

RESEARCH ARTICLE

Open Access



# Multi-omics sequencing provides insight into floral transition in *Catalpa bungei*. C.A. Mey

Zhi Wang<sup>1†</sup>, Wenjun Ma<sup>1†</sup>, Tianqing Zhu<sup>1</sup>, Nan Lu<sup>1</sup>, Fangqun Ouyang<sup>1</sup>, Nan Wang<sup>1</sup>, Guijuan Yang<sup>1</sup>, Lisheng Kong<sup>2</sup>, Guanzheng Qu<sup>3</sup>, Shougong Zhang<sup>1</sup> and Junhui Wang<sup>1\*</sup>

## Abstract

**Background:** Floral transition plays an important role in development, and proper time is necessary to improve the value of valuable ornamental trees. The molecular mechanisms of floral transition remain unknown in perennial woody plants. “Bairihua” is a type of *C. bungei* that can undergo floral transition in the first planting year.

**Results:** Here, we combined short-read next-generation sequencing (NGS) and single-molecule real-time (SMRT) sequencing to provide a more complete view of transcriptome regulation during floral transition in *C. bungei*. The circadian rhythm-plant pathway may be the critical pathway during floral transition in early flowering (EF) *C. bungei*, according to horizontal and vertical analysis in EF and normal flowering (NF) *C. bungei*. SBP and MICK-MADS-box were seemingly involved in EF during floral transition. A total of 61 hub genes were associated with floral transition in the METurquoise model with Weighted Gene Co-expression Network Analysis (WGCNA). The results reveal that ten hub genes had a close connection with the *GASA homologue gene* (Cbu.gene.18280), and the ten co-expressed genes belong to five flowering-related pathways. Furthermore, our study provides new insights into the complexity and regulation of alternative splicing (AS). The ratio or number of isoforms of some floral transition-related genes is different in different periods or in different sub-genomes.

**Conclusions:** Our results will be a useful reference for the study of floral transition in other perennial woody plants. Further molecular investigations are needed to verify our sequencing data.

**Keywords:** Floral transition, RNA sequencing, WGCNA, Early flowering, *Catalpa bungei*

## Background

Floral transition is the developmental process by which a plant transitions from vegetative growth to reproductive growth. During this process, inflorescence primordia instead of leaf primordia develop from the shoot apical meristem (SAM) [1–3]. Great progress has been made in understanding the factors that trigger floral transition

[4]. A set of floral transition-related genes, such as *SPL* (*Squamosa-promoter binding protein-like*) [5–7], *TOC* (*Timing of cab expression 1*) [8], *LUX* (*Luxarrhythm*) [8], *PIF* (*Phytochrome interacting factor*) [9], *CO* (*constans*) [10], *FRI* (*Frigida*) [11], *GA20ox* (*GA20oxidases*) [7], *GA3ox* (*GA3oxidases*) [12], *SOC1* (*Suppressor of overexpression of constans 1*) [13], have been detected, in addition to others [14, 15]. These genes are mainly categorized into five major pathways that regulate floral transition, including the age pathway, photoperiod and circadian clock pathway, autonomous pathway, vernalisation pathway and GA pathway [4]. These genes are

\* Correspondence: wangjh808@sina.com

†Zhi Wang and Wenjun Ma contributed equally to this work.

<sup>1</sup>State Key Laboratory of Tree Genetics and Breeding, Key Laboratory of Tree Breeding and Cultivation of State Forestry Administration, Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, PR China  
Full list of author information is available at the end of the article



© The Author(s). 2020 **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.

independent and closely related to each other, forming sophisticated gene regulatory networks (GRNs) [1, 16]. For example, *SPL* is involved in inducing the expression of flowering integrator genes, namely, *LEAFY (LFY)* and *APETALA1 (API)*, thereby triggering flowering [17]. *TOC* and *LUX* are the critical genes in circadian rhythms pathway [18, 19]. In term of the feed-back loop, *TOC1* can either directly or indirectly regulate *CCA1* and *LHY*, which in turn suppress *TOC1* expression by binding to its regulatory region [20, 21]. The circadian clock gene *LUX* affects flowering by forming the evening complex (EC) with *EARLY FLOWERING 3 (ELF3)* and *ELF* [22]. *FRI* controls flowering by regulating the expression of the floral transition of floral repressor *FLC*, which encodes a MADS-box protein [11]. *CO* promotes flowering by directly activating the expression of its downstream genes including *FT* and *SOC1* [23]. *SOC1* is also regulated by active GA in the gibberellin pathway and positively regulated by *SPL* in the age pathway [24]. However, most of these studies were focused on annual herbaceous model plants, such as *Arabidopsis* [25] and Rice [26]. In perennial woody plants, the studies involved in floral transition are still in their infancy [27, 28]. Few studies have been conducted on floral transition in trees, partly due to the long juvenile phase and the difficulty in distinguishing vegetative buds from flowering buds at the beginning of the budding phase of trees. *Catalpa bungei*. C.A. Mey (*C. bungei*, Family: Bignoniaceae) is an important ornamental tree species in China [29, 30]. *C. bungei* not only has good woody properties but is also famous for its beautiful flowers. The commercial value of this species is largely related to its flowering time. The optimum flowering time greatly affects the quality of *C. bungei*. *C. bungei* is a perennial tree that undergoes its first floral transition in the fifth year or more of planting. However, an early flowering (EF), the new natural variety of *C. bungei*, was found to undergo floral transition in the first planting year, and almost 100% of its buds were mixed buds, which is very rare for woody plants (<http://www.forestry.gov.cn/>). At present, the research on *C. bungei* mainly focuses on the development of wood and flower organs [29–32], and the study of the flowering of *C. bungei* is just beginning. The EF variety, which only develops mixed buds, solves the problem of material selection and provides an opportunity to evaluate the floral transition process in perennial ornamental woody plants.

Next-generation sequencing (NGS) technologies have become a powerful tool for describing gene expression levels. However, NGS is limited by the necessity of short reads during library construction [33]. Single-molecule real-time (SMRT) sequencing technology overcomes this limitation by generating kilobase-sized sequencing reads [34]. The

combination of NGS and SMRT approaches not only enables the overall transcript level of each gene to be analysed but also provides vital insight into alternative splicing (AS) events [35], which have fundamental roles in a wide range of plant growth and development processes [36–41]. In particular, the AS of genes, such as *FT*, *FLC*, and *PRR*, regulates floral transition [19, 20, 40, 42–46].

The NGS and SMRT sequencing platform was used to further investigate the genes involved in floral transition. In this study, we analysed the data from three perspectives, namely, horizontal analysis, vertical analysis and WGCNA. A total of 61 hub genes that may be associated with floral transition in *C. bungei* were mined. Several potential protein interactions were found by regulatory network analysis. The complexity of AS events in the EF and NF varieties was addressed via SMRT sequencing. More than 50% of the identified genes had multiple structures. This work provides a guideline for future studies on how woody plants regulate the expression of key genes during floral transition.

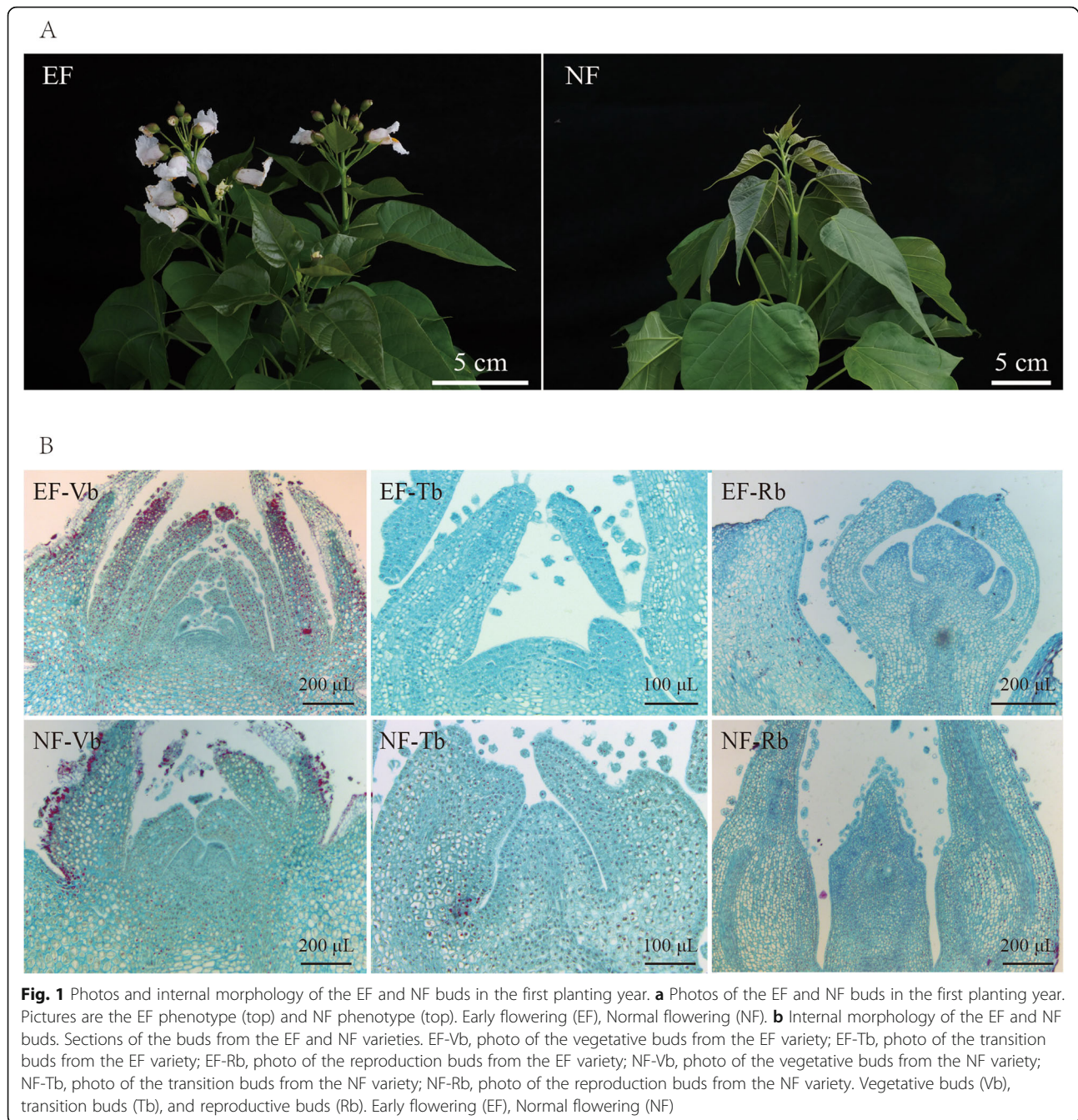
## Results

### Grouping of the buds from EF variety and NF variety

An EF variety was used to study floral transition. A NF variety was used as a control (Fig. 1a). The EF buds were subgrouped into three consecutive differentiation stages, namely, vegetative buds (Vb), transition buds (Tb), and reproductive buds (Rb), according to their anatomical structure (Fig. 1b). In the Vb, the reproductive shoot apex was still invisible. In the Tb, the reproductive shoot apex had initiated. In the Rb, the development of the reproductive shoot apex had completed, and the differentiated sepals, petals, pistils, etc. were observed. The NF buds were always Vb morphologically. However, we subgrouped them artificially into the three stages according to the corresponding collection date for the control. Since the molecular regulation of floral transition begins far before morphological changes occur, many critical molecular regulations should have already occurred in the Vb [29, 31, 47].

### Illumina-based RNA and SMRT sequencing and assembly

To explore the molecular regulation during floral transition in *C. bungei*, we carried out NGS and SMRT sequencing for the stem apical buds. The stem apical buds (Vb, Tb and Rb) from the EF and NF varieties were prepared for NGS. Each group had three biological replicates. A total of 18 mRNA samples were subjected to 2\*150 bp paired-end sequencing using the HiSeq 4000 platform, which produced more than 13G of clean reads (Table S1). Subsequently, the RNA samples were pooled according to EF and NF for SMRT sequencing. The full-length cDNAs of these samples were sequenced and



constructed using the PacBio RS II platform. In total, 13 SMRT cells and 16 SMRT cells were used for the EF and NF mixed samples, respectively, with three size fractions, namely, 1–2 kb, 2–3 kb, and > 3 kb. The mean ReadsOfInsert lengths produced in the EF and NF samples were 2702 bp and 4028 bp, respectively. ReadsOfInserts were composed of 261,651 full-length non-chimeric reads and 175,647 non-full-length reads in EF and 122,967 full-length reads and 339,065 non-full-length reads in NF. The average lengths of the full-length non-chimeric reads were 2592 bp and 2605 bp in

EF and NF, respectively. The non-full-length transcripts and the full-length transcripts were classified based on the presence of 5' primers, 3' primers and poly(A) tails reaching near-saturation of gene discovery (Table S2, Fig. S1, Fig. S2). The transcript length distributions generated by these two platforms showed that approximately 88% of the assembled transcripts from the Illumina platform and 11% of the transcripts from the SMRT reads were < 600 bases (Fig. S3A). A total of 22,934 annotated genes were detected by Illumina RNA-seq. In contrast, 14,753 EF and 15,212 NF annotated

genes were detected by SMRT sequencing. Of the annotated genes, 11,631 genes were found by both Illumina and SMRT. A total of 6628 genes were identified only by Illumina, and 1450 genes were identified only by SMRT, i.e., 383 EF-specific genes, 489 NF-specific genes and 578 common genes in both the EF and NF varieties (Fig. S3B). The high sensitivity of SMRT makes it possible to detect the alternative polyadenylation (APA) in the transcriptome high-throughput data. In our experiment, of the 36,935 genes detected by SMRT, 13,843 transcripts had one poly (A) site, while 1962 genes had at least five poly (A) sites (Fig. S3C). These APAs could increase transcriptome complexities, subsequently affecting post-transcriptional regulation.

