Genetics and Molecular Research

# Genome-wide analysis of TCP family in tobacco

**L. Chen[1,2], Y.Q. Chen[3], A.M. Ding[1], H. Chen[1,2], F. Xia[1,2], W.F. Wang[1,2] and Y.H. Sun[1]**

[1]Key Laboratory for Tobacco Gene Resources,
Tobacco Research Institute of Chinese Academy of Agricultural Sciences,
Qingdao, China
[2]Graduate School of Chinese Academy of Agricultural Sciences, Beijing, China
[3]Kunming Tobacco Corporation of China National Tobacco Company,
Kunming, China

*These authors contributed equally to this study.
Corresponding author: W.F. Wang
E-mail: sdwangweifeng@163.com

Genet. Mol. Res. 15 (2): gmr.15027728
Received September 25, 2015
Accepted January 18, 2016
Published May 23, 2016
DOI http://dx.doi.org/10.4238/gmr.15027728

**ABSTRACT.** The TCP family is a transcription factor family, members of which are extensively involved in plant growth and development as well as in signal transduction in the response against many physiological and biochemical stimuli. In the present study, 61 *TCP* genes were identified in tobacco (*Nicotiana tabacum*) genome. Bioinformatic methods were employed for predicting and analyzing the gene structure, gene expression, phylogenetic analysis, and conserved domains of TCP proteins in tobacco. The 61 *NtTCP* genes were divided into three diverse groups, based on the division of *TCP* genes in tomato and *Arabidopsis*, and the results of the conserved domain and sequence analyses further confirmed the classification of the *NtTCP* genes. The expression pattern of *NtTCP* also demonstrated that majority of these genes play important roles in all the tissues, while some special genes exercise their functions only in specific tissues. In brief, the comprehensive and thorough study of the TCP family in other plants

 ©FUNPEC-RP www.funpecrp.com.br

provides sufficient resources for studying the structure and functions of TCPs in tobacco.

**Key words:** Tobacco; TCP transcription factor; Phylogenetic analysis; Gene expression pattern analysis

## INTRODUCTION

The TCP family is named after four different genes: *TEOSINTE BRANCHED 1*(*TB1*) from *Zea mays*, *CYCLOIDEA* (*CYC*) from *Antirrhinum majus*, and *PROLIFERATING CELL FACTORS 1* and *2* (*PCF1* and *PCF2*) from *Oryza sativa* (Cubas et al., 1999; Martín-Trillo and Cubas, 2010). The sequences of TCP proteins contain a highly conservative TCP domain, which consists of a non-canonical basic helix-loop-helix (bHLH) at the N-terminus (Luo et al., 1996; Cubas et al., 1999). The bHLH domain includes highly conserved amino acid residues. The Basic region is rich in basic amino acid residues. The two amphipathic helical regions are abundant in hydrophobic amino acid residues (Ala, Leu, and Trp), whereas the Loop region contains the hydrophilic amino acids residues (Gly), and all the conserved amino acid residues possibly contribute to the function of DNA binding, protein-protein interactions, and protein nuclear localization (Kosugi and Ohashi, 2002; Yao et al., 2007; Viola et al., 2011). Bioinformatics analysis indicated that the bHLH domain exists in all TCP proteins; only a few of the TCP proteins contain the R domain, which include polar residues and is assumed to develop a hydrophilic α-helix involved in protein-protein interactions (Yao et al., 2007). Based on the variation in the TCP domain, 61 NtTCP proteins can be classified into two groups, known as class I and class II. The class II group can be further divided into two dissimilar subfamilies, the CIN and the CYC/TB1.In plants, the TCP transcription factors (TFs) are specific plant TFs, which act as key regulators, widely involved in the regulation of plant growth and development. Among other things, the TCP transcription factors regulate embryonic growth, floral organ morphogenesis, pollen development, leaf development, seed germination, senescence, cell cycle regulation, circadian rhythm, and hormone signaling (Martín-Trillo and Cubas, 2010; Danismanet al., 2012; Uberti-Manassero et al., 2013). The TCP family proteins are studied in some dicotyledonous and monocotyledonous plants and all class I and class II proteins are reported to function as transcriptional activators and repressors in the control of plant growth and development (Uberti-Manassero et al., 2013; De Paolo et al., 2015). Although, the function of class II TCP proteins is well known, the data on the class I have been lacking. The class II TCP proteins are responsible for inhibiting cell growth and development in different organs. For instance, the research on *TEOSINTE BRANCHED 1* (*mTB1*) in maize revealed that the *mTB1* prevents the growth of shoot branching (Doebley et al., 1997), and the rice *TB1* gene (*OsTB1*), which was identified based on the sequence ofmTB1, served as the negative regulator in controlling the lateral branching in rice (Takeda et al., 2003). Most of *TCP* genes of class II in *Arabidopsis* were classified and defined; the *BRC1* (*AtTCP18*) and *BRC2* (*AtTCP12*) belonged to CYC/TB1 clade as the closest homologues of *TB1*, which is known to control the architecture of plants. It was shown that *BRC1* might take part in the development of axillary outgrowth by influencing the auxin and strigolactone (SL) synthesis pathways (Aguilar-Martínezetal., 2007).Furthermore, the study on the function of *BRANCHED1* in pea (*PsBRC1*) also proved that the gene can control shoot branching by affecting the synthesis of SL and cytokinins (Braun et al., 2012). *AtTCP1* was a member of

CYC/TB1 clade, which could affect the synthesis of brassinosteroids and elongation of the stem (Guo et al., 2010; Koyama et al., 2010). Members of the CIN clades are targeted by microRNA miR319, and they affect cell differentiation in *Arabidopsis* (Koyama et al., 2007; Zhou and Luo, 2014). Some members perform the function of negative regulators that affect the growth of leaf margin (Palatniket al., 2003). Compared to the class II proteins, very little is known about the function of class I proteins. *AtTCP15* participates in the auxin pathway (Lucero et al., 2015), *AtTCP14* (*At3g47620*) and *AtTCP15* (*At1g69690*) are two similar genes in class I, which have overlapping functions in the regulation of plant branching and meristem development (Kieffer et al., 2011). The two genes simultaneously regulate plant stature by promoting cell proliferation in young internodes (Koyama et al., 2007; Steiner et al., 2012). *AtTCP16* gene could affect the floral organ development. The mutants of *tcp16* in *Arabidopsis* are played variation in pollen grains, including size, shape, and staining pattern (Takeda et al., 2006). However, the data indicated that there were tremendous overlapping functions between the class I and II proteins in cell growth and division. Considering the effects of TCP transcription factors on plant growth and development and the importance of tobacco in tobacco-producing countries, research on tobacco TCP family is of great significance. Previous data showed that a large number of TCP family TFs from tomato (Parapunova et al., 2014) and *Arabidopsis* have been identified. Therefore, the present study was designed on the basis of all these factors. In this study, we systematically analyzed the 61 TCP transcription factors in tobacco using bioinformatics methods including the gene structure, phylogenetic analyses, and expression patterns of the putative *TCP* genes. We believe that the results from this study would be helpful in research and in understanding the function of *TCP* genes in tobacco.

