

journal homepage: [www.elsevier.com/locate/febsopenbio](http://www.elsevier.com/locate/febsopenbio)

# Genome-wide analysis of the AP2/ERF gene family in *Salix arbutifolia*

Guodong Rao<sup>a,b,c</sup>, Jinkai Sui<sup>b</sup>, Yanfei Zeng<sup>a</sup>, Caiyun He<sup>a</sup>, Jianguo Zhang<sup>a,b,c,\*</sup><sup>a</sup>State Key Laboratory of Tree Genetics and Breeding, Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, China<sup>b</sup>Collaborative Innovation Center of Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing 210037, China<sup>c</sup>Key Laboratory of Tree Breeding and Cultivation, State Forestry Administration, Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, China

## ARTICLE INFO

### Article history:

Received 2 December 2014

Revised 15 February 2015

Accepted 18 February 2015

### Keywords:

Transcription factors

AP2/ERF

Transcriptome

*Salix arbutifolia*

## ABSTRACT

**AP2/ERF genes encode transcriptional regulators with a variety of functions in plant growth and development and in response to biotic and abiotic stresses. To date, there are no detailed classification and expression profiles for AP2/ERF genes in *Salix*. In this study, a comprehensive computational analysis identified 173 AP2/ERF superfamily genes in willow (*Salix arbutifolia*), by using *in silico* cloning methods with the use of the AP2/ERF conserved domain amino acid sequence of *Arabidopsis thaliana* as a probe. Based on the results of phylogenetic analyses and the number of AP2/ERF domains, the AP2/ERF genes were classified into four groups: AP2, RAV, ERF and Soloist. The expression profile was analyzed using transcriptome data from different tissues. A comparative analysis of AP2/ERF superfamily genes among *Salix*, *Populus* and *Arabidopsis* was performed. The *Salix* DREB, AP2 and RAV groups had a similar number to those in *Arabidopsis*, and the size of the ERF subfamily in *Salix* was about 1.4-fold that of *Arabidopsis*. The *Salix* DREB subfamily was smaller compared to *Populus*, while the other families were similar in size to those in *Populus*. These results will be useful for future functional analyses of the ERF family genes.**

© 2015 The Authors. Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

The genus *Salix*, including willows, sallows and osiers, contains around 400 species in the Northern hemisphere, making this genus a valuable forest resource and an important ecological species [1]. The utilization of and research into the use of willow for biomass production in short-rotation-intensive culture systems is growing worldwide, which has created increased interest in the breeding of high-yield and resistant clones adapted to different environments.

Plant cellular and developmental processes are regulated and controlled by gene expression of multiple different sets of transcription factors (TFs) [2]. TFs also play pivotal functions in signal transduction during the activation or suppression of defense response genes, especially when plants face abiotic and biotic stresses such as drought and high salinity [3]. The AP2/ERF transcription factor superfamily contains the most TFs in plants and these all include at least one APETALA2 (AP2) domain. The AP2/ERF superfamily is divided into the AP2, ERF (including ERF and DREB

subfamily), RAV and Soloist families according to the number of AP2 domains [4,5]. The AP2 family has duplicated AP2/ERF domains, the ERF family has a single AP2/ERF domain, while the RAV family has one B3 domain and one AP2/ERF domain [5–7]. In *Arabidopsis*, a total of 145 AP2/ERF genes have been characterized, including 17 in the AP2 family, 6 in the RAV family, 65 in the ERF subfamily (of the ERF family), 56 in the DREB subfamily (of the ERF family) and 1 in the Soloist family [8]. In *Populus*, a total of 200 AP2/ERF genes have been characterized, including 26 in the AP2 family, 5 in the RAV family, 91 in the ERF subfamily, 77 in the DREB subfamily and 1 in the Soloist family [9]. With more extensive genome sequences, the identification and characterization of the AP2/ERF superfamily have been conducted in various plants, such as grape [10], cucumber [11], rice [12], tomato [13], peach, sorghum [14], Chinese cabbage [15] and soybean [16]. However, no research has been performed for the identification and characterization of the AP2/ERF superfamily in *Salix*.

*Salix arbutifolia* Pallas (originally *Chosenia arbutifolia*) is a riparian dioecious tree species distributed in Northeast Asia [17]. Putative genes encoding proteins of the AP2/ERF family in *S. arbutifolia* were found, resulting in the identification of 173 genes in this superfamily, including 23 in the AP2 family, 4 in the RAV family, 88 in the DREB subfamily and 57 in the ERF subfamily. Protein motif structure analysis was performed, as well as a phylogenetic analysis

**Abbreviations:** TFs, transcription factors; Nr, non-redundant protein; COGs, Cluster of Orthologous Groups

\* Corresponding author at: State Key Laboratory of Tree Genetics and Breeding, Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, China. Tel.: +86 10 62889601; fax: +86 10 62872015.

E-mail address: [zhangjg@caf.ac.cn](mailto:zhangjg@caf.ac.cn) (J. Zhang).

