

Evolution of Lineage-Specific Gene Networks Underlying the Considerable Fruit Shape Diversity in Persimmon

Haruka Maeda¹, Takashi Akagi^{1,2,4,*}, Noriyuki Onoue³, Atsushi Kono³ and Ryutaro Tao¹

¹Division of Agronomy and Horticultural Science, Graduate School of Agriculture, Kyoto University, Kyoto, 606-8502 Japan

²Japan Science and Technology Agency (JST), PRESTO, Kawaguchi-shi, Saitama, 332-0012 Japan

³Institute of Fruit Tree and Tea Science, NARO, Higashihiroshima, 739-2494 Japan

⁴Present address: Graduate School of Environmental and Life Science, Okayama University, Tsushima-naka 1-1-1, Okayama, 700-8530 Japan

*Corresponding author: E-mail, takashia@okayama-u.ac.jp; Fax, +81-86-251-8337.

(Received March 7, 2019; Accepted July 7, 2019)

The shapes of plant organs reflect the evolution of each lineage and have been diversified according to lineage-specific adaptations to environment. Research on the molecular pathways responsible for organ shapes has traditionally been focused mainly on leaves or flowers. Thus, little is known about the pathways controlling fruit shapes, despite their diversity in some plant species. In this study, we analyzed oriental persimmon (*Diospyros kaki*), which exhibits considerable diversity in fruit shapes among cultivars, to elucidate the underlying molecular mechanism using transcriptomic data and quantitative evaluation. First, to filter the candidate genes associated with persimmon fruit shapes, the whole gene expression patterns obtained using mRNA-Seq analysis from 100 individuals, including a segregated population and various cultivars, were assessed to detect correlations with principal component scores for fruit shapes characterized with elliptic Fourier descriptors. Next, a gene co-expression network analysis with weighted gene co-expression network analysis (WGCNA) package revealed that class 1 KNOX family genes and *SEEDSTICK* function as integrators along with some phytohormone-related genes, to regulate the fruit shape diversity. On the other hand, the *OVATE* family genes also contribute to fruit shape diversity, of which pathway would be potentially shared with other plant species. Evolutionary aspects suggest that acquisition of a high lineage-specific and variable expression of class 1 KNOX gene, *knotted-like homeobox of Arabidopsis thaliana 1* (*KNAT1*), in young fruit is important for establishing the persimmon-specific mechanism that determines fruit shape diversity.

Keywords: Class 1 KNOX • Co-expression network analysis • *Diospyros* • Fruit shape • Image analysis • mRNA-seq.

Accession numbers: The sequence data generated in this study have been deposited in the DDBJ database (PRJDB8003).

Introduction

Plant shapes have long been a subject of interest. To date, studies on plant morphogenesis have been mainly focused on leaves and flowers, presumably because the sizes and shapes of

these plant parts are more complex than those of other organs such as stems and roots (Tsukaya 2003). Transitions in the shapes of plant organs have often been used as indices reflecting developmental evolution. The organ shapes of diverse plant species, from basic land plants to higher plants (angiosperms), have been extensively studied, including analyses of gene regulatory pathways or gene evolution (Nishiyama et al. 2003, Sakakibara et al. 2008). In plants from an early-diverging lineage, such as *Physcomitrella patens*, transcription factors that regulate water supply and the movement of flagellated sperm cells were identified (Koshimizu et al. 2018). Regarding angiosperms, mainly *Arabidopsis thaliana* and *Solanum* spp., some studies have unveiled molecular pathways affecting leaf shapes from the viewpoint of evolutionary developmental biology (Ichihashi et al. 2014, Vlad et al. 2014). Traditionally, distinct genes determining organ shapes have been identified (Hareven et al. 1996, Tsukaya 2006), whereas recent studies have focused on the conservation and diversity of genes and gene networks conferring differences among closely related species (Hay and Tsiantis 2006, Kimura et al. 2008, Gupta and Tsiantis 2018). For example, many plant species related to *A. thaliana* express *shoot meristemless* (*STM*) in the leaf primordia and have morphologically complex organs (Piazza et al. 2010). In tomato, the *blade-on-petiole* (*BOP*) transcription factor controls the core gene network by altering the concentration of the protein encoded by the *knotted-like* *HOMEBOX* gene *LeT6* (tomato ortholog of *STM*) (Ichihashi et al. 2014). Phytohormones, such as cytokinin and auxin, function cooperatively with these transcriptional control elements to also affect leaf shapes (Shani et al. 2010, Bilsborough et al. 2011).

Evolutionary developmental research involves a comparative study on development from an evolutionary viewpoint, and has been conducted on leaves as well as flowers (Fernández-Mazuecos and Glover 2017). By contrast, studies on fruit evolutionary development have been limited to only a few plant species such as tomato (van der Knaap et al. 2014, van der Knaap and Østergaard 2018), although horticultural crop species, such as grape, tomato, melon, cucumber or persimmon, show wide variations in their fruit shapes (Paran and van der Knaap 2007, Monforte et al. 2014, Shimomura et al. 2016, Maeda et al. 2018). For crop production, the size and shape of edible parts are commercially important traits.

Therefore, crop domestication has not only affected fruit weight, it has also often increased the diversity of fruit shapes within the same species, including tomato, melon and pepper (Paran and van der Knaap 2007, Rodríguez et al. 2011, Monforte et al. 2014). Tomato represents a model plant species for fleshy fruits, and several genes, such as *locule numbers* (LC), *fasciated* (FAS), *SUN* and *OVATE*, are considered to control fruit shapes (Tanksley 2004). The *OVATE* gene is thought to be a determinant of fruit shape also in pepper (Tsaballa et al. 2011). In melon, several QTLs involving fruit morphology have been investigated, although genes responsible for their fruit shape variations have not been identified (Monforte et al. 2014). To date, only a few of the genes responsible for fundamental changes in fruit shapes of tomato have been well characterized, often during crop domestication. However, the relationships among the determinant genes or their potential cascades have not been determined. To assess these relationships, a mutual gene networking approach involving multiple varieties exhibiting diverse fruit shapes is necessary.

Oriental persimmon (*Diospyros kaki* Thunb.) is a major fruit tree crop mainly cultivated in East Asia. Among the fruit-bearing *Diospyros* species, only oriental persimmon cultivars, including multiple landraces, show quantitatively continuous fruit shape diversity (Fig. 1a). The differences include several characteristics specific to this species, including the position of the fruit apex or grooves along the fruit side, which distinguish oriental persimmon from the fruits of other *Diospyros* species and other fruit plant species. However, their shapes have merely been empirically categorized mainly based on the ratio of length to diameter (L/D) as well as the characteristics of specific fruit parts (Fruit Tree Experiment Station of Hiroshima Prefecture 1979).

Recent developments in computer technology have contributed to the increased application of digital image processing for morphological analyses. For example, studies have been conducted in which elliptic Fourier descriptors (EFDs) and a principal component analysis (PCA) have been combined (Kuhl and Giardina 1982, Rohlf and Archie 1984). Quantitative analyses on shapes based on a combination of EFDs and a PCA using SHAPE program (Iwata and Ukai 2002) were applied not only for studies simply evaluating fruit shapes (Shimomura et al. 2016) but also for comparative studies on evolutionary development (Chitwood et al. 2016). Regarding persimmon, a quantitative characterization of developing fruit using SHAPE resulted in the determination of some principal components (PCs) for fruit shape among several varieties (Maeda et al. 2018). This implies that there are common gene networks integrating fruit shape components in persimmon varieties. These gene networks may be accessible by assembling transcriptomic data and the quantitative scores of the shape components. In this study, we aimed to identify gene networks by analyzing co-expression patterns, which have often been applied to integrate large transcriptional data set (Li et al. 2015, Liseron-Monfils and Ware 2015). In the ‘guide-gene approach’, which represents an effective strategy for a co-expression analysis (Serin et al. 2016), key components of a specific pathway are identified (Itkin et al. 2013, Serin et al. 2016). We used genes reportedly involved in *A.*

thaliana or tomato as the guide gene (or ‘bait gene’) for the analysis of co-expression networks. The objective of this study was to identify the gene networks underlying the fruit shape diversity of persimmon based on transcriptomic data and the quantitative scores of fruit shapes. Our results suggest a lineage-specific evolution of gene networks mediating the determination of fruit shape in persimmon, with common pathways in other plant species at some points as well as some pathways specific to persimmon. These insights may help to clarify the molecular mechanisms underlying persimmon fruit shapes and the diverse evolution of fruit shape determination in plants.

