

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

Yanhong Chen

Nantong University

Yuna Jiang

Nantong University

Guoyuan Liu

Nantong University

Chunmei Yu

Nantong University

Bolin Lian

Nantong University

Fei Zhong

Nantong University

Wenxiang Feng

Nantong University

Jian Zhang (✉ ntdxylzw@163.com)

Nantong University

Research article

Keywords: *Salix matsudana*, SmAP2/ERF, Genome-wide, Salt tolerance, Expression patterns

Posted Date: February 27th, 2020

DOI: <https://doi.org/10.21203/rs.2.24667/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at PeerJ on April 13th, 2021. See the published version at <https://doi.org/10.7717/peerj.11076>.

Abstract

Background: 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, including model plants, horticulture plants, crops, and trees. Many AP2/ERF TFs from various species have been functionally characterized. 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. The recently produced whole genome sequencing data of *Salix matsudana* allowed us to conduct an evolutionary analysis on the AP2/ERF genes during polyploidization and study the genome-wide expression profiles of AP2/ERF genes in *Salix matsudana* under salt stress.

Results: In this study, 364 AP2/ERF genes of *Salix matsudana* (SmAP2/ERF) were identified and renamed according to the chromosomal location of the SmAP2/ERF genes. After phylogenetic analysis with known categories of genes from other species, the AP2/ERF genes were divided into three subfamilies: AP2 (55 members), ERF (301 members), and RAV (six members). Two Soloist genes were also identified. Gene structure and conserved motifs were analyzed in AP2/ERF genes, and introns were not found in most genes of the ERF subfamily, although some unique motifs were found to be important for the function of AP2/ERF genes. Syntenic relationships between the SmAP2/ERF genes and AP2/ERF genes from other species were also investigated to elucidate their evolutionary relationships during polyploidization. Moreover, analyses on the expression profiles under salt stress were also conducted.

Conclusion: The progenitors of *Salix matsudana* underwent whole genome duplication not more than 10 Mya. Synteny analysis with other species showed macrosynteny between willow and poplar AP2/ERF genes and that *Salix matsudana* is genetically more closely related to *Populus trichocarpa* than to *Salix purpurea*. Further investigation of the AP2/ERF TFs demonstrated that differential expression patterns during salt stress and this information can help reveal the mechanism of salt tolerance regulation in *Salix matsudana*.

Background

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 [1]. 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 [1]. 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 [2].

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* [1–3]. Different members play various regulatory roles in plant growth and development, defense response, fruit ripening, and metabolism [4]. Several recent reports demonstrated functions of AP2/ERF2 TFs in plant development. For example, loss of DRNL function affects gynoecium development [5]; the function of *Populus* ERF139 (Potri.013G101100) in xylem cell expansion was characterized by transgenic overexpression and dominant repressor lines of ERF139 [6]; RhERF1 and RhERF4 play roles in petal abscission in rose [7]; 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 [8], and CitAP2.10 was characterized as a regulator of (+)-valencene synthesis in sweet orange fruit [9].

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 [4, 10]. Through extensive investigation on their regulatory mechanism, people want to elucidate their potential applications in crop improvement [10]. 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* promote tolerance to salt and drought stresses in overexpressing *Arabidopsis* lines [11, 12]. Overexpression of HARDY, an AP2/ERF gene from *Arabidopsis*, improves drought and salt tolerance by reducing transpiration and sodium uptake in transgenic *Trifolium alexandrinum* L [13]. A soybean DREB ortholog, GmDREB1, enhances the salt tolerance in transgenic alfalfa [14].

Comparative genomic analysis of model plants such as *Arabidopsis* have provided unprecedented advantages for gene discovery and functional annotation of newly sequenced plant genomes [15–17]. By exploring the available genomic data, AP2/ERF gene families from 50 species were discovered and classified, and provide critical guidance for functional analysis [10]. 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 [18–20].

Salix matsudana Koidz., a member of Salicaceae, is an important ornamental tree species native to northeastern China [21, 22]; 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 [21]. 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 [21]. 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, *Salix arbutifolia* and *Populus trichocarpa*, respectively [3, 23]. The *Salix matsudana* genome was recently sequenced and assembled (unpublished); as a tetraploid, 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.

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 S1.

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 the ERF and RAV families; alternatively, in the AP2 family, the conserved sequences from 58–60AA were converted into YLG elements [24]. 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 S2. 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 S2. The gene name of AP2/ERF genes from *Populus trichocarpa* and *Salix purpurea* are listed in Table S3. 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* [1], *Populus trichocarpa* [3], *Salix purpurea*, and *Salix arbutifolia* [23], 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, 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 or autotetraploid evolution of *Salix matsudana*. Many gene pairs (approximately 70%) maintained the same or similar gene structure during *Salix matsudana* formation, which indicated that the SmAP2/ERF genes were conserved at the DNA level during 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. 2c). 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 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 S4). 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 S5). 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 S4). 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).

Syntenic 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* and *Salix purpurea* (Fig. 6). Collinear AP2/ERF genes pairs between *Salix matsudana* and other two species are listed in Supplementary Table S6. 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 S6). 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. 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*.

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 to demonstrate the expression pattern change under salt stress (Fig. 7).

The expression patterns of 286 genes are illustrated in Fig. 7, and included 48 AP2, 108 DREB, and 130 ERF subgroup genes. In the AP2 subgroup, the FPKM 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. Thirteen Genes with upregulated expression patterns was verified by qRT-PCR (Real-time Quantitative PCR) (Fig. 8).

Discussion

Salix is one of the few woody plants with a large number of polyploid taxa, in *Salix matsudana*, both tetraploid and diploid individuals have been observed. [25]. In our experiment, we sequenced the genome of tetraploid *Salix matsudana*. Tetraploid *Salix* have value 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 [22]. The molecular mechanism of salinity response regulation is very complex, and AP2/ERF TFs are key regulators in plants [26]. 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 a tetraploid species, *Salix matsudana* has more AP2/ERF gene members than other plants selected for comparison, including three Salicaceae family relatives (Table S2). 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 S2). 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 [1].

