RESEARCH ARTICLE

Genomewide analysis of *TCP* transcription factor gene family in *Malus domestica*

RUIRUI XU[1], PENG SUN[2], FENGJUAN JIA[2], LONGTAO LU[3], YUANYUAN LI[2], SHIZHONG ZHANG[2]* and JINGUANG HUANG[2]*

[1]*Key Laboratory of Biology and Molecular Biology in University of Shandong, Weifang University, Weifang, Shandong 261061, People's Republic of China*
[2]*State Key Laboratory of Crop Biology, Shandong Agricultural University, Taian, Shandong 271018, People's Republic of China*
[3]*Weifang Traditional Chinese Medicine Hospital, Clinical Laboratory, Weifang, Shandong 261061, People's Republic of China*

Abstract

Teosinte branched1/cycloidea/proliferating cell factor1 (TCP) proteins are a large family of transcriptional regulators in angiosperms. They are involved in various biological processes, including development and plant metabolism pathways. In this study, a total of 52 *TCP* genes were identified in apple (*Malus domestica*) genome. Bioinformatic methods were employed to predicate and analyse their relevant gene classification, gene structure, chromosome location, sequence alignment and conserved domains of MdTCP proteins. Expression analysis from microarray data showed that the expression levels of 28 and 51 *MdTCP* genes changed during the ripening and rootstock–scion interaction processes, respectively. The expression patterns of 12 selected *MdTCP* genes were analysed in different tissues and in response to abiotic stresses. All of the selected genes were detected in at least one of the tissues tested, and most of them were modulated by adverse treatments indicating that the *MdTCP*s were involved in various developmental and physiological processes. To the best of our knowledge, this is the first study of a genomewide analysis of apple *TCP* gene family. These results provide valuable information for studies on functions of the *TCP* transcription factor genes in apple.

[Xu R., Sun P., Jia F., Lu L., Li Y., Zhang S. and Huang J. 2014 Genomewide analysis of *TCP* transcription factor gene family in *Malus domestica*. *J. Genet.* **93**, 733–746]

Introduction

Transcription factors (TFs) contain distinct types of DNA-binding domains and transcriptional regulation regions, and are capable of activating or repressing the transcription rates of multiple target genes (Riechmann *et al.* 2000; Wray *et al.* 2003). TFs are important regulators of diverse cellular processes and the complexity of living organisms necessitates a large number of TFs. Teosinte branched1/cycloidea/ proliferating cell factor1 (TCP) proteins constitute one of the largest families of plant-specific TFs (Martin-Trillo and Cubas 2010). The *TCP* gene family was named after the TCP domain from the first identified members: *TB1* (teosinte branched1 from maize), *CYC* (cycloidea from *Antirrhinum*) and *PCFs* (PCF proteins from rice) (Doebley *et al.* 1997; Kosugi and Ohashi 1997; Cubas *et al.* 1999; Luo *et al.* 1999). Sequence comparison of TCP proteins identified two conserved regions, including a TCP domain with a noncanonical basic helix-loop-helix (bHLH) structure and an R domain (Doebley *et al.* 1997; Kosugi and Ohashi 1997; Cubas *et al.* 1999; Luo *et al.* 1999). Although the bHLH domain is present in all TCP proteins, only some of the *TCP* genes had the R domain (Brameier 2010).

Recently, many TCPs were found to have important roles in regulating various developmental processes, such as branching, floral organ morphogenesis and leaf growth (Aguilar-Martinez *et al.* 2007; Nag *et al.* 2009; Martin-Trillo and Cubas 2010; Danisman *et al.* 2012). *OsTB1*, the rice homologue of the maize *TB1* gene, was considered as a negative regulator of lateral branching (Takeda *et al.* 2003). Functional analysis of the *TCP* genes *BRC1* and *BRC2* in

*For correspondence. E-mail: Jinguang Huang, jghuang@sdau.edu.cn; Shizhong Zhang, shizhongsdau@gmail.com.

Keywords. TCP; transcription factor; gene family; *Malus domestica*.

Arabidopsis demonstrated that these genes were involved in suppressing axillary bud outgrowth (Aguilar-Martinez *et al.* 2007). *TCP2*, *TCP3*, *TCP4*, *TCP10* and *TCP24* in *Arabidopsis* were all targeted by miR319, and have been implicated in regulating leaf morphogenesis (Palatnik *et al.* 2003). A group of AS2-binding TCP TFs (TCP3, TCP4, TCP10 and TCP24) were identified as regulators in leaf development by binding directly to the promoters of *BP* and *KNAT2* to repress their expression (Li *et al.* 2012a). Reporter gene analysis and the use of SRDX fusions suggested that *AtTCP14* and *AtTCP15* modulate cell proliferation in the developing leaf blade and specific floral tissues (Hiratsu *et al.* 2003; Kieffer *et al.* 2011). In addition, *AtTCP15* played an important role in regulating endoreduplication during *Arabidopsis* development (Li *et al.* 2012b). In *Arabidopsis*, *TCP16* functioned in pollen development, including variations in size, shape and staining patterns, suggesting that *TCP16* was important for early pollen development (Takeda *et al.* 2006). *AtTCP20*, acting upstream of *AtTCP9*, served as a pivotal link between the regulation of growth and cell division pathways, and controlled leaf development via the jasmonate signalling pathway (Li *et al.* 2005; Aguilar-Martinez *et al.* 2007; Nag *et al.* 2009; Martin-Trillo and Cubas 2010; Danisman *et al.* 2012). *LjCYC2*, a *CYC* homologue, was shown to function in establishing dorsal identity together with Keeled wings in Lotus 1 (*Kew1*), which regulated the control of lateral petal identity, suggesting a common molecular origin for the mechanisms controlling floral zygomorphy in *Lotus japonicus* (Feng *et al.* 2006). Two CYC-like TCP proteins were characterized in the genetic control of floral zygomorphy in *Pisum sativum* L. (Wang *et al.* 2008).

Evolution of the *TCP* gene family in the Asteridae suggested that gene duplication, followed by functional divergence, might be the mechanism responsible for regulatory gene family diversification and its impact on morphological evolution in the *Lamiales* (Reeves and Olmstead 2003). To date, various members of the *TCP* family from *Arabidopsis* (23 members) and rice (22 members) have been identified, which can be divided into three classes by phylogenetic analysis. Expression pattern analyses of the *AtTCP*s and *OsTCP*s in stems, leaves and flowers showed that approximately half of the genes (22 *TCP* genes) were expressed in all the three tissues tested, suggesting that TCPs may play regulatory roles at multiple developmental stages in *Arabidopsis* and rice (Brameier 2010) In contrast to the intensive research on *TCP* in model and crop plants such as *Arabidopsis* and rice, there are very limited reports on apple. Recently, the draft genome sequence of apple has been decoded, which provided an excellent opportunity for genomewide analyses of all the genes belonging to specific gene families (Velasco *et al.* 2010). The genomewide analysis of the *RING finger*, *DREB*, *dehydrin* and *Hsf* gene families have been reported in apple (Li *et al.* 2011; Giorno *et al.* 2012; Liang *et al.* 2012; Zhao *et al.* 2012). However, no genomewide information on the apple *TCP* gene family is currently available.

Given the importance of *TCP*s in diverse biological and physiological processes and their potential application for the development of stress-tolerant transgenic plants, a systematic analysis of the apple *TCP* gene family was carried out for the first time. The chromosomal location and gene structure of the putative *TCP* genes were analysed carefully. In addition, the *TCP*s were subjected to phylogenetic analyses with their *Arabidopsis* counterparts. These comparisons enabled the identification of gene orthologues and clusters of orthologous groups that can be subjected to further functional characterization. Further, we analysed the expression patterns using microarray and expressed sequence tag (EST) data. To our knowledge, this is the first genomewide analysis of the apple *TCP* family, which provides valuable information for understanding the classification and putative functions of TCPs. Ultimately, these findings will lead to potential applications for the improvement of apples via genetic engineering.

Materials and methods

Identification of TCPs in apple

To identify members of the *TCP* gene family, multiple database searches were performed. The *Arabidopsis*, rice and *Populus TCP* sequences were used as queries to perform repetitive blast searches against the GDR database (Jung *et al.* 2014). Additionally, all proteins sequences were then used as queries to perform multiple database searches against proteome and genome files downloaded from GDR database. Stand-alone versions of BLASTP and TBLASTN (basic local alignment search tool, http://blast.ncbi.nlm.nih.gov) available from NCBI (National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov) were used with the e-value cutoff set to 1e-003 (Mount 2007). Moreover, the predicted *TCP* gene family sequences were downloaded from the Apple GFDB database (Apple gene function and gene family database, http://www.applegene.org/) (Yao *et al.* 2011). All protein sequences derived from the candidate *TCP* genes were examined using the domain analysis programmes, Pfam (protein family, http://pfam.sanger.ac.uk/) and SMART (simple modular architecture research tool, http://smart.embl-heidelberg.de/) with the default cutoff parameters (Letunic *et al.* 2012; Finn *et al.* 2014) The pIs (isoelectric points) and molecular weights of the TCP TFs were obtained with the help of proteomics and sequence analysis tools on the ExPASy proteomics server (http://expasy.org/).

