Genome-wide investigation of the AP2/ERF superfamily and their expression under salt stress in Chinese willow (*Salix matsudana*) (#53462)

First submission

Guidance from your Editor

Please submit by **22 Nov 2020** for the benefit of the authors (and your $200 publishing discount) .

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Custom checks
Make sure you include the custom checks shown below, in your review.

Author notes
Have you read the author notes on the guidance page?

Raw data check
Review the raw data.

Image check
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files

Download and review all files from the materials page.

11 Figure file(s)
1 Table file(s)
3 Raw data file(s)

Custom checks

DNA data checks

- Have you checked the authors data deposition statement?
- Can you access the deposited data?
- Has the data been deposited correctly?
- Is the deposition information noted in the manuscript?

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Speculation is welcome, but should be identified as such.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Genome-wide investigation of the AP2/ERF superfamily and their expression under salt stress in Chinese willow (*Salix matsudana*)

Jian Zhang [Equal first author, 1], **Yuna Jiang** [Equal first author, 1], **Shi zheng Shi** [2], **Fei Zhong** [1], **Guoyuan Liu** [1], **Chunmei Yu** [1], **Bolin Lian** [1], **Yanhong Chen** [Corresp. 1]

[1] School of life science, Nantong University, Nantong, Jiangsu, China

[2] Jiangsu Academy of Forestry, Nanjing, Jiangsu, China

Corresponding Author: Yanhong Chen
Email address: chenyh@ntu.edu.cn

AP2/ERF transcription factors (TFs) play indispensable roles in plant growth, development, and especially in various abiotic stresses responses. The AP2/ERF TF family has been discovered and classified in more than 50 species. However, little is known about the *AP2/ERF* gene family of Chinese willow (*Salix matsudana*), which is a tetraploid ornamental tree species that is widely planted and is also considered a species that can improve the soil salinity of coastal beaches. In this study, 364 *AP2/ERF* genes of *Salix matsudana* (*SmAP2/ERF*) were identified depending on the recently produced whole genome sequencing data of *Salix matsudana*. These genes were renamed according to the chromosomal location of the *SmAP2/ERF* genes. The *SmAP2/ERF* genes included three major subfamilies: AP2 (55 members), ERF (301 members), and RAV (six members) and two Soloist genes. Genes' structure and conserved motifs were analyzed in SmAP2/ERF family members, and introns were not found in most genes of the ERF subfamily, some unique motifs were found to be important for the function of *SmAP2/ERF* genes. Syntenic relationships between the *SmAP2/ERF* genes and *AP2/ERF* genes from *Populus trichocarpa* and *Salix purpurea* showed that *Salix matsudana* is genetically more closely related to *Populus trichocarpa* than to *Salix purpurea*. Evolution analysis on paralog gene pairs suggested that progenitor of *S. matsudana* originated from hybridization between two different diploid salix germplasms and underwent genome duplication not more than 10 Mya. RNA sequencing results demonstrated the differential expression patterns of some *SmAP2/ERF* genes under salt stress and this information can help reveal the mechanism of salt tolerance regulation in *Salix matsudana*.

Genome-wide investigation of the AP2/ERF superfamily and their expression under salt stress in Chinese willow (*Salix matsudana*)

Running title: AP2/ERF superfamily in Chinese willow

Jian Zhang[1a], Yuna Jiang[1a], Shi zheng Shi[2a], Fei Zhong[1], Guoyuan Liu[1], Chunmei Yu[1], Bolin Lian[1], Yanhong Chen [1]*

[1]Lab of Landscape Plant Genetics and Breeding, School of Life Science, Nantong University, Nantong, Jiangsu Province, China
[2] Jiangsu Academy of Forestry, Nanjing 211153, China
[a]These authors contribute equally to this work
*Corresponding author. E-mail: chenyh@ntu.edu.cn ; FAX: 86-513-85012818; Tel: 86-513-85012818

E-mail address for other authors
Jian Zhang: 56071007@qq.com
Yuna Jiang :18252098108@163.com;
Shizheng Shi: shshzn@163.com
Fei Zhong: fzhong@ntu.edu.cn;
Guoyuan Liu: cjqm1989@126.com;
Chunmei Yu: ychmei@ntu.edu.cn;
Bolin Lian: lianziadd9@163.com;

Abstract

AP2/ERF transcription factors (TFs) play indispensable roles in plant growth, development, and especially in various abiotic stresses responses. The AP2/ERF TF family has been discovered and classified in more than 50 species. However, little is known about the *AP2/ERF* gene family of Chinese willow (*Salix matsudana*), which is a tetraploid ornamental tree species that is widely planted and is also considered a species that can improve the soil salinity of coastal beaches.

In this study, 364 *AP2/ERF* genes of *Salix matsudana* (*SmAP2/ERF*) were identified depending on the recently produced whole genome sequencing data of *Salix matsudana*. These genes were renamed according to the chromosomal location of the *SmAP2/ERF* genes. The *SmAP2/ERF* genes included three major subfamilies: AP2 (55 members), ERF (301 members), and RAV (six members) and two Soloist genes. Genes' structure and conserved motifs were analyzed in SmAP2/ERF family members, and introns were not found in most genes of the ERF subfamily, some unique motifs were found to be important for the function of *SmAP2/ERF* genes. Syntenic relationships between the *SmAP2/ERF* genes and *AP2/ERF* genes from *Populus trichocarpa* and *Salix purpurea* showed that *Salix matsudana* is genetically more closely related to *Populus trichocarpa* than to *Salix purpurea*. Evolution analysis on paralog gene pairs suggested that progenitor of *S. matsudana* originated from hybridization between two different diploid salix germplasms and underwent genome duplication not more than 10 Mya. RNA sequencing results demonstrated the differential expression patterns of some *SmAP2/ERF* genes under salt stress and this information can help reveal the mechanism of salt tolerance regulation in *Salix matsudana*.

52 **Abbreviations**

53 AP2/ERF: AP2-like ethylene-responsive transcription factor

54 FPKM: Fragments Per Kilobase of transcript per Million fragments mapped

55 HMM: Hidden Markov Model

56 Ka: Nonsynonymous substitution rate

57 Ks: Synonymous substitution rate

58 Mya: million years ago

59 qRT-PCR: Real-time Quantitative PCR

60 SDs: segmental duplication events

61 TDs: tandem duplication events

62 TFs: transcription factors

63

Introduction

APETALA 2/ethylene-responsive element binding factors (AP2/ERF) are important transcription factors (TFs) coded by genes from the AP2/ERF superfamily. All of the members of this superfamily possess AP2 domains and, according to the number and structure of AP2 domains, the superfamily is divided into several categories, including AP2, ERF, RAV, and Soloist (*Nakano et al., 2006*). Most of the AP2 gene family members have two AP2 domains and can be further divided into AP2 and ANT groups; ERF family members have only one AP2 domain and can also be subdivided into ERF and DREB subfamilies based on binding motifs in the promoter of downstream genes. Members of the ERF and DREB subfamilies are classified into 12 groups (groups A1–B6). DREB includes groups A1–A6, whereas ERF includes groups B1–B6 (*Nakano et al., 2006*). In addition to one AP2 domain, RAVs also have one B3 domain. The Soloist group contains a single AP2 domain with sequence divergence from the AP2 and ERF families and has less than three members in most species (*Song et al., 2016*).

The AP2/ERF superfamily is plant-specific and has more than 100 members in many plant species; for example, there are 147 members in *Arabidopsis*, 200 members in *Populus trichocarpa*, and more than 500 members in the tetraploid crop *Brassica napus* (*Nakano et al,. 2006; Song et al., 2016; Zhuang et al., 2008*). Different members play various regulatory roles in plant growth and development, defense response, fruit ripening, and metabolism (*Gu et al., 2017*). Several recent reports demonstrated functions of AP2/ERF2 TFs in plant development. For example, loss of DRNL function affects gynoecium development (*Duran-Medina et al., 2017*); the function of *Populus* ERF139 (Potri.013G101100) in xylem cell expansion was characterized by transgenic overexpression and dominant repressor lines of ERF139 (*Wessels et al., 2019*); RhERF1 and RhERF4 play roles in petal abscission in rose (*Gao et al., 2019*); and a maize AP2/ERF TF, ZmRAP2.7, is involved in brace root development. AP2/ERF TFs such as ZmEREB94 and CitAP2.10 also play important roles in plant metabolism; ZmEREB94 acts as a key regulator of starch synthesis in maize (*Li et al., 2017b*), and CitAP2.10 was characterized as a regulator of (+)-valencene synthesis in sweet orange fruit (S*hen et al., 2016*).

The AP2/ERF superfamily plays major and crucial roles in abiotic stress tolerance, which is why this superfamily has received special attention by plant scientists (*Gu et al., 2017; Srivastava & Kumar, 2018*). Through extensive investigation on their regulatory mechanism, people want to elucidate their potential applications in crop improvement (*Srivastava & Kumar, 2018*). Members of this superfamily (primarily ERFs and DREBs) have been prominently used to improve stress tolerance in plants. To improve salinity stress tolerance, many genes from different species were identified. IbRAP2-12, an AP2/ERF gene cloned from the salt-tolerant sweet potato, and LkERF-B2 from *Larix kaempferi* promotes tolerance to salt and drought stresses in overexpressing *Arabidopsis* lines (*Cao et al., 2019; Li et al., 2019b*). Overexpression of HARDY, an AP2/ERF gene from *Arabidopsis*, improves drought and salt tolerance by reducing transpiration and sodium uptake in transgenic *Trifolium al.exandrinum L* (*Abogadallah et al., 2011*). A soybean DREB ortholog, GmDREB1, enhances the salt tolerance in transgenic al.falfa (*Jin et al., 2010*).

Comparative genomic analysis of model plants such as *Arabidopsis* have provided unprecedented advantages for gene discovery and functional annotation of newly sequenced plant genomes (*Brendel et al., 2002; Hall et al., 2002; Schranz et al., 2007*). By exploring the available genomic data, AP2/ERF gene families from 50 species were discovered and classified, and provide critical guidance for functional analysis (*Srivastava & Kumar, 2018*). For example, in radish, cauliflower, and celery, whole genome identification and classification of AP2/ERF gene family members were carried out; additionally, expression patterns of different members under different stresses were revealed, and the function of candidate genes was verified (*Karanja et al., 2019; Li et al., 2017a; Li et al., 2019a*).

Salix matsudana Koidz., an allotetraploid member of Salicaceae, is an important ornamental tree species native to northeastern China (*Zhang et al., 2016; Zhang et al., 2017*); it is widely cultivated and considered an important economic plant because of its easy vegetative propagation, rapid growth, and substantial biomass yields. *Salix matsudana* also plays an important ecological role when grown along Chinese coastal beaches, where the salinity content

is high (*Zhang et al., 2016*). This species can improve the beach soil and alleviate salinization. Newly reclaimed beach soil has higher salinity and requires new germplasm with higher salinity tolerance (*Zhang et al., 2017*). Because the AP2/ERF gene family members have regulatory roles in salinity tolerance, whole genome characterization of the AP2/ERF gene family in *Salix matsudana* will reveal mechanisms underlying stress signal transmission and provide guidance for selection or creation of new germplasm with higher salinity tolerance. In total, 200 and 173 AP2/ERF superfamily genes were identified from two species, diploid salix germplasm *Salix arbutifolia* and *Populus trichocarpa*, respectively (*Rao et al., 2015, Zhuang et al., 2008*). The *Salix matsudana* genome was recently sequenced and assembled (In preparation of publication in Horticulture Research); as an allotetraploid, identification of the AP2/ERF gene family will reveal the evolutionary relationship with poplar and other members of *Salix*, and the molecular mechanisms responsible for salinity stress responses.

Materials and Methods

Plant material and salt stress treatment

The salinity stress treatments were carried out on Salix matsudana 'yanjiang' and Salix matsudana '9901'. Our previous research results showed that 'yanjiang' was salt-sensitive variety, while '9901' was salt-tolerant variety (*Zhang et al., 2016*). The two salix cultivars used in this study were collected from the botany garden of Nantong University (Nantong, China). The two salix cultivars were authorized for only scientific research purpose, and were deposited in school of life science in Nantong University. The stem cuttings (length, 8–10 cm; coarse, 2–3 mm) of two materials were selected for hydroponic rooting for 20 days. The stems with new generated roots were dipped into NaCl solution (150 mM) (only root and part of shoots were immersed in the solution) for 4 h. All root samples were divided into four categories with three biological replicate to do RNA sequencing: Sample1-1/Sample1-2/Sample1-3 ('yanjiang' without salt stress treatment, Yanjiang), Sample1N-1/Sample1N-2/Sample1N-3 ('yanjiang' with salt stress treatment, Yanjiang NT), Sample2-1/Sample2-2/Sample2-3 ('9901' without salt stress treatment, 9901), and Sample2N-1/Sample2N-2/Sample2N-3 ('9901' with salt stress treatment,

9901 NT).

RNA isolation and Real-time Quantitative PCR (qRT-PCR) analysis

Total RNA was extracted using TaKaRa MiniBEST Plant RNA Extraction Kit (Takara, Dalian, China) from roots according to the manufacturer's instruction. Four samples (Yanjiang, Yanjiang NT, 9901, 9901 NT) were collected from the salt treat experiment mimic RNA sequencing sample collection. For each sample, 3 μg of total RNA was used to synthesize first-strand cDNA with SuperScriptII reverse transcriptase (Takara, Dalian, China). For qRT-PCR, the reaction preparation, application parameter settings and quantitative analysis were performed as previously described (*Chen et al., 2018*). The reactions were performed using the ABI Prism 7000 Real-time PCR system (Applied Biosystems, USA). The Salix purpurea Actin1 gene (SapurV1A.0655s0050.1) were used as reference genes. The gene-specific primers for the 13 selected genes are listed in Table S1.

