RESEARCH ARTICLE

SABRAO Journal
of Breeding and Genetics
50 (3) 313-328, 2018

MOLECULAR DIVERSITY OF THE FLOWERING RELATED GENE (*LEAFY*) ON SHALLOT (*Allium cepa* var. *aggregatum*) AND ALLIUM RELATIVES

MARLIN[1,2]*, A. MAHARIJAYA[1,3], A. PURWITO[1] and SOBIR[1,3]

[1]Department of Agronomy and Horticulture, Faculty of Agriculture, Bogor Agricultural University
Jalan Raya Darmaga Bogor 16680, Indonesia
[2]Department of Agroecotechnology, Faculty of Agriculture, University of Bengkulu
Jalan WR. Supratman Kandang Limun Bengkulu 38371, Indonesia
[3]Center for Tropical Horticulture Studies, Bogor Agricultural University, Indonesia
*Corresponding author's email: marlin@unib.ac.id
Email addresses of coauthors: awang.maharijaya@gmail.com, apurwito@ipb.ac.id, rsobir@yahoo.com

SUMMARY

Flowering initiation in shallot (*Allium cepa* var. *aggregatum*) has predicticably been associated to the function of the *LEAFY* (*LFY*) gene. This gene has been reported as the flowering meristem identity, a potential indicator of plant flowering ability. This study identified diversity of the *shLFY* (shallot-*LFY*) gene in 5 shallot genotypes in relation to their flowering pattern compared to homolog sequence of *Allium* relatives. Genomic DNA from the 5 genotypes was amplified using novel primers (F/5'-GACCCAACAGACCCCTAC-3' and R/5'-TTAGAACATAGAGGACACAGAG-3'), resulting in *shLFY* gene sequences with fragment lengths between 1,161 and1,253 bp. These fragments were found to consist of 3 exons with 2 introns. Identification of the *shLFY* gene from the sequences of 5 shallot genotypes confirms the existence of genetic polymorphism among shallot genotypes. The result of sequence analysis based on genomic and amino acid sequences shows 3 groups of *shLFY* genes in shallot related to flowering pattern. The first group consists of a single member, *bm1LFY* (Bentanis), and was named the naturally flowering type. The second group contains *bm2LFY* (Bima Brebes) and *bm4LFY* (Tajuk), and was named the inducible flowering type. The third group is made up of *bm3LFY* (Ilokos) and *bm5LFY* (Sumenep) and was called the non-flowering shallot type. The *ShLFY* sequence encodes a putative protein of 363 amino acids, with ~ 99% homology to the *Allium cepa LEAFY* and > 95% homology to *LEAFY* proteins from other higher plants. The s*hLFY* protein in the 5 shallot genotypes showed homology with *FLORICAULA/LFY* protein from referenced *Allium* relatives.

Key words: BLAST analysis, flowering gene, genomic PCR, phylogenetic analysis, sequence aligment

Key findings: The *shLFY* gene confirms the existence of genetic polymorphism among shallot genotypes. The *shLFY* sequence encodes a putative protein of 363 amino acids, and shows homology with *LFY* protein from referenced *Allium*

relatives. Research base on genetic diversity underlying the flowering induction will become an important target in shallot breeding programs.

Manuscript received: December 15, 2017; Decision on manuscript: July 24, 2018; Accepted: August 14, 2018.

Communicating Editor: Dr. Ram Nair

INTRODUCTION

The flowering initiation is transitional process which is indicated by the reduced vegetative growth of plants into reproductive development. Flowering initiation on shallot plants (*Allium cepa* var. *aggregatum*) predicticably have been related to flowering pattern. The control of flowering in plants involves many environmental and intrinsic factors (Yang *et al*., 2016).The induction treatment is required to trigger the flowering process by activating a number of enzymes to improve the flowering gene function. Molecular diversity of flower-regulatory genes determines the plant flowering competency.

Some genes that control the flowering process of plants have been identified successfully. Jeon *et al*. (2000) findings on rice transgenic plants indicated that OsMADS14 plays a role in supporting flowering and determining flowering meristems. In the genome of rice plants identified 3 AP1-like MADS-box genes namely OsMADS14 / RAP1B, OsMADS15 / RAP1A and OsMADS18, which are derivatives of FRUITFULL (FUL) (Litt and Irish, 2003). William *et al*. (2004) show that the transition process of vegetative meristem to flowering is controlled by plant specific transcription factor and meristem identity regulator that is *LEAFY* gene.

LEAFY homologous genes (LFY) plays a central role to control flowering time and floral meristem identity in plants. *LFY* have been studied in angiosperms and gimnosperms (Frohlich and Parker 2000; Rudall *et al*., 2009). Studies on *Arabidopsis* plants indicated the presence of the specific flower-related genes, such as the *APETALA1* gene (Mandel *et al*., 1992), the *FRI/FLC* gene (Caicedo *et al*., 2004), and *LFY COTYLEDONE2* (*LEC2*) (Kim *et al*., 2014). *FLORICAULA/LFY* genes have been studied in *Anthirrinum majus* (Coen *et al*., 1990), *Oryza sativa* (Kyozuka *et al*., 1998), Citrus (Pena 2001), *Zea mays* (Bomblies *et al*., 2003), *Allium sativum* (Rotem *et al*., 2007), *Brassica rafa* subsp. Chinensis (Hong-xia *et al*., 2015) and *Allium cepa* (Yang *et al*., 2016), while Kim *et al*. (2014) had not been elucidated well the *LFY* gene identity in adzuki beans. In *Arabidopsis,* the *LFY* gene is expressed constitutionally throughout its life cycle, *LFY* expression is significantly high in reproductive tissue and tends to be reduce in vegetative tissue (Weigel *et al*., 1992; Blazquez *et al*., 1997).

The *LFY* gene is an important element switches process in the vegetative and reproductive phase (Blázquez *et al*., 1997). In *Arabidopsis thaliana*, the *LFY* gene is the main gene regulated the transition of shoot meristem to flower meristem (Weigel *et al*., 1992; 1997; Yamaguchi *et al*., 2014). The *LFY* gene is known as the primary control of flowering competencies in plant (Putterill *et al*., 2004). The *LFY* gene is a direct connection between the flowering

induction and all the processes related to flowering initiation (Blazquez *et al*., 1997). The floral meristem identity genes expression is determined by the *LFY* gene representation through the emergence of floral meristems (Parcy *et al*., 1998; Siriwardana and Lamb 2012). The *LFY* transcription factor in *Arabidopsis*was recognized as the main regulator for flower development (Engelhorn *et al*., 2014).

