bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Genome-Wide Analysis of *TCP* Family Genes and Their Constitutive Expression Pattern Analysis in the Melon (*Cucumis melo*)

Md Jahid Hasan Jone[1], Md Nure Adil Siddique[1], Manosh Kumar Biswas[2], and Mohammad Rashed Hossain[1,*]

[1]Plant Molecular Breeding and Bioinformatics Laboratory, Department of Genetics and Plant Breeding, Bangladesh Agricultural University, Mymensingh - 2202, Bangladesh

[2]Department of Genetics and Genome Biology, University of Leicester, United Kingdom

*Correspondence**:** m.r.hossain@bau.edu.bd

Abstract

TCP proteins are plant-specific transcription factors that play essential roles in various developmental processes, including leaf morphogenesis and senescence, flowering, lateral branching, hormone crosstalk, and stress responses. However, the specific functions of *TCP* genes in melon remain largely unknown. This study identified and characterized 29 putative *TCP* genes in melon. These genes were classified into two classes: Class-I (13 genes) and Class-II (16 genes). The chromosomal location, gene structure, conserved motifs, structural homology, cis-regulating elements, transcript expression patterns, and potential protein-protein interactions were further analyzed. The results revealed that the putative *CmTCP* genes are distributed across nine of the twelve melon chromosomes and exhibit diverse expression patterns in different tissues and during floral organ development. Phylogenetic analysis suggests that some *CmTCP* genes may have similar functions to their homologs in other plant species, while others may have undergone functional diversification. This study provides a valuable resource for future investigations into the specific roles of individual *CmTCP* genes in melon development and paves the way for elucidating the mechanisms by which *TCP* proteins regulate leaf elongation, floral development, and lateral branching.

Keywords: *Cucumis melo*, Expression pattern, Genome-wide, *TCP*, Transcription factors

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

INTRODUCTION

Melon (*Cucumis melo* L., 2n = 24) is an important economic crop grown worldwide for its sweet, edible and fleshy fruit. It is herbaceous, diploid and monoecious or andro-monoecious in nature. It produces male flowers in the main stem while lateral branches produce female or hermaphroditic flowers (Fukino et al. 2012). Well-maintained lateral branches result in good-quality fruits, higher yields and reduced cost of production as it is vital for light penetration, photosynthesis and energy partitioning. Lateral branching plays a significant role in the architecture of the melon plant (Fukino et al. 2012; Shahwar et al. 2023). Breaking the dominance of apical meristem through main stem pruning increases the lateral branching and leaf expansion, positively affecting fruit quality (Pereira et al. 2003; Widaryanto et al. 2020). However, the transcription factors responsible for lateral branching in melons were rarely studied before.

The *TCP* family genes are a group of plant-specific (Michael and Pilar 2016; Navaud et al. 2007) transcription factors encoding genes named after *TEOSINTE BRANCHED 1* (*TB1)*, *CYCLOIDEA* (*CYC*) and *PCF* proteins initially found in Maize, Garden snapdragon and Rice, respectively (Cubas et al. 1999; Finlayson 2007; Kosugi and Ohashi 2002). The conserved region of *TCP* is predicted to form a basic helix-loop-helix (bHLH) structure of 59 amino acid residues (Cubas et al. 1999). Based on sequence homology, *TCP* proteins are divided into two subfamilies namely, Class-I (conserved four-amino-acid deletion in basic region) and Class-II (Manassero et al. 2013). Class-I subfamily contains *PCF* (*TCP-P*) proteins whereas Class-II subfamily is again divided into two clades: *CIN* and *CYC/TB1* (Manassero et al. 2013).

The *TCP* proteins that have been characterized thus far function in meristem growth: TB1 influences the growth of axillary meristems, while *CYC* regulates the growth of floral meristems and primordia. *PCF* bind to the promoter of a gene involved in meristematic cell division (Cubas et al. 1999). A number of *CYC/TB1* genes are found to be involved in the growth of auxiliary meristems and buds, which produce lateral shoots (Michael and Pilar 2016).

The functions of *TCP* proteins in different plants were extensively studied. Members of the *TCP* family were found to be involved in the morphogenesis of shoot lateral organs (Aguilar-Martinez et al. 2007; Finlayson 2007), shoot branching (Martin-Trillo et al. 2011), cell growth and proliferation in leaf (Aguilar-Martinez and Sinha 2013; Kosugi and Ohashi

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

1997), hormone signalling (Daviere et al. 2014; Li and Zachgo 2013), controlling apical dominance (Doebley et al. 1997), leaf and floral development (Aguilar-Martinez and Sinha 2013). Overall, *TCP* proteins affect leaf shape and induce branching activity in plants (Michael and Pilar 2016).

Despite the importance of lateral branching in melon and its predictable relation with *TCP* family genes, a comprehensive genome-wide analysis of these transcription factors is yet to be conducted. The present study aimed to identify genome-wide *TCP* family genes and to analyze their structure, motif distribution, evolution and syntenic relation with other crops, to predict protein structure (homology modelling), sequence similarity, cis-regulatory elements, and to analyze protein-protein interaction and expression profiling. The findings may provide a deeper insight into the abundance, and evolutionary history of this gene family and may reveal a set of candidate genes responsible for lateral branching in melon.

MATERIALS AND METHODS

Sequence Retrieval and Confirmation

A comprehensive search of five databases, namely PlantTFDB (https://planttfdb.gao-lab.org/), iTAK (http://itak.feilab.net/), Melonomics v4.0 (https://www.melonomics.net/), CuGenDB (http://cucurbitgenomics.org/) and CuGenDBv2 (http://cucurbitgenomics.org/v2/), was conducted to identify melon *TCP* family genes and their corresponding proteins (Jin et al. 2017; Yu et al. 2022; Zheng et al. 2016). Consolidating the results from all five databases unique gene IDs for all *TCP* genes were recorded. To confirm the presence of *TCP* and other potential domains, these sequences were analyzed using NCBI Conserved Domain (www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) Search and GenomeNet Motif Search (www.genome.jp/tools/motif/). Following this analysis, sequences containing the *TCP* domain were selected for further investigation

Genomic information and protein sequences for *TCP* genes in Cucumber and Tomato were retrieved from published articles (Parapunova et al. 2014; Wen et al. 2020). Corresponding genome databases (www.arabidopsis.org/, https://rapdb.dna.affrc.go.jp/) and published articles were used to retrieve data for Arabidopsis and Rice (Shang et al. 2022; Su et al. 2021; Yao et al. 2007).

Characteristics of *CmTCP* Proteins

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Protein length (aa), Molecular weight (KDa), Isoelectric points (pI), and Grand Average of Hydropathy (GRAVY) values were determined from Expasy-Protparam (Gasteiger et al. 2003). Subcellular localizations were predicted by CELLO: Subcellular Localization Predictive System (Yu et al. 2006).

Chromosomal Localization and Gene Duplication Analysis

Chromosome lengths and chromosomal positions of *CmTCP* genes were retrieved from the CuGenDBv2 database (Yu et al. 2022). Genes were mapped on chromosomes using TBtools software (Chen et al. 2020a). Nine chromosomes that contain *CmTCP* gene(s) were visualized excluding the rest three. Duplicate gene pairs and the type of duplications were determined using MCScanX (Wang et al. 2012) and then visualized in TBtools software.

Genomic Structures, Motif Distribution, and Cis-Acting Elements Analysis

$GSDS_{2.0}$ web tools (Hu et al. 2015) were used to analyze the exon-intron structures of the *TCP* family genes. TBtools software (Chen et al. 2020a) and MEME web tools (Bailey et al. 2015) were used to analyze the motif distribution.

Upstream sequences of each *CmTCP* gene (2000 bp) were retrieved using PlantCARE (Lescot et al. 2002) and a heatmap was generated based on the numbers using MS Excel 2019. The Basic Biosequence view in TBtools was used to localize the cis-regulatory elements in the upstream region. The grouping of cis-regulating elements according to their functions was carried out as the procedures described in Wu et al (Wu et al. 2023).

Sequence Alignment and Phylogenetic Analysis

BioEdit Sequence Alignment Editor (Alzohairy 2011) was used to align *CmTCP* protein sequences (Supplementary File 1) via ClustalW and visualized to find the *TCP* domain region. Molecular Evolutionary Genetics Analysis (MEGA11 and MEGA7) software (Kumar et al. 2016; Tamura et al. 2021) was also used to align the *TCP* sequences of melon again via ClustalW with default settings. A rectangular phylogenetic tree was constructed with the aligned *CmTCP* sequences, via MEGA11, using the Maximum Likelihood method and 1000 replicates for bootstrap. A circular phylogenetic tree was constructed with the aligned sequences of Melon, Cucumber, Rice, Tomato and Arabidopsis, via MEGA7, using the Maximum Likelihood method and 1000 replicates for bootstrap. Finally, the tree was visualized using Interactive Tree Of Life (iTOL) v5 (Letunic and Bork 2021).

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Synteny Relationships

Genomic annotation (.gff) and genome sequence (.fa) files of Cucumber and Arabidopsis were retrieved from their genome database (CuGenDBv2 and TAIR Database) and Phytozome (Goodstein et al. 2012) was used to retrieve the files of Rice, and Tomato. The syntenic relationship of *CmTCP* genes with these genomes was analyzed using the MCScanX tool (Wang et al. 2012). TBtools software was used to construct the collinearity map (Chen et al. 2020a).

Expression Profiling Using RNA-Seq Data

Melonet DB (melonet-db.dna.affrc.go.jp/ap/mvw), a web tool for functional genomics research of melons (Yano et al. 2020) was used to analyze the expression of *CmTCP* genes based on the available RNA-seq data. FPKM values of 25 genes for 15 tissues and 45 different samples were available in the database. The tissues include- callus, seed (dry and imbibed), cotyledon, hypocotyl, root, flower (petal, anther, stigma, ovary), stem, shoot apex, leaves, tendril, fruit (flesh and epicarp). Heatmap of expression data was generated by R programming via the "pheatmap" package (Kolde 2019) using Euclidean distance and ward.D2 as a clustering method.

Protein Structure Prediction and Sequence Similarity

Protein homology modeling was conducted using the Swiss Model (Waterhouse et al. 2018) to generate predicted three-dimensional structures based on template (known homologous protein structure) identification. Additionally, sequence similarity analysis was performed using TBtools, facilitating the comparison of amino acid sequences and the construction of a similarity matrix.

Interaction Network of *CmTCP* Proteins

Protein-protein association networks of *CmTCP* proteins were predicted and visualized using the String Database tool (Szklarczyk et al. 2023). Three clusters of proteins were predicted using the K-means clustering method as well (Kalaivani et al. 2018).

RESULT

Identification of *TCP* genes in the Melon genome

A total of 29 *TCP* genes were identified in PlantTFDB and iTAK. Besides, 23 *TCP* genes were identified in Melonomics v4.0, while 50 and 47 genes were identified in CuGenDB and

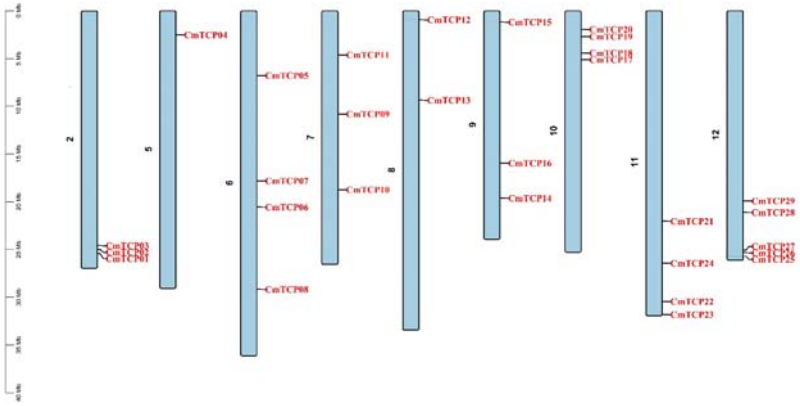
bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

CuGenDBv2, respectively. These altogether yielded 53 unique gene IDs which were further analyzed based on Conserved Domain and Motif analyses. Finally, 29 sequences containing the *TCP* domain were selected for subsequent analysis (Supplementary Table 1). These 29 sequences were named *CmTCP-01 to CmTCP-29* according to their location in various chromosomes. The size of *CmTCP* proteins ranges from 61 to 600 amino acids with a molecular weight of 7.11516 to 65.4692 kDa (Table 1). The highest size and molecular weight were found in the case of *CmTCP-10* and the lowest was for *CmTCP-04*.

The isoelectric point ranged from 5.62 to 9.92, where the pI value for 25 proteins was more than 7 indicating that *CmTCP* proteins are mostly basic in nature. Contrastingly, genes *CmTCP-11*, *CmTCP-24*, *CmTCP-25*, and *CmTCP-29* were found to be acidic in nature. Only one protein (*CmTCP-14*) is predicted to be localized at the cytoplasm whereas all other genes are predicted to be localized at the nucleus (Table 1). The GRAVY value for *CmTCP* protein ranges from -0.329 (*CmTCP*-08) to -1.002 (*CmTCP*-14). The negative value for every protein indicates that *CmTCP* proteins are non-polar and hydrophobic in nature (Table 1).

