biologia plantarum
an international journal for experimental botany

BIOLOGIA PLANTARUM (2023) 67: 88-99 DOI: 10.32615/bp.2023.003

Identification of *AP2/ERF* gene family of *Salicaceae* and their response to salt stress, abscisic acid, and gibberellic acid in *Populus euphratica* seeds

X.L. HAN[1], C. QIU[1], J.H. SUN[1], J.D. XU[3], X. ZHANG[1], J.T. ZHAI[1], S.H. ZHANG[1], Z.H. WU[2,*], and Z.J. LI[1,*]

[1] *College of Life Sciences, Tarim University/Key Laboratory of Protection and Utilization of Biological Resources in Tarim Basin, Xinjiang Production & Construction Corps/Research Center of Populus euphratica, Alar 843300, Xinjiang, P.R. China*
[2] *College of Chemistry and Life Sciences, Zhejiang Normal University, Jinhua 321004, Zhejiang, P.R. China*
[3] *Hubei Provincial Key Laboratory for Protection and Application of Special Plant Germplasm in Wuling Area of China, College of Life Sciences, South Central University for Nationalities, Wuhan 430040, P.R. China*

[*]Corresponding authors: E-mails: lizhijun0202@126.com; zhwu2022@126.com

Abstract

Populus euphratica belongs to *Salicaceae* family and grows in extreme desert environments. At present, the identification of the *AP2/ERF* gene family of transcription factors in *Salicaceae* is rare, and the role of the *AP2/ERF* gene family in *P. euphratica* under salt stress and exogenous hormones has not been reported. In this study, 197, 210, 231, 192, and 147 *AP2/ERF* genes were identified in *P. euphratica*, *Populus trichocarpa*, *Populus deltoides*, *Salix sinopurpurea*, and *Arabidopsis thaliana*, respectively. The 197 *AP2/ERF* gene family members of *P. euphratica* were divided into five subfamilies, namely, AP2 (35), RAV (5), ERF (96), DREB (65), and Soloist (1), by sequence alignment and phylogenetic analysis. In addition, these genes were scattered across 19 chromosomes. The detection of 10 motifs in the *P. euphratica AP2/ERF* gene family revealed that motif-8 and motif-9 only appeared in the ERF subfamily and DREB subfamily, respectively. Transcriptome data showed that *PeAP2/ERF* genes had different expression patterns under salt stress, abscisic acid (ABA) and gibberellic acid (GA_3) treatments, suggesting that the genes *PeERF002*, *PeERF037*, *PeERF082*, *PeERF090*, and *PeAP2-14* may play important roles under salt stress and exogenous hormone treatments. This study provides a reference for the functional study of the *PeAP2/ERF* gene, and it also lays a foundation for the breeding strategy to improve the salt tolerance of *P. euphratica*.

Keywords: abscisic acid, *AP2/ERF*, gene expression, gibberellic acid, *Populus euphratica*, *Salicaceae*, salt stress.

Introduction

As forest trees with important ecological value, most species of *Populus* and *Salix* (*Salicaceae*) are distributed in the northern hemisphere ranging from the cold zone to the temperate zone (Taylor 2002, Tuskan *et al*. 2006). Compared with other poplar and willow species, *Populus euphratica*, known as the 'hero of the desert', has stronger

Received 31 March 2022, *last revision* 21 February 2023, *accepted* 6 March 2023.
Abbreviations: aa - amino acid; ABA - abscisic acid; AP2 - APETALA2; DREB - dehydration response element binding; ERF - ethylene-responsive transcription factor; GA_3 - gibberellic acid; RAV - Related to ABI3/VP; TF - transcription factor.
Acknowledgements: The authors are grateful to Tarim University and Zhejiang Normal University for technical service and support, as well as the key Laboratory of Protection and Utilization of Biological Resources in Tarim Basin of Xinjiang Production and Construction Corps for providing equipment. This work was supported by the Regional Innovation Guidance Program Project of Xinjiang Production and Construction Corps (2021BB010), Innovative Team Building Plan for Key Areas of Xinjiang Production and Construction Corps (2018CB003), and Xinjiang Production and Construction Corps Key Laboratory of Protection and Utilization of Biological Resources in Tarim Basin (BRZD2003).
Conflicts of interest: The authors declare that they have no conflict of interest.

tolerance to salt-alkali, drought, and high temperature in desert environment. It is not only a rare tree species but also an excellent forest tree for wind-proof and sand-fixing, as well as soil and water conservation (Zhai *et al*. 2020). Therefore, *P. euphratica* has become an ideal species for studying the adaptation of woody plants to extreme environments (especially salinity and drought stress).

Seed germination is the first stage to start the plant life cycle, but it is susceptible to environmental conditions (Song *et al*. 2005). Salt can destroy the structure of organelles, enzymes, and other macromolecules through the toxic effects of Na^+ and Cl^- on embryo vitality, thereby inhibiting seed germination (Jahromi *et al*. 2008, Daszkowska-Golec 2011). As the NaCl concentration increases, seed germination is significantly inhibited (Sun *et al*. 2000). The relative germination rates of *P. euphratica* in 0.4% NaCl solution exceed 80%, whereas the relative germination rates in 2.4% NaCl solution are close to zero (Zhang *et al*. 2019).

When a plant responds to biotic or abiotic stress, it combines signal transduction with *cis*-acting elements to initiate gene expression at the molecular level. The gene product produced by the final transcription can participate in the regulation of plant response to adversity stress. The AP2/ERF superfamily, as one of the plant-specific transcription factor (TF) families, has been found to be widely involved in the regulation of plant abiotic stress response and plant hormones (abscisic acid, ethylene, jasmonic acid, and salicylic acid) (Agarwal *et al*. 2007, Mizoi *et al*. 2012). The AP2/ERF superfamily is defined by the existence of the AP2/ERF domain, which consists of about 60 - 70 amino acids (aa) involved in DNA binding (Sakuma *et al*. 2002), including AP2 (APETALA2), RAV (Related to ABI3/VP), DREB (dehydration response element binding) and ethylene-responsive transcription factor (ERF) subfamily. The AP2 family plays an important role in the regulation of plant growth and development, such as the recognition of leaf epidermal cells, and the development of flowers and ovules (Moose and Sisco 1996, Aukerman and Sakai 2003, Dinh *et al*. 2012). The RAV family is involved in the regulation of plant development and response to various stresses (Sohn *et al*. 2006, Li *et al*. 2011), Studies have shown that the RAV gene family can cope with high salt and low temperature stress, and this gene family has expanded from six members of the *Arabidopsis* genome to nine members of the salt mustard genome (Fu *et al*. 2014). The DREB subfamily contains an AP2 domain, which plays an important role in the response of plants to abiotic stress. For example, the overexpression of *LsDREB2A* increases salt tolerance in lettuce (Kudo *et al*. 2014). The ERF subfamily recognizes GCC *cis*-acting elements, and many members of this subfamily have been identified and discovered in many processes, such as metabolic regulation (Van der Fits and Memelink 2000, Dubouzet *et al*. 2003, Yu *et al*. 2012, Deng *et al*. 2017), response to biotic and abiotic stress (Stockinger *et al*. 1997, Yao *et al*. 2017), hormone signalling (MM SM-B 2015) and plant development. The conserved 14th alanine and 19th aspartic acid of the ERF subfamily are essential for the function of this subfamily.