Sequence alignment and phylogenetic analysis of TCP TFs in apple

TCP sequences were aligned using the program ClustalX with BLOSUM30 as the protein weight matrix. The MUSCLE (multiple sequence comparison by log-expectation) program (ver. 3.52) was also used to perform multiple sequence alignments to confirm the ClustalX data output

(http://www.clustal.org/) (Edgar 2004). Phylogenetic trees based on the protein sequences of the MdTCPs were constructed using the NJ (neighbour-joining) method of the program MEGA5.0 (molecular evolutionary genetics analysis) with p distance and using the complete deletion option parameters engaged (Tamura *et al.* 2011). The reliability of the trees obtained was tested using bootstrapping with 1000 replicates. Images of the phylogenetic trees were also drawn using MEGA5.0.

Chromosomal location and gene structure of MdTCPs

The chromosomal locations and gene structures were retrieved from the apple genome data that were downloaded from the GDR database. The remaining genes were selected using a Perl-based program and mapped to the chromosomes with MapDraw (Liu and Meng 2003), further the gene structures of the *MdTCP*s were generated with the GSDS (gene structure display server, http://gsds.cbi.pku.edu.cn/).

Expression analysis of TCP TFs in microarray

The microarray data of gene expression in apple fruits during fruit ripening process was downloaded from the gene expression omnibus database using the GSE series accession number GSE24523. The sequences of the identified MdTCP-containing genes were used as queries to blast against probe sequence (GPL11164) to find corresponding unigene IDs used in microarray data. The microarray data during rootstock–scion interactions process (GSE4762) was also downloaded from the gene expression omnibus database. And MdTCP-containing genes were used as queries to blast against probe platform (GPL3715) to find corresponding unigene IDs used in microarray data. Phylogenetic analysis was performed to determine the corresponding unigene IDs when sequences of high similarity were acquired. The microarray data were made into a database by Perl-based programs and then clustered using Cluster3.0 with Euclidean distances and the hierarchical cluster method of complete linkage clustering. The clustering tree was constructed and viewed in Java Treeview.

RNA extraction and cDNA synthesis

The young leaves and other tissues of the *M. hupehensis* (an excellent apple rootstock widely used for grafting in China) were used to determine the expression patterns of the apple *TCP* genes. The *M. hupehensis* trees were 12 years of age and planted in the experimental orchard of Shandong Institute of Fruit Tree Science (Taian, China).

The total RNA was extracted using the PureLink^{TM} RNA mini kit (Invitrogen, Carlsbad, USA) and treated with RNase-free DNase I. Two micrograms of the total RNA was used to synthesize the first-strand cDNA using the PrimeScript First Strand cDNA synthesis kit (Takara, Dalian, China).

Quantitative real-time PCR analysis

The qRT-PCR reaction was performed in 25 μL volumes containing 10 μM of each primer (table 1), 50 ng of cDNA and 12.5 μL of SYBR Premix Ex *Taq* II. The PCR amplification conditions included an initial heat denaturing step at 95°C for 3 min and then 40 cycles of 95°C for 20 s, 56°C for 20 s and 72°C for 20 s. The fluorescence was measured at the end of each cycle. A melting-curve analysis was performed by heating the PCR product from 55 to 95°C. The expression data for the apple *TCP* genes were presented as relative units after their normalization to the apple actin gene using the $2^{-\Delta\Delta CT}$ method. The qRT-PCR experiment was carried out at least three times under identical conditions using actin as an internal control. Details of primers are listed in table 1.

Results

Identification of TCP TFs in apple

To identify *TCP* genes from the apple genome, BLASTP searches of the entire apple genome database (GDR,

Table 1. Primers used for the qRT-PCR analysis.

Gene name	Forward primer 5′–3′	Reverse primer 5′–3′
MdTCP3	TCGTGTCCTGATGCTGGCTAT	TGGGATGACACGAACATGGTG
MdTCP4	TAGAACATCAGGGCCGCCAC	CGATGATAGAGGGCTCGGCATG
MdTCP8	CACCACGGACATCACAGTTCG	GTGCTGCTGATGGTGGTGGTG
MdTCP9	CACCACGGACATCACAGTTCG	GTGGTGGTGGTGATTCTCCATG
MdTCP10	GCAGCAGCACCGATGGATTACT	GTGATGAGGCATAAGAGACGGTG
MdTCP16	CCGTGGAGTGGTTGCTCATTC	CTCTCTTTGGCCGAGGGTTTG
MdTCP25	CAGAGGACTTGATGGTGGCTT	AGAGGAAGTGGAAGCCGTTCG
MdTCP31	GAACACTAGGGACCGCCATAC	TAGCGACTTGAGCGGTTGAAC
MdTCP33	GAGCAAGGGCAAGAGAGAGA	TCTGCACCACCACAATTACTAC
MdTCP40	GAGAGAAAGATTGTCCACAGGC	GCCACCATTGAATCAGGCTG
MdTCP45	CAAGAGCGCCATCAAGGAC	TCTGCCCACAATTTTCTGACC
MdTCP49	CTGCTTCCTCCGTTGACGAT	TACAAAGCCCCCATCAAGTCC
MdActin	TGACCGAATGAGCAAGGAAATTACT	TACTCAGCTTTGGCAATCCACATC

Genome database for Rosaceae, http://www.rosaceae.org) using well-studied plant (*Arabidopsis*, rice and *Populus*) TCPs as queries were first performed. The hidden Markov model (HMM) of the simple modular architecture research tool (SMART) and Pfam tool were then exploited as query to confirm the putative *TCP* genes. Finally, 52 typical *TCP* genes containing full open reading frame (ORF) were identified. Later, these genes were analysed manually using InterProScan (Jones *et al.* 2014) and ClustalX program to confirm the presence of a TCP domain. We provisionally named them *MdTCP1* through *MdTCP52*, based on their chromosomal locations (table 2).

Table 2. TCP transcription factors in apple.

Gene number	Gene accession no.	CDS	Exon number	Size/aa	Molecular weight/D (MW)	Isoelectric point (pI)	Chromosome location	*Arabidopsis* homologous gene
MdTCP1	MDP0000123919	1047	1	348	35970.18	6.97	chr1: 25131292..25132388	*AtTCP19*
MdTCP2	MDP0000594000	516	4	171	19521.38	8.39	chr2: 34636835..34640797	*AtTCP13*
MdTCP3	MDP0000681033	573	2	190	21190.64	10.65	chr4: 17051470..17055503	*AtTCP15*
MdTCP4	MDP0000393985	594	2	197	22003.1	7.17	chr4: 17076789..17078318	*AtTCP15*
MdTCP5	MDP0000182310	723	1	240	27586.5	6.31	chr5: 4983945..4984667	*AtTCP10*
MdTCP6	MDP0000259723	723	1	240	27586.5	6.31	chr5: 4986196..4986918	*AtTCP10*
MdTCP7	MDP0000534647	1155	6	384	42809.65	9.27	chr5: 6310689..6314838	*AtTCP13*
MdTCP8	MDP0000927314	1464	1	487	53092.36	8.65	chr5: 6897651..6899111	*AtTCP2*
MdTCP9	MDP0000920127	1464	1	487	53141.05	8.85	chr5: 6905767..6907227	*AtTCP2*
MdTCP10	MDP0000763497	1464	1	487	53141.05	8.85	chr5: 6906611..6908071	*AtTCP2*
MdTCP11	MDP0000287069	1455	1	484	52876.67	7.27	chr5: 7572414..7573868	*AtTCP2*
MdTCP12	MDP0000280252	1008	3	335	34901.43	5.58	chr6: 13199161..13200294	*AtTCP9*
MdTCP13	MDP0000877369	1281	1	426	45449.99	6.98	chr6: 20909349..20910626	*AtTCP14*
MdTCP14	MDP0000531313	1281	1	426	45425.74	6.98	chr6: 20930056..20931333	*AtTCP14*
MdTCP15	MDP0000120671	1125	1	374	41686.22	9.02	chr6: 21954058..21955182	*AtTCP5*
MdTCP16	MDP0000219838	1617	3	538	60965.41	9.28	chr6: 22600713..22602619	*AtTCP18*
MdTCP17	MDP0000199422	702	2	233	26244.63	9.6	chr7: 7533528..7535742	*AtTCP13*
MdTCP18	MDP0000260056	1089	1	362	37483.89	5.99	chr7: 23199732..23200820	*AtTCP19*
MdTCP19	MDP0000617746	447	1	148	16597.89	9.16	chr8: 24098713..24099156	*AtTCP13*
MdTCP20	MDP0000253526	444	1	147	16597.89	9.16	chr8: 24107900..24108343	*AtTCP13*
MdTCP21	MDP0000319266	1827	3	608	63870.51	8.87	chr8: 28824646..28827556	*AtTCP8*
MdTCP22	MDP0000523096	969	1	322	34308.18	8.67	chr9: 626759..627724	*AtTCP20*
MdTCP23	MDP0000130524	444	1	147	16680.19	10.08	chr9: 1402850..1403293	*AtTCP13*
MdTCP24	MDP0000692406	1155	1	384	42111.01	7.96	chr9: 4205989..4207140	*AtTCP13*
MdTCP25	MDP0000442611	1053	1	350	38217.12	6.64	chr9: 32452747..32453796	*AtTCP4*
MdTCP26	MDP0000264920	1134	2	377	39069.02	6.18	chr10: 4137787..4139257	*AtTCP9*
MdTCP27	MDP0000184743	1839	3	612	69007.24	7.36	chr10: 6942026..6944154	*AtTCP4*
MdTCP28	MDP0000238683	864	1	287	30365.65	8.86	chr10: 17761157..17762020	*AtTCP7*
MdTCP29	MDP0000189749	984	2	327	36374.5	8.13	chr10: 27694805..27696702	*AtTCP10*
MdTCP30	MDP0000243495	1797	3	598	65728.91	6.55	chr11: 29676059..29679200	*AtTCP4*
MdTCP31	MDP0000139807	669	1	222	24323.61	6.66	chr12: 15276397..15277065	*AtTCP14*
MdTCP32	MDP0000535805	909	1	302	32225.71	5.41	chr12: 19508030..19508935	*AtTCP9*
MdTCP33	MDP0000173048	1434	1	477	53214.59	8.96	chr13: 2901807..2903240	*AtTCP12*
MdTCP34	MDP0000242185	1200	1	399	42547.06	6.73	chr13: 4641471..4642670	*AtTCP15*
MdTCP35	MDP0000374900	756	1	251	26948.08	9.67	chr13: 8629824..8630576	*AtTCP20*
MdTCP36	MDP0000202241	669	4	222	25375.36	9.91	chr13: 25582032..25583228	*AtTCP13*
MdTCP37	MDP0000693146	741	3	246	27271.87	9.89	chr13: 34653532..34663244	*AtTCP4*
MdTCP38	MDP0000210785	1197	2	398	42180.84	9.71	chr14: 25705450..25706743	*AtTCP14*
MdTCP39	MDP0000155433	1116	1	371	41251.98	7.32	chr14: 26663505..26664620	*AtTCP5*
MdTCP40	MDP0000224810	1440	2	479	54019.35	6.56	chr14: 27236441..27238113	*AtTCP18*
MdTCP41	MDP0000617459	795	3	264	29182.19	10.64	chr15: 31621875..31625233	*AtTCP8*
MdTCP42	MDP0000247249	1005	6	334	36745.93	6.18	chr15: 41975515..41982669	*AtTCP24*
MdTCP43	MDP0000515080	1596	1	531	55578.99	6.37	chr15: 47289729..47291321	*AtTCP8*
MdTCP44	MDP0000608645	1596	1	531	55406.74	6.71	chr15: 47289793..47291385	*AtTCP8*
MdTCP45	MDP0000272980	1251	1	416	46624.44	9.01	chr16: 1586089..1587339	*AtTCP12*
MdTCP46	MDP0000319941	1209	1	402	42973.7	6.83	chr16: 3210527..3211735	*AtTCP15*
MdTCP47	MDP0000915616	948	1	315	33460.38	8.68	chr17: 492812..493756	*AtTCP20*
MdTCP48	MDP0000320363	1143	1	380	41775.72	6.81	chr17: 4673803..4674945	*AtTCP13*
MdTCP49	MDP0000916623	1068	1	355	38772	5.98	chr17: 19109090..19110154	*AtTCP4*
MdTCP50	MDP0000149841	348	1	115	12816.8	9.79	chr0: 109274102..109274449	*AtTCP4*
MdTCP51	MDP0000851695	846	1	281	29713.74	8.08	chr0: 4448042..4448884	*AtTCP7*
MdTCP52	MDP0000373350	384	2	127	14337.32	9.51	chr0: 15633466..15633909	*AtTCP13*

