BIOCELL
2022 46(8): 1947-1958

Tech Science Press

Genome-wide identification of *NAC* gene family and expression analysis under abiotic stresses in *Salvia miltiorrhiza*

Xin LI[1]; Jianmin PAN[1]; Faisal ISLAM[2]; Juanjuan LI[1]; Zhuoni HOU[1]; Zongqi YANG[1]; Ling XU[1,*]

[1] Zhejiang Province Key Laboratory of Plant Secondary Metabolism and Regulation, College of Life Sciences and Medicine, Zhejiang Sci-Tech University, Hangzhou, 310018, China
[2] Institute of Crop Science, Ministry of Agriculture and Rural Affairs Laboratory of Spectroscopy Sensing, Zhejiang University, Hangzhou, 310058, China

Key words: *Salvia miltiorrhiza*, *Arabidopsis thaliana*, *NAC* transcription factors, Genome-wide analysis, Abiotic stress

Abstract: *NAC* (NAM, ATAF, CUC) is a class of transcription factors involved in plant growth regulation, abiotic stress responses, morphogenesis and metabolism. *Salvia miltiorrhiza* is an important Chinese medicinal herb, but the characterization of *NAC* genes in this species is limited. In this study, based on the *Salvia miltiorrhiza* genomic databases, 82 *NAC* transcription factors were identified, which were divided into 14 groups. Meanwhile, phylogenetic analysis, gene structure, chromosomal localization and potential role of *SmNAC*s in abiotic stress conditions were also studied. The results revealed that some *SmNAC*s had different structures than others, which advised that these genes may have multiple/distinct functions. Real-time quantitative polymerase chain reaction (RT-qPCR) analysis showed that *SmNAC*s exhibited differential expression patterns under salt and drought stress. The NaCl induced salinity treatments modulated the expression of several *SmNAC* genes more in roots compared with leaves. Conversely, under drought stress conditions, more genes were upregulated in leaves compared with roots. These results will be useful for the further study involved in the functional characteristics of *SmNAC* genes, especially in response to salt and drought stresses, thereby may facilitate genetic breeding in *Salvia miltiorrhiza*.

Introduction

Salvia miltiorrhiza (Lamiaceae), as an important traditional Chinese medicinal herb that is widely used to treat cardiovascular diseases (CVD) (Hou *et al.*, 2020; Wang *et al.*, 2017). *Salvia miltiorrhiza* contains water-soluble phenolic compounds, such as salvianolic acid A, salvianolic acid B and fat-soluble substances, including tanshinone I, tanshinone IIA, dihydrotanshinone, etc (Jung *et al.*, 2020). Modern pharmacology investigations have shown that these bioactive compounds have diverse pharmacological effects that were mediated through antiapoptotic, antioxidative, and anti-inflammatory responses (Wu *et al.*, 2016; Zhou *et al.*, 2011).

The *NAC* family is a largest class of plant transcriptional factors, whereas, term NAC is derived from closely related three protein families, namely no apical meristem (NAM), *Arabidopsis* transcription activation factor (ATAF)1/2) and cup-shaped cotyledon (CUC2) (Olsen *et al.*, 2005). A typical NAC proteins contain a N-terminal that is highly homologous to the DNA-binding region of NAC domain, composed of 150 amino acids (Nakashima *et al.*, 2012). The NAC domain consists of five conserved regions, and are named as A to E. The binding of NAC protein to its target is controlled by subdomains C and D (Mathew *et al.*, 2016). The C-terminal interacts with DNA or other transcription factors and has a highly variable transcriptional regulatory region (Shen *et al.*, 2019). Like other TF families such as *MYB*, *WRKY* and *AP2/ERF*, *NAC* is also involved in regulating various physiological and developmental processes in plants. The *NAC* gene family have been identified in many plants, for example, around 71 *NAC* TF were in *Solanum muricatum* (Yang *et al.*, 2021), 102 in *Theobroma cacao* (Shen *et al.*, 2019), 106 in *Prunus dulcis* (Zafar *et al.*, 2021), 142 in *Actinidia* (Jia *et al.*, 2021) and 151 in *Oryza sativa* (Nuruzzaman *et al.*, 2010). However, the diversity and numbers of *NAC* genes in *Salvia miltiorrhiza* are rarely reported.

The chromosome numbers of *Salvia miltiorrhiza* are 2n = 2x = 16, and it is considered as a tetraploids (Zhao *et al.*, 2006). Sequencing of the whole genome of *Salvia miltiorrhiza* was completed in 2020 (Song *et al.*, 2020). Its genome size is about 594,750,066 base pairs. In the present study, genome-wide identification and expression analysis of *NAC* genes under salt and drought conditions were studied.

*Address correspondence to: Ling Xu, lxu@zstu.edu.cn
Received: 16 October 2021; Accepted: 27 December 2021

Doi: 10.32604/biocell.2022.019806

The genomic structure, phylogenetic relationships and chromosome location of a *NAC* protein in *Salvia miltiorrhiza* were identified and discussed. A total of 82 *NAC* genes were identified and their expression patterns were analyzed under salinity and drought stress using real-time quantitative (RT-qPCR) PCR. This study is to lay a foundation for understanding the regulatory mechanism of *NAC* genes and may help to elucidate the molecular mechanisms underlying salt/drought stress responses, and to design optimal genetic improvement strategies in *Salvia miltiorrhiza*.

Materials and Methods

Database search and sequence analysis of NAC genes in Salvia miltiorrhiza

The *Salvia miltiorrhiza* genome sequence data was obtained from the China National Center for Bioinformation (https://bigd.big.ac.cn). The nucleotide sequences and amino acid data of 96 *AtNAC* genes were downloaded from The *Arabidopsis* Information Resource (TAIR, https://www.arabidopsis.org/index.jsp). The Hidden Markov Model (HMM, http://hmmer.org/) search was performed against the *Salvia miltiorrhiza* protein database using the NAM domain (PF02365, *e*-value ≤ 1.1e-10). The candidate sequences were submitted in the Pfam database (http://pfam.xfam.org) and NCBI Conserved Domains Database for verification (https://www.ncbi.nlm.nih.gov/cdd). The results showed that 82 NAC proteins exist the *Salvia miltiorrhiza* genomes (Table 1). The isoelectric point and protein molecular weight of *SmNAC* proteins were predicted using the pI/Mw tool on the online software ExPASy (http://web.expasy.org/protparam).

Multiple sequence alignment, and phylogenetic analysis

The ClustalW analysis was used to analyze the multiple sequence alignment on the *NAC* domain sequences of *Arabidopsis* and *Salvia miltiorrhiza NAC* genes. The MEGA 7.0 software and Maximum Likelihood method (ML, bootstrap = 1000) were used to construct a phylogenetic tree (Sudhir *et al*., 2018). The sequences for phylogenetic trees are shown in Table S1.

Chromosomal locations

The Perl script was used to extract mRNA location information from the *Salvia miltiorrhiza* GFF file to determine the *SmNAC* location, and TBtools 1.007 software was used to map 9 chromosomes in the *Salvia miltiorrhiza* genome.

Protein properties and sequence analysis

MEME online software (http://meme-suite.org/tools/meme) was used to analyze the conversed motif prediction of 82 *SmNAC* with the following conditions: zoops, 10 motifs with an optimum motif width between 10 and 50 residues, and any number of repetitions. The CDS and UTR location information in the mRNA of *SmNAC*s was selected from the GFF file of *Salvia miltiorrhiza* genomes for complete gene structure identification.

Plant materials and abiotic stress treatments

Salvia miltiorrhiza seeds were obtained in Shanxi Province, China. Healthy seeds were grown in potting soil composed of peat, perlite and vermiculite (4:3:1). When *Salvia miltiorrhiza* plants were five months old, uniform plants were selected for abiotic stress treatments. The plants were carefully uprooted and were directly transformed into the pots containing Hoagland solution with desired levels of salinity and drought treatments. For salt stress treatments, 100 mM, and 200 mM of NaCl induced salinity, and for drought 5%, and 10% polyethylene glycol (PEG6000) induced drought stress treatments were selected, whereas control treatment was without addition of any salt (NaCl) or PEG additions. Leaves and roots were harvested after 24 h of stress exposure, and were frozen in liquid nitrogen and stored at –80°C for total RNA extraction. All experiments were repeated with three biological replicates.