Genome sequence retrieval

The *Populus trichocarpa* and *Salix purpurea* sequences were downloaded from JGI (http://www.phytozome.net/). The *Salix matsudana* sequences were obtained from our sequencing, and assembly results were obtained by Roche/454 and Illumina/HiSeq-2000 sequencing technologies (In preparation of publication in Horticulture Research).

Identification of *AP2/ERF* genes in *Salix matsudana* and *Salix purpurea*

The Pfam accession number of AP2 domain is PF00847.16. We downloaded the Hidden Markov Model (HMM) profile for the AP2/ERF TFs from the Pfam database (http://pfam.xfam.org/) with Pfam accession number PF00847.16 as the search keyword. An alternative HMM profile was built by sequence alignment using ClustalW (*Larkin, et al., 2007)*. Two HMM profile files were provided as supplemental file S1 and file S2. Using an in-house Perl script with two HMM profiles as queries, hmmsearch was carried out by searching the *Salix matsudana* and *Salix purpurea* protein databases with default parameters. To validate the putative accuracy of two HMM search results, the candidate protein sequences were checked in three websites: SMART (http://smart.embl.de/#), CDD (https://www.ncbi.nlm.nih.gov/cdd/), and Pfam

(http://pfam.xfam.org/). Candidate proteins with positive results from all three websites were selected as AP2/ERF family members of *Salix matsudana* and *Salix purpurea*. Additionally, putative AP2/ERF protein characteristics, including length, molecular weight, and isoelectric point, were calculated by the ExPasy site (http://au.expasy.org/tools/pi_tool.html). The genes CDS sequences were listed in supplemental file S3.

Phylogenetic analysis and classification of AP2/ERF genes

Using an in-house Perl script (domain_xulie.pl), the conserved AP2 core domains of putative SmAP2/ERF proteins were obtained and subjected to multiple sequence alignment using ClustalW (*Larkin et al., 2007*). To better classify these SmAP2/ERF proteins, 48 AP2 domains from known categories of *Arabidopsis* and *Populus trichocarpa* AP2 genes were selected to carry out multiple sequence alignment with SmAP2/ERF proteins, and a phylogenetic tree based on this alignment was built by MEGA 7.0 with the neighbor-joining method with default parameters (*Kumar et al., 2016)*. Bootstrap value was set to 1000. Depending on the phylogenetic tree constructed by SmAP2, PtAP2, and AtAP2 domains, these SmAP2/ERF proteins were classified into different subfamilies and subgroups.

Gene structure and conserved motif structure analysis

The UTR–exon–intron structures of the *SmAP2/ERF* genes were obtained based on the gene annotation gff3 files we assembled. Using the online website tool Gene Structure Display Server (http://gsds.cbi.pku.edu.cn/), we obtained the gene structure diagrams (*Hu et al., 2014*).

To characterize the structures of SmAP2/ERF proteins, the online tool MEME (http://meme-suite.org/tools/meme) was used to search for conserved motifs (*Bailey et al., 2009*). The optimized parameters were employed as follows: any number of repetitions, maximum number of motifs = 10, and the optimum width of each motif was 6–50 residues. The search result file meme.xml was downloaded from the website and opened by TBtools v0.66831 to obtain the gene structure diagram (*Chen et al., 2020*).

Gene position on chromosomes, and gene tandem and segmental duplication analysis

Using the “Amazing Gene Location from GFF3/GTF File” tool of TBtools, the *SmAP2/ERF*

genes were mapped on 38 chromosomes of *Salix matsudana*. Because some scaffolds were not assembled onto the chromosomes, not all *SmAP2* genes mapped onto the chromosomes (*Chen et al., 2020*).

Salix matsudana is a tetraploid willow. Gene duplication events are a common phenomenon in the genome. There are two kinds of gene duplications in the genome: tandem duplication events (TDs) and segmental duplication events (SDs). TDs refer to two or more adjacent homologous genes located within 200 Kb on a single chromosome; SDs refer to homologous gene pairs between different chromosomes (*Cannon et al., 2004*). The gene duplication pairs were identified in TBtools by the "Blast compare 2 Seq [sets] <Big File>" and "Quick McscanX Wrapper" tools. The candidate duplicated genes should have ≥80% coverage and ≥65% similarity. The TDs of *SmAP2* genes were revealed on a chromosome using the "Amazing Gene Location from GFF3/GTF File" tool of TBtools. The SDs of *SmAP2* genes were visualized by the "Amazing Super Circos" tool of TBtools (*Chen et al., 2020*).

Divergence time calculation of duplicated genes

After BLASTn analysis of CDS sequences and obtaining duplicated gene pairs, the nonsynonymous substitution rate (Ka) and Synonymous substitution rate (Ks) of gene pairs were calculated by the "Simple Ka/Ks calculator (NG)" tool of TBtools. The divergence time was estimated with the formula: $T = Ks/2\lambda$. The clock-like rate λ value (9.1×10^{-9}) from *Populus* was used in the calculation. (*Lynch & Conery 2000*)

Collinearity analysis between *Salix matsudana* and the representative species

To demonstrate the syntenic relationships of the orthologous *SmAP2/ERF* genes obtained from *Salix matsudana* and other two selected plants (*Populus trichocarpa*, and *Salix purpurea*,), the syntenic analysis maps were constructed using the "Amazing Super Circos" tool of TBtools (*Chen et al., 2020*).

RNA sequencing and a heat map generated by hierarchical clustering

Transcriptome sequencing data of 12 samples were obtained by Illumina HiSeq sequencing. Using TopHat2 software, the clean reads were mapped to the reference genome sequence of *S.*

matsudana (*Kim et al., 2013*). Gene expression levels were estimated by fragments per kilobase of transcript per million fragments mapped (FPKM). The FPKM values of all genes from RNA sequencing were available as supplemental file S4. Differential expression analysis of two conditions/groups was performed using the DESeq R package (1.10.1). To identify DEGs, fold change ≥ 2 and false discovery rate (FDR) < 0.01 were used as screening criteria. Using the "Amazing HeatMap" tool of TBtools, a graph of the expression level of *SmAP2/ERF* family genes with hierarchical clustering was generated (*Chen et al., 2020*).

Results

Identification, phylogenetic analysis, and classification of 364 AP2/ERF TF family members in *Salix matsudana*

By HMM profile search against the *Salix matsudana* protein database, a total of 364 full-length AP2/ERF family proteins containing at least one AP2/ERF domain were identified as AP2/ERF superfamily members of *Salix matsudana* (Fig. 1). The name, protein length, molecular weight, and isoelectric point of individual genes are listed in Supplementary Table S2.

The phylogenetic relationships of SmAP2/ERF proteins were inferred by multiple sequence alignment of the AP2 domain, which included approximately 50–60 amino acids. The sequence alignment of all *AP2/ERF* genes showed some conserved amino acids at specific positions, as previously reported (Fig. S1). For example, the WLG element (58th–60th amino acids; 58–60AA) was highly conserved in in the ERF and RAV families; alternatively, in the AP2 family, the conserved sequences from 58–60AA were converted into YLG elements (*Liu et al., 2019*). In many species, these conserved amino acid profiles contribute to convincing classification of *AP2/ERF* genes. Basing on multiple sequence alignments of 48 AP2/ERF proteins from *Arabidopsis* and *Populus trichocarpa* with known categories and 364 *Salix matsudana* AP2/ERF proteins, we constructed a phylogenetic tree using the neighbor-joining method to explore the phylogenetic relationships of *Salix matsudana* AP2/ERF proteins. The phylogenetic tree showed that there were 55 *AP2/ERF* genes that belong to the AP2 family, with 47 genes that encode proteins with two AP2 domains and eight genes (*SmAP2-20*, *SmAP2-25*, *SmAP2-29*, *SmAP2-35*,

SmAP2-36, *SmAP2-40*, *SmAP2-41* and *SmAP2-55*) that encode proteins with a single AP2 domain (Fig. 1). Additionally, 301 genes that were predicted to encode proteins with a single AP2 domain were members of the ERF family. The ERF family could be further classified into two subfamilies, ERF and DREB. Of the 301 members, 166 and 135 genes belonged to the ERF and DREB subfamilies, respectively. The ERF family genes from *Salix matsudana* were distributed in B1–B6 subgroups; the DREB family genes from *Salix matsudana* were classified into A1–A6 subgroups. The gene number and percentage of each subgroup are listed in Fig. 2 and Table S3. Six putative genes were classified as RAV subgroup genes that encode proteins containing one *AP2/ERF* domain and one B3 domain (Fig. 1). Two genes were designated as Soloist genes, whose *AP2/ERF*-like domain sequences had lower homology compared with other *AP2/ERF* genes (Fig. 1).

The *AP2/ERF* genes number, classification and percentage of different subgroups from five plant species, including the model plant *Arabidopsis*, *Populus*, and two *Salix* plants, are listed in Table S3. The gene name of *AP2/ERF* genes from *Populus trichocarpa* and *Salix purpurea* are listed in Table S4. As a tetraploid plant, the total number (364) of *AP2/ERF* genes was much larger in *Salix matsudana* than in the other four species. The number of *AP2/ERF* genes in *Salix matsudana* was 2.5-, 1.8-, 1.9-, and 2.1-fold higher than those in *A. thaliana* (*Nakano et al., 2006*), *Populus trichocarpa (Zhuang et al., 2008)*, *Salix purpurea*, and *Salix arbutifolia* (*Rao et al., 2015)*, respectively. For DREB and ERF subfamilies, the percentage of all *AP2/ERF* genes in *Salix matsudana* was similar to those of *A. thaliana*, *Populus trichocarpa*, and *Salix purpurea*, and the percentages of DREB and ERF subfamilies were 38% and 45%, respectively. In *Salix arbutifolia*, the percentage of DREB (33%) was lower than that of the other four species, whereas the percentage of ERF (50.8%) was higher. In *Salix matsudana*, the percentage of the AP2 subgroup was highest among all five species (15%) and the numbers of most of gene sub-classifications were doubled, including the Soloist gene; there were two Soloist genes in the *Salix matsudana* genome. However, no duplications were observed in the RAV subgroup, and only six RAV genes were found in the *Salix matsudana* genome.

Gene structure and conserved motif analysis

To understand the structural diversity of *SmAP2/ERF* genes in different clades, a different form of phylogenetic tree of SmAP2/ERF family was constructed (Fig.3A). The intron and exon structures of *SmAP2/ERF* genes were revealed by inputting Gff3 files into TBtools (Fig.3B). A total of 55 genes of the AP2 subfamily had more exons than ERF and other subfamilies. Apart from three exons in the *SmAP2-29* and four exons in the *SmAP2-20*, other members of the AP2 subfamily contained more than seven exons. The intron number was less than three in many members of the ERF and RAV subfamilies. In total, 215 gene members did not have introns (Fig. 3B). The exon/intron structures of genes that were classified in the same clade were similar. Many gene pairs were found in the phylogenetic tree that potentially originated from allotetraploid evolution of *Salix matsudana*. Many gene pairs (approximately 70%) maintained the same or similar gene structure during *Salix matsudana* evolution, which indicated that the *SmAP2/ERF* genes were conserved at the DNA level after polyploidization.

TF proteins always contain many conserved motifs to activate gene expression. A total of 10 conserved motifs were detected in 364 SmAP2/ERF proteins using the online MEME software, and a block diagram was constructed to characterize SmAP2/ERF protein structure (Fig. 3C, Fig. S2). Motif-4, Motif-1, Motif-2, Motif-3, Motif-5, Motif-7, and Motif-9 were found in the AP2 domain regions. The Motif-5 region covered the region of Motif-4 and Motif-1, whereas Motif-7 included Motif2 and Motif3. Motif-9 is a specific motif that is only found in the second AP2 domain of the AP2 subgroup. Motif-1, Motif-2, Motif-3, and Motif-4 were detected in 90% percent of the ERF subfamily proteins. Thirty proteins of the ERF subfamily lacked one or two motifs of Motif-1–4. An extreme example is *SmERF B2-13*, which only had Motif-2. Motif-6, Motif-8, and Motif-10 are motifs located outside of the AP2 domain. Motif-6 was primarily found in the AP2 subfamily with only one exception, *SmERF B4-4*, which was in the ERF-B4 clade. In the AP2 subfamily, members with two AP2 domains had Motif-6 located between the two AP2 domains. Motif-8 was found in 69 proteins of the AP2/ERF family, and its location was adjacent to the carboxyl terminal of Motif-3. Many proteins from the DREB-A1, DREB-A4,

DREB-A5 clades had Motif-8. Motif-10 was found in 62 proteins of the AP2/ERF family, with 61 proteins from the ERF subfamily and only one from the AP2 subfamily. Motif-10 was mostly distributed on the proteins from the ERF-B3, DREB-A2, and DREB-A4 clades. The functions of these three motifs need to be elucidated by further experimental analysis.

Besides protein SmAP2-20, the entire AP2 domain was distributed in the amino terminal or in the middle position of the proteins. In the two Soloist genes, only one motif, Motif-2, was found. The conserved motif composition and gene structure of the same subfamily were similar, thus verifying the reliability of the phylogenetic tree classification..

Chromosome distribution and duplication of *SmAP2/ERF* superfamily genes

The chromosome location of the identified SmAP2/ERF genes was constructed using TBtools. In total, 310 genes from the AP2/ERF superfamily were unevenly distributed on 38 chromosomes (Fig. 4); 54 other genes located on scaffolds were not illustrated in Fig. 4. The chromosome with the largest number of AP2/ERF genes was Chr21, which had 22 genes. Only one *AP2/ERF* gene each was located on Chr14 and Chr36. On the four chromosomes Chr1, Chr3, Chr22, and Chr27, only two *AP2/ERF* genes were found. In 38 chromosomes, most of the *AP2/ERF* genes from different subgroups were arbitrarily distributed, such as five of six RAV genes located on Chr15, Chr37, Chr34, Chr31, and Chr11. Moreover, the two Soloist genes were distributed on Chr29 and Chr5. However, *SmERF* B3 subgroup members clustered together with three genes as a cluster unit. We found 12 clusters in 12 chromosomes (Fig. 4), which accounted for 62% of the whole *SmERF* B3 subgroup.