Chromosomal localization

The chromosomal distribution of *CmTCP* family genes illustrates that they are distributed throughout the 9 (out of 12) chromosomes in the melon genome while chromosomes 1, 3 and 4 contain no *TCP* gene (Figure 1). The highest number (five) of *CmTCP* genes were found in chromosome 12. Chromosomes 6, 10, and 11 contain 4 *CmTCP* genes whereas chromosomes 2, 7, and 9 contain 3 *CmTCP* genes in each. Chromosomes 8 and 5 contain 2 and 1 genes respectively. Most of the genes are located far from the centromere with a tendency to be located toward the telomere regions. Only a few genes such as *CmTCP-06, CmTCP-07, and*



CmTCP-09 are located in the centromeric regions (Figure 1).

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Fig. 1 Chromosomal distribution of 29 *CmTCP* genes in melon. The length of the chromosomes is indicated by the scale on the left.

Table 1. Characteristics of the 29 *CmTCP* family gene in melon.

Gene name	Gene ID	Chromosomal location	Strand	Gene (bp)	CDS (bp)	Protein (aa)	Molecular Weight (kDa)	pI	GRAVY	Exon	Intron	Subcellular Localization
CmTCP-01	MELO3C017168	02 : 25419111-25420541	Forward	1433	1232	409	43.99313	7.15	-0.728	1	0	Nuclear
CmTCP-02	MELO3C017230	02 : 24983902-24986084	Reverse	2185	1250	415	45.76948	8.86	-0.729	1	0	Nuclear
CmTCP-03	MELO3C017286	02 : 24576035-24577997	Reverse	1965	1070	341	38.82925	9.11	-0.930	2	1	Nuclear
CmTCP-04	MELO3C014426	05 : 2513719-2513904	Reverse	188	188	61	7.115160	8.79	-0.410	1	0	Nuclear
CmTCP-05	MELO3C006863	06 : 6767274-6768731	Reverse	1460	1043	346	36.69640	8.91	-0.492	1	0	Nuclear
CmTCP-06	MELO3C014903	06 : 20554511-20555522	Reverse	1014	791	262	27.86820	7.94	-0.510	1	0	Nuclear
CmTCP-07	MELO3C023772	06 : 17801439-17802578	Forward	1142	1142	379	42.64302	9.42	-0.742	1	0	Nuclear
CmTCP-08	MELO3C025357	06 : 29167155-29168260	Forward	1108	800	265	27.60193	9.77	-0.329	1	0	Nuclear
CmTCP-09	MELO3C010371	07 : 10795619-10797541	Reverse	1925	929	307	34.22900	9.30	-0.792	1	0	Nuclear
CmTCP-10	MELO3C016092	07 : 18723693-18726567	Reverse	2877	1805	600	65.46920	9.92	-0.876	2	1	Nuclear
CmTCP-11	MELO3C025629	07 : 4601921-4604309	Forward	2391	1292	429	47.01419	6.11	-0.660	1	0	Nuclear
CmTCP-12	MELO3C007121	08 : 922389-929076	Forward	6690	1361	452	49.11905	9.02	-0.764	1	0	Nuclear
CmTCP-13	MELO3C024459	08 : 9353286-9354143	Forward	860	761	207	22.19292	8.45	-0.446	1	0	Nuclear
CmTCP-14	MELO3C005358	09 : 19614471-19615388	Forward	920	770	255	29.61013	9.63	-1.002	2	1	Cytoplasm
CmTCP-15	MELO3C022091	09 : 1173102-1177531	Reverse	4432	1190	396	44.92324	8.50	-0.803	2	1	Nuclear
CmTCP-16	MELO3C025049	09 : 15948957-15950155	Reverse	1199	788	261	27.49579	9.13	-0.380	1	0	Nuclear
CmTCP-17	MELO3C011754	10 : 5105118-5106803	Reverse	1688	950	256	26.48040	9.92	-0.429	1	0	Nuclear
CmTCP-18	MELO3C011835	10 : 4410767-4411981	Forward	1217	1217	404	41.21548	9.26	-0.467	1	0	Nuclear
CmTCP-19	MELO3C012086	10 : 2683489-2685269	Forward	1783	1100	365	38.96212	9.09	-0.658	1	0	Nuclear
CmTCP-20	MELO3C012212	10 : 1934213-1937023	Forward	1813	1442	309	35.37687	9.69	-0.690	2	1	Nuclear
CmTCP-21	MELO3C019745	11 : 22032106-22033227	Forward	1124	902	299	32.64721	8.67	-0.619	1	0	Nuclear
CmTCP-22	MELO3C022331	11 : 30437155-30440954	Forward	3802	887	294	31.69923	7.97	-0.687	1	0	Nuclear
CmTCP-23	MELO3C022520	11 : 31804195-31805184	Reverse	992	992	329	35.41270	9.07	-0.502	1	0	Nuclear
CmTCP-24	MELO3C026920	11 : 26425626-26426186	Reverse	563	561	186	20.52901	5.62	-0.650	1	0	Nuclear
CmTCP-25	MELO3C001972	12 : 25710988-25713753	Forward	2768	896	297	32.34194	6.63	-0.701	2	1	Nuclear
CmTCP-26	MELO3C002038	12 : 25334307-25335161	Reverse	857	545	180	20.50137	9.27	-0.642	1	0	Nuclear
CmTCP-27	MELO3C002076	12 : 25077880-25079727	Forward	1850	1079	271	30.97712	9.91	-0.934	2	1	Nuclear
CmTCP-28	MELO3C002636	12 : 21111900-21112494	Forward	597	398	165	18.05653	9.85	-0.491	2	1	Nuclear
CmTCP-29	MELO3C002754	12 : 19903500-19909976	Forward	6479	1418	471	51.33053	6.71	-0.734	1	0	Nuclear

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Sequence Alignment and *TCP* Domain

The alignment of *CmTCP* proteins (Supplementary File 1) revealed the *TCP* domain with 59 amino acid residues consisting of a 'basic' region and a helix-loop-helix (HLH) structure (Figure 2). There are 19 residues that have at least 80% similarity (Figure 2). Based on the similarity of residues, the 'basic' region is the most conservative followed by the 'helix' region. The loop region shows a great variation compared to the other two regions. Based on the sequence characteristics of the 'basic' region, *CmTCP* proteins are divided into two classes: class I and class II. In addition, the *TCP* genes were again divided into three groups namely, PCF, CYC/TB1 and CIN based on the last three residues of the first 'helix' region. Proteins with TRE amino acids were put in the *PCF* group while proteins containing QDM and QDK/QDR were assigned to *CYC/TB1* and *CIN* groups, respectively. All Class I genes fell within the *PCF* group as well. Class I (*PCF*) contains 13 *CmTCP* proteins which have 4 amino acid residues less than that of class II proteins. The *CIN* group (class II) has 10 *CmTCP* proteins and the *CYC/TB1* group (class II) has 6 *CmTCP* proteins, respectively.

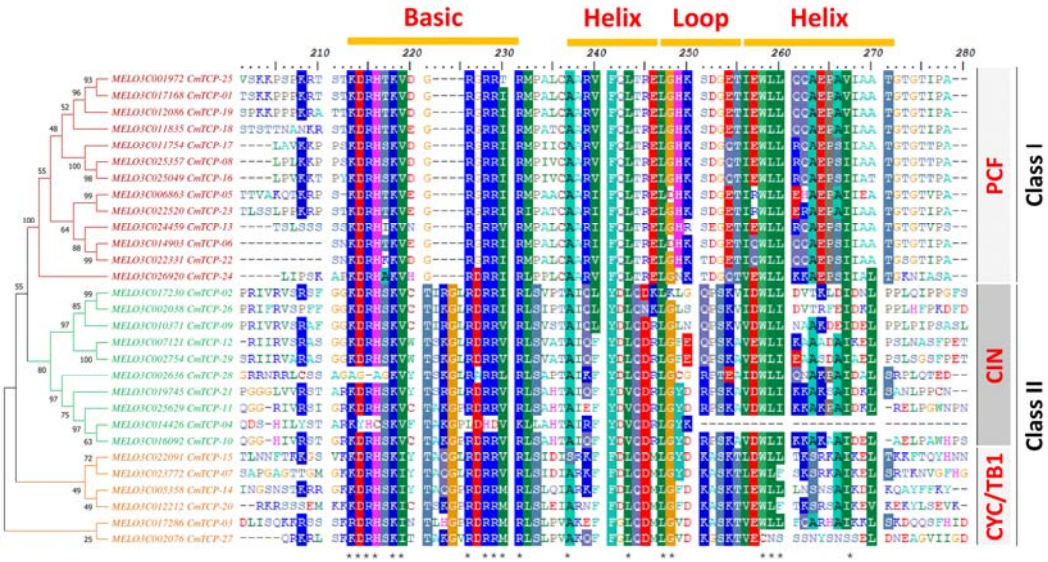


Fig. 2 Alignment of 29 *CmTCP* protein sequences. Similar amino acids are shaded using 50% as the threshold value. Conserved residues (at least 80%) are indicated with an Asterisk (*) at the bottom. The alignment of domain-related regions of *CmTCP* protein sequences is shown here. The complete alignment is shown in Supplementary File 2.

Phylogeny, Gene structure and Motif distribution

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

A phylogeny was constructed using 29 *CmTCP* proteins (Figure 3A). A total of 10 homologue gene pairs were found among the three groups. The *PCF* group contains four pairs (*CmTCP-25/CmTCP-01*, *CmTCP-08/CmTCP-16*, *CmTCP-05/CmTCP-23* and *CmTCP-06/CmTCP-22*) while *CIN* group consists three pairs (*CmTCP-02/CmTCP-26*, *CmTCP-12/CmTCP-29*, and *CmTCP-04/CmTCP-10*) and *CYC/TB1* group contains another three pairs of proteins (*CmTCP-15/CmTCP-07*, *CmTCP-14/CmTCP-20* and *CmTCP-03/CmTCP-27*). Closely related proteins generally contain similar structure and motif composition with some exceptions.

A total of 12 motifs were identified in these 29 *TCP* genes (Figure 3B) of which Motif 1 and Motif 2 are widely distributed in every protein except *CmTCP-28* (lacks Motif 2) and *CmTCP-04* (lacks Motif 1). Alignment of *CmTCP* protein sequences (Figure 2) shows that the sequences of *CmTCP-28* and *CmTCP-04* are either mutated or incomplete. Hence, it can be assumed that Motif 1 and Motif 2 are probably located within the *TCP* domain region. There is a visual difference in motif distribution between the two classes. Motif 3 is a part of every Class-I (*PCF*) protein (except *CmTCP-24*). Motif 4 and Motif 7 are found in Class-II while Motif 5, Motif 6, Motif 10 and Motif 11 are found only in Class-I proteins. Motif 8, Motif 9 and Motif 12 are found in both classes. Interestingly, *CmTCP-15* contains two segments of Motif 9 side-by-side (Figure 3B).

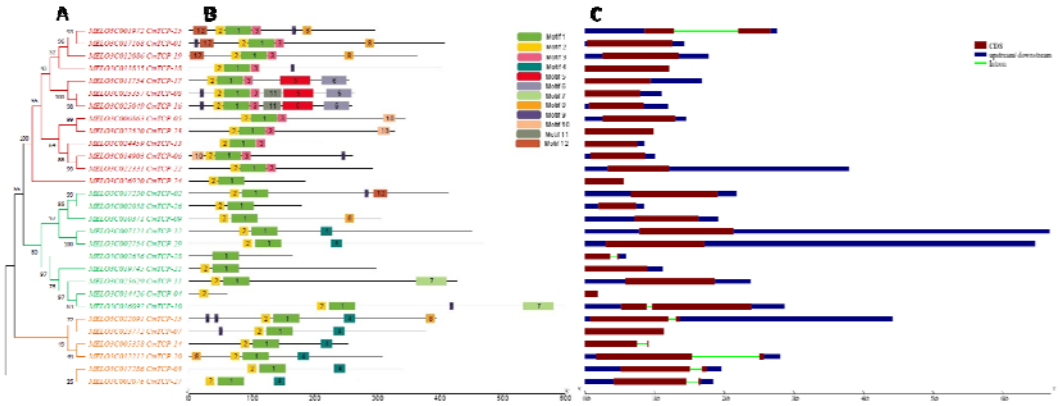


Fig. 3 *CmTCP* sequence analysis- **(A)** Phylogenetic tree. *PCF*, *CIN* and *CYC/TB1* groups are indicated in red, green and gold colour, respectively; **(B)** Motif distribution. Motif numbers are indicated with coloured boxes as per the legends shown on the right; **(C)** Gene structure analysis. dark red, dark blue and bright green colours illustrate exons, UTRs and introns, respectively. Scales below the graphs are proportional to the length of the chromosomes, motifs and exons/introns.

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

The exon-intron structures of 29 *TCP* genes were analized to understand the characteristics of *TCP* gDNA which revealed that *TCP* genes have a maximum of two exons (Figure 3C). Twenty-one *CmTCP* genes do not have any introns while 8 genes (*CmTCP-03, CmTCP-10, CmTCP-14, CmTCP-15, CmTCP-20, CmTCP-25, CmTCP-27, and CmTCP-28*) contain a single intron (Figure 3C). The size of *CmTCP* genes ranges from 188 bp to 6690 bp (Table 1). *CmTCP-12* is found to be the largest gene followed by *CmTCP-29* while the smallest gene was *CmTCP-04*. Among six genes, 5 *CYC/TB1* group genes contain an intron except *CmTCP-07* (Figure 3C).