RNA extraction and cDNA synthesis

RNA was extracted using RNAprep Pure Plant Plus Kit (TIANGEN Kit, DP441) from the leaves and roots in *Salvia miltiorrhiza* for tissue-specific expression. Determination of RNA concentration and purity (OD260/OD280 = 1.8–2.1) was measured using Nucleic acid micrometer (Thermo, NANODROP 2000). The RNA was reversely transcribed into cDNA using the TaKaRa reverse transcription kit (Dalian, China) for qRT-PCR analysis.

Gene expression analyses by real-time quantitative PCR (RT-qPCR)

RT-qPCRs were performed using the ChamQ Universal SYBR qPCR Master Mix (Vazyme, China). Primers were designed using online tool: https://www.genscript.com/tools/real-time-pcr-taqman-primer-design-tool. The length of amplicons is between 150 bp and 200 bp. SmActin was used as the reference gene (Han *et al*., 2020). The gene expression was calculated using $2^{-\Delta\Delta CT}$ method. Initially, eighty-two genes were selected for RT-qPCR assay with three biological replicates. However, expression levels of thirty-two genes were low, and therefore fifty genes were selected for further analysis.

Statistical analysis

The gene expression data were presented as mean values of three biological replicates with standard error (SE). The fisher least significant difference (LSD) test to determine significant differences ($P \leq 0.05$). Heat map was performed using the OmicStudio tools at https://www.omicstudio.cn/tool.

Results

Identification of NAC gene family in Salvia miltiorrhiza

A sequencing data of 96 *NAC* genes were downloaded from TAIR. A total of 82 *NAC* proteins were identified from the *Salvia miltiorrhiza* genome after comparison with the Hidden Markov Model (HMM) search using the *NAM* domain (PF02365). For convenience, the 82 *NAC* genes were successively named *SmNAC1-SmNAC82*. The characteristic parameters of the predicted protein encoded by *SmNAC* are shown in Table 1. The approximate length of NAC protein ranged from 120 amino acid (*SmNAC36*) to 608 (*SmNAC25*) amino acids, with an average length of 311aa. The isoelectric point (pIs) of *NAC* proteins ranged from 4.53 (*SmNAC43*) to 9.73 (*SmNAC26*), and 55 pIs < 7.00, 27 pIs > 7.00. The molecular weights (MWs) ranged from 14014.88 Da (*SmNAC26*) to 68033.50 Da (*SmNAC2*).

TABLE 1

Sequence characteristics of *NAC* gene in *Salvia miltiorrhiza*

Name	Annotation ID	AA Len	pI	MW	CDS number	Chr	Group
SmNAC1	EVM0000014.1	436	4.93	36864.91	6	Chr5	Group7
SmNAC2	EVM0000553.1	572	4.86	68033.50	6	Chr6	Group5
SmNAC3	EVM0000946.1	182	5.21	19240.10	3	Chr4	Group11
SmNAC4	EVM0001254.1	192	5.33	37564.97	3	Chr8	Group11
SmNAC5	EVM0001825.1	338	8.55	31348.62	3	Chr7	Group10
SmNAC6	EVM0001826.1	330	6.46	27097.76	3	Chr3	Group8
SmNAC7	EVM0002275.1	246	5.12	21936.14	1	Chr4	Group14
SmNAC8	EVM0002346.1	209	9.61	28865.15	3	Chr3	Group13
SmNAC9	EVM0002598.1	195	8.29	21629.45	3	Chr4	Group14
SmNAC10	EVM0002612.1	562	4.84	22152.92	3	Chr4	Group5
SmNAC11	EVM0002774.1	333	5.79	63721.62	3	Chr6	Group1
SmNAC12	EVM0003692.2	324	8.84	24567.72	3	Chr3	Group1
SmNAC13	EVM0004898.1	329	7.13	33674.93	4	Chr1	Group9
SmNAC14	EVM0005048.1	213	5.70	38590.04	3	Chr8	Group8
SmNAC15	EVM0005269.1	376	5.34	32474.77	6	Chr7	Group7
SmNAC16	EVM0006891.1	289	6.75	35863.28	3	Chr5	Group1
SmNAC17	EVM0007240.1	301	5.50	21313.84	3	Chr4	Group1
SmNAC18	EVM0007342.1	342	6.37	22312.28	3	Chr4	Group1
SmNAC19	EVM0007446.1	535	5.79	37965.51	5	Chr8	Group3
SmNAC20	EVM0007949.1	193	6.21	44820.53	3	Chr6	Group13
SmNAC21	EVM0008601.1	208	9.51	33762.01	3	Chr1	Group6
SmNAC22	EVM0009029.1	191	9.46	44083.88	3	Chr6	Group13
SmNAC23	EVM0010196.1	306	5.20	31157.14	3	Chr7	Group10
SmNAC24	EVM0010890.1	554	5.66	32591.64	6	Chr1	Group4
SmNAC25	EVM0011250.1	608	4.84	28099.50	6	Chr3	Group2
SmNAC26	EVM0011295.1	194	9.73	14014.88	3	Chr9	Group10
SmNAC27	EVM0011989.1	551	4.83	24025.74	6	Chr4	Group2
SmNAC28	EVM0012415.1	285	8.82	31648.34	3	Chr7	Group7
SmNAC29	EVM0012594.1	354	6.10	20876.47	4	Chr4	Group4
SmNAC30	EVM0013110.1	444	5.00	31118.37	4	Chr7	Group10
SmNAC31	EVM0013308.1	286	7.69	30333.32	3	Chr2	Group8
SmNAC32	EVM0013738.1	194	9.59	16709.99	3	Chr4	Group10
SmNAC33	EVM0013784.1	297	8.47	63640.53	3	Chr6	Group9
SmNAC34	EVM0014582.1	346	4.57	38567.91	11	Chr8	Group12
SmNAC35	EVM0014742.1	363	7.81	37473.45	3	Chr8	Group8
SmNAC36	EVM0014951.1	120	6.90	30929.89	3	Chr2	Group8
SmNAC37	EVM0015253.1	396	6.21	40678.39	3	Chr8	Group1
SmNAC38	EVM0015811.1	287	8.85	33048.86	3	Chr1	Group10
SmNAC39	EVM0016132.1	285	6.76	35497.98	3	Chr5	Group10
SmNAC40	EVM0016762.1	268	8.85	35762.93	3	Chr5	Group9
SmNAC41	EVM0017168.1	321	8.31	31541.38	3	Chr7	Group1
SmNAC42	EVM0017178.1	187	7.01	19243.12	3	Chr4	Group8
SmNAC43	EVM0017275.1	184	4.53	64290.51	3	Chr6	Group11
SmNAC44	EVM0017545.1	269	8.56	37071.81	4	Chr5	Group10
SmNAC45	EVM0017672.1	265	5.12	37626.44	3	Chr8	Group3

(Continued)

Table 1 (continued).