In addition, we also analyzed the tandem duplication events (TDs) of the *AP2/ERF* genes located within in the 200-kb range of chromosomal regions of the *Salix matsudana* genome. Eleven TD regions, which included 23 *SmAP2/ERF* genes, clustered into 11 linkage groups (LGs) of the *Salix matsudana* genome (Fig. 4). LGs that contained cluster repeat genes were Chr7, Chr8, Chr10, Chr13, Chr17, Chr19, Chr21, Chr24, Chr28, Chr31, and Chr33. All genes of the repeat clusters were localized within a genomic segment of less than 20 Kb; for example, SmDREB A4-20 and SmDREB A4-19 were localized on a 3.6-Kb segment of Chr24. On Chr8, three genes

clusters (SmERF B3-6, SmERF B3-7, and SmERF B3-8) located on a less than 12-Kb segment. In 11 tandem repeats, six came from the SmERF B3 subgroup, two came from the SmDREB A4 subgroup, and one each came from the SmDREB A1 and AP2 subgroups. SmERF B3-40 and SmERF B3-39 tandem repeat pairs had 97% protein sequence identity.

In addition to tandem duplications, many segmental duplication events (SDs) were found in *Salix matsudana* by MCScanX (Fig. 5, Table S5). We found a total of 28,348 collinear gene pairs (not shown) in the *Salix matsudana* genome, from which 298 *AP2/ERF* collinear gene pairs were identified. Then, Ka, Ks, and Ka/Ks ratios of these 298 *AP2/ERF* collinear gene pairs were calculated to estimate the divergence time (T value) and selection pressure among duplicated *SmAP2/ERF* gene pairs (Table S6). All of the Ka/Ks values were below 1, which indicated that these genes might have experienced strong purifying selective pressure during evolution. Among the 298 *AP2/ERF* collinear gene pairs, 198 were located on duplicated segments on 38 chromosomes in *Salix matsudana* (Fig. 5 and Table S3). The collinear gene pairs in the *Salix matsudana* genome were visualized by Circos, and the gene pairs were linked by lines (grey lines indicated all gene pairs, red lines indicated *AP2/ERF* collinear gene pairs).

Synteny analysis of *AP2/ERF* genes between *Salix matsudana* and two related Salicaceae species, *Populus trichocarpa* and *Salix purpurea*

To further infer the phylogenetic mechanisms of the SmAP2/ERF family, we constructed two comparative syntenic maps of *Salix matsudana* with two related species, *Populus trichocarpa* (Fig. 6A) and *Salix purpurea* (Fig. 6B). Collinear *AP2/ERF* genes pairs between *Salix matsudana* and other two species are listed in Supplementary Table S7. A total of 263 *SmAP2/ERF* genes showed syntenic relationships with 183 genes from *Populus trichocarpa*, and 248 *SmAP2/ERF* genes showed syntenic relationships with 144 genes from *Salix purpurea*. The number of orthologous pairs between *Salix matsudana* and *Populus trichocarpa*, and *Salix matsudana* and *Salix purpurea* were 423 and 292, respectively (Supplementary Table S7). Some *PtAP2/ERF* and *SpAP2/ERF* genes were found to be associated with at least four syntenic gene pairs. Interestingly, the number of collinear gene pairs identified between *Salix matsudana* and

Salix purpurea were less than that between *Salix matsudana* and *Populus trichocarpa*.

In the comparative syntenic map between *Salix matsudana* and *Populus trichocarpa*, syntenic links were found between all 19 *Populus trichocarpa* chromosomes and all 38 *Salix matsudana* chromosomes (Fig. 6A). Alternatively, in the comparative syntenic map between *Salix matsudana* and *Salix purpurea*, there were no syntenic links between Chr1, Chr12, and Chr36 from *Salix matsudana*, and Chr15Z and Chr15W from *Salix purpurea* (Fig. 6B).

Specific expression of AP2/ERF superfamily genes under salt stress

To investigate the physiological roles of *SmAP2/ERF* genes in salt stress tolerance, we identified the expression patterns of SmAP2, SmERF, and SmDREB subgroup genes from the RNA sequencing data. By inputting the FPKM values (Fragments Per Kilobase of transcript per Million fragments mapped) of these genes in TBtools, three heatmaps were constructed using Log10-transformed values of the FPKM values to demonstrate the expression pattern change under salt stress (Fig.7).

The expression patterns of 285 genes are illustrated in Fig. 7, and included 47 AP2 (Fig. 7A), 108 DREB (Fig. 7B), and 130 ERF subgroup genes (Fig. 7C). In the AP2 subgroup, the Log10-transformed values of 31 genes were < 3, which indicated lower expression in the root and no response to salt stress. Five genes had differential expression patterns. The expression levels of four genes (*SmAP2-38*, *SmAP2-4*, *SmAP2-3*, and *SmAP2-33*) were induced by salt stress, whereas the expression of gene *SmAP2-15* decreased after salt stress. In the DREB subgroup, 108 genes were present in the heatmap, and expression levels of 10 genes, such as *SmDREB A1-10*, *SmDREB A1-9*, and *SmDREB A1-7*, were induced by salt stress and remained higher. In the ERF subgroup heatmap, 130 genes were included. The expression levels of 13 genes were upregulated by salt stress, but only the expression of *SmERF B4-1* was higher. Three genes were downregulated by salt stress, including *SmERF B3-52*. In many paralog gene pairs, we found one gene with higher expression, whereas the other gene had lower expression, such as *SmDREB-A9*/*SmDREB-A10*, *SmAP2-33*/*SmAP2-39* and *SmERF-9*/*SmERF-10* gene pairs. Fourteen genes with upregulated expression patterns were verified by qRT-PCR (Real-time Quantitative PCR) (Fig.8).

From the results, we found that most genes' expression patterns were consistent with the FRKM values except one gene, *SmERF B3-42*. Both in 'Yanjiang' and '9901' samples, the expression level of thirteen genes was induced after salt treatment, but the expression level of seven genes (*SmAP2-33*, *SmDREBA4-24*, *SmDREBA1-4*, *SmDREBA1-7*, *SmDREBA5-23*, *SmERF B3-45* and *SmERF B4-1*) was much higher induced in '9901' than that in 'Yanjiang', which was not found in RNA sequencing results.

Discussion

Salix is one of the few woody plants with a large number of polyploid taxa, in *S. matsudana*, both tetraploid and diploid individuals have been observed (*Guo et al., 2016*). In our previous experiment, we sequenced the *Salix matsudana,* an allotetraploid salix. Tetraploid *Salix* is valuable because they have higher tolerance to abiotic stress than their diploid relatives; therefore, they can be planted beachside to alleviate soil salinity and improve the ecological environment (*Zhang et al., 2016*). The molecular mechanism of salinity response regulation is very complex, and AP2/ERF TFs are key regulators in plants (*Xie et al., 2019*). Here, we identified 364 *AP2/ERF* gene members in *Salix matsudana*, and characterized their classification, chromosome location, gene structure, and syntenic relationships of these genes within the genome and between other species. We also revealed the expression patterns under salt stress. These efforts can serve as a first step in comprehensive functional characterization of *AP2/ERF* genes by reverse genetic approaches and molecular genetics research.

As an allotetraploid species, *Salix matsudana* has more *AP2/ERF* gene members than other plants selected for comparison, including three Salicaceae family relatives (Table S3). The total number of genes is approximately double compared with poplar and two willow relatives, but the proportions of some subgroups were slightly different. *Salix arbutifolia* had a higher percentage (50.8% > 45%) of ERF subfamily members, but a lower proportion (33% < 38%) of DREB subfamily members compared with other species (Table S3). For DREB-A1 and ERF-B2 subgroups, the highest percentage was found in *Salix purpurea*, and there were the same number of or more members of these two subgroups compared with other species, including the

tetraploid *Salix matsudana*. For ERF-B3, *Salix arbutifolia* had the highest subgroup percentage (18.5%). In *Salix matsudana*, ERF-B6 had the lowest percentage (5.7%), whereas the subfamily AP2 had the highest percentage (15.1%). These data indicated that, during evolution, AP2/ERF family subgroup members probably underwent gene duplication or loss and therefore evolved into the specific AP2/ERF subgroup proportions in each species.

A phylogenetic tree that included 364 genes from *Salix matsudana* and 48 genes from *A. thaliana* and *Populus trichocarpa* was constructed (Fig.1). All subgroups were clustered together. Eight genes with one AP2 domain were classified into the AP2 family because of a close phylogenetic relationship. This classification was similar to that in *Arabidopsis*, in which four genes involved in the AP2 family contained a single AP2 domain (*Nakano et al., 2006*).

The gene intron/exon structure and conserved motifs were identified in the 364 SmAP2/ERF members. Similar to that of the AP2/ERF genes from other species, such as cauliflower and radish, the AP2 subfamily had more introns and the ERF subfamily had fewer (*Karanja et al., 2019; Li et al., 2017a*). Previous studies found that intron number and distribution are related to plant evolution, and introns of the ERF family genes were probably lost during evolution in higher plants (*Tang et al., 2016; Zhang & Li, 2018*). In total, 215 of the 301 members (70%) of the ERF family had no introns, which was a little less than that of tartary buckwheat (*Liu et al., 2019*) and also consistent with previous finding.

Through the conserved domains and motifs, TFs play roles in gene expression regulation by promoter binding, transcription activation, and protein–protein interactions (*Liu et al., 2010*). Motif analysis showed that Motif-6, Motif-8, and Motif-10 were specifically detected in different groups of the AP2/ERF subfamily; seven other motifs were all related to the AP2 domain (Fig. 3b). Motif-8 was specifically found in the DREB subgroup, such as in the DREB-A1, DREB-A4, and DREB-A5 clades. Motif-10 was mostly found distributed on proteins from the ERF-B3, DREB-A2, and DREB-A4 clades. Motif-6 was specifically located between the two AP2 domains of AP2 subgroup members. These results indicate that, although some motifs of the AP2/ERF family genes were highly conserved and involved in DNA binding, such as motifs

443 from the AP2 domain, the functions of other subgroup-specific motifs are still unknown, and
444 more work is required to clarify their regulatory functions.
445 Based on the genome assembly data, 301 genes were anchored on the 38 chromosomes (LGs),
446 but they were unevenly distributed. Eleven TDs were found on 11 chromosomes, and seven
447 tandem duplication gene pairs came from the SmERF B3 subgroup, which included three
448 duplicated genes (*SmERF B3-6*, *SmERF B3-7*, and *SmERF B3-8*) that clustered together. Apart
449 from the tandem duplication cluster, SmERF B3 members typically clustered on a chromosome,
450 with three genes as a unit. In 12 clusters, 37 *SmERF B3* genes were found. This phenomenon
451 were also found in *Populus trichocarpa*, thirteen PtERF B3 genes located in 4 clusters, which
452 indicated that in the evolution of *Salix matsudana*, apart from the chromosome duplication,
453 segmental duplication were also happened.
454 Using MCScanX, we found a total of 28,348 collinear gene pairs in the *Salix matsudana* genome,
455 from which 299 *AP2/ERF* collinear gene pairs were identified; this indicated that, during
456 evolution, the *Salix matsudana* genome experienced a whole genome duplication event.
457 Population genetic theory predicts that, after duplication, some redundant duplicate copies will
458 be silenced and eliminated, and other retained paralogs will obtain sub- or neofunction by DNA
459 mutation in coding or regulatory sequences (*Adams et al., 2003; Hou et al; 2019; Xionglei &*
460 *Jianzhi, 2005*).
461 Then, we calculated Ka, Ks, and Ka/Ks ratios of these 298 AP2/ERF collinear gene pairs to
462 estimate the divergence time and selection pressure. All Ka/Ks values were below 1, which
463 indicated that these genes might have experienced strong purifying selective pressure during
464 evolution. It was previously reported that purifying selection would lead to the loss of redundant
465 genes (*Kondrashov et al., 2002*). Based on the gene number of most subgroups, we did not find
466 any obvious evidence of gene loss, but in the RAV subgroup, there was an exception; there were
467 only six members in *Salix matsudana*, which is identical to the gene number in *Arabidopsis*.
468 Based on the gene loss hypothesis, the duplication paralogs of RAVs may have been lost during
469 genome evolution because of their rapid evolutionary rate.

Approximately 52–59 million years ago (Mya), willow and poplar, which are two modern taxa, originated from a diploid progenitor, but when and how *Salix matsudana* originated and experienced chromosome duplication remains largely unknown (Hou et al., 2016). The allotetraploid *Salix matsudana* may originated from hybridization between two diploid salix germplam and subsequently genome duplication. In our study, the divergence time (T Value) of gene pairs can be classified mainly into two time periods, 2–8 Mya and 20–36 Mya (Table S6). Gene pairs with a divergence time of 20–36 Mya were probably paralogs from two diploid hybrid parents; whereas 2–8 Mya is probably the divergence time of paralogs after whole genome duplication events. These data indicated that the two diploid progenitors of *Salix matsudana* underwent hybridization and genome duplication not more than 10 Mya.

