Annals of Forest Science (2020) 77: 82
https://doi.org/10.1007/s13595-020-00983-9

RESEARCH PAPER

Genome-wide analysis of evolution and expression profiles of NAC transcription factor gene family in *Juglans regia* L.

Hanif Khan[1] · Feng Yan[1] · Yujie Yan[1] · Pengpeng Chen[1] · Ruimin Xi[1] · Irfan Ullah[1] · Xiaobang Peng[2] · Xiang Luo[3] · Ming Yue[4] · Peng Zhao[1]

Received: 28 December 2019 / Accepted: 14 July 2020 / Published online: 17 August 2020

Abstract

• ***Key message*** ***NAC*** **transcription factors may play important roles in the biological processes in Persian walnut. A total of 102** ***JrNAC*****s were identified in Persian walnut. The conserved domains, transcriptome profile, expression analysis, and interaction network suggest that** ***JrNAC1-4*** **plays potential roles in Persian walnut flowering development.**

• ***Context*** *NACs* are plant-specific transcription factors that participate in various plant developmental processes such as flowering, plant growth regulation, and development. We identified and analyzed the evolution and expression profiles of *NAC* genes in *Juglans regia*.

• ***Aims*** The main objectives were to identify the *NAC* transcription factors and verify the expression levels in different tissues and female flowers in developmental stages in Persian walnut.

• ***Methods*** We identified *NAC* transcription factors in *J. regia* based on the genome-wide analysis. We analyzed the phylogenetic relationships, conserved domain, chromosome location, gene structure, and gene collinearity of *JrNACs*. We also verified the *JrNAC* expression levels based on transcriptome analysis and qRT-PCR.

• ***Results*** We identified 102 *NAC* genes in *J. regia* and divided them into ten subfamilies. A total of 30 pairs of *JrNAC* genes were expanded by whole-genome duplications (WGDs) and one pair of genes (*JrNAC2-10* and *JrNAC9-8*) as a tandem duplication in Persian walnut. Collinearity analysis results indicate that a large number of syntenic relationship events existed between *J. regia* and *Populus trichocarpa*. We found that *JrNAC1-4* and *JrNAC2-6* were expressed significantly higher in female flowers based on both transcriptome and qPCR analysis. We further identified that *JrNAC2-9* and *JrNAC9-6* were highly expressed at the end period of flowering stages.

Hanif Khan and Feng Yan contributed equally to this work.

Handling Editor: Ricardo Alia

Contribution of the co-authors Conceptualization: Peng Zhao and Ming Yue
Methodology: Hanif Khan and Feng Yan
Software: Hanif Khan and Feng Yan
Validation: Yujie Yan, Pengpeng Chen, and Ruimin Xi
Formal Analysis: Hanif Khan and Feng Yan
Investigation: Hanif Khan, Feng Yan, and Xiaobang Peng
Resources: Hanif Khan, Feng Yan, and Ming Yue
Data curation: Peng Zhao, Feng Yan, and Hanif Khan
Writing (original draft): Hanif Khan, Feng Yan, and Peng Zhao
Writing (review and editing): Hanif Khan, Feng Yan, Irfan Ullah, Xiaobang Peng, Xiang Luo, Ming Yue, and Peng Zhao
Visualization: Ming Yue and Peng Zhao
Supervision: Peng Zhao
Project administration: Peng Zhao and Ming Yue
Funding acquisition: Ming Yue and Peng Zhao

✉ Ming Yue
mingyue@nwu.edu.cn

✉ Peng Zhao
pengzhao@nwu.edu.cn

Extended author information available on the last page of the article

• ***Conclusion*** A total of 102 *JrNAC*s were identified in the Persian walnut genome. These genes were conserved in plants for collinearity analysis, which was performed within the genome and other genomes (*P. trichocarpa*, *Olea europaea*, and *Quercus robur*). A total of 24 *NAC* transcription factors were highly expressed in female and male flowers, and these transcription factors play a role in *J. regia* flowering.

Keywords *Juglans regia* · *NAC* transcription factor · Expression profile · q-PCR · Flowering

Abbreviations

NAC	*NAM*, *ATAF*, and *CUC*
qRT-PCR	Quantitative real-time PCR
NAM	No apical meristem
TAR	The non-conserved transcriptional activation region
WGD	Whole-genome duplication
TF	Transcription factor
Pfam	Protein family
GO	Gene ontology
NF-Y	Nuclear transcription factor Y
HMM	Hidden Markov model
NJ	Neighbor-joining
MEGA	Molecular Evolutionary Genetics Analysis
SMART	Simple modular architecture research tool
CDD	Conserved domain database
CDS	Coding sequences
MEME	Multiple motif alignment for motif elicitation
KEGG	Kyoto encyclopedia of genes and genomes
CTAB	Cetyltrimethylammonium bromide
BLAST	Basic Local Alignment Search Tool
RNA-seq	Ribose nucleic acid sequencing

1 Introduction

The *NAC* (NAM, ATAF1/2, and CUC) transcription factor (TF) family contains about ~ 150 amino acid residues and is the most important and largest TF family in plants that contains a conserved *NAM* domain at their N-terminus (Souer et al. 1996; Tran et al. 2010; Zhu et al. 2012). *NAM* was first identified in *petunia* and determines the position of the primordium and shoot apical meristem (Souer et al. 1996; Duval et al. 2002; Ooka et al. 2003). The *NAC* genes at their C-terminal end also contain a-helical transmembrane motifs (TMs) (Puranik et al. 2012). The C-terminal domain contains a non-conserved transcriptional activation region (TAR) in plants and acts as a transcription repressor or activator, sometimes displaying protein-binding activity. In addition, an α-helical transmembrane motif (named NTL) is present in some *NACs* that are required for plasma membrane or endoplasmic reticulum membrane anchoring the C-terminal region (Ooka et al. 2003; Tran et al. 2004). To date, *NAC* transcription factors in many plants have been detected, such as in the model plant *Arabidopsis*, where a total of 105 *NAC* genes have been identified (Ooka et al. 2003); in important annual crops *Brassica pekinensis*, *Glycine max*, *Oryza sativa*, *Cajanus cajan*, *Setaria italica*, *Medicago truncatula*, and *Triticum turgidum* (Nuruzzaman et al. 2010; Le et al. 2011; Puranik et al. 2013; Liu et al. 2014; Ling et al. 2017; Saidi et al. 2017); and perennial woody plants including *Populus trichocarpa*, *Vitis vinifera*, *Prunus mume*, *Morus notabilis*, and *Malus domestica* (Hu et al. 2010; Su et al. 2013; Satheesh et al. 2014; Baranwal and Khurana 2016; Zhuo et al. 2018).

NAC genes are involved in many biological processes, including stress responses (Wang et al. 2013; Satheesh et al. 2014), apical meristem development (Souer et al. 1996; Wang et al. 2009), hormone signaling (Ooka et al. 2003), leaf senescence (Guo and Gan 2006), fruit ripening (Duval et al. 2002; Shan et al. 2012), and flower formation (Souer et al. 1996; Sablowski and Meyerowitz 1998; Liu et al. 2009). Expression analysis showed that the *NAC* gene family is involved in both reproductive and vegetative tissues (Hu et al. 2003; Hennig et al. 2004). In woody plants, for example, a microarray transcriptomic analysis showed six of these genes to be expressed during the development and ripening of the *Fragaria x ananassa* fruit (Moyano et al. 2018), *P. trichocarpa NAC* genes have putative functional roles in wood-forming and secondary cell wall biosynthesis (Hu et al. 2010), and the grapevine *NAC* genes play a potential role in response to stress (Wang et al. 2013). *MdNAC1* overexpressing apple plants maintained a higher photosynthetic rate under drought conditions and accumulated lower levels of reactive oxygen species under drought conditions (Su et al. 2013). Sixteen *PmNAC*s (*Prunus mume*) exhibited downregulation during flower bud opening in apricot (Zhuo et al. 2018). For the *NAC* in flowering development, *SlNAM2* participates in the establishment of tomato flower whorl and sepal boundaries (Han et al. 2012). There are some previous studies on woody plants (Hu et al. 2010; Su et al. 2013; Wang et al. 2013; Moyano et al. 2018; Zhuo et al. 2018), including the plant development and flowering process; however, there are no reports of *NAC* genes in Persian walnut (*Juglans regia*) that focus on the flowering and development.

Persian walnut is a diploid ($2n = 32$), large, wind-pollinated, monoecious, dichogamous, enduring, perennial tree and is the most monetarily vital nut tree on earth belonging to the family Juglandaceae (Han et al. 2016; Martínez-García et al. 2016; Feng et al. 2018). It has been an important tree species

since ancient times, valued for both wood and nuts (Feng et al. 2018; Zhao et al. 2018; Yan et al. 2019a). As we know, the flowering development is important for the Persian walnut nuts, and the *NAC* gene is involved in flower formation (Souer et al. 1996; Sablowski and Meyerowitz 1998; Liu et al. 2009). In this study, to better understand the potential role and characteristics of *NAC* transcription factors in Persian walnut flowering, we performed phylogenetic analysis and analyzed the tandem and segmental duplications and intron–exon structures of *JrNAC* genes. To better comprehend whether the *NAC* gene might play an important role in flowering, we analyzed the transcriptome expression level in reproductive and vegetative tissues and carried out qRT-PCR analysis for three genes (*JrNAC1-4*, *JrNAC2-6*, and *JrNAC13-5*) in male and female flowers and leaves. This study provides the first genome-wide analysis of the Persian walnut *NAC* transcription factor family, and these findings will be useful for understanding the putative functions of Persian walnut *NAC* genes.

2 Materials and methods

2.1 Identification of *NAC* transcription factors in *J. regia*

The Persian walnut whole protein sequence was downloaded from National Center for Biotechnology Information (NCBI) (Martínez-García et al. 2016). Members of the *Arabidopsis NAC* gene family were downloaded from The Arabidopsis Information Resource (TAIR) website (Garcia-Hernandez et al. 2002). To search against Persian walnut protein sequences, we used *Arabidopsis NAC* protein sequences as a query using a local alignment search tool Basic Local Alignment Search Tool (BLAST), considering those with an E value less than 1×10^{-10}. We implemented a profile hidden Markov model (HMM) in HMMER v.3.2.1 for the window (Prakash et al. 2017) with default parameters to search for *NAC* proteins and *NAC* domains in the protein family (Pfam) database (El-Gebali et al. 2018).