Results

Characterization of the fruit shape diversity in persimmon

Fruit shapes of 50 individuals of YTF1 population, which was F₁ progeny derived from *D. kaki* cultivars, ‘Yamatogoshō’ and ‘Taishū’ (Fig. 1b), 48 *D. kaki* cultivars and one accession each from two wild *Diospyros* species, *Diospyros lotus* and *Diospyros rhombifolia* (Supplementary Table S1), were assessed using SHAPE program, in the two developmental stages, stage 2 and 4, which are important for fruit shape determination, and in the maturing stage, stage 10 (Maeda et al. 2018, Supplementary Fig. S1). The contours for the final fruit shapes in stage 10, which were reconstructed based on the standardized EFD coefficients, revealed significantly different PCs (Fig. 1c, d for the YTF1 population and cultivars, respectively). Although fruit shapes were quantified in both longitudinal and transverse sections (Supplementary Fig. S2), in this study, we focused only on the diversity in the longitudinal section, which was determined synchronously across multiple persimmon cultivars (Maeda et al. 2018), and is regarded as a major feature of the diversity of persimmon fruit shapes (Fruit Tree Experiment Station of Hiroshima Prefecture 1979, Maeda et al. 2018). In the YTF1 population, a visual interpretation of these reconstructed contours implied that PC1 and PC2 for the longitudinal section were indicators of the L/D ratio with peripheral variations (Supplementary Fig. S3). By contrast, PC3 did not reflect prominent persimmon fruit shape features. In addition, PC2 was most consistent with the empirical shape evaluation based on some categories (Fruit Tree Experiment Station of Hiroshima Prefecture 1979, Supplementary Fig. S3), and could be used to appropriately evaluate the elongated shapes, such as the shape of YTF1-84 fruits in the YTF1 population (see Fig. 1b; Supplementary Fig. S3). Among the 50 cultivars/accessions, PC1 for the longitudinal direction could explain 96.2% of the whole fruit shape variance (in the longitudinal direction), which was a good indicator of not only the L/D ratio but also co-existing peripheral variations, such as dent on calyx (downward of fruit) side (Fig. 1d). Across the YTF1 population and the cultivars, each PC1 to PC3 and PC1 were significant PCs, respectively (contribution rate > 2.5%, or 1/40 contribution in SHAPE software, see Materials and Methods section).

Histological observation in representative individuals showing distinct fruit shape with fundamental differences

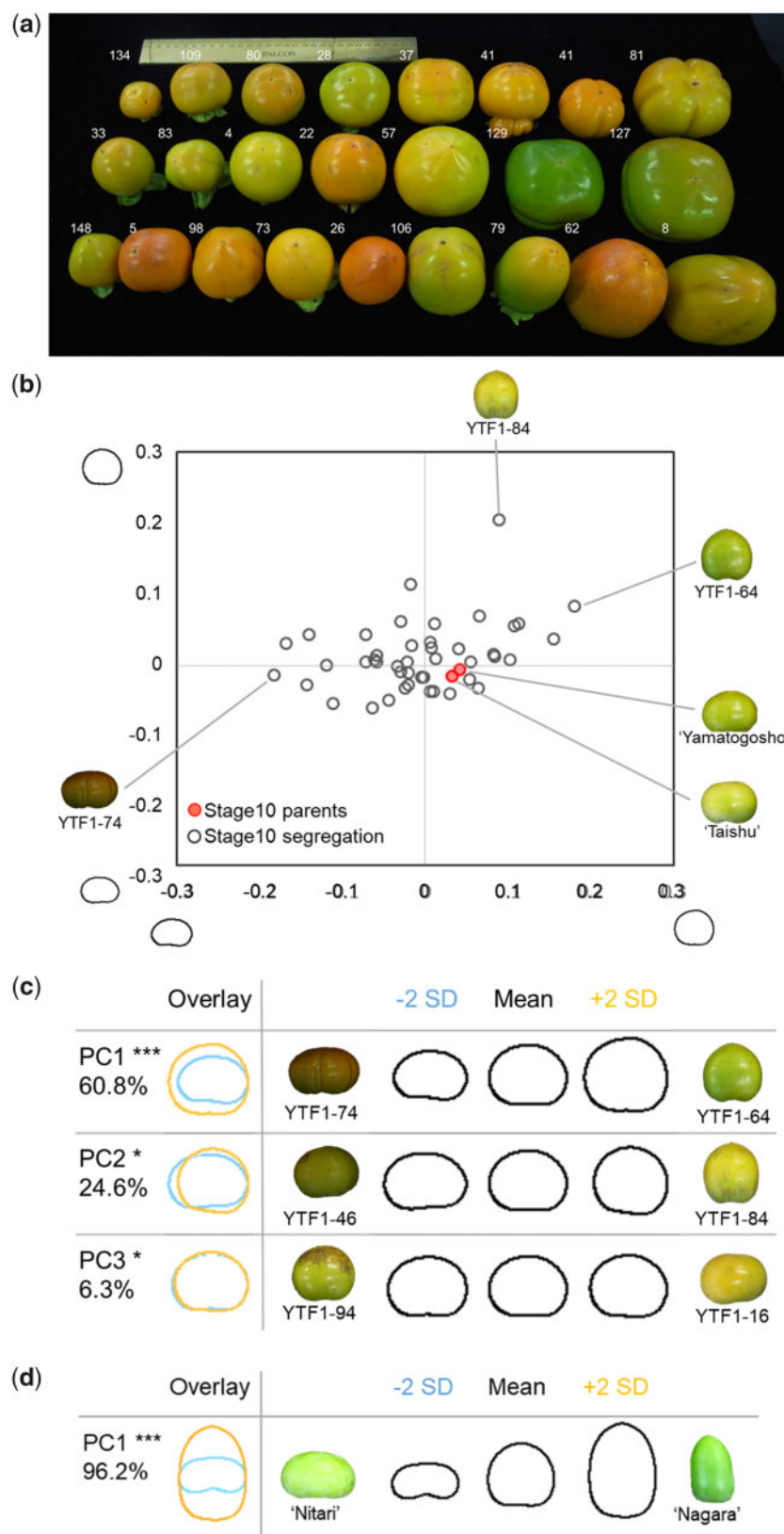


Fig. 1 Two-dimensional quantitative evaluation of the diversity in persimmon fruit shapes. (a) Diversity in fruit shapes among representative persimmon (*D. kaki*) cultivars. Annotated numbers refer to the cultivar numbers listed in Supplementary Table S1. (b) Distribution of the first two PCs (PC1 and PC2) of the fruit shapes in the YTF1 population, as determined by SHAPE. The proportion of variance explained by each PC is provided in parentheses along each axis. (c, d) Fruit shape variations in a longitudinal fruit section of the YTF1 population (c) and cultivars (d) analyzed in this study, visualized by contour images based on the significantly influential components. Each shape was reconstructed from the EFD coefficients, which were calculated with the score for a PC equal to zero and the mean \pm 2.

in PCs, indicated no difference in cell sizes/shapes (Supplementary Fig. S4), suggesting that differences in cell numbers or cell proliferation rates/directions are the morphological factors to determine fruit shape in persimmon. The PC2 in the YTF1 population showed a significantly negative correlation to fruit weight ($r = -0.31$, $P < 0.05$), which indicated that lower PC1, or flatter fruit, tended to be heavier (Supplementary Fig. S5). Considering together, for the PC1, it was suggested that cell proliferations in horizontal/transverse directions were differentiation in fruit shape in persimmon.

Transcriptome profiles and identification of the fruit shape-correlated genes

To identify the genes correlated to the quantitative indexes of the fruit shapes, transcriptomic data was obtained from 50 YTF1 individuals and their parents, and from 50 cultivars/accessions, for stages 2 and 4. The gene expression data for stages 2 and 4 (RPKM > 1) of the YTF1 population (Fig. 2a) and the cultivars (Fig. 2b) were clustered according to a PCA. For the YTF1 population, PC1 and PC2 represented 38.1% and 10.4% of the variance of gene expression level, respectively (Fig. 2a). Significant differences in PC1 and PC2 were detected between the two stages ($P = 5.4e^{-15}$ and $P = 0.0004$ for PC1 and PC2, respectively). For the cultivars, PC1 and PC2 represented 23.2% and 13.0% of the whole expression variances, respectively (Fig. 2b). The PCA results for the YTF1 population and cultivars suggested that changes in dynamic gene expression occurred between developmental stages, while the expression patterns among individuals exhibiting substantial fruit shape diversity were not considerably different. These results implied that the developmental stages were not substantially different across individuals with various fruit shapes, and specific regulatory pathways are likely to be responsible for fruit shape differentiation. Meanwhile, the PCA for the *D. kaki* cultivars with two wild relatives (*D. lotus* and *D. rhombifolia*) revealed clear differences in the gene expression data for PC4 (Supplementary Fig. S6). A gene ontology (GO) analysis of the top 5% of genes with the highest loading factors in the PC4 in the transcriptomic data indicated that the GO terms related to polarity, cation-transporting ATPase activity, and galactose catabolic process were significantly enriched (Supplementary Fig. S6).

Pearson's single correlation test revealed some candidate genes with expression patterns in the YTF1 population or in the cultivars that were significantly correlated with the quantitative values of the main fruit shape index, PC2, in SHAPE ($r > 0.37$, $P < 0.01$; Fig. 2c–f; Supplementary Table S3). Genes in the YTF1 population and in the cultivars were significantly correlated with fruit shape. Some of the candidates positively correlated with persimmon fruit shapes were orthologs of the representative genes involved in leaf shape determination, such as *asymmetric leaves 2 like 1* (ASL1), *bel1-like homeodomain 1* (BLH1) and *knotted-like homeobox of Arabidopsis thaliana 1* (KNAT1), as well as phytohormone-related genes, such as *auxin response factor 19* (ARF19) and *jasmonate zim-domain* (JAZ3).

In 50 cultivars, we exploited the genes involved in representative binary traits, such as warts and groves (see Fig. 1a), with

differentially expressed gene (DEG) analysis using DESeq2 (Love et al. 2014, see Materials and Methods section). We could detect 175 and 204 DEGs specifically biased to each trait, respectively (false discovery rate; FDR < 0.1), in stage 2 (Supplementary Table S4). Representative DEGs, potentially related to the traits, were xylem serine peptidase and xyloglucan endotransglucosylase /hydrolase, which were involved in cell wall formation/loosening and the resultant cell expansion/proliferation (Cosgrove 2005), although further assessment of the DEGs would be difficult due to the lack of quantitative evaluations of these traits, and also due to small numbers of cultivars with warts or groves ($N = 3$ and 11, respectively).