The *LFY* gene is a flowering meristem indentity genes together with APETALA1 (AP1) is an important transcription factor for the flowering initiation. The *LFY* gene is an *AP1* regulatory gene that directly stimulates the transcription of the *AP1* gene (Weigel *et al*., 1992; Hempel *et al*., 2000; Chahtane *et al*., 2013). The presence of *LFY* and *AP1* genes expressed in primordia buds indicates the specification on the orchestra switches of shoot buds development to flower formation (Weigel and Nilsson, 1995), upregulation of *LFY* and *AP1* push the flowering mode initiation (Putterill *et al*., 2004). *LFY* and *AP1* organize the expression of genes concerned in the hormone pathways (O'Maoileidigh *et al*., 2014). The formation of floral meristems controlled by the *LFY* gene through auxin activity (MONOPTEROS auxin response factor) increases the expression of the *LFY* gene (Yamaguchi *et al*., 2013). *LFY* stimulates flower primordium initiation, ensuing the floral organ initiation and development, and presenting with polar auxin transport (Yamaguchi *et al*., 2014).

Therefore, identification of *LFY* gene in shallot plant is necessary to promote flower induction and seeds production. *LFY* gene remains uncharacterized in agronomicaly important shallot genotypes. Shallot *LFY* gene remains poorly characterized in shallot thus far. This study objected to identify the molecular diversity of shallot *LEAFY* gene (*shLFY*) sequences on 5 genotypes of shallot collected in Indonesia related to flowering pattern and compare them to homolog sequence of *Allium* relatives. The diversity of *LFY* genes in 5 shallot genotypes can be utilized for development of gene-based markers in order to observe and identify the genetic diversity underlying the flowering induction in shallot. Research base on flower initiation will become an important target in shallot breeding programs.

MATERIALS AND METHODS

DNA extraction

Local genotypes of shallot with different pattern of flowering were selected as plant material. Shallot genotype 'Bentanis' confirms as the naturally flowering shallot, while shallot genotypes Bima Brebes, Ilokos, Tajuk and Sumenep identify as the inducible flowering shallot. Shallot fresh leaf of 5 genotypes (5 weeks after planting) were cutted directly for DNA extraction. DNA isolation method of *Cetyl Trimethyl Ammonium Bromide* (CTAB) Doyle dan Doyle (1987) have been used with modification. Amount of 0.5 g sample was grind with 0.1 g PVP. The crude sample were mixed with 300 µL of extraction buffer (tris-HCl 1 M pH 8.0, NaCl 5 M, EDTA 0.5 M, and CTAB 10%. The samples were incubated at 65 ^{o}C for 60 min, 500 µL amount of *chloroform: isoamyl alcohol* = 24 : 1 was added, and gently mixed for 1 min. Sample were centrifuge at 11000 rpm, temperature 4^{o}C, 10 min. 500 µL of the aqueous supernatant

was carefully transferred to new microtube. DNA precipitation was prepared by mixed the aqueous supernatant with 1 mL of isopropanol (twice of volume) and incubated in the freezer *overnight*. The precipitated DNA were then pelletized by a centrifugation step (11.000 rpm, temperature 4 °C, 10 min), and the supernatant was gently removed. The DNA pellet was washed with 200 µL of ethanol 70% and was centrifugated at 11000 rpm, 5 min. Residual ethanol was removed by maintaining the sample at room temperature for 2 hours. The DNA pellet was finally fixed in100 µL buffer TE.

DNA solution (5 microliters of DNA template combined with 2 µL loading dye) were separated by agarose gel electrophoresis (0.8%), stained with *ethidium bromide* solution (0.01%), and perceived with transilluminator UV light. The DNA concentration was examined on a 1% agarose gel and then estimated using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, Del., USA).

Primer design and DNA amplification

The primer design of *shLFY* gene based on GenBank database of some *Allium* species that is accessed by using the BLAST (*basic local alligment search tool*) program at the NCBI *(National Center for Biotechnological Information)* website of http: //ncbi. nlm.nih.gov/). Plants used for primer designing such as *LFY* flowering gene of *Allium fistulosum* (KF270626.1), *Allium fistulosum* (KF270625.1), *Allium sativum* (AY563105.1), *Allium sativum* (AY672745.1), *Allium cepa* var. *aggregatum* (JX275963.1), *Allium cepa* (JX275962.1), and *Allium sativum* (AY563104.1). The primer has matching a conserved region of the *LFY* genes from *Allium* plants. Alignment of the selected gene using Geneiuos program 9.0.2. Primer pairs selected based online Primer3plus program. Primer pair subsequently consited of forward primer (F/5'-GACCCAACAGACCCCTAC-3') and reverse primer (R/5'-TTAGAACATAGAGGACACAGAG-3') and used to amplify specific nucleotide sequences of *shLFY* genes.

Isolated DNA was applied directly in *polymerase chain reaction* (PCR) amplifications. Amplification reaction was used 1 µM of DNA template, 1 µM of each primer forward and reverse, 6 µl PCR mix *go taq green master* Promega and 5 µl *nuclease water* were combined in a total volume of 14 µL. After an initial denaturation at 94 °C for 5 min, 35 cycles at 94 °C for 30 s, annealing at 53 °C for 1 min, and extension at 72 °C for 25 s were operated. For a terminal elongation, the reaction batches were heated for 5min at 72 °C. Reactions were executed in a thermocycler (*Applied Biosystem 2720 Thermal Cycler).*

PCR product (5 µL of DNA template combined with 2 µL loading dye) were separated by agarose gel electrophoresis (1.2%), 47 min, 50 volt, stained with *ethidium bromide* solution (0.01%), and detected with transilluminator UV light.

Gene structure

Representative gene structure of *shLFY* gene from 5 genotypes *Allium cepa* var. *aggregatum*, Bentanis genotype (*bm1LFY*), Bima Brebes (*bm2LFY*), Ilokos (*bm3LFY*), Tajuk (*bm4LFY*), and Sumenep (*bm5LFY*) were described by matching transcript

sequences with genomic regions. Each regions were aligned manually using Geneious Pro 5.6.6 (Biomatters, USA). Numbers above exons and under introns resembled to their sizes, in base pairs.