At present, the whole genome identification and analysis of the *AP2/ERF* gene family has been extensively studied in *Arabidopsis* (Nakano *et al*. 2006), pear (Li *et al*. 2018), sesame (Dossa *et al*. 2016), rice (Rashid *et al*. 2012), pepper (Jin *et al*. 2018), cauliflower (Li *et al*. 2017), and Chinese cabbage (Song *et al*. 2013). However, studies on the identification and characteristics of the *AP2/ERF* gene family in *Salicaceae* plants, such as *P. euphratica*, are few. Here, genome data of *P. euphratica* (Zhang *et al*. 2020), *P. trichocarpa* (Tuskan *et al*. 2006), *S. sinopurpurea* (Guo *et al*. 2021), *P. deltoides* and *A. thaliana* (as reference species, https://phytozome.jgi.doe.gov/) were used to identify and characterize the *AP2/ERF* gene family among these species. Furthermore, the chromosome location and gene structure of the *AP2/ERF* genes in *P. trichocarpa* were analysed, and phylogenetic relationships were compared. The response of the *AP2/ERF* genes to salt stress and exogenous hormones (ABA and GA_3) during seed germination was further analysed in *P. euphratica* at the transcriptome level. Our research will provide an important reference for the improvement of *P. euphratica* salt-tolerant strains.

Materials and methods

Physiological and morphological changes during seed germination: The seed samples were collected in the Aksu Awat County (Xinjiang, China) and stored at 4°C. In the salt stress experimental treatment, vigorous seeds were imbibed in distilled water (control) and 0.1, 0.2, 0.3, and 0.4 M NaCl for germination. In the hormone treatment, seeds were imbibed in distilled water with an equal amount of methanol (control), 100 µM abscisic acid (ABA) solution with methanol, distilled water with an equal amount of ethanol and 300 mg dm^{-3} gibberellic acid (GA_3) solution with ethanol and then germinated on wet filter paper in 9 cm diameter Petri dishes in a plant growth incubator (16 h light/8 h dark photoperiod, 30/25°C light/dark temperature). The germination rate was measured using the Chinese national standard test (GB2772-1999). Each sample contained 100 seeds and had three replicates. The germinating seeds were scanned and photographed using a stereo microscope (*Nikon SMZ1500 Z1500*, Tokyo, Japan) to record their morphology. After 4 d, the seedlings were collected and stored in the refrigerator at -80°C for further analyses.

Identification of *AP2/ERF* gene family in *Salicaceae*: Two different methods were used to identify the *AP2/ERF* domain sequences in *P. euphratica*, *P. trichocarpa*, *P. deltoides*, and *S. sinopurpurea*: *1*) Using the *Pfam* protein family database (https://www.ebi.ac.uk/interpro/), the hidden Markov model (HMM) file of the corresponding AP2 domain (PF00847) was used for HMM-search scanning; *2*) Homologous search of the AP2/ERF transcription factor conserved domain of

Arabidopsis was performed against the *P. euphratica*, *P. trichocarpa*, *P. deltoides*, and *S. sinopurpurea* genome database with *BLAST* (*BLASTN*, *BLASTP*) (http://itak.feilab.net/cgi-bin/itak/index.cgi). The threshold of similar AP2 domain value (E value) was set to $1e^{-5}$, and the sequences with less than 100 amino acid residues were deleted. *Smart* and *Pfam* were used to confirm the candidate *AP2/ERF* gene members containing complete AP2 conserved domains.

Classification of *AP2/ERF* genes: Neighbor-joining (NJ) method with pairwise deletion option was used for the construction of phylogenetic trees for all four *Salicaceae* using domain peptide sequences of *AP2/ERF*. The reliability of the constructed trees was assessed by bootstraps with 5 000 replicates. The conserved motifs of *AP2/ERF* were predicted using the standalone version of motif-based sequence analysis tool (*MEME*) (http://meme-suite.org/) with default parameters, number of motifs set at 10, optimum width of 5 - 200 amino acids and any number of repetitions of a motif. Gene structure prediction was made using online Gene Structure Display Server (http://gsds.gao-lab.org/). *ExPASy* (http://ExPASy.org/) was used to obtain protein structural features, including isoelectric point, molecular mass, and gravy score.

Phylogenetic analysis and classification of the *PeAP2/ERF* gene family: Using the amino acid sequence of the AP2 conserved domain of the *PeAP2/ERF* transcription factor as the material, the *Muscle* program was used for multiple sequence alignment, and *MEGA7.0* software was used to construct the NJ tree. The parameters were set as follows: Poisson model, gap (GA3ps): pairwise deletion, verification parameter (test of phylogeny): bootstrap (1 000).

Identification of orthologous *PeAP2/ERF* genes, their distribution and duplication: Generic feature format files for the genomes of four *Salicaceae* were used to mark the location of each gene on their physical maps. The distribution of *AP2/ERF* genes was visualized using *MapChar*. Orthologous genes with respect to *P. euphratica* were predicted using the *JCVI* approach with e-value threshold of $1e^{-10}$. The *PeAP2/ERF* genes were used as query against the *AP2/ERF* genes of *P. trichocarpa*, *P. deltoides*, *S. sinopurpurea*, and *A. thaliana*. The *AP2/ERF* genes in each *Salicaceae* species were searched for duplication events at an e-value was $\leq 1e^{-1}$, identity of comparable regions of $\geq 70\%$ and distance < 200 kb.

Gene expression analysis: Gene expression patterns of the *AP2/ERF* genes in *P. euphratica* were studied using RNA-seq data generated from seed germination stages with three biological replicates for each sample. *HISAT2* (*version 2.0.4*) software was used to compare the obtained high-quality clean reads to the reference genome of *P. euphratica*. *String Tie* software was used to perform quantitative expression analysis of each sample gene, and FPKM (fragments per kilobase per million) was used to quantify the gene expression levels. Moreover, DESeq2 was used to identify differentially expressed genes (DEGs) with *P*-value ≤ 0.05 and absolute log2fold change ≥ 2. DEGs were analyzed according to the four treatments of NaCl *versus* the control (NaCl *vs.* CK), ABA *versus* the control (ABA *vs.* CK), GA_3 *versus* the control (GA_3 *vs.*CK), and ABA *versus* GA_3.

Validation of gene expression profiles using RT-qPCR: To further verify the accuracy of the sequencing data, we selected four DEGs treated with salt stress, ABA, and GA_3 for RT-qPCR experimental verification. Reverse transcription of the extracted RNA samples was performed using the reverse transcription kit of *Beijing Novazin Biotech* (Beijing, China) to obtain cDNA. The candidate gene-specific fluorescent quantitative RT-qPCR primers were designed and obtained on *NCBI Prime-BLAST* (https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi) (Table 5 Suppl.). *Applied Biosystems 7500* real-time PCR system was used and the $2^{-\Delta\Delta CT}$ method for calculation of gene expression. RT-qPCR experiments for each gene had three replicates. The housekeeping *Actin* gene of *P. euphratica* was used as the internal reference gene.

Results

To identify the *AP2/ERF* TF gene family of *Salicaceae* species, all possible *AP2/ERF* genes were excavated from the *P. euphratica*, *P. trichocarpa*, *P. deltoides*, *S. sinopurpurea*, and *A. thaliana* genomes using two methods. The results showed that the number of *AP2/ERF* genes in the *Salicaceae* family species was similar (197 of *P. euphratica*, 210 of *P. trichocarpa*, 231 of *P. deltoides*, and 192 of *S. sinopurpurea*), and significantly higher than in *A. thaliana* (147) (Table 1). Among the five subfamilies of the *AP2/ERF* gene family, the ERF subtype was the most numerous, followed by the DREB superfamily. In *A. thaliana*, the six members of the RAV subfamily were classified as *AP2/ERF* transcription factors because they contained both the B3 and AP2 domains (Fu *et al.* 2014). Unlike rice and potato, which have two to three members of the RAV subfamily, we found four to five members of the RAV superfamily in *Salicaceae*. Additional information on the isoelectric point, molecular mass, and amino acid sequence of proteins of *P. trichocarpa*, *P. deltoides*, *S. sinopurpurea*, and *A. thaliana* is provided in Tables 1 - 4 Suppl.

To explore the phylogenetic relationship of *AP2/ERF* between *P. euphratica* and other plants, phylogenetic tree was performed. Most members of the AP2 subfamily of *P. euphratica* were clustered with the three other *Salicaceae* plants (Fig. 1). The phylogenetic tree of the RAV subfamily showed that, except for *PeRAV3*, other four *PeRAV* members were more closely to the two other *Populus* species than to *A. thaliana* and *S. sinopurpurea* (Fig. 1 Suppl.).

