Mol Genet Genomics (2013) 288:111–129
DOI 10.1007/s00438-013-0733-0

ORIGINAL PAPER

# Genome-wide identification, phylogeny and expression analysis of *SUN*, *OFP* and *YABBY* gene family in tomato

**Zejun Huang · Jason Van Houten · Geoffrey Gonzalez · Han Xiao · Esther van der Knaap**

Received: 12 October 2012 / Accepted: 9 January 2013 / Published online: 31 January 2013

**Abstract** Members of the plant-specific gene families *IQD*/*SUN*, *OFP* and *YABBY* are thought to play important roles in plant growth and development. *YABBY* family members are involved in lateral organ polarity and growth; *OFP* members encode transcriptional repressors, whereas the role of *IQD*/*SUN* members is less clear. The tomato fruit shape genes *SUN*, *OVATE*, and *FASCIATED* belong to *IQD*/*SUN*, *OFP* and the *YABBY* gene family, respectively. A gene duplication resulting in high expression of *SUN* leads to elongated fruit, whereas a premature stop codon in *OVATE* and a large inversion within *FASCIATED* control fruit elongation and a flat fruit shape, respectively. In this study, we identified 34 *SlSUN*, 31 *SlOFP* and 9 *SlYABBY* genes in tomato and identified their position on 12 chromosomes. Genome mapping analysis showed that the *SlSUN*, *SlOFP*, and *SlYABBY* genes were enriched on the top and bottom segments of several chromosomes. In particular, on chromosome 10, a cluster of *SlOFPs* were found to originate from tandem duplication events. We also constructed three phylogenetic trees based on the protein sequences of the IQ67, OVATE and YABBY domains, respectively, from members of these families in *Arabidopsis* and tomato. The closest putative orthologs of the *Arabidopsis* and tomato genes were determined by the position on the phylogenetic tree and sequence similarity. Furthermore, expression analysis showed that some family members exhibited tissue-specific expression, whereas others were more ubiquitously expressed. Also, certain family members overlapped with known QTLs controlling fruit shape in *Solanaceous* plants. Combined, these results may help elucidate the roles of *SUN*, *OFP* and *YABBY* family members in plant growth and development.

**Keywords** Tomato · *SUN* · *OPF* · *YABBY* · Phylogenetic analysis · Expression analysis

## Abbreviations

| | |
|---|---|
| AtIQD | IQ67 domain protein in *Arabidopsis thaliana* |
| AtOFP | *Arabidopsis thaliana* OVATE family proteins |
| AtYABBY | YABBY protein in *Arabidopsis thaliana* |
| *CaOvate* | *OVATE*-like gene of *Capsicum annuum* |
| DBA | Days before anthesis |
| DPA | Days post anthesis |
| DUF623 | Domain-of-unknown-function 623, Pfam accession PF04844 |
| FAS | FASCIATED |
| IQD/SUN | IQ67 domain protein, SUN-like protein |
| ITAG | International Tomato Annotation Group |
| KaFTom | Kazusa Full-Length Tomato cDNA Database |
| Myr | Million years |
| NCBI | National Center for Biotechnology Information |
| QTL | Quantitative Trait Locus |
| RPKM | Reads per kilobase of exon model per million mapped reads |
| SGN | Sol Genomics Network |
| SlOFP | *Solanum lycopersicum* OVATE family proteins |

---

Communicated by C. Gebhardt.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00438-013-0733-0) contains supplementary material, which is available to authorized users.

Z. Huang · J. Van Houten · G. Gonzalez · H. Xiao · E. van der Knaap (✉)
Department of Horticulture and Crop Science,
The Ohio State University/OARDC, 217A Williams Hall,
1680 Madison Avenue, Wooster, OH 44691, USA
e-mail: vanderknaap.1@osu.edu

| | |
|---|---|
| SlSUN | *Solanum lycopersicum* SUN-like proteins |
| SlYABBY | *Solanum lycopersicum* YABBY proteins |
| TAIR | The Arabidopsis Information Resource |
| TALE | 3-Amino acid loop extension |
| WGS | Whole Genome Shotgun Sequencing |

## Introduction

Tomato (*Solanum lycopersicum*) is one of the most important vegetable crops in the world due to its great nutritive and commercial value. It is also a model organism for studying fleshy fruit development and ripening (Klee and Giovannoni 2011), compound leaf development, floral system and plant architecture (Kimura and Sinha 2008), as well as defense response against abiotic and biotic stresses (Kennedy 2003; Sun et al. 2011). Tomato belongs to the family Solanaceae, which includes vegetable crops such as pepper (*Capsicum annuum*), eggplant (*Solanum melongena*) and potato (*Solanum tuberosum*). The tomato genome is considered a reference for solanaceous species because it is one of the smallest diploid genomes within the family and, in particular, for species within the *Solanum* genus, it shows high conservation of gene order among each other (Tomato-Genome-Consortium 2012). Therefore, the study of tomato genes is important because the knowledge obtained may be easily applied to other Solanaceae species.

Gene families are groups of similar genes that arise from a common ancestor through duplication and divergence. Many genes belong to gene families. In *Arabidopsis thaliana*, 41 % of the predicted proteins belong to gene families containing at least five members (The-Arabidopsis-Genome-Initiative 2000). In rice (*Oryza sativa*), 77 % of the predicted genes are found to have at least one paralog (Goff et al. 2002). The *IQD/SUN*, *OFP* (*OVATE* family protein) and *YABBY* gene families are characterized by the IQ67, OVATE and YABBY domain, respectively (Golz and Hudson 1999; Bowman 2000; Abel et al. 2005; Hackbusch et al. 2005). In tomato, three quantitative trait loci (QTLs) controlling fruit shape have been cloned: *SUN*, *OVATE* and *FASCIATED* (*FAS*) belonging to the *IQD/SUN*, *OFP* and *YABBY* gene families, respectively (Liu et al. 2002; Hackbusch et al. 2005; Cong et al. 2008; Xiao et al. 2008).

The cloning of *SUN* revealed that the elongated fruit phenotype is caused by a 24.7-kb gene duplication that caused *SUN* to be controlled by the promoter of a defensin (*DEFL1*) gene leading to high expression in the fruit (Xiao et al. 2008; Jiang et al. 2009; Wu et al. 2011). Phenotypic analysis of *SUN* near isogenic lines shows that high *SUN* expression leads to fruit elongation by increased cell number in the longitudinal direction and reduced cell number in the transverse direction of the fruit. Overexpression of *SUN* results in slender cotyledons and leaflets as well as extremely elongated, seedless fruits (Wu et al. 2011). *SUN* encodes a protein containing the IQ67 domain (Abel et al. 2005). There are 33 and 29 genes encoding proteins with the IQ67 domain in *Arabidopsis* and rice, respectively (Abel et al. 2005). Over-expression of *AtIQD1* (At3g09710) leads to glucosinolate accumulation in *Arabidopsis* (Levy et al. 2005). It was recently found that AtIQD1 interacts with both kinesin light chain-related protein-1 (KLCR1) and also CaM/CMLs and recruits those proteins to the microtubules (Buerstenbinder et al. 2012). However, the function of other members of this family is unknown.

*OVATE* also controls tomato fruit elongation (Liu et al. 2002). A single mutation leading to a premature stop codon in the *OVATE* gene results in the transition of tomato fruit from round to pear-shaped. Over-expression of *OVATE* reduces the size of floral organs and leaflets; therefore, OVATE is considered to be a negative regulator of plant growth (Liu et al. 2002). *CaOvate*, an *OVATE*-like gene of *Capsicum annuum*, may play a similar role in fruit shape determination because it expresses higher in cv. "Mytilini Round" than cv. "Piperaki Long". Down-regulation of *CaOvate* through virus-induced gene silencing in cv. "Mytilini Round" changes its fruit to a more oblong shape (Tsaballa et al. 2011). *OVATE* encodes a protein with a 60–70 amino acid C-terminal domain termed the OVATE domain (Liu et al. 2002; Wang et al. 2007). In *Arabidopsis*, 18 genes encode OVATE domain-containing proteins, and are named *Arabidopsis thaliana* OVATE family proteins (AtOFPs) (Hackbusch et al. 2005; Wang et al. 2007). Most AtOFPs appear to function as transcriptional repressors in the transient *Arabidopsis* protoplast expression system (Wang et al. 2011). In a yeast two-hybrid screen, nine AtOFPs are found to interact with three-amino acid loop extension (TALE) homeodomain proteins (Hackbusch et al. 2005). AtOFP1 and AtOFP5 control the subcellular localization of one of the TALE homeodomain proteins, BLH1. When coexpressed with AtOFP1 and AtOFP5 in *Nicotiana benthamiana* leaves, BLH1 is relocated from the nucleus to the cytoplasmic space (Hackbusch et al. 2005). These results imply that the effect on growth is controlled by interactions of OFP with TALE homeodomain transcription factors and also by direct transcriptional repression of target genes. One such target gene is *AtGA20ox1* (gibberellin 20-oxidase1, a gibberellin biosynthetic gene) whose expression is reduced by *AtOFP1* overexpression. The reduced length of above ground organs is partially restored by application of gibberellin (Hackbusch et al. 2005; Wang et al. 2007). Besides interaction with TALE homeodomain proteins, AtOFP1 also interacts with AtKu, which is involved in DNA double-strand break repair

(Wang et al. 2010). In another study, AtOFP5 acts as a negative regulator of BLH1-KNAT3 activity during early embryo sac development (Pagnussat et al. 2007) and AtOFP4 plays a role in secondary cell wall formation through its interaction with KNAT7 (Li et al. 2011a). Contrary to tomato *OVATE*, the analysis of loss-of-function alleles of *OFPs* in *Arabidopsis* suggests that these genes have redundant functions because single knock out mutants of *AtOFP1*, *AtOFP4*, *AtOFP8*, *AtOFP10*, *AtOFP15* and *AtOFP16* do not show morphological defects (Wang et al. 2011). In all, the OFP proteins might regulate plant growth and development by affecting transcriptional regulation of target genes either directly or indirectly.

In contrast to *SUN* and *OVATE*, which control elongated fruit shape, a mutation in *FAS* results in a flat tomato due to an increase in locule number (Lippman and Tanksley 2001; Barrero et al. 2006). The mutation is the result of an inversion that knocks out the likely ortholog of *Arabidopsis YABBY2*, and this mutation is found in several tomato accessions with a high locule number and flat fruit shape (Cong et al. 2008; Huang and van der Knaap 2011; Rodriguez et al. 2011). YABBY proteins have conserved roles in specifying abaxial cell fate in lateral organs such as leaves, floral organs and ovules, and establishing the proper boundaries in meristems (Golz and Hudson 1999; Bowman 2000). *Arabidopsis* has six *YABBY* gene family members (Golz and Hudson 1999; Bowman 2000). Four of them, *FILAMENTOUS FLOWER* (*FIL*, also called *YAB1*), *YABBY2* (*YAB2*), *YABBY3* (*YAB3*) and *YABBY5* (*YAB5*), have overlapping functions in *Arabidopsis* leaf development based on the phenotype of their loss-of-function mutants (Stahle et al. 2009; Sarojam et al. 2010). The other two *Arabidopsis YABBY* genes, *CRC* and *INO*, are only expressed in floral organs (Bowman and Smyth 1999; Villanueva et al. 1999; Schmid et al. 2005). *CRC* is required for nectary specification and carpel polarity (Alvarez and Smyth 1999; Bowman and Smyth 1999), and *INO* is essential for development of the outer integument (Villanueva et al. 1999). A deletion mutant of the *INO* ortholog in sugar apple (*Annona squamosa*) was found in a spontaneous seedless mutant (*Thai seedless*; *Ts*) (Lora et al. 2011). There are eight YABBY genes in rice (Toriba et al. 2007). *DROOPING LEAF* has diverse roles in rice leaf development and homeotic transformations of floral organs (Yamaguchi et al. 2004; Ohmori et al. 2011; Li et al. 2011b). *TONGARI-BOUSHI1* (*OsYABBY5*) is reported to control lateral organ development and regulation of meristem organization in the rice spikelet (Tanaka et al. 2012). Moreover, sorghum has three different mutations in the *YABBY* gene *Shattering1* (*Sh1*), which result in the loss of seed shattering in domesticated sorghum (Lin et al. 2012).

