Transcriptome-wide mining of the differentially expressed transcripts for natural variation of floral organ size in *Physalis philadelphica*

Li Wang¹, Zhichao Li¹,² and Chaoying He¹,*

¹ State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Nanxincun 20, Xiangshan, 100093 Beijing, China
² Graduate University, Chinese Academy of Sciences, Yuquan Road 19, 100049 Beijing, China

* To whom correspondence should be addressed. E-mail: chaoying@ibcas.ac.cn

Received 31 July 2012; Revised 1 October 2012; Accepted 3 October 2012

**Abstract**

Natural phenotypic variation, a result of genetic variation, developed during evolution in response to environmental selections. *Physalis philadelphica*, known as tomatillo in the Solanaceae, is rich in floral and post-floral organ size diversity. However, its genetic variation is unknown. Here *P. philadelphica* was classified into three groups with large, intermediate, and small reproductive organ size, and a positive correlation was observed between floral organ and berry sizes. Through cDNA-amplified fragment length polymorphism (AFLP) analyses, 263 differentially expressed transcript-derived fragments (TDFs) were isolated from two accessions with different floral organ sizes. The genes encode various transcription factors, protein kinases, and enzymes, and they displayed multiple expression patterns during floral development, indicating a complexity in the genetic basis of phenotypic variation. Detailed expression analyses revealed that they were differentially expressed during floral and post-floral development, implying that they have roles in the development of flowers and fruits. Expression of three genes was further monitored in 26 accessions, and in particular the expression variation of *Pp30*, encoding an AP2-like transcription factor, correlates well with the observed phenotypic variations, which strongly supports an essential role for the gene in the natural variation of floral and post-floral organ size in *Physalis*. The results suggest that alteration in the expression pattern of a few key regulatory genes in the developmental process may be an important source of genetic variations that lead to natural variation in morphological traits.

**Key words:** Berry, cDNA-AFLP, flower, natural variation, organ size, *Physalis*.

**Introduction**

Natural variation is caused by spontaneous genetic mutations that are maintained during evolution. It is in general determined by multiple genes and their regulatory networks. The locations of these genes on chromosomes are referred to as quantitative trait loci (QTLs). Linking genetic changes to small perturbations in the network may allow us to understand how tuning of the network can produce phenotypic variations. Genetic changes controlling natural variation could be inferred by several approaches. In QTL mapping, phenotypic variation is associated with allelic variations in molecular markers that segregate in experimental mapping populations derived from the directed crosses (Clerkx et al., 2004; Song et al., 2007). The causal QTLs accounting for phenotypic variation are mapped on certain genomic regions. Further analyses of these regions with a combination of functional strategies allow identification of the genes, and the nucleotide polymorphisms altering gene function (Koomnef...
Association mapping by screening for phenotype–genotype associations in a general population of individuals, whose degree of relatedness or pedigree is unknown, is also becoming more popular and useful in plant systems (Myles et al., 2009). Recent genetic analyses of large data sets generated by high-throughput profiling of mapping populations with various ‘omics’ procedures allow genetic integration of several levels of molecular mechanisms regulating phenotypic variation (Chevalier et al., 2004; Manning et al., 2006; West et al., 2007). The integration of systems biology with quantitative genetic analyses of natural variation is a promising research field (Benfey and Mitchell-Olds, 2008).

Evaluation and analysis of natural genetic variation has been performed successfully in wild species, such as Arabidopsis thaliana (Mitchell-Olds and Schmitt, 2006). Arabidopsis, as a plant model, has provided the largest number of genes and nucleotide polymorphisms underlying natural variation (Mitchell-Olds and Schmitt, 2006; Alonso-Blanco et al., 2009). These analyses are helpful in elucidating the molecular basis of phenotypic differences related to plant adaptation to distinct natural environments, and to determine the ecological and evolutionary processes that maintain this variation (Mitchell-Olds and Schmitt, 2006).

Natural genetic variation has also been evaluated in cultivated crops (Purugganan and Fuller, 2009), such as Oryza sativa (Sang, 2009) and Solanum lycopersicum (Frary et al., 2000; Cong et al., 2008; Orsi and Tanksley, 2009). These analyses enable identification of genes and nucleotide polymorphisms involved in domestication, yield, biotic and abiotic stress, and quality traits (Doebley et al., 2006; Sang, 2009). However, the specific ecological niche and life history of these plants limit the plant traits and processes that can be studied in a single or a few species. Therefore, new plant models such as Helianthus (Rieseberg et al., 2003), Ipomoea (Clegg and Durbin, 2003), Physalis (He et al., 2004), Mimulus (Wu et al., 2007), and Aquilegia (Kramer, 2009) are beginning to be used in studies of natural variation and speciation. Around 100 genes and allelic variants participating in plant adaptation to different natural and agricultural environments through modifications in plant development and physiology have been identified thus far from both wild species and domesticated crops (Alonso-Blanco et al., 2009).