<http://dx.doi.org/10.1016/j.fob.2015.02.002>

2211-5463/© 2015 The Authors. Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

of the AP2/ERF genes. Moreover, using transcriptome data, we demonstrate the expression of AP2/ERF genes in different tissues. The resulting classification of groups and identification of putative functional motifs will be useful in studies on the biological functions of each gene in the *Salix* AP2/ERF families. The transcriptome reported here significantly enhances our knowledge of TFs and stress response genes in *Salix*.

## 2. Materials and methods

### 2.1. Gene isolation and analysis

AP2/ERF genes were isolated from the *S. arbutifolia* genome sequence (paper in press) using annotation and BLAST. First, sequences indicated as belonging to AP2/ERF were isolated from the annotation. Then, using *Arabidopsis* and *Populus* AP2/ERF genes with TBALSTN and BLASTP, a batch of related sequences were obtained. All of the sequences that had AP2/ERF domains were selected as candidates. To avoid the inclusion of redundant sequences, the characterized AP2/ERF genes were assembled again using the CAP3 Sequence Assembly Program [18]. The motif identification of *Salix* AP2/ERF protein sequences was performed using a motif-based sequence analysis tool, MEME Suite version 4.9.1 [19]. The optimum width of amino acid sequences was set from 6 to 25. The maximum number of motifs was set to 35.

### 2.2. Plant material, library construction and deep sequencing

Stems, leaves and Blossom buds from a natural population of *S. arbutifolia* were harvested in spring 2013 in Liaoning Province in China. To ensure that there were minimal differences in growth conditions, all materials were sampled within an area of  $30 \times 30 \text{ m}^2$ . Samples were frozen in liquid nitrogen immediately for the extraction of total RNA.

For each tissue type of sample, equal numbers of samples from five individual plants were pooled for total RNA extraction with TRIZOL reagent according to the manufacturer's instructions (Invitrogen). Transcriptome sequencing was performed using total RNA from the pooled samples. After eliminating traces of genomic DNA by treatment with DNase I, an automated capillary gel electrophoresis was used to assess RNA quality and quantity using a Bioanalyzer 2100 with RNA 6000 Nano Labchips (Agilent Technologies Ireland, Dublin, Ireland). Transcriptome sequencing libraries were prepared using an Illumina Small RNA Sample Prep Kit and an Illumina TruSeq RNA Sample Prep Kit. After the three libraries had been prepared, raw reads generated by using the Illumina HiSeq™ 2000 were initially processed to get clean reads. Then, all the clean reads were assembled using the *de novo* assembly program Trinity [20].

### 2.3. Phylogenetic analysis

Phylogenetic and molecular evolutionary analyses were conducted using the software of MEGA version 6 [21]. Complete AP2/ERF predicted amino acids sequences was performed with a gap open penalty of 10 and gap extension penalty of 0.2 using ClustalW implemented in MEGA. Neighbor-joining method was used to construct phylogenetic tree by using, and the reliability of the obtained trees was tested using bootstrapping with 1000 replicates.

### 2.4. Expression pattern analysis of AP2/ERF genes

An *in silico* analysis was used to compare the relative transcript abundance of the AP2/ERF genes in stems, leaves and Blossom buds, based on transcriptome data. The gene expression level was normalized to the values of RPKM (Reads per kb per million

reads) [22]. The random selected AP2/ERF genes were analyzed by real-time quantitative PCR. PCR were conducted using the SYBR Green Perfect (Takara) and StepOnePlus™ System (Applied Biosystems). All of the PCR products were sequenced and the dissociation curve was analyzed to verify the amplification specificity. The purified PCR products were used to make standard curve for establishing a quantitative correlation between the CT values and the transcript copy numbers. Each qRT-PCR reaction was repeated at least three times, and each standard curve contains at least 5 points. *Populus* actin gene was used for reference gene for transcriptome genes expression validation, respectively.

## 3. Results

### 3.1. Identification of AP2/ERF genes in *S. arbutifolia*

In total, one hundred and seventy-three AP2/ERF genes were identified from the *S. arbutifolia* genome sequence (paper in press). Twenty-two genes were predicted to encode proteins containing two AP2-domains based on the similarity of their amino acid sequences with the AP2 family of *Populus trichocarpa* and *Arabidopsis thaliana* proteins, and were therefore assigned to the AP2 family. One hundred forty-five genes were predicted to encode proteins containing a single AP2/ERF domain. These 145 genes could be further classified into two groups on the basis of similarity of the amino acid sequences. Fifty-seven genes were identified as possibly encoding members of the DREB subfamily, and eighty-eight genes were predicted to encode the ERF subfamily. Four genes were predicted to encode one AP2-domain and one B3-domain, and were assigned to the RAV family. The remaining gene, which was obtained by homology of the Soloist from *Arabidopsis* (At4g13040), included an AP2/ERF-like domain sequence, but its homology appears quite low in comparison with the other AP2/ERF factors. Therefore, this gene was assigned to the Soloist family (Table 1).

