-class MADS-box protein AP3-2 mRNA, partial cds	<i>Phaius tancarvilleae</i>	363	363	66%	89.51	EU444051.1
Brassavola nodosa B-class MADS-box protein AP3-3 mRNA, partial cds	<i>Brassavola nodosa</i>	363	363	66%	89.51	EU444023.1
Dendrobium moniliforme AP3-related protein B mRNA, complete cds	<i>Dendrobium moniliforme</i>	363	363	66%	89.51	EU056328.1
Cymbidium goeringii clone ep123.comp26182_c0_seq1 MADS-box protein-like protein mRNA, partial cds	<i>Cymbidium goeringii</i>	361	361	84%	84.93	MF462073.1
Cymbidium ensifolium AP3-like MADS-box 3 protein mRNA, complete cds	<i>Cymbidium ensifolium</i>	361	361	84%	84.93	JQ326260.1

Table 14 BLAST® results of final partial DEF-like clade 4 nucleotide sequence in *R. retusa*. The result was shown all sequences which had similarity from NCBI database.

Description	Scientific Name	Max Score	Total Score	Query Cover	Per. ident	Accession
PREDICTED: Phalaenopsis equestris MADS-box transcription factor 16-like (LOC110030201), mRNA	<i>Phalaenopsis equestris</i>	641	641	85%	91.94	XM_020732801.1
Phalaenopsis equestris MADS box transcription factor (MADS4) mRNA, complete cds	<i>Phalaenopsis equestris</i>	617	617	82%	91.86	AY378147.1
Phalaenopsis equestris MADS box transcription factor 4 (MADS4) gene, promoter region and complete cds	<i>Phalaenopsis equestris</i>	422	561	73%	90.57	KJ127931.1
PREDICTED: Dendrobium catenatum MADS-box transcription factor 16 (LOC110096227), mRNA	<i>Dendrobium catenatum</i>	363	363	57%	87.82	XM_020822071.2
Oncidium hybrid cultivar B-class MADS-box protein AP3-3 mRNA, partial cds	<i>Oncidium</i> hybrid cultivar	363	363	54%	89.04	EU4444043.1
Dendrobium moniliforme AP3-related protein 4 (MADS4) mRNA, complete cds	<i>Dendrobium moniliforme</i>	348	348	55%	87.84	GU132995.1
Dendrobium hybrid cultivar B-class MADS-box protein AP3-3 mRNA, partial cds	<i>Dendrobium</i> hybrid cultivar	346	346	54%	88.01	EU4444027.1
Cymbidium goeringii DEF4 mRNA, complete cds	<i>Cymbidium goeringii</i>	305	305	55%	85.28	KU058678.1
Cymbidium goeringii clone ep456.comp52927_c0_seq1 MADS-box protein-like protein mRNA, partial cds	<i>Cymbidium goeringii</i>	259	259	58%	81.76	MF462092.1

Table 15 BLAST® results of final partial *AGL6*-like clade 1 nucleotide sequence in *R. retusa*. The result was shown first 10 highest total scores.

Description	Scientific Name	Max Score	Total Score	Query Cover	Per. ident	Accession
PREDICTED: Phalaenopsis equestris MADS-box transcription factor 6-like (LOC110035879), partial mRNA	<i>Phalaenopsis equestris</i>	1116	1116	100%	96.72	<u>XM_020740192.1</u>
<i>Cymbidium ensifolium</i> SEP-like MADS-box protein mRNA, complete cds	<i>Cymbidium ensifolium</i>	896	896	100%	90.9	<u>JN613148.1</u>
<i>Cymbidium goeringii</i> AGL6 mRNA, complete cds	<i>Cymbidium goeringii</i>	896	896	100%	90.9	<u>HM208533.1</u>
<i>Cymbidium goeringii</i> AGL6-like protein 1 mRNA, complete cds	<i>Cymbidium goeringii</i>	896	896	100%	90.9	<u>GQ265900.1</u>
<i>Cymbidium goeringii</i> clone ep123.comp42499_c0_seq1 MADS-box protein-like protein mRNA, complete cds	<i>Cymbidium goeringii</i>	891	891	100%	90.75	<u>MF474236.1</u>
<i>Cymbidium faberi</i> AGL6 mRNA, complete cds	<i>Cymbidium faberi</i>	891	891	100%	90.75	<u>HM208534.1</u>
PREDICTED: <i>Dendrobium catenatum</i> MADS-box transcription factor 6 (LOC11011596), mRNA	<i>Dendrobium catenatum</i>	880	880	100%	90.45	<u>XM_020843535.2</u>
<i>Oncidium</i> Gower Ramsey MADS box transcription factor 7 (MADS7) mRNA, complete cds	<i>Oncidium</i> hybrid cultivar	774	774	100%	87.61	<u>HM140845.1</u>
<i>Erycina pusilla</i> MADS-box protein 3 mRNA, complete cds	<i>Erycina pusilla</i>	652	652	100%	84.33	<u>KJ002728.1</u>
<i>Phyllostachys edulis</i> MADS-box protein mRNA, complete cds	<i>Phyllostachys edulis</i>	316	316	99%	75.52	<u>EU327784.1</u>

Table 16 BLAST® results of final partial AGL6-like clade 2 nucleotide sequence in *R. retusa*. The result was shown first 10 highest total scores.

Description	Scientific Name	Max Score	Total Score	Query Cover	Per. ident	Accession
Cymbidium goeringii clone CL11237Contig1 MADS-box protein-like protein mRNA, complete cds	Cymbidium goeringii	468	468	80%	84.46	<u>MF474255.1</u>
Dendrobium hybrid cultivar AGL6 (Agl6) mRNA, complete cds	Dendrobium hybrid cultivar	374	374	70%	82.8	<u>KF550139.1</u>
PREDICTED: Dendrobium catenatum MADS-box transcription factor 6-like (LOC110095336), mRNA	Dendrobium catenatum	311	311	46%	86.74	<u>XM_020820844.2</u>
Oncidium Gower Ramsey MADS box transcription factor 1 (MADS1) mRNA, complete cds	Oncidium hybrid cultivar	296	296	68%	80	<u>HM140843.1</u>
Erycina pusilla MADS-box protein 5 mRNA, complete cds	Erycina pusilla	279	279	46%	84.75	<u>KJ002730.1</u>
Erycina pusilla MADS5 (MADS5) gene, complete cds	Erycina pusilla	71.3	71.3	7%	95.45	<u>KJ715213.1</u>

BIOGRAPHY

Name-Surname Sudarat Kasemcholathan

Date of Birth

Address

Education Background B.Sc (Biology)