The cultivated solanaceous crops from Solanum, Capsicum, and Physalis are important for human consumption and well-being (Pratt et al., 2008). Five of the leading economical species are potato (Solanum tuberosum), tomato (S. lycopersicum), eggplant (S. melongena), pepper (Capsicum spp.), and husk tomato (Physalis spp.). The latter has >70 species, and features a post-floral morphological novelty—the Chinese lantern, also named the inflated calyx syndrome (ICS), and has become a new model to study evolution and development of morphological novelty (He et al., 2004). The most widely grown and distributed species is Physalis philadelphica (syn. Physalis ixocarpa), known as tomatillo. It has a rich diversity in floral morphology and berry size. The genetic bases underlying these trait variations are not known yet. QTL analysis to determine DNA variation directly is not easy because of difficulty in making an experimental mapping population. In this study, cDNA-amplified fragment length polymorphism (AFLP) technology was therefore used to dissect genetic variations underlying the floral diversity in P. philadelphica. A total of 263 differentially expressed transcript-derived fragments (TDFs) were identified. Expression variation of three of them was further characterized with respect to variation in floral morphology and berry size in a collection of P. philadelphica, and expression variation of one TDF (Pp30), encoding an AP2-like transcription factor, correlates well with natural variation in final flower and berry size. These results, focusing on alterations in gene expression during evolution, contribute to our understanding of the genetic basis of plant development and evolution.

Materials and methods

Plant material and growth conditions

Seeds of P. philadelphica were provided by BGN, Botanical and Experimental Garden of Radboud University, Nijmegen, The Netherlands. In 2009, 26 accessions (Supplementary Table S1 available at JXB online) were grown in a greenhouse under long-day conditions until flowering. For a better fruiting rate, they were moved to the experimental fields in the Institute of Botany, Chinese Academy of Sciences, Beijing. The age of mature flowers was set as 0 d old, and the age of flower buds was counted back as −10, −7, −2, and 0 d old, followed by berries which were 14, 20, and 30 d old. All developmental stages including buds, flowers and berries were harvested, immediately frozen in liquid nitrogen, and stored at −80 °C. Three independent biological samples were collected and total RNA was isolated using the TRIzol® reagent (Invitrogen).

Visualization of chromosome number

Root tips from 1-week-old seedlings of P58 and P64 were collected to count the chromosome number. They were fixed in Carnoy’s fixative (ethanol/glacial acetic acid 3:1) at room temperature for at least 4 h before storing at −20 °C. Chromosome spreads were then stained with 4′,6-diamidino-2-phenylindole and slides were examined under a fluorescence microscope (Axioskop40 with HBO100, Zeiss).

The cDNA-AFLP analysis

The cDNA-AFLP-based transcript profiling was performed as previously described (Vuysteke et al., 2007). Total RNA was first treated with RNase-free DNase (Promega) to remove DNA contamination. About 5.0 µg of total RNA was used as the starting material for the first- and second-strand cDNA syntheses and cDNA-AFLP template preparation. Restriction enzymes used were BstYI and Msel (New England BioLabs). For pre-amplifications, a non-selective BstYI primer containing a T or C at the 3′ end. Amplified products were diluted 500-fold and 5.0 µl was used in the final selective amplifications with Msel+NN primer in combination with BstYI+CN or BstYI+TN primer in a 50.0 µl reaction volume. N represents A, T, C, and G. All possible primer combinations (2 × 8 × 4 × 4 = 128) were used (Supplementary Table S2 at JXB online). Selective amplification products were separated on a 6.0% polyacrylamide denaturing gel at 100 W until bromophenol blue reached the bottom of the gel.

Cloning the differentially expressed fragments

After staining the gels with silver nitrate, all bands of differential expression were cut with a razor blade and boiled in 100 µl of ddH2O until bromophenol blue reached the bottom of the gel. After purification using the High Pure PCR Product Purification Kit (Roche). The fragments were cloned into the pGEM®-T easy Vector (Promega) and transformed into competent Escherichia coli DH5α cells (Tiangen) by electroporation. For each TDF, plasmid DNA from three independent...
clones was extracted using the AxyPrep™ Plasmid Miniprep Kit (Axygen). Sequencing of the inserts was carried out commercially by BGI (Beijing Genomics Institute). Sequences were annotated using the BLAST program and are deposited in GenBank with accession numbers JK998347–JK998611.

Gene ontology classification

Each TDF was functionally annotated into one of four categories: (i) biological process; (ii) molecular function; (iii) cellular component; and (iv) non-functional protein according to the Gene Ontology Database (Ashburner et al., 2000). The proportion of each category to the total TDFs cloned and of subcategories to its category was calculated.

Expression analyses

To confirm the differential expression in cDNA-AFLP analyses, reverse transcription–PCR (RT–PCR) analysis was performed using M-MLV reverse transcriptase (Invitrogen) with the ACTIN gene as an internal control. PCR products were separated on a 1.2% agarose gel. Primers used to amplify the 24 differentially expressed transcripts used are presented in Supplementary Table S3 at JXB online.

