J Plant Biotechnol (2022) 49:15-29
DOI:https://doi.org/10.5010/JPB.2022.49.1.015

ISSN 1229-2818 (Print)
ISSN 2384-1397 (Online)

Research Article

Genome-wide analysis of *Solanum lycopersicum* L. cyclophilins

Khadiza Khatun · Arif Hasan Khan Robin · Md. Rafiqul Islam · Subroto Das Jyoti · Do-Jin Lee · Chang Kil Kim · Mi-Young Chung

Received: 16 March 2022 / Revised: 26 March 2022 / Accepted: March 26 2022
ⓒ Korean Society for Plant Biotechnology

Abstract Cyclophilins (CYPs) are highly conserved ubiquitous proteins belong to the peptidyl prolyl cis/trans isomerase (PPIase) superfamily. These proteins are present in a wide range of organisms; they contain a highly conserved peptidyl-prolyl cis/trans isomerase domain. A comprehensive database survey identified a total of 35 genes localized in all cellular compartments of *Solanum lycopersicum* L., but largely in the cytosol. Sequence alignment and conserved motif analyses of the SlCYP proteins revealed a highly conserved CLD motif. Evolutionary analysis predicted the clustering of a large number of gene pairs with high sequence similarity. Expression analysis using the RNA-Seq data showed that the majority of the *SlCYP* genes were highly expressed in mature leaves and blooming flowers, compared with their expression in other organs. This study provides a basis for the functional characterization of individual *CYP* genes in the future to elucidate their role(s) in protein refolding and long-distance signaling in tomatoes and in plant biology, in general.

Keywords Genome-wide analysis, cyclophilins, *Solanum lycopersicum* L., PPIase domain, expression profiles, evolutionary relation

K. Khatun · D.-J. Lee
Department of Biotechnology, Patuakhali Science and Technology University, Dumki, Patuakhali, Bangladesh-8602

A. H. K. Robin · S. D. Jyoti
Department of Genetics and Plant Breeding, Bangladesh Agricultural University, Mymensingh, Bangladesh-2202

Md. R. Islam
Department of Biotechnology, Sher-e-Bangla Agricultural University, Dhaka, Bangladesh

C. K. Kim (✉)
Department of Horticultural Science, Kyungpook National University, Daegu, South Korea
e-mail: ckkim@knu.ac.kr

M.-Y. Chung (✉)
Department of Agricultural Education, Sunchon National University, Suncheon, South Korea
e-mail: queen@sunchon.ac.kr

Introduction

Cyclophilins (CYPs), are highly conserved ubiquitous proteins found in all types of living organism including bacteria, fungi, mammals, plants and insects (Galat 1999). Cyclophilin was first identified in 1984 as a receptor of drug cyclosporin A in mammalian cells (Handschumacher et al. 1984). In plant, the *CYP* cDNA sequences were first identified from tomato, maize and oilseed rape (Gasser et al. 1990). Cyclophilins together with FK506-binding proteins (FKBPs) and parvulin-like proteins belong to a subgroup of protein called immunophilins (Lescot et al. 2002). CYPs, FKBPs and parvulin-like protein carry out the peptidyl prolyl cis-trans activity (PPIase) but they show dissimilarity in sequence and structure. Cyclophilins and FKBPs bind to the drug cyclosporin A (CsA) and FK506 (tacrolimus), and rapamycin respectively whereas parvulins bind to juglone (Grebe et al. 2000; Wang and Heitman 2005). Since the cells do not produce these drug naturally, therefore these drug treatment have clinical but no physiological relevance (Barik 2006). Based on the presence of types of domain, cyclophilins are classified as a member of the single PPIase group or protein groups with cyclophilin-like domain (CLD) and multidomain (MD) cyclophilin contain several functional domains such as RRM, Zinc Finger, internal repeats domain (RPT), tetratricopeptide repeat (TPR), WD40 and, coiled-coil domain (CCD) along with CLD (Stangeland et al. 2005).

CYPs are a prominent and abundant class of proteins involved in various fundamental biological process including signalling, protein folding, trafficking, transcription, RNA-binding, apoptosis, pathogen and plant stress responses (Allain et al. 1994; Anderson et al. 2002; Aumüller et al.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

2010; Baker et al. 1994; Brazin et al. 2002; Bukrinsky 2002; Dubourg et al. 2004; Jing et al. 2015; Kern et al. 1995; Klappa et al. 1995; Krzywicka et al. 2001; Li et al. 2007; Lin and Lechleiter 2002; Pogorelko et al. 2014; Schiene-Fischer and Yu 2001; Zander et al. 2003). Rice OsCYP2 participate in auxin signaling pathways by interacting with a zinc finger protein (OsZEP) that controls lateral root development (Cui et al. 2017). Soybean GmCYP1 interacts with GmMYB176, an isoflavonoid regulator which is stimulated by several abiotic stresses (Mainali et al. 2017). Two cyclophilin genes *ROC1-1* and *ROC1-2* of *Brassica rapa* related to light induction response (Yan et al. 2018). CYPs in *Arabidopsis* showed important roles in different cellular pathways; AtCYP59 coupled with RNA polymerase *II* control transcription and pre-mRNA processing, AtCYP20-3 is sensitive to oxidative stress; and AtCYP40 mutant reduces the number of juvinile leaves without making any change in inflorescence and flowering time etc. (Deng et al. 1998; GULLEROVA et al. 2006). Cyclophilins have been identified variable in numbers in different crops through comparative genomics approach. There are 29 members of cyclophilins in *Arabidopsis* (Li et al. 2007), 27 in rice (Li et al. 2007), 23 in *Chlamydomonas* (Li et al. 2007), 8 in *Saccharomyces cerevisiae* (Arévalo-Rodríguez and Heitman 2005), 16 in *Caenorhabditis elegans*, 11 in *Aspergillus nidulans* (Pemberton 2006) and 24 in humans (Galat 2003). In general, the plants contain higher number of cyclophilins compared to other eukaryotes, specially soybean crop is reported to bear the highest 62 members (Mainali et al. 2014).

Genome-wide studies of *cyclophilin* gene family in various species help understand the evolution and function of this gene family. However, comprehensive genome-wide characterization of this gene family has completed only in a few plant species including *Arabidopsis*, rice, soybean, cotton, *Brassica* etc. Tomato is an economically important vegetable and model fruit plant for genetic and molecular studies. Despite an earlier evidence of the existence of cyclophilins in the cDNA sequences, none of the later studies explored this family genes in tomato (Gasser et al. 1990). The present study was therefore conducted the genome-wide identification and characterization of *cyclophilin* gene family in tomato and compared their phylogenetic relatedness with other plant species. A detailed analysis on conserved domain architecture, genome structure, motif analysis, syntenic relationships, subcellular localization, comparative phylogenetic analysis and RNA-seq data analysis provide an in depth insight into the functional diversity of *cyclophilin* gene family in solanaceous crops.

Materials and Methods

Sequence analysis

Tomato *cyclophilin* genes and corresponding protein sequences were searched from Sol Genomics Network (SGN) (http://solgenomics.net/) database using “cyclophilin” as query. Preliminary 37 *CYP* genes were identified using BLASTP and TBLASTN algorithms techniques that matched with *cyclophilin* gene sequences of *Arabidopsis*. Finally, 35 genes were identified using HMM (Hidden Markov Model) profile of the PPIase domain (PF00160) from the Pfam website (http://pfam.xfam.org/) and was used as the query to identify all possible cyclophilin-like sequences with HMMER software (http://hmmer.org) considering the E-value <1E-5. The *cyclophilin* gene sequences of *Arabidopsis* and rice were obtained from TAIR (https://www.arabidopsis.org/), and TIGR-Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/), respectively. The maize *cyclophilin* gene sequences were obtained from maize genome database (http://www.maizesequence.org/index.html). Soybean *cyclophilin* gene sequences were obtained from Phytozome v11.0 (http://www.phytozome.net/) and *B. napus cyclophilin* gene sequences were from Hanhart *et al*. 2017. The presence of CLD and potential additional domains of all identified proteins were analyzed using NCBI domain search tool (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and hmmer algorithms (https://prosite.expasy.org/scanprosite/) tool. The protein length, molecular weight and isoelectric point were determined by the Expasy-Protopram (https://web.expasy.org/protparam/) web tool. Subcellular localization was predicted by CELLO v.2.5: sub cellular localization predictor (http://cello.life.nctu.edu.tw/). ORF finder (https://www.bioinformatics.org/sms2/orf_find.html) was employed to analyze open reading frames (ORFs). The *Arabidopsis thaliana* homologs for the individual *SlCYP*s were identified using SlCYPs amino acid sequences as queries for a BLAST search on UniProtKB (https://www.uniprot.org/). GSDS (http://gsds.cbi.pku.edu.cn/) and MEME (http://memesuite.org/) web tools were used to analyze the exon/intron structures and motif distribution of *cyclophilin* gene family respectively.