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 [18, 19]. 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 [27, 28]. 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 [24] and also consistent with previous findings.

Through the conserved domains and motifs, TFs play roles in gene expression regulation by promoter binding, transcription activation, and protein–protein interactions [29]. 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 from the AP2 domain, the functions of other subgroup-specific motifs are still unknown, and more work is required to clarify their regulatory functions.

Based on the genome assembly data, 301 genes were anchored on the 38 chromosomes (LGs), but they were unevenly distributed. Eleven TDs were found on 11 chromosomes, and seven tandem duplication gene pairs came from the SmERF B3 subgroup, which included three duplicated genes (SmERF B3-6, SmERF B3-7, and SmERF B3-8) that clustered together. Apart from the tandem duplication cluster, SmERF B3 members typically clustered on a chromosome, with three genes as a unit. In 12 clusters, 37 SmERF B3 genes were found. This phenomena were also found in *Populus trichocarpa*, thirteen PtERF B3 genes located in 4 clusters, which indicated that in the evolution of Sm, apart from the chromosome duplication, segmental duplication were also happened.

Using MCScanX, we found a total of 28,348 collinear gene pairs in the *Salix matsudana* genome, from which 299 AP2/ERF collinear gene pairs were identified; this indicated that, during evolution, the *Salix matsudana* genome experienced a whole genome duplication event. Population genetic theory predicts that, after duplication, some redundant duplicate copies will be silenced and eliminated, and other retained paralogs will obtain sub- or neofunction by DNA mutation in coding or regulatory sequences [30–32].

Then, we calculated K_a , K_s , and K_a/K_s ratios of these 298 AP2/ERF collinear gene pairs to estimate the divergence time and selection pressure. All K_a/K_s values were below 1, which indicated that these genes might have experienced strong purifying selective pressure during evolution. It was previously reported that purifying selection would lead to the loss of redundant genes [33]. Based on the gene number of most subgroups, we did not find any obvious evidence of gene loss, but in the RAV subgroup, there was an exception; there were only six members in *Salix matsudana*, which is identical to the gene number in *Arabidopsis*. Based on the gene loss hypothesis, the duplication paralogs of RAVs may have been lost during 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* experienced chromosome duplication

remains largely unknown [34]. The divergence time (T Value) of gene pairs mainly occurred during two time periods, 2–8 and 20–36 Mya. Gene pairs with a divergence time of 20–36 Mya were probably paralogs before whole genome duplication events, whereas 2–8 Mya is probably the divergence time of paralogs after whole genome duplication events. These data indicated that the progenitor of *Salix matsudana* underwent whole genome duplication not more than 10 Mya.

Similar to the findings of a previous report [35], alignment of a *Salix* linkage map to the *Populus* genomic sequence revealed macrosynteny between willow and poplar genomes (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 [36]; willow may evolve faster, which would lead to them being more diverse. Researchers proposed that *Populus* might be evolutionarily more primitive than *Salix* [30, 37]. 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; 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* underwent whole 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*.

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*.

Methods

Material and salt stress treatment

The salinity stress treatments were carried out on *Salix matsudana* ‘yanjiang’ and *Salix matsudana* ‘9901’. The 2 salix cultivars used in this study were collected from the botany garden of Nantong University (Nantong, China). The 2 salix cultivars were authorized for only scientific research purpose, and were deposited in College of life science in Nantong University. The stem cuttings (length, 8–10 cm; coarse, 2–3 cm) of two materials were selected for hydroponic rooting for 20 days and then treated with NaCl solution (150 mM) for 4 h. All root samples were divided into four categories with three biological replicates: Sample1-1/Sample1-2/Sample1-3 (‘yanjiang’ without salt stress treatment), Sample1N-1/Sample1N-2/Sample1N-3 (‘yanjiang’ with salt stress treatment), Sample2-1/Sample2-2/Sample2-3 (‘9901’ without salt stress treatment), and Sample2N-1/Sample2N-2/Sample2N-3 (‘9901’ with salt stress treatment).

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 (unpublished).

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 Clustal

W[38]. 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).

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 proteins were obtained and subjected to multiple sequence alignment using ClustalW [38]. To better classify these SmAP2 genes, 48 AP2 domains from known categories of *Arabidopsis* and *Populus trichocarpa* AP2 genes were selected to carry out multiple sequence alignment with SmAP2 proteins, and a phylogenetic tree based on this alignment was built by MEGA 7.0 with the neighbor-joining method [39] with default parameters. Bootstrap value was set to 1000. Depending on the phylogenetic tree constructed by SmAP2, PtAP2, and AtAP2 domains, these SmAP2 genes were classified into different subfamilies and subgroups.

Gene Structure And Conserved Motif Structure Analysis

The UTR–exon–intron structures of the SmAP2 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 [40].

To characterize the SmAP2 protein structure, the online tool MEME (<http://meme-suite.org/tools/meme>) was used to search for conserved motifs [41]. 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 [42].

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 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.

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 [43]. 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 $\geq 80\%$ coverage and $\geq 65\%$ 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.

Divergence Time Calculation Of Duplicated Genes

After BLASTn analysis of Coding sequences and obtaining duplicated gene pairs, the nonsynonymous (Ka) and synonymous (Ks) 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. [44]

Collinearity analysis between *Salix matsudana* and the representative species

To demonstrate the syntenic relationships of the orthologous AP2 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.

RNA sequencing and a heat map generated by hierarchical clustering

Transcriptome sequencing data of 12 samples were obtained by Illumina HiSeq sequencing. The expression levels of genes in different samples were calculated using FPKM values. Using the “Amazing HeatMap” tool of TBtools, a graph of the expression level of AP2 family genes with hierarchical clustering was generated.

RNA Isolation And 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. 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 [45]. 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 S7.

Abbreviations

AP2/ERF

AP2-like ethylene-responsive transcription factor
FPKM
Fragments Per Kilobase of transcript per Million fragments mapped
qRT-PCR
Real-time Quantitative PCR
Mya
million years ago
HMM
Hidden Markov Model
TDs
tandem duplication events
SDs
segmental duplication events
TFs
transcription factors

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The *Salix matsudana* genomic sequence datasets used during the current study are available from the corresponding author on reasonable request. Other data generated or analysed during this study are included in this published article and its supplementary information files.