According to the distribution principle of DREB subfamily members in *A. thaliana* (Nakano *et al.* 2006), a phylogenetic tree was constructed for DREB subfamily

Table 1. Distribution of *AP2/ERF* superfamily members in the whole genome of five species.

	Populus euphratica	*Populus trichocarpa*	*Populus deltoides*	*Salix sinopurpurea*	*Arabidopsis thaliana*
AP2	30	32	29	26	18
DREB	65	69	69	53	57
ERF	96	102	125	68	65
RAV	5	5	5	5	6
Soloist	1	1	3	3	1
Total *AP2/ERF* gene family	197	210	231	192	147
Total genes in genome	36 603	63 498	57 249	57 462	35 386
AP2/ERF transcription factor genes [%]	0.54	0.33	0.40	0.33	0.42

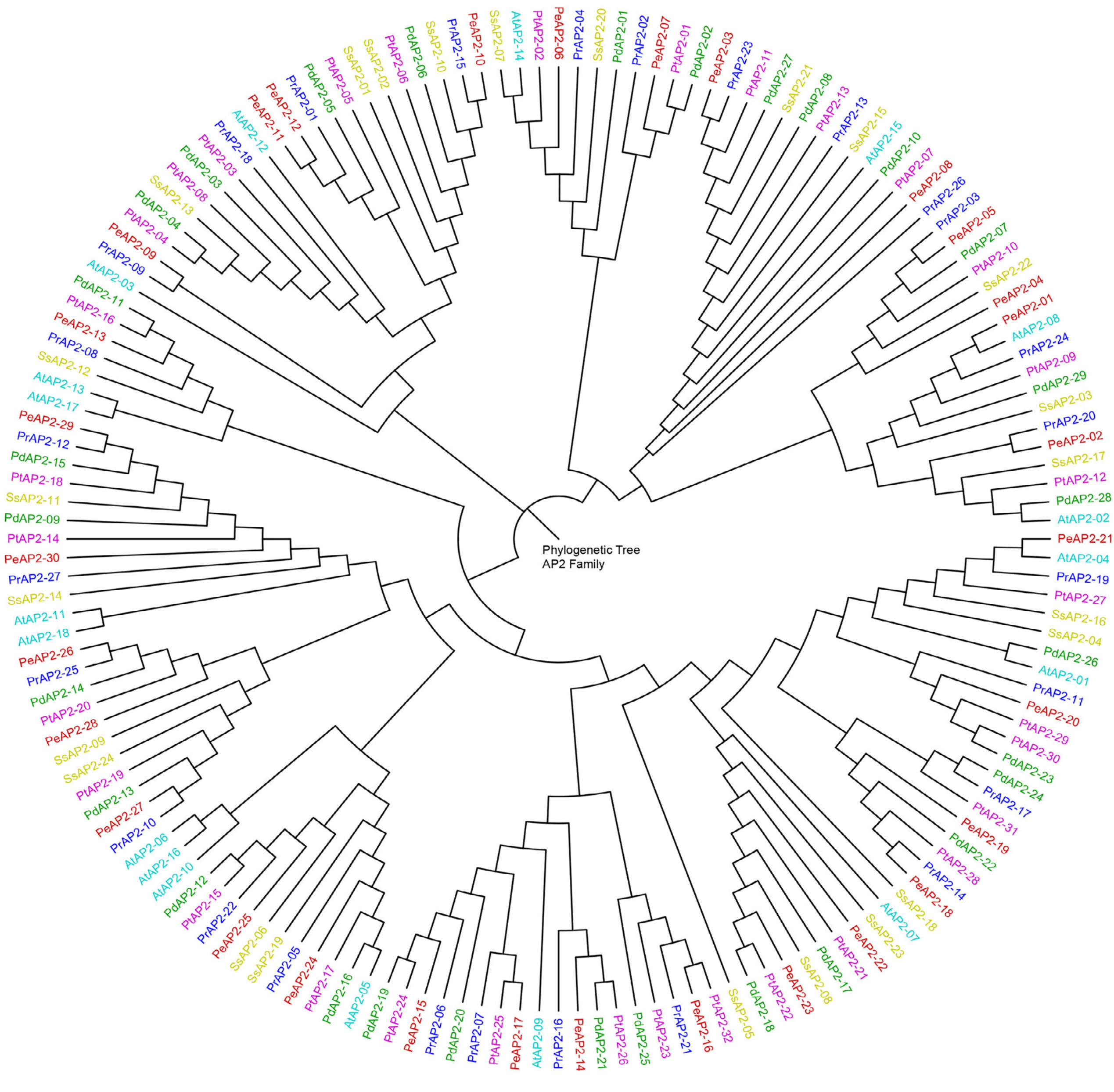

Fig. 1. Phylogenetic analysis of the AP2 subfamily of *Populus* species.

members of the five species. Ten motifs were detected in the DREB subfamily of five species (group a-l, Fig. 2 Suppl.). All members contained motif-1, and the genes clustered together contained similar motifs. However, some unique motifs were found in *P. euphratica*. For example, *PeDREB6* and *PeDREB7* in panel k had two motif-8s but different positions from the two motif-8s in *SsDREB54* (group k, Fig. 2 Suppl.). In addition, *PeDREB43* in group a lacked motif-5, and *PeDREB13* in group h lacked motif-2 and motif-3 compared with other species.

PeERF genes were divided into 14 groups (groups a - n, Fig. 3 Suppl.). A total of 10 motifs of the ERF subfamily were detected in all plants, and all members included motif-2 and motif-3. Among them, *PeERF077* in group f only had motif-5, *PeERF016* in group k only had motif-1 and *PeERF004* in group b lacked motif-1. In general, *AP2/ERF* genes between *P. euphratica* and *P. trichocarpa* had a closer relationship than that between *P. euphratica* and other plants.

To further analysis of the phylogenetic relationship of *PeAP2/ERF* members with *A. thaliana*, we constructed the phylogenetic tree of 197 *PeAP2/ERFs* and 147 *AP2/ERFs* in *A. thaliana* (Fig. 2). *PeAP2/ERF* contained 30 AP2 subfamily members, 65 DREB subfamily members, 96 ERF subfamily members, 5 RAV subfamily members and 1 Soloist subfamily member (Table 1). In addition, the DREB and ERF subfamilies were divided into four groups (I - IV) and six groups (V - X), respectively. The DREB subfamily had the most gene members (32) in group III, whereas the ERF subfamily had the most gene members (26) in group IX.

To understand the distribution of *AP2/ERF* gene family members, *PeAP2/ERF* genes were mapped to the chromosomes of *P. euphratica*. We found that chromosome 1 had the most *PeAP2/ERF* genes (27), whereas chromosome 19 had the least *PeAP2/ERF* genes (2). In addition, ERF subfamily genes (16) were mainly distributed on chromosome 1; AP2 subfamily genes were mainly distributed on chromosomes 5 and 10; DREB subfamily genes were distributed on chromosomes 1 and 5 (Fig. 3). Amongst these *PeAP2/ERF* genes, *PeERF-VIIIb-151* encodes the smallest protein with 112 amino acids, and *PeERF-Vb-104* encodes the largest protein with 990 amino acids. The average protein molecular mass was 34 306.12 Da, and the average isoelectric point was 7.073 (Table 2).

To further analyze the phylogenetic relationship and explore the potential functions of the *AP2/ERF* genes, the conserved domains in the amino acid sequence of the *AP2/ERF* genes were analyzed (Fig. 4 Suppl.). The results showed that motif-l and motif-2 were widely present in ERF and DREB subfamily members, and motif-l and motif-7 were widely present in the RAV and AP2 subfamily. In addition, motif-3, motif-4 and motif-10 were specific to the AP2 subfamily; motif-8 only appeared in the ERF subfamily; and motif-9 only appeared in the DREB subfamily. These results indicated that these motifs were related to the unique functions of the subfamily. The distribution of conserved domains between different subfamilies varied, suggesting that the phylogenetic classification of the *PeAP2/ERF* gene family was reliable.