We also analyzed transcriptomic data from developing fruits of 11 cultivars/accessions treated with synthetic cytokinin (CPPU) at full bloom, which also substantially regulated fruit shape (Supplementary Fig. S7, Table S1; Itai et al. 1995). It is empirically suggested that CPPU treatment has no effect on seed size and numbers, and that there is no visible tendency in fruit shapes relating to the presence or absence of seeds in fruitlet. With paired EdgeR (Robinson et al. 2010) analysis in comparison to the controls, we detected 1,490 and 1,216 DEGs specifically biased to CPPU-treated samples in stage 2 and 4, respectively (FDR < 0.1), where some are potentially related to dynamic cytokinin metabolism or signaling (Supplementary Fig. S7). These observations suggested that the cytokinin signaling genes during the full flowering stage can substantially contribute to the fruit shapes.

Gene networks associated with fruit shapes in the YTF1 population

We attempted to construct gene networks reflecting the relationships between gene expression patterns and fruit shapes based on the weighted gene co-expression network analysis (WGCNA) method (Zhang and Horvath 2005), in the YTF1 population (see Materials and Methods section). To assess the gene expression patterns among individuals during the early fruit development stage, we used the transcriptomic data for stage 2 samples (total number of samples = 50; Supplementary Table S1) to construct co-expression networks. In addition to the quantitative values of the fruit shape indices, PC1 to 3, we used *OVATE* and *SUN* orthologs, which are involved in major fruit shape changes in tomato, as the 'bait' genes, to assess the correlations with the expression patterns of eigengenes. A phylogenetic analysis revealed that there are two and one orthologs of *OVATE* and *SUN* in the persimmon genome, respectively (Supplementary Fig. S8). In the YTF1 population, one of the two *OVATE*-like genes was not substantially expressed in developing fruit (average RPKM < 1). Thus, the other *OVATE*-like and *SUN*-like orthologs that were significantly expressed (RPKM > 1) were used as the 'bait' genes.

A total of 22,492 genes (RPKM > 1) were classified into 26 modules (Fig. 3a). Amongst the three PCs in fruit shape index, the PC2 showed the highest correlation ($r = 0.6$) to an 'expression module', which is a group of genes clustered into same expression pattern), M7 ($P = 1e^{-5}$; Fig. 3c). On the other hand, the other two PCs in fruit shape index, PC1 and PC3, showed

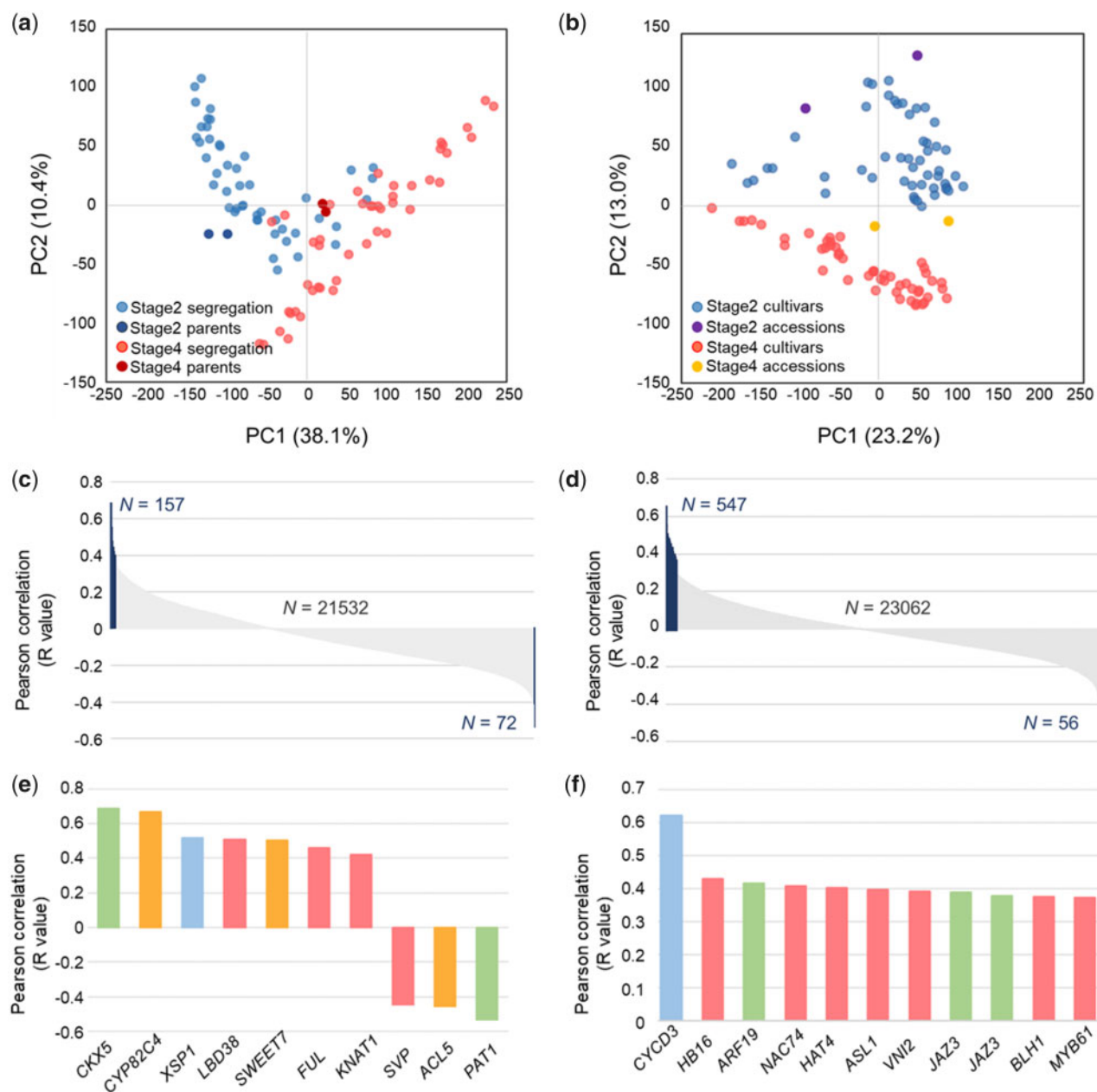


Fig. 2 Transcriptome profiles and identification of the fruit shape-correlated genes. (a, b) Characterization of the expression dynamics in the fruits of the YTF1 population (a) and the analyzed cultivars (b) by a PCA of gene expression during stages 2 and 4. Significant differences in PC1 ($P = 5.4 \times 10^{-15}$) and PC2 ($P = 0.0004$) were detected between the two developmental stages. (c, d) Distribution of the Pearson correlation coefficients between gene expression patterns and fruit shapes in the YTF1 population (gene nos = 21,761) (c) and the cultivars (gene nos = 23,665) (d) in stage 2. For the fruit shape scores, PC2 for the YTF1 population and PC1 for the cultivars, as determined by SHAPE (Fig. 1c, d), were used for the correlation analysis. Genes significantly correlated with fruit shapes ($P < 0.01$) are highlighted in blue (listed in Supplementary Table S3). (e, f) Representative genes significantly correlated with fruit shapes in the YTF1 population (e) and the cultivars (f). Putatively phytohormone- or metabolite-related as well as transcription factor or cell cycle/proliferation genes are presented in green, orange, pink or blue bars, respectively.

relatively weak correlations ($r < 0.43$) to the expression modules (Fig. 3c). Thus, we here focused on the M7 module. This module also showed significant correlation to *OVATE family protein (OFP)-like 1* ($r = 0.51$; $P = 3 \times 10^{-4}$; Fig. 3c), although *OFP-like 1* belonged to the M17 module. The M7 module included 49 genes (Fig. 3b, d; Supplementary Table S5). This module included three transcription factors, *KNAT1*, *anthocyaninless2* (*ANL2*) and *NAC-regulated seed morphology1* (*NARS1/NAC2*),

and one phytohormone-related gene, *Cytokinin oxidase/dehydrogenase* (*CKX5*), which is related to cytokinin signaling (Rupp et al. 1999, Eviatar-Ribak et al. 2013) (Fig. 3e). This was consistent with the fact that *ARR17* and *ARR11* gene expression patterns were changed by treatment with synthetic cytokinin during the flowering stages, described earlier. Cell wall-related genes such as *xylem serine peptidase 1* (*XSP1*) and *xyloglucan endotransglucosylase/hydrolase 10* (*XTH10*), which are

reportedly involved in the construction of the secondary cell wall were also included in this module. Among the transcription factors listed above, the expression level of *KNAT1* was substantially correlated with the fruit shape, and in individuals with high expression levels, the fruit shape tended to be elongated (Fig. 3f). The fact that *KNAT1* is a key gene regulating cell proliferation and the resultant leaf shape in *A. thaliana* supports our hypothesis that *KNAT1* might be important for determining persimmon fruit shapes. Furthermore, our histological analyses, described earlier (Supplementary Figs. S4, S5), suggesting that differentiation in cell proliferation is a key in fruit shape diversity, would be consistent with this hypothesis.

Putative orthologs of *SUN*, which is a determinant of tomato fruit shape (van der Knaap et al. 2004), were correlated to the M23 and M22 modules. The M23 module included 2,688 genes, and was divided into seven modules (Supplementary Fig. S9). This module included *SUN-like* enriched transcription factors as well as the *SUN-like IQD11* gene. Although this module was not significantly related to fruit shape, IQD may have some influence on persimmon fruit shapes.