Competing Interests

The authors declare that they have no conflict of interest.

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). The funding body has no role in the study design and data collection, analysis, interpretation of data and in writing the manuscript.

Authors' contributions

Y.C. and J.Z. conceived the research plans; Y.C. and Y.J. performed most of the experiments; G.L., C.Y., B.L., F.Z. and W.F. performed the rest of the experiments; Y.C. analyzed the data; Y.C. and J.Z. wrote the article. All authors have read and approved the manuscript.

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.

References

1. Nakano T, Suzuki K, Fujimura T, Shinshi H: **Genome-wide analysis of the ERF gene family in Arabidopsis and rice.** *Plant Physiol* 2006, **140**(2):411-432.
2. Song X, Wang J, Ma X, Li Y, Lei T, Wang L, Ge W, Guo D, Wang Z, Li C *et al*: **Origination, Expansion, Evolutionary Trajectory, and Expression Bias of AP2/ERF Superfamily in Brassica napus.** *Front Plant Sci* 2016, **7**:1186.
3. Zhuang J, Cai B, Peng RH, Zhu B, Jin XF, Xue Y, Gao F, Fu XY, Tian YS, Zhao W *et al*: **Genome-wide analysis of the AP2/ERF gene family in Populus trichocarpa.** *Biochem Biophys Res Commun* 2008, **371**(3):468-474.
4. Gu C, Guo ZH, Hao PP, Wang GM, Jin ZM, Zhang SL: **Multiple regulatory roles of AP2/ERF transcription factor in angiosperm.** *Botanical studies* 2017, **58**(1):6.
5. Duran-Medina Y, Serwatowska J, Reyes-Olalde JI, de Folter S, Marsch-Martinez N: **The AP2/ERF Transcription Factor DRNL Modulates Gynoecium Development and Affects Its Response to Cytokinin.** *Front Plant Sci* 2017, **8**:1841.
6. Wessels B, Seyfferth C, Escamez S, Vain T, Antos K, Vahala J, Delhomme N, Kangasjarvi J, Eder M, Felten J *et al*: **An AP2/ERF transcription factor ERF139 coordinates xylem cell expansion and secondary cell wall deposition.** *New Phytol* 2019.
7. Gao Y, Liu Y, Liang Y, Lu J, Jiang C, Fei Z, Jiang CZ, Ma C, Gao J: **Rosa hybrida RhERF1 and RhERF4 mediate ethylene- and auxin-regulated petal abscission by influencing pectin degradation.** *Plant J* 2019, **99**(6):1159-1171.
8. Li H, Xiao Q, Zhang C, Du J, Li X, Huang H, Wei B, Li Y, Yu G, Liu H *et al*: **Identification and characterization of transcription factor ZmEREB94 involved in starch synthesis in maize.** *J Plant Physiol* 2017, **216**:11-16.
9. Shen SL, Yin XR, Zhang B, Xie XL, Jiang Q, Grierson D, Chen KS: **CitAP2.10 activation of the terpene synthase CsTPS1 is associated with the synthesis of (+)-valencene in 'Newhall' orange.** *J Exp Bot* 2016, **67**(14):4105-4115.
10. Srivastava R, Kumar R: **The expanding roles of APETALA2/Ethylene Responsive Factors and their potential applications in crop improvement.** *Briefings in functional genomics* 2018, **18**(4):240-254.

11. Li Y, Zhang H, Zhang Q, Liu Q, Zhai H, Zhao N, He S: **An AP2/ERF gene, IbRAP2-12, from sweetpotato is involved in salt and drought tolerance in transgenic Arabidopsis.** *Plant Sci* 2019, **281**:19-30.
12. Cao B, Shu L, Li A: **Functional characterization of LkERF-B2 for improved salt tolerance ability in Arabidopsis thaliana.** *3 Biotech* 2019, **9**(7):263.
13. Abogadallah GM, Nada RM, Malinowski R, Quick P: **Overexpression of HARDY , an AP2/ERF gene from Arabidopsis, improves drought and salt tolerance by reducing transpiration and sodium uptake in transgenic Trifolium alexandrinum L.** *Planta* 2011, **233**(6):1265-1276.
14. Jin T, Chang Q, Li W, Yin D, Li Z, Wang D, Bao L, Liu L: **Stress-inducible expression of GmDREB1 conferred salt tolerance in transgenic alfalfa.** *Plant Cell Tissue & Organ Culture* 2010, **100**(2):219-227.
15. Brendel V, Kurtz S, Walbot V: **Comparative genomics of Arabidopsis and maize: prospects and limitations.** *Genome Biol* 2002, **3**(3):. reviews1005.1.
16. Hall AE, Fiebig A, Preuss D: **Beyond the Arabidopsis genome: opportunities for comparative genomics.** *Plant Physiol* 2002, **129**(4):1439-1447.
17. Schranz ME, Song BH, Windsor AJ, Mitchell-Olds T: **Comparative genomics in the Brassicaceae: a family-wide perspective.** *Curr Opin Plant Biol* 2007, **10**(2):168-175.
18. Karanja BK, Xu L, Wang Y, Tang M, M'Mbone Muleke E, Dong J, Liu L: **Genome-wide characterization of the AP2/ERF gene family in radish (Raphanus sativus L.): Unveiling evolution and patterns in response to abiotic stresses.** *Gene* 2019, **718**:144048.
19. Li H, Wang Y, Wu M, Li L, Li C, Han Z, Yuan J, Chen C, Song W, Wang C: **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* 2017, **8**:946.
20. Li MY, Liu JX, Hao JN, Feng K, Duan AQ, Yang QQ, Xu ZS, Xiong AS: **Genomic identification of AP2/ERF transcription factors and functional characterization of two cold resistance-related AP2/ERF genes in celery (Apium graveolens L.).** *Planta* 2019, **250**(4):1265-1280.
21. Zhang J, Yuan H, Yang Q, Li M, Wang Y, Li Y, Ma X, Tan F, Wu R: **The genetic architecture of growth traits in Salix matsudana under salt stress.** *Hortic Res* 2017, **4**:17024.
22. Zhang J, Yuan H, Li M, Li Y, Wang Y, Ma X, Zhang Y, Tan F, Wu R: **A High-Density Genetic Map of Tetraploid Salix matsudana Using Specific Length Amplified Fragment Sequencing (SLAF-seq).** *PLoS One* 2016, **11**(6):e0157777.
23. Rao G, Sui J, Zeng Y, He C, Zhang J: **Genome-wide analysis of the AP2/ERF gene family in Salix arbutifolia.** *FEBS open bio* 2015, **5**:132-137.
24. Liu M, Sun W, Ma Z, Zheng T, Huang L, Wu Q, Zhao G, Tang Z, Bu T, Li C *et al*: **Genome-wide investigation of the AP2/ERF gene family in tartary buckwheat (Fagopyum Tataricum).** *BMC Plant Biol* 2019, **19**(1):84.
25. Guo W, Hou J, Yin T, Chen Y: **An analytical toolkit for polyploid willow discrimination.** *Scientific reports* 2016, **6**:37702-37702.