To explore the distribution of introns and exons of *PeAP2/ERF* gene family members, the gene structure of *AP2/ERF* gene family members was analyzed. The results showed that 47.9% (46) and 70.7% (46) genes in the ERF and DREB subfamilies of *P. euphratica* did not contain introns (Fig. 4 Suppl.). Among them, 38.5% (37) of the ERF subfamily members had only one intron, and *PeERF009* had the most introns (21). In the DREB subfamily, 25% (16) members contained 1 intron, and *PeDREB25* and *PeDREB65* had more introns (4). A member of the RAV subfamily (*PeRAV5*) did not contain introns, whereas genes in the AP2 subfamily and Soloist subfamily contained a large number of introns, with *PeAP2-27* having the most introns (10). Moreover, *Pe-soloist* contained 6 introns. Gene structure analysis showed that the intron and exon structures of the *PeAP2/ERF* gene family were relatively diversified, indicating that the evolutionary trend may lead to the diversification of gene functions.

To explore the effects of salt and plant hormones on seed germination, we performed seed germination under NaCl treatment with different concentrations and exogenous hormones (100 µM ABA and 300 mg dm^{-3} GA_3). As the salt concentration increased, the germination indexes of seeds gradually decreased (Fig. 4*A-D*), and the seedling morphology gradually changed (Fig. 4*I*). When the NaCl concentration was 0.3 M, these indexes began to show significant differences. Therefore, 0.3 M NaCl was used as the inflection point for *P. euphratica* seeds to resist salt stress in this study. For ABA and GA_3 treatment, we found that the seed germination index after ABA treatment was lower, while the seed germination index after GA_3 treatment was higher compared with the control (Fig. 4*E-H*). These results suggested that GA_3 can promote the seed germination of *P. euphratica*, whereas ABA can inhibit the seed germination of *P. euphratica*.

Gene expression patterns can provide important indicators for gene function. DEGs of 197 *AP2/ERF* gene family members in *P. euphratica* seedlings under the treatment of NaCl, ABA, and GA_3 with their corresponding controls were analysed. On the basis of the threshold of |fold change| > 2 and *P*-value < 0.05, we identified the 79 DEGs (38 up-regulated and 41 down-regulated) in GA_3 treatment, 45 DEGs (0 up-regulated and 45 down-regulated) in ABA treatment, 93 DEGs (40 up-regulated and 53 down-regulated) in NaCl treatment with their corresponding controls and 48 DEGs (13 up-regulated and 35 down-regulated) in ABA treatment compared with GA_3 (Fig. 5). We also analysed the expression patterns of 10 *AP2/ERF* gene family members that were highly homologous to the reported functional genes in *Arabidopsis*, such as *PeDREB57*, *PeDREB59*, *PeAP2-14*, *PeERF002*, *PeERF037*, *PeERF044*, *PeERF059*, *PeERF082*, *PeERF087*, and *PeERF090*. Amongst them, *PeDREB57*, *PeDREB59*, *PeERF002*, *PeERF037*, and *PeERF087* were up-regulated in ABA *vs.* CK; *PeAP2-14*, and *PeERF090* were up-regulated in GA_3 *vs.* CK and

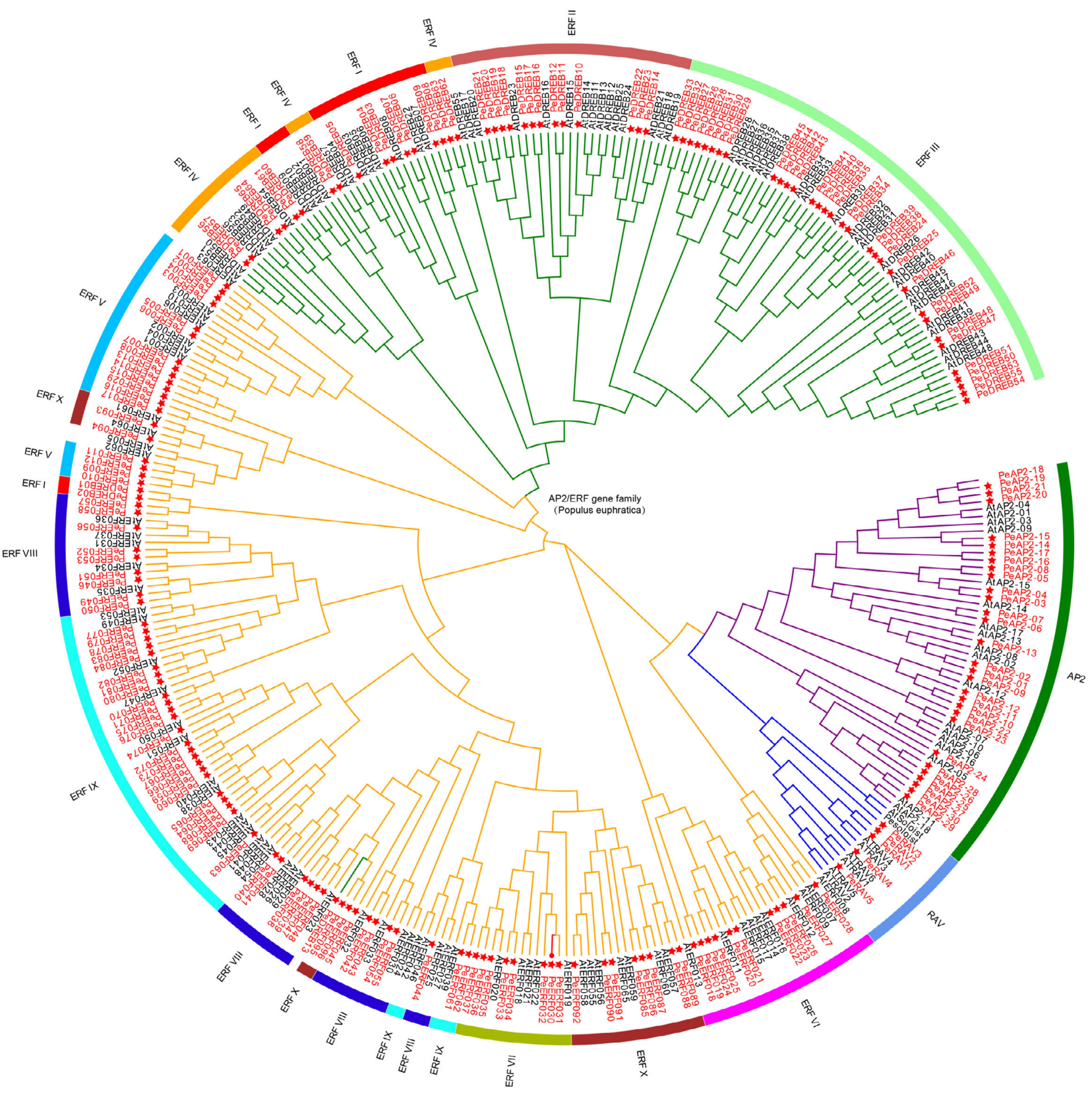

Fig. 2. Phylogenetic analysis of *AP2/ERF* gene families from *P. euphratica* and *A. thaliana.*

ABA *vs.* GA_3; *PeERF044* was up-regulated in ABA *vs.* CK and NaCl *vs.* CK; and *PeERF059* and *PeERF082* were up-regulated in ABA *vs.* GA_3.

To verify the accuracy of the data, the RT-qPCR analysis of three genes (*PeERF037*, *PeERF082*, and *PeERF002*) was performed (Fig. 5 Suppl.). The results showed that the expression patterns of these genes under salt stress, ABA, and GA_3 treatments were consistent with the results of RNA-Seq data, indicating that our RNA-Seq data were relatively accurate and the conclusions were relatively reliable.