Taken together, members of *IQD/SUN*, *OFP* and *YABBY* gene families play important roles in plant growth and development and may also underlie additional fruit shape genes in tomato and other Solanaceae plants. However, except for *SUN*, *OVATE* and *FAS*, virtually no information is available about the members of these three gene families in tomato. In this study, we identified 34 *Solanum lycopersicum SUN* (*SlSUN*) genes, 31 *Solanum lycopersicum OVATE family protein* (*SlOFP*) genes and 9 *Solanum lycopersicum YABBY* (*SlYABBY*) genes, and determined their closest orthologs in *Arabidopsis* based on phylogenetic relationships. We also investigated their expression pattern in 11 different tissues from tomato's closest wild relative, *Solanum pimpinellifolium*, from which it is thought to be domesticated (Peralta et al. 2008; Tomato-Genome-Consortium 2012). Our results may provide important clues for understanding the roles of the *SlSUN*, *SlOFP* and *SlYABBY* genes in tomato growth and development, and this information could be extended to other plants.

## Materials and methods

### Plant material and tissue collection for expression analysis

Seeds of *S. pimpinellifolium* accession LA1589 were obtained from the C.M. Rick Tomato Genetics Resource Center, Davis, California, USA. Plants were grown under standard conditions with supplemental lighting in the greenhouse in Wooster, OH, USA. Over the span of a month, seven different tissue types from 17 separate LA1589 tomato plants were collected in a greenhouse between 9:00 a.m. and 10:00 a.m. and were pooled for each tissue type. The collected tissues were immediately frozen in liquid nitrogen. The tissues collected were newly developed leaves around 5 mm long, mature green leaflets, flower buds younger than or equal to 10 days before anthesis (DBA), flowers at anthesis, 10 days post anthesis (DPA) fruit, 20 DPA fruit and 33 DPA ripening fruit. The following tissues were collected from seeds that germinated and grew for 7 days in a petridish under growing lights: whole root, hypocotyl from below the cotyledons to above the root zone, cotyledons, and vegetative meristems (including leaf primordia).

### Identification of *SUN*, *OFP* and *YABBY* genes in tomato

The IQ67 domain (Abel et al. 2005) of SUN was used to identify the members of this family in tomato; the OVATE domain (Liu et al. 2002; Hackbusch et al. 2005; Wang et al. 2007), also known as DUF623 domain (Domain-of-Unknown-Function 623, Pfam accession PF04844), was used to identify *OFP* genes; the YABBY domain (Pfam accession PF04690) of FAS was used to identify *YABBY* genes (Cong et al. 2008; Punta et al. 2012). With these domains as initial queries, systematic BLAST searches were performed on all sequences in the International Tomato Annotation Group (ITAG) Release 2.3 predicted proteins (2.40) (BLASTP, *E* value $\leq$1e−5), and tomato WGS chromosomes (2.40) (TBLASTN, OVATE domain and YABBY domain *E* value $\leq$1e−5; IQ67 domain *E* value $\leq$100) (SGN http://solgenomics.net). We identified nine genes that were not in database ITAG Release 2.3 but appear to have protein coding potential based on annotation by FGENESH (http://linux1.softberry.com/berry.phtml). Initial evidence of transcription of all genes was based on the identification in the Lycopersicon Combined (Tomato) Unigenes, and the *Solanum peruvianum* de novo transcriptome available at SOL Genomics Network (SGN, http://solgenomics.net), and full-length cDNA sequences in the KaFTom database (http://www.pgb.kazusa.or.jp/kaftom/). Further evidence of transcription, including that was not annotated in the latest release of the tomato genome, was based on expression analysis shown in this research. Only genes with at least one average RPKM value from all 11 tissues $\geq$2 in this study were considered to be expressed. The chromosomal location of *SUN*, *OFP* and *YABBY* genes was initially based on both their genetic map position using segregating populations (van der Knaap and Tanksley 2001) as well as their position on the tomato WGS Chromosomes (SL2.40) (SGN http://solgenomics.net). The sequences of AtIQD, AtOFP and AtYABBY proteins were downloaded from the *Arabidopsis thaliana* TAIR10 Protein database (ftp://ftp.arabidopsis.org/home/tair/Proteins/TAIR10_protein_lists/TAIR10_pep_20101214). Moss (*Physcomitrella patens*) IQD/SUN and OFP sequences, and grape (*Vitis vinifera*), poplar (*Populus trichocarpa*) YABBY sequences were downloaded from Phytozome v9.0 (http://www.phytozome.net/). The cucumber (*Cucumis sativus*) YABBY sequences were downloaded from Cucumber Genome DataBase (http://cucumber.genomics.org.cn/page/cucumber/index.jsp). The potato (*S. tuberosum*) YABBY sequences were downloaded from Solanaceae Genomics Resource (http://solanaceae.plantbiology.msu.edu/pgsc_download.shtml). The sitka spruce tree (*Picea sitchensis*) YABBY sequences were downloaded from Genbank (http://www.ncbi.nlm.nih.gov/genbank/).

### Multiple alignment and phylogenetic analysis

The IQ67 domain was defined as described (Abel et al. 2005). The OVATE and YABBY domains were defined using the Pfam program (http://pfam.sanger.ac.uk/). Multiple alignments of the three conserved domain sequences were performed by ClustalX 2.1 (Larkin et al. 2007) with default setting. The alignment results were exported to MEGA 5.0 (Tamura et al. 2011). Unrooted phylogenetic trees were constructed with neighbor-joining (NJ) method, JTT model and 1,000 replicates. The identification of paralogous and orthologous relationships was based on their phylogenies, sequence similarity and all-against-all bidirectional best hits using SSEARCH (Smith and Waterman 1981; Pearson 1991).

### RNA library construction

Total RNA was extracted with Trizol (Invitrogen Inc. USA) as described by the manufacturer or using a hot borate method (only for fruit at 20 DPA or 33 DPA) (Pang et al. 2011). RNA quantity and quality were assessed using a Qubit 2.0 fluorometer RNA Assay Kit (Invitrogen Inc. USA) and an Agilent 2100 Bioanalyzer RNA 6000 Nano kit (Agilent, USA). Strand-specific RNA-seq libraries of approximately 250 bp fragments were prepared using 10 μg total RNA (Zhong et al. 2011). Libraries were barcoded and pooled to represent six libraries from different tissues per lane on the flowcell. Sequences of 51 bp were generated on an Illumina HiSeq2000 at the Genomics Resources Core Facility at Weill Cornell Medical College (New York, NY, USA).

### Alignment and analysis of illumina reads

After illumina reads were quality checked, demultiplexed and trimmed, they were clustered per library. The reads were aligned to ribosomal RNA sequences using Bowtie (Langmead et al. 2009) allowing for two mismatches to identify rRNA contamination. The ribosomal filtered reads were then aligned with TopHat (Trapnell et al. 2009) against the *S. lycopersicum* genome allowing for maximum intron lengths of 5,000 bp, segment lengths of 22 bp and 1 mismatch per segment. All other parameters were set to default. Reads that mapped up to 20 genes were counted as 1 for each match. Aligned sequences were then separated into sense and antisense, and the count of aligned reads for each tomato gene model and from each sample was derived using an in-house perl script. This script also counted reads that partially mapped to the UTRs. Reads per kilobase of exon model per million mapped reads (RPKM) were calculated using an in-house script based on both the ITAG 2.3 exon lengths and also the total number of reads that mapped to the tomato genome. For the expression analysis of selected genes in different tissues, the average RPKM values for each tissue type was shown. All raw reads were deposited in the NCBI sequence read archive with accession number SRA061767. The average RPKM values per sample for all genes can be found at http://ted.bti.cornell.edu/cgi-bin/TFGD/digital/home.cgi.

## Results

### The *SUN* genes in tomato

#### *Identification of SUN genes in tomato*

Twenty-nine genes encoding the entire IQ67 domain were identified in the ITAG database version 2.3. Four additional genes that potentially encoded other members of the *SUN* family were found in tomato WGS Chromosomes (SL2.40) (SGN http://solgenomics.net) and evaluated using FGENESH program (http://linux1.softberry.com/berry.phtml). Three of them consisted of a different predicted CDS of Solyc01g009340 (*SlSUN2*), Solyc01g097490 (*SlSUN4*) and SL1.00sc00090_96 (*SlSUN6*) (Table 1; Fig. 1; Online source 1). The *SUN* gene on chromosome 7, which controls elongated fruit shape, was identified as *SUN*. The original copy of *SUN* on chromosome 10 (Xiao et al. 2008) was referred to as *SlSUN1*. The other members were named *SlSUN2*–*SlSUN33* according to their position from the top to the bottom on chromosomes 1–12. Twenty-five *SlSUN* genes were supported by unigenes or full-length cDNA sequences, and 28 *SlSUN* genes demonstrated expression in this study (Table 1). Evidence for the expression of the five remaining *SUN*-like genes either was not found or was below the cut-off in the RNA-seq dataset developed for this study.

All *SlSUN*-like genes had multiple introns including one that disrupted the IQ67 domain between codons 16 and 17 (Table 1). This has also been noted for most *Arabidopsis IQD* genes (Abel et al. 2005). *SlSUN6* was the smallest member of this family. It had two exons and was predicted to encode a 128 amino acid protein (Table 1). Whereas *SUN* is located on chromosome 7 (Xiao et al. 2008), none of the other 33 *SUN* family members were located on this chromosome. *SlSUN19* (Solyc08g007920.1.1) and *SlSUN20* (Solyc08g007930.1.1) were close to each other, within a segment of 15 kb on chromosome 8 (Table 1; Fig. 1).

#### *Phylogenetic analysis of Arabidopsis IQD genes and tomato SUN genes*

To uncover the phylogenetic relationships between *Arabidopsis IQD* and tomato *SUN* genes, we constructed a dendrogram based on their IQ67 domain sequences (Fig. 2; Online source 2). The phylogenetic trees illustrate that the *AtIQD* and *SlSUN* genes could be divided into ten subgroups (Fig. 2). The detailed information of closest ortholog pairs between AtIQDs and SlSUNs was listed in Online source 2. *SUN* and *SlSUN1* were paralogs of *SlSUN12*, and their ortholog was likely represented by *AtIQD12* (Online source 2) as reported previously (Xiao et al. 2008). Several AtIQD and SlSUN proteins showed a one-to-one orthologous relationship, such as SlSUN6 and AtIQD20, SlSUN14 and AtIQD32, SlSUN22 and AtIQD6, and SlSUN31 and AtIQD5, which implied there was a common ancestor for these pairs, respectively (Fig. 2; Online source 2).

#### *The expression pattern of SlSUN genes in wild tomato*

To gain insights into the role of the *SlSUN* genes in tomato growth and development, we analyzed their expression patterns in both different tissues and also developmental stages using an RNA-seq approach. Twenty-eight *SlSUN* genes were expressed in this study. The average of the highest RPKM values in the 11 tissues of the 28 *SlSUN* genes is 135.18, and *SlSUN29* demonstrates the highest gene expression of this family with an RPKM of 836.66 in one of the 11 tissues (Table 1; Online source 3). *SlSUN1* was expressed slightly higher in the hypocotyl, flower at anthesis and fruit at 10 and 20 DPA (Fig. 3a; Online source 3). Some *SlSUN* genes were specifically expressed in certain tissues. For example, *SlSUN2* was specifically expressed in the vegetative meristem, young leaf and young flower bud; *SlSUN5*, *SlSUN21* and *SlSUN27* were specifically expressed in the root; *SlSUN11* and *SlSUN22* were specifically expressed in the young leaf and young flower bud; *SlSUN12* and *SlSUN26* were specifically expressed in the hypocotyl; *SlSUN24* was specifically expressed in the vegetative meristem and young flower bud; *SlSUN28* was specifically expressed in ripening fruit (33 DPA fruit); *SlSUN33* was specifically expressed in fruit at 20 DPA (Fig. 3; Online source 3).