26. Xie Z, Nolan TM, Jiang H, Yin Y: **AP2/ERF Transcription Factor Regulatory Networks in Hormone and Abiotic Stress Responses in Arabidopsis.** *Front Plant Sci* 2019, **10**:228.
27. Tang Y, Qin S, Guo Y, Chen Y, Wu P, Chen Y, Li M, Jiang H, Wu G: **Genome-Wide Analysis of the AP2/ERF Gene Family in Physic Nut and Overexpression of the JcERF011 Gene in Rice Increased Its Sensitivity to Salinity Stress.** *Plos One* 2016, **11**(3):e0150879.
28. Zhang Z, Li X: **Genome-wide identification of AP2/ERF superfamily genes and their expression during fruit ripening of Chinese jujube.** *Scientific reports* 2018, **8**(1):15612.
29. Liu L, ., White MJ, Macrae TH: **Transcription factors and their genes in higher plants functional domains, evolution and regulation.** *Febs Journal* 1999, **262**(2):247-257.
30. Hou J, Wei S, Pan H, Zhuge Q, Yin T: **Uneven selection pressure accelerating divergence of Populus and Salix.** *Horticulture Research* 2019, **6**(1):37.
31. Xionglei H, Jianzhi Z: **Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution.** *Genetics* 2005, **169**(2):1157-1164.
32. Adams KL, Richard C, Ryan P, Wendel JF: **Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing.** *P Natl Acad Sci USA* 2003, **100**(8):4649-4654.
33. Kondrashov FA, Rogozin IB, Wolf YI, Koonin EV: **Selection in the evolution of gene duplications.** *Genome Biology* 2002, **3**(2):1-9.
34. Hou J, Ye N, Dong Z, Lu M, Li L, Yin T: **Major Chromosomal Rearrangements Distinguish Willow and Poplar After the Ancestral "Salicoid" Genome Duplication.** *Genome Biology & Evolution* 2016, **8**(6):1868-1875.
35. Hanley SJ, Mallott MD, Karp A: **Alignment of a Salix linkage map to the Populus genomic sequence reveals macrosynteny between willow and poplar genomes.** *Tree Genetics & Genomes* 2006, **3**(1):35-48.
36. Argus GW: **Infrageneric Classification of Salix (Salicaceae) in the New World.** *Systematic Botany Monographs* 1997, **52**:1-121.
37. Dai X, Hu Q, Cai Q, Kai F, Ning Y, Tuskan GA, Milne R, Chen Y, Wan Z, Wang Z: **The willow genome and divergent evolution from poplar after the common genome duplication.** *Cell Research* 2014, **24**(10):1274-1277.
38. Larkin M, Blackshields G, Brown N, Chenna R, Mcgettigan P, Mcwilliam H, Valentin F, Wallace I, Wilm A, R: **Clustal W and clustal X version 2.0.** *Bioinformatics* 2007, **23**(21):2947-2948.
39. Kumar S, Stecher G, Tamura K: **MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets.** *Molecular Biology & Evolution* 2016, **33**(7):1870.
40. Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G: **GSDS 2.0: an upgraded gene feature visualization server.** *Bioinformatics* 2014, **31**(8):1296.
41. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS: **MEME SUITE: tools for motif discovery and searching.** *Nucleic Acids Res* 2009, **37**(Web Server issue):W202-208.

42. Chen C, Xia R, Chen H, He Y: **TBtools, a Toolkit for Biologists integrating various HTS-data handling tools with a user-friendly interface.** *bioRxiv* 2018:doi: 10.1101/289660.
43. Cannon SB, Mitra A, Baumgarten A, Young ND, May G: **The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*.** *BMC Plant Biology* 2004, **4**(1):10.
44. Lynch M, Conery JS: **The evolutionary fate and consequences of duplicate genes.** *Science* 2000, **290**(5494):1151-1155.
45. Chen YH, Cao YY, Wang LJ, Li LM, Yang J, Zou MX: **Identification of MYB transcription factor genes and their expression during abiotic stresses in maize.** *Biol Plantarum* 2018, **62**(2):222-230.