Sequence alignment and phylogenetic analysis

Protein sequences were aligned using Genedoc (http://www.nrbsc.org/gfx/genedoc/ebinet.htm) multiple sequence alignment tool. The phylogenetic tree was calculated by using the multiple alignment from ClustalOmega, and subsequently processed with MEGA 6.0 in the Neighbor-Joining (NJ) algorithm method (Tamura et al. 2013).

Chromosomal localization and gene duplication analysis

Start and end positions of *cyclophilin* genes were identified using the SGN database (https://solgenomics.net/) and their positions along the 12 chromosomes were mapped using the MapGene2Chromosome2 (http://mg2c.iask.in/). Any segmental or tandem duplication of tomato *cyclophillin* genes were analyzed using NCBI BLAST search tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi) based on the query coverage percentage and identity of each gene according to Kong et al. (2013) and Wang et al. (2010), respectively (Kong et al. 2013; Wang et al. 2010).

Synteny relationships of *cyclophilin* genes

Synteny analysis of *cyclophilin* genes between tomato, potato and *Arabidopsis* was performed using the MCScanX tool. TBtools software (Chen et al. 2018) was utilized to construct a syntenic analysis map. Circos (Krzywinski et al. 2009) was used to visualize the syntenic relationships between the genomes. The gene sequences, genome location, and chromosome length of each *cyclophilin* gene of *Arabidopsis* were downloaded from the TAIR database (https://www.arabidopsis.org/). Potato *cyclophilin* gene sequences, genome location and length of individual chromosome were downloaded from the phytozome database (https://phytozome.jgi.doe.gov/pz/portal.html).

Expression profiling using RNA-seq data

The expression of 35 *SlCYP* (species *S. pimpinellifolium)* genes was analysed based on RNA-seq data from 12 tissues available at Tomato Functional Genomics Database (http://ted.bti.cornell.edu/cgi-bin/TFGD/digital/expression.cgi) using Illumina RNA-seq RPKM values. The types of tissue include-developing organs of seeds (cotyledon, hypocotyl), vegetative organs (roots, young leaves, mature leaves, and vegetative meristem), reproductive organs (young floral bud, flowers at anthesis) and fruits (developing fruit at 10 and 20 days post anthesis and ripening fruit). Heatmaps of normalized expression values of *SlCYP* genes of tomato were generated using the online tool Heatmapper (http://www.heatmapper.ca/expression/).

Interaction network of *SlCYP* proteins

String web-based software (https://string-db.org/) was used to predict the protein interactions of tomato cyclophilin proteins and the *Arabidopsis* cyclophilin proteins. The tomato cyclophilin proteins corresponded with the *Arabidopsis* tomato cyclophilin are enlisted below the *Arabidopsis* cyclophilin proteins in figure. The online program ran with default parameters.

Results

Identification of cyclophilins in tomato genome

A total of 37 tomato cyclophilin sequences syntenic to *A. thaliana* CYPs were identified in tomato genome database by BLASTp search. Finally, 35 sequences were selected for subsequent analysis like domain search, homology analysis with *Arabidopsis* etc. The identified 35 tomato *cyclophilin* genes were named following Romano et al. (2004) and He et al. (2004) based on the relative molecular weight (He et al. 2004; Romano et al. 2004). Thus, the proteins having comparatively similar molecular weights were numbered chronologically. Exception was Solyc01g111170.2.1 that was previously named as SlCYP1 which is a potential mutant of *diageotropica* gene (*dgt-SlCyp1*) (Oh et al. 2006). But following the naming method of AtCYPs by (He et al. 2004; Romano et al. 2004) the sequence was named as SlCYP5-3, in the study (He et al. 2004; Romano et al. 2004). The length of SlCYP proteins ranged between 61 and 809 amino acids with molecular weight ranging between 7.169 and 91.028 kDa (Table 1). The isoelectric point ranged from 4.72 to 11.59 indicating that tomato cyclophilin proteins are both acidic and basic in nature. Out of the 35 SlCYPs, 16 proteins are predicted to be localized in cytoplasm and 10 in nucleus (Table 1). The remaining proteins are predicted to be localized in mitochondria, chloroplast and extracellular tissues. Unlike *Arabidopsis* and rice, the SlCYPs were localized neither in endoplasmic reticulum nor in plasma membrane (Table 1).

Cyclophilin-like domain (CLD)

CLDs are highly conserved among the members of Cyclophilin. Majority of the identified sequences have full length CLDs but some of the CYPs (e.g. SlCYP1, SlCYP2, and SlCYP3) contained partial CLDs, missing one or two essential residues or complete secondary structure and thus might be lacking PPIase activity (Fig. 1).

Figure 1 represents the conserved sequence of all identified tomato CLDs. CLD sequences of SlCYP aligned with the secondary structure of human cyclophilin A (hCYPA) as an external reference (Fig. 1). The hCYPA

Table 1 All the identified CYP family members in the tomato plant and their nomenclature, locus name, molecular weight, protein sequence length, chromosomal location, subcellular localization, theoretical isoelectric point, and predicted exons

Gene name	Locus name	ORF	Chromosome location	Exon	Protein					
					Length (aa)	MW (kDa)	pI	Domain information	CLD position	Subcellular localization
SlCYP1	>Solyc12g038110	186	SL4.0ch12:48323666..48323851	1	61	7.169	9.30	SD	2-61	Extracellular
SlCYP2	>Solyc12g089200	228	SL4.0ch12:63881747..63881974	1	75	8.149	5.39	SD	1-75	Nuclear
SlCYP3	>Solyc12g038030	282	SL4.0ch12:48064472..48064862	2	93	10.549	9.12	SD	2-52	Mitochondrial
SlCYP4	>Solyc11g006070	462	SL4.0ch11:904408..904869	1	153	16.489	6.06	SD	1-149	Cytoplasmic
SlCYP5-1	>Solyc08g006090	483	SL4.0ch08:848486..854332	4	160	17.530	7.01	SD	2-153	Cytoplasmic
SlCYP5-2	>Solyc10g054910	519	SL4.0ch10:55122558..55123076	1	172	17.873	8.59	SD	7-170	Cytoplasmic
SlCYP5-3	>Solyc01g111170	516	SL4.0ch01:89876816..89877826	1	171	17.910	8.83	SD	1-164	Cytoplasmic
SlCYP6-1	>Solyc09g010190	495	SL4.0ch09:3614647..3619878	6	164	18.175	8.58	SD	11-162	Cytoplasmic
SlCYP6-2	>Solyc12g038070	477	SL4.0ch12:48149484..48150037	2	158	18.403	6.72	SD	2-74	Nuclear
SlCYP6-3	>Solyc12g038150	498	SL4.0ch12:48601797..48602468	2	165	18.894	4.72	SD	1-97	Nuclear
SlCYP7	>Solyc01g096520	573	SL4.0ch01:79859565..79864679	7	190	20.510	8.42	SD	26-189	Cytoplasmic
SlCYP8-1	>Solyc06g076970	624	SL2.50ch06:47833488..47837300	7	207	22.249	9.19	SD	41-204	Cytoplasmic
SlCYP8-2	>Solyc12g038010	594	SL4.0ch12:48040757..48041494	3	197	22.340	5.17	SD	35-125	Cytoplasmic
SlCYP9-1	>Solyc06g051650	678	SL4.0ch06:32962272..32966959	8	225	24.469	8.91	SD	59-222	Cytoplasmic
SlCYP9-2	>Solyc12g038000	645	SL2.50ch12:49574555..49575486	4	214	24.485	5.15	MD	156-211	Nuclear
SlCYP9-3	>Solyc01g111360	687	SL4.0ch01:89991208..89995745	7	228	24.933	6.65	SD	49-213	Mitochondrial
SlCYP10	>Solyc01g010590	687	SL4.0ch01:5632492..5638269	8	228	25.763	8.68	SD	35-192	Cytoplasmic
SlCYP11-1	>Solyc10g083930	693	SL4.0ch10:62791267..62795709	7	230	26.055	9.30	SD	76-226	Mitochondrial
SlCYP11-2	>Solyc01g009990	747	SL4.0ch01:4610206..4614220	6	248	26.535	9.20	SD	84-244	Chloroplast
SlCYP11-3	>Solyc09g008410	711	SL4.0ch09:1895141..1903823	7	236	26.881	6.76	SD	82-232	Cytoplasmic
SlCYP12	>Solyc07g007110	894	SL4.0ch07:1829661..1833568	2	297	32.380	8.68	SD	89-252	Chloroplast
SlCYP13	>Solyc02g061800	882	SL4.0ch02:31311295..31312755	2	293	33.831	4.76	MD	2-165	Cytoplasmic
SlCYP14	>Solyc03g119860	954	SL4.0ch03:62855114..62856604	2	317	34.575	8.74	SD	97-288	Chloroplast
SlCYP15	>Solyc08g077790	1032	SL4.0ch08:59811302..59816547	5	343	37.275	5.15	MD	167-323	Extracellular
SlCYP16-1	>Solyc02g090480	1086	SL4.0ch02:50052680..50058430	8	362	40.293	5.66	MD	7-172	Cytoplasmic
SlCYP16-2	>Solyc01g108340	1089	SL4.0ch01:87996043..88000791	8	362	40.347	6.05	MD	7-172	Cytoplasmic
SlCYP17	>Solyc12g049430	1149	SL4.0ch12:60709814..60710962	1	382	43.879	5.59	MD	2-164	Nuclear
SlCYP18-1	>Solyc12g013580	1356	SL4.0ch12:4454376..4461071	12	451	49.078	5.95	SD	277-443	Chloroplast
SlCYP18-2	>Solyc02g086910	1356	SL4.0ch02:47512351..47516135	7	451	49.287	5.00	SD	151-308	Chloroplast
SlCYP19	>Solyc08g062700	1479	SL4.0ch08:49890458..49910307	10	492	54.985	8.40	SD	14-168	Nuclear
SlCYP20	>Solyc02g092380	1791	SL4.0ch02:51497813..51503293	11	596	65.980	7.29	MD	260-443	Cytoplasmic
SlCYP21	>Solyc07g066420	1764	SL4.0ch07:67698634..67708435	14	587	68.095	5.87	MD	2-161	Nuclear
SlCYP22	>Solyc11g067090	1869	SL4.0ch11:50864889..50872734	13	622	70.038	6.62	MD	468-619	Cytoplasmic
SlCYP23	>Solyc09g065720	1983	SL4.0ch09:60106675..60114638	13	660	73.006	10.69	SD	10-174	Nuclear
SlCYP24	>Solyc08g067090	2430	SL4.0ch08:54099820..54111045	13	809	91.028	11.59	SD	9-175	Nuclear