### The *OFP* genes in tomato

#### *Identification of OFP genes in tomato*

Twenty-five putative *SlOFP* genes encoding the OVATE domain were found in the ITAG database version 2.3 (Table 2; Fig. 1; Online source 4). Six putative additional genes that were predicted to encode the OVATE domain were found in tomato WGS Chromosomes (SL2.40) (SGN http://solgenomics.net) using FGENESH program (http://linux1.softberry.com/berry.phtml). Two of them were found in the previous genome annotation, ITAG version 1.0: SL1.00sc02618_4 (*SlOFP4*) and SL1.00sc03540_201 (*SlOFP31*) (Table 2; Fig. 1; Online source 4). The gene locus Solyc09g065350 (*SlOFP18*) in the reference genome of cultivar Heinz1706 had a one-nucleotide deletion causing a nonsense mutation and the loss of the OVATE domain-coding region. The allele in *S. pimpinellifolium*, LA1589 and *S. peruvianum* had longer CDS (coding sequence) encoding the OVATE domain (Table 2; Online source 4). In this study, the tomato *OVATE* gene was referred to as *SlOFP1* and the other genes were named

**Table 1** *SUN* gene family in tomato

| Gene name | Gene locus[a] | Position[b] | CDS (bp) | Intron[c] | Protein (aa) | Unigene | cDNA | RNAseq (RPKM)[d] |
|---|---|---|---|---|---|---|---|---|
| *SUN* | | SL2.40ch07:2395262 | 1,266 | 4 | 421 | SGN-U569959 | EU491503 | |
| *SlSUN1* | Solyc10g079240.1.1 | SL2.40ch10:60140568…60142797 (+) | 1,266 | 4 | 421 | SGN-U569959 | EU491503 | 8.68 |
| *SlSUN2*[e] | Solyc01g009340 | SL2.40ch01:3537789…3541754 (+) | 1,521 | 6 | 506 | | | 54.49 |
| *SlSUN3* | Solyc01g088250.2.1 | SL2.40ch01:74824669…74827406 (+) | 1,377 | 4 | 458 | SGN-U567883 | | 40.38 |
| *SlSUN4*[e] | Solyc01g097490 | SL2.40ch01:80081666…80082520 (+) | 1,233 | 2 | 410 | SGN-U575716 | | 1.12 |
| *SlSUN5* | Solyc02g077260.2.1 | SL2.40ch02:36828296…36831764 (−) | 909 | 4 | 302 | SGN-U566701 | | 23.74 |
| *SlSUN6*[e] | SL1.00sc00090_96 | SL2.40ch02:41309977…41310589 (+) | 387 | 1 | 128 | | | 1.11 |
| *SlSUN7* | Solyc02g087760.2.1 | SL2.40ch02:44643248…44648496 (+) | 1,671 | 5 | 556 | SGN-U570588 | AK320299 | 98.97 |
| *SlSUN8* | Solyc03g026110.2.1 | SL2.40ch03:7899519…7903299 (−) | 1,461 | 3 | 486 | SGN-U586572 | AK327068 | 37.14 |
| *SlSUN9*[e] | | SL2.40ch03:9052195…9054923 (+) | 561 | 3 | 186 | | | 0.02 |
| *SlSUN10* | Solyc03g083100.2.1 | SL2.40ch03:46466257…46469172 (−) | 1,410 | 3 | 469 | SGN-U576764 | AK325058 | 106.36 |
| *SlSUN11* | Solyc03g121760.2.1 | SL2.40ch03:63883128…63885625 (−) | 1,290 | 4 | 429 | | | 10.49 |
| *SlSUN12* | Solyc04g016480.2.1 | SL2.40ch04:7305326…7308804 (−) | 1,233 | 4 | 410 | SGN-U585221 | AK320616 | 29.07 |
| *SlSUN13* | Solyc04g050050.2.1 | SL2.40ch04:44467392…44470126 (−) | 1,185 | 3 | 394 | SGN-U603215 | | 14.30 |
| *SlSUN14* | Solyc04g081210.2.1 | SL2.40ch04:62803374…62809165 (+) | 2,589 | 5 | 862 | SGN-U563761 | | 146.20 |
| *SlSUN15* | Solyc05g007130.2.1 | SL2.40ch05:1694652…1699497 (+) | 1,656 | 5 | 551 | SGN-U569068 | AK322457 | 66.48 |
| *SlSUN16* | Solyc06g052010.1.1 | SL2.40ch06:32163038…32167768 (+) | 1,194 | 3 | 397 | SGN-U598310 | AK323901 | 5.23 |
| *SlSUN17* | Solyc06g053450.2.1 | SL2.40ch06:32744659…32751251 (+) | 1,779 | 4 | 592 | SGN-U581234 | AK321552 | 511.99 |
| *SlSUN18* | Solyc06g066430.2.1 | SL2.40ch06:38090484…38092311 (−) | 1,179 | 2 | 392 | SGN-U604798 | | 33.88 |
| *SlSUN19* | Solyc08g007920.1.1 | SL2.40ch08:2425707…2427454 (+) | 705 | 2 | 234 | | | 1.19 |
| *SlSUN20* | Solyc08g007930.1.1 | SL2.40ch08:2436580…2438211 (+) | 684 | 2 | 227 | | | 1.72 |
| *SlSUN21* | Solyc08g014280.2.1 | SL2.40ch08:3966331…3970317 (+) | 1,620 | 4 | 539 | SGN-U581070 | AK321872 | 268.94 |
| *SlSUN22* | Solyc08g062940.2.1 | SL2.40ch08:49582645…49584764 (+) | 930 | 5 | 309 | SGN-U602929 | | 137.36 |
| *SlSUN23* | Solyc08g080470.2.1 | SL2.40ch08:60928845…60932810 (+) | 1,500 | 3 | 499 | SGN-U569631 | AK247102 | 248.36 |
| *SlSUN24* | Solyc08g083240.2.1 | SL2.40ch08:62923263…62925544 (−) | 1,470 | 4 | 489 | SGN-U569480 | | 90.36 |
| *SlSUN25* | Solyc09g007410.2.1 | SL2.40ch09:985216…988218 (−) | 1,452 | 4 | 483 | SGN-U575982 | AK328336 | 28.99 |
| *SlSUN26* | Solyc09g082560.2.1 | SL2.40ch09:63677616…63679791 (−) | 1,404 | 3 | 467 | SGN-U581815 | AK322916 | 674.78 |

**Table 1** continued

| Gene name | Gene locus[a] | Position[b] | CDS (bp) | Intron[c] | Protein (aa) | Unigene | cDNA | RNAseq (RPKM)[d] |
|---|---|---|---|---|---|---|---|---|
| *SlSUN27* | Solyc10g005000.2.1 | SL2.40ch10:4134…5923 (−) | 1,185 | 4 | 394 | SGN-U565477 | | 63.59 |
| *SlSUN28* | Solyc10g008790.2.1 | SL2.40ch10:2859728…2865652 (−) | 915 | 4 | 304 | SGN-U582866 | AK321732 | 26.51 |
| *SlSUN29* | Solyc10g084280.1.1 | SL2.40ch10:63223672…63226246 (+) | 1437 | 4 | 478 | SGN-U575980 | BT013378 | 836.66 |
| *SlSUN30* | Solyc10g086060.1.1 | SL2.40ch10:64348811…64350952 (−) | 1,416 | 4 | 471 | SGN-U575981 | AK325367 | 133.49 |
| *SlSUN31* | Solyc11g071840.1.1 | SL2.40ch11:52274253…52280254 (+) | 1,347 | 5 | 448 | SGN-U576265 | | 9.95 |
| *SlSUN32* | Solyc12g008520.1.1 | SL2.40ch12:1931971…1934807 (+) | 1,230 | 2 | 409 | | | 12.36 |
| *SlSUN33* | Solyc12g014130.1.1 | SL2.40ch12:4940650…4946027 (+) | 786 | 4 | 261 | | | 66.31 |

[a] Gene locus from ITAG2.3

[b] Physical position on tomato WGS chromosomes (2.40)

[c] The number of intron in coding region

[d] RNAseq data in this study, maximum average valule in the 11 tissues

[e] The gene predicted by FGENESH

from *SlOFP2* to *SlOFP31* based on their position on the chromosome (Table 2; Fig. 1). There was a cluster of eight *SlOFP* genes on chromosome 10: *SlOFP21–SlOFP28* (Table 2; Fig. 1). The expression of 20 *SlOFP* genes was supported by unigene, full-length cDNA, *S. peruvianum* de novo transcriptome and/or RNA-seq results from this study (Table 2, Online source 3). Expression for the 11 remaining *SlOFP* genes was below the threshold level of 2 RPKM.

### *Phylogenetic analysis of OFP genes in Arabidopsis and tomato*

A dendrogram based on the OVATE domain was constructed to uncover the phylogenetic relationships between *Arabidopsis* and tomato *OFPs* (Fig. 4). The phylogenetic tree illustrated that the AtOFP and SlOFP proteins were divided into three subfamilies (Fig. 4). The detailed information of closest ortholog pairs between AtOFPs and SlOFPs was listed in Online source 5. OVATE was a paralog of SlOFP6, and their ortholog was likely represented by AtOFP7. In some subfamilies, *SlOFP* genes appeared to have expanded in tomato compared to *Arabidopsis*. For example, within subfamily 1, there were eight SlOFP proteins (from SlOFP22 to SlOFP29) and only one ortholog AtOFP13 in *Arabidopsis*. On the other hand, several AtOFP and SlOFP proteins demonstrated a one-to-one orthologous relationship, such as SlOFP5 and AtOFP5, SlOFP7 and AtOFP14, and SlOFP15 and AtOFP9 (Fig. 4; Online source 5).

### *The expression pattern of SlOFP genes in wild tomato*

We examined seventeen *SlOFP* genes expressed in the wild tomato tissues for this study. *SlOFP20* is the highest expressed gene of this family with 175.05 RPKM in one of the 11 tissues combined (Table 2; Online source 3). *OVATE* was expressed slightly higher in the vegetative meristem, young flower bud, flower at anthesis and fruit at 33 DPA (Fig. 5a; Online source 3). Several *SlOFP* genes were specifically expressed in one or more tissue. *SlOFP7* was specifically expressed in fruit at 20 DPA; *SlOFP8* and *SlOFP20* were specifically expressed in anthesis-stage flower; *SlOFP10* was specifically expressed in the root and hypocotyl; *SlOFP13* was specifically expressed in the root; *SlOFP14* was specifically expressed in fruit at 10 and 20 DPA; *SlOFP18* was specifically expressed in young flower buds; *SlOFP22* was specifically expressed in young leafs; *SlOFP29* was specifically expressed in fruit at 10 DPA. On the other hand, *SlOFP30* demonstrated similar expression in all tissues that were evaluated (Fig. 5; Online source 3).

