Plant Physiology and Biochemistry 136 (2019) 1–12

Contents lists available at ScienceDirect

Plant Physiology and Biochemistry

journal homepage: www.elsevier.com/locate/plaphy

Research article

Comprehensive analysis of the three-amino-acid-loop-extension gene family and its tissue-differential expression in response to salt stress in poplar

Kai Zhao[a], Xuemei Zhang[a], Zihan Cheng[a], Wenjing Yao[a,b], Renhua Li[a], Tingbo Jiang[a], Boru Zhou[a,*]

[a] State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, 51 Hexing Road, Harbin, 150040, China
[b] Bamboo Research Institute, Nanjing Forestry University, 159 Longpan Road, Nanjing, 210037, China

ARTICLE INFO

Keywords:
Poplar TALE gene family
Protein interaction
Salt stress
Tissue-differential expression

ABSTRACT

The three-amino-acid-loop-extension (TALE) transcription factor gene family is widely present in plants and plays an important role in its growth and development. However, studies on the gene family are limited in poplar. In this study, we investigated 35 TALE gene family members in terms of their evolutionary relationship, classification, physicochemical properties, gene structures, and protein motifs. We divided the genes into four classes, based on their protein sequences similarity. The members from each class share similar gene structures and motif compositions. Evidence from transcript profiling indicated that the majority of the TALE genes exhibited distinct expression patterns over leaf, stem, and root tissues. Out of the 35 genes, 17 genes are highly expressed in stems, suggesting that the TALE gene family may play an important role in secondary growth and wood formation. Furthermore, out of the 35 genes, 11 genes are responsive to salt stress, and the spatio-temporal expression patterns of these 11 genes under salt stress were analysed using RT-qPCR. Yeast two-hybridization analysis indicated that poplar TALE proteins from different classes can form heterodimers. These results lay the foundation for future studies on biological functions of poplar TALE genes.

1. Introduction

Homeobox genes are widely present in eukaryotes and play an important role in the growth and development of animals and plants. The homeobox was first discovered in Drosophila in 1984 (McGinnis et al. 1984a, 1984b; Scott and Weiner, 1984), and the first plant homeobox gene *Knotted-1* was identified in maize in 1991 (Vollbrecht et al., 1991). The homeobox gene family encodes a highly conserved homeodomain consisting of approximately 60 amino acids that form three alpha-helices, of which the first and second helixes are linked by a loop structure (Billeter et al., 1993; W J Gehring et al., 1994). Similarly, the second and third helixes make a helix-turn-helix motif (Billeter et al., 1993; W J Gehring et al., 1994). In previous studies, plant homeobox genes were divided into 11 classes, including HD-ZIP (with four subclasses: I to IV), WOX, NDX, PHD, PLINC, LD, DDT, SAWADEE, PINTOX, KNOX and BEL (Burglin and Affolter, 2016). Among them, the KNOX and BEL genes belong to the TALE gene family, as they encode an atypical homeodomain that consists of 63 amino acids (Burglin, 1997). The three extra amino acid residues (P—Y—P) connect the region between the first and the second helixes (Burglin, 1997).

The KNOX (KNOTTED1-LIKE HOMEOBOX) proteins contain domains including KNOX1, KNOX2, ELK, and homeodomain, except for few new members without the homeodomain (Kimura et al., 2008; Magnani and Hake, 2008). In addition, both the KNOX1 and KNOX2 domains merge to form a MEINOX domain that plays a role in mediating interactions of KNAT proteins with BEL1 (Bellaoui et al., 2001; Hamant and Pautot, 2010). In the previous study, the KNOX proteins were divided into two classes, based on their sequence similarities within the homeodomain, intron position, expression pattern and phylogenetic analysis (Hay and Tsiantis, 2009). With the discovery of the KNOX proteins that lack the homeodomain in dicotyledons, such as Arabidopsis, the proteins were divided into three classes (Magnani and Hake, 2008; Hamant and Pautot, 2010). The KNOX Class I genes are mainly expressed in meristem and play an important role in plant growth and development. For example, Arabidopsis KNOX gene *STM* participates in the shoot apical meristem formation during embryogenesis (Long et al., 1996). The TALE gene *KNAT2* expresses in the internal parts of the vegetative shoot apical meristem and plays a role in

Abbreviations: TALE, three-amino-acid-loop-extension; DEGs, differentially expressed genes; RT-qPCR, reverse transcription quantitative real-time PCR; SD, synthetically defined; AbA, Aureobasidin A; FPKM, fragments per kilo-bases per million mapped reads

* Corresponding author.
E-mail address: boruzhou@yahoo.com (B. Zhou).

https://doi.org/10.1016/j.plaphy.2019.01.003
Received 15 October 2018; Received in revised form 31 December 2018; Accepted 2 January 2019
Available online 03 January 2019

carpel development in *Arabidopsis thaliana* (Pautot et al., 2001). In rice, *OSH1* is essential for the shoot apical meristem maintenance after germination (Tsuda et al., 2011), and *OSH15* is associated with the architecture of internodes (Sato et al., 1999). The KNOX gene *POTH1* in potato plays an important role in vegetative development (Rosin et al., 2003). In addition, besides participating to growth and development, *GmSBH1* can also respond to high temperature and humidity stress in soybean (Shu et al., 2015). The KNOX Class II genes are expressed in multiple tissues. At present, studies on the KNOX Class II members focused on Arabidopsis *KNAT7*, which is involved in the regulation of secondary wall formation (Li et al., 2012). To date, the KNOX Class III members are only found in dicotyledons (Tsuda and Hake, 2015). For instance, *KNATM* is expressed in proximal-lateral domains of organ primordia and the boundary of mature organs and associated with leaf development (Magnani and Hake, 2008). *PTS* gene is responsible for the natural variation of leaf shape in *S. pimpinellifolium* (Kimura et al., 2008). Therefore, the KNOX genes significantly impact plant growth and development.

BELL proteins have such domains as SKY, BELL and homeodomain (Bellaoui et al., 2001). Furthermore, the SKY and BELL can merge into a complex MEINOX interacting domain (Hamant and Pautot, 2010). The BELL and KNOX proteins can form heterodimers to affect their cellular localization (Rutjens et al., 2009). Like the KNOX genes, the BELL genes also play important roles in plant growth and development. For instance, Arabidopsis *BEL1* participates in the developmental process of ovule (Brambilla et al., 2007). *ATH1* can control floral competency and stem growth in *Arabidopsis thaliana* (Proveniers et al., 2007; Gomez-Mena and Sablowski, 2008). In addition, several genes, such as *ATH1*, *PNY*, and *PNF*, are required for the shoot apical meristem initiation and maintenance in *Arabidopsis thaliana* (Rutjens et al., 2009). *BLH1* can cooperate with *KNAT3* to regulate seed germination and seedling development of Arabidopsis (Kim et al., 2013). Tomato BELL gene *SlBEL11* is involved in chloroplast development and chlorophyll synthesis in fruit (Meng et al., 2018). *StBEL5* in potato plays an important role in enhancing tuber yields (Lin et al., 2013); in contrast, *StBEL11* and *StBEL29* may suppress tuber growth by repressing the target genes of *StBEL5* (Ghate et al., 2017). Recent studies reported that the soybean *GmBLH4* can specifically interact with *GmSBH1*, and plays an important role in modulating plant development and high temperature and humidity stress responses (Tao et al., 2018).

Large-scale yeast two-hybrid experiments demonstrated that each KNOX protein has at least one interactive BELL protein and vice versa in *Arabidopsis thaliana* (Hackbusch et al., 2005). The results indicated that KNOX and BELL proteins can generally form heterodimers (Hackbusch et al., 2005).

The TALE gene family plays an important role in plant growth and development. However, previous studies on poplar are limited to few genes that are associated with secondary growth and wood formation (Du et al., 2009; Liu et al. 2015a, 2015b; Petzold et al., 2018). Although there were limited previous studies on the homeobox gene family in several species (Mukherjee et al., 2009; Burglin and Affolter, 2016), however, systematic analysis of poplar TALE gene family is still lacking. In this study, we obtained 35 poplar TALE transcription factors and characterized their physicochemical properties, evolutionary relationships, gene structures, and protein motifs. We then explored their tissue-differential expression patterns and salt stress responses. Evidence from yeast two-hybrid experiments indicated that the proteins of the poplar TALE gene family over-expressed in stems can also form heterodimers. These studies lay the foundation for future studies on biological functions of poplar TALE genes.

2. Materials and methods

2.1. Database search and TALE gene family analysis

By searching the PlantTFDB database, we obtained 34 poplar TALE members that contain either the homeodomain/BELL domains or the homeodomain/ELK domains (Jin et al., 2017). To be comparable to Arabidopsis containing a TALE gene family member without homeodomain, we added a poplar gene that is homologous to *KNATM* (Magnani and Hake, 2008), and verified it by BLASTP (Goodstein et al., 2012). The amino acid sequences of the poplar TALE gene family were downloaded from the Phytozome database, based on Joint Genome Institute (JGI) Ptri version 3.0 (Goodstein et al., 2012; Tuskan et al., 2006). To further verify the accuracy of the results, we then used the SMART database to examine the conserved domains contained in the TALE gene family (Letunic et al., 2015; Letunic and Bork, 2018). Subcellular localization of the TALE proteins was predicted by WoLF PSORT (Horton et al., 2007). The distribution of genes on the chromosomes was from the PopGenIE v3.0 database (Sjodin et al., 2009; Sundell et al., 2015).

We downloaded the amino acid sequences of Arabidopsis TALE gene family from Phytozome, which used the Arabidopsis Information Resource (Goodstein et al., 2012; Lamesch et al., 2012). Multiple sequence alignment of Arabidopsis and poplar TALE protein sequences was performed using ClustalX 1.83 (Thompson et al., 1997). Then we imported the resulting file of the alignment into MEGA5.05 to construct a phylogenetic tree with the neighbor method, 1000 repetitions of bootstrap tests, and Poisson model (Tamura et al., 2011). In addition, we also used PhyML 3.0 to build the phylogenetic tree, based on the maximum likelihood (ML) method and the best model selected by the SMS software (Guindon et al., 2010; Lefort et al., 2017). Protein physicochemical properties were predicted by ProtParam (Gasteiger et al., 2005). Multiple sequence alignment of proteins of different classes was carried out using the BioEdit software (Hall, 1999). We identified the domains included in the TALE proteins by use of the SMART database and the information in previous studies (Bellaoui et al., 2001; Letunic et al., 2015; Letunic and Bork, 2018).