The complete genomes of *P. trichocarpa* and *A. thaliana* provide ideal systems for comparison of the AP2/ERF group among these plant species. The total number of genes in AP2/ERF in *S. arbutifolia* (173) was 1.19-times those genes in *A. thaliana* (145). The total number in AP2/ERF of *P. trichocarpa* (200) was 1.16-times those genes from *S. arbutifolia*. The proportion of genes found in the DREB, ERF, AP2 and RAV groups from *S. arbutifolia* were similar to the proportion in *P. trichocarpa* and *A. thaliana* (Fig. 1).

### 3.2. Phylogenetic reconstruction of AP2/ERF superfamily in *S. arbutifolia*

To study the phylogenetic relationships of the AP2/ERF superfamily in *S. arbutifolia*, a multiple alignment analysis was performed using the full amino acid sequences of the AP2/ERF superfamily between *S. arbutifolia* and *A. thaliana*. The resulting phylogenetic tree resolved fifteen clades (Fig. 2), containing five group, named DREB, ERF, AP2, RAV and Soloist, which are congruent with previous studies [6,11]. DREB genes were distributed into A1, A2, A3, A4, A5 and A6 subgroups, and ERF genes were also classified into six subgroups, namely, B1, B2, B3, B4, B5 and B6. *Salix* and *Populus* AP2/ERF genes share much amino acid sequence identity, and have similar properties in gene structure. A phylogenetic tree showed that most of the *Salix* AP2/ERF genes had counterparts in *Populus*, suggesting *Salix* and *Populus* AP2/ERF genes descended from the same ancestor. However, not every *Salix* AP2/ERF gene has a counterpart in *Populus*, suggesting that *Salix* and *Populus* have separately undergone different expansions (Fig. S2).

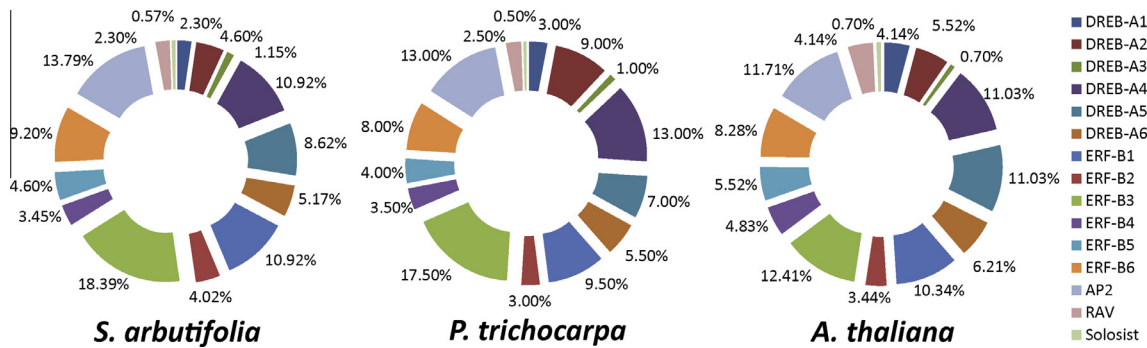
### 3.3. Transcript expression levels of AP2/ERF genes in *S. arbutifolia*

Transcription levels of AP2/ERF genes were statistical analyzed from the transcriptome sequenced using RNA-seq and *de novo*

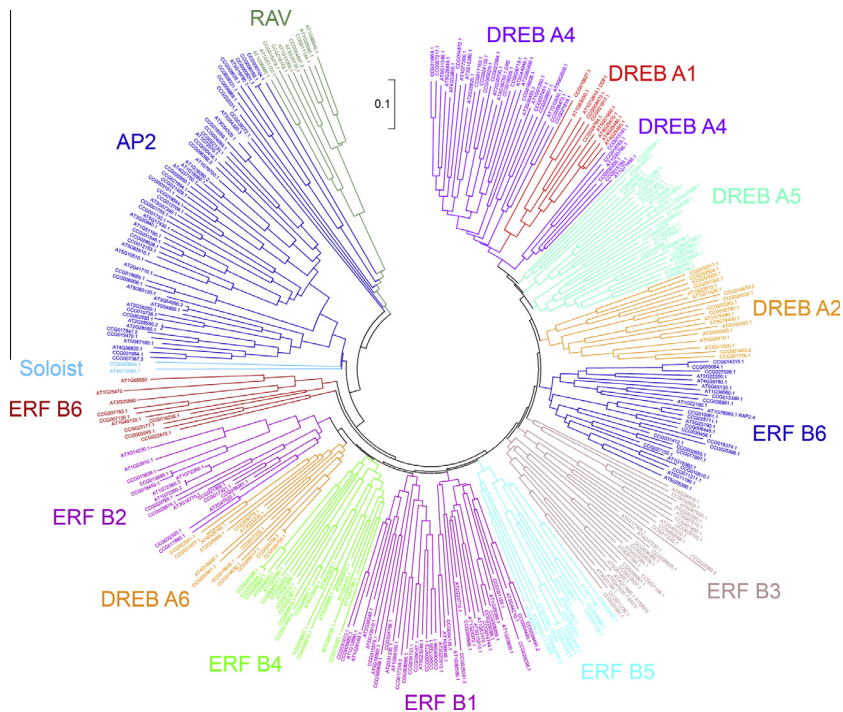
**Table 1**  
Summary of the AP2/ERF family of *S. arbutifolia*, *P. trichocarpa* and *A. thaliana*.