Table 2 also shows the gene identifier, genomic position, pI, number of amino acids (aa), protein size, exon numbers and homologous genes. The ORF lengths ranged from 348 (*MdTCP50*) to 1839 bp (*MdTCP27*), with an average of 1049 bp. The corresponding proteins contained aa from 115 to 612 (average 348 aa), with a predicted molecular mass range of 12,816–69,007 Da; the pIs ranged from 5.41 (*MdTCP32*) to 10.65 (*MdTCP3*).

Table 3. Number of *TCP*s in *Arabidopsis*, rice and apple.

Species	Class 1	Class 2	Class 3	Total
Arabidopsis	13	2	8	23
Rice	8	3	9	22
Apple	22	4	26	52

Phylogenetic relationships and gene structure analysis of TCP gene family in apple

To evaluate the evolutionary relationship among the MdTCP proteins, the full-length aa sequences of 52 MdTCPs and 24 TCPs from *Arabidopsis* were subjected to a multiple sequence alignment using the MEGA5.0 program. The multiple sequence alignment file was then used to construct an unrooted phylogenetic tree. As shown in figure 1, the MdTCPs were divided into three classes (class 1, 2 and 3) as monophyletic clades with at least 50% bootstrap support, containing 22, 4 and 26 members, respectively. The small number of proteins in class 2 was consistent with the classifications of *Arabidopsis* and rice (table 3). In addition, 12 sister pairs of paralogous *MdTCP* genes were identified (figure 2), with strong bootstrap support (> 90%). The gene structures of the *MdTCP*s were analysed using the gene structure display server 2.0 (GSDS). As shown in figure 2, most *MdTCP*

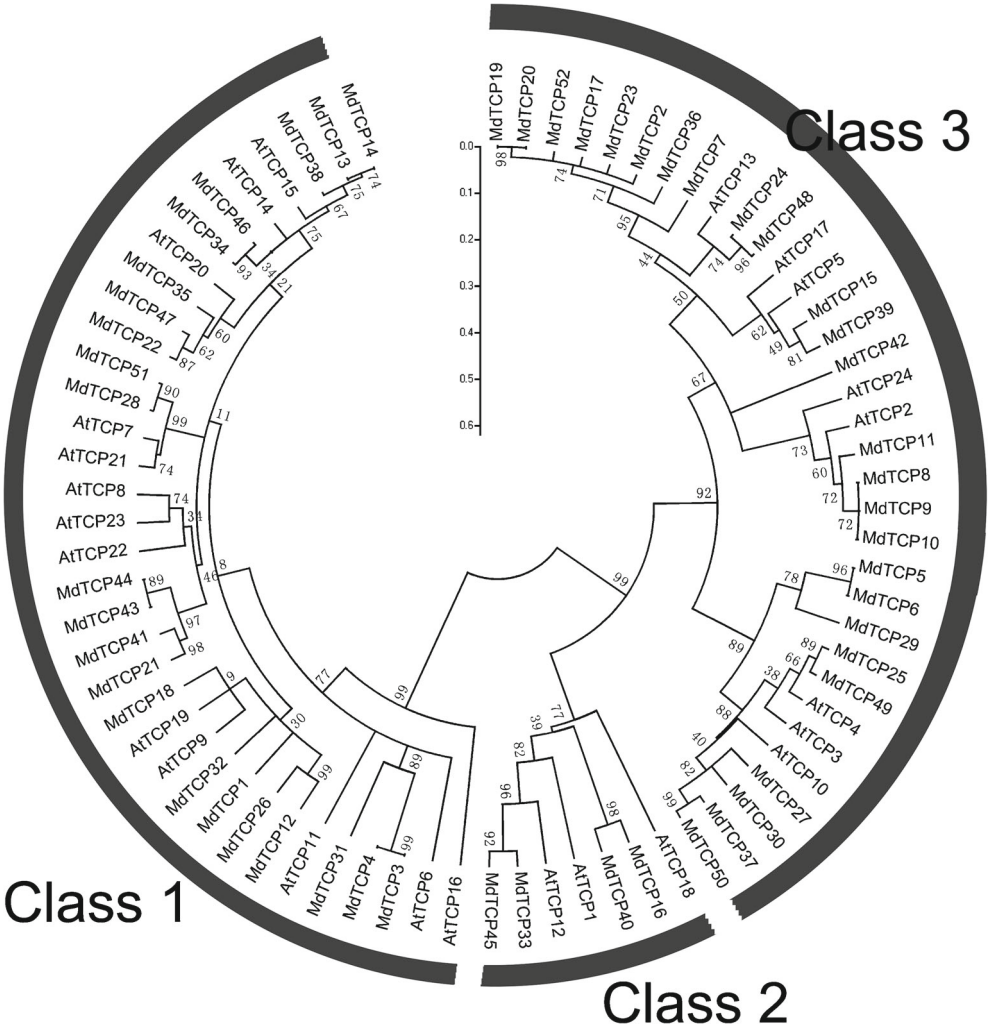

Figure 1. Phylogenetic relationships of *Arabidopsis* and apple *TCP* genes. The phylogenetic tree was constructed based on a complete protein sequence alignment of TCPs in *Arabidopsis* and apple by the (NJ) method with bootstrapping analysis (1000 replicates). The classes are marked by blue fragments. Scale bar represents 0.1 aa substitution per site.

