J. Trop. Agric. and Fd. Sc. 46(2)(2018): 75 – 89

Isolation and phylogenetic characterisation of LdSVP, SHORT VEGETATIVE PHASE (SVP) homologous gene from *Lansium domesticum*

[Pemencilan dan pencirian filogenetik LdSVP, gen homolog *SHORT VEGETATIVE PHASE* (SVP) daripada *Lancium domesticum*]

A.C.K. Ling[1], L. Rozano[1] and U.K. Abu Bakar[2]

[1]Biotechnology and Nanotechnology Research Centre, MARDI Headquarters, Persiaran MARDI-UPM, 43400 Serdang, Selangor, Malaysia
[2]MARDI Directorate, MARDI Headquarters, Persiaran MARDI-UPM, 43400 Serdang, Selangor, Malaysia

Abstract

SHORT VEGETATIVE PHASE (SVP), encoding a MADS-domain protein, is one of the central floral repressors in *Arabidopsis thaliana* besides FLOWERING LOCUS C. It is expressed in vegetative tissues and floral primordia but not in developed flowers, and acts in a dose-dependent manner. SVP homologous gene from *Lansium domesticum*, namely LdSVP, was isolated from young leaf tissues of 3-month old plants. Deduced amino acid sequence of LdSVP contains MADS-box, I-box, K-box and C-terminal domains, conforming to MIKC-type protein, and highly similar to SVP. LdSVP cDNA consists of 684 nucleotides of coding sequence, 226 nucleotides of 5' untranslated region (UTR) and 266 nucleotides of 3' UTR. Phylogenetic study of SVP and SVP homologous proteins of 23 plant species indicates that LdSVP is closely related to PtSVP, a SVP homologous protein of trifoliate orange (*Poncirus trifoliata*), and both share the same ancestor. It is phylogenetically further from PtAGL24, a MADS-box gene from *P. trifoliata* that is closely related to AGAMOUS-LIKE 24 (AGL24) of *Arabidopsis*, than PtSVP. Therefore, LdSVP is more likely to take on the functions of SVP than AGL24. Protein phylogenetic analysis and domain organisation study showed that LdSVP may have a similar function to PtSVP in flowering time regulation. PtSVP is highly expressed in dormant tissues and vegetative meristems, and induces late flowering when transformed into wild-type *A. thaliana*, suggesting that LdSVP may be involved in seasonal flowering of duku, and SVP homologous genes may be the main floral repressors in seasonal flowering regulation of tropical plants. Although SVP homologues have evolved among plant species, the functional domains are well conserved between *Arabidopsis* and *L. domesticum*.

Keywords: MADS-box, flowering, duku, tropical plant

Introduction

Durian (*Durio* sp.) cempedak (*Artocarpus integer*), various *Lansium domesticum* (duku, langsat and dokong), mango (*Mangifera indica*), mangosteen (*Garcinia mangostana*), jackfruit (*Artocarpus heterophylla*) and rambutan (*Nephelium lappaceum*) are popular local fruits in

Article history
Received: 19.04.2017
Accepted: 8.09.2017

Authors' full names: Adrain Ling Chieng Kuang, Lina Rozano and Umi Kalsom Abu Bakar
E-mail: adrain@mardi.gov.my

Malaysia and they are all seasonal. Seasonal fruits have the disadvantage of inconsistency of fruit supply as the supply will be limited to none during off-season. However, over-supply during fruiting season is common which in turn leads to a drop in price and reduces the profit. To spur on and expand the seasonal fruit industry to one that is economically viable, much has to be done to address the seasonality problem.

Lansium domesticum Correa is a tropical tree species that belongs to the family Meliaceae. It can still be found in the wild and is one of the important indigenous fruit trees of Malaysia that is widely cultivated (Hanum et al. 2013; Yee and Rao 2013). Malaysia, Thailand, the Philippines and Indonesia are the major *L. domesticum* producing countries (Hanum et al. 2013). In Malaysia, langsat and duku are two well-known commercial varieties. *Lansium domesticum* flower has receptive stigma with glandular surface and ample vascular tissues but pollen degeneration is common. Therefore, the fruit develops apomictically as pollen degenerates completely at anthesis (Yee and Rao 2013). The species being an octoploid with 72 chromosomes may also contribute to the pollen being unviable. Parts of *L. domesticum* have been reported to have medicinal properties such as anticancer (Manosroi et al. 2013) and antimalarial (Yapp and Yap 2003).

Flowering in model plant *Arabidopsis thaliana* has been extensively studied. *Arabidopsis* flowering has been postulated as a default developmental programme that needs to be suppressed by one or more floral repressors to enable vegetative growth (Hartmann et al. 2000). Hence, floral repressor genes seem to determine the duration of the vegetative phase. FLOWERING LOCUS C (FLC) is a central flowering regulator in *Arabidopsis* and it is a floral repressor that regulates flowering antagonistically with CONSTANS (CO), a floral activator (Michaels and Amasino 1999; Samach et al. 2000; Lee and Lee 2010). The underlying molecular processes of the suppression of FLC during vernalisation are of epigenetic nature which enables vegetative to reproductive transition (Michaels and Amasino 1999; Henderson and Dean 2004; Searl et al. 2006). However, the extent of FLC conservation outside of Brassicaceae is rare with the one notable exception of BvFL1 (Reeves et al. 2007) of sugar beet (*Beta vulgaris*) from the family Amaranthaceae.

SHORT VEGETATIVE PHASE (SVP) is another central floral repressor in *A. thaliana* that encodes a MADS domain protein (Hartmann et al. 2000; Li et al. 2008). SVP loss-of-function mutant showed early flowering (Méndez-Vigo et al. 2013). It acts in a dose-dependent manner, and it is expressed in vegetative tissues and floral primordia but not in developed flowers. It is of particular interest in fruit tree breeding due to its role in the maintenance of juvenility (Hanke et al. 2007). SVP also controls flowering time in response to ambient temperature (Lee et al. 2007), an environmental cue that may also be responsible for the seasonality of tropical fruit trees. In *Arabidopsis*, ambient temperature is perceived through a thermosensory pathway involving FCA and FVE (Blázquez et al. 2003). SVP mediates temperature signalling and functions primarily within the thermosensory pathway, downstream from FCA and FVE. It also acts independently of FLC in regard to the transcriptional level within the thermosensory pathway (Lee et al. 2007). SVP-like genes in grapevine were speculated to play a similar role in ambient temperature signalling due to the involvement of high temperature in the promotion of flowering transition in grapevine latent buds (Díaz-Riquelme et al. 2009). Since the conservation of FLC outside of Brassicaceae is uncommon, SVP may well be the main floral repressor in the flowering pathways of most seasonal fruit trees. AGAMOUS-LIKE 24 (AGL24), encoding a MADS-box

transcription factor, is a close homologue of SVP. However, AGL24 promotes flowering and is repressed by the floral meristem identity genes LEAFY and APETALA1 (Yu et al. 2002; Yu et al. 2004).

The fundamental mechanisms of flowering time genes may be different between perennial plants and *Arabidopsis* as they have different flowering characteristics, such as the long juvenile phase and seasonal flowering (Sun et al. 2016). A partial sequence of a MADS-box gene closely related to SVP was identified and obtained from a transcriptome next-generation sequencing (NGS) database of *L. domesticum* in a previous study on duku flowering (unpublished). In view of the importance of SVP in flowering regulation, it was our interest to isolate and characterise the SVP homologous gene from *L. domesticum* based on the partial transcript sequence. The sequence of SVP homologous gene from *L. domesticum*, namely LdSVP, was analysed using bioinformatics tools for the prediction of its putative functions.