Gene networks associated with fruit shape diversity among cultivars

A WGCNA involving all 50 cultivars in stages 2 and 4, including the CPPU-treated samples (total number of samples = 124; Supplementary Table S1), classified 29,236 genes into 35 modules (Fig. 4a, b). In this study, we assessed the correlation between eigengenes and fruit shape indices (PC1 during the early fruit development stages), the CPPU treatment as a binary phenotype, and the expression of *KNAT1* and *CKX5*, which were considered as key genes in the YTF1 population (Fig. 3d–g), and *CYCD3*, which were highly correlated genes in correlation test (Fig. 2f), as ‘bait’ genes (Fig. 4c).

Nine modules (M17, M19, M20, M22, M24, M25, M26, M27 and M34) showed significantly positive correlation to the fruit shape index, PC1 ($P < 0.05$) (Fig. 4c). Although the M26 module was the most highly correlated with the PC1 amongst the nine modules ($r = 0.52$, $P = 2e^{-8}$; Supplementary Fig. S10), this module showed no correlation to the expression pattern in CPPU treatment. The M26 module contained downstream genes involved in fruit shape control such as *CYCD3* and thus this module was considered to have the highest correlation value with PC1. On the other hand, only M19 and M20 modules showed significant correlations to all five traits/expressions ($r > 0.29$, $P < 0.01$) (Fig. 4c). Here, we focused on the M19 module including a ‘bait’ gene, *KNAT1* ($r = 0.70$, $P = 2e^{-16}$), which was described as a candidate core gene for determining fruit shapes in the YTF1 population (Fig. 3d; Supplementary Table S5). This module was also significantly correlated with the CPPU treatment ($r = 0.32$; $P = 9e^{-4}$). The M19 module included four class 1 *KNOX* genes, five other transcription factors (subsequently described in greater detail), several plant hormone-related genes, especially those in cytokinin, auxin and gibberellin metabolizing pathways, and cell wall-related genes, such as *XSP1* and *XTH32* (Fig. 4d, e). Four GO terms, ‘oxidation-reduction process’ (GO:0055114), ‘adenine phosphoribosyltransferase activity’ (GO:0003999), ‘adenine

salvage’ (GO:0006168) and ‘gibberellin 20-oxidase activity’ (GO:0045544), were significantly enriched in the M19 module (FDR < 0.1). These terms clearly represent cytokinin and gibberellin signaling/metabolic activities. With the exception of the four class I *KNOX* genes, the transcription factors in the M19 module were the orthologs of one of the five *YABBY* classes (Bowman and Smyth 1999, Yamada et al. 2004, Lee et al. 2005), namely, *inner no outer* (*INO*), which regulates ovule development (Villanueva et al. 1999); *NAC2*, which regulates seed morphogenesis (Kunieda et al. 2008); *OFPP9*, which is one of the *OFPP* genes associated with the *KNOX* and *BELL* class homeodomain genes regulating ovule and/or fruit development (Wang et al. 2016). *OFPP* proteins interact with microtubules, which are proposed to lead to changes in cell division patterns, resulting in the different growth forms (Wu et al. 2018). *Homeobox 21* (*HB21*), which controls the ABA signal-regulated branching pattern (González-Grandío et al. 2017); and a *MADS* box *seedstick/ago-mous-like 11* (*STK/AGL11*), which is responsible for seed and ovule development (Favaro et al. 2003, Pinyopich et al. 2003, Singh et al. 2013). Importantly, *STK/AGL11* was the central gene that linked most of the transcription factors to the plant hormone-, cell wall- and metabolism-related genes (Fig. 4d; Supplemental Table S6). Although the other transcription factors were paralleled in the network, their representative functions in model plants were potentially linked (e.g. regarding ovule or carpel development), which was consistent with the fact that this module is substantially correlated with persimmon fruit shapes.

Evolution and changes in expression patterns of the class 1 KNOX family genes

The number of class 1 *KNOX* family genes tended to be higher in the order Ericales (e.g. 12 in kiwifruit, eight in blueberry and nine in persimmon) than in other species (e.g. two in grape, four in *A. thaliana* and three in tomato) (Fig. 5a). In addition, the *KNAT1* clade of the phylogenetic tree comprised one gene from each of *A. thaliana*, grape, tomato, blueberry and persimmon as well as two genes from kiwifruit (Fig. 5a, b). During the evolution of *KNAT1* genes, there was no relaxation of purifying selection ($dN/dS < 0.2$), indicating a lack of changes to protein functions (Fig. 5b). Furthermore, we did not detect significant site-branch-specific positive selection ($dN/dS \gg 1$) with PAML (Yang 1997) in any evolutionary branch (posterior probability < 0.81; Bayes empirical Bayes test). In persimmon, the *KNAT1* ortholog was expressed in young developing fruits (average RPKM = 8.0), with considerable diversity among cultivars (RPKM = 0.4–30.3). Published transcriptomic data indicated that in *A. thaliana*, *KNAT1* is not substantially expressed in young ovaries (RPKM = 1.94) (Klepikova et al. 2016). In tomato and kiwifruit, the expression levels of *KNAT1* orthologs in young fruits, corresponding to young persimmon fruits, as examined in this study, were very low (RPKM < 1).

Discussion

Molecular pathways regulating fruit shape diversity have been mainly characterized in tomato, where *OVATE* and *SUN* are the

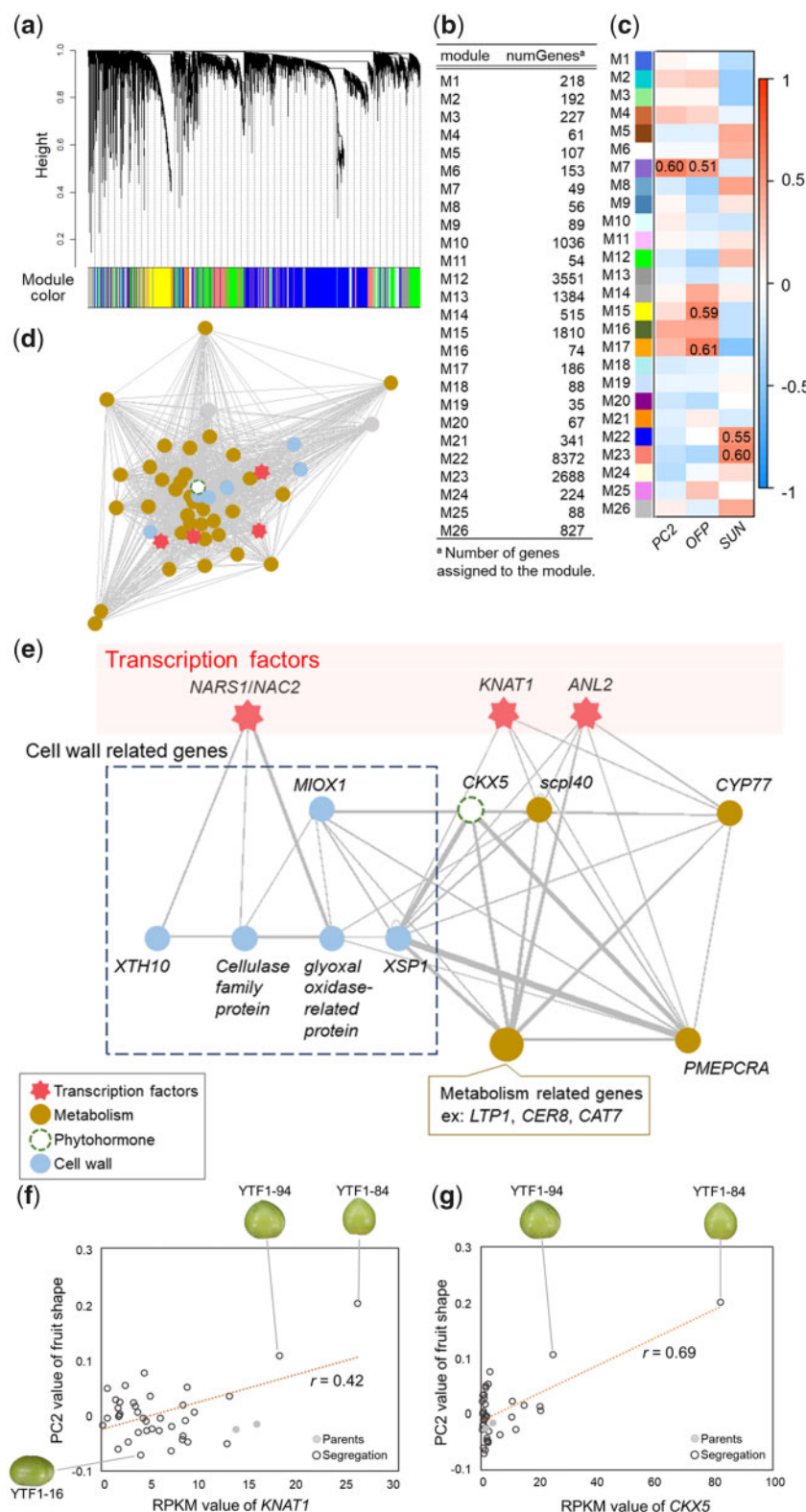


Fig. 3 Gene networks associated with fruit shapes in the YTF1 population. (a) Clustering of the gene expression patterns in stage 2, as determined with the WGCNA package, resulting in 26 modules, M1–M26. (b) The number of genes in the 26 co-expression modules. (c) Heat map visualization of the correlations of the eigengenes of 26 modules with fruit shape indexes (PC1 to PC3) and the expression of two 'bait' genes, *OVATE-like* and *SUN-like*. The M7 module is the most highly correlated with PC2. (d, e) Visualization of the M7 module network. The genes were clustered based on their putative functions and are presented in different nodes. Transcription factors are indicated by pink stars (see Supplementary Table S5). (f, g) Correlation between *KNAT1* (f) or *CKX5* (g) expression patterns and the fruit shape index, PC2. Both genes were significantly correlated ($r = 0.42$ and 0.69 , respectively) with fruit shapes.