## The *YABBY* genes in tomato

### *Identification of YABBY genes in tomato*

Nine *YABBY* genes were identified in the tomato genome. They were named by their likely orthologous relationship with *Arabidopsis YABBY* genes (Table 3; Fig. 6). *SlYABBY2b* was renamed as *FAS* because its mutation

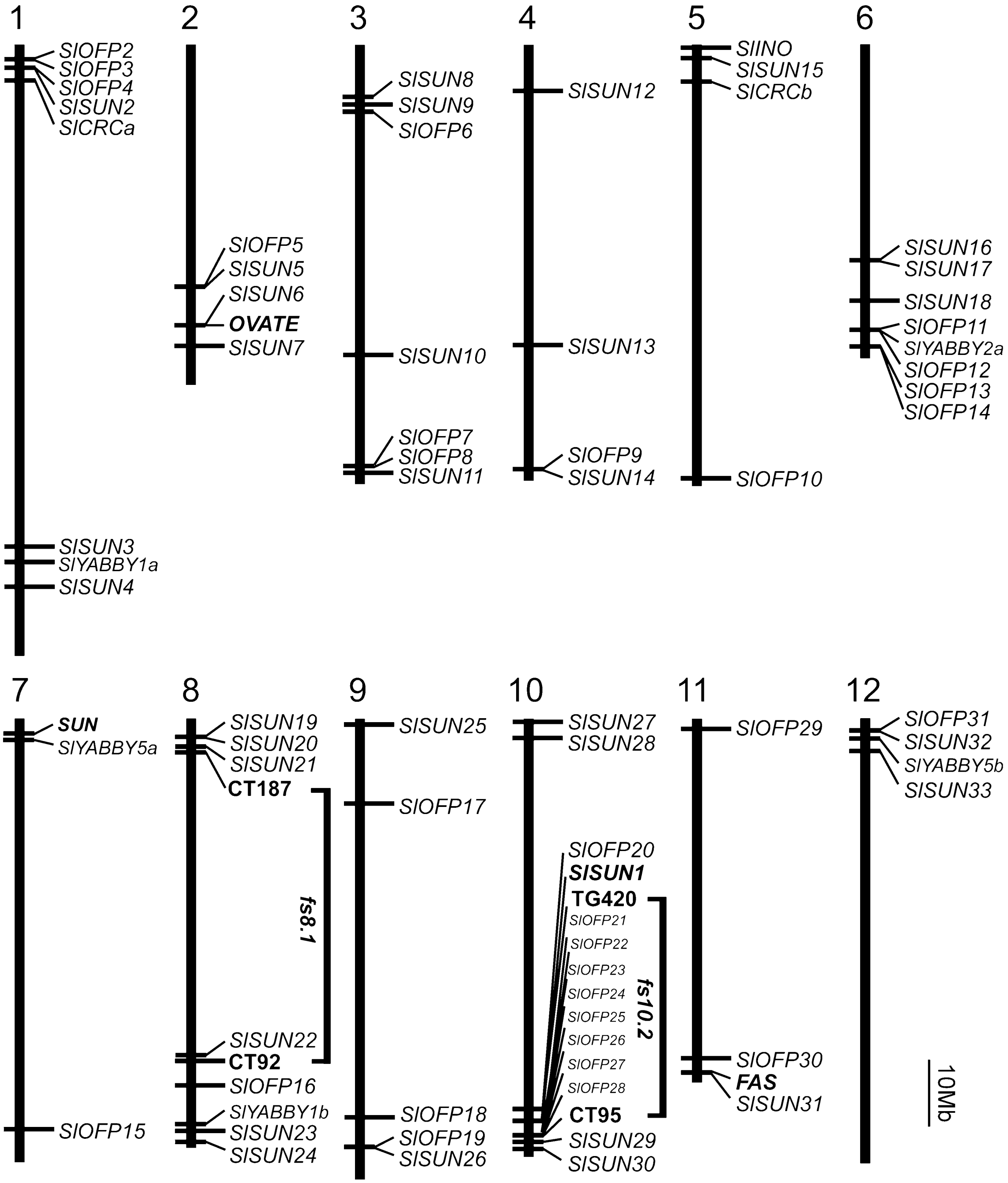


**Fig. 1** Chromosomal distribution of tomato *SUN*, *OFP* and *YABBY* genes. The position of *SlSUN*, *SlOFP* and *SlYABBY* genes on the chromosome was based on tomato WGS chromosome (SL2.40). The region of *fs8.1* locus was modified from the paper (Ku et al. 2000), and the region of *fs10.2* locus was modified from the review (Grandillo et al. 1999)

underlies the *FASCIATED* phenotype (Cong et al. 2008). The nine *YABBY* genes were distributed on 7 chromosomes, *SlCRCa* and *SlYABBY1a* were located on chromosome 1, *SlINO* and *SlCRCb* were located on chromosome 5, *SlYABBY2a* was located on chromosome 6, *SlYABBY5a* was located on chromosome 7, *SlYABBY1b* was located on chromosome 8, *FAS* was located on chromosome 11 and *SlYABBY5b* was located on chromosome 12 (Fig. 1; Table 3). Full-length cDNA or unigene sequences were available for six of these genes. All *YABBY* genes demonstrated expression in the tissues examined in this study (Table 3).

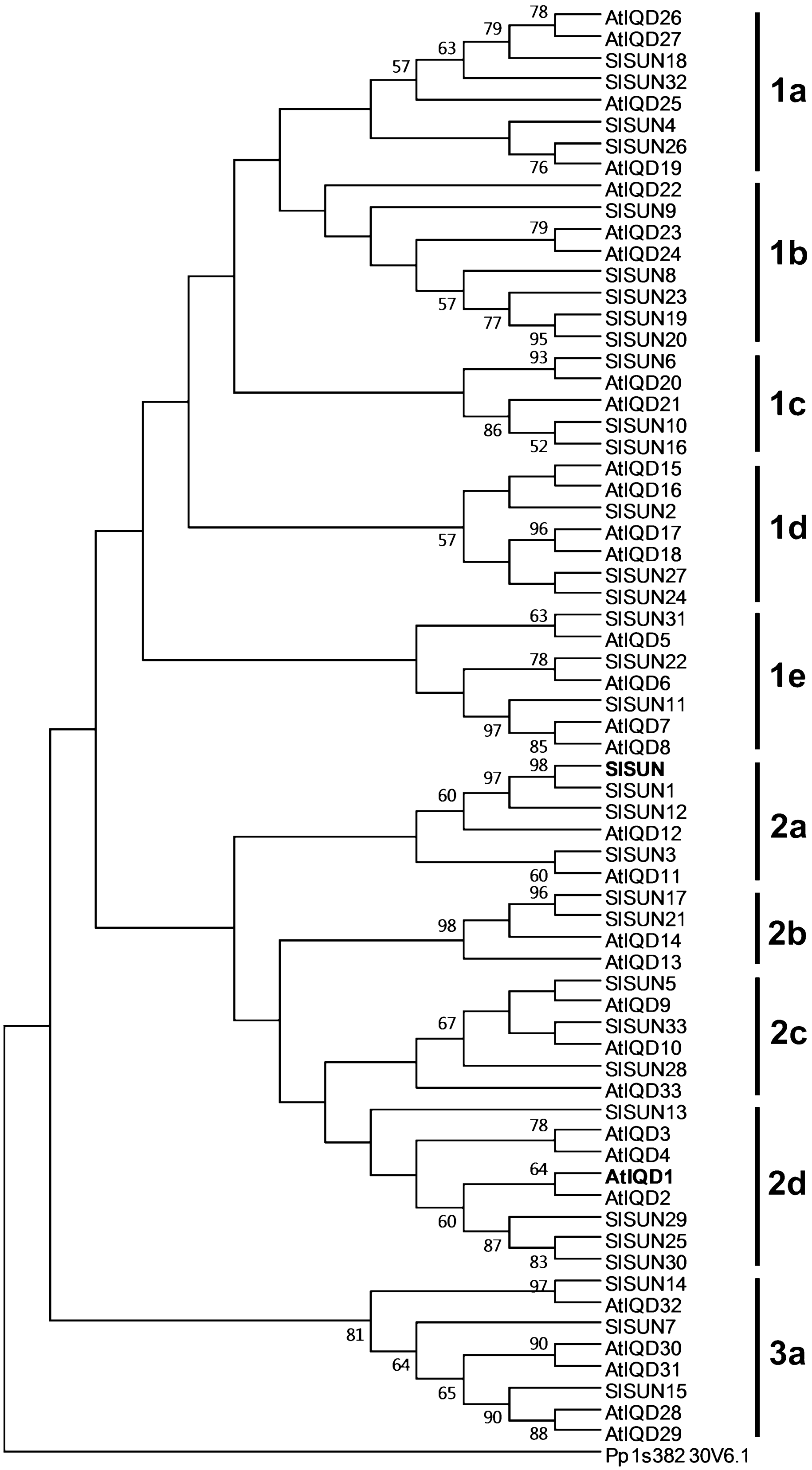


**Fig. 2** Phylogenetic tree of the *AtIQDs* and *SlSUNs* based on their IQ67 domain sequence. This tree is unrooted tree and is illustrated using gene *Pp1s382_30V6.1* in *Physcomitrella patens* subsp. *Patens* as an outgroup. Low bootstrap support (<50 %) was not reported

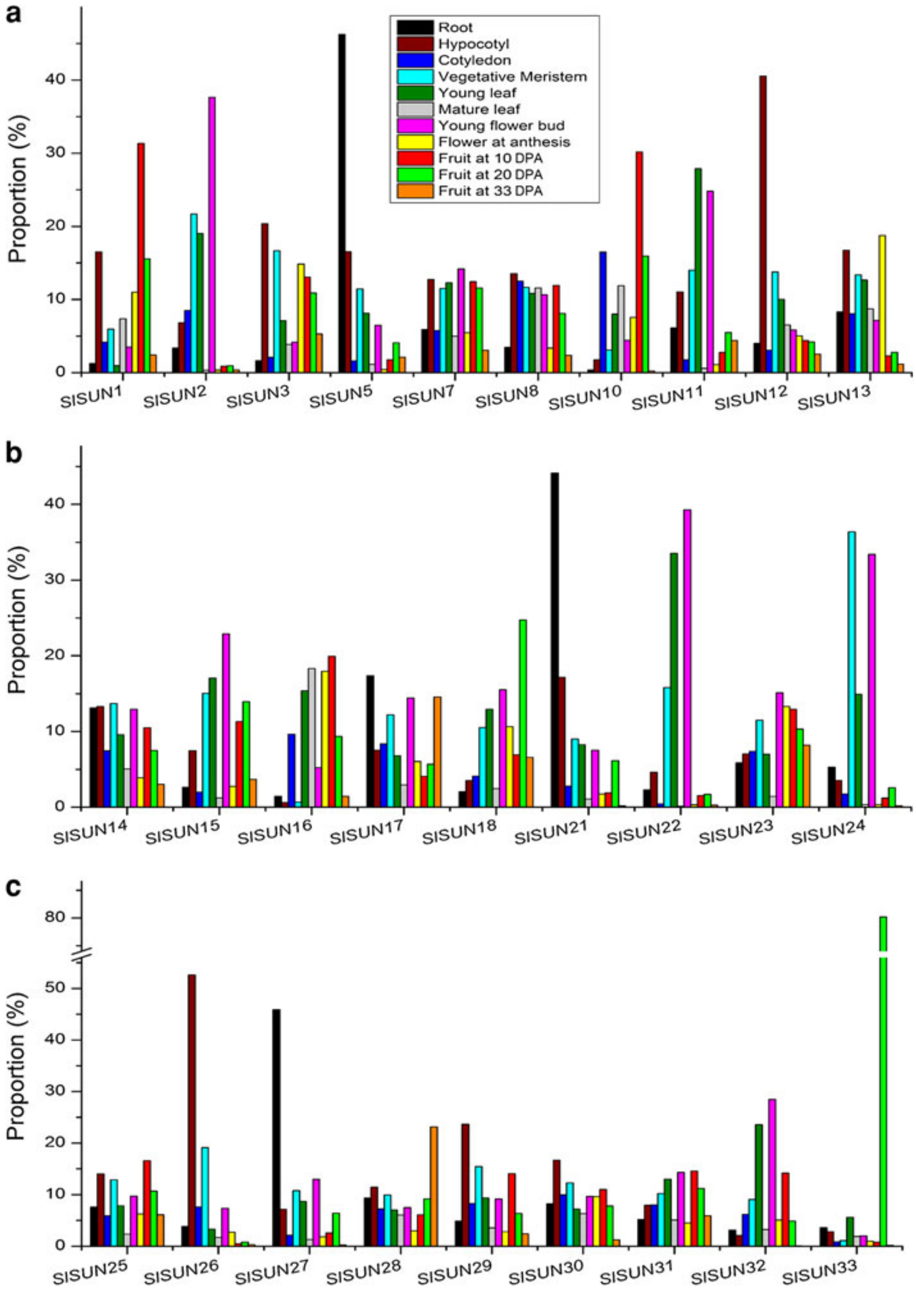


**Fig. 3** Expression pattern of *SlSUN* genes in tomato LA1589. **a** Genes from *SlSUN1* to *SlSUN13*, **b** genes from *SlSUN14* to *SlSUN24*, **c** genes from *SlSUN25* to *SlSUN33*

### *Phylogenetic analysis of YABBY genes in Arabidopsis and tomato*

To understand the phylogenetic relationships between YABBY proteins in *Arabidopsis* and tomato, we constructed a dendrogram based on the YABBY domain (Fig. 6). The phylogenetic tree showed that the AtYABBY and SlYABBY proteins were divided into five groups: INO, CRC, YAB2, YAB1/YAB3 and YAB5 (Fig. 6; Online source 6). The pattern of the tree was largely consistent with a previously reported tree (Toriba et al. 2007). Among the five orthologous groups, AtINO and SlINO in the INO group showed a one-to-one orthologous relationship; AtFIL, AtYABBY3, SlYABBY1a and SlYABBY1b in the YAB1/YAB3 group showed a two-to-two orthologous relationship; AtCRC, SlCRCa and SlCRCb in CRC group, AtYABBY2, SlYABBY2a and FAS (SlYABBY2b) in YAB2 group, AtYABBY5, and SlYABBY5a and SlYABBY5b in YAB5 group showed a one-to-two orthologous relationship (Fig. 6; Online source 6).

