

RESEARCH ARTICLE

Open Access



Diversification of *FT*-like genes in the PEBP family contributes to the variation of flowering traits in Sapindaceae species

Xing Huang^{1,2,3}, Hongsen Liu^{1,2,3}, Fengqi Wu^{1,2,3}, Wanchun Wei^{1,2,3}, Zaohai Zeng^{1,2,3}, Jing Xu^{1,2,3}, Chengjie Chen^{1,2,3}, Yanwei Hao^{1,2,3*}, Rui Xia^{1,2,3*} and Yuanlong Liu^{1,2,3,4*} 

Abstract

Many species of Sapindaceae, such as lychee, longan, and rambutan, provide nutritious and delicious fruit. Understanding the molecular genetic mechanisms that underlie the regulation of flowering is essential for securing flower and fruit productivity. Most endogenous and exogenous flowering cues are integrated into the florigen encoded by *FLOWERING LOCUS T*. However, the regulatory mechanisms of flowering remain poorly understood in Sapindaceae. Here, we identified 60 phosphatidylethanolamine-binding protein-coding genes from six Sapindaceae plants. Gene duplication events led to the emergence of two or more paralogs of the *FT* gene that have evolved antagonistic functions in Sapindaceae. Among them, the *FT1*-like genes are functionally conserved and promote flowering, while the *FT2*-like genes likely serve as repressors that delay flowering. Importantly, we show here that the natural variation at nucleotide position – 1437 of the lychee *FT1* promoter determined the binding affinity of the SVP protein (LcSVP9), which was a negative regulator of flowering, resulting in the differential expression of *LcFT1*, which in turn affected flowering time in lychee. This finding provides a potential molecular marker for breeding lychee. Taken together, our results reveal some crucial aspects of *FT* gene family genetics that underlie the regulation of flowering in Sapindaceae.

Keywords Sapindaceae, Lychee (*Litchi chinensis*), Flowering, *FT1*-like, *Cis*-regulatory element

*Correspondence:

Yanwei Hao
yanweihao@scau.edu.cn
Rui Xia
rxia@scau.edu.cn
Yuanlong Liu
yuanlongliu@mail.hzau.edu.cn

¹ State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, South China Agricultural University, Guangdong Guangzhou 510642, China

² South China Agricultural University, Guangdong Laboratory for Lingnan Modern Agriculture, South China Agricultural University, Guangdong Guangzhou 510642, China

³ South China Agricultural University, Key Laboratory of Biology and Germplasm Enhancement of Horticultural Crops in South China, Ministry of Agriculture and Rural Affairs, South China Agricultural University, Guangdong Guangzhou 510642, China

⁴ College of Plant Science and Technology, Huazhong Agricultural University, Hubei Wuhan 430070, China

Core

The natural variation at nucleotide position – 1437 of the lychee *FT1* promoter determined the binding affinity of the SVP protein (LcSVP9), a negative regulator of flowering, resulting in the differential expression of *LcFT1*, which in turn affected flowering time in lychee. This finding provides a potential molecular marker for lychee breeding.

Gene and accession numbers

Sequence data of Sapindaceae plants from this article can be found in the SapBase database (<http://www.sapindaceae.com/Download.html>) (Li et al. 2024). The remaining sequences were obtained from the Phytozome, GenBank, and Arabidopsis Information Resource (TAIR) databases. The accession numbers can be found in Table S2 and Table S6.



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Lychee (*Litchi chinensis*), longan (*Dimocarpus longan*), and rambutan (*Nephelium lappaceum*) are commercially important fruit trees in the Sapindaceae family, which are extensively cultivated in tropical and subtropical areas worldwide (Menzel et al. 1995; Shahrajabian et al. 2020). These species are closely related and possess valuable fruits with an edible aril (Zee et al. 1998). Although they are closely related, their flowering traits, which are directly linked to production, are not entirely identical. Lychee and longan are day-neutral plants that require prolonged exposure to low temperatures to consistently flower annually (Subhadrabandhu et al. 2000; Chen et al. 2003). In contrast, rambutan flowering is triggered by water scarcity, which occurs biannually during March to May and July to August in reaction to two short periods of arid conditions followed by intermittent rainfall (Shaari et al. 1983; Tindall, 1994). Longan exhibits a unique floral induction mechanism in response to potassium chlorate (KClO₃), enabling off-season and year-round fruit production, which is unparalleled in other fruit crops (Matsumoto et al. 2007). Understanding the molecular genetic processes that control flowering would help the Sapindaceae fruit industry.

Plants undergo significant physiological changes during the transition from vegetative phase to reproductive stage, which is triggered by various internal and external signals that ultimately lead to flowering. The regulatory mechanisms of flowering in model plants have been elucidated by identifying at least five genetically defined pathways (Srikanth et al. 2011). The components of these pathways can vary across species, but most internal and external signals are integrated into a few central hubs that are conserved, the most well-known of which is the florigen encoded by *Flower Locus T (FT)*, which belongs to the phosphatidylethanolamine-binding protein (PEBP) family (Kardailsky et al. 1999; Liu et al. 2016). In plants, FT serves as a small, mobile signaling molecule that is synthesized in the leaves and moves to the shoot apex, enabling the transition to reproductive development and flowering (Corbesier et al. 2007; Mathieu et al. 2007). Despite the considerable conservation observed in flowering regulators among plant species, recent research has revealed cases where FT homologs have evolved to inhibit flowering (Blackman et al. 2010; Pin et al. 2010; Harig et al. 2012; Lee et al. 2013; Nan et al. 2014; Zhai et al. 2014). The presence of four FT homologs has been identified in the Sunflower (*Helianthus annuus*), with one of them exhibiting a novel repressor function attributed to a mutation (Blackman et al. 2010). Similarly, sugar beet contains the BvFT1 and BvFT2 homologs of FT with antagonistic functions (Pin et al. 2010). External loop formation of BvFT caused by the divergence of the

amino acids within segment B may elucidate the opposite functions (Pin et al. 2010). The Y134 and W138 residues within segment B act to change the FT function (Pin et al. 2010). The MADS-box genes, *SHORT VEGETATIVE PHASE (SVP)* and *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1)*, encode essential transcription factors that play a crucial role in regulating the integration of flowering signals (Ng and Yanofsky 2011; Moon et al. 2003; Kou et al. 2022). SVP directly binds to the CARG motifs of FT and its homolog, repressing their expression in leaves (Jang et al. 2009). In contrast, SOC1 activates FT transcription in the leaf vasculature of soybean by directly binding to the CARG motif within its promoter; thus, ensuring the induction of flowering (Kou et al. 2022). SOC1 and SVP bind to CARG boxes, but their binding preferences vary. SOC1 predominantly binds to the SRP-type CARG box consensus sequence, while SVP binds to the intermediate type (Tao et al. 2012).

In contrast to mutations that impair gene function in protein-coding regions, variations occurring in either *cis*-regulatory elements (CREs) or regulatory regions result in alterations of transcriptional level that impact the phenotype. Variations in the *MdIPT5b* gene promoter region of apples results in a marked increase in cellular cytokinin levels (Feng et al. 2021). This can be attributed to the deletion of a 42-bp sequence within the promoter region, which impairs the *cis*-element ProRE (an ACTCAT motif), leading to stable gene expression under salt stress conditions. A variation (single nucleotide polymorphism (SNP)13 T/C) in the regulatory region of PbCPK28, a SNP-13, leads to variations in the fructose content of pear (Li et al. 2023). Similar regulatory mechanisms have been confirmed in the regulation of cold tolerance in tomatoes (Zhu et al. 2023), grain width and weight in rice (Ruan et al. 2020), and low-temperature tolerance in maize (Jiang et al. 2022). The variations in the promoter region have the potential to drive evolution, generating genetic variability that can be harnessed for domestication and breeding advances (Swinnen et al. 2016).

Despite the mechanisms of floral regulation have been thoroughly investigated in model species, such as *Arabidopsis thaliana* (Jang et al. 2009), tomato (*Solanum lycopersicum*) (Huang et al. 2021; Chong et al. 2022), and rice (*Oryza sativa*) (Pan et al. 2022; Tang et al. 2023), they remain poorly understood in the Sapindaceae family. Therefore, it is essential to comprehend the regulatory mechanisms that govern flowering in Sapindaceae fruit trees. Understanding the regulation of flowering time in Sapindaceae is crucial for efficient breeding and improving commercial production. Thus, in our study, we used genomic data from six Sapindaceae species and massive RNA-seq data to carry out a genome-wide

characterization of the FT family in Sapindaceae and explore the variations in the coding and promoter regions of the *FT* genes, which contribute to the variation of flowering traits in Sapindaceae species.

Results

Gene member variations of the PEBP gene family in Sapindaceae

To investigate the composition and evolutionary relationships in the PEBP proteins from Sapindaceae, we screened the genomes of six Sapindaceae species with available genomes, including lychee (*Litchi chinensis*), longan (*Dimocarpus longan*), rambutan (*Nephelium lappaceum*), soapberry (*Sapindus mulorossi*), yellowhorn (*Xanthoceras sorbifolium*), and balloon vine (*Cardiospermum halicacabum*). Sixty PEBP proteins with complete open reading frames were identified in Sapindaceae (Fig. 1a, b, Fig. S1 and Table S1). We performed a phylogenetic analysis with 96 functionally reported PEBP proteins from 46 other major angiosperm lineages (Wickland et al. 2015; Table S2), which indicated that the PEBP genes were separated into three primary clusters

of MFT-like, TFL1-like, and FT-like clades (Fig. 1a). The number of PEBP genes in Sapindaceae varied, ranging from nine to twelve copies (Fig. 1b). Lychee and longan possessed four *FT* genes, but there were seven *FT*-like genes in rambutan, which is a close relative of lychee and longan (Fig. 1b), indicating that *FT* gene expansion occurs in rambutan. Fewer *FT*-like genes were detected in balloon vine than in the other five woody Sapindaceae species, however, the *MFT*-like genes were expanded (Fig. 1b).