Name	Annotation ID	AA Len	pI	MW	CDS number	Chr	Group
SmNAC46	EVM0018134.1	164	8.59	32851.13	3	Chr1	Group14
SmNAC47	EVM0018804.1	292	6.67	61672.44	3	Chr6	Group11
SmNAC48	EVM0019152.1	313	6.40	37405.16	3	Chr8	Group12
SmNAC49	EVM0019301.1	386	5.22	25275.60	4	Chr3	Group2
SmNAC50	EVM0019406.1	307	6.02	49339.63	6	Chr6	Group1
SmNAC51	EVM0019979.1	341	6.05	32567.78	3	Chr1	Group1
SmNAC52	EVM0020291.1	266	7.19	30622.40	3	Chr2	Group1
SmNAC53	EVM0020339.1	271	7.23	35553.87	3	Chr5	Group1
SmNAC54	EVM0020736.2	310	5.85	32064.98	3	Chr7	Group1
SmNAC55	EVM0020746.1	292	5.19	35623.88	4	Chr5	Group3
SmNAC56	EVM0020924.1	290	6.54	32022.66	3	Chr7	Group11
SmNAC57	EVM0021031.1	220	9.45	38023.63	3	Chr8	Group13
SmNAC58	EVM0021188.1	237	6.61	42427.36	3	Chr8	Group8
SmNAC59	EVM0021282.1	325	8.86	36225.81	3	Chr5	Group1
SmNAC60	EVM0021410.1	577	4.77	22100.61	3	Chr4	Group5
SmNAC61	EVM0021481.1	319	8.59	31374.54	3	Chr7	Group11
SmNAC62	EVM0021657.1	249	9.12	62367.35	3	Chr6	Group8
SmNAC63	EVM0021731.1	241	5.02	34851.76	1	Chr5	Group14
SmNAC64	EVM0021962.1	272	6.27	43287.60	3	Chr6	Group1
SmNAC65	EVM0021981.1	280	5.70	61488.46	3	Chr6	Group1
SmNAC66	EVM0022633.1	331	6.81	32325.15	3	Chr7	Group1
SmNAC67	EVM0022741.1	297	6.67	51152.96	3	Chr6	Group1
SmNAC68	EVM0023726.1	327	6.79	32982.62	3	Chr1	Group9
SmNAC69	EVM0023821.1	330	6.05	33719.79	3	Chr1	Group1
SmNAC70	EVM0024451.1	163	6.52	23096.39	3	Chr4	Group14
SmNAC71	EVM0024696.1	281	6.39	22413.27	3	Chr4	Group1
SmNAC72	EVM0025216.1	336	6.47	37203.93	3	Chr5	Group9
SmNAC73	EVM0025354.1	400	5.21	23621.98	4	Chr4	Group2
SmNAC74	EVM0025395.1	560	4.58	26454.96	6	Chr3	Group2
SmNAC75	EVM0025749.1	145	6.84	38620.42	2	Chr8	Group8
SmNAC76	EVM0025950.1	283	5.86	32376.21	3	Chr7	Group11
SmNAC77	EVM0025955.1	275	6.12	27122.66	3	Chr3	Group1
SmNAC78	EVM0026433.2	280	9.16	34734.17	3	Chr5	Group7
SmNAC79	EVM0026994.1	326	8.39	38406.84	3	Chr8	Group1
SmNAC80	EVM0027150.1	241	5.91	39696.27	3	Chr8	Group8
SmNAC81	EVM0027814.1	275	6.97	60681.96	2	Chr6	Group1
SmNAC82	EVM0027895.1	458	6.38	50578.20	6	Chr6	Group7

Note: pI, proteins' isoelectric point; MW, molecular weight.

SmNAC gene structure and conserved sequence analysis

The structural diagrams of *SmNAC* genes were established by the results of MEME motif analysis. We divided *SmNAC* into 14 groups, as G1-G14. Among them, motif 1, motif 2, motif 3, motif 4, motif 5, and motif 6 are the most widely distributed (Fig. 1b).

The results have shown that the motif composition across different members of each group was similar, this represents that the gene structure is conserved in a specific subfamily. In addition, each *SmNAC* has its own specificity, with motif 10 specific to *SmNAC23* and *SmNAC30* this may be related to the important function of *NAC* transcription factors. The function of most conserved motifs remains to be elucidated. Exon identification was performed on the identified *SmNAC* members to further understand the diversity among *NAC* family members in *Salvia miltiorrhiza* (Fig. 1c). All *SmNAC* genes had one to 11 protein-coding DNA sequence (CDS).

Phylogenetic tree analysis of SmNAC gene family

Phylogenetic analysis of *Arabidopsis thaliana* and *Salvia miltiorrhiza NAC* genes was carried out, in order to further understand the evolutionary relationship. A total of 261

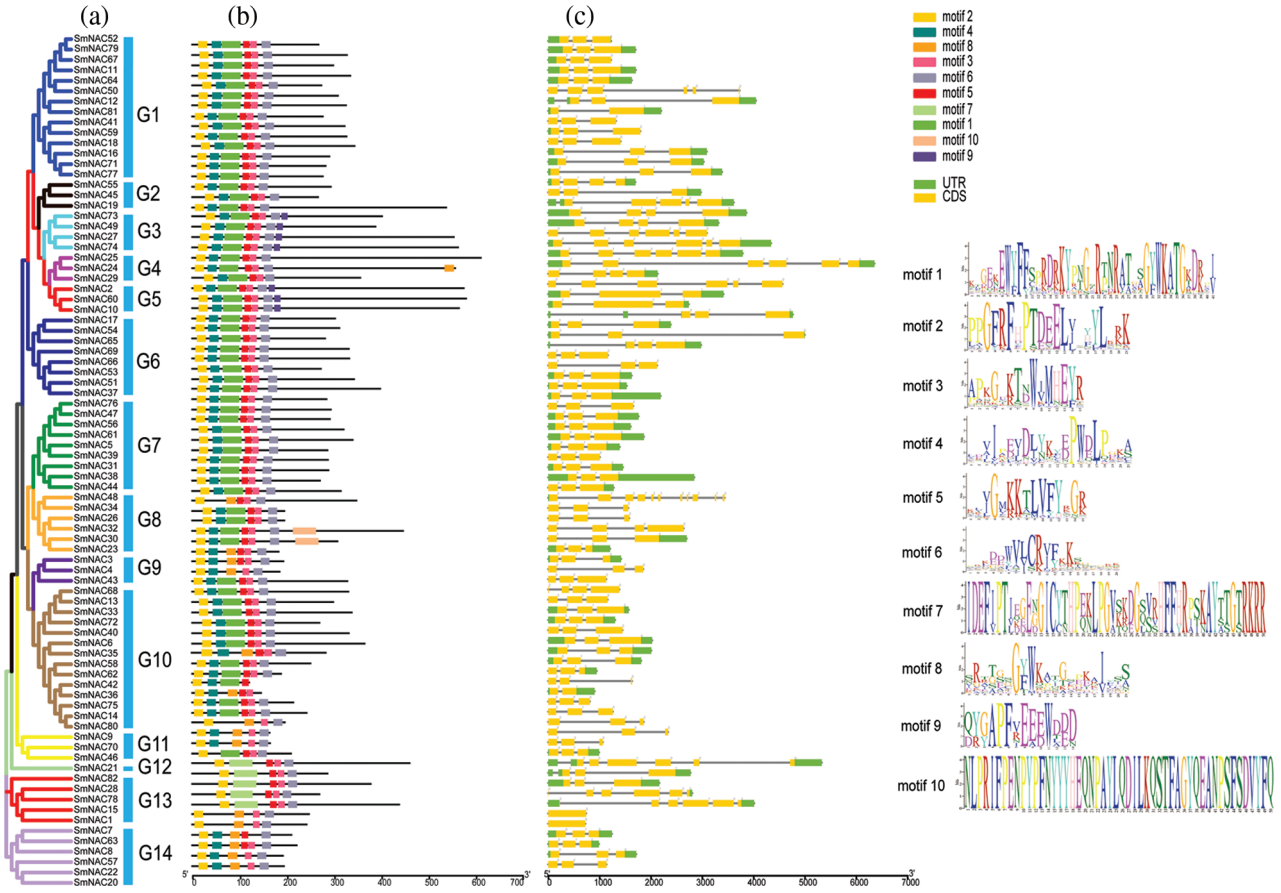

FIGURE 1. Phylogenetic relationship, conserved protein motif structure and gene structure of *SmNAC* gene in *Salvia miltiorrhiza*. (a) Phylogenetic tree was constructed using Mega 7.0 software based on the domains of *Salvia miltiorrhiza* NAC proteins, with branching details shown in different colors. (b) The motif composition of *Salvia miltiorrhiza* NAC proteins. (c) The motifs, numbers 1 to 10, are displayed in different colored boxes.

NAC's, of which 179 were from *Arabidopsis thaliana* and 82 were from *Salvia miltiorrhiza*, were used to construct the phylogenetic tree (Fig. 2). They have divided them into 14 subgroups (G1 to G14) according to their evolutionary branches. The results also showed that the distribution of *SmNAC* members are uneven among all subgroups. Group1 (G1) contained the highest (22) *SmNAC* members, followed by Group8 with 10 members, and Group6 with only one member.

