Physiol Mol Biol Plants (October–December 2015) 21(4):465–478
DOI 10.1007/s12298-015-0325-z

RESEARCH ARTICLE

Identification and expression analyses of *MYB* and *WRKY* transcription factor genes in *Papaver somniferum* L

Tayebeh Kakeshpour[1] · **Shadi Nayebi**[1] · **Sajad Rashidi Monfared**[1] · **Ahmad Moieni**[1] · **Ghasem Karimzadeh**[1]

Received: 17 August 2015 / Revised: 30 September 2015 / Accepted: 5 October 2015 / Published online: 13 October 2015

Abstract *Papaver somniferum* L. is an herbaceous, annual and diploid plant that is important from pharmacological and strategic point of view. The cDNA clones of two putative *MYB* and *WRKY* genes were isolated (GeneBank accession numbers KP411870 and KP203854, respectively) from this plant, via the nested-PCR method, and characterized. The MYB transcription factor (TF) comprises 342 amino acids, and exhibits the structural features of the R2R3MYB protein family. The WRKY TF, a 326 amino acid-long polypeptide, falls structurally into the group II of WRKY protein family. Quantitative real-time PCR (qRT-PCR) analyses indicate the presence of these TFs in all organs of *P. somniferum* L. and *Papaver bracteatum* L. Highest expression levels of these two TFs were observed in the leaf tissues of *P. somniferum* L. while in *P. bracteatum* L. the espression levels were highest in the root tissues. Promoter analysis of the 10 co-expressed gene clustered involved in noscapine biosynthesis pathway in *P. somniferum* L. suggested that not only these 10 genes are co-expressed, but also share common regulatory motifs and TFs including MYB and WRKY TFs, and that may explain their common regulation.

Keywords MYB, and WRKY TFs · Nested-PCR · *P. somniferum* L. · Promoter analysis · qRT-PCR

Tayebeh Kakeshpour and Shadi Nayebi contributed equally to this work.

✉ Sajad Rashidi Monfared
Rashidims@modares.ac.ir

[1] Plant Breeding and Biotechnology Department, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran

Abbreviations

BAC	Bacterial Artificial Chromosome
E.F.Iα	*Elongation Factor I α*
ESTs	Expressed Sequence Tags
HTH	Helix-Turn-Helix
ORF	Open Reading Frame
qRT-PCR	Quantitative Real-Time PCR
TFBS	Transcription Factor Binding Site
TSS	Transcription Start Site

Introduction

BIAs are low molecular weight secondary metabolites that are produced in higher plants often in response to environmental abiotic and biotic stresses, and have been pharmaceutically useful. They are derived from L-tyrosine, and diversified to generate 2500 known structures (Stranska et al. 2013). Opium poppy, *P. somniferum* L., one of the ancient, herbal medicines, is the only commercial source of many precious BIAs such as noscapine, morphine (Choe et al. 2011), codeine, papaverine, and sanguinarine which have been proven to have anticancer, analgesic, cough suppressant, muscle drelaxants, and antimicrobial activities, respectively (Holková et al. 2010; Dang et al. 2012). Despite the apparent vital roles in various physiological programs, specially, regulation of secondary metabolites, MYB and WRKY TFs in *P. somniferum* L. are ill-explored.

MYB proteins are among the largest TF superfamilies in eukaryotes (Riechmann et al. 2000). They function in diverse biological processes including regulation of primary/secondary metabolisms, biotic/abiotic stress responses, hormone syntheses, signal transduction and control of cell cycle as well as its development (Petroni et al. 2008; Gomez-Gomez et al. 2012; Chen et al. 2014). COLORED1 (C1) was the first identified gene encoding a MYB domain protein found in plants. It is involved

in anthocyanin synthesis in the aleurone layer of *Zea mays* kernels (Paz-Ares et al. 1987). Functionally, MYBs are Sequence specific DNA-binding TFs that regulate gene expressions at the mRNA level through interaction with specific motifs in promoter regions of their target genes (Gantet and Memelink 2002; Rabara et al. 2013). Members of the MYB family are characterized by possession of a MYB domain located near the amino terminus, and are composed of one to three imperfect tandem repeats. The MYB repeat comprises a region of 50 to 53 amino acids containing three regularly spaced tryptophan/phenylalanine residues. Typically, each repeat is capable of forming three α-helices. The two that are located at the C-terminus adopt a variation of the HTH conformation that binds to the DNA major groove at the specific recognition site. The three tryptophan residues together form a hydrophobic core that participates in protein–DNA interactions (Jia et al. 2004). According to the number of repeats in the MYB domain, MYB proteins can be divided into four subfamilies: 4R-MYB with four adjacent repeats, 3R-MYB (R1R2R3-MYB) with three adjacent repeats, R2R3-MYB with two adjacent repeats, and MYB-related proteins with one or two separated repeats (Dubos et al. 2010).

WRKY proteins also represent a superfamily of plant-specific TFs (Rushton et al. 2010). They play crucial roles in a variety of biological processes including wounding, responses to biotic/abiotic stresses, and consequently secondary metabolisms (Li et al. 2012; Liu et al. 2013; Qin et al. 2013). Several WRKYs are involved in the regulation of plant growth and developmental processes such as leaf senescence, seed and trichome developments, embryogenesis, and germination (Rushton et al. 2010; Li et al. 2014). The first WRKY protein, SPF1, which takes part in the regulation of sucrose induced genes, was isolated from *Ipomoea batatasin* in 1994 (Ishiguro and Nakamura 1994; Xiong et al. 2013). WRKY TFs recognize the conserved consensus motif TGAC(C/T) of the W-box which has been identified in the promoter sequence of many genes such as stress-inducible ones, in addition to WRKY genes themselves, and regulate expression of target genes (Ando et al. 2014; Li and Luan 2014). They contain a 60 residue domain with the conserved WRKYGQK core sequence in their N-terminus, which is followed by either a C2H2 (Cx4–5Cx22–23HxH), or C2HC (Cx7Cx23HxC) zinc finger motif (Liu et al. 2013; Qin et al. 2013). Based on the number of WRKY domains and type of zinc-finger, WRKY proteins are classified into three distinct groups. The first group contains two WRKY domains with a C2H2 zinc-finger motif, the second group contains one WRKY domain with a C2H2 zinc-finger motif, and the third group contains one WRKY domain with a C2HC zinc-finger motif (Eulgem et al. 2000; Liu et al. 2013).

Herein, we report the isolation and characterization of MYB and WRKY TFs, and study their expression profiles in various tissues of *P. somniferum* L. and *P. bracteatum* L. Furthermore, to investigate whether the members of MYB and WRKY family are present in promoters of the genes involved in BIAs biosynthesis, promoter analysis of 10-gene clustered involved in noscapine biosynthesis in *P. somniferum* L. are performed.

Materials and methods

EST assembly and identification of *MYB* and *WRKY* genes

EST sequences for *P. somniferum* L. were downloaded from several RNAseq projects (http://www.ncbi.nlm.nih.gov/sra/). Then, 468 ESTs and 943 ESTs that had high similarity respectively to the MYB and WRKY families were selected using the Offline BLAST software, and were subsequently used to build up two consensus sequences. After that, the EST sequences were assembled via the "align-then-assemble" approach utilizing the Codon Code Aligner v. 5.0.1. Program, and then two consensus sequences for each group of ESTs were created. Next, ORF of each chosen consensus sequence was found via the ORF finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). Finally, according to the bioinformatics analysis, gene-specific PCR primer pairs were designed for PCR-amplification of *MYB* and *WRKY* genes, separately, based on the complete ORF cDNA sequences (Table 1). The primers were confirmed by the Oligo Analyzer v.3.1 (http://eu.idtdna.com/calc/analyzer).

Sequence alignment and phylogenetic construction

After a BLASTp search on the NCBI database, the MYB and WRKY protein sequences from different species that had more than 50 % identity with the coding region of the consensus sequences were selected. For determining the relationship between the identified *MYB* and *WRKY* TFs and the downloaded proteins from the BLASTp search, multiple alignments were run utilizing the web-based ClustalW2 program (http://www.ebi.ac.uk/Tools/msa/clustalw2/). The maximum likelihood method in MEGA 6.06 was used to conduct the phylogenetic trees (Tamura et al. 2013), and 1000 iterations were used for calculating bootstrap values.