2.2 Phylogenetic, chromosome location, domain analysis, motif, and gene structure analysis of *JrNAC* transcription factors

We constructed a neighbor-joining (NJ) tree of 102 *JrNAC* transcription factors using MEGA v.7.0 software (Kumar et al. 2008; Yan et al. 2019b) with the pairwise deletion of 1000 bootstraps and a Poisson model (Lescot et al. 2002) by using the Pfam webserver (El-Gebali et al. 2018) to search for the presence of potential domains. The simple modular architecture research tool (SMART) program (Schultz et al. 2000) also detected the same domains obtained from Pfam with an E value cutoff of 1.0 to validate the final result. The chromosomal locations of *NAC* transcription factors were searched against *J. regia* whole-genome sequence using BLASTN. A conserved domain database (CDD) search was conducted in NCBI (Marchler-Bauer et al. 2016). The whole coding sequence (CDS) database was downloaded (https://treegenesdb.org/FTP/Genomes/Jure/v1.4/annotation/). The exon and intron structures were displayed using the online gene structure display server (Hu et al. 2014). The genome browser was used to search for related Persian walnut gene sequences. The motif identification used the MEME program with default parameters, the maximum number of motifs (20), and the optimum motif width (30–50) (Bailey et al. 2015).

2.3 Synteny analysis and calculating K_a, K_s, and K_a/K_s values of duplicated gene pairs

Identification of potential pairs of homologous genes across multiple genomes ($E < 1 \times 10^{-5}$, top 5 matches) was performed using BLASTP. We used homologous gene pairs to identify syntenic chains through MCScanX (Wang et al. 2012). We detected duplicate gene pairs by using MCScanX, which included whole-genome duplication (WGD), tandem duplication, segmental, and other types of gene pairs. To evaluate the type of *NAC* gene selected, we used the ratio of non-synonymous substitutions to synonymous substitutions (K_a/K_s) using DnaSP software (Librado and Rozas 2009).

2.4 Cis-element analysis and GO annotation

To conduct cis-element analysis, 1500 bp upstream of the *NAC* genes of each species were analyzed using PlantCARE (plant cis-acting regulatory element) with default parameters (Lescot et al. 2002). Blast2GO v2.5 with a cutoff E value of 1×10^{-6} was used to performed gene ontology (GO) annotations (Conesa et al. 2005). First, we performed BLASTP analysis with an E value of 1e-05 using the *JrNAC* protein sequence. The analysis was then carried out with the GO mapping. After that, the GO annotation program was used to obtain the GO annotation of the *JrNAC* members. Finally, the GO enrichment analysis was carried out through the online GO enrichment program on the OmicShare website (https://www.omicshare.com/ tools/Home/Sof/gogsea) (Table 1).

2.5 Plant materials, treatments, and collections

In this study, the first opening of female flowers occurred on 10 April, 15 April, and 22 April, and full opening of female flowers occurred specifically on 15 April and 22 April, when the stigma was not fully developed, and 1 May was the end date. A total of 26 samples were collected from Persian walnut in this study, including 15 female flowers at different stages, 3

Table 1 The synonymous (K_s) and non-synonymous (K_a) substitution rates for each gene pair and the estimated time of the tandem and segmental replication events of *NAC* genes

Gene 1	Gene 2	K_a	K_s	K_a/K_s	Negative selection
JrNAC9-8	*Jure_20254.t1*	0.283565708	0.364839514	0.777234092	Yes
JrNAC13-5	*Jure_05048.t1*	0.129975351	0.236269542	0.55011471	Yes
JrNAC1-5	*Jure_13612.t1*	0.259477046	0.538664133	0.481704702	Yes
JrNAC10-5	*Jure_17221.t1*	0.259477046	0.538664133	0.481704702	Yes
JrNAC4-2	*Jure_19897.t1*	0.165094392	0.377652223	0.437159857	Yes
JrNAC2-11	*Jure_20271.t1*	0.165094392	0.377652223	0.437159857	Yes
JrNAC9-7	*Jure_20272.t1*	0.22262165	0.516172086	0.43129347	Yes
JrNAC10-8	*Jure_28343.t1*	0.22262165	0.516172086	0.43129347	Yes
JrNAC13-5	*Jure_05048.t1*	0.134350041	0.329749428	0.407430703	Yes

male flowers, 3 leaves, and 3 hulls. The female flowers were collected on 23 March, 1 April, 8 April, 16 April, and 23 April as 3 replicates (Table 2), and the male flowers were collected on 10 April, 11 April, and 2 May. The leaves were collected on 23 April as 3 replicates, and the hulls were downloaded from https://treegenesdb.org/FTP/Genomes/Jure.v1.0/transcriptome/rawreads/ (Martínez-García et al. 2016). After harvesting, the pericarp was immediately dissected, and the flesh was frozen in liquid nitrogen and stored at − 80 °C. The three leaves downloaded from the website and the other 23 tissues' total RNA were isolated using RNA-prep Pre-Plant Kit (Tiangen, Beijing, China) (Zhu et al. 2011). Finally, 23 libraries were constructed and sequenced on an Illumina HiSeq 2500 platform. Analysis of differential gene expression (DESeq) was carried out using the package DESeq R v.1.1.1. Genes found by DESeq with adjusted P values > 0.05 were allocated as differentially expressed (Khan et al. 2020).

2.6 Quantitative real-time PCR

To verify the expression pattern, we used the tissues of female flowers, male flowers, and leaves (for details, see Table 2). Upon dilution to 1:10 with sterile water, the synthetic cDNA was used as the qRT-PCR template. iQ™ SYBR® Green Supermix was used to performed qRT-PCR (Cat. 170-8880AP; Bio-Rad). PCR was conducted on a Light Cycler 480 Real-Time PCR system (Roche Diagnostics, Laval, QC, Canada). For internal control of gene, 18S rRNA was used (Xu et al. 2012). Details of primer information are listed in Table 3. Before the experiment, primer specificities and corresponding melting curves were verified. Each experiment was conducted in triplicate.

2.7 Interaction network of *JrNAC* proteins

Persian walnut *NAC* protein matched a homologous *Arabidopsis NAC* protein in the BLASTP program with an E value of 1e-05 (Aoki et al. 2007, Camacho et al. 2009). Of the *Arabidopsis NAC* proteins that represent the walnut *NAC* proteins, 102 were uploaded to the STRING website to predict protein interactions (https://string-db.org/) using the input proteins of *J. regia* and six predicted input proteins (Szklarczyk et al. 2016).

3 Results

3.1 Identification, phylogenetic relationship, and chromosome location of *NAC* transcription factor family in *J. regia*

We identified a total of 102 *NAC* genes based on the *J. regia* whole reference genome (Fig. 1; Khan et al. 2020). The neighbor-joining (NJ) phylogenetic tree showed that the *JrNAC* genes are divided into ten subfamilies in Persian walnut (Fig. 1). In the phylogenetic tree, subfamilies IV and VII with 16 *NAC* family members were the largest clades. III and X, with five *NAC* family members, were the smallest clades. Moreover, the numbers of subfamilies I, II, V, VI, VIII, and IX were 15, 15, 10, 6, 7, and 6, respectively. The *JrNAC* genes were then renamed according to their location on the chromosome (Khan et al. 2020). A map of the *NAC* genes' physical positions was created based on the Persian walnut genome physical location information (Khan et al. 2020). Our results show that the *NAC* genes were distributed unevenly on 16 chromosomes of Persian walnut (Khan et al. 2020). A maximum number of 14 *JrNAC* genes were present on chromosome 10, followed by 13 on chromosomes 1, representing 27/102 (26.5%) of the total *JrNAC* genes, followed by 11 each on chromosomes 2 and 11; 2 each on chromosomes 8, 11, and 12; and a minimum of 1 *JrNAC* on chromosome 5 (Khan et al. 2020).

Table 2 A total of 26 samples of common walnut used for expression profiling in this study

Sample name	Tissue	Date	Source
F1-1	Female flower	23 March 2019	In this study
F1-2	Female flower	23 March 2019	In this study
F1-3	Female flower	23 March 2019	In this study
F2-1	Female flower	1 April 2019	In this study
F2-2	Female flower	1 April 2019	In this study
F2-3	Female flower	1 April 2019	In this study
F3-2	Female flower	8 April 2019	In this study
F3-2	Female flower	8 April 2019	In this study
F3-3	Female flower	8 April 2019	In this study
F4-1	Female flower	16 April 2019	In this study
F4-2	Female flower	16 April 2019	In this study
F4-3	Female flower	16 April 2019	In this study
F5-1	Female flower	23 April 2019	In this study
F5-2	Female flower	23 April 2019	In this study
F5-3	Female flower	23 April 2019	In this study
M1-1	Male flower	10 April 2019	Yan et al. (2019a)
M1-2	Male flower	11 April 2019	Yan et al. (2019a)
M1-3	Male flower	2 May 2019	Yan et al. (2019a)
L1-1	Leaves	23 April 2019	In this study
L1-2	Leaves	23 April 2019	In this study
L1-3	Leaves	23 April 2019	In this study
H1-1	Hull	10 May 2019	Martínez-García. (2016)
H1-2	Hull	10 May 2019	Martínez-García. (2016)
H1-3	Hull	10 May 2019	Martínez-García. (2016)

3.2 Conserved motifs and gene structure analysis

A total of 20 different motifs were detected (Fig. 2 a and b). Ninety-eight of 102 *NAC* genes contained at least four main motifs (motifs 1, 2, 3, and 5) (Fig. 2b). Motifs 11, 12, 17, and 19 were present in 7 *JrNACs* (*JrNAC13-3, JrNAC16-2*, *JrNAC16-3*, *JrNAC12-2*, *JrNAC7-3*, *JrNAC2-2*, and *JrNAC9-3*) (Fig. 2b). Structural analysis of exons–introns indicates that the number of exons varies from 1 (on *JrNAC3-7*) to 22 (on *JrNAC1-6*) (Fig. 2c). Moreover, two exons were found on one *NAC*, three exons on 62 *NAC* genes, four exons on 12 *NAC* genes, and five exons on six *NACs* (Fig. 2c). The results show that genes on the same branch might show similar organization of exons and introns (Fig. 2c).