Discussion

AP2/ERF proteins are widely involved in many processes, such as plant growth, flower development, fruit development, seed development and biotic/abiotic stress response (Faraji *et al*. 2020). For example, in chickpea, *Ca_02170* and *Ca_16631*, members of the *AP2/ERF* gene family, were potential candidate genes in response to *Fusarium* sp. wilt and sterility mosaic disease stress (Agarwal *et al*. 2016). In the halophyte *Halostachys caspica*, *HcTOE3*, a member of the *AP2/ERF* gene family,

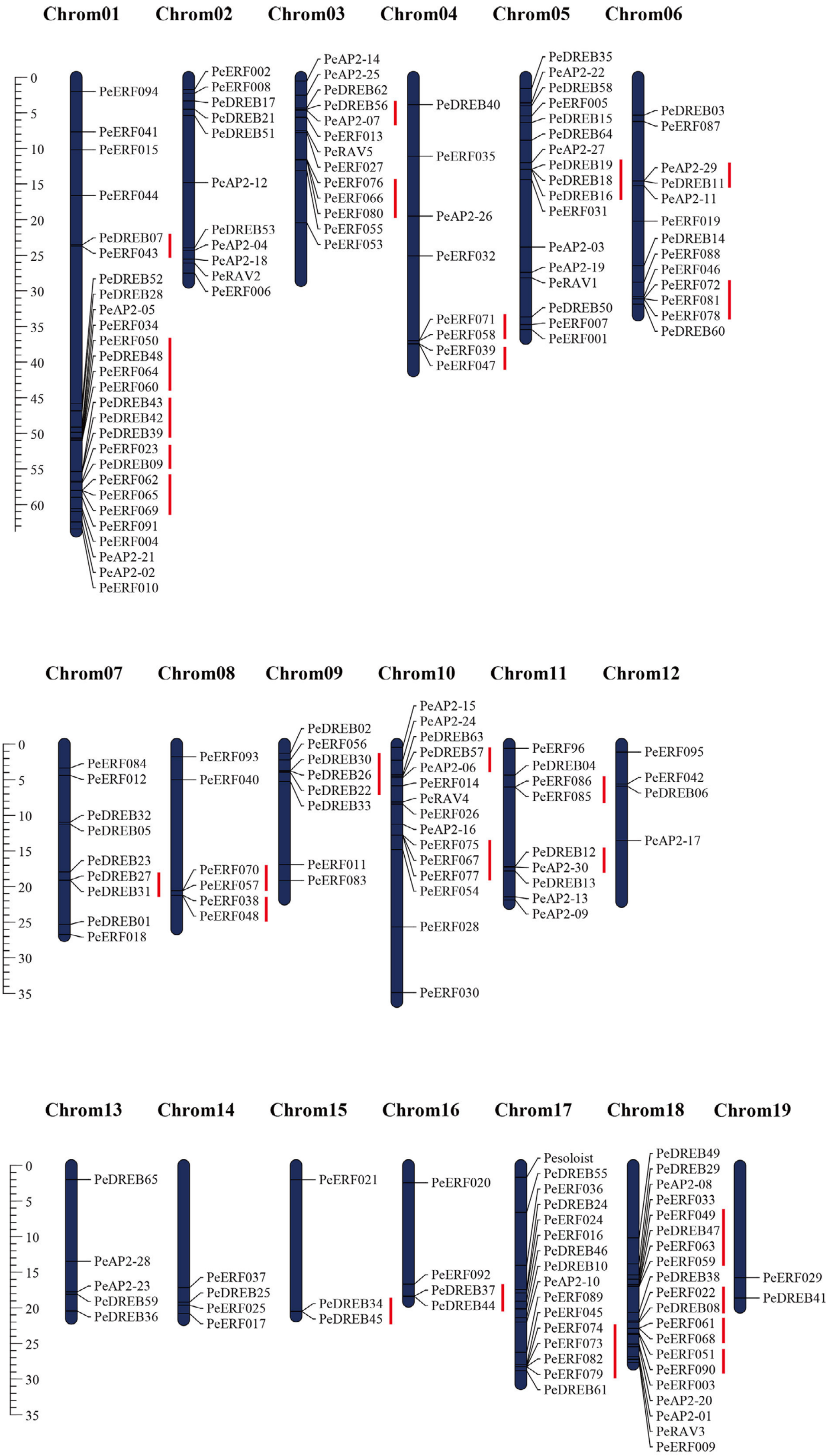

Fig. 3. Chromosomal distribution of *PeAP2/ERF* genes. *The red lines* indicate tandem duplicated *AP2/ERF* genes.

Table 2. Physical and chemical properties of *PeAP2/ERF* protein sequence.

Subfamily	Quantity	Numbering	Length	Mr	pI
AP2	30	*AP2-001-030*	339 - 720	38 575.5 - 78 767.1	5.51 - 9.40
DREB	65	*PeERF-031-095*	138 - 843	15 522.7 - 94 539.9	4.56 - 11.87
ERF	96	*PeERF-096-191*	112 - 990	12 728.1 - 110 611.3	4.29 - 12.23
RAV	5	*PeRAV-192-196*	293 - 547	33 469.9 - 61 042.6	8.18 - 9.69
Soloist	1	*Pe-soloist-197*	232.00	26 101.40	9.750
Average			309.26	34 306.12	7.073

