Title

Evolutionary developmental transcriptomics reveals a gene network module regulating interspecific diversity in plant leaf shape

Journal

Proceedings of the National Academy of Sciences of the United States of America, 111(25)

ISSN

0027-8424

Authors

Ichihashi, Yasunori
Aguilar-Martínez, José Antonio
Farhi, Moran
et al.

Publication Date

2014-06-24

DOI

10.1073/pnas.1402835111

Peer reviewed

Evolutionary developmental transcriptomics reveals a gene network module regulating interspecific diversity in plant leaf shape

Yasunori Ichihashi[a,b], José Antonio Aguilar-Martínez[a], Moran Farhi[a], Daniel H. Chitwood[a,1], Ravi Kumar[a,2], Lee V. Millon[c], Jie Peng[d], Julin N. Maloof[a], and Neelima R. Sinha[a,3]

[a]Department of Plant Biology, [c]School of Veterinary Medicine, and [d]Department of Statistics, University of California, Davis, CA 95616; and [b]Center for Sustainable Resource Science, RIKEN, Yokohama, Kanagawa 230-0045, Japan

Edited by Philip N. Benfey, Duke University, Durham, NC, and approved May 14, 2014 (received for review February 14, 2014)

Despite a long-standing interest in the genetic basis of morphological diversity, the molecular mechanisms that give rise to developmental variation are incompletely understood. Here, we use comparative transcriptomics coupled with the construction of gene coexpression networks to predict a gene regulatory network (GRN) for leaf development in tomato and two related wild species with strikingly different leaf morphologies. The core network in the leaf developmental GRN contains regulators of leaf morphology that function in global cell proliferation with peripheral gene network modules (GNMs). The *BLADE-ON-PETIOLE* (*BOP*) transcription factor in one GNM controls the core network by altering effective concentration of the KNOTTED-like HOMEOBOX gene product. Comparative network analysis and experimental perturbations of *BOP* levels suggest that variation in *BOP* expression could explain the diversity in leaf complexity among these species through dynamic rewiring of interactions in the GRN. The peripheral location of the *BOP*-containing GNM in the leaf developmental GRN and the phenotypic mimics of evolutionary diversity caused by alteration in *BOP* levels identify a key role for this GNM in canalizing the leaf morphospace by modifying the maturation schedule of leaves to create morphological diversity.

Solanum species | RNA-seq | bioinformatics

Although morphological diversity abounds, its underlying causes remain largely unexplored at the molecular level. Comparative studies of development in an evolutionary context [evolutionary developmental biology (evo-devo)] have been used to understand the developmental mechanisms that are modulated over time to generate morphological diversity. In attempting to elucidate how developmental regulation was modulated over time, these studies have relied on quantitative genetics and candidate gene expression further informed by functional analyses (1–4). Even though pursued on a gene-by-gene level, these evo-devo studies underscore the importance of gene expression regulation, suggesting the rewiring of developmental gene regulatory networks (GRNs) is a crucial causal factor driving morphological evolution (5). Recent development of genomic tools enabled generation of large datasets that can be used for understanding complex biological processes such as development (6). These data can be used to determine exactly how developmental gene modules are organized into a network hierarchy that generates morphological diversity. However, studies using genome-wide data to generate GRNs within an evo-devo context are lacking.

Angiosperms exhibit a wide diversity of leaf shapes. Leaf development has been characterized in several species, making the leaf a model organ for analyzing the mechanisms underlying natural morphological variation in plants. Leaf complexity, the degree to which a leaf is subdivided into smaller segments, is the most conspicuous characteristic of leaf shape. One important regulator of leaf complexity is the homeobox family of transcription factors KNOTTED-like HOMEOBOX (KNOX). *KNOX* factors maintain the shoot apical meristem (SAM), a pluripotent cell population, which generates the entire aboveground body of vascular plants (7, 8). Current evo-devo studies suggest that *KNOX* expression was recruited repeatedly to generate natural variation in leaf shape, including leaf complexity, in several plant lineages (2, 7–10). Defining the position of *KNOX* in the leaf developmental GRN hierarchy (i.e., *KNOX* regulation, and its downstream targets in a network context) will allow us to determine whether the *KNOX*-containing gene module was an evolutionary hot spot that was repeatedly recruited for generating the natural variation in leaf shape.

Domesticated tomato (*Solanum lycopersicum*) is an excellent model species for the analysis of natural variation in leaf complexity because a number of tomato wild relatives show tremendous variation in this trait (11, 12), and their genome and transcriptome data are now available (13–15). *S. lycopersicum* and its wild relatives, *Solanum pennellii* and *Solanum habrochaites* (15, 16), show a varying number of leaflets on the main leaf axis (total number of primary and intercalary leaflets, termed leaf complexity in this study; Fig. 1*A*). Morphometric analyses of leaf development suggest that the duration of leaf maturation correlates with the leaf complexity of these species (*SI Appendix*, Fig. S1).

Significance

Ever since Darwin's pioneering research, a major challenge in biology has been to understand the genetic basis of morphological evolution. Utilizing the natural variation in leaf morphology between tomato and two related wild species, we identified a gene network module that leads to a dynamic rewiring of interactions in the whole leaf developmental gene regulatory network. Our work experimentally validates the hypothesis that peripheral regions of network, rather than network hubs, are more likely to contribute to evolutionary innovations. Our data also suggest that, likely due to their bottleneck location in the network, the regulation in KNOX homeobox genes was repeatedly manipulated to generate natural variation in leaf shape.

Author contributions: Y.I. and N.R.S. designed research; Y.I., J.A.A.-M., M.F., D.H.C., R.K., and L.V.M. performed research; J.P. and J.N.M. contributed new reagents/analytic tools; Y.I. analyzed data; and Y.I. and N.R.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The sequences reported in this paper have been deposited in the Sequence Read Archive (http://trace.ncbi.nlm.nih.gov/Traces/sra/) (accession nos. SRP041563 and SRP041731).

[1]Present address: Donald Danforth Plant Science Center, St. Louis, MO 63132.

[2]Present address: Novozymes, Inc., Davis, CA 95618.

[3]To whom correspondence should be addressed. E-mail: nrsinha@ucdavis.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1402835111/-/DCSupplemental.