To investigate the evolutionary relationships among the *FT*-like genes, a syntentic analysis was performed on the Sapindaceae family. The *FT1* genes maintained a good syntentic relationship across Sapindaceae, indicating that *FT1* is functionally conserved. The *FT2* genes lost their syntentic relationship in the basal Sapindaceae plant, yellowhorn, but they were conserved in the other Sapindaceae plants. The *FT1*, *FT2*, and *FT3* loci maintained good collinearity in lychee, longan, rambutan, and soapberry (Fig. 1c). In particular, the *FT1* gene was tandemly duplicated into four copies in rambutan, which led to the expansion of the *FT* genes (Fig. 1c). A similar

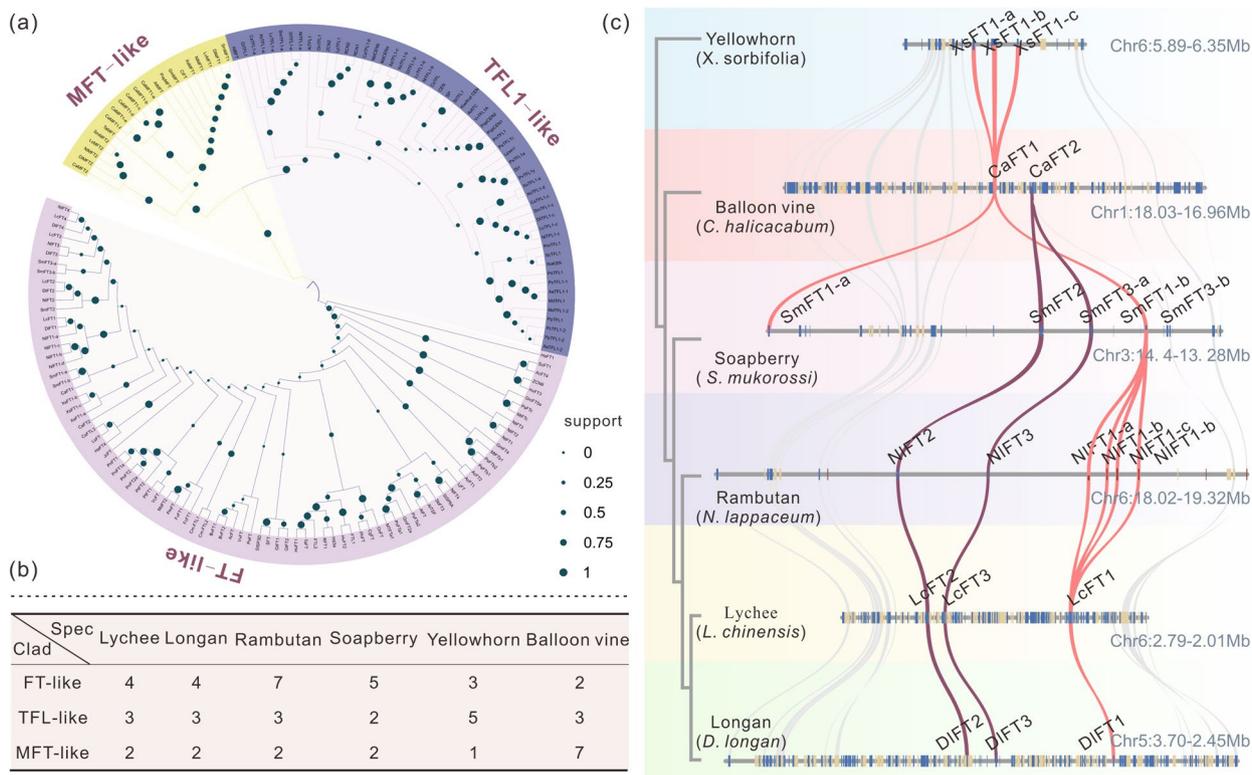


Fig. 1 The variation in gene numbers within the PEBP gene family across six Sapindaceae plants. **a** Phylogenetic analysis of the PEBP protein family in 52 flowering plant species. The construction of the tree was performed utilizing the Neighbor Joining in MEGA X (Kumar et al. 2018). **b** The PEBP gene copy numbers in each species were mapped in this table. “Clad” and “spec” are short for “clades” and “species” respectively. **c** Macro-synteny across lychee, longan, rambutan, soapberry, yellowhorn and balloon vine for *FT*-like genes. Syntenic blocks are connected by lines. The gene pairs of *FT1*-like are linked with red curves, while those of *FT2*-like/*FT3*-like are connected by purple curves

tandem replicate event occurred in *FT1* loci of yellowhorn (Fig. 1c), and *MFT1* loci of balloon vine (Fig. S2). These results suggest that tandem duplication events contribute to PEBP gene expansion in Sapindaceae.

Functional antagonism of *FT* homologs in floral induction

Florigen *FT* induces flowering in angiosperms, but some *FT* homologs repress flowering (Wickland et al. 2015). A previous study demonstrated that the antagonistic function of *FT* homologs is caused by changes in the external loop of the *FT* protein in sugar beets, which is encoded by segment B of the fourth exon (Pin et al. 2010). To identify the potential flowering inducer and repressor of *FT* homologs in Sapindaceae, the *FT* amino acid sequences of the segment B fragments from the six Sapindaceae species and other functionally characterized species were aligned (Fig. 2a). Tryptophan at position 138 (W138) was conserved across Sapindaceae species. This observation suggests that the amino acid substitution in position 134, either a Y (tyrosine) or a non-Y residue, may be more important for the functional divergence of *FT*-like homologs in Sapindaceae species (Fig. 2a). Based on the amino acid variation in position 134, we found that there are two types of *FT* gene in Sapindoideae, Y in *FT1/FT3/FT4*-like and N in *FT2*-like (Fig. 2a). Subsequently, we performed a folding prediction and our analysis revealed that the external loops in segment B of longan *FT1* (the main floral inducer) and *FT2* were conformationally different (Fig. 2b). This finding suggests that differences, particularly at residue Y134/N134, may confer functional antagonism between these two genes. We then detected the changes in expression of *DIFT1* and *DIFT2* in the shoot apices of longan before the initiation of floral primordia. The results revealed that the expression levels of *DIFT1* and *DIFT2* were nearly absent (Fig. S3a), suggesting that *FT* in longan is barely expressed in the shoot

apices. We further explored the expression profiles of *DIFT1* and *DIFT2* in leaves during natural flowering, and found that a dramatic increase in *DIFT1* from –28 to –7 days before flowering, while *DIFT2* exhibited a slight decrease (Fig. 2c). Additionally, the expression of *LcFT1* in lychee was enhanced during reactive oxygen species (ROS) and low-temperature treatments to induce flowering, while *LcFT2* decreased expression in leaves (Lu et al. 2020a; Zhang et al. 2017) (Fig. S3c, d), suggesting their possible functional antagonism. Taken together, we infer that the *FT1*-like genes may be conserved flowering inducers in Sapindaceae, while the *FT2*-like genes may function as flowering repressors due to the key substitution at *FT* residue 134.

To validate whether *FT1* and *FT2* promote or inhibit floral transition among Sapindaceae species, *DIFT1* and *DIFT2* were constitutively overexpressed and driven by the CaMV 35S promoter in *Arabidopsis*. Under long-day conditions, *DIFT1* overexpression plants started to flower when there were 7.5 rosette leaves, much earlier than WT control which flowered with 14.3 leaves (Fig. 2d, e, Fig. S3b). Conversely, the floral initiation of *DIFT2* transgenic plants was dramatically delayed until bearing 31.0 leaves (Fig. 2d and e, Fig. S3b). Thus, ectopic expression of *DIFT1* dramatically accelerated floral initiation, while ectopic expression of *DIFT2* substantially postponed floral initiation, confirming their functional antagonism in flowering induction.

Specific insertion in the *DIFT1* promoter may be involved in flowering diversification in longan

As *FT1*-like genes are the flowering-promoting genes in Sapindaceae, we next studied the regulatory differences in the *FT1*-like genes among the Sapindaceae species. Nucleotide sequences of the *FT1*-like genes (including 3.0 kilobases (kb) upstream and 1.0 kb downstream regions)

(See figure on next page.)

Fig. 2 Investigating the crucial regions contributing to the antagonistic functions of *FTs*. **a** The alignment of the amino acid sequence in the fourth exon for the *FT* inducer and inhibitor from onion (Lee et al. 2013), sugar beet (Pin et al. 2010), tobacco (Harig et al. 2012), sunflower (Blackman et al. 2010), soybean (Nan et al. 2014; Zhai et al. 2014; Kong et al. 2010), *Arabidopsis* (Kardailsky et al. 1999; Kobayashi et al. 1999), lychee (Ding et al. 2015), longan (Winterhagen et al. 2013), rambutan, soapberry, yellowhorn and balloon vine. Red dots indicate inducers that have been functionally validated, while blue triangles indicate repressors that have undergone functional validation. Pink dots represent speculated activators, while light blue triangles represent speculated inhibitors. **b** The prediction of protein folding for *DIFT1* and *DIFT2*. The molecular structures were visualized using solid three-dimensional traces in a diverse color scheme for alignment. **c** The expression pattern of *FTs* during flower induction in longan leaves was analyzed using quantitative real-time polymerase chain reaction (qRT-PCR). The $-\Delta\Delta CT$ calculation was employed to normalize expression levels. The data presented represent the mean \pm standard deviation (SD) from three independent biological replicates, with *Actin-100* (Dil.11g013910.1) used as the internal control gene. **d** The phenotype of *Arabidopsis* with the overexpression of longan *FT1* and *FT2* genes. All plants were grown under long-day conditions with 16 h of light and 8 h of darkness at 23°C. The left image depicts the phenotype after 26 days of sowing, while the right image showcases the phenotype after 30 days of sowing. Col (WT) represents the wild-type *Arabidopsis* with a Col-0 background. OE represents the transgenic lines with *DIFT* overexpression. T₂ generation lines with *DIFT1* and *DIFT2* genes are represented by numbers 1 to 3 and 16 to 18, respectively, with a scale of 1 cm. **e** The statistics of the number of rosette leaves during flowering for both wild-type and transgenic *Arabidopsis* plants, with six plants used for each line. The statistics for wild-type *Arabidopsis* began approximately 30 days after sowing, while the OE-*DIFT1* lines started around 26 days after sowing. The OE-*DIFT2* lines commenced around 36 days after sowing. The data is presented as mean \pm SD. One-way analysis of variance (ANOVA) was performed. *P* value: **P* < 0.05, ***P* < 0.01, ****P* < 0.001