3.3 Paralogous *NACs*, gene duplication, and synteny analysis of *JrNAC* genes

We identified nine pairs of paralogous *NAC* genes in Persian walnut (Fig. 3a and Table 1). The ratio of K_a/K_s of nine pairs of paralogous *JrNAC* gene pairs was less than 1 based on the synonymous (K_s) and non-synonymous (K_a) estimation (Table 1), indicating that these genes are under negative selection. The MCScanX analysis showed that a total of 30 pairs of genes undergo whole-genome duplications (WGDs), while tandem duplication was observed for one pair of genes (*JrNAC2-10* and *JrNAC9-8*), and 60 genes were dispersed (Fig. 3b; Khan et al. 2020).

Table 3 Primers used for quantitative real-time PCR

Protein ID	Gene	Forward primer (5′-3′)	Reverse primer (5′-3′)
Jure_18043.t1	*JrNAC13-5*	GGGATCCAGCATCTCCATCG	ATTCCTCTCTGCACCACAGC
Jure_17198.t1	*JrNAC1-4*	AAGATGGCAAACTC GCCAGA	ATGTGTAATTGGTC GCCGGT
Jure_13238.t1	*JrNAC2-6*	TTATTGGGTGCTGGCCTTGT	ATCATCTCCGCCACTATCGC

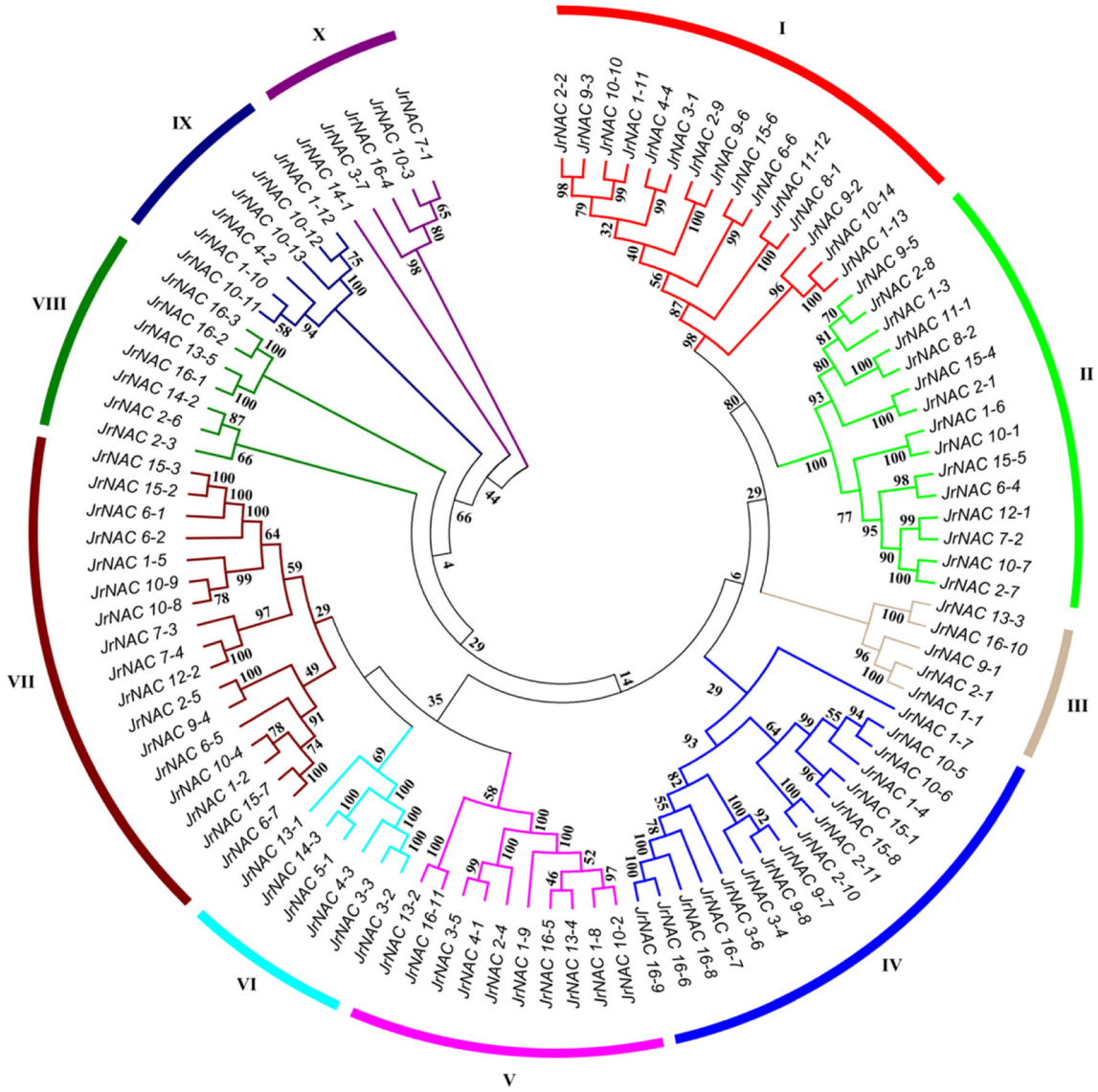
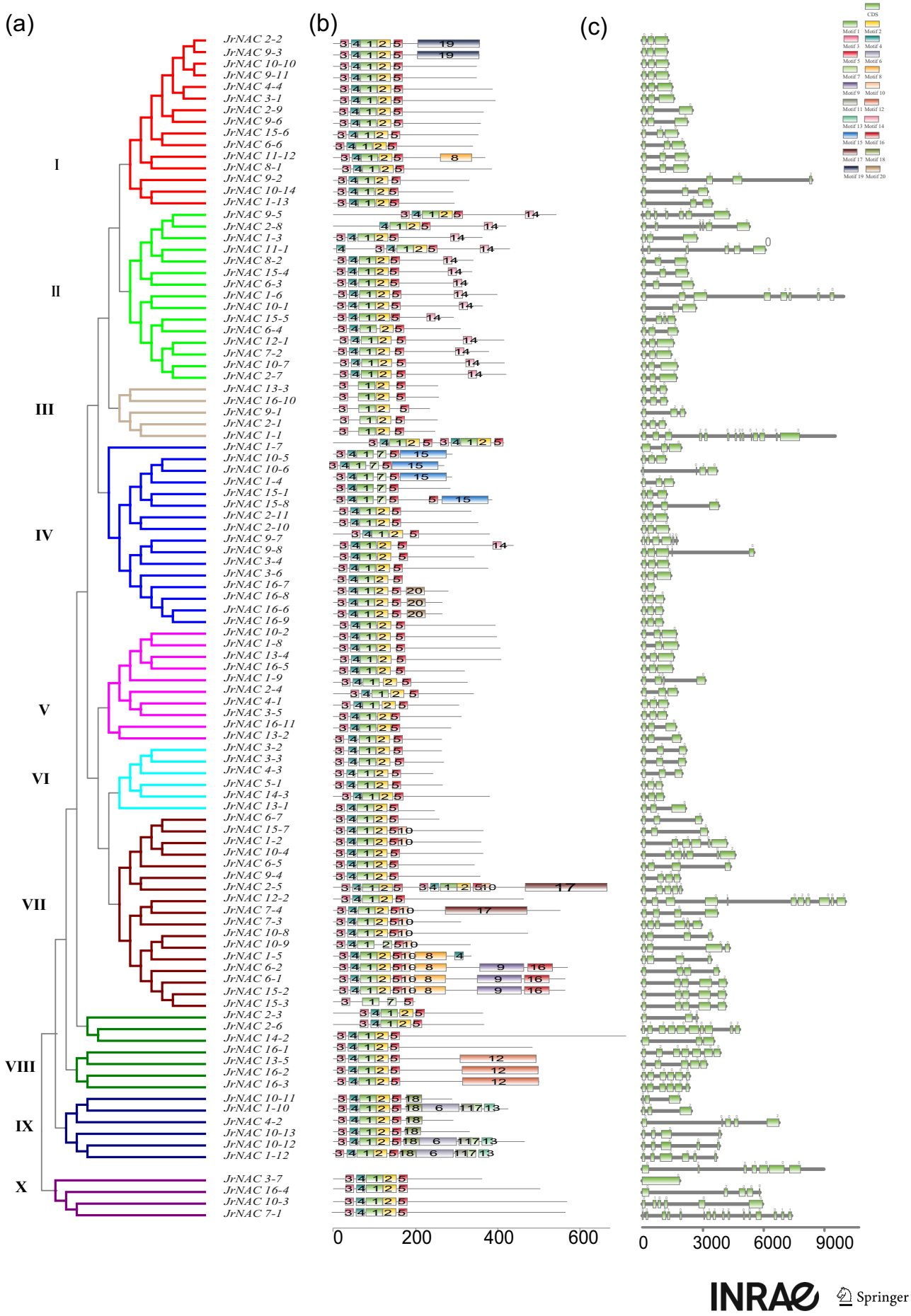
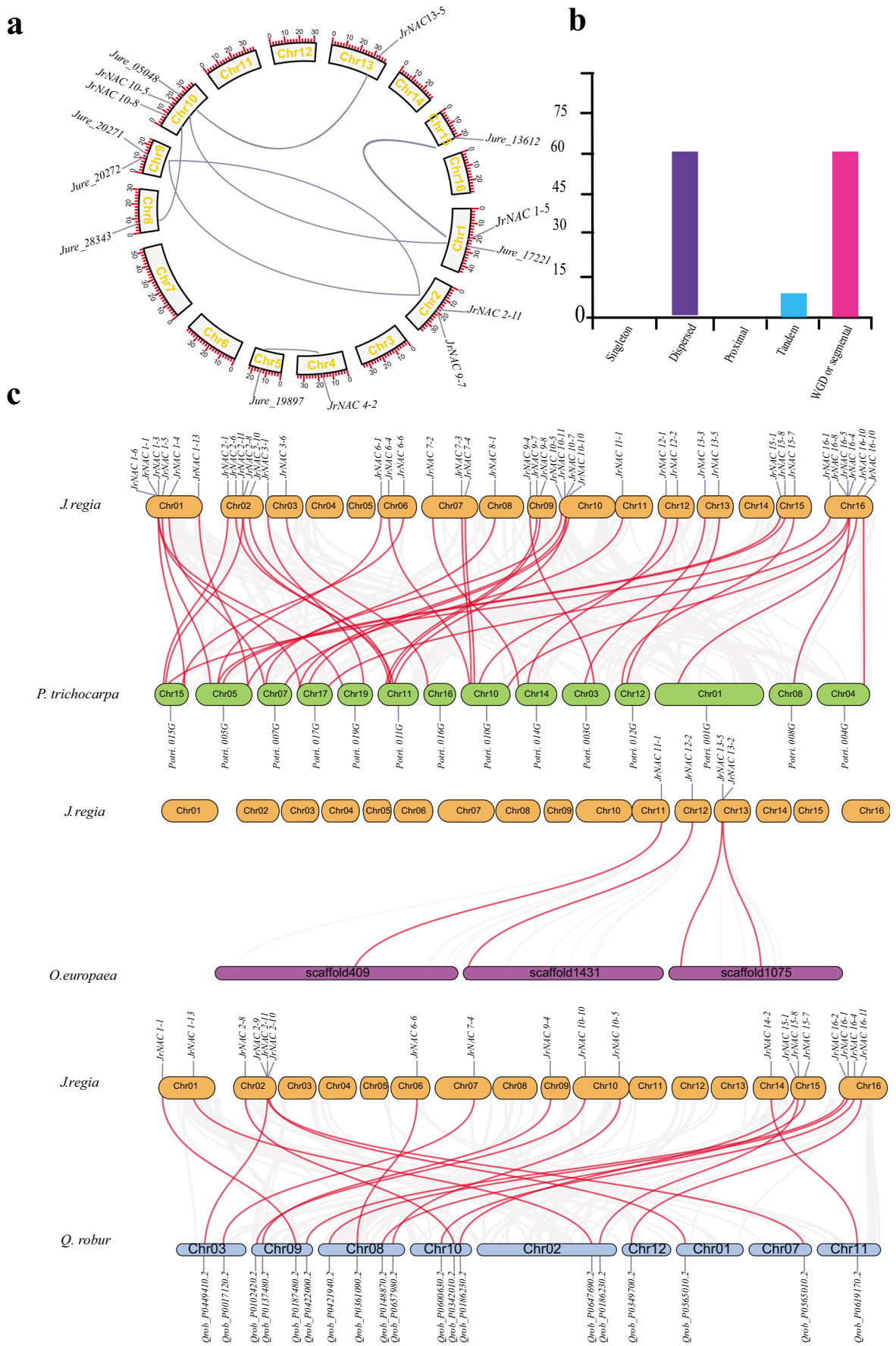