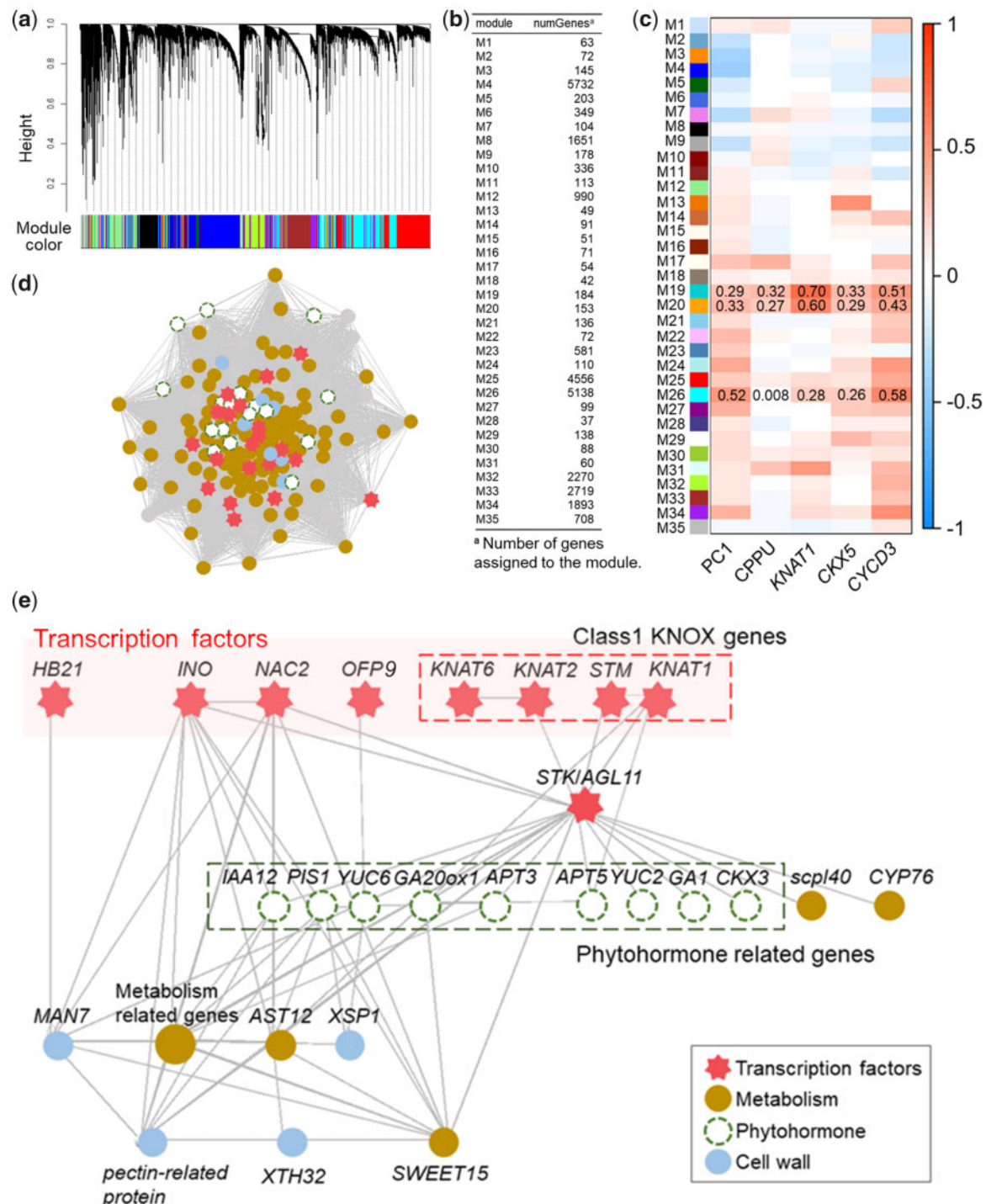


Fig. 4 Gene networks associated with fruit shapes in cultivars. (a) Clustering of the gene expression patterns in stage 2, as determined with the WGCNA package, resulting in 35 modules, M1–M35. (b) The number of genes in the 35 co-expression modules. (c) Heat map visualization of the correlations of the module eigengenes with fruit shape index (PC1), CPPU treatment and the expression of three 'bait' genes, *KNAT1*, *CKX5* and *CYCD3*. The M19 module was significantly correlated with all of the objective traits/genes, and was the most highly correlated with *KNAT1* expression patterns. (d, e) Visualization of the M19 module network. The genes were clustered based on their putative functions and are presented in different nodes. Transcription factors are indicated by pink stars (see Supplementary Table S7).

main players for the production of elongated fruits (Tanksley 2004). On the other hand, *LC* and *FAS* can play major roles in the determination of carpel/locule numbers and production of flat fruits (Rodríguez et al. 2011). These pathways developed via

lineage-specific domestication, implying that the genetic determinants of the diversity among cultivars might be specific to tomato. Consistent with this prediction, our results in persimmon revealed gene pathways likely to be responsible for fruit



shape diversity that differed from those in tomato (Fig. 6). Substantial diversity in fruit shape among persimmon cultivars are not dependent on the carpel/locule numbers, but on cell proliferations in early developmental stages (Maeda et al. 2018). Thus, the mechanism driving fruit shape diversity would be distinct from the tomato *LC* or *FAS* genes. Also for the other two genes, *OVATE* and *SUN*, which would affect cell proliferation, our results suggested no direct involvement of their expression pattern in fruit shape determination among the persimmon cultivars, although a few *OFP* family gene, to which tomato *OVATE* is nested, might have certain significance, as discussed later. In persimmon, one of the class I *KNOX* genes, *KNAT1*, and cytokinin signals appear to play important roles in a segregated population. An analysis of gene networks in multiple cultivars implied that *KNAT1* and the related transcription factors significantly correlated with fruit shapes were integrated by the variable expression of a core gene, *STK*, which is specific to persimmon fruit shape determination. In *A. thaliana*, the expression balance between *STK* and the *BEL1* homeobox

The contribution of *KNAT1* to the determination of fruit shape may also be specific to persimmon. Although the expression of *KNAT1* in the *A. thaliana* shoot apical meristem reportedly helps regulate leaf shape (Hav and Tsiantis 2010), there has

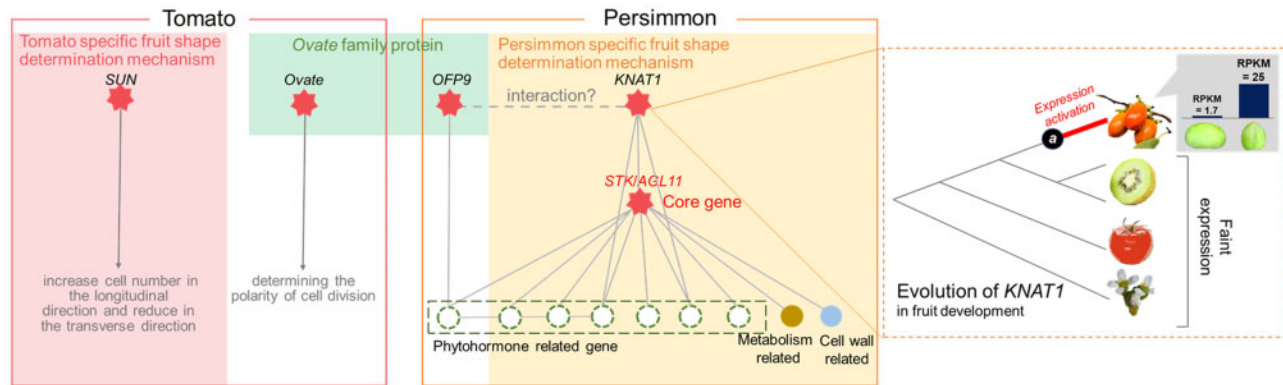


Fig. 6 Model for the lineage-specific evolution of fruit shape determination based on a comparison between tomato and persimmon. In persimmon, the lineage-specific acquisition of *KNAT1* expression, and its variation in young fruit, may be important for the considerable fruit shape diversity. Furthermore, *STK/AGL11* may function as a central gene that mediates the expression of transcription factors, including *KNAT1*, to regulate many phytohormone-related and cell proliferation-related genes, which represents a novel mechanism. However, *OFP*, which helps regulate fruit shapes in tomato (Tanksley 2004), was also included in the same network that determines persimmon fruit shapes. This *OFP*-dependent pathway that regulates fruit shapes may be common in the two lineages.

been no report describing its variations in expression of fruits and determination of fruit shape. The persimmon-specific *KNAT1*-mediated mechanism underlying the regulation of fruit shape may be explained from an evolutionary perspective by two points. The evolutionary tree of representative angiosperm *KNAT1* genes suggested that *KNAT1* is mostly monophyletic, with no significant signs of protein functional changes (Fig. 5b). In addition, the *KNAT1* expression level in young fruit was substantially higher in persimmon than in other species (Fig. 6). These observations suggest that the relatively high *KNAT1* expression levels in persimmon fruit and the varying expression levels among cultivars/accessions might represent the persimmon-specific mechanism that determines fruit shape diversity. Although the average *KNAT1* expression level in the YTF1 population was similar to that in the parental cultivars, a few segregants with considerably higher *KNAT1* expression levels produced fruit shapes that were consistently different than those observed for the accessions/cultivars assessed in this study (Fig. 3f), implying persimmon fruit shape is continuously evolving. This lineage-specific high expression of *KNAT1* in young fruit was also observed in wild relatives, such as the *D. rhombifolia* or *D. lotus* included in this study (data not shown), but there has been no reported diversity in fruit shapes. Thus, the considerable diversity in fruit shapes specific to oriental persimmon is likely due to the variations in *KNAT1* expression, which may affect the genes expressed in an oriental persimmon-specific manner, as indicated in our transcriptome analysis (Supplementary Fig. S6).