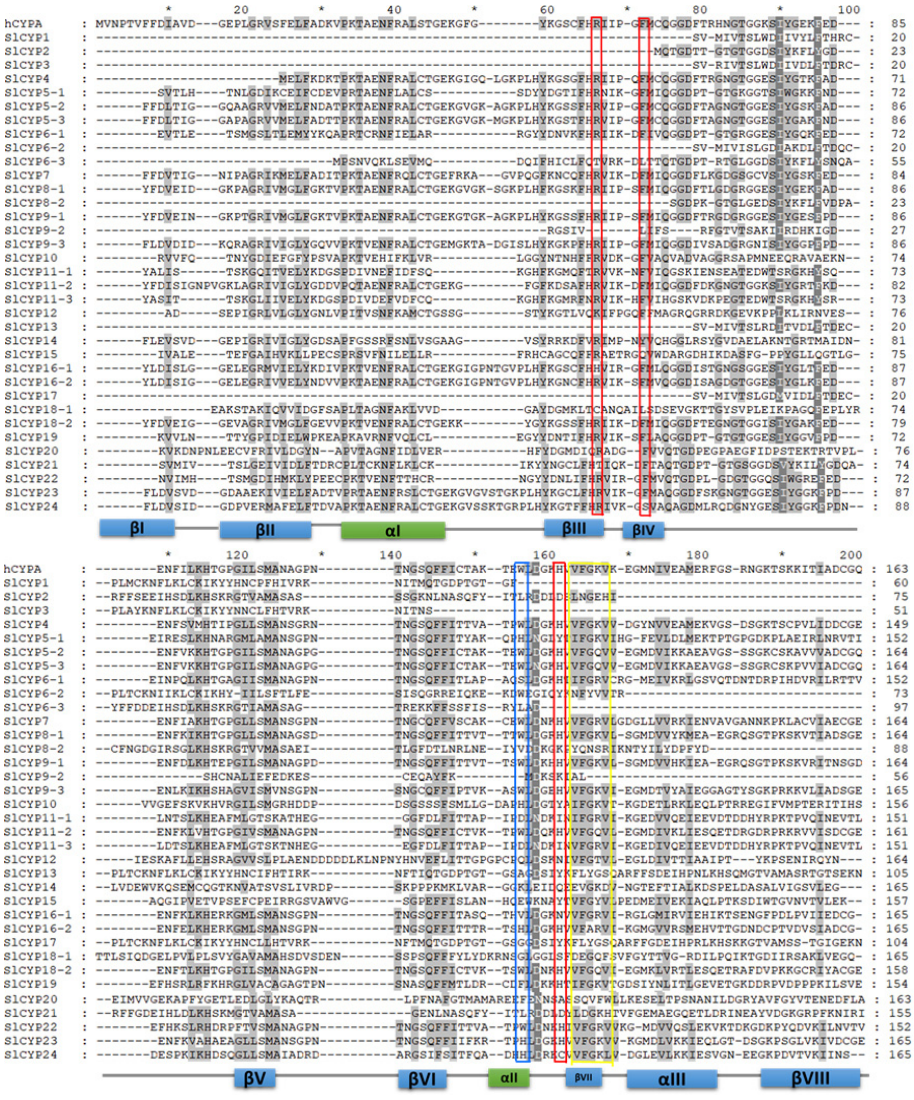

Fig. 1 Multiple alignment of the CLD sequences of the members of the CYP family in tomatoes with those of human CYPA (hCYPA) (whole protein sequences were cropped to the predicted CLD sequence). Conserved residues are highlighted; the yellow box indicates the amino acids of the VXGXV motif discovered in *Arabidopsis* CYPs; the red box indicates the amino acids important for PPIase activity; the blue box indicates the amino acids required for binding with CsA (Cyclosporin A) . The backgrounds indicate the percentage of amino acid similarity: black, 100%; dark gray, 80%; light gray, 60%

often referred to as the "archetypal" CYP consists of a β -barrel (eight antiparallel strands, β1 - β8) and two α -helices (on the top and bottom), respectively (Fig. 1).

Amino acid residues present in the structure are highly conserved and important for CYP function whereas gaps in the conserved regions denote insertions in individual residues of this family Exceptional insertion and deletion of amino acids were observed in plant CYPs in this study (Fig. 1, Supplementary Fig. 1).

Amino acid insertions from 8 to 11 molecules were observed in SlCYP8-1, SlCYP8-1, SlCYP9-1, SlCYP16-1, SlCYP16-2, SlCYP23, SlCYP24 in between α-helix-I and β-sheet-III (Fig. 1). Within the β-V and β-VI junction, SlCYP12, SlCYP14, SlCYP15, and SlCYP18-1 possessed 3-10 amino acids insertion. Romano et al. (2004) described the additional insertion of amino acids between α -helix-I and β-sheet-III, where 8 to 11 amino acids were inserted in several AtCYPs (Romano et al. 2004). Deletion of amino acids was also observed in some cases, for example- SlCYP12, SlCYP14, SlCYP15, and SlCYP24 showed this deletion (Fig. 1).

The amino acid residues that are essential for structure and function of CYP are highly conserved in hCYPA and other species. Among the conserved essential amino acids W121 (W158 in Fig. 1) was present in the 11 out of 35 SlCYP proteins. Further, R55 (R67 in Fig. 1), F60 (F72 in Fig. 1), H126 (H117 in Fig. 1) are the other highly conserved fundamental amino acids were present in 18,

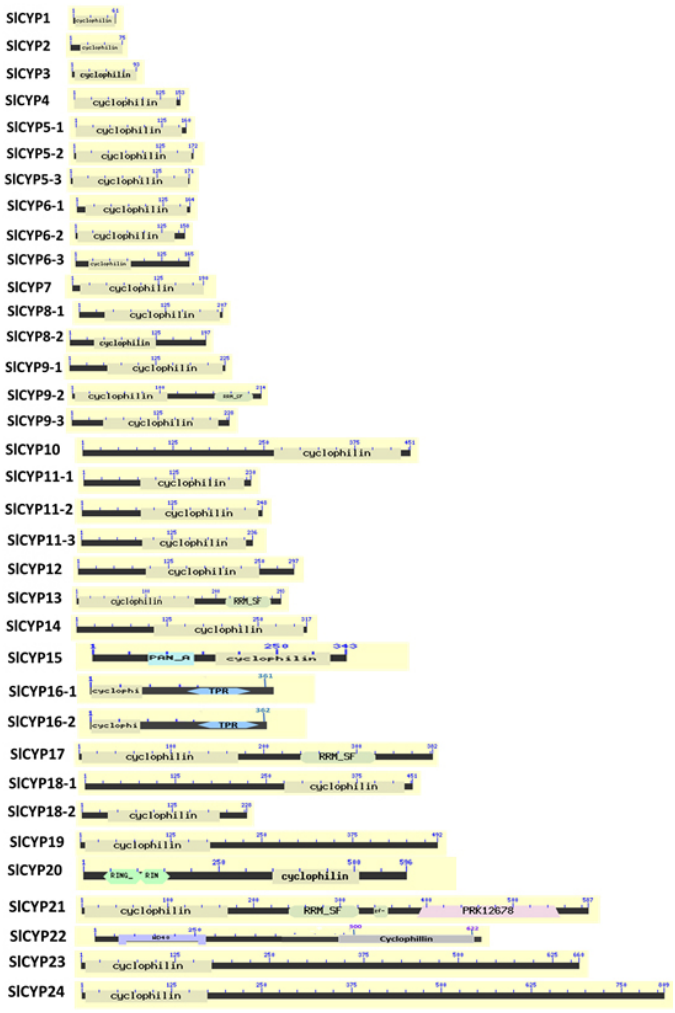

Fig. 2 Domain architecture of the tomato cyclophilin proteins. The beginning and ending amino acids number of each protein are shown at each end of the diagram. The cyclophilin domains are represented by grey boxes. The other functional domains, such as tetratricopeptide repeats (TPRs), RRM (RNA recognition motif)/ PAN_A, WD40 (tryptophan-aspartate repeat), PRK12668, and Ring_RIN, are indicated separately. The numbers above the boxes denote the positions of amino acids in the functional domains

22, 14 SLCYP proteins, respectively (Fig. 1). Another conserved motif VXGXV reported to be highly conserved in all AtCYP proteins. A total of 17 SlCYP proteins out of 35 contained this important VXGXV motif (Fig. 1).