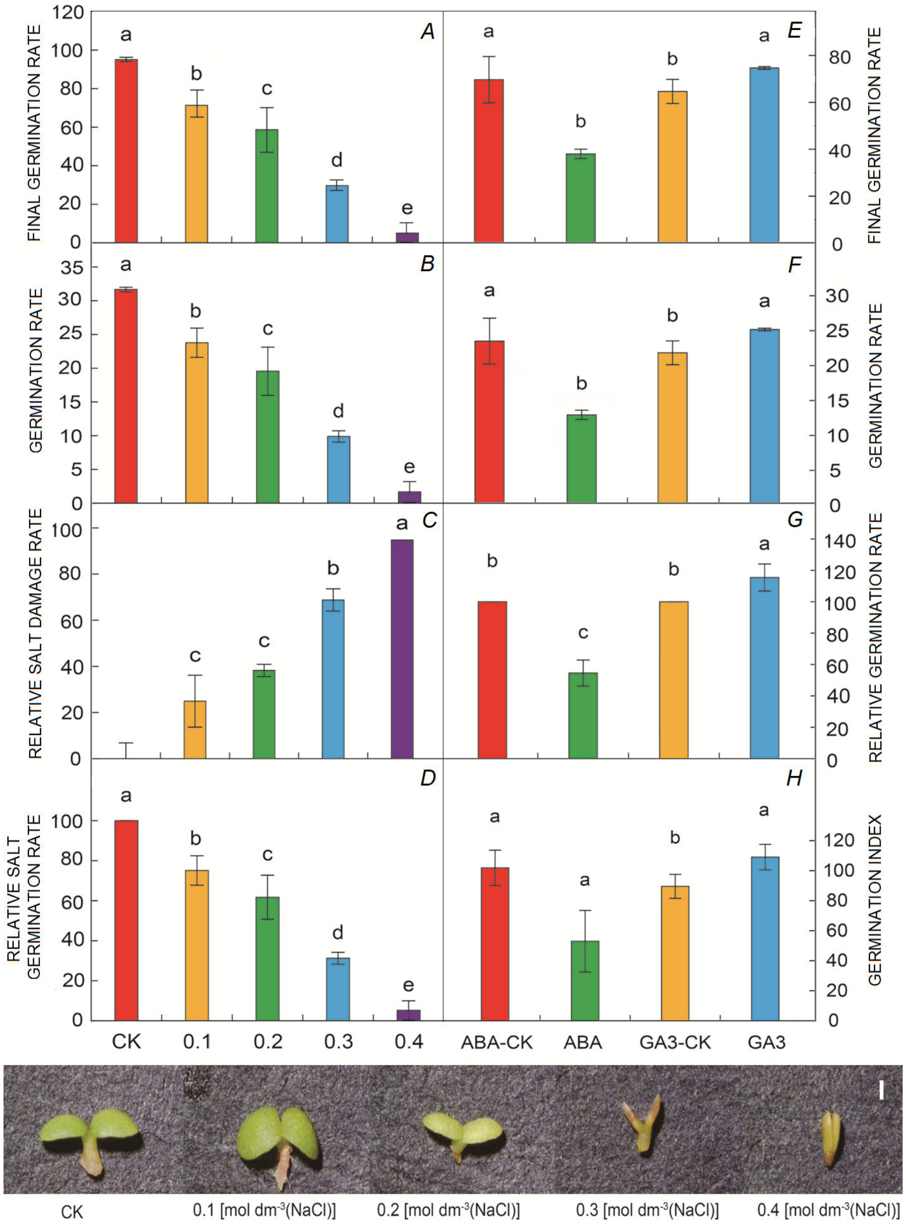

Fig. 4. Comparison of physiological indices for *P. euphratica* seeds under salt stress (*A*,*B*,*C*,*D*,*I*) and exogenous hormones (*E*,*F*,*G*,*H*): *A* - final germination rate, *B* - germination rate, *C* - relative salt damage rate, *D* - relative germination rate; *E* - final germination rate, *F* - germination rate, *G* - relative germination rate, *H* - germination index; *I* - the phenotype of seed germination: cotyledons and hypocotyls under salt stress.

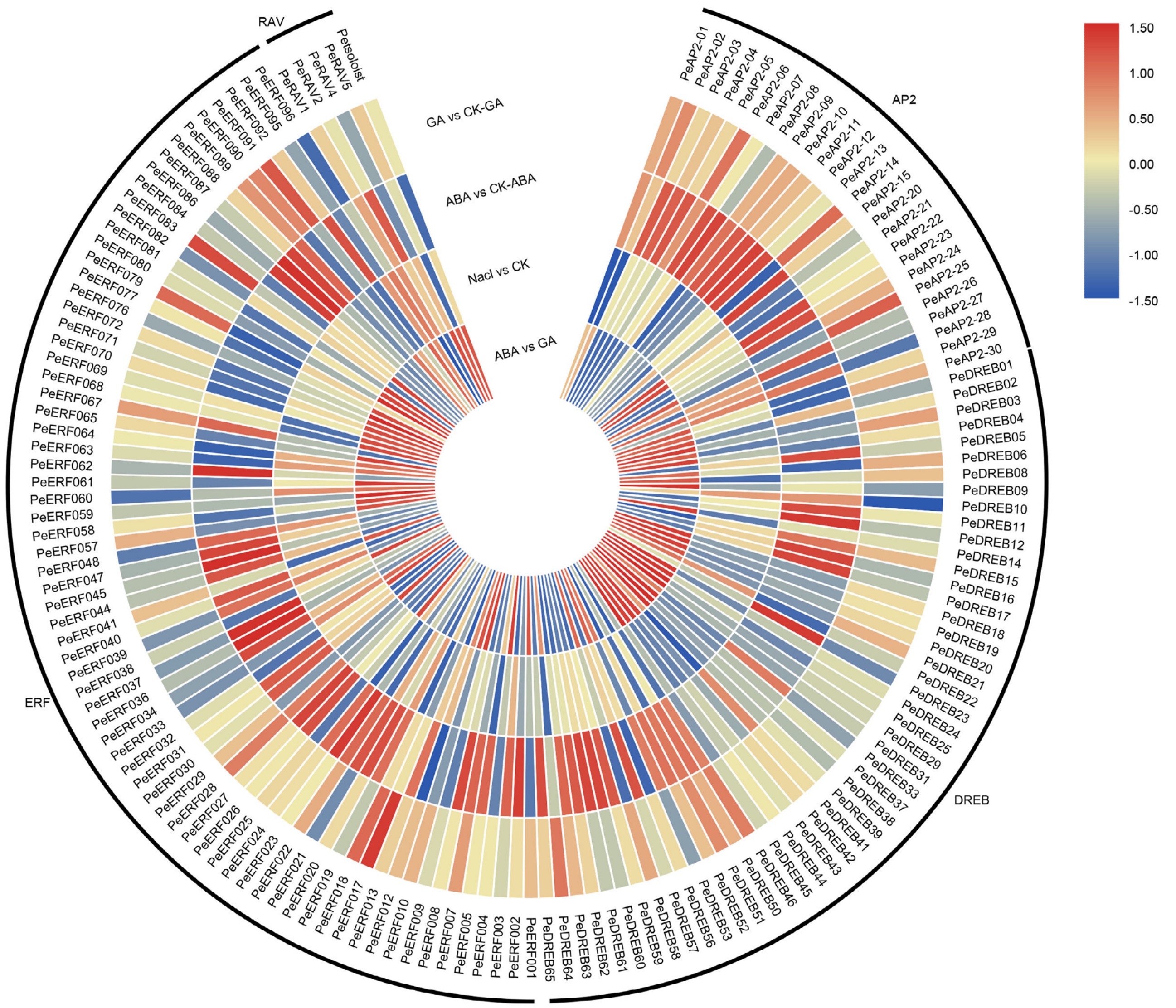

Fig. 5. The gene expression heat map of *P. euphratica AP2/ERF* gene family.

can regulate the tolerance of transgenic *Arabidopsis* to different abiotic stresses, such as high salinity, drought and extreme temperature (heat, chilling and freezing), as well as ABA and methyl viologen treatments (Yin *et al.* 2021). In *Catharanthus roseus*, the *AP2/ERF* transcription factor *CrERF5* is induced by ethylene and jasmonic acid and activates the transcriptional activity of tryptophan decarboxylase promoter. Transient overexpression of *CrERF5* in petals causes a significant increase in the expressions of key genes in the upstream and downstream pathways of monoterpene indole alkaloids (MIA) biosynthesis, whereas silencing *CrERF5* results in a decrease in them (Pan *et al.* 2019). However, few studies have investigated the *AP2/ERF* genes in *P. euphratica*.

In our study, based on the high-quality chromosome-level genome data of *P. euphratica*, we identified the 197 *AP2/ERF* gene family members, furthermore, a total of 210, 231, 192 and 147 *AP2/ERF* gene family members were identified in *P. trichocarpa*, *P. deltoides*, *S. sinopurpurea*, and *A. thaliana*, respectively. Among the four *Salicaceae* species, *P. deltoides* had the largest number of *AP2/ERF* gene family members, whereas *S. sinopurpurea* had the least number, but far more than *A. thaliana*. This distribution showed that *AP2/ERF* gene family members had expanded or contracted in the evolution of different families and species. Concurrently, the genome sizes of the four *Salicaceae* plants were different, with 580 Mb in *P. euphratica*, 439 Mb in *P. trichocarpa*, 452 Mb in *P. deltoides*, and 333 Mb in *S. sinopurpurea*. These values suggested that the number of AP2/ERF superfamily members was relatively stable, with no absolute correlation with genome size. *Populus* species had more protein-coding genes than *A. thaliana*, with an average of 1.4 - 1.6 putative poplar homologous genes per *A. thaliana* gene (Tuskan *et al*. 2006). In this study, the ratio of the number of genes in each subfamily to *Arabidopsis* was also about 1.1 - 1.6 times.

Bioinformatics analysis, such as family member classification, multiple sequence alignment and conserved motifs, can be helpful to study the biological function of each gene in the poplar *AP2/ERF* gene family. The phylogenetic analysis of the conserved domains of the DREB, ERF, AP2, and RAV subfamily revealed that the *AP2/ERF* genes of four *Salicaceae* plants were divided into 12 categories. Taking *P. euphratica* as an example, the *AP2/ERF* gene family is divided into I - IV group (DREB subfamily), V - X group (ERF subfamily), AP2 subfamily, RAV subfamily, and a Soloist member (*Pe-soloist*). The number of *AP2/ERF* genes is determined by the number of ERF subfamily members to a certain extent (Zhao *et al*. 2014). The quantitative difference between AP2 and RAV in *Salicaceae* plants is not very significant. We speculated that ERF and DREB of the *AP2/ERF* gene family in *P. euphratica* may play an important role in adapting to drought and high salt environment, so it can be amplified in large numbers.

