Article

Identification and Analysis of the AP2 Subfamily Transcription Factors in the Pecan (*Carya illinoinensis*)

Zhengfu Yang [†], Hongmiao Jin [†], Junhao Chen, Caiyun Li [ID], Jiani Wang, Jie Luo and Zhengjia Wang *

State Key Laboratory of Subtropical Silviculture, Zhejiang A&F University, Lin'an District, Hangzhou 311300, China; zafuyzf@163.com (Z.Y.); jhm@stu.zafu.edu.cn (H.J.); jlocke@163.com (J.C.); licaiyun@stu.zafu.edu.cn (C.L.); wangjiani@stu.zafu.edu.cn (J.W.); luojie0103@Outlook.com (J.L.)
* Correspondence: wangzj@zafu.edu.cn
† These authors contributed equally to this work.

Abstract: The AP2 transcriptional factors (TFs) belong to the APETALA2/ ethylene-responsive factor (AP2/ERF) superfamily and regulate various biological processes of plant growth and development, as well as response to biotic and abiotic stresses. However, genome-wide research on the AP2 subfamily TFs in the pecan (*Carya illinoinensis*) is rarely reported. In this paper, we identify 30 AP2 subfamily genes from pecans through a genome-wide search, and they were unevenly distributed on the pecan chromosomes. Then, a phylogenetic tree, gene structure and conserved motifs were further analyzed. The 30 AP2 genes were divided into euAP2, euANT and basalANT three clades. Moreover, the *cis*-acting elements analysis showed many light responsive elements, plant hormone-responsive elements and abiotic stress responsive elements are found in *CiAP2* promoters. Furthermore, a qPCR analysis showed that genes clustered together usually shared similar expression patterns in euAP2 and basalANT clades, while the expression pattern in the euANT clade varied greatly. In developing pecan fruits, *CiAP2-5*, *CiANT1* and *CiANT2* shared similar expression patterns, and their expression levels decreased with fruit development. *CiANT5* displayed the highest expression levels in developing fruits. The subcellular localization and transcriptional activation activity assay demonstrated that CiANT5 is located in the nucleus and functions as a transcription factor with transcriptional activation activity. These results help to comprehensively understand the pecan AP2 subfamily TFs and lay the foundation for further functional research on pecan AP2 family genes.

Keywords: pecan; AP2 family; transcription factor; *CiANT5*

1. Introduction

The APETALA2/ethylene-responsive factor (AP2/ERF) is a superfamily of transcription factors (TFs), which were characterized by conserved AP2 domains. AP2/ERF superfamily TFs are mainly divided into the ERF (ethylene-responsive-element-binding protein), DREB (dehydration responsive element binding), AP2 (APETALA2), RAV (Related to ABI3/VP) and Soloists five major subfamilies [1,2]. Both the ERF and DREB subfamily members only contain one AP2 domain, and both are further subdivided into six subgroups [3]. The AP2 subfamily members containing two consecutive AP2 domains are further subdivided into euAP2, basalANT and euANT clades depending on the nuclear localization and amino acid sequences as well as other conserved motifs [4]. The RAV subfamily members contain a single AP2 domain and a B3 DNA-binding domain [5]. Although the Soloist subfamily with a single AP2 domain is displayed in most plants, its sequence and structure distinguish it from other AP2/ERF superfamily TFs [2,3]. The AP2/ERF domain consisted of a 60–70 amino acid functioning DNA binding domain; however, each of its subfamily recognized specific *cis*-acting elements. For example, ERF subfamily members usually bind to GCC-box responding to ethylene, pathogens and wounding [6], while DREB subfamily TFs usually bind to the DRE/CRT(A/GCCGAC) cis-element to

activate drought and salt responsive genes [7]. The second AP2 domain in the AP2 subfamily binding to T/A-rich elements positively regulates floral organ identity [8], while in RAV1 it binds to the CAACA motif involved in ABA signaling [9,10]. Taken together, the diversification of gene structure and DNA-binding affinities may contribute to the different functions of the AP2/ERF superfamily members.

The AP2/ERF superfamily members have been reported to serve as regulators of stress responses and development processes, especially in reproductive organ development. *APETALA2* (*AP2*), expressed in floral organs and developing ovules as well as nonfloral organs, was initially identified as controlling flower and seed development in *Arabidopsis* [11]. Further research revealed that AP2 could bind to the AT-rich target sequence of *AGAMOUS* (*AG*) and directly activate its expression through the second AP2 domain [8]. *AINTEGUMENTA* (*ANT*), which belongs to the euANT clade, regulated cell division and organ size. The ectopic expression of *ANT* enlarged embryonic and all shoot organs by increasing cell numbers in *Arabidopsis* [12,13]. Moreover, seven high sequence similarity genes, *AINTEGUMENTA-like* (*AIL*), were identified, and they primarily expressed in young actively dividing tissues. The expression analysis by in situ hybridization demonstrated that *AIL5*, *AIL6* and *AIL7* expressed distinct spatial and temporal patterns in inflorescence meristems and flowers [14]. The overexpression of *AIL5* resulted in similar large floral patterns with ectopic expressions of *ANT*, while *AIL6*'s functioned redundancy with *ANT* took place during flower development [14,15]. In rice, the AP2 factor gene *RSR1* was identified as a negative regulator of type I starch syntheses genes, and the loss of function of *RSR1* resulted in an altered amylopectin structure and increased gelatinization temperature [16]. Interestingly, the same genetic locus was identified as the candidate of major QTL *qHD5*, controlling the heading date in rice [17]. *FRIZZY PANICLE* (*FZP*) was required to establish rice floral meristem identity and determine grain size [18,19]. *SUPERNUMERARY BRACT* (*SNB*) encoding *APETALA2-like* TF controlled rice seed shattering and seed size. A point mutation in *SNB* resulted in alternative splicing and reduced shattering in its mutant lines [20]. *EXCESSIVE NUMBEROF FLORAL ORGANS* (*ENO*) was identified as a tomato fruit regulator and regulating *WUSCHEL* expression. The disruption of *ENO* expanded the expression domain of *SlWUS* in a flower-specific manner, resulting in enlarged fruit size [21]. These studies demonstrated that the AP2 family TFs participated in the regulation of development processes, such as flower development, heading date, spikelet development and seed development.

With advances in sequencing technology, whole genomic sequences of more plants are available. The AP2 family genes were genome-wide identified and analyzed in increasingly more plant species, such as barley [22], common wheat [23] and *Dendrobium officinale* [24], which contributed to the understanding of the molecular function of AP2 family genes. However, the AP2 family TFs have not been reported in pecan (*Carya illinoinensis* (Wangenh.) K. Koch), which is considered a famous nut and woody oil tree species rich in unsaturated fatty acids [25]. In this study, we identify 30 AP2 subfamily TFs from pecans through a genome-wide search and mapped them on pecan chromosomes. Then, we construct a phylogenetic tree, analyzed gene structure and conserved motifs as well as promoter *cis*-acting elements. Furthermore, a qPCR is conducted to analyze their expression pattern, then subcellular localization and transcriptional activation activity are investigated. These results enrich our understanding of the AP2 family in pecans and lay the foundation for further functional research on pecan AP2 family genes.

2. Results

2.1. Identification of AP2 Subfamily Genes in the Pecan

To identify AP2/ERF family genes, an HMM search was employed to ascertain the genes containing the AP2 domain (Pfam ID:PF00847) from rice, *Arabidopsis* and pecan protein datasets. Then, 207 putative genes were retrieved from the pecan, and an evolutionary tree of AP2 family genes was initially constructed (Figure S1). According to the number of AP2 domains contained, AP2/ERF family genes can be usually divided into ERF, AP2

and RAV subfamilies [4]. Furthermore, we found that both *Arabidopsis AINTEGUMENTA* (*ANT*) [26] and rice *qHD5* [17], containing two AP2 domains, were grouped into the AP2 subfamily. In this study, we specially focused on the pecan AP2 subfamily genes and further checked the candidate sequences through Smart (http://smart.embl-heidelberg.de, accessed on 12 July 2021). Interestingly, although *OF21942-RA*, *OF26251-RA*, *OF27743-RA* and *OF31010-RA* were clustered into the AP2 subfamily (Figure S1), they do not contain two AP2 domains. In this study, we preferentially focused on members containing two AP2 domains, and finally obtained 30 AP2 subfamily genes from pecan renamed from *CiAP2-1* to *CiAP2-7* and *CiANT1* to *CiANT23*, depending on their structural characteristics (Table 1). Furthermore, we analyzed the genomic distribution of AP2 genes, and results demonstrated that *CiAP2* genes were mainly distributed on chromosome 1, 6, 8 and 13 (Figure 1). Unfortunately, *CiANT13*, *16*, *19*, *20* and *22* were not anchored to the corresponding chromosomes.