An alignment of full-length protein sequences of SlCYPs with previously characterized CYPs from several different plants including human revealed considerable sequence identity among CYP proteins (Supplementary Fig. 1). The amino acid residues (arginine (R), phenylalanine (F) and histidine (H), which correspond to positions -62, -67 and -133 in AtCyp19-3 that critically essential for PPIase activity were present in most of the SlCYP proteins (Supplementary Fig. 1). The tryptophan (W121) is required for CsA binding activity was present in ten SlCYPs out of thirty five (Supplementary Fig. 1). Our observations are consistent with with previous reports in other species (Romano et al. 2004).

Homology and domain structure of cyclophilin

Most of the SlCYPs (24 out of 35) with a comparatively lower molecular weight encoded a protein with a single cyclophilin-like domain (SD) (Fig. 2; Table 1, Supplementary Table 1). The remaining 9 SlCYPs with comparatively high molecular weight possessed MD i.e., they contained CLD together with other functional domains, such as- TPR, RRM_SF, PAN_A, RING, RIN, WD40 and zinc finger domains (Fig. 2; Table 1, Supplementary Table 1). While there are mostly two SlCYPs homologs corresponding to each AtCYP with high sequence identity (80-86%), there are only a few exceptions with only one corresponding SlCYP homolog (Supplementary Table 2). As for example, SlCYP5-1 which is one of the *S. lycopersicum* homolog of AtCYP18-2 with 93.8% sequence identity (Supplementary Table 2).

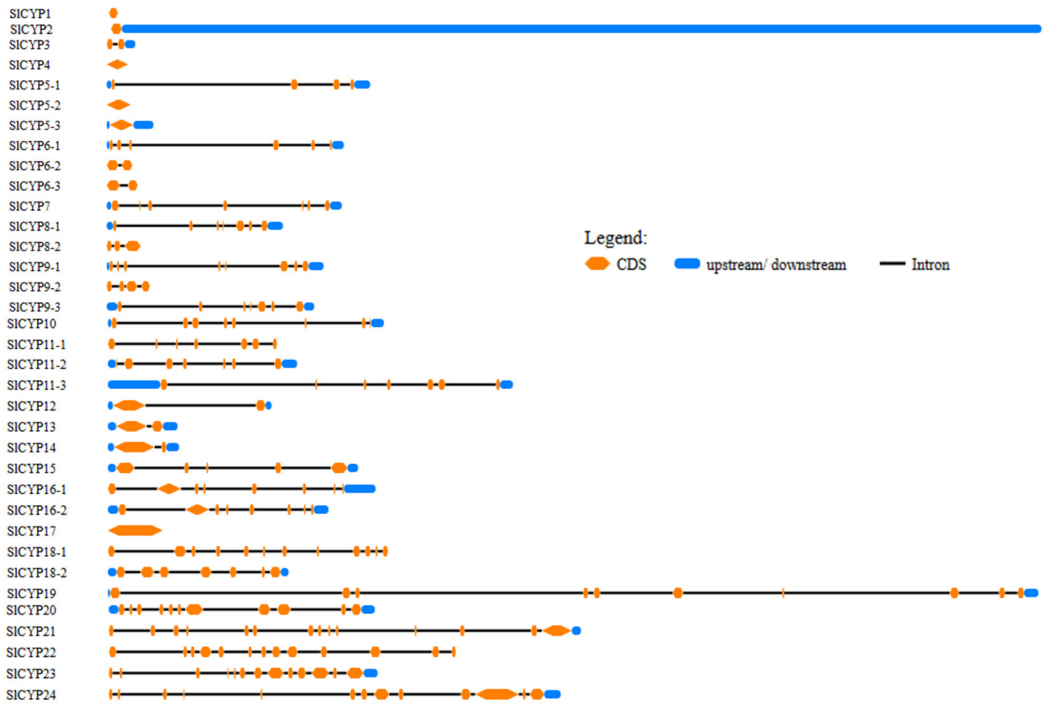

Fig. 3 Schematic representation of the exon–intron structure of *SlCYP*s. Exon–intron structures of *SlCYP*s were analyzed using GSDS 2.0 (http://gsds.cbi.pku.edu.cn/) database. The yellow boxes, green boxes, and lines indicate the exons, UTRs, and introns, respectively. The scale below can be used to estimate the lengths of the exons/introns

As shown in Figure 2, the MD CYPs, SlCYP16-1 and SlCYP16-2 are characterized by a TPR motif with CLD (Fig. 2). SlCYP16-1 and SlCYP16-2 being the homologs of AtCYP40 showed 68.3 and 76.4 identity, respectively (Supplementary Table 2). SlCYP22 contained a tryptophan-aspartic acid at the N-terminus and SlCYP22 showed 84.3% sequence identity to its *Arabidopsis* homolog AtCYP71 (Fig. 2; Supplementary Table 2). Four SlCYPs including SlCYP9-2, SlCYP13, SlCYP17 and SlCYP21 contain a RNA recognition motif (RRM). SlCYP21 and corresponding AtCYP homolog, AtCYP59 are the members containing an additional zinc finger motif (Fig. 2; Supplementary Table 2).

Gene structure and motif distribution of cyclophilin

Variable number of exon-intron distribution were observed among the members of *SlCYP*s. Six *SlCYP*s (*SlCYP1*, *SlCYP2*, *SlCYP4*, *SlCYP5-2*, *SlCYP5-3*, and *SlCYP17)* had no intron in their ORF (open reading frame) region. Variable number of introns were observed in the other 29 *SlCYP*s gene. Intron number varied from 1 to 13, where *SlCYP3, SlCYP6-2, SlCYP6-3, SlCYP12, SlCYP13, SlCYP14* contained single intron and *SlCYP21* have the largest numbers of introns (Fig. 3). Intron size varied considerably among the different *SlCYP* genes ranging from 32 bp (*SlCYP9-2*) to 4678 bp (*SlCYP19*). Similar exon-intron organization patterns of *SlCYP*s were observed in majority of cases in the phylogenetic tree such as *SlCYP11-2, SlCYP16-1, SlCYP16-2, SlCYP5-2, SlCYP17, SlCYP3,* and *SlCYP13* (Supplementary Fig. 1). The putative motif distribution of *SlCYP* was also analysed to detect potential conserved motifs of *SlCYP* (Supplementary Fig. 1). The closely related members in the phylogenetic tree generally had similar motif composition such as *SlCYP11-1, SlCYP11-3; SlCYP5-2, SlCYP5-3; SlCYP16-1, SlCYP16-2* (Fig. 4).

Phylogenetic relationships and subcellular localization

Phylogenetic analysis of tomato, *Arabidopsis* and rice cyclophilin family members exhibited several clusters of proteins with high sequence similarity (Fig. 4). Generally cyclophilin from *Arabidopsis* and tomato were found to be more closely related to each other as compared with rice. Several SlCYPs clustered together in pairs with high bootstrap value with AtCYPs. Often, they possessed the same supplementary domains, consistent subcellular localization, and the same AtCYP homolog (Table 1; Fig. 4; Supplementary Table 1).

