forests

MDPI

Article

Characterization of the Expansin Gene Promoters in *Populus trichocarpa*

Junkang Zhang †, Xiaoyu Li †, Lei Wang, Longfeng Gong, Mengtian Li and Jichen Xu *

State Key Laboratory of Tree Genetics and Breeding, College of Biological Sciences and Technology, Beijing Forestry University, Beijing 100083, China; wangleibfu@bjfu.edu.cn (L.W.); gonglongfeng@bjfu.edu.cn (L.G.)
* Correspondence: xujichen@bjfu.edu.cn
† These authors contributed equally to this work.

Abstract: The expansin genes are commonly expressed in plant cells, and the encoded proteins influence plant growth and stress resistance by loosening the structure and increasing the flexibility of the cell wall. The objective of this study was to characterize expansin gene promoters in *Populus trichocarpa* to clarify the regulatory mechanisms underlying gene expression and evolution. Sequence alignments revealed that the similarity among 36 poplar expansin genes was greater for the coding sequences than for the promoter sequences, which suggested these promoter sequences evolved asynchronously. The bases flanking the start codon exhibited a usage bias, with sites +3, +4, and +5 biased toward GC, whereas the other sites were biased toward AT. The flanking sites were significantly correlated with gene expression, especially sites −10 and −17, in which C and G are the bases positively associated with gene expression. A total of 435 regulatory elements (61 types) were identified on the promoters of the poplar expansin genes; *Skn-1* was the most common element in 23 promoters. Some expansin genes had more regulatory elements on their promoters (e.g., *PtrEXPA4*, *PtrEXPA3*, *PtrEXPB3*, and *PtrEXPB1*), whereas some others had less (e.g., *PtrEXLA2*, *PtrEXLB1*, and *PtrEXPA23*). Furthermore, 26 types of elements were involved in expansin gene expression, 25 of which positively affected expression in all analyzed samples. The exception was the endosperm expression-related element *Skn-1*, which negatively regulated expression in four tissues or treatments. Expression analysis showed that the expansin genes in *Populus trichocarpa* performed much differently under regular and abiotic stress conditions, which well matched the diversity of their promoter sequences. The results show that expansin genes play an important role in plant growth and development and stress resistance through expression adjustment.

Keywords: *Populus trichocarpa*; expansin; promoter; expression; *cis*-element

Citation: Zhang, J.; Li, X.; Wang, L.; Gong, L.; Li, M.; Xu, J. Characterization of the Expansin Gene Promoters in *Populus trichocarpa*. *Forests* **2024**, *15*, 1485. https://doi.org/10.3390/f15091485

Academic Editor: Tadeusz Malewski

Received: 8 July 2024
Revised: 31 July 2024
Accepted: 18 August 2024
Published: 24 August 2024

1. Introduction

A promoter is necessary for initiating gene transcription, and *cis*-elements within the promoter help regulate gene expression levels. Promoters can be classified into three types (i.e., constitutive, tissue-specific, and inducible) based on gene expression patterns. Constitutive promoters, which regulate gene expression levels in all tissues of an organism and under various environmental conditions, include the virus-derived CaMV 35S promoter [1] and the plant-derived ubiquitin promoter [2]. Some promoters contain tissue-specific regulatory elements that control gene expression only in specific cells. Examples include the seed-specific promoter of the small subunit ADP-glucose pyrophosphorylase gene in rice [3], the seed-specific promoter of *Pro-at5g54000* in *Arabidopsis* [4,5], the root-specific promoter *Os03g01700* promoter in rice [6], the root-specific *SlREO* promoter in tomato [7], and the root-specific promoter of the *FaRB7* gene in strawberry [8]. Inducible promoters regulate gene expression only in response to external stimuli such as light, temperature, and hormones. For instance, the promoter of the *AtNUDT5* gene in *Arabidopsis thaliana* is activated after infection by pathogens [9], and the promoters of *SD9-2* and *SD18-1* were significantly promoted under water-deficit conditions [10–12].

Promoter sequence variations may significantly influence gene expression. A previous study examining the apple β-tubulin gene revealed it was expressed at a significantly lower level in the '*Wuzi*' variety than in the '*Yanfu3*' variety. A gene sequence alignment indicated the coding regions were identical in both varieties, but six bases differed in the promoters [13]. Specifically, the promoter region contains a series of regulatory elements, including the *TATA-box*, *TFIIB*-recognition element, initiator element, and motif ten elements [14], which ensure that transcription is initiated precisely. Some *cis*-elements within promoters bind to transcription factors, which further influence gene transcription. These *cis*-elements include abscisic acid response elements (ABREs), ethylene response elements, jasmonate response elements, low-temperature response elements, and drought response elements. In *Ziziphus jujube*, MYB41 specifically binds to the upstream site CAACCA of the *PAL* gene and induces its expression, whereas bHLH engages with the upstream site CACGTG of the *DFR* gene and induces its expression. Both *DFR* and *PAL* thereby enhanced the accumulation of flavonoids in jujube [15]. The promoter sequence of the *PgFAR40* gene in *Panax ginseng* harbors 10 MeJA *cis*-acting elements. When subjected to 200 mM methyl jasmonate treatment, *PgFAR40* was significantly up-regulated [16]. Hendelman reported four chromatin opening regions in the *wox9* promoter in tomato (*Solanum lycopersicum*) using the ATAC-seq technique. Mutation in *pro-Reg1* adjacent to the start codon resulted in abnormal vegetative growth phenotypes, such as embryonic death, and stopped the growth of the apical meristem. Mutations in *pro-Reg2* have abnormal reproductive growth phenotypes, such as excessive branching of inflorescence development. It was suggested that the tomato *wox9* gene had pleiotropic function regulated by *cis*-elements in the gene promoter region [17]. Therefore, analyses of promoter sequences will be useful for characterizing gene expression patterns and mechanisms underlying plant responses to various environmental conditions.

Expansin could promote cell wall stretching and loosening by disrupting non-covalent linkages between the cellulose and hemicellulose in cell walls. Thus, expansin is important for plant growth and development, as well as resistance to abiotic and biotic stresses [18]. For example, overexpression of the *Salix matsudana* expansin gene *SmEXPA13* enhances plant salt tolerance [19]. The *PtEXLA1* transgenic tobacco plants had a larger corolla and strong resistance to drought, high temperature, and salt stress in comparison to the wild-type [20]. A total of 36 expansin genes have been identified in the poplar genome [21]. A previous study revealed that *PtrEXLA2*, *PtrEXPA23*, and *PtrEXPA6* were specifically expressed in mature stem segments as well as in the primary stem growth region and during the mid-stem growth stage [22]. *PtEXPA2*, *PtEXPA4*, *PtEXPA7*, and *PtEXPA11* exhibited similar expression patterns, being predominantly expressed in buds and seeds of poplar. *PtEXLB3* is primarily highly expressed in seeds, with significantly lower expression levels observed in other tissues. Notably, *PtEXPA22* and *PtEXPA25* displayed no expression across various tissues [23]. *PtrEXPA3* and *PtrEXPA4* genes were highly expressed in poplar buds and leaves during periods of active growth but were not expressed during the dormant period [24]. *PtrEXPA3*, *PtrEXPA4*, and *PtrEXPA14* were expressed in female flowers but not in male flowers. *PttEXP1* and *PttEXP5* were differentially expressed in the stem segments of normal wood and tension wood [25]. Therefore, the expression of poplar expansin genes is obviously regulated by various mechanisms depending on the tissues and environmental conditions due to the diversity of promoters. The objective of this study was to characterize the poplar expansin gene promoters regarding their sequences, regulatory elements, and evolution to understand their roles or status in plant cells and provide the basis for manipulating expansin gene expression patterns.

2. Materials and Methods

2.1. Plant Materials and Treatments

The tissue culture plantlets of *Populus tomentosa* 84K were transplanted into a 10 × 10 cm pot with vermiculite as the substrate and grew at 25 °C with a 16 h light/8 h dark cycle. The plants were irrigated with water every four days and with Hoagland nutrient solution

once a week. Three-month-old plants were treated with 200 mM NaCl for 3 d, 4 °C for 2 d, 42 °C for 3 d, and water deficiency for 7 d, respectively. The plants under regular growth conditions were used as a control for the next analysis.

2.2. Expression Analysis of Expansin Genes

Total RNA was extracted from the leaf samples using the RNAprep Pure Polysaccharide Polyphenol Plant Total RNA Extraction Kit (Beijing TIANGEN, Beijing, China, DP441). The cDNA was synthesized using the FastKing cDNA First Strand Synthesis Kit (Beijing TIANGEN, Beijing, China, KR116). Based on the poplar genome sequence (http://www.phytozome.net/poplar, accessed on 29 May 2023), the specific primers were designed for each expansin gene (Table S1). The expression test was conducted by RT-PCR with 35 cycles at 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min. *UBQ* was the reference gene. The absolute expression amount of each gene was determined via Image J version 1.54 software, and the expression values under various stresses were standardized using the control expression value of 1.

2.3. Expansin Gene and Promoter Sequences

The sequences of the poplar expansin gene and promoters were obtained from the poplar genome database (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Ptrichocarpa, version 13, accessed on 29 May 2023) and the extended protein research website (http://www.Personal.Psu.Edu/fsl/ExpCentral/index.htm, accessed on 29 May 2023). Based on our previous study, the promoter activity of the expansin genes was 1000 bp [26]. The upstream 1000 bp from the start codon of the gene was then selected for promoter element analysis.