Table 1. AP2 subfamily genes in Pecan.

No.	Gene ID	Gene Name	Gene Location	Predicted Protein Length (aa)	Isoelectric Point (pI)	Clade
1	OF00880-RA	CiAP2-1	Chr13: 16145978-16149358	476	8.28	
2	OF05517-RA	CiAP2-2	Chr16:22489793-22494373	504	8.82	
3	OF07993-RA	CiAP2-3	Chr01:62982646-62985262	504	7.43	
4	OF11627-RA	CiAP2-4	Chr01:65432895-65435508	468	6.46	euAP2
5	OF14885-RA	CiAP2-5	Chr05:11211192-11213844	502	7.26	
6	OF17041-RA	CiAP2-6	Chr01:1638337-1641093	436	7.10	
7	OF32680-RA	CiAP2-7	Chr14:27110789-27113491	510	7.08	
8	OF02559-RA	CiANT1	Chr06:6672431-6675340	556	6.85	
9	OF02759-RA	CiANT2	Chr06:4414801-4417762	509	7.78	
10	OF03547-RA	CiANT3	Chr08:12993211-12997175	509	7.04	
11	OF09391-RA	CiANT4	Chr06:25934540-25938961	511	6.57	
12	OF10594-RA	CiANT5	Chr08:36743529-36746728	681	7.01	
13	OF10749-RA	CiANT6	Chr08:38789911-38792954	511	6.90	
14	OF13068-RA	CiANT7	Chr08:24346944-24350910	712	6.40	euANT
15	OF18987-RA	CiANT8	Chr06:34173511-34177430	754	6.88	
16	OF22559-RA	CiANT9	Chr13:29470995-29473785	547	6.68	
17	OF22570-RA	CiANT10	Chr13:29876966-29879708	547	6.68	
18	OF32746-RA	CiANT11	Chr14:28189896-28192903	585	6.37	
19	OF32974-RA	CiANT12	Chr01:7995878-7999758	546	6.72	
20	OF08566-RA	CiANT13	ContigUN:49571770-49574715	350	8.28	
21	OF13593-RA	CiANT14	Chr11:34909386-34912135	358	7.64	
22	OF18890-RA	CiANT15	Chr02:51565314-51568192	366	7.39	
23	OF22034-RA	CiANT16	ContigUN:48433777-48436671	370	8.30	
24	OF24164-RA	CiANT17	Chr01:6433685-6436815	356	7.02	
25	OF24522-RA	CiANT18	Chr02:53441378-53444256	330	8.99	basalANT
26	OF28856-RA	CiANT19	ContigUN:3940772-3949092	274	6.26	
27	OF28858-RA	CiANT20	ContigUN:3974783-3976648	368	9.25	
28	OF29459-RA	CiANT21	Chr16:14006190-14016673	888	7.16	
29	OF30183-RA	CiANT22	ContigUN:15665370-15668325	343	8.06	
30	OF31987-RA	CiANT23	Chr13:7041337-7053094	608	4.68	

2.2. Phylogenetic Analysis of Pecan AP2 Subfamily Genes

To further reveal the evolutionary relationships among pecan AP2 subfamily proteins, we constructed a phylogenetic tree using the sequences of 30 pecan, 17 rice and 14 *Arabidopsis* AP2 subfamily proteins (Figure 2). The results were credible with high bootstrap values and demonstrated that pecan AP2 subfamily proteins could also be classified into the euAP2, euANT and basalANT three subgroups consistent with the previous study on AP2 family genes in *Arabidopsis* and rice [4]. We noted that *Arabidopsis ANT* (*AT4G37750*) and rice *qHD5* (*Os05g03040*) were clustered into euANT and euAP2, respectively. *ANT* is

involved in ovule and female gametophyte development and controlled plant organ size by regulating growth and cell numbers during organogenesis in *Arabidopsis* [12,13]. *qHD5* was reported as a candidate gene for heading date in rice [17]. Therefore, although all AP2 subfamily genes contain two AP2 domains, they may regulate different biological processes.
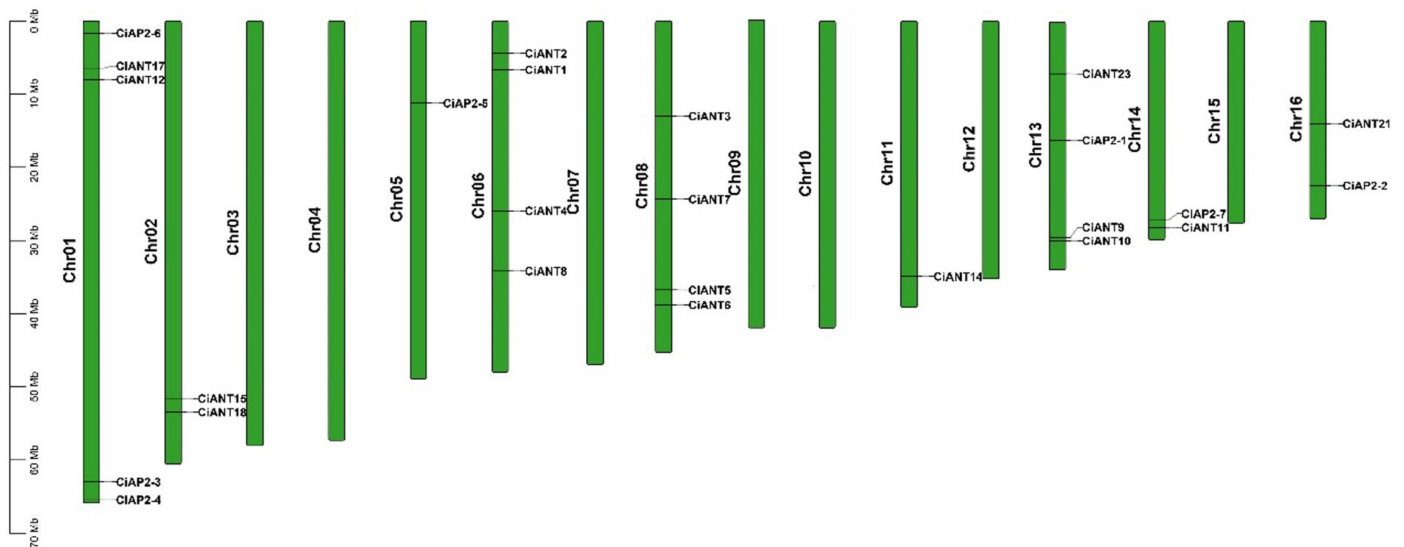

Figure 1. The distribution of the AP2 subfamily genes on pecan chromosomes. The position of each *CiAP2* was mapped according to its physical position on the pecan genome. The chromosome number is labeled at the right of each chromosome, and the scale is in mega bases (Mb).
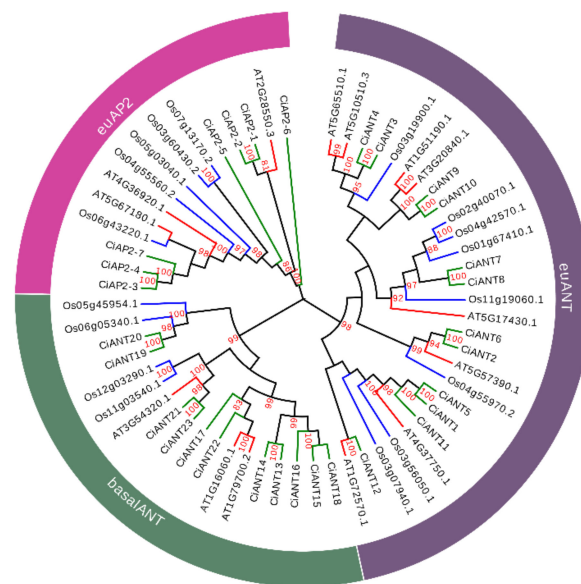

Figure 2. Phylogenetic analysis of AP2 subfamily proteins in the pecan, rice and *Arabidopsis*. A total of 61 AP2 proteins retrieved from pecan (30), rice (14) and *Arabidopsis* (17) were adopted to construct the phylogenetic tree. The tree was constructed using FastTree according to the maximum likelihood method. The AP2 subfamily proteins were clustered into three groups: euAP2, euANT and basal ANT. Proteins from pecan, rice and *Arabidopsis* are marked by green, blue and red colors, respectively.