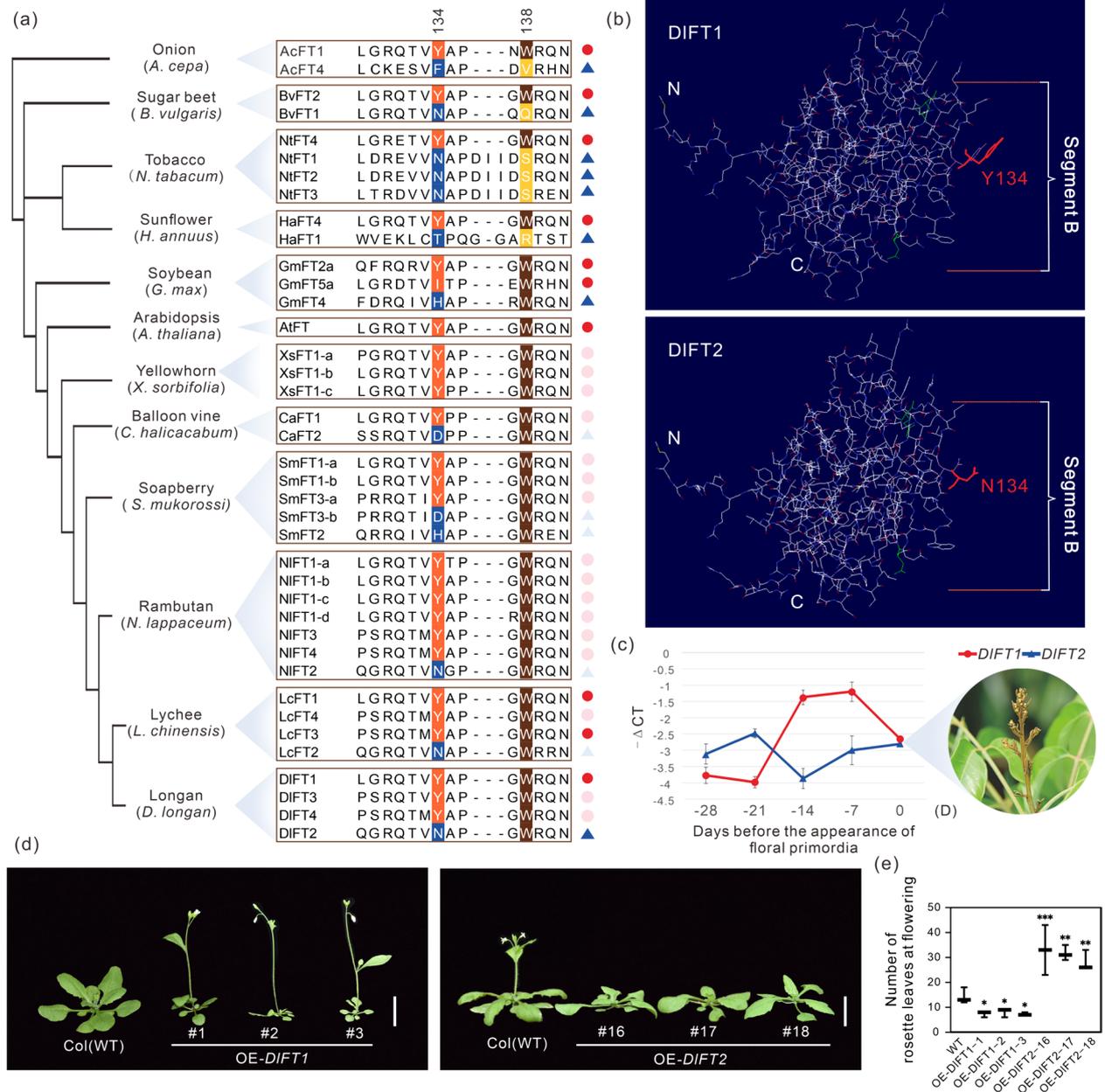


Fig. 2 (See legend on previous page.)

were extracted from three longan varieties ('SX', 'HML', and 'JDB'), three lychee varieties ('FZX', 'GW', and 'HML'), as well as rambutan, soapberry, yellowhorn, and balloon vine for comparative analysis (Adrian et al. 2010). Interestingly, the alignment revealed the specific 321-bp insertion located between 619 and 940 nucleotides upstream from the *DIFT1* start codon in the three longan cultivars (Fig. 3a, highlighted in red box). To investigate whether this 321-bp insertion harbored longan-specific CREs, we compared the CREs among the complete promoter regions of *LcFT1*, *NIFT1*, the 321-bp insertion, and the

DIFT1 promoter region and excluding the 321-bp insertion (Fig. 3b). Consequently, two specific CREs within the 321-bp insertion of the *DIFT1* promoter were identified (Fig. 3b, Table S3, Xu et al. 2011; Yin et al. 2016). One was an SRF-type CArG-box (CC[A/T]6GG) (Fig. 3c), which is a binding motif for the MADS domain proteins associated with flower formation (Tao et al. 2012). For example, the MADS-box protein SOC1, which activates FT transcription by directly binding to the CArG motif within its promoter (Kou et al. 2022), has the potential to bind to the SRF-type CArG-box, suggesting that a specific

regulation pathway may be involved in flowering of longan (Fig. 3d).

Promoter diversity of tandemly duplicated *NIFT1s* likely contributes to their subfunctionalization in rambutan

We characterized four tandemly replicated *FTI*-like paralogs in rambutan (Fig. 1c). To explore whether these *NIFT1s* are functionally different, we analyzed their coding sequences and determined that *NIFT1-a* shared 95.98%, 94.25%, and 94.25% amino acid identity with *NIFT1-b*, *NIFT1-c*, and *NIFT1-d*, respectively (Fig. S4a), suggesting that their functions are likely similar. Additionally, we investigated the expression patterns across various tissues and noted that *NIFT1-a* showed a distinct expression pattern while the remaining three *NIFT1* shared similar expression profiles, with their main expression occurring in flowers (Fig. 4a, Fig. S4c). We inferred that the differential gene expression pattern between *NIFT1-a* and *NIFT1-b/c/d* was due to differences in the promoter region. Therefore, we employed the phylogenetic shadowing method to investigate the similarity in the *NIFT1-a* and *NIFT1-b/c/d* promoter regions. As results, the *NIFT1-b/c/d* promoter region showed similar changes compared with *NIFT1-a* (Fig. S4b). The evolutionary phylogeny analysis revealed that the *NIFT1-a* promoter sequence was different from the other *NIFT1s*, and the *NIFT1-b*, *NIFT1-c*, and *NIFT1-d* promoter regions shared higher sequence similarity (Fig. 4b). Thus, *NIFT1* genes may be subfunctionalized into two subgroups.

To investigate the variations between the *NIFT1-a* and *NIFT1-b/c/d* promoters, the *cis*-elements in the 2.5-kb *NIFT1-a/b/c/d* promoter sequences were analyzed. In total, 29, 35, 33, and 31 types of CREs were predicted from *NIFT1-a*, *NIFT1-b*, *NIFT1-c*, and *NIFT1-d*, respectively (Fig. 4c). The *NIFT1-b/c/d* promoter sequences

shared four identical CREs (targeted by WUN, two MYBs, and ARE) that were absent in the *NIFT1-a* promoter region. All of these shared CREs in *NIFT1-b/c/d* were stress-response relative regulatory elements (Fig. 4d, e, Table S4, Xu et al. 2011; Yin et al. 2016), suggesting that *NIFT1-b/c/d* may adopt additional regulation in response to abiotic stress.

SNP in the *LcFT1* promoter affects flowering time by influencing the binding affinity of SVP proteins

After comparing the promoter region of *FTI*-like genes among the Sapindaceae species and different *FTI* loci in rambutan, we determined that some *FTI*-like genes may have been subfunctionalized by the evolution of their promoter sequences. Then, we examined whether the change in a single locus of the promoter region could contribute to the regulatory diversity of the *FTI* gene within a single species. Based on previously published resequencing data from 47 lychee germplasms (Hu et al. 2022), 67 SNPs were detected in the *LcFT1* promoter sequences (Fig. 5a). We predicted two SVP binding motifs (intermediate type CA_nG box) in the highly conserved 2.5-kb sequence of the *LcFT1* promoter (Fig. 5a). Notably, we discovered a SNP situated within one of these binding motifs, which was widely shared across 47 distinct lychee varieties (Fig. 5b, c, Table S5). The consensus sequence of the binding motif (CTATACAAA AAGGGA(G/A)ATAA) was located from 1,452 to 1,433 nucleotides upstream from the *LcFT1* start codon, with a G/A SNP positioned at 1,437-bp (Fig. 5b). Interestingly, almost all of the extremely early-maturing lychee varieties (see purple shadow in Fig. 5c) exhibited the G allele, whereas the early-ripening lychee varieties (see blue shadow in Fig. 5c) possessed both genotypes (R represents a heterozygous G/A genotype). In contrast, all late-ripening lychee varieties (see yellow shadow in Fig. 5c) exclusively exhibited the A genotype (Fig. 5c). According

(See figure on next page.)