### *The expression pattern of YABBY genes in wild tomato*

The *SlYABBY* genes were either not expressed or were they expressed at very low levels in the root (Fig. 5c; Online source 3). *SlCRCa*, *SlCRCb* and *SlINO* were highly

**Table 2** *OFP* gene family in tomato

| Gene name | Gene locus[a] | Position[b] | CDS (bp) | Intron[c] | Protein (aa) | Unigene | cDNA | | RNAseq (RPKM)[d] |
|---|---|---|---|---|---|---|---|---|---|
| *OVATE* | Solyc02g085510.1.1 | SL2.40ch02:42945361…42947025 (+) | 858 | 1 | 285 | SGN-U582169 | AK247861 | a73864 | 18.63 |
| *SlOFP2* | Solyc01g007800.2.1 | SL2.40ch01:1955327…1956217 (+) | 549 | 0 | 182 | SGN-U602443 | AK319748 | a14189 | 37.54 |
| *SlOFP3* | Solyc01g007810.1.1 | SL2.40ch01:1971082…1971978 (−) | 897 | 0 | 298 | | | a87232 | 19.26 |
| *SlOFP4*[e] | SL1.00sc02618_4.1.1 | SL2.40ch01:3532060…3532833 (+) | 774 | 0 | 257 | | | | 0.00 |
| *SlOFP5* | Solyc02g072030.1.1 | SL2.40ch02:35908214…35909347 (−) | 1,134 | 0 | 377 | | | a401730 | 6.42 |
| *SlOFP6* | Solyc03g034100.2.1 | SL2.40ch03:10070572…10072340 (+) | 1,176 | 2 | 391 | | | | 1.34 |
| *SlOFP7* | Solyc03g120190.2.1 | SL2.40ch03:62702469…62703598 (+) | 828 | 1 | 275 | | | | 2.04 |
| *SlOFP8* | Solyc03g120790.1.1 | SL2.40ch03:63116563…63117252 (+) | 690 | 0 | 229 | | | | 22.93 |
| *SlOFP9* | Solyc04g080210.1.1 | SL2.40ch04:62055156…62055956 (−) | 801 | 0 | 266 | | | | 0.43 |
| *SlOFP10* | Solyc05g055220.1.1 | SL2.40ch05:64106469…64107158 (−) | 690 | 0 | 229 | SGN-U584716 | | | 2.42 |
| *SlOFP11*[e] | Solyc06g073040 | SL2.40ch06:41391235…41391522 (−) | 288 | 0 | 95 | | | | 2.16 |
| *SlOFP12* | Solyc06g074020.2.1 | SL2.40ch06:42152025…42153093 (−) | 477 | 1 | 158 | | | | 0.97 |
| *SlOFP13* | Solyc06g082450.1.1 | SL2.40ch06:44545426…44546437 (−) | 567 | 1 | 188 | | | | 2.90 |
| *SlOFP14* | Solyc06g082460.1.1 | SL2.40ch06:44548662…44549717 (+) | 1,056 | 0 | 351 | SGN-U603533 | AK323647 | | 82.98 |
| *SlOFP15* | Solyc07g055240.1.1 | SL2.40ch07:60663461…60663967 (−) | 507 | 0 | 168 | | | a191825 | 0.58 |
| *SlOFP16* | Solyc08g068170.1.1 | SL2.40ch08:54495235…54496749 (−) | 1,515 | 0 | 504 | | | a128970 | 4.22 |
| *SlOFP17* | Solyc09g018200.1.1 | SL2.40ch09:13537071…13537826 (-) | 756 | 0 | 251 | | | a382136 | 0.28 |
| *SlOFP18*[e] | Solyc09g065350 | SL2.40ch09:59030019…59030899 (−) | 687 | 1 | 228 | | | a106617 | 3.15 |
| *SlOFP19* | Solyc09g082080.1.1 | SL2.40ch09:63313130…63313765 (−) | 636 | 0 | 211 | | | | 1.10 |
| *SlOFP20* | Solyc10g076180.1.1 | SL2.40ch10:58330861…58331826 (+) | 966 | 0 | 321 | SGN-U573115 | | | 175.05 |
| *SlOFP21* | Solyc10g082050.1.1 | SL2.40ch10:62268995…62269531 (+) | 537 | 0 | 178 | | | a302968 | 1.17 |
| *SlOFP22* | Solyc10g082060.1.1 | SL2.40ch10:62274760…62275518 (−) | 759 | 0 | 252 | SGN-U576698 | | | 6.17 |
| *SlOFP23* | Solyc10g083070.1.1 | SL2.40ch10:62291057…62292046 (−) | 990 | 0 | 329 | | | | 0.52 |
| *SlOFP24*[e] | | SL2.40ch10:62294354…62293690 (−) | 609 | 1 | 202 | | | | 0.04 |
| *SlOFP25*[e] | | SL2.40ch10:62301706…62302128 (−) | 423 | 0 | 140 | | | | 0.06 |
| *SlOFP26* | Solyc10g083080.1.1 | SL2.40ch10:62296279…62297088 (−) | 810 | 0 | 269 | | | | 0.94 |
| *SlOFP27* | Solyc10g083090.1.1 | SL2.40ch10:62299100…62300089 (−) | 990 | 0 | 329 | | | | 0.23 |
| *SlOFP28* | Solyc10g083100.1.1 | SL2.40ch10:62304379…62305080 (−) | 702 | 0 | 233 | | | | 0.31 |
| *SlOFP29* | Solyc11g006670.1.1 | SL2.40ch11:1276863…1277597 (−) | 735 | 0 | 244 | | | | 30.58 |

**Table 2** continued

| Gene name | Gene locus[a] | Position[b] | CDS (bp) | Intron[c] | Protein (aa) | Unigene | cDNA | RNAseq (RPKM)[d] |
|---|---|---|---|---|---|---|---|---|
| *SlOFP30* | Solyc11g068780.1.1 | SL2.40ch11:50496070…50496483 (−) | 414 | 0 | 137 | SGN-U600438 | a197013 | 21.48 |
| *SlOFP31*[e] | SL1.00sc03540_201.1.1 | SL2.40ch12:1524186…1525400 (−) | 825 | 1 | 274 | | | 6.09 |

[a] Gene locus from ITAG2.3

[b] Physical position on tomato WGS chromosomes (2.40)

[c] The number of intron in coding region

[d] RNAseq in this study, maximum average valule in the 11 tissues

[e] The gene predicted by FGENESH

expressed in reproductive tissues. *SlCRCa* was specifically expressed in young flower buds; *SlCRCb* was specifically expressed in young flower buds and flowers at anthesis; *SlINO* was specifically expressed in flowers at anthesis (Fig. 5c; Online source 3). To study the three genes in more detail in reproductive tissues, we evaluated their expression pattern in floral and fruit tissues at different developmental stages using semi-quantitative RT-PCR (Online source 6). *SlCRCa* transcripts were only detected during the early stage of flower development, namely 10 days before anthesis (DBA) and 5 DBA. *SlCRCb* transcripts were detected in flowers at 10 DBA until 2 DPA in the developing fruit. The peak of *SlCRCb* expression was in anthesis-stage ovaries. *SlINO* transcripts were detected in flowers at 5 DBA until 2 DPA of the developing fruit. The peak of the *SlINO* transcripts was also found in anthesis-stage ovaries (Fig. 5c; Online source 6).

The other *SlYABBY* genes also showed different expression patterns even though they belonged to the same phylogenetic group. For example, *SlYABBY1a* was expressed in young flower bud at level of 419.3 RPKM and in flower at anthesis at level of 121.0 RPKM, whereas *SlYABBY1b* was expressed in young flower bud at level of 121.5 RPKM and in flower at anthesis at level of 37.8 RPKM. *SlYABBY2a* was expressed at much higher levels than *FAS* (*SlYABBY2b*) in all productive tissues. In young flower bud, flower at anthesis, fruit at 10, 20 and 33 DPA, *SlYABBY2a* were expressed at levels of 146.1, 578.4, 392.6, 191.2, 206.1 RPKM, respectively, whereas, *SlYABBY2b* was expressed at levels of 105.3, 81.2, 38.2, 16.0, 11.9 RPKM, respectively. *SlYABBY5a* was expressed at higher levels than *SlYABBY5b* in all tissues we detected in this study (Online source 3).

## Discussion

### The *SlSUN* genes

Orthologs are genes that originate from a single ancestral gene in the last common ancestor of the species and are likely to have equivalent functions (Fitch 1970; Koonin 2005). Four pairs of putative one-to-one orthologous genes were found between *SlSUN* and *AtIQD* genes (Fig. 2; Online source 2). Three of these pairs had a similar expression pattern in tomato and *Arabidopsis*: *SlSUN14* and *AtIQD32*, and *SlSUN31* and *AtIQD5* are almost ubiquitously expressed, whereas *SlSUN22* and *AtIQD6* are highly expressed in young flower buds (Fig. 3; Online source 2, Online source 3) (Schmid et al. 2005). Their similar expression patterns suggest that these orthologous pairs may play equivalent roles in growth and development.

Paralogs are genes originating from duplication within one organism and may have more divergent functions (Fitch 1970; Koonin 2005). Eleven pairs of putative paralogs were found in *SlSUN* gene family (Online source 2). Several pairs of paralogs showed a similar expression pattern, which suggests that they might share a common or similar function. For example, *SlSUN11* and *SlSUN22* were highly expressed in both young leaves and also young flower buds, *SlSUN25*, *SlSUN29* and *SlSUN30* were expressed almost equally (Fig. 3). Several pairs of paralogs have a different expression pattern, suggesting they play a diverse role in tomato development. For example, *SlSUN1* demonstrated highest expression in fruit at 10 DPA but *SlSUN12* demonstrated greatest expression in the hypocotyls; *SlSUN5* showed greatest expression in the root but *SlSUN28* had greatest expression in ripening fruit; *SlSUN17* was evenly expressed in almost all tissues, yet *SlSUN21* demonstrated highest expression in the root; *SlSUN24* had greater expression in both vegetative meristems and also young flower buds but *SlSUN27* showed much greater expression in the root (Fig. 3; Online source 2, Online source 3).