2.3. Conserved Motif Analysis in AP2 Proteins

The motifs in 30 AP2 protein sequences were analyzed through MEME. The results displayed the first ten conserved motifs, and they varied among euAP2, basalANT and euANT evolution branches (Figure 3). In euAP2, motifs 1, 2, 3, 4, 5 and 6 were found almost in all members. Among them, motif 6 was located in the first AP2 domain and motifs 1

and 3 were parts of the first AP2 domain. Motifs 2 and 4 were located in the second AP2 domain and motif 5 was part of the second AP2 between two AP2. Compared with the euAP2 group, basalANT and euANT contained the unique motif 10 in the first AP2 domain. BasalANT contained the unique motifs 7 and 8 on the C terminal, while euANT contained the unique motif 9 on the N terminal. Overall, the whole AP2 domain is presented in the middle of the euAP2 group, while near the N terminal in the basalANT group and in a diverse position in the euANT group. Moreover, the gene structures were visualized to further reveal the variations between *CiAP2* genes. As shown in Figure 3, the genes in the euANT group exhibited highly similar gene structures with 8 or 9 exons. Although the exon number ranged from 8 to 11 in euAP2 and 5 to 14 in the basalANT group, most genes showed similar gene structures in their respective groups except for *CiAP2-2* and *CiANT19, 21, 23*.
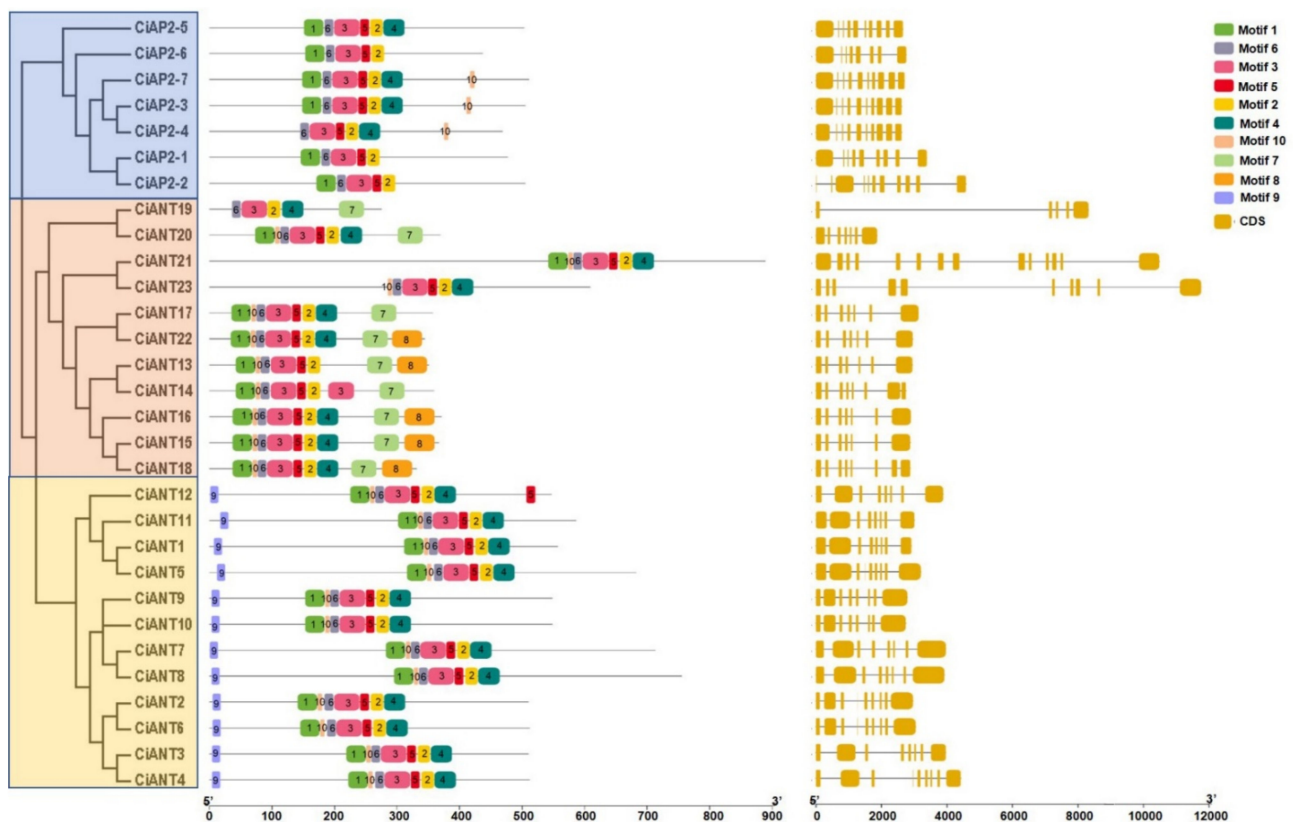

Figure 3. The motif distribution and gene structure of AP2 subfamily proteins in pecan. The middle panel represents the distribution of conserved motifs in CiAP2 proteins. The color blocks represent different motifs on proteins. The right panel represents the structure of CiAP2 genes. The green blocks represent exons.

2.4. cis-Element Analysis of AP2 Subfamily Genes

AP2 family genes have been demonstrated to contribute to the processes of plant growth and development as well as various environmental stimuli. Therefore, detecting the characteristic of the *cis*-element in the promoter will deepen our understanding of the expression pattern of pecan AP2 genes. We analyzed the promoter through the PlantCARE server and clustered all motifs into eight groups depending on the function annotation (Figure 4). Evidently, promotor and enhancer elements accounted for the largest proportion, with TATA-box and CAAT-box occurring the most frequently. Light responsive elements and plant hormone-responsive elements ranked as the second and third largest proportion, respectively. Box4, G-box and GT1-motif were the top three in light responsive elements. We further predicted the plant hormone-responsive elements and found that abscisic acid, salicylic acid, MeJA, gibberellin and auxin responsive elements were identified in their