#### Differential gene expression during floral transition

To characterize the expression profiles of the 14,231 EF DEGs and 7378 NF DEGs, the expression data  $v$  (from Vb to Tb and Tb to Rb) were normalized to 0,  $\log_2^{(Tb/Vb)}$ , and  $\log_2^{(Rb/Vb)}$ . In total, all the DEGs clustered into eight profiles based on STEM analysis (Fig. 2a and Fig. S4A). It was assumed that the DEGs obtained from the vertical analysis between EF-Vb and EF-Tb were mainly associated with floral transition. In our data, genes belonging to Profile 3 and Profile 4 showed no significant difference between EF-Vb and EF-Tb. Therefore, Profiles 0, 1, 2, 5, 6, and 7 were chosen for subsequent analyses (Fig. 2b). Profiles 0, 1, and 2 were downregulated between Vb and Tb in the EF buds and contained 427, 568 and 4286 DEGs, respectively. Profiles 5, 6, and 7 were upregulated between Vb and Tb in the EF buds and contained 4268, 627 and 272 DEGs, respectively.

All the DEGs in EF buds that belonged to profiles 0, 1, 2, 5, 6 and 7 were subjected to KEGG pathway enrichment analysis (Table S3). The KEGG pathways associated with plant floral transition are listed in Fig. 2c. Plant-pathogen interaction (ko04626), plant hormone signal transduction (ko04075), microbial metabolism in diverse environments (ko01120), starch and sucrose metabolism (ko00500) and circadian rhythm-plant (ko04712) were significantly enriched in all six profiles. Plant-pathogen interaction (ko04626) was significantly enriched in Profile 5, plant hormone signal transduction was significantly enriched in Profile 2, starch and sucrose metabolism was significantly enriched in Profile 7 and circadian rhythm-plant was significantly enriched in Profile 0. Most of the pathways, such as photosynthesis (ko00195), brassinosteroid biosynthesis, and anthocyanin biosynthesis, were not enriched in all six profiles. The photosynthesis and anthocyanin biosynthesis pathways were obviously enriched obviously in Profile 7. Brassinosteroid biosynthesis was obviously enriched in Profile 6. In

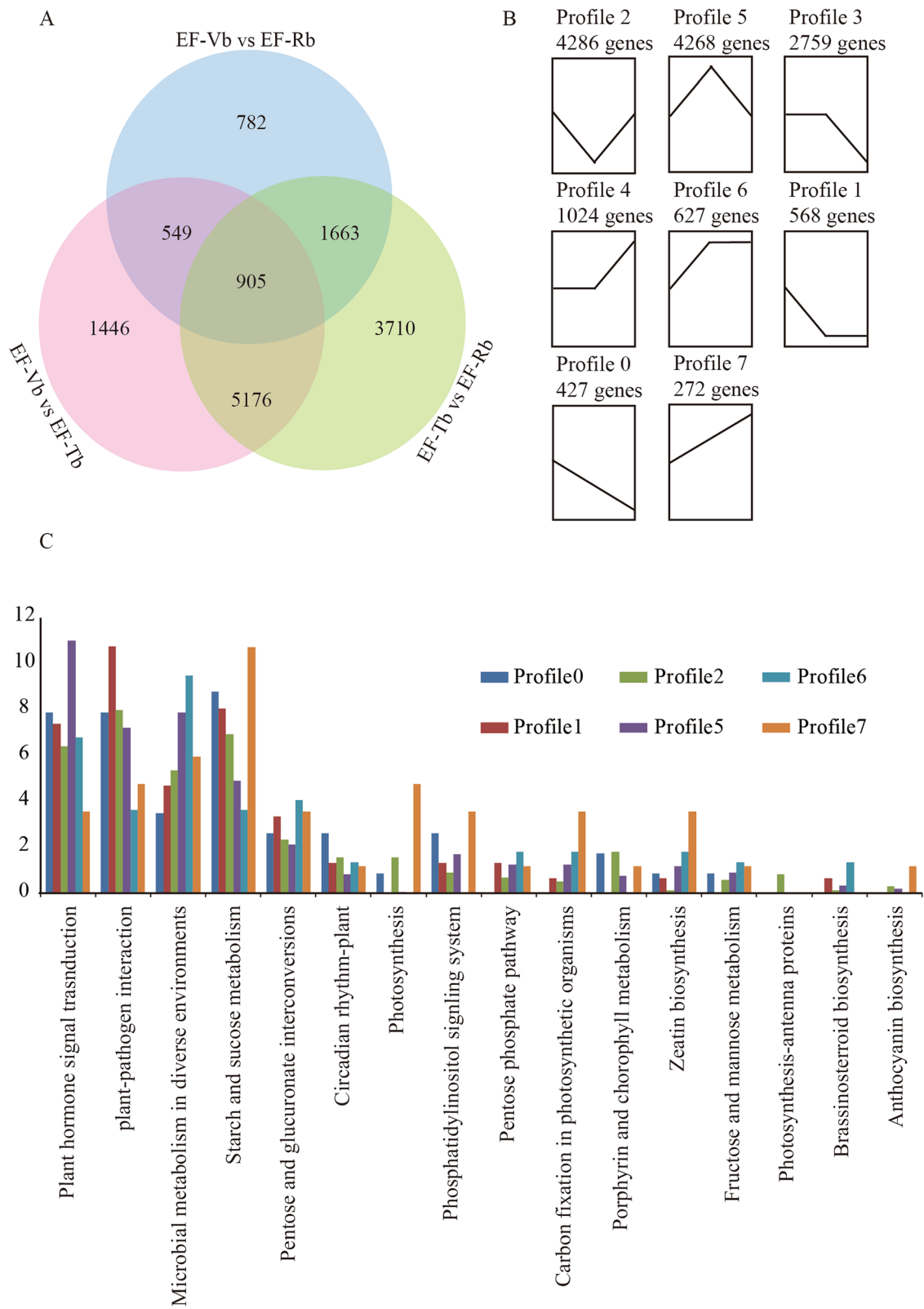
addition, the photosynthesis-antenna proteins pathway was only enriched in Profile 2. The high expression of the circadian rhythm-plant pathway in EF-Vb implied that circadian rhythm-related genes may promote the activation of related downstream pathways, eventually leading to early flowering. In addition, the KEGG pathway enrichment results of DEGs in NF buds were mainly related to carbohydrate metabolism and energy metabolism, and no related plant floral transition pathways were found (Fig. S4B).

#### Gene sets differentially expressed between the EF and NF buds

To investigate the DEGs that might lead to floral transition, horizontal analysis was performed between EF and NF. In total, 4584 genes exhibited significantly higher expression and 4351 genes exhibited significantly lower expression at different stages in EF compared to NF. There were 1905 DEGs between EF-Vb and NF-Vb (including 65 upregulated and 34 downregulated TFs). There were 5438 DEGs between EF-Tb and NF-Tb (including 217 upregulated and 235 downregulated TFs). There were 1593 DEGs between EF-Rb and NF-Rb (including 14 upregulated and 23 downregulated TFs) (Fig. 3a).

TFs are critical for development transition in plants [48, 49]. In our data, 58 TF families were significantly differentially expressed in EF compared to NF during floral transition (Table S4). Thirteen of the 58 TF families, such as *B3* [50], *bHLH* [51], *GRAS* [52, 53], *ARF* [54], *AP2* [55], *SBP* [6] have been reported as important developmental regulators (Fig. 3b). *GRAS*, *HSE*, *NAC* and *MYB-related genes* showed significant enrichment in EF/NF-Tb-UP. *MYB*, *bHLH*, and *GATA* showed significant enrichment in EF/NF-Rb-UP. In addition, *C3H* and *SBP* showed significant enrichment in EF/NF-Vb. Furthermore, all *SBPs* were only enriched via upregulation in the EF compared to NF in vegetative buds. This implies that the *SBP* family might relate with the early floral transition in EF, similar to the function of *SBP* in other plants during floral transition [7, 31, 56–62].

The DEGs were assigned to 67 KEGG pathways. The top 20 pathways are presented in (Table S5). Enrichment analysis suggested circadian rhythm-plant (ko04712) and ubiquitin mediated proteolysis (ko04120) were significantly enriched in Vb, while photosynthesis-antenna proteins (ko00196), nitrogen metabolism (ko00910) and plant-pathogen interaction (ko04626) were significantly enriched in Tb (Fig. 3c). These results combined with data from the vertical analysis, further supported the idea that the circadian rhythm-plant pathway was critical during floral transition.



**Fig. 2** (See legend on next page.)

(See figure on previous page.)

**Fig. 2** Analysis of differential gene expression during floral transition of the EF variety. **a** Venn diagram analysis of the number of DEGs between EF-Vb vs EF-Tb, EF-Tb vs EF-Rb and EF-Vb vs EF-Rb. EF-Vb, the data of vegetative buds from the EF variety; EF-Tb, the data of the transition buds from the EF variety; EF-Rb, the data of the reproduction buds from the EF variety. **b** The 8 significant expression profiles during floral transition of EF. **c** Partial KEGG pathways associated with floral transition of EF. The longitudinal axis represents the percent of the number of genes. The horizontal axis represents the pathway names. The dark blue rectangle indicates the data were from Profile 0. The red rectangle indicates the data were from Profile 1. The green rectangle indicates the data were from Profile 2. The purple rectangle indicates the data were from Profile 5. The light blue rectangle indicates the data were from Profile 6. The orange rectangle indicates the data were from Profile 7

### Identification of conserved and/or divergent gene co-expression modules

WGCNA was performed to obtain a comprehensive understanding of genes expressed in the successive developmental stages of EF and NF and to identify the genes that might be associated with floral transition. After filtering out the genes with low expression (FPKM < 0.05), 34,483 genes were retained for WGCNA. Co-expression networks were constructed on the basis of pair-wise correlations of gene expression across all samples. Modules were defined as clusters of highly interconnected genes, and genes within the same cluster had high correlations.

Correlated expression profiles imply that the genes operate in collaboration or in related pathways and that they contribute together to a given phenotype [63]. Our analysis identified 11 distinct modules (labelled with different colours), which are defined by major tree branches (Fig. S5). The number of genes in the modules ranged from 81 to 11,700. Four modules were highly expressed in one sample: MEdarkturquoise was highly associated with EF-Vb; MElightgreen was highly associated with NF-Vb; METurquoise was highly associated with EF-Tb, and MEdarkgrey was highly associated with NF-Rb (Fig. 4a).

To explore the significance of the modules, correlations between the MEs and the three developmental periods were analysed. As the molecular regulation of floral transition starts before morphology changes occur, genes should have already changed in the vegetative stage to direct floral transition. The genes associated with floral transition should exhibit differential expression in Vb (Fig. 4b). Based on this principle, MEdarkturquoise was considered the main module of interest. In total, 1223 genes were included in the MEdarkturquoise module, among which 677 genes were known genes, and 564 genes were new genes (Table S6). To validate the accuracy of the transcriptome analysis results, 8 unigenes were selected for qRT-PCR confirmation. The expression profiles of the candidate unigenes revealed using qRT-PCR data were consistent with those derived from sequencing (Fig. S6).