Chromosomal localization and distribution of SmNAC genes

The TB tools software was used to visualize the location of the acquired genes on the chromosomes. We mapped 82 *SmNAC* genes to GWHA0SJ00000006, GWHA0SJ00000010, GWHA0SJ00000013, GWHA0SJ00000020, GWHA0SJ00000023, GWHA0SJ00000040, GWHA0SJ00000057, GWHA0SJ00000084, and GWHA0SJ00000114 were all on nine chromosomes, (Fig. 3). A total of 14 genes were located at Chr4 and Chr6, 13 genes on Chr8, 11 gens on Chr5 and Chr7, and so on (Fig. 3).

Differential expressions of SmNAC genes under salt and drought stress

We selected 44 *SmNAC* genes out of 82 genes to verify the effect of abiotic stresses on *SmNAC* genes of *Salvia miltiorrhiza*. The transcription levels of these genes were relatively high in different tissues (Fig. 4). We divided 44 *SmNAC* genes into six groups. In groups I–II, the expression patterns of 18 *SmNAC* genes were significantly modulated due to abiotic stresses in different tissues of *Salvia miltiorrhiza*. On the one hand, under NaCl stress, 10 *SmNAC* genes were Upregulated in roots and eight *SmNAC* genes were Upregulated in leaves. The results also showed that the expression of *SmNAC* genes in roots of *S. miltiorrhiae* was more upregulated than that in leaves under NaCl stress. In addition, under drought stress, 15 *SmNAC* genes were upregulated in leaves and six *SmNAC* genes were upregulated in roots.

In groups III–V, most of the *SmNAC* genes were upregulated in only roots. Among them, 22 *SmNAC* genes were up-regulated under salt stress, and 20 *SmNAC* genes were up-regulated under drought stress conditions. The results showed that in abiotic stress environment (salt or drought), *NAC* gene transcripts were mainly induced in roots. With the increase of Na^+ induced salinity and the increased drought level, the expression of some *SmNAC* genes in roots were abundantly expressed. For example, *SmNAC8* and *SmNAC13* (Fig. 5b) were significantly up-regulated under salt conditions. The genes *SmNAC24*, *SmNAC25*, and *SmNAC54* (Fig. 5d) were significantly up-regulated under drought conditions, indicating that *NAC* genes play an important role in coping with abiotic stress environment. In addition, with the increase of stress intensity, the expression of some genes was down-regulated or inhibited by high concentration of NaCl and drought stress. In this study, 18 genes were significantly

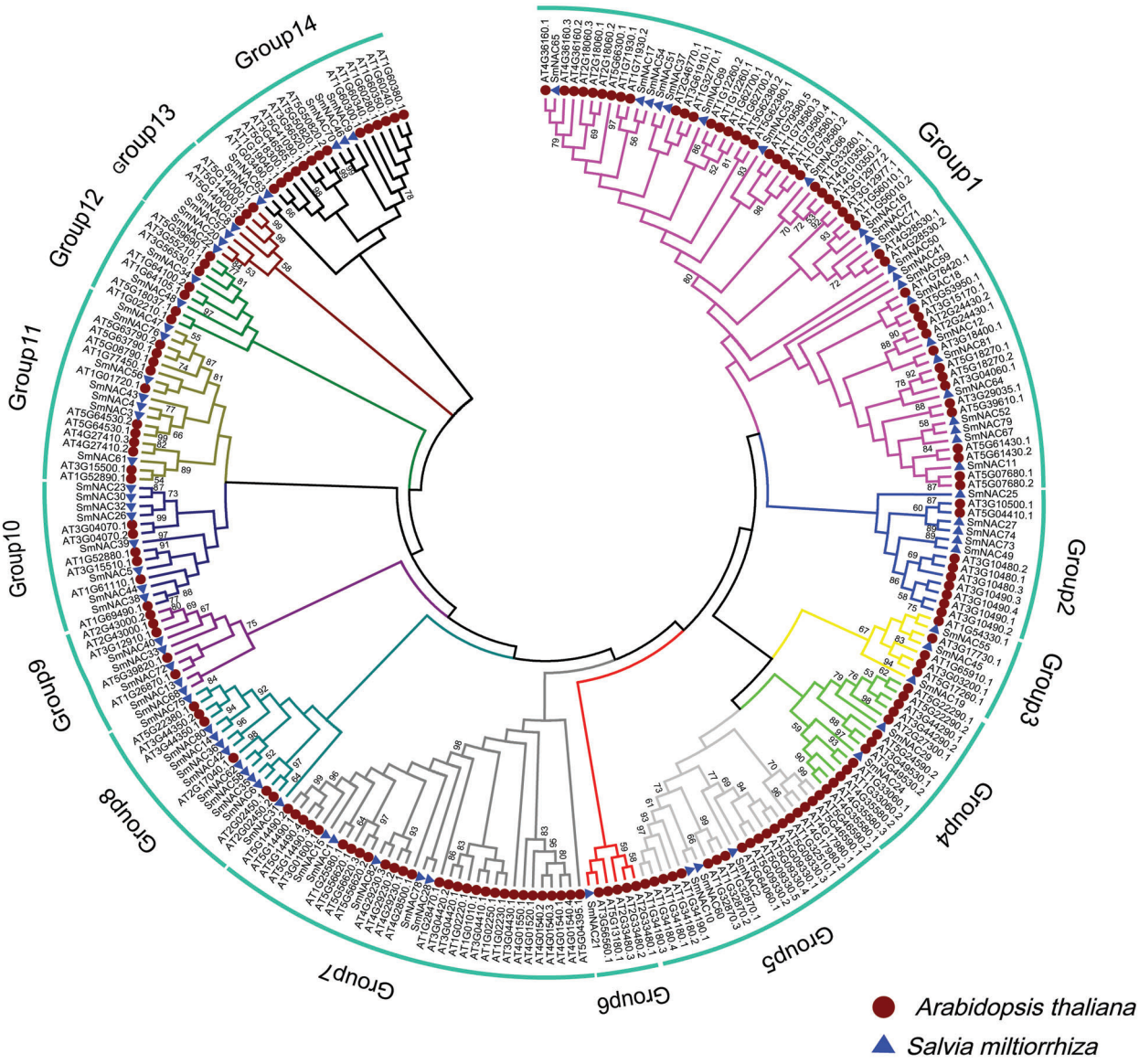

FIGURE 2. Phylogenetic analysis of *NAC* in *Arabidopsis thaliana* and *Salvia miltiorrhiza*. The tree amplified a total of 1000 Bootstrap replicates using the maximum likelihood (ML) method using MEGA 7.0. The tree divided these NAC proteins into 14 groups, named Group1 to Group14. NAC protein members of *A. thaliana* and *Salvia miltiorrhiza* are labeled with red circles and blue triangles, respectively, with different colored branches representing different group.

expressed in different tissues under abiotic stresses (salt/drought). *SmNAC22* gene was significantly up-regulated by 5.83 times in roots under 100 mM NaCl stress (Fig. 5b). Under 200 mM NaCl stress, *SmNAC8* gene was up to 13.05 times significantly up-regulated in leaves (Fig. 5a). Moreover, under 5% PEG stress, *SmNAC8* gene was up 16.65 times significantly in leaves (Fig. 5c), while under 10% PEG stress, *SmNAC8* gene was up 22.53 times significantly in roots (Fig. 5d). *SmNAC8* gene expression was gradually induced, by the increase of Na^+ concentration (Figs. 5a and 5b), while *SmNAC25*, *SmNAC54* gene expression was gradually induced by the increase of PEG concentration (Fig. 5d). In addition, with the increase of Na^+ concentration, the expression levels of *SmNAC1* in leaves and *SmNAC22* and *SmNAC34* in roots were significantly down-regulated. Similarly, with the increase of PEG concentration, *SmNAC1*, *SmNAC2*, *SmNAC19*, *SmNAC25*, *SmNAC54*, *SmNAC60* in leaves and *SmNAC35*, *SmNAC81* in roots were also significantly down-regulated (Figs. 5a and 5c). Conversely, *SmNAC57* expression in leaves was significantly down-regulated. These results indicated that *SmNAC* gene had different regulatory mechanisms when *Salvia miltiorrhiae* responded to abiotic stresses.