Fig. 1 Phylogenetic analysis of *NAC* proteins in Persian walnut (102). A neighbor-joining (NJ) tree was constructed using 102 *NAC* sequences. The tree further clustered into 10 subfamilies

Syntenic analysis was conducted for three plants (*Populus trichocarpa*, *Olea europaea*, and *Quercus robur*) to infer the *NAC* genes' evolutionary relationship among these three species (Fig. 3c). We identified 42 pairs of orthologous *NAC* genes between *J. regia* and *P. trichocarpa*, a total of 23 orthologous gene pairs between *J. regia* and *Q. robur*, and only 4 orthologous gene pairs between *J. regia* and *O. europaea* (Fig. 3c). Interestingly, 16 collinear gene pairs identified between *J. regia* and *Q. robur* were not found between *J. regia* and *O. europaea*/*P. trichocarpa*, 7 collinear gene pairs identified between *J. regia* and *P. trichocarpa* were not found between *J. regia* and *O. europaea*/*Q. robur*, and 4 collinear gene pairs identified between *J. regia* and *O. europaea* were not found between *J. regia* and *P. trichocarpa*/*Q. robur* (Fig. 3c). However, one collinear gene pair was found in three species, *J. regia*, *P. trichocarpa*, and *O. europaea*, and three collinear gene pairs were found in three species, *J. regia*, *P. trichocarpa*, and *Q. robur* (Fig. 3c).

Fig. 2 The phylogenetic relationship, motif compositions, and gene structure of *JrNAC* transcription factors. **a** The phylogenetic relationships of *JrNACs* based on the NJ method. The various colors characterize the ten subfamilies. **b** Motif compositions of *JrNACs*. Gray lines indicate non-conserved sequences, and colored boxes represent conserved motifs. The motifs are displayed proportionally in each protein. **c** Gene structure of *JrNACs*. CDS represents coding sequence, green boxes indicate CDS, and gray lines represent introns; 0, 1, and 2 represent different types of phase. Phase 0, located between two consecutive codons; phase 1, splitting codon between the first and second nucleotides; Phase 2, between the second and third nucleotides of a codon


(a)
(b)
(c)
JrNAC 2-2
JrNAC 9-3
JrNAC 10-10
JrNAC 9-11
JrNAC 4-4
JrNAC 3-1
JrNAC 2-9
JrNAC 9-6
JrNAC 15-6
JrNAC 6-6
JrNAC 11-12
JrNAC 8-1
JrNAC 9-2
JrNAC 10-14
JrNAC 1-13
JrNAC 9-5
JrNAC 2-8
JrNAC 1-3
JrNAC 11-1
JrNAC 8-2
JrNAC 15-4
JrNAC 6-3
JrNAC 1-6
JrNAC 10-1
JrNAC 15-5
JrNAC 6-4
JrNAC 12-1
JrNAC 7-2
JrNAC 10-7
JrNAC 2-7
JrNAC 13-3
JrNAC 16-10
JrNAC 9-1
JrNAC 2-1
JrNAC 1-1
JrNAC 1-7
JrNAC 10-5
JrNAC 10-6
JrNAC 1-4
JrNAC 15-1
JrNAC 15-8
JrNAC 2-11
JrNAC 2-10
JrNAC 9-7
JrNAC 9-8
JrNAC 3-4
JrNAC 3-6
JrNAC 16-7
JrNAC 16-8
JrNAC 16-6
JrNAC 16-9
JrNAC 10-2
JrNAC 1-8
JrNAC 13-4
JrNAC 16-5
JrNAC 1-9
JrNAC 2-4
JrNAC 4-1
JrNAC 3-5
JrNAC 16-11
JrNAC 13-2
JrNAC 3-2
JrNAC 3-3
JrNAC 4-3
JrNAC 5-1
JrNAC 14-3
JrNAC 13-1
JrNAC 6-7
JrNAC 15-7
JrNAC 1-2
JrNAC 10-4
JrNAC 6-5
JrNAC 9-4
JrNAC 2-5
JrNAC 12-2
JrNAC 7-4
JrNAC 7-3
JrNAC 10-8
JrNAC 10-9
JrNAC 1-5
JrNAC 6-2
JrNAC 6-1
JrNAC 15-2
JrNAC 15-3
JrNAC 2-3
JrNAC 2-6
JrNAC 14-2
JrNAC 16-1
JrNAC 13-5
JrNAC 16-2
JrNAC 16-3
JrNAC 10-11
JrNAC 1-10
JrNAC 4-2
JrNAC 10-13
JrNAC 10-12
JrNAC 1-12
JrNAC 3-7
JrNAC 16-4
JrNAC 10-3
JrNAC 7-1
I
II
III
IV
V
VI
VII
VIII
IX
X
0
200
400
600
0
3000
6000
9000
CDS
Motif 1
Motif 2
Motif 3
Motif 4
Motif 5
Motif 6
Motif 7
Motif 8
Motif 9
Motif 10
Motif 11
Motif 12
Motif 13
Motif 14
Motif 15
Motif 16
Motif 17
Motif 18
Motif 19
Motif 20


a
b
c
Singleton
Dispersed
Proximal
Tandem
WGD or segmental
J.regia
P. trichocarpa
O.europaea
Q. robur
scaffold409
scaffold1431
scaffold1075

◀ **Fig. 3** Paralogous genes, duplicate type, and synteny analysis of *NAC* genes between *J. regia* and three groups of representative plant species. **a** Schematic representation for the chromosomal distribution and interchromosomal relationships of Persian walnut *NAC* genes. Gray lines indicate duplicated *NAC* gene pairs. The chromosome number is indicated at the bottom of each chromosome. **b** The 102 *JrNAC* member duplicate type in Persian walnut. **c** Synteny analysis of *NAC* genes between *J. regia* and three plant species. Gray lines in the background indicate the collinear blocks within Persian walnut and other plant genomes, while the red lines highlight the syntenic *NAC* gene pairs. The species names with the prefixes "*P. trichocarpa*," "*O. europaea*," and "*Q. robur*" indicate *Populus trichocarpa*, *Olea europaea*, and *Quercus robur*, respectively

3.4 Promoter region analysis and conserved domains of *JrNACs*

We found that *JrNAC1-4* contained auxin response component, while the gene *JrNAC2-6* has a protein-binding site based on in silico analysis (Fig. 4a). The *JrNAC* promoter regions contain 77% of light-responsive elements, 13% of site-binding-related elements, 7% of environmental stress-related elements, and 3% of the response to plant growth (Fig. 4b).

All the *JrNAC*s contained the *NAM* conserved domain, and six *JrNAC* transcription factors contained special domains at their C-terminal (PEP_TPR_lipo, FBA_1, BAR, SET, Rubis-subs-bind, and CBF5) based on conserved domains analysis (Khan et al. 2020). The analysis of GO enrichment was divided into three categories. In the category of biological processes, bio-regulation, metabolic processes, cellular processes, and stimulus response are significantly enriched terms. In the cellular component category, cell, organelle, and cell parts are significant. GO terms for the transcription factor activity of nucleic acid binding were highly represented in the molecular function category. The most enriched GO term in members of the *JrNAC* is GO: 003674 (molecular function).