Figures

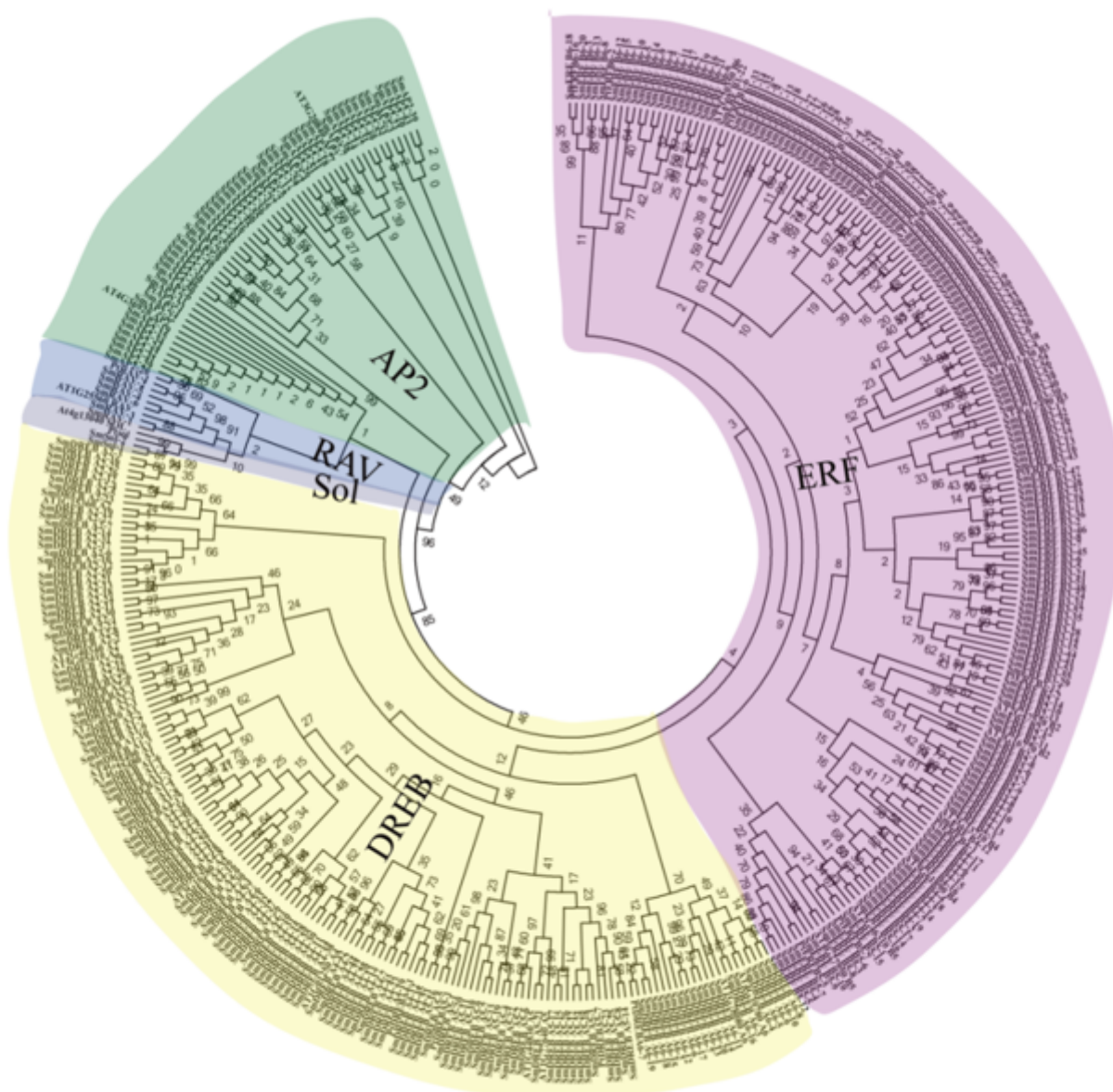


Figure 1

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.

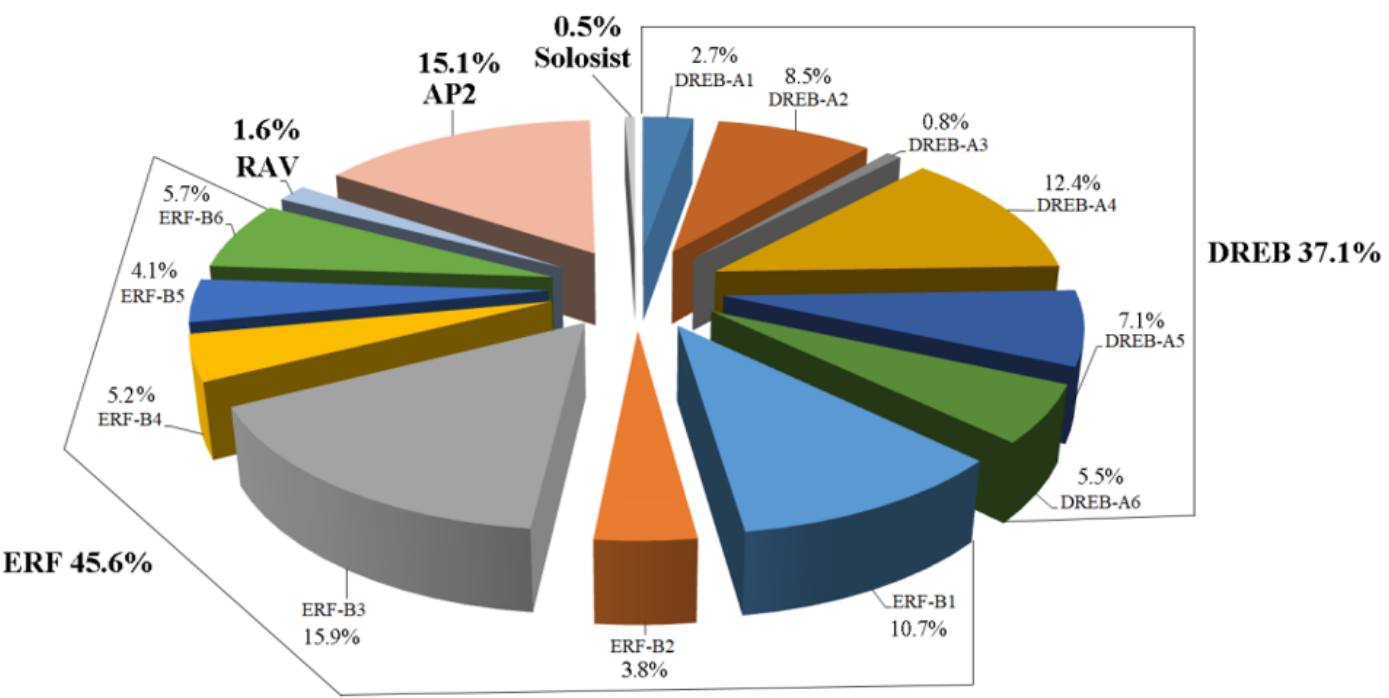


Figure 2

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.