2.4. Phylogenetic Analysis

The phylogenetic tree was constructed based on the nucleotide sequences of the coding or promoter region (1000 bp) of the poplar expansin genes, respectively, by using maximum likelihood and Bayesian outperformed neighbor-joining and maximum parsimony [27]. The multiple sequence alignment was conducted through the default parameters of sequence weighting, position-specific gap penalties, and weight matrix choice by ClustalW, MEGA7 [28]. Both trees with the highest log likelihood were shown. The best substitution matrix for the trees was drawn in General Time Reversible (GTR) mode. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) was shown next to the branches [29]. The initial trees for the heuristic search were obtained by applying neighbor-join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach and then selecting the topology with a superior log likelihood value. A discrete Gamma distribution was used to model phylogenetic rate differences among sites.

2.5. Sequence Characterization of the Expansin Gene Promoters

The upstream 20 bases and the downstream 3 bases flanking the start codon ATG were counted, and the occurrence probability of four bases at each position was calculated. *TATA-box*, *CAAT-box*, and other regulatory elements in the upstream 1000 bp of the gene transcription initiation site (TSS) were determined using the online software TSSP-TCM [30] and the plant *cis*-acting element database PlantCARE.

2.6. Association Analysis between Expansin Gene Expression and Promoter Elements

The expression data of the poplar expansin genes in different tissues, organs, or treatments were obtained from the phytozome website (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Ptrichocarpa_er, version 13, accessed on 29 May 2023). Association analysis between the gene expression level and the base usage frequency, or the number of regulatory elements in the promoters, was conducted based on correlation

coefficient calculation by SPSS 20.0 software. The significant associated base/site and cis-element were determined at a *p*-value of 0.05.

3. Results

3.1. Expression Pattern of the Expansin Genes in Populus tomentosa under Stress Treatments

The poplar expansin family contains 36 members, which were classified into four subfamilies: *PtrEXPA* (27), *PtrEXPB* (3), *PtrEXLA* (2), and *PtrEXLB* (4). RT-PCR results showed that 36 expansin genes had expression patterns that varied widely under different stress conditions. For example, *PtrEXPALA2*, *PtrEXPLB4*, *PtrEXPA7*, and *PtrEXPA12* were expressed highly in all samples, but *PtrEXPA18* was always expressed weakly. *PtrEXPA15* and *PtrEXPA22* were up-regulated under all four stress conditions. Some genes were specifically associated with one or several stress conditions, such as *PtrEXPA14*, *PtrEXPA15,* and *PtrEXPA3,* which were more sensitive to salt treatment; *PtrEXPA20* was sensitive to heat stress; and *PtrEXPA24* was sensitive to cold stress. Several genes were insensitive to stress, such as *PtrEXPLB4*, *PtrEXPA7*, *PtrEXA12*, *PtrEXPB3*, *PtrEXPA27*, *PtrEXPALA1*, and *PtrEXPA17*, which were always expressed stably with or without stress treatment (Figure 1). Therefore, each expansin gene was regulated in a specific way according to the specific components of its promoter.

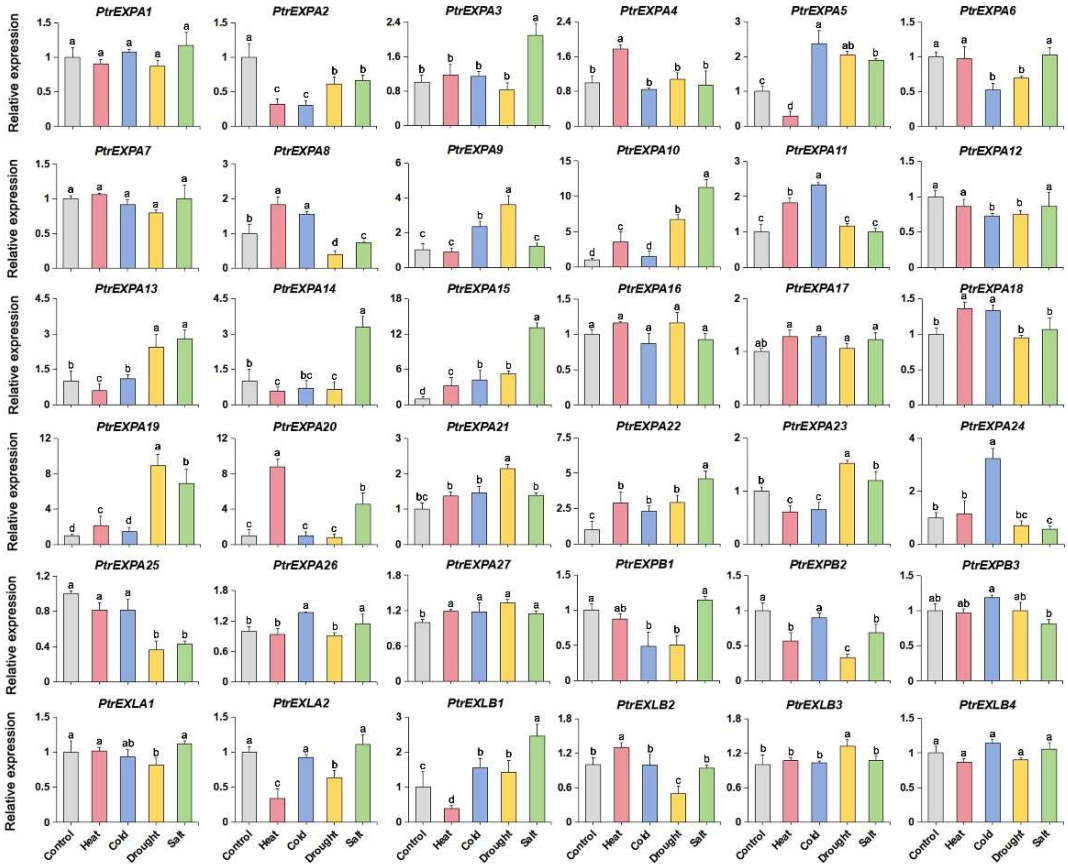

Figure 1. Relative expression of the expansin genes in *Populus trichocarpa* under different stress treatments. Values represent the mean of three replicates ± SD. Different letters denote statistically significant differences resulting from Duncan's test following a one-way ANOVA.

3.2. Base Usage Flanking the Start Codon

We detected obvious biases in the base usage, flanking the start codon (−20 to +6). Specifically, more than 75% of the genes had an A at site −1, 61.11% of the genes had a T at site −12, 66.67% of the genes had a G at site +4, and 58.33% of the genes had a C at site +5 (Table 1). The start codon of all expansin genes was ATG.

Table 1. Usage ratio of the bases flanking start codon ATG.

Location	A	C	G	T
−20	36.11%	16.67%	11.11%	36.11%
−19	27.78%	16.67%	25.00%	30.56%
−18	30.56%	2.78%	16.67%	50.00%
−17	44.44%	5.56%	16.67%	33.33%
−16	38.89%	22.22%	16.67%	22.22%
−15	33.33%	16.67%	8.33%	41.67%
−14	19.44%	33.33%	19.44%	27.78%
−13	27.78%	22.22%	22.22%	27.78%
−12	16.67%	22.22%	0.00%	61.11%
−11	25.00%	25.00%	16.67%	33.33%
−10	27.78%	22.22%	22.22%	27.78%
−9	25.00%	5.56%	25.00%	44.44%
−8	25.00%	19.44%	27.78%	27.78%
−7	44.44%	19.44%	22.22%	13.89%
−6	33.33%	22.22%	16.67%	27.78%
−5	50.00%	13.89%	16.67%	19.44%
−4	25.00%	25.00%	27.78%	22.22%
−3	50.00%	2.78%	33.33%	13.89%
−2	38.89%	25.00%	5.56%	30.56%
−1	75.00%	11.11%	8.33%	5.56%
+1	100.00%	0.00%	0.00%	0.00%
+2	0.00%	0.00%	0.00%	100.00%
+3	0.00%	0.00%	100.00%	0.00%
+4	16.67%	13.89%	66.67%	2.78%
+5	11.11%	58.33%	16.67%	13.89%
+6	44.44%	5.56%	8.33%	41.67%

The frequency of AT and GC at different sites flanking the start codon was calculated. Most sites tended to have AT. The exceptions were sites +3, +4, and +5, which were biased toward GC (Figure 2A). Non-coding sites near the start codon (−1 to −8) tended to comprise purines (A/G), whereas those far from the start codon (−11 to −15) were mainly pyrimidines (T/C), especially at site −12 (Figure 2B).

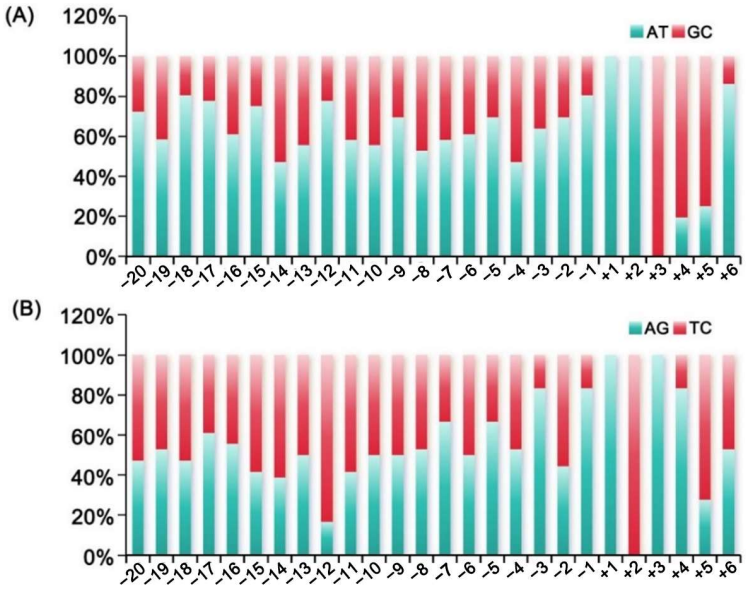

Figure 2. Four bases frequency in the flanking sites of the initial codon ATG (−20~+6). (**A**). frequency of AT and GC; (**B**). frequency of AG and TC. The *X*-axis means flanking sites of the initial codon. The *Y*-axis means the percentage of the detected bases at each site.