To study the relationship between these genes and floral transition more accurately, the top 10% of the genes were selected according to the correlation results. Sixty-one of these genes were annotated as hub genes

involved in floral transition (Table 2, Table S7). The 61 hub genes were classed into the five floral regulation pathways, namely, the age pathway (Cbu.gene.9773 and Cbu.gene.16991, *SPL* homologous genes), autonomous pathway Cbu.gene.669, *FCA* homologous gene; Cbu.gene.14804, *FY* homologous genes), verbalization pathway (TCONS\_00014487, *FRI* homologous genes), GA pathway (Cbu.gene.15447, *GA20ox* homologous genes; Cbu.gene.1698, *GA3ox* homologous genes) and photoperiod and circadian clock pathway (Cbu.gene.21497 *PIF* homologous gene; Cbu.gene.12567, *LUX* homologous genes; Cbu.gene.7628, *CO* homologous genes). In addition, several floral integrators, such as *SOC1* and *AP2*-like, and several hormone relation factors, including Cbu.gene.26092 and Cbu.gene.26299 (*ARF* homologous genes) (Fig. 5, Table 1), were detected. Subsequently, we analysed the regulatory network of the 61 hub genes in the METurquoise module. Thirty-eight TFs were annotated from the regulatory network. Accordingly, the MIKCMADS-box was shown to be highly related to floral transition [64–67].

Interestingly, 10 out of the 61 hub genes had a close connection with Cbu.gene.18280, which was annotated as a *GASA* homologous gene (Fig. S7). According to WGCNA analysis, *GASA* was predicted to have high connectivity with *CbuSPL* (age pathway), *CbuFCA* and *CbuFY* (autonomous pathway), *CbuGA3ox* and *CbuG20ox* (GA pathway) and *CbuTOC1* and *CbuLUX* (photoperiod and circadian clock pathway). In addition, *CbuPIF4* (photoperiod pathway) and *CbuGA20ox* (GA pathway) can affect the floral transition by promoting the expression of *CbuSOC1* (Fig. 6). However, floral transition is a very complicated process in *C. bungei* and needs to be further verified.

To verify the intersection results of *GASA*, we performed protein-protein interaction analysis ([http://www.iitm.ac.in/bioinfo/PPA\\_Pred/prediction.html#](http://www.iitm.ac.in/bioinfo/PPA_Pred/prediction.html#)). The dissociation constants (Kd), as well as on- and off-rates ( $k_{on}$  and  $k_{off}$ ) less than  $10^{-9}$ , were set to predict protein binding. The protein interaction prediction results were highly consistent with the WGCNA results (Table 2).

To further study the correlation of *CbuGASA* and the 8 known hub genes (Table 2), we analysed the correlation coefficients of these mRNAs between the EF and NF samples during three developmental periods. Based

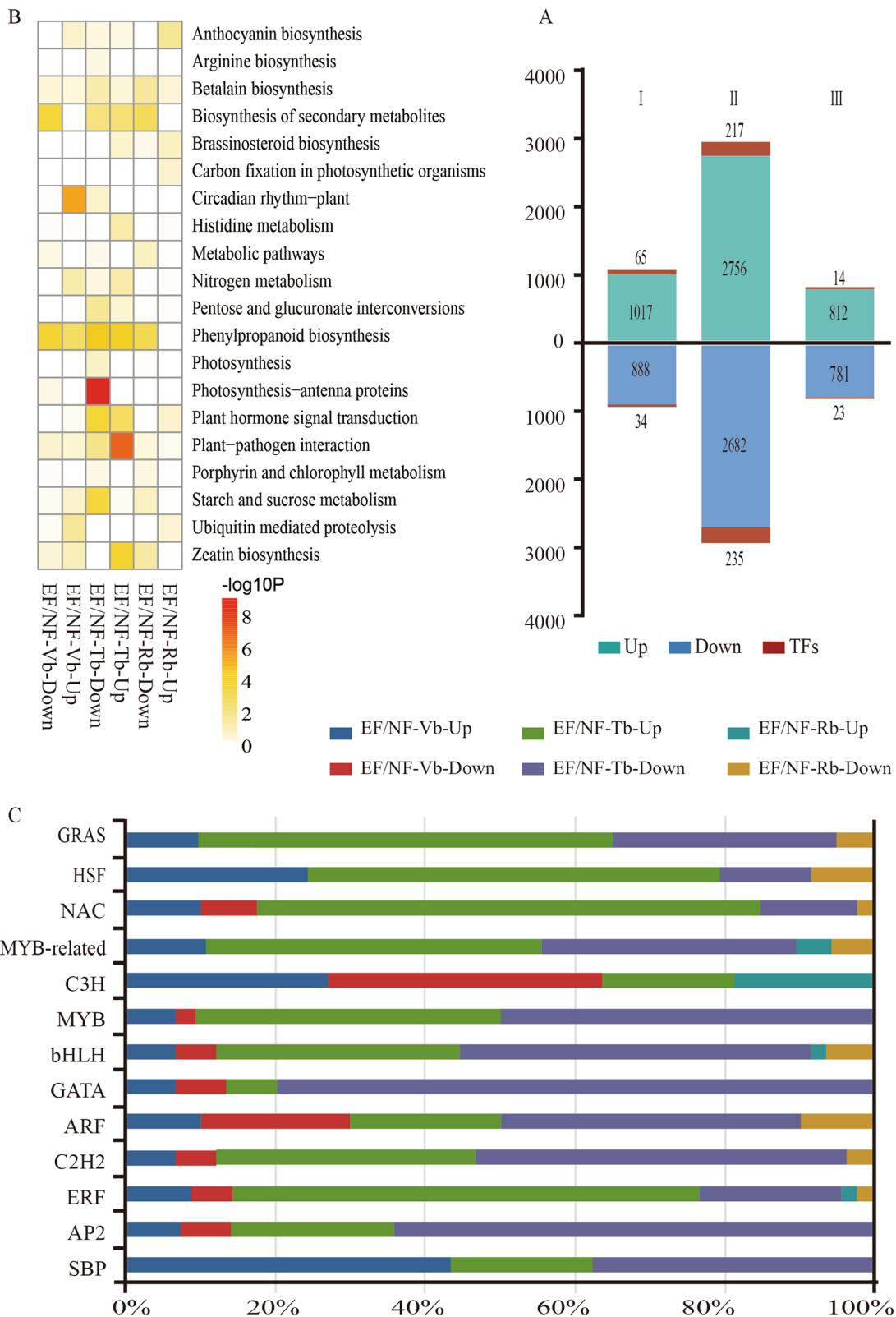


Fig. 3 (See legend on next page.)

(See figure on previous page.)

**Fig. 3** Differential gene expression in EF compared to NF at different stages of floral transition. **a** The number of upregulated (upper bars) and downregulated (lower bars) genes at each stage of floral transition in EF compared to NF is given. I indicates EF-Vb vs NF-Vb; II indicates EF-Tb vs NF-Tb; III indicates EF-Rb vs NF-Rb. The number of TFs up- or downregulated at each stage of floral transition is also given. EF-Vb, the number of genes of vegetative buds from the EF variety; EF-Tb, the number of genes of the transition buds from the EF variety; EF-Rb, the number of genes of the reproduction buds from the EF variety; NF-Vb, the number of genes of vegetative buds from the NF variety; NF-Tb, the number of genes of the transition buds from the NF variety; NF-Rb, the number of genes of the reproduction buds from the NF variety. **b** KEGG analysis of at different stages of floral transition in down- and upregulated genes in EF. The colour scale at the bottom represents significance (corrected *P*-value). **c** Partial TF families showing up-regulation or down-regulation at different stages during floral transition between EF and NF

on all trends, *CbuGASA* with 6 of 8 known hub genes (*CbuGA3ox*, *CbuSPLs*, *CbuLUXs*, *CbuFCA*) exhibited a positive correlation ( $r > 0.75$ ) (Table S8). During the flowering process, the expression levels of mRNAs were significantly higher in the EF-Vbs than in the NF-Vbs, and gradually decreased with age. These results were consistent with the expression patterns of the homologous genes in *Arabidopsis*.

#### Analysis of AS events from hub genes during floral transition development in *C. bungei*

Understanding AS events plays an extremely important role in understanding protein diversity [35, 68, 69]. On the basis of obtaining high-quality full-length isoforms, we performed a systematic analysis of AS in *C. bungei*. A total of 79,356 AS events from 25,662 mRNAs and 69,775 AS events from 25,046 mRNAs were detected in the two pools by SMRT. These AS events could be classified into five major types, namely, SE, IR, A5, A3, and AE [70]. However, no AE type was detected by Illumina (Fig. 7a). For both datasets, IR events showed the highest proportion in the EF and NF buds; SE events were the least frequent AS type in both the EF and NF buds. In addition, the percentage of the other more complicated AS types was greater in EF than in NF (Fig. 7a, Table S9). The AS event data from EF were compared with those from NF. A total of 23,851 and 22,936 common AS events from 9521 genes were detected in the EF and the NF buds. A total of 46,839 AS events associated with 15,526 genes were identified only in NF, and 55,505 AS events from 16,141 genes were detected only in EF (Table S10, Table S11). This result showed that SMRT is more accurate than Illumina for AS detection.

One of the most important features of SMRT is the ability to identify AS events by directly comparing isoforms of the same gene [33, 68, 71–74]. We randomly selected ten genes to evaluate the accuracy of AS events using RT-PCR (Table S12). The size of each amplified fragment was consistent with that of the predicted fragment (Fig. 7b, Fig. S8). These amplified fragments were then cloned for Sanger sequencing, and the amplified sequences were consistent with the SMRT data. These genes were divided into three groups according to the number of isoforms in EF and NF. The genes that had

fewer isoforms in the EF than in NF were classified into genre-I, the genes with no obvious difference in the number of isoforms in EF vs NF were classified into genre-II, and the genes that had less isoforms in NF compared to EF were classified into genre-III (Fig. 7c). Further analysis of the AS events during floral transition was performed based on the SMRT data.

To evaluate the novel isoforms that may be involved in floral transition, we analysed the AS events of 200 genes (61 hub genes and their 10% target genes) in EF and NF. Among them, 41 genes were in genre-I, 36 genes were in genre-II and 28 genes were in genre-III. To study the specificity of the AS events in different stages, we verified the isoforms from three developmental periods in EF and NF by RT-PCR. For example, *Cbu.gene.25040* (*GRF* homologous gene) was found in two isoforms in EF-Vb, NF-Vb and NF-Rb, and one *Cbu.gene.25040* isoform was found in the rest of the buds. *Cbu.gene.7628* (*CO* homologous gene) had different numbers of isoforms between EF-Vb and NF-Vb. Two *Cbu.gene.7628* isoforms were found in EF-Vb, and one *Cbu.gene.7628* isoform was found in NF-Vb. *Cbu.gene.16991* (*SPL* homologous gene) had different numbers of isoforms between EF-Vb and NF-Vb. Two *Cbu.gene.16991* isoforms were found in EF-Vb, and one *Cbu.gene.16991* isoform was found in NF-Vb. In addition, there were different numbers of isoforms between Vb and Tb. There were three *Cbu.gene.16991* isoforms in Tb from both EF and NF. (Fig. 7d, Fig. S9). The results showed that the genes in *C. bungei* exhibit divergent structures of isoform splicing in different bud stages.

#### Discussion

Perennial ornamental woody plants have irreplaceable economic value. In this study, we analysed the data from three perspectives, namely, horizontal analysis (EF vs NF), vertical analysis (EF buds at different developmental stages) and WGCNA. We applied this strategy to the EF and NF samples, thereby enabling the correlation of specific expression information from transcriptional data to EF vegetative buds that formed during floral transition. Floral transition is a very complex regulatory system [4, 25, 75]. Multiple biological pathways, such as the age