For real-time RT–PCR, about 2.0 µg of total RNA was used for cDNA synthesis with the SuperScript® II Reverse Transcriptase (Invitrogen). All reactions were repeated three times. Assays were performed with 1.0 µl of cDNA, diluted 2-fold, using the SYBR® Premix Ex Taq™ (TaKaRa). The relative quantification method (ΔΔCT) was used to evaluate variation between replicates. The ACTIN gene was used as an internal control to normalize the data. PCR conditions consisted of an initial denaturation at 95 °C for 3 min, followed by 40 cycles at 95 °C for 30 s, 60 °C for 1 min, and 72 °C for 30 s, and a final extension at 72 °C for 5 min. Amplification efficiencies of the gene-specific primers were between 95% and 105%. The primers were: for Pp30, 5'-TAG TGG TGG CAC CAT GAT GAA T-3' and 5'-AAA TGG GGC AGA CTC T-3'; for Pp97, 5'-TAG TGG TGG CAC CAT GAT GAA T-3' and 5'-AAA TGG GGC AGA CTC T-3'; and for ACTIN, 5'-GGG TAT GCC CTT CCA CAT GCC A-3' and 5'-GAG CCT CCA ATC CAG ACA CTG TAT T-3'.

Analysis of morphological traits

The size of the flower bud or flower represents the length between the receptacle and the tip of the flower buds (or petal tip). Berry and ovary size were defined either by the weight or the distance from the receptacle to the place at which the stigma was located. For each accession, the size of 20 flowers and 20 berries was quantified. All pictures were taken under a microscope equipped with a Nikon camera. The correlation coefficient between flower size, berry size, and gene expression was evaluated using SPSS 13.0.

Results

Size diversity of flowers and berries in Physalis philadelphica

Plants of 26 accessions (provenances) of P. philadelphica that were collected feature a rich diversity in the size of floral organs, berries, and leaves, flowering time, and response to viruses. In this study, floral diversity in particular was evaluated. Basic floral structure and morphology in these accessions are similar, with dramatic variations in the size of mature flowers (Supplementary Table S1 at JXB online). The distance between the receptacle and tip of the corolla was recorded for flower size, which ranges from 3.9 mm to 11.9 mm among different provenances. Based on flower size, they can be classified into three groups: large (>10.0 mm), intermediate (5.0–10.0 mm), and small (<5.0 mm), indicating a huge variation in the trait (Fig. 1A; Supplementary Table S1). The weight of the mature berries ranges from 1.2 g to 11.2 g and is also classified into three groups: the large group with berries heavier than 8.0 g; the intermediate group with fruits ranging from 4.0 g to 8.0 g; and the small group with berries lighter than 4.0 g. Notably, berry weights of all seven accessions in the small group are ~1.0–2.0 g (Fig. 1A; Supplementary Table S1). These findings suggest a significant variation in berry size. Correlation is observed between the flower and berry sizes within the small (s-R^2=0.6279), intermediate (i-R^2=0.8864), and large groups (l-R^2=0.8209), and a high correlation is observed among the whole collection (w-R^2=0.9386). These correlations indicate that the same set of genes may regulate both floral and post-floral growth, considered as different developmental processes in P. philadelphica.

A comparison of floral phenotypic differences between the two accessions P58 and P64 which are representative of the large and intermediate group is presented. At the same developmental stage, the size of flower buds in P64 is smaller than in P58 (Fig. 1B). Similarly, the size of mature flowers (Fig. 1C) and ovaries (Fig. 1D) also differs between P64 and P58, and berry size in P64 is two-thirds of that in P58 (Fig. 1A; Supplementary Table S1 at JXB online). Since polyploidy, which normally influences organ size, often occurs in Solanaceae, the chromosome number in the two accessions was checked. The chromosome number of P58 is 22.8±1.14 and that of P64 is 23.0±1.04 (Supplementary Fig. S1), indicating the occurrence of diploidy.

Fig. 1. Diversity of floral morphology and berry size in Physalis philadelphica. (A) Size distribution of flowers and berries. Blue represents flower size and red represents berry weight. The size correlation coefficient between flowers and berries in mature plants is given as s-R^2, i-R^2, and l-R^2 for small, intermediate, and large groups as indicated, and the whole collection, respectively. (B) Flower buds. (C) Mature flowers. (D) Carpels. The corolla, stamen, and part of the calyx were removed from mature flowers to show the carpel. On the left is P64 and on the right is P58. Scale bars=5 mm. Measurement of organ size is indicated with red lines and white arrows.
(2n=24). Therefore, these two accessions are ideal to determine genetic variations underlying the floral phenotypic variations. Because structural variations in the genome could result in differential expression of genes, which in turn could create variations in floral phenotype, the focus was on identifying the differentially expressed TDFs.