2.2. Gene structure and protein sequence motif analysis

The intron/exon structure of poplar TALE gene family members was mapped by Gene Structure Display Server Program (Hu et al., 2015). The respective coding sequences and genomic sequences were downloaded from Phytozome (Tuskan et al., 2006; Goodstein et al., 2012). The conserved motifs harbor in the TALE proteins were identified by the MEME Suite (Bailey et al., 2009). The motif logos were drawn by use of the Tbtools (Chen et al., 2018). The annotation information was obtained from the SMART database and InterProScan application (Jones et al., 2014; Letunic et al., 2015; Finn et al., 2017; Letunic and Bork, 2018).

2.3. Gene expression analysis

In order to explore the tissue-differential expression patterns of the poplar TALE family, we used the RNA-Seq data from 12 samples, as reported in our previous study (Zhao et al., 2018). The data has two biological replicates with approximately ten-fold sequencing depths. Among the 12 samples, six are leaf, stem, and root tissues collected after being treated with 150 mM NaCl for 24 h, and the rest are controls. Pearson correlation coefficients between biological replicates were calculated using R software (https://www.r-project.org/). The differentially expressed genes (DEGs) were identified using DESeq software and the thresholds were fold change ≥ 2 and padj (p-value adjusted for multiple testing) < = 0.05 (Anders and Huber, 2010).

2.4. Verification by reverse transcription quantitative real-time PCR

In order to verify the results of RNA-Seq, we performed reverse transcription quantitative real-time PCR (RT-qPCR) experiments using 32 genes, except for three genes that were barely expressed. Primer sequences are listed in Table S1. We used actin as the reference gene

Table 1
Summary of PtTALE sequences and *A. thaliana* homologs.

Gene Name	Accession number	AA[a]	MW[b] (Da)	pI[c]	II[d]	AI[e]	GRAVY[f]	AtTALE Ortholog	AtTALE Name
PtTALE1	Potri.001G100800.2	657	73618.74	6.32	47.86	60.56	−0.771	AT5G41410.1	*BEL1*
PtTALE2	Potri.001G112200.1	301	33920.22	6.23	61.68	80.03	−0.639	AT1G62990.1	*KNAT7*
PtTALE3	Potri.001G216600.1	732	79768.87	7.18	49.23	68.74	−0.582	AT2G35940.2	*BLH1*
PtTALE4	Potri.002G030900.1	448	50632.31	6.38	51.23	81.85	−0.436	AT1G75430.1	*BLH11*
PtTALE5	Potri.002G031000.1	704	77874.8	5.98	48.85	70.38	−0.609	AT2G16400.1	*BLH7*
PtTALE6	Potri.002G113300.1	368	42459.43	5.98	51.54	59.16	−0.954	AT4G08150.1	*KNAT1*
PtTALE7	Potri.003G131300.1	664	74280.56	6.29	51.32	60.11	−0.752	AT5G41410.1	*BEL1*
PtTALE8	Potri.004G004700.1	369	41189.27	6.14	48.76	60.62	−0.667	AT1G62360.1	*STM*
PtTALE9	Potri.004G159300.1	676	73917.68	5.97	42.35	60.37	−0.68	AT4G34610.1	*BLH6*
PtTALE10	Potri.004G213300.1	833	91420.77	6.05	50.68	67.41	−0.543	AT2G27990.1	*BLH8*
PtTALE11	Potri.005G014200.1	317	36324.55	5.31	47.73	58.83	−0.754	AT1G23380.1	*KNAT6*
PtTALE12	Potri.005G017200.1	316	36277.52	5.31	46.58	60.54	−0.752	AT1G23380.1	*KNAT6*
PtTALE13	Potri.005G129500.1	825	89423.93	6.65	52.86	63.54	−0.664	AT2G23760.3	*BLH4*
PtTALE14	Potri.005G232000.1	462	51915.26	6.05	60.96	77.08	−0.486	AT1G75430.1	*BLH11*
PtTALE15	Potri.006G190000.1	338	38563.14	5.81	54.76	74.44	−0.782	AT5G25220.1	*KNAT3*
PtTALE16	Potri.006G203000.1	678	73697.11	6.49	43.65	70.37	−0.581	AT2G35940.2	*BLH1*
PtTALE17	Potri.006G230700.2	567	62775.47	6.13	50.96	79.63	−0.366	AT4G32980.1	*ATH1*
PtTALE18	Potri.006G259400.1	432	48185.5	5.91	55.42	69.79	−0.707	AT5G25220.1	*KNAT3*
PtTALE19	Potri.007G032700.1	824	89326.9	7.08	54.04	62.44	−0.67	AT2G23760.2	*BLH4*
PtTALE20	Potri.008G061000.1	477	53956.89	9.22	43.11	71.53	−0.627	AT5G02030.1	*BLH9*
PtTALE21	Potri.008G188700.1	341	38685.48	5.51	50.12	75.01	−0.668	AT1G23380.1	*KNAT6*
PtTALE22	Potri.009G009800.1	835	91972.47	6.69	50.95	69.57	−0.575	AT2G27990.1	*BLH8*
PtTALE23	Potri.009G009900.1	811	89741.65	6.17	48.96	68.8	−0.674	AT5G02030.1	*BLH9*
PtTALE24	Potri.009G017400.1	735	80239.45	8.48	52.35	67.7	−0.621	AT2G35940.3	*BLH1*
PtTALE25	Potri.009G120800.1	675	73963.79	6.1	44.87	61.6	−0.675	AT4G34610.1	*BLH6*
PtTALE26	Potri.010G043500.1	309	35200.65	4.88	53	78.71	−0.606	AT1G23380.1	*KNAT6*
PtTALE27	Potri.010G197300.1	479	53581.56	9.01	38.19	75.91	−0.494	AT5G02030.1	*BLH9*
PtTALE28	Potri.011G011100.1	373	41549.57	6.13	50.22	60.48	−0.67	AT1G62360.1	*STM*
PtTALE29	Potri.012G043400.1	181	20176.92	5.6	35.01	79.67	−0.382	AT1G14760.2	*KNATM*
PtTALE30	Potri.012G087100.1	340	37951.31	4.95	48.63	65.44	−0.61	AT1G23380.1	*KNAT6*
PtTALE31	Potri.013G008600.1	320	36167.52	5.15	38.63	71	−0.582	AT1G23380.1	*KNAT6*
PtTALE32	Potri.015G079100.1	347	38787.08	5.03	42.05	64.15	−0.634	AT1G23380.1	*KNAT6*
PtTALE33	Potri.016G069700.1	679	73938.28	6.67	43.57	70.12	−0.59	AT2G35940.3	*BLH1*
PtTALE34	Potri.018G022700.1	426	47762.87	5.88	53.29	73.29	−0.729	AT5G25220.1	*KNAT3*
PtTALE35	Potri.018G054700.1	586	64596.32	6.07	50.52	77.61	−0.399	AT4G32980.1	*ATH1*

[a] Length of the amino acid sequence.
[b] Molecular weight of the protein.
[c] Theoretical pI of the protein.
[d] Instability index (II) of the protein.
[e] Aliphatic index of the protein.
[f] Grand average of hydropathicity of the protein.

(Regier and Frey, 2010). The detailed experimental steps refer to our previous study (Yao et al., 2016). The expression level of each gene was calculated relative to its expression level in leaf or stem control. An one-way analysis of variance followed by LSD test ($P = 0.05$) was performed to examine the significance of gene expression levels.

2.5. Plant materials and stress treatments

The plant materials used in this study are di-haploid *Populus simonii* × *Populus nigra* growing in the culture bottle with 70 ml 1/2 MS culture medium. The culture conditions are 16/8-h light/dark cycles and average temperature of 25 °C. For gene cloning, we collected the stem tissues of seedling growing for one month and immediately froze them in liquid nitrogen. In order to explore the gene expression patterns under different salt concentrations, the 40-day-old plants were treated with 100 mM, 150 mM, or 200 mM NaCl for 0 h, 3 h, 6 h, 12 h, or 24 h. Then the leaves, roots, and stems were harvested and frozen in liquid nitrogen. All treatments were performed with three biological replicates. Then we stored the materials in the refrigerator at −80 °C.

2.6. Gene cloning and vector construction

Total RNA was extracted using TaKaRa MiniBEST Plant RNA Extraction Kit (Takara, Dalian, China). The cDNA was then synthesized by use of PrimeScript™ RT reagent Kit with gDNA Eraser (Takara, Dalian, China). The procedures are described in the instructions of the kits. We designed the primers based on the transcript sequences (Tuskan et al., 2006; Goodstein et al., 2012). The gene fragments were obtained using reverse transcription PCR with TaKaRa Ex Taq® DNA Polymerase (Takara, Dalian, China), and then sent to Sangon Biotech Company (http://www.sangon.com/) for sequencing and validation. In order to develop fusion vectors of pGBKT7-PtTALE4, pGBKT7-PtTALE20, pGADT7-PtTALE8, and pGADT7-PtTALE12, we introduced restriction sites at both ends of the genes by PCR and performed double digestion experiments with genes and vectors. The digestion products were linked by T4 DNA Ligase (Takara, Dalian, China). The connection products were transferred to *E. coli*, and positive clones were confirmed by PCR and sequencing. The primers and restriction sites used in this section are shown in Table S1.

2.7. Yeast two-hybrid experiment

In order to verify the interaction between members of the TALE gene family, we performed the yeast two-hybridization with the constructed vectors. In yeast transformation we used Yeastmaker™ Yeast Transformation System 2 (Takara, Dalian, China), following the user manual. First, we transferred the vectors of empty pGBKT7 (negative control), pGBKT7-PtTALE4, pGBKT7-PtTALE20, and pGBKT7-53/pGADT7-T (positive control) into the Y2HGold yeast strain respectively, and then spread them onto synthetically defined (SD)/-Trp solid media (the positive control onto SD/-Trp/-Leu solid media) to select transformed cells at 30 °C for 3–5 days. Transformed yeast cells were inoculated onto SD/-Trp and SD/-Trp/-His/-Ade/X-α-Gal to test auto-activation.