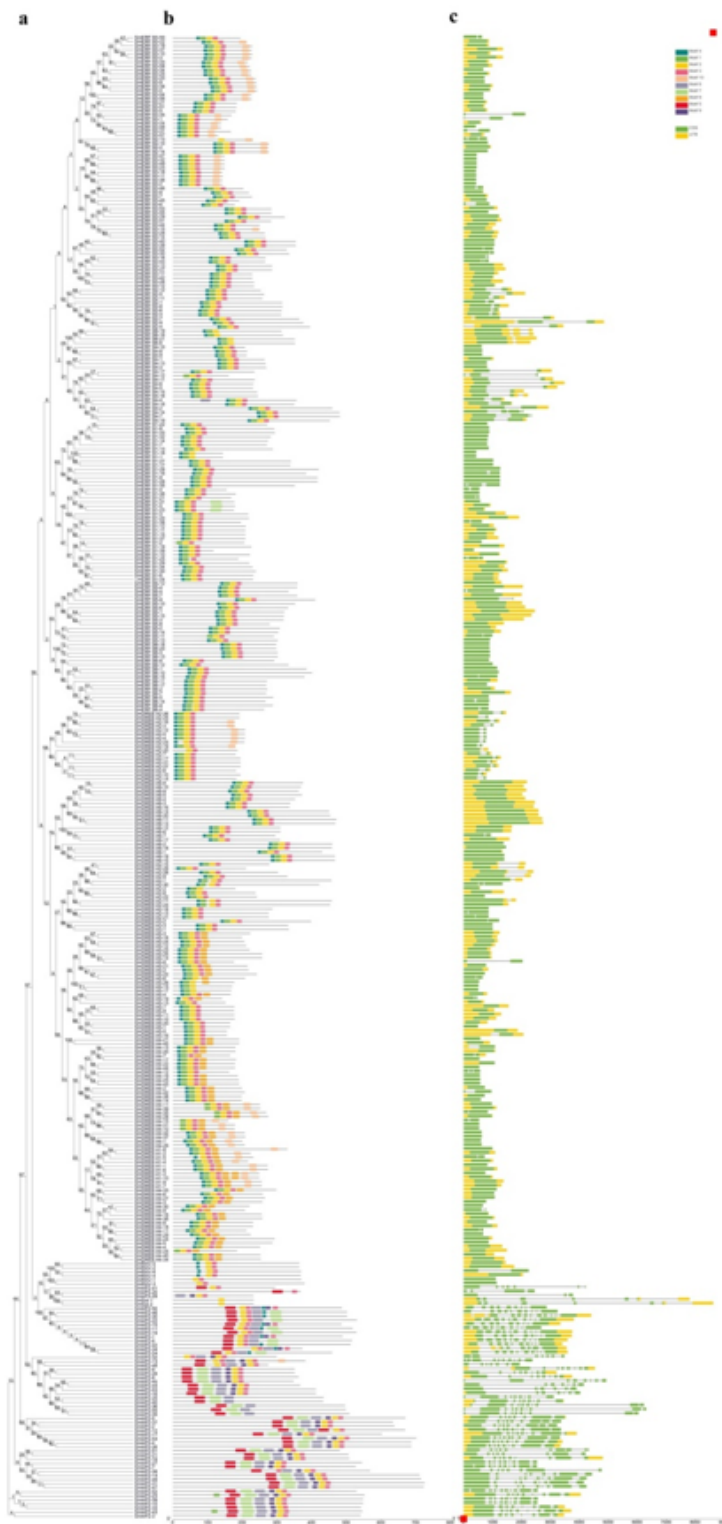


Figure 3

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

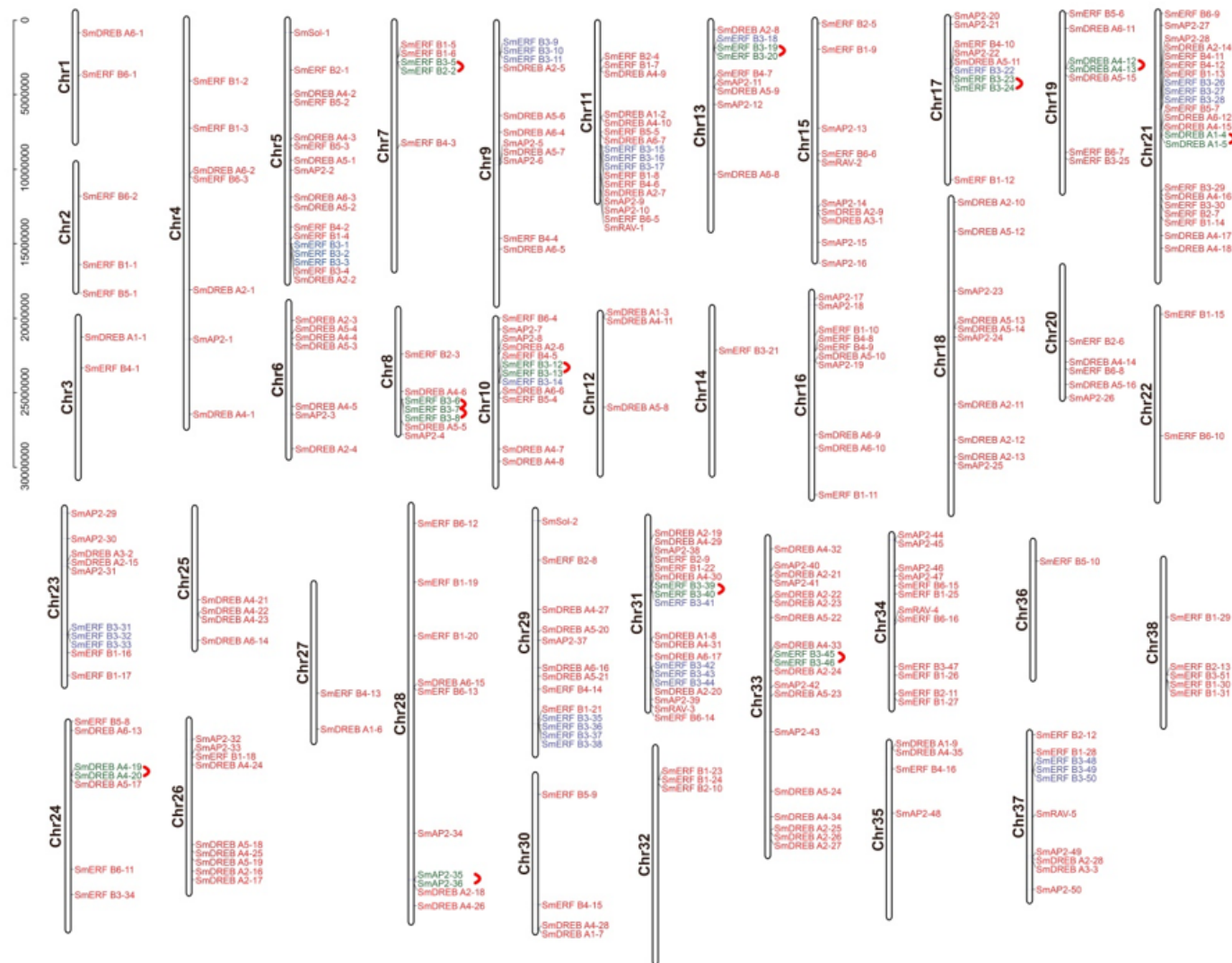


Figure 4

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.