Discussion

In the present study, we identified and characterized 82 *SmNACs* genes in *Salvia miltiorrhiza* plants (Fig. 2). Gene replication can increase the number of genes, and it is speculated that gene replication (WGD) events occur in different lineages (Prince and Bryan, 2002). For example, four large repeating events have been found in the *Arabidopsis thaliana* genome (Vision *et al.*, 2000; Blanc *et al.*, 2003), and approximately 83–123 Mai of GWD events have been occurred in *tomatoes* (Tomato Genome Consortium, 2012).

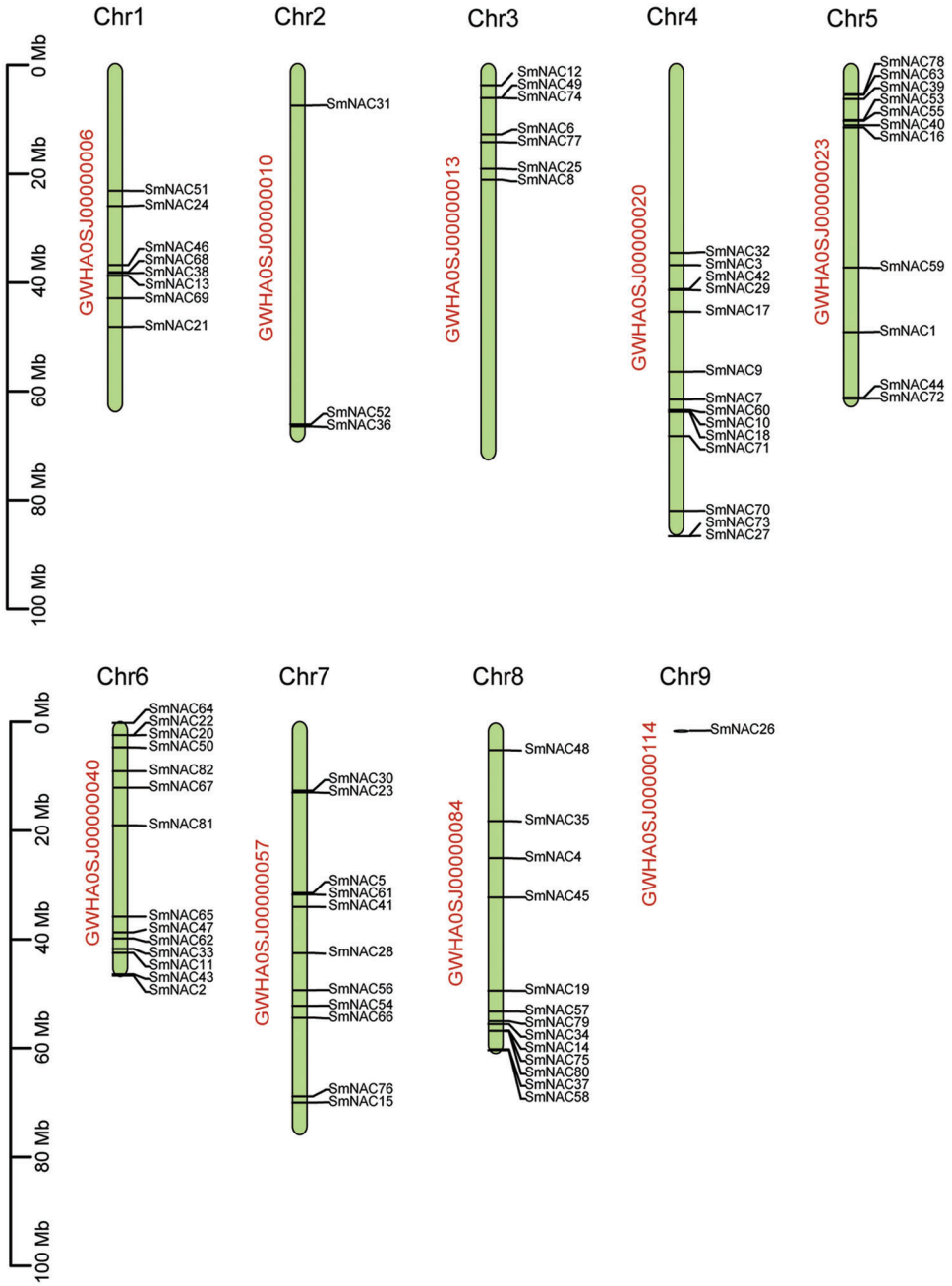

FIGURE 3. Distribution of *SmNAC* genes in chromosomes (Chrs). The vertical bars represent the chromosomes in the *Salvia miltiorrhiza* genome, with the Chr number at the top.

Therefore, the decrease in the number of *SmNAC* genes may be due to the lack of response to gene replication (WGD) events during the evolutionary process of *Salvia miltiorrhiza*.

Based on the protein motif and gene structure analysis of *Salvia miltiorrhiza*, the *SmNAC* family was divided into 14 groups. The motif and CDS analysis showed that the most closely related members of the phylogenetic tree shared a common motif composition (Fig. 1a). Despite the length, molecular weight and PI of *SmNAC* genes varied greatly, the gene structure was relatively conserved. Most *SmNAC* have similar motifs (motif 1, motif 2, motif 3, motif 4, motif 5, and motif 6). Motif 2 is common in all *SmNAC* genes, suggesting that this motif is necessary for *NAC* protein functioning. At the same time, there are also changes in genetic structure in a given cluster of genes. For example, the motif structure of *SmNAC23*, *SmNAC30*, *SmNAC34* in G8 and *SmNAC43* in G9 is different, indicating that *SmNACs* have diversity in the evolutionary process. In previous studies, a similar phenomenon was found in tomato (Jin *et al.*, 2020). In addition, we found that *SmNAC21* was isolated in a single group, and its protein function needs to be further explored. The *SmNAC* gene has three CDS (Fig. 1c), and the number of CDS ranged from 1 to 11, indicating that the *NAC* proteins had functional similarities. At the same time, phylogenetic tree analysis revealed the separation of different types of *NAC* transcription factors in *Salvia miltiorrhiza*, and we divided the *SmNAC* gene family into 14 groups (Fig. 2). The *NAC* family members of *Arabidopsis thaliana* and *Salvia miltiorrhiza* are not only homologous, but they may be

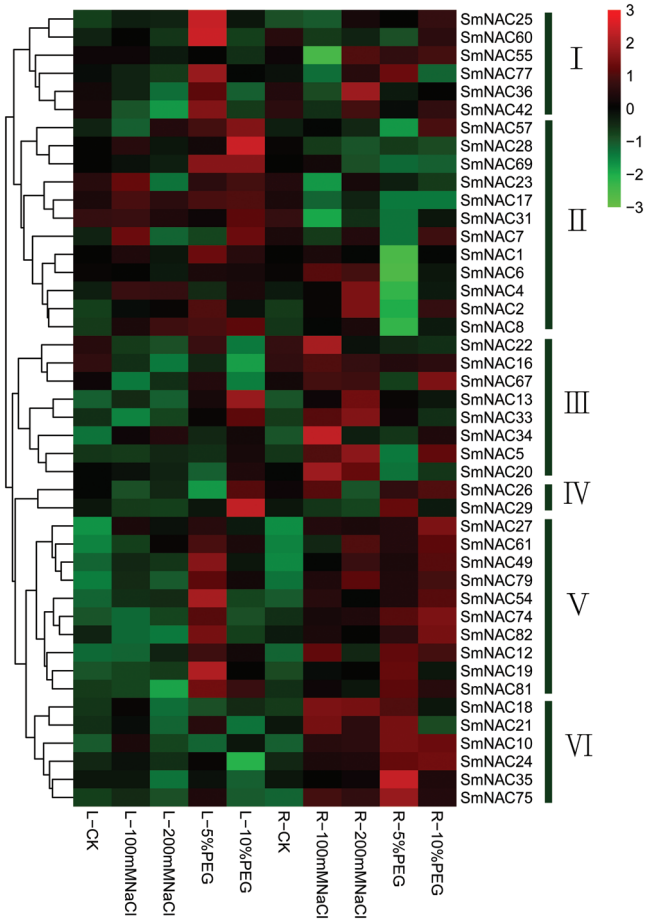

FIGURE 4. Expression patterns of 44 *SmNAC* genes in leaves and roots of *Salvia miltiorrhiza*. L/R-CK (control leaf/root), L/R-100 mM NaCl (*Salvia miltiorrhiza* leaves/roots under 100 mmol L^{-1} NaCl treatment), L/R-200 mM NaCl (*Salvia miltiorrhiza* leaves/roots under 200 mmol L^{-1} NaCl treatment), L/R-5% PEG (*Salvia miltiorrhiza* leaves/roots under 5% PEG treatment), L/R-10% PEG (*Salvia miltiorrhiza* leaves/roots under 10% PEG treatment), Colored blocks represent low/low expression (green), high/up expression (red), and no expression/no change (black).