## MATERIAL AND METHODS

### Identification of TCPs in tobacco

The genome sequences of tobacco were downloaded from the National Center for Biotechnology Information website (NCBI, http://www.ncbi.nlm.nih.gov/). Information on previously reported TCP proteins of *Arabidopsis thaliana* (AtTCP) and *Solanum lycopersicum* (SlTCP) were retrieved from *Arabidopsis* Information Resource (http://www.arabidopsis.org/) and Phytozome (http://www.phytozome.net/) (Goodstein et al., 2012), respectively. The domain analysis programs, Pfam (http://pfam.sanger.ac.uk/) and SMART (http://smart.embl-heidelberg.de/), were used to examine the protein sequences of all the candidate *TCP* genes of tobacco with default cut-off parameters (Letunic et al., 2012; Finn et al., 2014). Physical and chemical characteristics of all the *TCP* genes, including pIs (isoelectric points), extinction coefficient, molecular weights, and hydrophilicity were obtained with the help of proteomics and sequence analysis tools on the ExPASy Proteomics Server (http://expasy.org/; Artimo et al., 2012).

### Phylogenetic analysis and conserved motif identification

We performed multiple sequence alignment of 115TCP sequences, which consisted of 24 AtTCPs from *Arabidopsis thaliana*, 30 SlTCPs from tomato, and 61 NtTCPs from tobacco. Multiple-sequence alignments were carried out using the Cluster X (version1.83) program (Thompson et al., 1997). Phylogenetic analysis was performed using MEGA5.0 (Tamura et al., 2011; http://www.megasoftware.net) by neighbor-joining method (Saitou and Nei, 1987)

and the reliability of the phylogenetic trees obtained was examined using 1,000 bootstrap replicates. In addition, Maximum Likelihood, Minimal Evolution, and PhyML methods were also employed in the tree construction to verify the results. We identified the TCP domains and conserved sites using DNAMAN software (http://www.lynnon.com/). The intron and exon structures of the *TCP* genes were obtained by GSDS (Gene Structure Display Server http://gsds.cbi.pku.edu.cn/). MEME (Multiple EM for motif elicitation, http://meme.sdsc.edu/meme/) was used to search for the conserved motifs in all the TCP proteins (Bailey et al., 2006), with the optimum width from 8 to 50 and the maximum number of motifs 10.

## RNA extraction and cDNA synthesis

Total RNA from roots, stems, leaves, and flower was extracted from adult greenhouse-grown HonghuaDajinyuan tobacco variety, using the Trizol reagent kit (Invitrogen, Germany) according to the manufacturer instructions. The total RNA quantity and purity were assessed by a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). The first cDNA strand was generated by reverse transcription of 5 mg total RNA (50 μL reaction volume) using AMV reverse transcriptase (Takara Biotechnology, Japan), at 42°C for 1 h.

## Expression analysis of TCP in tobacco

The real-time RT-PCR was performed with a 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) according to the manufacturer protocols. Each reaction mixture contained 6 μL DNase/RNase free water, 10 μL 2X Real-Time SYBR Premix ExTaqII, 0.4 μL 50X ROX Reference Dye II, 2 μL diluted cDNA product from reverse transcription PCR, and 0.8 μL gene-specific primers. Three biological replicates for each tissue were used and each reaction was technically repeated three times. The thermal cycle employed was as follows: 95°C for 2 min followed by 40 cycles of 95°C for 10 s, 60°C for 30 s, and 60°C for 45 s. The fluorescence was measured at the end of each cycle. The expression data for the tobacco *TCP* genes were presented in relative units after their normalization with the actin gene expression using the $2^{-\Delta\Delta Ct}$ method. All the gene-specific primers were designed based on the sequences of the 22 TCP proteins (**S1 Table**).

# RESULTS

## Identification of TCP proteins in tobacco

We scored 61 candidates *TCP* genes from tobacco. Since there were no standard or principles to name the *NtTCP* genes, we named all the 61 *NtTCPs* from *NtTCP1* to *NtTCP61* according to the nomenclature system of *Arabidopsis* and sequence similarity with AtTCPs. All the 61 predicted TCP proteins were divided into three groups, which consisted of 26 in class I, 23 in CIN subclass, and 12 in CYC/TB1 subclass (Table 1). Physical and chemical characteristics of all the 61 TCP proteins of tobacco were analyzed. The amino acid in the 61 *TCP* genes ranged from 210 (*NtTCP11a/b*) to 815 (*NtTCP19d*) with an average 348 amino acids (**S2 Table**). Other physical and chemical characteristics of the 61 *NtTCP* genes, such as isoelectric point (pI), molecular weight (Mw) and hydropathicity, are listed in **S2 Table**. We observed that the isoelectric point of about 36% NtTCP proteins was less than 7; these proteins

were rich in acidic amino acids. The hydropathicity values of all the NtTCP proteins were less than zero, suggesting that the NtTCP proteins are hydrophilic. The values of Instability index for all the NtTCP proteins were higher than 40, with an average of 54.66, indicating that all the TCP transcription factors are unstable proteins. In general, all the NtTCP proteins had similar physical and chemical characteristics, however, slight differences were still observed. These differences might have been due to the different amino acid residues in non-conservative region.