Promoter Analysis

Gene expression patterns are largely determined by the cis-regulating elements present in the promoter region (Yamaguchi-Shinozaki and Shinozaki 2005). A total of 32 cis-acting elements were found involved in response to light responses (6), phytohormone responses

	Function	Elements	CmTCP-01	CmTCP-02	CmTCP-03	CmTCP-04	CmTCP-05	CmTCP-06	CmTCP-07	CmTCP-08	CmTCP-09	CmTCP-10	CmTCP-11	CmTCP-12	CmTCP-13	CmTCP-14	CmTCP-15	CmTCP-16	CmTCP-17	CmTCP-18	CmTCP-19	CmTCP-20	CmTCP-21	CmTCP-22	CmTCP-23	CmTCP-24	CmTCP-25	CmTCP-26	CmTCP-27	CmTCP-28	CmTCP-29
Light Responsive	Light	Box 4	7	3	2	5	4	3	1	2	11		2			4	2	9	3	3	5	6	2	4	5			2	3	5	5
		G-Box	3	4	3	5	4	2	5	2	2	1	4	4		3		1	1	2	1	3		3	2	2	3	1	4	1	
		GT1-motif	1	1			3	1	3	1	2	4	2	1	1	5	1	1	2	3		4		2	2	3	1	2		2	1
		I-box	1		1		1		1		1										1		1	1			1		2		
		TCCC-motif	2		3	2			1			1			2	1			3	1	1		1				1				
		TCT-motif				1	2	2				1				2	1		1	1			3			2	1				1
Phytohormone Responsive	MeJA	CGTCA-motif		1			2		1	2	1	1	2	1					1	1	1		1	1			2				
		TGACG-motif		1			2		1	2	1	1	2	1					1	1	1		1	1			2				
	ABA	ABRE	2	3	2	4	4	2	4	2	4	1	4	1		4		1	1	2	1	3		2	2	2	3	1	3		1
		DRE1																											1		
	Auxin	AuxRR-core																	1						1						
		TGA-element				1			3							1											1				
	Ethylene	ERE					1	1	4	3		1	1		1	1	1	2	2	1		5	3	3	1	3					1
	SA	TCA-element	2	2	1	1	3	5	3		1					1				1		2		2	1	1	1		2	1	2
	Gibberellin	GARE-motif	2	1	1	1	1	2	1			1		1	1	1		1		1	2		1		1			1	1	2	
		P-box			2							1	1		2		2		1	1	1	1	3			1		1	1	1	
		TATC-box																											1		
Stress Responsive	Drought	as-1		1			2		1	2	1	1	2	1					1	1	1		1	1			2				
		MBS		1			1						2		1	1	1	1	2				2				1				
		MYC	7	2	3	7	5	4	8	3	6	1	6	4	3	6	6	6	4	3	2	2	5	2	5	5	4	4	1	4	2
	Wound/Pathogen	WRE3				1		1		1	1	2	1		1			1	1	1					3	1	1			4	2
	Wound	WUN-motif						2	3	2			2		1				2	2		2		1	1	3	2	1	3		1
	Anaerobic induction	ARE		5	3	2	1	1	2	1	4	2	1	1	2	4	2	1	1	3		2	3	4	1	2	2	2	2	3	4
	Cold/Dehydration	DRE core		1																										1	
	Anoxic Specific inducibility	GC-motif				1																									
	Low Temperature	LTR			1		1							2	1	1				1			1			1		1			
	Stress	STRE	5	2	2		1	4	1	1	1	5	3	1	1	1	3	1	4	5	8		1	1	4		3		1		3
	Defence/Stress	TC-rich repeats	1	2					1					1	1		1	1	2			1		1				1			
Growth & Development	Specific Day Length	A-box																						1							
	Zein Metabolism	O2-site	1		2	2	1				1									1									1		
	Meristem Specific gene activation	CCGTCC-box																						1							
	Seed Specific Expression	AAGAA-motif	5	4	4	1	4	3	3		2	6	2	4	1	2	7	1	2	3	2	2	3	1	2	2	4	5	3	1	

(11), stress responses (11), and growth and development (4) (Wu et al. 2023).

Fig. 4 Cis-acting elements identified in the promoter region of *CmTCP* genes.

We found Box 4, G-Box, GTI-motif, I-box, TCCC-motif and TCT-motif for light responsiveness. The elements for phytohormone responsiveness were divided into six groups, such as Methyl Jasmonate (CGTCA-motif and TGACG-motif), ABA (ABRE and DRE1), Auxin (AuxRR-core and TGA-element), Ethylene (ERE), Salicylic Acid (TCA-element) and Gibberellin (GARE-motif, P-box and TATC-box). Defence and stress response elements

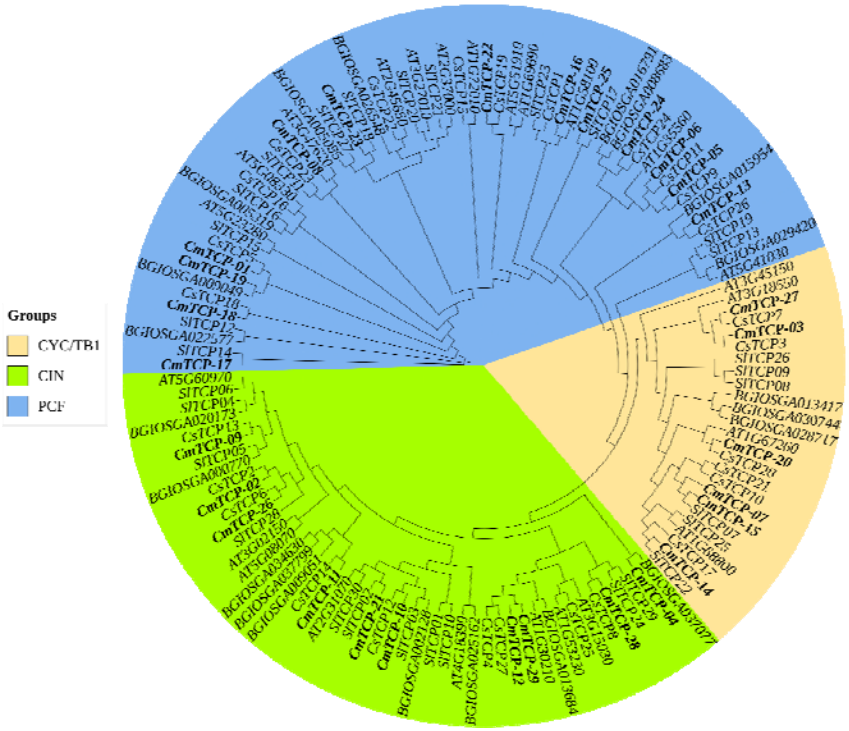
bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

included as-1, MBS, MYC, WRE3, WUN-motif, ARE, DRE core, GC-motif, LTR, STRE and TC-rich repeats. Moreover, several other elements related to growth and development, such as meristem-specific gene activation, zein metabolism, circadian control element, and seed-specific expressions were found (Figure 4).

Each gene consists of multiple cis-regulating elements from each group. Drought-responsive element, MYC were found in the promoter region of every *CmTCP* gene whereas DRE1 and TATC-box were only found in *CmTCP-27* and A-box and CCGTCC-box in *CmTCP-22* (Figure 4). These results indicated that *CmTCP* genes might be involved in the response to various stresses, light, and phytohormones along with growth and development.

Evolutionary Analysis of the *CmTCP* Gene Family

A total of 131 *TCP* proteins from 5 different plants namely, Melon (29 proteins), Cucumber (27 proteins), Rice (21 proteins), Arabidopsis (24 proteins), and Tomato (30 proteins) were



used to study the evolution of *CmTCP* proteins in the melon genome (Figure 5) (Supplementary File 2).

Fig. 5 Phylogenetic analysis of *TCP* proteins in Melon (*CmTCP*), Cucumber (*CsTCP*), Rice (*BGIOSGA*), Arabidopsis (*At*), and Tomato (*SlTCP*). The phylogenetic tree was constructed

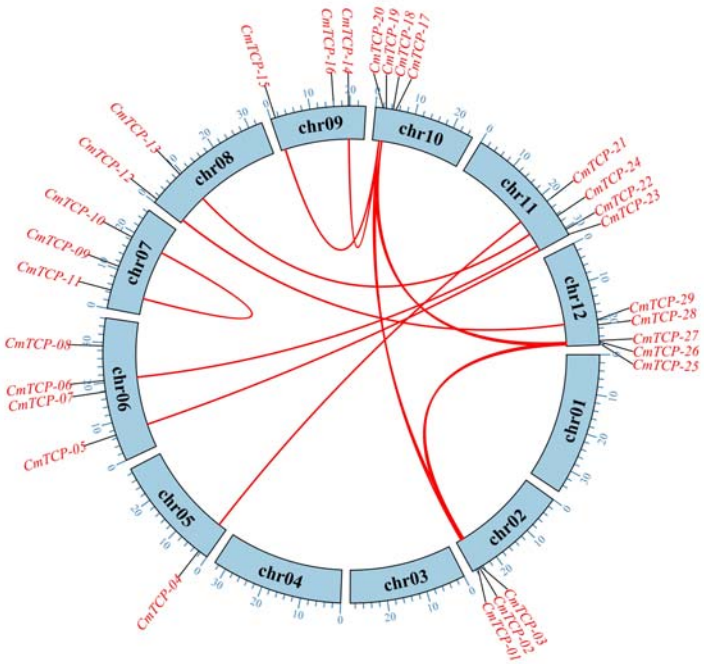
bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

by the Maximum likelihood method using 1000 bootstrap replicates. *CmTCP* proteins are highlighted in black bold letters.

Despite a huge difference in the size of the genome and the number of genes in these plants, the *TCP* encoding genes in each genome are more or less similar in number (Supplementary Table 2) in these genomes. The phylogenetic analysis also demonstrated that several of the *CmTCP* proteins were discovered to be more closely related to those of Cucumber followed by Tomato (Figure 5). For instance, *CmTCP09/CsTCP13, CmTCP26/CsTCP6, CmTCP28/CsTCP8, CmTCP21/SlTCP02, CmTCP14/SlTCP22,* and *CmTCP25/SlTCP17.* All the Arabidopsis *TCP*s in this phylogenetic tree belong to the same group as reported previously (Martin-Trillo and Cubas 2010).

Gene Duplication and Syntenic Relationship

There are 15 whole genome duplications or segmental duplications (syntenic pairs) found among *CmTCP* genes with no evidence of tandem duplication (Figure 6, Supplementary Table 3). Out of 10 homologous gene pairs, seven showed gene duplication events except *CmTCP-08/CmTCP-16, CmTCP-04/CmTCP-10* and *CmTCP-15/CmTCP-07*. Moreover, eight syntenic pairs, *CmTCP-01/CmTCP-19, CmTCP-03/CmTCP-20, CmTCP-04/CmTCP-21, CmTCP-11/CmTCP-10, CmTCP-13/CmTCP-24, CmTCP-15/CmTCP-20, CmTCP-*



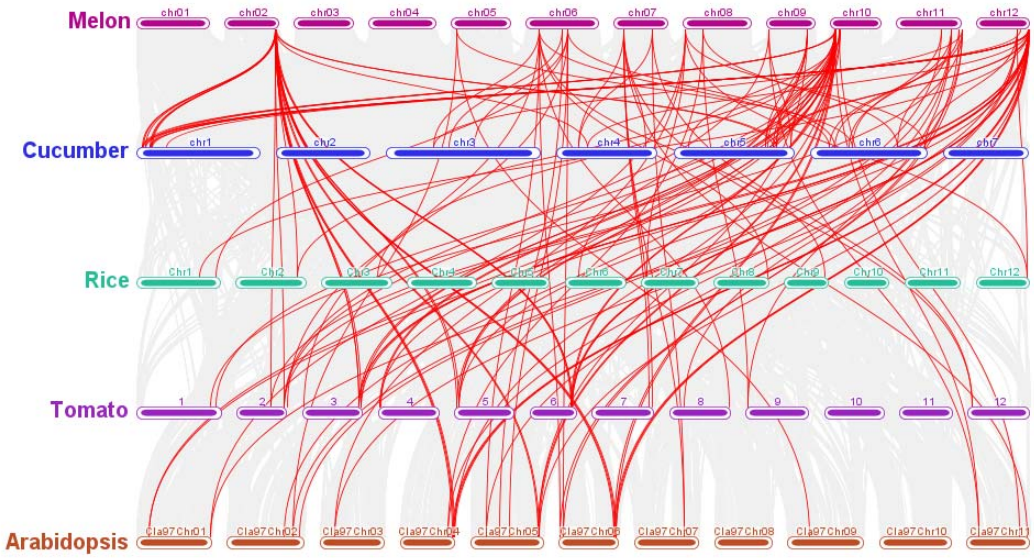
19/CmTCP-25 and *CmTCP-20/CmTCP-27* are non-homologous gene pairs. These results depict that whole genome duplication or segmental duplication probably plays an essential

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

role in the evolution of *CmTCP* genes together with other factors including random insertion events.

Fig. 6 Gene duplication among 29 *CmTCP* genes. Duplicate genes are indicated using red lines.

To understand the evolution and abundance of *TCP* genes, syntenic relations of *CmTCP* genes with other crops were studied. A total of 151 syntenic pairs of *CmTCP* genes were calculated for these four species (Figure 7, Supplementary Table 3). Among them, the



highest number of pairwise syntenic blocks were found in Cucumber (49) followed by Arabidopsis (48) and Tomato (43) whereas only 11 pairwise syntenic blocks were found between melon and rice (Figure 6, Supplementary Table 3). *CmTCP-20* showed the maximum number of syntenic blocks (14) followed by *CmTCP-19* (11) while no syntenic blocks were found in the case of *CmTCP-08, CmTCP-09* and *CmTCP-16* (Figure 7, Supplementary Table 3).