3.3. Base Usage in the Conserved Elements of Promoters

The TSSP-TCM online program was used to predict the transcription start site (TSS) in the promoters. The frequency of the bases A, C, T, and G at the TSS was 58.33%, 30.56%, 8.33%, and 2.78%, respectively, indicating there was an obvious base usage bias. The TSS was separated from the start codon of genes [5′ untranslated region (UTR)] by 55–888 bp (Table 2), with an average of 243 bp.

Table 2. TSS and 5′ UTR length in the poplar expansin gene promoters.

Gene	Gene #	TSS	5′ UTR Length (bp)	Gene	Gene #	TSS	5′ UTR Length (bp)
PtrEXLA1	Potri.004G181700	C	77	*PtrEXPA13*	Potri.016G135200	A	85
PtrEXLA2	Potri.007G083400	C	88	*PtrEXPA14*	Potri.006G108000	C	120
PtrEXLB1	Potri.003G083200	C	120	*PtrEXPA15*	Potri.010G202500	C	65
PtrEXLB2	Potri.001G151500	A	76	*PtrEXPA16*	Potri.006G086100	C	61
PtrEXLB3	Potri.001G147200	A	56	*PtrEXPA17*	Potri.002G017900	A	659
PtrEXLB4	Potri.003G087000	A	55	*PtrEXPA18*	Potri.005G244100	A	58
PtrEXPA1	Potri.001G240900	C	368	*PtrEXPA19*	Potri.002G184700	T	888
PtrEXPA2	Potri.013G154700	A	91	*PtrEXPA20*	Potri.001G401700	T	633
PtrEXPA3	Potri.010G167200	A	168	*PtrEXPA21*	Potri.009G169500	A	356
PtrEXPA4	Potri.008G088300	A	117	*PtrEXPA22*	Potri.004G208300	T	713
PtrEXPA5	Potri.009G031800	C	596	*PtrEXPA23*	Potri.017G140000	A	195
PtrEXPA6	Potri.004G123200	C	199	*PtrEXPA24*	Potri.001G112866	A	320
PtrEXPA7	Potri.008G057100	A	595	*PtrEXPA25*	Potri.016G097500	A	90
PtrEXPA8	Potri.004G080200	A	573	*PtrEXPA26*	Potri.016G098700	G	96
PtrEXPA9	Potri.017G085300	C	185	*PtrEXPA27*	Potri.017G092700	A	88
PtrEXPA10	Potri.001G001100	A	230	*PtrEXPB1*	Potri.014G066300	A	68
PtrEXPA11	Potri.013G060800	C	162	*PtrEXPB2*	Potri.019G101900	A	207
PtrEXPA12	Potri.019G057500	A	72	*PtrEXPB3*	Potri.013G134300	A	205

Except for *PtrEXPA21*, all the other poplar expansin genes were predicted to have a *TATA box* in the −35 to −10 region of their promoters. The bases in this conserved motif were determined as $C_{0.64}T_{0.86}A_{0.94}T_{0.94}A_{0.92}A_{0.69}A_{0.94}T_{0.72}A_{0.81}C_{0.61}$ (the subscript values represent the frequency of the corresponding base). In contrast, the *CAAT-box* was detected in the −200 to −38 bp region, except in *PtrEXPLA1*, *PtrEXPLB4*, *PtrEXPA1*, *PtrEXPA3*, *PtrEXPA5*, *PtrEXPA11*, *PtrEXPA12*, *PtrEXPA13*, *PtrEXPA14*, *PtrEXPA16*, *PtrEXPA19*, and *PtrEXPA26*. The base usage patterns in the *CAAT-box* were detected as $T_{0.50}A_{0.50}T_{0.50}C_{0.50}C_{1.00}A_{1.00}A_{1.00}A_{1.00}T_{1.00}T_{0.50}T_{0.50}$ for *PtrEXPA2*, *PtrEXPA9*, *PtrEXPA17*, and *PtrEXPA21*, and $A_{0.50}A_{0.35}(T_{0.35})G_{0.40}A_{0.30}(C_{0.30})(T_{0.30})C_{1.00}A_{1.00}A_{1.00}T_{1.00}A_{0.40}A_{0.50}$ for the other genes.

3.4. Sequence Similarity of the Promoters

A sequence alignment analysis revealed that the nucleotide similarity was 23.20%–55.00% (an average of 30.78%) for the promoters in the same gene subfamily, whereas it was 24.50%–39.20% (an average of 31.14%) for the promoters belonging to different subfamilies (Table 3). Comparably, the sequence similarity of the coding regions was 47.60%–94.30% (an average of 60.20%) in the same gene subfamily but was 33.10%–53.40% (an average of 39.46%) in different gene subfamilies. Clearly, the coding regions of the expansin genes showed significant sequence similarity in the same subfamily but not in a different subfamily and exhibited certain subfamily characteristics. The promoters performed similarly in the same subfamily and in different subfamilies and exhibited nonsubfamily characteristics. The promoter sequences may have changed more than the coding sequences during evolution.

Table 3. Sequence similarity of the promoters (up) and coding sequences (down) of the poplar expansin genes (%) (the average value is in brackets).

Subfamily	*EXPA*	*EXPB*	*EXLA*	*EXLB*
EXPA	23.20–39.00(30.87) 47.60–94.30(59.05)			
EXPB	26.00–38.50(32.47) 35.60–43.20(39.92)	32.80–42.60(36.47) 48.40–92.20(63.10)		
EXLA	26.10–39.20(31.90) 35.80–42.60(38.42)	32.10–34.90(33.03) 45.00–46.80(45.82)	55.00(55.00) 89.50(89.50)	
EXLB	24.50–36.00(30.20) 33.10–42.60(37.90)	29.10–33.60(31.58) 41.40–46.50(44.01)	29.40–33.50(31.00) 49.80–53.40(51.31)	27.50–35.10(31.82) 52.40–89.70(68.93)

Some genes seemed to have identical phylogenetic tendencies in both the promoter and coding regions. For example, the sequence similarity between *PtrEXLA1* and *PtrEXLA2* was high both in the promoter sequence (55.00%) and in the coding sequence (89.50%), the same as the sequence similarity between *PtrEXPB2* and *PtrEXPB3* (42.60% in the promoter sequence and 92.20% in the coding sequence). The sequence similarity between *PtrEXPA5* and *PtrEXPA8* was low both in the promoter sequence (31.90%) and in the coding sequence (48.00%). However, most of the other analyzed genes performed differently in the promoters and coding sequences. For example, the similarity of the coding sequence between *PtrEXPA25* and *PtrEXPA26* was relatively high (94.30%) in contrast to the sequence similarity of the promoters (30.40%). These results suggested that the evolution of the poplar expansin genes varied.

A phylogenetic tree constructed based on the expansin gene coding sequences revealed that all members of the same subfamily clustered together (Figure 3, right). However, the genetic relationships between the promoter sequences sometimes differed from those between the coding sequences. Thus, the phylogenetic tree based on the promoter sequences (Figure 3, left) varied slightly from that constructed based on the coding sequences. For example, the distance between the promoter sequences of *PtrEXPA8* and *PtrEXPLB2* was much closer in contrast to their coding sequences. The distance between the coding sequences of *PtrEXPA3* and *PtrEXPA4* was much closer, unlike their promoter sequences. Meanwhile, some genes were genetically closely related in both the promoter and coding sequences, including the gene pairs of *PtrEXPA7/PtrEXPA15*, *PtrEXPA25/PtrEXPA26*, *PtrEXPA4/PtrEXPA9*, *PtrEXPLA1/PtrEXPLA2*, and *PtrEXPB1/PtrEXPB3*.

3.5. Regulatory Elements in Poplar Expansin Gene Promoters

A total of 435 regulatory elements were identified in the poplar expansin gene promoters based on the PlantCARE database. These regulatory elements were classified into 61 types according to their functions (e.g., related to stress responses, hormones, or light) and their binding site specificity (Figure 4A). Some of these elements occurred in several promoters or repeatedly in one promoter. For example, the endosperm expression-related element *Skn-1* motif was present in 23 promoters, with 10 promoters comprising two copies. The anaerobic responsive element was present in 16 promoters, with the *PtrEXPA3*, *PtrEXPB3*, and *PtrEXPA20* promoters consisting of two, two, and three copies, respectively.

Each poplar expansin gene promoter contained an average of 12 regulatory elements. The *PtrEXPA4*, *PtrEXPA3*, *PtrEXPB3*, and *PtrEXPB1* promoters had the most, with 20, 19, 19, and 18 elements, respectively. In contrast, *PtrEXLA2*, *PtrEXLB1*, and *PtrEXPA23* had only 6 regulatory elements each. Additionally, the examined promoter had an average of 10 types of regulatory elements. The *PtrEXPA3* promoter contained a maximum of 17 types of regulatory elements, including the antioxidant response element (2), photoresponsive element (e.g., Box 4) (2), heat shock element (1), ABRE (1), jasmonic acid response element (TGACG-motif) (1), and other elements (10). It implied that *PtrEXPA3* is involved in responses to various environmental stimuli (Figure 4C). Regarding the distribution of the

regulatory elements within promoters, 100-bp fragments underwent a sequential interception analysis. The data indicated the regulatory elements were less enriched adjacent to the gene coding region and were less abundant further away from the coding region. The −800–−700 bp segment contained the maximum of 60 regulatory elements, whereas the −200–−100 bp segment contained the fewest elements, with only 25 (Figure 4B).