Materials and methods

Plant materials and sample collections

Duku (*L. domesticum*) seedlings of about 3 months old were obtained from the Gene Bank and Seed Centre, MARDI. The seedlings were propagated from mother plants at MARDI Sintok. New fully expanded young leaves were collected for total ribonucleic acid (RNA) extraction. Petioles and mid veins were removed before the leaves were frozen in liquid nitrogen and stored at –80 °C.

Primer design and sequence alignment

Gene-specific primers were designed based on a partial SVP homologous sequence obtained from a transcriptome NGS database of *L. domesticum* (unpublished data) using Primer3 (Rozen and Skaletsky 2000). The designed primers were 20 nucleotides long with melting temperature (Tm) in the range of 67 – 70 °C. Sequence alignment was performed using ClustalW (Higgins and Sharp 1988; Thompson et al. 1994; Larkin et al. 2007).

RNA extraction, first-strand cDNA synthesis and isolation of full-length LdSVP

Total RNA was extracted per reported method (Chang et al. 1993; Accerbi et al. 2010) with some modifications. Before extraction, mortars, pestles and spatulas were wrapped in aluminium foil and baked at 180 °C for 5 h. RNase-free pipette tips and microcentrifuge tubes were used. All solutions were prepared with sterile RNase-free water. The sampling of leaves were placed in a mortar, frozen with liquid nitrogen and ground to fine powder with a pestle. Ground tissue was transferred to 2 ml tubes at 100 mg tissue per tube. To each tube, 600 μl of 65 °C preheated extraction buffer [2% (w/v) hexadecyltrimethylammonium bromide (CTAB); 2% (w/v) polyvinylpyrrolidone (PVP-40); 100 mM tris(hydroxymethyl) aminomethanehydrochloride (Tris-HCl), pH8.0; 25 mM ethylenediaminetetraacetic acid (EDTA), pH 8.0; 2 M sodium chloride (NaCl); 0.05% (w/v) spermidine; 2% (v/v) β-mercaptoethanol] was added before vortex mixing for 2 min. An equal volume of acid phenol:chloroform:isoamyl alcohol (125:24:1) was then added to the mixture and incubated at room temperature for 20 min with 3 times of 2 min vortex mixing at 5 min intervals during the incubation period. The mixture was then centrifuged at 10,000 x g for 10 min at 4 °C. The supernatant was transferred to a 1.5 ml tube before adding an equal volume of chloroform:isoamyl alcohol (24:1), mixed and centrifuged at 10,000 x g for 15 min at 4 °C. The aqueous layer was collected in a new 1.5 ml tube, mixed with a third volume of 8 M lithium chloride (LiCl) and kept at –20 °C overnight. The mixture was subjected to 30 min incubation on ice then centrifuged at 12,000 x g for 30 min at 4 °C. RNA pellet was washed twice with 700 μl chilled 70% (v/v)

ethanol and centrifuged at 12,000 x g for 10 min at 4 °C. The pellet was air dried for 5 – 10 min and re-suspended in RNase-free water. The quantity and quality of extracted RNA was determined by a NanoDrop® ND-1000 spectrophotometer (Thermo Scientific, USA) with its default settings for RNA quantification.

The first-strand complementary DNA (cDNA) was synthesised from total RNA according to the manufacturer's instruction using a QuantiTect® Reverse Transcription Kit (Qiagen, USA). RNA ligase mediated rapid amplification of cDNA ends (RLM-RACE) was performed using FirstChoice® RLM-RACE Kit (Ambion, USA) for both 5' and 3' ends as per manufacturer's instructions. After gel electrophoresis, desired polymerase chain reaction (PCR) products were extracted from agarose gel using a QIAquick® Gel Extraction Kit (Qiagen, Germany). PCR products were cloned into pGEM®-T Easy Vector System I (Promega, USA) and transformed into TOP10 chemically competent *Escherichia coli* cells. Blue/white screening on indicator plates and colony PCR were carried out to screen the colonies for insert. Plasmid DNAs were extracted using a QIAprep® Spin Miniprep Kit (Qiagen, Germany) and sequenced to identify the sequence of LdSVP. Ten clones for each insert were selected for sequencing.

Bioinformatics analysis of LdSVP

SVP, SVP homologous gene and other related gene sequences of both monocots and dicots other than *L. domesticum* were selected based on their availability in the public databases (*Table 1*). Since *L. domesticum* is a tropical plant, representatives of other tropical plants were also selected to predict the function of LdSVP and gene evolution from geographical distributions.

Protein sequences were analysed for domain organisation using NCBI-CD searches (http://www.ncbi.nlm.gov/Structure/cdd/wrpsb.cgi) against CDD v3.13 database. The low-complexity filter was turned off and the expected value was set at 0.01 to detect short domains and/or regions that were less conserved. Full result mode was selected. Domains were also verified and named according to the SMART (http://smart.embl-heidelberg.de), Pfam (http://pfam.xfam.org) and InterProScan (http://www.ebi.ac.uk/Tools/pfa/iprscan5)_databases.

Multiple sequence alignments of 29 SVP and SVP homologous gene sequences from 23 plant species were performed using ClustalW. Phylogenetic analysis was performed using MEGA7 (Kumar et al. 2016). The phylogenetic tree was constructed with the following settings: Tree inference as Neighbour Joining; Include Sites as pairwise deletion option for total sequence analysis and complete deletion option for each class analysis; Substitution Model; Poisson correction; and Bootstrap test of 1,000 replicates for internal branch reliability.

Results and discussion

Identification and annotation of LdSVP

A partial sequence of SVP homologous gene was identified and obtained from a transcriptome NGS database of *L. domesticum* in a previous study (unpublished). From the partial sequence, primers were designed for 5' and 3' rapid amplification of cDNA ends (RACE) to obtain full-length LdSVP (*Figure 1*). The full-length LdSVP was 1,176 bp long with 684 bp of coding sequence, a 5' untranslated region (UTR) of 226 bp and 266 bp of 3' UTR (GenBank accession number KY404063) (*Figure 2*). BLASTx search from the National Centre for Biotechnology Information (NCBI) revealed that LdSVP was highly similar to SVP. It is a type II lineage of MADS-domain transcription factor that has the intervening domain/I-box (I), a keratin-like domain/K-box (K) and a carboxyl-terminal/C-terminal domain (C) after the MADS-domain (M) (Nam et al. 2003). LdSVP encoded a MADS-box protein of 56 amino acids,