Fig 7. Syntenic relationships of 29 Melon *CmTCP* genes with that of Cucumber, Rice, Tomato and Arabidopsis. The grey lines in the background represent the collinear blocks within Melon and other plant genomes, while the red lines highlight the syntenic *TCP* gene pairs.

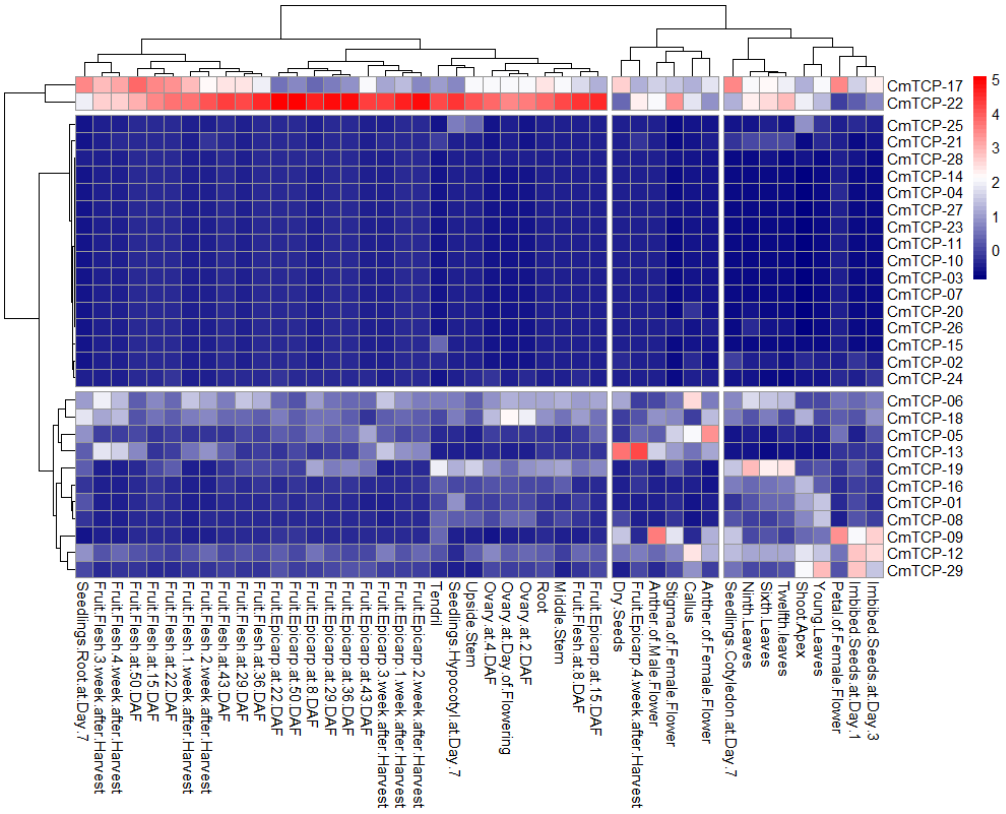
Expression pattern analysis of *CmTCP* genes

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

1 RNA expression data of 29 *CmTCP* genes from 45 samples of 25 vegetative and reproductive

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

tissues were extracted from Melonet DB (in FPKM unit) (Figure 8). Among them, the highest number of genes (twenty-four) were expressed in shoot apex followed by tendril and young leaves where 21 and 20 genes were expressed, respectively. The lowest number of genes (nine) were expressed in the fruit flesh after a week of harvesting (Supplementary Table 4). This result indicates that *CmTCP* genes are more expressed in the vegetative stage compared to the reproductive stage. No expression records were found for five genes (*CmTCP-03*, *CmTCP-07*, *CmTCP-10*, *CmTCP-11*, and *CmTCP-23*) whereas eight genes (*CmTCP-05*, *CmTCP-06*, *CmTCP-12*, *CmTCP-13*, *CmTCP-17*, *CmTCP-18*, *CmTCP-22*, and *CmTCP-29*) were expressed in every samples (Supplementary Table 4). The expression patterns indicate that *CmTCP-17* and *CmTCP-22* were highly expressed in every tissue



compared to others, and may have an influence on overall growth and development. According to the clusters, 11 genes were moderately expressed and 16 genes were either expressed at a very low level or did not show any expression. Several genes are expressed in a specific tissue much higher than in other tissues indicating their regulation of the tissue-specific phenotypes. For example, *CmTCP-09 was* expressed higher in anther and petals whereas *CmTCP-13 was* expressed in dry seed and fruit epicarp. Such expression data illustrates the possible function and involvement of *CmTCP* genes in various stages of plant growth and development.

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

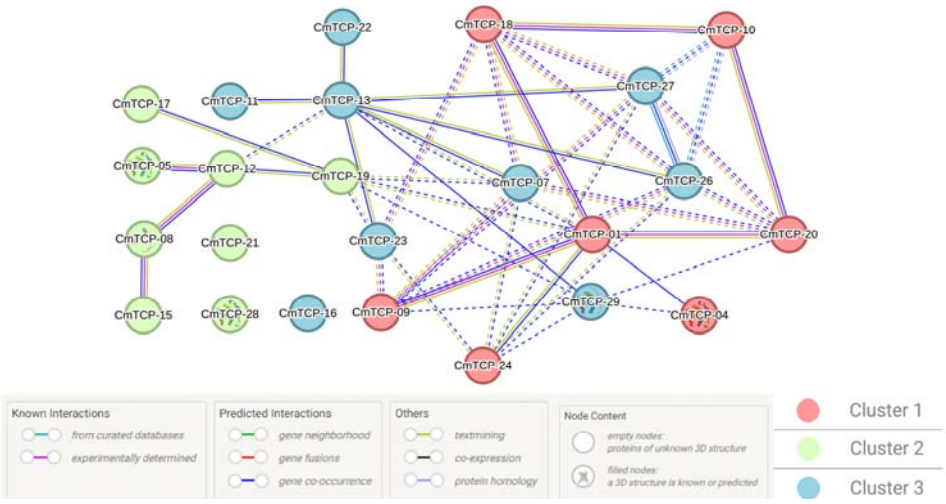
Fig. 8 Expression pattern of *CmTCP* genes in different melon tissues and developmental stages. The heatmap was generated by the RNA expression data (FPKM) from Melonet-DB.

Protein Structure Prediction and Sequence Similarity

Protein homology modeling is a powerful technique used to predict the three-dimensional structure of a protein based on its amino acid sequence similarity to known protein structures. In this study, the Swiss Model, a widely used homology modeling tool was used to generate structural models for the target proteins. Predicted structures of *CmTCP* proteins along with template information are presented in Supplementary File 3. GMQE (Global Model Quality Estimation) is a measure used in protein homology modeling to estimate the reliability of a generated model. The highest GMQE value was found in the case of *CmTCP-26* (0.65) and the lowest was for *CmTCP-10* (0.36). Proteins of each group showed a similar pattern in the number of alpha helix and beta sheets (Supplementary File 3). The sequence similarity among the *CmTCP* proteins ranged from 4.69 (*CmTCP-0* vs *CmTCP-18)* to 93.58 (*CmTCP-08* vs *CmTCP-16)*. For the PCF group of proteins, the range is 21.14 (*CmTCP-24* vs *CmTCP-18)* to 93.58 (*CmTCP-08* vs *CmTCP-16)* which is 7.05 (*CmTCP-04* vs *CmTCP-29)* to 65.05 (*CmTCP-12* vs *CmTCP-29)* and 27.21 (*CmTCP-15* vs *CmTCP-27)* to 45 (*CmTCP-07* vs *CmTCP-15)* for CIN and CYC/TB1 groups respectively. Comparatively, proteins from similar groups showed a higher similarity percentage (Supplementary Table 5).

Predicted network of *CmTCP* proteins

The protein-protein interaction network includes three clusters of 24 genes predicted with String-db. Cluster 1 includes seven proteins, whereas Cluster 2 includes eight proteins and the rest belong to Cluster 2 (Figure 8). Several proteins like *CmTCP-01, CmTCP-07, CmTCP-09, CmTCP-10, CmTCP-13, CmTCP-18, CmTCP-26, CmTCP-27,* and *CmTCP-29* are highly connected to other proteins and bear main nodes. The result suggested that these



bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

proteins are possibly responsible for diverse functions in the melon genome.

Fig. 9 Predicted network of protein-protein interaction of *CmTCP* proteins. Each node indicates a *CmTCP* protein and lines indicate association among them. Different colours indicate clusters predicted using the K-means method. Each node represents a protein and the lines are for interactions among them.

DISCUSSION

The *TCP* gene family has been found to be involved in a range of biological processes (Martin-Trillo and Cubas 2010) including light response (Chen et al. 2024), reproductive development (Viola and Gonzalez 2023), growth and development in plants (Jiang et al. 2023) and environmental adaptation (Wangdan Xiong and Juan 2022). We employed *in silico* bioinformatical analyses to identify and analyze the genome-wide *TCP* family genes, in a way to understand the genomic features, evolutionary processes and functional characteristics of those in melon.

TCP Family Genes: Numbers & Chromosomal Distribution

The numbers of *TCP* genes in different plant species generally ranges from 17 to 75 genes (Li et al. 2022) and varies widely among species such as Arabidopsis (24), Cucumber (27), Potato (23), Rice (22), and Cotton (74) etc. (Bao et al. 2019; Li 2015; Wen et al. 2020; Yao et al. 2007; Zheng et al. 2018). We have identified 29 *TCP* genes in the melon genome (Table 1). The probable reason for the variation in the number of *TCP* genes might be the differences in genome size, chromosome numbers, gene duplication and evolutionary pathway etc. (Huang et al. 2023). Most of the genes seemed to be distributed throughout the chromosomes and mostly toward the telomeric region (Figure 1). Similar patterns of chromosomal distributions of *TCP* family genes were observed in the case of *Melastoma candidum* (Li et al. 2022), Ginger (Jiang et al. 2023), Tea (Shang et al. 2022), and *Dendrobium chrysotoxum* (Huang et al. 2023). All of the *CmTCP* proteins are predicted to be localized in the nucleus except *CmTCP-14* which is localized in cytoplasm (Table 1). As *CmTCP* proteins are located at the nucleus, they can regulate gene expression as transcription factors binding with DNA target sequences (Xiao et al. 2022).

Gene Structure, Motif and Protein Domains

The structure of most of the *CmTCP* genes is very simple with either only one or no introns at all (Figure 3C). Of the 29 *CmTCP* genes, only 8 genes have introns. This may be because

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

introns were lost during the evolution of gene structures (Ma et al. 2014). Intron-free genes often respond fast to stress, even though they are not suitable for recombination or species evolution (Sang et al. 2016). Furthermore, as shown in Figure 3, *CmTCP* genes within the same group displayed similarity in exon-intron structures and conserved motifs, indicating a strong evolutionary relationship among these genes. Every protein group (Figure 3A), with a few notable exceptions, has a distinct variation in the distribution of their motifs, such as the unique distribution of motifs 3 and 4 (Figure 3B). The complex structure of *TCP* proteins revealed that they contain a distinct DNA-binding motif (Zhang et al. 2023b).

Despite some variations, proteins of each group also exhibited a similar pattern in their structures, especially in the number of alpha helix and beta sheets (Supplementary File 3). The functional differentiation of the *CmTCP* protein is supported by the unique motif distribution and structural homology among different groups. A basic-helix-loop-helix (bHLH) structure with 59 amino acids residue was found to form the *TCP* domain in melon (Figure 2) which is similar to the findings of previous studies (Aggarwal et al. 2010; Shang et al. 2022). Sequence alignment of 29 *CmTCP* genes exhibits a highly conserved *TCP* domain which can be divided into two classes (Li 2015). The classification is based on 4 residual differences within the basic region (Figure 2). Li (Li 2015) explained the event as a residual insertion in Class-II genes. Through the *bHLH* structure and sequence similarity with other species (Jiang et al. 2023; Li 2015; Ma et al. 2014; Shang et al. 2022), it can be assumed that the *TCP* domain region is relatively conserved in the melon evolutionary process.

Evolutionary Relationship

The sequence differences of *TCP* genes throughout the species could be the result of diverse plant selection processes such as natural selection, mutation or evolution. Using whole-genome data from several species, synteny analysis can be used to hypothesize the locations of both orthologous and paralogous genes (Cao et al. 2016; Lin et al. 2014). Moreover, Gene duplication can produce new genetic material that can be used by selection, drift, and mutation to produce specialized or novel gene functions (Magadum et al. 2013). The variation within the *TCP* genes indicated that *CmTCP* genes might be involved during the process of evolution from its common and close ancestors (Figure 6) and gene duplication results in abundance and functional differences of *CmTCP* genes. Phylogenetic analyses of homologous genes across multiple genomes can provide comprehensive insights into the extent of evolutionary and functional diversification within a specific gene family

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

(Stavrinides and Ochman 2009).The phylogenetic analysis showed that *CmTCP* genes are closely related with *CsTCP* genes indicating the close evolutionary relationship between melon and cucumber. Some of the *CmTCP*s were found to be closely clustered with homologs of cucumber, this suggested that these proteins might serve similar roles in melon and cucumber (Pearson 2013).