On the basis of the four fusion vectors, we have four combinations (pGBKT7-PtTALE4/pGADT7-PtTALE8; pGBKT7-PtTALE4/pGADT7-PtTALE12; pGBKT7-PtTALE20/pGADT7-PtTALE8; pGBKT7-PtTALE20/pGADT7-PtTALE12). We then co-transferred the combinatory fusion vectors into the Y2HGold yeast strain, followed by selection on the SD/-Trp/-Leu solid media. The co-transformed yeast cells with pGBKT7-Lam/pGADT7-T and pGBKT7-53/pGADT7-T were used as negative and positive control respectively. Transformed yeast cells were subsequently cultured in the SD/-Trp/-Leu liquid media and then spotted on the SD/-Trp/-Leu and SD/-Trp/-Leu/-His/-Ade/X-α-Gal/Aureobasidin A (AbA, 125 ng/ml) solid media to test interactions.

3. Results

3.1. Identification of the TALE gene family

By searching the PlantTFDB database, we obtained 34 poplar TALE gene family members, which contain BELL or ELK domains besides the homeodomain (Jin et al., 2017). In addition, we added the homologous gene of *KNATM* (Magnani and Hake, 2008), a KNOX class III member without homeodomain, to the poplar TALE gene family. The E-value of BLASTP between the homologous gene and *KNATM* is 1.8E-8, which also supports our result. We named them *PtTALE1* to *PtTALE35* according to their gene coordinate (Table 1). Then we used the SMART database to verify our results (Letunic et al., 2015; Letunic and Bork, 2018), and it showed that except for PtTALE29, the remaining members contain such domains as homeodomain, POX or homeodomain, KNOX1, KNOX2, and ELK. In contrast, PtTALE29 harbors only KNOX1 and KNOX2 domains (Table S2). To explore the subcellular localization of TALE family proteins, we used the WoLF PSORT program for prediction (Horton et al., 2007). As expected, all proteins are located in the nucleus (Table S2). We also used the PopGenIE v3.0 database to explore the distribution of genes on chromosomes (Sjodin et al., 2009; Sundell et al., 2015). As shown in Table S2, the poplar TALE genes are unevenly distributed on 16 out of 19 chromosomes (except chromosomes 14, 17, and 19). Among them, each of chromosomes 5, 6, and 9 contains a maximum number of four genes.

3.2. Phylogenetic trees and protein physicochemical properties

To explore the evolutionary relationship between members of the poplar TALE gene family, we constructed an unrooted phylogenetic tree using the amino acid sequences of Arabidopsis and poplar (Fig. 1). According to the previous classification in Arabidopsis (Hamant and Pautot, 2010), we divided the poplar TALE proteins into four classes, that is KNOX Class I (ten members), KNOX Class II (four), KNOX Class III (one), and BELL (twenty) (Fig. 1). The analysis of SMART database also shows that members of the same class have congruent protein domains (Table S2). In addition, the ML-phylogenetic tree, which uses the maximum likelihood method and best model of JTT + G + I + F selected by the SMS software, also confirms classification results (Fig. S1) (Guindon et al., 2010; Lefort et al., 2017). We then performed multiple sequence alignments of the classes and identified the various domains (Fig. S2; Fig. S3). As shown in the figures, all of the TALE members share the homeodomain with three inserted residues (P—Y—P), except for PtTALE29.

By predicting the physicochemical properties of poplar TALE family members, we found that their lengths and molecular weights vary substantially. The average length of the 35 proteins is 526.43 amino acids (181–835 amino acids) and the average molecular weight is 58.396 kDa, ranging from 20.177 kDa to 91.972 kDa. And the average length (667.35 amino acids) and average molecular weight (73.532 kDa) of BELL class members are much larger than others (340 amino acids and 38.459 kDa, 374.25 amino acids and 42.108 kDa, 181 amino acids and 20.177 kDa). The theoretical isoelectric points of these proteins also differ significantly (4.88–9.22), with the maximum variation in the BELL class (5.97–9.22). The theoretical isoelectric points of other classes are between 4.88 and 6.14 (KNOX Class I), 5.81 and 6.23 (KNOX Class II), or 5.6 (KNOX Class III). The instability index (II) of the proteins ranged from 35.01 to 61.68, with the KNOX Class II members possessed a highest mean (56.2875). Only three proteins from the other classes have the indexes of less than 40, which suggests they may be stable in the test tube. The aliphatic indexes are between 58.83 and 81.85. The grand average of hydropathicity (GRAVY) ranged from −0.954 to −0.366, indicating that these proteins are all hydrophilic proteins (Table 1).

3.3. Sequence structure features analysis

We characterized the gene intron/exon structures of the TALE genes (Fig. 2A). As we expected, members of the same class share similar gene structures. For example, the BELL class members contain three introns and four exons, except for *PtTALE17* with four introns and five exons. The majority of the KNOX Class I members harbor four introns and five exons. We also observed that close members within a class have similar length of exons.

In order to explore the conserved motifs contained in the poplar TALE family, we used MEME software to search the protein sequences from the 35 genes (Fig. 2B) (Bailey et al., 2009). We have obtained a total of 12 conserved protein motifs, in which motif 1 is annotated as homeodomain, motifs 2, 7, 8 as POX domain, motif 3 as KNOX2 domain, motif 4 as ELK domain, and motif 5 as KNOX1. The rest have no annotation (Table S3). As we expected, all of the TALE proteins have motif 1 except PtTALE29. All members in the BELL class contain motif 2 and motif 7 annotated as POX domain, which is also present in the *Arabidopsis thaliana* BLH1 protein. In addition, motif 8, annotated as POX domain, is also present in the majority of the BELL members except for PtTALE17 and PtTALE35. All KNOX Class I and KNOX Class II members contain domains, which were annotated as KNOX1 (motif 5), KNOX2 (motif 3), and ELK (motif 4) domains. KNOX Class III member harbors only KNOX1 (motif 5) and KNOX2 (motif 3) domains. These results are consistent with previous studies in Arabidopsis (Hamant and Pautot, 2010). For those motifs that have no annotation, motif 6 exists only within the BELL members. Motif 9 exists only in several BELL members. Motif 10 distributes within the BELL members except for four proteins. Motif 11 appears in the TALE family members except for three proteins. Motif 12 exists only in all of the members of KNOX Class II. It is worth noting that the proteins of the same class, especially those have closer relationship, harbor similar conserved motifs, suggesting that they may have similar functions.

3.4. Tissue-differential gene expression without treatment

We calculated correlation of the RNA-seq data used in this study and found that the pearson correlation coefficients between biological replicates are between 0.969 and 0.997, which are significant ($p < 2.2e\text{-}16$).

In order to dig into the differential expression patterns of the poplar TALE family genes, we compared gene expression between pairwise tissues; that is, leaf-root, leaf-stem, stem-root pairs. We found 19 DEGs between leaves and roots, 24 between leaves and stems, and 20 between stems and roots (Data S1). There is a distinct gene expression pattern over the three tissues. Out of the 35 TALE genes, 17 genes are over-expressed in stems and lowly expressed in other tissues (Data S1; Fig. S4). In contrast, nine genes are highly expressed in leaves and lowly expressed in stems (Data S1; Fig. S4). We then compared these sets of DEGs and extracted the shared genes. In the leaf-root and leaf-stem combinations, we obtained 14 shared genes, meaning that these genes

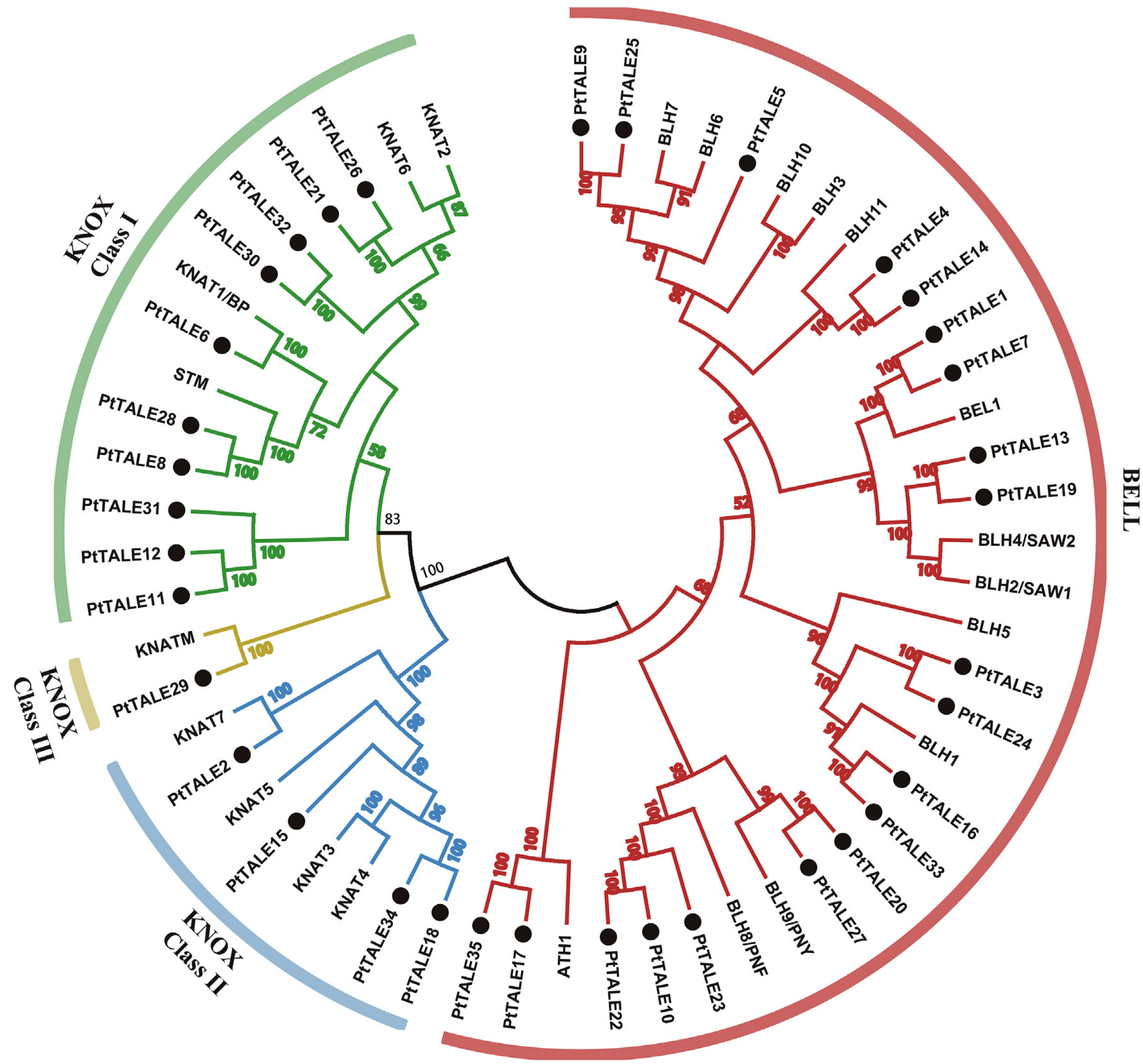

Fig. 1. Dendrogram of the TALE gene family proteins. The dendrogram was constructed by MEGA5.05 with the neighbor method and Poisson model. Poplar TALE proteins are marked with solid black circle, and each color represents a special class. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

are differentially expressed in leaves relative to stem and root tissues. Similarly, we obtained 13 shared genes in the root-leaf and root-stem combinations; 15 shared genes in the stem-leaf and stem-root combinations (Fig. 3A–C). These genes are all DEGs in one tissue relative to the others. We then compared the three sets of shared genes and obtained eight shared genes, meaning that these genes are differentially expressed in any two of the three tissues (Fig. 3D; Table S4).