3.5 Expression profiles and qRT-PCR analysis of *JrNACs* among different tissues and flowering development stages

We investigated the transcriptome expression profiles in different tissues of the Persian walnut to further provide information on the function of the *JrNAC* transcription factor family. The *JrNACs* were differently expressed in different tissues, indicating that these genes have a function variation. Twenty-four of 102 Persian walnut *NAC* members were expressed highly in female flowers compared with other tissues (M, L, and H), particularly *JrNAC1-4*, *JrNAC16-8*, *JrNAC16-9*, *JrNAC16-1*, *JrNAC6-5*, *JrNAC10-13*, *JrNAC9-6*, and *JrNAC2-6*. Two *JrNACs* (*JrNAC2-9* and *JrNAC9-6*) were highly expressed in male flowers, four genes in leaves and two genes highly expressed in the hull; these results show that the *NAC* genes in different tissues have a different expression pattern (Fig. 5a, Khan et al. 2020). The expression of some genes increases as the flowers grow, such as *JrNAC1-13*, *JrNAC16-9*, *JrNAC16-5*, *JrNAC16-8*, *JrNAC7-1*, *JrNAC2-9*, and *JrNAC9-6* (Fig. 5b).

The transcription levels of *JrNAC1-4* and *JrNAC2-6* in flowering and vegetative tissues were analyzed using qRT-PCR. *JrNAC1-4* and *JrNAC2-6* were highly expressed in female compared with male flowers, while *JrNAC13-5* was expressed highly in leaves of *J. regia* (Fig. 5c). These genes are differentially expressed in female and male flowers and leaves (Fig. 5c), which could be subsequently prioritized in plant functional studies for further analysis.

3.6 The interaction network of *JrNAC* proteins

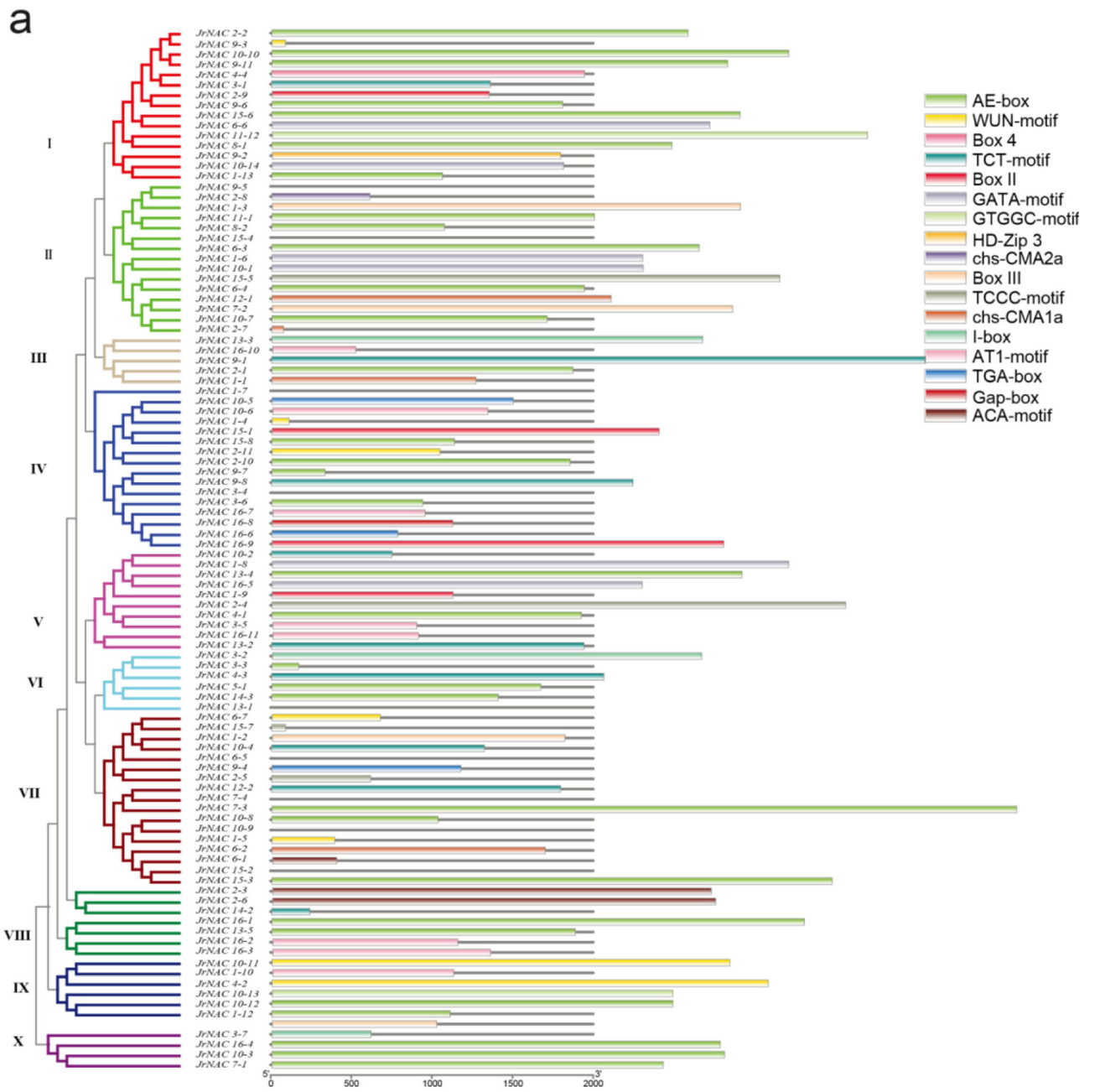
Each *JrNAC* protein was in close association with at least one *NAC* protein from *Arabidopsis*. Some *JrNAC* proteins were closely aligned with the same *NAC* protein in *Arabidopsis*. We downloaded *NAC* proteins from the *Arabidopsis* to detect the predicted role of highly expressed genes in the flowering of Persian walnut. The previous study claimed that these genes regulate the development of the flower. Therefore, we detected the interaction relationship between these genes, and the results indicate a strong relationship between the *JrNAC1-4* proteins and *AtWRKY12* (Fig. 6) (Li et al. 2016).

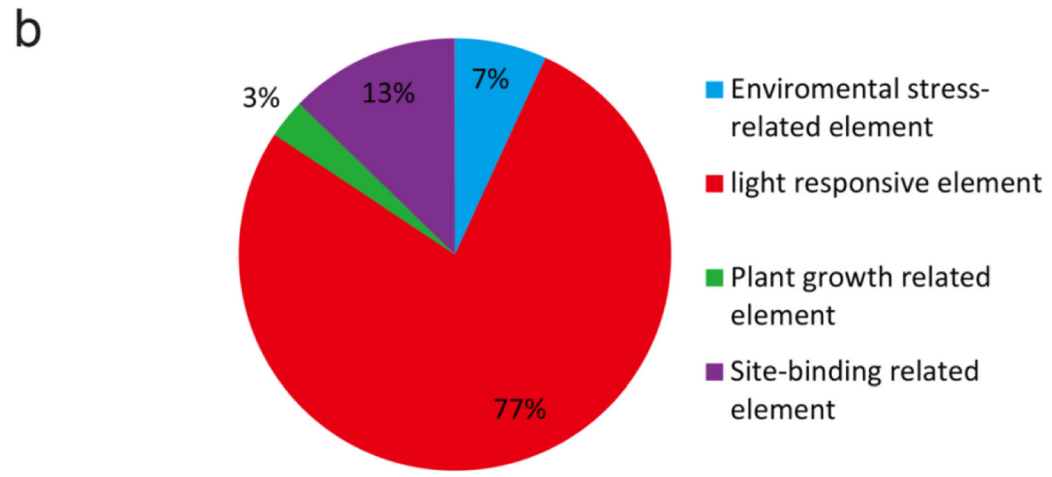
4 Discussion

4.1 The number, phylogenetic relationships, and location of *JrNACs* in Persian walnut

In this study, we identified a total of 102 *NAC* genes in Persian walnut. Previously, a large number of *NAC* genes identified in other plants and contained over 100 members (Ooka et al. 2003; Nuruzzaman et al. 2010; Liu et al. 2014; Peng et al. 2015; Saidi et al. 2017). The number of *NAC* genes in Persian walnut is lower as compared with other plants including *B. pekinensis* (204) (Liu et al. 2014), *O. sativa* (151) (Nuruzzaman et al. 2010), *G. max* (152) (Le et al. 2011), *P. trichocarpa* (163) (Hu et al. 2010), *Arabidopsis* (105) (Ooka et al. 2003), *P. mume* (113) (Zhuo et al. 2018), and *M. domestica* (180) (Jia et al. 2019).

Based on phylogenetic analysis, the *JrNACs* were divided into 10 distinct subfamilies. The phylogenetic tree obtained in this study mainly aligned with previous reports (Ooka et al. 2003; Shen et al. 2009). In Persian walnut, all *NAC* gene motifs 1, 2, 3, and 5 were frequent. The gene structure analysis showed that exon numbers vary from 1 to 22, and this number is greater than in *M. acuminata*, in which the number of exons varies from 0 to six (Cenci et al. 2014); in *O. sativa*, where the number ranges from 0 to 16; and in *G. hirsutum*, where the number is 0 to nine (Nuruzzaman et al. 2010; Zhu et al. 2011). The Persian walnut 62


a
I
II
III
IV
V
VI
VII
VIII
IX
X
AE-box
WUN-motif
Box 4
TCT-motif
Box II
GATA-motif
GTGGC-motif
HD-Zip 3
chs-CMA2a
Box III
TCCC-motif
chs-CMA1a
I-box
AT1-motif
TGA-box
Gap-box
ACA-motif
0
500
1000
1500
2000


b
7%
13%
3%
77%
Enviromental stress-related element
light responsive element
Plant growth related element
Site-binding related element

◀ **Fig. 4** Cis-element analysis in the *JrNAC* gene family. **a** Each gene cis-element in the phylogenetic tree of walnut. **b** The percentage of responsive elements: hormone-responsive elements, environmental stress-related elements, plant growth responsive elements, and site-binding responsive elements (except for TATA and CAAT binding sites) in all *JrNAC* genes

NAC genes mainly contain three exons and two introns. In *Arabidopsis*, *G. hirsutum*, *M. acuminata*, and *O. sativa*, this phenomenon was observed, and three exons were present in the majority of *NAC* genes (Nuruzzaman et al. 2010; Zhu et al. 2011; Cenci et al. 2014). The motif compositions, exon–intron structure, and intron phase 0 of *JrNACs* indicate that *NAC* was highly conserved in each subfamily.