Sequence alignment and phylogenetic analysis

Forward and reverse sequence fragments for each shallot genotypes were edited and assembled in contigs using Sequence scanner software v1.0 and CLC Sequence Viewer version 6.7.1. Ambiguous bases were adjusted by visual examination of chromatograms to generate consensus sequences. Consensus sequences for each region were aligned manually using Geneious Pro 5.6.6 (Biomatters, USA). A search of the GenBank sequences, (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was accomplished using the BLASTn and BLASTp program for the purpose of grouping the sequence in each genotypes. Protein sequences of different *shLFY* homologs were repossessed from Genbank (http://www.ncbi. nlm.nih.gov/) and used for phylogenetic tree construction. Phylogenetic analysis were created by using a neighbour-joining method of Mega 6 software. The nearest neighbour-joining method was applied to create the tree. Bootstrap values were derived from 1,000 replicates and the scale bar indicates the length of the branch equivalent to the average substitution base per site.

RESULTS

Genomic DNA and amplification

We found a good quality of genomic DNA of 5 shallot genotypes with the absorbance ratio (A_{260}: A_{280}) between 2.05-2.19. The concentration of genomic DNA is difference among shallot genotypes. Genomic DNA concentration for the Bentanis genotype is 233.9 ng.µl^{-1}, Bima Brebes genotype is 194.2 ng.µl^{-1}, Ilokos genotype is 290.1 ng.µl^{-1}, Sumenep genotype is 96.3 ng.µl^{-1}, whereas the Tajuk genotype is the highest concentration of genomic DNA, 635.4 ng.µl^{-1}.

The specific primer for *shLFY* gene, (F/5'-GACCCAACAGACCCCTAC-3') and (R/5'-TTAGAACATAGAGGACACAGAG-3'), were able to amplify all *shLFY* genomic DNA fragments from 5 shallot genotypes. PCR amplification using a single primer pair (forward and reverse) produced an amplicon band size of ~ 1,200 bp, respectively (Figure 1).

***ShLFY* gene structure**

Advances in molecular DNA technology are able to detect certain genes sequence, including flower-related genes in shallot plants. Sequence amplification of *shLFY* genes from 5 shallot genotypes constructs a simple structure with fragment lengths between 1,161 and 1,253 bp. The fragment contains open reading frame between 1,062 and 1,096 bp. The *shLFY* gene sequence from the 5 genotypes of *Allium* putatively consists of 3 exons and 2 introns. The exon fragments has fragment lengths between 173 and 512 bp, while the intron fragment lengths between 62 and 89 bp (Figure 2). We illustrated the structure of *shLFY* fragment using sample sequence of

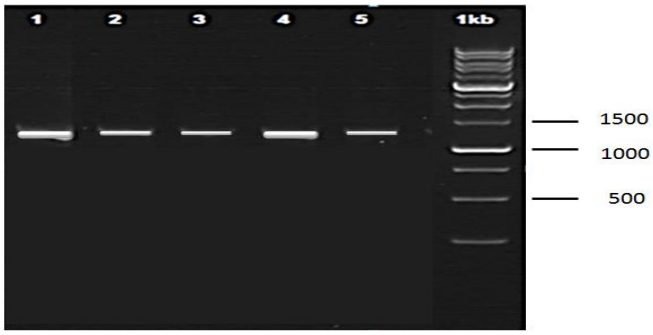

Figure 1. PCR-agarose product visualization of 5 DNA genom of shallot *Allium cepa* var. *aggregatum*, Bentanis genotype (1), Bima Brebes genotype (2), Ilokos genotype (3), Tajuk genotype (4), and Sumenep genotype (5).

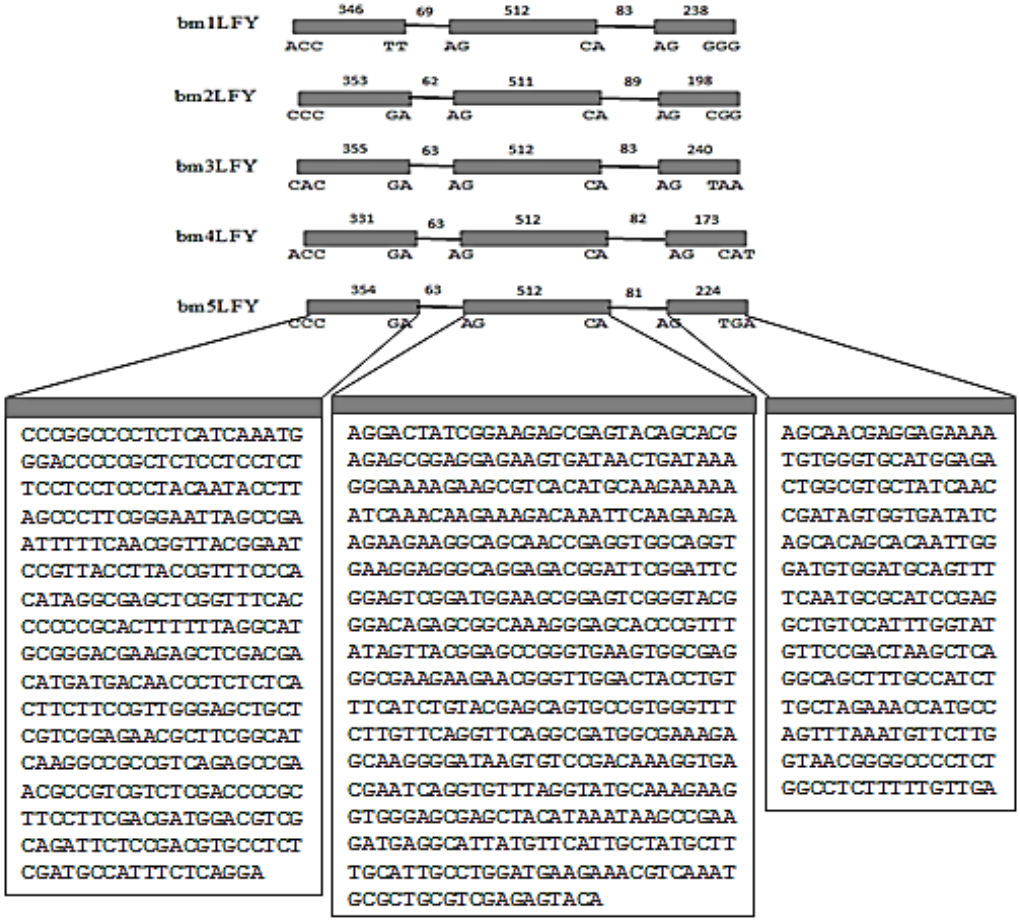

Figure 2. Representative gene structure of *shLFY* gene from 5 genotypes *Allium cepa* var. *aggregatum*, Bentanis genotype (*bm1LFY*), Bima Brebes (*bm2LFY*), Ilokos (*bm3LFY*), Tajuk (*bm4LFY*), and Sumenep (*bm5LFY*). The structure were described by comparing transcript sequences with genomic regions. Numbers above exons and under introns correspond to their sizes, in base pairs.

shLFY of Sumenep genotype (*bm5LFY)*, as presented in Figure 2.