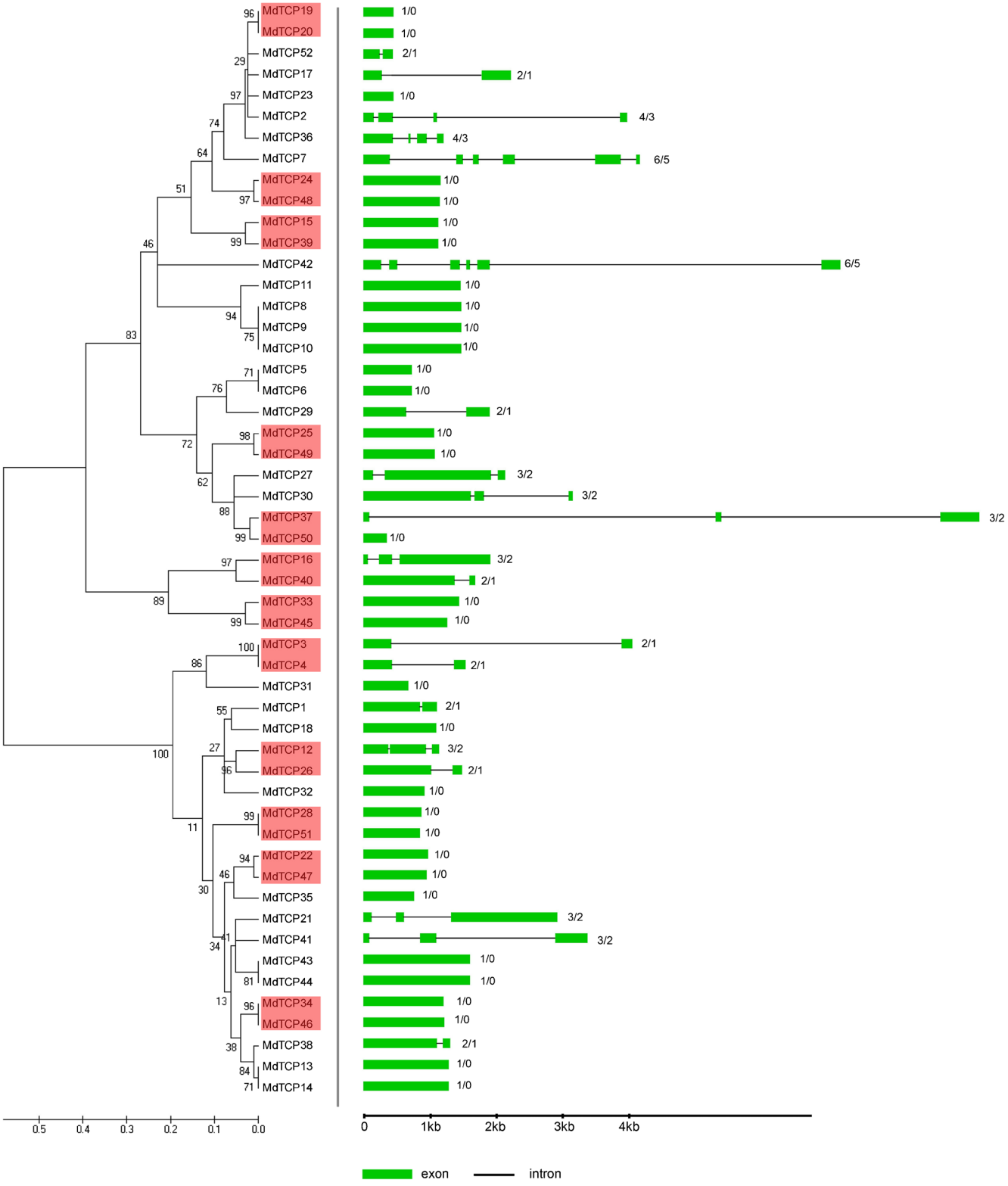

Figure 2. Phylogenetic relationships and exon/intron structure of *TCP* genes in apple. The phylogenetic tree was constructed based on a complete protein sequence alignment of TCPs by the (NJ) method with bootstrapping analysis (1000 replicates). Sister paralogous gene pairs are indicated by red line. The exons and introns are represented by the green boxes and black lines, respectively, and the number of the exons and introns are marked behind the gene structure. The sizes of exons, introns and untranslated regions are drawn to scale as indicated at the bottom of the figure.

genes (32 members, 61.5%) have no introns. Nine genes have one intron (17.3%), seven have two introns (13.5%), two have three introns (3.8%) and two have five introns (3.8%). The sequence alignment analysis of these TCP proteins showed that the specific TCP domain with bHLH was highly conserved and the R domain was not present in apple TCP proteins (figure 3). We also compared the *TCP* gene structures between *Arabidopsis* and apple (figure 4). In *Arabidopsis*, the exon numbers ranged from one to four, 82% genes contained only one exon and the average number is 1.30. In *MdTCP*s, the exon numbers ranged from one to six, 63% genes contained only one exon and the average number is 1.73. These results suggested that the *TCP* genes did not exhibit gene structure diversification independently in different organisms.

Chromosomal location of TCPs on apple genomes

Chromosomal location analyses showed that 49 *MdTCP* genes were located on 16 chromosomes, dispersed throughout the genome (figure 5). No *TCP* gene was found on chromosome 3, while chromosome 5 had the most *TCP* members

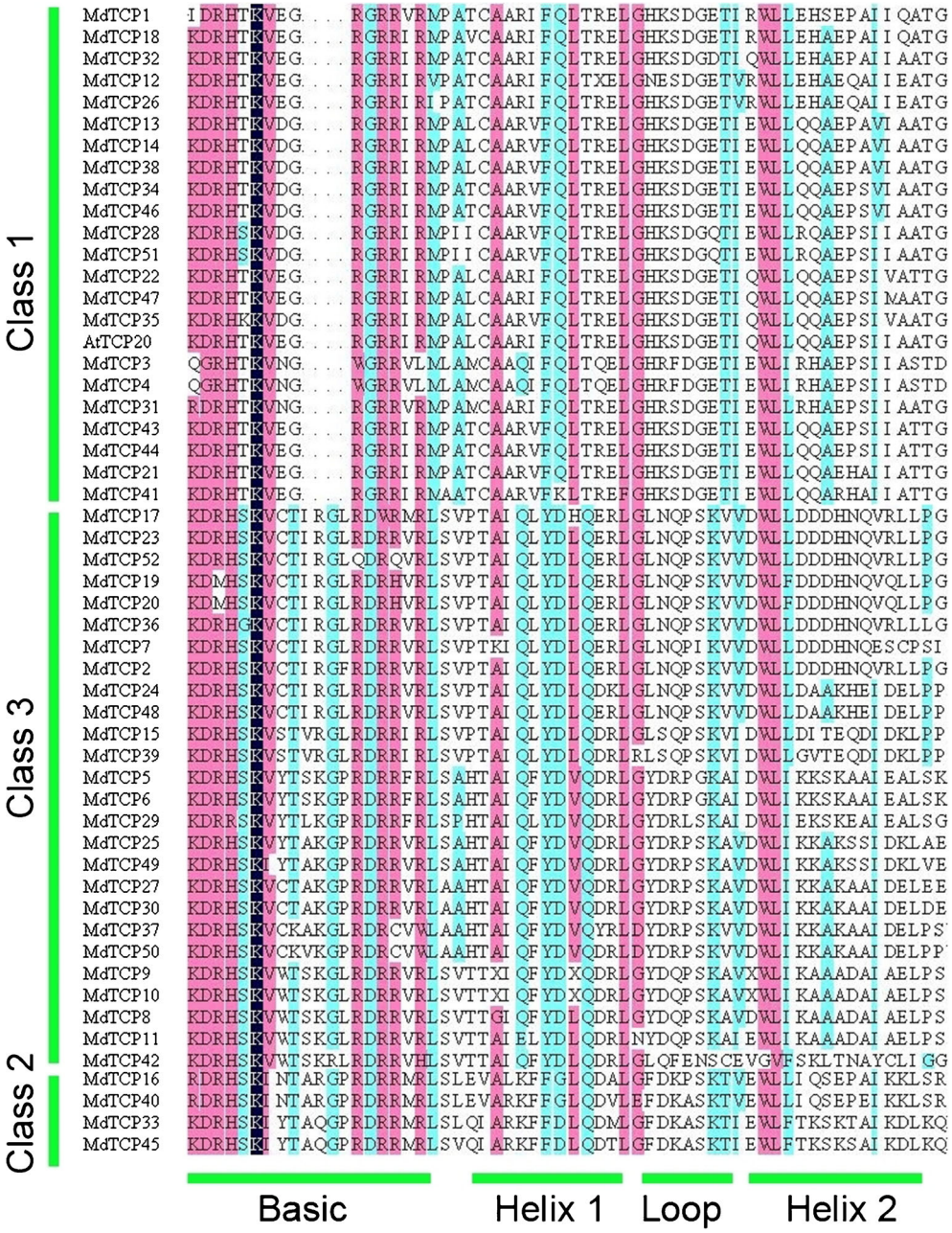

Figure 3. Sequence alignment of the conserved motif of TCP proteins in apple. Conserved aa are shown in colour. Dots denote gaps.