RNA extraction, and cDNA synthesis

Total RNA of poppy's capsule, leaf, root and stem tissues was extracted using a commercial RNA extraction kit (Top Plant and Fungi RNA Purification Kit, TOPAZ GENE RESEARCH., cat No.: TGK2004, Iran). The quantity and quality of the extracted RNA was determined via a NanoDrop spectrophotometer (BioTek, EPOCH, serial: 121004C, USA), and confirmed by agarose gel electrophoresis. RNase-free DNaseI (TIANGEN Biotech(Beijing)Co., Ltd., cat No.: RT411, China) was used to eliminate genomic DNA content. TIANScript M-MLV (TIANGEN

Table 1 PCR primers

Use	Gene name	Primer name	Sequence	Position
Nested-PCR	*MYB*	Forward.2	5′-atgggccaacattcttgttgta-3’	Outside ORF
		Forward.1	5′-ttaaaaagatgatgaagaaggataa-3’	Outside ORF
		Forward.ORF	5′-ggaggcaaaggataaacagata-3’	ORF
		Reverse.ORF	5′-agggatgtaatgccgtttctta-3’	ORF
		Reverse.1	5′-tttgttgtttcaaatcccttagtt-3’	Outside ORF
	WRKY	Forward.1	5′-cagtacagagatatgactattga-3’	Outside ORF
		Forward. ORF	5′-atgtattccgatcatccgacg-3’	Inside ORF
		Reverse.ORF	5′-ctaatactttgtgttgccaagag-3’	ORF
		Reverse.1	5′-atgacctaacttcaaagaacca-3’	Outside ORF
		Reverse.2	5′-tggtaatggtcgaaactcatag-3’	Outside ORF
qRT-PCR	*MYB*	Forward	5′-gttaatgataaatgacagtgagg-3’	Inside ORF
		Reverse (Reverse.ORF)	5′-ttaaaaagatgatgaagaaggataa-3’	ORF
	WRKY	Forward	5′-gcccataaatctaactctgatga-3’	Inside ORF
		Reverse (Reverse.ORF)	5′-ctaatactttgtgttgccaagag-3’	ORF
	EF1α	Forward	5′-ctggtggttttgaagctggt-3’	
		Reverse	5′-tgttgtcaccctcgaatcca-3’	

Biotech(Beijing)Co.,Ltd., cat No.: ER104, China) was used to synthesize the first-strand cDNA.

Gene isolation, and nested-PCR

The expected lengths of amplification fragments are shown (Fig. 1). Nested primers were designed based on different positions on consensus sequences to amplify fragments which overlap with each other to confirm the characterized ORFs. The following program was used for nested-PCR: 10 pmol.μl^{-1} forward primer, and 10 pmol.μl^{-1} reverse primer, 10 mmol.L^{-1} dNTPs, 50–100 ng.μl^{-1} cDNA, 5 U.μl^{-1} Taq DNA Polymerase (TOPAZ GENE RESEARCH., Cat No.: TGT3001.1, Iran), 10 X Taq Buffer and 10 X $MgCl_2$. The total volume for each reaction was 20 µl. The following thermal cycler was used for amplification: 3 min at 94 °C; 35 cycles of 30 s at 94 °C, 45 s at 47–52 °C, 1 min at 72 °C; followed by 10 min at 72 °C. The PCR products of forward.ORF and reverse.ORF (for *MYB* primers), and forward.1 and reverse.ORF (for *WRKY* primers) were cleaned using a commercial gel recovery kit (Top Gel Recovery Kit, TOPAZ GENE RESEARCH., cat No.:

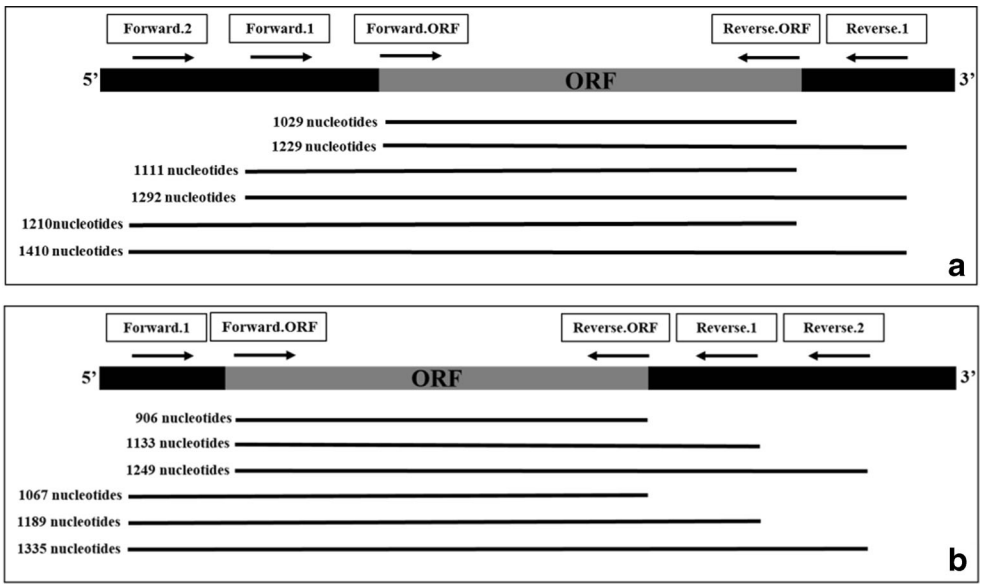

Fig. 1 **a.** Positions of nested-PCR primers, and the expected lengths of PCR products containing *MYB* gene, **b**. and *WRKY* gene. Arrows and lines denote primers and PCR products of each primer pairs, respectively. The graph is built using the Microsoft Power Point software

TGK1006, Iran). The cleaned PCR products were cloned into pTG19- T vector (Vivantis, Malaysia) in accordance with the manufacturer's instructions, and transformed into *E. coli* strain DH5α. Plasmid DNA templates were purified using plasmid purification kit (Top Plasmid DNA Purification Kit, TOPAZ GENE RESEARCH., cat No.: TGK1008, Iran), and subjected to sequencing. The cDNA sequences of both *MYB*, and *WRKY* genes were determined using the ClustalW2 alignment analysis with their corresponding consensus sequences obtained from the aforementioned ESTs assembly.

Quantitative real-time PCR analysis

qRT-PCR was conducted with genes' specific primers designed based on the non-conserved region of gene sequences to ensure specific amplification (Table 1). The amplification fragments of both genes were 168 bp. qRT-PCR was performed using a BioRad system with the fluorescent dye SYBR®Green Master Mix 2X (Ampliqon, cat No.: A325402, Denmark) in accordance with the manufacturer's recommendation**.** The following program was used to perform qRT-PCR: 1 μL of cDNA as template, 1 μL (10 pmol. μl^{-1}) of each primer, 10 μL SYBR®Green Master Mix 2X in a total volume of 20 μL. qRT-PCR was run at 95 °C (15 min), 35 cycles at 95 °C (30 s), 51 °C (30), 72 °C (20 s), followed by gradient: 60–75 °C by 5 °C (5 s). Three technical replicates were performed per sample. For quantifying transcription levels, the reference gene *E.F.Iα* (GeneBank Accession No.: JN399225.1) was used as an internal control. Cycle thresholds (C(t)s) were analyzed using Livak method ($2^{-\Delta\Delta Ct}$) and relative expression levels were calculated using the Microsoft Excel software.

Promoter analysis and TFBS identification

The promoter sequences of *PSSDR1, PSCXE1, CYP82X1, CP82X2, PSAT1, PSMT2, CYP82Y1, PSMT3, CYP719A21* and *PSMT1* genes (accession numbers of JQ659007, JQ659006, JQ659002, JQ659004, JQ659008, JQ659000, JQ659005, JQ659001, JQ659003 and JQ658999, respectively) involved in noscapine biosynthesis pathway in poppy were analyzed. Since poppy is a non-model plant, its complete genome sequence is not available, therefore four BACs with accession numbers of JQ659009, JQ659010, JQ659011 and JQ659012 contained these 10 genes were obtained from the NCBI database. These BACs were merged to build up a BAC containing all 10 genes. Afterwards, pairwise sequence alignment between individual genes and the BAC was performed to locate each gene. Then, orientation, introns, exons, and start and end codons of these 10 genes were obtained from NCBI database. Subsequently, considering orientation of each gene, 1500 bp upstream of the TSS was isolated by version 7.0.4.1 of the BioEdit software (Hall 1999) as promoter for analysis. After promoter regions were isolated, cis-regulatory elements were found in silico using the MotifSampler program (Thijs et al. 2002). In order to identify overrepresented motifs in the cluster, the POBO program was employed. The web-based POBO program, which compares the frequency of a given motif in a set of promoter sequences with a random set of Arabidopsis promoter sequences, confirmed whether general consensus promoter motifs are statistically overrepresented in set of promoter sequences or not. (Kankainen and Holm 2004) (ekhidna.biocenter.helsinki.fi/poxo/pobo). After overrepresented motifs were identified, putative trans-regulatory elements which interact to the motifs were found computationally using the TRANSFAC (Wingender et al. 1996), PLACE (Higo et al. 1999) and Soft Berry (NSITE) (Solovyev et al. 2010) databases. To determine the number and type of TFBSs, the whole promoter sequence of each gene was also screened using TRANSFAC database.