To explore the expression patterns of these eight genes, we drew the heat map based on the RNA-Seq data; that is, fragments per kilo-bases per million mapped reads (FPKM) (Fig. 4). It is clear that the eight genes have different expression patterns. In general, these genes can be divided into three clusters. The members of cluster 1 and 3 are over expressed in leaves, moderately expressed in roots, and lowly expressed in stems. Cluster 2 genes display an opposite pattern to the genes of cluster 1 and 3, and are highly, moderately, and lowly expressed in stems, roots, and leaves (Fig. 4).

3.5. Gene expression in response to salt stress

In order to explore the expression pattern of TALE genes under salt stress, we analyzed gene expression in poplar with or without stress. The results showed that there are six DEGs (two up-regulated genes (*PtTALE4, 5*) and four down-regulated genes (*PtTALE7, 14, 20, 27*)) in the treated leaves compared to the control. Similarly, we obtained two DEGs in roots that are up-regulated (*PtTALE5, 16*), and six DEGs (one up-regulated (*PtTALE21*) and five down-regulated (*PtTALE6, 8, 20, 27, 28*)) in stems (Data S2). Subsequently, we calculated fold changes of the genes (Fig. S5). The magnitude of fold changes in down-regulated DEGs is at least two to four times that in the up-regulated DEGs. Expression of down-regulated genes exhibits greater variations than that of up-regulated genes in stems and leaves, especially in leaves.

Distribution of DEGs in different tissues is shown in the Venn diagram (Fig. 3E–G). Among the 11 DEGs that are responsive to salt stress, most of them are specific to a single tissue. Three of them respond to salt stress only in leaves, four genes only in stems, and one gene only in roots. In addition, two down-regulated genes respond to salt stress in both leaves and stems, and one up-regulated gene responds to salt stress in leaves and roots (Table S4).

3.6. Expression analysis of TALE genes under different salt concentrations

To investigate the temporal and spatial expression patterns of TALE genes under different salt stress levels, we analyzed the expression of the 11 salt stress response genes in different tissues treated with 100 mM, 150 mM, or 200 mM NaCl (Fig. 5). The results showed that almost each gene could respond to the salt stresses in different tissues,

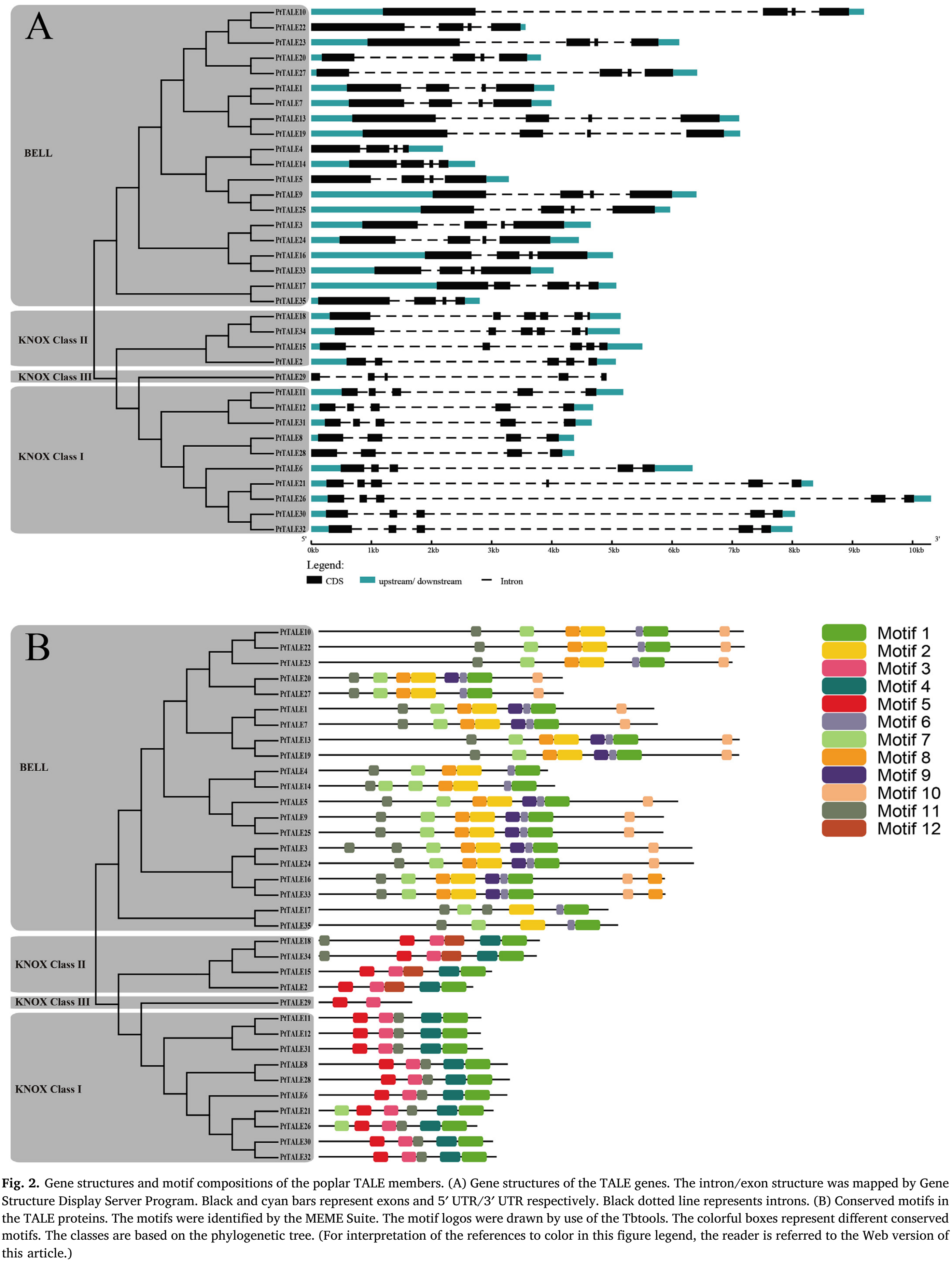

Fig. 2. Gene structures and motif compositions of the poplar TALE members. (A) Gene structures of the TALE genes. The intron/exon structure was mapped by Gene Structure Display Server Program. Black and cyan bars represent exons and 5′ UTR/3′ UTR respectively. Black dotted line represents introns. (B) Conserved motifs in the TALE proteins. The motifs were identified by the MEME Suite. The motif logos were drawn by use of the Tbtools. The colorful boxes represent different conserved motifs. The classes are based on the phylogenetic tree. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

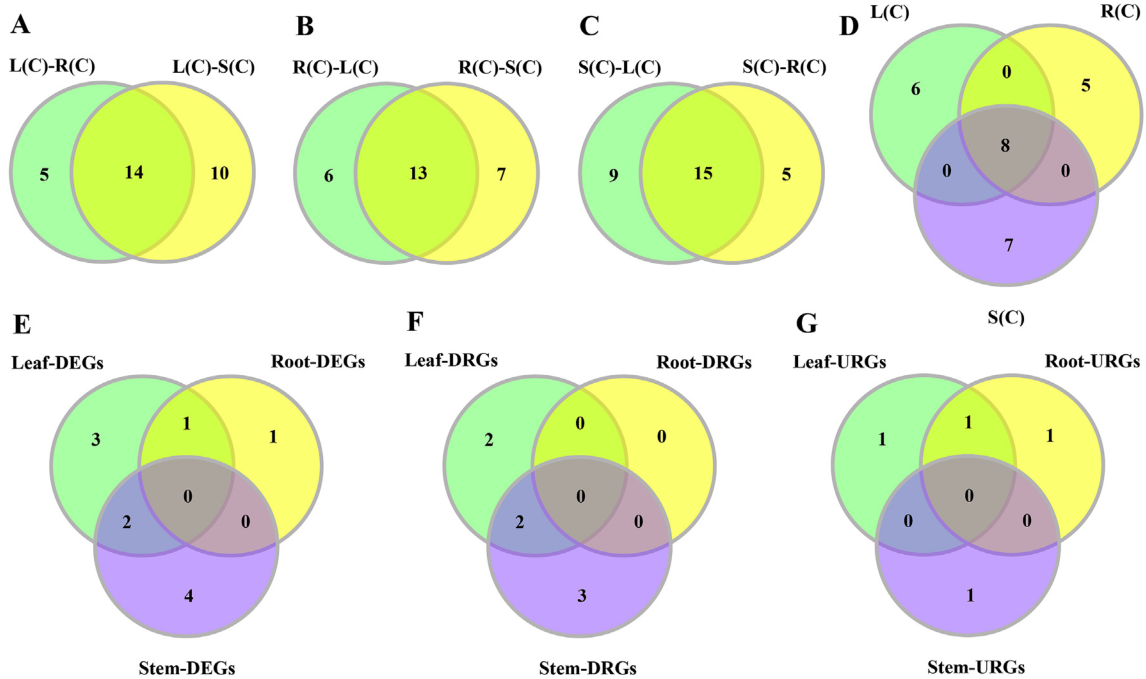

Fig. 3. Venn diagrams of differentially expressed TALE genes over tissues and in response to salinity. (A–C) Number of genes displaying distinct and shared expression in tissue pairs without treatment. L(C): leaves without treatment; S(C): stems without treatment; R(C): roots without treatment. (A) Comparison between leaf-root and leaf-stem pairs. (B) Comparison between root-leaf and root-stem pairs. (C) Comparison between stem-leaf and stem-root pairs. (D) Comparison among the three sets of shared genes from A to C. (E–G) Number of DEGs, down-regulated DEGs (DRGs), or up-regulated DEGs (URGs) in response to salt stress in each tissue.