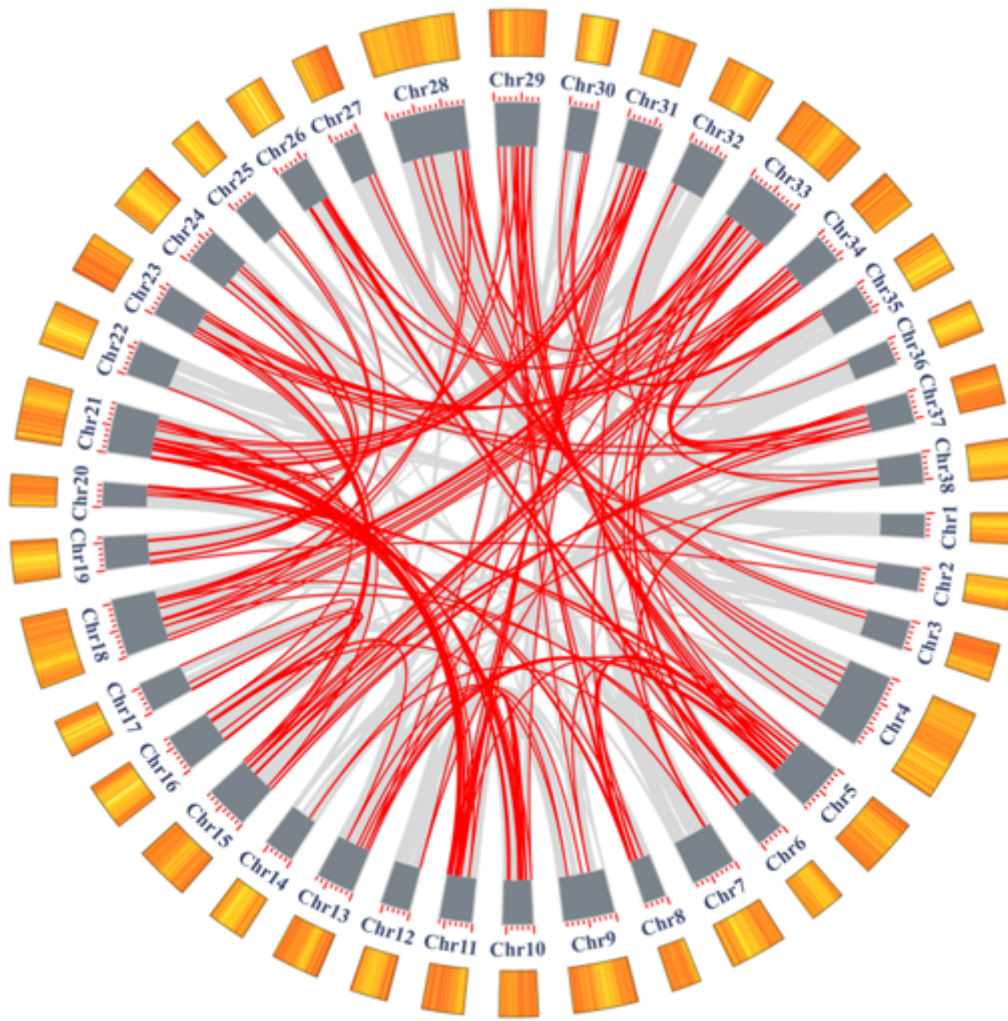


Figure 5

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 columns outside of the circle indicated the gene density on each 38 chromosomes. The deeper color means the higher density of genes.