Similar to the findings of a previous report, alignment of a Salix linkage map to the Populus genomic sequence revealed macrosynteny between willow and poplar genomes (Hanley et al. 2006) (Fig 6A, 6B). Synteny analysis of *Salix matsudana* vs *Populus trichocarpa*, and *Salix matsudana* vs *Salix purpurea* revealed 423 and 292 orthologous pairs, respectively. In total, 263 *SmAP2/ERF* genes had syntenic relationships with 183 genes in *Populus trichocarpa*, whereas 248 *SmAP2/ERF* genes showed syntenic relationships with 144 genes in *Salix purpurea*. Interestingly, the collinear gene pairs identified between *Salix matsudana* and *Salix purpurea* were less than that from *Salix matsudana* and *Populus trichocarpa*. Syntenic links were found between all 19 *Populus trichocarpa* chromosomes and all 38 *Salix matsudana* chromosomes, but there were no syntenic links between Chr1, Chr12, and Chr36 from *Salix matsudana* and Chr15Z and Chr15W from *Salix purpurea*. *Salix* has 300–500 species and considerable variation, ranging from shrubs to trees (Argus, 1997); willow may evolve faster, which would lead to them being more diverse. Researchers proposed that *Populus* might be evolutionarily more primitive than *Salix* (Dai et al. 2014, Hou et al. 2019). From our results, we could infer the evolutionary relationships of three Salicaceae species (*Populus trichocarpa*, *Salix matsudana*, and *Salix purpurea*); *Populus trichocarpa* was the most primitive taxon, *Salix purpurea* was the most derived taxon, and *Salix matsudana* was located between them but was genetically more closely

related to *Populus trichocarpa* than *Salix purpurea*.

Plants must adapt to various biotic and abiotic stresses because they are immobile in their life cycles. For example, *Salix matsudana* must adapt to the soil salinity when grown along coastal beaches. Consequently, some *AP2/ERF* TFs play important roles in plants by facilitating defense against stress and improving resistance. From the RNA sequencing data, we extracted the expression FPKM values and constructed expression heatmaps to show the expression patterns under salt stress (Fig. 7). The expression levels of four genes from the AP2 subgroup, 10 genes from the DREB subgroup, and 13 genes from the ERF subgroup were induced by salt stress, but only the expression levels of four genes were downregulated after salt stress. The expression patterns were verified by qRT-PCR. The expression pattern of many *AP2/ERF* gene pairs with evolutionary relationships differed, which indicated that the *AP2/ERF* gene family may have changed at the transcriptional regulation level following polyploidization. That finding provides additional evidence that redundant duplicated gene pairs experienced functional divergence based on expression pattern change. These differentially expressed *SmAP2/ERF* genes could be selected as candidate genes; such as SmDREB A1-4 and SmERF B3-45, further exploration on their roles under salt stress will reveal molecular mechanisms responsible for salinity stress responses in *Salix matsudana*.

In conclusion, 364 *AP2/ERF* TFs were identified in *Salix matsudana*. Clustering and phylogenetic analysis were conducted to classify these TFs into 15 subgroups. Chromosome location, gene structure, and conserved motifs were identified for 364 *AP2/ERF* TFs. Evolutionary relationships of these genes were revealed by tandem and segmental duplication gene pair identification, divergence time estimation, and T value calculation, which indicated that the progenitor of *Salix matsudana*, two diploid salix germplasms, underwent hybridization and genome duplication not more than 10 Mya. Synteny analysis with other species showed macrosynteny between willow and poplar *AP2/ERF* genes, and *Salix matsudana* was genetically more closely related to *Populus trichocarpa* than *Salix purpurea*. The *AP2/ERF* TFs were also confirmed to exhibit differential expression patterns during salt stress. The functions of these

genes should be investigated in future studies to better clarify the mechanism of salt tolerance regulation in *Salix matsudana*, which will be helpful for breeders in salt tolerance varieties selection.

Conclusion

In this study, 364 *SmAP2/ERF* genes of *Salix matsudana* were identified and renamed according to the chromosomal location of the SmAP2/ERF genes. Gene classification, gene structure and conserved motifs were analyzed in detail. Investigation results on syntenic relationships between the *SmAP2/ERF* genes and AP2/ERF genes from other species elucidated that the progenitors of *Salix matsudana* underwent whole genome duplication not more than 10 Mya and *Salix matsudana* is genetically more closely related to *Populus trichocarpa* than to *Salix purpurea*. Moreover, analyses on the differential expression patterns of *SmAP2/ERF* genes during salt stress can help to reveal the mechanism of salt tolerance regulation in *Salix matsudana*.

Acknowledgments We thank Mallory Eckstut, PhD, from Liwen Bianji, Edanz Editing China (www.liwenbianji.cn/ac), for editing the English text of a draft of this manuscript.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by the National Natural Science Foundation of China (31971681), Jiangsu Province Forestry Science and Technology Innovation and Promotion Project (LYKJ [2018]36) and Nantong University Scientific Research Start-up project for Introducing Talents (18R08).

Grant Disclosures

The following grant information was disclosed by the authors:

National Natural Science Foundation of China: 31971681.

Jiangsu Province Forestry Science and Technology Innovation and Promotion Project: LYKJ [2018]36.

Nantong University Scientific Research Start-up project for Introducing Talents: 18R08 .

Competing Interests

The authors declare there are no competing interests.

Author Contributions

Yanhong Chen conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.

Jian Zhang, Yuna Jiang conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.

Shizheng Shi conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, approved the final draft.

Fei Zhong, Guoyuan Liu, Chunmei Yu, Bolin Lian performed the experiments, contributed reagents/materials/analysis tools, authored or reviewed drafts of the paper, approved the final draft.

Data Availability

The following information was supplied regarding data availability:
The latest S. matsudana genome including annotation gff file, coding sequences cds file and protein sequences pep file are available at our laboratory website (https://js-garden.cn/?list_6/99.html); the FPKM values of all genes from RNA sequencing were available as supplemental file S4; the raw data for quantitative real-time PCR were supplied as supplemental file S5.

REFERENCES

Abogadallah GM, Nada RM, Malinowski R, Quick P. 2011. Overexpression of HARDY , an AP2/ERF gene from Arabidopsis, improves drought and salt tolerance by reducing transpiration and sodium uptake in transgenic Trifolium al.exandrinum L. *Planta* **233 (6)**: 1265-1276 DOI 10.1007/s00425-011-1382-3.

Adams KL, Richard C, Ryan P, Wendel JF. 2003. Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proceedings of the National Academy of Sciences of the United States of America* **100(8)**: 4649-4654 DOI 10.1073/pnas.0630618100.

Argus GW. 1997. Infrageneric classification of Salix (Salicaceae) in the new world. Systematic *Botany Monographs* **52**: 1-121 DOI 10.2307/25096638.

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research* **37**: W202-208 DOI 10.1093/nar/gkp335.

Brendel V, Kurtz S, Walbot V. 2002. Comparative genomics of Arabidopsis and maize:

prospects and limitations. *Genome Biology* **3 (3)**: reviews1005.1–1005.6 DOI 10.1186/gb-2002-3-3-reviews1005.

Cannon SB, Mitra A, Baumgarten A, Young ND, May G. 2004. The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana. *BMC Plant Biology* **4**: 10 DOI 10.1186/1471-2229-4-10.

Cao B, Shu L, Li A. 2019. Functional characterization of LkERF-B2 for improved salt tolerance ability in Arabidopsis thaliana. *3 Biotech* **9 (7)**: 263 DOI 10.1007/s13205-019-1793-6.

Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. 2020. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Molecular Plant* **13(8)**:1194-1202 DOI 10.1016/j.molp.2020.06.009.

Chen YH, Cao YY, Wang LJ, Li LM, Yang J, Zou MX. 2018. Identification of MYB transcription factor genes and their expression during abiotic stresses in maize. *Biologia Plantarum* **62**: 222-230 DOI 10.1007/s10535-017-0756-1.

Dai X, Hu Q, Cai Q, Feng K, Ye N, Tuskan GA, Milne R, Chen Y, Wan Z, Wang Z, Luo W, Wang K, Wan D, Wang M, Wang J, Liu J, Yin T. 2014. The willow genome and divergent evolution from poplar after the common genome duplication. *Cell Research* **24 (10)**: 1274-1277 DOI 10.1038/cr.2014.83.

Duran-Medina Y, Serwatowska J, Reyes-Olalde JI, de Folter S, Marsch-Martinez N. 2017. The AP2/ERF transcription factor DRNL modulates gynoecium development and affects Its response to cytokinin. *Frontier in Plant Science* **8**: 1841 DOI 10.3389/fpls.2017.01841.

Gao Y, Liu Y, Liang Y, Lu J, Jiang C, Fei Z, Jiang CZ, Ma C, Gao J. 2019. Rosa hybrida RhERF1 and RhERF4 mediate ethylene- and auxin-regulated petal abscission by influencing pectin degradation. *Plant Journal* **99 (6)**: 1159-1171 DOI 10.1111/tpj.14412.

Gu C, Guo ZH, Hao PP, Wang GM, Jin ZM, Zhang SL. 2017. Multiple regulatory roles of AP2/ERF transcription factor in angiosperm. *Botanical Studies* **58 (1)**: 6 DOI 10.1186/s40529-016-0159-1.

Guo W, Hou J, Yin T, Chen Y. 2016. An analytical toolkit for polyploid willow discrimination. *Scientific Reports* **6**: 37702 DOI 10.1038/srep37702.

Hall AE, Fiebig A, Preuss D. 2002. Beyond the Arabidopsis genome: opportunities for comparative genomics. *Plant Physiology* **129 (4)**: 1439-1447 DOI 10.2307/4280576.

Hanley SJ, Mallott MD, Karp A. 2006. Alignment of a Salix linkage map to the Populus genomic sequence reveals macrosynteny between willow and poplar genomes. *Tree Genetics & Genomes* **3**: 35-48 DOI 10.1007/s11295-006-0049-x.

Hou J, Wei S, Pan H, Zhuge Q, Yin T. 2019. Uneven selection pressure accelerating divergence of Populus and Salix. *Horticulture Research* **6**: 37 DOI 10.1038/s41438-019-0121-y.

Hou J, Ye N, Dong Z, Lu M, Li L, Yin T. 2016. Major chromosomal rearrangements distinguish willow and poplar after the ancestral "Salicoid" genome duplication. *Genome Biology and Evolution* **8 (6)**: 1868-1875 DOI 10.1093/gbe/evw127.

Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G. 2015. GSDS 2.0: an upgraded gene feature

visualization server. *Bioinformatics* **31 (8)**: 1296-1297 DOI 10.1093/bioinformatics/btu817.

Jin T, Chang Q, Li W, Yin D, Li Z, Wang D, Liu B, Liu L. 2010. Stress-inducible expression of GmDREB1 conferred salt tolerance in transgenic alfalfa. *Plant Cell, Tissue and Organ Culture* **100**: 219-227 DOI 10.1007/s11240-009-9628-5.

Karanja BK, Xu L, Wang Y, Tang M, M'mbone Muleke E, Dong J, Liu L. 2019. Genome-wide characterization of the AP2/ERF gene family in radish (Raphanus sativus L.): Unveiling evolution and patterns in response to abiotic stresses. *Gene* **718**: 144048 DOI 10.1016/j.gene.2019.144048

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* **14(4)**: R36 DOI 10.1186/gb-2013-14-4-r36.

Kondrashov FA, Rogozin IB, Wolf YI, Koonin EV. 2002. Selection in the evolution of gene duplications. *Genome Biology* **3 (2)**: Research0008 DOI 10.1186/gb-2002-3-2-research0008.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution* **33 (7):** 1870-1874 DOI 10.1093/molbev/msw054.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and clustal X version 2.0. *Bioinformatics* **23(21)**: 2947-2948 DOI 10.1093/bioinformatics/btm404.

Li H, Wang Y, Wu M, Li L, Li C, Han Z, Yuan J, Chen C, Song W, Wang C. 2017a. Genome-Wide Identification of AP2/ERF Transcription Factors in Cauliflower and Expression Profiling of the ERF Family under Salt and Drought Stresses. *Frontier in Plant Science* **8**: 946. DOI 10.3389/fpls.2017.00946.

Li H, Xiao Q, Zhang C, Du J, Li X, Huang H, Wei B, Li Y, Yu G, Liu H, Hu Y, Liu Y, Zhang J, Huang Y. 2017b. Identification and characterization of transcription factor ZmEREB94 involved in starch synthesis in maize. *Journal of Plant Physiology* **216**: 11-16 DOI 10.1016/j.jplph.2017.04.016.

Li MY, Liu JX, Hao JN, Feng K, Duan AQ, Yang QQ, Xu ZS, Xiong AS. 2019a. Genomic identification of AP2/ERF transcription factors and functional characterization of two cold resistance-related AP2/ERF genes in celery (Apium graveolens L.). *Planta* **250 (4)**: 1265-1280 DOI 10.1007/s00425-019-03222-2.

Li Y, Zhang H, Zhang Q, Liu Q, Zhai H, Zhao N, He S. 2019b. An AP2/ERF gene, IbRAP2-12, from sweetpotato is involved in salt and drought tolerance in transgenic Arabidopsis. *Plant Science* **281**: 19-30 DOI 10.1016/j.plantsci.2019.01.009.

Liu L, White MJ, Macrae TH. 2010. Transcription factors and their genes in higher plants functional domains, evolution and regulation. *Febs Journal* **262(2)**: 247-257 DOI 10.1046/j.1432-1327.1999.00349.x.

Liu M, Sun W, Ma Z, Zheng T, Huang L, Wu Q, Zhao G, Tang Z, Bu T, Li C, Chen H.