### The *SlOFP* genes

The tomato *OVATE* gene is the founding member of the *OFP* family. Its loss-of-function mutation results in an elongated tomato fruit. It is both thought to be a plant-growth suppressor and expressed in the reproductive organs

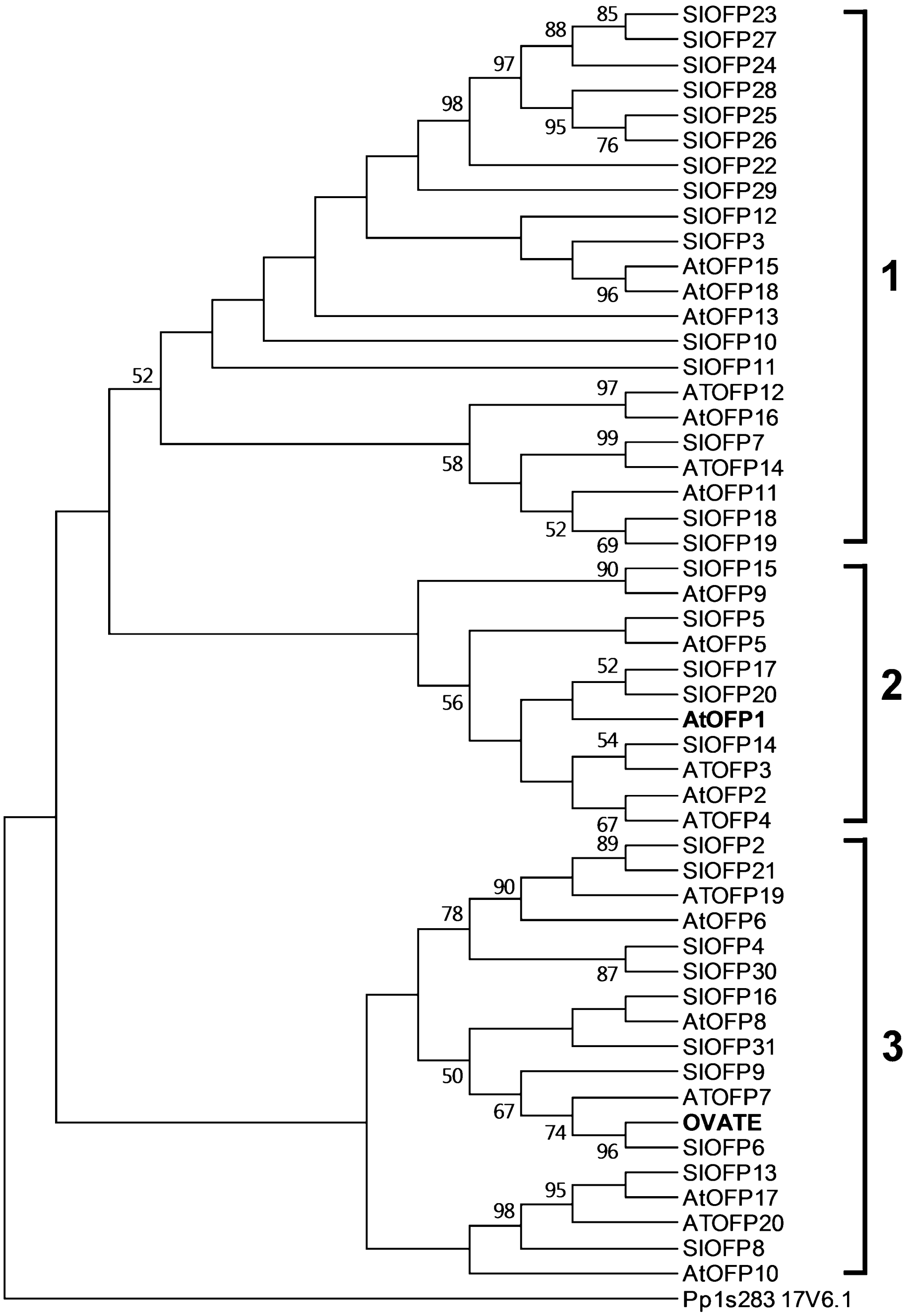


**Fig. 4** Phylogenetic tree of the AtOFPs and SlOFPs based on OVATE domain sequence. This tree is unrooted tree and is illustrated using gene *Pp1s283_17V6.1* in *Physcomitrella patens* subsp. *Patens* as an outgroup. Low bootstrap support (<50 %) was not reported. AtOFP19 (AT2G36026.1), AtOFP20 (AT1G06923.1)

in the early stages of flower and fruit development as determined by real-time PCR analysis (Liu et al. 2002). In this study, we found that *OVATE* was indeed expressed in vegetative meristem, but its expression in the reproductive organs showed a different pattern from what was previously reported. In this study, *OVATE* demonstrated high gene expression in young flower buds and decreased expression in 20 DPA fruit. *OVATE* expression also increased at the time of fruit ripening. A similar expression pattern of the *OVATE* gene was found for both the tomato cultivar Heinz1706 and the same wild tomato *S. pimpinellifolium* accession as was used in this study (Tomato-Genome-Consortium 2012). It might be interesting to further investigate the role of *OVATE* at the fruit ripening stage.

Several pairs of orthologs between *SlOFPs* to *AtOFPs* were shown to have a similar expression pattern, suggesting that they might share common functions. For example, *SlOFP7* and *AtOFP14* demonstrated greater expression in fruit/silique; *SlOFP13* and *AtOFP17* were expressed much higher in the root (Fig. 5a, Online source 3, Online source 5) (Schmid et al. 2005; Wang et al. 2011).

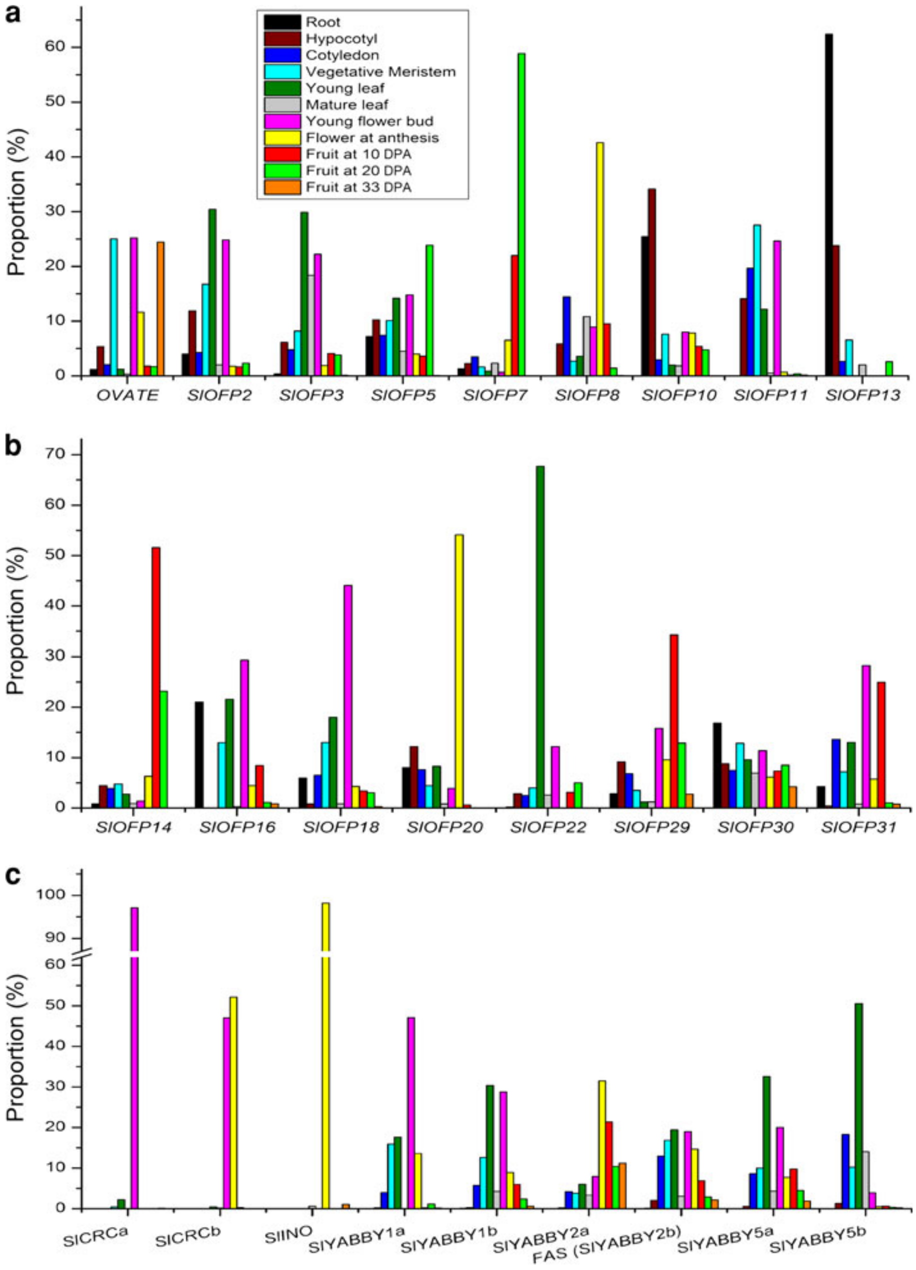


**Fig. 5** Expression pattern of *SlOFP* and *SlYABBY* genes in tomato LA1589. **a** Genes from *OVATE* to *SlOFP13*, **b** genes from *SlOFP14* to *SlOFP31*, **c** *SlYABBY* genes

Fourteen *SlOFP* genes were not expressed or were expressed at very low levels. The other members, except *SlOFP30*, were expressed at high levels in one or a few tissues. This suggests they have a specialized function in plant development. For example, *SlOFP8* and *SlOFP20* demonstrated much greater expression in anthesis-stage flowers; *SlOFP10* and *SlOFP13* were specifically expressed in the root and hypocotyl; *SlOFP14* and *SlOFP29* were expressed much higher in 10 DPA fruit. *SlOFP18* was specifically expressed in young flower buds; *SlOFP22* was expressed much higher in young leaves (Fig. 5; Online source 3).

## The *SlYABBY* genes

The expression pattern of tomato *YABBY* genes was similar to that of *Arabidopsis YABBY* genes. The *Arabidopsis YABBY* genes are divided into two classes based on their expression pattern: the reproductive and the vegetative *YABBY* genes. The reproductive *YABBY* genes of *Arabidopsis* include *CRC* and *INO*, which express exclusively in floral organs (Bowman and Smyth 1999; Villanueva et al. 1999). In contrast, the vegetative *YABBY* genes of *Arabidopsis*, including *FIL* (*YAB1*), *YAB2*, *YAB3*, and *YAB5*, are expressed in the leaf-derived organs, such as

**Table 3** *YABBY* gene family in tomato

| Gene name | Gene locus[a] | Position[b] | CDS (bp) | Intron[c] | Protein (aa) | Unigene | cDNA | RNAseq (RPKM)[d] |
|---|---|---|---|---|---|---|---|---|
| *SlCRCa* | Solyc01g010240.2.1 | SL2.40ch01:5028897…5031428 (+) | 507 | 5 | 168 | | | 42.06 |
| *SlCRCb* | Solyc05g012050.2.1 | SL2.40ch05:5275528…5277367 (−) | 477 | 6 | 158 | SGN-U572646 | | 10.04 |
| *FASCIATED (SlYABBY2b)* | Solyc11g071810.1.1 | SL2.40ch11:52249470…52255656 (−) | 534 | 5 | 177 | SGN-U578286 | AK248039 | 107.61 |
| *SlINO* | Solyc05g005240.1.1 | SL2.40ch05:191257…193043 (−) | 579 | 6 | 192 | | | 2.56 |
| *SlYABBY1a* | Solyc01g091010.2.1 | SL2.40ch01:76475369…76478983 (−) | 657 | 6 | 218 | SGN-U583546 | | 419.33 |
| *SlYABBY1b* | Solyc08g079100.2.1 | SL2.40ch08:59908783…59911941 (−) | 651 | 6 | 216 | SGN-U583545 | AK326840 | 128.31 |
| *SlYABBY2a* | Solyc06g073920.2.1 | SL2.40ch06:42038601…42044688 (+) | 579 | 5 | 192 | SGN-U580931 | AK328263 | 578.39 |
| *SlYABBY5a* | Solyc07g008180.2.1 | SL2.40ch07:2916878…2921216 (+) | 543 | 6 | 180 | SGN-U577176 | AK246138 | 166.38 |
| *SlYABBY5b* | Solyc12g009580.1.1 | SL2.40ch12:2837633…2843798 (−) | 576 | 6 | 191 | | | 63.33 |

[a] Gene locus from ITAG2.3

[b] Physical position on tomato WGS chromosomes (2.40)

[c] The number of intron in coding region

[d] RNAseq in this study, maximum average valule in the 11 tissues

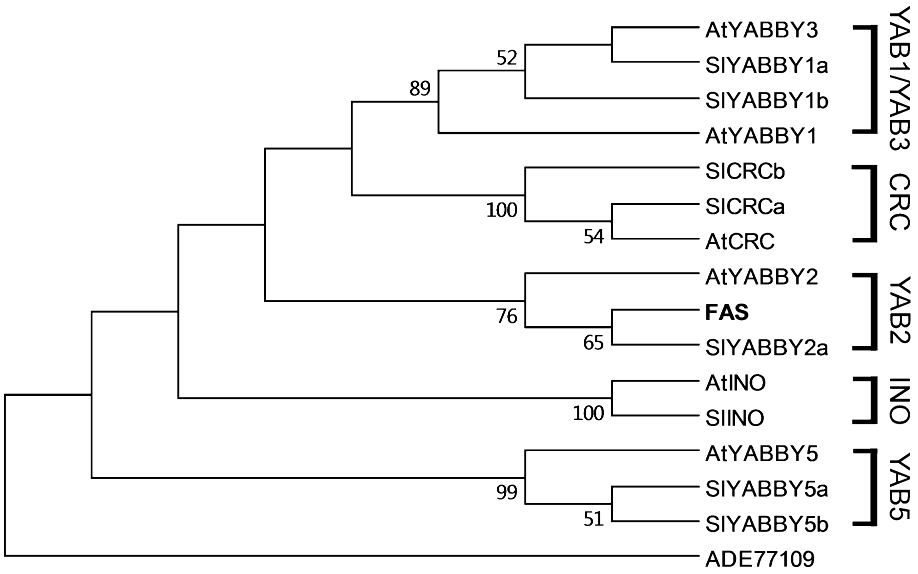
cotyledons, leaves, and floral organs (Sawa et al. 1999; Siegfried et al. 1999; Watanabe and Okada 2003; Stahle et al. 2009; Sarojam et al. 2010). The tomato *CRCa*, *CRCb* and *INO* genes, the orthologs of *Arabidopsis* reproductive *YABBY* genes, were expressed in flower and the early stage of fruit development (Fig. 5c; Online source 6). On the other hand, and as expected, tomato *FAS*, *YABBY2a*, *YABBY1a*, *YABBY1b*, *YABBY5a* and *YABBY5b* genes were also expressed in vegetative tissues (Fig. 5c).