**Table 1.** Phylogenetic analysis of the 61 *TCP* genes in *Nicotiana tabacum* genome.

| *TCP* subclass | Gene name | TCP domain | Homologous gene in *Arabidopsis* |
|---|---|---|---|
| CIN | *NtTCP2* | 59-209 | *AtTCP2* |
| | *NtTCP4a* | 17-129 | *AtTCP4* |
| | *NtTCP4b* | 21-146 | *AtTCP4* |
| | *NtTCP4c* | 21-134 | *AtTCP4* |
| | *NtTCP4d* | 168-293 | *AtTCP4* |
| | *NtTCP4e* | 29-155 | *AtTCP4* |
| | *NtTCP4f* | 21-118 | *AtTCP4* |
| | *NtTCP4g* | 29-151 | *AtTCP4* |
| | *NtTCP5a* | 52-133 | *AtTCP5* |
| | *NtTCP5b* | 52-149 | *AtTCP5* |
| | *NtTCP5c* | 51-148 | *AtTCP5* |
| | *NtTCP5d* | 52-149 | *AtTCP5* |
| | *NtTCP10a* | 111-245 | *AtTCP10* |
| | *NtTCP10b* | 93-194 | *AtTCP10* |
| | *NtTCP10c* | 93-157 | *AtTCP10* |
| | *NtTCP10d* | 93-157 | *AtTCP10* |
| | *NtTCP10e* | 113-256 | *AtTCP10* |
| | *NtTCP13a* | 51-196 | *AtTCP13* |
| | *NtTCP13b* | 51-197 | *AtTCP13* |
| | *NtTCP17a* | 51-156 | *AtTCP17* |
| | *NtTCP17b* | 55-140 | *AtTCP17* |
| | *NtTCP24a* | 61-213 | *AtTCP24* |
| | *NtTCP24b* | 58-150 | *AtTCP24* |
| CYC/TB1 | *NtTCP1a* | 90-214 | *AtTCP1* |
| | *NtTCP1b* | 90-212 | *AtTCP1* |
| | *NtTCP1c* | 82-225 | *AtTCP1* |
| | *NtTCP1d* | 83-220 | *AtTCP1* |
| | *NtTCP12a* | 122-221 | *AtTCP12* |
| | *NtTCP12b* | 110-230 | *AtTCP12* |
| | *NtTCP12c* | 123-267 | *AtTCP12* |
| | *NtTCP12d* | 131-251 | *AtTCP12* |
| | *NtTCP18a* | 102-250 | *AtTCP18* |
| | *NtTCP18b* | 123-281 | *AtTCP18* |
| | *NtTCP18c* | 121-278 | *AtTCP18* |
| | *NtTCP18d* | 102-243 | *AtTCP18* |
| Class I | *NtTCP7a* | 34-131 | *AtTCP7* |
| | *NtTCP7b* | 29-121 | *AtTCP7* |
| | *NtTCP7c* | 33-124 | *AtTCP7* |
| | *NtTCP7d* | 34-131 | *AtTCP7* |
| | *NtTCP8a* | 148-252 | *AtTCP8* |
| | *NtTCP8b* | 153-256 | *AtTCP8* |
| | *NtTCP8c* | 39-160 | *AtTCP8* |
| | *NtTCP9a* | 367-453 | *AtTCP9* |
| | *NtTCP9b* | 71-138 | *AtTCP9* |
| | *NtTCP11a* | 52-120 | *AtTCP11* |
| | *NtTCP11b* | 53-127 | *AtTCP11* |
| | *NtTCP15a* | 63-180 | *AtTCP15* |
| | *NtTCP15b* | 92-219 | *AtTCP15* |
| | *NtTCP15c* | 89-210 | *AtTCP15* |
| | *NtTCP15d* | 81-248 | *AtTCP15* |
| | *NtTCP15e* | 81-247 | *AtTCP15* |
| | *NtTCP15f* | 55-173 | *AtTCP15* |
| | *NtTCP19a* | 69-175 | *AtTCP19* |
| | *NtTCP19b* | 90-185 | *AtTCP19* |
| | *NtTCP19c* | 69-173 | *AtTCP19* |
| | *NtTCP19d* | 547-646 | *AtTCP19* |
| | *NtTCP20a* | 48-175 | *AtTCP20* |
| | *NtTCP20c* | 28-131 | *AtTCP20* |
| | *NtTCP20b* | 28-110 | *AtTCP20* |
| | *NtTCP23* | 114-210 | *AtTCP23* |
| | *NtTCP20d* | 48-129 | *AtTCP20* |
| Total | 61 | - | - |

 ©FUNPEC-RP www.funpecrp.com.br

**Table 2.** Summary of TCP transcription factors of the 13 species.

| Classification | Species | Number of TCP | Genome size (Mb) |
|---|---|---|---|
| Dicotyledoneae | *Nicotiana tabacum* | 61 | 4500.0 |
| | *Malus x domestica*[a] | 52 | 881.3 |
| | *Populus trichocarpa*[a] | 60 | 422.9 |
| | *Carica papaya*[a] | 22 | 135.0 |
| | *Theobroma cacao*[a] | 31 | 346.0 |
| | *Vitis vinifera*[a] | 15 | 487.0 |
| | *Cucumis sativus* | 22 | 203.0 |
| | *Glycine max* (L.) Merr. | 54 | 1126.4 |
| | *Gossypium raimondii* | 38 | 737.8 |
| | *Arabidopsis thaliana* | 24 | 135.0 |
| | *Solanum lycopersicum* | 30 | 760.0 |
| Monocotyledoneae | *Oryza sativa* | 23 | 372.0 |
| | *Zea mays* | 52 | 2500.0 |

[a]Woody plant.

## Phylogenetic relationships and gene structure

To analyze the evolutionary and phylogenetic relationships among the TCP transcription factor families in three different plants, we constructed the phylogenetic tree of the tomato, tobacco, and *Arabidopsis* using MEGA5.0. There were a total 115 TCPs, including 24 in *Arabidopsis* (AtTCP), 30 in tomato (SlTCP), and 61 in tobacco (Figure 1). Class I contained 52 members, representing 45.2% of the total TCPs, Class II which was divided into two subclasses, CIN and CYC/TB1 contained 63 members, representing 54.8% of the total TCPs. In Figure 2A, it is evident that the 61 NtTCPs possess the non-canonical bHLH domain, and some members of the CIN and CYC/TB1subclasses included the R-domain (Figure 2B).

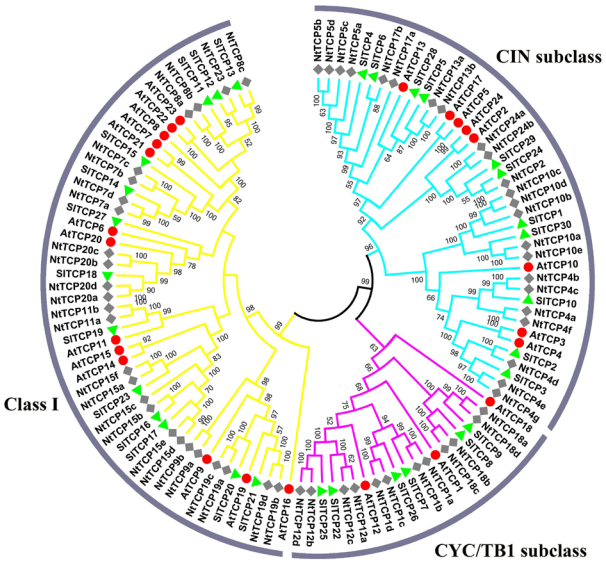


**Figure 1.** Phylogenetic tree of TCP transcription factor proteins in tobacco, *Arabidopsis*, and tomato constructed by the neighbor-joining method. The phylogenetic tree was constructed using MEGA5.0. The numbers are bootstrap values based on 1000 iterations. The three subclasses are indicated with different colors.