Table 1. List of amino acid sequences used in bioinformatics analysis of LdSVP

Plant species	Accession number	Reference
Actinidia chinesis	AFA37967.1	Uniprot
Arabidopsis lyrata subsp. *lyrata*	EFH54881.1	Uniprot
Arabidopsis thaliana	AFU85636.1	Uniprot
Brassica rapa subsp. *campestris*	AAQ55452.1	Uniprot
Carica papaya	evm.model.supercontig_55.31	Phytozome
Eucalyptus grandis	AAP33087.1	Uniprot
Glycine max	ABY78023.1	Uniprot
Gossypium raimondii	orange1.1g027190m	Phytozome
Hordeum vulgare subsp. *vulgare*	CAB97350.1	Uniprot
Malus x *domestica*	MDP0000209705	Phytozome
Manihot esculenta	Carubv10023953m	Phytozome
Medicago truncatula	Medtr5g032520.1	Phytozome
Musa acuminata	GSMUA_Achr10P19290_001	BananaGenome
Oryza sativa	LOC_Os02g52340.1	Phytozome
Panicum virgatum	Pavirv00052107m	Phytozome
Pisum sativum	AAX47170.1	Uniprot
Poncirus trifoliata (PtSVP)	ACJ09170.1	Uniprot
Poncirus trifoliata (PtAGL24)	A0A1I9ZJQ3	Uniprot
Prunus persica	ABJ96361.2	Uniprot
Solanum lycopersicum	Solyc11g010570.1.1	Phytozome
Thellungeilla halophila	Thhalv10000308m	Phytozome
Theobroma cacao	Thecc1EG000992t1	Phytozome
Vitis vinifera	GSVIVP00009443001	PlantGDB
Zea mays	GRMZM2G046885_T01	Phytozome

I-box, K-box and C-terminal domains (*Figure 2*), similar to SVP which has a clear recognisable MIKC domain structure (Hartmann et al. 2000).

It has been reported that SVP is involved in the maintenance of juvenility (Hanke et al. 2007) and control of flowering time in response to ambient temperature (Lee et al. 2007), an environmental cue that may also be responsible for the seasonality of tropical fruit trees. In *Arabidopsis*, FLC is the major floral repressor that suppresses floral initiation prior to vernalisation (Michaels and Amasino 1999; Henderson and Dean 2004; Searl et al. 2006). However, FLC orthologues have not been widely found in plants outside of Brassicaceae, the family to which *Arabidopsis* belongs. Hence, SVP homologous gene may play the role of the floral repressor where FLC homologue is absent, since SVP was found to be involved in juvenility maintenance and control of flowering time in *Arabidopsis*. As evident in this study, LdSVP was expressed in 3-month old duku plants which made it a potential floral repressor for duku and indicated its possible involvement in maintenance of juvenility.

Phylogenetic characterisation of LdSVP

SVP homologous genes from 10 plant species were found in UniProt databases while 14 others were identified by carrying out BLAST searches based on genome sequence data from Phytozome and other databases. Although *L. domesticum* is a dicot plant, it was useful to include SVP homologous sequences from monocot plants (*Musa acuminata*, *Hordeum vulgare*, *Oryza sativa* and *Zea mays*) for comparative

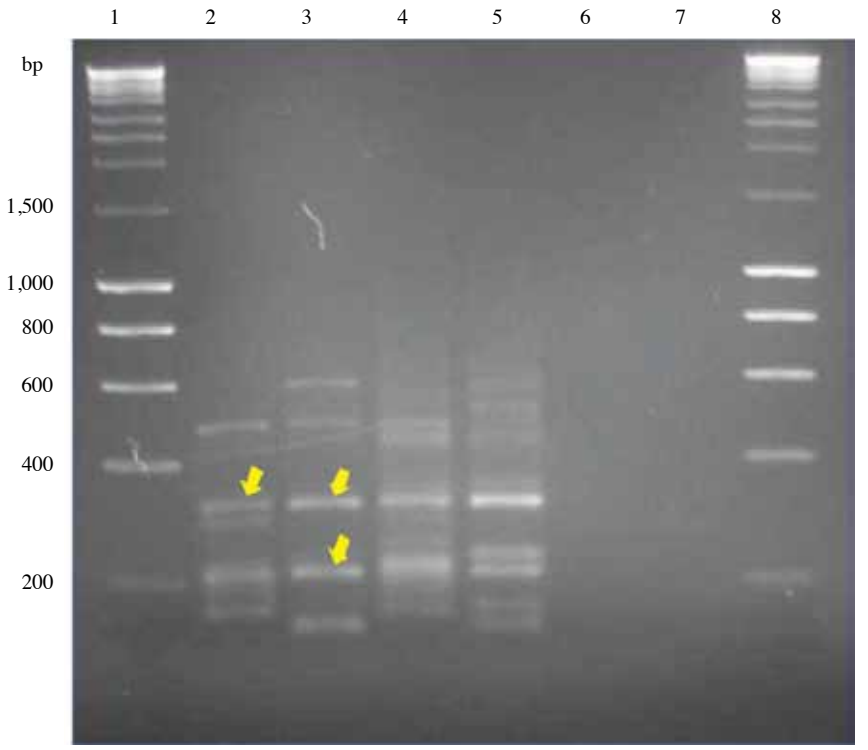


Figure 1. Gel electrophoresis of 5' inner rapid amplification of cDNA ends (RACE) PCR. Lane 1 and 8: DNA marker; Lanes 2-5: 5' inner RACE PCR using 5' outer RACE PCR products from different primer combinations as DNA template, with (Lanes 2 and 3) and without (Lanes 4 and 5) hot start PCR; Lane 6: Negative control without hot start PCR; Lane 7: Negative control with hot start PCR. Yellow arrows indicate PCR products cloned and sequenced

studies as monocots and dicots separated about 150 million years ago (Theißen et al. 2000). Most SVP homologous genes from tropical plant species were obtained from Phytozome as research on SVP homologous gene in tropical plants is relatively scarce.

Multiple sequence alignment of LdSVP and its homologous gene sequences showed the presence of MIKC domains, where M- and K-domains were highly conserved in LdSVP and its homologous genes of all plant species that were analysed (*Figure 3*). More differentiation was observed in C-terminal than M- and K-domains, rendering it as a less conserved domain in plant species. The presence of MIKC groups in LdSVP as shown in *Figure 3* indicates a type II MADS-domain or the MYOCYTE-ENHANCER-FACTOR2 (MEF2-like) classification which is related to different types of gametocyte/sporophyte development (Shore and Sharrocks 1995).

MADS-domain, a well conserved sequence of about 60 amino acids, is positioned at the N-terminus and is involved in DNA binding at consensus recognition sequences known as CArG boxes [CC(A/T)6GG] (Riechmann et al. 1996). There are nine classes of MIKC-type MADS-domain genes and SVP belongs to Class T, the class that is involved in flowering repression in *A. thaliana* (Nam et al. 2003). The name MADS originates from four members of gene family MINICHROMOSOME MAINTENANCE 1, AGAMOUS, DEFICIENS AND SERUM RESPONSE FACTOR, which were discovered from various organisms. It encodes transcription