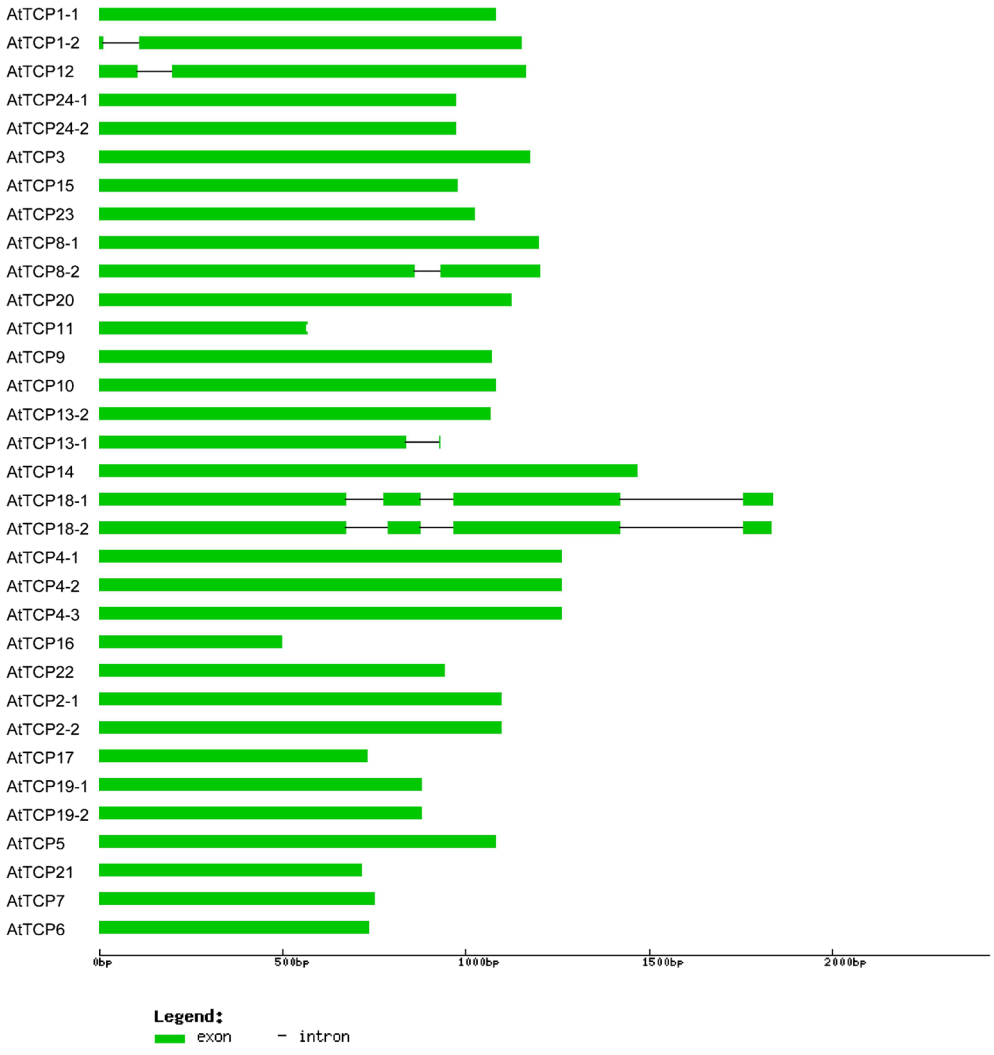

Figure 4. *TCP* gene structures in *Arabidopsis*.

(seven genes). Interestingly, among the 12 sister pairs of paralogous genes, two pairs *MdTCP19/20* and *MdTCP3/4* were tightly colocated in the apple genome and seven sister pairs (*MdTCP24/48*, *MdTCP15/39*, *MdTCP25/49*, *MdTCP16/40*, *MdTCP33/45*, *MdTCP22/47* and *MdTCP34/46*) were linked to at least one of the 15 potential chromosomal/segmental duplications combined with genomewide duplication on different chromosomes. Therefore, it was suggested that segmental duplication and transposition events both played roles in the evolution of *TCP* gene family in apple.

Microarray analysis of the expression patterns of TCPs

Gene expression pattern analysis can provide important information for understanding the roles of genes. The microarray data of GEO (GSE24523) showed the transcriptional variation of apple genes from four weeks before ripening (WBR) to ripening. By BLAST searching against the apple unigene database from GDR, 28 *MdTCP*s were identified (figure 6). Based on hierarchical clustering, the expression patterns of these 28 *MdTCP* genes were divided into two groups, namely, groups I and II (figure 6). The expression of group II genes, such as *MdTCP11*, *MdTCP13*, *MdTCP14*, *MdTCP25* and *MdTCP49*, increased slightly from four WBR to ripening, whereas *MdTCP16*, *MdTCP22*, *MdTCP32* and *MdTCP47*, showed slightly decreased expression among the apple cultivar, Honey Crisp (HC). In addition, the expression of *MdTCP16* (group II) decreased slightly from four WBR to ripening, whereas genes in group I, such as *MdTCP1*, *MdTCP8* and *MdTCP40*, increased in the apple cultivar, Cripps Pink (CP) (figure 6). However, the expression of other genes did not show significant changes during the ripening process.

Additionally, using BLAST searches against the probe sequences (GPL3715), the expression patterns of 51 *MdTCP*s (except MdTCP29) during the rootstock–scion interaction process were identified. Based on hierarchical clustering, the expression patterns of the *MdTCP* genes were

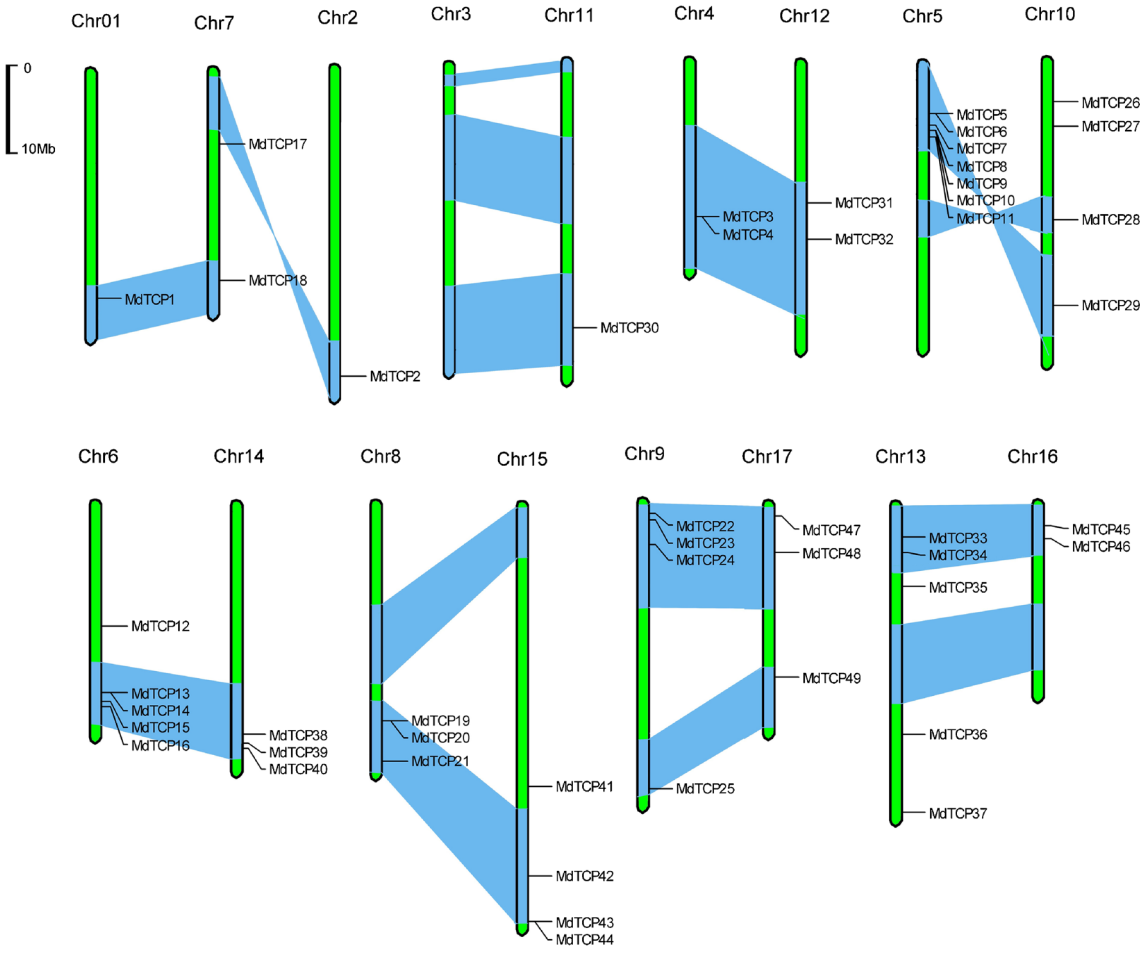

Figure 5. Chromosomal mapping analysis of the *TCP* gene family in apple. The chromosome number is indicated at the top of each chromosome representation. Scale measures a 10 Mb chromosomal distance. The gene names on the right side of each chromosome correspond to the approximate locations of each *TCP* gene. Segmented duplicate homologous blocks are indicated with a blue shadow.

divided into three groups, namely, groups I, II and III, with 22, 15 and 14 *TCP* genes, respectively (figure 7). The expression of group I decreased slightly in Ambrosia/ B9, Gala/B9 and Melrose/B9, but increased in the other seven rootstock–scion types. The transcripts of group II increased slightly in Ambrosia/B9 and Gala/B9, while most decreased in other rootstock–scion types. Intriguingly, the expression of group III showed significant differences during the rootstock–scion interaction processes. Taken together, the different expression patterns of *MdTCP* genes indicated that they might be directly or indirectly involved in the fruit development and rootstock–scion interaction processes.

Expression profiles of MdTCP genes

Based on the sequence similarity comparisons and phylogenetic analyses, we randomly selected 12 apple *TCP* genes for further study. Among them, three, four and five genes were selected from classes 1, 2 and 3, respectively. RT-PCR was used to detect their expression patterns in different tissues and in response to abiotic stresses. As shown in figure 8, 10 genes were differentially expressed in all of the examined tissues. However, tissue-specific expression was also observed. For example, the expression of *MdTCP8* was too low to be detected in fruits, while *MdTCP45* could be only detected in the stem, leaf and fruit (figure 8). When treated with stressed conditions, the expression levels of most genes (except *MdTCP3*, *MdTCP25* and *MdTCP49*) were apparently modulated (figure 9). Among them, seven genes (*MdTCP4*, *MdTCP8*, *MdTCP9*, *MdTCP10*, *MdTCP16*, *MdTCP31* and *MdTCP40*) were suppressed by heat (37°C), cold (4°C), high salinity (150 mM NaCl) and osmotic stress (10% PEG). Intriguingly, *MdTCP33* and *MdTCP45* were significantly induced by heat, salt and osmotic stress, respectively (figure 9). Overall, the expression profile analysis indicated that *MdTCP* genes were widely distributed and might play multiple roles in apple development and abiotic stress tolerance.