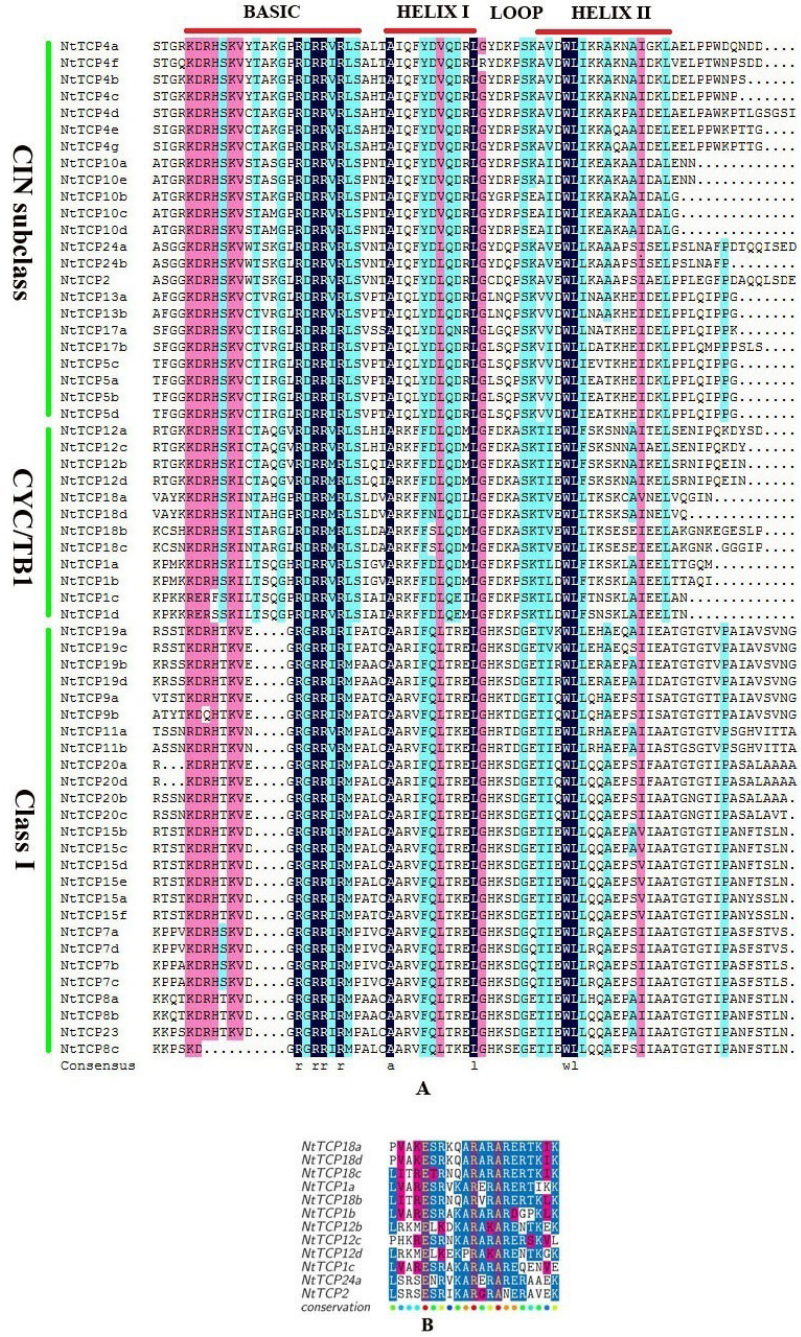


**Figure 2.** Multiple sequence alignment of conserved domain of the 61 *TCP* proteins of tobacco. **A.** Alignment of TCP domain of the 61 NtTCP proteins of tobacco. Amino acids that are conserved throughout are shaded in black. Conserved domains, including BASIC, HELIXI, LOOP and HELIXII are shown at the top. **B.** Alignment of R-domain of the class II subfamily members. Conserved amino acids in class II are shaded in different colors.

Moreover, we analyzed 484 TCP proteins from 13 species including tobacco (Table 2). The results of the comparative genomic analysis indicated that the number of TCP transcription factors in tobacco (61) was almost equal to that of *Populus trichocarpa* (60), *Malus* x *domestica* (52), *Glycine max* (L.) Merr (54), and *Zea mays* (52) and was higher than that of *Carica papaya* (22), *Theobroma cacao* (31), *Vitis vinifera* (15), *Cucumis sativus* (22), *Gossypium raimondii* (38), *Arabidopsis thaliana* (24), *Solanum lycopersicum* (30), and *Oryza sativa* (23). The genome size of tobacco (4500.0Mb) is more than that of all the other plants analyzed; it is about 10.6 times that of *populous trichocarpa* (422.9 Mb). However, there are many factors, such as the rapid genomic revolution and diploidization, which can affect genome size. Although the genome size of *V. vinifera* (487.0 Mb) was more than that of *P. trichocarpa* (60 Mb), *C. papaya* (135.0 Mb), *T. cocoa* (346.0 Mb), *C. sativus* (203.0 Mb), *A. thaliana* (135.0 Mb), and *O. sativa* (372.0 Mb), the number of TCP in *V.* vinifera (15) was the least. Therefore, we can infer that there is no direct correlation between the number of the TCP transcription factors and genome size, in both dicotyledons and monocotyledons, and it was same for woody as well as herbal plants. Earlier studies have shown that the recent whole-genome duplication (WGD) of tobacco has accelerated the expansion of the TCP family. Among the woody plants, *Malus x domestica* (Velasco et al. 2010) and *P. trichocarpa* (Tuskan et al., 2006) have also undergone the genome duplication events.

## Gene structure and conserved motifs

The preliminary analysis of the structures of the *NtTCP* genes was done using the gene structure display server 2.0 (GSDS). Phylogenetic tree (Figure 3A) revealed that most NtTCP proteins in the same group have similar genetic structure, including the length and number of exon. As shown in Figure 3B, most *NtTCP* genes (41 members, 67.2%) have one exon, 15 genes had two exons (24.6%), three had three exons (4.9%), and one had seven exons (1.6%), while *NtTCP9a* had eight exons (1.6%). In *Arabidopsis*, the numbers of exon ranged from one to four and 82% genes contained only one exon. Compared to *Arabidopsis*, *TCP* genes in tobacco exhibited gene structure diversification.

Based on the results of the MEME analyses, we deduced that two different groups had different conserved motifs. As shown in Figure 3C, eight conserved motifs were identified and named as motif 1 to motif 8 (**S1 Figure**), and the class I and class II had slightly disparate conserved motifs. Firstly, it can be seen that both groups possessed the motif 1 (Figure 3C), suggesting that motif 1 might be essential for the NtTCP proteins to perform their important function. The NtTCPs in subfamily CIN had motif3 and all members of the CYC/TB1 subfamily contained the motif 6, while the class I included motif 2. Thus, we inferred that the specific functions of the TCP proteins might depend on the specific conserved motifs.