Sequence analysis of 5 fragments of *shLFY* gene

Analysis of sequences homology of *shLFY* gene were compared to accession of *LFY* gene in GenBank database using BLASTn analysis. The BLASTn analysis shows that nucleotide sequences of 5 *shLFY* gene fragments have a high identity (82%-99%) with nucleotide sequences from other *LFY* genes of *Allium* deposited at GenBank, ie with accession number KF270625.1, KF270626.1, JX275962.1, JX275963.1, AY563104.1, AY672745.1, AB829896.1, and AY563105.1 as data listed in Table 1.

A pairwise distance method were estimated to determine the evolutionary divergence between *shLFY* genes of 5 shallot genotypes. The result in Table 2 shows a high genetic similarity among *shLFY* gene. This similarity is indicated by low genetic distance value among 5 *shLFY* gene. The lowest genetic differences (at 0.0044) is the *bm2LFY* (Bima Brebes) and *bm4LFY* (Tajuk) genotypes. While the *bm5LFY* (Sumenep genotype) is close to the *bm3LFY* (Ilokos genotype). The *bm1LFY* gene (Bentanis genotype) has a considerable genetic difference to *shLFY* genes of the other genotype, that is range between 0.0967 and 0.1429. Genes with close genetic distance indicate high genetic compatibility, and locate within the same cluster.

Figure 3 shows the results of phylogenetic analysis between *shLFY* gene from 5 shallot genotypes and other accession *LFY* gene sequence from GenBank database. The phylogenetic tree was constructed by using a neighbour-joining method. There are 3 groups of 5 *shLFY* genes, *bm1LFY* (Bentanis genotype) in the first group, *bm2LFY* (Bima Brebes genotype) and *bm4LFY* (Tajuk genotype) in the second group, *bm3LFY* (Ilokos genotype) and *bm5LFY* (Sumenep genotype) in the third group. The phylogenetic analysis of *shLFY* flowering gene fragments shows that *bm2LFY* (Bima Brebes genotype) and *bm4LFY* (Tajuk genotype) are in the same clade with *Allium cepa* and *Allium cepa* var. *aggregatum.* However, *bm3LFY* (Ilokos genotype) and *bm5LFY* (Sumenep genotype) are in the same clade with *Allium fistulosum.*

Sequence analysis of translated amino acid of 5 shallot genotype *shLFY* gene fragments

We have identified preliminary alignment analysis of the predicted amino acid of *shLFY* gene fragments from 5 shallot genotypes and amino acid of 3 other *Allium* accessions from the GenBank database (Figure 4). Black blocks represented 100% conservation (identical) motifs where there is no amino acid changes occurred. Gray block represented 75% conservation motifs, and white blocks indicated unconserved motifs where there is changes in amino acid residues.

The sequence of *shLFY* amino acids shows a sequence identity with other accession *LFY*amino acid sequence from GenBank database.The analysis of predicted amino acids shows that *shLFY* genes from 5 shallot genotypes have similar identity with *LFY* mRNA from *Allium cepa* (identity 99%), *Allium cepa* var. *aggregatum* (98%), and *Allium fistulosum*

(96%). This observation is exemplified by the data listed in Table 3.

Phylogenetic analysis were constructed by neighboring joining method shows that the predicted amino acid of *shLFY* genes clustered into 3 groups. The first group was a group of*bm1LFY* (Bentanis), the second group was *bm2LFY* (Bima Brebes) and *bm4LFY* (Tajuk), which were shared the same clade with *Allium cepa* AFR67540.1, *Allium cepa* var. *aggregatum* AFR67541.1, and the third group was consisted *bm3LFY* (Ilokos), and *bm5LFY* (Sumenep), which were shared the same clade with *Allium fistulosum* AGU12799.1 (Figure 5).

Table 1. Sequence identity of nucleotide sequences of *shLFY* gene fragments with related accessions of *Allium* in GenBank database.

Accession information	Genbank accession no.	*shLFY* genotype	Quiry (%)	Identity (%)	E-value
Allium cepa	JX275962.1	*bm1LFY*	90	81	0
		bm2LFY	98	99	0
		bm3LFY	100	93	0
		bm4LFY	100	99	0
		bm5LFY	97	95	0
Allium cepa var. *aggregatum*	JX275963.1	*bm1LFY*	92	90	0
		bm2LFY	98	99	0
		bm3LFY	100	93	0
		bm4LFY	100	98	0
		bm5LFY	97	94	0
Allium fistulosum	KF270625.1	*bm1LFY*	91	89	0
		bm2LFY	98	97	0
		bm3LFY	100	96	0
		bm4LFY	100	97	0
		bm5LFY	99	97	0
Allium fistulosum	KF270626.1	*bm1LFY*	91	90	0
		bm2LFY	98	97	0
		bm3LFY	100	99	0
		bm4LFY	100	96	0
		bm5LFY	99	99	0
Allium sativum	AY563105.1	*bm1LFY*	44	82	7e-52
		bm2LFY	47	88	0
		bm3LFY	46	89	0
		bm4LFY	49	88	1e-180
		bm5LFY	46	88	1e-180

Table 2. Estimates of evolutionary divergence between *shLFY* genes of 5 shallot genotypes. The base substitutions per site of 5 shallot genotypes sequences number were analyzed using the maximum composite likelihood model (Tamura *et al*., 2004). Evolutionary analyses of *shLFY* genes of 5 shallot genotypes were run in MEGA6 (Tamura *et al*.,2013).

	bm1LFY	*bm2LFY*	*bm3LFY*	*bm4LFY*	*bm5LFY*
bm1LFY					
bm2LFY	0.0967				
bm3LFY	0.1429	0.0577			
bm4LFY	0.0997	0.0044	0.0606		
bm5LFY	0.1385	0.0547	0.0360	0.0576	