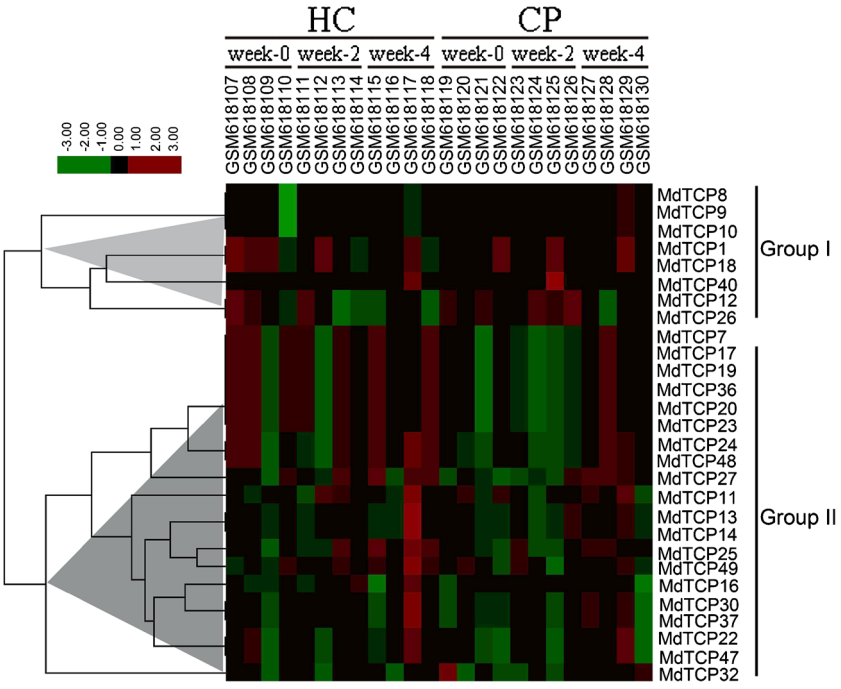

Figure 6. Expression patterns of the *MdTCP*s from microarray during fruit development in apple. The colour scale representing the relative signal values is shown above (green, low expression; black, medium expression; red, high expression). HC, Honey Crisp; CP, Cripps Pink; week-4, four weeks before ripening; week-2, two weeks before ripening; week-0, ripening. Twenty-four samples are shown as GSM618107–GSM618130 (GEO accession numbers).

Discussion

TCP TFs play important roles in diverse processes, especially in developmental programmes (Cubas *et al.* 1999; Li *et al.* 2005; Takeda *et al.* 2006; Aguilar-Martinez *et al.* 2007; Herve *et al.* 2009; Nag *et al.* 2009; Giraud *et al.* 2010; Martin-Trillo and Cubas 2010; Danisman *et al.* 2012). To date, only a limited number of TCP TFs have been identified and functionally characterized even in model plants such as *Arabidopsis* and rice, and no dataset of apple TCP TFs are available. In this study, we identified 52 *TCP*s containing full ORFs in apple using genomewide analysis. However, the missing ones may reside in the apple genome due to the screening methods that we utilized. Compared with the *TCP*s from *Arabidopsis* and rice, the number in apple is much larger, indicating that the *TCP* gene family in apple has expanded (Brameier 2010). We speculate that the presence of more *TCP* genes in apple genome may reflect the greater need for the involvement of these genes in the complicated transcriptional regulations in the woody perennial species.

We constructed a phylogenetic tree based on the full length of the TCP proteins and three classes were clustered, which was consistent with the previous studies (Brameier 2010). *MdTCP8–MdTCP11*, *MdTCP25* and *MdTCP49* belonged to class 3 and shared the highest homology with the miR319-targetd *TCP* genes and AS2-binding genes, which implies that these *MdTCP* genes may be involved in leaf development through similar mechanisms (Palatnik *et al.* 2003; Li *et al.* 2012a). Likewise, *MdTCP13*, *MdTCP14*, *MdTCP34*, *MdTCP38* and *MdTCP46* in class 1, which resembled *TCP14* and *TCP15* in *Arabidopsis*, might be involved in regulating cell proliferation in the developing leaf blade and specific floral tissues (Hiratsu *et al.* 2003; Kieffer *et al.* 2011). *MdTCP22*, *MdTCP35* and *MdTCP47* in class 1 might act as key regulators with similar function to *AtTCP20* in flower development (Li *et al.* 2005; Kieffer *et al.* 2011). As shown in figure 2, most *TCP* genes from the same subfamily shared relatively similar exon/intron structures in terms of intron number and exon length between *Arabidopsis* and apple, which provided an excellent reference to explore the functions of the *MdTCP*s. The structural analyses of apple TCP TFs will mirror the diverse functions of *TCP* genes and encourage future functional research. However, the detailed biological functions of most *TCP* genes require further investigation.

Gene duplications, including segmental and tandem duplications, play important roles not only in genomic rearrangement and expansion but also in the diversification of gene functions, implicating them as the primary driving forces throughout the evolutionary process of genomes (Moore and Purugganan 2003; Cannon *et al.* 2004). It has been reported

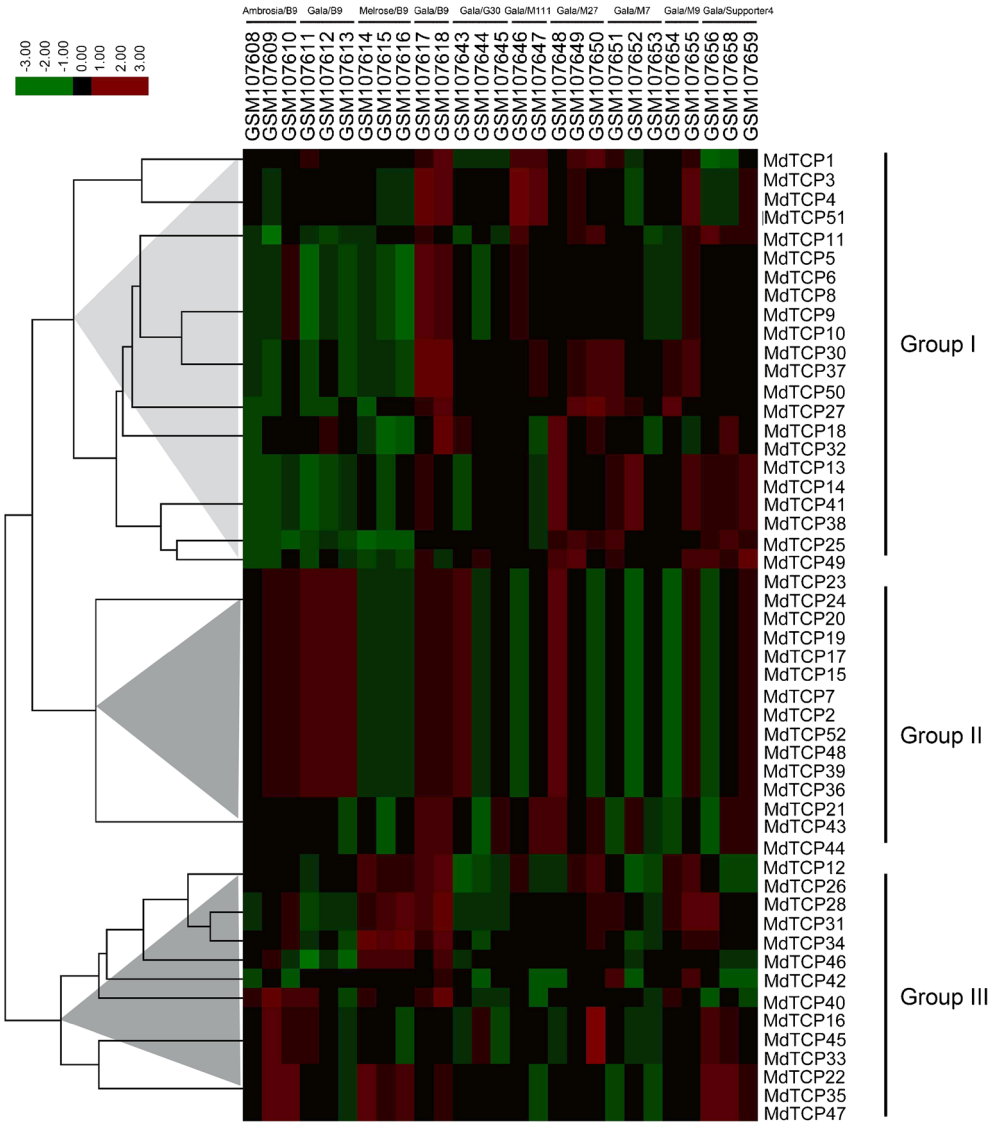

Figure 7. Expression profile of the *MdTCP*s during rootstock–scion interactions. The colour scale represents the relative signal value from weak (green) to strong (red). Ten types of rootstock–scion interactions in altogether 27 samples were analysed.

that a recent genomewide duplication event in apple occurred 60–65 million years ago, which resulted in gene expansion (Velasco *et al.* 2010). In this study, chromosomal location analysis showed that the apple *TCP* genes were dispersed throughout 16 out of 17 chromosomes with different densities, and mutiple sister pairs were linked to chromosomal segmental duplications. Consistent with this there was a clear paralogous pattern of gene family divergence by gene duplication in apple. Evolutionary divergence analysis suggested that the whole genome duplication, chromosomal segment duplications and transposition events, might have contributed to the expansion of apple TCP TFs.