## Analysis of the expression of the *TCP* genes

Based on the hierarchical clustering, we obtained the *NtTCP* genes expression profile in nine different plant organs such as dry capsule, root, stem, young flower, mature flower, senescent flower, young leaf, mature leaf, and senescent leaf (Figure 4). The color bar scale represented the gene expression levels. There were 32 *NtTCP* genes that were expressed in all the nine tissues, and the levels of the transcript of 18 *NtTCP* genes (*NtTCP9b*, *NtTCP12b*,

*NtTCP12d*, *NtTCP9a*, *NtTCP11b*, *NtTCP8c*, *NtTCP18a*, *NtTCP7c*, *NtTCP13a*, *NtTCP7a*, *NtTCP8a*, *NtTCP8b*, *NtTCP7b*, *NtTCP1b*, *NtTCP1c*, *NtTCP18d*, *NtTCP7d*, and *NtTCP2*) were very high in all of them. It was also noted that the expression pattern of same gene was different in different organs. For instance, *NtTCP12a* and*NtTCP12c* were detected in senescent flower, young leaves, and young flower while the expression of *NtTCP20a* was high in root and senescent flower.

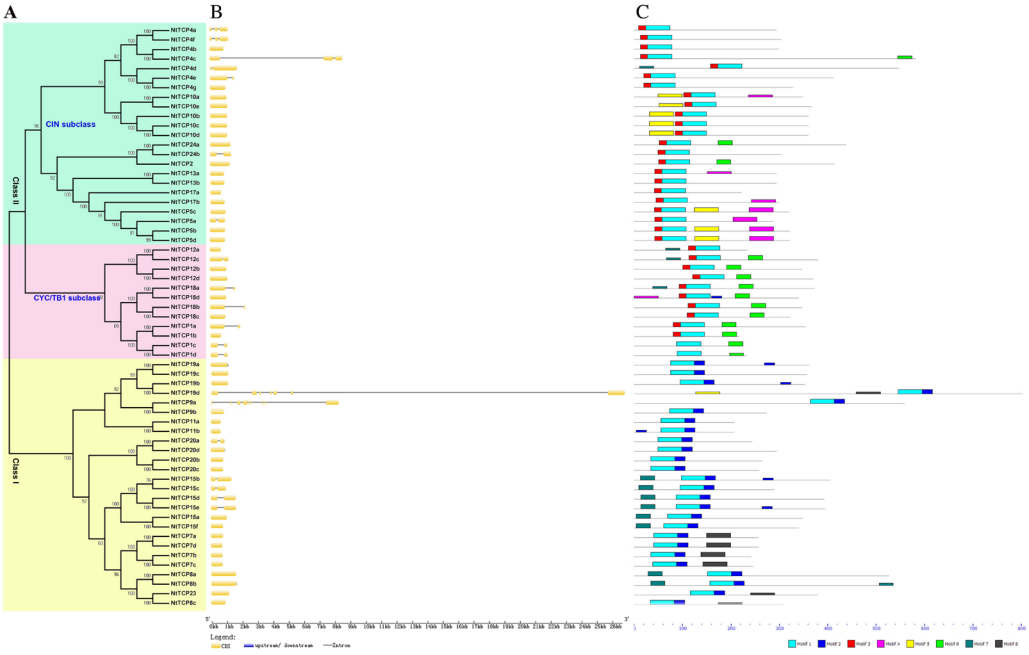


**Figure 3.** Phylogenetic analysis, gene structure, and conserved motifs of *TCP* family in tobacco. **A.** Phylogenetic tree of all *TCP* transcription factors in tobacco. **B.** Exon/intron organization of *TCP* genes of tobacco. Yellow boxes represent exons and black lines indicate introns. **C.** Conserved protein motifs in *TCP* family were identified using MEME program. Each motif is indicated with a specific color and named as motif 1-8.

Some *NtTCP* genes apparently exhibit tissue-specific expression; *NtTCP15d* was highly expressed only in young flower. We also observed that the expression of *NtTCP24a* was negative. This could be because *NtTCP24a* might have lost its function in the process of evolution and had spatial and temporal expression patterns. We also noted that even though *NtTCP13a* and *NtTCP7a* belonged to different groups, they shared similar expression patterns. This could be due to continual gene duplication and divergence, which affected the development of TCP family.

For the purpose of examining the transcript levels of 22 different NtTCPs in tobacco, we performed quantitative real time PCR (qRT-PCR) analysis. From Figure 5, we inferred that the expression levels of *NtTCP* genes exhibited variations in different organs, which mean that *NtTCP* genes play a variety of functions during the plant development. For example, *NtTCP7b*, *NtTCP8a*, *NtTCP11a*, *NtTCP15b*, *NtTCP19d*, *NtTCP20a,* and *NtTCP23*were present in all the tested tissues, implying that these genes play crucial roles related to plant

morphology. It was probable that the tissue-specific and other special conditions limited the expression of *NtCP9a* in roots (Figure 5). The genes *NtTCP4f*, *NtTCP5b*, *NtTCP13a,* and *NtTCP17b* shared identical expression patterns, and were expressed in both leaf and flower at very high levels. Furthermore, *NtTCP4d* and *NtTCP24b* shared the similar expression profile. In the CYC/TB1 subfamily, the expression of *NtTCP18b* and *NtTCP18c* was high in leaf while almost no expression was observed in root and flower.