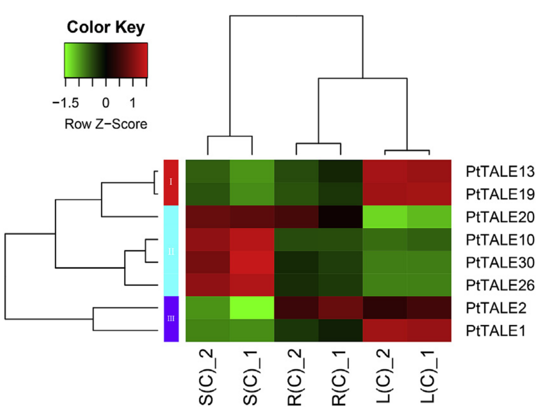

Fig. 4. A heatmap of eight genes shared by the pairwise tissue comparisons without treatment. L(C): leaves without treatment; S(C): stems without treatment; R(C): roots without treatment. The gene expression levels are square-root transformed FPKM values. We used the Z-score as standardization method for each gene. The colorful vertical bars on the left side denote gene clusters.

but their expression patterns were different. In the leaves treated with 100 mM NaCl, three genes (*PtTALE14, 20, 27*) were down-regulated during the early stages of stress and recovered at 24 h; three genes (*PtTALE6, 7, 8*) maintained low expression levels after stress; four genes (*PtTALE5, 16, 21, 28*) reached the highest expression level at 6 h (Fig. 5A; 5D). In the roots treated with 100 mM NaCl, almost every gene was up-regulated after stress and reached the highest expression level at 24 h (Fig. 5B; 5E). In the stems treated with 100 mM NaCl, six genes (*PtTALE4, 6, 7, 16, 21, 28*) reached the highest expression level at 6 h; four genes (*PtTALE8, 14, 20, 27*) kept low expression levels after stress (Fig. 5C; 5F). In the leaves treated with 150 mM NaCl, five genes (*PtTALE6, 7, 14, 20, 27*) maintained low expression levels after stress (Fig. 5G; 5J). In the roots treated with 150 mM NaCl, five genes (*PtTALE4, 5, 14, 21, 28*) reached higher expression levels in the early stages of stress (Fig. 5H; 5K). In the stems treated with 150 mM NaCl, three genes (*PtTALE6, 20, 27*) kept low expression levels after stress, and two genes (*PtTALE5, 16*) were opposite with them (Fig. 5I; 5L). In the leaves treated with 200 mM NaCl, eight genes (*PtTALE4, 6, 7, 14, 16, 20, 27, 28*) were down-regulated after stress (Fig. 5M; 5P). In the stem treated with 200 mM NaCl, six genes (*PtTALE6, 8, 14, 20, 27, 28*) maintained low expression levels after stress, and three genes (*PtTALE4, 16, 21*) reached the highest expression levels at 6 h (Fig. 5O; 5R). It is worth noting that some genes belonging to the same group had similar expression patterns. For example, *PtTALE14, 20, 27*, the members of BELL group, had similar expression patterns. In addition, *PtTALE5* was generally up-regulated after salt stress, suggesting that it may play an important role in salt stress response.

3.7. Examination of TALE gene expression by RT-qPCR

In order to verify the results of RNA-Seq, we performed RT-qPCR experiments using 32 TALE genes. As shown in Fig. 6, the trends of relative gene expression by the two methods are comparable. For example, tissue-differentially expressed genes, such as *PtTALE1, PtTALE13, and PtTALE19,* are down-regulated over four fold changes in roots, compared to leaves in both RT-qPCR and RNA-Seq (Fig. 6). On the other hand, *PtTALE3* and *PtTALE24* are up-regulated with over 5.5 fold changes in leaves and stems, compared to roots. *PtTALE23* is up-regulated over two fold changes in stems, compared to leaves.

3.8. Protein interaction analysis

To test the hypothesis that proteins encoded by the poplar TALE gene family can form heterodimers, we performed yeast two-hybrid experiments, using four genes that are up-regulated in the stems compared to the leaves. The amplified coding region sequences are shown in Table S5. Autoactivation test experiments showed that the yeast strains containing the pGBKT7-PtTALE4 or the pGBKT7-PtTALE20 vector can grow on the SD/-Trp solid medium, but not on the SD/-Trp/-His/-Ade/X-α-Gal solid medium (Fig. 7A). The positive control can grow on both media and turns blue on the SD/-Trp/-His/-Ade/X-α-Gal

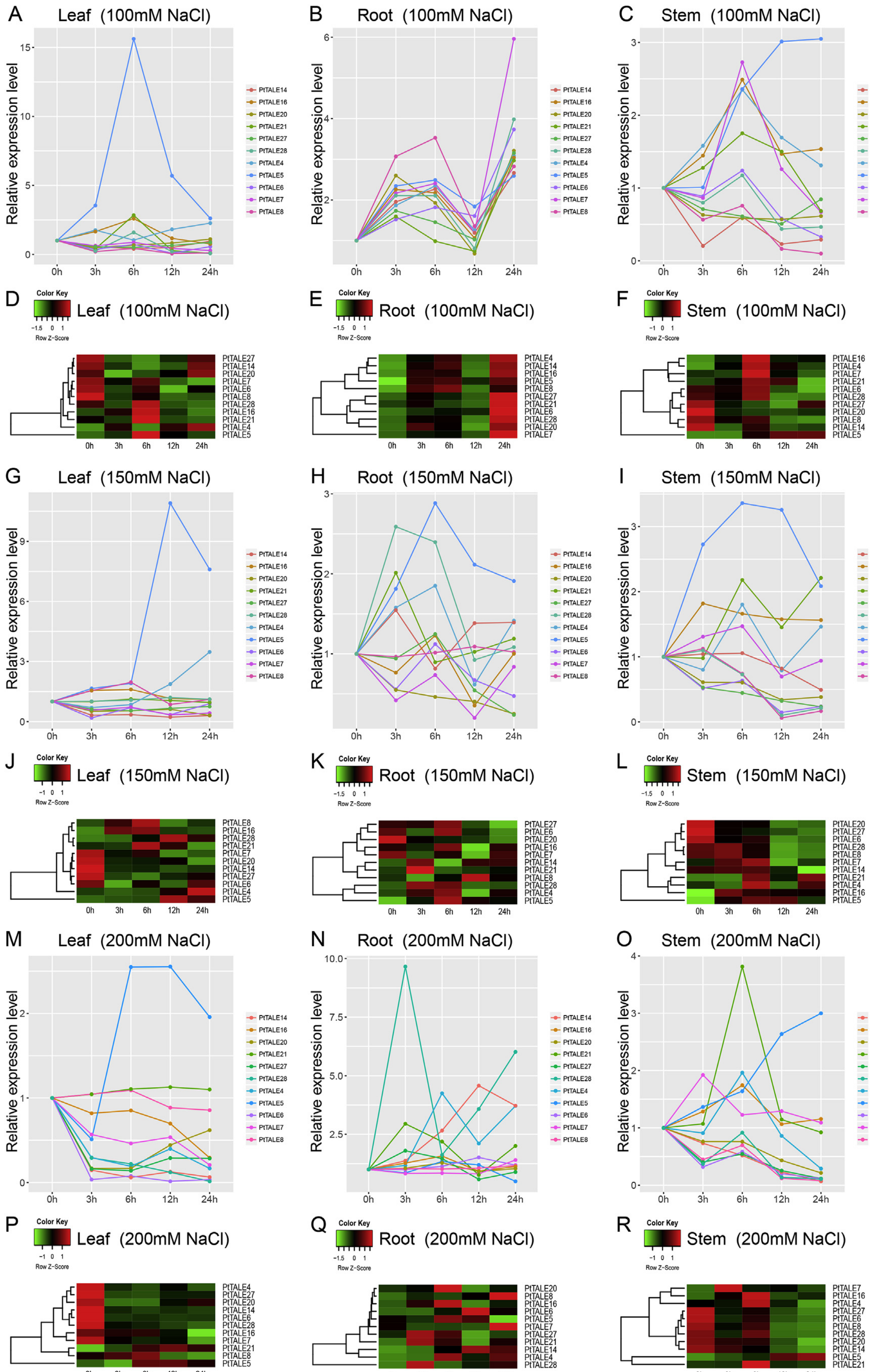

Fig. 5. Spatio-temporal expression patterns of 11 TALE genes under different salt stress levels. A–F: Lineplots and heatmaps of relative gene expression levels treated with 100 mM NaCl. G–L: Lineplots and heatmaps of relative gene expression levels treated with 150 mM NaCl. M–R: Lineplots and heatmaps of relative gene expression levels treated with 200 mM NaCl. The expression levels of each gene were calculated relative to its expression level at 0 h. We used the Z-score as standardization method for each gene in heatmaps.