Identifying differentially expressed floral TDFs via cDNA-AFLP analysis

To obtain differentially expressed TDFs, a cDNA-AFLP analysis was performed on RNA samples of flower buds and mature flowers from the two *P. philadelphica* accessions. In total, four samples, P58 buds, P58 flowers, P64 buds, and P64 flowers, were compared using different primer combinations. Each primer combination formed one group and a total of 128 primer combinations was used in the study. An image of the seven groups (g1–g7) is presented in Supplementary Fig. S2A at *JXB* online. About 59 TDFs, ranging in size from 100 bp to 1000 bp, were visualized as bands per group. Thus ~7500 transcripts were presented in 128 groups. The reproducibility of these profiles was determined by repeating the experiments using independent biological samples, and the differentially expressed TDFs that appear in both replicates were analysed further (Supplementary Fig. S2A). The focus was on bands with on/off patterns among four samples in the two accessions. In total, 263 differentially expressed TDFs were cloned and sequenced (Supplementary Fig. S2B at *JXB* online). Therefore, ~62% of these TDFs are unique to one of the accessions, ~26% are differentially expressed between flower buds and mature flowers among accessions, and only ~12% are differentially expressed between floral buds and flowers in one of the accessions. In a total of 263 differentially expressed transcripts, the percentages which are down-regulated (types 2, 5, 7, 9, and 11) and up-regulated (types 3, 6, 8, 10, and 12) during flower development are similar at ~36.5% and 36.1%, respectively.

Therefore, multiple types of differential expression are observed in a substantial number of genes involved in floral development in P58 and P64.

Functional categories of the differentially expressed transcripts

A complete list of the differentially expressed TDFs isolated from the two accessions is available as *Pp1*–*Pp263*. Their lengths range from 108 bp to 840 bp. When compared with the tomato database (http://mips.helmholtz-muenchen.de/), 193 TDFs were shared between the two species, while 70 TDFs with unknown functions are unique to *P. philadelphica*. Each transcript was functionally annotated based on the Gene Ontology Database (Ashburner *et al*., 2000). Figure 2 shows the percentages of differentially expressed TDFs between P58 and P64 in the four categories of gene ontology. Approximately 38.02% of the annotated sequences are categorized under ‘molecular function’, with 11.00% being transporters and 89% enzymes. These enzymes are either involved in signal transduction (12.00% kinase) or have primary metabolic roles particularly in protein

![Fig. 2. Functional category of the cloned differentially expressed genes. The cloned TDFs are classified into four categories of gene ontology as indicated in the pie diagram (middle). The percentage for each category in the TDFs is given in the corresponding sector. Molecular function includes transporter and enzymes as shown on the left-hand chart. The percentage of enzyme type in the total enzymes is given in parentheses. The right-hand chart indicates biological process, and these genes participate in cell growth and maintenance, transcriptional regulation, and purine metabolism. Percentages in this category are given in the corresponding sector.](image-url)
Expression verification and developmental regulation of candidate genes

In order to confirm the reliability of cDNA-AFLP analyses, some TDFs were randomly selected for further analyses. Based on their expression patterns (Supplementary Fig. S2B at JXB online), 24 genes were selected, with two TDFs in each type, to investigate their mRNA expression. These include transcription factors, various kinases, enzymes, and the sequences unique to Physalis with unknown function (Supplementary Table S3), representing 9.1% of the cloned TDFs.

For further expression analyses, developing flower buds (–10, –7, and –2 d old), mature flowers (set as 0 d old), and developing berries (including the 14-, 20-, and 30-day-old berries) were harvested from P58 and P64 and expression of the 24 selected TDFs was monitored by RT–PCR. Differential expression was observed during flower development, consistent with the observations in cDNA-AFLP analyses (Supplementary Figs S2B, S3 at JXB online). It was also found that homologous genes in the various functional categories have different expression patterns between the two accesses. For example, \(Pp170\) and \(Pp178\), encoding different isoforms of pyruvate kinase, are type 2 and type 4, while \(Pp105\) and \(Pp206\), which are the homologous genes of zinc finger C3HC4-type RING finger family proteins, are type 3 and type 11, and \(Pp62\) and \(Pp174\) (the transcripts of serine-threonine protein kinase-like proteins) are type 8 and type 9 (Supplementary Table S4; Figs S2B, S3). However, the robustness in differential expression between floral buds and mature flowers suggests their potential roles in floral development. During fruit development, expression of these 24 genes is also developmentally regulated with different patterns. They are either up-regulated (like \(Pp30\) and \(Pp90\), down-regulated (like \(Pp97\) and \(Pp206\), or constitutively expressed (like \(Pp58\) and \(Pp64\) after fertilization (Supplementary Fig. S3). These results suggest their role in post-floral development.