Fig. 3 Conservation analysis of *FTI*-like promoter sequences in Sapindaceae species. **a** Pairwise alignment of *FTI* nucleotide sequences linked to upstream of the ATG, spanning 3.0 kilobases, and downstream of stop codons, spanning 1.0 kilobase from six Sapindaceae species by using mVISTA (Burdno et al. 2003). The graphical output displays the base pair identity within a sliding window of 75-bp, ranging from 50 to 100%. The upper shows the *FTI*-like gene structure and the bottom with red box shows longan-specific conserved segments in the multiple sequence alignment of different Sapindaceae species by ClustalW2 (Larkin et al. 2007). The approximately 2.5-kb upstream sequence of the ATG in lychee, longan, and rambutan showed a high degree of conservation, marked by blue dotted box **b** Venn diagram assessing the count of CREs disparately among *DIFT1_{pro}^{Del-321}* (2.5-kb upstream from the ATG of *DIFT1* out of longan-specific 321-bp insertion), *DIFT1_{pro}^{In-321}* (longan-specific 321-bp insertion located between 619 and 940 nucleotides upstream from the ATG of *DIFT1*), *LcFT1_{pro}* (2.5-kb upstream from the ATG of *LcFT1*), *NIFT1_{pro}* (2.5-kb upstream from the ATG of *NIFT1-a*). The red triangle highlights the unique CREs of longan-specific conserved segments in *DIFT1_{pro}^{Del-321}*, *DIFT1_{pro}^{In-321}*, *LcFT1_{pro}*, and *NIFT1_{pro}*. **c** *Cis*-regulatory sequences of longan-specific 321-bp insertion are highlighted in b with red triangle. **d** *Cis*-element prediction on the longan-specific 321-bp insertion and red rectangle highlight the TF binding motif (TFBM) of SOC1 using MAST (Timothy et al. 1998), by which sequences with an *E*-value less than 10 are included in the output. The motif logo derived from JASPAR TF binding profile associated with SOC1(MA0054.1) is provided in bottom black box

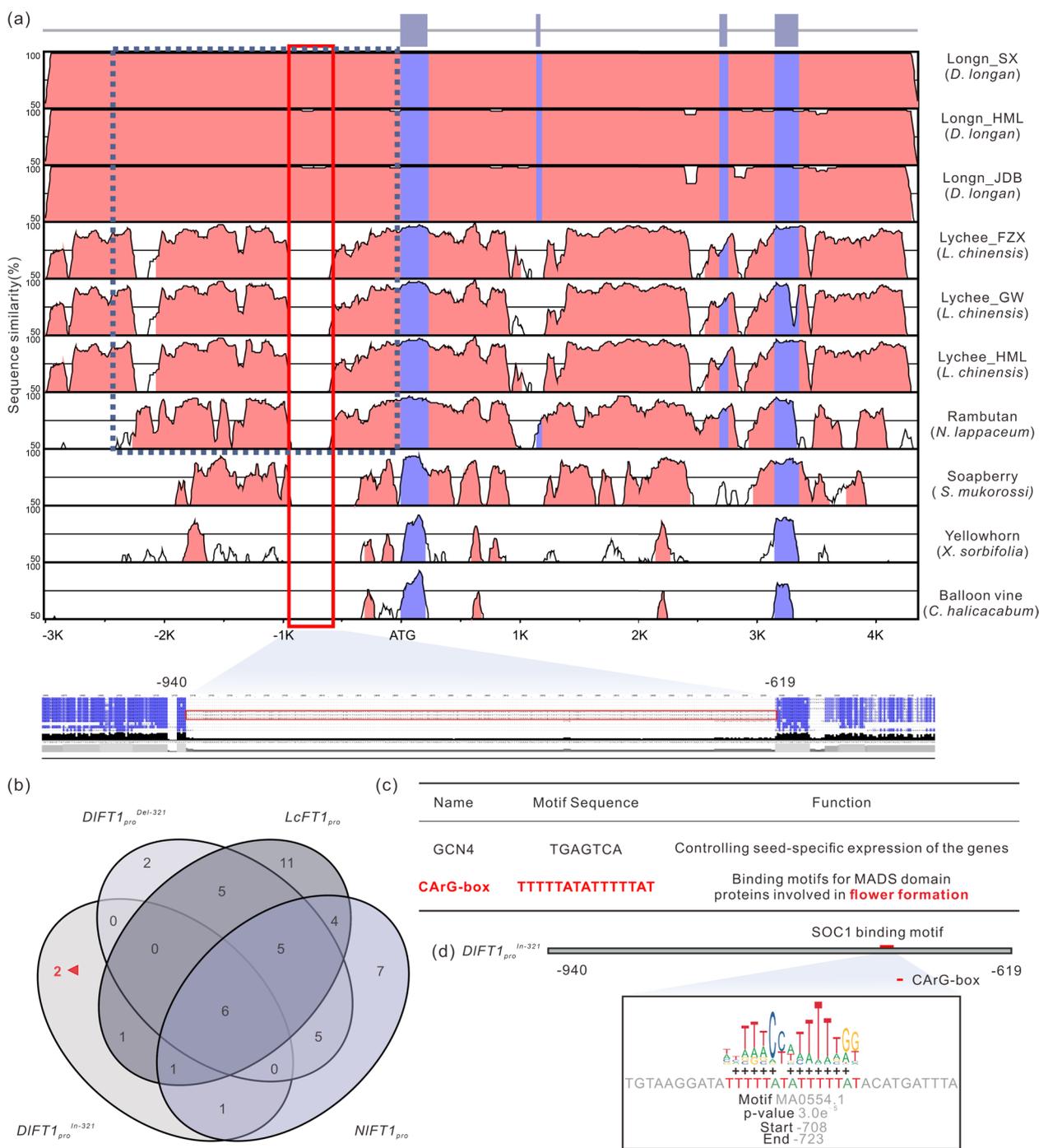


Fig. 3 (See legend on previous page.)

to the single binding site prediction using position weight matrix (PWM) scanning, the A-type motif had a higher binding affinity to SVP than the G-type motif (Fig. 5b).

To confirm whether the G/A SNP affects the binding affinity of the SVP protein, we assessed the binding ability of these two types of motifs compared to lychee

SVP. In total, there were ten SVP homologs in the lychee genome. Among them, *LcSVP9* belongs to the *SVP1* clade, which is associated with delayed flowering (Fig. S6, Liu et al. 2018) and exhibits preferential expression in lychee leaves (Fig. S5) similar to the *Arabidopsis SVP* (Hartmann et al. 2000). Thus, we presume that *LcSVP9*

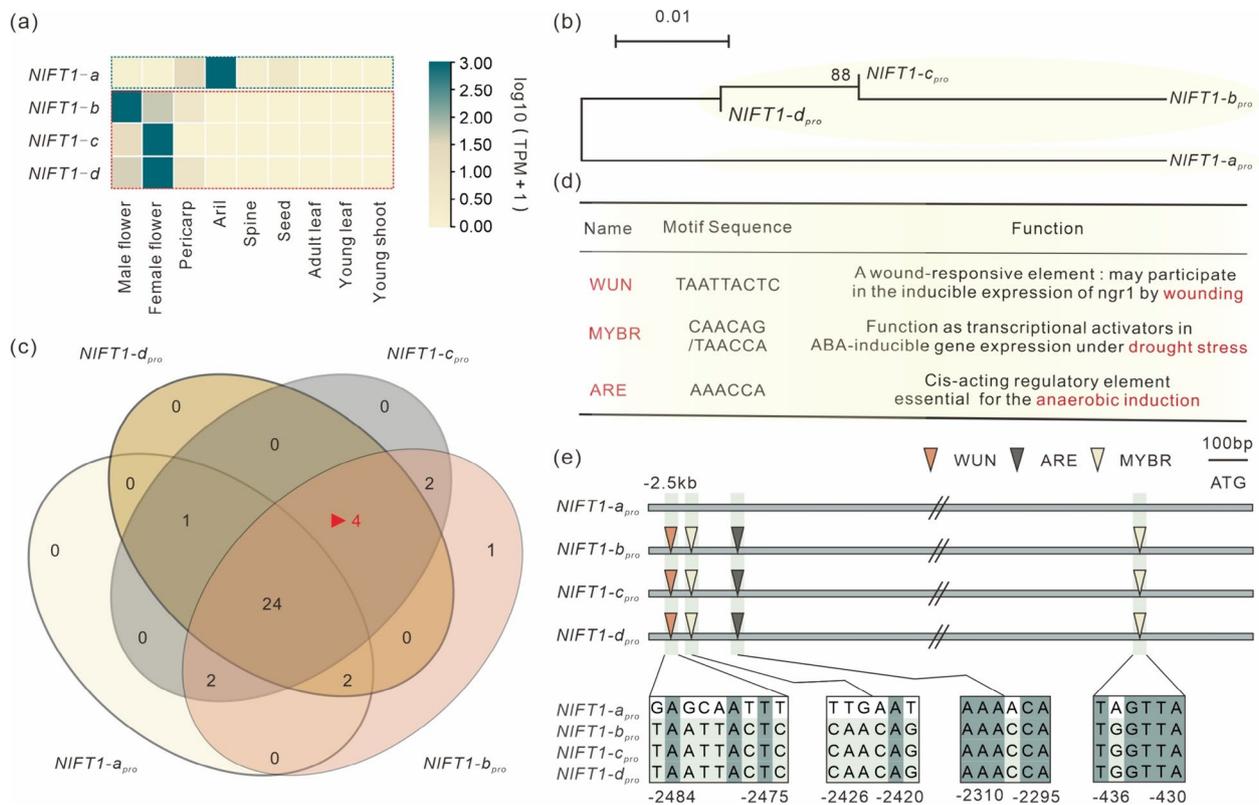


Fig. 4 Scanning the promoter sequence of *FT1s* in rambutan. **a** The gene expression patterns of *FT*-like genes across various tissues in rambutan. **b** A Neighbor Joining phylogenetic analysis of promoter sequence associated with 4 *NIFT1s* in rambutan. **c** Venn diagram compares the number of CREs differentially among 4 *NIFT1s* linked to promoter sequence. The red triangles indicate the count of CREs in the promoter regions of *NIFT1-b*, *NIFT1-c*, and *NIFT1-d* that are different from those in *NIFT1-a*. **d** The promoter regions of *NIFT1-b*, *NIFT1-c*, and *NIFT1-d* contain identical CREs that are distinct from those present in the *NIFT1-a* gene. **e** Identification of CREs in the promoter regions of *NIFT1s*. CREs are depicted as triangles with different colors. The box displays the sequence of each CRE position in the four *NIFT1* genes. In the box there displayed the forward (WUN/ MYB/ARE) or reverse complementary motif (One of the two MYB on the right) sequences within *NIFT1* gene promoter regions

(See figure on next page.)