Cis-regulating Elements

Several important cis-regulating elements were found to be present in the promoter analysis. These elements included those related to the expression of meristems and endosperms, responses to light, hormones (such as gibberellin, auxin, salicylic acid, MeJA, and auxin), stress (anaerobic, low temperature), and drought response. These results suggest that the *CmTCP* gene family may have roles in regulating hormone levels to govern melon growth and development in addition to participating in responses to abiotic stresses.

Putative Functions of *TCP* gene family

Class II CYC/TB1 clade *TCP* genes are primarily involved in controlling the development of lateral branches or flowers by regulating axillary meristems (Zhang et al. 2023a). The proteins are known to control the transition from cell division to leaf expansion (Kazama et al. 2010; Qi et al. 2017). The size, shape, flatness, and complexity of the leaf lamina are significantly determined by the time of the transition. The *CYC/TB1* group of proteins essentially assist in the regulation of axillary meristem development, which leads to the formation of either lateral branches or flowers (Tähtiharju et al. 2011; Yu et al. 2021). Insights into gene function are frequently obtained by analyzing the temporal and spatial expression patterns of a gene within the species . In order to find out if *CmTCP* plays any crucial role in growth and development, especially in lateral branching, we consequently looked at the expression patterns to understand gene expression and phylogeny of the proteins to predict functional homology by comparing with the genes of other species with known roles.

Expressions of *TCP* family genes in melon

The expression pattern reveals that almost every *CmTCP* gene, except *CmTCP-23* from the *PCF* group (Class-I) is moderately to highly expressed in every tissue (Figure 7, Supplementary Table 4). It indicates the possible involvement of class-I genes in the overall growth and development of various organs including leaves, flowers, fruits and others (Viola

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

et al. 2023). In previous studies, it was found that class-I genes are involved in the elongation of leaves and internode (Kieffer et al. 2011; Koyama et al. 2010), pollen and flower bud development (Balsemao-Pires et al. 2013; Takeda et al. 2006), trichome branching (Camoirano et al. 2020), and fruit ripening (Parapunova et al. 2014). *SlTCP18* a homolog of *CmTCP-18* was predicted to control the fruit ripening in tomatoes (Parapunova et al. 2014). Though the genes were closely related to each other, no expression-data in fruit tissue was found for *CmTCP-18* gene (Figure 7, Supplementary Table 4).

The *CIN* group of *TCP* genes are found to be expressed highly in flower buds and control plant flowering (Lan and Qin 2020; Wen et al. 2020). *Petunia* PhLA, a *CIN-TCP* gene, has been shown to be involved in flower morphogenesis (Chen et al. 2020b). Most of the *CsTCP* genes from the *CIN* group in cucumber were found to be involved in the regulation of female flower bud development. Homologs of these genes like *CmTCP-09, CmTCP-12,* and *CmTCP-29* showed higher expression in floral tissue indicating their potential involvement in floral development. As *CmTCP-12,* and *CmTCP-29* are expressed in every type of tissues, these genes might have roles in the overall growth and development (Figure 7, Supplementary Table 4). In addition, *CmTCP-09* might be responsible for floral growth and development as it was found to be expressed in mostly floral parts.

Previous studies suggest that *CYC/TB1* genes are responsible for the regulation of branching in different plants (Bai et al. 2012; Poza-Carrion et al. 2007; Shang et al. 2022; Wen et al. 2020; Zhao et al. 2018). *CsTCP21,* a homolog of *CmTCP-15* was found to be involved in the regulation of tendril growth and development. Similar expression data indicates the involvement of *CmTCP-15* in tendril growth and development as the gene was found to be expressed only in tendril (Figure 7, Supplementary Table 4). *CsTCP3* and *CsTCP7, an* ortholog of *AtTCP18* (AT3G18550), were found to be expressed in auxiliary buds of cucumber and *CsTCP3* was predicted to be responsible for lateral branching in cucumber (Aguilar-Martinez et al. 2007; van Es et al. 2019; Wen et al. 2020). Interestingly, *CmTCP-03* and *CmTCP-27* were closely related with *CsTCP3, CsTCP7,* and *AtTCP18* (Figure 4). The expression data indicated that *CmTCP-03* and *CmTCP-27* were not expressed in any of the tissues (Figure 7, Supplementary Table 4), probably due to the lack of expression data from the auxiliary bud. Though the Melonet DB does not contain any expression-related data from the auxiliary buds, based on the homologs, we can predict that these genes might be organ-specific (expressed in auxiliary bud) and probably regulate lateral branching in melon as well. The primary endogenous regulators of plant growth and development are GA and ethylene

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

(Katyayini et al. 2020; Wang et al. 2010). The presence of TATC-motif also justifies the involvement of *CmTCP-27* in lateral branching as TATC-motif is a GA-related cis-element and was found to affect the growth and development of lateral organs (Wen et al. 2020).

CONCLUSION

The present study identifies 29 *TCP* genes in the melon genome, distributed mainly in the telomeric regions of 9 chromosomes. Structure, motif, domain distribution and sequence alignment indicated that the genes are highly conserved in this species. Gene duplication and synteny analysis depict that the genes are developed during evolution as well as from duplication within the genome. Phylogeny and expression data indicated that *CmTCP-09* might be responsible for floral development, and *CmTCP-03* and *CmTCP-27* probably regulate lateral branching in melon. The findings paves the way for any future functional genomic study aiming at elucidating the exact functions of the *TCP* genes in melon which then can be used for improvement of melon via breeding and biotechnological means.

Acknowledgements

Not applicable

Author Contributions

MJHJ, MNAS and MRH were responsible for conceptualization. MJHJ and MNAS was responsible for data collection and curation. MJHJ was responsible for analysis and writing the original draft. MKB was responsible for synteny analysis. MRH were responsible for fund acquisition, supervising the work, reviewing and finalizing the draft. All authors have read and agreed to the final version of the manuscript.

Funding

This study was supported by the Bangladesh Agricultural University, Mymensingh-2202 (Project No: 2021/1042/BAU) and the Ministry of Science and Technology, Bangladesh (Project No: SRG-231143). The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Ethical approval

Not applicable.

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Informed consent

Not applicable.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

Aggarwal P, Das Gupta M, Joseph AP, Chatterjee N, Srinivasan N, Nath U (2010) Identification of Specific DNA Binding Residues in the TCP Family of Transcription Factors in Arabidopsis The Plant Cell 22:1174-1189

Aguilar-Martinez JA, Poza-Carrion C, Cubas P (2007) Arabidopsis BRANCHED1 acts as an integrator of branching signals within axillary buds. The Plant Cell 19:458-472

Aguilar-Martinez JA, Sinha N (2013) Analysis of the role of Arabidopsis Class I TCP genes AtTCP7, AtTCP8, AtTCP22, and AtTCP23 in leaf development. Frontiers in Plant Science 4:406

Alzohairy A (2011) BioEdit: An important software for molecular biology. GERF Bulletin of Biosciences 2:60-61

Bai F, Reinheimer R, Durantini D, Kellogg EA, Schmidt RJ (2012) TCP transcription factor, BRANCH ANGLE DEFECTIVE 1 (BAD1), is required for normal tassel branch angle formation in maize. Proc Natl Acad Sci U S A 109:12225-12230

Bailey TL, Johnson J, Grant CE, Noble WS (2015) The MEME Suite. Nucleic Acids Res 43:W39-49

Balsemao-Pires E, Andrade LR, Sachetto-Martins G (2013) Functional study of TCP23 in Arabidopsis thaliana during plant development. Plant Physiol Biochem 67:120-125

Bao S, Zhang Z, Lian Q, Sun Q, Zhang R (2019) Evolution and expression of genes encoding TCP transcription factors in Solanum tuberosum reveal the involvement of StTCP23 in plant defence. BMC Genetics 20:91

Camoirano A, Arce AL, Ariel FD, Alem AL, Gonzalez DH, Viola IL (2020) Class I TCP transcription factors regulate trichome branching and cuticle development in Arabidopsis. J Exp Bot 71:5438-5453

Cao Y, Han Y, Jin Q, Lin Y, Cai Y (2016) Comparative Genomic Analysis of the GRF Genes in Chinese Pear (Pyrus bretschneideri Rehd), Poplar (Populous), Grape (Vitis vinifera), Arabidopsis and Rice (Oryza sativa). Front Plant Sci 7:1750

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R (2020a) TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. Mol Plant 13:1194-1202

Chen C, Zhang Y, Chen Y, Chen H, Gong R (2024) Sweet cherry TCP gene family analysis reveals potential functions of PavTCP1, PavTCP2 and PavTCP3 in fruit light responses. BMC Genomics 25

Chen HW, Lee PL, Wang CN, Hsu HJ, Chen JC (2020b) Silencing of PhLA, a CIN-TCP gene, causes defected petal conical epidermal cell formation and results in reflexed corolla lobes in petunia. Botanical Studies 61:24

Cubas P, Lauter N, Doebley J, Coen E (1999) The TCP domain: a motif found in proteins regulating plant growth and development. Plant J 18:215-222

Daviere JM, Wild M, Regnault T, Baumberger N, Eisler H, Genschik P, Achard P (2014) Class I TCP-DELLA interactions in inflorescence shoot apex determine plant height. Curr Biol 24:1923-1928

Doebley J, Stec A, Hubbard L (1997) The evolution of apical dominance in maize. Nature 386:485-488

Finlayson SA (2007) Arabidopsis Teosinte Branched1-like 1 regulates axillary bud outgrowth and is homologous to monocot Teosinte Branched1. Plant Cell Physiol 48:667-677

Fukino N, Ohara T, Sugiyama M, Kubo N, Hirai M, Sakata Y, Matsumoto S (2012) Mapping of a gene that confers short lateral branching (slb) in melon (Cucumis melo L.). Euphytica 187:133-143

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31:3784-3788

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N et al. (2012) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res 40:D1178-1186

Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G (2015) GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31:1296-1297

Huang Y, Zhao X, Zheng Q, He X, Zhang MM, Ke S, Li Y, Zhang C, Ahmad S, Lan S et al. (2023) Genome-Wide Identification of TCP Gene Family in Dendrobium and Their Expression Patterns in Dendrobium chrysotoxum. Int J Mol Sci 24

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Jiang Y, Jiang D, Xia M, Gong M, Li H, Xing H, Zhu X, Li HL (2023) Genome-Wide Identification and Expression Analysis of the TCP Gene Family Related to Developmental and Abiotic Stress in Ginger. Plants (Basel) 12:3389

Jin J, Tian F, Yang DC, Meng YQ, Kong L, Luo J, Gao G (2017) PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Res 45:D1040-D1045

Kalaivani S, Ramyachitra D, Manikandan P (2018) K-means Clustering: An Efficient Algorithm for Protein Complex Detection. In: Progress in Computing, Analytics and Networking. Advances in Intelligent Systems and Computing. pp 449-459

Katyayini NU, Rinne PLH, Tarkowská D, Strnad M, van der Schoot C (2020) Dual Role of Gibberellin in Perennial Shoot Branching: Inhibition and Activation. Front Plant Sci 11:736

Kazama T, Ichihashi Y, Murata S, Tsukaya H (2010) The mechanism of cell cycle arrest front progression explained by a KLUH/CYP78A5-dependent mobile growth factor in developing leaves of Arabidopsis thaliana. Plant Cell Physiol 51:1046-1054

Kieffer M, Master V, Waites R, Davies B (2011) TCP14 and TCP15 affect internode length and leaf shape in Arabidopsis. Plant J 68:147-158

Kolde R (2019) Pheatmap: pretty heatmaps. 1:726

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

Kosugi S, Ohashi Y (2002) DNA binding and dimerization specificity and potential targets for the TCP protein family. Plant J 30:337-348

Koyama T, Sato F, Ohme-Takagi M (2010) A role of TCP1 in the longitudinal elongation of leaves in Arabidopsis. Biosci Biotechnol Biochem 74:2145-2147

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular Biology and Evolution 33:1870-1874

Lan J, Qin G (2020) The Regulation of CIN-like TCP Transcription Factors. Int J Mol Sci 21

Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouze P, Rombauts S (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res 30:325-327

Letunic I, Bork P (2021) Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res 49:W293-W296

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Li H, Wen X, Huang X, Wei M, Chen H, Yu Y, Dai S (2022) Genome-Wide Identification and Characterization of TCP Gene Family Members in Melastoma candidum. Molecules 27

Li S (2015) The Arabidopsis thaliana TCP transcription factors: A broadening horizon beyond development. Plant Signal Behav 10:e1044192

Li S, Zachgo S (2013) TCP3 interacts with R2R3-MYB proteins, promotes flavonoid biosynthesis and negatively regulates the auxin response in Arabidopsis thaliana. Plant J 76:901-913

Lin Y, Cheng Y, Jin J, Jin X, Jiang H, Yan H, Cheng B (2014) Genome duplication and gene loss affect the evolution of heat shock transcription factor genes in legumes. PLoS One 9:e102825

Ma J, Wang Q, Sun R, Xie F, Jones DC, Zhang B (2014) Genome-wide identification and expression analysis of TCP transcription factors in Gossypium raimondii. Scientific Reports 4:6645

Magadum S, Banerjee U, Murugan P, Gangapur D, Ravikesavan R (2013) Gene duplication as a major force in evolution. J Genet 92:155-161

Manassero NG, Viola IL, Welchen E, Gonzalez DH (2013) TCP transcription factors: architectures of plant form. Biomol Concepts 4:111-127