669 **2019.** Genome-wide investigation of the AP2/ERF gene family in tartary buckwheat
670 (Fagopyum Tataricum). *BMC Plant Biology* **19 (1)**: 84 DOI 10.1186/s12870-019-1681-6.
671 **Lynch M, Conery JS. 2000.** The evolutionary fate and consequences of duplicate genes.
672 *Science* **290 (5494)**: 1151-1155 DOI 10.1126/science.290.5494.1151.
673 **Nakano T, Suzuki K, Fujimura T, Shinshi H. 2006.** Genome-wide analysis of the ERF gene
674 family in Arabidopsis and rice. *Plant Physiology* **140 (2)**: 411-432 DOI
675 10.1104/pp.105.073783.
676 **Rao G, Sui J, Zeng Y, He C, Zhang J. 2015.** Genome-wide analysis of the AP2/ERF gene
677 family in Salix arbutifolia. FEBS open bio **5**: 132-137 DOI 10.1016/j.fob.2015.02.002.
678 **Schranz ME, Song BH, Windsor AJ, Mitchell-Olds T. 2007.** Comparative genomics in the
679 Brassicaceae: a family-wide perspective. *Current Opinion in Plant Biology* **10 (2)**: 168-
680 175 DOI 10.1016/j.pbi.2007.01.014.
681 **Shen SL, Yin XR, Zhang B, Xie XL, Jiang Q, Grierson D, Chen KS. 2016.** CitAP2.10
682 activation of the terpene synthase CsTPS1 is associated with the synthesis of (+)-
683 valencene in 'Newhall' orange. *Journal Of Experimental Botany* **67 (14)**: 4105-4115 DOI
684 10.1093/jxb/erw189.
685 **Song X, Wang J, Ma X, Li Y, Lei T, Wang L, Ge W, Guo D, Wang Z, Li C, Zhao J, Wang**
686 **X. 2016.** Origination, Expansion, Evolutionary Trajectory, and Expression Bias of
687 AP2/ERF Superfamily in Brassica napus. *Frontier in Plant Science* **7**: 1186 DOI
688 10.3389/fpls.2016.01186.
689 **Srivastava R, Kumar R. 2018.** The expanding roles of APETALA2/Ethylene Responsive
690 Factors and their potential applications in crop improvement. *Briefings in Functional*
691 *Genomics* **18 (4)**: 240-254 DOI 10.1093/bfgp/elz001.
692 **Tang Y, Qin S, Guo Y, Chen Y, Wu P, Chen Y, Li M, Jiang H, Wu G. 2016.** Genome-
693 wideanalysis of the AP2/ERF gene family in physic Nut and overexpression of the
694 JcERF011 gene in rice increased its sensitivity to salinity stress. *Plos One* **11 (3)**:
695 e0150879 DOI 10.1371/journal.pone.0150879.
696 **Wessels B, Seyfferth C, Escamez S, Vain T, Antos K, Vahala J, Delhomme N, Kangasjärvi**
697 **J, Eder M, Felten J, Tuominen H. 2019.** An AP2/ERF transcription factor ERF139
698 coordinates xylem cell expansion and secondary cell wall deposition. *New Phytologist*
699 **224 (4)**: 1585-1599 DOI 10.1111/nph.15960.
700 **Xie Z, Nolan TM, Jiang H, Yin Y. 2019.** AP2/ERF transcription factor regulatory networks in
701 hormone and abiotic stress responses in Arabidopsis. *Frontier in Plant Science* **10**: 228
702 DOI 10.3389/fpls.2019.00228.
703 **He X, Zhang J. 2005.** Rapid subfunctionalization accompanied by prolonged and substantial
704 neofunctionalization in duplicate gene evolution. *Genetics* **169(2)**: 1157-1164 DOI
705 10.1534/genetics.104.037051.
706 **Zhang J, Yuan H, Li M, Li Y, Wang Y, Ma X, Zhang Y, Tan F, Wu R. 2016.** A high-density
707 genetic map of tetraploid Salix matsudana using specific length amplified fragment
708 sequencing (SLAF-seq). *PLoS One* **11(6)**: e0157777 DOI 10.1371/journal.pone.0157777.
709 **Zhang J, Yuan H, Yang Q, Li M, Wang Y, Li Y, Ma X, Tan F, Wu R. 2017.** The genetic

architecture of growth traits in Salix matsudana under salt stress. *Horticulture Research* **4**: 17024 DOI 10.1038/hortres.2017.24.

Zhang Z, Li X. 2018. Genome-wide identification of AP2/ERF superfamily genes and their expression during fruit ripening of Chinese jujube. *Scientific Reports* **8 (1)**: 15612 DOI 10.1038/s41598-018-33744-w

Zhuang J, Cai B, Peng RH, Zhu B, Jin XF, Xue Y, Gao F, Fu XY, Tian YS, Zhao W, Qiao YS, Zhang Z, Xiong AS, Yao QH. 2008. Genome-wide analysis of the AP2/ERF gene family in Populus trichocarpa. *Biochemical and Biophysical Research Communications* **371 (3)**: 468-474 DOI 10.1016/j.bbrc.2008.04.087.

720 **Figure Legends**

721 **Figure. 1 Unrooted phylogenetic tree and classification of 364 *SmAP2/ERF* genes and their**
722 **representative orthologs from Arabidopsis and Populus.**
723 The amino acid sequences of AP2 domains from 364 *SmAP2/ERF* proteins and 48 orthologs
724 from Arabidopsis and Populus were aligned by ClustalW, and the neighbor-joining tree was
725 constructed using MEGA 7.0 with 1000 bootstrap replicates. The evolutionary distances were
726 computed using the p-distance method. In total, 364 *SmAP2/ERF* members were classified into
727 15 smaller subgroups, and their names are labeled beside the tree.
728 **Figure. 2 Classification and subgroup proportions of *SmAP2/ERF* family genes.**
729 The size of each piece is proportional to the relative abundance to the *SmAP2/ERF* genes
730 assigned to this group.
731 **Figure. 3 Phylogenetic relationships, gene structure, and architecture of conserved protein**
732 **motifs in *SmAP2/ERF* superfamily memebers.**
733 A, The phylogenetic tree was constructed based on the amino acid sequences of the AP2 domain
734 from 364 *SmAP2/ERF* proteins using MEGA7.0. B, Motif composition of tartary buckwheat
735 *AP2/ERF* proteins. Motifs 1–10 are displayed in different colored boxes. The sequence
736 information for each motif is provided in Fig. S2. C, Exon/intron structure of *SmAP2/ERF* genes.
737 Yellow boxes indicate untranslated 5′- and 3′-regions; green boxes indicate exons; black lines
738 indicate introns. The protein length can be estimated using the scale at the bottom.
739 **Figure. 4 Schematic representations for the chromosomal distribution and tandem**
740 **duplication of SmAP2/ERF genes.**
741 The red lines indicate tandem duplicated AP2/ERF gene pairs, which are indicated in green. The
742 SmERF-B3 subgroup members labeled with blue clustered on the same chromosome. The
743 chromosome number is indicated to the left of each chromosome. The orange color columns
744 outside the circle indicate gene density, the deeper color in the column, the higher gene density
745 in this chromosome region part.
746 **Figure. 5 Schematic representations of the segmental duplication and interchromosomal**
747 **relationships of *SmAP2/ERF* genes.**
748 Grey lines indicate all syntenic gene pairs in the *Salix matsudana* genome, red lines indicate
749 syntenic relationships between *SmAP2/ERF* genes. The orange color columes outside of the
750 circle indicated the gene density on each 38 chromosomes. The deeper color means the higher
751 density of genes.
752 **Figure. 6 Synteny analysis of *AP2/ERF* genes between *Salix matsudana* and two related**
753 **Salicaceae species, *Populus trichocarpa* and *Salix purpurea*.**
754 A, Synteny analysis of AP2 genes between *Salix matsudana* and *Populus trichocarpa*. B.

Synteny analysis of AP2 genes between *Salix matsudana* and *Salix purpurea*
Gray lines in the background indicate the collinear blocks within *Salix matsudana* and other plant genomes, whereas red lines highlight syntenic *AP2/ERF* gene pairs.

Figure. 7 Hierarchical clustering of AP2 genes and heatmap that demonstrates the differential expression patterns of *SmAP2/ERF* genes in roots before and after salt stress.
The Log10-transformed expression values of the FPKM values of 12 samples were used to draw the heat map. The color bar indicates the gene expression level. A, Heatmap and hierarchical clustering representation of 47 AP2 members. B, Heatmap and hierarchical clustering representation of 108 DREB members. C, Heatmap and hierarchical clustering representation of 130 ERF members.

Figure. 8 Verification of the *SmAP2/ERF* genes with differentially expressed patterns under salt stress by quantitative real-time PCR
For salt stress, 'Yanjiang' and '9901' roots that were subjected to 20 days of hydroponic culture and then treated with 150 mM sodium chloride for 4 h. The control was an untreated 'Yanjiang' sample. Three biological replicates for each sample were performed, and bars represent the standard deviations of the mean. Asterisks on top of the bars indicate statistically significant differences between stress treatment and the control (*$0.01 < P < 0.05$; **$p<0.01$, Student's t-test). Gene expression profiles were evaluated using the $2^{-\Delta\Delta Ct}$ method, and the control value was normalized to 1. YJ-CK and YJ-NT indicate the 'Yanjiang' roots treated with water and 150 mM NaCl respectively; 9901-CK and 9901-NT indicate '9901' roots treated with water and 150 mM NaCl respectively.

Figure 1

Figure. 1 Unrooted phylogenetic tree and classification of 364 *SmAP2/ERF* genes and their representative orthologs from Arabidopsis and Populus.

The amino acid sequences of AP2 domains from 364 *SmAP2/ERF* proteins and 48 orthologs from Arabidopsis and Populus were aligned by ClustalW, and the neighbor-joining tree was constructed using MEGA 7.0 with 1000 bootstrap replicates. The evolutionary distances were computed using the p-distance method. In total, 364 *SmAP2/ERF* members were classified into 15 smaller subgroups, and their names are labeled beside the tree.

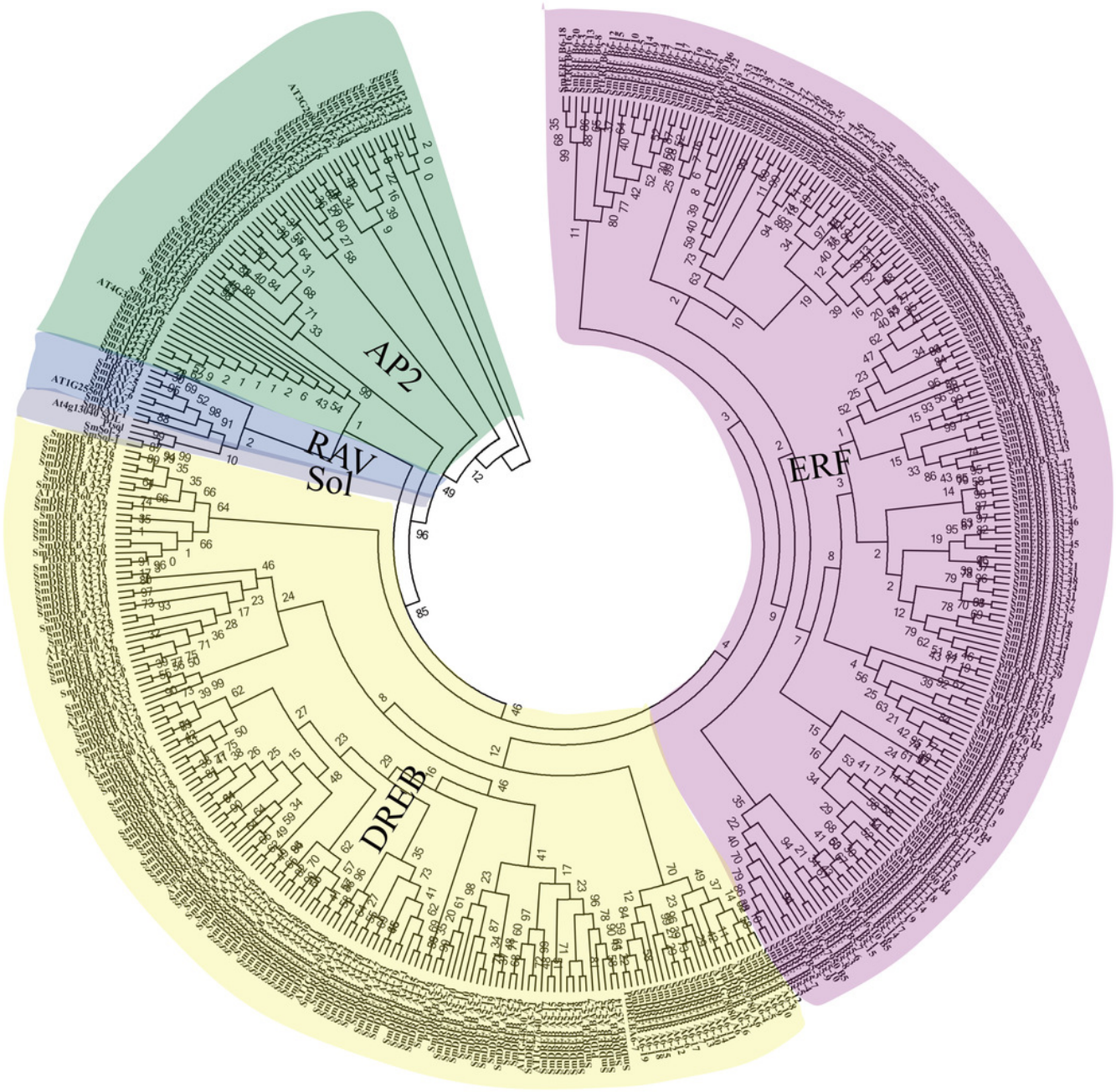

AP2
RAV
Sol
ERF
DREB

Figure 2

Figure. 2 Classification and subgroup proportions of *SmAP2/ERF* family genes.

The size of each piece is proportional to the relative abundance to the *SmAP2/ERF* genes assigned to this group.