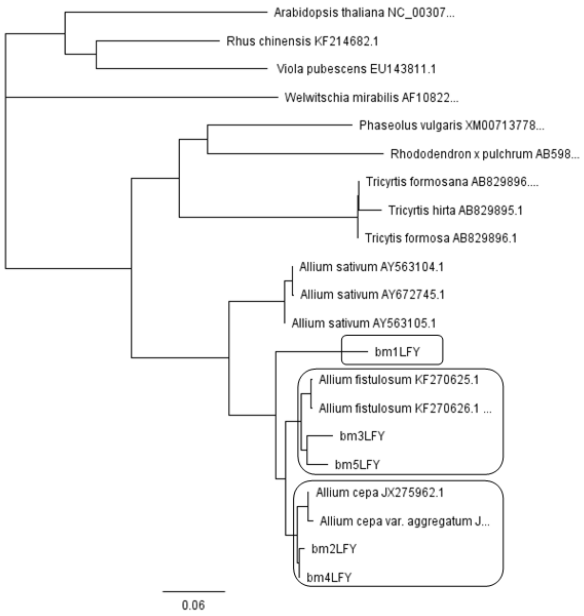

Figure 3. Phylogenetic analysis of *shLFY* flowering gene fragments from 5 shallot genotypes.The phylogenetic tree was constructed by using a neighbour-joining method of Mega 6 software. The branch number is the percentage of repetition using the bootstrap method (1000 bootstrap). The 0.06 scale under the tree represents the length of the branch equivalent to the average substitution base per site. Bentanis genotype (*bm1LFY*), Bima Brebes (*bm2LFY*), Ilokos (*bm3LFY*), Tajuk (*bm4LFY*), and Sumenep (*bm5LFY*).

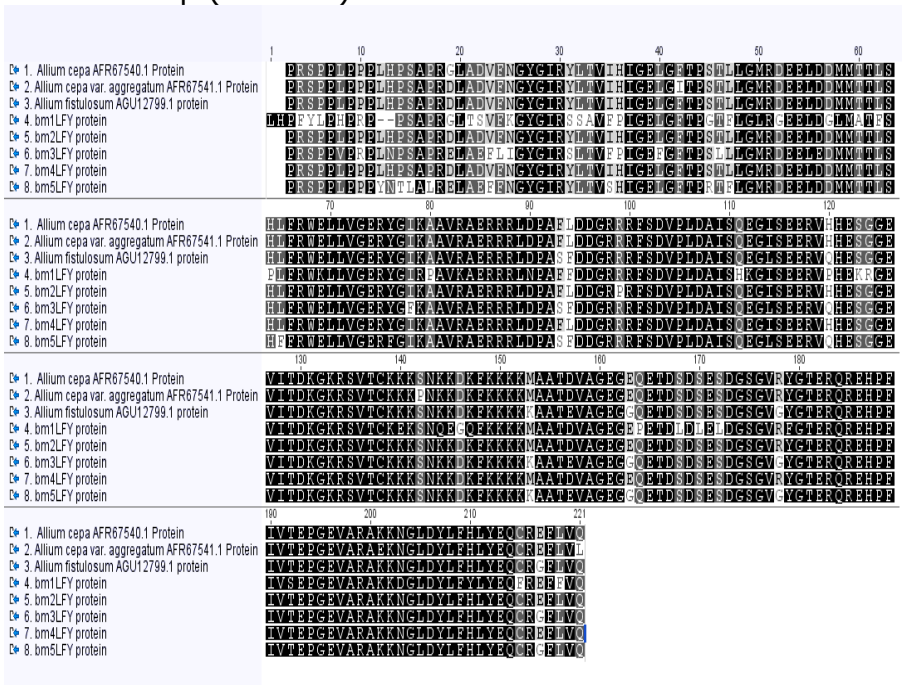

Figure 4. Alignment of the conserved *shLFY* genes from 5 shallot genotypes and 3 accessions of *Allium* plants. The identical amino acid residues were indicated with black background, while 75% conservation was marked as gray.

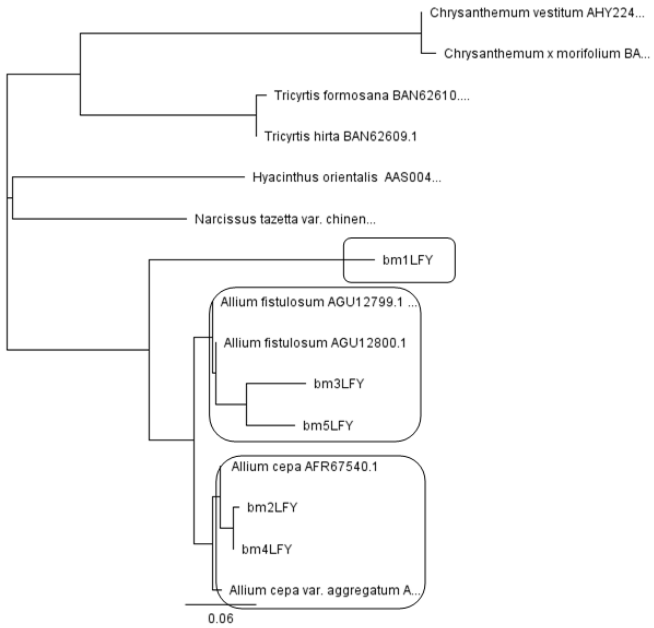

Figure 5. Phylogenetic of the sequence variation of the predicted amino acid of the *shLFY* gene from 5 shallot genotypes and homologs. The tree was constructed by using a neighbour-joining method of Mega 6 software. The number on the branching axis is the percentage of repetition using the bootstrap method (1000 bootstrap). The scale under the tree (0.06) represents the length of an equivalent branch with an average substitution rate per site.

Table 3. Sequence identity between translated amino acid of *shLFY* gene fragments from 5 shallot genotypes with related accessions of *Allium* in GenBank database.