Results and Discussion

Leaf Developmental GRN in the Tomato Species Complex. To predict leaf developmental GRN within an evo-devo context, we used cross-species, tissue-specific, and large-scale RNA-seq on two different regions of leaf primordia (proximal and distal regions, corresponding to the leaflet-forming region and terminal differentiating/differentiated leaflet, respectively) at four developmental stages (meristem + P1–P3; P4; P5; P6) across three species (*S. pennellii*, *S. lycopersicum*, and *S. habrochaites*; *SI Appendix*, Fig. S1 and Table S1 and Dataset S1). To extract genes functionally relevant to leaf development, we generated a self-organizing map (SOM) from the read counts and partitioned the resulting SOM clusters using principal-component analysis (PCA) (17). This analysis revealed three major SOM gene clusters whose expression patterns varied along leaf developmental stages in our dataset: clusters 1 and 2 show increased and decreased expression changes, respectively, along the leaf development gradient, and cluster 3 shows specific expression at the proximal region of P5 leaf primordia (Fig. 1*B*). Genes in clusters 1 and 3 are enriched for "photosynthesis" and "peptidase regulation" gene ontology (GO) terms, respectively ($P < 0.05$; Fig. 1*B* and *SI Appendix*, Table S2). We found that *S. pennellii* orthologs are dominant in cluster 3 (*SI Appendix*, *SI Methods*) and suggest that this species might have unique gene expression at the P5 stage, which involves peptidase regulation. GO enrichment analysis identified genes in cluster 2 as enriched for "transcription"-related GO terms. Genes in this cluster show higher expression in younger leaf morphogenesis stages, which is consistent with a role in leaf development. Given this expression pattern and the importance of transcriptional regulation in generating morphology, we focused on cluster 2 genes to construct a coexpression network. The resultant network is scale-free and reveals three communities (C1–C3) composing a core network (with hub genes with >200 edges; *SI Appendix*, Table S3) connected to two other communities (C4 and C5) on the periphery (cutoff values for Pearson correlation coefficient, adjusted $P < 1.0 \times 10^{-8}$; Fig. 2*A*, *SI Appendix*, Fig. S2, and Dataset S2). The core network has hub genes defined as highly interconnected genes in each network. These hub genes likely function in cell proliferation [including *AINTEGUMENTA* (*ANT*) and *GROWTH REGULATING FACTOR2* (*GRF2*)], chromatin remodeling, and cell growth (18–21) based on annotation in *Arabidopsis thaliana* (Dataset S3). *ANT* (Solyc04g077490) is also the most highly interconnected hub gene in the gene coexpression network constructed using all expressed genes (*SI Appendix*, Fig. S3 and Dataset S4). This supports the current idea that *ANT* genes have central roles in various developmental processes (22). The expression patterns of these hub genes in the core network are significantly negatively correlated with the differences in leaf length between the tomato species (Spearman's rank correlation, adjusted $P < 0.05$; Dataset S5), suggesting that the core network is involved in global cell proliferation and elongation driving leaf growth. Promoters of genes in cluster 2 (Fig. 1*B*) are enriched for E2F and M-specific activator (MSA) element binding sites, which are usually found in cell cycle genes (23, 24) (Fisher's exact test, $P < 0.01$), and a majority of the genes with these binding sites are located in the core network (*SI Appendix*, Fig. S4), further supporting the function of this core network in cell proliferation. Several regulators of leaf complexity such as *GOBLET* (*GOB*) (Solyc07g062840; tomato ortholog *CUP-SHAPED COTYLEDON2*) (25), *LYRATE* (*LYR*) (Solyc05g009380; tomato ortholog *JAGGED*) (26), and auxin efflux carrier *PIN1* (Solyc10g080880) (27) are also placed in the core network.

A majority of developmental genes exhibit broadly conserved function across plant species, and this is particularly true of genes relating to simple and compound leaf development (28). To define key gene network modules (GNMs), we focused on literature-curated genes that are known to regulate leaf and SAM development in the model species *Arabidopsis thaliana* (29–31) and genes whose expression patterns are correlated with these genes in our dataset (LC+ = the literature-curated plus coexpressed genes; *SI Appendix*, Fig. S5 and Datasets S6 and S7). Of a total of 1,329 LC+ genes, 147 are in the SOM cluster 2 generated network (significant enrichment, Fisher's exact test, $P < 1.0 \times 10^{-15}$; *SI Appendix*, *SI Methods* and Fig. S6). Of the peripheral and core communities in the leaf developmental GRN, one peripheral community (C4) shows significant enrichment of literature-curated leaf developmental genes (Fisher's exact test, $P < 0.05$; *SI Appendix*, Table S4). We then extracted a statistically significant leaf GNM by constructing a network weighted by bootstrap selection frequency on each edge (Fig. 2*B*). Utilization of LC+ genes and statistical tests allowed us to capture a key GNM in the leaf developmental GRN. This GNM retained the hub genes (with >25.8 the averaged edges across bootstrapped networks) *PETROSELINUM* (*PTS*) (tomato ortholog *KNATM*, Solyc06g072480), *KLUH* (*KLU*) (Solyc03g114940), and *VERNALIZATION1* (*VRN1*) (Solyc02g021260) of the community C4 (Fig. 2*B*) and has *BLADE-ON-PETIOLE* (Solyc04g040220; designated as *BOPa*) and *LIGHT-DEPENDENT SHORT HYPOCOTYLS* (*LSH3a/b* and *LSH6*, Solyc06g083860, Solyc09g025280, and Solyc09g090180, respectively) that are annotated to function in leaf development of *Arabidopsis thaliana* (29–31), as well as the *gibberellin 2-oxidase4* (*GA2OX4*) (Solyc01g058030) ortholog that is responsive to KNOX activity (32). Two genes in this GNM, *LSH6* (*TERMINATING FLOWER, TMF*) and *KLU* (underlies the *fruit weight3.2* locus), were recently cloned from tomato and shown to affect vegetative/reproductive transition and lateral organ growth, respectively (33, 34). *PTS* was earlier described as a key regulator of natural variation in leaf complexity in the Galapagean tomatoes and regulates *KNOX LeT6* (tomato ortholog of *SHOOT MERISTEMLESS*) at the protein level by competing with BEL1-LIKE HOMEODOMAIN (BELL)/BIPINNATA (BIP) and preventing entry of the LeT6/BIP complex into the nucleus (Fig. 2*C*) (2). Thus, the *PTS*-containing GNM, identified using transcriptional data, is not expected to include *KNOX* and *BIP*. However, the output of this GNM should include potential targets of KNOX, and *GA2OX4* is one such functionally validated target (32) (Fig. 2*B*). Because of its central role in this module and in regulating leaf development, hereafter we designate this GNM as the *PTS* GNM.