Fig. 5 SNP-1,437 affected the transcriptional regulation of *LcSVP9* on *LcFT1*. **a** Distribution of SNP in *LcFT1* promoter sequence of 47 lychee germplasms. The dark purple square exhibits the binding motif of SVP protein predicted by MAST. **b** The specific binding sites of SVP in two genotypes were displayed, and the *p*-value indicated the comparison of their binding ability. Motif logo derived from JASPAR TF binding profile associated with SVP (MA0555.1) is provided above and below of the gray background, respectively. **c** Homozygous and heterozygous SNPs from each accession in SVP binding site. The red background G illustrates the G homozygous genotype at the SVP-bound SNP site, while the green background A represents the A homozygous genotype. The gray background R shows the G/A heterozygous genotype. EEM is short for "extremely early maturing"; EM is short for "early maturing"; LM is short for "late maturing." **d** The phenotype of *Arabidopsis* with the overexpression of lychee *LcSVP9* genes. All plants were grown under long-day conditions with 16 h of light and 8 h of darkness at 23°C. The image depicts the phenotype after 24 days of sowing. Col (WT) represents the wild-type *Arabidopsis* with a Col-0 background. OE represents the transgenic lines with *LcSVP9* overexpression. T₂ generation lines with *LcSVP9* gene are represented by numbers 1, 4 and 7, with a scale of 1 cm. **e** The statistics of the number of rosette leaves during flowering for *svp-31* mutants, wild-type and *LcSVP9* transgenic *Arabidopsis* plants, with five plants used for each line. The statistical analysis for *svp-31* mutants commenced around 24 days post-sowing, while wild-type *Arabidopsis* started approximately 27 days after sowing, and the OE-*LcSVP9* lines began between 28 and 42 days after sowing. The data is presented as mean ± SD. One-way analysis of variance (ANOVA) was performed. *P* value: ****P* < 0.001, *****P* < 0.0001. **f** Expression levels of *LcSVP9* and *LcFT1* in response to ROS treatment (Lu et al. 2020a) and low temperature (Zhang et al. 2017) during floral induction according to the RNA-seq data in leaves of lychee. TPM (transcripts per million) was used to indicate gene expression levels or transcript accumulated levels. The \log_2 ratio of fold change ($\log_2\text{FC}$) of the gene expression value between control and treatment was calculated according to the treatment timepoint. **g** EMSA was performed to compare the binding affinity of *LcSVP9* in the promoter region of *LcFT1^{A type}* and *LcFT1^{G type}* containing the SNP-1,437. **h** Transient dual-luciferase expression assay. The control group included an empty vector, and the data are presented as the mean ± SD derived from three biologically independent samples (black dots). One-way ANOVA *p* value: ns = 0.097, ****P* < 0.001b

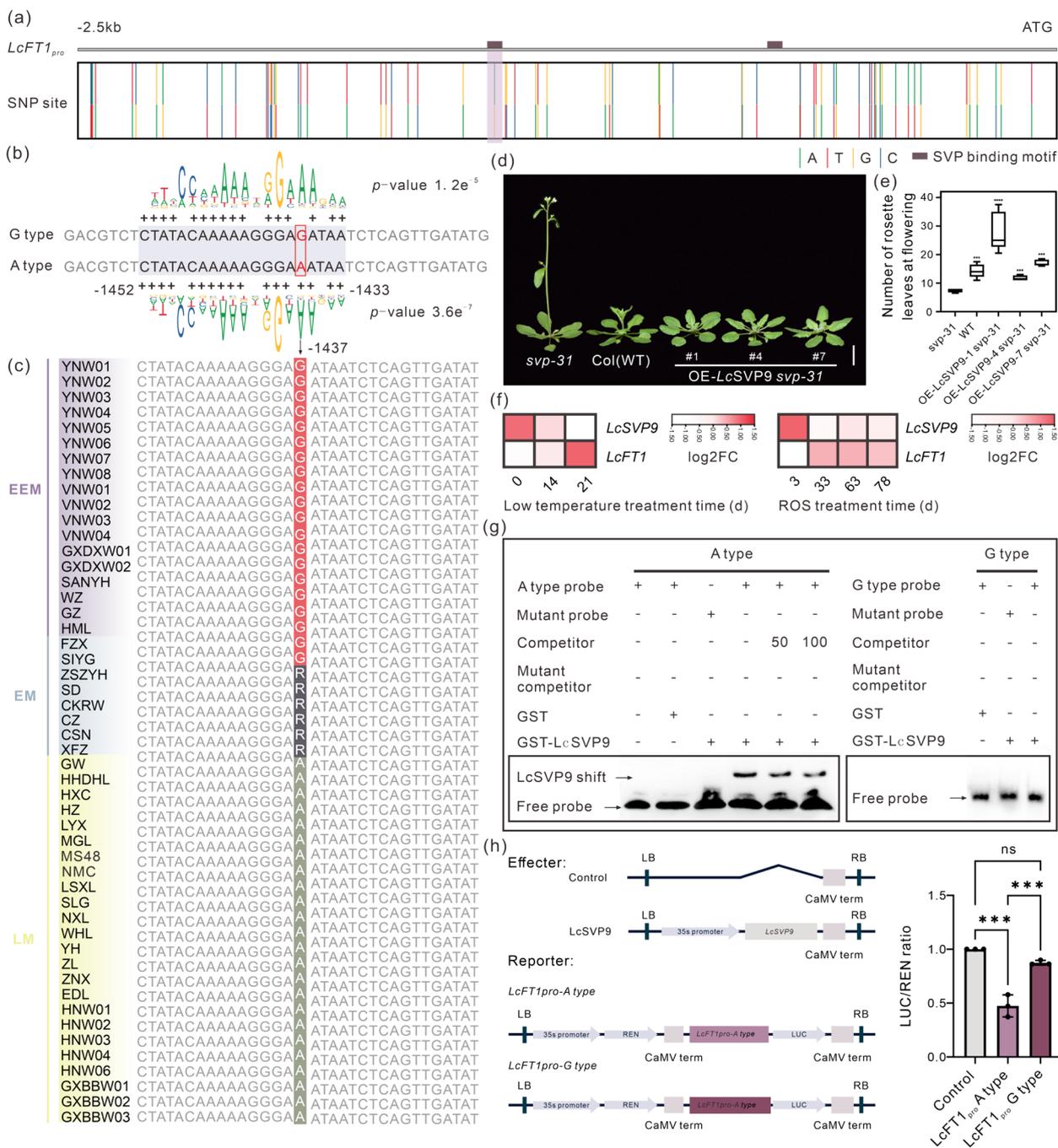


Fig. 5 (See legend on previous page.)

may function similarly to *AtSVP* in delaying the flowering time of plants. Also, *LcSVP9* exhibited the highest expression level in lychee leaves among these ten *SVP* genes (Fig. S5a). We hypothesize that *LcSVP9* could be a key regulator within the *SVP* gene family. To confirm the function of *LcSVP9* in regulating flowering, we introduced the *LcSVP9* gene into *svp-31* mutant *Arabidopsis*

under the control of the CaMV 35S promoter. Under long-day conditions, plants overexpressing *LcSVP9* initiated flowering with 11.8 to 28.6 rosette leaves, showing a significantly delayed flowering time compared with the *svp-31* mutant control (7.4 rosette leaves) and a timing similar to the WT control (14.3 rosette leaves) (Fig. 5d, e, Fig. S5c). This suggests that the expression of *LcSVP9*

impacted flowering time, functioning as a floral repressor like *AtSVP*. We attempted to determine the expression levels of known flowering time genes in *LcSVP9* over-expression plants. We found that the expression level of *AtFT* were significantly reduced, indicating that *LcSVP9* regulates flowering via repressing *AtFT* (Fig. S5d). This finding is consistent with the observation that levels of *AtFT* expression were elevated in *svp-31* mutants (Fig. S5d).

Furthermore, we observed a reduction in *LcSVP9* levels in leaves subjected to reactive oxygen species (ROS) and low-temperature treatments, while *LcFT1* exhibited contrasting expression patterns (Fig. 5f). Thus, we propose that *LcSVP9* may repress the expression of *LcFT1*. Initially, we performed an electrophoretic mobility shift assay (EMSA) to investigate the potential impact of the -1,437 SNP on the binding affinity between *LcSVP9* and the *LcFT1* promoter. The results showed that the *LcSVP9* recombinant protein bound to the A-type motif but did not bind to the G-type motif (Fig. 5g), confirming that the natural variation of nucleotide position -1,437 in the *LcFT1* promoter are correlated with differential binding of *LcSVP9*. Subsequently, we conducted a dual-luciferase reporter assay to investigate the transcriptional activity of *LcSVP9* on *LcFT1*^{A-type} and *LcFT1*^{G-type} promoters. The results showed *LcSVP9* significantly reduced the expression of the LUC reporter by interacting with *LcFT1*^{A-type} promoter (Fig. 5h). In contrast, *LcSVP9* lost its transcriptional repression on LUC driven by the *LcFT1*^{G-type} promoter (Fig. 5h). Therefore, the presence of the SVP-binding motif of *LcFT1*^{A-type} promoters is crucial for the inhibition of *LcFT1* expression by *LcSVP9*. We then analyzed data from ten lychee leaf samples collected in the first half of December 2014, sourced from public transcriptome data (Lu et al. 2022). The results indicated that the expression of *LcFT1* was significantly higher in the early-maturing variety (EM) compared to the late-maturing variety (LM) during the flowering induction period in most lychee (Fig. 5Sb). These findings are consistent with our initial expectations. The collective data suggest that the natural variation at nucleotide position -1,437 of the lychee *FT1* promoter influences the binding affinity of *LcSVP9*, a negative regulator, thereby leading to differential expression of *LcFT1*, which is likely implicated in regulating lychee flowering time.