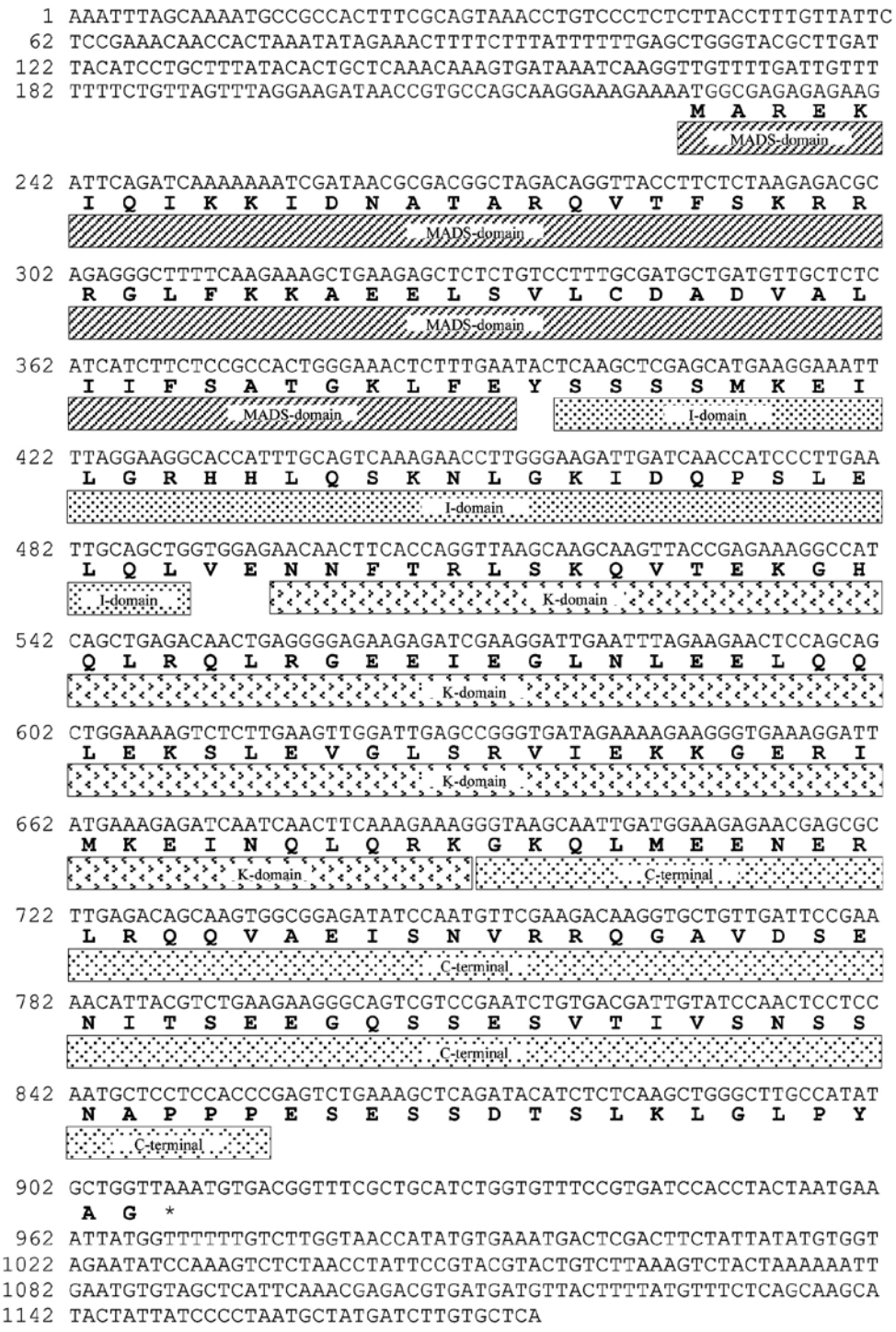


Figure 2. Nucleotide sequence of LdSVP cDNA isolated (GenBank accession number KY404063) and amino acid sequence of its deduced protein product. Amino acids are printed bold and boxes below indicate the MADS-, I-, K- and C-terminal domains

factors with important regulatory roles at various stages of plant development and there are more than 100 MADS-domain genes that have been discovered in flowering plants (Gramzow and TheiBen 2010). I-box is a short variable domain of about 30 amino acids which determines the selective formation of DNA-binding dimers (Pařenicová et al. 2003). I-box, as well as K-box, is the main feature that diverges MIKC-type proteins (Henschel et al. 2002). K-box, which is responsible for the dimerisation of MADS-box protein, spans about seven amino acids and is an exclusive evolutionary acquirement of plant MADS proteins due to its absence from animal and fungi MADS proteins (Davis et al. 1996; Theißen et al. 2000). It is characterised as having regularly spaced hydrophobic amino acids (Henschel et al. 2002).

C-terminal domain of LdSVP varies in length and sequence. It is thought to mediate protein-protein interactions. It has been shown to obtain a trans-activation domain or contribute to the formation of ternary or quaternary MADS-box protein complexes (García-Maroto et al. 2003). Most plant MADS proteins share a similar structure designated as MIKC which also appears in LdSVP.

The phylogenetic tree, consisting of 29 SVP and its homologous gene sequences from 23 plant species, was constructed with five distantly related MADS-box genes from three plant species and served as an outgroup (*Figure 4*). The out-group genes belong to type I lineage of MADS-domain with main differences in its K-box. From the ancestor, two branches were generated. Starting from the original speciation event,

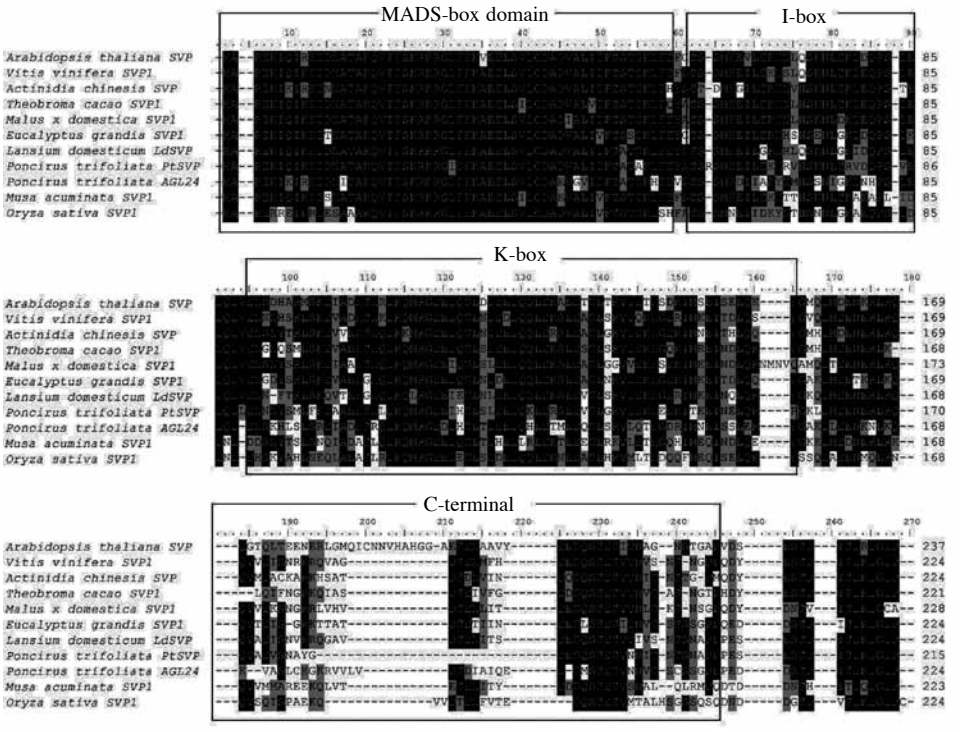


Figure 3. Sequence alignment and comparison of Lansium domesticum SVP (LdSVP)with related SVP and SVP homologous proteins from Arabidopsis thaliana, Vitis vinifera, Actinidia chinesis, Theobroma cacao, Malus x domestica, Eucalyptus grandis, Poncirus trifoliata, Musa acuminata and Oryza sativa. Identical amino acids are shaded in black. The MIKC domains are indicated in respective boxes