However, the bootstrap value of the parent nodes as well as the nodes of the individual clades was found to

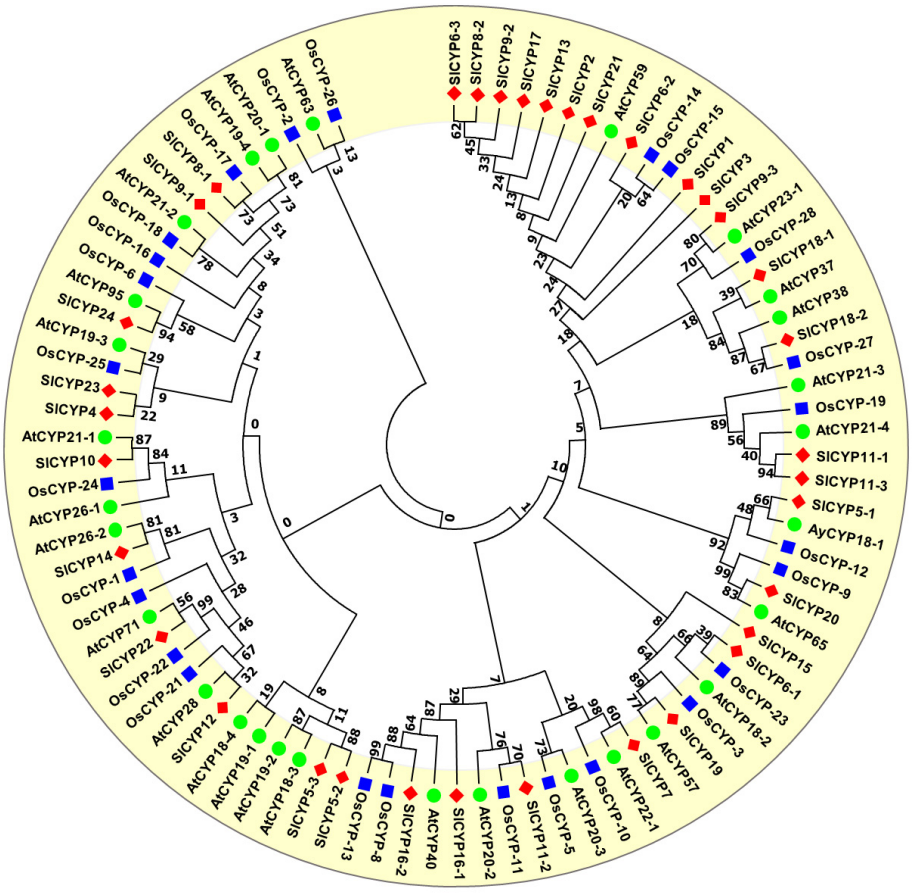

Fig. 4 Phylogenetic analysis of *cyclophilin* genes in tomato, rice, and *Arabidopsis* to analyze the relationship between the *cyclophilin* genes in the genomes from these three plants. The phylogenetic tree was constructed using the neighbor-joining method using 1000 bootstrap replicates

be very low for some members (Fig. 4). According to the phylogenetic analysis all proteins from tomato, *Arabidopsis* and rice were categorized into three clades A, B and C (Fig. 4). Clades A contained the highest number of SlCYP proteins while no SlCYP clustered into clade C (Fig. 4). Clade A included 24 and clade B contained 11 out of 35 SlCYPs (Fig. 4). The phylogenetic analysis also showed that some of *SlCYP* proteins were found to be more closely related to those of *AtCYP* and *OsCYP*, such as *SlCYP24* with *AtCYP94*, *SlCYP10* with *AtCYP21-1*, *SlCYP14* with *AtCYP26-2*, *SlCYP20* with *AtCYP65*, *SlCYP9-3* with *AtCYP23-1*, *SlCYP14* with *osCYP-1*, *SlCYP22* with *OsCYP-22*, *SlCYP16-2* with *OsCYP-8*, *SlCYP19* with *OsCYP-3* (Fig. 4).

As previously described for *Arabidopsis thaliana* homologs all of the SlCYPs except SlCYP1 and SlCYP15 are targeted to intercellular organelles. Majority (16 out of 35) of the SlCYPs predicted to be located in cytosol which is comparable to the subcellular distribution of AtCYPs where 14 of 29 are either have experimentally proved or predicted localization in the cytosol (Table 1). Furthermore, 9 SlCYPs are predicted to be located in nucleus, 5 in chloroplasts, and 2 in mitochondria (Table 1). None of the SlCYPs are predicted to be located in the endoplsmic reticulum or golgi or plasma membrane unlike *Arabidopsis* and rice CYPs (Trivedi et al. 2012).

Genomic distribution and gene duplication analysis

Genomic distribution of *SlCYPs* revealed that *SlCYPs* are distributed on 9 out of the 12 chromosomes of *Solanum lycopersicum* genome and the gene density per chromosome is unequal (Fig. 5). Chromosome 12 contain the highest number with 9 *SlCYP* genes while chromosome 03 has only one *SlCYP* gene (Fig. 5). The remaining chromosomes contained between 2 and 6 genes (Fig. 5). Most of the *SlCYP*s located towards the end of chromosomes but *SlCYP13* and *SlCYP17* were found closer to the centromere (Fig. 5). The evolutionary analysis among the tomato CYPs showed that most of the *SlCYP*s clustered together in pairs with high sequence homologies.

The duplication analysis of *SlCYP*s showed that two pairs of genes namely *SlCYP5-2* and *SlCYP5-3*, *SlCYP8-1*

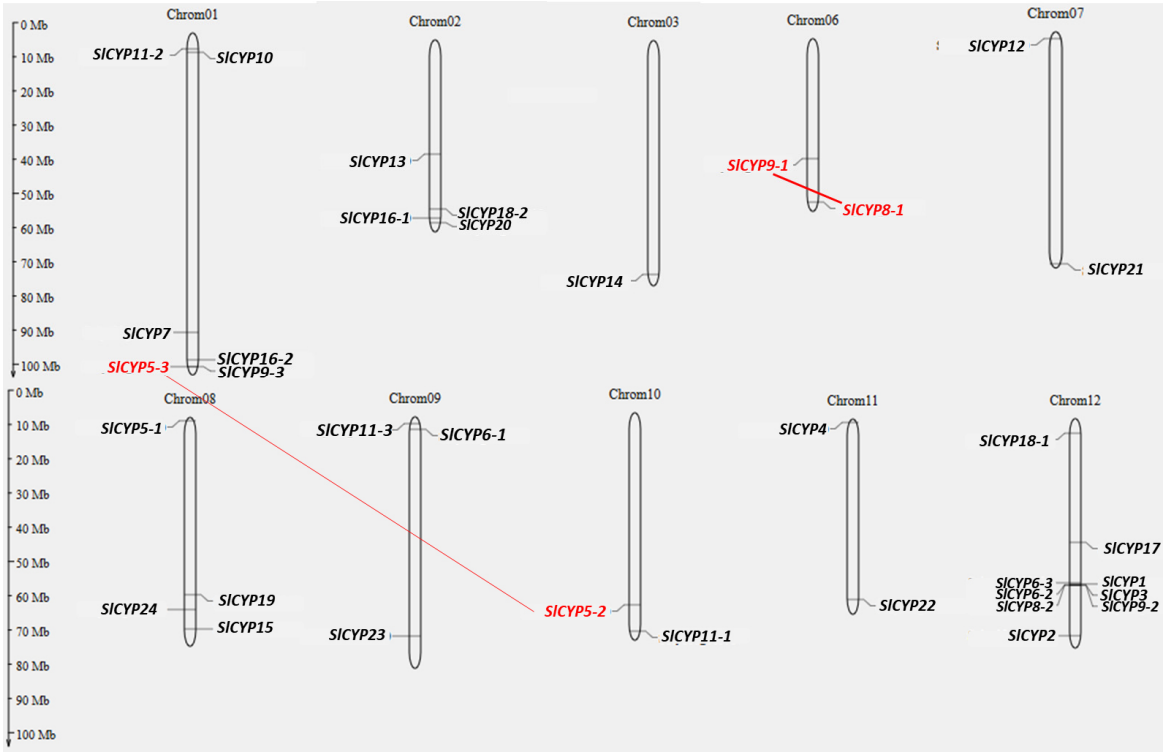

Fig. 5 The chromosomal locations of tomato *cyclophilin* genes are indicated based on the information provided by the tomato Genome Browser Solgenomics network (https://solgenomics.net/feature/17742813/details). The chromosomes are drawn to scale and the chromosome numbers are shown above each chromosome. The *SlCYP*s that are speculated to have undergone segmental duplication are indicated by the same color and are connected to each other by a line

and *SlCYP9-1* originated through segmental duplication and we did not find any evidence of tandem duplication event among the *SlCYP* genes (Fig. 5). These results illustrate that segmental and tandem duplications did not play important role in the evolution of *SlCYP*s rather these genes may arose through random insertion events. However, the *A. thaliana* homologs of *SlCYP5-2* and *SlCYP5-3, SlCYP8-1* and *SlCYP9-1* are *AtCYP18-4* and *AtCYP19-1* which have important role in various activity of plant growth and pathogen defense predicted the possible function of these protein in tomato.

Synteny relationships of *cyclophilin* genes

Microsynteny analysis of 95 *cyclophilin* genes from tomato, potato and *Arabidopsis* showed that there are many syntenic blocks between tomato, *Arabidopsis* and potato (Fig. 6). Among these blocks, 26 tomato *cyclophilin* genes showed pairwise synteny with genes of potato genome (red line), and 25 *cyclophilin* genes showed pairwise synteny with genes of *Arabidopsis* genome (blue line). Additionally, it has been identified that the 25 tomato *cyclophilin* genes have orthologous genes within *Arabidopsis* and potato genome simultaneously (Fig. 6).