Plant		<i>S. arbutifolia</i>		<i>P. trichocarpa</i>		<i>A. thaliana</i>	
Classification	Group	Number	Percent	Number	Percent	Number	Percent
DREB	A1	4	2.30	6	3.00	6	4.14
	A2	8	4.60	18	9.00	8	5.52
	A3	2	1.15	2	1.00	1	0.70
	A4	19	10.92	26	13.00	16	11.03
	A5	15	8.62	14	7.00	16	11.03
	A6	9	5.17	11	5.50	9	6.21
	Total		57	32.76	77	38.50	56
ERF	B1	19	10.92	19	9.50	15	10.34
	B2	7	4.02	6	3.00	5	3.44
	B3	32	18.39	35	17.50	18	12.41
	B4	6	3.45	7	3.50	7	4.83
	B5	8	4.60	8	4.00	8	5.52
	B6	16	9.20	16	8.00	12	8.28
	Total		88	50.57	91	45.50	65
AP2		22	13.79	26	13.00	17	11.71
RAV		4	2.30	5	2.50	6	4.14
Soloist		1	0.57	1	0.50	1	0.70
Total		173		200		145	

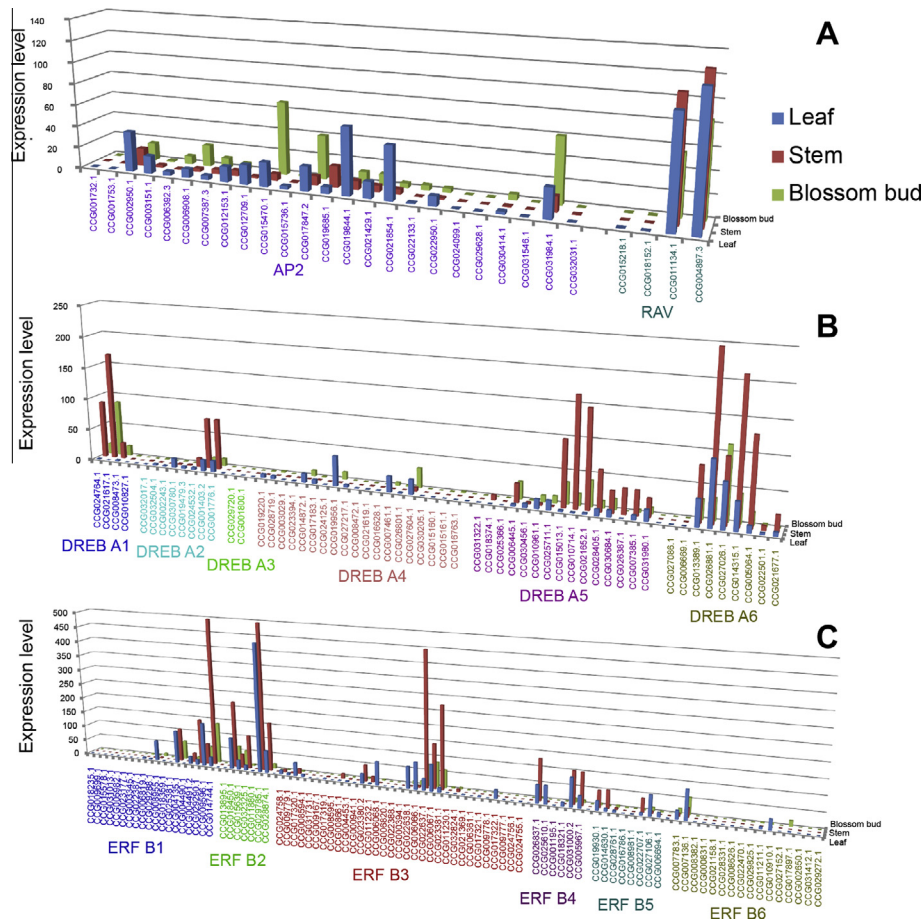
The number of AP2/ERF family from *P. trichocarpa* and *A. thaliana* was according to Zhuang et al. [9].