Discussion

The number of PEBP genes varied greatly across the Sapindaceae species, ranging from nine to twelve copies (Fig. 1b). Most eudicot species typically have fewer than ten PEBP genes, with a few exceptions, including turnips (*Brassica rapa*), soybean (*Glycine max*), tomato (*Solanum lycopersicum*), and lotus (*Nelumbo nucifera*),

where the number exceeds ten due to additional whole genome duplication (WGD) events (Liu et al. 2016). No further recent WGD events have been observed in Sapindaceae, with the exception of the two WGD events following the γ event that are shared among all dicot plants (Zheng et al. 2022; Xue et al. 2022; Hu et al. 2022; Wang et al. 2022). Therefore, local duplications are more noticeable in the expansion of the PEBP-family genes among the Sapindaceae genomes. Phylogenetic analysis indicated that *FT1*-like, *FT2*-like, and *FT3*-like were present in lychee, longan, rambutan, and soapberry, suggesting that duplication occurred before the species diverged (Fig. 1a). Despite that the genes located upstream and downstream of *FT2*-like were highly collinear across all six Sapindaceae species, a homologous *FT2* gene was not identified in yellowhorn. This observation indicates that the *FT2*-like clade may be traced back to a common ancestor in the subfamily Sapindoideae, which encompasses the other five Sapindaceae species, excluding yellowhorn, which is in the subfamily Dodonoideae. No *FT2*-like flowering inhibitory factor has been identified in yellowhorn (Fig. 2a), which may be linked to the absence of juvenile phase characteristics in this species (Yao et al. 2013). The *MFT*-like was amplified in balloon vine compared to the other five Sapindaceae species (Fig. 1b). Gene collinearity analysis showed that balloon vine *MFT1*-like was arranged in tandem with six copies (Fig. S2a). The expression analysis of the six tandemly duplicated *CaMF1* genes demonstrated a predominant expression pattern in the seeds (Fig. S2b). Previous studies have reported that the *CaMF1* homolog *MFT* is involved in promoting seed germination (Danilevskaya et al. 2008). Thus, we hypothesized that the function of *CaMF1* in balloon vine may also be associated with seed germination, specifically the ability of balloon vine seeds to maintain their vitality and germination capacity, even after prolonged storage, in contrast to the limited preservation capacity of lychee and longan seeds (Zhu et al. 2019; Johnston et al. 1979). This distinct characteristic is likely attributed to the amplification of *CaMFT1*.

A comparative analysis revealed that the 321-bp insertion in the longan promoter region, which was absent in the closely related species lychee and rambutan, contained an SRF-type CARG-box domain which would be bound by SOC1 (Fig. 3d). Furthermore, previous research has suggested that SOC1 functions as a transcriptional activator of *FT* in soybean leaves by directly binding to the CARG motif within the *FT* promoter (Kou et al. 2022). We speculate that SOC1 regulates *DIFT1* expression in longan leaves by binding to the CARG box on the *DIFT1* promoter within the

321-bp insertion specific to longan, thereby controlling flowering time. Longan exhibits unique floral induction in response to $KClO_3$ application, a capability unparalleled in other Sapindaceae species (Matsumoto et al. 2007). Therefore, we speculate that SOC1 may respond to $KClO_3$ treatment and bind to a species-specific motif in the *DIFT1* promoter, accelerating the expression of *DIFT1* to promote the transition to reproductive development and flowering. Further analysis should be performed to validate this possibility.

The *NIFT1* gene undergoes tandem duplication in rambutan, resulting in four copies (Fig. 1c). Their regulatory sequences and expression patterns in different biological processes were in two distinct branches and one *NIFT1-a* clade, and the other three genes (*NIFT1-b*, *NIFT1-c*, and *NIFT1-d*) clustered in a single clade (Fig. 4a, b, Fig. S4c). This result suggests that the expression of the *NIFT1-b*, *NIFT1-c*, and *NIFT1-d* genes may be regulated by common CREs. We found that *NIFT1-b*, *NIFT1-c*, and *NIFT1-d* contained three types of identical CREs that are distinct from those found in the *NIFT1-a* gene. These CREs were associated with stress responses, including the WUN motif for wound responsiveness, the MYB-target motifs, and the anaerobic response (Fig. 4c, d). A previous study showed that the WUN-motif elements are bound by NAC transcription factors (Huang et al. 2017). In response to drought, the NAC-type transcription factor VASCULAR PLANT ONE-ZINC FINGER 1 (SIVOZ1) directly binds to the promoter of the major flowering-integrator gene *SINGLE FLOWER TRUSS* (*SFT*), an FT orthologue, thereby promoting the transition to flowering in tomato (Chong et al. 2022). Thus, we speculate that the three rambutan *NIFT1* genes (*NIFT1-b*, *NIFT1-c*, and *NIFT1-d*), resulting from gene replication, regulate flowering through stress response pathways.

A comparative analysis of the regulatory regions was conducted to investigate inter-specific variations in different lychee germplasms. Variations in the CREs were identified within the regulatory regions of the *FTI*-like gene in lychee. Our experiments showed that natural variation occurring at position -1,437 within the *LcFTI* promoter correlates with differential binding of LcSVP9 transcription factors, potentially influencing lychee flowering time. This variation in the CREs offers a potential molecular marker for lychee breeding. Modifications in CREs are linked to alterations in diverse agronomic traits (Rodríguez-Leal et al. 2017; Hendelman et al. 2021). For example, a recent study utilized *cis*-regulatory editing to alter the transcription of *WOX9* in tomatoes, which subsequently impacted floral development and mitigated undesired outcomes (Hendelman et al. 2021). The recent application of CRISPR/Cas9-mediated genome editing

technology to improve lychee varieties (Wang et al. 2023a) offers promising possibilities for biotechnological engineering of Sapindaceae species. For example, by modifying the regulatory *cis*-element that controls *FT*-like gene expression, the flowering period and the period of fresh fruit supply in commercially important fruit trees in Sapindaceae can be extended.

Materials and methods

Plant materials

This study utilized eight-year-old *Dimocarpus longan* cv. Shixia trees. The experimental trees were cultivated at South China Agricultural University in Guangzhou, China (lat. 23.1568° N, lon. 113.3537° E). The terminal shoots of three trees were sampled to obtain adult leaves, with three biological replicates collected for each sample.

Identification of PEBP family members from six Sapindaceae species

The conserved domain (PF01161) of the PEBP was acquired from <http://pfam.xfam.org/> (Mistry et al. 2021; Wang et al. 2023b). TBtools (Chen et al. 2020) was utilized to perform HMMER analysis, allowing for the retrieval of protein data and the identification of potential members of the PEBP family. The identified genes were screened and validated using two online tools: <https://www.ebi.ac.uk/Tools/pfa/pfamscan/> (Madeira et al. 2019) and NCBI Conserved [tps://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi](https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (Lu et al. 2020b). Genes without complete PEBP domains were eliminated for further analysis.

Phylogenetic analysis and multiple sequence alignment

The PEBP protein sequences in other 46 seed plants species characterized by transgenic approaches, were retrieved from the Phytozome, GenBank and Arabidopsis Information Resource (TAIR) database (Table S2). A Neighbor Joining phylogenetic analysis was conducted on a total of 156 PEBP homologs across 52 flowering plants in MEGA X (V.10.2.6) (Kumar et al. 2018). The clades were assigned based on the observed clustering pattern among genes in Arabidopsis. The Taxonomy tool (Schoch et al. 2020) [available at https://www.ncbi.nlm.nih.gov/](https://www.ncbi.nlm.nih.gov/) was utilized to conduct a species phylogenetic tree. Pairwise alignments of *FTI*-like promoter sequences from the six Sapindaceae species were created using mVISTA [Shuffl://genome.lbl.gov/vista](http://genome.lbl.gov/vista) (Brudno et al. 2003). The conserved regions were aligned using ClustalW2 (V.2.1) (Larkin et al. 2007).

Syntenic analysis among six Sapindaceae plants

Syntenic gene pairs were identified among six plants from the Sapindaceae family using [JCVI://github.com/](https://github.com/JCVI/)

tanghaibao/jcvi/wiki/MCscan-(Python-version))(Tang et al. 2008). The identification of syntenic blocks for each pair of species was performed using the 'jcvi.compara.catalog.ortholog -cscore=0.7' parameter.

Analysis of CREs in the lychee, longan, and rambutan

The promoter sequence of the *FT1*-like gene was obtained by extracting the 2500-bp sequence upstream of the translation initiation site. To identify potential *cis*-regulatory elements, promoter sequences were predicted using Plformatcs.ps-b.ugent.be/webtools/plantcare/html/ (Lescot et al. 2002) [aps://meme-suite.org/meme/tool-s/mast](https://meme-suite.org/meme/tool-s/mast) (Timothy et al. 1998), by which sequences with an *E*-value less than 10 were included in the output. The motif profile of SVP protein was derived from [://jaspar.genereg.net/](http://jaspar.genereg.net/) (Wasserman et al. 2004).

RNA-Seq data analysis

All raw pair-ends reads were trimmed utilizing trimmomatic (V.0.36) (Bolger et al. 2014) to remove the adapters and low-quality bases. Subsequently, cleaned reads were aligned to the reference genomes using the STAR alignment tool (V.2.7.10b) (Dobin et al. 2013). The read counts and Transcripts Per Million Reads (TPM) for the genes were computed using the StringTie (V.2.2.1) (Kovaka et al. 2019).

Folding prediction for the DIFT1 and DIFT2 proteins

The DeepView/Swiss-pdb viewer (V.4.0.1) (Schwede et al. 2003) was used to predict the structure of DIFT1 and DIFT2 proteins, based on the Arabidopsis FT structure (PDB ID: 1WKP). The three-dimensional protein structures of DIFT1 and DIFT2 were obtained by utilizing the SWISSmodel.expasy.org (Waterhouse et al. 2018).

Constructs for transgenic plants

The coding sequences of *DIFT1*, *DIFT2* and *LcSVP9* were cloned and inserted into the pEarleyGate 201 vector, which harbored the CaMV 35S promoter. The recombinant plasmids pEarleyGate 201-*DIFT1*, pEarleyGate 201-*DIFT2* and pEarleyGate 201-*LcSVP9* were introduced into *Agrobacterium tumefaciens* strain GV3101. Plants were transformed through the floral dip method and subsequently screened for resistance to BASTA. The T-DNA insertion line *svp-31* (SALK_026551) was acquired from the Arabidopsis Biological Resource Center (ABRC) (Alonso et al. 2003).