The domains and motifs of transcription factors are often related to protein interaction, transcriptional activity, and DNA binding (Liu *et al*. 1999). Motif analysis showed that motif-2 was widely distributed in ERF and DREB subfamilies, motif-7 was widely distributed in RAV and AP2 subfamilies and motif-1 was widely distributed in all subfamilies. By contrast, most conserved motifs were specific and only appeared in specific subfamilies. For example, motif-3, motif-4, and motif-10 were specific to the AP2 subfamily; motif-9 only appeared in the DREB subfamily. These results showed that these motifs of the *AP2/ERF* gene family were highly conserved, but different motifs may play different functions in plants. However, this conclusion needs to be further verified (Liu *et al*. 2019).

Structural analysis of *AP2/ERF* showed that it did not contain introns and accounted for about 61.5% of the ERF and DREB subfamily. This characteristic was also noted in the *AP2/ERF* genes of *Arabidopsis*, bamboo, and banana (Kamath *et al.* 2003, Wu *et al*. 2015, Lakhwani *et al*. 2016), whereas the AP2 subfamily genes contained more than 10 introns. This highly diverse gene structure suggested that vast differentiation may occur during genome evolution. The exon–intron position pattern provided important clues to the evolutionary relationship. A previous study indicated the possibility that a longer piece of DNA makes a transposition and duplication event less likely, which is consistent with the small number of members of AP2 and RAV subfamilies (Magnani *et al.* 2004).

Gene duplication events are one of the main evolutionary mechanisms for the production of new genes, and they are also important means to help plants cope with environmental changes during growth and development (Ren *et al*. 2018). The fragment replication events found in 104 *SsAP2/ERF* in sugarcane confirmed that the gene may produce gene clusters through tandem duplication and fragment replication, thereby producing homologous genes and expanding the total number of genes (Li *et al.* 2020). Our results showed that *PeAP2/ERF* replication events mainly occur on chromosomes other than chromosomes 2, 12, 13, 14, and 19. A total of 67/197 *AP2/ERF* members were caused by replication events, including 27 pairs of tandem repeat events. This result may explain the uneven distribution of *AP2/ERF* gene family on the chromosomes of *P. euphratica*. Therefore, fragment duplication may produce many homologous *AP2/ERF* genes on different chromosomes, expanding the number of the *AP2/ERF* gene family in different species.

The functions of genes can be preliminarily predicted by analysis of the gene expression patterns. In this research, all *AP2/ERF* family members were involved in gene expression with ABA *vs.* CK, NaCl *vs.* CK, GA_3 *vs.* CK, and ABA *vs.* GA_3 group. The results indicated that most *AP2/ERF* gene family members participate in the regulation of seed germination in response to stress through different pathways. Overexpressing DREB subfamily genes *OsDREB1F* and *OsDREB2A* in *A. thaliana* and *Oryza sativa* can improve their drought and salt tolerance (Mallikarjuna *et al*. 2011). Meanwhile, the deletion mutant of *Arabidopsis* ERF family member *AtERF71/HRE2* (*AT2G47520.1*) shows salt tolerance (Park *et al.* 2011). The phylogenetic tree revealed that *P. euphratica* DREB subfamily genes (*PeDREB57* and *PeDREB59*) and ERF subfamily gene *PeERF037* were highly homologous with *A. thaliana DREB2A* (*AT5G05410.1*) and *AtERF71/HRE2* (*AT2G47520.1*), respectively. Therefore, *PeDREB57*, *PeDREB59*, and *PeERF037* may participate in plant response to drought and salt stress, and their products play an important role in the expression of dehydration response genes.

Some proteins of the ERF subfamily can also be used as negative regulatory proteins. *AtERF4* can regulate stomatal size or plant response to drought, salinity, cold, and ABA stress through the ABA-dependent signalling pathway (Song *et al*. 2005). According to our phylogenetic tree analysis, *PeERF044* was highly homologous with *AtERF7*, and high expression in ABA *vs.* CK and NaCl *vs.* CK. Therefore, *PeERF044* may also participate in the regulation of plant salt and drought tolerance through the ABA-dependent signalling pathway. Studies have shown that GA_3 can also regulate abiotic stresses. For example, low content of GA_3 will slow down the growth of plants under abiotic stresses, such as cold, salinity, and low osmotic potential (Colebrook *et al*. 2014). In general, abiotic stresses cause a reduction in GA_3 content and signalling through the inhibition of the expression of GA_3 biosynthesis enzymes mediated by *Arabidopsis AP2/ERFs* (Xie *et al*. 2019). In *P. euphratica*, *PeAP2-14*, *PeERF090*, *PeERF059*, and *PeERF082* were up-regulated in GA_3 *vs.* CK, and *PeDREB59* was still up-regulated in ABA *vs.* CK. The results showed that these candidate genes had become key regulators of stress response. In summary, some genes respond to salt stress, and some genes help activate ABA-dependent and independent stress response genes. Some genes were related to GA_3-mediated growth and development. In addition to participating in specific stress, the *AP2/ERF* gene family was interrelated, enabling them to form stress regulatory networks that helped *P. euphratica* seedlings adapt to extreme environment and improve their seedling survival rate.

References

Agarwal G., Garg V., Kudapa H. *et al.*: Genome-wide dissection of AP2/ERF and HSP90 gene families in five legumes and expression profiles in chickpea and pigeonpea. - Plant Biotechnol. Rep. **14**: 1563-1577, 2016.

Agarwal P., Agarwal P.K., Nair S. *et al.*: Stress-inducible DREB2A transcription factor from *Pennisetum glaucum* is a phosphoprotein and its phosphorylation negatively regulates its DNA-binding activity. - Mol. Genet. Genomics **277**: 189-198, 2007.

Aukerman M.J., Sakai H.: Regulation of flowering time and floral organ identity by a microRNA and its *APETALA2*-like target genes. - Plant Cell **15**: 2730-2741, 2003.

Colebrook E.H., Thomas S.G., Phillips A.L., Hedden P.: The role of gibberellin signalling in plant responses to abiotic stress. - J. Exp. Biol. **217**: 67-75, 2014.

Daszkowska-Golec A.: *Arabidopsis* seed germination under abiotic stress as a concert of action of phytohormones. - Omics **15**: 763-774, 2011.

Deng B., Huang Z., Ge F. *et al.*: An AP2/ERF family transcription factor PnERF1 raised the biosynthesis of saponins in *Panax notoginseng*. - J. Plant Growth Regul. **36**: 691-701, 2017.

Dinh T.T., Girke T., Liu X. *et al.*: The floral homeotic protein APETALA2 recognizes and acts through an AT-rich sequence element. - Development **139**: 1978-1986, 2012.

Dossa K., Wei X., Li D.H. *et al.*: Insight into the AP2/ERF transcription factor superfamily in sesame and expression profiling of DREB subfamily under drought stress. - BMC Plant Biol. **16**: 171, 2016.

Dubouzet J.G., Sakuma Y., Ito Y. *et al.*: *OsDREB* genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. - Plant J. **33**: 751-763, 2003.

Faraji S., Filiz E., Kazemitabar S.K. *et al.*: The *AP2/ERF* gene family in *Triticum durum*: genome-wide identification and expression analysis under drought and salinity stresses. - Genes **11**: 1464, 2020.

Fu M., Kang H.K., Son S.H. *et al.*: A subset of *Arabidopsis* RAV transcription factors modulates drought and salt stress responses independent of ABA. - Plant Cell Physiol. **55**: 1892-1904, 2014.

Guo F.Y., Liu K.J., Wang Y.C. *et al.*: Complete chloroplast genome sequence of *Salix sinopurpurea* (*Salicaceae*). - Mitochondr. DNA B **6**: 718-719, 2021.

Jahromi F., Aroca R., Porcel R., Ruiz-Lozano J.M.: Influence of salinity on the *in vitro* development of *Glomus intraradices* and on the *in vivo* physiological and molecular responses of mycorrhizal lettuce plants. - Microb. Ecol. **55**: 45-53, 2008.