Expression analysis of *SlCYP* genes

The expression data represents the RNA-seq reads from various tissues across several stages such as seed development (cotyledon, hypocotyl), vegetative (roots, young leaves, mature leaves, vegetative meristem), reproductive (young floral bud, anthesis flower) and fruit tissues 10 days post anthesis fruit1 (10 DPA1), 10 days post anthesis fruit2 (10 DPA2), 20 days post anthesis fruit (20 DPA), ripening fruit (33 DPA)}. The expression pattern indicates tissue-specific expression of most of the *SlCYP* genes (Fig. 7). Out of the 35 *SlCYP* genes the highest number of 18 genes were expressed in young leaves followed by 17 in vegetative meristem and 14 in young flower buds, respectively (Fig. 7). About 10 genes were expressed at the different stages of fruit development (Fig. 7).

Interaction network of SlCYP proteins

Protein interaction network of tomato cyclophilin proteins was constructed with the *Arabidopsis* cyclophilin proteins based on the homologous interaction relationship of *Arabidopsis* cyclophilin homologs (Fig. 8). The protein interaction analysis indicated that several SlCYPs (such as

Fig. 6 Synteny analysis of the *cyclophilin* genes from tomato, *Arabidopsis*, and potato. The chromosomes from different species are depicted as differently colored segments. The syntenic counterparts of the conserved *cyclophilin* genes between the genomes are interconnected by colored lines (tomato-*Arabidopsis*: blue line, tomato-potato: red line, and potato-*Arabidopsis*: yellow line)

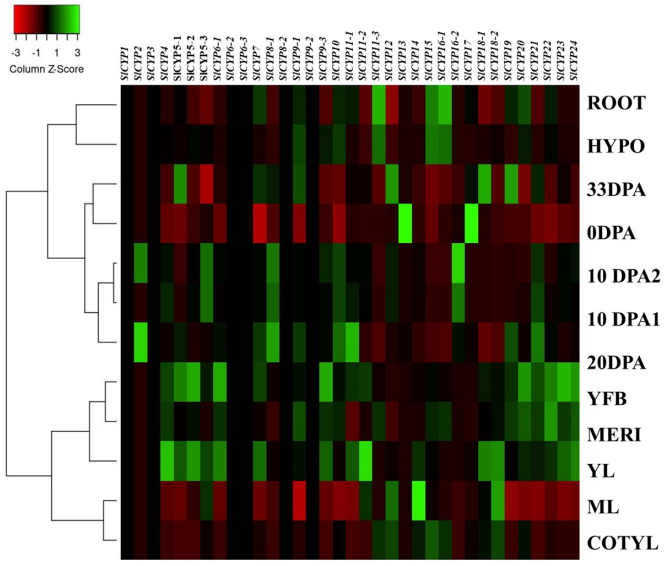

Fig. 7 Predicted expression analyses of tomato *CYP* genes. The RNA-seq data for different tissues and developmental stages of the tomato were obtained from the Tomato Functional Genomics Database (http://ted.bti.cornell.edu/cgi-bin/TFGD/digital/home.cgi) for heat map generation. The color scale indicates the expression values; the green color indicates low transcript abundance and the red color indicates high levels of transcript abundance. Cotyl: cotyledon, Hypo: hypocotyl, roots, YL: young leaves, ML: mature leaves, Meri: vegetative meristem, YFB: young floral bud, 0DPA: anthesis flower, 10 DPA1:10 days post anthesis fruit1, 10 DPA2:10 days post anthesis fruit2, 20 DPA2: 20 days post anthesis fruit, 33 DPA: ripening fruit. The heat map was generated via the online tool Heatmapper (http://www.heatmapper.ca/expression/) using the clustering method (Complete linkage) and distance measurement method (Spearman rank correlation)

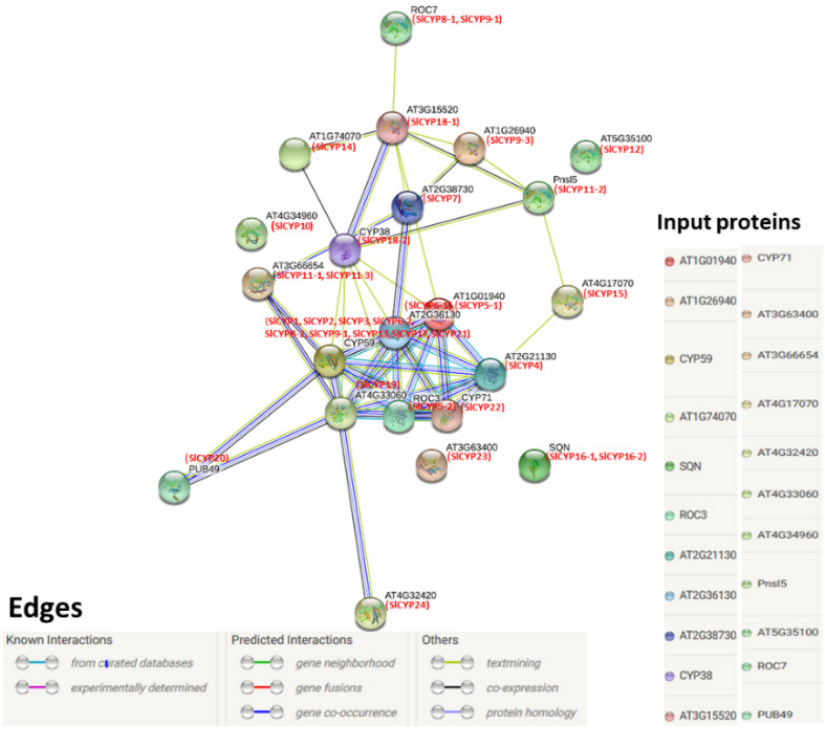

Fig. 8 Predicted protein–protein association networks of SlCYP proteins. The network nodes represent the proteins and the lines represent the protein–protein associations. The 3D structure of the proteins is shown inside the nodes and the colors of the line indicate different data sources. The light blue and purple lines indicate the known interactions from experimentally determined database; the green, red, and blue lines indicating the proteins are the predicted interactions; the yellow, black, and light sky-blue lines represent textmining, co-expression, and protein homology, respectively

SlCYP19/SlCYP4) were predicted to bear core nodes in the network, suggested their potential participation in diverse functions by interacting with other proteins. The result showed that AT1G26940 (SlCYP9-3) interact with AT3G15520 (SlCYP18-1), AT2G38730 (SlCYP7), Pnsl5 (SlCYP11-2) and predicted to be had a co-expression phenomenon. AT2G21130 (SlCYP4) was predicted to associate with CYP71 (SlCYP22) and AT4G17070 (SlCYP15) showing high protein homology (Fig. 8).

Discussion

Cyclophilins are ubiquitous proteins found in a wide range of organisms those contain the conserved PPIase (peptidyl prolyl cis-trans isomerase) domain (Gasser et al. 1990). In this study, all of the 35 identified tomato cyclophilin proteins have the conserved PPIase domain but considerable variation was observed in their molecular weight, isoelectric point as well as sequence at the N and the C-terminal regions (Wang and Heitman 2005). The amino acid residues that are essential for structure and function of CYP are highly conserved in hCYPA. It has been demonstrated previously that W121 (W158 in Fig. 1) of hCYP is essential for CsA (Cyclosporin A) binding but that is not essential for PPIase activity (Peptidyl-prolyl isomerases) (Liu et al. 1991; Zydowsky et al. 1992). Further, R55 (R67 in Fig. 1), F60 (F72 in Fig. 1), H126 (H117 in Fig. 1) are the highly conserved fundamental amino acids for the PPIase activity of hCYPA (Zydowsky et al. 1992).

Since some of the SlCYPs proteins did not have all of these three essential amino acids therefore their functional mechanism on PPIase activity is a subject of further investigation. The cyclophilin proteins exhibited a huge diversity at the N and the C-terminal region of protein sequences whereas the amino and carboxyl terminal ends were most divergent (Supplementary Fig. 1). Most conserved sequence was seen at the central position with highly conserved PPIase domain supporting an observation made previously (Wang and Heitman 2005).

Cyclophilins are grouped as both single-domain (SD) and multi-domain (MD) forms based on the presenence of domain within the protein (He et al. 2004). Other than the cyclophilin domain (SD), several tomato cyclophilin possessed additional domains (MD) those are involved in mRNA splicing, RNA stabilization and processing and cell morphogenesis. It has been demonstrated that AtCYP18-2 is involved in the regulation of pre-mRNA splicing and

because of the high homology its tomato homolog SlCYP6-1 might play a similar role in the nucleus of tomato cell (Romano et al. 2004). TPR motifs mediate protein-protein interactions and help in the assembly of multi-protein complexes (Blatch and Lässle 1999). AtCYP40 with its TPR motif forms a complex with small RNA duplex-bound AGO1 (ARGONAUTE family protein1) and HSP90 (heat shock protein 90) (Earley and Poethig 2011; Iki et al. 2012). The complex of AtCYP40 and HSP90-bound AGO1 plays a unique and important role in the assembly of plant RISC (RNA-induced silencing complex) where the activity of CYP40-HSP90 complex that facilitate RISC assembly is conserved between different species. Therefore, a similar function can be expected for the AtCYP40 homolog SlCYP16-1 and SlCYP16-2. WD40 usually developes a four stranded anti-parallel β -sheet and multiple copies of those sheets build a circular β-propeller structure promoting protein-protein interactions. SlCYP22 showed 84.3% sequence identity to its *Arabidopsis* homolog AtCYP71. AtCYP71 that bears a WD40 domain and locates in the nucleus, is involved in the regulation of gene expression and organogenesis. The WD40 domain enables AtCYP71 to interact with the chromatin histone H3 that affect the level of histone methylation (Li et al. 2007). Besides, AtCYP71 also interact with FASCIATA1 (a subunit of Chromatin Assembly Factor-1) and LHP1 (Like Heterochromatin Protein1), indicating the involvement of AtCYP71 in histone modification and chromatin assembly (Li and Luan 2011). Thus, a similar function for the highly identical CYP SlCYP22 can be assumed.