evolved from the same ancestor. However, these genes were unevenly distributed on the chromosomes of *Salvia miltiorrhiza*, with the number of each chromosome varying from 1 to 14 (Fig. 3), indicating that there was no positive correlation between chromosome length and the number of *NAC* genes, a situation similar to that of peanut (Yuan *et al.*, 2020).

Generally, gene expression patterns provide clues to gene function. In previous studies, *AgNaAC63* and *AgNaAC47* in celery can be highly expressed in leaves under adverse conditions (heat, cold, drought and salt) (Duan *et al.*, 2020). We performed a phylogenetic analysis of *Salvia miltiorrhiza* using *Arabidopsis thaliana* to study evolutionary relationships and predict drought or salt-responsive genes. We determined the expression levels of 82 *SmNAC* genes in the leaves and root tissues of *Salvia miltiorrhiza*. As shown in Fig. 4, 44 *SmNACs* genes were detected in leaves and roots showing tissue and development-specific expression patterns. These genes may play an important role in the growth and development of *Salvia miltiorrhiza*, and their exact functions still need to be clarified in future studies. In previous studies, expressions of *ANAC55* (At3G15500) and *ANAC72* (At4G27410) in *Arabidopsis thaliana* were also induced by drought and high salinity (Tran *et al.*, 2004). Interestingly, we found that *SmNAC4* was in Group11 with *ANAC55* (At3G15500) and *ANAC72* (At4G27410). Meanwhile, *SmNAC4* was detected to be highly expressed in different tissues. We speculated that *SmNAC4*, as a key gene, plays a critical role in response to abiotic stress. In addition, *SmNAC2*, *SmNAC8* and *SmNAC25* were also significantly expressed in different tissues (Fig. 5), suggesting that they may also be involved in specific activities in the growth and development of *Salvia miltiorrhiza* under adverse stress conditions. In the future, more studies are needed to examine the specific function of the *SmNAC* family gene in *Salvia miltiorrhiza* to explore their dynamic role in plant growth and development under normal and environmentally stressed conditions.

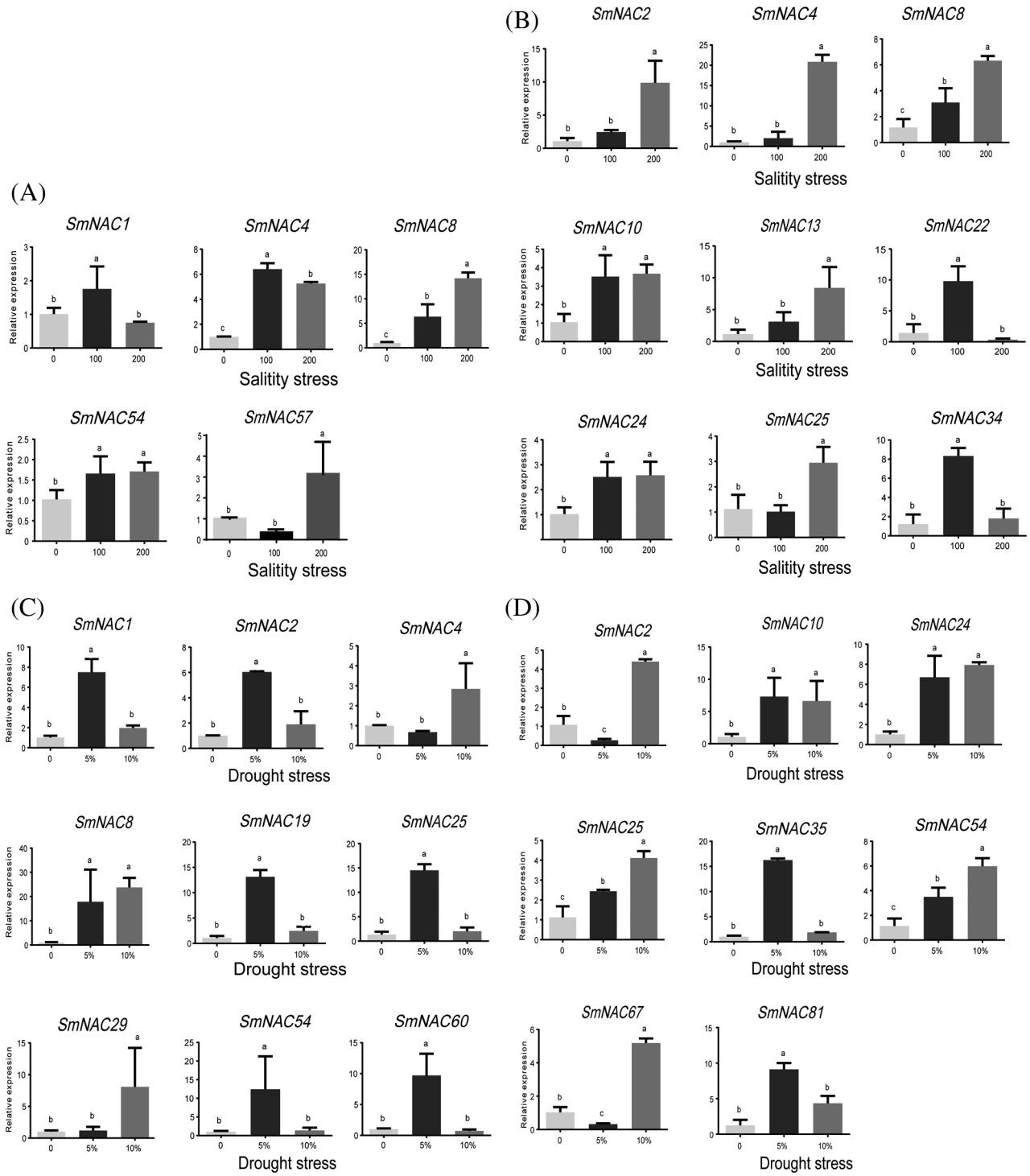

FIGURE 5. Expression pattern of *SmNAC* gene responding to NaCl and PEG in different tissues of *Salvia miltiorrhiza*. (a) Expression pattern of some genes in leaves under salt stress. (b) Expression patterns of some genes in roots of under salt stress. (c) Expression patterns of some genes in leaves under drought stress. (d) Expression patterns of some genes in roots under drought stress.

Availability of Data and Materials: All data generated or analysed during this study are included in this article.

Authors' Contribution: Ling Xu and Xin Li conceived the study. Xin Li, Juanjuan Li and Jianmin Pan conducted the experiment. Xin Li and Jianmin Pan analyzed the data. Xin Li and Juanjuan Li wrote the manuscript. Faisal Islam and Ling Xu revised the manuscript. Ling Xu, Zhuoni Hou and Zongqi Yang supported and supervised the experiment. All authors revised and approved the manuscript.

Ethics Approval: Not applicable.

Funding Statement: This work was supported by the National Natural Science Foundation of China (31871694, 31800255), the Fundamental Research Funds of Zhejiang Sci-Tech University (14042216-Y).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

Blanc G, Hokamp K, Wolfe KH (2003). A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome. *Genome Research* **13**: 137–144. DOI 13.doi 10.1101/gr.751803.