promotors (Figure S2). Among them, the abscisic acid-responsive element ABRE and MeJA-responsiveness CGTCA-motif appeared the most frequently. We also found that abiotic stress responsive elements, such as the MYB binding site (MBS), were involved in drought inducibility, which implied *CiAP2* genes may participate in drought-stress response through the MYB transcription factor. Interestingly, we found tissue-specific elements were all *cis*-regulatory elements involved in seed-specific regulation, which was consistent with the function of AP2 genes participating in embryo development. Furthermore, we only identified one circadian-responsive and two-cell-cycle regulation elements in all AP2 subfamily genes. Overall, our results indicate that AP2 subfamily genes may participate in growth and development processes by responding to plant hormones, light and abiotic stress.
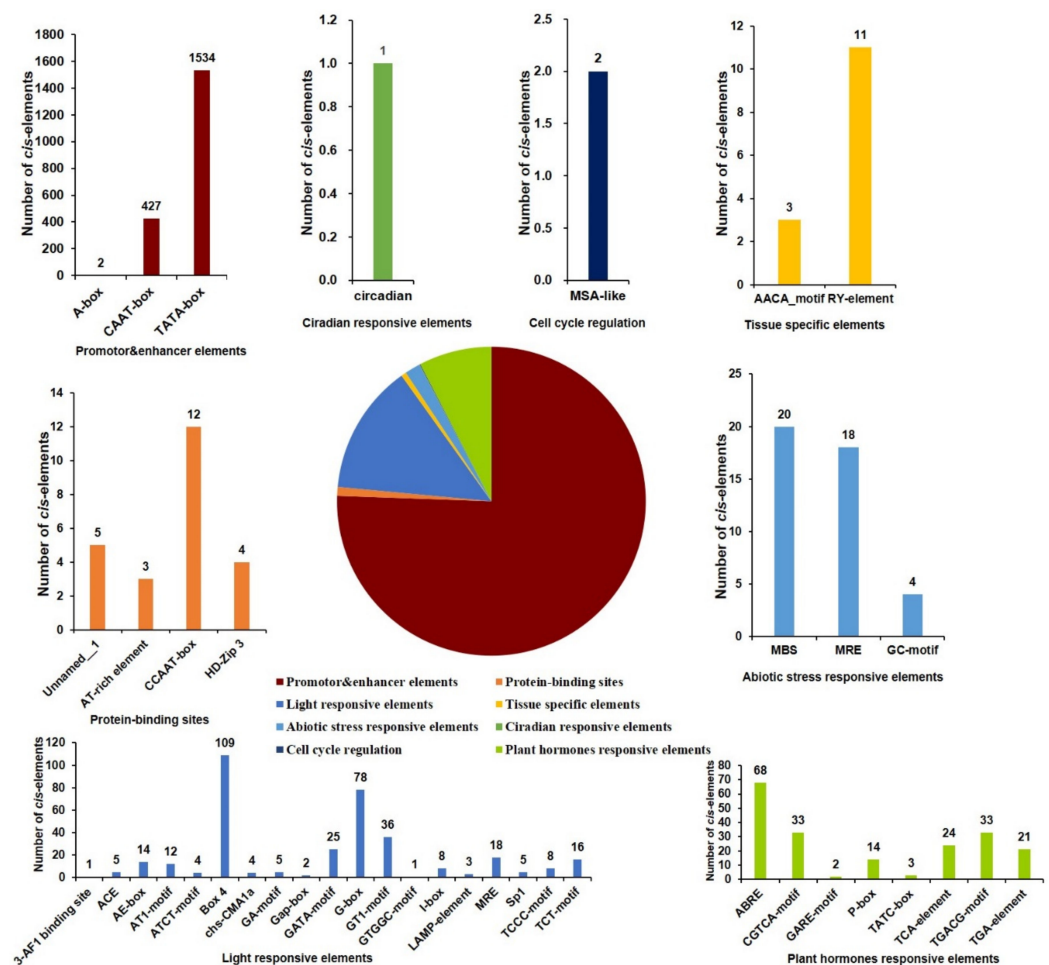

Figure 4. The analysis of *cis*-regulatory elements in the putative promoter of the AP2 subfamily genes. The pie chart shows the classification of *cis*-regulatory elements, and the corresponding color chart displays the frequency of different cis-elements.

2.5. Expression Analysis of CiAP2 Genes

To investigate the putative roles of pecan AP2 subfamily genes, a qPCR assay was employed to analyze their expression patterns in different tissues. Among the euAP2 clade, *CiAP2-5* had the highest relative expression level in leaves, stamens and pistils. The expression level of *CiAP2-7* was second in stamens and pistils. *CiAP2-1* and *CiAP2-2* were clustered together and presented consistent expression patterns (Figure 5a). Genes in basalANT clades, such as *CiANT15*, *16* and *18*, displayed similar expression patterns with relatively high expression levels in pistils, while *CiANT17*, *21*, *22* and *23* showed higher expression levels in stamens, especially *CiANT23* (Figure 5b). The expression pattern in the euANT clade varied greatly (Figure 5c,d). Although *CiANT1*, *5*, *11* and *12* were clustered

together, *CiANT1* and *5* expressed higher, while *CiANT11* and *12* showed almost no expression in the tested tissues. The expression levels of *CiANT2* and *CiANT6* were higher in different tissues, while *CiANT3* displayed almost no expression in the tested tissues and *CiANT4* showed trace expression in leaves. Previous studies have demonstrated that some AP2 family genes participate in seed development in *Arabidopsis*. To identify the potential genes that are involved in pecan fruit development, we selected genes with a relatively high expression in the pistil for further quantitative analysis in the developing pecan fruits. As the results show (Figure 5e), *CiAP2-5*, *CiANT1* and *2* shared similar expression patterns in developing fruits, and their expression levels decreased with fruit development. The expression of *CiANT6* was lowest in developing fruits. *CiANT5* and *CiANT23* displayed the highest expression levels in developing fruits; *CiANT5* expressed higher in the first three weeks, while *CiANT23* expressed higher in the last three weeks during pecan fruit development. The difference in the expression levels implied that *CiANT5* and *CiANT23* were likely to function diversely during pecan fruit development.
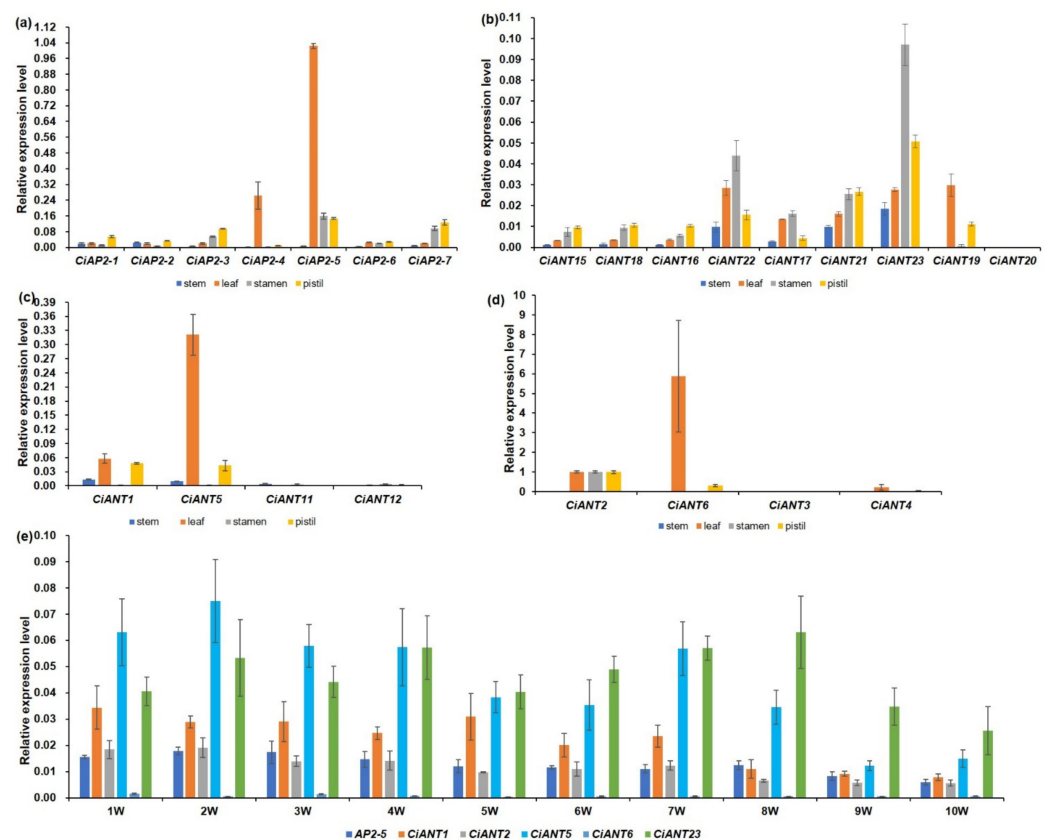

Figure 5. The expression analysis of pecan AP2 subfamily genes. (**a**–**d**) The expression of AP2 subfamily genes in pecan stems, leaves, stamens and pistils. (**e**) The expression of *CiAP2-5*, *CiANT1*, *CiANT5*, *CiANT6* and *CiANT23* in developing pecan fruits after pollened 1 to 10 weeks, respectively. Pecan *actin* gene (*OF23284*) was used as the internal reference gene. Error bars indicate SD. Each reaction represents three technical repeats.