The group of proteins that encoded an RRM in addition to the CLD is called cyclophilin-RNA interacting protein (CRIP) (Krzywicka et al. 2001). SlCYP9-2, SlCYP13, SlCYP17 and SlCYP21 contain a RNA recognition motif (RRM) in addition to CLD. Furthermore, AtCYP59 in *A. thaliana,* homologs to those four SlCYPs also found in the genomes of Drosophila *melanogaster,* Paramecium *tetraurelia, Schizosaccharomyces pombe, Homo sapiens and Caenorhabditis elegans* (Krzywicka et al. 2001). AtCYP59 affects transcription by interacting with the CTD (C-terminal domain) of the largest subunit of RNA polymerase II and thus influences the phosphorylation of the CTD (Aumüller et al. 2010). Additionally, binding of specific RNAs prevent the PPIase activity of AtCyp59 which might modulate the activity of RNA polymerase II (Aumüller et al. 2010). Therefore, it is suggested that AtCYP59 might have an important function in pre-mRNA processing and transcription regulation. The similar function can be speculated for the four SlCYPs bearing a RRM, since similar domain structures exist in the corresponding *A. thaliana* homologs.

Intron is a major component of eukaryotic genomes and there is a correlation of intron size and genome size suggesting the possible evolution of some component of genome size within genes (McLysaght et al. 2000). It has also been found that intron size varies substantially between species, within species, among different genes even within a single gene which may reflect different functional properties they possess for the evolution of genomic and phenotypic traits (Zhang and Edwards 2012). Intron size varied considerably among the different *SlCYP* genes indicating their potential role in the evolution and function of tomato. Moreover, similar exon-intron organization patterns of *SlCYP*s were observed in majority of cases in the phylogenetic tree such as *SlCYP11-2, SlCYP16-1, SlCYP16-2, SlCYP5-2, SlCYP17, SlCYP3, SlCYP13* suggesting that there is a high conservation in the evolutionary process.

The putative motif distribution of *SlCYP* was also analysed to detect potential conserved motifs of *SlCYP* (Supplementary Fig. 1). The closely related members in the phylogenetic tree generally had similar motif composition such as *SlCYP11-1, SlCYP11-3; SlCYP5-2, SlCYP5-3; SlCYP16-1, SlCYP16-2* suggesting that there might have functional similarities among the *SlCYP*s (Fig. 4). However, different motif composition was also observed among the closely related members in phylogenetic tree indicating the structural divergence of *SlCYP* genes.

The phylogenetic analysis showed that the three clades included genes from both monocotyledons and dicotyledons indicated that the *CYP* is an ancient gene family originated before the divergence between the monocotyledon (rice) and dicotyledon (Arabidopsis and tomato) plants. Close clustering of tomato and *Arabidopsis* cyclophilins with high bootstrap value suggesting their close evolutionary relation. However, the bootstrap value of the parent nodes as well as the nodes of the individual clades was found to be very low for some members, possibly indicating a high sequence dissimilarity in the proteins. Since the plant kingdom arise from a common ancestor, the sequence dissimilarity among the tomato and rice CYPs may be due to differential selection process on the two plant species. However, the close clustering of some of the SlCYPs with those of Arabidopsis and rice homologs indicated that these SlCYP proteins may have the same or similar functions in tomato to those in Arabidopsis and/or rice.

Microsynteny analysis can be used to speculate the location of both orthologous genes and paralogous genes based on the whole-genome data of different species (Cao et al. 2016; Lin et al. 2014). These results speculated that

cyclophilin genes might have evolved from the common ancestor in various plant species. Genome comparison using orthologous genes from well-studied plant species may provide a valuable clue for the newly identified genes (Koonin 2005). *Arabidopsis thaliana* is widely used as a *model plant* and many *cyclophilin* genes have been functionally well characterized in *Arabidopsis*. The high number of orthologous gene pairs between tomato and *Arabidopsis* speculated that the genes may share a common ancestor and their functions could be conserved during evolution. However, additional research is needed to determine the specific function of each of the individual gene.

RNAseq expression analysis of the tomato *cyclophilin* gene families revealed distinct tissue-specific expression pattern where majority of the *SlCYP*s (51%) were expressed in the young leaves (Fig. 7). Furthermore, several chloroplastic protein SlCYP11-2, SlCYP14, SlCYP18-1, SlCYP18-2 were the most abundant CYP proteins in young leaf tissue suggesting their possible involvement in photosynthesis (Mainali et al. 2014). Expression of several *SlCYP* genes at the different stages of fruit suggesting their possible involvement in fruit development. Recently, emerging evidences have revealed the involvement of CYPs in organismal development and growth in plants (He et al. 2004; Romano et al. 2004; Vasudevan et al. 2015). Several studies have unveiled the putative roles of AtCYPs in a wide range of cellular processes including photosynthetic and hormone signaling pathways, transcriptional regulation, organogenesis, stress adaptation and defense responses (Vasudevan et al. 2015).

The role of cyclophilins in the regulation of different aspects of plant growth and development has been demonstrated by various recent studies. AtCYP19-1 (ROC3) was implicated in seed development, AtCYP71, resulted in compromised lateral organ formation and apical meristem activity, AtCYP40 was identified as a regulator of vegetative growth in *Arabidopsis*. (Grebe et al. 2000; Li et al. 2007; Stangeland et al. 2005). Several SlCYP proteins interact with AtCYP proteins with high sequence similarity and these interactions indicated that SlCYP proteins might be involved in plant growth and development.

Conclusions

The present study performed a genome-wide identification of CYPs in an important vegetable crop *S. lycopersicum*. Conserved CLD domain analysis, sequence alignment and sequence similarity searches with known *Arabidopsis* and rice CYPs identified a total of 35 CYP-coding genes, most of them are single-domain in nature. Subsequently, a comprehensive phylogeny, conserved motifs and localization analysis and expression analysis explained their function attributes. The results of the study provide a valuable resource for future investigation focusing the functional characterization of *SlCYP* genes in order to utilize them for crop improvement.

Acknowledgments

This work was carried out with the support of Sunchon National University Research Fund in 2021 (Grant number: 2021-0293) and "Cooperative Research Program for Agriculture Science & Technology Development (Project No. PJ01485801)" Rural Development Administration, Republic of Korea.

References

Allain F, Denys A, Spik G (1994) Characterization of surface binding sites for cyclophilin B on a human tumor T-cell line. Journal of Biological Chemistry 269(24):16537-16540

Anderson M, Fair K, Amero S, Nelson S, Harte PJ, Diaz MO (2002) A new family of cyclophilins with an RNA recognition motif that interact with members of the trx/MLL protein family in Drosophila and human cells. Development genes and evolution 212(3):107-113

Arévalo-Rodríguez M, Heitman J (2005) Cyclophilin A is localized to the nucleus and controls meiosis in Saccharomyces cerevisiae. Eukaryotic Cell 4(1):17-29

Aumüller T, Jahreis Gn, Fischer G, Schiene-Fischer C (2010) Role of prolyl cis/trans isomers in cyclophilin-assisted Pseudomonas syringae AvrRpt2 protease activation. Biochemistry 49(5): 1042-1052

Baker EK, Colley NJ, Zuker CS (1994) The cyclophilin homolog NinaA functions as a chaperone, forming a stable complex in vivo with its protein target rhodopsin. The EMBO journal 13(20):4886-4895

Barik S (2006) Immunophilins: for the love of proteins. Cellular and Molecular Life Sciences CMLS 63(24):2889-2900

Blatch GL, Lässle M (1999) The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. Bioessays 21(11):932-939

Brazin KN, Mallis RJ, Fulton DB, Andreotti AH (2002) Regulation of the tyrosine kinase Itk by the peptidyl-prolyl isomerase cyclophilin A. Proceedings of the National Academy of Sciences 99(4):1899-1904

Bukrinsky MI (2002) Cyclophilins: unexpected messengers in intercellular communications. Trends in immunology 23(7): 323-325

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). Frontiers in Plant Science 7:1750

Chen C, Chen H, He Y, Xia R (2018) TBtools, a toolkit for biologists integrating various biological data handling tools with a user-friendly interface. BioRxiv:289660