**Fig. 1.** The proportion of AP2/ERF genes in *S. arbutifolia*, *P. trichocarpa* and *A. thaliana*.



**Fig. 2.** The phylogenetic tree of the entire AP2/ERF superfamily from *A. thaliana* and *S. arbutifolia*.



**Fig. 3.** Transcript levels of AP2/ERF groups in Blossom buds, stems and leaves. (A) the expression level of the AP2 and RAV families; (B) the expression level of the DREB subfamily; (C) the expression level of the ERF subfamily.

assembly. The RNA-seq libraries were prepared from the total RNA extracted from *S. matsudana* subjected to pair-end read (PE) sequencing using the Illumina platform. We removed reads with adaptors, reads with unknown nucleotides larger than 5%, and low quality reads to obtain clean PE reads. Gene annotation was performed using BLAST2GO to find sequence homologies using the databases of NCBI non-redundant protein (Nr), NCBI nucleotide sequence (Nt), Kyoto Encyclopedia of Genes and Genomes (KEGG), Swiss-Prot protein, Protein family, Gene Ontology (GO), and Cluster of Orthologous Groups (COGs). A quality control test for the data assembled from each cDNA library confirmed that they were suitable for statistical analysis for differentially expressed gene identification.

*S. arbutifolia* is strongly affected by environmental cues and pathogens during development. Since AP2/ERF genes play a role in plant development and in response to biotic and abiotic stresses [23,24], they are ideal candidates to investigate the molecular regulation of these processes. Transcription levels of AP2/ERF genes were statistically analyzed from the stem, leaf and Blossom bud transcriptomes. AP2 genes had the most abundant transcript in leaves, and were also expressed at moderate levels in stems and buds. The four members of RAV had a different expression pattern; two of the genes were highly expressed in all tissues examined, while the other two had very low transcription levels (Fig. 3A). Six DREB genes also exhibited a diverse expression pattern. DREB A1, DREB A5 and DREB A6 had very high expression levels, while DREB A2 and DREB A3 had a low amount of transcript copies (Fig. 3B). Transcripts of ERF genes had the highest expression level compared to other AP2/ERF groups (Fig. 3C). To validate

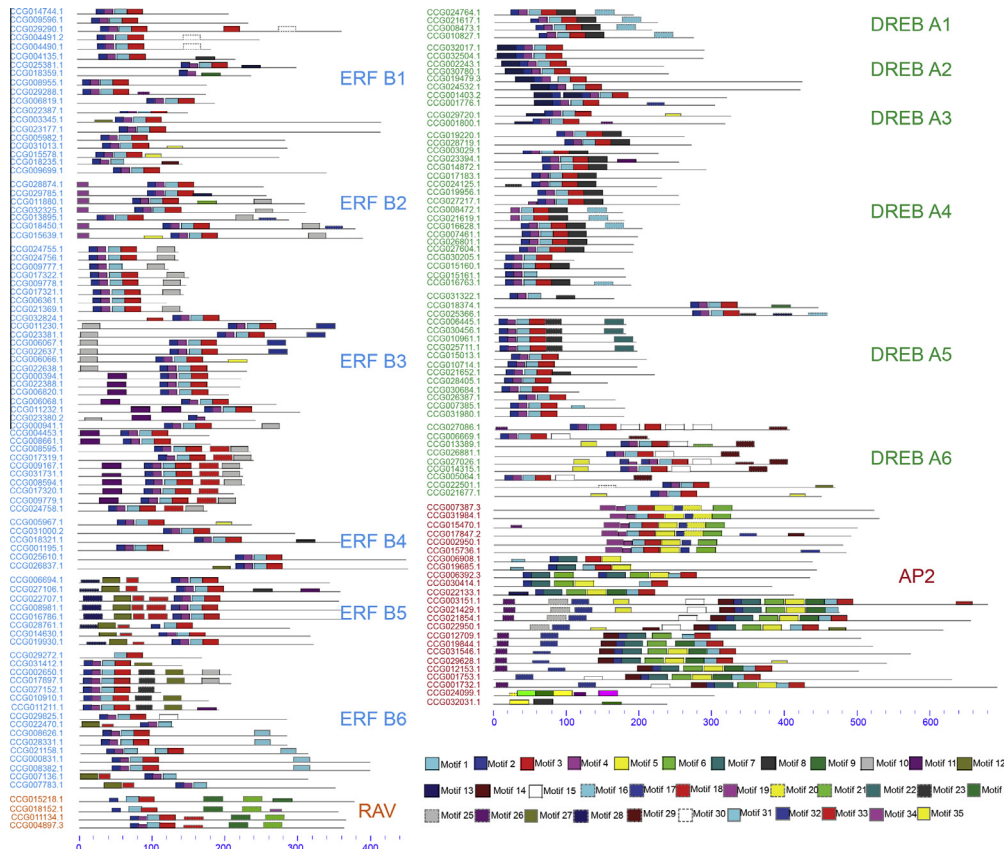
the AP2/ERF genes expression levels, twenty genes were randomly selected for a real-time qPCR experiment. The results showed that the expression of AP2/ERF genes changed trends as compared with the transcriptome sequencing results (Fig. S1).