Based on these expression data, three TDFs (\(Pp30\), \(Pp58\), and \(Pp97\)) were further selected as candidates to determine the transcriptional regulation of reproductive organ development and their size variation. \(Pp30\), \(Pp58\), and \(Pp97\) encode a putative transcription factor with two AP2 domains, a protein phosphatase-2C and a GAMYB-like transcription factor, respectively (Supplementary Table S4 at JXB online). However, analysis of gene expression between the two accesses is insufficient to draw a conclusion on variations in organ size. Therefore, an association analysis of gene expression with organ development and phenotypes was performed in a larger collection/population.

Expression of candidate genes correlates to organ development and growth

In order to approach this goal, three accesses of P58, P64, and P59 were selected as representatives of three groups with large, intermediate, and small size flowers and fruits (Fig. 1A; Supplementary Table S1 at JXB online). First, their development and growth curves (Supplementary Fig. S4) were documented. All developmental stages including the mature flower set as 0 d old, –10, –7, and –2-day-old flower buds, 0-d-old mature flowers (Supplementary Fig. S4A), 0-day-old ovaries, and 14-, 20-, 30-, and 40-day-old berries (Supplementary Fig. S4B) were harvested and their sizes were quantified. The curves reflect a size difference during the development of both floral organs and berries (Supplementary Fig. S4), resulting in a correlation in final organ sizes in the collection of \(P. philadelphica\) (Fig. 1A).

Expression of the three characterized TDFs, \(Pp30\), \(Pp58\), and \(Pp97\) was determined by real-time RT–PCR analyses during the development of flowers and berries in all three accesses. During floral development, the expression level of \(Pp30\) was the highest in P58, followed by lower levels in P64, and extremely low levels in P59. (Fig. 3A), suggesting a possible correlation between its expression and flower size. Thus, \(Pp30\) is likely to promote floral organ growth. During fruit development, its expression is triggered by fertilization and is up-regulated developmentally. In addition, the difference in expression levels in all three accesses remains stable (Fig. 3A), indicating involvement of \(Pp30\) in fruit development. Taken together, it is likely that \(Pp30\) might be a promoter of the reproductive organ size. Similarly, during flower development, \(Pp58\) is steadily up-regulated in P58 at high levels, while it is triggered 2 d before anthesis in P64 and P59. Higher expression is maintained during berry development (Fig. 3B), implying its role in the development of both flowers and berries. In all three accesses, \(Pp97\) is down-regulated during flower development. Expression is not observed after fertilization, and extremely low expression is observed during berry development (Fig. 3C), which indicates a major role in flower development. The roles of these genes in causing natural variation in organ size were revealed by association analyses, although their roles in flower and fruit development in \(Physalis\) should be confirmed by genetic manipulations.

\(Pp30\) is associated with natural variation in reproductive organ size

In order to reveal their putative roles in phenotypic variations of reproductive organ size, association of the expression of the three TDFs (\(Pp30\), \(Pp58\), and \(Pp97\)) with the diversity of floral and post-floral organ size was evaluated in a tomatillo collection. Mature flowers were harvested from each accession and total RNAs were subjected to real-time RT–PCR analyses. Correlation of the gene expression with flower and berry size was estimated. The correlation coefficient of \(Pp30\) expression with flower and
berry size is 0.8754 and 0.8951, respectively (Fig. 4), indicating that expression of \( Pp30 \) is well associated with the observed size variation in flowers and the berries. This also suggests that it might play a role in natural variation of reproductive organ size. The correlation coefficient of \( Pp58 \) and \( Pp97 \) gene expression with flower/berry size is 0.0446/0.0334 and –0.2714/–0.2179, respectively (Fig. 4). These results suggest that expression of both \( Pp58 \) and \( Pp97 \) does not correlate with organ size. Although these three genes are developmentally regulated, only \( Pp30 \) has a clear association with natural variations in organ size, indicating that not all genes with a role in organ development are recruited in the evolution of natural variation.

**Discussion**

The size of floral organs and berries varies dramatically in *P. philadelphica*. In this study, a clear correlation was found between flower size and berry size. A total of 263 differentially expressed TDFs were identified by cDNA-AFLP analyses. Association between gene expression and organ size variation suggests that, during evolution, changes in expression of a few genes in organ development may be an important way to evolve natural variation.

*Physalis* is a new model for evolution, natural variation, and ecology

The genus *Physalis* belongs to the family Solanaceae, with attractive characteristics. First of all, *Physalis* is distinguished within Solanaceae because of its novel post-floral morphological trait called the ‘Chinese lantern’ or ICS (He et al., 2004). This novelty is an intriguing evolutionary question to address. Currently, it has been demonstrated that heterotopic expression of an MPF2-like gene is critical to the origin of this trait (He and Saedler, 2005; Khan et al., 2009). Divergence of MPF2-like proteins is correlated to the evolution of ICS within Solanaceae (Zhang et al., 2012). MPF2 interacts with protein products of...
many floral expressed MADS-box genes, by affecting calyx growth, to control ICS development (He et al., 2007), indicating that ICS is an apparently simple but complex trait. To understand its complete genetic architecture is a challenge. In addition, fertilization signals, which could be replaced by hormones such as cytokinin and gibberellins, trigger ICS development (He and Saedler, 2007). MPF2 is also involved in male fertility in Physalis (He and Saedler, 2005; He et al., 2007), indicating that fertility may be an important selective pressure for the origin of ICS. However, its selective values, such as ecological function, need further investigation.