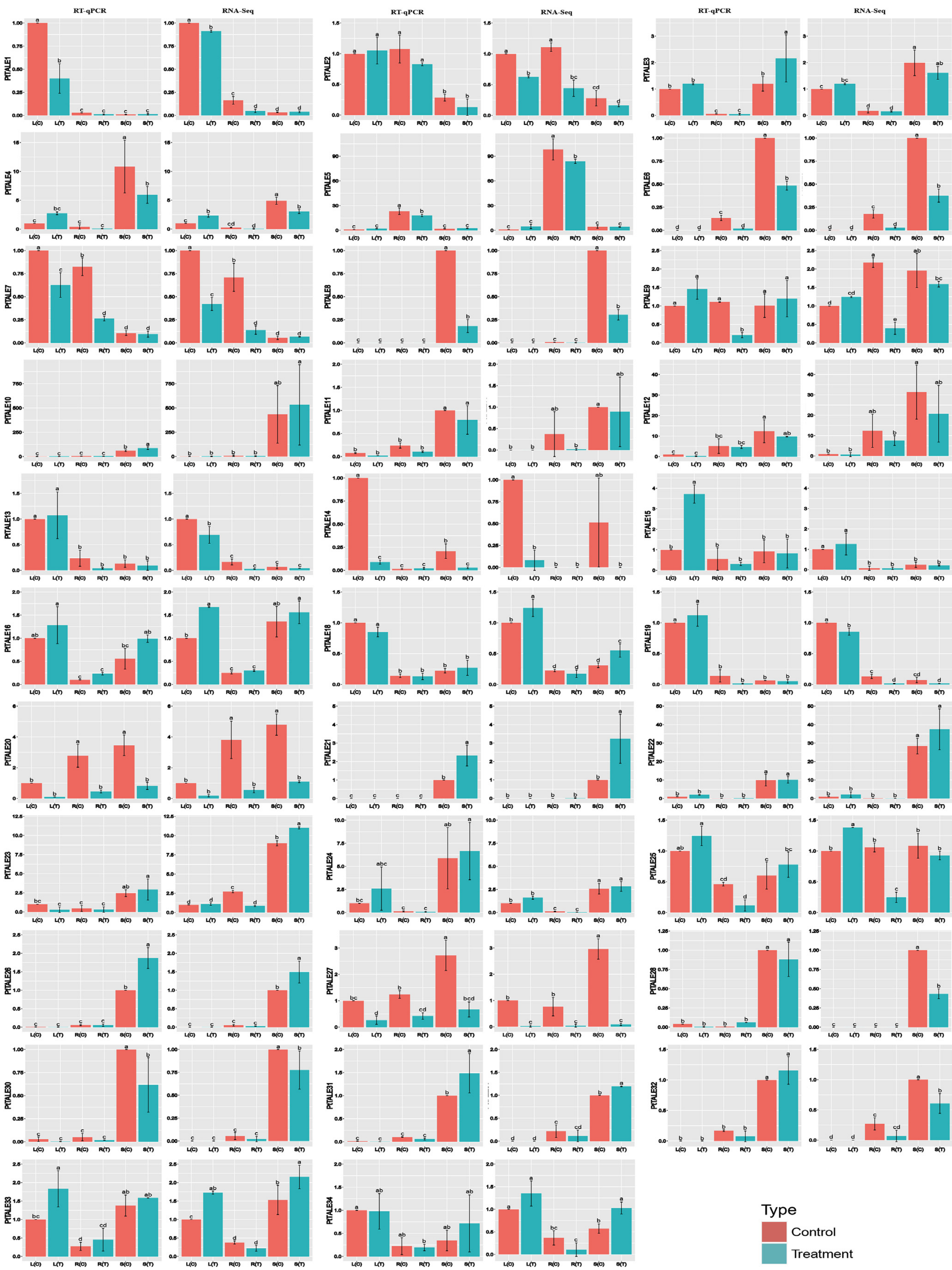

Fig. 6. Barplots of relative gene expression levels using RT-qPCR and RNA-Seq. L(C): leaves without treatment; S(C): stems without treatment; R(C): roots without treatment; L(T): leaves with salt treatment; S(T): stems with salt treatment; R(T): roots with salt treatment. The expression levels of each gene were calculated in relevance to corresponding gene expression in leaf or stem without treatment. Error bars represent standard deviation of biologic replicates. Significant differences ($P < 0.05$) are indicated by different lower case letters.

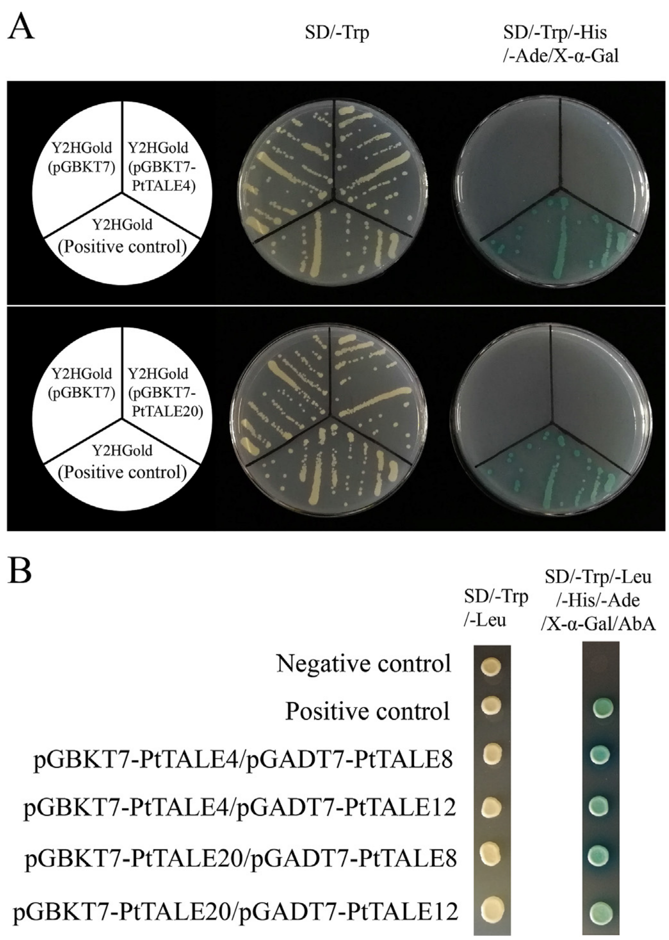

Fig. 7. Autoactivation and interaction assay. (A) Autoactivation test experiments of PtTALE4 and PtTALE20. The empty pGBKT7 vector transformed yeast cells were used as negative control. The pGBKT7-53/pGADT7-T co-transformed yeast cells were used as positive control. (B) Yeast two-hybrid experiments. The pGBKT7-Lam/pGADT7-T and pGBKT7-53/pGADT7-T co-transformed yeast cells were used as negative and positive control respectively.

solid medium (Fig. 7A). It indicates that the two transcription factors have no autoactivation. Yeast two-hybrid experiments demonstrated that co-transfection of gene combinations, including respective pGBKT7-PtTALE4/pGADT7-PtTALE8, pGBKT7-PtTALE4/pGADT7-PtTALE12, pGBKT7-PtTALE20/pGADT7-PtTALE8, pGBKT7-PtTALE20/pGADT7-PtTALE12 complex vectors, can grow on the SD/-Trp/-Leu and SD/-Trp/-Leu/-His/-Ade/X-α-Gal/AbA solid media, and turn blue on SD/-Trp/-Leu/-His/-Ade/X-α-Gal/AbA solid media (Fig. 7B). This is comparable to the positive control (Fig. 7B). It clearly indicates that the pairwise transcription factors can interact with one another in gene regulation.

4. Discussion

The poplar TALE gene family is widely found in eukaryotes and plays an important role in regulating growth and development. Previous studies on plants were focused on the whole homeobox genes across multiple species (Mukherjee et al., 2009). In this study, we investigated 35 poplar TALE gene family members (15 KNOX proteins, 20 BELL proteins), of which 34 genes were obtained from the PlantTFDB database (Jin et al., 2017), and one is the poplar gene homologous to the Arabidopsis *KNATM* gene without the homeodomain (Magnani and Hake, 2008). As expected, they are all predicted to locate in the nucleus. And these genes are unevenly distributed on 16 chromosomes.

To explore the evolutionary relationships and classification of the proteins encoded by the 35 genes, we constructed an unrooted phylogenetic tree based on their amino acid sequences. We divided these proteins into four classes that are similar to previous studies in Arabidopsis (Hamant and Pautot, 2010). The KNOX genes are distributed in three classes, but BELL genes are included in one class. The annotations of the SMART database and the ML-phylogenetic tree constructed using the best model also support our classification results. Each of the classes is represented by specific domains or domain combinations.

Evidence from physicochemical properties indicated that the poplar TALE proteins are all hydrophilic proteins, but they differ significantly in other aspects. Among them, amino acid lengths and molecular weights of BELL proteins are much larger than the others. And the theoretical isoelectric points of BELL proteins also vary a lot. The KNOX Class II proteins have a high coefficient of instability. The results of gene structure analysis showed that all TALE genes contain introns. The members of the same class share similar gene structures. And the close members in the phylogenetic tree have similar exon lengths. Analysis and annotation of protein motifs revealed that members of the same class contain similar protein motifs, which is consistent with previous studies of the poplar bHLH family (Zhao et al., 2018). It is worth noting that there are several unannotated motifs specific to one class, which suggests that they may have different functions compared to other proteins. This requires future explorations.

Previous studies have shown that the TALE gene family plays an important role in plant growth and development, so we explored the tissue-differential expression of poplar TALE genes based on RNA-Seq data. There is a distinct gene expression pattern over the three tissues. Out of the 35 TALE genes, 17 genes are highly expressed in stems and lowly expressed in other tissues (Data S1; Fig. S4). In contrast, nine genes are highly expressed in leaves and lowly expressed in stems (Data S1; Fig. S4).

We screened DEGs between tissues, and identified eight genes that are differentially expressed and shared by the three sets of pairwise tissue comparisons. These include five genes of the BELL class, two KNOX Class I genes, and one KNOX Class II gene.

We then annotated the eight genes onto the Arabidopsis genome to better understand their function by use of Phytozome (Table S6) (Goodstein et al., 2012; Lamesch et al., 2012). *SAW2*, the best-hit homologous gene of *PtTALE19* and *PtTALE13*, is associated with development of leaf shape (Kumar et al., 2007). The homologous gene of *PtTALE26* and *PtTALE30* in Arabidopsis is *KNAT6*, which is involved in meristem activity and organ separation (Belles-Boix et al., 2006). *BEL1*, the homologous gene of *PtTALE1*, is related to the development of ovule in Arabidopsis (Robinson-Beers et al., 1992). Both *PtTALE2* and its homologous gene (*KNAT7*) are related to secondary wall formation (Li et al., 2012). *BLH9*, the homologous gene of *PtTALE20*, regulates inflorescence stem growth in Arabidopsis (Bhatt et al., 2004). *BLH8*, homologous to *PtTALE10*, plays an important role in the transition from vegetative to reproductive development in Arabidopsis (Smith et al., 2004). In all, the DEGs may play a significant role in poplar growth and development.

Since tolerance to abiotic stresses, such as high salinity, is closely related to plant growth and development, we then investigated gene expression levels of the 35 poplar TALE genes under salt stress. The results indicated that 11 genes are responsive to the treatment, and the majority of them represent only in a single tissue. Only three remaining genes can respond to salt stress in two tissues. The number and variation of down-regulated DEGs are greater than that of up-regulated ones. It is interesting that all of the salt responsive genes belong to either the BELL class (seven genes) or the KNOX Class I (four genes), which suggests that genes in these two classes are related to salt stress response.

Furthermore, we then annotated these salt-responsive genes onto the Arabidopsis genome (Table S6). As we expected, all but the functionally unknown genes are associated with growth and development of Arabidopsis. For example, *BLH1*, homologous to *PtTALE16*, is associated with seed germination and early seedling development (Kim et al., 2013). *STM*, homologous to *PtTALE8 and PtTALE28*, participates in shoot apical meristem formation (Long et al., 1996). *KNAT1*, homologous to *PtTALE6*, is related to leaf development (Chuck et al., 1996).