2.6. CiANT5 Function as a Transcription Activator in Yeast

As mentioned above, *CiANT5* demonstrated a high expression level in pecan leaves, pistil and developing fruits. Therefore, we selected it for further analysis. To explore whether CiANT5 is located in the nucleus as expected of a transcription factor, we conducted a subcellular localization analysis first. We obtained the full-length CDS of CiANT5 without a stop code through PCR amplification and fused CiANT5 to the N terminal of GFP. Then, the CiANT5-GFP and nuclear maker Ghd7-mCherry were transiently co-expressed in tobacco leaves. The results demonstrated that the CiANT5-GFP fluorescence signals

localized in the nucleus overlapped perfectly with the red fluorescence of the nuclear maker Ghd7-mCherry, while free GFPs were detected throughout the nucleus and cytoplasm (Figure 6a,b). Subsequently, the transcriptional activation activity was analyzed in yeast cells. The full-length of CiANT5 as well as seven truncated fragments were ligated into the *pGBKT7* vector to generate BD-CiANT5(1-682aa), BD-N(1-324aa), BD-N1AP2(1-397aa), BD-ΔC(1-491aa), BD-C(492-682aa), BD-C2AP2(428-682aa), BD-ΔN(325-682aa) and BD-AP2(325-491aa) fusion constructions. These constructions as well as the empty *pGBKT7* were transformed into yeast *AH109* and cultured on selective medium SD/-Trp and SD/-Trp/-His/-Ade, respectively. As the result show, CiANT5 displayed transcriptional activation activity in yeast (Figure 6c). Among the six truncated fragments, except C2AP2, ΔN and AP2, others showed transcriptional activity. Furthermore, the results were confirmed by quantitatively measuring β-galactosidase (β-gal) activity. When the truncated fragments contained the AP2 domain, β-gal activity decreased and remained lower than the negative control in the BD-AP2 fragment. Taken together, these results demonstrate that CiANT5 is a transcription factor localized in the nucleus with transcriptional activity, while the AP2 domain has a negative effect on transcriptional activity.
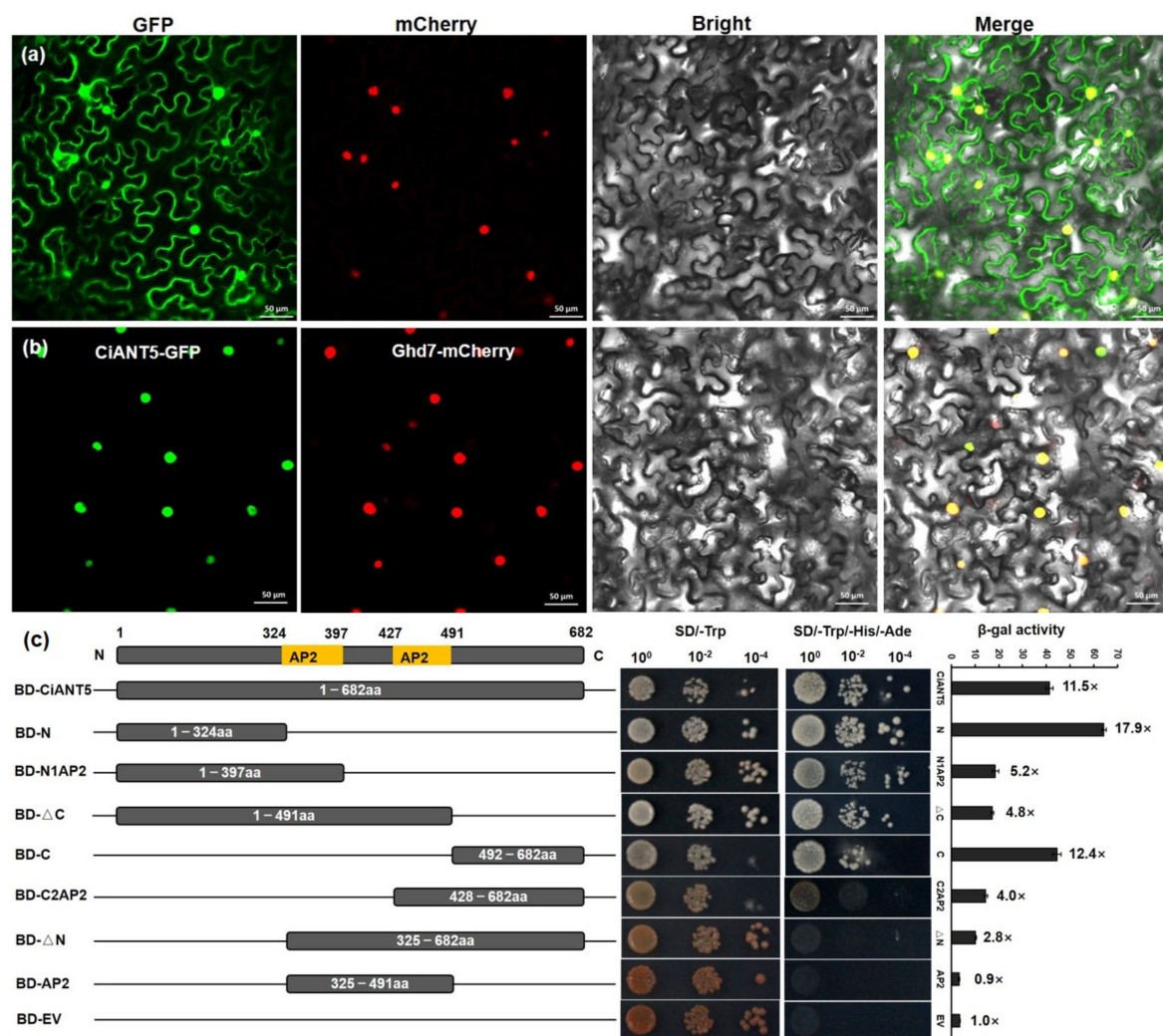

Figure 6. The subcellular and transcriptional analysis of CiANT5. (**a**) GFP fluorescence signal of GFP protein. (**b**) GFP fluorescence signal of the CiANT5-GFP fused protein. Ghd7-mCherry was used as nuclear maker. Bar = 50 μm. (**c**) The transcriptional activity analysis of CiANT5 in yeast *AH109* strain. The left diagram shows different constructions used for transcriptional assay. BD, GAL4-DNA binding domain; EV, empty *pGBKT7* vector used as a negative control. The middle panel displayed different constructions cultured on selective medium SD/-Trp and SD/-Trp/-His/-Ade. The right panel displays β-galactosidase (β-gal) activity in the various constructions. Mean values ± *SE* were calculated from three repeats.

3. Discussion

3.1. Characterization of Pecan AP2 Subfamily Genes

In this study, we initially identified 207 AP2/ERF family genes in the pecan, which were classified into three subfamilies, ERF, RAV and AP2, depending on the number of AP2 domains (Figure S1). Of them, both ERF and RAV members contain a single AP2 domain, while members in the AP2 subfamily contain two tandem AP2 domains. Further, we verified the sequences and finally obtained 30 AP2 members with two AP2 domains, which was higher than that in *Pinus massoniana* (7) [27], Pineapple (24) [28], *Dendrobium officinale* (14) [24], while less than that in *Elaeis guineensis* (34) [29] and *Saccharum spontaneum* (43) [30]. Pecan AP2 members were also grouped into three clades, euAP2, euANT and basalANT, which is consistent with previous research reported in angiosperms, such as wheat and *Dendrobium officinale* [4,22,24]. Furthermore, the analysis of genomic distribution displayed that most genes were distributed on chromosomes 1, 6, 8 and 13 (Figure 1). Unfortunately, five genes in the basalANT clade, namely *CiANT13*, *16*, *19*, *20* and *22*, failed to be anchored on corresponding chromosomes. Similarly, three genes were not mapped to any chromosomes in pineapple [28], perhaps due to the imprecise chromosome assembly.