Results

Identification, and characterization of the two TFs

Nested-PCR was performed using designed primers to amplify fragments containing the coding region of MYB and WRKY proteins. The full-length *MYB* and *WRKY* were cloned from *P. somniferum* L. using the nested-PCR method. As a result of nested-PCR, the lengths of amplified fragments were exactly the same as expectation from bioinformatics analysis (Fig. 2). The cloned cDNAs of *MYB* and *WRKY* were 1513 and 1330 bp, which contained ORFs of 1029 and 981 bp encoding 342 and 326 amino acids, with nuclear localization scores larger than 0.3, and 0.4 examined via the Predict NLS website, (Rost et al. 2004), respectively. These values of nuclear localization signal indicate their transcriptional regulatory nature (submitted to GenBank as accession numbers; *MYB*: KP411870 *WRKY*: KP203854). Multiple alignments between isolated MYB, and WRKY proteins and other MYB and WRKY proteins by the ClustalW2 v.2.1 showed that the MYB protein contains two imperfect tandem repeats, and therefore belongs to the R2R3-MYB subfamily, and the WRKY protein possesses one WRKY domain with a C2H2 zinc-finger motif, hence belongs to the group II WRKY family (Fig. 3). Phylogenetic analyses suggest that the MYB is more closely related to the D.myb-like.Zea, E.MYB.Zea, C.myb-related.Setaria and F.SsMYB7. *Saccharum* proteins, whereas the WRKY relates to the PsWRKY58 protein (Fig. 4).

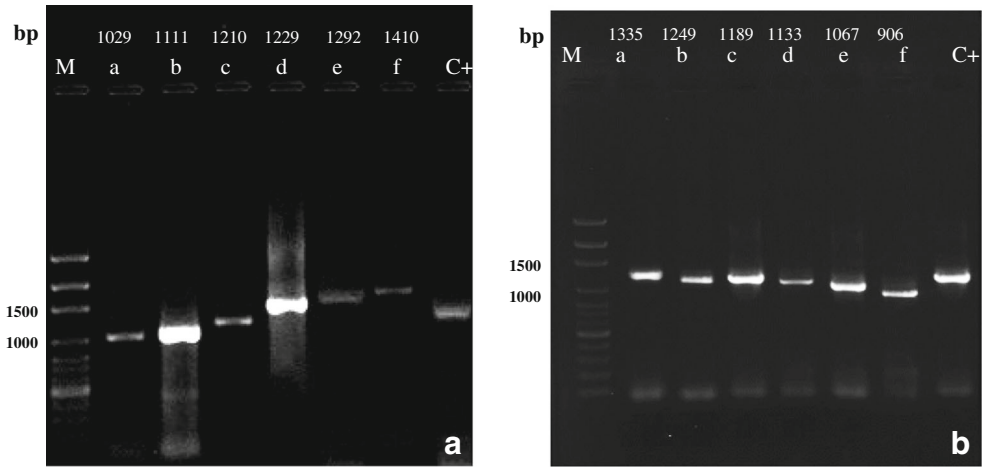

Fig. 2 Nested-PCR analysis of *P. somniferum* L. using **a**. a: forward.ORF and reverse.ORF, b: forward.1 and reverse.ORF, c: forward.2 and reverse.ORF, d: forward.ORF and reverse.1, e: forward.1 and reverse.1, f: forward.2 and reverse.1 of *MYB* gene. The target fragments containing *MYB* gene are 1029, 1111, 1210, 1229, 1292, 1410 bp, respectively. Positive control (C^+): *Arabidopsis thaliana* At5g43350 primers and **b**. a: forward.1and reverse.2, b: forward.ORF and reverse.2, c: forward.1 and reverse.1, d: forward.ORF and reverse.1, e: forward.1 and reverse.ORF and f: forward.ORF and reverse.ORF primers of *WRKY* gene. The target fragments containing *WRKY* gene are 1335, 1249, 1189, 1133, 1067 and 906 bp, respectively. Positive control (C^+): *Arabidopsis thaliana* At5g43350 primers. M: Ladder 100 bp

Transcriptome expression analyses of *MYB* and *WRKY* genes

Distribution of proteins can show their function in tissues (Shui et al. 2008). To determine whether *MYB* and *WRKY* transcripts are expressed in all tissues, mRNA levels of *MYB*, *WRKY* and *E.F.1α* were determined in various tissues of *P. somniferum* L. and *P. bracteatum* L. via the qRT-PCR method. Expression analysis of *P. somniferum* L. showed that transcription levels of *MYB* in leaves, capsules and stems were respectively 123, 24 and 4 times more than its transcription level in roots, and transcripts of *WRKY* in leaves, stems and capsules were respectively 406, 91 and 3 times more than its transcripts in roots. Therefore, the highest expression levels of both TFs were detected in leaves of *P. somniferum* L., and the lowest expression levels were detected in roots of *P. bracteatum* L. (Fig. 5a). The highest transcripts of MYB and WRKY TFs observed in the leaf organ have also been reported for various species like transcripts of IbMYB1 in *Ipomoea batatas* (Kim et al. 2013). SpWRKY2 which was isolated from *Solanum pimpinellifolium* also shows the highest expression level in leaves (Li et al. 2014). Unlike, *P. somniferum* L. the highest transcription levels of *MYB* and *WRKY* were detected in the roots of *P. bracteatum* L. The transcripts of *MYB* in roots, stems and leaves were respectively 24, 12 and 4 times more than its expression level in stems, and transcripts of *WRKY* in roots, leaves and capsules were respectively 94, 7 and 3 times more than its expression level in stems (Fig. 5b). Transcripts of GmMYB12B2 which was isolated from *Glycine max*, also shows the highest expression level in roots (Li et al. 2013). In some species such a *Brassica oleracea*, *Nicotiana tabacum*, *Arabidopsis thaliana and Oryza sativa* also the highest expression levels of most WRKY TFs were observed in root tissues (Hu et al. 2012; Liu et al. 2013; Nuruzzaman et al. 2014; Yao et al. 2014).

Comparison of transcript levels in both species, showed that expression levels of *MYB* in leaves, capsules and stems of *P. somniferum* L. were respectively 141, 135 and 1.9 times more than their expression in same tissues of *P. bracteatum* L., however, the expression level of *MYB* in roots of *P. bracteatum* L. was 4 times more than *P. somniferum* L. Transcripts of *WRKY* in stems, leaves and capsules of *P. somniferum* L. were respectively 323, 209 and 4 times more than transcripts in same tissues of *P. bracteatum* L., but the expression level of *WRKY* in the roots of *P. bracteatum* L. was 26 times more than *P. somniferum* L. (Fig. 5c).

Promoter analysis

Transcriptomic analysis of a high noscapine–producing poppy, bacterial artificial chromosome sequencing and F2 mapping analysis showed that 10-gene clustered involved in noscapine biosynthesis are co-expressed. Genes in the same cluster might be co-expressed but not necessarily co-regulated. Co-regulated genes have similar regulatory motifs. TFs are one of the major components in the complex molecular network of

Fig. 3 **a** The protein sequence of MYB TF was aligned with related MYB proteins from other plants. Proteins accession numbers are as follows: MYB-related308-*Beta vulgaris*: XP_010676421; MYB3-*Amboralla trichopoda*: XP_011629059; MYB-*Zea mays*: AFW57370; MYB-like-*Zea mays*: NP_001151578; MYB-related-*Setaria italic*: XP_004972853; ScMYB7-*Saccharum hybrid*: CCU64159; MYB86-like-*Brassica rapa*: XP_009112787; MYB103-*Arabidopsis thaliana*: NP_17657; Papaver-MYB: KP411870 and MYB2-*Pinus taeda*: ABD60283, **b**. The protein sequence of WRKY TF was aligned with related WRKY proteins from other plants. Proteins accession numbers are as follows: *Populus euphratica* WRKY40: XP_011020606.1; *Populus trichocarpa* WRKY40: EEE78831.2; *Jatropha curcas* hypothetical protein: KDP41972.1; *Theobroma cacao* WRKY40: EOY26053.1; *Morus notabilis* WRKY40: EXB50423.1; *Coffea canephora* unnamed protein: CDO98227.1; *Citrus clementina* hypothetical protein: ESR40862.1; *Nelumbo nucifera* WRKY40: XP_010275778.1; *Elaeis guineensis* WRKY40: XP_010929158.1; *P. somniferum* WRKY58: AFU81787.1 and *P. somniferum WRKY*: KP203840. The ClustalW2 program with default parameters was used for multiple alignment

gene expression regulation. Complete genome sequences and expression profiling technology have made the study of transcriptional regulation networks easier (Winzer et al. 2012). 28 motifs that were commonly shared by 100 % of genes were derived as result of promoter analysis by the MotifSampler (Table 2), and checked with the POBO program. There was a significant (t-test: *p*-value >0.0001) enrichment in the frequency of the all 28 consensus motifs identified by MotifSampler in 1.5 kb upstream promoter sequences of the selected genes compared to the *A. thaliana* background (Table 3). Graphs of some of motifs are presented in Fig. 6. There is a significant difference between background and cluster of genes in two motifs. Number of occurrences of the two motifs in cluster of genes is larger than background, which suggests that these motifs are tightly related to noscapine biosynthesis. As a result of TFBSs identification obtained using the TRANSFAC, PLACE and SoftBerry databases, multiple unique TFBSs of many known TFs in the promoter regions of these 10 co-expressed genes were found (Table 4). WRKY and MYB TFs are two of the main identified TF families.