In our study, the *NAC* genes were unevenly distributed on 16 chromosomes but most commonly observed on chromosome 10 (14 *NAC* genes) with comparatively few *NAC*s on chromosome 5 (1 *NAC* gene) (Khan et al. 2020). In the previous studies, for example, in *O. sativa* and *Z. mays NAC* genes, uneven distribution on chromosomes was also reported (Nuruzzaman et al. 2010; Peng et al. 2015).

4.2 The evolution and expansion of *JrNACs*

Tandem duplication, segmental duplication, and WGD were most likely chosen by gene families as forms of expansion (Cannon et al. 2004; Dong et al. 2017). However, a total of 30 *JrNAC*s genes pair were duplicated by WGD, and only one gene pair (*JrNAC2-10* and *JrNAC9-8*) experienced tandem duplication events. These results indicate that the evolutionary expansion patterns of *NAC* transcription factor family members were duplicated by WGD events (Khan et al. 2020; Cannon et al. 2004). This finding contrasts with several previous reports in which a similar phenomenon was analyzed (Zhu et al. 2011; Satheesh et al. 2014). For example, the expansion of the *Populus* and *Gossypium NAC* genes were tandem duplication events (Zhu et al. 2011; Satheesh et al. 2014).

There are collinear genes between the Persian walnut and *P. trichocarpa*, *O. europaea*, and *Q. robur*; these results suggest that the *NAC* genes may have evolved from a common ancestor in different plants (Khan et al. 2020; Satheesh et al. 2014). The K_a/K_s values of the paralogous gene pairs were calculated to estimate their evolution history, and we found that *JrNAC* genes evolved through negative selection as the K_a/K_s values were less than 1. We found that the *JrNAC* promoter regions contain 77% of light-responsive elements, 7% of environmental stress-related elements, and 3% of the response to plant growth, while the *JrNAC1-4* gene promotes the auxin response component based on in silico analysis (Khan et al. 2020). These results indicate that *NAC* members of Persian walnut experienced the selection of ecological factors, especially for light and environmental stress (Tran et al. 2004; Tran et al. 2010; Puranik et al. 2012; Satheesh et al. 2014).

4.3 The expression profile of *NAC* members of Persian walnut

The *NAC* transcription factor family has diverse functions and plays important roles in plant development and physiological processes (Souer et al. 1996; Sablowski and Meyerowitz 1998; Ooka et al. 2003; Guo and Gan 2006; Liu et al. 2009; Wang et al. 2009; Shan et al. 2012; Wang et al. 2013; Satheesh et al. 2014), involving in at least four types of processes, such as

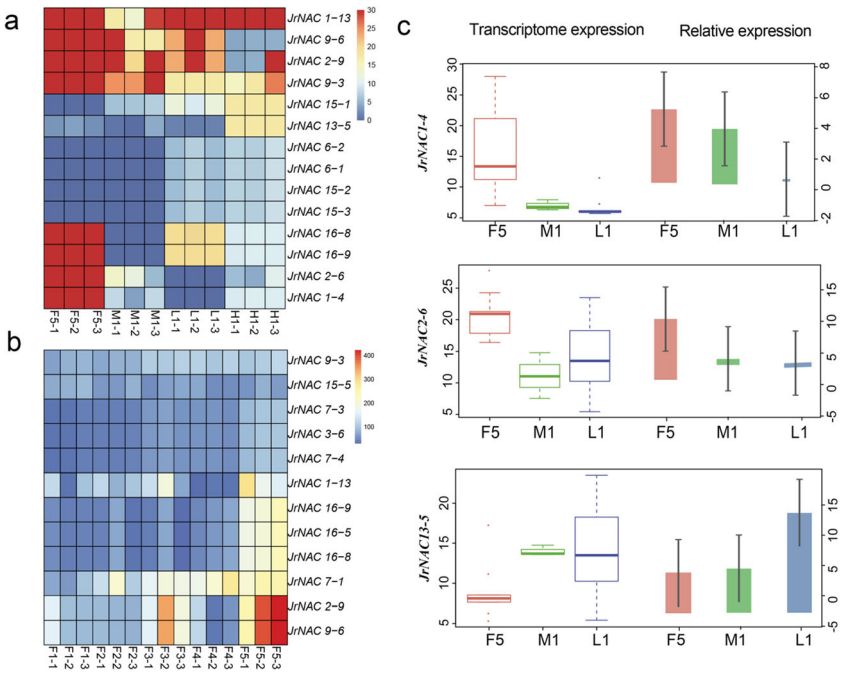


Fig. 5 Expression profiles of the Persian walnut *NAC* genes. **a** Hierarchical clustering of expression profiles of 14 Persian walnut *NAC* genes in 12 samples including F, M, L, and H represents female flower, male flower, leaves, and hull as 3 biological replicates. **b** The heatmap exhibits the ratio of the expression levels of 12 *NAC* genes in five developmental stages; 1, 2, 3, 4, and 5 represent different stages of Persian walnut flower. **c** Expression analysis of 3 *JrNACs* in three representative samples by qRT-PCR. Data were normalized to the β-actin gene, and vertical bars indicate standard deviation

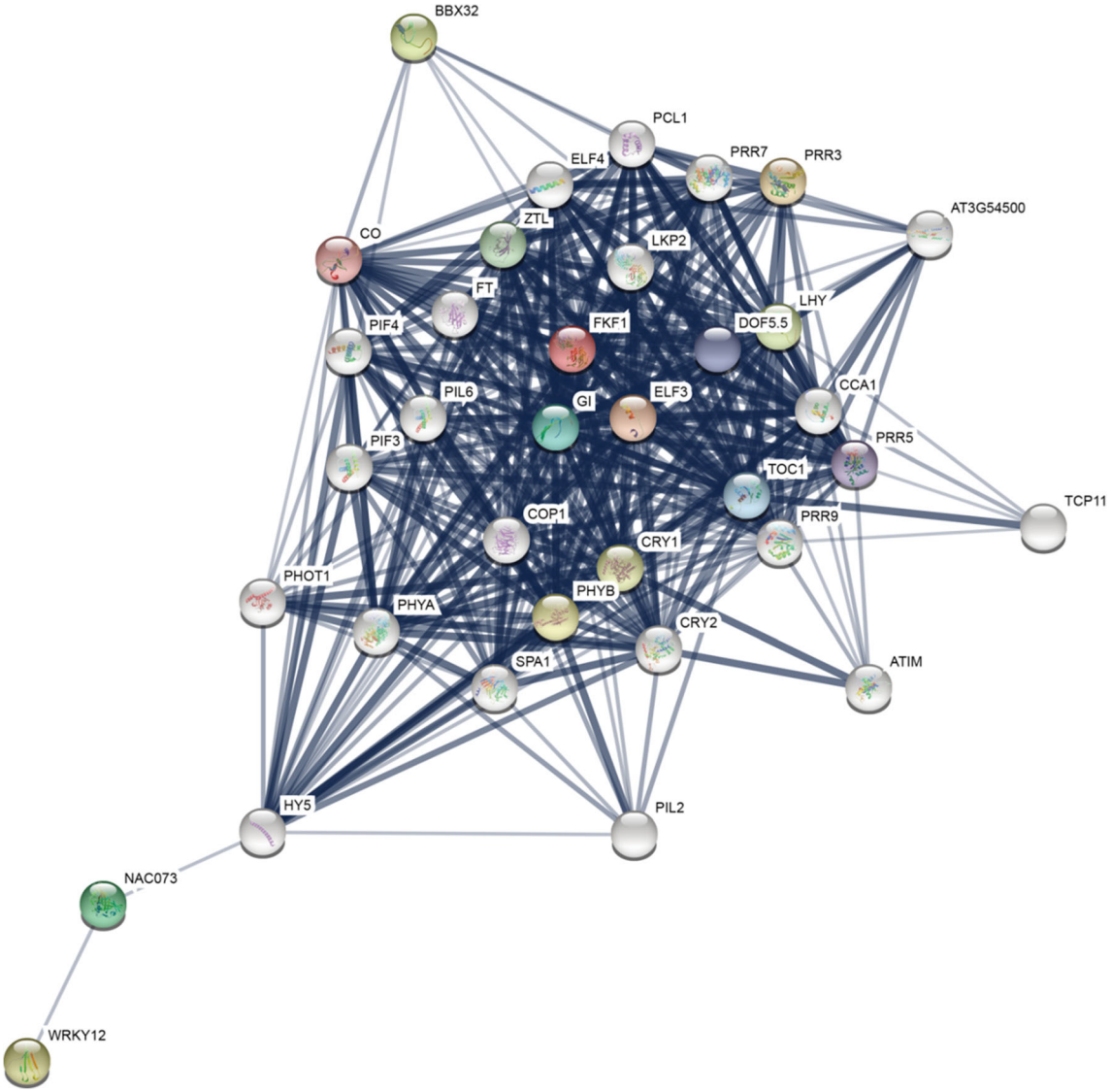


Fig. 6 *JrNAC* protein interaction network. The *JrNAC* protein interaction network was constructed using *Arabidopsis* homologous *NAC* proteins. The corresponding relationship between the walnut *NAC* protein and the *Arabidopsis NAC* protein is shown at the top left of the figure. Proteins are represented by network nodes. The 3D protein structure is displayed inside the nodes. Edges represent associations of proteins

flower formation, cotyledon development, shoot apical meristem maintenance, and subsequent development of the root (Sablowski and Meyerowitz 1998; Takada et al. 2001; He et al. 2005). In this study, a total of seven conserved domains, including one conserved domain (*NAM*) and six special domains, were identified (Khan et al. 2020). In *Z. mays*, 13 distinct conserved domains and at least six conserved domains were identified in *M. notabilis* (Peng et al. 2015; Baranwal and Khurana 2016). Our results are similar to those for *O. sativa* and *S. officinarum*, which also have seven conserved domains (Ooka et al. 2003; Ramaswamy et al. 2017). All the *JrNACs* contain the *NAM* domain, which is consistent with previous studies and is mainly associated with DNA binding and flower development (Hussain et al. 2017; Ramaswamy et al. 2017). In *Petunia*, *G. max*, and *P. trichocarpa*, the *NAM* domain has a role in flower formation, primordia, and embryo development (Souer et al. 1996; Hu et al. 2010; Hussain et al. 2017), indicating that the *JrNACs* may play a potential role in the flowering of *J. regia*. A comprehensive analysis was conducted to evaluate the expression patterns of the *JrNAC* gene family in vegetative and reproductive tissues of Persian walnut. Based on the expression analysis, a total of 24 *JrNAC* genes showed higher expression levels in female flowers, indicating that these genes play a key role in the development of flowers (Ikeda et al. 2004; Tran et al. 2009; Zhou et al. 2010; Su et al. 2013; Singh et al. 2013; Kou et al. 2014; Zhuo et al. 2018), and the results for two genes (*JrNAC1-4* and *JrNAC2-6*) were also supported by qRT-PCR.