Proteins in the same family usually share a common characterized domain; however, members may contain diverse domains and motifs contributing to a variety of functions. A motif analysis showed that the top ten conserved motifs in the pecan AP2 subfamily TFs, and motifs 1, 2, 3, 4, 5 and 6 were found in almost all members (Figure 3), while motifs 7 and 8 were specific to basalANT and motif 9 was specific to euANT. This difference may lead to a diversity of gene functions. As reported, *miR172* binding to specific motif sites regulates euAP2 TF expression through transcript cleavage and translational repression [31,32]. In *Arabidopsis*, AP2 family members, such as *AP2*, *TOE1*, *TOE2*, *TOE3*, *SMZ* and *SNZ*, were *miR172* target genes. AP2 participated in floral organ formation, while others mainly suppressed floral formation [31,33]. Moreover, the gene structure reflected gene function and evolution to some extent. Our results show that the euANT group exhibited a highly similar gene structure with 8 or 9 exons, and the basalANT group had more diversified structures with 5 to 14 exons, which is consistent with previous findings in wheat [23]. In short, our results implied that different conserved motifs and gene structures may result in a diversity of functions responding to evolutionary selection pressures.

3.2. Promoter cis-Acting Elements and Expression Analysis Implied the Functions of Pecan AP2 TFs

AP2/ERF, as a plant-specific TF family, has been proven to play important roles during plant growth, abiotic stress and hormone response [2,34]. As expected, we found many tissue-specific elements, abiotic stress-responsive elements and plant hormone-responsive elements in the promoters of AP2 family genes (Figure 4). The tissue-specific elements mainly are the RY-element and AACA_motif involving seed and endosperm-specific expression regulation. Previous studies have demonstrated that petunia AP2-like genes PhAp2A, PhAp2B and PhAp2C are strongly expressed in endosperm and function in flower and seed development [35]. Wang et al. identified 57 putative TFs during maize seed development by expressed sequence tag (EST) analysis and found that three AP2/ERF members, ZmAP2-*1*, ZmAP2-*2* and ZmAP2-*2*, played important regulating roles in maize seed development [36]. In *Brassica napus*, *BABY BOOM* (*BBM*) belongs to the AP2/ERF family, preferentially expressed in developing embryos and seeds, and participates in cell proliferation and morphogenesis during embryogenesis [37]. Our results demonstrate that *CiAP2-5*, *CiANT1*, *2*, *5* and *23* displayed relatively higher expression in pistils and developing pecan fruits (Figure 5); they may participate in pecan fruit development. Besides, abiotic stress-responsive elements are another common element in the AP2 gene promoters, which may be related to the AP2 members regulating drought, heat, salt and freezing. *Arabidopsis* DREB1A/CBF1, DREB1B/CBF2 and DREB1C/CBF3 and their orthologous genes *OsDREB1A* and *OsDREB1B* in rice were involved in a low-temperature and drought response [7,38–40]. Zeng et al. [24] also found that the expression

of *DoAP2* genes was changed to different degrees by being treated with cold, PEG and NaCl. Furthermore, we found that many elements respond to hormones in the promoters of pecan AP2 genes, of which ABA-responsive elements (ABRE) were the largest group (Figures 4 and S2). This implied that the expression of *CiAP2* genes may be induced by ABA and that they may participate in various stress regulatory networks. Interestingly, MeJA-responsive elements were the largest group in *DoAP2* promoters [24]. This difference may be related to the optimum growing environment required by different species.

3.3. CiANT5 as a Member of AP2 TFs Dispayed Transcription Activity

Transcription factors can bind to specific DNA sequences and activate or repress the transcription of corresponding genes in the nucleus. Nuclear localization is the basic feature of transcription factors. In this study, we selected a homologous gene of *Arabidopsis ANT*, *CiANT5*, to conduct a subcellular localization assay. As expected, CiANT5, in the manner of SNB [20] and DoAP2-1, DoAP2-2, DoAP2-6 and DoAP2-11 [24], was part of the AP2 family TFs located in the nucleus (Figure 6b). Furthermore, we analyzed the transcriptional activity and found ANT5 functioned as a transcriptional activator in yeast. A truncated fragment analysis revealed that the activating domain was located in BD-N(1-324aa) and BD-C(492-682aa) (Figure 6c). While the two tandem AP2 or respective domains showed no activating function, perhaps they acted as binding domains bound to specific DNA motifs. Similarly, APETALA2-like TF SNB presented transcriptional activity in yeast, but the activating domain was located in residues 278 to 324 [20]. The C-terminal region (107-318) of FZP is indispensable for transcriptional activation activity by LUC activity in rice protoplasts [19]. These results implied the activating domain may be located in a diverse position in different AP2 TFs. Moreover, some AP2 TFs displayed transcriptional inhibitory activity. *Arabidopsis* AP2 negatively regulated the homeotic gene *AGAMOUS*, which is involved in floral organ development [41]. DoAP2-1, DoAP2-2, DoAP2-6 and DoAP2-11 displayed transcriptional repression activity in tobacco leaves [24]. Whether transcriptional activation or repression, both are the precise regulation of TFs to its target genes. Thus, the related transcriptional characteristics of the remaining pecan AP2 TFs still need additional research.

4. Materials and Methods

4.1. Identification of AP2 Subfamily Genes in Pecan

The genome sequences of the pecan were derived from our re-sequenced data. The AP2 family sequences of *Arabidopsis* and rice were downloaded from Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html, accessed on 25 March 2021). The Hidden Markov Model of AP2 was downloaded from Pfam (http://pfam.xfam.org/, accessed on 27 March 2021), according to the Pfam id (PF00847). The HMM search from the HMMER program (v3.2.1) was employed to ascertain the presence of the AP2 domain in these three species. Furthermore, all candidate sequences were checked by Smart (http://smart.embl-heidelberg.de, accessed on 12 July 2021) to remove sequences without two AP2 domains and redundant sequences. The amino acid sequences of AP2 subfamily members are listed in Supplementary File S1.

4.2. Phylogenetic Analysis and Chromosome Localization of AP2 Subfamily Genes

The sequences were aligned using MAFFT (v7.271) [42], and the evolutionary tree was obtained by FastTree (v2.1.8) [43]. Then, the evolutionary tree was added colourful visualization plots through the online website evolview (www.evolgenius.info/evolview, accessed on 10 August 2021) [44]. The position of pecan AP2 subfamily genes on chromosomes were visualized by a localization tool in tbtools [45].

4.3. Gene Structure and Conserved Motif Analysis

According to the annotation files, the coding sequences and corresponding genomic sequences were extracted from the pecan genome. The conserved motif structure of AP2

subfamily genes was found by MEME (version 5.0.2) [46]. Then, all data were input into tbtools to visualize the gene structure and conserved motif.

4.4. cis-Element Analysis

A total of 2.0 Kb upstream sequences from initiation codon were retrieved as promoter sequences, and then the cis-regulatory elements were predicted through the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/, accessed on 22 August 2021). Furthermore, we clustered all motifs into different groups depending on the function annotation and counted the frequency distribution of promotor and enhancer elements, protein-binding sites, light responsive elements, tissue specific elements, abiotic stress responsive elements, circadian responsive elements, cell cycle regulation and plant hormone responsive elements.

4.5. Plant Material and Sample Collection

The pecan cultivar Western used as plant material was grown in Lin'an City (30° N, 119° E), Zhejiang Province, China. The leaves, stems, stamens and pistils were collected from Western in late April 2021, when the stigmas were acceptable. Moreover, we collected developing fruits one to ten weeks after controlling for pollination. All samples were immerged in liquid nitrogen immediately after collection and stored at −80 °C in a refrigerator for its use.