### 3.4. Analysis of conserved motifs in AP2/ERF

All AP2/ERF proteins were subjected to MEME motif analysis to identify conserved motifs among proteins in the families and subfamilies. Conserved motifs can provide evidence for further classification as it is likely that identical motifs exhibit similar functions [25]. The MEME motif analysis revealed that different AP2/ERF genes had different conserved motifs (Fig. 4). Four conserved domains (domains 1–4) are a key characteristic of AP2/ERF. Most of the *Salix* AP2/ERF genes had the full AP2/ERF domain, with all 4 of the motifs, although some had only partial domains (Table 2). Aside from the AP2/ERF domain, other domains can also play an important role in transcriptional regulation. In contrast to the conserved AP2/ERF domain, the other domains were only found in some subfamilies or distributed among some genes. For example, the AP2 family had a conserved domain motif 5, most of the DREB subfamily had motif 8, and all of the RAV family had motif 21. Some other motifs were distributed among various families and subfamilies.

## 4. Discussion

Increasing concern for climate change and energy security has resulted in a focus on the economic importance of *Salix* species due to their utility for bioenergy production. However, plants have



**Fig. 4.** Distribution of conserved motifs in AP2/ERF protein sequences. The motifs (1–35) were identified using the MEME tool. Subfamily/family names are indicated on the right; each color represents a distinct motif and the length of each motif is indicated.

**Table 2**  
Four conserved domains (motif 1–4) in AP2/ERF subfamilies.

Classification	Group	Number	Motif 1	Motif 2	Motif 3	Motif 4
DREB	A1	4	4	4	4	4
	A2	8	8	8	8	8
	A3	2	2	2	2	2
	A4	19	19	19	19	19
	A5	15	15	15	14	15
	A6	9	9	9	9	9
	Total	57	57	57	57	57
ERF	B1	19	19	18	18	19
	B2	7	7	7	7	7
	B3	32	31	32	31	32
	B4	6	6	6	6	6
	B5	8	8	8	8	7
	B6	16	16	15	13	15
	Total	88	88	88	88	88
AP2		23	22	23	22	22
RAV		4	4	4	4	2
Total		172	170	170	165	169

to tolerate many abiotic and biotic stresses, which are serious threats to agriculture and forestry. These common environmental stresses include salinity, extreme temperatures, drought, chemical toxicity, oxidative stress, pests and pathogen infection. To exploit their potential for renewable energy, willows have to be kept free of pests and diseases and the yield should be improved without significantly increasing the requirement for fertilizers and water [26,27]. To improve *Salix* tolerance to abiotic and biotic stresses, research has been focused on the physiological and molecular mechanisms of the response of *Salix* to these stresses. An important aim in identifying and isolating AP2/ERF gene family members is to understand the molecular genetic basis useful for the

improvement of *Salix* and for providing the functional genetic resources required for transgenic research.

Transcription factors play a key role in modulating the acclimatization response of plants to various internal or external cues. They potentially control downstream gene expression in stress signal transduction pathways through activation and repression of genes after exposure to stress. Plant genomes contain a large number of transcription factors. It has been estimated that the Arabidopsis genome codes for at least 1533 transcription factors, consisting of over 5.9% of its total estimated genes [28,29]. Based on the presence of conserved AP2-like domains, 173 proteins were identified as belonging to the AP2/ERF superfamily in the *Salix* genome in this analysis, compared to the previously reported 200 genes in *Populus* [9], 145 genes in Arabidopsis [30] and 147 in rice [6]. Eighty-eight genes were predicted to encode for the ERF subfamily, 57 genes for the DREB subfamily, 23 for the AP2 family, four genes for the RAV family and one for the Soloist family.

The availability of the complete genome sequence of *Populus* and Arabidopsis enabled a comparison of individual families and groups using strict criteria [6]. Several DREB groups (B1, B2, B5 and B6), ERF groups (A3 and 14) and the AP2 family contained more members in *Salix* than in Arabidopsis, and many DREB groups (A1, A2 and A6), ERF groups (B3 and B4), and the AP2 family contained less members in *Salix* compared to *Populus* (Table 1). There were also several different family members among *Salix* and other species, such as *Vitis vinifera* [10], *Medicago truncatula* [31], *Oryza sativa* [32] and *Triticum aestivum* [33]. The number of AP2 and ERF genes were very similar between *Salix* and *Populus* (24 AP2 in *Salix*, 26 AP2 in *Populus*; 88 ERF in *Salix* and 91 ERF in *Populus*), and were greater than that of Arabidopsis (17 AP2 and 65 ERF). The similar number of AP2/ERF genes in the *Populus* genome is likely to be a consequence of recent whole-genome duplication in

the *Populus* lineage. The number of *Salix* DREB genes (57) is similar to the number of *Arabidopsis* DREB genes (56), and less than that of *Populus* (77), suggesting that the DREB subfamily experienced another gene duplication event after the whole-genome duplication in the *Populus* lineage. Genes belonging to the RAV family are highly conserved among the species examined in this study, with six in *Arabidopsis*, five in *Populus* and rice, and four in *Salix*.