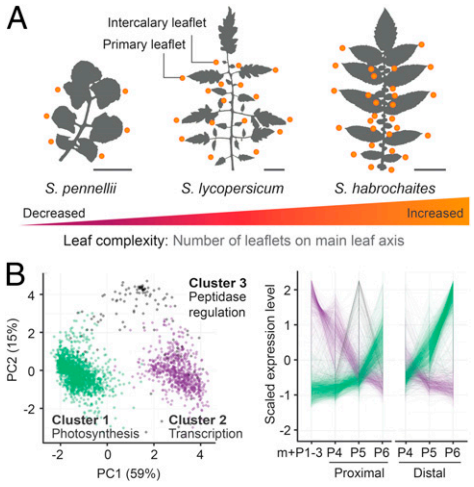

Fig. 1. Leaf complexity and transcriptome profiling of tomato and its wild relatives. (*A*) Mature leaves from *S. pennellii*, *S. lycopersicum*, and *S. habrochaites*. The dots indicate leaflets on the main leaf axis; the sum of them is here defined as leaf complexity. (Scale bars: 5 cm.) (*B*) PCA with SOM clustering of gene expression. The expression profile of each gene is represented, and genes belonging to different clusters are indicated by color (clusters 1–3 are green, purple, and black, respectively) and separated by PCs (*Left*), and their scaled expression patterns plotted during four stages of leaf development are separated by proximal and distal regions (*Right*).

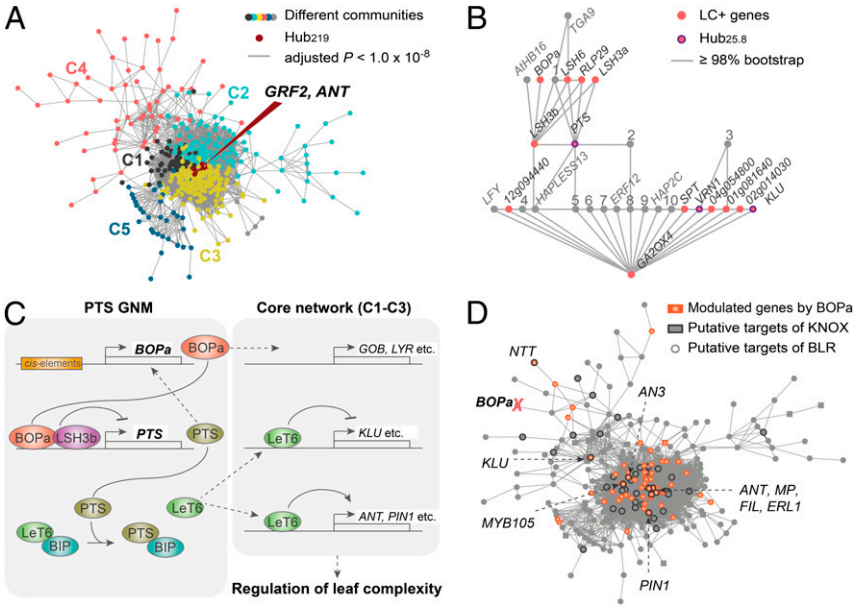

Fig. 2. Prediction and validation of the leaf developmental GRN. (*A*) Gene coexpression network for the genes in cluster 2 in Fig. 1*B*. Nodes represent genes. Only nodes with at least one edge (403 nodes and 18,964 total edges) are represented. The five communities determined by the Fast Greedy modularity optimization algorithm are represented by different colored nodes (C1–C5). The dark red nodes represent hub genes (>219 edges). (*B*) The *PETROSELINUM* (*PTS*) GNM, defined by the literature-curated plus coexpressed (LC+) genes and high confidence strong interactions (≥98% selection frequencies) in community C4 of Fig. 2*A*. *PTS* was earlier described as a key regulator of natural variation in leaf complexity in the Galapagean tomatoes (2). Hub genes are determined by the averaged connectivity of this gene across the bootstrapped networks (>25.8). Genes of unknown function are shown by numbers: 1, Solyc07g065920; 2, Solyc04g074810; 3, Solyc12g044920; 4, Solyc07g049400; 5, Solyc10g081700; 6, Solyc03g120870; 7, Solyc07g006380; 8, Solyc08g078890; 9, Solyc01g108170; 10, Solyc07g055100. (*C*) Model of leaf developmental GRN under *PTS* GNM regulation. In the *PTS* GNM, *BOPa* gene expression is *cis*-regulated. BOPa interacts with LSH3b and the complex directly regulates *PTS* expression. *PTS* regulates LeT6 at the protein level by competing with BIP. The released LeT6 regulates many genes in the core network (C1–C3) that influence leaf complexity. *PTS* shows feedback regulation of *BOPa* gene expression. *BOPa* also regulates genes in the core network independently of *PTS*. This GRN regulates leaf complexity as a phenotypic output. (*D*) *BOPa* (marked with a red cross), genes whose expression can be modulated by *BOP* (orange color), putative direct targets of KNOX homeobox gene (black outlines), genes with binding site of BELLRINGER (BLR) (circles) are mapped on the gene coexpression network from *A*. Genes lacking BLR sites are represented as squares. Genes with all of the above features are named.