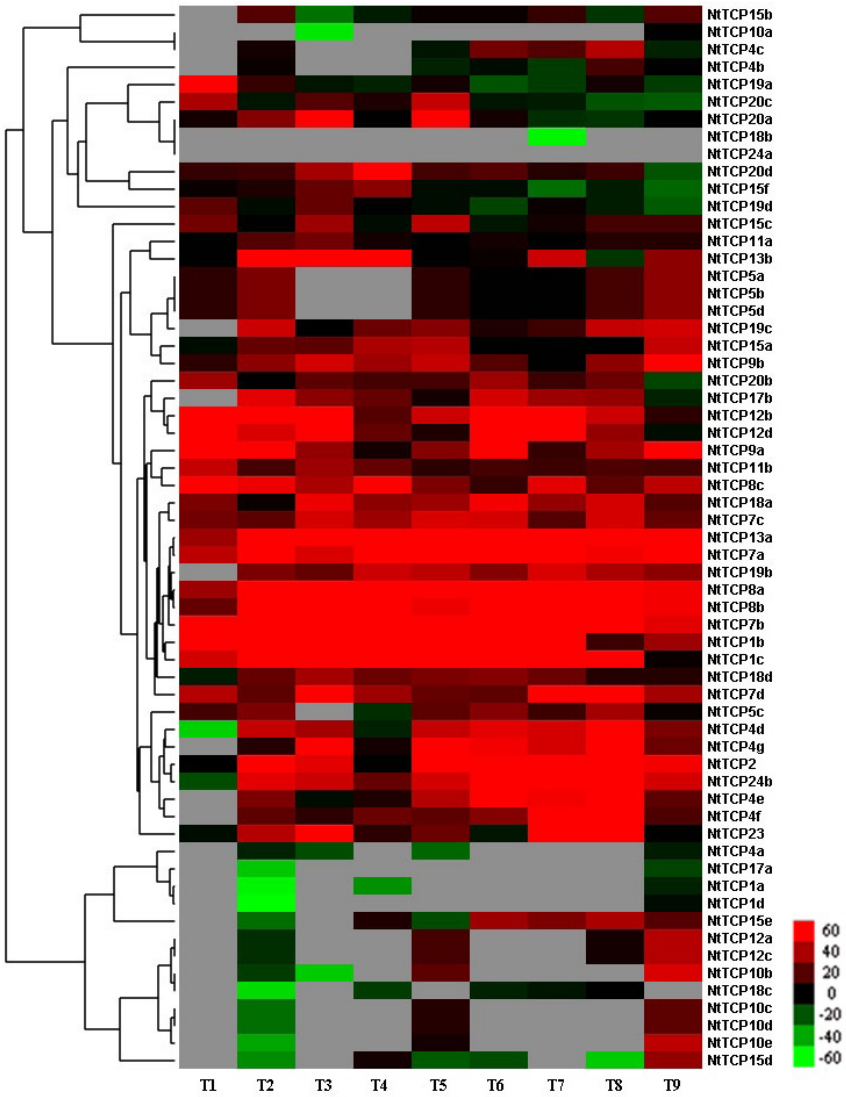


**Figure 4.** Heat map representation of the expression patterns of *NtTCP* genes, across different tissues. The color scale representing the relative signal values is shown at the bottom right (green refers to low expression; black refers to medium expression, red refers to high expression and gray refers to without expression). T1, dry capsule; T2, mature flower; T3, root; T4, stem; T5, senescent flower; T6, mature leaf; T7, senescent leaf; T8, young leaf; T9, young flower.

It is shown that 9.1 and 13.6% of *NtTCP* genes were highly expressed in root and stem, respectively. Most of *NtTCP* genes (72.7%) were detected in leaf, and only *NtTCP1c* gene was expressed only in flower. We found that in *Arabidopsis*, *AtTCP15* was expressed mainly in growing leaf tissues and the mutants with loss of function of *AtTCP15* genes resulted in altered leaf developmental traits (Kieffer et al., 2011). This led us to conclude that *NtTCP15* play important roles in the growth of leaves.

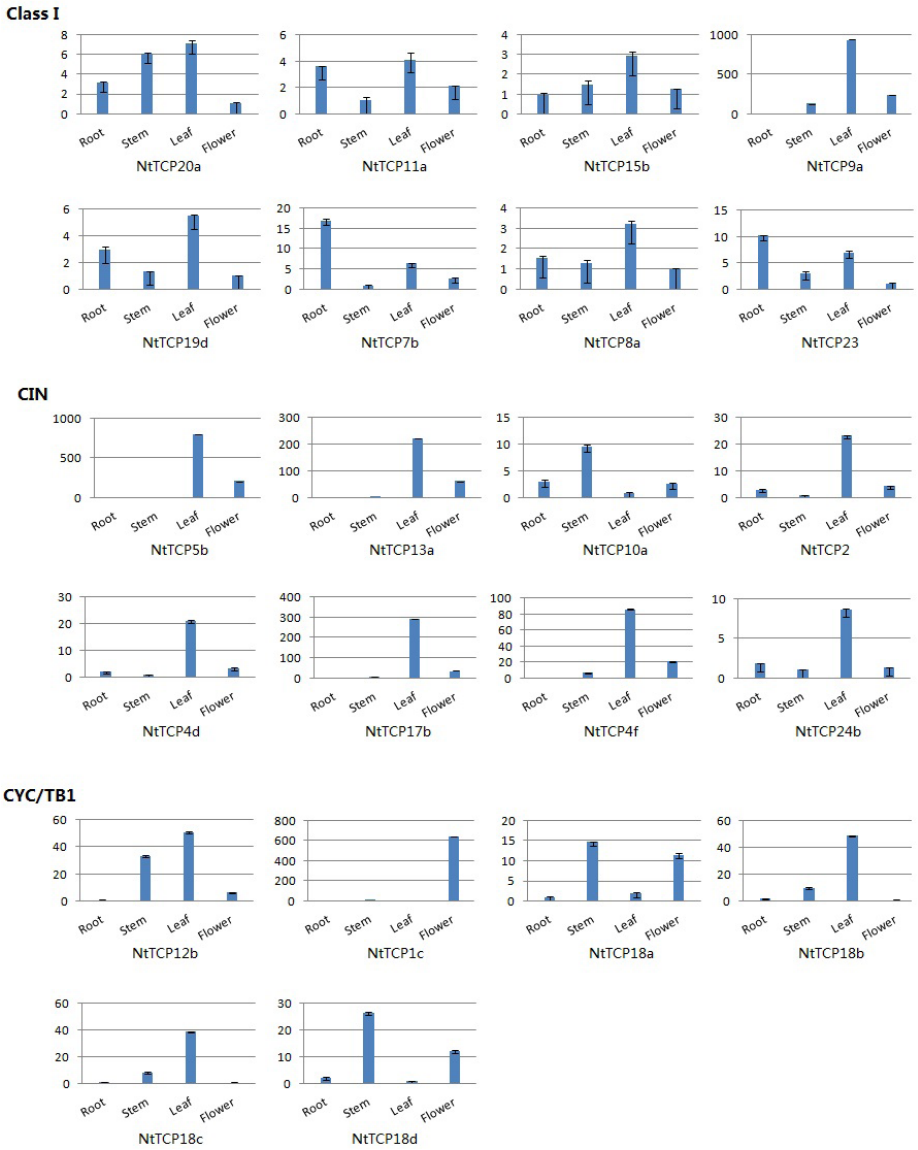


**Figure 5.** Expression levels of the selected *NtTCP* genes in tobacco tissues. The y-axis represents the relative expression levels of *NtTCP* genes against reference gene *NtL25*. Mean expression value was calculated from three independent replicates. The vertical bars indicate the standard deviation.