4.6. RNA Isolation and qPCR Analysis

The RNA was isolated from pecan samples (leave, stems, stamens and pistils, as well as developing fruits) with modified CTAB combining TRIzol reagent (Invitrogen, Grand Island, NY, USA), as we used previously in [47]. The cDNA was synthesized using the Prime-Script™ RT Reagent Kit (TaKaRa, Kyoto, Japan), according to the manufacturer's instructions.

qPCR was performed on a Bio-Rad CFX 92 Touch device using SYBR Green PCR Kit (TaKaRa). The procedure was as follows: 95 °C, 90 s; 40 cycles of 95 °C, 5 s; 55 °C, 20 s; 72 °C, 10 s. The relative expression level was calculated using the $2^{-\Delta\Delta Ct}$ method. The pecan actin gene (OF23284) was used as the internal reference gene. All primers used for the qPCR analysis are listed in supplementary Table S1.

4.7. Subcellular Localization Analysis

The full-length CDS of CiANT5 without a stop code was amplified and inserted into the vector pYBA1132 between the XbaI and HindIII sites to form the fusion construct CiANT5-GFP. Ghd7-mCherry was used as nuclear maker. Different combinations of fusion constructs were co-transformed in *N. benthamiana* leaves. The EGFP signals were detected with a laser scanning confocal microscope (LSM880, Karl Zeiss, Oberkochen, Germany) 48 h after transfection. The primers used in the analysis are listed in Supplementary Table S1.

4.8. Transcriptional Activity Analysis

The transcriptional activity assay was performed using the Matchmaker GAL4 Yeast Two-Hybrid System 3 (Clontech, CA, USA). The full-length CDS and seven CiANT5 cDNA fragments were amplified and fused with pGBKT7 plasmid linearized by EcoRI. Empty pGBKT7 and fused plasmids, which were marked as BD-ANT5, BD-N, BD-N1AP2, BD-ΔC, BD-C, BD-C2AP2, BD-ΔN and BD-AP2 were transformed into the yeast strain AH109. All transformants were grown in a tryptophan-negative dropout medium (SD/-Trp), then incubated in selective media plates (SD/-Trp/-His/-Ade) at 30 °C for four days to assay transcriptional activity. Furthermore, chlorophenol red-β-D-galactopyranoside (CPRG, Roche, Basel, Switzerland) was used as a substrate to measure β-galactosidase activity via liquid culture assays, as we did previously in [48]. All the primers used for the vector construction are listed in supplementary Table S1.

References

1. Nakano, T.; Suzuki, K.; Fujimura, T.; Shinshi, H. Genome-wide analysis of the ERF gene family in Arabidopsis and rice. *Plant Physiol.* **2006**, *140*, 411–432. [CrossRef]
2. Licausi, F.; Ohme-Takagi, M.; Perata, P. APETALA2/Ethylene Responsive Factor (AP2/ERF) transcription factors: Mediators of stress responses and developmental programs. *New Phytol.* **2013**, *199*, 639–649. [CrossRef]
3. Sakuma, Y.; Liu, Q.; Dubouzet, J.G.; Abe, H.; Shinozaki, K.; Yamaguchi-Shinozaki, K. DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. *Biochem. Biophys. Res. Commun.* **2002**, *290*, 998–1009. [CrossRef]
4. Kim, S.; Soltis, P.S.; Wall, K.; Soltis, D.E. Phylogeny and domain evolution in the APETALA2-like gene family. *Mol. Biol. Evol.* **2006**, *23*, 107–120. [CrossRef] [PubMed]
5. Swaminathan, K.; Peterson, K.; Jack, T. The plant B3 superfamily. *Trends Plant Sci.* **2008**, *13*, 647–655. [CrossRef] [PubMed]
6. Ohme-Takagi, M.; Shinshi, H. Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. *Plant Cell* **1995**, *7*, 173–182.
7. Stockinger, J.E.; Gilmour, S.J.; Thomashow, M.F. *Arabidopsis thaliana CBF1* encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 1035–1040. [CrossRef]
8. Dinh, T.T.; Girke, T.; Liu, X.; Yant, L.; Schmid, M.; Chen, X. The floral homeotic protein APETALA2 recognizes and acts through an AT-rich sequence element. *Development* **2012**, *139*, 1978–1986. [CrossRef] [PubMed]
9. Kagaya, Y.; Ohmiya, K.; Hattori, T. RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. *Nucleic Acids Res.* **1999**, *27*, 470–478. [CrossRef]
10. Feng, C.Z.; Chen, Y.; Wang, C.; Kong, Y.H.; Wu, W.H.; Chen, Y.F. *Arabidopsis* RAV1 transcription factor, phosphorylated by SnRK2 kinases, regulates the expression of *ABI3*, *ABI4*, and *ABI5* during seed germination and early seedling development. *Plant J.* **2014**, *80*, 654–668. [CrossRef]
11. Jofuku, K.D.; den Boer, B.G.; Van Montagu, M.; Okamuro, J.K. Control of *Arabidopsis* flower and seed development by the homeotic gene *APETALA2*. *Plant Cell* **1994**, *6*, 1211–1225.
12. Klucher, K.M.; Chow, H.; Reiser, L.; Fischer, R.L. The *AINTEGUMENTA* gene of *Arabidopsis* required for ovule and female gametophyte development is related to the floral homeotic gene *APETALA2*. *Plant Cell* **1996**, *8*, 137–153. [PubMed]
13. Mizukami, Y.; Fischer, R.L. Plant organ size control: *AINTEGUMENTA* regulates growth and cell numbers during organogenesis. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 942–947. [CrossRef] [PubMed]
14. Nole-Wilson, S.; Tranby, T.L.; Krizek, B.A. *AINTEGUMENTA-like* (*AIL*) genes are expressed in young tissues and may specify meristematic or division-competent states. *Plant Mol. Biol.* **2005**, *57*, 613–628. [CrossRef] [PubMed]
15. Krizek, B. *AINTEGUMENTA* and *AINTEGUMENTA-LIKE6* act redundantly to regulate Arabidopsis floral growth and patterning. *Plant Physiol.* **2009**, *150*, 1916–1929. [CrossRef]
16. Fu, F.F.; Xue, H.W. Coexpression analysis identifies *Rice Starch Regulator1*, a rice AP2/EREBP family transcription factor, as a novel rice starch biosynthesis regulator. *Plant Physiol.* **2010**, *154*, 927–938. [CrossRef] [PubMed]
17. Sun, B.; Zhan, X.D.; Lin, Z.C.; Wu, W.X.; Yu, P.; Zhang, Y.X.; Sun, L.P.; Cao, L.Y.; Cheng, S.H. Fine mapping and candidate gene analysis of *qHD5*, a novel major QTL with pleiotropism for yield-related traits in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* **2017**, *130*, 247–258. [CrossRef] [PubMed]
18. Komatsu, M.; Chujo, A.; Nagato, Y.; Shimamoto, K.; Kyozuka, J. *FRIZZY PANICLE* is required to prevent the formation of axillary meristems and to establish floral meristem identity in rice spikelets. *Development* **2003**, *130*, 3841–3850. [CrossRef] [PubMed]

19. Ren, D.; Hu, J.; Xu, Q.; Cui, Y.; Zhang, Y.; Zhou, T.; Rao, Y.; Xue, D.; Zeng, D.; Zhang, G.; et al. *FZP* determines grain size and sterile lemma fate in rice. *J. Exp. Bot.* **2018**, *69*, 4853–4866. [CrossRef]

20. Jiang, L.; Ma, X.; Zhao, S.; Tang, Y.; Liu, F.; Gu, P.; Fu, Y.; Zhu, Z.; Cai, H.; Sun, C.; et al. The APETALA2-Like Transcription Factor SUPERNUMERARY BRACT Controls Rice Seed Shattering and Seed Size. *Plant Cell* **2019**, *31*, 17–36. [CrossRef] [PubMed]