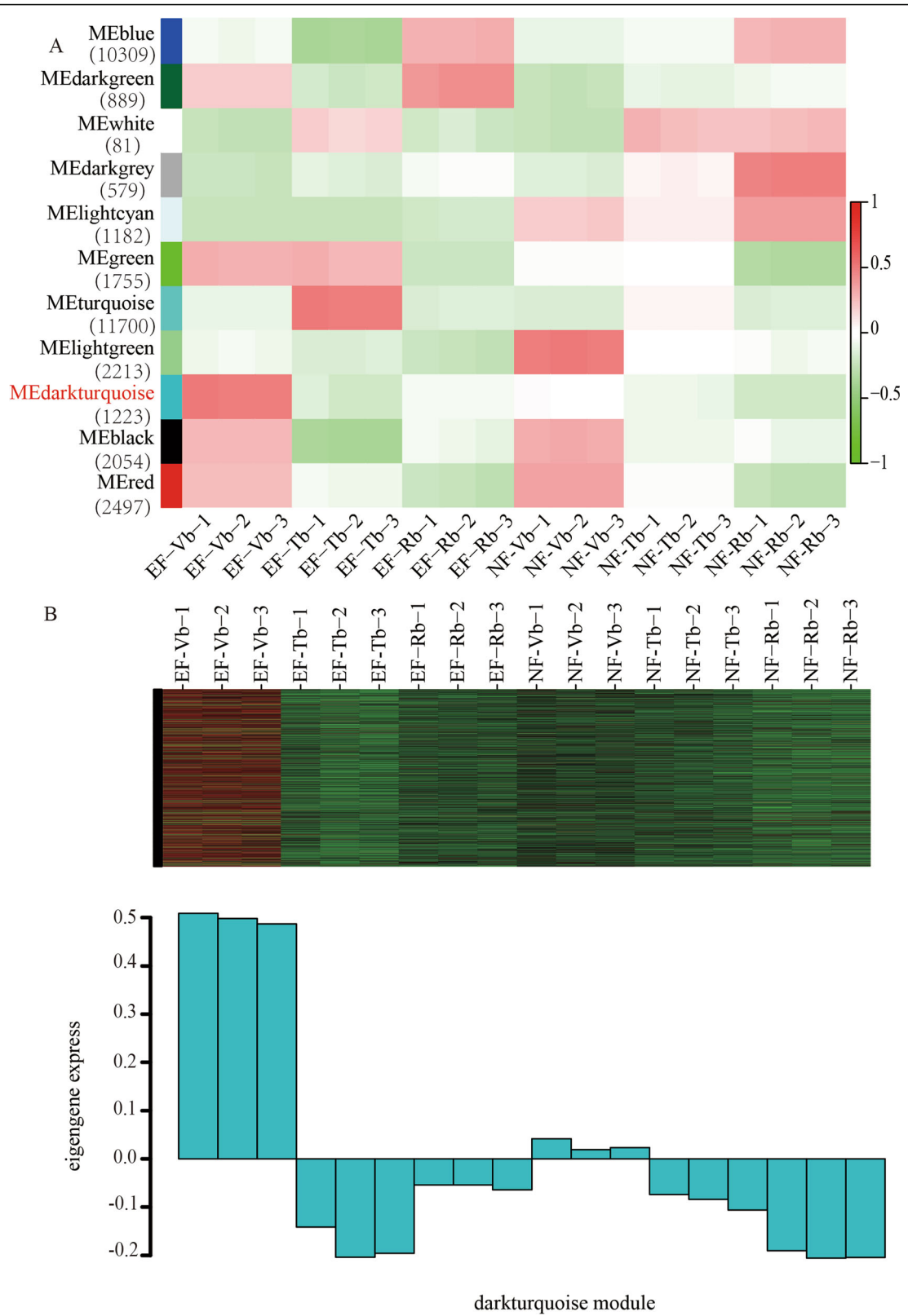


Fig. 4 (See legend on next page.)

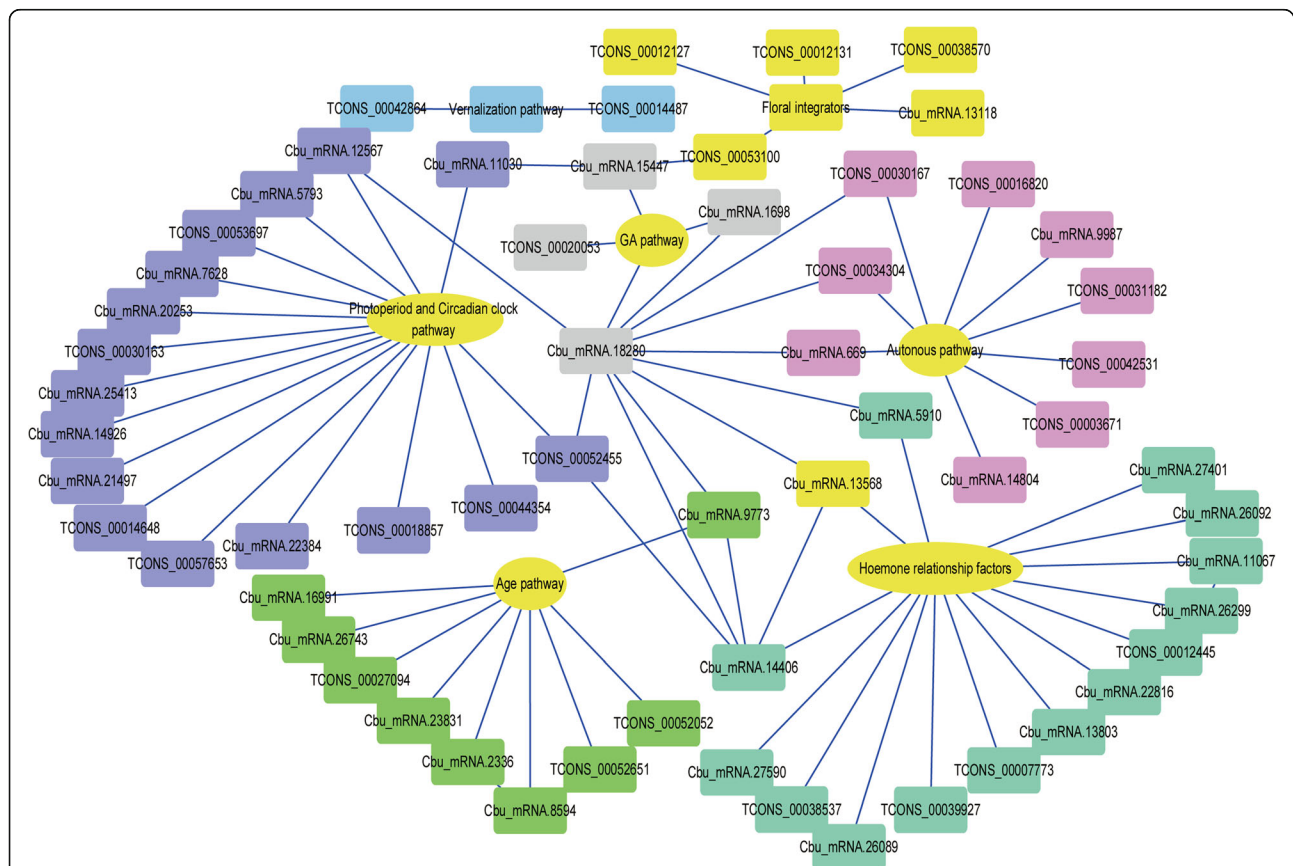
(See figure on previous page.)

**Fig. 4** Co-expression networks during floral transition in EF and NF. **a** Module-trait associations were evaluated by correlations between MEs, and traits are shown. The left panel shows the 11 modules and the number of member genes. The colour scale on the right shows module-trait correlations from -1 (Green) to 1 (Blue). EF-Vb represents the vegetative buds from early flowering, as a trait; EF-Tb represents the transition buds from early flowering, as a trait; EF-Rb represents the reproductive buds from early flowering, as a trait; NF-Vb represents the vegetative buds from natural flowering, as a trait; NF-Tb represents the transition buds from natural flowering, as a trait; NF-Rb represents the reproductive buds from natural flowering, as a trait. A high degree of correlation between a specific module and the trait is indicated when the module name is highlighted in red. **b** Heatmaps showing genes in the modules that were significantly over-represented in EF-Vb

pathway, GA pathway, and photoperiod pathway, are involved in this process. The circadian rhythm pathway was significantly enriched in the whole floral transition cycle of EF compared with NF (Fig. 2c and Fig. 3b). This suggested that the circadian pathway might be critical for the early floral transition of EF. In addition, the photosynthesis pathway was mainly concentrated in Profile 7 (Fig. 2c), showing that this pathway gradually increased with the completion of floral transition. The circadian rhythm and photosynthesis pathways are important pathways that affect the reproductive transformation in Arabidopsis [18, 26, 76]. Circadian rhythm

genes are regulated by the photosynthesis pathway and are used in related downstream photosynthesis pathways to regulate flowering. In the EF variety, the enrichment of circadian pathway genes may promote the activation of related downstream pathways, leading to early flowering.

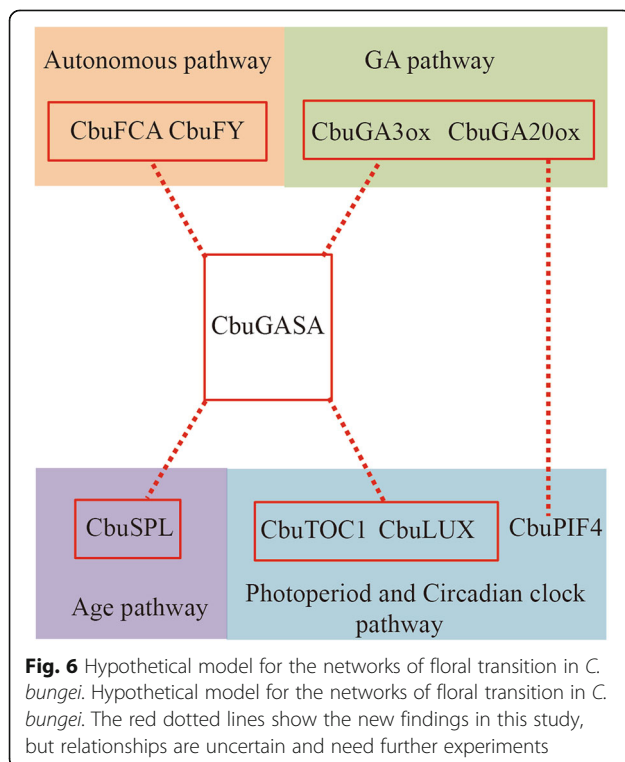
TFs play important roles during development transition. Several TF families, such as *SBP*, *MIKC type MADS-BOX*, *ARF*, *bHLH*, and *G2-like*, were identified as floral transition-related in our study (Fig. 3a, Table S4). The floral transition process was regulated by multiple TFs in the EF variety, similar to the case of other plants.



**Fig. 5** Correlation networks analysis of the top hub genes in MEdarkturquoise module. Yellow ovals represent the pathway names and rectangles of different colours represent individual genes. The genes in the purple shade were assigned to the photoperiod and circadian clock pathway. The genes in the green shade may belong to the age pathway. The genes in the brown shade may belong to the autonomous pathway. The genes in the grey shade may belong to the GA pathway. The genes in the blue shade may belong to the vernalisation pathway. The genes in the dark green shade may belong to the hormone relations factors. The genes in the yellow shade may belong to the floral integrator pathway

**Table 1** Summary of the 61 hub genes involving floral transition from the MEdarkturquoise module

Pathway	Gene ID		
Age pathway	TCONS_00052651	TCONS_00057653	TCONS_00027094
	Cbu_mRNA.23831	Cbu_mRNA.26743	TCONS_00052052
	Cbu_mRNA.9773	Cbu_mRNA.16991	Cbu_mRNA.2336
	Cbu_mRNA.8594		
Photoperiod and Circadian clock pathway	Cbu_mRNA.14926	TCONS_00052455	TCONS_00053697
	Cbu_mRNA.7628	Cbu_mRNA.5793	TCONS_00014648
	Cbu_mRNA.11030	TCONS_00018857	Cbu_mRNA.21497
	TCONS_00044354	TCONS_00030163	Cbu_mRNA.20253
	Cbu_mRNA.12567	Cbu_mRNA.22384	
Autonomous pathway	Cbu_mRNA.14804	TCONS_00016820	TCONS_00031182
	TCONS_00003671	TCONS_00034304	TCONS_00042531
	Cbu_mRNA.669	TCONS_00030167	Cbu_mRNA.9987
Verbalization pathway	TCONS_00014487	TCONS_00042864	
Hormone related pathway	Cbu_mRNA.1698	Cbu_mRNA.26092	Cbu_mRNA.27590
	Cbu_mRNA.15447	Cbu_mRNA.26089	Cbu_mRNA.11067
	Cbu_mRNA.13568	Cbu_mRNA.5910	Cbu_mRNA.27401
	Cbu_mRNA.14406	TCONS_00007773	TCONS_00038537
	TCONS_00039927	Cbu_mRNA.26299	Cbu_mRNA.13803
	TCONS_00012445	Cbu_mRNA.25413	TCONS_00012127
	Cbu_mRNA.22816		
Others	Cbu_mRNA.18280	Cbu_mRNA.9758	TCONS_00053100
	Cbu_mRNA.13118	TCONS_00020053	TCONS_00038570
	TCONS_00012131		



In this study, *SBP*, *ERF*, *bHLH*, *C3H* and *NAC* were the top 5 TFs by number in vegetative buds, and all *SBPs* were only enriched via upregulation in EF compared to NF in vegetative buds (Fig. 3c). In the EF/NF-Tb-UP groups, *SBP* was the top TF by number. Four *SBP* family members were identified in the 61 hub genes, which were determined by WGCNA (Fig. 5). *SBPs* are important floral transition regulators related with the age pathway [17, 77]. In Arabidopsis, *SPL*, which is negatively regulated by miR156, promotes floral transition by activating the expression of several other genes, such as *SOC1*, miR172, and *LFY* [5, 78–81]. Combined with the traits of the EF, the age pathway seems to be an important active pathway in floral transition, which implies that *SBPs* are important TFs affecting floral transition in EF.