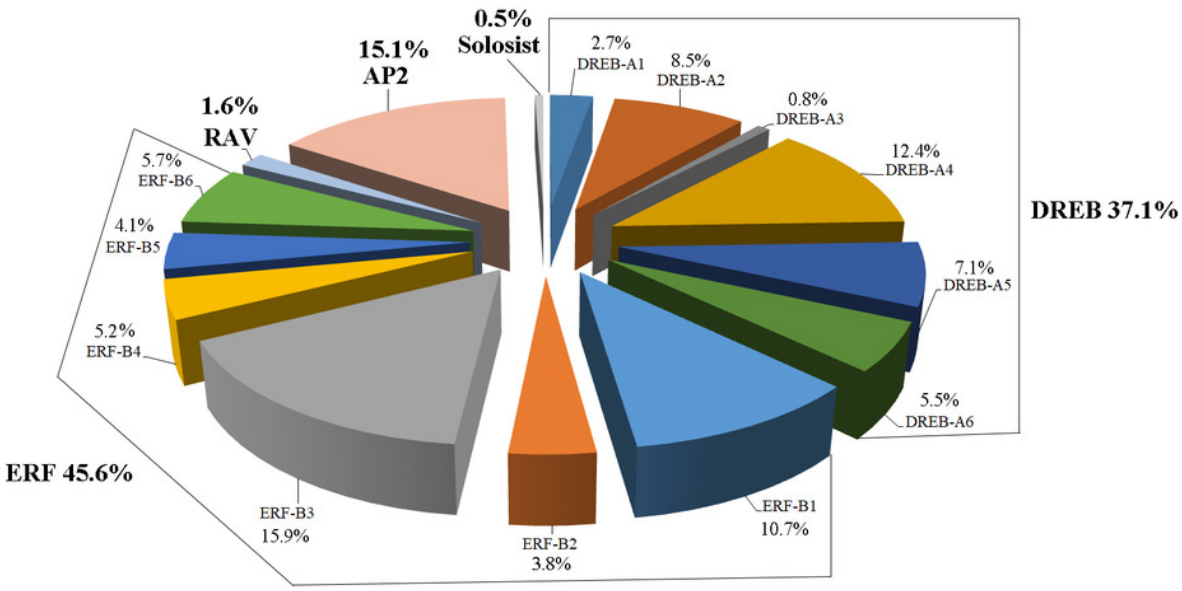


Figure 3

Figure. 3 Phylogenetic relationships, gene structure, and architecture of conserved protein motifs in *SmAP2/ERF* superfamily memebers.

A, The phylogenetic tree was constructed based on the amino acid sequences of the AP2 domain from 364 *SmAP2/ERF* proteins using MEGA7.0. B, Motif composition of tartary buckwheat *AP2/ERF* proteins. Motifs 1–10 are displayed in different colored boxes. The sequence information for each motif is provided in Fig. S2. C, Exon/intron structure of *SmAP2/ERF* genes. Yellow boxes indicate untranslated 5′- and 3′-regions; green boxes indicate exons; black lines indicate introns. The protein length can be estimated using the scale at the bottom.

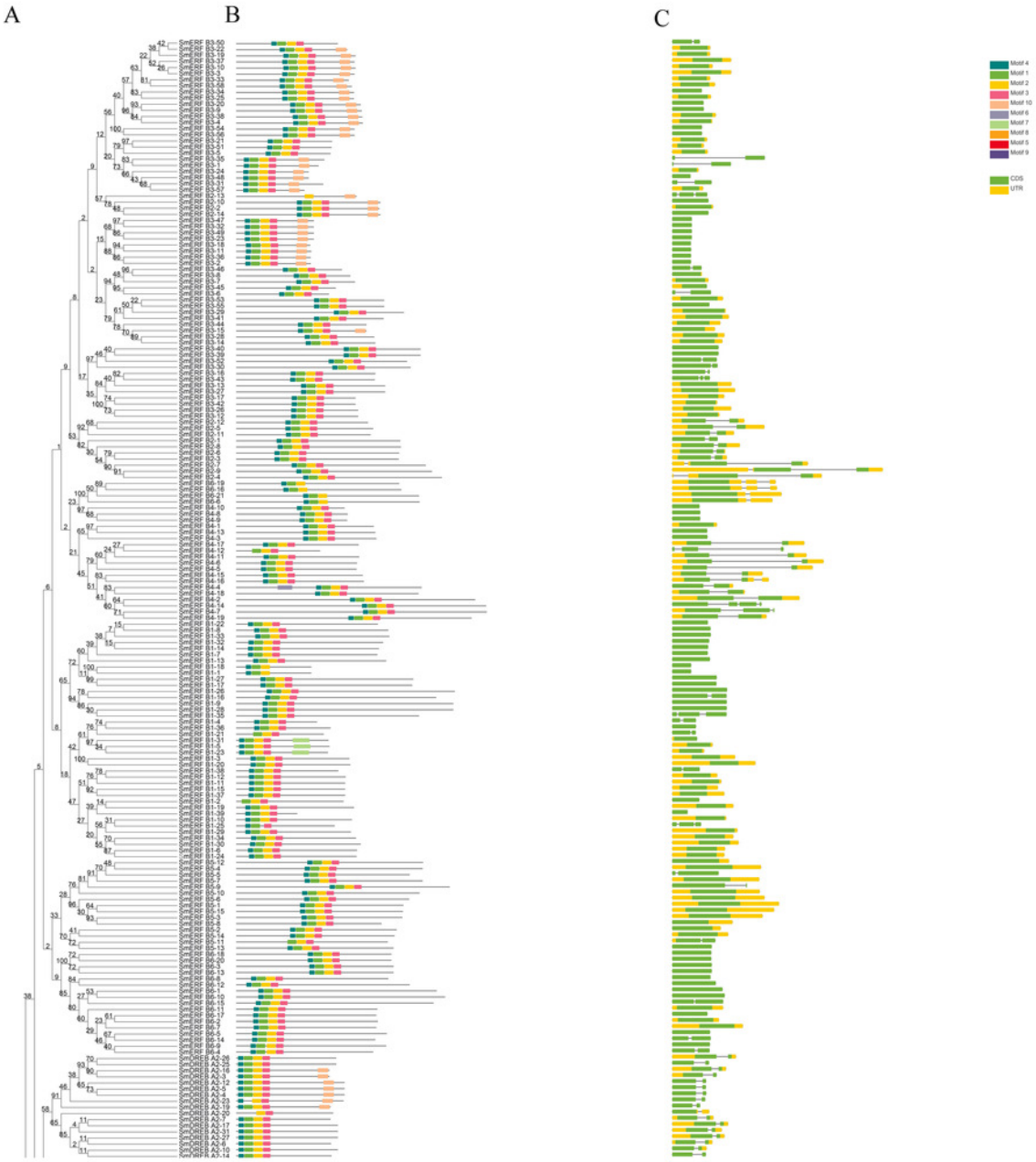

A
B
C
Motif 4
Motif 1
Motif 2
Motif 3
Motif 10
Motif 6
Motif 7
Motif 8
Motif 5
Motif 9
CDS
UTR

Figure 4

Figure 3 Phylogenetic relationships, gene structure, and architecture of conserved protein motifs in *SmAP2/ERF* superfamily memebers.(CONTINUED)

A, The phylogenetic tree was constructed based on the amino acid sequences of the AP2 domain from 364 *SmAP2/ERF* proteins using MEGA7.0. B, Motif composition of tartary buckwheat *AP2/ERF* proteins. Motifs 1–10 are displayed in different colored boxes. The sequence information for each motif is provided in Fig. S2. C, Exon/intron structure of *SmAP2/ERF* genes. Yellow boxes indicate untranslated 5′- and 3′-regions; green boxes indicate exons; black lines indicate introns. The protein length can be estimated using the scale at the bottom.

Fig. 3 Continued

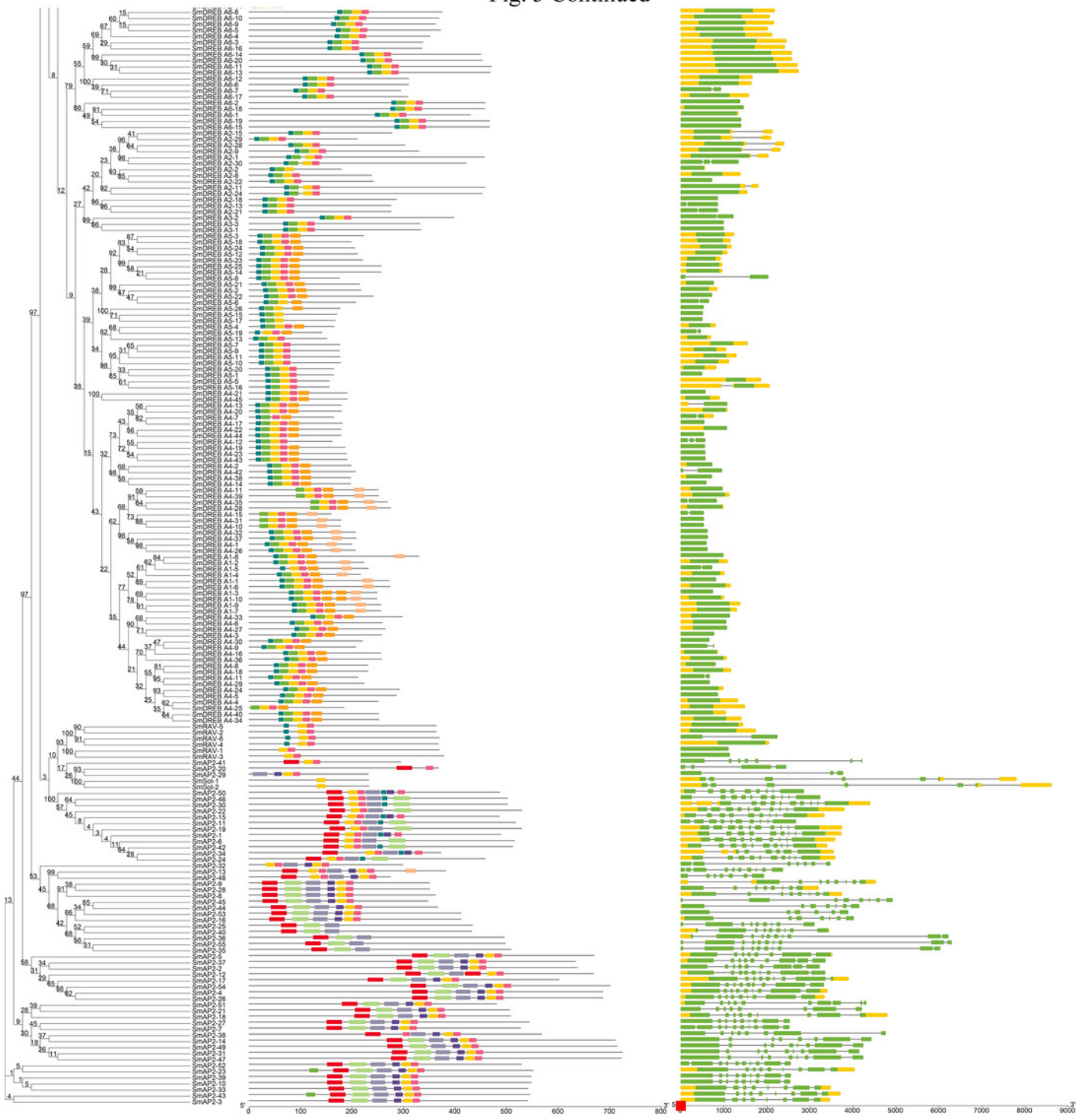


Figure 5

Figure. 4 Schematic representations for the chromosomal distribution and tandem duplication of SmAP2/ERF genes.

Figure. 4 Schematic representations for the chromosomal distribution and tandem duplication of SmAP2/ERF genes. The red lines indicate tandem duplicated AP2/ERF gene pairs, which are indicated in green. The SmERF-B3 subgroup members labeled with blue clustered on the same chromosome. The chromosome number is indicated to the left of each chromosome. The orange color columns outside the circle indicate gene density, the deeper color in the column, the higher gene density in this chromosome region part.