In contrast to the putatively lineage-specific pathways, *OFP* expression and its associated effects on fruit shapes may be common in tomato and persimmon (Fig. 6). In persimmon, variations in the expression of *OFP* as well as other transcription factors in the same expression module may regulate fruit shape, whereas distinct mutations in *OVATE* have significant consequences in fruit shapes in tomato (Liu et al. 2002). Although this signifies that the causal mutations that drive the diversity of fruit shapes are distinct in these two species, the gene

commonality suggests that the *OFP* gene potentially mediates the general development of fruit shapes. This possibility is supported by the fact that *OVATE*-like proteins regulate fruit shapes in pepper and melon (Tsaballa et al. 2011, Monforte et al. 2014). Collectively, our results shed light on lineage-specific rate-limiting genetic factors regulating fruit shape determination, whereas the components of the relevant pathways may be somewhat common among different plant species.

Materials and Methods

Plant materials and treatments

Oriental persimmon (*D. kaki*) cultivars, 'Yamatogoshō' and 'Taishū', and 50 of their F_1 progeny (YTF1 population), which exhibited considerable fruit shape differences, were included in this study (Fig. 1b). Trees were maintained in the experimental orchard of the Grape and Persimmon Research Station, NARO Institute of Fruit Tree and Tea Science, Japan. To assess fruit shape developmental patterns, 1–3 fruits were sampled from each individual tree on June 13, July 5 and October 3, 2017, which corresponded to stages 2, 4 and 10 (mature) of the fruit shape development period (Maeda et al. 2018; Supplementary Fig. S1). The stages 2 to 4 are supposed to be important for the determination of fruit shape in persimmon cultivars (Maeda et al. 2018). Forty-eight *D. kaki* cultivars, and one accession each from two wild *Diospyros* species, *D. lotus* and *D. rhombifolia*, which were maintained in the orchard of Kyoto University, Kyoto, Japan, were also analyzed in this study (Supplementary Table S1). Three fruits were harvested on June 1, June 27 and September 20, 2016, corresponding to stages 2, 4 and 10. A synthetic cytokinin, CPPU, reportedly influences persimmon fruit shape (Itai et al. 1995). Thus, 10 cultivars of *D. kaki*, *D. lotus* and *D. rhombifolia* at full bloom were treated with 100 ppm CPPU (Fulmet; Kyowa Hakko Bio, Tokyo, Japan) prepared in a solution of 0.05% Tween-20 as a surfactant.

To assess the fruit development from histological analysis, frozen sections were prepared with the fruits of four individuals from the YTF 1 population, using Cryostat (Leica CM1869, Leica, Osaka, Japan), according to the Kawamoto method (Kawamoto 2003). Twenty-micrometer sections were stained with toluidine blue.

Quantification of fruit shape

Fruit shapes were evaluated quantitatively based on EFDs and a PCA using SHAPE program (Iwata and Ukai 2002) as previously described (Maeda et al.

2018). The PCA was completed using the EFDs resulting from the first 20 harmonics of Fourier coefficients. Thus, 80 (4×20) standardized EFDs were calculated per fruit. The 80 coefficients were classified into two groups related to symmetrical and asymmetrical variations (Iwata et al. 1998, Yoshioka et al. 2004); the contour reconstructed only by the former coefficients is symmetrical with respect to the central fruit axis, whereas the latter coefficients represent the asymmetrical aspects of fruit shape. Because the asymmetrical elements of fruits are affected by non-heritable factors, we used 40 coefficients related to symmetrical elements in the following analysis. The sampled fruits were weighed, frozen in liquid nitrogen and stored at -80°C for a subsequent RNA extraction.

Transcriptome library construction and Illumina sequencing

Total RNA was extracted from 222 fruit samples (50 YTF1 individuals and their parents, 50 cultivars/accessions and 11 CPPU-treated samples for stages 2 and 4) according to the hot borate method (Wan and Wilkins 1994), after which mRNA was extracted using the DynabeadsTM mRNA Purification Kit (Ambion, Austin, TX, USA). Illumina sequencing libraries were prepared from 2 μg total RNA as previously described (Akagi et al. 2014, Akagi et al. 2016) and then purified with AMPureXP (Beckman Coulter Life Science) to remove fragments shorter than 300 bp. All 222 libraries were multiplexed and analyzed with the Illumina HiSeq 4000 system at the QB3 Genomic Sequencing Laboratory of UC Berkeley (<http://qb3.berkeley.edu/gsl/>) to generate 50-bp single-end reads. All Illumina sequencing experiments were conducted at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, and the resulting raw sequencing reads were processed with custom Python scripts developed in the Comai laboratory and available online (http://comailab.genomecenter.ucdavis.edu/index.php/Barcoded_data_preparation_tools) as previously described (Akagi et al. 2014). In brief, reads were split based on the index information and trimmed for quality (average Phred sequence quality > 20) and to eliminate adapter sequences. Next, a read length cutoff of 35 bp was applied. All samples used to generate Illumina sequences are listed in Supplementary Table S1. The sequence data generated in this study have been deposited in the DDBJ database (PRJDB8003).

Transcriptomic data mining, correlation analysis and extraction of DEGs

The mRNA-seq Illumina reads were aligned to the reference CDS of *D. lotus* (Akagi et al. 2019; <http://persimmon.kazusa.or.jp/index.html>) using the default parameters of the Burrows–Wheeler Aligner program (version 0.7.12) (Li and Durbin 2009; <https://github.com/lh3/bwa>). The read counts per coding sequence were generated from the aligned SAM files using a custom R script (Akagi et al. 2014). The expression level of each gene was calculated as the reads per kilobase of exon per million sequenced reads (RPKM). Genes with an average RPKM < 1 for the aligned samples were excluded from the subsequent co-expression network analysis.

For characterizing the expression dynamics in datasets, a PCA was completed with `prcomp` in R (version 3.5.1) (R Core Team 2018). To assess single correlations between a gene expression pattern and quantitative values for fruit shapes, a Pearson's product moment correlation analysis was completed among individuals, using `cor.test` with the 'Pearson' argument in R, with gene expression levels during each stage as explanatory variables and PC values for mature fruit shapes as objective variables. Pearson's correlation test was also used to assess expression similarities among individuals based on gene expression patterns.

The genes correlated to the warts and groves (Fig. 1a), which were evaluated as binary traits, were assessed using DESeq2 package (Love et al. 2014) to determine DEGs across the cultivars, with the threshold value of FDR set to 0.1.

A GO enrichment analysis was completed using the default parameters of the `goseq` package (Young et al. 2010) in R. The annotations of genes fulfilling the RPKM cutoff were used as a reference for the GO analysis. A threshold for the significance of enriched GO terms was set as a FDR < 0.1 . The contigs were annotated with Blast2GO (version 4.0.2), using a Blastx search of the NCBI non-redundant protein database, with an E-value cutoff of $1e^{-5}$. Up to 20 hits were retained for each gene to assign GO terms, effective concentration (EC) numbers and possible descriptions.

Co-expression network construction

The WGCNA method (Zhang and Horvath 2005) was used to generate modules of highly correlated genes based on the RNA-seq expression data. Gene modules were identified by implementing the WGCNA package in R (Langfelder and Horvath 2008). The soft-thresholding power and `treecut` parameters used for the WGCNA analysis were 7 and 0.25, respectively. In the co-expression network, genes were represented by nodes, and the correlation values (weight) between two genes were calculated by increasing Pearson's correlation coefficient. Gene expression profiles for the modules were summarized as module eigengenes by calculating the first PC of the expression profile of genes assigned to the module. To identify the modules influencing fruit shape diversity, correlations between the module eigengene and PC values for mature fruit shapes or 'bait' gene expression patterns in the YTF1 population or in cultivars were assessed using WGCNA package.

Connections between genes in a module were first visualized using VisANT program (Hu et al. 2008). A TOM cutoff for the co-expression in the visualization was determined based on the strength of the correlation with the node degree distribution, and was set as 0.05.