A total of 1223 genes were included in the MEdarkturquoise module. To more precisely analyse the genes related to floral transition, we selected the top 10% of these genes according to connectivity. Sixty-one of 123 genes were annotated as specifically related to floral transition. These 61 genes were distributed into five pathways, i.e., the age pathway (e.g., *CbuSPL*), the photoperiod pathway (e.g., *CbuPIF*, *CbuCO* and *CbuLUX*), the autonomous pathway (e.g., *CbuFCA* and *CbuFY*), and

**Table 2** Protein interactions predicted with CbuGASA online

Gene ID	Annotation	Gene ID	Annotation	Delta G (binding free energy)/kcal/mol	Kd (dissociation constant)/M
Cbu_mRNA.18280	GASA	Cbu_mRNA.15447	GA20ox	-14.76	1.98E-12
Cbu_mRNA.18280	GASA	Cbu_mRNA.9773	SBP	-15.85	2.39E-12
Cbu_mRNA.18280	GASA	TCONS_00030167	FY	-15.40	5.08E-12
Cbu_mRNA.18280	GASA	TCONS_00018857	TOC	-14.57	2.08E-11
Cbu_mRNA.18280	GASA	Cbu_mRNA.12567	LUX	-14.56	2.09E-11
Cbu_mRNA.18280	GASA	Cbu_mRNA.16991	SBP	-13.40	1.48E-10
Cbu_mRNA.18280	GASA	Cbu_mRNA.20253	LUX	-13.18	2.17E-10
Cbu_mRNA.18280	GASA	Cbu_mRNA.669	FCA	-12.99	2.99E-10
Cbu_mRNA.18280	GASA	Cbu_mRNA.14804	FY	-12.66	5.19E-10
Cbu_mRNA.18280	GASA	Cbu_mRNA.1698	GA3ox	-12.16	1.21E-09
Cbu_mRNA.11030	PIF4	Cbu_mRNA.15447	GA20ox	-11.05	7.88E-09

the vernalisation pathway (e.g., *CbuFRI*), and *CbuGA20ox*, *CbuARF*, and *CbuERF* are hormone-related genes. These genes were highly expressed in EF-Vb. With the completion of floral transition, the expression of these genes decreased gradually, indicating that these genes may promote floral transition in *C. bungei*. In addition, several DEGs were annotated as *CbuSOC1* and *CbuAP2-LIKE*, which are critical genes in flowering. Most of our RNA-seq results were consistent with those in *Arabidopsis*, so the identified genes already had clear pathways in *Arabidopsis*. In addition to *SBP*, which regulates floral transition via the age pathway, several pathways were reflected in these data [82]. For example, *TOC1* and *LUX* are important factors in the three interlocking feedback loops in the circadian pathway, which mainly act on the upstream of the photoperiod pathway and ultimately regulate floral transition via the positive regulation of *CO* [8]. *FCA* and *FY* are autonomous pathway genes that have been studied in depth [83, 84]. *FCA* can inhibit the accumulation of *FLC*, which contains RNA binding proteins containing RNA recognition motifs (RRM) [8, 85]. *FLC* interacts with *FY* through the *FCA* WW domain, and the *FCA*/*FY* complex may negatively regulate *FLC* at the mRNA level [86]. *GA20ox* and *GA3ox* are critical for the synthesis of active GA [87]. The GA pathway influences floral transition mainly through two branches of active GA and *DELTA*. At the early stage of flower development, the active GA is involved in promoting the expression of *SOC1* and *LFY* and then the expression of downstream flowering genes. At the later stage of flower development, the active GA leads to the degradation of the *DELTA* protein and thus relieves the inhibition of flowering. Therefore, floral transition is guaranteed [87, 88]. Interestingly, several studies have revealed that *DELTA* regulates hypocotyl elongation by interacting with PIFs [75], contributing to floral transition by interacting with *SPL* [7], *FD* [10] and *SOC1* [24, 62]. *SOC1* plays an important role in

regulating floral transition by integrating signals involved in all pathways, as well as interacting with many other genes to regulate floral transition [24]. However, all floral transition-related pathways still need to be further explored in *C. bungei*. These genes should be further studied to determine whether they are related to floral transition in the EF variety.

Furthermore, 41 highly connected gene pairs in the 61 hub genes were predicted by WGCNA (Fig. 5). Among these genes, Cbu.gene.18280 (*GASA* homologous gene) is particularly interesting. In this study, *GASA* had a strong interaction with *GA20ox*. *GA20ox* is an important gene for synthesizing GA3. These predictions were further verified by protein-protein interaction analysis online (Table 2). The *GASA* family is named for its GA3-induced expression in *Lycopersicon esculentum* [89–91]. In *Arabidopsis*, *GASA4* regulates floral meristem identity [90] and *GASA5* can extend flowering time by promoting the expression of *FLC* and *FT* [91]. GA plays an important role in the floral transition of plants, but more specific pathways need to be further studied [92]. We also found that CbuGASA was correlated with CbuFCA, CbuFY, CbuTOC1, CbuLUX and CbuSPL. The expression correlations of *CbuGASA* and its co-expressed genes suggest that *CbuGASA* is a positive regulator and involved in flowering regulation. These proteins/genes are important for floral transition by their respective pathways. This result provides important information for our subsequent research. Additional experiments need to be performed to verify the hypotheses about the role of CbuGASA in floral transition in *C. bungei*. To ensure more accurate transcriptome data, analysis of the genetic differences between the EF and NF varieties is one of the most important experiments we will undertake in future research.

In eukaryotes, AS greatly contributes to transcriptional diversity [33, 71, 74, 93]. AS produces multiple transcripts from a single gene and gives rise to

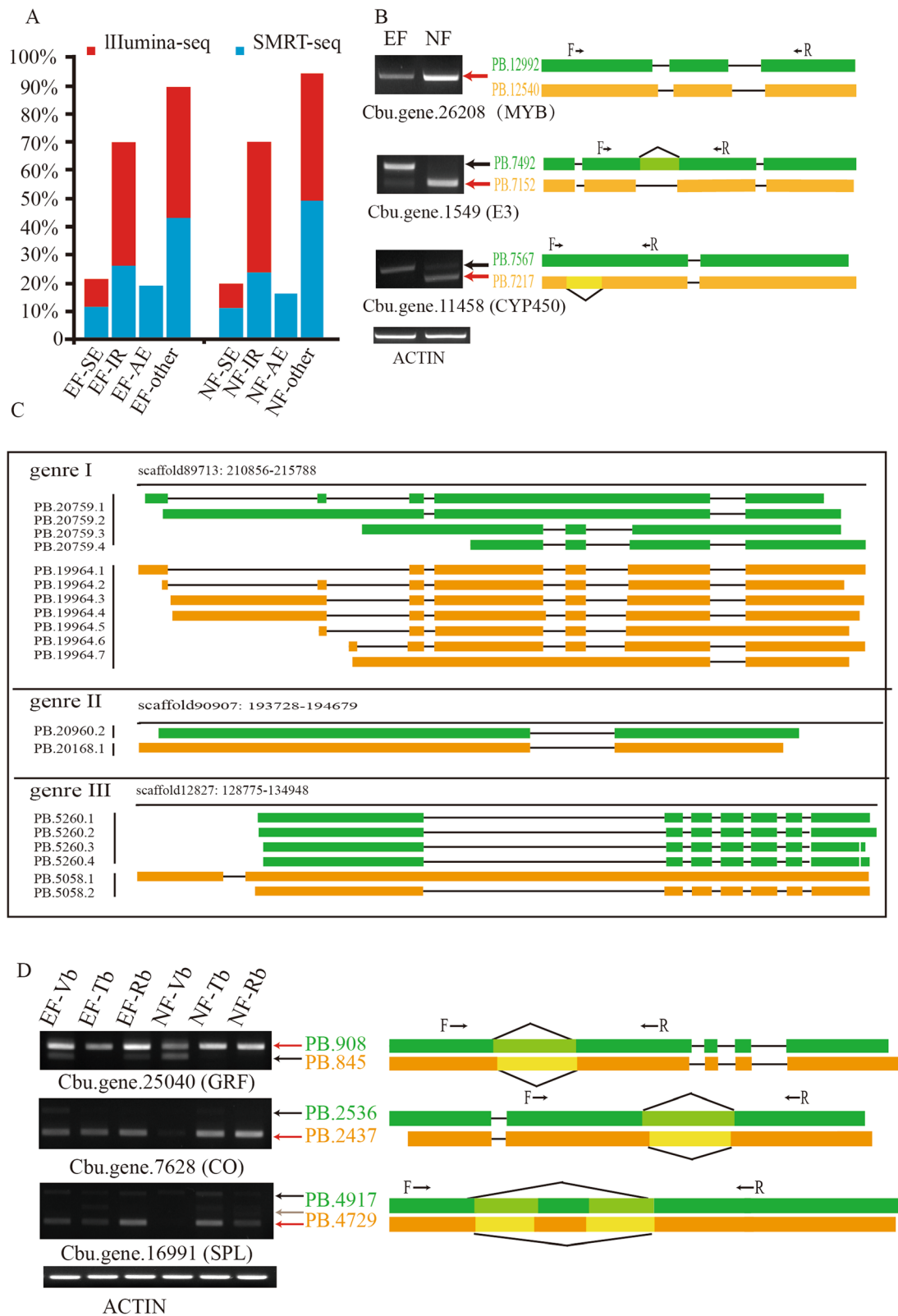


Fig. 7 (See legend on next page.)

(See figure on previous page.)

**Fig. 7** Characterization of AS events and validation of isoforms using reverse transcription polymerase chain reaction (RT)-PCR. **a** The proportion of different types of AS events detected by Illumina-seq or SMRT-seq in EF and NF. Classification of AS events: exon skipping (SE), intron retention (IR), and alternative exon (AE). **b** RT-PCR validation of AS events for five genes. Gel bands in each figure show PCR results in NF and EF. The transcript structure of each isoform is shown in the right panel. Boxes show exons in each transcript model. Green boxes show the isoforms from EF, and yellow boxes show the isoforms from NF. PCR primers (F, forward and R, reverse) are shown on the first isoform of each gene. **c** The number of genes belonging to the three genres. Genre-I contains the genes that have less isoforms in the EF variety than in the NF variety; genre-II contains the genes with similar numbers of isoforms in the EF and NF varieties. Genre-III contains the genes with less isoforms in the NF variety than in the EF variety. **d** The number isoforms of homologous genes in EF and NF. Gene names in the reference annotation (in parentheses) and corresponding names in the *C. bungei* genome are shown

proteins with various structures, subcellular localizations, stabilities and functions. AS has fundamental roles in a wide range of plant growth and development processes [41, 94–96]. Previous reports showed that isoforms have tissue specificity [93], and the ratio of the isoforms changes during the different growth periods [44]. For example, flowering integrator *FT* is also subjected to AS events in temperate grasses. The ratio of the two AS events of *FT* was progressively reduced during development, indicating that one of the two AS events is regulated by endogenous cues rather than an external cue for flowering. In our study, similar results were found. For example, the key genes in the floral transition of *C. bungei*, *CbuCO* (Cbu.gene.7628) and *CbuSPL* (Cbu.gene.16991) are also subjected to AS during the flowering process. In Vb of EF and NF, the ratio of AS in Vb (2:1) was higher than that of the other two periods of EF and NF (1:1). In addition, *CbuSPL* had more isoforms in Tb than in Vb in both the EF and NF buds. The above results imply that AS events may have an important role in floral transition in *C. bungei*. These observations indicate that AS greatly increases the complexity of gene transcription in *C. bungei*, and more experiments are needed to test this hypothesis. For example, we should not only consider how genes are transcribed in development periods but also investigate the functional differences of homologous genes in *C. bungei*. The multi-omics data were integrated to explore the floral transition mechanism in *C. bungei*.

## Conclusions

This study expands our view of the transcriptomes of *C. bungei* during floral transition. A number of DEGs were detected in vegetative to reproductive growth buds following WGCNA analysis. These genes belonged to pathways that collectively regulate floral transition, and the results enhance our understanding of gene regulation during floral transition in perennial woody plants. Furthermore, SMRT analysis provided the first insights into AS events in *C. bungei*. Frame usage in the same transcript further increases the genetic complexity of *C. bungei*. These results will facilitate future functional genomics studies.

## Methods

### Plant material and experimental procedures

*C. bungei* is a perennial plant that flowers after 5 years of forestation (<http://www.forestry.gov.cn/>). A natural EF variety that flowers after 1-year forestation was found in Henan province, China (Fig. 1a). This EF has been applied to produce new varieties of *C. bungei*, which are named 'bairihua'. From 28 February to 31 March 2016, we collected the buds from the first round to the axillary buds of EF and NF varieties at an interval of every 1–2 days. All samples were collected from 9:00–12:00 and transferred immediately to liquid nitrogen for SMRT- and Illumina-based RNA sequencing and reverse transcription-polymerase chain reaction (RT-PCR). Using paraffin section analysis, vegetative buds (Vb), transition buds (Tb) and reproduction buds (Rb) were identified for transcriptome sequencing.