## DISCUSSION

Rapid development in bioinformatics tools facilitated the analysis of the TFs of *Arabidopsis*, poplar, and rice. Sixty-fourplant TF families are ascertained in *Arabidopsis* (DATF; Guo et al., 2005) and poplar (*Populus* spp; Zhu et al., 2007), while the rice (*Oryza sativa*) contains only 63 families (Gao et al., 2006) without the STERILE APETALA (SAP) family. Total 2576 TFs are confirmed in poplar, 1922 in *Arabidopsis*, 2025 in *indica* rice, and 2384 in *japonica* rice; the 64 plant TF families in tobacco together contain 2,513 TFs (Rushton et al., 2008). The large number of transcription factors in tobacco could be due to the gene duplication and inter specific hybridization. TCP proteins are one of the important types of plant transcription factors, and they play a pivotal role in the control of various aspects of plant development and growth.

Overall, 61 NtTCP proteins were detected in the tobacco genome, and the number of TCP proteins in tobacco is more than that in rice (22), *Arabidopsis* (24), tomato (30), and *G. raimondii* (38) (Martín-Trillo and Cubas, 2010; Ma et al., 2014). The analysis demonstrated that all the *TCP* genes in tobacco had two or more counterparts in *Arabidopsis* and tomato except *NtTCP2* and *NtTCP23*. Possibly, whole-genome duplication (WGD) events resulted in the expansion of NtTCP family. The TCP domain was found in all the 61NtTCP proteins, which were divided into two classes according to the differences in their domains. Expression analysis showed that the 22 selected genes were expressed at higher levels in at least one of the tested tissues, and the result obtained from the expression detection is consistent with that from the cluster analysis, indicating that the NtTCPs participate in the regulation of various aspects of plant growth and development in tobacco. Although slight differences were observed in the results of two different experiments, we concluded that there were some selected genes from TCP family, which were highly expressed in special tissues. Ample data indicated that TCP family proteins in the same group had similar function in *Arabidopsis*. For example, some members of the class I play an important role in promoting cell division, growth, and differentiation (Li et al., 2005). In contrast, class II genes act as negative regulators and control the cell growth and proliferation (Palatnik et al., 2003; Martin-Trillo and Cubas, 2010). In CYC/TB1 group from *Arabidopsis*, the *AtTCP12* and *AtTCP18* (renamed *BRC2* and *BRC1*, respectively) were involved in signals controlling branching; they were integrated within axillary buds (Aguilar-Martínez et al., 2007). Earlier studies have proven that the two *SlBRC1* (*SlTCP9* and *SlTCP7*) of tomato were the homologous genes of the *BRC* in *Arabidopsis* genes, There was high consistency between *SlBRC* and *AtBRC*, which mean that *SlBRC* could be involved in effecting the growth of the shoot branches. Besides, it is reported that *SlBRC1*is expressed highly in basal nodes (Martín-Trillo et al., 2011). Generally, the TCP family proteins belonging to the same group exhibited similar structures and functions. Because of the close relationship between tobacco and tomato, highly homologous genes between the two species were identified, and their functions in tomato were taken into account while predicting their functions in tobacco. The phylogenetic tree analysis showed that *AtTCP18* (*BRC1*), *SlTCP9* (*SlBRC1a*), and *NtTCP18* belong to the CYC/TB1 group, and the expression of *NtTCP18* was high in leaf, therefore we inferred that *NtTCP18* affected the development and growth of the leaf. In conclusion, in absence of concrete studies of the TCP transcription factor family in tobacco, this study will help in the future studies of NtTCP proteins in tobacco growth and development.

## Conflicts of interest

The authors declare no conflict of interest.

## ACKNOWLEDGMENTS

Research supported by the National high technology research and development program of China ("863" Program, #2012AA021801). The authors thank all the staff at the Department of Biotechnology of TRIC for their help in providing plant materials, giving technical assistance in bioinformatics, and providing advice on the manuscript.

## REFERENCES

Aguilar-Martínez JA, Poza-Carrión C and Cubas P (2007). *Arabidopsis BRANCHED1* acts as an integrator of branching signals within axillary buds. *Plant Cell* 19: 458-472. http://dx.doi.org/10.1105/tpc.106.048934

Artimo P, Jonnalagedda M, Arnold K, Baratin D, et al. (2012). ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* 40: W597-W603. http://dx.doi.org/10.1093/nar/gks400

Bailey TL, Williams N, Misleh C and Li WW (2006). MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* 34: W369-W373. http://dx.doi.org/10.1093/nar/gkl198

Braun N, de Saint Germain A, Pillot JP, Boutet-Mercey S, et al. (2012). The pea TCP transcription factor *PsBRC1* acts downstream of Strigolactones to control shoot branching. *Plant Physiol.* 158: 225-238. http://dx.doi.org/10.1104/pp.111.182725

Cubas P, Lauter N, Doebley J and Coen E (1999). The TCP domain: a motif found in proteins regulating plant growth and development. *Plant J.* 18: 215-222. http://dx.doi.org/10.1046/j.1365-313X.1999.00444.x

Danisman S, van der Wal F, Dhondt S, Waites R, et al. (2012). *Arabidopsis* class I and class II TCP transcription factors regulate jasmonic acid metabolism and leaf development antagonistically. *Plant Physiol.* 159: 1511-1523. http://dx.doi.org/10.1104/pp.112.200303

De Paolo S, Gaudio L and Aceto S (2015). Analysis of the *TCP* genes expressed in the inflorescence of the orchid *Orchis italica. Sci. Rep.* 5: 16265. http://dx.doi.org/10.1038/srep16265

Doebley J, Stec A and Hubbard L (1997). The evolution of apical dominance in *maize. Nature* 386: 485-488. http://dx.doi.org/10.1038/386485a0

Finn RD, Bateman A, Clements J, Coggill P, et al. (2014). Pfam: the protein families database. *Nucleic Acids Res.* 42: D222-D230. http://dx.doi.org/10.1093/nar/gkt1223

Gao G, Zhong Y, Guo A, Zhu Q, et al. (2006). DRTF: a database of rice transcription factors. *Bioinformatics* 22: 1286-1287. http://dx.doi.org/10.1093/bioinformatics/btl107

Goodstein DM, Shu S, Howson R, Neupane R, et al. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40: D1178-D1186. http://dx.doi.org/10.1093/nar/gkr944

Guo A, He K, Liu D, Bai S, et al. (2005). DATF: a database of *Arabidopsis* transcription factors. *Bioinformatics* 21: 2568-2569. http://dx.doi.org/10.1093/bioinformatics/bti334

Guo Z, Fujioka S, Blancaflor EB, Miao S, et al. (2010). *TCP1* modulates brassinosteroid biosynthesis by regulating the expression of the key biosynthetic gene DWARF4 in *Arabidopsis thaliana. Plant Cell* 22: 1161-1173. http://dx.doi.org/10.1105/tpc.109.069203