Genome-wide survey of the core candidates in gene networks

To examine the evolutionary relationships among core candidates in the gene network putatively involved in determining persimmon fruit shapes, peptide sequences orthologous to those encoded by *OFP* and the class 1 *KNOX* genes in the Persimmon Genome Database (<http://persimmon.kazusa.or.jp/index.html>) were analyzed using Basic Local Alignment Search Tool algorithm (BLASTP; $e\text{-value} < 1e^{-10}$ with default parameters), with the *OVATE* domain (Liu et al. 2002, Wang et al. 2007) and class 1 *KNOX* gene (Hay and Tsiantis 2010) sequences from *A. thaliana* as the queries. The *OFP* orthologs in *A. thaliana*, tomato (*Solanum lycopersicum*) and hornwort (*P. patens*) were obtained from Phytozome (version 9.0) (<http://www.phytozome.net/>) according to Huang et al. (2013) (Supplementary Table S7). For the class 1 *KNOX* genes, orthologs in grapevine (*Vitis vinifera*) and tomato were obtained from Phytozome (<https://phytozome.jgi.doe.gov/pz/portal.html>), with the *A. thaliana* class 1 *KNOX* genes (Hay and Tsiantis 2010) as queries. Regarding species of the order Ericales, in which persimmon is nested, orthologs in kiwifruit (*Actinidia chinensis*), blueberry (*Vaccinium corymbosum*) and diploid Caucasian persimmon (*D. lotus*) were obtained from the Kiwifruit Genome Database (<http://bioinfo.bti.cornell.edu/cgi-bin/kiwi/home.cgi>), the Genome Database for *Vaccinium* (<https://www.vaccinium.org/>) and the Persimmon Genome Database described above, respectively.

To estimate the evolutionary topology, protein sequences were aligned according to the default parameters of MAFFT program (version 7) (Katoh and Standley 2013), and then manually pruned with Seaview (Gouy et al. 2010). The maximum likelihood (ML) and neighbor-joining (NJ) methods were applied using MEGA (version 10), with 1,000 bootstrap replicates, to determine phylogenetic relationships (Kumar et al. 2018). In the ML method, the Whelan and Goldman (WAG) model with invariant sites and gamma-distributed rates (three categories) was considered, with nearest-neighbor interchange (NNI) as the tree-searching heuristic. All sites including missing and gap data were used to construct a phylogenetic tree. In the NJ approach, the Poisson matrix with gamma-distributed rates (number of discrete gamma categories: 3) was used, with a pairwise deletion for missing data.

To determine evolutionary rates (dN/dS), the nucleotide sequences of the analyzed genes were subjected to an in-frame alignment using MAFFT and Pal2Nal programs (Suyama et al. 2006). The ancestral sequences were estimated using MEGA (version 6) (Tamura et al. 2013), and then pairwise dN/dS values were determined from the in-frame alignments with DnaSP (version 5) (Librado and Rozas 2009). We also detected codon-based site-specific and branch-specific positive selection using PAML (Yang 1997), with the same in-frame nucleotide alignment. The significance of the positive selection on the foreground branches was evaluated according to the likelihood ratio test, with a null hypothesis of $dN/dS = 1$. Site-specific positive selection was assessed using a Bayes empirical Bayes analysis.

Supplementary Data

Supplementary data are available at PCP online.

Funding

Some of this work was performed at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley supported by NIH S10 OD018174 Instrumentation Grant. PRESTO, the Japan Science and Technology Agency (to T.A.), and a Grant-in-Aid for Young Scientists (A) [26712005 to T.A.], Grant-in-Aid for Scientific Research (B) [18H02199 to T.A.] and Grant-in-Aid for Scientific Research on Innovative Areas [J16H06471 to T.A.] from JSPS, and a Grant-in-Aid for Challenging Research (Pioneering) [19H05539 to R.T.] from JSPS.

Acknowledgments

We thank Dr. Akihiko Sato for development and management of experimental persimmon trees. We also thank Edanz Group (www.edanzediting.com/ac) for editing a draft of this manuscript.

Disclosures

The authors have no conflicts of interest to declare.

References

- Akagi, T., Henry, I.M., Kawai, T., Comai, L. and Tao, R. (2016) Epigenetic regulation of the sex determination gene *MeGI* in polyploid persimmon. *Plant Cell* 28: 2905–2915.
- Akagi, T., Henry, I.M., Tao, R. and Comai, L. (2014) A y-chromosome-encoded small RNA acts as a sex determinant in persimmons. *Science* 346: 646–650.
- Akagi, T., Shirasawa, K., Nagasaki, H., Hirakawa, H., Tao, R., Comai, L., et al. (2019) The persimmon genome reveals clues to the evolution of a lineage-specific sex determination system in plants. *bioRxiv*. doi: 10.1101/628537.
- Bencivenga, S., Simonini, S., Benková, E. and Colombo, L. (2012) The transcription factors BEL1 and SPL are required for cytokinin and auxin signaling during ovule development in *Arabidopsis*. *Plant Cell* 24: 2886–2897.
- Bilsborough, G.D., Runions, A., Barkoulas, M., Jenkins, H.W., Hasson, A., Galinha, C., et al. (2011) Model for the regulation of *Arabidopsis thaliana* leaf margin development. *Proc. Natl. Acad. Sci. USA* 108: 3424–3429.
- Bowman, J.L. and Smyth, D.R. (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.
- Brambilla, V., Battaglia, R., Colombo, M., Masiero, S., Bencivenga, S., Kater, M.M., et al. (2007) Genetic and molecular interactions between BELL1 and MADS box factors support ovule development in *Arabidopsis*. *Plant Cell* 19: 2544–2556.
- Chitwood, D., Klein, L., O'Hanlon, R., Chacko, S., Greg, M., Kitchen, C., et al. (2016) Latent developmental and evolutionary shapes embedded within the grapevine leaf. *New Phytol.* 210: 343–355.
- Cosgrove, D.J. (2005) Growth of the plant cell wall. *Nat. Rev. Mol. Cell Biol.* 6: 850–861.
- Eviatar-Ribak, T., Shalit-Kaneh, A., Chappell-Maor, L., Amsellem, Z., Eshed, Y. and Lifschitz, E. (2013) A cytokinin-activating enzyme promotes tuber formation in tomato. *Curr. Biol.* 23: 1057–1064.
- Favaro, R., Pinyopich, A., Battaglia, R., Kooiker, M., Borghi, L., Ditta, G., et al. (2003) MADS-box protein complexes control carpel and ovule development in *Arabidopsis*. *Plant Cell* 15: 2603–2611.
- Fernández-Mazuecos, M. and Glover, B.J. (2017) The evo-devo of plant speciation. *Nat. Ecol. Evol.* 1: 110.
- Fruit Tree Experiment Station of Hiroshima Prefecture. (1979) Showa 53-Nendo Shubyo-Tokusei-Bunrui-Chosa-Hokokusho (Kaki). Fruit Tree Experiment Station of Hiroshima Prefecture, Hiroshima, Japan. In Japanese.
- González-Grandío, E., Pajoro, A., Franco-Zorrilla, J.M., Tarancón, C., Immink, R.G.H. and Cubas, P. (2017) Absciscic acid signaling is controlled by a BRANCHED1/HD-ZIP I cascade in *Arabidopsis* axillary buds. *Proc. Natl. Acad. Sci. USA* 114: E245–E254.
- Gouy, M., Guindon, S. and Gascuel, O. (2010) SeaView version 4: a multi-platform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* 27: 221–224.
- Gupta, M.D. and Tsiantis, M. (2018) Gene networks and the evolution of plant morphology. *Curr. Opin. Plant Biol.* 45: 82–87.
- Hareven, D., Gutfinger, T., Parnis, A., Eshed, Y. and Lifschitz, E. (1996) The making of a compound leaf: genetic manipulation of leaf architecture in tomato. *Cell* 84: 735–744.
- Hay, A. and Tsiantis, M. (2006) The genetic basis for differences in leaf form between *Arabidopsis thaliana* and its wild relative *Cardamine hirsute*. *Nat. Genet.* 38: 942–947.
- Hay, A. and Tsiantis, M. (2010) KNOX genes: versatile regulators of plant development and diversity. *Development* 137: 3153–3165.
- Hu, Z., Snitkin, E.S. and DeLisi, C. (2008) VisANT: an integrative framework for networks in systems biology. *Brief. Bioinform.* 9: 317–325.
- Huang, B.W., Routaboul, J.M., Liu, M.C., Deng, W., Maza, E., Mila, I., et al. (2017) Overexpression of the class D MADS-box gene *Sl-AGL11* impacts fleshy tissue differentiation and structure in tomato fruits. *J. Exp. Bot.* 68: 4869–4884.
- Huang, Z.J., Houten, J.V., Gonzalez, G., Xiao, H. and van der Knaap, E. (2013) Genome-wide identification, phylogeny and expression analysis of *SUN*, *OFD* and *YABBY* gene family in tomato. *Mol. Genet. Genomics* 288: 111–129.
- Ichihashi, Y., Aguilar-Martinez, J.A., Farhi, M., Chitwood, D.H., Kumar, R., Millon, L.V., et al. (2014) Evolutionary developmental transcriptomics reveals a gene network module regulating interspecific diversity in plant leaf shape. *Proc. Natl. Acad. Sci. USA* 111: 2616–2621.
- Itai, A., Tanabe, K., Tamura, F., Susaki, S., Yonemori, K. and Sugiura, A. (1995) Synthetic cytokinins control persimmon fruit shape, size and quality. *J. Hortic. Sci.* 70: 867–873.
- Itkin, M., Heinig, U., Tzfadia, O., Bhide, A.J., Shinde, B., Cardenas, P.D., et al. (2013) Biosynthesis of antinutritional alkaloids in solanaceous crops is mediated by clustered genes. *Science* 341: 175–179.
- Iwata, H., Niikura, S., Matsuura, S., Takano, Y. and Ukai, Y. (1998) Evaluation of variation of root shape of Japanese radish (*Raphanus sativus* L.) based on image analysis using elliptic Fourier descriptors. *Euphytica* 102: 143–149.
- Iwata, H. and Ukai, Y. (2002) SHAPE: a computer program package for quantitative evaluation of biological shapes based on elliptic Fourier description. *J. Hered.* 93: 384–385.
- Katoh, K. and Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30: 772–780.
- Kawamoto, T. (2003) Use of a new adhesive film for the preparation of multi-purpose fresh-frozen sections from hard tissues, whole-animals, insects and plants. *Arch. Histol. Cytol.* 66: 123–143.
- Kimura, S., Koenig, D., Kang, J., Yoong, F.Y. and Sinha, N. (2008) Natural variation in leaf morphology results from mutation of a novel KNOX gene. *Curr. Biol.* 18: 672–677.