### Library preparation and PacBio sequencing

To construct libraries for Pacific Biosciences (PacBio) sequencing, equal amounts of EF and NF buds from each stage (vegetative stage, transition stage and reproduction stage) were pooled together. Total RNA from the NF buds for three periods, NF-Vb, NF-Tb and NF-Rb, was mixed to provide the 'NF' sample for comparison with the EF sample. For EF buds, total RNA from three periods, EF-Vb, EF-Tb and EF-Rb, was mixed to provide the 'EF' sample. The two mixed RNA samples from buds were reverse transcribed using the SMRT<sup>er</sup> PCR cDNA Synthesis Kit. PCR amplification was carried out using the KAPA HiFi PCR Kit. The product was separated by agarose gel-based size selection into cDNA fractions 1–2 kb, 2–3 kb and > 3 kb in length. These SMRT libraries were generated using the PacBio 1.0 Template Preparation Kit (Menlo Park, CA, USA, part #001–322–716) according to the standard protocol. The 1–2 kb library was sequenced using five SMRT cells, and the other two libraries were sequenced using four SMART cells. The cDNA products were purified for library construction using the SMRTbell Template Prep Kit 1.0. Libraries were sequenced using P6C4 polymerase (PacBio, P/N 100–372–700) and chemistry on the PacBio RS II platform with 240-min movie times. Each size fraction for each sample was run through the Iso-Seq pipeline

included in the SMRT analysis software package individually. First, ROIs (previously known as circular consensus sequences) were generated using the minimum filtering requirement of 0 or more passes of the insert and a minimum read quality of 75. This requirement allows for the highest yield going into the subsequent steps, while creating higher-accuracy consensus sequences when possible. The pipeline then classified the ROIs as full-length non-chimeric or non-full-length reads. Full-length reads with lengths of at least 300 bp were determined by detecting poly(A) tails, 5' primers and 3' primers. All full-length reads were aligned to the *C. bungei* genome using  $G_{MAP}$  software.

### Illumina RNA sequencing of EF and NF buds

The samples stages were used to construct 18 libraries for Illumina-based RNA sequencing, which were named NF-Vb, NF-Tb, NF-Rb, EF-Vb, EF-Tb and EF-Rb. Each stage had three biological replicates. Total samples were sent to the Beijing Genomics Institute for strand-specific library construction and sequencing on an Illumina HiSeq 4000 platform. In total, 13G of 150-bp paired-end reads were generated. Raw sequence data of the libraries for differentially expressed gene (DEG) profiling analysis were filtered to remove reads containing adapters, reads with an unknown nucleotide content exceeding 10% unknown nucleotides, and reads with a low-quality base (value  $\leq 5$ ) content greater than 50%. Clean reads were mapped into the transcriptome reference database using SOAP software. No more than 2 mismatched bases were permitted, and unique mapped reads were obtained. Fragments per kilobase of exon per million fragments mapped (FPKM) was used to obtain the relative expression levels. A differential expression analysis of the two groups was performed using the DESeq package. The resulting  $P$ -values were adjusted using Benjamini and Hochberg's approach for controlling the false discovery rate. DEGs with a  $|\text{fold change}| \geq 1.2$  and an  $\text{FDR} < 0.05$  were identified between each comparison. The DEG expression data  $v$  (from Vb to Tb and Tb to Rb) were normalized to 0,  $\log_2^{(Tb/Vb)}$ , and  $\log_2^{(Rb/Vb)}$ . DEGs were clustered using the Short Time-series Expression Miner (STEM). To analyse gene co-expression patterns based on mRNA profiles in buds from the EF and NF varieties, WGCNA was performed according to Langfelder and Horvath (2008) [97]. Here, we chose a power of four so that the resulting networks exhibited approximate scale-free topology (model fitting index  $R^2 = 0.80$ ). The resulting gene dendrogram was used for module detection using the Dynamic Tree Cut method (minimum module size = 50) [63]. Co-regulated genes are grouped into modules based on the corresponding genes' information. In addition, the intra-modular hub genes refer to highly connected genes in a module. They can be determined by calculating the Pearson correlation between the expression level and the

module eigengene. In this study, the top 10% of genes with high correlation were considered as hub genes for a given module. Finally, the modules of interest were input into Cytoscape to determine network information [44].

With the Gene Ontology (GO; <http://www.geneontology.org>) and Kyoto Encyclopedia of Genes and Genomes (KEGG; <https://www.kegg.jp/>) pathway annotation results, we classified mRNAs according to official classifications, and we also performed GO and pathway functional enrichment using Phyper, a function of R software. The parameters for Phyper were set as  $P$ -values 0.05 after Bonferroni correction. The phyper function in R was used to analyse the  $P$ -value for each function theme:

$$P = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}}$$

Smaller  $P$ -values were associated with greater enrichment of the candidate genes in a given function theme ([https://en.wikipedia.org/wiki/Hypergeometric\\_distribution](https://en.wikipedia.org/wiki/Hypergeometric_distribution)). The generation of Venn diagrams and hierarchical clustering heat maps in this study were conducted using the gmodels, Venn diagram and Pheatmap packages in R (<https://www.r-project.org/>) based on the gene list and the gene expression levels for each type.

### Pipeline for isoform sequencing analysis

To classify AS events, the tool AStalavista was employed using the raw.gtf file assembled from the Illumina RNA-seq and SMRT sequencing data. AS analysis was conducted using SpliceGrapher by converting the detected splice isoforms into splice graphs. Introns fully subsumed by an exon were labelled as retained. Overlapping exons that differed at their 5' or 3' splice junctions were considered alternative 5' or 3' splicing events, respectively. Finally, exons absent in other isoforms were considered exon skipping events.

MATS was used to call the differentially spliced events between the EF and NF samples at the three development periods, using the aligned.bam files as input with default settings. For comparison, the merged.gtf file derived from the SMRT and new Illumina data was used as a reference. The examined events included skipped exon (SE), alternative 5' splice site (A5), alternative 3' splice site (A3), alternative exons (AE), and retained intron (IR) events.

### RNA extraction, quantification, and RT-PCR

Total RNA was extracted using RNA Reagent (RN38; Aidlab Biotechnology, Beijing, China) according to the manufacturer's protocol and treated with RNase-free DNase I (Takara, Dalian, China) to remove genomic DNA contamination. First-strand cDNA was generated

from 1 µg of total RNA isolated from buds using the superscript first-strand synthesis system (Invitrogen, USA). The specific primers were designed using Primer 3plus and synthesized by Majorbiogene Co., Ltd. (Beijing, China). The melting temperature of the primers was 60 °C, and the amplicon lengths were 100–200 bp. Real-time qRT-PCR was performed on a Roche Light-Cycle 480 Real-Time PCR System (Roche Applied Science, Germany) using a SYBR Premix Ex Taq™ Kit (TaKaRa, Dalian, China) according to the manufacturer's instructions. *Cbu-actin* was used as an internal control, and each reaction was conducted in triplicate [31]. All the primers are shown in Table S12. Each reaction was performed in a total reaction mixture volume of 20 µL containing 2 µL of first-strand cDNA as template. The amplification program was as follows: 3 min at 95 °C and 30 cycles of 15 s at 95 °C, 30 s at 60 °C, 1 min at 72 °C and 10 min 72 °C. Each reaction was performed with three replicates. The expression levels of candidate genes were determined by CT values and calculated by the  $2^{-\Delta\Delta Ct}$  method. We test the correlation of expression (CEG) between mRNAs by using the Pearson correlation coefficient. The Pearson correlation coefficient was calculated by COR() using the average relative expression of three replicates in R [98].

### Transcription factor prediction and protein and protein interaction analysis

We found the ORF of each DEG by using getorf (version: EMBOSS: 6.5.7.0, parameters: -minsize 150, //www.bioinformatics.nl/cgi-bin/emboss/help/getorf). We aligned ORFs to transcription factor (TF) domains from PlntfDB (<http://plntfdb.bio.uni-potsdam.de/v3.0/>) by using hmmsearch (<http://hmmer.org>). We used DIAMOND (version: v0.8.31, parameters: --evaluate le-5 --outfmt 6 --max-target-seqs 1 --more-sensitive, <https://github.com/bbuchfink/diamond>) to map the DEGs to the STRING (version: v10, <http://string-db.org/>) database to obtain interactions between DEG-encoded proteins using homology with known proteins.

### Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12864-020-06918-y>.

**Additional file 1: Fig. S1.** Read length and quality in EF. (A) Read length distribution in the 1–2 kb library. (B) Read quality distribution in the 2–3 kb library. (C) Read length distribution in the above 3 kb library.

**Additional file 2: Fig. S2.** Read length and quality in NF. (A) Read length distribution in the 1–2 kb library. (B) Read quality distribution in the 2–3 kb library. (C) Read quality distribution in the 3–6 kb library. (D) Read quality distribution in the above 6 kb library.

**Additional file 3: Fig. S3.** Characterization of the *C. bungei* transcriptome by SMRT-seq. (A) Distribution of transcript lengths from different sequencing platforms. (B) Venn diagram showing the common

and unique annotated genes detected by SMRT and Illumina. (C) Distribution of the number of APA sites per gene.

**Additional file 4: Fig. S4.** Analysis of differential gene expression of the NF. (A) The 8 significant expression profiles of NF. (B) Partial KEGG pathways associated with the NF.

**Additional file 5: Fig. S5.** Hierarchical cluster tree showing all modules.

**Additional file 6: Fig. S6.** qRT-PCR validation of mRNA for 14 genes.

**Additional file 7: Fig. S7.** qRT-PCR validation of mRNA for 10 genes.

**Additional file 8: Fig. S8.** RT-PCR validation of AS events for 5 genes. This figure is supplemental to Fig. 7b.

**Additional file 9: Fig. S9.** RT-PCR validation of AS events in three development periods. This figure is supplemental to Fig. 7d.

**Additional file 10: Table S1.** Summary of Illumina short reads.

**Additional file 11: Table S2.** Statistics of the SMRT sequencing data.

**Additional file 12: Table S3.** Top 20 KEGG pathways significantly enriched with DEGs during floral transition in EF.

**Additional file 13: Table S4.** All predicted transcription factors in EF during floral transition.

**Additional file 14: Table S5.** Top 20 KEGG pathways significantly enriched with DEGs in three different stages between EF and NF.

**Additional file 15: Table S6.** Summary of all genes from the MEdarkturquoise module.

**Additional file 16: Table S7.** Summary of DEGs classified into five pathways.

**Additional file 17: Table S8.** The correlation coefficients of expression of mRNAs between the EF and NF during three developmental periods. *r* is the correlation coefficient of expression of the mRNAs.

**Additional file 18: Table S9.** Summary of identification of splicing junctions using NGS and SMRT data.

**Additional file 19: Table S10.** Summary of AS events from DEGs in EF.

**Additional file 20: Table S11.** Summary of AS events from DEGs in NF.

**Additional file 21: Table S12.** Summary of primers used in this study.

### Abbreviations

NGS: Short-read next-generation sequencing; SMRT: Single-molecule real-time; EF: Early flowering; NF: Normal flowering; WGCNA: Weighted Gene Co-expression Network Analysis; AS: Alternative splicing; SAM: Shoot apical meristem (SAM); SPL: Squamosa-promoter binding protein-like; TOC: Timing of cab expression 1; LUX: Luxarrhythm; PIF: Phytochrome interacting factor; CO: Constans; FRI: Frigida; GA20ox: GA20oxidases; GA3ox: GA3oxidases; SOCT1: Suppressor of overexpression of constans 1; Vb: Vegetative bud; Tb: Transition bud; Rb: Reproductive bud; APA: Alternative polyadenylation; STEM: Short Time-series Expression Miner; DEG: Differentially expressed gene; GASA: GA-Stimulated in Arabidopsis; HSP: Heat shock protein; NAC: Nascent polypeptide associated complex; MYB: v-myb avian myeloblastosis viral oncogene homolog; bHLH: basic helix-loop-helix protein; FPKM: Fragments per kilobase million; AP2: Apetala 2; Kd: Dissociation constants; SE: Exon Skipping; IR: Intron Retention; A5: Alternative 5' Splice Site; A3: Alternative 3' Splice Site; AE: Alternative Exon; GRF: Growth regulating factor; ARF: Auxin-responsive factor; G2-like: Golden2-Like; ERF: Ethylene-responsive factor; LFY: Leafy; FD: Flowering locus D

### Acknowledgements

The authors would like to appreciate the financial support from the Fundamental Research Funds of the Chinese Academy of Forestry.