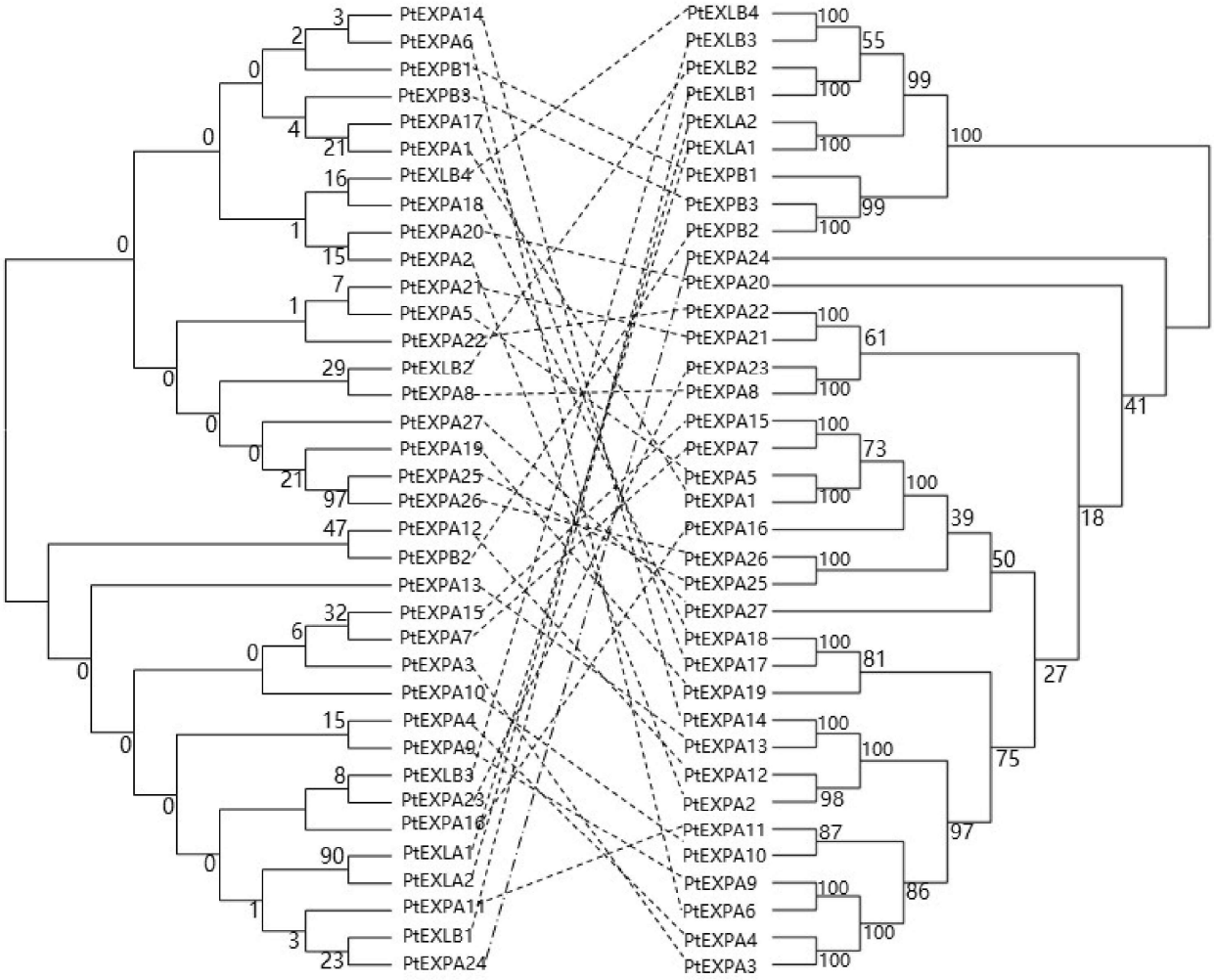

Figure 3. Phylogenetic trees based on the poplar expansin gene promoter sequences (left) and coding sequences (right) by the neighbor-joining algorithm, MEGA7. The Bootstrap test was used to verify phylogenetic tree reliability (replications = 1000).

3.6. Promoter Components in Response to Plant Tissues and Growth Conditions

The bases flanking the start codon and the regulatory elements in the promoter were analyzed in terms of their relationships to the gene expression level in tissues or in response to stress treatments. The results indicated that site −10 was significantly associated with gene expression in 16 samples, with C representing the most common base at this position (positive correlation in all 16 samples) and with G, T, and A being the non-associated bases. Site −17 was significantly associated with gene expression in 11 samples, with T (negative correlation in 5 samples) and G (positive correlation in 6 samples) being the most frequent bases, and with C (positive correlation in 1 sample) and A (non-correlation) being the less associated bases. In contrast, sites −13 and −9 were not related to gene expression in any sample. Moreover, gene expression levels in some samples were related to multiple sites. For example, gene expression levels in the *BESC423.ZL7* female early, *GW9911.ZK51* male mid, and predormant bud I samples were related to the base preferences at 10, 9, and 7 sites, respectively (Figure 5).

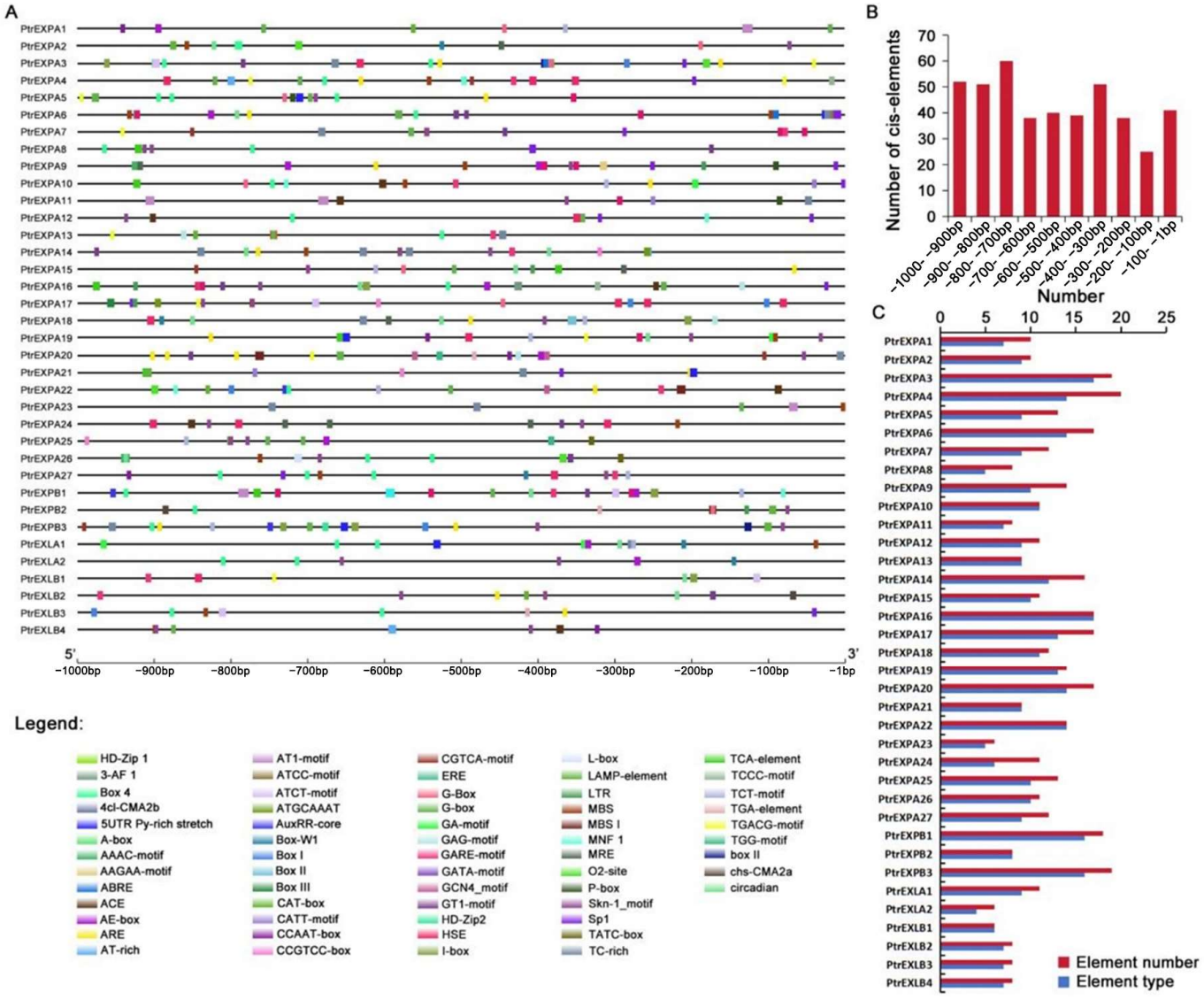

Figure 4. Distribution of the regulatory elements in 100 bp sequential interception segment along the poplar expansin gene promoters. The *X*-axis means the segments of the promoters started from the initial codon. The *Y*-axis means the number of regulatory elements in the relative segment. (**A**) Distribution of the *cis*-elements on the promoters; (**B**) distribution of the *cis*-elements on the promoter regions; (**C**) comparison of the *cis*-elements among the expansin genes. *PtrEXPA14*, *PtrEXPA15,* and *PtrEXPA3*.