it branched out into SVP related branch and the outgroup. The SVP related branch consisted of all the SVP related sequences from all species while the sequences of APETALA1, FLC and SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 (SOC1) from *A. thaliana*, MADS-box gene and SVP homologous gene of *Hordeum vulgare* and *Panicum virgatum* were in the outgroup. The outgroup served as a control for the formation of the SVP related branch. The phylogenetic tree split into two main clades with bootstrap values of 0.474 and 0.567 and *L. domesticum* appeared in the larger clade together with 18 other plant species. The clade with bootstrap value of 0.567 also represented the dicots and most of the selected tropical plants. The monocots were nested in the clade with bootstrap value of 0.474. Based on the phylogenetic analysis, it was observed that LdSVP belong in the clade containing SVP homologous genes of *Poncirus trifoliata* (PtSVP) and *Malus* x *domestica* which was nested within the larger clade with *Theobroma cacao*, *Carica papaya* followed by *Vitis vinifera*. These data suggest that LdSVP may share similar functions of SVP homologues in citrus and apple.

Malus x *domestica* SVP homologous genes accumulate in the buds during winter, and then peak during the coldest time of the year, followed by a transient decline before budbreak (Wu et al. 2017). In *Arabidopsis*, SVP involves in flowering repression and modulates flowering time at different ambient temperatures (Lee et al. 2007).

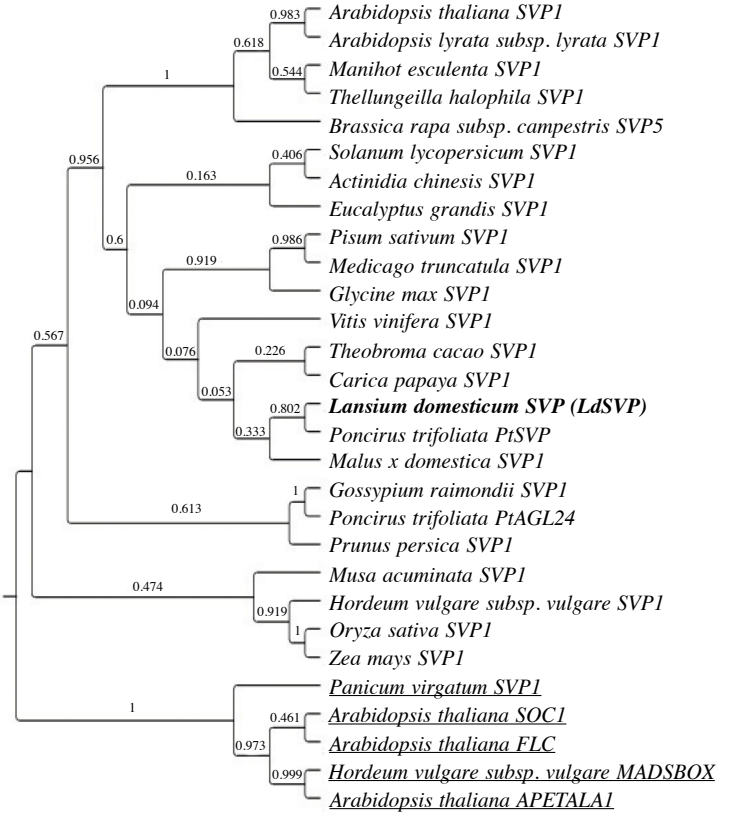


Figure 4. Phylogenetic relationship of LdSVP and other SVPs from various plants. In bold is LdSVP with the outgroup underlined. The phylogenetic tree was constructed with the neighbour-joining method of the MEGA7 programme. The numbers at the branches are bootstrap values from 1000 replicates

This may be taken into consideration in regards to the more distinctive temperature change between dry and wet seasons in tropical climate. Successive rainless weather and larger difference between daily minimum and maximum air temperature are among the conditions during low night-time temperature events which coincide with flowering of some dipterocarp trees (Yasuda et al. 1999) suggesting that temperature can be an important environmental cue for flowering in tropical trees. Hence, it is tempting to speculate that LdSVP could play the similar role in duku as SVP in *Arabidopsis*. However, experimental validation is needed to confirm the functions of LdSVP.

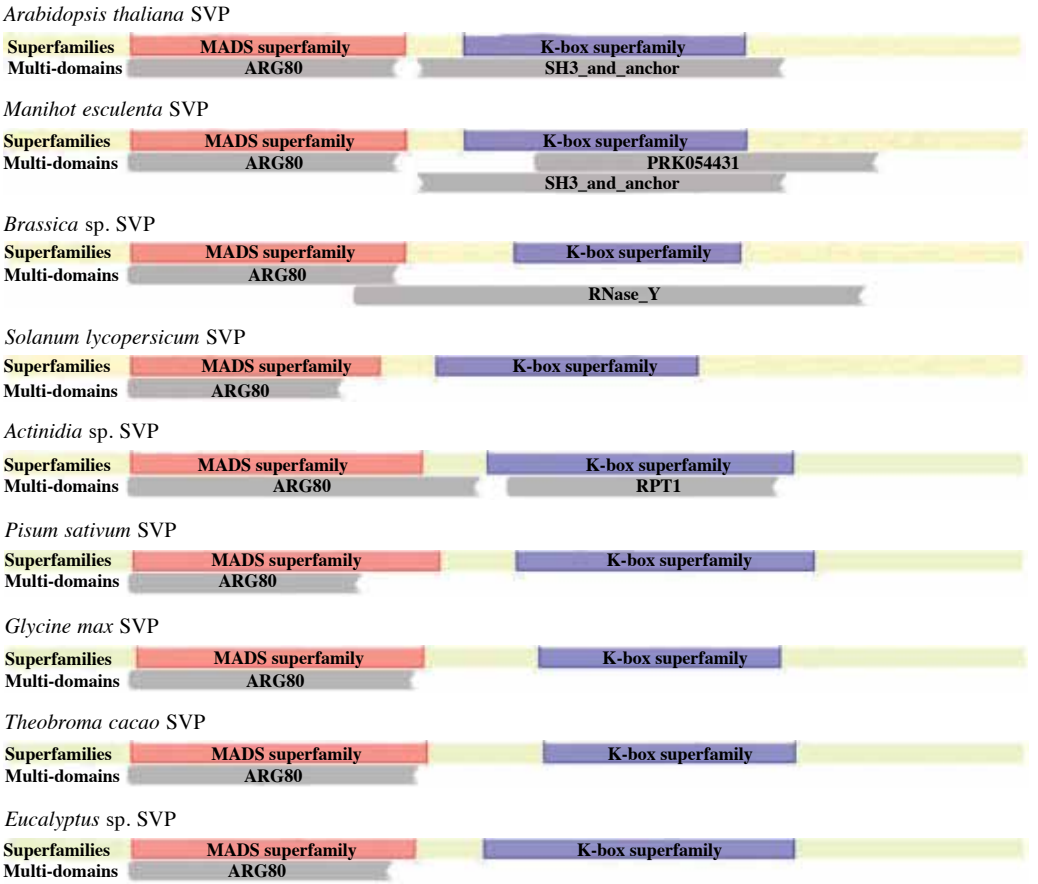
After the separation of SVP homologous gene in monocots from the dicots, the majority of the tropical plants, including *L. domesticum*, were the first to diverge from the rest of the dicots SVP homologous genes (*Figure 4*). From the constructed phylogenetic tree, MIKC-type genes can be further subdivided into gene clades based on the plant species. Most clade members share highly related functions and similar expression patterns during vegetative or reproductive growth (Theißen et al. 2000). Evolution study showed MIKC-type protein domain facilitated an efficient functional diversification through gene duplication, sequence divergence and fixation of gene regulatory networks. This is reflected in the diverse roles of MADS-box genes in flowering plant development where interestingly, they cover both floral integrators and repressors (Michaels and Amasino 1999; Hartmann et al. 2000). Therefore, understanding their phylogeny may lead to better understanding of the origin and evolution of flowering plants. Phylogenetic analysis also predicted that LdSVP may be orthologues of *Malus* x *domestica* SVP and PtSVP respectively. However, LdSVP may not share the same dormancy regulatory role as *Prunus persica* dormancy-associated MADS-box gene group, which are most closely related to SVP and AGL24 (Bielenberg et al. 2004; Wells et al. 2015), as their amino acid sequences are highly different (*Figure 4*).