Accession no.	Accession information	Quiry (%)	Identity (%)	E-value
AFR67540.1	*Allium cepa*	99	99	0
AFR67541.1	*Allium cepa* var. *aggregatum*	99	98	0
AGU127999.1	*Allium fistulosum*	99	96	0
AGU12800.1	*Allium fistulosum*	98	96	0
ADR78683.1	*Narcissus tazetta* var. *chinensis*	96	72	4e-176
AAS00458.1	*Hyacinthus orientalis*	95	68	5e-166
BAN62609.1	*Tricyrtis hirta*	98	64	1e-151
BAN62610.1	*Tricyrtis formosana*	98	63	4e-151
AHY22450.1	*Crysanthemum vestivum*	97	59	5e-150
AHY22448.1	*Crysanthemum x morifolium*	98	59	5e-150

DISCUSSION

DNA isolation is the first step in the study of DNA sequences from genomic DNA populations addressed for analysis of gene structure and polymorphism. DNA isolation techniques vary on each species to produce the desired target DNA. The CTAB buffers that have been used for DNA extraction of the 5 shallot genotypes were able to generate a high quality of whole genome products. A high quality of whole genome isolation products represent with a thick and clear DNA band when

visualized using a gel image of electrophoresis. The high molecular weight of the DNA should have clean bands, while the smear band at the bottom indicated the isolate impurity. The smear band indicates the presence of RNA contamination that blocks the migration process. The quality of preference genomic DNA was indicated by the absorbance ratio value (A_{260}: A_{280}) might be between 1.8-2.0. Lower absorbance values (A_{260}: A_{280}) indicated poor genomic DNA quality due to protein contamination. Effective and efficient DNA isolation and purification techniques could produce good and less DNA contamination. The quality of genomic DNA is signified by clear band pattern. PCR is a process of enzymatic synthesis to amplify DNA nucleotides outside the living cell, and can increase the number of DNA sequences multiplication in million copies. The specific PCR amplification of gene sequence should show the presence of a single DNA band of the targeted size based on previous information. The PCR process is a recurrent cycle process including denaturation, annealing and extension by DNA polymerase enzymes. A specific oligonucleotide primer pair is used to make a hybrid with the 5-tip 'toward the 3-end' of the target DNA strand and duplicates the targeted sequence in certain length. PCR reactions are very sensitive, using asmall amount of DNA template, and able to multiply very large amounts of DNA (Jonas, 2003).

The *shLFY* gene is successfully cloned from a single plant isolate, that migrated as a single band during agarose gel electrophoresis. Figure 1 shows the results of sequence amplification of *shLFY* genes from 5 shallot genotypes.The sequence amplification has an open reading frame of 1,161-1,253 bp, and has been registered in GenBank (access number: KY985382, KY985383, and KY985384). The amplified products of *shLFY* gene consider similar to amplified product of *LFY* gene in another plants. Yang *et al*. (2016) reported that amplified product of *LFY* gene of *Allium cepa* (*AcLFY*) contains a 1,119 bp open reading frame. Chaidamsari *et al*. (2009) obtained amplified product for *LFY* gene forcocoa contains a 1,200 bp, and Hong-Xia *et al*. (2015) obtained amplified product for *BrcLFY* of 1,260 bp.

Figure 2 shows that the *shLFY* gene of 5 shallot genotypes has exons and introns in a conserved position with the *LFY* gene of previous reported *Allium cepa* from reference. Exon3 lengthis shorter than exon1 and exon2. While intron1 is shorter than intron2. The *LFY* gene has a relatively simple structure (Peng *et al*., 2010) which has 3 exons and 2 introns, with the structure and location of the exact introns domain (Frohlich and Meyerowitz, 1997; Ma *et al*., 2008). Gene structure of 5 genotypes shallot are consistent with the comparison of the size of introns and exons of LFY gene structure. In the process of gene expression, the intron part will be removed (splicing) from the cDNA before the translation stage. The high sequence identity of *shLFY* gene fragments indicated that the obtained *shLFY* gene has similar function as the *LFY* genes of other plant species. The *LFY* gene function is as an important element in the vegetative and reproductive phase transition process. The *LFY* gene is a meristem identity gene, which regulates the flowering differentiation process of the inflorescence branch (William *et al*.,

2004). In *Arabidopsis* plants the flowering process is consistent with upregulation of the *LFY* gene (Blázquez *et al*., 1997).

Full-length sequence on *shLFY* gene of 5 genotypes shallot revealed that the *shLFY* gene to be more closely related to *LFY* gene in *Allium* plant. The results showed that the *shLFY* gene fragment of 5 shallot genotypes comprised a high sequence identity (99%) with *LFY* gene from *A. cepa, A. cepa* var. *aggregatum* and *A. fistulosum*, while the sequence identity with *A. sativum* was 82-89%, as listed in Table 1. This suggested that the *shLFY* gene sequence of 5 shallot genomes has similar structural and functional to *LFY* genes present in the *Allium* species. Yang *et al*. (2016) showed that *LFY* gene in *Allium cepa* homologs to shallot (97%), Chinese narcissus (74.9%), MAO point grass (70.8%), and more than 60% highly conserved between different plants. The *LFY* gene is specific transcription factor that plays an important role in the development of plant meristem identity. The *LFY* gene stimulates the initiation of primordial interest by activating the downstream AUXIN RESPONSE FACTOR5/ MONOPTEROS (Yamaguchi *et al*., 2014). Plants with *LFY* gene activity increased sensitivity response to auxin transport damage.

The evolutionary divergence between *shLFY* genes of 5 shallot genotypes estimated that there were a total of 1,160 positions in the final dataset. All positions containing gaps and missing data were eliminated. Bima Brebes genotype (*bm2LFY*) and Tajuk genotype (*bm4LFY*) are in the same clade with the lowest genetic difference at 0.0044. These genotypes are typically as the induceble flowering of shallot plants. While the Sumenep genotype (*bm5LFY*) is closely to the Ilokos genotype (*bm3LFY*), typically as non-flowering of shallot plants. The Bentanis genotype (*bm1LFY*) considers to have genetic difference to *shLFY* genes of the other genotype. This *shLFY* gene of Bentanis genotype is suggested to have an important control to drive the flowering mode initiation.

The *shLFY* gene clustering information implied the diversity of *shLFY* genes in *Allium* plants and related organism. There is *bm1LFY* (Bentanis genotype) in the first group reffered as naturally flowering pattern. In the second group, there are *bm2LFY* (Bima Brebes genotype) and *bm4LFY* (Tajuk genotype) reffer as sensitive inducible flowering type. While in the third group are *bm3LFY* (Ilokos genotype) and *bm5LFY* (Sumenep genotype) reffer as non-sensitive flowering type as shown in Figure 3. *ShLFY* gene including in the second group may being active due to exogenous inducement such as vernalization treatments. Ream *et al*. (2014) classified 13 accessions of *Brachypodium distachyon* into 6 flowering ability groups, ie extremely rapid flowering, rapid flowering, intermediate rapid flowering, intermediate delayed flowering, delayed flowering, and extremely delayed flowering. Identification of *shLFY* gene on 5 shallot genotypes shows that flowering pattern in shallot is related to *LFY* gene function. Homologous flowering genes in 5 genotypes shallot are often clustered on the same group, and their exon-intron architectures and key domains are generally conserved. *LFY* sequence in 5 shallot genotypes revealed the genes polymorphism.