Network Validation. To experimentally validate the bioinformatically predicted gene interactions within the *PTS* GNM (Fig. 2*B*), we analyzed interactions between key genes in the GNM. Expression differences in *PTS* between species of Galapagean tomatoes positively correlate with leaf complexity, and are shown to be caused by *PTS* promoter alterations (2). Our GNM showed a connection between *PTS* and *LSH3b.* We tested this prediction and found that LSH3b directly binds to the *PTS* promoter (*SI Appendix*, Fig. S7). The PTS promoter fragment used has three repeats of ACATTTTT and sequences upstream of the repeats. We found that the upstream region is necessary, and the first ACATTTTT is necessary but not sufficient for the interaction with LSH3b (*SI Appendix*, Fig. S7). In addition, we validated the predicted BOPa/LSH3b interaction via yeast two-hybrid assays (*SI Appendix*, Fig. S8), consistent with the finding that interacting proteins are more likely to coexpress than noninteracting proteins (35). Previous studies in inflorescence development in tomato have shown BOPa and LSH6 interactions (33). Our experiments suggest that BOPa can physically interact with LSH, and the complex may directly regulate *PTS* expression, leading to alterations in expression of genes within the *PTS* GNM as well as in the core network (i.e., C1–C3, Fig. 2*C*). Because *PTS* regulates *KNOX* at the protein level (2), *BOPa* might affect *KNOX* targets via regulation of *PTS* expression. To examine this idea, we identified genes whose expression can be modulated by *BOPa* using RNA-seq of *BOPa* knockdown and overexpression transgenic tomato lines (*SI Appendix*, Figs. S9 and S10 and Table S5 and Datasets S8–S10). In total, 413 and 210 genes are regulated positively and negatively by *BOPa*, respectively (false discovery rate < 0.05); these genes are overrepresented in SOM cluster 2 (Fisher's exact test, $P < 1.0 \times 10^{-15}$) (*SI Appendix*, Table S6). To further validate the GNM, we confirmed that *BOPa* regulates *PTS* and *KLU* gene expression (Datasets S9 and S10). Using wild tomato species, we show that *PTS* regulates *KLU* and *GA2OX4* expression, and feedback regulation of *BOPa* gene expression (*SI Appendix*, Figs. S11 and S12). In addition, we found that *BOPa* regulates a number of genes in the core network including the hub genes *GRF2* and *ANT*, as well as leaf complexity regulators *GOB*, *LYR*, and *PIN1* (Fig. 2*D* and Datasets S9 and S10). Consistent with the idea that *BOPa* may modulate *KNOX* targets, we found that putative KNOX direct targets (36) were significantly enriched in the genes modulated in expression by *BOP* (Fisher's exact test, $P < 0.05$; Fig. 2*D* and Dataset S11). *KLU*, *PIN1*, and the hub gene *ANT* were putative KNOX direct targets under *BOP* regulation (Fig. 2*D*). In addition, the promoters of potential targets of *BOPa* were enriched for homeobox gene BELLRINGER binding site 1 and 3 (Fisher's exact test, $P < 0.05$; Fig. 2*D*). Our RNA-seq data showed that *BOPa* regulates the expression of four *BELL* genes (Solyc04g080780, Solyc08g081400, Solyc10g086640, and Solyc11g069890; Datasets S9 and S10). Thus, *BOPa* controls the expression of both *BELL*-type homeobox genes and *KNOX* targets via *BELL* and *PTS* level modulation, a fact borne out by the enrichment of KNOX and BELL targets in the leaf developmental GRN. Our molecular experiments and hypothesis testing validate the bioinformatically predicted *PTS* GNM, and suggest that this GNM could explain leaf developmental variation across the tomato species complex (Fig. 2*C*).

Dynamic Rewiring of Interactions in the GRN Across Tomato Species. To determine the causal differences in the leaf developmental GNM that account for the interspecific morphological variation in *Solanum* leaf complexity, we constructed gene coexpression networks based on bootstrapped datasets for each species separately (Fig. 3*A*) using the genes in SOM cluster 2 (Fig. 1*B*). Connectivity (total number of edges of the network), modularity (based on Fast Greedy modularity optimization algorithm; how modular is a given partition of a network into sub networks), and centralization (based on degree centralization; the number of edges incident upon a gene) were calculated to define network properties. We found that these network properties were significantly different between the three species (Fig. 3*A*), suggesting a dynamic rewiring of interactions in the leaf developmental GRNs (i.e., defined by SOM cluster 2) between the tomato species. Because the *PTS* GNM regulates many genes in the leaf developmental GRN (Fig. 2*C*), we examined detailed network properties of the *PTS* GNM across the species. Consistent with proposed hypotheses that complex developmental traits require installation of additional regional network modules (5), the *PTS* GNM showed dramatic differences in the number of

intramodular hub genes (with more than four averaged edges across bootstrapped networks) across species and this number correlated with the degree of leaf complexity (Fig. 3*B*). Furthermore, although expression changes in most genes in the *PTS* GNM are not correlated with leaf phenotypes, the expression levels of *BOPa* and *PTS* respectively correlated negatively and positively with the degree of leaf complexity among the three species (Fig. 3*C* and *SI Appendix*, Fig. S13). Our molecular data (*SI Appendix*, Figs. S7 and S8) show that *BOPa* regulates *PTS* as an upstream factor, whereas *PTS* has feedback regulation on *BOPa* gene expression (*SI Appendix*, Fig. S12). It is therefore likely that structural changes in the *PTS* GNM, coordinated by *BOPa-PTS* gene expression, have altered expression/interactions in the leaf developmental GRN during the course of evolution of leaf complexity in these species. Such structural changes in human and chimpanzee brain GRNs were shown to account for differences in brain organization in the two species (37). Although hub genes could be internally modular (e.g., have multiple *cis*-regulatory regions and functional domains), and tweaks in the expression patterns of the hub input by the peripheral genes can be important for evolution, mutations in genes located in the peripheral regions of GRNs have greater potential to contribute to evolutionary innovations because mutations that disrupt hubs may have more drastic phenotypic changes and difficulty propagating in the population (38). Thus, network changes that might influence the evolution of leaf morphology in the tomato species likely involve genes in the peripheral GNMs (like the *PTS* GNM) and not the core network (i.e., C1–C3). Because changes in *PTS* expression itself can cause such leaf morphological variation in Galapagean tomatoes (2), we decided to investigate *BOP*, a transcription factor upstream of *PTS*, to determine its function in evolutionary network rewiring in the *Solanum* species.