A total of 26 types of elements were associated with gene expression, 25 of which were positively correlated with gene expression. The exception was the endosperm expression-related element *Skn-1* motif, which was negatively correlated with gene expression in four samples. Some elements (e.g., *ABRE*, *3-AF1* binding site, *G-box*, and *TCCC-motif*) were related to gene expression in various tissues, organs, or stress responses. The gene expression levels in a sample were related to several types of elements. For example, six types of regulatory elements influenced the expansin gene expression level in fully open bud samples, whereas seven types of regulatory elements affected expansin gene expression in standard stem nodes (Figure 6).

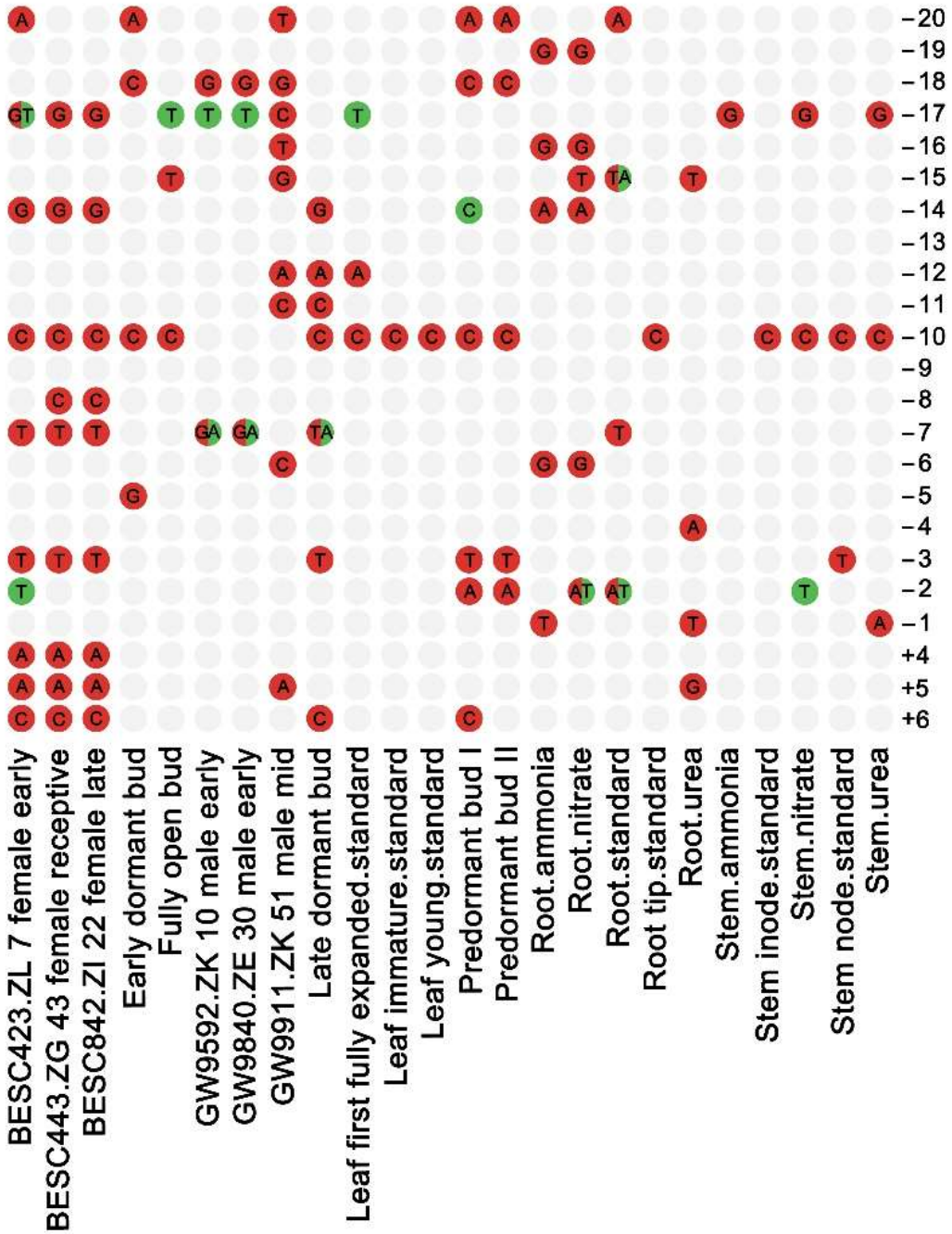

Figure 5. The gene expression in response to the start codon flanking sites. The circles represent the significant correlation (p = 0.05) between the expression level of the poplar expansin genes and the base usage frequency of the site on the promotor based on THE Pearson correlation coefficient calculation by SPSS 20.0. The relevant base was marked inside the circle. The red circle indicates a positive correlation between the site and the gene expression of the poplar expansin genes in the tissue, while the green circle indicates a negative correlation.

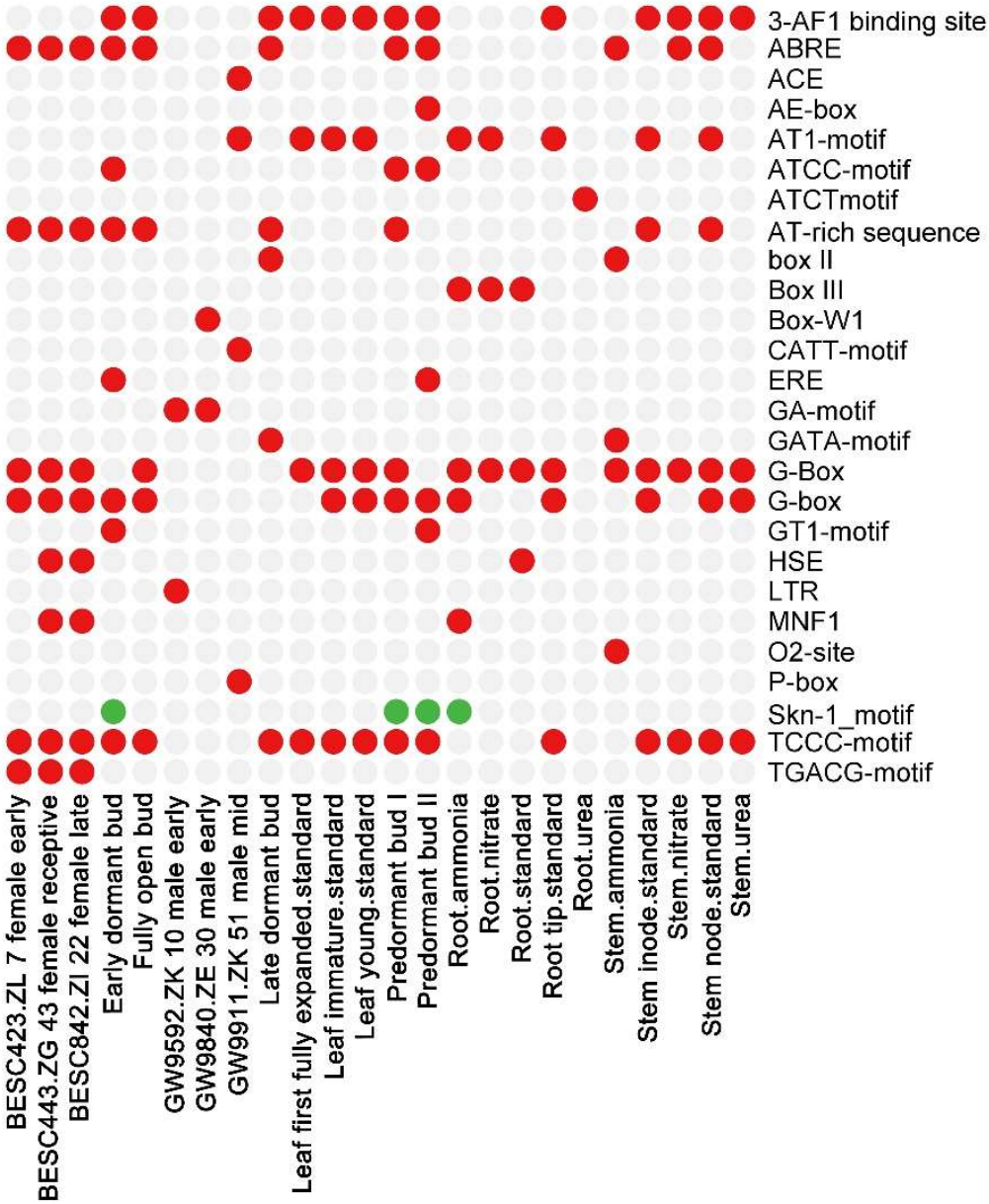

Figure 6. The gene expression in response to the cis-elements on the promoters. The circles represent the significant correlation (p = 0.05) between the expression level of the poplar expansin genes and the number of the specific cis-element on the promotor based on the Pearson correlation coefficient calculation by SPSS 20.0. The red circle indicates a positive correlation between the element and the gene expression of the poplar expansin genes in the tissue, while the green circle indicates a negative correlation.

4. Discussion

4.1. Phylogeny of Promoters and Coding Sequences

Plant expansin genes are classified into four sub-families, which originated from 12 ancestral *EXPA* genes, 2 ancestral *EXPB* genes, and 1–3 *EXLA/EXLB* genes in the *Arabidopsis* and rice genomes [31]. It is also applicable to the other plant genomes and displays that some clades have remained relatively constant in gene number, whereas a few clades have expanded [32]. This leads to the diversity of expansins and their importance in plant genomes.