Previous studies have demonstrated that the AP2/ERF family genes are differentially expressed in different tissues. The expression pattern of three tissues – stems, leaves and Blossom buds – were examined using transcriptome data. For the 173 AP2/ERF genes, 19 had no transcript level in any tissue, and the other genes were more stem-specific (Fig. 3). AP2 genes are important during the process of determination of fully developed organs from meristematic tissue. However, with the exception of four AP2 genes, the expression level of AP2 genes were not restricted to a single organ, with many genes being mainly expressed in the Blossom bud, suggesting other functions of AP2 genes.

Although many studies have reported the results of transcriptome analysis of *Salix*, there have been no studies on gene expression profiles in *S. arbutifolia*. In this study, RNA-seq technology was used to generate the transcriptome and to examine the global gene expression profile of stems, leaves and Blossom buds in *S. arbutifolia*. This work facilitates the analysis of gene expression levels even in the absence of a genomic database. Moreover, it demonstrates that RNA-seq is a useful and effective tool for *de novo* transcriptome assembly. We generated mRNA libraries from stem, leaf and Blossom bud tissue of *Salix* to substantially increase available data on *Salix* mRNAs in order to build genomic resources for further investigation of candidate genes, especially stress resistance genes, from various metabolic pathways in *Salix*.

## Acknowledgements

This work was supported by grants from the National Natural Science Foundation of China (31400569), Collaborative Innovation Plan of Jiangsu Higher Education, the Fundamental Research Funds for the Central Non-profit Research Institution of CAF (RIF2013-11), Beijing Co-building Plan for Scientific Research and Postgraduate Education (2013K0140, 2013K0141, 2014K0151, 2014K0152), and We thank LetPub ([www.letpub.com](http://www.letpub.com)) for its assistance with the manuscript preparation. GR and JZ designed the project, JS and GR acquired the data, GR, FZ and CH analysis the data, GR wrote the paper.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.fob.2015.02.002>.