Duan AQ, Yang X, Feng K, Liu JX, Xu ZS, Xiong AS (2020). Genome-wide analysis of *NAC* transcription factors and their response to abiotic stress in celery (*Apium graveolens* L). *Computational Biology and Chemistry* **84**: 107186. DOI 10.1016/j.compbiolchem.2019.107186.

Han LM, Hua WP, Cao XY, Yan JA, Chen C, Wang ZZ (2020). Genome-wide identification and expression analysis of the superoxide dismutase (SOD) gene family in *S. miltiorrhiza*. *Gene* **742**: 144603. DOI 10.1016/j.gene.2020.144603.

Hou Z, Li Y, Su F, Chen J, Zhang X, Xu L, Yang D, Liang Z (2020). Application of ^{1}H-NMR combined with qRT-PCR technology in the exploration of rosmarinic acid biosynthesis in hair roots of *Salvia miltiorrhiza* Bunge and *Salvia castanea* f. *tomentosa* Stib. *Planta* **253**: 2. DOI 10.1007/s00425-020-03506-y.

Jia D, Jiang Z, Fu H, Chen L, Liao G, He Y, Huang C, Xu X (2021). Genome-wide identification and comprehensive analysis of *NAC* family genes involved in fruit development in kiwifruit (*Actinidia*). *BMC Plant Biology* **21**: 44. DOI 10.1186/s12870-020-02798-2.

Jin JF, Wang ZQ, He QY, Wang JY, Li PF, Xu JM, Zheng SJ, Fan W, Yang JL (2020). Genome-wide identification and expression analysis of the *NAC* transcription factor family in tomato (*Solanum lycopersicum*) during aluminum stress. *BMC Genomics* **21**: 288. DOI 10.1186/s12864-020-6689-7.

Jung I, Kim H, Moon S, Lee H, Kim B (2020). Overview of *Salvia miltiorrhiza* as a potential therapeutic agent for various diseases: An update on efficacy and mechanisms of action. *Antioxidants (Basel)* **9**: 857. DOI 10.3390/antiox9090857.

Mathew IE, Das S, Mahto A, Agarwal P (2016). Three rice NAC transcription factors heteromerize and are associated with seed size. *Front Plant Science* **7**: 1638. DOI 10.3389/fpls.2016.01638.

Nakashima K, Takasaki H, Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K (2012). *NAC* transcription factors in plant abiotic stress responses. *Biochimica et Biophysica Acta* **1819**: 97–103. DOI 10.1016/j.bbagrm.2011.10.005.

Nuruzzaman M, Manimekalai R, Sharoni AM, Satoh K, Kondoh H, Ooka H, Kikuchi S (2010). Genome-wide analysis of *NAC* transcription factor family in rice. *Gene* **465**: 30–44. DOI 10.1016/j.gene.2010.06.008.

Olsen AN, Ernst HA, Leggio LL, Skriver K (2005). *NAC* transcription factors: Structurally distinct, functionally diverse. *Trends in Plant Science* **10**: 79–87. DOI 10.1016/j.tplants.2004.12.010.

Prince VE, Bryan PF (2002). Splitting pairs: The diverging fates of duplicated genes. *Nature Reviews Genetics* **3**: 827–837. DOI 10.1038/nrg928.

Shen S, Zhang Q, Shi Y, Sun Z, Zhang Q, Hou S, Wu R, Jiang L, Zhao X, Guo Y (2019). Genome-wide analysis of the *NAC* domain transcription factor gene family in *Theobroma cacao*. *Genes* **11**: 35. DOI 10.3390/genes11010035.

Song Z, Lin C, Xing P, Fen Y, Jin H, Zhou C, Gu YQ, Wang J, Li X (2020). A high-quality reference genome sequence of *Salvia miltiorrhiza* provides insights into tanshinone synthesis in its red rhizomes. *Plant Genome* **13**: e20041. DOI 10.1002/tpg2.20041.

Sudhir K, Glen S, Michael CK, Koichiro T (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35**: 1547–1549. DOI 10.1093/molbev/msy096.

Tomato Genome Consortium (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* **485**: 635–641. DOI 10.1038/nature11119.

Tran LS, Nakashima K, Sakuma Y, Simpson SD, Fujita Y, Maruyama K, et al. (2004). Isolation and functional analysis of *Arabidopsis* stress-inducible *NAC* transcription factors that bind to a drought-responsive *cis*-element in the early responsive to dehydration stress 1 promoter. *Plant Cell* **16**: 2481–2498. DOI 10.1105/tpc.104.022699.

Vision TJ, Brown DG, Tanksley SD (2000). The origins of genomic duplications in *Arabidopsis*. *Science* **290**: 2114–2117. DOI 10.1126/science.290.5499.2114.

Wang L, Ma R, Liu C, Liu H, Zhu R et al. (2017). *Salvia miltiorrhiza*: A potential red light to the development of cardiovascular diseases. *Current Pharmaceutical Design* **23**: 1077–1097. DOI 10.2174/1381612822666161010105242.

Wu CF, Bohnert S, Thines E, Efferth T (2016). Cytotoxicity of *Salvia miltiorrhiza* against multidrug-resistant cancer cells. *The American Journal of Chinese Medicine* **44**: 871–894. DOI 10.1142/S0192415X16500488.

Yang S, Zhu H, Huang L, Zhang G, Wang L, Jiang X, Zhong Q (2021). Transcriptome-wide and expression analysis of the *NAC* gene family in pepino (*Solanum muricatum*) during drought stress. *PeerJ* **9**: e10966. DOI 10.7717/peerj.10966.

Yuan C, Li C, Lu X, Zhao X, Yan C, Wang J, Sun Q, Shan S (2020). Comprehensive genomic characterization of *NAC* transcription factor family and their response to salt and drought stress in peanut. *BMC Plant Biology* **20**: 454. DOI 10.1186/s12870-020-02678-9.

Zafar Z, Fatima S, Bhatti MF (2021). Comprehensive analyses of *NAC* transcription factor family in almond (*Prunus dulcis*) and their differential gene expression during fruit development. *Plants* **10**: 2200. DOI 10.3390/plants10102200.

Zhao HX, Zhang L, Fan X, Yang RW, Ding CB, Zhou YH (2006). Studies on chromosome numbers of *Salvia miltiorrhiza, S. flava and S. evansiana*. *Zhongguo Zhong Yao Za Zhi* **31**: 1847–1849.

Zhou Y, Li W, Xu L, Chen L (2011). In *Salvia miltiorrhiza*, phenolic acids possess protective properties against amyloid β-induced cytotoxicity, and tanshinones act as acetylcholinesterase inhibitors. *Environmental Toxicology and Pharmacology* **31**: 443–452. DOI 10.1016/j.etap.2011.02.006.