Domain organisation showed high similarity of MADS-box domain pattern except for domain organisation of K-box, which is different in both length and position among the selected SVP and SVP homologous genes (*Figure 5*). Domain organisation for K-box in LdSVP and PtSVP showed similar length but there is a shift in location from MADS-domain (*Figure 5*). K-box of LdSVP was closer to MADS-domain than that of PtSVP. The closest neighbour for LdSVP was PtSVP, with proof of divergence from *L. domesticum* ancestor as shown in *Figure 4*. I-box and C-terminal are weakly conserved domains (Kaufmann et al. 2005). The domains, which were being considered under K-box region, were obtained from InterproScan analysis and not from Pfam analysis. Diversity in the distribution of K-box in the phylogenetic tree was contributed by the presence of multiple-domains under the K-box superfamily such as SH3-anchor, PRK05431, PRK05771, RNase Y, RPT1 and PTZ00108 (*Figure 5*). All SVP and SVP homologous proteins contained MADS superfamily with ARG80 domain (*Figure 5*).

It was observed that different domain architectures of K-box generated diverse protein sizes. ARG80 regulates the metabolism of arginine in association with arginine sensor (Amar et al. 2000). The main difference between K-box of LdSVP and PtSVP was the presence of four multi-domains, conjTIGR03752, PTZ00108, SMC prok A and SH3-anchor, in LdSVP as compared to two multi-domains, RPT1 and PRK05771, in PtSVP (*Figure 5*). SVP and SVP homologous protein of *M. trunculata* contained similar SH3-anchor multi-domain as LdSVP while PTZ000108 multi-domain was found in the K-box of *Oryza sativa* SVP. Conservation of amino acids and their properties has been used to predict functions for the conserved MIKC domains. Studies on the expression

of PtSVP revealed its close relation to floral bud initiation in *P. trifoliata*, the trifoliate orange. PtSVP expressed intensively in vegetative meristems and dormant tissues, and the expression had an annual fluctuation pattern which coincided with the flowering competence of shoots (Li et al. 2010), a typical pattern of yearly phase transition in woody plants that is absent in annual herbaceous plants. Down regulation of PtSVP led to initiation of floral bud differentiation. Although AGL24 is a close homologue of SVP, LdSVP is phylogenetically closer to PtSVP than PtAGL24, a MADS-box gene from *P. trifoliata* that is closely related to AGL24 (*Figure 4*). Ectopic expression of PtAGL24 in *Arabidopsis* caused early flowering (Sun et al. 2016). Hence, LdSVP is more likely a floral suppressor than a floral inducer.

SVP group seems to have expanded functionally as well as phylogenetically in perennials and the functions may include regulation of floral transition and seasonal growth, dormancy cycle, and/or juvenile-mature transition (Jiménez et al. 2009). Perennial flowering genes orthologues have been shown to have similar functions as their *Arabidopsis* name-sakes in general (Wilkie et al. 2008). MADS transcription factors, particularly SVP, appear as important candidate genes attributed to the natural variation of flowering time in plant families that are phylogenetically distant from *Arabidopsis* (Bielenberg et al. 2008; Méndez-Vigo et al. 2013). Studies have shown that SVP homologous genes display partially conserved functions in different families of both mono- and dicotyledonous plants (Trevaskis et al. 2007;



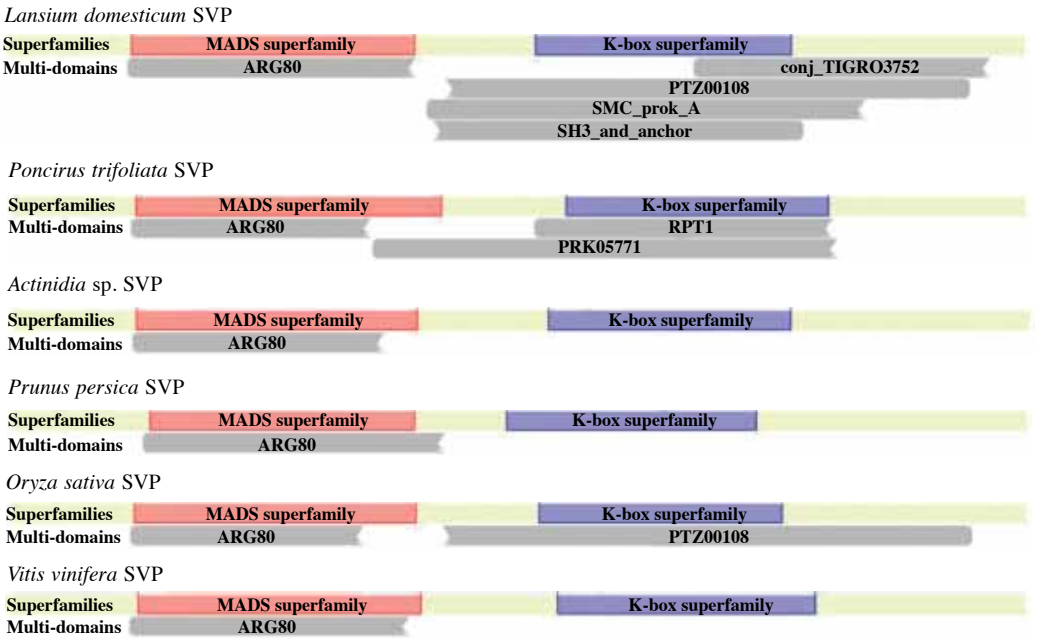


Figure 5. Domain organisation of selected species clades in the phylogenetic tree. Superfamilies and multi-domains are indicated

Li et al. 2010; Cohen et al. 2012; Lee et al. 2012). The predicted LdSVP protein conformed to MIKC-type protein and had similar functional domains as SVP. The phylogenetic analysis also showed that LdSVP was evolutionarily close to PtSVP. Hence, LdSVP is a possible candidate SVP homologous gene in *L. domesticum* and may share similar functions of SVP group, especially PtSVP in regulating annual floral transition.