The results of the sequence analysis implied that there are conserved motifs and non-

conservation motifs based on *shLFY* genes alignment analysis. *ShLFY* encodes a putative protein of 363 amino acids. Conserved motifs proved that the predicted amino acid of *shLFY* has similar sequences with amino acid of other *Allium*. The deletions on certain sites along the *LFY* gene sequence affected the amount of amino acid residues differences. Identified *LFY* genes have conserved structures and functions (Lu *et al*., 2011). Domain analysis may play a more important role in prediction of flowering orthologous proteins in monocot crop species. Yang *et al*. (2016) showed that *AcLFY* encodes a putative protein of 372 amino acids, with ~ 70% homology to the daffodils *LEAFY* and > 50% homology to *LEAFY* proteins from other higher plants. *LEAFY* protein correlated with controlling regulation of vegetative phase transition process to the reproductive phase of the plant (Weigel *et al*., 1992; Weigel and Nilsson 1995).

The phylogenetic relationship between the predicted amino acid sequence of *shLFY* gene from 5 shallot genotypes with other *LFY-like* proteins of other accession deposited on the GenBank database was constructed. All of the sequences from the same taxa are clustered together. The amino acids analysis indicated that *shLFY* genes from 5 shallot genotypes have identical identity with *LFY* mRNA from *Allium cepa* (identity 99%), *Allium cepa* var. *aggregatum* (98%), and *Allium fistulosum* (96%). The translated amino acid of *shLFY* from 5 genotypes of shallot has a closed identity to the mRNA/amino acid from other species, such as *Narcissus tazetta* var. *chinensis* and *Hyacinthus orientalis*that is exemplified in Table 3. This indicated that there are significant similarity between *shLFY* gene and the related *LFY* in other species. It shows that the predicted amino acid of *shLFY* gene also contributes to the high identity with *LFY* gene protein deposited on the GenBank. Obtained predicted amino acid of *shLFY* has a higher protein homology revealed that *LEAFY* genes are highly conserved between different plants. This results consistent to the result from Yang *et al*. (2016) that further verified a closer genetic relationship between them. In *Allium cepa AcLFY* has 97% homology with shallot, 74.9% homology with *Chinese narcissus,* 70.8% homology MAO point grass, and >60% homology with *A. thaliana*, mango, and walnut (Yang *et al*., 2016).

Nilsson *et al*. (1998) findings revealed that there is a class of flowering-time genes that affect the activity of the *LFY* gene, and there are other gene classes that greatly affect the transcriptional induction of *LFY* genes. The translated amino acid of the *shLFY* gene from 5 shallot genotypes represented identical to the amino acids mRNA reference in other plants based on open source reference frames (ORF) in European molecular laboratory protein (Pfam) database and were confirmed with BLASTx NCBI software (Figure 4).

The *FLORICAULA/LFY* protein is included in the SAM protein domain, which consists of various proteins regulated plant development. Protein *FLORICAULA/LFY* is a floral meristem identity protein. Mutations in this protein sequence affect the development of leaves and flowers. The results of this study indicated that the *shLFY* gene obtained from 5 shallot genotypes is identical with*LFY/FLO* gene family. *ShLFY* gene showed a high homology with *LFY*gene

from *Allium* plant and monocotyl plant such as *Tricyrtis formosana*. The results of Weigel *et al*. (1992) showed that the *LFY* gene in *Arabidopsis* is a homolog of the *FLORICAULA* gene that controls the flowering meristem identity of *Antirrhinum majus*, and Yang *et al*. (2016) that *AcLFY* putative amino acid sequence has typical structure characteristics of the *LEAFY* (FLO) family, with a proline rich region at the N-terminal, a leucine zipper structure, and a central acid region. Identification of *shLFY* gene diversity from 5 shallot genotypes revealed that *shLFY* gene correlates to flowering pattern in shallot. The results of *shLFY* sequence alignment analysis clearly identified the naturally flowering genotype, sensitive inducing flowering genotypes, and non-sensitive inducing flowering genotypes. The presence of *shLFY* sequences obtained by genomic DNA amplification from 5 shallot genotypes indicates the nucleotide and amino acid diversity between the *shLFY* genes within 5 shallot genotypes, and also relates organism. Overall, *in-silico* analysis based on sequence information, and homology modeling of *shLFY* protein revealed that it has considerable similarity to orthologous *LFY* proteins in plant species. These data indicates that we successfully isolated and sequenced the shallot *LFY* homolog. This is the first study to identify and characterize *LFY* gene in shallot *Allium cepa* var *aggregatum*. This molecular identify will be a great resource for the scientific community interested in studying shallot and other members of *Allium* plant in corresponding to flowering gene competency.

ACKNOWLEDGEMENT

This researchs is supported by grants from National Innovation System Research Insentive 2016, Ministry of Research, Technology, and Higher Education, Indonesia.

REFERENCES

Blázquez MA, Soowal LN, Lee I, Weigel D (1997). *LFY* expression and flower initiation in *Arabidopsis*. *Develop.* 124: 3835–3844.

Bomblies K, Wang RL, Ambrose BA, Schmidt RJ, Meeley RB, Doebly J (2003). Duplicate Floricaula / *LFY* homologs zfl1 and zfl2 control inflorescence architecture and flower patterning in maize. *Develop.* 130: 2385–2395.

Caicedo AL, Stinchcombe JR, Olsen KM, Schmitt J, Purugganan MD (2004). Epistatic interaction between the *Arabidopsis FRI* and *FLC* flowering time genes generates a latitudinal cline in a life history trait. *Proceed. Natl. Acad. Sci. USA* 101: 15670-15675.

Chahtane H, Vachon G, Masson ML, Thévenon E, Périgon S, Mihajlovic N, Kalinina A, Michard R, Moyroud E, Monniaux M, Sayou C, Grbic V, Parcy F, Tichtinsky G (2013). A variant of *LFY* reveals its capacity to stimulate meristem development by inducing RAX1. *The Plant J.* 74: 678–689.

Chaidamsari T, Hayati R, Syarief A. Anwar A, Santoso D (2009). Cloning the LFY gene of cocoa from the active flower bearing tissue. *Menara Perkebunan* 77(2): 79-88 (*in Indonesian languange*).