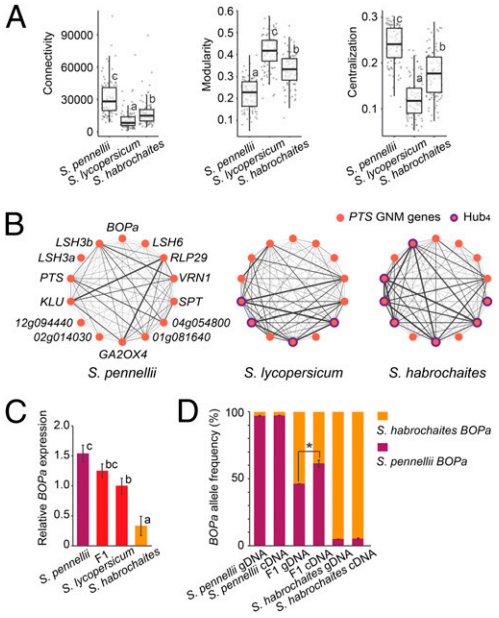

Fig. 3. Comparative GRNs regulating leaf development in different tomato species. (*A* and *B*) Comparison of network properties between gene coexpression networks of different tomato species. Gene coexpression networks built on bootstrapped data sets using the data from the proximal (leaflet forming) region of P4, P5, and P6 leaf developmental stages for genes in cluster 2 (Fig. 1*B*) were compared. Connectivity, modularity, and centralization in GRN of different tomato species are shown (*A*). Error bars indicate SE over 100 bootstrapped replicates. Different letters indicate significant differences between networks determined by Dunnett's test ($P < 0.001$). *PTS* GNM of different tomato species are shown in *B*. Hub genes are determined by the averaged connectivity of this gene across the bootstrapped networks (>4), and the number of hub genes varies across the species networks. (*C*) qRT-PCR analysis of *BOPa* expression in *S. lycopersicum*, *S. pennellii*, *S. habrochaites*, and the F1 hybrid between *S. pennellii* and *S. habrochaites*. Error bars indicate SE over three biological replicates (each replicate, 6–10 pooled seedlings). Different letters indicate significant differences determined by Dunnett's test ($P < 0.05$). (*D*) Allele-specific expression assays using pyrosequencing. F1 hybrids between *S. pennellii* and *S. habrochaites* and parental species were tested for the percentage of SNPs on *BOPa*. The significant deviation of allele frequency in cDNA from gDNA in F1 ($P < 0.01$, Mann–Whitney *U* test) indicates *cis* regulation in *BOPa* (detail explanation in *SI Appendix*, Fig. S17). Error bars indicate SE over three biological replicates (each replicate, 6–10 pooled seedlings).

The tomato genome has three *BOP* orthologs, but our network analysis specifically includes only one of these (*BOPa*) in the *PTS* GNM (Fig. 2*B*), and the expression of *BOPa* alone is significantly different between species (*SI Appendix*, Fig. S14). BOPa protein sequences of tomato and its wild relatives show high similarity (over 95%; *SI Appendix*, Figs. S15 and S16). Thus, regulatory changes in *BOPa* expression, rather than protein structural changes, might be more relevant to leaf morphological variation. To test whether *cis*-regulatory changes in *BOP* were important, we used an F1 hybrid between *S. pennellii* and *S. habrochaites* (showing intermediate levels of both leaf complexity and *BOPa* expression; Fig. 3*C* and *SI Appendix*, Fig. S17) to look at allele-specific expression. Despite their sharing the same *trans*-acting factors in F1 hybrid individuals, we found a higher level of *S. pennellii BOPa* than *S. habrochaites BOPa* transcripts in F1 hybrid seedlings (Fig. 3*D* and *SI Appendix*, Fig. S17). Thus, differences in *cis*-regulation of *BOPa* appear important for regulating gene expression between these species. Characterization of *cis*/*trans* regulation at the genome-wide level should provide further insights into regulation of gene expression in the F1 generation. Collectively, our evolutionary developmental transcriptomics provides arguments to formulate the hypothesis that the evolutionary changes in *BOPa* gene expression trigger dynamic rewiring of interactions in the leaf developmental GRN through regulation of the hub gene *PTS* in the *PTS* GNM, eventually generating the morphological diversity seen in the tomato species complex.

Transcriptional Regulation of *BOP* Generates Leaf Morphological Diversity. To test the hypothesis that variation in expression of genes in the periphery of a GRN contributes to morphological variation (38), we manipulated *BOPa* gene expression by genetic transformation in the selected tomato species. The high degree of sequence similarity between the three BOP paralogs (over 89%; *SI Appendix*, Fig. S15) does not allow individual down-regulation experiments. We generated transgenic plants of tomato *S. lycopersicum* in which *BOP* expression was reduced to less than 40% by RNA interference (RNAi) (*SI Appendix*, Fig. S10). These transgenic plants show an increase in leaf complexity (number of leaflets on the main leaf axis; Fig. 4*A* and *SI Appendix*, Table S7) as well as defects in reproductive traits including bracteate inflorescences (*SI Appendix*, Fig. S18), mimicking the leaf complexity and inflorescence phenotypes of *S. habrochaites*, which shows low *BOPa* expression (Fig. 3*C*). Next, to recapitulate the effects of high *BOPa* expression seen in *S. pennellii* (Fig. 3*C*), we overexpressed *BOPa* (and separately, *BOPc*) in *S. lycopersicum*. Both expression manipulations led to a decrease in leaf complexity (Fig. 4*B* and *SI Appendix*, Tables S8 and S9). These results establish the role of *BOP* in repressing leaflet formation. We compared lines overexpressing the *S. lycopersicum BOPa* and *S. pennellii BOPa*, and saw similar leaf phenotypes (*SI Appendix*, Fig. S19), indicating that the differences in BOPa protein sequences among tomato and *S. pennellii* have little effect on leaf phenotype. Thus, transcriptional regulation of *BOPa* levels, rather than the sequence of the BOPa protein, regulates the morphospace of leaf complexity in the tomato species.