Conclusion

LdSVP is a potential gene for future study to understand the floral transition in duku based on the predicted functions suggested in this study. It is phylogenetically closer to PtSVP, a transcription factor correlated with the floral initiation in trifoliate orange than SVP and functional study will confirm the functions predicted through phylogenetic analysis. This will provide better understanding on how SVP homologous gene regulates or is being regulated by other genes in the flowering pathways of tropical plants in general, and duku in particular. The gene sequence will be useful in developing molecular marker for marker-assisted breeding and germplasm screening. Flowering time related traits can be identified at a much earlier stage of plant development which will shorten the time for selecting desired plants in breeding programmes of perennial woody plants.

Acknowledgements

We thank Nor Hidayah Mohktar and Nur Anisza Hanoum Naserom for their contributions to the laboratory work. Financial support was provided by MARDI Reserve Fund (WRM) JP-RB-0325.

References

Accerbi, M., Schmidt, S.A., De Paoli, E., Park, S., Jeong, D.H. and Green, P.J. (2010). Methods for isolation of total RNA to recover miRNAs and other small RNAs from diverse species. In: *Plant MicroRNAs: Methods and Protocols* (Meyers, B.C. and Green, P.J., ed.) p. 31 – 50. Humana Press

Amar, N., Messenguy, F., El Bakkoury, M. and Dubois, E. (2000). ArgRII, a component of the ArgR-Mcm1 complex involved in the control of arginine metabolism in *Saccharomyces cerevisiae*, is the sensor of arginine. *Mol. Cell. Biol.* 20: 2087 – 2097

Bielenberg, D.G., Wang, Y., Fan, S., Reighard, G.L., Scorza, R. and Abbott, A.G. (2004). A deletion affecting several gene candidates is present in Evergrowing peach mutant. *J. Heredity* 95: 436 – 444

Bielenberg, D.G., Wang, Y., Li, Z., Zhebentyayeva, T., Fan, S., Reighard, G.L., Scorza, R. and Abbott, A.G. (2008). Sequencing and annotation of the evergrowing locus in peach [*Prunus persica* (L.) Batsch] reveals a cluster of six MADS-box transcription factors as candidate genes for regulation of terminal bud formation. *Tree Genet. Genomes* 4: 495 – 507

Blázquez, M.A., Ahn, J.H. and Weigel, D. (2003). A thermosensory pathway controlling flowering time in *Arabidopsis thaliana*. *Nat. Genet.* 33: 168 – 171

Chang, S., Puryear, J. and Cairney, J. (1993). A simple and efficient method for isolating RNA from pine trees. *Plant Mol. Biol. Rep.* 11: 113 – 116

Cohen, O., Borovsky, Y., David-Schwartz, R. and Paran, I. (2012). *CaJOINTLESS* is a MADS-box gene involved in suppression of vegetative growth in all shoot meristems in pepper. *J. Exp. Bot.* 63: 4947 – 4957

Davis, B., Egea-Cortines, M., de Andrade Silva, E., Saedler, H. and Sommer, H. (1996). Multiple interactions amongst floral homeotic MADS box proteins. *EMBO J.* 15: 4330 – 4343

Díaz-Riquelme, J., Lijavetzky, D., Martínez-Zapater, J.M. and Carmona, M.J. (2009). Genome-wide analysis of MIKC^{C}-type MADS box genes in grapevine. *Plant Physiol.* 149: 354 – 369

García-Maroto, F., Carmona, M.-J., Garrido, J.-A., Vilches-Ferrón, M., Rodríguez-Ruiz, J. and López Alonso, D. (2003). New roles for MADS-box genes in higher plants. *Biol. Plant.* 46: 321 – 330

Gramzow, L. and Theißen, G. (2010). A hitchhiker's guide to the MADS world of plants. *Genome Biol.* 11: 214

Hanke, M-V., Flachowsky, H., Peil, A. and Hättasch, C. (2007). No flower no fruit - genetic potentials to trigger flowering in fruit trees. *G3* 1: 1 – 20

Hanum, L., Kasiamdari, R.S., Santosa and Rugayah. (2013). The phylogenetic relationship among varieties of *Lansium domesticum* Correa based on ITS rDNA sequences. *Indonesian J. Biotech.* 18: 123 – 132

Hartmann, U., Höhmann, S., Nettesheim, K., Wisman, E., Saedler, H. and Huijser, P. (2000). Molecular cloning of *SVP*: a negative regulator of the floral transition in *Arabidopsis*. *Plant J.* 21: 351 – 360

Henderson, I.R. and Dean, C. (2004). Control of *Arabidopsis* flowering: the chill before the bloom. *Development* 131: 3829 – 3838

Henschel, K., Kofuji, R., Hasebe, M., Saedler, H., Münster, T. and Theißen, G. (2002). Two ancient classes of MIKC-type MADS-box genes are present in the moss *Physcomitrella patents*. *Mol. Biol. Evol.* 19: 801 – 814

Higgins, D.G. and Sharp, P.M. (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* 73: 237 – 244

Jiménez, S., Lawton-Rauh, A.L., Reighard, G.L., Abbott, A.G. and Bielenberg, D.G. (2009). Phylogenetic analysis and molecular evolution of the dormancy associated MADS-box genes from peach. *BMC Plant Biol.* 9: 81

Kaufmann, K., Melzer, R. and Theißen, G. (2005). MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. *Gene* 347: 183 – 198

Kumar, S., Stecher, G. and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33: 1870 – 1874

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. and Higgins, D.G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947 – 2948

Lee, J. and Lee, I. (2010). Regulation and function of SOC1, a flowering pathway integrator. *J. Exp. Bot.* 61: 2247 – 2254

Lee, J.H., Park, S.H. and Ahn, J.H. (2012). Functional conservation and diversification between rice OsMADS22/OsMADS55 and *Arabidopsis* SVP proteins. *Plant Sci.* 185 – 186: 97 – 104

Lee, J.H., Yoo, S.J., Park, S.H., Hwang, I., Lee, J.S. and Ahn, J.H. (2007). Role of *SVP* in the control of flowering time by ambient temperature in *Arabidopsis*. *Genes Dev*. 21: 397 – 402

Li, D., Liu, C., Shen, L., Wu, Y., Chen, H., Robertson, M., Helliwell, C.A., Ito, T., Meyerowitz, E. and Yu, H. (2008). A repressor complex governs the integration of flowering signals in Arabidopsis. *Dev. Cell* 15: 110 – 120

Li, Z.M., Zhang, J.Z., Mei, L., Deng, X.X., Hu, C.G. and Yao, J.L. (2010). *PtSVP*, an *SVP* homolog from trifoliate orange (*Poncirus trifoliate* L. Raf.), shows seasonal periodicity of meristem determination and affects flower development in transgenic *Arabidopsis* and tobacco plants. *Plant Mol*. *Biol*. 74: 129 – 142

Manosroi, A., Chankhampan, C., Manosroi, W. and Manosroi, J. (2013). Anti-proliferative and matrix metalloproteinase-2 inhibition of Lonkong (*Lansium domesticum*) extracts on human mouth epidermal carcinoma. *Pharm*. *Biol*. 51: 1311 – 1320

Méndez-Vigo, B., Martínez-Zapater, J.M. and Alonso-Blanco, C. (2013). The flowering repressor *SVP* underlies a novel *Arabidopsis thaliana* QTL interacting with the genetic background. *PLOS Genet*. 9: e1003289

Michaels, S.D. and Amasino, R.M. (1999). *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell* 11: 949 – 956