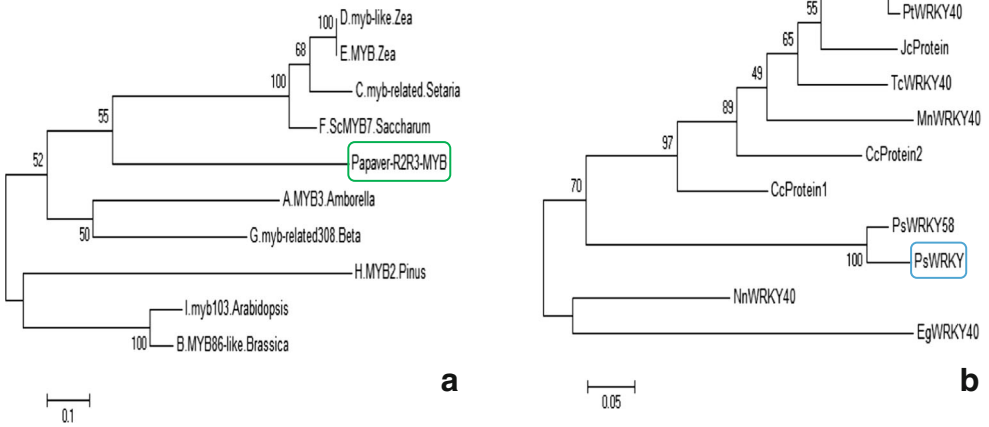

Fig. 4 **a**. Phylogenetic analysis of identified *P. somniferum* L. MYB protein, and other MYB proteins from various species using the Mega 6.06 software. Proteins accession numbers are as follows: MYB-related308-*Beta vulgaris*: XP_010676421; MYB3-*Amboralla trichopoda*: XP_011629059; MYB-*Zea mays*: AFW57370; MYB-like-*Zea mays*: NP_001151578; MYB-related-*Setaria italic*: XP_004972853; ScMYB7-*Saccharum hybrid*: CCU64159; MYB86-like-*Brassica rapa*: XP_009112787; MYB103-*Arabidopsis thaliana*: NP_17657; Papaver-MYB: KP411870 and MYB2-*Pinus taeda*: ABD60283, **b**. Phylogenetic analysis of identified *P. somniferum* L. WRKY protein, and other WRKY proteins from various species using the Mega 6.06 software. Proteins accession numbers are as follows: *Populus euphratica* WRKY40: XP_011020606.1; *Populus trichocarpa* WRKY40: EEE78831.2; *Jatropha curcas* hypothetical protein: KDP41972.1; *Theobroma cacao* WRKY40: EOY26053.1; *Morus notabilis* WRKY40: EXB50423.1; *Coffea canephora* unnamed protein: CDO98227.1; *Citrus clementina* hypothetical protein: ESR40862.1; *Nelumbo nucifera* WRKY40: XP_010275778.1; *Elaeis guineensis* WRKY40: XP_010929158.1; *P. somniferum* WRKY58: AFU81787.1 and *P. somniferum WRKY*: KP203840

As a result of TFBSs identification of promoters of noscapine gene cluster, 18 TFs that were commonly shared by 50 to 100 % of promoters were found using the TRANSFAC database (Table 5). According to Table 5, members of GT-1 TF family were found widely in all promoters. Occurrence of this family of TFs has widely been reported in different plant families such as *Solanaceae*. These TFs are similar to the SBF-1 TF. Unlike most TFs, these groups are interacting to the inhibitory cis elements in promoter sequences (Argüello-Astorga and Herrera-Estrella 1996). It has been reported that they play negative roles in gene expression regulation of most transgenic plants. GT-1, GT-1a and GT-1b TFs, which are presented in Table 5, are members of this family. PBF TFs were also occurred in all 10 genes. They are activator TFs that have been found in potato (Desveaux et al. 2000). In addition to these groups, some of TFs such as WRKY, MYB and BZIP were also found in most promoters. There are 2 WRKYs, and no MYB listed in the plant TFs database for *P. somniferum* L. (Eulgem et al. 2000). All promoters also carry a number of putative cis-acting regulatory elements involved in response to various external factors (Table 5). Results of two studies on promoter regulatory network of noscapine gene cluster including motifs and TFBSs identifications were similar. For instance, GT-1, MYB, WRKY, and PBF TFs were observed widely in both studies. Distributions of TFs on promoter sequences of 10 genes, identified by the TRANSFAC database were obtained, and the schematic view of them was shown in Fig. 7. The position of same TFBSs differs from one promoter to another, and there is not a regular pattern for their distribution. In fact, there is not a definite position for a specific TF on different promoters. For instance, SEF TF in *PSSDR1* promoter is located on position 82, whereas in the *CYP82Y1* promoter, it is located on position 1280.

Discussion

Two novel MYB and WRKY TFs were identified and cloned from *P. somniferum* L. using integrated bioinformatics and experimental data. *P. somniferum* L. was chosen for this study due to its valuable medicinal properties. Transcript levels of TFs were determined by means of qRT-PCR in capsules, leaves, roots, and stems of *P. somniferum* L. and *P. bracteatum* L. The two TFs were detected in all the studied tissues of both species. Transcripts of the two TFs in leaves of *P. somniferum* L. were much higher than other tissues, so in addition to their role in secondary metabolite biosynthesis, they might be involved in leaf developmental processes such as trichome development. The role of MYB TFs with a complex of other TFs including WRKY TFs, in trichome development of *Arabidopsis thaliana* has been reported (Du et al. 2009; Rushton et al. 2010). The highest expression level of both TFs in *P. bracteatum* L. was observed in roots. Unlike

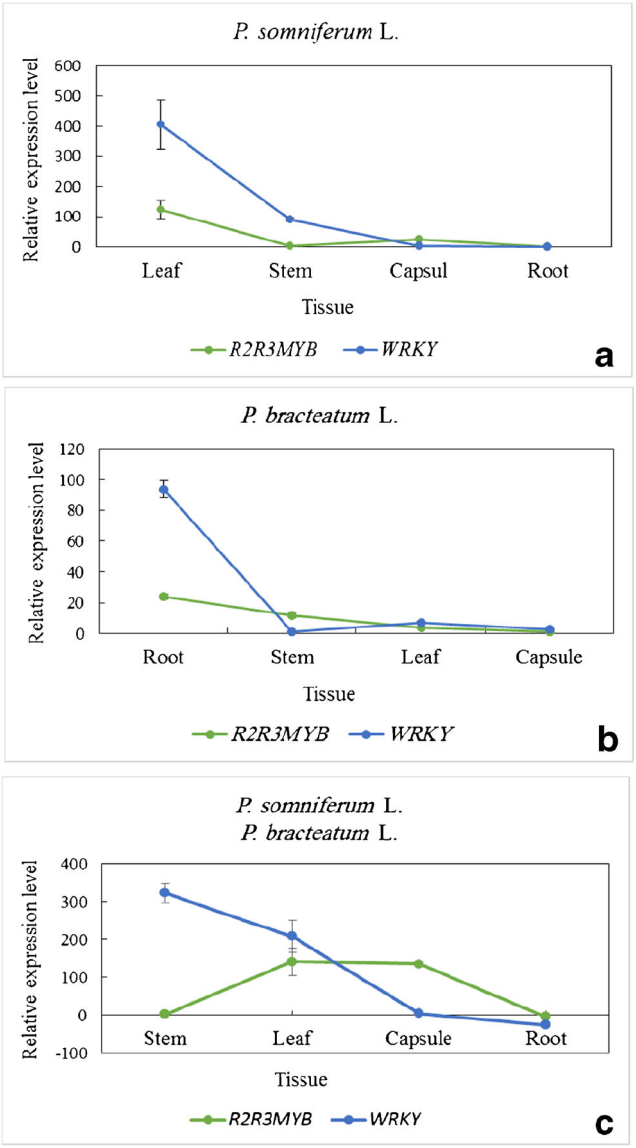

Fig. 5 **a**. Relative expression patterns of *MYB* and *WRKY* in leaf, stem and capsule tissues of *P. somniferum* L. in comparison with root tissue. **b.** Relative expression patterns of *MYB* in root, stem and leaf tissues of *P. bracteatum* L. in comparison with capsule tissue, and *WRKY* in root, leaf and capsule in comparison with stem tissue. **c**. Relative expression patterns of *MYB* and *WRKY* in leaf, stem, capsule and root tissues of *P. somniferum* L. in comparison with the same tissues of *P. bracteatum* L