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

Martin-Trillo M, Grandio EG, Serra F, Marcel F, Rodriguez-Buey ML, Schmitz G, Theres K, Bendahmane A, Dopazo H, Cubas P (2011) Role of tomato BRANCHED1-like genes in the control of shoot branching. Plant J 67:701-714

Michael N, Pilar C (2016) Chapter 16 - The Role of TCP Transcription Factors in Shaping Flower Structure, Leaf Morphology, and Plant Architecture. In: Gonzalez DH (ed) Plant Transcription Factors: Evolutionary, Structural and Functional Aspects. pp 249-267

Navaud O, Dabos P, Carnus E, Tremousaygue D, Herve C (2007) TCP transcription factors predate the emergence of land plants. J Mol Evol 65:23-33

Parapunova V, Busscher M, Busscher-Lange J, Lammers M, Karlova R, Bovy AG, Angenent GC, de Maagd RA (2014) Identification, cloning and characterization of the tomato TCP transcription factor family. BMC Plant Biol 14:157

Pearson WR (2013) An introduction to sequence similarity ("homology") searching. Curr Protoc Bioinformatics Chapter 3:3.1.1-3.1.8

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Pereira FHF, Nogueira IC, Pedrosa JF, Negreiros MZ, Bezerra Neto FJHB (2003) Poda da haste principal e densidade de cultivo na produção e qualidade de frutos em híbridos de melão. Horticultura Brasileira 21:192-197

Poza-Carrion C, Aguilar-Martinez JA, Cubas P (2007) Role of TCP Gene BRANCHED1 in the Control of Shoot Branching in Arabidopsis. Plant Signal Behav 2:551-552

Qi J, Wu B, Feng S, Lu S, Guan C, Zhang X, Qiu D, Hu Y, Zhou Y, Li C et al. (2017) Mechanical regulation of organ asymmetry in leaves. Nat Plants 3:724-733

Sang Y, Liu Q, Lee J, Ma W, McVey DS, Blecha F (2016) Expansion of amphibian intronless interferons revises the paradigm for interferon evolution and functional diversity. Sci Rep 6:29072

Shahwar D, Khan Z, Park Y (2023) Molecular Marker-Assisted Mapping, Candidate Gene Identification, and Breeding in Melon (Cucumis melo L.): A Review. Int J Mol Sci 24:15490

Shang X, Han Z, Zhang D, Wang Y, Qin H, Zou Z, Zhou L, Zhu X, Fang W, Ma Y (2022) Genome-Wide Analysis of the TCP Gene Family and Their Expression Pattern Analysis in Tea Plant (Camellia sinensis). Front Plant Sci 13:840350

Stavrinides J, Ochman H (2009) Phylogenetic Methods. In: Schaechter M (ed) Encyclopedia of Microbiology (Third Edition). Academic Press, Oxford, pp 247-260

Su X, Wang B, Geng X, Du Y, Yang Q, Liang B, Meng G, Gao Q, Yang W, Zhu Y et al. (2021) A high-continuity and annotated tomato reference genome. BMC Genomics 22:898

Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S et al. (2023) The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res 51:D638-D646

Tähtiharju S, Rijpkema AS, Vetterli A, Albert VA, Teeri TH, Elomaa P (2011) Evolution and Diversification of the CYC/TB1 Gene Family in Asteraceae—A Comparative Study in Gerbera (Mutisieae) and Sunflower (Heliantheae). Molecular Biology and Evolution 29:1155-1166

Takeda T, Amano K, Ohto MA, Nakamura K, Sato S, Kato T, Tabata S, Ueguchi C (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, Stecher G, Kumar S (2021) MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol 38:3022-3027

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

van Es SW, van der Auweraert EB, Silveira SR, Angenent GC, van Dijk ADJ, Immink RGH (2019) Comprehensive phenotyping reveals interactions and functions of Arabidopsis thaliana TCP genes in yield determination. Plant J 99:316-328

Viola IL, Alem AL, Jure RM, Gonzalez DH (2023) Physiological Roles and Mechanisms of Action of Class I TCP Transcription Factors. Int J Mol Sci 24

Viola IL, Gonzalez DH (2023) TCP Transcription Factors in Plant Reproductive Development: Juggling Multiple Roles. Biomolecules 13:750

Wang DH, Li F, Duan QH, Han T, Xu ZH, Bai SN (2010) Ethylene perception is involved in female cucumber flower development. Plant J 61:862-872

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

Wangdan Xiong YZHGYLWTLMGY, Juan S (2022) Genomic characterization and expression analysis of TCP transcription factors in Setaria italica and Setaria viridis. Plant Signaling & Behavior 17:2075158

Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L et al. (2018) SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46:W296-W303

Wen H, Chen Y, Du H, Zhang L, Zhang K, He H, Pan J, Cai R, Wang G (2020) Genome-Wide Identification and Characterization of the TCP Gene Family in Cucumber (Cucumis sativus L.) and Their Transcriptional Responses to Different Treatments. Genes (Basel) 11

Widaryanto E, Putri MR, Yamika WSD, Saitama A, Zaini AH (2020) The Effect of Leaf Bud Trimming and Fruit Position Arrangement on the Quality of Golden Melon *(Cucumis melo L.)*. Acta Agrobotanica 73:1

Wu X, Cheng C, Ma R, Xu J, Ma C, Zhu Y, Ren Y (2023) Genome-wide identification, expression analysis, and functional study of the bZIP transcription factor family and its response to hormone treatments in pea (Pisum sativum L.). BMC Genomics 24:705

Xiao X, Lin W, Feng E, Wu C, Ou X (2022) Genome-Wide Identification of Binding Sites for SmTCP7a Transcription Factors of Eggplant during Bacterial Wilt Resistance by ChIP-Seq. Int J Mol Sci 23

Yamaguchi-Shinozaki K, Shinozaki K (2005) Organization of cis-acting regulatory elements in osmotic- and cold-stress-responsive promoters. Trends Plant Sci 10:88-94

bioRxiv preprint doi: https://doi.org/10.1101/2024.07.30.605410; this version posted August 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Yano R, Ariizumi T, Nonaka S, Kawazu Y, Zhong S, Mueller L, Giovannoni JJ, Rose JKC, Ezura H (2020) Comparative genomics of muskmelon reveals a potential role for retrotransposons in the modification of gene expression. Commun Biol 3:432

Yao X, Ma H, Wang J, Zhang D (2007) Genome□Wide Comparative Analysis and Expression Pattern of TCP Gene Families in Arabidopsis thaliana and Oryza sativa. Journal of Integrative Plant Biology 49:885-897

Yu CS, Chen YC, Lu CH, Hwang JK (2006) Prediction of protein subcellular localization. Proteins 64:643-651

Yu J, Wu S, Sun H, Wang X, Tang X, Guo S, Zhang Z, Huang S, Xu Y, Weng Y et al. (2022) CuGenDBv2: an updated database for cucurbit genomics. Nucleic Acids Research 51:D1457-D1464

Yu S, Li P, Zhao X, Tan M, Ahmad MZ, Xu Y, Tadege M, Zhao J (2021) CsTCPs regulate shoot tip development and catechin biosynthesis in tea plant (Camellia sinensis). Hortic Res 8:104

Zhang M, Qin S, Yan J, Li L, Xu M, Liu Y, Zhang W (2023a) Genome-wide identification and analysis of TCP family genes in Medicago sativa reveal their critical roles in Na(+)/K(+) homeostasis. BMC Plant Biol 23:301

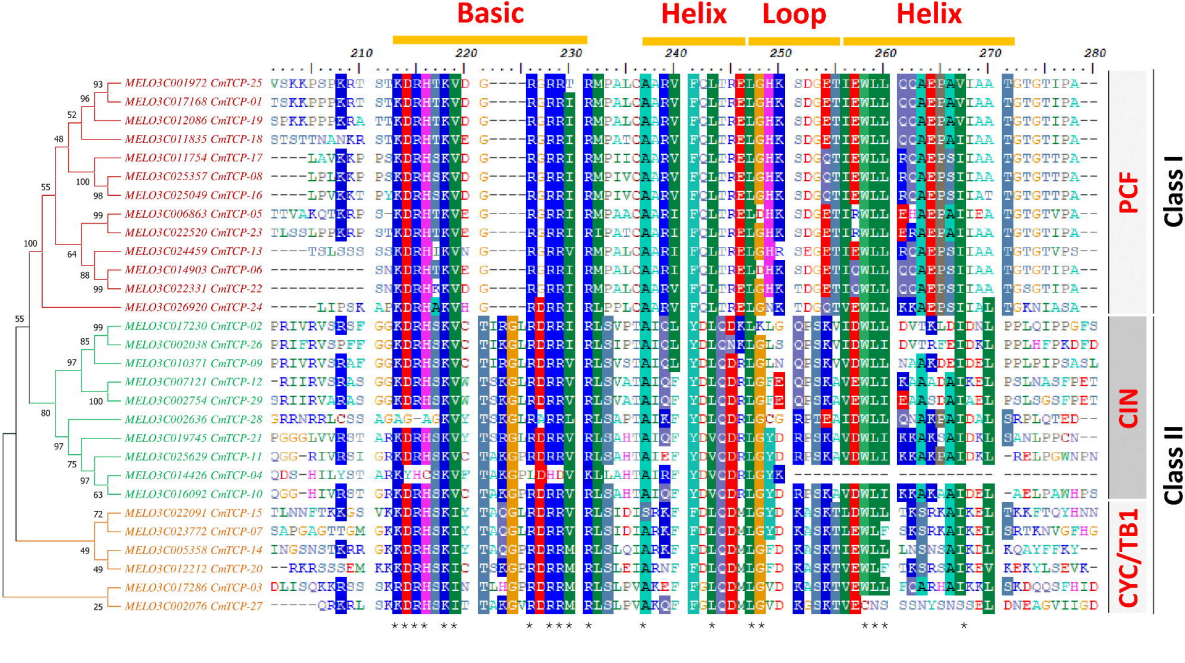
Zhang Y, Xu YP, Nie JK, Chen H, Qin G, Wang B, Su XD (2023b) DNA-TCP complex structures reveal a unique recognition mechanism for TCP transcription factor families. Nucleic Acids Res 51:434-448

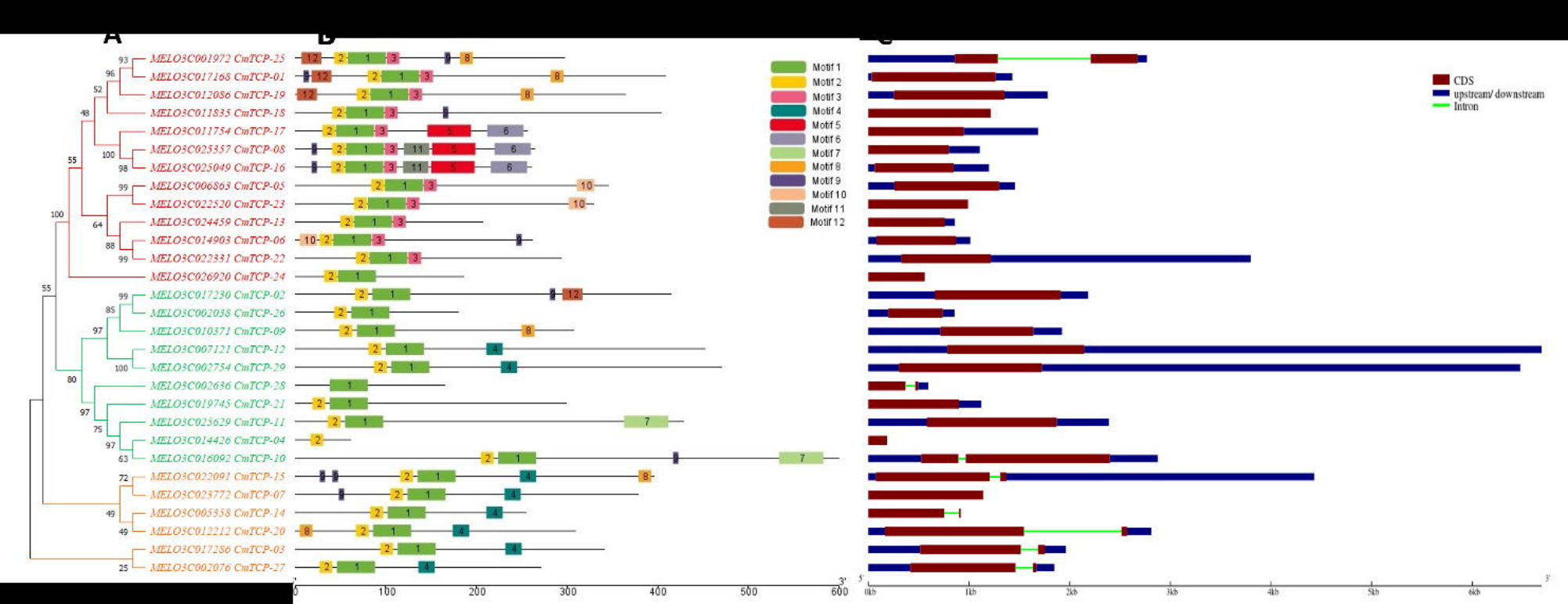
Zhao Y, Pfannebecker K, Dommes AB, Hidalgo O, Becker A, Elomaa P (2018) Evolutionary diversification of CYC/TB1-like TCP homologs and their recruitment for the control of branching and floral morphology in Papaveraceae (basal eudicots). New Phytol 220:317-331