Jin J.H., Wang M., Zhang H.X. *et al.*: Genome-wide identification of the AP2/ERF transcription factor family in pepper (*Capsicum annuum* L.). - Genome **61**: 663-674, 2018.

Kamath R.S., Fraser A.G., Dong Y. *et al.*: Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. - Nature **421**: 231-237, 2003.

Kudo K., Oi T., Uno Y.: Functional characterization and expression profiling of a *DREB2*-type gene from lettuce (*Lactuca sativa* L.). - Plant Cell Tiss. Org. Cult. **116**: 97-109, 2014.

Lakhwani D., Pandey A., Dhar Y.V. *et al.*: Genome-wide analysis of the AP2/ERF family in *Musa* species reveals divergence and neofunctionalisation during evolution. - Sci. Rep.-UK **6**: 18878, 2016.

Li C.W., Su R.C., Cheng C.P. *et al.*: Tomato RAV transcription factor is a pivotal modulator involved in the AP2/EREBP-mediated defense pathway. - Plant Physiol. **156**: 213-227, 2011.

Li H., Wang Y., Wu M. *et al.*: Genome-wide identification of *AP2/ERF* transcription factors in cauliflower and expression profiling of the ERF family under salt and drought stresses. - Front Plant Sci. **8**: 946, 2017.

Li P.T., Chai Z., Lin P.P. *et al.*: Genome-wide identification and expression analysis of AP2/ERF transcription factors in sugarcane (*Saccharum spontaneum* L.). - BMC Genomics **21**: 685, 2020.

Li X.L., Tao S.T., Wei S.W. *et al.*: The mining and evolutionary investigation of *AP2/ERF* genes in pear (*Pyrus*). - BMC Plant Biol. **18**: 46, 2018.

Liu L., White M.J., MacRae T.H.: Transcription factors and their genes in higher plants: functional domains, evolution and regulation. - Eur. J. Biochem. **262**: 247-257, 1999.

Liu M.Y., Sun W.J., Ma Z.T. *et al.*: Genome-wide investigation of the AP2/ERF gene family in tartary buckwheat (*Fagopyum tataricum*). - BMC Plant Biol. **19**: 84, 2019.

Magnani E., Sjölander K., Hake S.: From endonucleases to transcription factors: evolution of the AP2 DNA binding domain in plants. - Plant Cell **16**: 2265-2277, 2004.

Mallikarjuna G., Mallikarjuna K., Reddy M.K., Kaul T.: Expression of *OsDREB2A* transcription factor confers enhanced dehydration and salt stress tolerance in rice (*Oryza sativa* L.). - Biotechnol. Lett. **33**: 1689-1697, 2011.

Mizoi J., Shinozaki K., Yamaguchi-Shinozaki K.: *AP2/ERF* family transcription factors in plant abiotic stress responses. - BBA-Gene Regul. Mech. **1819**: 86-96, 2012.

Moose S.P., Sisco P.H.: *Glossy15*, an APETALA2-like gene from maize that regulates leaf epidermal cell identity. - Gene. Dev. **10**: 3018-3027, 1996.

Nakano T., Suzuki K., Fujimura T., Shinshi H.: Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. - Plant Physiol. **140**: 411-432, 2006.

Pan Q.F., Wang C.Y., Xiong Z.W. *et al.*: CrERF5, an AP2/ERF transcription factor, positively regulates the biosynthesis of bisindole alkaloids and their precursors in *Catharanthus roseus*. - Front. Plant Sci. **10**: 931, 2019.

Park H.Y., Seok H.Y., Woo D.H. *et al.*: *AtERF71/HRE2* transcription factor mediates osmotic stress response as well as hypoxia response in *Arabidopsis*. - Biochem. Biophys. Res. Commun. **414**: 135-141, 2011.

Rashid M., Guangyuan H., Guangxiao Y. *et al.*: *AP2/ERF* transcription factor in rice: genome-wide canvas and syntenic relationships between monocots and eudicots. - Evol. Bioinform. **8**: 321-355, 2012.

Ren R., Wang H.F., Guo C.C. *et al.*: Widespread whole genome duplications contribute to genome complexity and species diversity in angiosperms. - Mol. Plant **11**: 414-428, 2018.

Sakuma Y., Liu Q., Dubouzet J.G. *et al.*: 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, 2002.

Sohn K.H., Lee S.C., Jung H.W. *et al.*: Expression and functional roles of the pepper pathogen-induced transcription factor RAV1 in bacterial disease resistance, and drought and salt stress tolerance. - Plant Mol. Biol. **61**: 897-915, 2006.

Song C.P., Agarwal M., Ohta M. *et al.*: Role of an *Arabidopsis* AP2/EREBP-type transcriptional repressor in abscisic acid and drought stress responses. - Plant Cell **17**: 2384-2396, 2005.

Song X.M., Li Y., Hou X.L.: Genome-wide analysis of the AP2/ERF transcription factor superfamily in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). - BMC Genomics **14**: 573, 2013.

Stockinger E.J., 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. - PNAS **94**: 1035-1040, 1997.

Sun X.F., Zheng Q.S., Liu Y.L.: [Salinity injury to germination and growth of cotton (*Gossypium hirsutum* L.) at emergence and seedling stages.] - J. Plant Resour. Environ. **9**: 22-25, 2000. [In Chinese]

Taylor G.: *Populus*: *Arabidopsis* for forestry. Do we need a model tree? - Ann. Bot.-London **90**: 681-689, 2002.

Tuskan G.A., Difazio S., Jansson S. *et al.*: The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). - Science **313**: 1596-1604, 2006.

Van der Fits L., Memelink J.: *ORCA3*, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. - Science **289**: 295-297, 2000.

Wu H.L., Lv H., Li L. *et al.*: Genome-wide analysis of the AP2/ERF transcription factors family and the expression patterns of *DREB* genes in Moso bamboo (*Phyllostachys edulis*). - PLoS ONE **10**: e0126657, 2015.

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. **10**: 228, 2019.

Yao Y., He R.J., Xie Q.L. *et al.*: *ETHYLENE RESPONSE FACTOR 74* (*ERF74*) plays an essential role in controlling a respiratory burst oxidase homolog D (RbohD)-dependent mechanism in response to different stresses in *Arabidopsis*. - New Phytol. **213**: 1667-1681, 2017.

Yin F.L., Zeng Y.L., Ji J.Y. *et al.*: The halophyte *Halostachys caspica* AP2/ERF transcription factor HcTOE3 positively regulates freezing tolerance in *Arabidopsis*. - Front. Plant Sci. **12**: 638788, 2021.

Yu Z.X., Li J.X., Yang C.Q. *et al.*: The jasmonate-responsive AP2/ERF transcription factors AaERF1 and AaERF2 positively regulate artemisinin biosynthesis in *Artemisia annua* L. - Mol. Plant **5**: 353-365, 2012.

Zhai J.T, Li Y.L, Han Z.J, Li Z.J.: Morphological, structural and physiological differences in heteromorphic leaves of Euphrates poplar during development stages and at crown scales. - Plant Biol. **22**: 366-375, 2020.

Zhang C.H., Luo W.C., Li Y.D. *et al.*: Transcriptomic analysis of seed germination under salt stress in two desert sister species (*Populus euphratica* and *P. pruinosa*). - Front. Genet. **10**: 231, 2019.

Zhang Z.Y., Chen Y., Zhang J.L. *et al.*: Improved genome assembly provides new insights into genome evolution in a desert poplar (*Populus euphratica*). - Mol. Ecol. Resour. **20**: 781-794, 2020.

Zhao T., Xia H., Liu J.Y., Ma F.W.: The gene family of dehydration responsive element-binding transcription factors in grape (*Vitis vinifera*): genome-wide identification and analysis, expression profiles, and involvement in abiotic stress resistance. - Mol. Biol. Rep. **41**: 1577-1590, 2014.