Supplementary Materials

TABLE S1

RT-qPCR primer sequence of *S. miltiorrhiza SmNAC* genes

Name	Forward (5' → 3')	Reverse (5' → 3')
Smactin	AGGAACCACCGATCCAGACA	GGTGCCCTGAGGTCCTGTT
SmNAC1	GGTGGCAGAAACAAGTGAGG	TTCTGTCCTTGACCGTCTCC
SmNAC2	CCCAAACTTCTGGGATGCAG	CAGCATGTCTTCATCGGCAA
SmNAC3	GCTGCTACCCTACCATCCAG	TCGACTCCTATTGGTTGCCA
SmNAC4	CTGGTGGTCCATTTCCTCCA	CCAGTATCCACTCCCACCAA
SmNAC5	GGTACTTCTTCAGCCCTCGT	GATCCCTTTGGGTGGCTTTC
SmNAC6	GGCGGAAATCTCTCTTTGCC	GGATAAACGTGGAGGAGGCT
SmNAC7	GGACATTGGAACATGCACGA	CTATGTCTCCGCCACCTCTC
SmNAC8	TGGAGAGGGCAAGAAGTGAG	TCCTTCAACTCACCAGGCAA
SmNAC9	GGAGATTGATGCAACCGAGT	TGCCTTCTCCAACTTGTCTCA
SmNAC10	AGATGCCAGGATGCTAAGGG	ATCGACAGGCATGGTATCGT
SmNAC11	GGTGCCGATTCCCAATCTTC	GGAAGTCGCGGTTCTTTGAC
SmNAC12	GCCAGACAACTTCCACTTCC	TTCAGCCCGGATATGGTGAA
SmNAC13	GATTTGCTGCCAGGGTTCAG	CGGTCTCTTGGGCAGTAGAA
SmNAC14	GCTTCCACAATTTGCAGGAGA	ACATCACAAGGAGAGCCAGT
SmNAC15	GCCGTAGGGTTCAGGATCAA	TTCATCGCTCCCAAGATGGT
SmNAC16	TACAGCCAAAGGGACCGAAA	TGCATGACCCAGTCTGTCTT
SmNAC17	ATCTTTGAGCTCCCGAACCT	GGCTAAGGGCATGTTTGTGT
SmNAC18	CCAAGTGGGTTATGCACGAG	ATGTGGGTCAAGCAATGCAG
SmNAC19	TCAACTCGTCAGACACCCAA	CCTCACCACACAATCTCCCT
SmNAC20	AGCTTGCAGGTGATTCTTCG	GTTGAAACGCCGGATCTCAT
SmNAC21	ATACCGCCTCACTACCAAGG	TGCTTTCTTCTTCGCGATCC
SmNAC22	TTGCCTCCTGGGTTCAGATT	ACATGTCTTGCCCTCCATCA
SmNAC23	ACTTGGAGACGACGAATGGT	ACACCTTTCTGAGGCTTTCCT
SmNAC24	ACAAGGCAGATGAGAGGCTT	GACAGTCGGCTGGATCAGTA
SmNAC25	CACCCGATGGAAAGAGGACT	AACTTCCGGAGGAGAAGCAA
SmNAC26	GACACGGGTTATTGGAAGGC	CTTCTTGCGTGCAACTCCAT
SmNAC27	CAGGGCCAAAGAATGGTGAG	CCATCTGAGCTGACAGGAGT
SmNAC28	CGAAACAAGGTGGCACAAGA	TACCTTGGAGAGCACGAACT
SmNAC29	TATCTGGGCTTGTGGATGGG	TTCCTCGTTGAGGCTGCTTA
SmNAC30	GCAAAGGAGCAAACCGATGA	AGCCAGCAAATGTGGATTGC
SmNAC31	GCTTCCACCAAACGTCATCA	AGCTCCGAATAGGCTGATCC
SmNAC32	GACACGGGTTATTGGAAGGC	ATCTGCACTCGTCTTGTTGC
SmNAC33	ACAGAGGTCTCTCACACACG	GCGGGTTGGTTGAAGAACAT
SmNAC34	GACAACAAAGCCAGCAAAGC	GCTGCACCTGTAGTTGACTG
SmNAC35	CGCTTCAACGTCGAACTCAT	TCCAGTAGCCAGATGTCGTC
SmNAC36	CATCTACCAGCTGGATCCGT	ATCACCTTGTTCTCCGACGA
SmNAC37	AGTTTGGAGGAAATGCTG	GTAGTTGGTTTGGGTGCG
SmNAC38	CGTGATCGGAAGTACCCGAA	TTGAGACGGTACTCGTGCAT
SmNAC39	CTTCACGGCTGGATGATTGG	TGAACTTGCCATTCCTTGGC
SmNAC40	GAAGAGGAGGAAGACGCAGA	GCCGTTGGACTTTGGAAGAT
SmNAC41	AGGATGAGTGGGTGATCTCG	GCCATGGAGAAACAGGACAC
SmNAC42	CATCTACGAGCTGGATCCGT	ATCACCTTGTTCTCCGACGA

(Continued)

Table S1 (continued).

Name	Forward (5' → 3')	Reverse (5' → 3')
SmNAC43	AGAGGGCGTTTCGGAGTAAA	AGTATGAACTCCATCGCCACA
SmNAC44	AAGGCTCCTGAGCAAAGAGT	AGGGCGCATTGTATGCATTT
SmNAC45	CCCACAAGGGATAAGGACTGA	CCCACAAGGGATAAGGACTGA
SmNAC46	TCACGACGGATGTGAGGAAA	CCATGGACACCGTGAAACAA
SmNAC47	TGAGTTTGTGTGCGACAAGG	CGGAGAAGAGAGGGAACTGG
SmNAC48	GGAAGCATCGGAAGGGATTC	GGTGATCACGATGAGCTGTG
SmNAC49	CACGAGTTGAATGCACGGAT	ATTTGCATGGCTCGAGTTGG
SmNAC50	TGCCTACAACCTCTCCCATC	ACCAACTCAGCCCATCTTCA
SmNAC51	ACGGCGTTCTGGATCAACTA	GTTCCCATTTCGTGCGTGAT
SmNAC52	GCCTTGGAGAGCGAAATTAGG	GGGCTCTCCCTTGGTAGAAA
SmNAC53	CGCTCCTCCCAATGAAATCC	TTCGGGACTGAGTAGTTGGG
SmNAC54	CTCAGGAAGAAGGGTGGGTT	GTAGCTAGGCAGGTTGGGAT
SmNAC55	GATTCCGGTTCCATCCAACG	GGCCCGTAGAAGTACCACTC
SmNAC56	CTGGATGATTGGGTGCTGTG	GAAGTCGTCGTACACCATCG
SmNAC57	AGGGAGGTGAATCATGGCAA	TGTGAGAGGGCAATGCAGTA
SmNAC58	GCATCATCCTACCGAGTTGC	GTCTTCTTCACTCCCAACGC
SmNAC59	CAGGACACCTCGTCGTCTTA	AGTGCCAGACATTTGATCGC
SmNAC60	CGCACCTGAGGTTTCTTCTG	AACAGTTCAAGCTCGTGCAG
SmNAC61	TTTCCCGCTGCAGATCATTG	GCTTTCCAGTAACCCGAACC
SmNAC62	AAATTCCGCGGCGATGTTAT	TCGCCTTCCAGAATCCCTTC
SmNAC63	AGCTTGCAGTGGGTTTCAAG	GCGCAGATAGCCAAGATTCA
SmNAC64	AGTAGTTGAGCTGCCTCCTG	CTCGCTGGCAGAAGAAGAAC
SmNAC65	CGATGAAGAGCTGGTCGGAT	CCCGCTCGGGTACTTCTTAT
SmNAC66	GGTTCTCGGACAAACAGAGC	AATCCTCGAGCCTGTACTCG
SmNAC67	GACCGCAACTTCTCCACAAA	GCCTTCCAGTATCCAGCACT
SmNAC68	CTGCCAGGGTTCAGATTCCA	CGGTCTCTTGGGCAGTAGAA
SmNAC69	GCCATGGGATCTCCAAGAGT	TCCTCATCCCGATCAAGTCG
SmNAC70	CTACACCTACGAGCCCTGTC	CCGATGACTTCTCCGGCTAT
SmNAC71	TCACTCATTGCCACCAGGAT	TCCTTTGAGTTCAGCTTCGC
SmNAC72	CCGATTGGATGATGCACGAG	CGTTGCAGAAGCTCATCGTA
SmNAC73	TGTGGGAAGCCATTCAGGTT	CCAGTCGCTTTCCAATACCC
SmNAC74	CATTTGTGCTTTGCCGTGTC	TGAGCCAAGGATCTGCTCAA
SmNAC75	ATCCGACGGAGGAAGAACTG	CTTCCCGCTGGTTGAAGAAG
SmNAC76	GAGGAGCTCGTGGTTCACTA	GTGAACCGTTCGGGTACTTG
SmNAC77	ATTTCGACGATCCCTCGGTT	ATGAGAGGCCGAGCAGTAAG
SmNAC78	CACGAAAGCGACGTAAGGTG	TTGGTGCATCACCCAGTTTG
SmNAC79	CCATGGAGAGCGAAATTGGG	CACGGTAGAAGACGAGGGTT
SmNAC80	GAAGTGAGCTTGTGCCGAAT	TCAGTTCCCACAGCTGGATT
SmNAC81	ACTTCTTCAGCCTGCGAGAT	GTGCATGACCCAGTTGGTTT
SmNAC82	CCGGAGTAACAAGAGACGGA	TCTTGCACCCTTTCTGCTTG