Coen ES, Romero JM, Doyle S, Elliott R, Murphy G, Carpenter R (1990). Floricaula: a homeotic gene required for flower development in *Antirrhinum majus*. *Cell* 63: 1311–1322.

Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19: 11–15.

Engelhorn J, Moreau F, Fletcher JC, Carles CC (2014). ULTRAPETALA1 and LEAFY pathways function independently in specifying identity and determinacy at the *Arabidopsis* floral meristem. *Ann. Bot.* 114: 1497–1505.

Frohlich MW, Meyerowitz EM (1997). The search for flower homeotic gene homologs in basal angiosperms and Gnetales: a potential new source of data on the evolutionary origin of flowers. *Int. J. Plant Sci*. S131–S142.

Frohlich MW, Parker DS (2000). The mostly male theory of flower evolutionary origins: from genes to fossils. *Syst. Bot.* 25: 155–170.

Hempel FD, Welch DR, Feldman LJ (2000). Floral induction and determination: where is flowering controlled? *Trends in Plant Sci.* 5(1): 17-21.

Hong-Xia S, Chao F, Li-Hui Y, Lei-Ping H, Xiao-Yong X, Mei-Lan L (2015). Cloning and expression analysis of *LFY* homologue in Pak Choi (*Brassica rapa* subsp. *chinensis*). *Biotechnol. Biotechnolog. Equip.* 29(6): 1035-1042.

Jeon JS, Lee S, Jung KH, Yang WS, Yi GH, Oh BG, An G (2000). Production of transgenic rice plants showing reduced heading date and plant height by ectopic expression of rice MADS-box genes. *Mol. Breed.* 6: 581–592.

Jonas NL (2003). PCR: Principles, procedures, and parameters. *In*: Theophilus BDM, Ralpley R (*eds)*. Methods in Moleculer Biology: PCR Mutation Detection Protocols. New Jersey: *Humana Press* 37-46.

Kim MY, Kang YJ, Lee T, Lee S (2014). Circadian clock and photoperiodic flowering genes in adzuki bean (*Vigna angularis* [Willd.] Ohwi & H. Ohashi). *Plant Genet Resour-C* 12(S1): 49–53.

Kyozuka J, Konishi S, Nemoto K, Izawa T, Shimamoto K (1998). Down-regulation of RFL, the FLO/*LFY* homolog of rice, accompanied with panicle branch initiation. *Proceed. Natl. Acad. Sci., USA* 95: 1979–1982.

Litt A, Irish VF (2003). Duplication and diversification in the *APETALA1*/*FRUITFULL* floral homeotic gene lineage: Implications for the evolution of floral development. *Genet*. 165: 821–833.

Lu LL, Sun GM, Liu YG (2011). Progress of *LFY* homologous gene study. *Acta Bot. Boreali-Occidentalia Sin.* 31(1): 0197-0203.

Ma YP, Fang XH, Chen F, Dai SL (2008). DFL, a FLORICAULA/*LFY* homologue gene from *Dendranthema lavandulifolium* is expressed both in the vegetative and reproductive tissues. *Plant Cell. Reports* 27: 647–654.

Mandel MA, Gustafson-Brown C, Savidge B, Yanofsky MF (1992). Molecular characterization of the *Arabidopsis* floral homeotic gene *APETALA1*. *Nature* 360: 273–277.

Nilsson O, Lee I, Blázquez MA, Weigel D (1998). Flowering-time genes modulate the response to *LEAFY* activity. *Genet.* 150: 403–410.

O'Maoileidigh DS, Graciet E, Wellmer F (2014). Gene networks controlling *Arabidopsis thaliana* flower development. *New Phytol.* 201: 16–30.

Parcy F, Nilsson O, Busch MA, Lee I, Weigel D (1998). A genetic framework for floral patterning. *Nature* 395: 561–566.

Pena (2001). Constitutive expression of *Arabidopsis LFY* or *APETALA1* genes in citrus reduces their generation time. *Nature Biotechnol.*19: 236-267.

Peng Y, Wei Y, Baum BR, Yan Z, Lan X, Dai S, Zheng Y (2010). Phylogenetic inferences in Avena based on analysis of FL intron2

sequences. *Theor. Appl. Genet*. 121: 985–1000.

Putterill J, Laurie R, Macknight R (2004). It's time to flower:the genetic control of flowering time. *Bio-Essays* 26: 363–373.

Rotem N, Shemesh E, Peretz Y, Akad F, Edelbaum O, Rabinowitch HD, Sela I, Kamenetsky R (2007). Reproductive development and phenotypic differences in garlic are associated with expression and splicing of *LFY* homologue ga*LFY*. *J. Expt. Bot.* 1-9.

Rudall PJ, Remizowa MV, Prenner G,Prychid CJ, Tuckett RE, Sokoloff DD (2009). Nonflowers near the base of extant angiosperms? Spatiotemporal arrangement of organs in reproductive units of hydatellaceae and its bearing on the origin of the flower. *Am. J. Bot.* 96: 67–82.

Siriwardana NS, Lamb RS(2012). The poetry of reproduction: the role of *LFY* in *Arabidopsis thaliana* flower formation. *Int. J. Develop. Biol.* 56: 207–221.

Tamura K, Nei M, Kumar S(2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceed. Natl. Acad. Sci.,* USA 101: 11030-11035.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30: 2725-2729.

Weigel D, Alvarez J, Smyth DR,Yanofsky MF, Meyerowitz EM (1992). *LFY* controls floral meristem identity in *Arabidopsis*. *Cell* 69: 843–859.

Weigel D, Nilsson O (1995). A developmental switch sufficient for flower initiation in diverse plants. Nature 377: 495–500.

William DA, Su Y, Smith MR, Lu M, Baldwin DA, Wagner D (2004). Genomic identification of direct target genes of LEAFY. *Proceed. Natl. Acad. Sci.,* USA 101(6): 1775–1780.

Yamaguchi N, Wu M, Winter CM, Wagner D (2014). *LFY* and polar auxin transport coordinately regulate *arabidopsis* flower development. *Plants* 3: 251-265.

Yamaguchi N, Wu MF, Winter CM, Berns MC, Nole-Wilson S, Yamaguchi A, Coupland G, Krizek BA, Wagner D (2013). A molecular framework for auxin-mediated initiation of flower primordia. *Develop. Cell* 24: 271–282.

Yang C, Ye Y, Song C, Chen D, BJiang B, Wang Y (2016). Cloning and functional identification of the AcLFY gene in *Allium cepa*. *Biochem. Biophy. Res. Commun.* 473: 1100-1105.