Next, we manipulated *BOP* gene expression in *S. pennellii* and *S. habrochaites* to determine whether the phenotypic effects of altering *BOP* transcript levels in the wild relatives approximates the morphology of the cultivated species. Due to high native *BOPa*

expression in *S. pennellii*, we down-regulated the expression of the *BOP* genes by RNAi (*SI Appendix*, Fig. S10). The transgenic plants show increased leaf complexity (Fig. 4*C* and *SI Appendix*, Table S10). Because of low *BOPa* expression in *S. habrochaites*, we overexpressed *BOPa* in this species, but we could not obtain any transformants. Because *BOPa* and *BOPc* share similar functions in leaf complexity regulation in tomato (*SI Appendix*, Tables S8 and S9), instead we overexpressed *S. lycopersicum BOPc* and *S. habrochaites BOPc* in the *S. habrochaites* background. Both transgenic lines showed decreased leaf complexity (Fig. 4*D* and *SI Appendix*, Fig. S20). All *BOP* transgenic plants, regardless of species, showed changes in the number of leaflets, not leaflet shape (Fig. 4 *A–D*). Given that LSH3b interacts with BOPa and also binds to the *PTS* promoter, we queried the effects of down-regulation of *LSH3b* in transgenic plants and showed that these have increased leaf complexity, consistent with the function of the BOP/LSH3b complex in down-regulation of *PTS* expression (Fig. 2*C* and *SI Appendix*, Fig. S21). Thus, *BOP*, and by extension the *PTS* GNM, might act as a genetic switch in leaflet initiation in a species-dependent context.

Window of Morphogenetic Competence in the Leaf Maturation Schedule. Finally, we investigated how the *PTS* GNM might regulate leaflet development leading to the control of leaf complexity. Accumulation of auxin, via transport by the auxin efflux carrier PIN1, delineates leaflet position (27). *BOP* regulates the transcription of *PIN1* as well as other auxin regulators such as *ANT* and *MONOPTEROS*, which are placed on the core network (Dataset S10). We expressed *AtPIN1:PIN1-GFP* as a marker of auxin response (27) in *BOP RNAi* tomato. Although wild-type and *BOP RNAi* plants showed a similar distribution of GFP at early leaf developmental stages, the P7 leaf primordia of *BOP RNAi* plants, unlike wild type, continued to develop auxin foci as evidenced by GFP between the developing leaflets (Fig. 4*E*). We also expressed a synthetic auxin response element *pDR5rev:3XVENUS-N7* (39) in *BOP* RNAi tomato and observed a similar readout of auxin foci compared with those seen in the *AtPIN1:PIN1-GFP* plants (*SI Appendix*, Fig. S22). In addition, the tomato *entire* mutant, which fails to down-regulate auxin response between developing leaflets (27), completely suppresses the iterative pattern of leaflet formation in the *BOP* RNAi background (*SI Appendix*, Fig. S23). Thus, the modulation of leaf complexity by *BOP* relies on existing patterns of auxin distribution. Even though we drove *BOP* RNAi and overexpression using the *Cauliflower Mosaic Virus* 35S promoter, which expresses ectopically and constitutively throughout the plant body, including all leaf cells (40), the spatial iterative pattern of the leaflets was still maintained in these transgenic plants (Fig. 4 *A–D*). Taken together, these data suggest that the *PTS* GNM could have a role in defining the temporal window of morphogenetic competence specified by auxin, without changing the spatial distribution of auxin, to modulate leaf complexity.

Given that the *PTS* GNM regulates effective KNOX concentration, this observation is consistent with the classical leaf maturation schedule model, which states that *KNOX* influences the temporal progression of a leaf cell through developmental states (8, 41). This agrees with our morphometric analyses of leaf development in the selected tomato species: the duration of leaf maturation reflects their leaf complexity (*SI Appendix*, Fig. S1). To further confirm the idea, we expressed *pDR5rev:3XVENUS-N7* in the three species (*SI Appendix*, Fig. S24). P4 leaf primordia of *S. pennellii* maintained only one pair of auxin foci marking the most proximal pair of leaflet sites and their terminal leaflet was expanded. P4 leaf primordia of *S. lycopersicum* maintained two pairs of auxin foci marking the two proximal pair of leaflet sites and were still developing the terminal leaflet vasculature. P4 leaf primordia of *S. habrochaites* continued to develop three or four pairs of auxin foci at the leaf base marking leaflet initiation and showed delayed terminal leaflet expansion with developing lobes. Thus, auxin response is correlated with the differential leaf maturation seen in the three tomato species, further corroborating the role of differential *BOP* expression in setting the temporal window of morphogenetic competence specified by auxin (Fig. 4*F*).

Conclusion

We have shown that evolutionary transcriptomics and network construction have predictive power to identify GNM regulating morphological evolution. Dynamic rewiring of interactions in the leaf developmental GRN under *PTS* GNM regulation is causal in generating evolutionary morphological diversity in leaf complexity. This is mediated by altering the window of morphogenetic competence in the leaf maturation schedule. Considering that previous quantitative genetics revealed polygenic contributions to the natural variation in *Solanum* leaf shape (12, 42), other modules might also underlie leaf shape evolution through modification of the leaf developmental GRN. Similar to homeobox genes contributing to animal body plan evolution (43), our data support the idea that the regulation of homeobox genes such as *KNOX* was recruited repeatedly to influence leaf diversity (2, 7–10). This regulation can use multiple inputs, including promoter variation at *KNOX* (10), alterations in effective KNOX concentration (2), alterations in *KNOX* expression patterns (7), as well as modulation of the entire GRN regulating *KNOX* (this study). Interestingly, we found that *KNOX* serves as a bridge connecting a peripheral GNM to the core network within the leaf developmental GRN. This finding suggests that *KNOX* is repeatedly coopted to generate plant morphological diversity by virtue of its bottleneck location in the GRN. We also show that evolutionary