### Authors' contributions

J.H.W., S.G.Z., G.Z.Q., L.I.S., Z.W. and W.J.M. designed the experiments. Z.W. and W.J.M. analysed the RNA-seq data and wrote the manuscript. Z.W., and T.Q.Z. detected the expression of genes using qRT-PCR. Z.W., N.L., F.Q.O.Y., N.W. and G.J.Y. collected the samples used in the experiment. All the authors have read the paper and agreed to list their names as co-authors.



### Funding

This work was supported by Fundamental Research Funds of Chinese Academy of Forestry (CAFYBB2017ZY002) and Fundamental Research Funds of Chinese Academy of Forestry (CAFYBB2017ZA001–8). The funders had no role in the study design, data analysis and interpretation, and manuscript writing, and only provided financial support.

### Availability of data and materials

The sequencing data have been submitted to the Sequence Read Archive (SRA) at the National Centre for Biotechnology Information. The accession number is PRJNA414524. The data supporting the conclusions of the article have been uploaded as additional files.

### Ethics approval and consent to participate

This article does not contain any studies with human participants or animals performed by any of the authors. All the plant materials used in this study were provided by the Research Institute of Forestry, Chinese Academy of Forestry. The field experiments were conducted under local legislation and permissions.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Author details

<sup>1</sup>State Key Laboratory of Tree Genetics and Breeding, Key Laboratory of Tree Breeding and Cultivation of State Forestry Administration, Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, PR China.

<sup>2</sup>Department of Biology Centre for Forest Biology, University of Victoria, Victoria, BC 11, Canada. <sup>3</sup>State Key Laboratory of Tree Genetics and Breeding (Northeast Forestry University), 26 Hexing Road, Harbin 150040, PR China.

Received: 9 November 2019 Accepted: 15 July 2020

Published online: 22 July 2020

### References

- Poethig RS. Phase Change and the Regulation of Developmental Timing in Plants. *Science* (80- ). 2003;301:334–6.
- Ung N, Smith HMS, Ung N, Smith HMS. Regulation of shoot meristem integrity during *Arabidopsis* vegetative development. *Plant Signal Behav*. 2011;6:1250–2.
- Tooke F, Tooke F, Battey N. Models of shoot apical meristem function. *New Phytol*. 2003;159:37–52.
- Bouché F, Lobet G, Tocquin P, Périlleux C. FLOR-ID: An interactive database of flowering-time gene networks in *Arabidopsis thaliana*. *Nucleic Acids Res*. 2016;44:D1167–71.
- Wang J. Regulation of flowering time by the miR156-mediated age pathway. *J Exp Bot*. 2014;65:4723–7430.
- Wang JW, Czech B, Weigel D. miR156-Regulated SPL Transcription Factors Define an Endogenous Flowering Pathway in *Arabidopsis thaliana*. *Cell*. 2009;138:738–49.
- Yu S, Galvao VC, Zhang Y-C, Horrer D, Zhang T-Q, Hao Y-H, et al. Gibberellin Regulates the *Arabidopsis* Floral Transition through miR156-Targeted SQUAMOSA PROMOTER BINDING-LIKE Transcription Factors. *Plant Cell*. 2012; 24:3320–32. <https://doi.org/10.1105/tpc.112.101014>.
- Hazen SP, Schultz TF, Pruneda-Paz JL, Borevitz JO, Ecker JR, Kay SA. LUX ARRHYTHMO encodes a Myb domain protein essential for circadian rhythms. *Proc Natl Acad Sci U S A*. 2005;102:10387–92. <https://doi.org/10.1073/pnas.0503029102>.
- Hayama R, Sarid-Krebs L, Richter R, Fernández V, Jang S, Coupland G. PSEUDO RESPONSE REGULATORS stabilize CONSTANS protein to promote flowering in response to day length. *EMBO J*. 2017;36:904–18. <https://doi.org/10.15252/embj.201693907>.
- Abe E, Fujino K, Masuda K, Yamaguchi Y, Abe E. Isolation and Expression Profiling of a CONSTANS-Like Gene and Two FLOWERING LOCUS T-Like Genes from *Spinacia oleracea* L. *Am J Plant Sci*. 2014;5:4018–28. <https://doi.org/10.4236/ajps.2014.526420>.
- Johanson U, West J, Lister C, Michaels S, Amasino R, Dean C. Molecular Analysis of FRIGIDA, a Major Determinant of Natural Variation in *Arabidopsis* Flowering Time. *Science* (80- ). 2000;290:13.
- Reinecke DM, Wickramaratna AD, Ozga JA, Kurepin LV, Jin AL, Good AG, et al. Gibberellin 3-oxidase Gene Expression Patterns Influence Gibberellin Biosynthesis, Growth, and Development in Pea. *Plant Physiol*. 2013;163:929–45. <https://doi.org/10.1104/pp.113.225987>.
- Tang M, Tao Y Bin, Fu Q, Song Y, Niu L, Xu ZF. An ortholog of LEAFY in *Jatropha curcas* regulates flowering time and floral organ development. *Sci Rep*. 2016;6:1–15. <https://doi.org/10.1038/srep37306>.
- Yamaguchi A, Kobayashi Y, Goto K, Abe M, Araki T. TWIN SISTER OF FT (TSF) acts as a floral pathway integrator redundantly with FT. *Plant Cell Physiol*. 2005;46:1175–89.
- Carmona MJ, Calonje M, Martínezzapater JM. The FT/TFL1 gene family in grapevine. *Plant Mol Biol*. 2007;63:637–50.
- Irish VF. The flowering of *Arabidopsis* flower development. *Plant J*. 2010;61: 1014–28.
- Jung JH, Lee HJ, Ryu JY, Park CM. SPL3/4/5 Integrate Developmental Aging and Photoperiodic Signals into the FT-FD Module in *Arabidopsis* Flowering. *Mol Plant*. 2016;9:1647–59. <https://doi.org/10.1016/j.molp.2016.10.014>.
- McClung CR. Plant Circadian Rhythms. *Plant Cell Online*. 2006;18:792–803. <https://doi.org/10.1105/tpc.106.040980>.
- Pose D, Pose LV, Verhage L, Ott F, Yant L, Mathieu J, et al. Temperature-dependent regulation of flowering by antagonistic FLM variants. *Nature*. 2013;503:414.
- Marshall CM, Tartaglio V, Duarte M, Harmon FG. The *Arabidopsis* sickle Mutant Exhibits Altered Circadian Clock Responses to Cool Temperatures and Temperature-Dependent Alternative Splicing. *Plant Cell*. 2016;28:2560–75.
- Niwa Y, Ito S, Nakamichi N, Mizoguchi T, Niinuma K, Yamashino T, et al. Genetic Linkages of the Circadian Clock-Associated Genes, TOC1, CCA1 and LHY, in the Photoperiodic Control of Flowering Time in *Arabidopsis thaliana*. *Plant cell physiol*. 2007;48:925–37.
- Lu SX, Webb CJ, Knowles SM, Kim SHJ, Wang Z, Tobin EM. CCA1 and ELF3 Interact in the Control of Hypocotyl Length and Flowering Time in *Arabidopsis*. *Plant Physiol*. 2012;158:1079–88. <https://doi.org/10.1104/pp.111.189670>.
- Zhang X, An L, Nguyen TH, Liang H, Wang R, Liu X, et al. The cloning and functional characterization of peach CONSTANS and FLOWERING LOCUS T homologous genes PpCO and PpFT. *PLoS One*. 2015;10:1–16.
- Lee J, Lee I. Regulation and function of SOC1, a flowering pathway integrator. *J Exp Bot*. 2010;61:2247–54.
- Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, et al. A gene expression map of *Arabidopsis thaliana* development. *Nat Genet*. 2005; 37:501–6.
- Shrestha Roshi, Gómez-Ariza J, Brambilla V, Fabio F. Molecular control of seasonal flowering in rice, *Arabidopsis* and temperate cereals. *Ann Bot*. 2014;114:1445–1458.
- Hsu C, Adams JP, Kim H, No K, Ma C, Strauss SH, et al. FLOWERING LOCUS T duplication coordinates reproductive and vegetative growth in perennial poplar. *PNAS*. 2011;108.
- Koskela EA, Mouhu K, Albani MC, Kurokura T, Rantanen M, Sargent DJ, et al. Mutation in TERMINAL FLOWER1 Reverses the Photoperiodic Requirement for Flowering in the Wild Strawberry *Fragaria vesca*. *Plant Physiol*. 2012;159: 1043–54.
- Wang Z, Zhu T, Ma W, Wang N, Qu G, Zhang S, et al. Genome-wide analysis of long non-coding RNAs in *Catalpa bungei* and their potential function in floral transition using high-throughput sequencing. *BMC Genet*. 2018;19:1–16.
- Jing D, Xia Y, Chen F, Wang Z, Zhang S, Wang J. Ectopic expression of a *Catalpa bungei* (Bignoniaceae) PISTILLATA homologue rescues the petal and stamen identities in *Arabidopsis* pi-1 mutant. *Plant Sci*. 2015;231:40–51. <https://doi.org/10.1016/j.plantsci.2014.11.004>.
- Wang Z, Zhu T, Ma W, Fan E, Lu N, Ouyang F, et al. Potential function of CbuSPL and gene encoding its interacting protein during flowering in *Catalpa bungei*. *BMC Plant Biol*. 2020;20:1–14.
- Xiao Y, Yi F, Ling J, Wang Z, Zhao K, Lu N, et al. Transcriptomics and proteomics reveal the cellulose and pectin metabolic processes in the tension wood (Non-g-layer) of *Catalpa bungei*. *Int J Mol Sci*. 2020;21:1–19.
- Li Y, Dai C, Hu C, Liu Z, Kang C. Global identification of alternative splicing via comparative analysis of SMRT- and Illumina-based RNA-seq in strawberry. *Plant J*. 2017;90:164–76.