Promoters contain a lot of regulatory elements and operate at the attached gene expression level. A sequence alignment revealed that the similarity among poplar expansin

genes was greater for the coding sequences than for the promoters; for example, the similarity of the coding sequence between *PtrEXPA25* and *PtrEXPA26* was 94.30%, but the similarity of the promoter sequence was only 30.40%. It suggested that the expansin gene promoter and coding sequences evolved differently. The promoters (i.e., the regulatory part) underwent more changes than the coding sequences, which could enable adaptations to diverse environments as well as appropriate growth and development. Similar observations were also reported in other studies. For example, Deng cloned an approximately 2-kb α-farnesene synthase gene promoter from '*Fuji*' apple and '*Fuyue*' pear (GenBank accession numbers: KM676083 and KM676082, respectively) [33]. Their sequence similarity was only 48.31%, which was much lower than the 97% similarity of the corresponding coding regions.

4.2. Promoter Characteristics in Association with Gene Expression

The examined expansin genes were generally expressed in an orderly and rigorous manner because of their promoters, which contain specific regulatory elements. Firstly, the base usage at the TSS considerably affects transcription efficiency [34]. Joshi studied the genomic sequences of 79 higher plant species and observed that most TSSs are adenine (A), whereas the flanking sites mostly consist of pyrimidine bases [35]. Our study also coincides with the results that the base usage frequencies at the TSS of poplar expansin gene promoters were 58.33% (A), 30.56% (C), 8.33% (T), and 2.78% (G). Particularly, an association analysis revealed that only base C was positively correlated with the expression of the expansin gene (12 samples). It indicated the special meaning of cytosine at the TSS for poplar genes. According to a previous study, the switch between a purine and a pyrimidine at the TSS might considerably affect the precise initiation of transcription [36].

The distribution of the TSS also defines the 5′ UTR length, which is highly variable in poplar expansin gene promoters (55–888 bp). Increases in the 5′ UTR length are likely associated with increased regulation of gene expression because of the potential for more regulatory elements than in shorter 5′ UTRs [37,38]. However, relatively long 5′ UTRs may also increase the energy consumed during transcription or lead to the formation of RNA secondary structures that are not conducive to the initiation of translation [39,40]. Thus, the diversity in the length and sequence of the 5′ UTR of poplar expansin gene promoters may be responsible for the various regulatory processes and functions of these genes in tissues or environmental conditions.

Transcription is initiated after RNA polymerase binds to the *TATA box* region. Thus, the *TATA-box* substantially influences gene expression efficiency along with its number and sequence [41]. A previous study on apple trees indicated that a *TATA-box* inserted in the upstream sequence of the iron transporter gene (*IRT1*) promoter enhances expression, thereby increasing the uptake of iron. Even so, promoter activity increased along with the number of inserted *TATA boxes* [42]. Joshi evaluated the promoter sequences of genes in 79 higher plants and determined that the base frequency of the *TATA-box* motif is $C_{0.47}T_{0.96}A_{0.97}T_{0.90}A_{0.94}T_{0.53}A_{0.95}T_{0.63}A_{0.71}G_{0.41}$ [43]. Bucher then analyzed the base usage frequency of the *TATA-box* in 502 unrelated RNA polymerase II promoter regions of eukaryotic plants and concluded the *TATA-box* sequence is G(C)TA(T)TA(T)A(T)AA(T)G(A)G(C) [44]. This indicates the conservatism of the *TATA-box* sequence is not so permanent. In the current study, we analyzed 36 poplar expansin gene promoters and concluded that the *TATA-box* motif comprises $C_{0.64}T_{0.86}A_{0.94}T_{0.94}A_{0.92}A_{0.69}A_{0.94}T_{0.72}A_{0.81}C_{0.61}$. Obviously, the core bases in the *TATA-box* sequence are consistent, which ensures normal gene expression. Also, some bases changed, including the sixth (T to A) and tenth bases (G to C). Some genes (e.g., *PtrEXPA1*, *PtrEXPA9*, *PtrEXPA23*, and *PtrEXPB2*) even had four differences among the conserved bases. The possible effects of these changes on gene expression levels were exciting and need to be investigated in future studies.

Meanwhile, the presence of the *CAAT-box* contributes to a substantial increase in gene expression [45]. Bezhani reported that the *CAAT-box* motif in a spinach photosynthesis-related gene (*AtpC*) promoter is AAAATTCAATGGC, which is targeted by ATPC-2 to promote transcription [46]. However, changing the third base (A to G) adversely affected

the *CAAT-box* motif. Among the poplar expansin gene promoters, the *CAAT-box* motif is represented by the following two sequences: AA(T)GA(C/T)CAATAA and TATCAAAATTT. Therefore, the differences in these two *CAAT-box* motifs do not appear to significantly influence poplar expansin gene expression.

In eukaryotes, the sequence flanking the start codon (ATG) is important for initiating translation. Kozak analyzed these sequences in 699 eukaryotic mRNAs [47]. Most of the sequences included the conserved motif A(G)NNATGG. Similarly, the poplar expansin genes described herein also followed this pattern. The frequency of A/G at site −3 was 83.33%, and the frequency of G at site +4 was 66.67%, indicative of an obvious base preference at these sites. Kozak explained that guanine at site +4 enhances the recognition of the initiation codon by the ribosome [48]. Furthermore, some of the other bases close to the initiation codon were also conserved in the poplar expansin genes, with a motif of ANAA(T)AATGGCA(T), which is similar to the AAAA(C)AATGGCT sequence in dicotyledonous plants [34]. However, a strong preference for G or C in this region, except at the start codon and at site −4, was revealed in rice, which is consistent with the CA(C)A(G)A(C)CATGGCG sequence in monocotyledonous plants. Such species characteristics in the region flanking the start codon are noteworthy and may be helpful for clarifying heterologous gene expression.

Further analysis showed that 21 of the 23 bases flanking the start codon of poplar expansin genes were related to gene expression. Primarily, the base at site −10 was associated with the gene expression in 16 samples, including various tissues, organs, developmental stages, and responses to stress treatments. Although the four base usage frequencies at site −10 were almost the same (Table 1), C was the most common base in the significantly expressed genes (positive correlation) (Figure 5). Additionally, site −17 was associated with gene expression in 11 samples, and bases G, T, and C were detected in 6, 4, and 1 samples, respectively. The other sites were also significantly associated with gene expression, including sites −19, −12, −11, −8, −5, −4, −3, +4, and +6. These sites exhibited a preference for one base (Figure 5). In Liu's report, nucleotide mutations flanking the ATG codon (i.e., sites −18, −16, −15, −9, −7, −1, and +6) in allergenic genes could downregulate gene expression by 63.3% in rice [49]. Thus, a more thorough characterization of these sites will help to elucidate the regulatory process underlying gene expression.

Of the 435 regulatory elements in poplar expansin gene promoters (Figure 6), 212 were related to photoreactions, especially *Box4* in 30 elements. The genes with the most photoreaction-related elements, such as *PtrEXPA3* (11), *PtrEXPA16* (11), and *PtrEXPB1* (12), were highly expressed in the photosynthesis-related *GW9592.ZK10* male early, leaf first fully expanded standard, leaf immature standard, and leaf young standard samples. Moreover, 46 plant-specific regulatory elements were identified in the poplar expansin gene promoters, with the *Skn-1* motif (endosperm-related regulatory element) being the most abundant (33). A previous study proved that the expression of *LeEXP4* in tomato plants can weaken the endosperm during seed germination, possibly due to the *Skn-1* motif and *GCN4* elements in its promoter [50]. This *Skn-1* element was also found in the promoter sequence of the *FAH12* gene in relation to lipid metabolism in *Ricinus communis* L. A protein interacting with *Skn-1*, RcWRKY48, was screened from the yeast library [51]. In addition, the expansin genes *PtrEXPLB2*, *PtrEXPA14*, *PtrEXPA17*, *PtrEXPA20*, *PtrEXPA25*, and *PtrEXPB3* each contained three or four tissue-specific regulatory elements and were accordingly expressed in the *GW9592.ZK10* male early, *GW9840.ZE30* male early, and *GW9911.ZK51* male mid-poplar samples. The promoter of the expansin gene *GmEXPB2*, which contributes to hairy root elongation, contains the meristem-related *CAT-box* motif, the core of the auxin response region, and ABRE [52]. Thus, we can hereby understand that each expansin gene has different or specific functions in diverse plant tissues and organs, as well as during various growth stages.

Some regulatory elements related to responses to environmental stresses were detected in the poplar expansin gene promoters, indicating the expansin genes are widely involved in plant stress resistance. These elements include 20 *ARE*, 16 *MBS*, 16 *HSE*, and 15 TC-

rich repeat elements. Some studies revealed an association between the elements and gene expression or resistance. In potatoes, the expression levels of *SpEXPA1*, *SpEXPA2*, *SpEXPA3*, and *SpEXPA4*, which have promoters with the *ARE* sequence, are reasonably up-regulated in the absence of O_2 and CO_2 [53]. In maize, *ZmEXPA4*, *ZmEXPA9*, and *ZmEXPB2* expression levels are down-regulated under water stress conditions since the promoters of these maize genes include the *MBS* element [54]. The turf grass *AsEXP1* gene, which has a promoter with the *HSE* sequence, exhibits up-regulated expression following exposure to high-temperature stress [55]. The promoter sequences of *TaEXPA1A* and *TaEXPB23B* harbor the drought-responsive element DRERTCOREAT (G/ACCGAC), which exhibits significant up-regulation in response to drought stress [56]. In this report, *PtrEXPA14*, *PtrEXPA15,* and *PtrEXPA3* were more sensitive to salt treatment. Their promoters all contained several salt resistance-related elements (*ARE* and *circadian*), while the promoters of the stress-insensitive genes *PtrEXPLB4*, *PtrEXPA7*, *PtrEXPA27*, *PtrEXPALA1*, and *PtrEXPA17* did not have these elements.