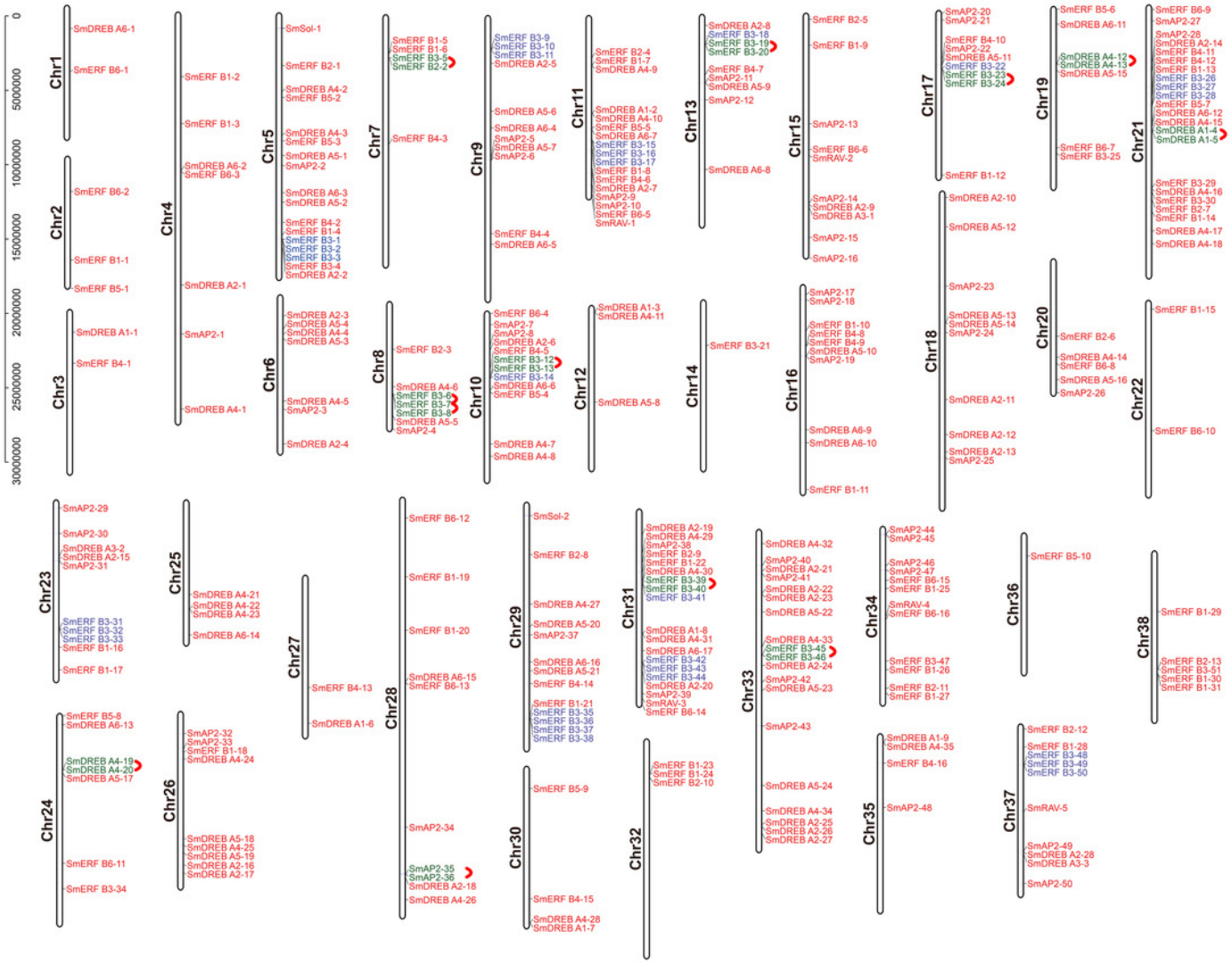

0
5000000
10000000
15000000
20000000
25000000
30000000
Chr1
SmDREB A6-1
SmERF B6-1
Chr2
SmERF B6-2
SmERF B1-1
SmERF B5-1
Chr3
SmDREB A1-1
SmERF B4-1
Chr4
SmERF B1-2
SmERF B1-3
SmDREB A6-2
SmERF B6-3
SmDREB A2-1
SmAP2-1
SmDREB A4-1
Chr5
SmSol-1
SmERF B2-1
SmDREB A4-2
SmERF B5-2
SmDREB A4-3
SmERF B5-3
SmDREB A5-1
SmAP2-2
SmDREB A6-3
SmDREB A5-2
SmERF B4-2
SmERF B1-4
SmERF B3-1
SmERF B3-2
SmERF B3-3
SmERF B3-4
SmDREB A2-2
Chr6
SmDREB A2-3
SmDREB A5-4
SmDREB A4-4
SmDREB A5-3
SmDREB A4-5
SmAP2-3
SmDREB A2-4
Chr7
SmERF B1-5
SmERF B1-6
SmERF B3-5
SmERF B2-2
SmERF B4-3
Chr8
SmERF B2-3
SmDREB A4-6
SmERF B3-6
SmERF B3-7
SmERF B3-8
SmDREB A5-5
SmAP2-4
Chr9
SmERF B3-9
SmERF B3-10
SmERF B3-11
SmDREB A2-5
SmDREB A5-6
SmDREB A6-4
SmAP2-5
SmDREB A5-7
SmAP2-6
SmERF B4-4
SmDREB A6-5
Chr10
SmERF B6-4
SmAP2-7
SmAP2-8
SmDREB A2-6
SmERF B4-5
SmERF B3-12
SmERF B3-13
SmERF B3-14
SmDREB A6-6
SmERF B5-4
SmDREB A4-7
SmDREB A4-8
Chr11
SmERF B2-4
SmERF B1-7
SmDREB A4-9
SmDREB A1-2
SmDREB A4-10
SmERF B5-5
SmDREB A6-7
SmERF B3-15
SmERF B3-16
SmERF B3-17
SmERF B1-8
SmERF B4-6
SmDREB A2-7
SmAP2-9
SmAP2-10
SmERF B6-5
SmRAV-1
Chr12
SmDREB A1-3
SmDREB A4-11
SmDREB A5-8
Chr13
SmDREB A2-8
SmERF B3-18
SmERF B3-19
SmERF B3-20
SmERF B4-7
SmAP2-11
SmDREB A5-9
SmAP2-12
SmDREB A6-8
Chr14
SmERF B3-21
Chr15
SmERF B2-5
SmERF B1-9
SmAP2-13
SmERF B6-6
SmRAV-2
SmAP2-14
SmDREB A2-9
SmDREB A3-1
SmAP2-15
SmAP2-16
Chr16
SmAP2-17
SmAP2-18
SmERF B1-10
SmERF B4-8
SmERF B4-9
SmDREB A5-10
SmAP2-19
SmDREB A6-9
SmDREB A6-10
SmERF B1-11
Chr17
SmAP2-20
SmAP2-21
SmERF B4-10
SmAP2-22
SmDREB A5-11
SmERF B3-22
SmERF B3-23
SmERF B3-24
SmERF B1-12
Chr18
SmDREB A2-10
SmDREB A5-12
SmAP2-23
SmDREB A5-13
SmDREB A5-14
SmAP2-24
SmDREB A2-11
SmDREB A2-12
SmDREB A2-13
SmAP2-25
Chr19
SmERF B5-6
SmDREB A6-11
SmDREB A4-12
SmDREB A4-13
SmDREB A5-15
SmERF B6-7
SmERF B3-25
Chr20
SmERF B2-6
SmDREB A4-14
SmERF B6-8
SmDREB A5-16
SmAP2-26
Chr21
SmERF B6-9
SmAP2-27
SmAP2-28
SmDREB A2-14
SmERF B4-11
SmERF B4-12
SmERF B1-13
SmERF B3-26
SmERF B3-27
SmERF B3-28
SmERF B5-7
SmDREB A6-12
SmDREB A4-15
SmDREB A1-4
SmDREB A1-5
SmERF B3-29
SmDREB A4-16
SmERF B3-30
SmERF B2-7
SmERF B1-14
SmDREB A4-17
SmDREB A4-18
Chr22
SmERF B1-15
SmERF B6-10
Chr23
SmAP2-29
SmAP2-30
SmDREB A3-2
SmDREB A2-15
SmAP2-31
SmERF B3-31
SmERF B3-32
SmERF B3-33
SmERF B1-16
SmERF B1-17
Chr24
SmERF B5-8
SmDREB A6-13
SmDREB A4-19
SmDREB A4-20
SmDREB A5-17
SmERF B6-11
SmERF B3-34
Chr25
SmDREB A4-21
SmDREB A4-22
SmDREB A4-23
SmDREB A6-14
Chr26
SmAP2-32
SmAP2-33
SmERF B1-18
SmDREB A4-24
SmDREB A5-18
SmDREB A4-25
SmDREB A5-19
SmDREB A2-16
SmDREB A2-17
Chr27
SmERF B4-13
SmDREB A1-6
Chr28
SmERF B6-12
SmERF B1-19
SmERF B1-20
SmDREB A6-15
SmERF B6-13
SmAP2-34
SmAP2-35
SmAP2-36
SmDREB A2-18
SmDREB A4-26
Chr29
SmSol-2
SmERF B2-8
SmDREB A4-27
SmDREB A5-20
SmAP2-37
SmDREB A6-16
SmDREB A5-21
SmERF B4-14
SmERF B1-21
SmERF B3-35
SmERF B3-36
SmERF B3-37
SmERF B3-38
Chr30
SmERF B5-9
SmERF B4-15
SmDREB A4-28
SmDREB A1-7
Chr31
SmDREB A2-19
SmDREB A4-29
SmAP2-38
SmERF B2-9
SmERF B1-22
SmDREB A4-30
SmERF B3-39
SmERF B3-40
SmERF B3-41
SmDREB A1-8
SmDREB A4-31
SmDREB A6-17
SmERF B3-42
SmERF B3-43
SmERF B3-44
SmDREB A2-20
SmAP2-39
SmRAV-3
SmERF B6-14
Chr32
SmERF B1-23
SmERF B1-24
SmERF B2-10
Chr33
SmDREB A4-32
SmAP2-40
SmDREB A2-21
SmAP2-41
SmDREB A2-22
SmDREB A2-23
SmDREB A5-22
SmDREB A4-33
SmERF B3-45
SmERF B3-46
SmDREB A2-24
SmAP2-42
SmDREB A5-23
SmAP2-43
SmDREB A5-24
SmDREB A4-34
SmDREB A2-25
SmDREB A2-26
SmDREB A2-27
Chr34
SmAP2-44
SmAP2-45
SmAP2-46
SmAP2-47
SmERF B6-15
SmERF B1-25
SmRAV-4
SmERF B6-16
SmERF B3-47
SmERF B1-26
SmERF B2-11
SmERF B1-27
Chr35
SmDREB A1-9
SmDREB A4-35
SmERF B4-16
SmAP2-48
Chr36
SmERF B5-10
Chr37
SmERF B2-12
SmERF B1-28
SmERF B3-48
SmERF B3-49
SmERF B3-50
SmRAV-5
SmAP2-49
SmDREB A2-28
SmDREB A3-3
SmAP2-50
Chr38
SmERF B1-29
SmERF B2-13
SmERF B3-51
SmERF B1-30
SmERF B1-31

Figure 6

Figure. 5 Schematic representations of the segmental duplication and interchromosomal relationships of *SmAP2/ERF* genes.

Figure. 5 Schematic representations of the segmental duplication and interchromosomal relationships of *SmAP2/ERF* genes. Grey lines indicate all syntenic gene pairs in the *Salix matsudana* genome, red lines indicate syntenic relationships between *SmAP2/ERF* genes. The orange color columes outside of the circle indicated the gene density on each 38 chromosomes. The deeper color means the higher density of genes.

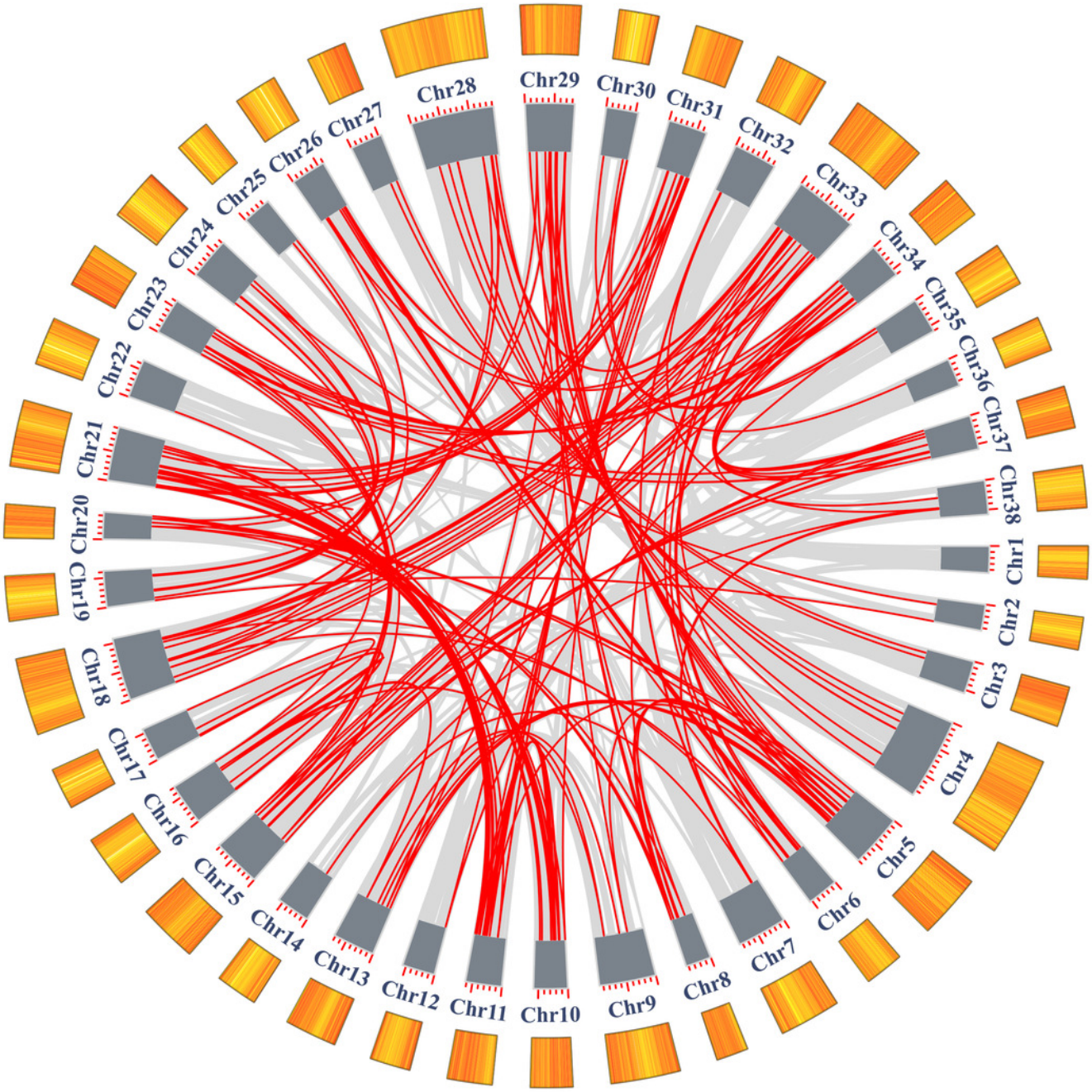

Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Chr8
Chr9
Chr10
Chr11
Chr12
Chr13
Chr14
Chr15
Chr16
Chr17
Chr18
Chr19
Chr20
Chr21
Chr22
Chr23
Chr24
Chr25
Chr26
Chr27
Chr28
Chr29
Chr30
Chr31
Chr32
Chr33
Chr34
Chr35
Chr36
Chr37
Chr38

Figure 7

Figure. 6 Synteny analysis of *AP2/ERF* genes between *Salix matsudana* and two related Salicaceae species, *Populus trichocarpa* and *Salix purpurea*.