The analysis of *YABBY* mutants suggests that their function has diversified during evolution, despite belonging to the same group in the phylogenetic tree (Yamaguchi et al. 2004; Cong et al. 2008). *Arabidopsis CRC* and *O. sativa DL* belong to the *CRC* group, and they both play a role in carpel development. However, *O. sativa DL* is also involved in leaf development, whereas *Arabidopsis CRC* expresses exclusively in floral organs (Bowman and Smyth 1999; Yamaguchi et al. 2004). Two *CRC* genes, *SlCRCa* and *SlCRCb*, were identified in tomato (Table 3). They were only expressed in reproductive tissues but showed a different expression pattern. *SlCRCa* was specifically expressed at the early stage of flower development (flower buds at 10 days or more before anthesis). *SlCRCb* is equally expressed at very young floral stages as well as the anthesis stage (Fig. 5c; Online source 6). The different expression pattern of *SlCRCa* and *SlCRCb* suggests that they might play different roles in reproductive tissues development. Similarly, two *YABBY2* genes, *FAS* (*SlYABBY2b*) and *SlYABBY2a*, have been identified in tomato, and only one *YABBY2* gene in *Arabidopsis*. *FAS* and *SlYABBY2a* showed different expression patterns in tomato tissues. *SlYABBY2a* demonstrated higher expression level than *FAS* did in all productive tissues we detected in this study (Online source 3). The knockout of the *FAS* gene results in an increase of carpel and locule number in tomato (Cong et al. 2008). However, there is no evidence that the *Arabidopsis YABBY2* gene is involved in regulating carpel number. This suggests that the members in the *YABBY2* group of tomato may have gained a new function during evolution.

## Duplication mechanisms accounting for the expansion of *SUN*, *OFP* and *YABBY* families

We noted that certain subfamilies of the *SlSUN*, *SlOFP* and *SlYABBY* families showed gene expansion. Gene family member expansions usually result from duplications, such as tandem duplications, segmental duplications and polyploidization or whole-genome duplications (Sankoff 2001; Adams and Wendel 2005). Whole genome duplication has occurred in tomato, and most of collinear blocks were located at the top and bottom part of the chromosomes (Song et al. 2012). Most of *SlSUN*, *SlOFP* and *SlYABBY* genes were also located at the top and bottom part of

**Fig. 6** Phylogenetic tree of the YABBY proteins in tomato and *A. thaliana* based on YABBY domain sequence. This tree is unrooted tree and is illustrated using protein ADE77109 in *Picea sitchensis* as an outgroup. Low bootstrap support (<50 %) was not reported



the chromosomes (Fig. 1), which suggests that whole-genome duplication may have played a significant role in the expansion of the three families.

Other types of duplication may also explain the expansion of the three families. *SUN* on chromosome 7 arose from a gene on chromosome 10 through a retrotransposon-mediated gene duplication (Xiao et al. 2008). The cluster of *SlSUN19* and *SlSUN20*, and the cluster of *SlOFPs* (from *SlOFP22 to SlOFP*28) might have arisen from tandem duplication, because they were close to each other on the chromosome and appeared in a close phylogenetic relationship as demonstrated by the dendogram. There was just one *SUN*-like gene, *PGSC0003DMG400005774* (Transcript_ID, *PGSC0003DMT400014796*), in potato in the homologous genomic region of tomato *SlSUN19* and *SlSUN20*. Using the divergence ratio $r = 6.5 \times 10^{-9}$ mutations per synonymous site per year (Gaut et al. 1996), the estimated divergence time of *SlSUN19* and *PGSC0003DMT400014796* was ~8.2 million years (Myr). The estimated divergence time of *SlSUN19* and *SlSUN20* was ~3.3 Myr (Online source 2). Therefore, *SlSUN19* and *SlSUN20* might have arisen from tomato-specific tandem duplication. However, tomato and potato might share the same kind of tandem duplication that results in the cluster of *SlOFPs* (Online source 4).

Semental duplication most likely explains the expansion of the tomato *YABBY2* subfamily. In the *YABBY2* subfamily, *Arabidopsis*, cucumber, and poplar had one member *AtYABBY2*, *Csa007814* and *Potri.016G067300.1*, respectively. Grape has two members *GSVIVG01022586001* (Transcript name, *GSVIVT01022586001*) and *GSVIVG01037533001* (Transcript name, *GSVIVT01037533001*); potato has two members *PGSC0003DMG400002988* (Transcript_ID, *PGSC0003DMT400007731*) and *PGSC0003DMG400005936* (Transcript_ID, *PGSC0003DMT400015197*); and tomato has two members *FAS* (*SlYABBY2b*) and *SlYABBY2a* (Online source 6). In this study, the estimated divergence time of tomato gene *SlYABBY2a* and potato gene *PGSC0003DMT400015197* was ~5.9 Myr, and the estimated divergence time of *SlYABBY2b* and *PGSC0003DMT400007731* was ~10.6 Myr. Their divergence time was close to what has been reported for these two species (~7.3 Myr ago) (Tomato-Genome-Consortium 2012). The genomic regions around these orthologous pairs *SlYABBY2a* and *PGSC0003DMT400015197*, *SlYABBY2b* and *PGSC0003DMT400007731* were also very similar; however, the tomato genes *SlYABBY2a* and *SlYABBY2b* diverged ~50.7 Myr ago, and the potato genes *PGSC0003DMT400007731* and *PGSC0003DMT400015197* diverged ~41.0 Myr ago. These results indicate that the gene expansion of the tomato and potato subfamily might arise from a segmental duplication, and this duplication already existed before the differentiation of potato and tomato (Online source 6); however, this duplication might be independent to the duplication resulting in gene expansion of *V. vinifera YABBY2* subfamily. In this study, potato and tomato is estimated to separate from grape ~76.2 Myr ago. Whereas, grape genes *GSVIVT01022586001* and *GSVIVT01037533001* separated ~60.9 Myr ago, and the duplication in potato and tomato YABBY2 subfamilies arose ~50.7 Myr ago. Therefore, after tomato and potato diverged from grape, they duplicated in the YABBY2 subfamily separately (Online source 6).

After duplication, the genes may have evolved to acquire new functions in a process called neofunctionalization. A good example of this is *SUN* on chromosome 7 after it was inserted into *DEFL1* showing a different expression pattern compared to its ancestral copy on chromosome 10 (Fig. 3a) (Xiao et al. 2008; Xiao et al. 2009). This change resulted in a new function, even though the gene sequence did not change resulting in an elongated tomato fruit (Xiao et al. 2008).

### *SUN*, *OFP*, *YABBY* genes and fruit shape loci

Nearly 30 loci control tomato fruit shape (Grandillo et al. 1999). Four genes underlying these loci, namely *OVATE*, *SUN*, *FAS* and *LC* (*Locule Number*), have been cloned (Liu et al. 2002; Cong et al. 2008; Xiao et al. 2008; Munos et al. 2011). Identification of the *SUN*, *OFP* and *YABBY* gene family members may help to uncover the genes underlying the other tomato fruit shape loci. For example, *fs8.1* is a major locus controlling elongation fruit in tomato, and it is located in the centromeric region of chromosome 8 (Grandillo et al. 1996; Ku et al. 2000) and *SlSUN22* gene maps to this region (Fig. 1). *SlSUN22* was highly expressed in young flowers (Online source 3), suggesting that it might be a candidate gene of *fs8.1*. There was a cluster of *SlOFPs* on the bottom part of chromosome 10 (Fig. 1; Table 2) which overlaps with the tomato *fs10.2* region (Grandillo et al. 1999).

Varying levels of synteny exist among members of the Solanaceae family (Livingstone et al. 1999; Doganlar et al. 2002a; Tomato-Genome-Consortium 2012). QTL analysis has shown the existence of several overlapping fruit shape loci in eggplant, pepper and tomato (Doganlar et al. 2002b; Frary et al. 2003; Zygier et al. 2005; Paran and van der Knaap 2007; Borovsky and Paran 2011). Down regulation of *CaOvate* changes the shape of a round pepper into a more oblong shape (Tsaballa et al. 2011), suggesting that the *CaOvate* and *OVATE* might play a similar role in fruit shape determination. Thus, the identification of *SUN*, *OFP* and *YABBY* genes may also help to uncover the genes underlying the fruit shape loci in other Solanaceae species.

In summary, we identified 34 *SlSUN*, 31 *SlOFP* and 9 *SlYABBY* genes in tomato. Genome sequence analysis shows that some *SlSUNs* and *SlOFPs* mapped within several known fruit shape loci. The closest putative orthologs in the families between *Arabidopsis* and tomato were determined through their phylogenetic relationship and sequence similarity. Furthermore, some family members exhibited tissue-specific expression based on the RNA-seq analysis. Our results will pave the way to study the roles of *SlSUN*, *SlOFP* and *SlYABBY* genes in tomato growth and development and further understanding of these families in plant biology in general.

**Acknowledgments** We thank Dr. Zhangjun Fei from Boyce Thompson Institute for Plant Research for assistance with the RNA-seq analysis, the Molecular and Cellular Imaging Center with depositing the RNA-seq data in the small read archive, and Dr. Dean Fraga for assistance with the phylogenetic analysis. This work was supported by the National Science Foundation grant (IOS 0922661).