The poplar expansin gene promoters contain many hormone response-related regulatory elements. For example, 13 expansin gene promoters contain 18 methyl jasmonate-related CGTCA motifs and TGACG motifs. A total of 12 expansin gene promoters contain 15 gibberellin-related regulatory elements, including the *TATC-box*, *GARE-motif*, and *P-box*. These data indicate that the expansin genes might be involved in hormone regulation. For example, *AsEXP1* expression was regulated by GA and ABA in *Agrostis stolonifera* L. [26], *LeEXPA1* and *LeEXPA2* were expression-induced by ethylene and auxin in tomato [57,58], and *OsEXPA7* significantly responded to methyl jasmonate, brassinolide, and gibberellin in rice [59].

Overall, the abundance of regulatory elements in expansin gene promoters is consistent with the diverse functions of the encoded proteins regarding plant growth and development, as well as resistance to various stresses. The variability in the promoters in terms of the number and type of regulatory elements reflects the specific roles of the genes. Some genes with relatively few regulatory elements are insensitive to external factors. For example, the promoters of *PtrEXLA2* and *PtrEXLB1*, which were expressed at low levels in all examined samples, contain six regulatory elements each. In contrast, *PtrEXPA4* and *PtrEXPA3* were highly expressed in almost all analyzed samples; the promoters of these genes contain 20 (14 types) and 19 (17 types) regulatory elements, respectively. The expression levels of the other genes vary depending on tissue performance and environmental conditions. Therefore, characterizing the expansin gene promoter regulatory elements may be useful for revealing the status of the genes for plant growth and development as well as for species evolution.

5. Conclusions

Gene promoters contain multiple *cis*-elements that are important for gene expression through binding to transcription factors. The characterization of the poplar expansin gene promoters revealed that they have more diversity in cis-elements and have specific bases flanking the start codon. This is satisfied with the abundant expression pattern of the expansin genes and with their wide roles in plant growth and development and stress resistance. It was also recognized the significance of gene families, in which the members have complex regulation processes and function diversely, although they seem to have gene redundancy. Further, the promoter sequences of the poplar expansin genes varied greatly in comparison with their coding sequences. This asynchrony in evolution is significant for higher-living organisms, which have a stable size genome but perform a variety of life activities through promoter regulation. More efforts would be expected in the artificial modification of the promoters to meet our needs in a breeding program.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/f15091485/s1, Table S1: The specific primers of the poplar expansin genes.

Author Contributions: J.Z. and X.L.: investigation and writing—original draft. L.W., L.G., and M.L.: methodology and conceptualization. J.Z. and X.L.: writing—review and editing. J.X.: supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Special Fund for Forestry–Scientific Research in the Public Interest [grant number 201504101].

Data Availability Statement: The sequences of the poplar expansin gene and promoters were obtained from the poplar genome database (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Ptrichocarpa, version 13, accessed on 29 May 2023).

Acknowledgments: We thank Lesley Benyon for editing the English text of a draft of this manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kang, T.J.; Kwon, T.H.; Kim, T.G.; Loc, N.H.; Yang, M.S. Comparing constitutive promoters using CAT activity in transgenic tobacco plants. *Mol. Cells* **2003**, *16*, 117–122. [CrossRef]
2. Christensen, A.H.; Sharrock, R.A.; Quail, P.H. Maize polyubiquitin genes: Structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Mol. Biol.* **1992**, *18*, 675–689. [CrossRef] [PubMed]
3. Russell, D.A.; Fromm, M.E. Tissue-specific expression in transgenic maize of four endosperm promoters from maize and rice. *Transgenic Res.* **1997**, *6*, 157–168. [CrossRef]
4. Yeon, M.H.; Park, C.H.; Lee, Y.E.; Roh, J.; Kim, S.K. Seed-specifically overexpressed *Arabidopsis cytochrome P450 85A2* promotes vegetative and reproductive growth and development of *Arabidopsis thaliana*. *J. Plant Biol.* **2022**, *65*, 75–86. [CrossRef]
5. Roh, J.; Lee, Y.E.; Park, C.H.; Kim, S.K. Usefulness and molecular mechanism of seed-specificity introduced by *AtBZR1* and *AtBES1* to improve seed yield and quality in *Arabidopsis thaliana*. *J. Plant Biol.* **2023**, *66*, 233–242. [CrossRef]
6. Li, Y.Y.; Li, C.X.; Cheng, L.Z.; Yu, S.S.; Shen, C.J.; Pan, Y. Over-expression of *OsPT2* under a rice root specific promoter *Os03g01700*. *Plant Physiol. Biochem.* **2019**, *136*, 52–57. [CrossRef] [PubMed]
7. Jenny, P.; Sakure, A.A.; Yadav, A.; Kumar, S. Molecular cloning and characterization of root-specific *SlREO* promoter of the Indian tomato (*Solanum lycopersicum* L.) cultivar. *Funct. Plant Biol.* **2024**, *51*, FP24063. [CrossRef]
8. Vaughan, S.P.; James, D.J.; Lindsey, K.; Massiah, A.J. Characterization of *FaRB7*, a near root-specific gene from strawberry (*Fragaria* × *Ananassa* Duch.) and promoter activity analysis in homologous and heterologous hosts. *J. Exp. Bot.* **2006**, *57*, 3901–3910. [CrossRef]
9. Zhang, X.C.; Li, M.Y.; Ruan, M.B.; Xia, Y.J.; Wu, K.X.; Peng, M. Isolation of *AtNUDT5* gene promoter and characterization of its activity in transgenic *Arabidopsis thaliana*. *Appl. Biochem. Biotechnol.* **2013**, *169*, 1557–1565. [CrossRef]
10. Yang, Y.; Lee, J.H.; Poindexter, M.R.; Shao, Y.; Liu, W.; Lenaghan, S.C.; Ahkami, A.H.; Blumwald, E.; Stewart, C.N. Rational design and testing of abiotic stress-inducible synthetic promoters from poplar *cis*-regulatory elements. *Plant Biotechnol. J.* **2021**, *19*, 1354–1369. [CrossRef] [PubMed]
11. Yang, Y.; Shao, Y.; Chaffin, T.A.; Lee, J.H.; Poindexter, M.R.; Ahkami, A.H.; Blumwald, E.; Stewart, C.N. Performance of abiotic stress-inducible synthetic promoters in genetically engineered hybrid poplar (*Populus tremula* × *Populus alba*). *Front. Plant Sci.* **2022**, *13*, 1011939. [CrossRef]
12. Yang, Y.G.; Tagaloguin, P.; Chaffin, T.A.; Shao, Y.H.; Mazarei, M.; Millwood, R.J.; Stewart, C.N. Drought stress-inducible synthetic promoters designed for poplar are functional in rice. *Plant Cell Rep.* **2024**, *43*, 69. [CrossRef] [PubMed]
13. Zhang, H.Y.; Xiao, Y.X.; Tian, Y.K.; Wang, C.H.; Yang, S.L.; Li, Y. Sequence characters analysis of a β-Tubulin gene and its promoter in apple. *Acta Agric. Boreali Sin.* **2018**, *33*, 78–85. [CrossRef]
14. Juven-Gershon, T.; Kadonaga, J.T. Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev. Biol.* **2010**, *339*, 225–229. [CrossRef]
15. Jin, J.; Li, L.L.; Fan, D.Y.; Du, Y.W.; Jia, H.C.; Yang, L.; Jia, W.S.; Hao, Q. Budding mutation reprogrammed flavonoid biosynthesis in jujube by deploying MYB41 and bHLH93. *Plant Physiol. Biochem.* **2024**, *211*, 108665. [CrossRef]
16. Jiang, Y.; Zeng, Z.X.; He, G.H.; Liu, M.N.; Liu, C.; Liu, M.M.; Lv, T.T.; Wang, A.M.; Wang, Y.; Zhao, M.Z. Genome-wide identification and integrated analysis of the *FAR1/FHY3* gene family and genes expression analysis under methyl jasmonate treatment in *Panax ginseng* C. A. Mey. *BMC Plant Biol.* **2024**, *24*, 549. [CrossRef]
17. Hendelman, A.; Zebell, S.; Rodriguez-Leal, D.; Dukler, N.; Lippman, Z.B.; Robitaille, G.; Wu, X.L.; Kostyun, J.; Tal, L.; Wang, P.P.; et al. Conserved pleiotropy of an ancient plant homeobox gene uncovered by *cis*-regulatory dissection. *Cell* **2021**, *184*, 1724–1739. [CrossRef] [PubMed]
18. Cosgrove, D.J. Loosening of plant cell walls by expansins. *Nature* **2000**, *407*, 321–326. [CrossRef]
19. Zhang, J.; Yang, R.; Wang, Y.; Wang, X.; Wang, L.; Xu, J. The expansin gene *SmEXPA13* in *Salix matsudana* in association with plant salt tolerance. *Plant Cell Tiss. Org.* **2023**, *154*, 219–225. [CrossRef]
20. Liu, J.; Wang, Y.; Yang, L.; Wang, X.; Zhang, J.; Xu, J. Characterization and functional analysis of the *PtEXLA1* gene from poplar. *Plant Biotechnol. Rep.* **2024**, *18*, 119–128. [CrossRef]