Table 2 Consensus common motifs identified in promoter region of 10 gene cluster involved in noscapine biosynthesis from *P. somniferum* L. using MotifSampler program

Consensus motifs	cs[a]	ic[b]	ll[c]
GTGCACsAs	1.47	1.72	159.26
ACsAGAAGC	1.45	1.57	118.88
AnGTGGGGT	1.32	1.48	140.08
CTCGTGywC	1.51	1.71	188.76
GwrCACGAG	1.45	1.66	193.54
TsGTGCACs	1.41	1.65	210.55
kCTTTTTCT	1.46	1.37	159.65
AnGyTGTGG	1.23	1.39	133.80
GCAGGTnGG	1.31	1.69	143.23
CCCAACmwG	1.35	1.55	168.86
AnCTAGCnA	1.19	1.24	162.22
GCTTCTsGT	1.45	1.58	148.28
nTTTAGkGG	1.26	1.42	150.06
CwkGTTGGG	1.34	1.57	155.53
TyTCTTTTT	1.65	1.41	151.79
GnTGTGGTk	1.30	1.40	157.02
wTGnTCACC	1.31	1.33	166.96
ATATATATm	1.57	1.22	133.36
mwkCTCGTG	1.44	1.60	158.12
nCTACCnAT	1.29	1.26	156.34
sTsGTGCAC	1.36	1.63	205.46
CACGAGAAG	1.34	1.45	173.79
sGTGCACsA	1.48	1.73	185.48
CTTCTnGTG	1.43	1.61	130.24
GnwGTGGTk	1.31	1.44	136.77
wkTTATGGC	1.27	1.30	144.17
TTTCTTCwA	1.38	1.25	155.12
wTmTGGTTs	1.35	1.35	167.49

a Consensus Score/b Information Content Score/c LogLikelihood Score

P. somniferum L., which is an annual plant, *P. bracteatum* L. is a perennial plant with stable roots. The highest expression level of TFs in roots of *P. bracteatum* L. might be related to strong roots of *P. bracteatum* L. and their high metabolisms. There has been no report on MYB TFs in *P. somniferum* L. yet, and the identified *MYB* in this work was the first one submitted to the NCBI database, whereas a WRKY TF named WRKY58 was isolated from *P. somniferum* L. by Mishra et al. (2013). WRKY58 showed the highest transcript levels in capsules, followed by roots, stems, and buds. Multiple alignments and phylogenetic analysis showed that both WRKYs belong to the group II of WRKY family, suggested by their structural similarity, they might have close functions. WRKY58 is responsible for response to wounding, and also binds to *TYDC* promoter to regulate BIAs biosynthesis pathway (Mishra et al. 2013). As a result, identified *WRKY* might also be able to bind to *TYDC* promoter, so its overexpression might result in increasing of BIAs accumulation.

Noscapin gene cluster was selected for promoter analysis, since it is the only pathway of secondary metabolites in *P. somniferum* L., which its promoters are identified and available. Noscapine mostly accumulates in capsule and leaf tissues (Winzer et al. 2012; Mishra et al. 2013; Nuruzzaman et al. 2014). Results of promoter analysis showed that not only these genes are neighboring and co-expressed (Winzer et al. 2012),

Table 3 Frequency of occurrence of overrepresented promoter motifs of 10 gene cluster involved in noscapine biosynthesis from *P. somniferum* L. using POBO program

Consensus motif	Background			Cluster			*p*-value
	Cluster mean	Promoter mean	StdDev	Cluster mean	Promoter mean	StdDev	
GTGCACsAs	0.56	0.01	0.95	74.81	1.50	10.30	0.0001**
ACsAGAAGC	1.13	0.02	1.08	44.95	0.90	7.36	0.0001**
AnGTGGGGT	0.92	0.02	0.94	34.92	0.70	8.52	0.0001**
CTCGTGywC	0.93	0.02	0.97	75.42	1.51	10.58	0.0001**
GwrCACGAG	0.90	0.02	0.94	74.67	1.49	10.60	0.0001**
TsGTGCACs	0.51	0.01	0.91	74.33	1.49	10.62	0.0001**
kCTTTTTCT	9.70	0.19	3.18	49.55	0.99	6.86	0.0001**
AnGyTGTGG	3.90	0.08	2.07	40.00	0.80	5.27	0.0001**
GCAGGTnGG	0.66	0.01	0.81	45.13	0.90	7.01	0.0001**
CCCAACmwG	2.09	0.04	1.48	5.02	0.10	2.09	0.0001**
AnCTAGCnA	7.87	0.16	3.08	49.56	0.99	6.21	0.0001**
GCTTCTsGT	1.12	0.02	1.09	45.10	0.90	7.57	0.0001**
nTTTAGkGG	4.80	0.10	2.32	30.00	0.60	5.83	0.0001**
CwkGTTGGG	1.52	0.03	1.24	35.02	0.70	8.16	0.0001**
TyTCTTTTT	24.31	0.49	5.22	84.96	1.70	10.34	0.0001**
GnTGTGGTk	5.49	0.11	2.36	34.86	0.70	4.52	0.0001**
wTGnTCACC	3.92	0.08	2.04	29.99	0.60	4.65	0.0001**
ATATATATm	47.75	0.95	13.98	80.84	1.62	14.22	0.0001**
mwkCTCGTG	5.54	0.11	2.30	75.34	1.51	10.59	0.0001**
nCTACCnAT	4.86	0.10	2.27	75.90	1.52	25.22	0.0001**
sTsGTGCAC	0.59	0.01	0.94	74.64	1.49	10.40	0.0001**
CACGAGAAG	0.45	0.01	0.67	34.86	0.70	5.38	0.0001**
sGTGCACsA	0.51	0.01	0.86	74.86	1.50	10.71	0.0001**
CTTCTnGTG	3.01	0.06	1.74	45.32	0.91	6.13	0.0001**
GnwGTGGTk	9.78	0.20	3.20	60.32	1.21	8.47	0.0001**
wkTTATGGC	2.14	0.04	1.49	45.00	0.90	4.84	0.0001**
TTTCTTCwA	8.83	0.18	3.00	50.16	1.00	5.40	0.0001**
wTmTGGTTs	24.93	0.50	5.11	14.86	0.30	3.22	0.0001**

but also they can be co-regulated, and share common regulatory TFs such as MYB and WRKY protein families. Chen and Facchini have shown that the 10-gene clustered which is present in noscapine-accumulating opium poppy, is absent from the genome of noscapine-free chemotype. The presence or absence of transcripts corresponding to clustered genes is also observed. Tough, more information is needed, existing evidence suggests coordinated regulation (Chen and Facchini 2014). Suppression

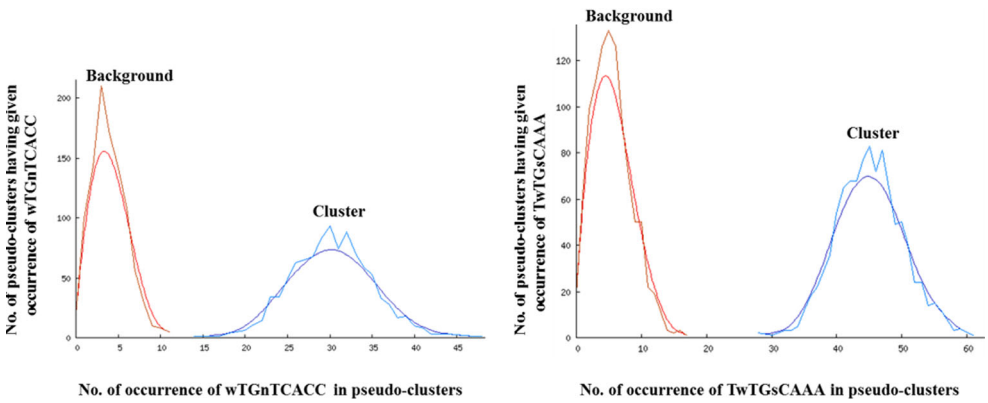

Fig. 6 Frequency of occurrence of the three motifs in artificial clusters generated in POBO for *A. thaliana* background promoters compared to the promoters of genes involved in noscapin biosynthesis. wTGnTCACC and TwTGsCAAA are motis that bind to MYB and WRKY TFs, respectively