Kieffer M, Master V, Waites R and Davies B (2011). *TCP14* and *TCP15* affect internode length and leaf shape in *Arabidopsis. Plant J.* 68: 147-158. http://dx.doi.org/10.1111/j.1365-313X.2011.04674.x

Kosugi S and Ohashi Y (2002). DNA binding and dimerization specificity and potential targets for the TCP protein family. *Plant J.* 30: 337-348. http://dx.doi.org/10.1046/j.1365-313X.2002.01294.x

Koyama T, Furutani M, Tasaka M and Ohme-Takagi M (2007). TCP transcription factors control the morphology of shoot lateral organs via negative regulation of the expression of boundary-specific genes in *Arabidopsis. Plant Cell* 19: 473-484. http://dx.doi.org/10.1105/tpc.106.044792

Koyama T, Sato F and Ohme-Takagi M (2010). A role of *TCP1* in the longitudinal elongation of leaves in *Arabidopsis. Biosci. Biotechnol. Biochem.* 74: 2145-2147. http://dx.doi.org/10.1271/bbb.100442

Letunic I, Doerks T and Bork P (2012). SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res.* 40: D302-D305. http://dx.doi.org/10.1093/nar/gkr931

Li C, Potuschak T, Colón-Carmona A, Gutiérrez RA, et al. (2005). *Arabidopsis TCP20* links regulation of growth and cell division control pathways. *Proc. Natl. Acad. Sci. USA* 102: 12978-12983. http://dx.doi.org/10.1073/pnas.0504039102
Lucero LE, Uberti-Manassero NG, Arce AL, Colombatti F, et al. (2015). *TCP15* modulates cytokinin and auxin responses during gynoecium development in *Arabidopsis*. *Plant J.* 84: 267-282. http://dx.doi.org/10.1111/tpj.12992
Luo D, Carpenter R, Vincent C, Copsey L, et al. (1996). Origin of floral asymmetry in *Antirrhinum*. *Nature* 383: 794-799. http://dx.doi.org/10.1038/383794a0
Ma J, Wang Q, Sun R, Xie F, et al. (2014). Genome-wide identification and expression analysis of TCP transcription factors in *Gossypium raimondii*. *Sci. Rep.* 4: 6645. http://dx.doi.org/10.1038/srep06645
Martín-Trillo M and Cubas P (2010). *TCP* genes: a family snapshot ten years later. *Trends Plant Sci.* 15: 31-39. http://dx.doi.org/10.1016/j.tplants.2009.11.003
Martín-Trillo M, Grandío EG, Serra F, Marcel F, et al. (2011). Role of tomato BRANCHED1-like genes in the control of shoot branching. *Plant J.* 67: 701-714. http://dx.doi.org/10.1111/j.1365-313X.2011.04629.x
Palatnik JF, Allen E, Wu X, Schommer C, et al. (2003). Control of leaf morphogenesis by microRNAs. *Nature* 425: 257-263. http://dx.doi.org/10.1038/nature01958
Parapunova V, Busscher M, Busscher-Lange J, Lammers M, et al. (2014). Identification, cloning and characterization of the tomato TCP transcription factor family. *BMC Plant Biol.* 14: 157. http://dx.doi.org/10.1186/1471-2229-14-157
Rushton PJ, Bokowiec MT, Laudeman TW, Brannock JF, et al. (2008). TOBFAC: the database of tobacco transcription factors. *BMC Bioinformatics* 9: 53. http://dx.doi.org/10.1186/1471-2105-9-53
Saitou N and Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406-425.
Steiner E, Yanai O, Efroni I, Ori N, et al. (2012). Class I TCPs modulate cytokinin-induced branching and meristematic activity in tomato. *Plant Signal. Behav.* 7: 807-810. http://dx.doi.org/10.4161/psb.20606
Takeda T, Suwa Y, Suzuki M, Kitano H, et al. (2003). The *OsTB1* gene negatively regulates lateral branching in rice. *Plant J.* 33: 513-520. http://dx.doi.org/10.1046/j.1365-313X.2003.01648.x
Takeda T, Amano K, Ohto MA, Nakamura K, et al. (2006). RNA interference of the *Arabidopsis* putative transcription factor TCP16 gene results in abortion of early pollen development. *Plant Mol. Biol.* 61: 165-177. http://dx.doi.org/10.1007/s11103-006-6265-9
Tamura K, Peterson D, Peterson N, Stecher G, et al. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28: 2731-2739. http://dx.doi.org/10.1093/molbev/msr121
Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, et al. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25: 4876-4882. http://dx.doi.org/10.1093/nar/25.24.4876
Tuskan GA, Difazio S, Jansson S, Bohlmann J, et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (*Torr. & Gray*). *Science* 313: 1596-1604. http://dx.doi.org/10.1126/science.1128691
Uberti Manassero NG, Viola IL, Welchen E and Gonzalez DH (2013). *TCP* transcription factors: architectures of plant form. *Biomol. Concepts* 4: 111-112.
Velasco R, Zharkikh A, Affourtit J, Dhingra A, et al. (2010). The genome of the domesticated apple (*Malus* × *domestica Borkh.*). *Nat. Genet.* 42: 833-839. http://dx.doi.org/10.1038/ng.654
Viola IL, Uberti Manassero NG, Ripoll R and Gonzalez DH (2011). The *Arabidopsis* class I TCP transcription factor AtTCP11 is a developmental regulator with distinct DNA-binding properties due to the presence of a threonine residue at position 15 of the TCP domain. *Biochem. J.* 435: 143-155. http://dx.doi.org/10.1042/BJ20101019
Yao X, Ma H, Wang J and Zhang D (2007). Genome-wide comparative analysis and expression pattern of *TCP* gene families in *Arabidopsis thaliana* and *Oryza sativa*. *J. Integr. Plant Biol.* 49: 885-897. http://dx.doi.org/10.1111/j.1744-7909.2007.00509.x
Zhou M and Luo H (2014). Role of microRNA319 in creeping bentgrass salinity and drought stress response. *Plant Signal. Behav.* 9: e28700. http://dx.doi.org/10.4161/psb.28700
Zhu QH, Guo AY, Gao G, Zhong YF, et al. (2007). DPTF: a database of poplar transcription factors. *Bioinformatics* 23: 1307-1308. http://dx.doi.org/10.1093/bioinformatics/btm113

## Supplementary material

**S1 Table.** Sequences of the primers used for qRT-PCR analysis.

**S2 Table.** Characteristics of the 61 *TCP* genes from tobacco.

**S1 Figure.** Conserved motifs of NtTCP protein.