Under different salt stress levels, the 11 salt stress response genes had different spatio-temporal expression patterns. Overall, gene expressions in stems or leaves were similar under different salt stress levels compared to roots. For example, in leaves, five genes (*PtTALE6, 7, 14, 20, 27*) were down-regulated in the early stages of stresses. In stems, four genes (*PtTALE8, 14, 20, 27*) maintained low expression levels after stress. We propose that the complexity of gene expression in roots may be due to the fact that roots are the primary tissues exposed to salt stress. It is worth noting that some genes belonging to the same group (*PtTALE14, 20, 27*) had similar expression patterns, and *PtTALE20* and *PtTALE27* are in the closest position of the phylogenetic tree, which suggests that they may have similar functions. In addition, *PtTALE5* was generally up-regulated after stress, suggesting that it has important functions in response to salt stress. But the gene functions need to be further verified by transgenes.

Previous study by large-scale yeast two-hybrid experiments indicated that many proteins encoded by the Arabidopsis TALE gene family can interact with one another (Hackbusch et al., 2005). This intrigued us to testify the interactions in poplar. Previous studies on poplar indicated that few members from the TALE gene family are associated with secondary growth and wood formation (Du et al., 2009; Liu et al. 2015a, 2015b; Petzold et al., 2018). We then selected two BELL genes and two KNOX genes, which are over-expressed in poplar stems, for yeast two-hybridization. Results indicated that the poplar BELL proteins can interact with the KNOX proteins (Fig. 7). Therefore, proteins of poplar TALE gene family can interact with one another to form heterodimers. But this is a preliminary verification, and more interactions between the poplar TALE proteins need to be explored using multiple assays in further studies.

5. Conclusion

In this study, we focus our studies on the 35 TALE gene family members in poplar. All of the TALE members are predicted to be located in nucleus and are unevenly distributed on 16 chromosomes. They can be divided into four classes with distinct gene structures and motif compositions. Evidence from transcript profiling indicated that the majority of the TALE genes exhibited distinct expression patterns over leaf, stem, and root tissues. Out of the 35 genes, 17 genes are highly expressed in stems, suggesting that the TALE gene family may play an important role in growth and wood formation. Furthermore, out of the 35 genes, 11 genes are responsive to salt stress. RT-qPCR analysis revealed spatio-temporal expression patterns of the 11 genes under different salt concentrations. Our experiments in yeast two-hybridization indicated that poplar TALE proteins from different classes can form heterodimers. These studies lay the foundation for future studies on biological functions of poplar TALE gene family.

Author contribution statement

BZ and TJ designed research. KZ conducted experiments and data analysis, and wrote the manuscript. XZ, ZC, and WY conducted data analysis. RL revised the manuscript. All authors read and approved the manuscript.

Conflicts of interest

The authors declare that they have no conflict of interest.

Acknowledgements

This work was supported by the Fundamental Research Funds for the Central Universities (2572018AA13) and the 111 Project (B16010).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.plaphy.2019.01.003.

References

Anders, S., Huber, W., 2010. Differential expression analysis for sequence count data. Genome Biol. 11 (10), R106. https://doi.org/10.1186/gb-2010-11-10-r106.
Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., Noble, W.S., 2009. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 37, W202–W208. https://doi.org/10.1093/nar/gkp335. Web Server issue).
Bellaoui, M., Pidkowich, M.S., Samach, A., Kushalappa, K., Kohalmi, S.E., Modrusan, Z., Crosby, W.L., Haughn, G.W., 2001. The Arabidopsis BELL1 and KNOX TALE homeodomain proteins interact through a domain conserved between plants and animals. Plant Cell 13 (11), 2455–2470.
Belles-Boix, E., Hamant, O., Witiak, S.M., Morin, H., Traas, J., Pautot, V., 2006. KNAT6: an Arabidopsis homeobox gene involved in meristem activity and organ separation. Plant Cell 18 (8), 1900–1907. https://doi.org/10.1105/tpc.106.041988.
Bhatt, A.M., Etchells, J.P., Canales, C., Lagodienko, A., Dickinson, H., 2004. VAAMANA–a BEL1-like homeodomain protein, interacts with KNOX proteins BP and STM and regulates inflorescence stem growth in Arabidopsis. Gene 328, 103–111. https://doi.org/10.1016/j.gene.2003.12.033.
Billeter, M., Qian, Y.Q., Otting, G., Muller, M., Gehring, W., Wuthrich, K., 1993. Determination of the nuclear magnetic resonance solution structure of an Antennapedia homeodomain-DNA complex. J. Mol. Biol. 234 (4), 1084–1093. https://doi.org/10.1006/jmbi.1993.1661.
Brambilla, V., Battaglia, R., Colombo, M., Masiero, S., Bencivenga, S., Kater, M.M., Colombo, L., 2007. Genetic and molecular interactions between BELL1 and MADS box factors support ovule development in Arabidopsis. Plant Cell 19 (8), 2544–2556. https://doi.org/10.1105/tpc.107.051797.
Burglin, T.R., 1997. Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. Nucleic Acids Res. 25 (21), 4173–4180.
Burglin, T.R., Affolter, M., 2016. Homeodomain proteins: an update. Chromosoma 125 (3), 497–521. https://doi.org/10.1007/s00412-015-0543-8.
Chen, C., Xia, R., Chen, H., He, Y., 2018. TBtools, a Toolkit for Biologists integrating various HTS-data handling tools with a user-friendly interface. bioRxiv, 289660.
Chuck, G., Lincoln, C., Hake, S., 1996. KNAT1 induces lobed leaves with ectopic meristems when overexpressed in Arabidopsis. Plant Cell 8 (8), 1277–1289. https://doi.org/10.1105/tpc.8.8.1277.
Du, J., Mansfield, S.D., Groover, A.T., 2009. The Populus homeobox gene ARBORKNOX2 regulates cell differentiation during secondary growth. Plant J. : for cell and molecular biology 60 (6), 1000–1014. https://doi.org/10.1111/j.1365-313X.2009.04017.x.
Finn, R.D., Attwood, T.K., Babbitt, P.C., Bateman, A., 2017. InterPro in 2017-beyond protein family and domain annotations 45 (D1), D190–d199. https://doi.org/10.1093/nar/gkw1107.
Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, Se, Wilkins, M.R., Appel, R.D., Bairoch, A., 2005. Protein identification and analysis tools on the ExPASy server. In: Walker, J.M. (Ed.), The Proteomics Protocols Handbook. Humana Press, Totowa, NJ, pp. 571–607. https://doi.org/10.1385/1-59259-890-0:571.
Gehring, W.J., Affolter a, M., Burglin, T., 1994. Homeodomain proteins. Annu. Rev. Biochem. 63 (1), 487–526. https://doi.org/10.1146/annurev.bi.63.070194.002415.
Ghate, T.H., Sharma, P., Kondhare, K.R., Hannapel, D.J., Banerjee, A.K., 2017. The mobile RNAs, StBEL11 and StBEL29, suppress growth of tubers in potato. Plant Mol. Biol. 93 (6), 563–578. https://doi.org/10.1007/s11103-016-0582-4.
Gomez-Mena, C., Sablowski, R., 2008. ARABIDOPSIS THALIANA HOMEOBOX GENE1 establishes the basal boundaries of shoot organs and controls stem growth. Plant Cell 20 (8), 2059–2072. https://doi.org/10.1105/tpc.108.059188.
Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., Rokhsar, D.S., 2012. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. 40, D1178–D1186. https://doi.org/10.1093/nar/gkr944. Database issue.
Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59 (3), 307–321. https://doi.org/10.1093/sysbio/syq010.
Hackbusch, J., Richter, K., Muller, J., Salamini, F., Uhrig, J.F., 2005. A central role of Arabidopsis thaliana ovate family proteins in networking and subcellular localization of 3-aa loop extension homeodomain proteins. Proc. Natl. Acad. Sci. U. S. A 102 (13), 4908–4912. https://doi.org/10.1073/pnas.0501181102.
Hall, T.A., 1999. BioEdit : a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41 (41), 95–98.
Hamant, O., Pautot, V., 2010. Plant development: a TALE story. Comptes Rendus Biol. 333 (4), 371–381. https://doi.org/10.1016/j.crvi.2010.01.015.

Hay, A., Tsiantis, M., 2009. A KNOX family TALE. Curr. Opin. Plant Biol. 12 (5), 593–598. https://doi.org/10.1016/j.pbi.2009.06.006.

Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., Nakai, K., 2007. WoLF PSORT: protein localization predictor. Nucleic Acids Res. 35, W585–W587. https://doi.org/10.1093/nar/gkm259. Web Server issue).

Hu, B., Jin, J., Guo, A.Y., Zhang, H., Luo, J., Gao, G., 2015. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31 (8), 1296–1297. https://doi.org/10.1093/bioinformatics/btu817.

Jin, J., Tian, F., Yang, D.C., Meng, Y.Q., Kong, L., Luo, J., Gao, G., 2017. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Res. 45 (D1), D1040–D1045. https://doi.org/10.1093/nar/gkw982.

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.Y., Lopez, R., Hunter, S., 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30 (9), 1236–1240. https://doi.org/10.1093/bioinformatics/btu031.

Kim, D., Cho, Y.H., Ryu, H., Kim, Y., Kim, T.H., Hwang, I., 2013. BLH1 and KNAT3 modulate ABA responses during germination and early seedling development in Arabidopsis. Plant J. : for cell and molecular biology 75 (5), 755–766. https://doi.org/10.1111/tpj.12236.

Kimura, S., Koenig, D., Kang, J., Yoong, F.Y., Sinha, N., 2008. Natural variation in leaf morphology results from mutation of a novel KNOX gene. Curr. Biol. : CB 18 (9), 672–677. https://doi.org/10.1016/j.cub.2008.04.008.

Kumar, R., Kushalappa, K., Godt, D., Pidkowich, M.S., Pastorelli, S., Hepworth, S.R., Haughn, G.W., 2007. The Arabidopsis BEL1-LIKE HOMEODOMAIN proteins SAW1 and SAW2 act redundantly to regulate KNOX expression spatially in leaf margins. Plant Cell 19 (9), 2719–2735. https://doi.org/10.1105/tpc.106.048769.