Table 4 Putative trans/cis-acting regulatory elements presented in identified consensus motif patterns of promoters of noscapine biosynthesis pathway genes from *P. somniferum* L. by TRANSFAC, PLACE and SoftBerry databases

Consensus motif	Identifier	Position	Score	Predicted plant TF	Site binding sequence
GTGCACsAs	UPRE1AT	1 (+)	0.59	UPRE1AT	ATTGGTCCACG
	GLUTEBOX2OSGT2	3 (+)	0.59	GLUTEBOX2OSGT2	TCCGTGTACCA
	GLUTECOREOS	2 (+)	0.59	GLUTECOREOS	CTTTCGTGTAC
ACsAGAAGC	B2GMAUX28	1 (-)	0.65	B2GMAUX28	CTTGTCGTCA
	ANAERO2CONSENSUS	3 (-)	0.83	ANAERO2CONSENSUS	AGCAGC
AnGTGGGGT	RICE$EM_01	1 (+)	90.00	OSBZ8, TRAB1	ACGTG
CTCGTGywC	GLUTECOREOS	2 (+)	0.64	GLUTECOREOS	CTTTCGTGTAC
GwrCACGAG	GLUTECOREOS	2 (+)	0.64	GLUTECOREOS	CTTTCGTGTAC
kCTTTTTCT	NT$PR1A_03	3 (-)	100.00	GT-1	GAAAAA
AnGyTGTGG	V$AML1_01	4 (-)	0.83	V$AML1_01	TGTGGT
GCAGGTnGG	PV$PHASL_02	2 (-)	91.67	CAN	CACCTG
	AT$COR15A_01	5 (-)	90.00	ANT, CBF2	CCGAC
	AT$RD29A_01	5 (-)	90.00	DREB1A	CCGAC
CCCAACmwG	RBENTGA3	0	0.59	RBENTGA3	TCCAACTTGGA
AnCTAGCnA	**MYB1LEPR**	**2 (-)**	**0.75**	**MYB1LEPR**	**CTAACCA**
	MYBATRD22	**2 (-)**	**0.75**	**MYBATRD22**	
GCTTCTsGT	ELK4	0	0.78	ELK4	ACCGGAAGY
	V$NRF2_01	0	0.68	V$NRF2_01	ACCGGAAGNS
nTTTAGkGG	V$MZF1_02	1 (+)	0.50	V$MZF1_02	KNNNKAGGGGNAA
CwkGTTGGG	RBENTGA3	0	0.59	RBENTGA3	TCCAACTTGGA
TyTCTTTTT	CTRMCAMV35S	0	0.72	CTRMCAMV35S	TCTCTCTCT
	SURE1STPAT21	1 (-)	0.72	SURE1STPAT21	AATAGAAAA
TwTGsCAAA	**POT$PR10a_01**	**3 (-)**	**90.00**	**PBF**	**TGACA**
	AT$RLK4_01	**4 (-)**	**90.00**	**WRKY18**	**TTGAC**
GnTGTGGTk	GLUTECOREOS	2 (+)	0.64	GLUTECOREOS	CTTTCGTGTAC
wTGnTCACC	**NT$CHN50_01**	**3 (+)**	**90.00**	**WRKY1, WRKY3, WRKY4**	**GGTCA**
ATATATATm	I$CF2II_01	0	0.83	I$CF2II_01	RTATATRTA
	SORLREP3AT	0	0.83	SORLREP3AT	TGTATATAT
	I$CF2II_02	0	0.83	I$CF2II_02	GTATATATA
mwkCTCGTG	CELLCYCLESC	0	0.75	CELLCYCLESC	CACGAAAA
nCTACCnAT	**AMYBOX2**	**2 (-)**	**0.75**	**AMYBOX2**	**TATCCAT**
	MYBATRD22	**1 (-)**	**0.75**	**MYBATRD22**	**CTAACCA**
	NAPINMOTIFBN	**2 (-)**	**0.75**	**NAPINMOTIFBN**	**TACACAT**
sTsGTGCAC	UPRE1AT	1 (+)	0.59	UPRE1AT	ATTGGTCCACG
	GLUTEBOX2OSGT2	3 (+)	0.59	GLUTEBOX2OSGT2	TCCGTGTACCA
	GLUTECOREOS	2 (+)	0.59	GLUTECOREOS	CTTTCGTGTAC
CACGAGAAG	CELLCYCLESC	0	0.88	CELLCYCLESC	CACGAAAA
sGTGCACsA	UPRE1AT	2 (+)	0.64	UPRE1AT	ATTGGTCCACG
	GLUTEBOX2OSGT2	2 (+)	0.64	GLUTEBOX2OSGT2	TCCGTGTACCA
CTTCTnGTG	Broad-complex_3	2 (+)	0.61	Broad-complex_3	TAAACWARAAG
GnwGTGGTk	V$AML1_01	2 (-)	0.92	V$AML1_01	TGTGGT
wkTTATGGC	WUSATAg	1 (-)	0.79	WUSATAg	TTAATGG
TTTCTTCwA	TL1ATSAR	0	0.59	TL1ATSAR	CTGAAGAAGAA
	GLUTECOREOS	1 (+)	0.59	GLUTECOREOS	CTTTCGTGTAC
	RGATAOS	0	0.59	RGATAOS	CAGAAGATA
wTmTGGTTs	P$RAV1_01	2 (+)	0.54	P$RAV1_01	NNGCAACAKAWN

Table 5 Identification of TFs which occurred in 50 to 100 % of promoters of noscapine gene cluster in *P. somniferum* L. using TRANSFAC database

TF	*PSSDR1*	*PSCXE1*	*CYP82X1*	*CYP82X2*	*PSAT1*	*PSMT2*	*CYP82Y1*	*PSMT3*	*CYP719A21*	*PSMT1*
SEF	1	2	0	1	1	5	1	4	1	1
GT-1	10	11	7	8	18	12	14	18	6	10
PBF	4	3	5	3	1	1	3	2	3	2
WRKY	**6**	**9**	**9**	**3**	**3**	**6**	**8**	**3**	**5**	**9**
CAN	1	0	1	3	0	3	1	0	1	1
Alfin1	1	0	1	0	2	1	1	1	0	0
MYB	**2**	**3**	**1**	**2**	**2**	**6**	**4**	**3**	**4**	**4**
GT-2	2	1	2	0	4	3	5	0	1	1
ARF1	2	0	0	2	0	1	0	1	2	1
OSBZ8	1	0	0	0	0	1	0	2	2	1
WZF1	2	0	0	0	0	2	0	1	2	1
Opaque-2	1	2	0	0	0	0	1	1	2	2
ABI3	1	0	1	1	0	0	0	3	0	1
CBF	1	2	0	0	1	0	4	2	1	1
DPBF	0	3	0	1	0	4	0	0	1	1
ESBF I	0	1	2	0	4	0	1	0	1	0
ROM	0	1	0	0	0	0	1	1	1	1
Dof	0	1	0	0	1	3	4	4	0	1

of a gene involved in noscapine biosynthesis pathway (*CYP82Y1*) resulted in reduction of noscapin, and changes in expression level of noscapine producing genes. Expressions of genes which are out of the gene cluster, and encode enzymes involved in noscapine biosynthesis were not affected by the silencing of *CYP82Y1*, while expression of clustered genes was affected. Co-suppression of only clustered genes is also indicates coordinated regulation (Dang and Facchini 2014). Overexpression of *CjWRKY* raised expression level of genes involved in BAIs biosynthesis which indicates although TFs are not the only regulators of co-expression in gene clusters, undoubtedly they are key regulatory elements (Agarwal et al. 2015). Most of the investigated promoters contain effective stress responsive elements, which states in addition to their similar expression pattern in noscapine biosynthesis, they might also show similar pattern in expression of stress response genes.