## References

- Berlin, S., Lagercrantz, U., von Arnold, S., Ost, T. and Ronnberg-Wastljung, A.C. (2010) High-density linkage mapping and evolution of paralogs and orthologs in *Salix* and *Populus*. *BMC Genomics* 11, 129.
- Dimova, D.K., Stevaux, O., Frolov, M.V. and Dyson, N.J. (2003) Cell cycle-dependent and cell cycle-independent control of transcription by the *Drosophila* E2F/RB pathway. *Genes Dev.* 17, 2308–2320.
- Sharma, M.K., Kumar, R., Solanke, A.U., Sharma, R., Tyagi, A.K. and Sharma, A.K. (2010) Identification, phylogeny, and transcript profiling of ERF family genes during development and abiotic stress treatments in tomato. *Mol. Genet. Genomics* 284, 455–475.
- Mizoi, J., Shinozaki, K. and Yamaguchi-Shinozaki, K. (1819) AP2/ERF family transcription factors in plant abiotic stress responses. *Biochim. Biophys. Acta* 2012, 86–96.
- Song, X., Li, Y. and Hou, X. (2013) Genome-wide analysis of the AP2/ERF transcription factor superfamily in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *BMC Genomics* 14, 573.
- Nakano, T., Suzuki, K., Fujimura, T. and Shinshi, H. (2006) Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. *Plant Physiol.* 140, 411–432.
- Liu, Z., Kong, L., Zhang, M., Lv, Y., Liu, Y., Zou, M., Lu, G., Cao, J. and Yu, X. (2013) Genome-wide identification, phylogeny, evolution and expression patterns of AP2/ERF genes and cytokinin response factors in *Brassica rapa* ssp. *pekinensis*. *PLoS One* 8, e83444.
- Sakuma, Y., Liu, Q., Dubouzet, J.G., Abe, H., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2002) 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.* 290, 998–1009.
- Zhuang, J., Cai, B., Peng, R.H., Zhu, B., Jin, X.F., Xue, Y., Gao, F., Fu, X.Y., Tian, Y.S., Zhao, W., Qiao, Y.S., Zhang, Z., Xiong, A.S. and Yao, Q.H. (2008) Genome-wide analysis of the AP2/ERF gene family in *Populus trichocarpa*. *Biochem. Biophys. Res. Commun.* 371, 468–474.
- Licausi, F., Giorgi, F.M., Zenoni, S., Osti, F., Pezzotti, M. and Perata, P. (2010) Genomic and transcriptomic analysis of the AP2/ERF superfamily in *Vitis vinifera*. *BMC Genomics* 11, 719.
- Hu, L. and Liu, S. (2011) Genome-wide identification and phylogenetic analysis of the ERF gene family in cucumbers. *Genet. Mol. Biol.* 34, 624–633.
- Chen, J.Q., Dong, Y., Wang, Y.J., Liu, Q., Zhang, J.S. and Chen, S.Y. (2003) An AP2/EREBP-type transcription-factor gene from rice is cold-inducible and encodes a nuclear-localized protein. *Theor. Appl. Genet.* 107, 972–979.
- Nakano, T., Fujisawa, M., Shima, Y. and Ito, Y. (2014) The AP2/ERF transcription factor SIERF52 functions in flower pedicel abscission in tomato. *J. Exp. Bot.* 65, 3111–3119.
- Yan, H.W., Hong, L., Zhou, Y.Q., Jiang, H.Y., Zhu, S.W., Fan, J. and Cheng, B.J. (2013) A genome-wide analysis of the ERF gene family in sorghum. *Genet. Mol. Res.* 12, 2038–2055.
- Wang, F., Qiu, N., Ding, Q., Li, J., Zhang, Y., Li, H. and Gao, J. (2014) Genome-wide identification and analysis of the growth-regulating factor family in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *BMC Genomics* 15, 807.
- Zhang, G., Chen, M., Chen, X., Xu, Z., Guan, S., Li, L.C., Li, A., Guo, J., Mao, L. and Ma, Y. (2008) Phylogeny, gene structures, and expression patterns of the ERF gene family in soybean (*Glycine max* L.). *J. Exp. Bot.* 59, 4095–4107.
- Kadis, I. (2005) *Chosenia*: an amazing tree of Northeast Asia. *Arnoldia* 63, 8–17.
- Huang, X. and Madan, A. (1999) CAP3: a DNA sequence assembly program. *Genome Res.* 9, 868–877.
- Bailey, T.L. and Elkan, C. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 2, 28–36.
- Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., Macmanes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., Leduc, R.D., Friedman, N. and Regev, A. (2013) *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* 8, 1494–1512.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729.
- Wagner, G.P., Kin, K. and Lynch, V.J. (2012) Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.* 131, 281–285.
- Licausi, F., Ohme-Takagi, M. and Perata, P. (2013) APETALA2/Ethylene Responsive Factor (AP2/ERF) transcription factors: mediators of stress responses and developmental programs. *New Phytol.* 199, 639–649.
- Shi, X., Gupta, S. and Rashotte, A.M. (2014) Characterization of two tomato AP2/ERF genes, SICRF1 and SICRF2 in hormone and stress responses. *Plant Cell Rep.* 33, 35–45.
- Xie, X.L., Shen, S.L., Yin, X.R., Xu, Q., Sun, C.D., Grierson, D., Ferguson, I. and Chen, K.S. (2014) Isolation, classification and transcription profiles of the AP2/ERF transcription factor superfamily in citrus. *Mol. Biol. Rep.* 41, 4261–4271.
- Karp, A., Hanley, S.J., Trybush, S.O., Macalpine, W., Pei, M. and Shield, I. (2011) Genetic improvement of willow for bioenergy and biofuels. *J. Integr. Plant Biol.* 53, 151–165.
- Miao, H., Qin, Y., da Silva, J.A., Ye, Z. and Hu, G. (2013) Identification of differentially expressed genes in pistils from self-incompatible *Citrus reticulata* by suppression subtractive hybridization. *Mol. Biol. Rep.* 40, 159–169.
- Lezhneva, L. and Meurer, J. (2004) The nuclear factor HCF145 affects chloroplast *psaA-psaB-rps14* transcript abundance in *Arabidopsis thaliana*. *Plant J.* 38, 740–753.
- Rao, G., Sui, J., Zeng, Y., He, C., Duan, A. and Zhang, J. (2014) *De novo* transcriptome and small RNA analysis of two Chinese willow cultivars reveals stress response genes in *Salix matsudana*. *PLoS One* 9, e109122.
- Feng, J.X., Liu, D., Pan, Y., Gong, W., Ma, L.G., Luo, J.C., Deng, X.W. and Zhu, Y.X. (2005) An annotation update via cDNA sequence analysis and comprehensive profiling of developmental, hormonal or environmental responsiveness of the *Arabidopsis* AP2/EREBP transcription factor gene family. *Plant Mol. Biol.* 59, 853–868.
- Pre, M., Atallah, M., Champion, A., De Vos, M., Pieterse, C.M. and Memelink, J. (2008) The AP2/ERF domain transcription factor ORA59 integrates jasmonic acid and ethylene signals in plant defense. *Plant Physiol.* 147, 1347–1357.
- Sharoni, A.M., Nuruzzaman, M., Satoh, K., Shimizu, T., Kondoh, H., Sasaya, T., Choi, I.R., Omura, T. and Kikuchi, S. (2011) Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. *Plant Cell Physiol.* 52, 344–360.
- Zhuang, J., Chen, J.M., Yao, Q.H., Xiong, F., Sun, C.C., Zhou, X.R., Zhang, J. and Xiong, A.S. (2011) Discovery and expression profile analysis of AP2/ERF family genes from *Triticum aestivum*. *Mol. Biol. Rep.* 38, 745–753.