The berry, as the second post-floral organ, is an important agronomic trait. Natural variation in berry size is observed in Physalis. Fruit size has been extensively investigated in Solanum and Capsicum (Frary et al., 2000; Zygier et al., 2005; Cong et al., 2008), and a microsynteny of the loci controlling fruit size is revealed between these genera (Zygier et al., 2005). Although Physalis is phylogenetically related to Solanum, it remains to be determined whether the regulatory networks that control berry development and growth are conserved. Molecular improvement of Physalis will become feasible when these regulatory processes are determined.

As a post-floral trait, the ICS rapidly develops as berry size increases after a successful fertilization. It is likely that they share common regulators that control development and growth, which are yet to be determined. Actually, regulation of floral growth is not understood yet, although the molecular process of organ identity has been well studied (Theißen, 2001). Natural variation in the observed floral organ size will bridge the gap between floral organ identity and their growth.

Based on these observations, many platforms and molecular tools for the genus Physalis have been established. For example, gene silencing systems such as RNA interference (He and Saedler, 2005; He et al., 2010) and virus-induced gene silencing (VIGS) (J.S. Zhang, unpublished) have been established. Protein–protein interaction tools such as the yeast library (He et al., 2007) and BiFC assay (He et al., 2010) are available. All these facilitate functional analyses of a gene in Physalis. Thus Physalis is an ideal system to study ecology, evolution, and natural variation, to decipher molecular evolution of biodiversity, and to evaluate the ecological effects in response to different environments.

Co-evolution of floral organ and berry size in P. philadelphica

The size of floral organs varies dramatically in P. philadelphica, thus providing an opportunity to take advantage of this natural variation to decipher the genetic basis underlying floral organ size. The 263 cloned TDFs are the arsenal to approach this goal. A dramatic variation in berry size is also observed. Surprisingly, a good size correlation between floral organs and berries is also demonstrated. Synergistic size control in different organs is often observed by genetic manipulation of some genes (Park et al., 2007; Powell and Lenhard, 2012), which could also be used by nature to evolve a tissue-specific pattern by restricting gene expression in certain tissues, such as floral organs and fruits in Physalis. The present findings indicate that floral organ size and fruit size may have co-evolved in P. philadelphica. Although the genetic basis for this co-evolution is unknown, the same set of genes may be recruited in the growth of both flower and berry.

Genetic control of berry size is relatively well studied in S. lycopersicum. FAS and fw2.2 are repressors that are selected for the evolution of tomato size during its domestication (Frary et al., 2000; Cong et al., 2008). In Physalis, the orthologue of fw2.2 exerts a role in control of berry size, while the EAS orthologue does not seem to be recruited (Z.C. Li, unpublished), suggesting both conservation and diversification in the regulatory networks between different genera within Solanaceae. The cloned TDFs were developmentally regulated during flower and berry development in Physalis. They are proposed as promoters or repressors of these organ sizes. In particular, Pp30, encoding an AP2-like transcription factor, has been associated with size variation of both flowers and berries. Pp30 might be a key regulatory factor in determining genetic architecture for the co-evolution of observed natural variation. It represents a promoter which is a new type of regulator involved in the domestication of fruit size in solanaceous crops.

Complexity of genetic variation determining polymorphism of floral organ size

In Physalis, the two accessions P58 and P64 feature different sizes of floral organs. The size of mature berries also differs significantly. Variations are observed in the Physalis collection of 26 different provenances. Phenotypic variations in general result from alterations in gene activities; therefore, it is likely that the differentially expressed TDFs formed genetic variations causing natural phenotypic variations. Since genome or transcriptome data are not available for Physalis, cDNA-AFLP technology, which is an efficient method of global transcriptional analysis for any species without a reference transcriptome (Ganeshan et al., 2011; Yu et al., 2011; Yang et al., 2012), was exploited to search for genes causing natural variation in the observed organ size. The 263 cloned differentially expressed TDFs have complicated expression patterns and they encode regulatory factors, protein kinases, protein phosphatase, and various enzymes, which contribute to the complexity of the genetic basis for the different floral phenotypes. Since their expression is also regulated during berry development and it is known that berry size co-evolved with floral organ size, these TDFs might also play a role in fruit development. Future work will be to determine systematically the regulation of differential gene expression between accessions and regulation during floral and post-floral development, and to reveal their roles in natural variation in floral organs and berries. By evolving an altered gene expression, these genes apparently form an important genetic basis for the development and evolution of diverse traits. Genetics of development versus evolution