Cui P, Liu H, Ruan S, Ali B, Gill RA, Ma H, Zheng Z, Zhou W (2017) A zinc finger protein, interacted with cyclophilin, affects root development via IAA pathway in rice. Journal of integrative plant biology 59(7):496-505

Deng W, Chen L, Wood DW, Metcalfe T, Liang X, Gordon MP, Comai L, Nester EW (1998) Agrobacterium VirD2 protein interacts with plant host cyclophilins. Proceedings of the National Academy of Sciences 95(12):7040-7045

Dubourg B, Kamphausen T, Weiwad M, Jahreis G, Feunteun J, Fischer G, Modjtahedi N (2004) The human nuclear SRcyp is a cell cycle-regulated cyclophilin. Journal of Biological Chemistry 279(21):22322-22330

Earley KW, Poethig RS (2011) Binding of the cyclophilin 40 ortholog SQUINT to Hsp90 protein is required for SQUINT function in Arabidopsis. Journal of Biological Chemistry 286(44):38184-38189

Galat A (1999) Variations of sequences and amino acid compositions of proteins that sustain their biological functions: an analysis of the cyclophilin family of proteins. Archives of Biochemistry and Biophysics 371(2):149-162

Galat A (2003) Peptidylprolyl cis/trans isomerases (immunophilins): biological diversity-targets-functions. Current topics in medicinal chemistry 3(12):1315-1347

Gasser CS, Gunning DA, Budelier KA, Brown SM (1990) Structure and expression of cytosolic cyclophilin/peptidyl-prolyl cis-trans isomerase of higher plants and production of active tomato cyclophilin in Escherichia coli. Proceedings of the National Academy of Sciences 87(24):9519-9523

Grebe M, Gadea J, Steinmann T, Kientz M, Rahfeld J-U, Salchert K, Koncz C, Jürgensa G (2000) A conserved domain of the Arabidopsis GNOM protein mediates subunit interaction and cyclophilin 5 binding. The Plant Cell 12(3):343-356

GULLEROVA M, BARTA A, LORKOVIĆ ZJ (2006) AtCyp59 is a multidomain cyclophilin from Arabidopsis thaliana that interacts with SR proteins and the C-terminal domain of the RNA polymerase II. Rna 12(4):631-643

Handschumacher RE, Harding MW, Rice J, Drugge RJ, Speicher DW (1984) Cyclophilin: a specific cytosolic binding protein for cyclosporin A. Science 226(4674):544-547

He Z, Li L, Luan S (2004) Immunophilins and parvulins. Superfamily of peptidyl prolyl isomerases in Arabidopsis. Plant physiology 134(4):1248-1267

Iki T, Yoshikawa M, Meshi T, Ishikawa M (2012) Cyclophilin 40 facilitates HSP90-mediated RISC assembly in plants. The EMBO journal 31(2):267-278

Jing H, Yang X, Zhang J, Liu X, Zheng H, Dong G, Nian J, Feng J, Xia B, Qian Q (2015) Peptidyl-prolyl isomerization targets rice Aux/IAAs for proteasomal degradation during auxin signalling. Nature Communications 6(1):1-10

Kern G, Kern D, Schmid FX, Fischer G (1995) A kinetic analysis of the folding of human carbonic anhydrase II and its catalysis by cyclophilin. Journal of Biological Chemistry 270(2): 740-745

Klappa P, Freedman RB, Zimmermann R (1995) Protein disulphide isomerase and a lumenal cyclophilin-type peptidyl prolyl cis-trans isomerase are in transient contact with secretory proteins during late stages of translocation. European journal of biochemistry 232(3):755-764

Kong X, Lv W, Jiang S, Zhang D, Cai G, Pan J, Li D (2013) Genome-wide identification and expression analysis of calcium-dependent protein kinase in maize. BMC genomics 14(1):1-15

Koonin EV (2005) Orthologs, paralogs, and evolutionary genomics. Annu. Rev. Genet. 39:309-338

Krzywicka A, Beisson J, Keller AM, Cohen J, Jerka-Dziadosz M, Klotz C (2001) KIN241: a gene involved in cell morphogenesis in Paramecium tetraurelia reveals a novel protein family of cyclophilin-RNA interacting proteins (CRIPs) conserved from fission yeast to man. Molecular microbiology 42(1):257-267

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos: an information aesthetic for comparative genomics. Genome research 19(9): 1639-1645

Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic acids research 30(1):325-327

Li H, He Z, Lu G, Lee SC, Alonso J, Ecker JR, Luan S (2007) A WD40 domain cyclophilin interacts with histone H3 and functions in gene repression and organogenesis in Arabidopsis. The Plant Cell 19(8):2403-2416

Li H, Luan S (2011) The cyclophilin AtCYP71 interacts with CAF-1 and LHP1 and functions in multiple chromatin remodeling processes. Molecular plant 4(4):748-758

Lin D-T, Lechleiter JD (2002) Mitochondrial targeted cyclophilin D protects cells from cell death by peptidyl prolyl isomerization. Journal of Biological Chemistry 277(34): 31134-31141

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(7):e102825

Liu J, Chen CM, Walsh CT (1991) Human and Escherichia coli cyclophilins: sensitivity to inhibition by the immunosuppressant cyclosporin A correlates with a specific tryptophan residue. Biochemistry 30(9):2306-2310

Mainali HR, Chapman P, Dhaubhadel S (2014) Genome-wide analysis of Cyclophilin gene family in soybean (Glycine max). BMC plant biology 14(1):282

Mainali HR, Vadivel AKA, Li X, Gijzen M, Dhaubhadel S (2017) Soybean cyclophilin GmCYP1 interacts with an isoflavonoid

regulator GmMYB176. Scientific reports 7(1):1-12
McLysaght A, Enright AJ, Skrabanek L, Wolfe KH (2000) Estimation of synteny conservation and genome compaction between pufferfish (Fugu) and human. Yeast 17(1):22-36
Oh K, Ivanchenko MG, White T, Lomax TL (2006) The diageotropica gene of tomato encodes a cyclophilin: a novel player in auxin signaling. Planta 224(1):133-144
Pemberton TJ (2006) Identification and comparative analysis of sixteen fungal peptidyl-prolyl cis/trans isomerase repertoires. BMC genomics 7(1):244
Pogorelko GV, Mokryakova M, Fursova OV, Abdeeva I, Piruzian ES, Bruskin SA (2014) Characterization of three Arabidopsis thaliana immunophilin genes involved in the plant defense response against Pseudomonas syringae. Gene 538(1):12-22
Romano PG, Horton P, Gray JE (2004) The Arabidopsis cyclophilin gene family. Plant physiology 134(4):1268-1282
Schiene-Fischer C, Yu C (2001) Receptor accessory folding helper enzymes: the functional role of peptidyl prolyl cis/trans isomerases. FEBS letters 495(1-2):1-6
Stangeland B, Nestestog R, Grini PE, Skrbo N, Berg A, Salehian Z, Mandal A, Aalen RB (2005) Molecular analysis of Arabidopsis endosperm and embryo promoter trap lines: reporter-gene expression can result from T-DNA insertions in antisense orientation, in introns and in intergenic regions, in addition to sense insertion at the 5′ end of genes. Journal of experimental botany 56(419):2495-2505
Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Molecular biology and evolution 30(12):2725-2729
Trivedi DK, Yadav S, Vaid N, Tuteja N (2012) Genome wide analysis of Cyclophilin gene family from rice and Arabidopsis and its comparison with yeast. Plant signaling & behavior 7(12):1653-1666
Vasudevan D, Gopalan G, Kumar A, Garcia VJ, Luan S, Swaminathan K (2015) Plant immunophilins: a review of their structure-function relationship. Biochimica et Biophysica Acta (BBA)-General Subjects 1850(10):2145-2158
Wang L, Guo K, Li Y, Tu Y, Hu H, Wang B, Cui X, Peng L (2010) Expression profiling and integrative analysis of the CESA/CSL superfamily in rice. BMC plant biology 10(1):282
Wang P, Heitman J (2005) The cyclophilins. Genome biology 6(7):1-6
Yan H, Zhou B, He W, Nie Y, Li Y (2018) Expression characterisation of cyclophilin BrROC1 during light treatment and abiotic stresses response in Brassica rapa subsp. rapa 'Tsuda'. Functional Plant Biology 45(12):1223-1232
Zander K, Sherman MP, Tessmer U, Bruns K, Wray V, Prechtel AT, Schubert E, Henklein P, Luban J, Neidleman J (2003) Cyclophilin A interacts with HIV-1 Vpr and is required for its functional expression. Journal of Biological Chemistry 278(44):43202-43213
Zhang Q, Edwards SV (2012) The evolution of intron size in amniotes: a role for powered flight? Genome biology and evolution 4(10):1033-1043
Zydowsky LD, Etzkorn FA, Chang HY, Ferguson SB, Stolz LA, Ho SI, Walsh CT (1992) Active site mutants of human cyclophilin A separate peptidyl-prolyl isomerase activity from cyclosporin A binding and calcineurin inhibition. Protein Science 1(9):1092-1099