## References

Abel S, Savchenko T, Levy M (2005) Genome-wide comparative analysis of the *IQD* gene families in *Arabidopsis thaliana* and *Oryza sativa*. BMC Evol Biol 5:72

Adams KL, Wendel JF (2005) Polyploidy and genome evolution in plants. Curr Opin Plant Biol 8:135–141

Alvarez J, Smyth DR (1999) CRABS CLAW and SPATULA, two *Arabidopsis* genes that control carpel development in parallel with AGAMOUS. Development 126:2377–2386

Barrero LS, Cong B, Wu F, Tanksley SD (2006) Developmental characterization of the fasciated locus and mapping of *Arabidopsis* candidate genes involved in the control of floral meristem size and carpel number in tomato. Genome 49:991–1006

Borovsky Y, Paran I (2011) Characterization of fs10.1, a major QTL controlling fruit elongation in Capsicum. Theor Appl Genet 123(4):657–665

Bowman JL (2000) The YABBY gene family and abaxial cell fate. Curr Opin Plant Biol 3:17–22

Bowman JL, Smyth DR (1999) *CRABS CLAW*, a gene that regulates carpel and nectary development in *Arabidopsis*, encodes a novel protein with zinc finger and helix-loop-helix domains. Development 126:2387–2396

Buerstenbinder K, Savchenko T, Mueller J, Adamson AW, Stamm G, Kwong R, Zipp BJ, Dinesh DC, Abel S (2012) *Arabidopsis* calmodulin-binding iqd1 localizes to microtubules and interacts with kinesin light chain-related protein-1. J Biol Chem. doi:10.1074/jbc.M112.396200

Cong B, Barrero LS, Tanksley SD (2008) Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. Nat Genet 40:800–804

Doganlar S, Frary A, Daunay MC, Lester RN, Tanksley SD (2002a) A comparative genetic linkage map of eggplant (Solanum melongena) and its implications for genome evolution in the solanaceae. Genetics 161(4):1697–1711

Doganlar S, Frary A, Daunay MC, Lester RN, Tanksley SD (2002b) Conservation of gene function in the solanaceae as revealed by comparative mapping of domestication traits in eggplant. Genetics 161:1713–1726

Fitch WM (1970) Distinguishing homologous from analogous proteins. Syst Zool 19(2):99–113

Frary A, Doganlar S, Daunay MC, Tanksley SD (2003) QTL analysis of morphological traits in eggplant and implications for conservation of gene function during evolution of solanaceous species. Theor Appl Genet 107(2):359–370

Gaut BS, Morton BR, McCaig BC, Clegg MT (1996) Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene Adh parallel rate differences at the plastid gene rbcL. Proc Natl Acad Sci USA 93(19):10274–10279

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A, Briggs S (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science 296(5565):92–100

Golz JF, Hudson A (1999) Plant development: YABBYs claw to the fore. Curr Biol 9:R861–R863

Grandillo S, Ku HM, Tanksley SD (1996) Characterization of *fs8.1*, a major QTL influencing fruit shape in tomato. Mol Breed 2:251–260

Grandillo S, Ku HM, Tanksley SD (1999) Identifying the loci responsible for natural variation in fruit size and shape in tomato. Theor Appl Genet 99:978–987

Hackbusch J, Richter K, Muller J, Salamini F, Uhrig JF (2005) A central role of *Arabidopsis thaliana* ovate family proteins in networking and subcellular localization of 3-aa loop extension homeodomain proteins. Proc Natl Acad Sci USA 102:4908–4912

Huang ZJ, van der Knaap E (2011) Tomato fruit weight 11.3 maps close to fasciated on the bottom of chromosome 11. Theor Appl Genet 123(3):465–474

Jiang N, Gao D, Xiao H, van der Knaap E (2009) Genome organization of the tomato sun locus and characterization of the unusual retrotransposon Rider. Plant J 60:181–193

Kennedy GG (2003) Tomato, pests, parasitoids, and predators: tritrophic interactions involving the genus *Lycopersicon*. Annu Rev Entomol 48:51–72

Kimura S, Sinha N (2008) Tomato (*Solanum lycopersicum*): a model fruit-bearing crop. CSH Protoc. doi:10.1101/pdb.emo105

Klee HJ, Giovannoni JJ (2011) Genetics and control of tomato fruit ripening and quality attributes. Annu Rev Genet 45:41–59

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

Ku HM, Grandillo S, Tanksley SD (2000) fs8.1, a major QTL, sets the pattern of tomato carpel shape well before anthesis. Theor Appl Genet 101:873–878

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10(3):R25

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23(21):2947–2948

Levy M, Wang Q, Kaspi R, Parrella MP, Abel S (2005) *Arabidopsis* IQD1, a novel calmodulin-binding nuclear protein, stimulates glucosinolate accumulation and plant defense. Plant J 43:79–96

Li E, Wang S, Liu Y, Chen JG, Douglas CJ (2011a) OVATE FAMILY PROTEIN4 (OFP4) interaction with KNAT7 regulates secondary cell wall formation in *Arabidopsis thaliana*. Plant J 67(2):328–341

Li H, Liang W, Hu Y, Zhu L, Yin C, Xu J, Dreni L, Kater MM, Zhang D (2011b) Rice MADS6 interacts with the floral homeotic genes SUPERWOMAN1, MADS3, MADS58, MADS13, and DROOPING LEAF in specifying floral organ identities and meristem fate. Plant Cell 23(7):2536–2552

Lin Z, Li X, Shannon LM, Yeh CT, Wang ML, Bai G, Peng Z, Li J, Trick HN, Clemente TE, Doebley J, Schnable PS, Tuinstra MR, Tesso TT, White F, Yu J (2012) Parallel domestication of the Shattering1 genes in cereals. Nat Genet 44(6):720–724

Lippman Z, Tanksley SD (2001) Dissecting the genetic pathway to extreme fruit size in tomato using a cross between the small-fruited wild species *Lycopersicon pimpinellifolium* and *L. esculentum* var Giant Heirloom. Genetics 158:413–422

Liu J, Van Eck J, Cong B, Tanksley SD (2002) A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. Proc Natl Acad Sci USA 99:13302–13306

Livingstone KD, Lackney VK, Blauth JR, van Wijk R, Jahn MK (1999) Genome mapping in Capsicum and the evolution of genome structure in the Solanaceae. Genetics 152(3):1183–1202

Lora J, Hormaza JI, Herrero M, Gasser CS (2011) Seedless fruits and the disruption of a conserved genetic pathway in angiosperm ovule development. Proc Natl Acad Sci USA 108:5461–5465

Munos S, Ranc N, Botton E, Berard A, Rolland S, Duffe P, Carretero Y, Le Paslier MC, Delalande C, Bouzayen M, Brunel D, Causse M (2011) Increase in tomato locule number is controlled by two single-nucleotide polymorphisms located near WUSCHEL. Plant Physiol 156(4):2244–2254

Ohmori Y, Toriba T, Nakamura H, Ichikawa H, Hirano HY (2011) Temporal and spatial regulation of DROOPING LEAF gene expression that promotes midrib formation in rice. Plant J 65(1):77–86

Pagnussat GC, Yu HJ, Sundaresan V (2007) Cell-fate switch of synergid to egg cell in *Arabidopsis* eostre mutant embryo sacs arises from misexpression of the BEL1-like homeodomain gene BLH1. Plant Cell 19:3578–3592

Pang MX, Stewart JM, Zhang JF (2011) A mini-scale hot borate method for the isolation of total RNA from a large number of cotton tissue samples. Afr J Biotech 10(68):15430–15437

Paran I, van der Knaap E (2007) Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. J Exp Bot 58:3841–3852

Pearson WR (1991) Searching protein sequence libraries: comparison of the sensitivity and selectivity of the Smith–Waterman and FASTA algorithms. Genomics 11(3):635–650

Peralta IE, Spooner DM, Knapp S (2008) Taxonomy of tomatoes: a revision of wild tomatoes (*Solanum* section *Lycopersicon*) and their outgroup relatives in sections *Juglandifolia* and *Lycopersicoides*. Syst Bot Monogr 84:1–186

Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer EL, Eddy SR, Bateman A, Finn RD (2012) The Pfam protein families database. Nucleic Acids Res 40 (Database issue):D290–D301

Rodriguez GR, Munos S, Anderson C, Sim SC, Michel A, Causse M, Gardener BB, Francis D, van der Knaap E (2011) Distribution of SUN, OVATE, LC, and FAS in the tomato germplasm and the relationship to fruit shape diversity. Plant Physiol 156(1):275–285

Sankoff D (2001) Gene and genome duplication. Curr Opin Genet Dev 11:681–684

Sarojam R, Sappl PG, Goldshmidt A, Efroni I, Floyd SK, Eshed Y, Bowman JL (2010) Differentiating *Arabidopsis* shoots from leaves by combined YABBY activities. Plant Cell 22:2113–2130

Sawa S, Watanabe K, Goto K, Liu YG, Shibata D, Kanaya E, Morita EH, Okada K (1999) FILAMENTOUS FLOWER, a meristem and organ identity gene of *Arabidopsis*, encodes a protein with a zinc finger and HMG-related domains. Genes Dev 13:1079–1088

Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Scholkopf B, Weigel D, Lohmann JU (2005) A gene expression map of *Arabidopsis thaliana* development. Nat Genet 37(5):501–506

Siegfried KR, Eshed Y, Baum SF, Otsuga D, Drews GN, Bowman JL (1999) Members of the YABBY gene family specify abaxial cell fate in *Arabidopsis*. Development 126:4117–4128

Smith TF, Waterman MS (1981) Identification of common molecular subsequences. J Mol Biol 147(1):195–197

Song C, Guo J, Sun W, Wang Y (2012) Whole genome duplication of intra- and inter-chromosomes in the tomato genome. J Genet Genomics 39:361–368

Stahle MI, Kuehlich J, Staron L, von Arnim AG, Golz JF (2009) YABBYs and the transcriptional corepressors LEUNIG and LEUNIG_HOMOLOG maintain leaf polarity and meristem activity in *Arabidopsis*. Plant Cell 21:3105–3118

Sun JQ, Jiang HL, Li CY (2011) Systemin/jasmonate-mediated systemic defense signaling in tomato. Mol Plant 4:607–615

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28(10):2731–2739

Tanaka W, Toriba T, Ohmori Y, Yoshida A, Kawai A, Mayama-Tsuchida T, Ichikawa H, Mitsuda N, Ohme-Takagi M, Hirano

HY (2012) The YABBY gene TONGARI-BOUSHI1 is involved in lateral organ development and maintenance of meristem organization in the rice spikelet. Plant Cell 24(1):80–95

The-Arabidopsis-Genome-Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408(6814):796–815

Tomato-Genome-Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485(7400): 635–641

Toriba T, Harada K, Takamura A, Nakamura H, Ichikawa H, Suzaki T, Hirano HY (2007) Molecular characterization the YABBY gene family in Oryza sativa and expression analysis of OsYABBY1. Mol Genet Genomics 277(5):457–468

Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25(9):1105–1111

Tsaballa A, Pasentsis K, Darzentas N, Tsaftaris AS (2011) Multiple evidence for the role of an Ovate-like gene in determining fruit shape in pepper. BMC Plant Biol 11:46

van der Knaap E, Tanksley SD (2001) Identification and characterization of a novel locus controlling early fruit development in tomato. Theor Appl Genet 103:353–358

Villanueva JM, Broadhvest J, Hauser BA, Meister RJ, Schneitz K, Gasser CS (1999) INNER NO OUTER regulates abaxial- adaxial patterning in *Arabidopsis* ovules. Genes Dev 13:3160–3169

Wang S, Chang Y, Guo J, Chen JG (2007) *Arabidopsis* Ovate Family Protein 1 is a transcriptional repressor that suppresses cell elongation. Plant J 50:858–872

Wang YK, Chang WC, Liu PF, Hsiao MK, Lin CT, Lin SM, Pan RL (2010) Ovate family protein 1 as a plant Ku70 interacting protein involving in DNA double-strand break repair. Plant Mol Biol 74:453–466

Wang S, Chang Y, Guo J, Zeng Q, Ellis BE, Chen JG (2011) *Arabidopsis* ovate family proteins, a novel transcriptional repressor family, control multiple aspects of plant growth and development. PLoS One 6(8):e23896

Watanabe K, Okada K (2003) Two discrete cis elements control the Abaxial side-specific expression of the FILAMENTOUS FLOWER gene in *Arabidopsis*. Plant Cell 15:2592–2602

Wu S, Xiao H, Cabrera A, Meulia T, van der Knaap E (2011) SUN regulates vegetative and reproductive organ shape by changing cell division patterns. Plant Physiol 157(3):1175–1186

Xiao H, Jiang N, Schaffner E, Stockinger EJ, van der Knaap E (2008) A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. Science 319:1527–1530

Xiao H, Radovich C, Welty N, Hsu J, Li D, Meulia T, van der Knaap E (2009) Integration of tomato reproductive developmental landmarks and expression profiles, and the effect of SUN on fruit shape. BMC Plant Biol 9:49

Yamaguchi T, Nagasawa N, Kawasaki S, Matsuoka M, Nagato Y, Hirano HY (2004) The YABBY gene DROOPING LEAF regulates carpel specification and midrib development in *Oryza sativa*. Plant Cell 16:500–509

Zhong S, Joung JG, Zheng Y, Chen YR, Liu B, Shao Y, Xiang JZ, Fei Z, Giovannoni JJ (2011) High-throughput illumina strand-specific RNA sequencing library preparation. Cold Spring Harb Protoc 8:940–949

Zygier S, Chaim AB, Efrati A, Kaluzky G, Borovsky Y, Paran I (2005) QTLs mapping for fruit size and shape in chromosomes 2 and 4 in pepper and a comparison of the pepper QTL map with that of tomato. Theor Appl Genet 111:437–445