34. Abdel-Ghany SE, Hamilton M, Jacobi JL, Ngam P, Devitt N, Schilkey F, et al. A survey of the sorghum transcriptome using single-molecule long reads. *Nat Commun.* 2016;7:1–11. <https://doi.org/10.1038/ncomms11706>.
35. Keren H, Lev-Maor G, Ast G. Alternative splicing and evolution: Diversification, exon definition and function. *Nat Rev Genet.* 2010;11:345–55. <https://doi.org/10.1038/nrg2776>.
36. Mandadi KK, Scholthof KG. Genome-Wide Analysis of Alternative Splicing Landscapes Modulated during Plant-Virus Interactions in *Brachypodium distachyon*. *J Immunol.* 2015;27:71–85.
37. Schlaen RG, Mancini E, Sanchez SE, Perez-Santángelo S, Rugnone ML, Simpson CG, et al. The spliceosome assembly factor GEMIN2 attenuates the effects of temperature on alternative splicing and circadian rhythms. *PNAS.* 2015;112:9382–7.
38. Shikata H, Hanada K, Ushijima T, Nakashima M, Suzuki Y. Phytochrome controls alternative splicing to mediate light responses in Arabidopsis. *PNAS.* 2014;111:18781–6.
39. Staiger D, Brown JWS. Alternative Splicing at the Intersection of Biological Timing, Development, and Stress Responses. *Plant Cell.* 2013;25:3640–56.
40. Su C, Zhao H, Zhao Y, Ji H, Wang Y, Zhi L, et al. RUG3 and ATM synergistically regulate the alternative splicing of mitochondrial nad2 and the DNA damage response in *Arabidopsis thaliana*. *Sci Rep.* 2017;7: 2016:1–14. doi:<https://doi.org/10.1038/srep43897>.
41. Thatcher SR, Zhou W, Leonard A, Wang B-B, Beatty M, Zastrow-Hayes G, et al. Genome-Wide Analysis of Alternative Splicing in *Zea mays*: Landscape and Genetic Regulation. *Plant Cell.* 2014;26:3472–87. <https://doi.org/10.1105/tpc.114.130773>.
42. James AB, Syed NH, Bordage S, Marshall J, Nimmo GA, Jenkins GI, et al. Alternative Splicing Mediates Responses of the Arabidopsis Circadian Clock to Temperature Changes. *Plant Cell.* 2012;24:961–81. <https://doi.org/10.1105/tpc.111.093948>.
43. Hofmann NR. Alternative Splicing Links the Circadian Clock to Cold Tolerance. *Plant Cell.* 2012;24:2238.
44. Qin Z, Wu J, Geng S, Feng N, Chen F, Kong X, et al. Regulation of FT splicing by an endogenous cue in temperate grasses. *Nat Commun.* 2017;8: 1–12. <https://doi.org/10.1038/ncomms14320>.
45. Hartmann L, Drewe-Bob P, Wiebner T, Wagner G, Geue S, Lee H-C, et al. Alternative Splicing Substantially Diversifies the Transcriptome during Early Photomorphogenesis and Correlates with the Energy Availability in Arabidopsis. *Plant Cell.* 2016;28:2715–34. <https://doi.org/10.1105/tpc.16.00508>.
46. Mao Y, Sun J, Cao P, Zhang R, Fu Q, Chen S, et al. Functional analysis of alternative splicing of the FLOWERING LOCUS T orthologous gene in *Chrysanthemum morifolium*. *Nature.* 2016;3:1–8.
47. Guo X, Yu C, Luo L, Wan H, Zhen N, Xu T, et al. Transcriptome of the floral transition in *Rosa chinensis* ' Old Blush. *BMC Genomics.* 2017;18:1–18.
48. Schwechheimer C, Bevan M. The regulation of transcription factor activity in plants. *Trends Plant Sci.* 1998;3:378–83. [https://doi.org/10.1016/S1360-1385\(98\)01302-8](https://doi.org/10.1016/S1360-1385(98)01302-8).
49. Sheen J. Metabolic Repression of Transcription in Higher Plants. *Plant Cell Online.* 1990;2:1027–38. <https://doi.org/10.1105/tpc.2.10.1027>.
50. Liu L, White MJ, Macrae TH. Functional domains , evolution and regulation. *Eur j Biochem.* 1999;257:247–57.
51. Fan M, Bai M-Y, Kim J-G, Wang T, Oh E, Chen L, et al. The bHLH Transcription Factor HBI1 Mediates the Trade-Off between Growth and Pathogen-Associated Molecular Pattern-Triggered Immunity in Arabidopsis. *Plant Cell.* 2014;26:828–41. <https://doi.org/10.1105/tpc.113.121111>.
52. Tian C, Wan P, Sun S, Li J, Chen M. Genome-wide analysis of the GRAS gene family in rice and Arabidopsis. *Plant Mol Biol.* 2004;54:519–32. <https://doi.org/10.1023/B:PLAN.0000038256.89809.57>.
53. Bolle C. The role of GRAS proteins in plant signal transduction and development. *Planta.* 2004;218:683–92.
54. Zhang K, Wang R, Zi H, Li Y, Cao X, Li D, et al. AUXIN RESPONSE FACTOR3 Regulates Floral Meristem Determinacy By Repressing Cytokinin Biosynthesis and Signaling. *Plant Cell.* 2018;30:705. <https://doi.org/10.1105/tpc.17.00705>.
55. Gras DE, Vidal EA, Undurraga SF, Riveras E, Moreno S, Dominguez-Figueroa J, et al. SMZ/SNZ and gibberellin signaling are required for nitrate-elicited delay of flowering time in *Arabidopsis thaliana*. *J Exp Bot.* 2018;69:619–31.
56. Xing L, Zhang D, Li Y, Zhao C, Zhang S, Shen Y, et al. Genome-wide identification of vegetative phase transition-associated microRNAs and target predictions using degradome sequencing in *Malus hupehensis*. *BMC Genomics.* 2014;15:1125.
57. Xu Z, Sun L, Zhou Y, Yang W, Cheng T, Wang J, et al. Identification and expression analysis of the SQUAMOSA promoter-binding protein (SBP)-box gene family in *Prunus mume*. *Mol Genet Genomics.* 2015;290:1701–15.
58. Yu N, Cai W-J, Wang S, Shan C-M, Wang L-J, Chen X-Y. Temporal Control of Trichome Distribution by MicroRNA156-Targeted SPL Genes in *Arabidopsis thaliana*. *Plant Cell Online.* 2010;22:2322–35. <https://doi.org/10.1105/tpc.109.072579>.
59. Srikanth B, Subhakara Rao I, Surekha K, Subrahmanyam D, Voleti SR, Neeraja CN. Enhanced expression of OsSPL14 gene and its association with yield components in rice (*Oryza sativa*) under low nitrogen conditions. *Gene.* 2016;576:441–50. <https://doi.org/10.1016/j.gene.2015.10.062>.
60. Tanaka N, Itoh H, Sentoku N, Kojima M, Sakakibara H, Izawa T, et al. The COP1 Ortholog PPS Regulates the Juvenile-Adult and Vegetative-Reproductive Phase Changes in Rice. *Plant Cell.* 2011;23:2143–54.
61. Zhang S, Ling L. Evolution and divergence of SBP-box genes in land plants. *BMC Genomics.* 2015;16:787.
62. Wang H, Pan J, Li Y, Lou D, Hu Y, Yu D. The DELLA-CONSTANS Transcription Factor Cascade Integrates Gibberellic Acid and Photoperiod Signaling to Regulate Flowering. *Plant Physiol.* 2016;172:479–88.
63. Chen J, Wang X, Hu B, He Y, Qian X, Wang W. Candidate genes in gastric cancer identified by constructing a weighted gene co-expression network. *PeerJ.* 2018;6:1–19.
64. Liu Y, Cui S, Wu F, Yan S, Lin X, Du X, et al. Functional Conservation of MIKC-Type MADS Box Genes in Arabidopsis and Rice Pollen Maturation. *Plant Cell.* 2013;25:1288–303.
65. Zheng H, Zhang F, Wang S, Su Y, Ji X, Jiang P, et al. MLK1 and MLK2 Coordinate RGA and CCA1 Activity to Regulate Hypocotyl Elongation in *Arabidopsis thaliana*. *Plant Cell.* 2017;30:67–82.
66. Paolacci AR, Tanzarella OA, Porceddu E, Varotto S, Ciaffi M. Molecular and phylogenetic analysis of MADS-box genes of MIKC type and chromosome location of SEP -like genes in wheat (*Triticum aestivum* L.). *Mol Genet Genomics.* 2007;16:35020.
67. Adamczyk BJ, Fernandez DE. MIKC \* MADS Domain Heterodimers Are Required for Pollen Maturation and Tube Growth in Arabidopsis. *Plant Physiol.* 2009;149:1713–23.
68. Prunet N, Morel P, Champelovier P, Thierry A, Negruțiu I, Jack T, et al. SQUINT promotes stem cell homeostasis and floral meristem termination in Arabidopsis through APETALA2 and CLAVATA signalling. *J Exp Bot.* 2015;66: 6905–16.
69. Kornblihtt AR, Schor IE, Alló M, Dujardin G, Petrillo E, Muñoz MJ. Alternative splicing: A pivotal step between eukaryotic transcription and translation. *Nat Rev Mol Cell Biol.* 2013;14:153–65. <https://doi.org/10.1038/nrm3525>.
70. Nancy A. Eckardt. The Plant Cell Reviews Alternative Splicing. *Plant Cell.* 2013;25(10):3639.
71. Zhu FY, Chen MX, Ye NH, Shi L, Ma KL, Yang JF, et al. Proteogenomic analysis reveals alternative splicing and translation as part of the abscisic acid response in Arabidopsis seedlings. *Plant J.* 2017;91:518–33.
72. Melzer R. Regulation of flowering time: A splicy business. *J Exp Bot.* 2017;68: 5017–20.
73. Ning G, Cheng X, Luo P, Liang F, Wang Z, Yu G, et al. Hybrid sequencing and map finding ( HySeMaFi ): optional strategies for extensively deciphering gene splicing and expression in organisms without reference genome. *Sci Rep.* 2017;4:3793. doi:<https://doi.org/10.1038/srep43793>.
74. He L, Fu S, Xu Z, Yan J, Xu J, Zhou H, et al. Hybrid sequencing of full-length cDNA transcripts of stems and leaves in *Dendrobium officinale*. *Genes (Basel).* 2017;8:1–13.
75. Gol L, Tomé F, Von Korff M. Floral transitions in wheat and barley: Interactions between photoperiod, abiotic stresses, and nutrient status. *J Exp Bot.* 2017;68:1399–410.
76. Lazaro A, Mouriz A, Piñeiro M, Jarillo JA. Red Light-Mediated Degradation of CONSTANS by the E3 Ubiquitin Ligase HOS1 Regulates Photoperiodic Flowering in Arabidopsis. *Plant Cell.* 2015;27(9):2437.
77. Hyun Y, Richter R, Coupland G. Competence to flower: Age-controlled sensitivity to environmental cues. *Plant Physiol.* 2017;173:36–46.
78. Kim DH, Sung S. Coordination of the vernalization response through a VIN3 and FLC gene family regulatory network in Arabidopsis. *Plant Cell.* 2013;25: 454–69.
79. Ayako Yamaguchi, Miin-Feng Wu, Li Yang, Gang Wu, R. Scott Poethig and DW. The microRNA regulated SBP-box transcription factor SPL3 is a direct upstream activator of LEAFY, FRUITFULL and APETALA1. *Dev Cell.* 2010;17: 268–78.

80. Jung JH, Ju Y, Seo PJ, Lee JH, Park CM. The SOC1-SPL module integrates photoperiod and gibberellic acid signals to control flowering time in *Arabidopsis*. *Plant J*. 2012;69:577–88.
81. Wang J, Czech B, Weigel D. miR156-Regulated SPL Transcription Factors Define an Endogenous Flowering Pathway in *Arabidopsis thaliana*. 2009; 138(4):738–49.
82. Bäurle I, Dean C. The Timing of Developmental Transitions in Plants. *Cell*. 2006;125:655–64.
83. Simpson GG, Laurie RE, Dijkwel PP, Quesada V, Stockwell PA, Dean C, et al. Noncanonical Translation Initiation of the *Arabidopsis* Flowering Time and Alternative Polyadenylation Regulator FCA. *Plant Cell*. 2010;22:3764–77.
84. Simpson GG, Dijkwel PP, Quesada V, Henderson I, Dean C. FY is an RNA 3' end-processing factor that interacts with FCA to control the *Arabidopsis* floral transition. *Cell*. 2003;113:777–87.
85. Macknight R. Functional Significance of the Alternative Transcript Processing of the *Arabidopsis* Floral Promoter FCA. *Plant Cell Online*. 2002;14:877–88. <https://doi.org/10.1105/tpc.010456>.
86. Quesada V, Macknight I, Dean C, Simpson GG. Autoregulation of FCA pre-mRNA processing controls *Arabidopsis* flowering time. *EMBO J*. 2003;22:3142–52.
87. Eriksson S. GA4 Is the Active Gibberellin in the Regulation of LEAFY Transcription and *Arabidopsis* Floral Initiation. *Plant Cell Online* *Plant Cell Online*. 2006;18:2172–81. <https://doi.org/10.1105/tpc.106.042317>.
88. Murase K, Hirano Y, Sun TP, Hakoshima T. Gibberellin-induced DELLA recognition by the gibberellin receptor GID1. *Nature*. 2008;456:459–63.
89. Sun S, Wang H, Yu H, Zhong C, Zhang X, Peng J, et al. GASA14 regulates leaf expansion and abiotic stress resistance by modulating reactive oxygen species accumulation. *J Exp Bot*. 2013;64:1637–47.
90. Roxrud I, Lid SE, Fletcher JC, Schmidt EDL, Opsahl-Sorteberg HG. GASA4, one of the 14-member *Arabidopsis* GASA family of small polypeptides, regulates flowering and seed development. *Plant Cell Physiol*. 2007;48:471–83.
91. Zhang S, Yang C, Peng J, Sun S, Wang X. GASAs, a regulator of flowering time and stem growth in *Arabidopsis thaliana*. *Plant Mol Biol*. 2009;69:745–59.
92. Appleford NEJ, Evans DJ, Lenton JR, Gaskin P, Croker SJ, Devos KM, et al. Function and transcript analysis of gibberellin-biosynthetic enzymes in wheat. *Planta*. 2006;4:568–82.
93. Wang M, Wang P, Liang F, Ye Z, Li J, Shen C, et al. A global survey of alternative splicing in allopolyploid cotton: landscape, complexity and regulation. *New Phytol*. 2018;217:163–78.
94. Thatcher SR, Danilevskaya ON, Meng X, Beatty M, Zastrow-Hayes G, Harris C, et al. Genome-Wide Analysis of Alternative Splicing during Development and Drought Stress in Maize. *Plant Physiol*. 2016;170:586–99. <https://doi.org/10.1104/pp.15.01267>.
95. Filichkin SA, Priest HD, Givan SA, Shen R, Bryant DW, Fox SE, et al. Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. *Genome Res*. 2010;20:45–58.
96. Zhang G, Guo G, Hu X, Zhang Y, Li Q, Li R, et al. Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome. *Genome Res*. 2010;20:646–54.
97. Du J, Wang S, He C, Zhou B, Ruan Y, Shou H. Identification of regulatory networks and hub genes controlling soybean seed set and size using RNA sequencing analysis. *J Exp Bot*. 2017;68:1955–72.
98. Kang C, Liu Z. Global identification and analysis of long non-coding RNAs in diploid strawberry *Fragaria vesca* during flower and fruit development. *BMC Genomics*. 2015;16:1–15. <https://doi.org/10.1186/s12864-015-2014-2>.

## Publisher's Note

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

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