Microarray and EST data revealed the expression patterns of 28 *MdTCP* genes during fruit development, and 51 *MdTCP*s during rootstock–scion interaction process. For example, *MdTCP11*, *MdTCP13*, *MdTCP14*, *MdTCP25*, *MdTCP49*, *MdTCP22* and *MdTCP47*, increased or decreased slightly from four WBR to ripening among the apple cultivars HC and CP, respectively (figure 6). This result implied that the *MdTCP* genes might be involved in the apple ripening process, especially in leaf and flower development, which correlated with the phylogenetic relationship analysis. Additionally, most *MdTCP*s showed significant differences during rootstock–scion interaction process. According to the microarray and EST analyses, *MdTCP* genes might be intimately involved in fruit development and the rootstock–scion interaction processes. Moreover, nine out of 12 randomly selected *MdTCP* genes exhibit obvious responsiveness to different kinds of abiotic stresses, implying that *MdTCP*s are also involved in abiotic stress signalling or tolerance, which is worth of further investigation.

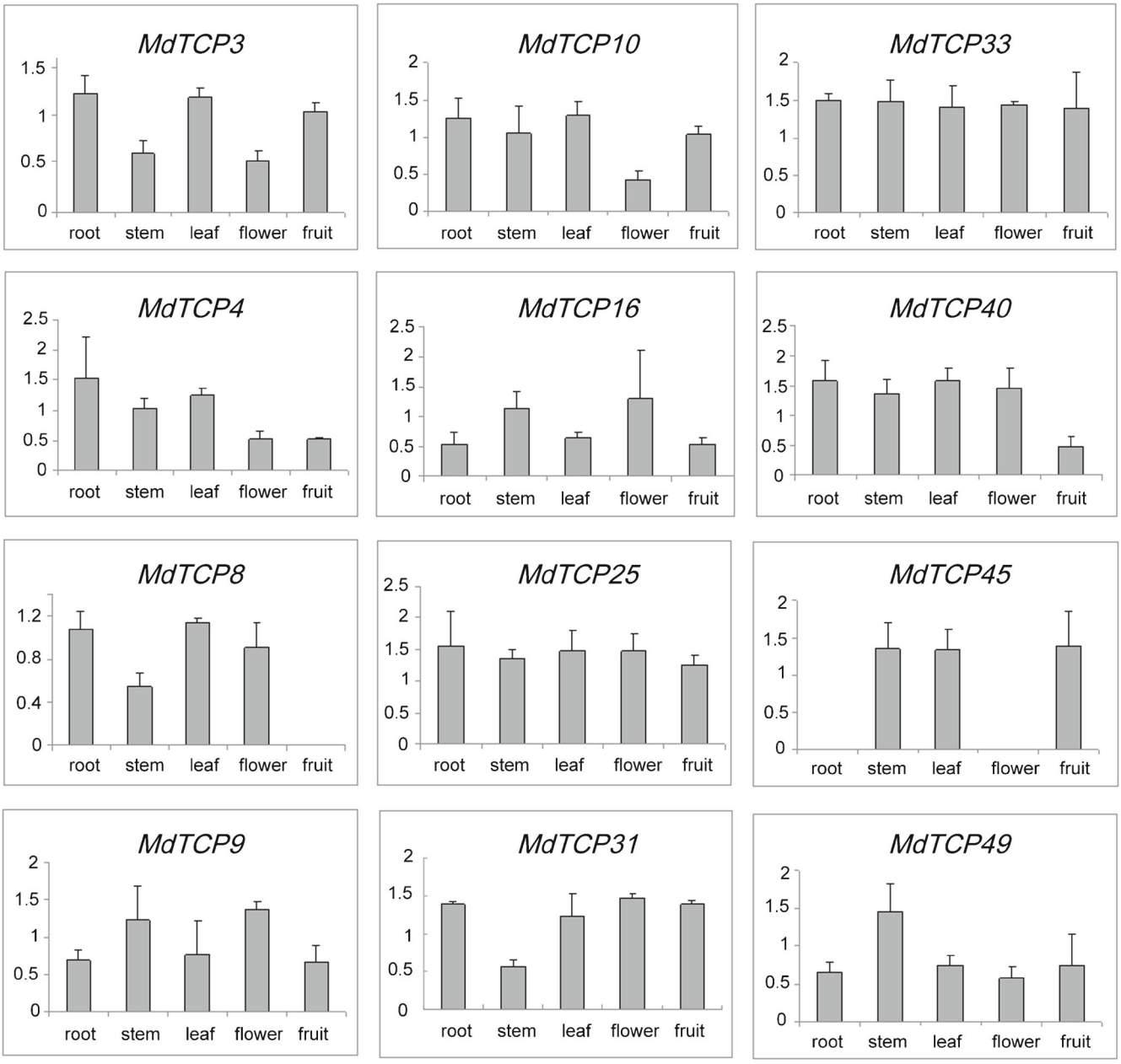

Figure 8. Expression profiles of the selected *MdTCP* genes in different apple tissues The data were normalized to the expression level of apple *Actin*. The mean expression value was calculated from three independent replicates. The vertical bars indicate the standard deviation (SD).

This study presented a systematic analysis of *TCP* gene family in apple, with special emphasis on fruit development and rootstock–scion interaction process. Our results lay the foundation for functional characterization of *MdTCP* genes and will lead to further understanding of the structure–function relationship between these family members. Additionally, our study provided comprehensive information and novel insights into the evolution and divergence of the *TCP* genes in plants. Potentially, this study will aid in the understanding of the molecular basis of many agronomically important traits of apple, such as fruit development, stress tolerance and other physiological processes.

Conclusion

In this study, 52 *TCP* genes were identified in the apple genome. The *MdTCP*s were divided into three classes and 49 *MdTCP* genes were distributed across 16 chromosomes, except Chr03, with different densities. Expression analysis showed that the *MdTCP*s were altered during the ripening process and rootstock–scion interaction process. Expression profile analyses of *MdTCP*s were performed in different tissues and in response to different stress conditions. All of the selected genes were expressed in at least one of the tissues tested and most of them were responsive to