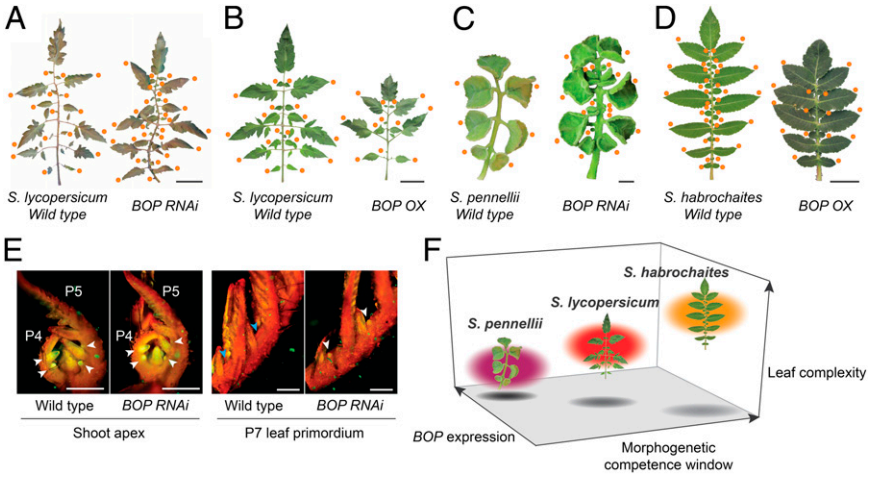

Fig. 4. Evolutionary role of the *BOP* gene in generation of leaf complexity. (*A–D*) Mature leaves of a wild-type and *BOP RNAi S. lycopersicum* (*A*), wild-type and *BOPa* overexpression *S. lycopersicum* (*B*), wild-type and *BOP RNAi S. pennellii* (*C*), wild-type and *BOPc* overexpression *S. habrochaites* (*D*). The dots indicate leaflets on the main leaf axis, which characterizes leaf complexity. (*E*) Distribution of the *AtPIN1:PIN1-GFP* marker in shoot apices and P7 leaf primordia of wild-type and *BOP RNAi* tomato. The white arrowheads indicate GFP expression in initiating leaflets. The blue arrowheads indicate the absence of GFP expression in the leaf primordium of wild type. [Scale bars: 5 cm (*A*, *B*, and *D*); 1 mm (*C*); 500 μm (*E*).] (*F*) Schematic diagram representing the leaf morphospace in the tomato species complex with three axes: *BOP* expression, the morphogenetic competence window, and leaf complexity.

changes in peripheral GNMs can generate dynamic but fine-tuned rewiring of interactions in the whole GRN, which may have been exploited to create morphological diversity during the course of evolution.

Materials and Methods

We made RNA-seq libraries using a high-throughput protocol (44) and sequenced the libraries on Illumina HiSEq 2000 platform. The reads of the three species were mapped to matched sets of reference cDNAs (17). Read-mapped datasets are provided as Datasets S1 and S8. We performed PCA with SOM clustering and GO enrichment analysis to select leaf developmental genes. For these genes, we constructed gene coexpression networks with the cutoff values for Pearson correlation coefficient (adjusted $P < 1.0 \times 10^{-8}$) to capture known interactions. EMSAs to assess the binding properties of LSH3 to *PTS* promoter were performed with the LightShift Chemiluminescent EMSA Kit (Pierce). Protein interactions were assayed using the GAL4 yeast two-hybrid system (Clontech). To validate gene expression, quantitative RT-PCR (qRT-PCR) analysis was performed and normalized to the reference gene Solyc04g064820 and/or Solyc03g111090 originally selected in this study for comparison across *Solanum* species. To compare networks of different species, we constructed bootstrapped networks (cutoff values for Pearson correlation coefficient, adjusted $P < 0.01$) based on 100 bootstrapped replications. All data analysis was performed on the iPlant cyberinfrastructure platform. *BOP* and *LSH* expression was manipulated by genetic transformation. Conserved sequences among three *BOP* genes inserted in *RNAi* constructs for knockdown and whole coding sequences of *BOP* genes for overexpression were driven by the ectopic and constitutive 35S promoter. Plasmids were transformed into *S. lycopersicum*, *S. pennellii*, and *S. habrochaites* at the Ralph M. Parsons Foundation Plant Transformation Facility (University of California, Davis, CA). The methods are detailed in *SI Appendix*. For primer sequences, see Dataset S12.

ACKNOWLEDGMENTS. We thank C. Kuhlemeier for providing the *AtPIN1:PIN1-GFP* tomato line; N. Ori for providing *pDR5rev:3XVENUS-N7* pART27 construct; D. Koenig, J. M. Jiménez-Gómez, A. Tat, M. F. Covington, U. K. Devisetty, and P. Y. Hsu for guidance in bioinformatics; D. M. Tricoli and K. J. Carney for insights in transformation in *S. pennellii* and *S. habrochaites*; T. Akagi, B. Townsley, and D. Fulop for guidance in phylogenetic and molecular evolution analyses; A. Ranjan for providing cDNA of developing leaves from *Solanum* species; C. C. Martinez for guidance in fluorescent microscopy; S. Ramireddy, L. Rynearson, J. Kao, and P. Starnes for plant care and technical assistance; and Tomato Genetics Resources Centre (University of California, Davis) for providing the seeds for the tomato accessions (http://tgrc.ucdavis.edu/). K. Nozue, J. M. Budke, D. West, and S. D. Rowland are also thanked for the critical comments on this project. This work was funded through National Science Foundation Grant IOS-0820854 (awarded to N.R.S., J.N.M., and J.P.), a postdoctoral fellowship from the Spanish Ministry of Science and Innovation (to J.A.A.-M.), and Research Fellow of Japan Society for the Promotion of Science (to Y.I.). This work used computational resources or cyberinfrastructure provided by the iPlant Collaborative (www.iplantcollaborative.org). The iPlant Collaborative is funded by National Science Foundation Grant DBI-0735191.