The results of the interaction relationship indicate a strong relationship between the *JrNAC1-4* proteins and the *AtWRKY12*

proteins (Li et al. 2016). Notably, *JrNAC1-4* contained part of an auxin-responsive element and has a systemic relationship with *P. trichocarpa* (*Potri.007G099400.1*) and *Q. robur* (*Qrob_P0657980.2*). *JrNAC1-4* was highly expressed in female and male flowers, with extremely low expression in leaves, which is consistent with the previous studies. For example, *O. sativa* (*ONAC300*) (Zhou et al. 2010), *G. max* (*GmNAC016* and *GmNAC14*) (Tran et al. 2009), *Arabidopsis* (*NAP*) (Sablowski and Meyerowitz 1998), *S. tuberosum* (*StNAC034* and *StNAC075*) (Singh et al. 2013), *M. domestica* (*MdNAC42*, *MdNAC110*, and *MdNAC138*) (Su et al. 2013), *S. lycopersicum* (*SNAC8*) (Kou et al. 2014), *Fragaria* (*FaNAC021*, *FaNAC022*, *FaNAC042*, and *FaNAC092*) (Moyano et al. 2018), and *P. mume* (*PmNAC*) (Zhuo et al. 2018) also showed higher expression in flowers and also showed high similarity with *AtWRKY12*, a gene that regulates the development of flowers (Li et al. 2016). These results show that *JrNAC1-4* might play a potential role in Persian walnut flowering. Taken together, our results suggest that *JrNAC1-4* may play important roles in Persian walnut flowering development.

5 Conclusion

In Persian walnut (*J. regia*), we identified a total of 102 *NAC* transcription factors. Phylogenetic analysis showed that the *NAC* transcription factors are clustered into 10 subfamilies. Based on the conserved domains, *NAC* transcription factors contain a conserved domain (NAM). The analysis of the expression profile showed that the *NAC* transcription factors reveal diverse patterns of expression in different Persian walnut tissues. Most of the Persian walnut *NAC* transcription factors are expressed highly in female and male flowers. A total of 24 *NAC* transcription factors were highly expressed in female and male flowers, which might play a role in *J. regia* flowering. The transcription data and qRT-PCR analysis indicated that two *NAC* transcription factors (*JrNAC1-4* and *JrNAC2-6*) were highly expressed in female and male flowers, while *JrNAC13-5* was expressed highly in leaves. In conclusion, these results provide a base for studying the potential function of Persian walnut *NAC* transcription factors.

Acknowledgment We thank Hui-Juan Zhou and Meng Dang for collecting the samples.

Funding information This study was funded by the Natural Science Foundation of Shaanxi Province of China (2019JM-008), Shaanxi Academy of Science Research Funding Project (2019K-06), the National Natural Science Foundation of China (No. 41471038; No. 31200500; No. 31860215), the Program for Excellent Young Academic Backbones funded by Northwest University, Shaanxi Academy of Science Research Funding Project (Y19Z604F12), and Opening Foundation of Key Laboratory of Resource Biology and Biotechnology in Western China (Northwest University), Ministry of Education (ZSK2018009), and the public health specialty in the Department of Traditional Chinese Medicine (2017-66, 2018-43, 2019-68).

Data availability The datasets generated and/or analyzed during the current study are available in the Zenodo repository (https://doi.org/10.5281/zenodo.3905995).

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

References

Aoki K, Ogata Y, Shibata D (2007) Approaches for extracting practical information from gene co-expression networks in plant biology. Plant Cell Physiol 48:381–390

Bailey TL, Johnson J, Grant CE, Noble WS (2015) The MEME suite. Nucleic Acids Res 43:W39–W49

Baranwal VK, Khurana P (2016) Genome-wide analysis, expression dynamics and varietal comparison of *NAC* gene family at various developmental stages in *Morus notabilis*. Mol Gen Genomics 291: 1305–1317

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinformatics 10:421

Cannon SB, Mitra A, Baumgarten A, Young ND, May G (2004) The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. BMC Plant Biol 4:10

Cenci A, Guignon V, Roux N, Rouard M (2014) Genomic analysis of *NAC* transcription factors in banana (*Musa acuminata*) and definition of *NAC* orthologous groups for monocots and dicots. Plant Mol Biol 85:63–80

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674–3676

Dong W, Xu C, Li W, Xie X, Lu Y, Liu Y, Jin X, Suo Z (2017) Phylogenetic resolution in *Juglans* based on complete chloroplast genomes and nuclear DNA sequences. Front Plant Sci 8:1148

Duval M, Hsieh TF, Kim SY, Thomas TL (2002) Molecular characterization of *AtNAM*: a member of the *Arabidopsis NAC* domain superfamily. Plant Mol Biol 50:237–248

El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A (2018) The Pfam protein families database in 2019. Nucleic Acids Res 47:D427–D432

Feng X, Yuan X, Sun Y, Hu Y, Zulfiqar S, Ouyang X, Dang M, Zhou H, Woeste K, Zhao P (2018) Resources for studies of iron walnut (*Juglans sigillata*) gene expression, genetic diversity, and evolution. Tree Genet Genom 14:51

Garcia-Hernandez M, Berardini T, Chen G, Crist D, Doyle A, Huala E, Knee E, Lambrecht M, Miller N, Mueller LA (2002) TAIR: a resource for integrated *Arabidopsis* data. Funct Integr Genomic 2: 239–253

Guo Y, Gan S (2006) *AtNAP*, a *NAC* family transcription factor, has an important role in leaf senescence. Plant J 46:601–612

Han Q, Zhang J, Li H, Luo Z, Ziaf K, Ouyang B, Wang T, Ye Z (2012) Identification and expression pattern of one stress-responsive *NAC* gene from *Solanum lycopersicum*. Mol Biol Rep 39:1713–1720

Han H, Woeste KE, Hu Y, Dang M, Zhang T, Gao XX, Zhou H, Feng X, Zhao G, Zhao P (2016) Genetic diversity and population structure of common walnut (*Juglans regia*) in China based on EST-SSRs and the nuclear gene phenylalanine ammonia-lyase (*PAL*). Tree Genet Genom 12:111

He XJ, Mu RL, Cao WH, Zhang ZG, Zhang JS, Chen SY (2005) *AtNAC2*, a transcription factor downstream of ethylene and auxin