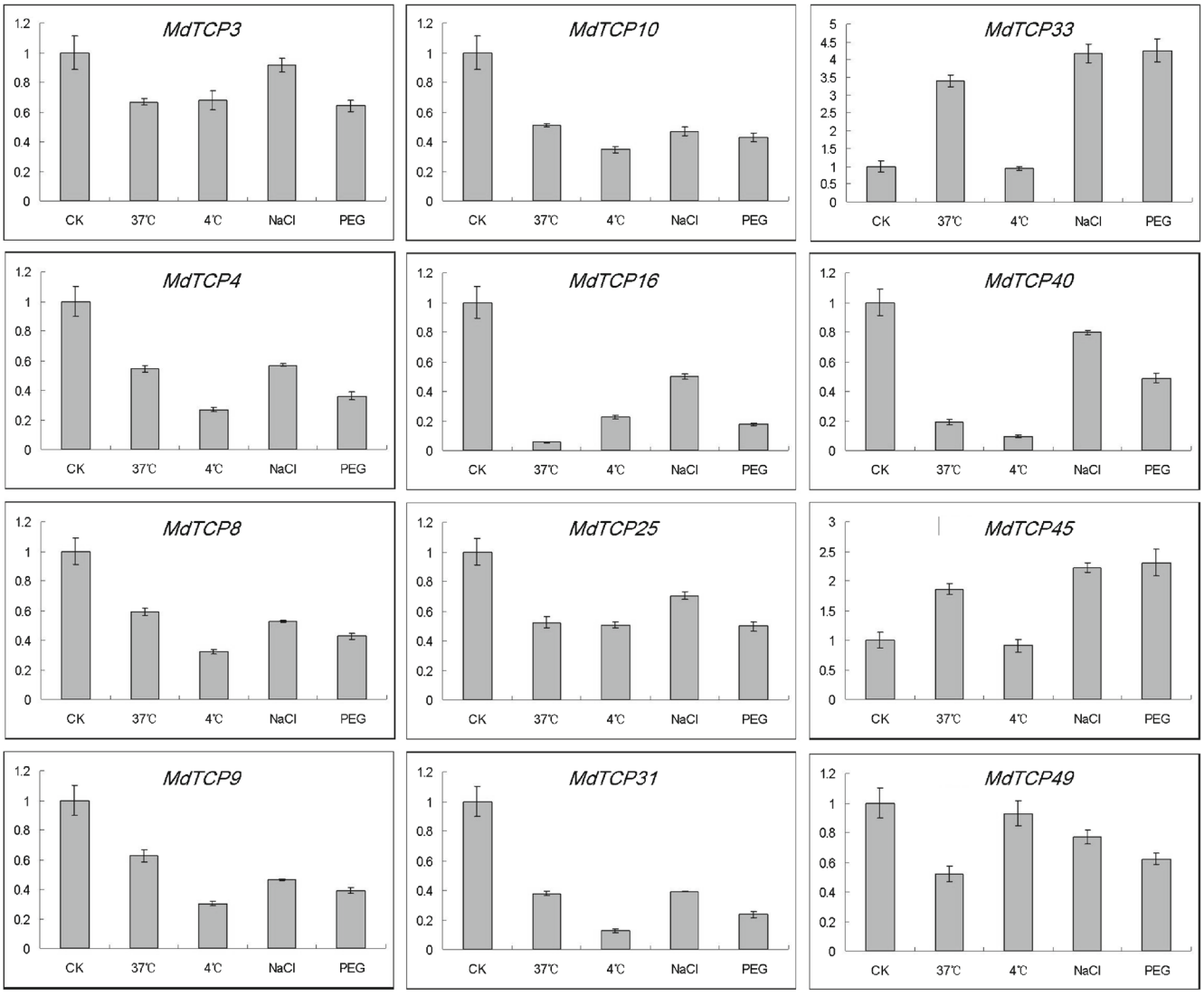

Figure 9. Expression analysis of *MdTCP* genes to different abiotic stresses. One-month-old seedlings were treated with 37°C, 4°C, 150 mM NaCl or 10% PEG6000 for 6 h, respectively. After treatment, the aerial part was collected for RNA extraction and real time PCR analysis. The data were normalized to the expression level of apple *Actin*. The mean expression value was calculated from three independent replicates. The vertical bars indicate the standard deviation (SD).

abiotic stress treatments, indicating that the *MdTCP*s were involved in various developmental and physiological processes in apples. To the best of our knowledge, this is the first report of a genomewide analysis of the apple *TCP* gene family. This study provides valuable information for understanding the classification and functions of the *TCP* genes in apples.

Acknowledgements

This work was supported by the Open Project Programme of the State Key Laboratory of Crop Biology (grant no. 2013KF07), and the Open Project Programme of Key Laboratory of Biology and Molecular Biology in University of Shandong (Weifang University) (grant no. 2012SWKF01) in China.

References

Aguilar-Martinez J. A., Poza-Carrion C. and Cubas P. 2007 Arabidopsis BRANCHED1 acts as an integrator of branching signals within axillary buds. *Plant Cell* **19**, 458–472.

Brameier M. 2010 Genome-wide comparative analysis of micro RNAs in three non-human primates. *BMC Res. Notes* **3**, 64.

Cannon S. B., Mitra A., Baumgarten A., Young N. D. and 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.

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.

Danisman S., van der Wal F., Dhondt S., Waites R., de Folter S., Bimbo A. *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.

Doebley J., Stec A. and Hubbard L. 1997 The evolution of apical dominance in maize. *Nature* **386**, 485–488.

Edgar R. C. 2004 MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797.

Feng X., Zhao Z., Tian Z., Xu S., Luo Y., Cai Z. *et al.* 2006 Control of petal shape and floral zygomorphy in *Lotus japonicus*. *Proc. Natl. Acad. Sci. USA* **103**, 4970–4975.

Finn R. D., Bateman A., Clements J., Coggill P., Eberhardt R. Y., Eddy S. R. *et al.* 2014 Pfam: the protein families database. *Nucleic Acids Res.* **42**, D222–D230.

Giorno F., Guerriero G., Baric S. and Mariani C. 2012 Heat shock transcriptional factors in *Malus domestica*: identification, classification and expression analysis. *BMC Genomics* **13**, 639.

Giraud E., Ng S., Carrie C., Duncan O., Low J., Lee C. P. *et al.* 2010 TCP transcription factors link the regulation of genes encoding mitochondrial proteins with the circadian clock in *Arabidopsis thaliana*. *Plant Cell* **22**, 3921–3934.

Herve C., Dabos P., Bardet C., Jauneau A., Auriac M. C., Ramboer A. *et al.* 2009 *In vivo* interference with AtTCP20 function induces severe plant growth alterations and deregulates the expression of many genes important for development. *Plant Physiol.* **149**, 1462–1477.

Hiratsu K., Matsui K., Koyama T. and Ohme-Takagi M. 2003 Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in *Arabidopsis*. *Plant J.* **34**, 733–739.

Jones P., Binns D., Chang H., Fraser M., Li W., McAnulla C. *et al.* 2014 InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240.

Jung S., Ficklin S. P., Lee T., Cheng C. H., Blenda A., Zheng P. *et al.* 2014 The genome database for rosaceae (GDR): year 10 update. *Nucleic Acids Res.* **42**, 1237–1244.

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.

Kosugi S. and Ohashi Y. 1997 PCF1 and PCF2 specifically bind to *cis* elements in the rice proliferating cell nuclear antigen gene. *Plant Cell* **9**, 1607–1619.

Letunic I., Doerks T. and Bork P. 2012 SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res.* **40**, 302–305.

Li C., Potuschak T., Colon-Carmona A., Gutierrez R. A. and Doerner P. 2005 *Arabidopsis* TCP20 links regulation of growth and cell division control pathways. *Proc. Natl. Acad. Sci. USA* **102**, 12978–12983.

Li Y., Wu B., Yu Y., Yang G., Wu C. and Zheng C. 2011 Genome-wide analysis of the RING finger gene family in apple. *Mol. Genet. Genomics* **286**, 81–94.

Li Z., Li B., Shen W. H., Huang H. and Dong A. 2012a TCP transcription factors interact with AS2 in the repression of class-I KNOX genes in *Arabidopsis thaliana*. *Plant J.* **71**, 99–107.

Li Z. Y., Li B. and Dong A. W. 2012b The *Arabidopsis* transcription factor AtTCP15 regulates endoreduplication by modulating expression of key cell-cycle genes. *Mol. Plant.* **5**, 270–280.

Liang D., Xia H., Wu S. and Ma F. 2012 Genome-wide identification and expression profiling of dehydrin gene family in *Malus domestica*. *Mol. Biol. Rep.* **39**, 10759–10768.

Liu R. H. and Meng J. L. 2003 MapDraw: a microsoft excel macro for drawing genetic linkage maps based on given genetic linkage data. *Yi Chuan* **25**, 317–321.

Luo D., Carpenter R., Copsey L., Vincent C., Clark J. and Coen E. 1999 Control of organ asymmetry in flowers of Antirrhinum. *Cell* **99**, 367–376.

Martin-Trillo M. and Cubas P. 2010 TCP genes: a family snapshot ten years later. *Trends Plant Sci.* **15**, 31–39.

Moore R. C. and Purugganan M. D. 2003 The early stages of duplicate gene evolution. *Proc. Natl. Acad. Sci. USA* **100**, 15682–15687.

Mount D. W. 2007 Using the basic local alignment search tool (BLAST). *CSH Protoc.* (doi: 10.1101/pdb.top17).

Nag A., King S. and Jack T. 2009 miR319a targeting of TCP4 is critical for petal growth and development in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **106**, 22534–22539.

Palatnik J. F., Allen E., Wu X., Schommer C., Schwab R., Carrington J. C. and Weigel D. 2003 Control of leaf morphogenesis by microRNAs. *Nature* **425**, 257–263.

Reeves P. A. and Olmstead R. G. 2003 Evolution of the TCP gene family in Asteridae: cladistic and network approaches to understanding regulatory gene family diversification and its impact on morphological evolution. *Mol. Biol. Evol.* **20**, 1997–2009.

Riechmann J. L., Heard J., Martin G., Reuber L., Jiang C., Keddie J. *et al.* 2000 *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. *Science* **290**, 2105–2110.

Takeda T., Suwa Y., Suzuki M., Kitano H., Ueguchi-Tanaka M., Ashikari M. *et al.* 2003 The OsTB1 gene negatively regulates lateral branching in rice. *Plant J.* **33**, 513–520.

Takeda T., Amano K., Ohto M. A., Nakamura K., Sato S., Kato T. *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.

Tamura K., Peterson D., Peterson N., Stecher G., Nei M. and Kumar S. 2011 MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739.

Velasco R., Zharkikh A., Affourtit J., Dhingra A., Cestaro A., Kalyanaraman A. *et al.* 2010 The genome of the domesticated apple (*Malus* × *domestica* Borkh.) *Nat. Genet.* **42**, 833–839.

Wang Z., Luo Y., Li X., Wang L., Xu S., Yang J. *et al.* 2008 Genetic control of floral zygomorphy in pea (*Pisum sativum* L.) *Proc. Natl. Acad. Sci. USA* **105**, 10414–10419.

Wray G. A., Hahn M. W., Abouheif E., Balhoff J. P., Pizer M., Rockman M. V. *et al.* 2003 The evolution of transcriptional regulation in eukaryotes. *Mol. Biol. Evol.* **20**, 1377–1419.

Yao Y. X., Li M., Zhai H., You C. X. and Hao Y. J. 2011 Isolation and characterization of an apple cytosolic malate dehydrogenase gene reveal its function in malate synthesis. *J. Plant Physiol.* **168**, 474–480.

Zhao T., Liang D., Wang P., Liu J. and Ma F. 2012 Genome-wide analysis and expression profiling of the DREB transcription factor gene family in *Malus* under abiotic stress. *Mol. Genet. Genomics* **287**, 423–436.

Received 27 February 2014, in revised form 9 May 2014; accepted 30 June 2014
Unedited version published online: 18 July 2014
Final version published online: 9 December 2014