1. Blein T, et al. (2008) A conserved molecular framework for compound leaf development. *Science* 322(5909):1835–1839.
2. Kimura S, Koenig D, Kang J, Yoong FY, Sinha N (2008) Natural variation in leaf morphology results from mutation of a novel KNOX gene. *Curr Biol* 18(9):672–677.
3. Loehlin DW, Werren JH (2012) Evolution of shape by multiple regulatory changes to a growth gene. *Science* 335(6071):943–947.
4. Rebeiz M, Pool JE, Kassner VA, Aquadro CF, Carroll SB (2009) Stepwise modification of a modular enhancer underlies adaptation in a *Drosophila* population. *Science* 326(5960):1663–1667.
5. Peter IS, Davidson EH (2011) Evolution of gene regulatory networks controlling body plan development. *Cell* 144(6):970–985.
6. Fischer AH, Smith J (2012) Evo-devo in the era of gene regulatory networks. *Integr Comp Biol* 52(6):842–849.
7. Bharathan G, et al. (2002) Homologies in leaf form inferred from KNOXI gene expression during development. *Science* 296(5574):1858–1860.
8. Hay A, Tsiantis M (2010) KNOX genes: Versatile regulators of plant development and diversity. *Development* 137(19):3153–3165.
9. Nakayama H, Yamaguchi T, Tsukaya H (2012) Acquisition and diversification of cladodes: Leaf-like organs in the genus *Asparagus*. *Plant Cell* 24(3):929–940.
10. Piazza P, et al. (2010) *Arabidopsis thaliana* leaf form evolved via loss of KNOX expression in leaves in association with a selective sweep. *Curr Biol* 20(24):2223–2228.
11. Chitwood DH, et al. (2012) The developmental trajectory of leaflet morphology in wild tomato species. *Plant Physiol* 158(3):1230–1240.
12. Holtan HE, Hake S (2003) Quantitative trait locus analysis of leaf dissection in tomato using *Lycopersicon pennellii* segmental introgression lines. *Genetics* 165(3):1541–1550.
13. Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485(7400):635–641.
14. Park SJ, Jiang K, Schatz MC, Lippman ZB (2012) Rate of meristem maturation determines inflorescence architecture in tomato. *Proc Natl Acad Sci USA* 109(2):639–644.
15. Koenig D, et al. (2013) Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato. *Proc Natl Acad Sci USA* 110(28):E2655–E2662.
16. Rodriguez F, Wu F, Ané C, Tanksley S, Spooner DM (2009) Do potatoes and tomatoes have a single evolutionary history, and what proportion of the genome supports this history? *BMC Evol Biol* 9:191.
17. Chitwood DH, Maloof JN, Sinha NR (2013) Dynamic transcriptomic profiles between tomato and a wild relative reflect distinct developmental architectures. *Plant Physiol* 162(2):537–552.
18. Mizukami Y, Fischer RL (2000) Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis. *Proc Natl Acad Sci USA* 97(2):942–947.
19. Gendrel AV, Lippman Z, Yordan C, Colot V, Martienssen RA (2002) Dependence of heterochromatic histone H3 methylation patterns on the Arabidopsis gene DDM1. *Science* 297(5588):1871–1873.
20. Minic Z, Jouanin L (2006) Plant glycoside hydrolases involved in cell wall polysaccharide degradation. *Plant Physiol Biochem* 44(7-9):435–449.
21. Kim JH, Choi D, Kende H (2003) The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in *Arabidopsis*. *Plant J* 36(1):94–104.
22. Horstman A, Willemsen V, Boutilier K, Heidstra R (2014) AINTEGUMENTA-LIKE proteins: Hubs in a plethora of networks. *Trends Plant Sci* 19(3):146–157.
23. Chabouté ME, Clément B, Sekine M, Philipps G, Chaubet-Gigot N (2000) Cell cycle regulation of the tobacco ribonucleotide reductase small subunit gene is mediated by E2F-like elements. *Plant Cell* 12(10):1987–2000.
24. Ito M, et al. (1998) A novel *cis*-acting element in promoters of plant B-type cyclin genes activates M phase-specific transcription. *Plant Cell* 10(3):331–341.
25. Berger Y, et al. (2009) The NAC-domain transcription factor GOBLET specifies leaflet boundaries in compound tomato leaves. *Development* 136(5):823–832.
26. David-Schwartz R, Koenig D, Sinha NR (2009) LYRATE is a key regulator of leaflet initiation and lamina outgrowth in tomato. *Plant Cell* 21(10):3093–3104.
27. Koenig D, Bayer E, Kang J, Kuhlemeier C, Sinha N (2009) Auxin patterns *Solanum lycopersicum* leaf morphogenesis. *Development* 136(17):2997–3006.
28. Townsley BT, Sinha NR (2012) A new development: Evolving concepts in leaf ontogeny. *Annu Rev Plant Biol* 63:535–562.
29. Ha CM, et al. (2003) The BLADE-ON-PETIOLE 1 gene controls leaf pattern formation through the modulation of meristematic activity in *Arabidopsis*. *Development* 130(1):161–172.
30. Takeda S, et al. (2011) CUP-SHAPED COTYLEDON1 transcription factor activates the expression of LSH4 and LSH3, two members of the ALOG gene family, in shoot organ boundary cells. *Plant J* 66(6):1066–1077.
31. Anastasiou E, et al. (2007) Control of plant organ size by KLUH/CYP78A5-dependent intercellular signaling. *Dev Cell* 13(6):843–856.
32. Jasinski S, et al. (2005) KNOX action in *Arabidopsis* is mediated by coordinate regulation of cytokinin and gibberellin activities. *Curr Biol* 15(17):1560–1565.
33. MacAlister CA, et al. (2012) Synchronization of the flowering transition by the tomato TERMINATING FLOWER gene. *Nat Genet* 44(12):1393–1398.
34. Chakrabarti M, et al. (2013) A cytochrome P450 regulates a domestication trait in cultivated tomato. *Proc Natl Acad Sci USA* 110(42):17125–17130.
35. Kemmeren P, et al. (2002) Protein interaction verification and functional annotation by integrated analysis of genome-scale data. *Mol Cell* 9(5):1133–1143.
36. Bolduc N, et al. (2012) Unraveling the KNOTTED1 regulatory network in maize meristems. *Genes Dev* 26(15):1685–1690.
37. Oldham MC, Horvath S, Geschwind DH (2006) Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *Proc Natl Acad Sci USA* 103(47):17973–17978.
38. Vidal M, Cusick ME, Barabási AL (2011) Interactome networks and human disease. *Cell* 144(6):986–998.
39. Heisler MG, et al. (2005) Patterns of auxin transport and gene expression during primordium development revealed by live imaging of the *Arabidopsis* inflorescence meristem. *Curr Biol* 15(21):1899–1911.
40. Zoulias N, Koenig D, Hamidi A, McCormick S, Kim M (2012) A role for PHANTASTICA in medio-lateral regulation of adaxial domain development in tomato and tobacco leaves. *Ann Bot (Lond)* 109(2):407–418.
41. Freeling M (1992) A conceptual framework for maize leaf development. *Dev Biol* 153(1):44–58.
42. Chitwood DH, et al. (2013) A quantitative genetic basis for leaf morphology in a set of precisely defined tomato introgression lines. *Plant Cell* 25(7):2465–2481.
43. Pick L, Heffer A (2012) Hox gene evolution: Multiple mechanisms contributing to evolutionary novelties. *Ann N Y Acad Sci* 1256:15–32.
44. Kumar R, et al. (2012) A high-throughput method for Illumina RNA-Seq library preparation. *Front Plant Sci* 3:202.