Quantitative gene expression analysis

TRIzol, a product from Thermo Fisher Scientific, was utilized for the isolation of total RNA from longan leaves. TransScript® One-Step gDNA Removal and cDNA Synthesis SuperMix (TransGen, AT311-02) was utilized to

synthesize cDNA from RNA samples with an amount of 0.5 µg. The qRT-PCR primers were synthesized by Sangon Co. Ltd. (Shanghai, China), using Primer 5.0 (Singh et al. 1998) for their design. The 2× GoTaq® qPCR Master Mix from Promega was utilized for conducting the qRT-PCR experiment. The $-\Delta CT$ and $2^{-\Delta\Delta CT}$ calculation was employed to normalize expression levels. The primer sequences utilized can be found in Table S8.

Dual-luciferase reporter assay

The coding sequence of *LcSVP9* was cloned and inserted into the pGreenII-62-SK vector, serving as the effector. The pGreen II-0800-Luc vector was utilized to insert the promoter fragments of *LcFT1^{A-type}* and *LcFT1^{G-type}* (Fig. 5e), which were subsequently employed as reporters. The empty vector served as the negative control. The effector and reporter plasmids were separately transformed into *Agrobacterium* strains GV3101 and subsequently co-transformed into *N. benthamiana* leaves. The OD600 of the bacterial cultures was adjusted to 1.0, with an effector to reporter ratio of 9:1. Dual luciferase reporter gene assay kit (YEASEN) was employed to measure the relative levels of luciferase activity, in accordance with the instructions provided by the manufacturer.

EMSA

The coding sequence of *LcSVP9* was cloned and inserted into the pGEX4T-1 vector. The recombinant protein GST-*LcSVP9* was purified in *Escherichia coli* with Glutathione and Nibeads. The oligonucleotide probes were biotin-labeled at the 3' end. The competition analysis reactions were conducted with the unlabeled probes. EMSA experiments were conducted utilizing the LightShift Chemiluminescent EMSA Kit (Thermo Fisher, 20148X).

Abbreviations

Bp	Base pair
CREs	<i>Cis</i> -Regulatory elements
EMSA	Electrophoretic mobility shift
<i>FT</i>	<i>Flower Locus T</i>
Kb	Kilobases
KClO ₃	Potassium chlorate
LUC	Firefly luciferase
<i>MFT</i>	<i>MOTHER OF FT AND TFL1</i>
NS	Not significant
OE	Overexpressing
PEBP	Phosphatidylethanolamine-binding protein
qRT-PCR	Quantitative real-time polymerase chain reaction
REN	Renilla luciferase
ROS	Reactive oxygen species
<i>SVP</i>	<i>SHORT VEGETATIVE PHASE</i>
<i>SOC1</i>	<i>SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1</i>
SNP	Single nucleotide polymorphism
SD	Standard deviation
TPM	Transcripts per million
<i>TFL1</i>	<i>TERMINAL FLOWER 1</i>
UTR	Untranslated Region
WGD	Whole genome duplication

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s43897-024-00104-4>.

Supplementary Material 1.

Acknowledgements

The authors express gratitude to numerous colleagues and collaborators for their contributions to the work described in this paper.

Authors' contributions

X.H., Y.L. and R.X. were responsible for the conception and design of this project. X.H. and W.W. carried out the experimental work, prepared the necessary materials, and drafted the manuscript. H.L. and F.W. contributed to data analysis. The manuscript was reviewed by Y.H., C.C., H.Z., and J.X.

Funding

This work is supported by funding from Science and Technology Projects in Guangzhou (Grant No. 2023A04J1455), the Key-Area Research and Development Program of Guangdong Province (#2022B0202070003). This work is also supported by the National Natural Science Foundation of China (#32072547, #32372665).

Availability of data and materials

All assemblies with annotations underlying this article were available in the SapBase <http://www.sapindaceae.com/Download.html>. Transcriptome data from three tissues (young shoot, young leaf, and adult leaf) of rambutan were downloaded from the NCBI SRA database (SRR14560235, SRR14560276, SRR14560287). RNA-Seq data from lychee leaves exposed to reactive oxygen species (ROS) and low-temperature treatments, along with lychee leaf samples collected in early December 2014, were deposited in the NCBI BioProjects PRJNA1045234, PRJNA1045227, and PRJNA766599, respectively. The remaining RNA-Seq were available at the SapBase <http://www.sapindaceae.com/Gene-Expression-V2/GeneExpression.html>. The lychee variants VCF files were downloaded from the Mendeley <https://data.mendeley.com/datasets/v37bv5jt6g/1>.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors have read and given their approval to the final version of the manuscript.

Competing interests

Prof. Rui Xia is a member of the Editorial Board for *Molecular Horticulture*. He was not involved in the journal's review of, and decisions related to, this manuscript. The other authors declare no competing interests.

Received: 26 February 2024 Accepted: 14 June 2024

Published online: 16 July 2024

References

- Adrian J, Farrona S, Reimer JJ, Albani MC, Coupland G, Turck F. Cis-Regulatory elements and chromatin state coordinately control temporal and spatial expression of FLOWERING LOCUS T in Arabidopsis. *Plant Cell*. 2010;22(5):1425–40.
- Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, et al. Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science*. 2003;301(5633):653–7.
- Blackman BK, Strasburg JL, Raduski AR, Michaels SD, Rieseberg LH. The role of recently derived FT paralogs in sunflower domestication. *Curr Biol*. 2010;20:629–35.
- Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–20.
- Brudno M, Malde S, Poliakov A, Do CB, Couronne O, Dubchak I, et al. Global alignment: finding rearrangements during alignment. *Bioinformatics*. 2003;19:i54–62.
- Chen HB, Huang HB. Low temperature requirements for floral induction in lychee. II International Symposium on Lychee, Longan, Rambutan and Other Sapindaceae Plants. 2003;665:195–202.
- Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol Plant*. 2020;13(8):1194–202.
- Chong L, Xu R, Huang P, Guo P, Zhu M, Du H, et al. The tomato OST1–VOZ1 module regulates drought-mediated flowering. *Plant Cell*. 2022;34:2001–18.
- Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I, et al. FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science*. 2007;316(5827):1030–3.
- Danilevskaia ON, Meng X, Hou Z, Ananiev EV, Simmons CR. A genomic and expression compendium of the expanded PEBP gene family from maize. *Plant Physiol*. 2008;146(1):250–64.
- Ding F, Zhang S, Chen H, Su Z, Zhang R, Xiao Q, et al. Promoter difference of LcFT1 is a leading cause of natural variation of flowering timing in different litchi cultivars (*Litchi chinensis* Sonn). *Plant Sci*. 2015;241:128–37.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15–21.
- Feng Y, Wang Y, Zhang G, Gan Z, Gao M, Lv J, et al. Group-C/S1 bZIP heterodimers regulate MdIPT5b to negatively modulate drought tolerance in apple species. *The Plant J*. 2021;107:399–417.
- Harig L, Beinecke FA, Oltmanns J, Muth J, Müller O, Rüping B, et al. Proteins from the FLOWERING LOCUS T-like subclade of the PEBP family act antagonistically to regulate floral initiation in tobacco. *The Plant J*. 2012;72:908–21.
- Hartmann U, Höhmann S, Nettesheim K, Wisman E, Saedler H, Huijser P. Molecular cloning of SVP: a negative regulator of the floral transition in Arabidopsis. *Plant J*. 2000;21(4):351–60.
- Hendelman A, Zebell S, Rodriguez-Leal D, Dukler N, Robitaille G, Wu X, et al. Conserved pleiotropy of an ancient plant homeobox gene uncovered by cis-regulatory dissection. *Cell*. 2021;184(7):1724–1739.e16.
- Hu G, Feng J, Xiang X, Wang J, Salojärvi J, Liu C, et al. Two divergent haplotypes from a highly heterozygous lychee genome suggest independent domestication events for early and late-maturing cultivars. *Nat Genet*. 2022;54(1):73–83.
- Huang Y, Li T, Xu ZS, Wang F, Xiong AS. Six NAC transcription factors involved in response to TYLCV infection in resistant and susceptible tomato cultivars. *Plant Physiol Biochem*. 2017;120:61–74.
- Huang X, Chen S, Li W, Tang L, Zhang Y, Yang N, et al. ROS regulated reversible protein phase separation synchronizes plant flowering. *Nat Chemical Biology*. 2021;17:549–57.
- Jang S, Torti S, Coupland G. Genetic and spatial interactions between FT, TSF and SVP during the early stages of floral induction in Arabidopsis. *The Plant J*. 2009;60:614–25.
- Jiang H, Shi Y, Liu J, Li Z, Fu D, Wu S, et al. Natural polymorphism of ZmICE1 contributes to amino acid metabolism that impacts cold tolerance in maize. *Nat Plants*. 2022;8:1176–90.
- Johnston S, Murray D, Williams J. Germination and emergence of balloonvine (*Cardiospermum halicacabum*). *Weed Sci*. 1979;27:73–6.
- Kardailsky I, Shukla VK, Ahn JH, Dagenais N, Christensen SK, Nguyen JT, et al. Activation tagging of the floral inducer FT. *Science*. 1999;286(5446):1962–5.
- Kobayashi Y, Kaya H, Goto K, Iwabuchi M, Araki T. A pair of related genes with antagonistic roles in mediating flowering signals. *Science*. 1999;286:1960–2.
- Kong F, Liu B, Xia Z, Sato S, Kim BM, et al. Two coordinately regulated homologs of FLOWERING LOCUS T are involved in the control of photoperiodic flowering in soybean. *Plant Physiol*. 2010;154:1220–31.
- Kou K, Yang H, Li H, Fang C, Chen L, Yue L, et al. A functionally divergent SOC1 homolog improves soybean yield and latitudinal adaptation. *Curr Biol*. 2022;32:1728–42.
- Kovaka S, Zimin AV, Pertea GM, Razaghi R, Salzberg SL, Pertea M. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol*. 2019;20(1):278.