Lamesch, P., Berardini, T.Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D.L., Garcia-Hernandez, M., Karthikeyan, A.S., Lee, C.H., Nelson, W.D., Ploetz, L., Singh, S., Wensel, A., Huala, E., 2012. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res. 40, D1202–D1210. https://doi.org/10.1093/nar/gkr1090. Database issue.

Lefort, V., Longueville, J.E., Gascuel, O., 2017. SMS: smart model selection in PhyML. Mol. Biol. Evol. 34 (9), 2422–2424. https://doi.org/10.1093/molbev/msx149.

Letunic, I., Bork, P., 2018. 20 years of the SMART protein domain annotation resource. Nucleic Acids Res. 46 (D1), D493–d496. https://doi.org/10.1093/nar/gkx922.

Letunic, I., Doerks, T., Bork, P., 2015. SMART: recent updates, new developments and status in 2015. Nucleic Acids Res. 43, D257–D260. https://doi.org/10.1093/nar/gku949. Database issue.

Li, E., Bhargava, A., Qiang, W., Friedmann, M.C., Forneris, N., Savidge, R.A., Johnson, L.A., Mansfield, S.D., Ellis, B.E., Douglas, C.J., 2012. The Class II KNOX gene KNAT7 negatively regulates secondary wall formation in Arabidopsis and is functionally conserved in Populus. New Phytol. 194 (1), 102–115. https://doi.org/10.1111/j.1469-8137.2011.04016.x.

Lin, T., Sharma, P., Gonzalez, D.H., Viola, I.L., Hannapel, D.J., 2013. The impact of the long-distance transport of a BEL1-like messenger RNA on development. Plant Physiol. 161 (2), 760–772. https://doi.org/10.1104/pp.112.209429.

Liu, L., Ramsay, T., Zinkgraf, M., Sundell, D., Street, N.R., Filkov, V., Groover, A., 2015a. A resource for characterizing genome-wide binding and putative target genes of transcription factors expressed during secondary growth and wood formation in Populus. Plant J. : for cell and molecular biology 82 (5), 887–898. https://doi.org/10.1111/tpj.12850.

Liu, L., Zinkgraf, M., Petzold, H.E., Beers, E.P., Filkov, V., Groover, A., 2015b. The Populus ARBORKNOX1 homeodomain transcription factor regulates woody growth through binding to evolutionarily conserved target genes of diverse function. New Phytol. 205 (2), 682–694. https://doi.org/10.1111/nph.13151.

Long, J.A., Moan, E.I., Medford, J.I., Barton, M.K., 1996. A member of the KNOTTED class of homeodomain proteins encoded by the STM gene of Arabidopsis. Nature 379 (6560), 66–69. https://doi.org/10.1038/379066a0.

Magnani, E., Hake, S., 2008. KNOX lost the OX: the Arabidopsis KNATM gene defines a novel class of KNOX transcriptional regulators missing the homeodomain. Plant Cell 20 (4), 875–887. https://doi.org/10.1105/tpc.108.058495.

McGinnis, W., Garber, R.L., Wirz, J., Kuroiwa, A., Gehring, W.J., 1984a. A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans. Cell 37 (2), 403–408.

McGinnis, W., Levine, M.S., Hafen, E., Kuroiwa, A., Gehring, W.J., 1984b. A conserved DNA sequence in homoeotic genes of the Drosophila Antennapedia and bithorax complexes. Nature 308 (5958), 428–433.

Meng, L., Fan, Z., Zhang, Q., Wang, C., Gao, Y., Deng, Y., Zhu, B., Zhu, H., Chen, J., Shan, W., Yin, X., Zhong, S., Grierson, D., Jiang, C.Z., Luo, Y., Fu, D.Q., 2018. BEL1-LIKE HOMEODOMAIN 11 regulates chloroplast development and chlorophyll synthesis in tomato fruit. Plant J. : for cell and molecular biology 94 (6), 1126–1140. https://doi.org/10.1111/tpj.13924.

Mukherjee, K., Brocchieri, L., Burglin, T.R., 2009. A comprehensive classification and evolutionary analysis of plant homeobox genes. Mol. Biol. Evol. 26 (12), 2775–2794. https://doi.org/10.1093/molbev/msp201.

Pautot, V., Dockx, J., Hamant, O., Kronenberger, J., Grandjean, O., Jublot, D., Traas, J., 2001. KNAT2: evidence for a link between knotted-like genes and carpel development. Plant Cell 13 (8), 1719–1734.

Petzold, H.E., Chanda, B., Zhao, C., Rigoulot, S.B., Beers, E.P., Brunner, A.M., 2018. Divaricata and Radialis Interacting Factor (DRIF) also interacts with WOX and KNOX proteins associated with wood formation in Populus trichocarpa. Plant J. : for cell and molecular biology 93 (6), 1076–1087. https://doi.org/10.1111/tpj.13831.

Proveniers, M., Rutjens, B., Brand, M., Smeekens, S., 2007. The Arabidopsis TALE homeobox gene ATH1 controls floral competency through positive regulation of FLC. Plant J. : for cell and molecular biology 52 (5), 899–913. https://doi.org/10.1111/j.1365-313X.2007.03285.x.

Regier, N., Frey, B., 2010. Experimental comparison of relative RT-qPCR quantification approaches for gene expression studies in poplar. BMC Mol. Biol. 11, 57. https://doi.org/10.1186/1471-2199-11-57.

Robinson-Beers, K., Pruitt, R.E., Gasser, C.S., 1992. Ovule development in wild-type Arabidopsis and two female-sterile mutants. Plant Cell 4 (10), 1237–1249. https://doi.org/10.1105/tpc.4.10.1237.

Rosin, F.M., Hart, J.K., Horner, H.T., Davies, P.J., Hannapel, D.J., 2003. Overexpression of a knotted-like homeobox gene of potato alters vegetative development by decreasing gibberellin accumulation. Plant Physiol. 132 (1), 106–117. https://doi.org/10.1104/pp.102.015560.

Rutjens, B., Bao, D., van Eck-Stouten, E., Brand, M., Smeekens, S., Proveniers, M., 2009. Shoot apical meristem function in Arabidopsis requires the combined activities of three BEL1-like homeodomain proteins. Plant J. : for cell and molecular biology 58 (4), 641–654. https://doi.org/10.1111/j.1365-313X.2009.03809.x.

Sato, Y., Sentoku, N., Miura, Y., Hirochika, H., Kitano, H., Matsuoka, M., 1999. Loss-of-function mutations in the rice homeobox gene OSH15 affect the architecture of internodes resulting in dwarf plants. EMBO J. 18 (4), 992–1002. https://doi.org/10.1093/emboj/18.4.992.

Scott, M.P., Weiner, A.J., 1984. Structural relationships among genes that control development: sequence homology between the Antennapedia, Ultrabithorax, and fushi tarazu loci of Drosophila. Proc. Natl. Acad. Sci. U. S. A 81 (13), 4115–4119.

Shu, Y., Tao, Y., Wang, S., Huang, L., Yu, X., Wang, Z., Chen, M., Gu, W., Ma, H., 2015. GmSBH1, a homeobox transcription factor gene, relates to growth and development and involves in response to high temperature and humidity stress in soybean. Plant Cell Rep. 34 (11), 1927–1937. https://doi.org/10.1007/s00299-015-1840-7.

Sjodin, A., Street, N.R., Sandberg, G., Gustafsson, P., Jansson, S., 2009. The Populus genome integrative explorer (PopGenIE): a new resource for exploring the Populus genome. New Phytol. 182 (4), 1013–1025. https://doi.org/10.1111/j.1469-8137.2009.02807.x.

Smith, H.M., Campbell, B.C., Hake, S., 2004. Competence to respond to floral inductive signals requires the homeobox genes PENNYWISE and POUND-FOOLISH. Curr. Biol. : CB 14 (9), 812–817. https://doi.org/10.1016/j.cub.2004.04.032.

Sundell, D., Mannapperuma, C., Netotea, S., Delhomme, N., Lin, Y.C., Sjodin, A., Van de Peer, Y., Jansson, S., Hvidsten, T.R., Street, N.R., 2015. The plant genome integrative explorer resource: PlantGenIE.org. New Phytol. 208 (4), 1149–1156. https://doi.org/10.1111/nph.13557.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28 (10), 2731–2739. https://doi.org/10.1093/molbev/msr121.

Tao, Y., Chen, M., Shu, Y., Zhu, Y., Wang, S., Huang, L., Yu, X., Wang, Z., Qian, P., Gu, W., Ma, H., 2018. Identification and functional characterization of a novel BEL1-LIKE homeobox transcription factor GmBLH4 in soybean. Plant Cell Tissue Organ Cult. 134 (2), 331–344. https://doi.org/10.1007/s11240-018-1419-4.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25 (24), 4876–4882.

Tsuda, K., Hake, S., 2015. Diverse functions of KNOX transcription factors in the diploid body plan of plants. Curr. Opin. Plant Biol. 27, 91–96. https://doi.org/10.1016/j.pbi.2015.06.015.

Tsuda, K., Ito, Y., Sato, Y., Kurata, N., 2011. Positive autoregulation of a KNOX gene is essential for shoot apical meristem maintenance in rice. Plant Cell 23 (12), 4368–4381. https://doi.org/10.1105/tpc.111.090050.

Tuskan, G.A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R.R., Bhalerao, R.P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G.L., Cooper, D., Coutinho, P.M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroeve, S., Dejardin, A., Depamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehlting, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjarvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leple, J.C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D.R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouze, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Sterky, F., Terry, A., Tsai, C.J., Uberbacher, E., Unneberg, P., Vahala, J., Wall, K., Wessler, S., Yang, G., Yin, T., Douglas, C., Marra, M., Sandberg, G., Van de Peer, Y., Rokhsar, D., 2006. The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). Science 313 (5793), 1596–1604. https://doi.org/10.1126/science.1128691.

Vollbrecht, E., Veit, B., Sinha, N., Hake, S., 1991. The developmental gene Knotted-1 is a member of a maize homeobox gene family. Nature 350 (6315), 241–243. https://doi.org/10.1038/350241a0.

Yao, W., Wang, S., Zhou, B., Jiang, T., 2016. Transgenic poplar overexpressing the endogenous transcription factor ERF76 gene improves salinity tolerance. Tree Physiol. 36 (7), 896–908.

Zhao, K., Li, S., Yao, W., Zhou, B., Li, R., Jiang, T., 2018. Characterization of the basic helix-loop-helix gene family and its tissue-differential expression in response to salt stress in poplar. PeerJ 6, e4502. https://doi.org/10.7717/peerj.4502.