Evolution of a morphological trait involves a modification in the developmental programmes or processes. Genetic variation resulting in heterochrony or heterotopy of expression of a gene in a developmental programme (Frary et al., 2000; Cong et al., 2002; He and Saedler, 2005), or neofunctionalization
of a regulatory protein (Cong et al., 2008), fulfils the evolution of development, leading to the morphological diversification including organ size and shape. The genetic architecture for developing a trait is complex. Does a change account for evolution of natural variation of the trait? Increasing evidence suggests that variations could be caused through alteration in gene expression, mutations in coding regions during evolution (Doebely et al., 2006; Orsi and Tanksley, 2009; Purugganan and Fuller, 2009; Sang, 2009), or epigenetic variation (Cubas et al., 1999; Manning et al., 2006; Shindo et al., 2006). Focusing on gene expression, this study sought to identify the genes playing an evolutionary role in natural variation of reproductive organ size. In Physalis, Pp30, Pp58, and Pp97 are all regulated during the development of flower organs and berries. The expression pattern is similar among the accessions. However, in a population association analysis, the expression of only Pp30 during development correlated with natural variation in floral organ and berry size, indicating an essential role in both development and evolution of floral organs and fruits. Among all TDFs identified, Pp30 is therefore a valuable candidate to understand size variation in floral organs and berries. The present findings allude to the possibility that genes involved in development are not necessarily recruited for evolution; however, genes recruited for the evolution of natural variation of a trait should be involved in trait development. To improve fitness, modifications in the key developmental regulators in particular could be selected and maintained during evolution.

Through transcriptome-wide screens, a series of differentially expressed TDFs between accessions with different sizes of floral organs were identified. It was discovered that Pp30 encoding an AP2-like transcription factor, which controls both flower and fruit size, is a valuable candidate to study the development and evolution of natural variation in reproductive organ size in Physalis.

Supplementary data

Supplementary data are available at JXB online.

Figure S1. Chromosome number in two accessions of Physalis philadelphica.

Figure S2. cDNA-AFLP analyses to identify the differentially expressed TDFs.

Figure S3. Expression verification of 24 candidate genes from the cloned TDFs.

Figure S4. Developmental and growth curves of flowers and berries.

Table S1. Diversity of organ size in Physalis philadelphica.

Table S2. Primers used for the cDNA-AFLP analyses.

Table S3. Primers used for RT–PCR analyses.

Table S4. Information on the TDFs verified.

Acknowledgements

We thank Dr H. Saedler for his generous support in initiating this project. We also thank the two anonymous reviewers for their constructive comments and suggestions to improve the manuscript. This work was supported by a grant (30870175) from the National Natural Science Foundation of China (to CYH) and by the Hundred Talents Project of the Chinese Academy of Sciences (to CYH).

References

Alonso-Blanco C, Aarts MG, Bentsink L, Keurentjes JJ, Reymond M, Vreugdenhil D, Koornneef M. 2009. What has natural variation taught us about plant development, physiology, and adaptation? The Plant Cell 21, 1877–1896.

Ashburner M, Ball CA, Blake JA, et al. 2000. The Gene Ontology Consortium: Gene Ontology: tool for the unification of biology. Nature Genetics 25, 25–29.

Benfey PN, Mitchell-Olds T. 2008. From genotype to phenotype: systems biology meets natural variation. Science 320, 495–497.

Chevalier F, Martin O, Rofidal V, Devauchelle AD, Barateau S, Sommerer N, Rossignol M. 2004. Proteomic investigation of natural variation between Arabidopsis ecotypes. Proteomics 4, 1372–1381.

Clegg MT, Durbin ML. 2003. Tracing floral adaptations from ecology to molecules. Nature Reviews Genetics 4, 206–215.

Clerkx EJ, EI-Lithy ME, Vierling E, Ruys GJ, Blankestijn-De Vries H, Groot SP, Vreugdenhil D, Koornneef M. 2004. Analysis of natural allelic variation of Arabidopsis seed germination and seed longevity traits between the accessions Landsberg erecta and Shakdara, using a new recombinant inbred line population. Plant Physiology 135, 432–443.

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

Cong B, Liu J, Tanksley SD. 2002. Natural alleles at a tomato fruit size quantitative trait locus differ by heterochronic regulatory mutations. Proceedings of the National Academy of Sciences, USA 99, 13606–13611.

Cubas P, Vincent C, Coen E. 1999. An epigenetic mutation responsible for natural variation in floral symmetry. Science 289, 85–88.

Doebely JF, Gaut BS, Smith BD. 2006. The molecular genetics of crop domestication. Cell 127, 1309–1321.

Frary A, Nesbitt TC, Grandillo S, Knaep E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tanksley SD. 2000. fw2.2: a quantitative trait locus key to the evolution of tomato fruit size. Science 289, 85–88.

Ganeshan S, Sharma P, Young L, Kumar A, Fowler DB, Chibbar RN. 2011. Contrasting cDNA-AFLP profiles between crown and leaf tissues of cold-acclimated wheat plants indicate differing regulatory circuits for low temperature tolerance. Plant Molecular Biology 75, 379–398.