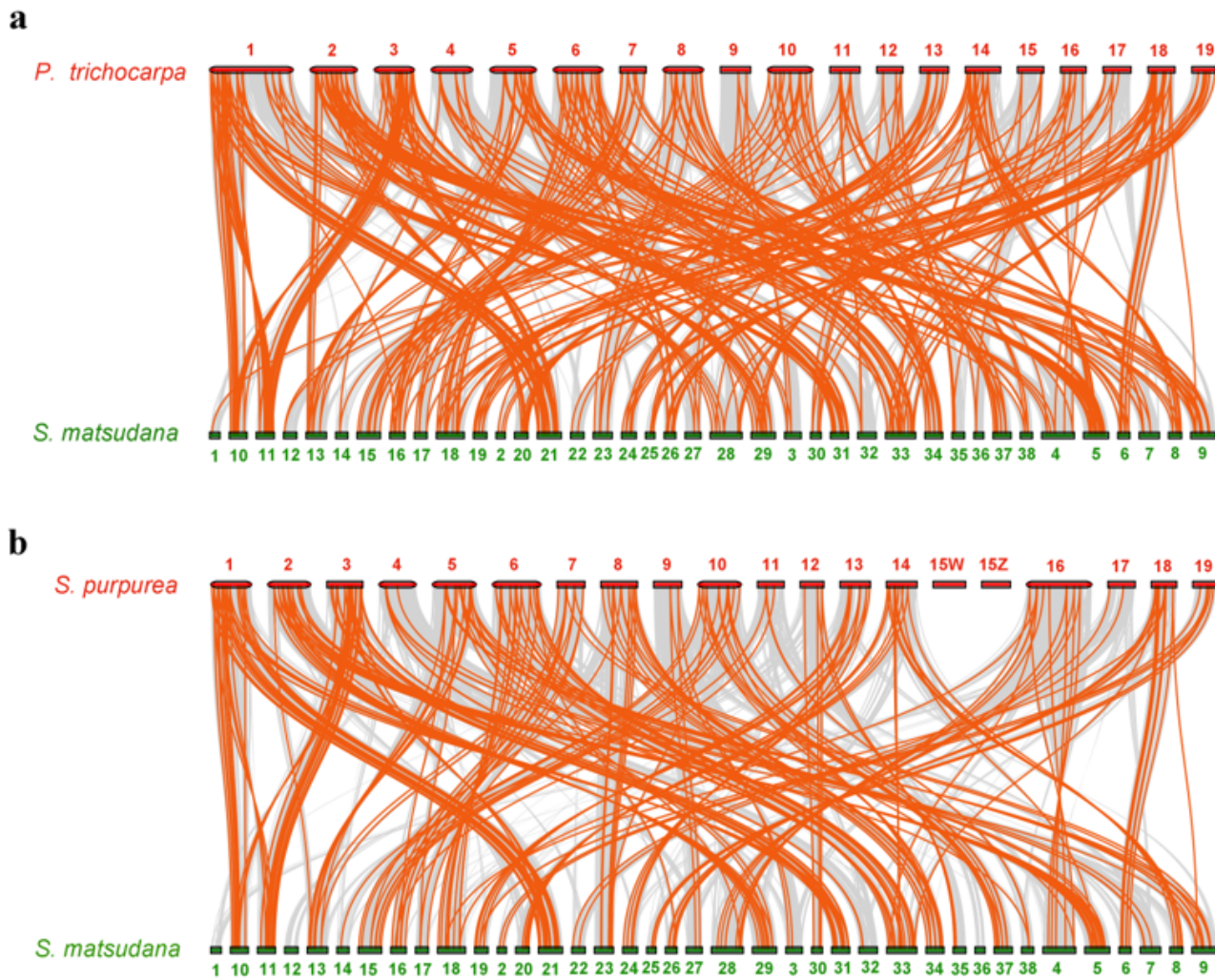


Figure 6

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.

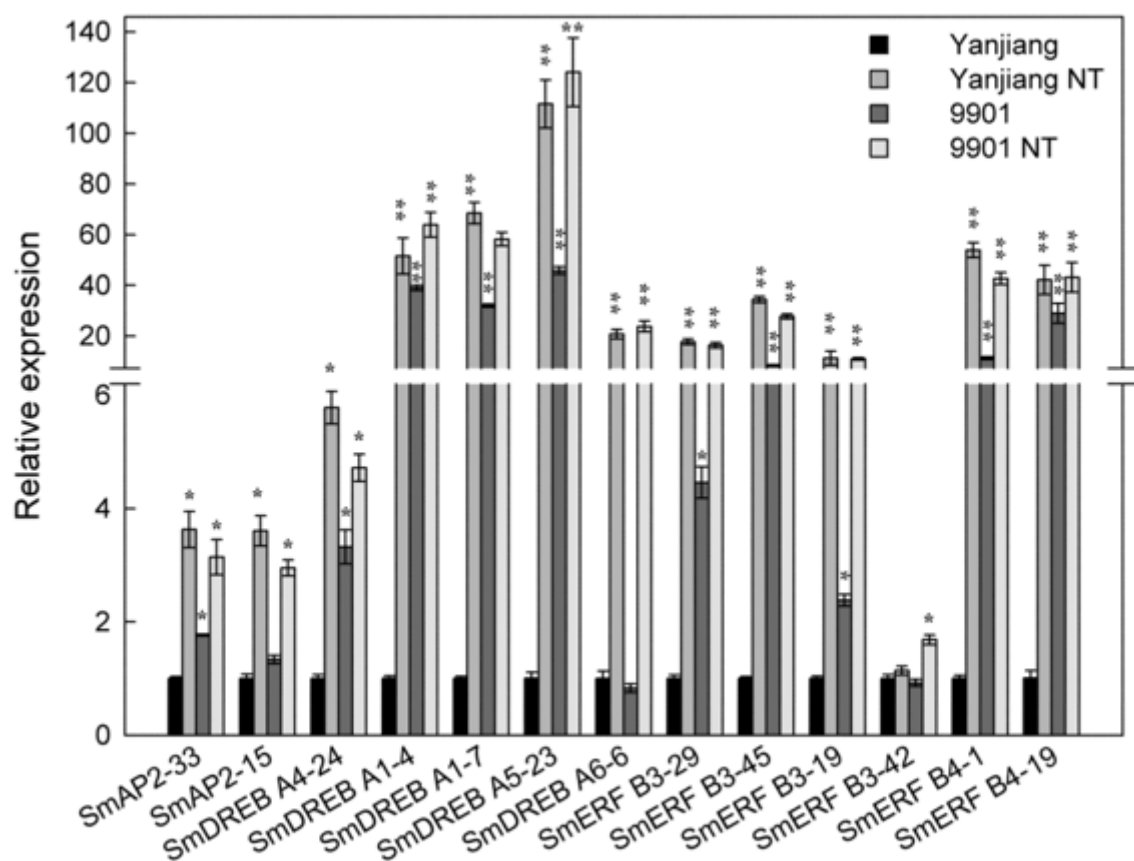


Figure 8

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 $2^{-\Delta\Delta C_t}$ method, and the control value was normalized to 1.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementalfilesTableS1S7FigureS1S2.docx](#)