21. Li, H.; Shi, Y.; Ding, Y.; Xu, J. Bioinformatics analysis of the expansin gene family in poplar genome. *J. Beijing For. Univ.* **2014**, *36*, 59–67. [CrossRef]
22. Dharmawardhana, P.; Brunner, A.M.; Strauss, S.H. Genome-wide transcriptome analysis of the transition from primary to secondary stem development in *Populus trichocarpa*. *BMC Genom.* **2010**, *11*, 150. [CrossRef]
23. Yin, Z.H.; Zhou, F.W.; Chen, Y.N.; Wu, H.T.; Yin, T.M. Genome-Wide analysis of the expansin gene family in *Populus* and characterization of expression changes in response to phytohormone (abscisic acid) and abiotic (low-temperature) stresses. *Int. J. Mol. Sci.* **2023**, *24*, 7759. [CrossRef]
24. Geisler-Lee, J.; Geisler, M.; Coutinho, P.M.; Segerman, B.; Nishikubo, N.; Takahashi, J.; Aspeborg, H.; Djerbi, S.; Master, E.; Andersson-Gunneras, S.; et al. Poplar carbohydrate-active enzymes.: Gene identification and expression analyses. *Plant Physiol.* **2006**, *140*, 946–962. [CrossRef] [PubMed]
25. Gray-Mitsumune, M.; Mellerowicz, E.G.; Abe, H.; Schrader, J.; Winzell, A.; Sterky, F.; Blomqvist, K.; McQueen-Mason, S.; Teeri, T.T.; Sundberg, B. Expansins abundant in secondary xylem belong to subgroup a of the α-expansin gene family. *Plant Physiol.* **2004**, *135*, 1552–1564. [CrossRef]
26. Zhou, P.; Zhu, Q.; Xu, J.; Huang, B. Cloning and characterization of a gene, *AsEXP1*, encoding expansin proteins inducible by heat stress and hormones in creeping bentgrass. *Crop Sci.* **2011**, *51*, 333–341. [CrossRef]
27. Ogden, T.H.; Rosenberg, M.S. Multiple sequence alignment accuracy and phylogenetic inference. *Syst. Biol.* **2006**, *55*, 314–328. [CrossRef] [PubMed]
28. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef]
29. Felsenstein, J. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **1985**, *39*, 783–791. [CrossRef] [PubMed]
30. Shahmuradov, I.A.; Solovyev, V.V.; Gammerman, A.J. Plant promoter prediction with confidence estimation. *Nucleic Acids Res.* **2005**, *33*, 1069–1076. [CrossRef]
31. Sampedro, J.; Lee, Y.; Carey, R.E.; dePamphilis, C.; Cosgrove, D.J. Use of genomic history to improve phylogeny and understanding of births and deaths in a gene family. *Plant J.* **2005**, *44*, 409–419. [CrossRef]
32. Sampedro, J.; Cosgrove, D.J. The expansin superfamily. *Genome Biol.* **2005**, *6*, 242. [CrossRef]
33. Deng, S.; Cheng, N.; Ding, R.; Liu, Y.; Zhang, Y.H. Cloning, sequence alignment and functional analysis of *AFS* gene promoter in apple and pear fruits. *Acta Hortic. Sin.* **2015**, *42*, 2353–2361. [CrossRef]
34. Massin, P.; Rodrigues, P.; Marasescu, M.; van der Werf, S.; Naffakh, N. Cloning of the chicken RNA polymerase I promoter and use for reverse genetics of influenza A viruses in avian cells. *J. Virol.* **2005**, *79*, 13811–13816. [CrossRef] [PubMed]
35. Joshi, C.P.; Hao, Z.; Huang, X.; Chiang, V.L. Context sequences of translation initiation codon in plants. *Plant Mol. Biol.* **1997**, *35*, 993–1001. [CrossRef] [PubMed]
36. Zhu, Q.; Dabi, T.; Beeche, A.; Yamamoto, R.; Lawton, M.A.; Lamb, C. Cloning and properties of a rice gene encoding phenylalanine ammonia-lyase. *Plant Mol. Biol.* **1995**, *29*, 535–550. [CrossRef] [PubMed]
37. Wray, G.A.; Hahn, M.W.; Abouheif, E.; Balhoff, J.P.; Pizer, M.; Rockman, M.V.; Romano, L.A. The evolution of transcriptional regulation in eukaryotes. *Mol. Biol. Evol.* **2003**, *20*, 1377–1419. [CrossRef]
38. Lynch, M.; Scofield, D.G.; Hong, X. The evolution of transcription-initiation sites. *Mol. Biol. Evol.* **2005**, *22*, 1137–1146. [CrossRef]
39. Beaudoin, J.D.; Perreault, J.P. 5′-UTR G-quadruplex structures acting as translational repressors. *Nucleic Acids Res.* **2010**, *38*, 7022–7036. [CrossRef]
40. Calvo, S.E.; Pagliarini, D.J.; Mootha, V.K. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 7507–7512. [CrossRef]
41. Jores, T.; Tonnies, J.; Wrightsman, T.; Buckler, E.S.; Cuperus, J.T.; Fields, S.; Queitsch, C. Synthetic promoter designs enabled by a comprehensive analysis of plant core promoters. *Nat. Plants* **2021**, *7*, 842–855. [CrossRef] [PubMed]
42. Zhang, M.L.; Lv, Y.D.; Wang, Y.; Rose, J.K.C.; Shen, F.; Han, Z.Y.; Zhang, X.Z.; Xu, X.F.; Wu, T.; Han, Z.H. TATA box insertion provides a selection mechanism underpinning adaptations to Fe deficiency. *Plant Physiol.* **2017**, *173*, 715–727. [CrossRef]
43. Joshi, C.P. An inspection of the domain between putative TATA box and translation start site in 79 plant genes. *Nucleic Acids Res.* **1987**, *15*, 6643–6653. [CrossRef] [PubMed]
44. Bucher, P. Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *J. Mol. Biol.* **1990**, *212*, 563–578. [CrossRef]
45. Sreenivasulu, G.; Senthilkumaran, B.; Sudhakumari, C.C.; Guan, G.; Oba, Y.; Kagawa, H.; Nagahama, Y. 20β-hydroxysteroid dehydrogenase gene promoter: Potential role for cyclic AMP and xenobiotic responsive elements. *Gene* **2012**, *509*, 68–76. [CrossRef] [PubMed]
46. Bezhani, S.; Sherameti, I.; Pfannschmidt, T.; Oelmuller, R. A repressor with similarities to prokaryotic and eukaryotic DNA helicases controls the assembly of the CAAT box binding complex at a photosynthesis gene promoter. *J. Biol. Chem.* **2001**, *276*, 23785–23789. [CrossRef] [PubMed]
47. Kozak, M. An analysis of 5′-noncoding sequences from 699 vertebrates messenger RNAs. *Nucleic Acids Res.* **1987**, *15*, 8125–8148. [CrossRef]
48. Kozak, M. Recognition of AUG and alternative initiator codons is augmented by G in position +4 but is not generally affected by the nucleotides in positions +5 and +6. *Embo J.* **1997**, *16*, 2482–2492. [CrossRef]

49. Liu, Q.P.; Tan, J.; Xue, Q.Z. Effect of the flanking sequence architecture of AUG, a initiator codon on gene expression level in rice. *Sci. Agric. Sin.* **2004**, *37*, 625–629.
50. Chen, F.; Bradford, K.J. Expression of an expansin is associated with endosperm weakening during tomato seed germination. *Plant Physiol.* **2000**, *124*, 1265–1274. [CrossRef]
51. Wang, Y. Construction of Castor (*Ricinus communis* L.) Yeast-One Hybrid cDNA Library and Screening of Skn-1 Binding Protein. Master's thesis, Inner Mongolia Minzu University, Inner Mongolia, China, 2018.
52. Guo, W.; Zhao, J.; Li, X.; Qin, L.; Yan, X.L.; Liao, H. A soybean β-expansin gene *GmEXPB2* intrinsically involved in root system architecture responses to abiotic stresses. *Plant J.* **2011**, *66*, 541–552. [CrossRef]
53. Ookawara, R.; Satoh, S.; Yoshioka, T.; Ishizawa, K. Expression of α-expansin and xyloglucan endotransglucosylase/hydrolase genes associated with shoot elongation enhanced by anoxia, ethylene and carbon dioxide in arrowhead (*Sagittaria pygmaea* Miq.) tubers. *Ann. Bot.* **2005**, *96*, 693–702. [CrossRef] [PubMed]
54. Muller, B.; Bourdais, G.; Reidy, B.; Bencivenni, C.; Massonneau., A.; Condamine, P.; Rolland, G.; Conéjéro, G.; Rogowsky, P.; Tardieu, F. Association of specific expansins with growth in maize leaves is maintained under environmental, genetic, and developmental sources of variation. *Plant Physiol.* **2007**, *143*, 278–290. [CrossRef] [PubMed]
55. Xu, J.C.; Tian, J.; Belanger, F.C.; Huang, B.R. Identification and characterization of an expansin gene *AsEXP1* associated with heat tolerance in C3*Agrostis* grass species. *J. Exp. Bot.* **2007**, *58*, 3789–3796. [CrossRef]
56. Chen, S.K.; Ren, H.Y.; Luo, Y.X.; Feng, C.Z.; Li, H.F. Genome-wide identification of wheat (*Triticum aestivum* L.) expansin genes and functional characterization of *TaEXPB1A*. *Environ. Exp. Bot.* **2021**, *182*, 104307. [CrossRef]
57. Rose, J.K.C.; Lee, H.H.; Bennett, A.B. Expression of a divergent expansin gene is fruit-specific and ripening-regulated. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 5955–5960. [CrossRef] [PubMed]
58. Catala, C.; Rose, J.K.; Bennett, A.B. Auxin-regulated genes encoding cell wall-modifying proteins are expressed during early tomato fruit growth. *Plant Physiol.* **2000**, *122*, 527–534. [CrossRef] [PubMed]
59. Zhang, X.W.; Wang, Y.; Liu, M.Y.; Yan, P.W.; Niu, F.; Ma, F.Y.; Hu, J.; He, S.C.; Cui, J.H.; Yuan, X.Y.; et al. *OsEXPA7* encoding an expansin affects grain size and quality traits in rice (*Oryza sativa* L.). *Rice* **2024**, *17*, 36. [CrossRef]