- Klepikova, A.V., Kasianov, A.S., Gerasimov, E.S., Logacheva, M.D. and Penin, A.A. (2016) A high resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq profiling. *Plant J.* 88: 1058–1070.
- Koshimizu, S., Kofuji, R., Sasaki-Sekimoto, Y., Kikkawa, M., Shimojima, M., Ohta, H., et al. (2018) Physcomitrella MADS-box genes regulate water supply and sperm movement for fertilization. *Nat. Plants* 4: 36–45.
- Kuhl, F.P. and Giardina, C.R. (1982) Elliptic Fourier features of a closed contour. *Comput. Graphics Image Process.* 18: 236–258.
- Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35: 1547–1549.
- Kunieda, T., Mitsuda, N., Ohme-Takagi, M., Takeda, S., Aida, M., Tasaka, M., et al. (2008) NAC family proteins NARS1/NAC2 and NARS2/NAM in the outer integument regulate embryogenesis in *Arabidopsis*. *Plant Cell* 20: 2631–2642.
- Langfelder, P. and Horvath, S. (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 9: 559.
- Lee, J.Y., Baum, S.F., Oh, S.H., Jiang, C.Z., Chen, J.C. and Bowman, J.L. (2005) Recruitment of CRABS CLAW to promote nectary development within the eudicot clade. *Development* 132: 5021–5032.
- Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754–1760.
- Li, Y., Pearl, S.A. and Jackson, S.A. (2015) Gene networks in plant biology: approaches in reconstruction and analysis. *Trends Plant Sci.* 20: 664–675.
- Librado, P. and Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.
- Liseron-Monfils, C. and Ware, D. (2015) Revealing gene regulation and associations through biological networks. *Curr. Plant Biol.* 4: 30–39.
- Liu, J., Van Eck, J., Cong, B. and Tanksley, S.D. (2002) A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. *Proc. Natl. Acad. Sci. USA* 99: 13302–13306.
- Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15: 550.
- Maeda, H., Akagi, T. and Tao, R. (2018) Quantitative characterization of fruit shape and its differentiation pattern in diverse persimmon (*Diospyros kaki*) cultivars. *Sci. Hortic.* 228: 41–48.
- Monforte, A.J., Diaz, A., Caño-Delgado, A. and van der Knaap, E. (2014) The genetic basis of fruit morphology in horticultural crops: lessons from tomato and melon. *J. Exp. Bot.* 65: 4525–4537.
- Nishiyama, T., Fujita, T., Shin-I, T., Seki, M., Nishide, H., Uchiyama, I., et al. (2003) Comparative genomics of *Physcomitrella patens* gametophytic transcriptome and *Arabidopsis thaliana*: implication for land plant evolution. *Proc. Natl. Acad. Sci. USA* 100: 8007–8012.
- Paran, I. and 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.
- Piazza, P., Bailey, C.D., Cartolano, M., Krieger, J., Cao, J., Ossowski, S., et al. (2010) *Arabidopsis thaliana* leaf form evolved via loss of KNOX expression in leaves in association with a selective sweep. *Curr. Biol.* 20: 2223–2228.
- Pinyopich, A., Ditta, G.S., Savidge, B., Liljegren, S.J., Baumann, E., Wisman, E., et al. (2003) Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature* 424: 85–88.
- R Core Team. (2018) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/> (July 25, 2019, date last accessed).
- Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139–140.
- Rodríguez, G.R., Muños, S., Anderson, C., Sim, S.C., Michel, A., Causse, M., et al. (2011) Distribution of SUN, OVATE, LC, and FAS in the tomato germplasm and the relationship to fruit shape diversity. *Plant Physiol.* 156: 275–285.
- Rohlf, F.J. and Archie, J.W. (1984) A comparison of Fourier methods for the description of wing shape in mosquitoes (Diptera: Culicidae). *Syst. Zool.* 33: 302–317.
- Rupp, H.M., Frank, M., Werner, T., Strnad, M. and Schumlling, T. (1999) Increased steady state mRNA levels of the STM and KNAT1 homeobox genes in cytokinin overproducing *Arabidopsis thaliana* indicate a role for cytokinins in the shoot apical meristem. *Plant J.* 18: 557–563.
- Sakakibara, K., Nishiyama, T., Deguchi, H. and Hasebe, M. (2008) Class 1 KNOX genes are not involved in shoot development in the moss *Physcomitrella patens* but do function in sporophyte development. *Evol. Dev.* 10: 555–566.
- Serin, E.A.R., Nijveen, H., Hilhorst, H.W.M. and Ligterink, W. (2016) Learning from co-expression networks: possibilities and challenges. *Front. Plant Sci.* 7: 444.
- Shani, E., Ben-Gera, H., Shleizer-Burko, S., Burko, Y., Weiss, D. and Ori, N. (2010) Cytokinin regulates compound leaf development in tomato. *Plant Cell* 22: 3206–3217.
- Shimomura, K., Horie, H., Sugiyama, M., Kawazu, Y. and Yoshioka, Y. (2016) Quantitative evaluation of cucumber fruit texture and shape traits reveals extensive diversity and differentiation. *Sci. Hortic.* 199: 133–141.
- Singh, R., Low, E.-T.L., Ooi, L.C.-L., Ong-Abdullah, M., Ting, N.-C., Nagappan, J., et al. (2013) The oil palm SHELL gene controls oil yield and encodes a homologue of SEEDSTICK. *Nature* 500: 340–344.
- Suyama, M., Torrents, D. and Bork, P. (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* 34: W609–W612.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) Mega 6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30: 2725–2729.
- Tanksley, S.D. (2004) The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. *Plant Cell* 16: 181–189.
- Tsaballa, A., Pasentsis, K., Darzentas, N. and Tsiftaris, A.S. (2011) Multiple evidence for the role of an Ovate-like gene in determining fruit shape in pepper. *BMC Plant Biol.* 11: 46.
- Tsukaya, H. (2003) Organ shape and size: a lesson from studies of leaf morphogenesis. *Curr. Opin. Plant Biol.* 6: 57–62.
- Tsukaya, H. (2006) Mechanism of leaf-shape determination. *Annu. Rev. Plant Biol.* 57: 477–496.
- van der Knaap, E., Chakrabarti, M., Chu, Y.H., Clevenger, J.P., Illa-Berenguer, E., Huang, Z., et al. (2014) What lies beyond the eye: the molecular mechanisms regulating tomato fruit weight and shape. *Front. Plant Sci.* 5: 227.
- van der Knaap, E. and Østergaard, L. (2018) Shaping a fruit: developmental pathways that impact growth patterns. *Sem. Cell Dev. Biol.* 79: 27–36.
- van der Knaap, E., Sanyal, A., Jackson, S.A. and Tanksley, S.D. (2004) High-resolution fine mapping and fluorescence in situ hybridization analysis of sun, a locus controlling tomato fruit shape, reveals a region of the tomato genome prone to DNA rearrangements. *Genetics* 168: 2127–2140.
- Villanueva, J.M., Broadhvest, J., Hauser, B.A., Meister, R.J., Schneitz, K. and Gasser, C.S. (1999) INNER NO OUTER regulates abaxial-adaxial patterning in *Arabidopsis* ovules. *Genes Dev.* 13: 3160–3169.
- Vlad, D., Kierzkowski, D., Rast, M.I., Vuolo, F., Dello Iorio, R., Galinha, C., et al. (2014) Leaf shape evolution through duplication, regulatory diversification, and loss of a homeobox gene. *Science* 343: 780–783.
- Wan, C. and Wilkins, T. (1994) A modified hot borate method significantly enhances the yield of high-quality RNA from cotton (*Gossypium hirsutum* L.). *Anal. Biochem.* 223: 7–12.
- Wang, S., Chang, Y. and Ellis, B. (2016) Overview of OVATE FAMILY PROTEINS, a novel class of plant-specific growth regulators. *Front. Plant Sci.* 7: 417.
- Wang, S., Chang, Y., Guo, J. and Chen, J.G. (2007) *Arabidopsis* ovate family protein 1 is a transcriptional repressor that suppresses cell elongation. *Plant J.* 50: 858–872.

- Wu, S., Zhang, B., Keyhaninejad, N., Rodríguez, G.R., Kim, H.J., Chakrabarti, M., et al. (2018) A common genetic mechanism underlies morphological diversity in plants. *Nat. Comm* 9: 4734.
- Yamada, T., Ito, M. and Kato, M. (2004) YABBY2-homologue expression in lateral organs of *Amborella trichopoda* (Amborellaceae). *Int. J. Plant Sci.* 165: 917–924.
- Yang, Z. (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* 13: 555–556.
- Yoshioka, Y., Iwata, H., Ohsawa, R. and Ninomiya, S. (2004) Analysis of petal shape variation of *Primula sieboldii* by elliptic Fourier descriptors and principal component analysis. *Ann. Bot.* 94: 657–664.
- Young, M.D., Wakefield, M.J., Smyth, G.K. and Oshlack, A. (2010) Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11: R14.
- Zhang, B. and Horvath, S. (2005) A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* 4: Article17.