He CY, Münster T, Saedler H. 2004. On the origin of floral morphological novelties. FEBS Letters 567, 147–151.

He CY, Saedler H. 2005. Heterotopic expression of MPF2 is the key to the evolution of the Chinese lantern of Physalis, a morphological novelty in Solanaceae. Proceedings of the National Academy of Sciences, USA 102, 5779–5784.

He CY, Saedler H. 2007. Hormonal control of the inflated calyx syndrome, a morphological novelty, in Physalis. The Plant Journal 49, 935–946.
He CY, Sommer H, Grosardt B, Huijser P, Saedler H. 2007. PFMAGO, a MAGO NASHII-like factor, interacts with the MADS-domain protein MPF2 from Physalis floridana. Molecular Biology and Evolution 24, 1229–1241.

He CY, Tian Y, Saedler E, Efremova N, Riss S, Khan MR, Yephremov A, Saedler H. 2010. The MADS-domain protein MPF1 of Physalis floridana controls plant architecture, seed development and flowering time. Planta 231, 767–777.

Khan MR, Hu J, Riss S, He CY, Saedler H. 2009. MPF2-like-A MADS-box genes control the inflated calyx syndrome in Withania (Solanaceae): roles of Darwinian’s selection. Molecular Biology and Evolution 26, 2463–2473.

Koornneef M, Alonso-Blanco C, Vreugdenhil D. 2004. Naturally occurring genetic variation in Arabidopsis thaliana. Annual Review of Plant Biology 55, 141–172.

Kramer EM. 2009. Aquilegia: a new model for plant development, ecology, and evolution. Annual Review of Plant Biology 60, 261–277.

Manning K, Tor M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ, Seymour GB. 2006. A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. Nature Genetics 38, 948–952.

Mitchell-Olds T, Schmitt J. 2006. Genetic mechanisms and evolutionary significance of natural variation in Arabidopsis. Nature 441, 947–952.

Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, Buckler ES. 2009. Association mapping: critical considerations shift from genotyping to experimental design. The Plant Cell 21, 2194–2202.

Orsi CH, Tanksley SD. 2009. Natural variation in an ABC transporter gene associated with seed size evolution in tomato species. PLoS Genetics 5, e1000347.

Park EJ, Jeknič Z, Chen TH, Murata N. 2007. The codA transgene for glycinebetaine synthesis increases the size of flowers and fruits in tomato. Plant Biotechnology Journal 5, 422–430.

Powell AE, Lenhard M. 2012. Control of organ size in plants. Current Biology 22, R360–R367.

Pratt RC, Francis DM, BarreroMeneses LS. 2008. Genomics of tropical solanaceous species: established and emerging crops. In: Moore PH, Ming R, eds. Genomics of tropical crop plants. Berlin: Springer, 453–467.

Purugganan MD, Fuller DG. 2009. The nature of selection during plant domestication. Nature 457, 843–848.

Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, Nakazato T, Durphy JL, Schwarzbach AE, Donovan LA, Lexer C. 2003. Major ecological transitions in annual sunflowers facilitated by hybridization. Science 301, 1211–1216.

Sang T. 2009. Genes and mutations underlying domestication transitions in grasses. Plant Physiology 149, 63–70.

Shindo C, Lister C, Crevillon P, Nordborg M, Dean C. 2006. Variation in the epigenetic silencing of FLC contributes to natural variation in Arabidopsis vernalization response. Genes and Development 20, 3079–3083.

Song XJ, Huang W, Shi M, Zhu MZ, Lin HX. 2007. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nature Genetics 39, 623–630.

Theißen G. 2001. Development of floral organ identity: stories from the MADS house. Current Opinion in Plant Biology 4, 75–85.

Vuylsteke M, Peleman JD, van Eijk MJT. 2007. AFLP-based transcript profiling (cDNA-AFLP) for genome-wide expression analysis. Nature Protocols 2, 1399–1413.

West MA, Kim K, Kliebenstein DJ, van Leeuwen H, Michelmore RW, Doerge RW, St Clair DA. 2007. Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in Arabidopsis. Genetics 175, 1441–1450.

Wu CA, Lowry DB, Cooley AM, Wright KM, Lee YW, Willis JH. 2007. Mimulus is an emerging model system for the integration of ecological and genomic studies. Heredity 100, 220–230.

Yu G, Duan J, Yan H, Song B, He Z. 2011. cDNA-AFLP analysis of gene expression differences between the flower bud and sprout-shoot apical meristem of Angelica sinensis (Oliv.) Diels. Genetics and Molecular Biology 34, 274–279.

Zhang JS, Khan MR, Tian Y, Li ZC, Simone R, He CY. 2012. Divergences of MPF2-like MADS-domain proteins have an association with the evolution of the inflated calyx syndrome within Solanaceae. Plants 236, 1247–1260.

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. Theoretical and Applied Genetics 111, 437–445.