Nam, J., dePamphilis, C.W., Ma, H. and Nei, M. (2003). Antiquity and evolution of the MADS-box gene family controlling flower development in plants. *Mol*. *Biol*. *Evol*. 20: 1435 – 1447

Pařenicová, L., de Folter, S., Kieffer, M., Horner, D.S., Favalli, C., Busscher, J., Cook, H.E., Ingram, R.M., Kater, M.M., Davies, B., Angenent, G.C. and Colombo, L. (2003). Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: new openings to the MADS world. *Plant Cell* 15: 1538 – 1551

Reeves, P.A., He, Y., Schmitz, R.J., Amasino, R.M., Panella, L.W. and Richards, C.M. (2007). Evolutionary conservation of the *FLOWERING LOCUS C*-mediated vernalization response: evidence from the sugar beet (*Beta vulgaris*). *Genetics* 176: 295 – 307

Riechmann, J.L., Wang, M. and Meyerowitz, E.M. (1996). DNA-binding properties of *Arabidopsis* MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA and AGAMOUS. *Nucleic Acids Res*. 24: 3134 – 3141

Rozen, S. and Skaletsky, H.J. (2000). Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (Krawetz, S. and Misener, S. eds.). p. 365 – 386 Humana Press, Totowa. N.J.

Samach, A., Onouchi, H., Gold, S.E., Ditta, G.S., Schwarz-Sommer, Z., Yanofsky, M.F. and Coupland, G. (2000). Distinct roles of CONSTANS target genes in reproductive development of *Arabidopsis*. *Science* 288: 1613 – 1616

Searle, I., He, Y., Turck, F., Vincent, C., Fornara, F., Kröber, S., Amasino, R.A. and Coupland, G. (2006). The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signalling in *Arabidopsis*. *Genes Dev*. 20: 898 – 912

Shore, P. and Sharrocks, A.D. (1995). The MADS-box family of transcription factors. *Eur. J*. *Biochem*. 229: 1 – 13

Sun, L.M., Zhang, J.Z. and Hu C.G. (2016). Characterization and expression analysis of *PtAGL24*, a *SHORT VEGETATIVE PHASE/ AGAMOUS-LIKE 24* (*SVP*/*AGL24*)-type MADS-box gene from trifoliate orange (*Poncirus trifoliata* L. Raf.). *Front*. *Plant Sci*. 7: 823

Theißen, G., Becker, A., Di Rosssa, A., Kanno, A., Kim, J.T., Münster, T., Winter, K.-U. and Saedler, H. (2000). A short history of MADS-box genes in plants. *Plant Mol*. *Biol*. 42: 115 – 149

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalty and weight matrix choice. *Nucleic Acids Res*. 22: 4673 – 4680

Trevaskis, B., Tadege, M., Hemming, M.N., Peacock, W.J., Dennis, E.S. and Sheldon, C. (2007). *Short vegetative phase*-like MADS-box genes inhibit floral meristem identity in barley. *Plant Physiol*. 143: 225 – 235

Wells, C.E., Vendramin, E., Torado, S.J., Verde, I. and Bielenberg, D.G. (2015). A genome-wide analysis of MADS-box genes in peach [*Prunus persica* (L.) Batsch]. *BMC Plant Biol*. 15: 41

Wilkie, J.D., Sedgley, M. and Olesen, T. (2008). Regulation of floral initiation in horticultural trees. *J. Exp. Bot.* 59: 3215 – 3228

Wu, R., Tomes, S., Karunairetnam, S., Tustin, S.D., Hellens, R.P., Allan, A.C., Macknight, R.C. and Varkonyi-Gasic, E. (2017). *SVP*-like MADS-box genes control dormancy and budbreak in apple. *Front. Plant Sci.* 8: 477

Yapp, D.T.T. and Yap, S.Y. (2003). *Lansium domesticum*: skin and leaf extracts of this fruit tree interrupt the lifecycle of *Plasmodium falciparum*, and are active towards a chloroquine-resistant strain of the parasite (T9) in vitro. *J. Ethnopharmacol.* 85: 145 – 150

Yasuda, M., Matsumoto, J., Osada, N., Ichikawa, S., Kachi, N., Tani, M., Okuda, T., Furukawa, A. Nik, A.R. and Manokaran, N. (1999). The mechanism of general flowering in Dipterocarpaceae in the Malay Peninsula. *J. Trop. Ecol.* 15: 437 – 449

Yee, T.F. and Rao, A.N. (2013). Comparative anatomy of duku and langsat - *Lansium domesticum*. *J. Trop. Med. Plants*. 14: 71 – 79

Yu, H., Ito, T., Wellmer, F. and Meyerowitz, E.M. (2004). Repression of AGAMOUS-LIKE 24 is a crucial step in promoting flowering development. *Nature Genet.* 36: 157 – 161

Yu, H., Xu, Y., Tan, E.L. and Kumar, P.P. (2002). AGAMOUS-LIKE 24, a dosage-dependent mediator of the flowering signals. *Proc. Natl. Acad. Sci. USA* 99: 16336 – 16341

Abstrak

Gen SHORT VEGETATIVE PHASE (SVP) yang mengekod domain protein MADS merupakan salah satu penindas pembungaan utama dalam *Arabidopsis thaliana* selain daripada FLOWERING LOCUS C (FLC). Ia bertindak mengikut dos dan pengekspresannya hanya di tisu vegetatif dan primodia bunga tetapi tidak pada bunga matang. Homolog SVP daripada *Lansium domesticum* yang dinamakan LdSVP telah dipencilkan daripada tisu daun muda anak pokok duku yang berusia 3 bulan. Jujukan amino asid LdSVP yang terhasil mengandungi domain MADS-box, I-box, K-box dan terminal-C yang memenuhi ciri-ciri protein jenis MIKC dan mempunyai homolog yang tinggi dengan SVP. cDNA LdSVP terdiri daripada 684 nukleotida kod jujukan, 226 nukleotida kawasan tidak tertranslasi 5' dan 266 nukleotida kawasan tidak tertranslasi 3'. Kajian filogenetik protein SVP dan homolognya melibatkan lebih daripada 30 spesies tumbuhan menunjukkan LdSVP mempunyai hubungan rapat dengan PtSVP, SVP homolog oren trifoliat (*Poncirus trifoliata*), di mana kedua-duanya berasal daripada leluhur yang sama. Jarak filogenetik LdSVP adalah lebih jauh daripada PtAGL24, gen MADS-box *P. trifoliata* yang berkaitan rapat dengan AGAMOUS-LIKE 24 (AGL24) *Arabidopsis*, berbanding dengan PtSVP. Oleh itu, adalah berkemungkinan LdSVP lebih berfungsi seperti SVP daripada AGL24. Analisis filogenetik protein dan kajian organisasi domain menunjukkan fungsi LdSVP mungkin serupa dengan PtSVP dalam penentuan masa pembungaan. Pengekspresan PtSVP adalah sangat tinggi di tisu dorman dan meristem vegetatif dan ia menyebabkan pembungaan lewat apabila ditranformasikan ke dalam *Arabidopsis* liar. Ini mencadangkan bahawa LdSVP mungkin terlibat dalam kawalan pembungaan bermusim untuk duku. Walaupun homolog SVP telah berkembang antara spesies tumbuhan tetapi domain berfungsinya adalah amat terpelihara di antara *Arabidopsis* dan *L. domesticum*.