Zheng K, Ni Z, Qu Y, Cai Y, Yang Z, Sun G, Chen Q (2018) Genome-wide identification and expression analyses of TCP transcription factor genes in Gossypium barbadense. Sci Rep 8:14526

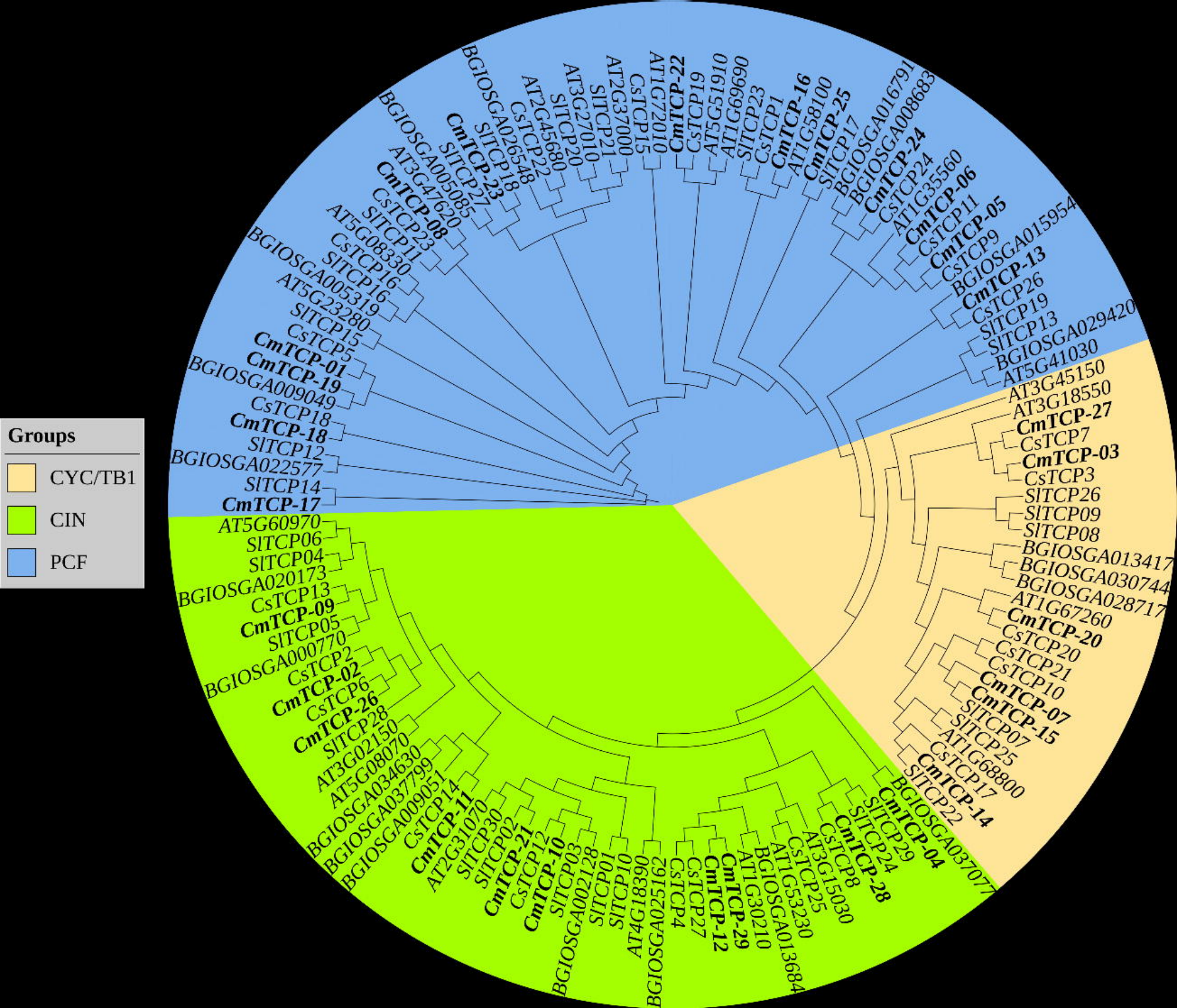
Zheng Y, Jiao C, Sun H, Rosli HG, Pombo MA, Zhang P, Banf M, Dai X, Martin GB, Giovannoni JJ et al. (2016) iTAK: A Program for Genome-wide Prediction and Classification of Plant Transcription Factors, Transcriptional Regulators, and Protein Kinases. Mol Plant 9:1667-1670

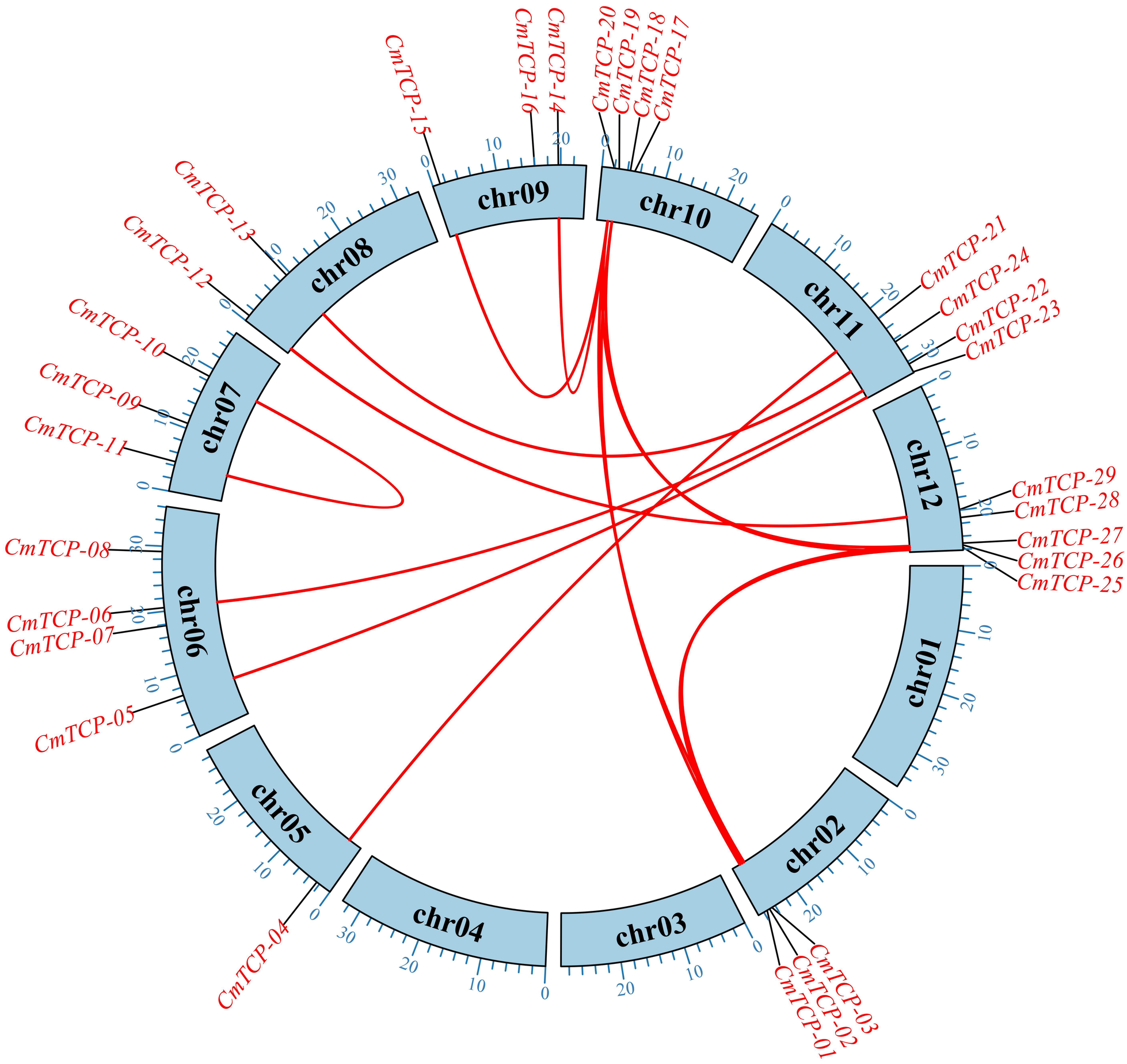
0 Mb
5 Mb
10 Mb
15 Mb
20 Mb
25 Mb
30 Mb
35 Mb
40 Mb
2
CmTCP03
CmTCP02
CmTCP01
5
CmTCP04
6
CmTCP05
CmTCP07
CmTCP06
CmTCP08
7
CmTCP11
CmTCP09
CmTCP10
8
CmTCP12
CmTCP13
9
CmTCP15
CmTCP16
CmTCP14
10
CmTCP20
CmTCP19
CmTCP18
CmTCP17
11
CmTCP21
CmTCP24
CmTCP22
CmTCP23
12
CmTCP29
CmTCP28
CmTCP27
CmTCP26
CmTCP25


Basic
Helix
Loop
Helix
210
220
230
240
250
260
270
280
MELO3C001972 CmTCP-25
MELO3C017168 CmTCP-01
MELO3C012086 CmTCP-19
MELO3C011835 CmTCP-18
MELO3C011754 CmTCP-17
MELO3C025357 CmTCP-08
MELO3C025049 CmTCP-16
MELO3C006863 CmTCP-05
MELO3C022520 CmTCP-23
MELO3C024459 CmTCP-13
MELO3C014903 CmTCP-06
MELO3C022331 CmTCP-22
MELO3C026920 CmTCP-24
MELO3C017230 CmTCP-02
MELO3C002038 CmTCP-26
MELO3C010371 CmTCP-09
MELO3C007121 CmTCP-12
MELO3C002754 CmTCP-29
MELO3C002636 CmTCP-28
MELO3C019745 CmTCP-21
MELO3C025629 CmTCP-11
MELO3C014426 CmTCP-04
MELO3C016092 CmTCP-10
MELO3C022091 CmTCP-15
MELO3C023772 CmTCP-07
MELO3C005358 CmTCP-14
MELO3C012212 CmTCP-20
MELO3C017286 CmTCP-03
MELO3C002076 CmTCP-27
PCF
CIN
CYC/TB1
Class I
Class II

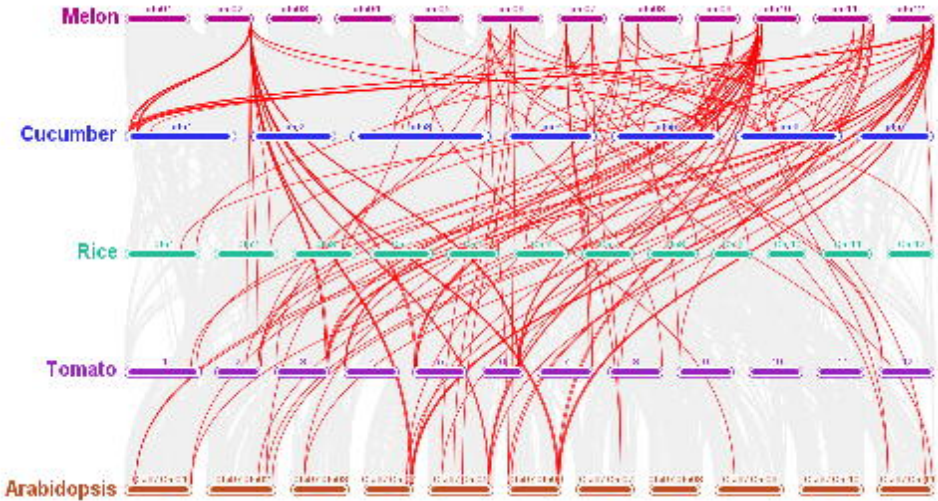

A
B
C
MELO3C001972 CmTCP-25
MELO3C017168 CmTCP-01
MELO3C012086 CmTCP-19
MELO3C011835 CmTCP-18
MELO3C011754 CmTCP-17
MELO3C025357 CmTCP-08
MELO3C025049 CmTCP-16
MELO3C006863 CmTCP-05
MELO3C022520 CmTCP-23
MELO3C024459 CmTCP-13
MELO3C014903 CmTCP-06
MELO3C022331 CmTCP-22
MELO3C026920 CmTCP-24
MELO3C017230 CmTCP-02
MELO3C002038 CmTCP-26
MELO3C010371 CmTCP-09
MELO3C007121 CmTCP-12
MELO3C002754 CmTCP-29
MELO3C002636 CmTCP-28
MELO3C019745 CmTCP-21
MELO3C025629 CmTCP-11
MELO3C014426 CmTCP-04
MELO3C016092 CmTCP-10
MELO3C022091 CmTCP-15
MELO3C023772 CmTCP-07
MELO3C005358 CmTCP-14
MELO3C012212 CmTCP-20
MELO3C017286 CmTCP-03
MELO3C002076 CmTCP-27
Motif 1
Motif 2
Motif 3
Motif 4
Motif 5
Motif 6
Motif 7
Motif 8
Motif 9
Motif 10
Motif 11
Motif 12
CDS
upstream/ downstream
Intron
0
100
200
300
400
500
600
0kb
1kb
2kb
3kb
4kb
5kb
6kb