Despite of the apparently crucial role of WRKY TFs in noscapine biosynthesis pathway in capsules which was observed as a result of promoter analysis, identified *WRKY* showed a low expression level in capsule tissue in comparison with leaf tissue, so it seems that the transcript level of *WRKY* in capsules should have been higher. Furthermore, transcripts of *MYB* in capsules were relatively higher in comparison with transcripts of *WRKY*, while their negative regulatory role in BIAs biosynthesis in *P. somniferum* L. has been reported. In fact, in *P. somniferum* L. there is a competition between BIAs and MIAs production. *STR* is the main enzyme in MIAs biosynthesis, and MYB TFs have positive role in its regulation. If expression of *MYB* is decreased, expression of *STR* will also be decreased. As a result, accumulation of MIAs will also be mitigated, and accumulation of BIAs, which they are more valuable alkaloids, will be increased (Zhou et al. 2010). Thereupon, overexpression of *WRKY* using capsule tissue specific promoters, and suppression of *MYB* might results in more BIAs and less MIAs accumulation in capsules. In a study conducted by Yang et al. (2015), the presence of two to four W-boxes in promoter region of several genes involved in artemisinin biosynthesis in *Artemisia annua* was reported. This study also suggests that AaWRKY is capable of enhancing expression of these genes (Yang et al. 2015). In another study on *Artemisia annua*, overexpression of *AaWRKY1* led to the increasing of artemisinin accumulation (1.8-fold)) Han et al. 2014).

It seems that genetic manipulation of TFs is a complicated, but more effective approach for controlling metabolites pathways compare to specific enzyme gene itself. Metabolic pathways in plants are very complicated and consist of different enzymatic steps (Gantet and Memelink 2002). Therefore, one of the main and practical applications of TFs is genetic engineering of these genes as valuable tools for plant manipulations. This will be only achieved by deep understanding of TFs genes, and regulatory networks of promoters (Grotewold et al. 1998).

Promoter is the central processor of transcription regulation. Understanding of how promoter networks are organized will shed light onto the control mechanism of specific genes. According to the promoter analysis method, co-regulatory

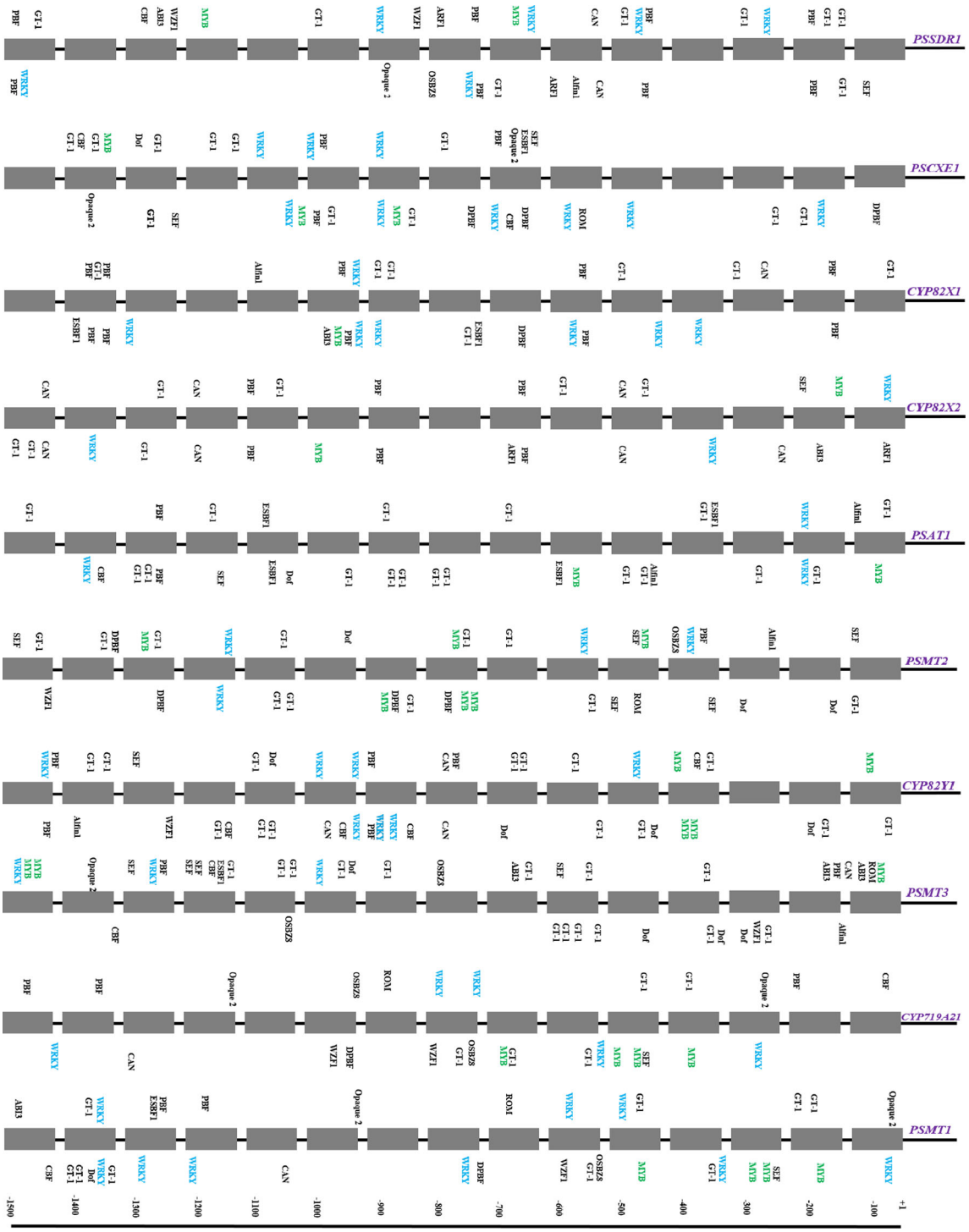

Fig. 7 Distributions of identified common TFs presented in promoter sequences of noscapine gene cluster by TRANSFAC database (Table 5). The graph is built using the Microsoft Power Point software

genes are categorized based on their promoter organization status. This is a powerful approach for description of co-regulated genes properties using their promoter structure. After a model or framework is developed, it may be used to search for possible co-regulatory genes. In this order, consideration of the difference between co-regulation and co-expression is crucial (Werener et al. 2003). Promoter structure Studies of co-regulate genes, and identification and isolation of common regulatory TFs are efficient strategies for breeding quantitative traits such as production of secondary

metabolites by engineering specific TFs, which is sometimes more effective than complicated enzymatic steps (Gantet and Memelink 2002; Yuan and Grotewold 2015).

Acknowledgments This work was funded by Tarbiat Modares University, Iran. We would like to acknowledge Mahdi Rezaei (University of Tehran, Iran) for providing tissues of *P. somniferum* and *P. bracteatum*.

Compliance with ethical standards

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

References

Agarwal P, Pathak S, Lakhwani D, Gupta P, Asif MH, Trivedi PK (2015) Comparative analysis of transcription factor gene families from *Papaver somniferum*: identification of regulatory factors involved in benzylisoquinoline alkaloid biosynthesis. Protoplasma. doi:10.1007/s00709-015-0848-8

Ando S, Obinata A, Takahashi H (2014) WRKY70 interacting with RCY1 disease resistance protein is required for resistance to cucumber mosaic virus in *Arabidopsis thaliana*. Physiol Mol Plant Pathol 85:8–14

Argüello-Astorga GR, Herrera-Estrella LR (1996) Ancestral multipartite units in light-responsive plant promoters have structural features correlating with specific phototransduction pathways. Plant Physiol 112:1151–1166

Chen N, Yang Q, Pan L, Chi X, Chen M, Hu D, Yang Z, Wang T, Wang M, Yu S (2014) Identification of 30 MYB transcription factor genes and analysis of their expression during abiotic stress in peanut (*Arachis hypogaea* L.). Gene 533:332–345

Chen X, Facchini PJ (2014) Short-chain dehydrogenase/reductase catalyzing the final step of noscapine biosynthesis is localized to laticifers in opium poppy. Plant J 77:173–184

Choe S, Kim S, Lee C, Yang W, Park Y, Choi H, Chung H, Lee D, Hwang BY (2011) Species identification of Papaver by metabolite profiling. Forensic Sci Int 211:51–60

Dang TT, Onoyovwi A, Farrow SC, Facchini PJ (2012) Biochemical genomics for gene discovery in benzylisoquinoline alkaloid biosynthesis in opium poppy and related species. Methods Enzymol 515:231–266

Dang TT, Facchini PJ (2014) CYP82Y1 is N-methylcanadine 1-hydroxylase, a key noscapine biosynthetic enzyme in opium poppy. J Biol Chem 289:2013–2026

Desveaux D, Després C, Joyeux A, Subramaniam R, Brisson N (2000) PBF-2 is a novel single-stranded DNA binding factor implicated in PR-10a gene activation in potato. Plant Cell 12:1477–1489