signaling pathways, is involved in salt stress response and lateral root development. Plant J 44:903–916
Hennig L, Gruissem W, Grossniklaus U, Köhler C (2004) Transcriptional programs of early reproductive stages in *Arabidopsis*. Plant Physiol 135:1765–1775
Hu W, Wang Y, Bowers C, Ma H (2003) Isolation, sequence analysis, and expression studies of florally expressed cDNAs in *Arabidopsis*. Plant Mol Biol 53:545–563
Hu R, Qi G, Kong Y, Kong D, Gao Q, Zhou G (2010) Comprehensive analysis of *NAC* domain transcription factor gene family in *Populus trichocarpa*. BMC Plant Biol 10:145
Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G (2014) GSDS2.0: an upgraded gene feature visualization server. Bioinformatics 31: 1296–1297
Hussain RM, Ali M, Feng X, Li X (2017) The essence of *NAC* gene family to the cultivation of drought-resistant soybean (*Glycine max* L. Merr.) cultivars. BMC Plant Biol 17:55
Ikeda K, Igic B, Ushijima K, Yamane H, Hauck NR, Nakano R, Sassa H, Iezzoni AF, Kohn JR, Tao R (2004) Primary structural features of the S haplotype-specific *F-box* protein, *SFB*, in *Prunus*. Sex Plant Reprod 16:235–243
Jia D, Jiang Q, van Nocker S, Gong X, Ma F (2019) An apple (*Malus domestica*) *NAC* transcription factor enhances drought tolerance in transgenic apple plants. Plant Physiol Biochem 139:504–512
Khan H, Yan F, Yan YJ, Chen PP, Xi RM, Ullah I, Peng X, Luo X, Yue M, Zhao P (2020) Genome-wide analysis of evolution and expression profiles of *NAC* transcription factor gene family in *Juglans regia* L. [Dataset]. Zenodo Repository V2. https://doi.org/10.5281/zenodo.3905995
Kou X, Wang S, Wu M, Guo R, Xue Z, Meng N, Tao X, Chen M, Zhang Y (2014) Molecular characterization and expression analysis of *NAC* family transcription factors in tomato. Plant Mol Biol Report 32:501–516
Kumar S, Nei M, Dudley J, Tamura K (2008) MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform 9:299–306
Le DT, Nishiyama R, Watanabe Y, Mochida K, Yamaguchi-Shinozaki K, Shinozaki K, Tran L-SP (2011) Genome-wide survey and expression analysis of the plant-specific *NAC* transcription factor family in soybean during development and dehydration stress. DNA Res 18:263–276
Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res 30:325–327
Li W, Wang H, Yu D (2016) *Arabidopsis WRKY* transcription factors *WRKY12* and *WRKY13* oppositely regulate flowering under short-day conditions. Mol Plant 9:1492–1503
Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452
Ling L, Song L, Wang Y, Guo C (2017) Genome-wide analysis and expression patterns of the *NAC* transcription factor family in *Medicago truncatula*. Physiol Mol Biol Plants 23:343–356
Liu YZ, Baig M, Fan R, Ye JL, Cao YC, Deng XX (2009) Identification and expression pattern of a novel *NAM*, *ATAF*, and *CUC*-like gene from *Citrus sinensis Osbeck*. Plant Mol Biol Report 27:292–297
Liu T, Song X, Duan W, Huang Z, Liu G, Li Y, Hou X (2014) Genome-wide analysis and expression patterns of *NAC* transcription factor family under different developmental stages and abiotic stresses in Chinese cabbage. Plant Mol Biol Report 32:1041–1056
Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR (2016) CDD/Sparcle: functional classification of proteins via subfamily domain architectures. Nucleic Acids Res 45:D200–D203
Martínez-García PJ, Crepeau MW, Puju D, Gonzale-lbeas D, Whalen J, Stevens KA, Paul P, Butterfield TS, Britton MT, Reagan RL, Chakraborty S, Walwage SL, Vasquez-Gross HA, Cardeno C, Famula RA, Pratt K, Kuruganti S, Aradhya MK, Leslie CA, Dandekar AM, Salzberg SL, Wegrzyn JL, Langley CH, Neale DB (2016) The walnut (*Juglans regia*) genome sequence reveals diversity in genes coding for the biosynthesis of non-structural polyphenols. Plant J 87:507–532
Moyano E, Martínez-Rivas FJ, Blanco-Portales R, Molina-Hidalgo FJ, Ric-Varas P, Matas-Arroyo AJ, Caballero JL, Muñoz-Blanco J, Rodríguez-Franco A (2018) Genome-wide analysis of the *NAC* transcription factor family and their expression during the development and ripening of the *Fragaria× ananassa* fruits. PLoS One 13: e0196953
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
Ooka H, Satoh K, Doi K, Nagata T, Otomo Y, Murakami K, Matsubara K, Osato N, Kawai J, Carninci P (2003) Comprehensive analysis of *NAC* family genes in *Oryza sativa* and *Arabidopsis thaliana*. DNA Res 10:239–247
Peng X, Zhao Y, Li X, Wu M, Chai W, Sheng L, Wang Y, Dong Q, Jiang H, Cheng B (2015) Genome wide identification, classification and analysis of *NAC* type gene family in maize. J Genet 94:377–390
Prakash A, Jeffryes M, Bateman A, Finn RD (2017) The HMMER web server for protein sequence similarity search. Curr Protoc Bioinformatics 60:3.15 11-13.15. 23
Puranik S, Sahu PP, Srivastava PS, Prasad M (2012) *NAC* proteins: regulation and role in stress tolerance. Trends Plant Sci 17:369–381
Puranik S, Sahu PP, Mandal SN, Parida SK, Prasad M (2013) Comprehensive genome-wide survey, genomic constitution and expression profiling of the *NAC* transcription factor family in foxtail millet (*Setaria italica* L.). PLoS One 8:e64594
Ramaswamy M, Narayanan J, Manickavachagam G, Athiappan S, Arun M, Gomathi R, Ram B (2017) Genome wide analysis of *NAC* gene family 'sequences' in sugarcane and its comparative phylogenetic relationship with rice, sorghum, maize and *Arabidopsis* for prediction of stress associated *NAC* genes. Agri Gene 3:1–11
Sablowski RW, Meyerowitz EM (1998) A homolog of no apical meristem is an immediate target of the floral homeotic genes *APETALA3*/*PISTILLATA*. Cell 92:93–103
Saidi MN, Mergby D, Brini F (2017) Identification and expression analysis of the *NAC* transcription factor family in durum wheat (*Triticum turgidum* L. ssp. durum). Plant Physiol Biochem 112:117–128
Satheesh V, Jagannadham PTK, Chidambaranathan P, Jain P, Srinivasan R (2014) *NAC* transcription factor genes: genome-wide identification, phylogenetic, motif and cis-regulatory element analysis in pigeonpea (*Cajanus cajan* (L.) Millsp). Mol Biol Rep 41:7763–7773
Schultz J, Copley RR, Doerks T, Ponting CP, Bork P (2000) SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res 28:231–234
Shan W, Kuang JF, Chen L, Xie H, Peng HH, Xiao YY, Li XP, Chen WX, He QG, Chen JY (2012) Molecular characterization of banana *NAC* transcription factors and their interactions with ethylene signalling component EIL during fruit ripening. J Exp Bot 63:5171–5187
Shen H, Yin Y, Chen F, Xu Y, Dixon RA (2009) A bioinformatic analysis of *NAC* genes for plant cell wall development in relation to lignocellulosic bioenergy production. BioEnergy Res 2:217–232
Singh AK, Sharma V, Pal AK, Acharya V, Ahuja PS (2013) Genome-wide organization and expression profiling of the *NAC* transcription factor family in potato (*Solanum tuberosum* L.). DNA Res 20:403–423
Souer E, van Houwelingen A, Kloos D, Mol J, Koes R (1996) The no apical meristem gene of *Petunia* is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries. Cell 85:159–170
Su H, Zhang S, Yuan X, Chen C, Wang XF, Hao YJ (2013) Genome-wide analysis and identification of stress-responsive genes of the

NAM-ATAF1, *2-CUC2* transcription factor family in apple. Plant Physiol Biochem 71:11–21

Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P (2016) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res:gkw937

Takada S, Hibara KI, Ishida T, Tasaka M (2001) The cup-shaped cotyledon1 gene of *Arabidopsis* regulates shoot apical meristem formation. Development 128:1127–1135

Tran LSP, Nakashima K, Sakuma Y, Simpson SD, Fujita Y, Maruyama K, Fujita M, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (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

Tran LSP, Quach TN, Guttikonda SK, Aldrich DL, Kumar R, Neelakandan A, Valliyodan B, Nguyen HT (2009) Molecular characterization of stress-inducible *GmNAC* genes in soybean. Mol Gen Genomics 281:647–664

Tran LSP, Nishiyama R, Yamaguchi-Shinozaki K, Shinozaki K (2010) Potential utilization of *NAC* transcription factors to enhance abiotic stress tolerance in plants by biotechnological approach. GM crops 1: 32–39

Wang XE, Basnayake BVS, Zhang H, Li G, Li W, Virk N, Mengiste T, Song F (2009) The *Arabidopsis ATAF1*, a *NAC* transcription factor, is a negative regulator of defense responses against necrotrophic fungal and bacterial pathogens. Mol Plant Microbe Interact 22: 1227–1238

Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H (2012) MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res 40:e49–e49

Wang N, Zheng Y, Xin H, Fang L, Li S (2013) Comprehensive analysis of NAC domain transcription factor gene family in Vitis vinifera. Plant Cell Rep 32:61–75

Xu F, Deng G, Cheng S, Zhang W, Huang X, Li L, Cheng H, Rong X, Li J (2012) Molecular cloning, characterization and expression of the phenylalanine ammonia-lyase gene from *Juglans regia*. Molecules 17:7810–7823

Yan F, Li H, Zhao P (2019a) Genome-wide identification and transcriptional expression of the *PAL* gene family in common walnut (*Juglans regia* L.). Genes 10:46

Yan F, Zhou H, Yue M, Yang G, Li H, Zhang S, Zhao P (2019b) Genome-wide identification and transcriptional expression profiles of the *F-box* gene family in common walnut (*Juglans regia* L.). Forests 10:275

Zhao P, Zhou HJ, Potter D, Hu YH, Feng XJ, Dang M, Feng L, Zulfiqar S, Liu WZ, Zhao GF (2018) Population genetics, phylogenomics and hybrid speciation of *Juglans* in China determined from whole chloroplast genomes, transcriptomes, and genotyping-by-sequencing (GBS). Mol Phylogen Evol 126:250–265

Zhou HL, Zhang HY, Jie Z (2010) Cloning and expression analysis of an *AP2/ERF* gene and its responses to phytohormones and abiotic stresses in rice. Rice Sci 17:1–9

Zhu H, Han X, Lv J, Zhao L, Xu X, Zhang T, Guo W (2011) Structure, expression differentiation and evolution of duplicated fiber developmental genes in *Gossypium barbadense* and *G. hirsutum*. BMC Plant Biol 11:40

Zhu Z, Shi J, He M, Cao J, Wang Y (2012) Isolation and functional characterization of a transcription factor *VpNAC1* from Chinese wild *Vitis pseudoreticulata*. Biotechnol Lett 34:1335–1342

Zhuo X, Zheng T, Zhang Z, Zhang Y, Jiang L, Ahmad S, Sun L, Wang J, Cheng T, Zhang Q (2018) Genome-wide analysis of the *NAC* transcription factor gene family reveals differential expression patterns and cold-stress responses in the woody plant *Prunus mume*. Genes 9:494

Affiliations

Hanif Khan[1] · Feng Yan[1] · Yujie Yan[1] · Pengpeng Chen[1] · Ruimin Xi[1] · Irfan Ullah[1] · Xiaobang Peng[2] · Xiang Luo[3] · Ming Yue[4] · Peng Zhao[1]

Hanif Khan
hanif.botany76@gmail.com

Feng Yan
Fengyan115@163.com

Yujie Yan
1531329726@qq.com

Pengpeng Chen
1454825096@qq.com

Ruimin Xi
174471982@qq.com

Irfan Ullah
uirfan904@gmail.com

Xiaobang Peng
441254924@qq.com

Xiang Luo
tjaulx@126.com

[1] Key Laboratory of Resource Biology and Biotechnology in Western China, Ministry of Education, College of Life Sciences, Northwest University, Xi'an 710069, Shaanxi, People's Republic of China

[2] Department of Biological and Medical Engineering, Shangluo University, Shangluo 726000, Shaanxi, People's Republic of China

[3] Zhengzhou Fruit Research Institute, Chinese Academy of Agricultural Sciences, Zhengzhou 450009, People's Republic of China

[4] Xi'an Botanical Garden of Shaanxi Province, Xi'an 710061, Shaanxi, People's Republic of China