A, Synteny analysis of AP2 genes between Salix matsudana *and* Populus trichocarpa*. B. Synteny analysis of AP2 genes between* Salix matsudana *and* Salix purpurea *Gray lines in the background indicate the collinear blocks within* Salix matsudana *and other plant genomes, whereas red lines highlight syntenic* AP2/ERF *gene pairs.*

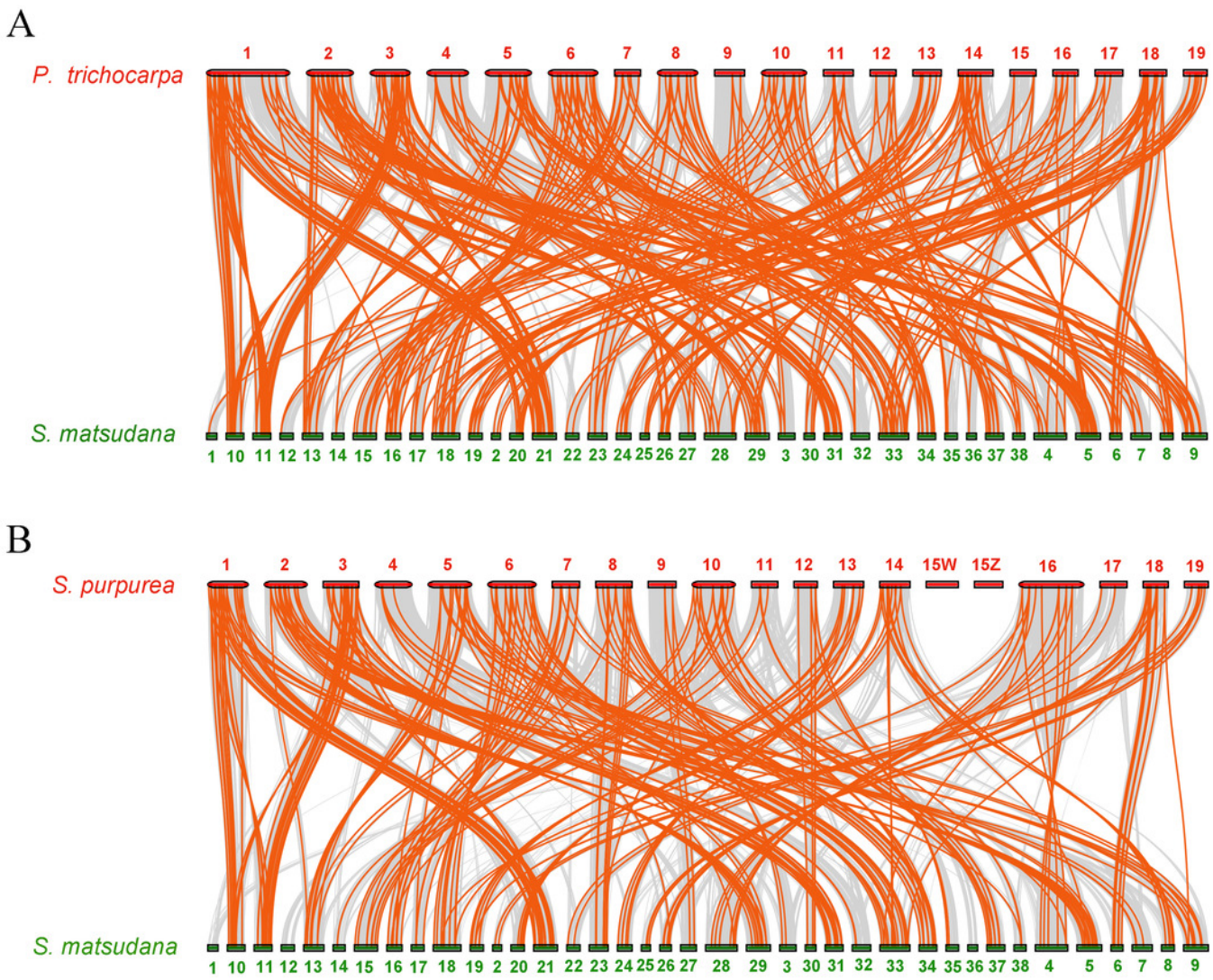


Figure 8

Figure. 7 Hierarchical clustering of AP2 genes and heatmap that demonstrates the differential expression patterns of *SmAP2/ERF* genes in roots before and after salt stress.

Figure. 7 Hierarchical clustering of AP2 genes and heatmap that demonstrates the differential expression patterns of *SmAP2/ERF* genes in roots before and after salt stress. The Log10-transformed expression values of the FPKM values of 12 samples were used to draw the heat map. The color bar indicates the gene expression level. a, Heatmap and hierarchical clustering representation of 47 AP2 members. b, Heatmap and hierarchical clustering representation of 108 DREB members. c, Heatmap and hierarchical clustering representation of 130 ERF members.

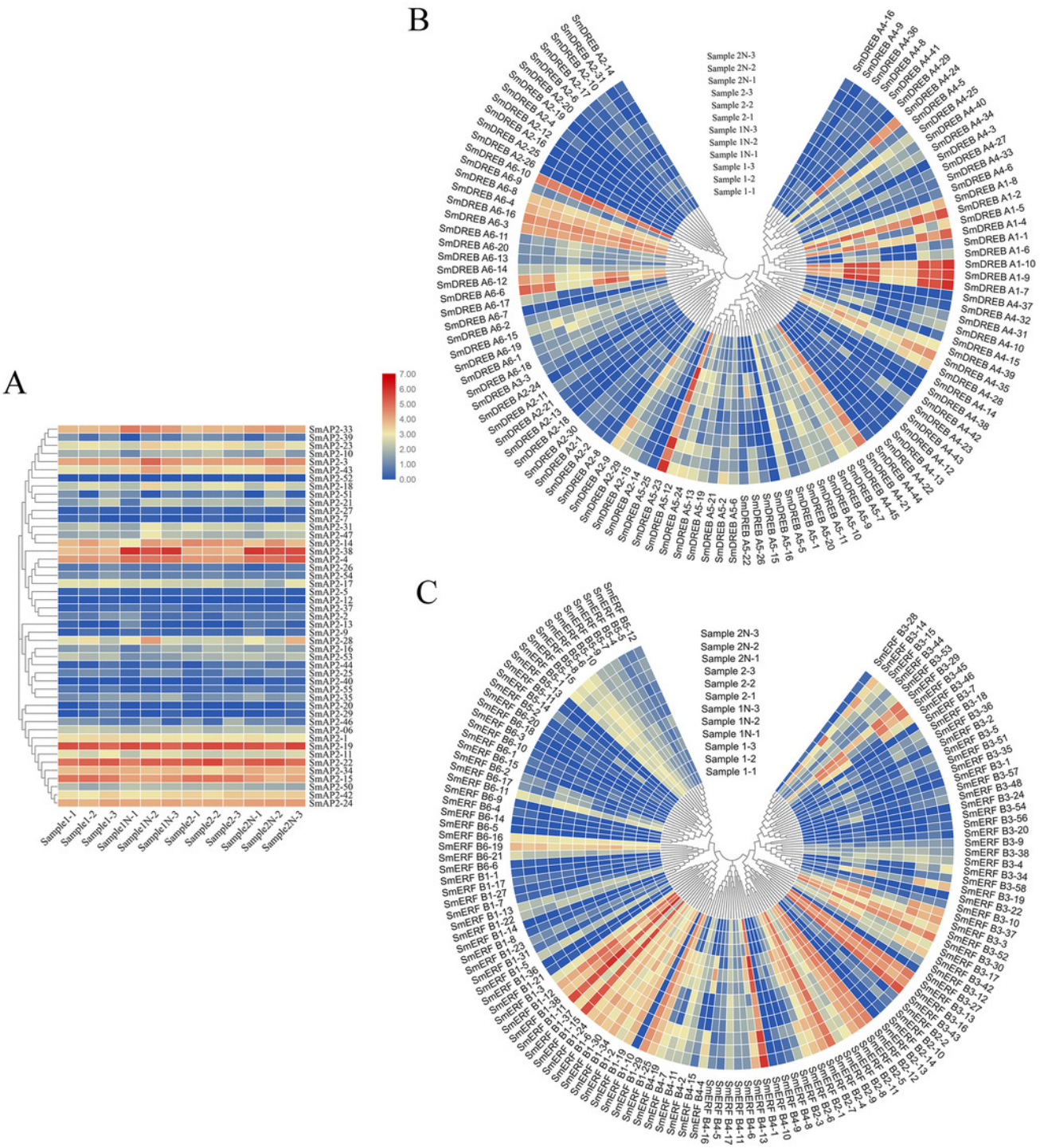
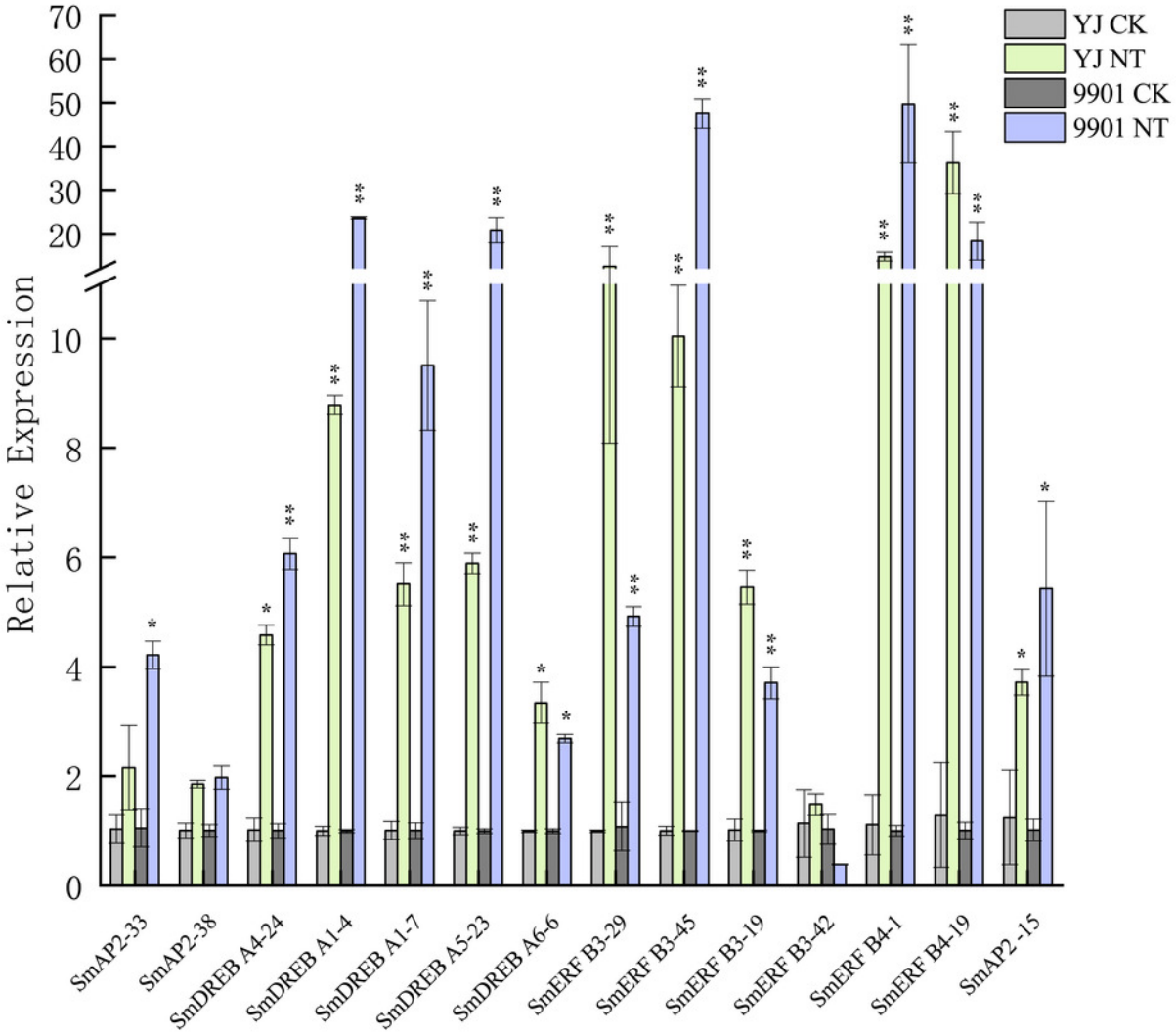

A
B
C
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
Sample 2N-3
Sample 2N-2
Sample 2N-1
Sample 2-3
Sample 2-2
Sample 2-1
Sample 1N-3
Sample 1N-2
Sample 1N-1
Sample 1-3
Sample 1-2
Sample 1-1

Figure 9

Figure. 8 Verification of the *SmAP2/ERF* genes with differentially expressed patterns under salt stress by quantitative real-time PCR

Figure. 8 Verification of the *SmAP2/ERF* genes with differentially expressed patterns under salt stress by quantitative real-time PCR For salt stress, ‘Yanjiang’ and ‘9901’ roots that were subjected to 20 days of hydroponic culture and then treated with 150 mM sodium chloride for 4 h. The control was an untreated Yanjiang sample. Three biological replicates for each sample were performed, and bars represent the standard deviations of the mean. Asterisks on top of the bars indicate statistically significant differences between stress treatment and the control (*0.01 < P < 0.05; **p<0.01, Student’s t-test). Gene expression profiles were evaluated using the $2^{-\Delta\Delta Ct}$ method, and the control value was normalized to 1.


Relative Expression
70
60
50
40
30
20
10
8
6
4
2
0
YJ CK
YJ NT
9901 CK
9901 NT
SmAP2-33
SmAP2-38
SmDREB A4-24
SmDREB A1-4
SmDREB A1-7
SmDREB A5-23
SmDREB A6-6
SmERF B3-29
SmERF B3-45
SmERF B3-19
SmERF B3-42
SmERF B4-1
SmERF B4-19
SmAP2 -15