- Kumar S, Stecher G, Li M, Nkryaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution*. 2018;35:1547.
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 20. *Bioinformatics*. 2007;23(21):2947–8.
- Lee R, Baldwin S, Kenel F, McCallum J, Macknight R. FLOWERING LOCUS T genes control onion bulb formation and flowering. *Nat Commun*. 2013;4:2884.
- Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res*. 2002;30(1):325–7.
- Li J, Zhu R, Zhang M, Cao B, Li X, et al. Natural variations in the PbCPK28 promoter regulate sugar content through interaction with PbTST4 and PbVHA-A1 in pear. *The Plant J*. 2023;114:124–41.
- Li, JW, Chen CJ, Zeng ZZ, Wu FQ, Feng JT, et al. SapBase: A central portal for functional and comparative genomics of Sapindaceae species. *J Integr Plant Biol*. 2024. <https://doi.org/10.1111/jipb.13680>.
- Liu YY, Yang KZ, Wei XX, Wang XQ. Revisiting the phosphatidylethanolamine-binding protein (PEBP) gene family reveals cryptic FLOWERING LOCUS T gene homologs in gymnosperms and sheds new light on functional evolution. *New Phytol*. 2016;212(3):730–44.
- Liu X, Sun Z, Dong W, Wang Z, Zhang L. Expansion and Functional Divergence of the SHORT VEGETATIVE PHASE (SVP) Genes in Eudicots. *Genome Biol Evol*. 2018;10(11):3026–37.
- Lu X, Yu S, Lü P, Chen H, Zhong S, Zhou B. Genome-wide transcriptomic analysis reveals a regulatory network of oxidative stress-induced flowering signals produced in litchi leaves. *Genes*. 2020a;11:324.
- Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, et al. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res*. 2020b;48(D1):D265–8.
- Lu X, Lü P, Liu H, Chen H, Pan X, Liu P, et al. Identification of chilling accumulation-associated genes for litchi flowering by transcriptome-based genome-wide association studies. *Front Plant Sci*. 2022;13:819188.
- Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res*. 2019;47(W1):W636–41.
- Mathieu J, Warthmann N, Küttner F, Schmid M. Export of FT protein from phloem companion cells is sufficient for floral induction in Arabidopsis. *Curr Biol*. 2007;17:1055–60.
- Matsumoto TK, Nagao MA, Mackey B. Off-season flower induction of longan with potassium chlorate, sodium chlorite, and sodium hypochlorite. *Hort Technology*. 2007;17(3):296–300.
- Menzel C. Lychee, its origin, distribution and production around the world. Australia: West Aust Nut Tree Crop Assoc Yearbook. 1995;19:41–8.
- Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, et al. Pfam: The protein families database in 2021. *Nucleic Acids Res*. 2021;49(D1):D412–9.
- Moon J, Suh SS, Lee H, Choi KR, Hong CB, Paek NC, et al. The SOC1 MADS-box gene integrates vernalization and gibberellin signals for flowering in Arabidopsis. *The Plant J*. 2003;35:613–23.
- Nan H, Cao D, Zhang D, Li Y, Lu S, Tang L, et al. GmFT2a and GmFT5a redundantly and differentially regulate flowering through interaction with and upregulation of the bZIP transcription factor GmFDL19 in soybean. *PLoS one*. 2014;9:e97669.
- Ng M, Yanofsky M. Function and evolution of the plant MADS-box gene family. *Nat Rev Genet*. 2011;2:186–95.
- Pan T, He M, Liu H, Tian X, Wang Z, Yu X, et al. Transcription factor bZIP65 delays flowering via suppressing Ehd1 expression in rice. *Mol Breeding*. 2022;42:63.
- Pin PA, Benlloch R, Bonnet D, Wremerth-Weich E, Kraft T, Gielen JJ, et al. An antagonistic pair of FT homologs mediates the control of flowering time in sugar beet. *Science*. 2010;330(6009):1397–400.
- Rodríguez-Leal D, Lemmon ZH, Man J, Bartlett ME, Lippman ZB. Engineering Quantitative Trait Variation for Crop Improvement by Genome Editing. *Cell*. 2017;171(2):470–480.e8.
- Ruan B, Shang L, Zhang B, Hu J, Wang Y, Lin H, et al. Natural variation in the promoter of TGW2 determines grain width and weight in rice. *New Phytol*. 2020;227:629–40.
- Schoch CL, Ciufo S, Domrachev M, Hotton CL, Kannan S, Khovanskaya R, et al. NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database (Oxford)*. 2020;2020:baaa062.
- Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res*. 2003;31(13):3381–5.
- Shaari AR, Osman MS, Abidin M Z. Aspects on research and production of rambutan in Malaysia. In: Annette S, editor. German: German Foundation for International Development; 1983.p.186–193.
- Shahrajabian MH, Sun W, Khoshkharam M, Cheng Q. Rambutan, a tropical plant with ethno-pharmaceutical properties. *Agrociencia*. 2020;54(1):121–8.
- Singh VK, Mangalam AK, Dwivedi S, Naik S. Primer premier: program for design of degenerate primers from a protein sequence. *Biotechniques*. 1998;24(2):318–9.
- Srikanth A, Schmid M. Regulation of flowering time: all roads lead to Rome. *Cell Mol Life Sci*. 2011;68:2013–37.
- Subhadrabandhu S, Yapwattanaphun C. Regulation off-season flowering of longan in Thailand. I International Symposium on Litchi and Longan. 2000;558:193–8.
- Swinnen G, Goossens A, Pauwels L. Lessons from Domestication, Targeting Cis-Regulatory Elements for Crop Improvement. *Trends Plant Sci*. 2016;21:506–15.
- Tang H, Bowers JE, Wang X, Ming R, Alam M, Paterson AH. Synteny and colinearity in plant genomes. *Science*. 2008;320(5875):486–8.
- Tang L, Li G, Wang H, Zhao J, Li Z, Liu X, et al. Exogenous abscisic acid represses rice flowering via SAPK8-ABF1-Ehd1/Ehd2 pathway. *J Adv Res*. 2023;59:35–47.
- Tao Z, Shen L, Liu C, Liu L, Yan Y, Yu H. Genome-wide identification of SOC1 and SVP targets during the floral transition in Arabidopsis. *The Plant J*. 2012;70:549–61.
- Timothy L, Bailey, Michael G. Combining evidence using p-values: application to sequence homology searches. *Bioinformatics*. 1998;14(1):48–54.
- Tindall, HD. Sapindaceae fruits: botany and horticulture. *Horticultural Review*. 1994;16:143–96.
- Wang J, Li J, Li Z, Liu B, Zhang L, Guo D, et al. Genomic insights into longan evolution from a chromosome-level genome assembly and population genomics of longan accessions. *Hortic Res*. 2022;9:uhac021.
- Wang S, Wang G, Li H, Li F, Wang J. Agrobacterium tumefaciens-mediated transformation of embryogenic callus and CRISPR/Cas9-mediated genome editing in 'Feixixiao' litchi. *Hortic Plant J*. 2023a;9:947–57.
- Wang J, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu S, Marchler GH, et al. The conserved domain database in 2023. *Nucleic Acids Res*. 2023b;51(D1):D384–8.
- Wasserman WW, Sandelin A. Applied bioinformatics for the identification of regulatory elements. *Nat Rev Genet*. 2004;5(4):276–87.
- Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018;46(W1):W296–303.
- Wickland DP, Hanzawa Y. The FLOWERING LOCUS T/TERMINAL FLOWER 1 gene family, functional evolution and molecular mechanisms. *Mol Plant*. 2015;8:983–97.
- Winterhagen P, Tiyyayon P, Samach A, Hegele M, Wünsche JN. Isolation and characterization of FLOWERING LOCUS T subforms and APETALA1 of the subtropical fruit tree Dimocarpus longan. *Plant Physiol Bioch*. 2013;71:184–90.
- Xu W, Yu Y, Zhou Q, Ding J, Dai L, Xie X, et al. Expression pattern, genomic structure, and promoter analysis of the gene encoding stilbene synthase from Chinese wild *Vitis pseudoreticulata*. *J Exp Bot*. 2011;62(8):2745–61.
- Xue T, Chen D, Zhang T, Chen Y, Fan H, Huang Y, et al. Chromosome-scale assembly and population diversity analyses provide insights into the evolution of *Sapindus mukorossi*. *Hortic Res*. 2022;9:uhac012.
- Yao ZY, Qi JH, Yin LM. Biodiesel production from *Xanthoceras sorbifolia* in China, Opportunities and challenges. *Renewable Sustainable Energy Rev*. 2013;24:57–65.
- Yin X, Singer SD, Qiao H, Liu Y, Jiao C, Wang H, et al. Insights into the Mechanisms Underlying Ultraviolet-C Induced Resveratrol Metabolism in Grapevine (*V. amurensis* Rupr.) cv. 'Tonghua-3'. *Front Plant Sci*. 2016;7:503.
- Zee F TP, Chan JHT, Yen CR. Lychee, longan, rambutan and pulasan. 1998.
- Zhai H, Lü S, Liang S, Wu H, Zhang X, Liu B, et al. GmFT4, a homolog of FLOWERING LOCUS T, is positively regulated by E1 and functions as a flowering repressor in soybean. *PLoS ONE*. 2014;9:e89030.

- Zhang H, Shen J, Wei Y, Chen H. Transcriptome profiling of litchi leaves in response to low temperature reveals candidate regulatory genes and key metabolic events during floral induction. *BMC Genomics*. 2017;18:1–14.
- Zheng J, Meinhardt LW, Goenaga R, Matsumoto T, Zhang D, Yin Y. The chromosome-level rambutan genome reveals a significant role of segmental duplication in the expansion of resistance genes. *Hortic Res*. 2022;9:uhac014.
- Zhu XR, Wang H, Sun J, Yang B, Duan XW, Jiang YM. Pericarp and seed of litchi and longan fruits: constituent, extraction, bioactive activity, and potential utilization. *J Zhejiang Univ Sci B*. 2019;20(6):503–12.
- Zhu Y, Zhu G, Xu R, Jiao Z, Yang J, Lin T, et al. A natural promoter variation of *SIBBX31* confers enhanced cold tolerance during tomato domestication. *Plant Biotechnol J*. 2023;21:1033–43.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.