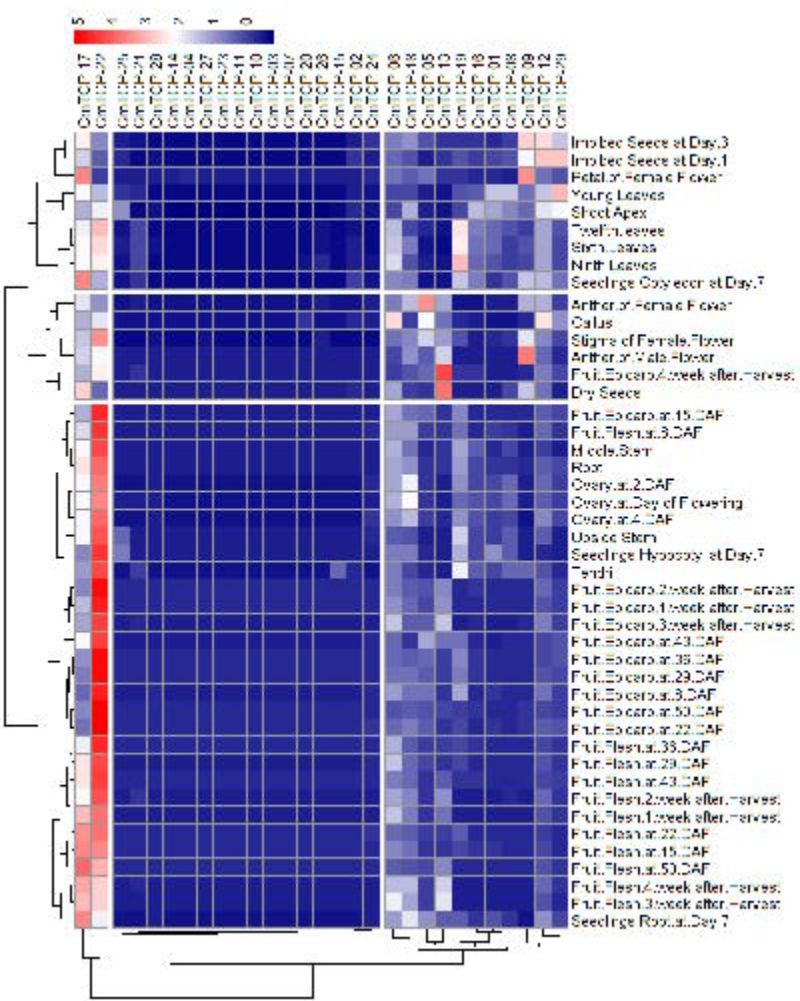
	Function	Elements	*CmTCP-01*	*CmTCP-02*	*CmTCP-03*	*CmTCP-04*	*CmTCP-05*	*CmTCP-06*	*CmTCP-07*	*CmTCP-08*	*CmTCP-09*	*CmTCP-10*	*CmTCP-11*	*CmTCP-12*	*CmTCP-13*	*CmTCP-14*	*CmTCP-15*	*CmTCP-16*	*CmTCP-17*	*CmTCP-18*	*CmTCP-19*	*CmTCP-20*	*CmTCP-21*	*CmTCP-22*	*CmTCP-23*	*CmTCP-24*	*CmTCP-25*	*CmTCP-26*	*CmTCP-27*	*CmTCP-28*	*CmTCP-29*
Light Responsive	Light	Box 4	7	3	2	5	4	3	1	2	11		2			4	2	9	3	3	5	6	2	4	5			2	3	5	5
		G-Box	3	4	3	5	4	2	5	2	2	1	4	4		3		1	1	2	1	3		3	2	2	3	1	4	1	
		GT1-motif	1	1			3	1	3	1	2	4	2	1	1	5	1	1	2	3		4		2	2	3	1	2		2	1
		I-box	1		1		1		1		1										1		1	1			1		2		
		TCCC-motif	2		3	2			1			1			2	1			3	1	1		1				1				
		TCT-motif				1	2	2				1				2	1		1	1			3			2	1				1
Phytohormone Responsive	MeJA	CGTCA-motif		1			2		1	2	1	1	2	1					1	1	1		1	1			2				
		TGACG-motif		1			2		1	2	1	1	2	1					1	1	1		1	1			2				
	ABA	ABRE	2	3	2	4	4	2	4	2	4	1	4	1		4		1	1	2	1	3		2	2	2	3	1	3		1
		DRE1																											1		
	Auxin	AuxRR-core																	1						1						
		TGA-element				1			3							1											1				
	Ethylene	ERE					1	1	4	3		1	1		1	1	1	2	2	1		5	3	3	1	3					1
	SA	TCA-element	2	2	1	1	3	5	3		1					1				1		2		2	1	1	1		2	1	2
	Gibberellin	GARE-motif	2	1	1	1	1	2	1			1		1	1	1		1		1	2		1		1			1	1	2	
		P-box			2							1	1		2		2		1	1	1	1	3			1		1	1	1	
		TATC-box																											1		
Stress Responsive	Drought	as-1		1			2		1	2	1	1	2	1					1	1	1		1	1			2				
		MBS		1			1						2		1	1	1	1	2				2				1				
		MYC	7	2	3	7	5	4	8	3	6	1	6	4	3	6	6	6	4	3	2	2	5	2	5	5	4	4	1	4	2
	Wound/Pathogen	WRE3				1		1		1	1	2	1		1			1	1	1					3	1	1			4	2
	Wound	WUN-motif						2	3	2			2		1				2	2		2		1	1	3	2	1	3		1
	Anaerobic induction	ARE		5	3	2	1	1	2	1	4	2	1	1	2	4	2	1	1	3		2	3	4	1	2	2	2	2	3	4
	Cold/Dehydration	DRE core		1																										1	
	Anoxic Specific inducibility	GC-motif				1																									
	Low Temperature	LTR			1		1							2	1	1				1			1			1		1			
	Stress	STRE	5	2	2		1	4	1	1	1	5	3	1	1	1	3	1	4	5	8		1	1	4		3		1		3
	Defence/Stress	TC-rich repeats	1	2					1					1	1		1	1	2			1		1				1			
Growth & Development	Specific Day Length	A-box																						1							
	Zein Metabolism	O2-site	1		2	2	1				1									1									1		
	Meristem Specific gene activation	CCGTCC-box																						1							
	Seed Specific Expression	AAGAA-motif	5	4	4	1	4	3	3		2	6	2	4	1	2	7	1	2	3	2	2	3	1	2	2	4	5	3	1	

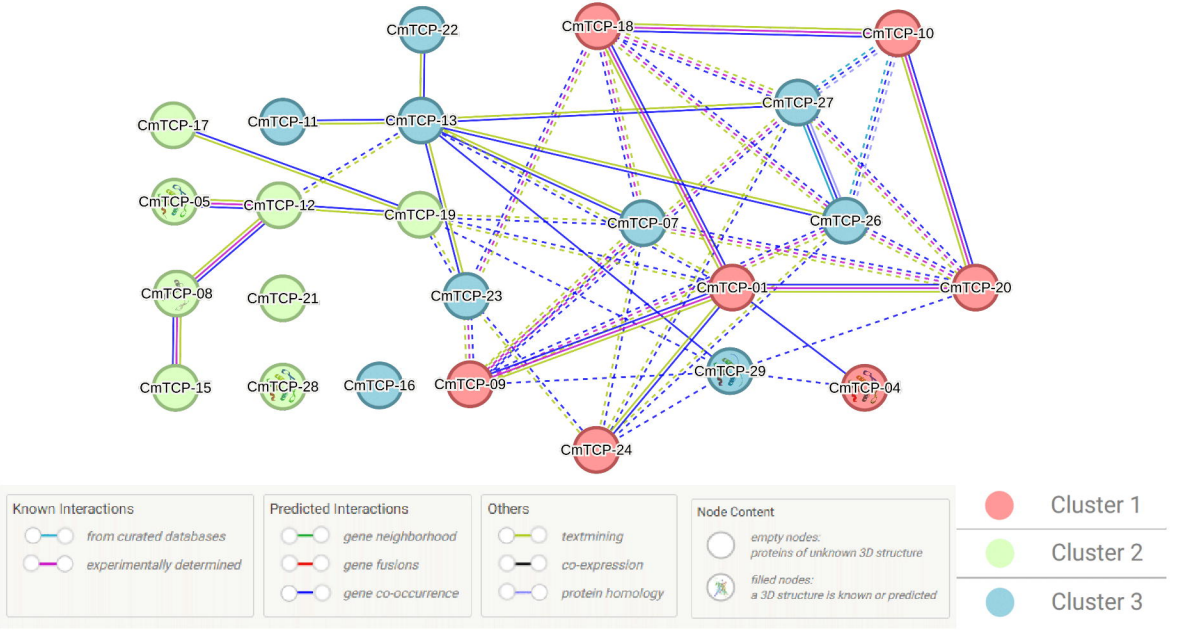

Groups
CYC/TB1
CIN
PCF
BGIOSGA016791
BGIOSGA008683
CmTCP-24
CsTCP24
AT1G35560
CmTCP-06
CsTCP11
CmTCP-05
CsTCP9
CmTCP-13
BGIOSGA015954
CmTCP-16
AT1G58100
CmTCP-25
SlTCP17
BGIOSGA029420
SlTCP13
BGIOSGA041030
AT5G41030
AT1G69690
SlTCP23
CsTCP1
AT5G51910
CsTCP19
CmTCP-22
AT1G72010
CsTCP15
AT2G37000
SlTCP21
AT3G27010
SlTCP20
AT2G45680
CsTCP22
BGIOSGA026548
SlTCP18
CmTCP-23
SlTCP27
BGIOSGA005085
AT3G47620
CmTCP-08
CsTCP23
SlTCP11
AT5G08330
CsTCP16
SlTCP16
BGIOSGA005319
AT5G23280
SlTCP15
CsTCP5
CmTCP-01
CmTCP-19
BGIOSGA009049
CsTCP18
CmTCP-18
SlTCP12
BGIOSGA022577
SlTCP14
CmTCP-17
AT3G45150
AT3G18550
CmTCP-27
CsTCP7
CmTCP-03
CsTCP3
SlTCP26
SlTCP09
SlTCP08
BGIOSGA013417
BGIOSGA030744
BGIOSGA028717
AT1G67260
CmTCP-20
CsTCP20
CsTCP21
CsTCP10
CmTCP-07
CmTCP-15
SlTCP07
SlTCP25
AT1G68800
CsTCP17
CmTCP-14
SlTCP22
BGIOSGA037077
CmTCP-04
SlTCP29
SlTCP24
CmTCP-28
CsTCP8
AT3G15030
CsTCP25
AT1G53230
BGIOSGA013684
AT1G30210
CmTCP-29
CmTCP-12
CsTCP27
CsTCP4
BGIOSGA025162
AT4G18390
SlTCP10
SlTCP01
BGIOSGA002128
SlTCP03
CmTCP-10
CsTCP12
CmTCP-21
SlTCP02
SlTCP30
AT2G31070
CmTCP-11
CsTCP14
BGIOSGA009051
BGIOSGA037799
AT5G08070
AT3G02150
SlTCP28
CmTCP-26
CsTCP6
CmTCP-02
CsTCP2
BGIOSGA000770
SlTCP05
CmTCP-09
CsTCP13
BGIOSGA020173
SlTCP04
SlTCP06
AT5G60970



bioRxiv preprint doi: https://doi.org/10.1101/2020.07.30.229012; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.


Melon
Cucumber
Rice
Tomato
Arabidopsis


Imrobec Seece at Day.3
Imrobec Seece at Day.1
Petalor Female Flower
Young Leaves
Shoot Apex
Twelfth Leaves
Stem Leaves
Ninth Leaves
Seedlings Cotyledon at Day.7
Anther of Female Flower
Calus
Stigma of Female Flower
Anther of Male Flower
Fruit Epicarp 4 week after Harvest
Dry Seece
Fruit Epicarp at 15 DAF
Fruit Flesh at 3 DAF
Middle Stem
Root
Ovary at 2 DAF
Ovary at Day of Flowering
Ovary at 4 DAF
Upper Stem
Seedlings Hypocotyl at Day.7
Tendril
Fruit Epicarp 2 week after Harvest
Fruit Epicarp 1 week after Harvest
Fruit Epicarp 3 week after Harvest
Fruit Epicarp at 43 DAF
Fruit Epicarp at 38 DAF
Fruit Epicarp at 29 DAF
Fruit Epicarp at 3 DAF
Fruit Epicarp at 50 DAF
Fruit Epicarp at 22 DAF
Fruit Flesh at 38 DAF
Fruit Flesh at 29 DAF
Fruit Flesh at 43 DAF
Fruit Flesh 2 week after Harvest
Fruit Flesh 1 week after Harvest
Fruit Flesh at 22 DAF
Fruit Flesh at 15 DAF
Fruit Flesh at 50 DAF
Fruit Flesh 4 week after Harvest
Fruit Flesh 3 week after Harvest
Seedlings Root at Day 7


CmTCP-22
CmTCP-18
CmTCP-10
CmTCP-27
CmTCP-17
CmTCP-11
CmTCP-13
CmTCP-05
CmTCP-12
CmTCP-19
CmTCP-07
CmTCP-26
CmTCP-08
CmTCP-21
CmTCP-23
CmTCP-01
CmTCP-20
CmTCP-15
CmTCP-28
CmTCP-16
CmTCP-09
CmTCP-29
CmTCP-04
CmTCP-24
Known Interactions
from curated databases
experimentally determined
Predicted Interactions
gene neighborhood
gene fusions
gene co-occurrence
Others
textmining
co-expression
protein homology
Node Content
empty nodes:
proteins of unknown 3D structure
filled nodes:
a 3D structure is known or predicted
Cluster 1
Cluster 2
Cluster 3