Du H, Zhang L, Liu L, Tang X, Yang W, Wu Y, Huang Y, Tang Y (2009) Biochemical and molecular characterization of plant MYB transcription factor family. Biochemistry 74:1–11

Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L (2010) MYB transcription factors in Arabidopsis. Trends Plant Sci 15:573–581

Eulgem T, Rushton PJ, Robatzek S, Somssich IE (2000) The WRKY superfamily of plant transcription factors. Trends Plant Sci 5:199–206

Gantet P, Memelink J (2002) Transcription factors tools to engineer the production of pharmacologically active plant metabolites. Trends Pharmacol Sci 23:563–569

Gomez-Gomez L, Trapero-Mozos A, Gomez MD, Rubio-Moraga A, Ahrazem O (2012) Identification and possible role of a MYB transcription factor from saffron (*Crocus sativus*). J Plant Physiol 169:509–515

Grotewold E, Chamberlin M, Snook M, Siame B, Butler L, Swenson J, Maddock S, St Clair G, Bowen B (1998) Engineering secondary metabolism in maize cells by ectopic expression of transcription factors. Plant Cell 10:721–740

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Han J, Wang H, Lundgren A, Brodelius PE (2014) Effects of overexpression of AaWRKY1 on artemisinin biosynthesis in transgenic *Artemisia annua* plants. Phytochemistry 102:89–96

Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database. Nucleic Acids Res 27:297–300

Holková I, Bezáková L, Bilka F, Balažová A, Vanko M, Blanáriková V (2010) Involvement of lipoxygenase in elicitor-stimulated sanguinarine accumulation in *Papaver somniferum* suspension cultures. Plant Physiol Biochem 48:887–892

Hu Y, Dong Q, Yu D (2012) Arabidopsis WRKY46 coordinates with WRKY70 and WRKY53 in basal resistance against pathogen *Pseudomonas syringae*. Plant Sci 185:288–297

Ishiguro S, Nakamura K (1994) Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5' upstream regions of genes coding for sporamin and beta-amylase from sweet potato. Mol Gen Genet 244:563–571

Jia L, Clegg MT, Jiang T (2004) Evolutionary dynamics of the DNA-bindingdomains in putative R2R3-MYB genes identified from rice subspecies indica and japonica genomes. Plant Physiol 134:575–585

Kankainen M, Holm L (2004) POBO, transcription factor binding site verification with bootstrapping. Nucleic Acids Res 32:W222.9

Kim JH, Nguyen NH, Jeong CY, Nguyen NT, Hong SW, Lee H (2013) Loss of the R2R3 MYB, AtMyb73, causes hyper-induction of the SOS1 and SOS3 genes in response to high salinity in *Arabidopsis*. J Plant Physiol 170:1461–1465

Li HL, Guo D, Yang ZP, Tang X, Peng SQ (2014) Genome-wide identification and characterization of WRKY gene family in *Hevea brasiliensis*. Genomics 104:14–23

Li J, Luan Y (2014) Molecular cloning and characterization of a pathogen-induced WRKY transcription factor gene from late blight resistant tomato varieties *Solanum pimpinellifolium* L3708. Physiol Mol Plant Pathol 87:25–31

Li JB, Luan YS, Jin H (2012) The tomato *SlWRKY* gene plays an important role in the regulation of defense responses in tobacco. Biochem Biophys Res Commun 427:671–676

Li XW, Li JW, Zhai Y, Zhao Y, Zhao X, Zhang HJ, Su LT, Wang Y, Wang QY (2013) A R2R3-MYB transcription factor, GmMYB12B2, affects the expression levels of flavonoid biosynthesis genes encoding key enzymes in transgenic Arabidopsis plants. Gene 532:72–9

Liu QL, Zhong M, Li S, Pan YZ, Jiang BB, Jia Y, Zhang HQ (2013) Overexpression of a chrysanthemum transcription factor gene, DgWRKY3, in tobacco enhances tolerance to salt stress. Plant Physiol Biochem 69:27–33

Mishra S, Triptahi V, Singh S, Phukan UJ, Gupta MM, Shanker K, Shukla RK (2013) Wound induced tanscriptional regulation of benzylisoquinoline pathway and characterization of wound inducible PsWRKY transcription factor from *Papaver somniferum*. PLoS One 8:e52784

Nuruzzaman M, Sharoni AM, Satoh K, Kumar A, Leung H, Kikuchi S (2014) Comparative transcriptome profiles of the WRKY gene family under control, hormone-treated, and drought conditions in near-isogenic rice lines reveal differential, tissue specific gene activation. J Plant Physiol 171:2–13

Paz-Ares J, Ghosal D, Wienand U, Peterson PA, Saedler H (1987) The regulatory c1 locus of Zea mays encodes a protein with homology to

myb proto-oncogeneproducts and with structural similarities to transcriptional activators. Embo j 6:3553–3558

Petroni K, Falasca G, Calvenzani V, Allegra D, Stolfi C, Fabrizi L, Tonelli C (2008) The *AtMYB11* gene from Arabidopsis is expressed in meristematic cells and modulates growth in plant and organogenesis *in vitro*. J Exp Bot 6:1201–1213

Qin Y, Tian Y, Hana L, Yang X (2013) Constitutive expression of a salinity-induced wheat WRKY transcription factor enhances salinity and ionic stress tolerance in transgenic *Arabidopsis thaliana*. Biochem Biophys Res Commun 441:476–481

Rabara RC, Tripathi P, Lin J, Rushton PJ (2013) Dehydration-induced WRKY genes from tobacco and soybean respond to jasmonic acid treatments in BY-2 cell culture. Biochem Biophys Res Commun 431:409–411

Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, Creelman R, Pilgrim M, Broun P, Zhang JZ, Ghandehari D, Sherman BK, Yu G (2000) Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science 290:2105–2110

Rost B, Yachdav G, Liu J (2004) The predict protein server. Nucleic Acids Res 32:321–326

Rushton PJ, Somssich IE, Ringler P, Shen QJ (2010) WRKY transcription factors. Trends Plant Sci 15:247–258

Shui Y, Guan ZB, Zhang JX, Zhang SQ (2008) CDNA cloning, genomic structure, expression analysis and biological activity of porcine a proliferation-inducing ligand (APRIL). Vet Immunol Immunopathol 125:190–197

Solovyev VV, Shahmuradov IA, Salamov AA (2010) Identification of promoter regions and regulatory sites. Methods Mol Biol 674:57–83

Stranska I, Skalicky M, Novak J, Matyasova E, Hejnak V (2013) Analysis of selected poppy (*Papaver somniferum* L.) cultivars: pharmaceutically important alkaloids. Ind Crop Prod 41:120–126

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729

Thijs G, Marchal K, Lescot M, Rombauts S, De Moor B, Rouzé P, Moreau Y (2002) A Gibbs sampling method to detect over-represented motifs in the upstream regions of co-expressed genes. J Comput Biol 9:447–464

Werener T, Fessele S, Maier P, Nelson J (2003) Computer modeling of promoter organization as a tool to study transcriptional coregulation. FASEB J 17:1228–1237

Wingender E, Dietze P, Karas H, Knüppel R (1996) TRANSFAC: a database on transcription factors and their DNA binding sites. Nucleic Acids Res 24:238–241

Winzer T, Gazda V, He Z, Kaminski F, Kern M, Larson TR, Graham IA (2012) A *Papaver somniferum* 10-gene cluster for synthesis of the anticancer alkaloid noscapine. Science 336:1704–1708

Xiong W, Xu X, Zhang L, Wu P, Chen Y, Li M, Jiang H, Wu G (2013) Genome-wide analysis of the WRKY gene family in physic nut (*Jatropha curcas* L.). Gene 524:124–132

Yang K, Rashidi Monafared S, Wang H, Lundgren A, Brodelius PE (2015) The activity of the artemisinic aldehyde Δ11(13) reductase promoter is important for artemisinin yield in different chemotypes of *Artemisia annua* L. Plant Mol Biol 88:325–340

Yao QY, Xia EH, Liu FH, Gao LZ (2014) Genome-wide identification and comparative expression analysis reveal a rapid expansion and functional divergence of duplicated genes in the WRKY gene family of cabbage, *Brassica oleracea* var capitata. Gene 557:35–42

Yuan L, Grotewold E (2015) Metabolic engineering to enhance the value of plants as green factories. Metab Eng 27:83–91

Zhou ML, Hou HL, Zhu XM, Shao JR, Wu YM, Tang YX (2010) Molecular regulation of terpenoidindole alkaloids pathway in the medicinal plant, *Catharanthus roseus*. Jmpr 4:2760–2772