21. Yuste-Lisbona, F.J.; Fernandez-Lozano, A.; Pineda, B.; Bretones, S.; Ortiz-Atienza, A.; Garcia-Sogo, B.; Muller, N.A.; Angosto, T.; Capel, J.; Moreno, V.; et al. ENO regulates tomato fruit size through the floral meristem development network. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 8187–8195. [CrossRef]

22. Guo, B.; Wei, Y.; Xu, R.; Lin, S.; Luan, H.; Lv, C.; Zhang, X.; Song, X.; Xu, R. Genome-Wide Analysis of APETALA2/Ethylene-Responsive Factor (AP2/ERF) Gene Family in Barley (*Hordeum vulgare* L.). *PLoS ONE* **2016**, *11*, e0161322. [CrossRef] [PubMed]

23. Zhao, Y.; Ma, R.; Xu, D.; Bi, H.; Xia, Z.; Peng, H. Genome-Wide Identification and Analysis of the AP2 Transcription Factor Gene Family in Wheat (*Triticum aestivum* L.). *Front. Plant Sci.* **2019**, *10*, 1286. [CrossRef]

24. Zeng, D.; Teixeira da Silva, J.A.; Zhang, M.; Yu, Z.; Si, C.; Zhao, C.; Dai, G.; He, C.; Duan, J. Genome-Wide Identification and Analysis of the APETALA2 (AP2) Transcription Factor in *Dendrobium officinale*. *Int. J. Mol. Sci.* **2021**, *22*, 5221. [CrossRef]

25. Huang, R.; Huang, Y.; Sun, Z.; Huang, J.; Wang, Z. Transcriptome Analysis of Genes Involved in Lipid Biosynthesis in the Developing Embryo of Pecan (*Carya illinoinensis*). *J. Agric. Food Chem.* **2017**, *65*, 4223–4236. [CrossRef]

26. Nole-Wilson, S.; Krizek, B.A. DNA binding properties of the Arabidopsis floral development protein AINTEGUMENTA. *Nucleic Acids Res.* **2000**, *28*, 4076–4082. [CrossRef] [PubMed]

27. Zhu, P.; Chen, Y.; Zhang, J.; Wu, F.; Wang, X.; Pan, T.; Wei, Q.; Hao, Y.; Chen, X.; Jiang, C.; et al. Identification, classification, and characterization of AP2/ERF superfamily genes in Masson pine (*Pinus massoniana* Lamb.). *Sci. Rep.* **2021**, *11*, 5441. [CrossRef] [PubMed]

28. Zhang, H.; Pan, X.; Liu, S.; Lin, W.; Li, Y.; Zhang, X. Genome-wide analysis of AP2/ERF transcription factors in pineapple reveals functional divergence during flowering induction mediated by ethylene and floral organ development. *Genomics* **2021**, *113*, 474–489. [CrossRef] [PubMed]

29. Zhou, L.; Yarra, R. Genome-Wide Identification and Characterization of AP2/ERF Transcription Factor Family Genes in Oil Palm under Abiotic Stress Conditions. *Int. J. Mol. Sci.* **2021**, *22*, 2821. [CrossRef]

30. Li, P.; Chai, Z.; Lin, P.; Huang, C.; Huang, G.; Xu, L.; Deng, Z.; Zhang, M.; Zhang, Y.; Zhao, X. Genome-wide identification and expression analysis of AP2/ERF transcription factors in sugarcane (*Saccharum spontaneum* L.). *BMC Genom.* **2020**, *21*, 685. [CrossRef]

31. Chen, X. A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. *Science* **2004**, *303*, 2022–2025. [CrossRef] [PubMed]

32. Wang, T.; Ping, X.; Cao, Y.; Jian, H.; Gao, Y.; Wang, J.; Tan, Y.; Xu, X.; Lu, K.; Li, J.; et al. Genome-wide exploration and characterization of miR172/euAP2 genes in *Brassica napus* L. for likely role in flower organ development. *BMC Plant Biol.* **2019**, *19*, 336. [CrossRef] [PubMed]

33. Aukerman, J.M.; Sakai, H. Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. *Plant Cell* **2003**, *15*, 2730–2741. [CrossRef]

34. Xie, Z.; Nolan, T.M.; Jiang, H.; Yin, Y. AP2/ERF Transcription Factor Regulatory Networks in Hormone and Abiotic Stress Responses in Arabidopsis. *Front. Plant Sci.* **2019**, *10*, 228. [CrossRef]

35. Maes, T.; Van de Steene, N.; Zethof, J.; Karimi, M.; D'Hauw, M.; Mares, G. Petunia Ap2-like genes and their role in flower and seed development. *Plant Cell* **2001**, *13*, 229–244. [CrossRef] [PubMed]

36. Wang, G.; Wang, H.; Zhu, J.; Zhang, J.; Zhang, X.; Wang, F.; Tang, Y.; Mei, B.; Xu, Z.; Song, R. An expression analysis of 57 transcription factors derived from ESTs of developing seeds in Maize (*Zea mays*). *Plant Cell Rep.* **2010**, *29*, 545–559. [CrossRef]

37. Boutilier, K.; Offringa, R.; Sharma, V.K.; Kieft, H.; Ouellet, T.; Zhang, L.; Hattori, J.; Liu, C.M. Ectopic expression of *BABY BOOM* triggers a conversion from vegetative to embryonic growth. *Plant Cell* **2002**, *14*, 1737–1749. [CrossRef] [PubMed]

38. Liu, Q.; Kasuga, M.; Sakuma, Y.; Abe, H.; Miura, S.; Yamaguchi-Shinozaki, K.; Shinozaki, K. Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell* **1998**, *10*, 1391–1406. [CrossRef]

39. Gilmour, S.J.; Zarka, D.G.; Stockinger, E.J.; Salazar, M.P.; Houghton, J.M.; Thomashow, M.F. Low temperature regulation of the *Arabidopsis* CBF family of AP2 transcriptional activators as an early step in cold-induced COR gene expression. *Plant J.* **1998**, *16*, 433–442. [CrossRef] [PubMed]

40. Dubouzet, J.G.; Sakuma, Y.; Ito, Y.; Kasuga, M.; Dubouzet, E.G.; Miura, S.; Seki, M.; Shinozaki, K.; Yamaguchi-Shinozaki, K. *OsDREB* genes in rice, *Oryza sativa* L. encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. *Plant J.* **2003**, *33*, 751–763. [CrossRef] [PubMed]

41. Drews, N.G.; Bowman, J.L.; Meyerowitz, E.M. Negative regulation of the *Arabidopsis* homeotic gene *AGAMOUS* by the APETALA2 product. *Cell* **1991**, *65*, 991–1002. [CrossRef]

42. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef] [PubMed]

43. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree 2—Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE* **2010**, *5*, e9490. [CrossRef] [PubMed]

44. Subramanian, B.; Gao, S.; Lercher, M.J.; Hu, S.; Chen, W.H. Evolview v3: A webserver for visualization, annotation, and management of phylogenetic trees. *Nucleic Acids Res.* **2019**, *47*, W270–W275. [CrossRef] [PubMed]

45. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools—An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **2020**, *13*, 1194–1202. [CrossRef]

46. Bailey, T.L.; Boden, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res.* **2009**, *37*, W202–W208. [CrossRef] [PubMed]

47. Wang, Z.J.; Huang, J.Q.; Huang, Y.J.; Li, Z.; Zheng, B.S. Discovery and profiling of novel and conserved microRNAs during flower development in *Carya cathayensis* via deep sequencing. *Planta* **2012**, *236*, 613–621. [CrossRef]

48. Yang, Z.; Liu, L.; Sun, L.; Yu, P.; Zhang, P.; Abbas, A.; Xiang, X.; Wu, W.; Zhang, Y.; Cao, L.; et al. OsMS1 functions as a transcriptional activator to regulate programmed tapetum development and pollen exine formation in rice. *Plant Mol. Biol.* **2019**, *99*, 175–191. [CrossRef] [PubMed]