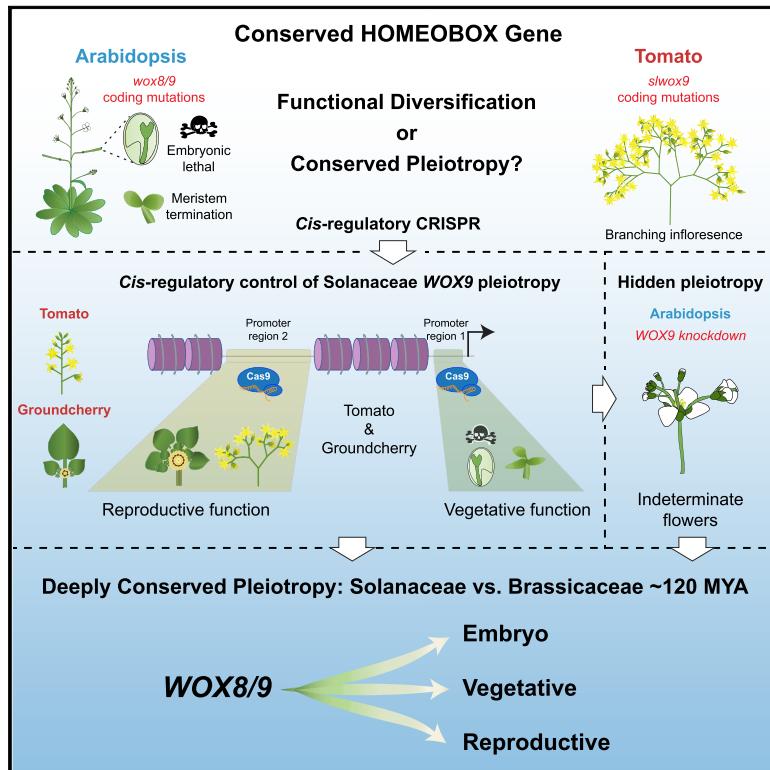


Conserved pleiotropy of an ancient plant homeobox gene uncovered by *cis*-regulatory dissection

Graphical Abstract



Authors

Anat Hendelman, Sophia Zebell,
Daniel Rodriguez-Leal, ..., Yuval Eshed,
Idan Efroni, Zachary B. Lippman

Correspondence

idan.efroni@mail.huji.ac.il (I.E.),
lippman@cshl.edu (Z.B.L.)

In brief

Functional dissections of *cis*-regulatory DNA suggest a conserved hidden pleiotropy widespread in plant evolution that has implications for the utilization of genome editing in agriculture

Highlights

- Targeted *cis*-regulatory mutagenesis uncovered hidden pleiotropy in a plant homeobox gene
- Different pleiotropic functions can be mapped to specific *cis*-regulatory regions
- Pleiotropy and *cis*-regulatory control sequences are conserved in related species
- Pleiotropy, but not *cis*-regulatory structure, is conserved in distantly related plants



Article

Conserved pleiotropy of an ancient plant homeobox gene uncovered by *cis*-regulatory dissection

Anat Hendelman,¹ Sophia Zebell,^{1,2} Daniel Rodriguez-Leal,^{1,7} Noah Dukler,¹ Gina Robitaille,^{1,2} Xuelin Wu,³ Jamie Kostyun,⁴ Lior Tal,^{5,8} Peipei Wang,⁶ Madelaine E. Bartlett,⁴ Yuval Eshed,⁵ Idan Efroni,^{6,*} and Zachary B. Lippman^{1,2,9,*}

¹Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA

²Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA

³The Salk Institute for Biological Research, San Diego, CA, USA

⁴Biology Department, University of Massachusetts Amherst, Amherst, MA, USA

⁵Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot, Israel

⁶Institute of Plant Sciences and Genetics in Agriculture, The Robert H. Smith Faculty of Agriculture, The Hebrew University, Rehovot, Israel

⁷Present address: Inari Agriculture, Cambridge, MA, USA

⁸Present address: Plant Biology Department, University of California, Davis, Davis, CA, USA

⁹Lead contact

*Correspondence: idan.efroni@mail.huji.ac.il (I.E.), lippman@cshl.edu (Z.B.L.)

<https://doi.org/10.1016/j.cell.2021.02.001>

SUMMARY

Divergence of gene function is a hallmark of evolution, but assessing functional divergence over deep time is not trivial. The few alleles available for cross-species studies often fail to expose the entire functional spectrum of genes, potentially obscuring deeply conserved pleiotropic roles. Here, we explore the functional divergence of *WUSCHEL* HOMEobox9 (*WOX9*), suggested to have species-specific roles in embryo and inflorescence development. Using a *cis*-regulatory editing drive system, we generate a comprehensive allelic series in tomato, which revealed hidden pleiotropic roles for *WOX9*. Analysis of accessible chromatin and conserved *cis*-regulatory sequences identifies the regions responsible for this pleiotropic activity, the functions of which are conserved in groundcherry, a tomato relative. Mimicking these alleles in *Arabidopsis*, distantly related to tomato and groundcherry, reveals new inflorescence phenotypes, exposing a deeply conserved pleiotropy. We suggest that targeted *cis*-regulatory mutations can uncover conserved gene functions and reduce undesirable effects in crop improvement.

INTRODUCTION

Divergence of gene function, shaped by mutation, gene duplication, and gene loss, is a major force in evolution (Holland et al., 2017; Panchy et al., 2016; Rebeiz and Tsiantis, 2017; Stern, 2000). However, assessment of gene functional divergence over deep time is challenging as large evolutionary distances make it difficult to distinguish between divergence of gene function and conserved function acting in different developmental contexts (Shubin et al., 2009; Stern, 2000). A major limitation in cross-species comparative genetic studies is the limited number of available alleles arising either from fortuitous natural mutations or from mutagenesis experiments (Paaby and Rockman, 2013; Rockman, 2012; Visscher and Yang, 2016; Wang et al., 2010). Such studies fail to capture the full breadth of gene function, often exposing only part of a gene's pleiotropic functions (Tyler et al., 2016; Wagner and Zhang, 2011), which could lead to misattribution of divergence. Furthermore, severe mutations,

such as those that cause embryonic lethality, can mask gene functions due to developmental epistasis, whereby an early phenotypic defect precludes evaluation of later phenotypes (Paaby and Rockman, 2013). In many cases, it remains unclear to what extent orthologs have undergone functional divergence or whether limited allelic diversity is hiding deeply conserved pleiotropy (Hodgkin, 1998; Paaby and Rockman, 2013; Stearns, 2010; Tyler et al., 2016; Wagner and Zhang, 2011).

WUSCHEL HOMEobox (*WOX*) genes comprise an important gene family controlling plant growth and development by regulating the production and maturation of populations of stem cells in meristems (Bäurle and Laux, 2005; Deveaux et al., 2008; Dolzblasz et al., 2016; Haecker et al., 2004). Most functional studies of *WOX* genes have been performed in the Brassicaceae plant *Arabidopsis thaliana* (*Arabidopsis*), where individual and combined mutations in multiple *WOX* genes have revealed diverse, often redundant developmental roles (Costanzo et al., 2014; Dolzblasz et al., 2016; Lin et al., 2013; Meng et al., 2019; Zhang



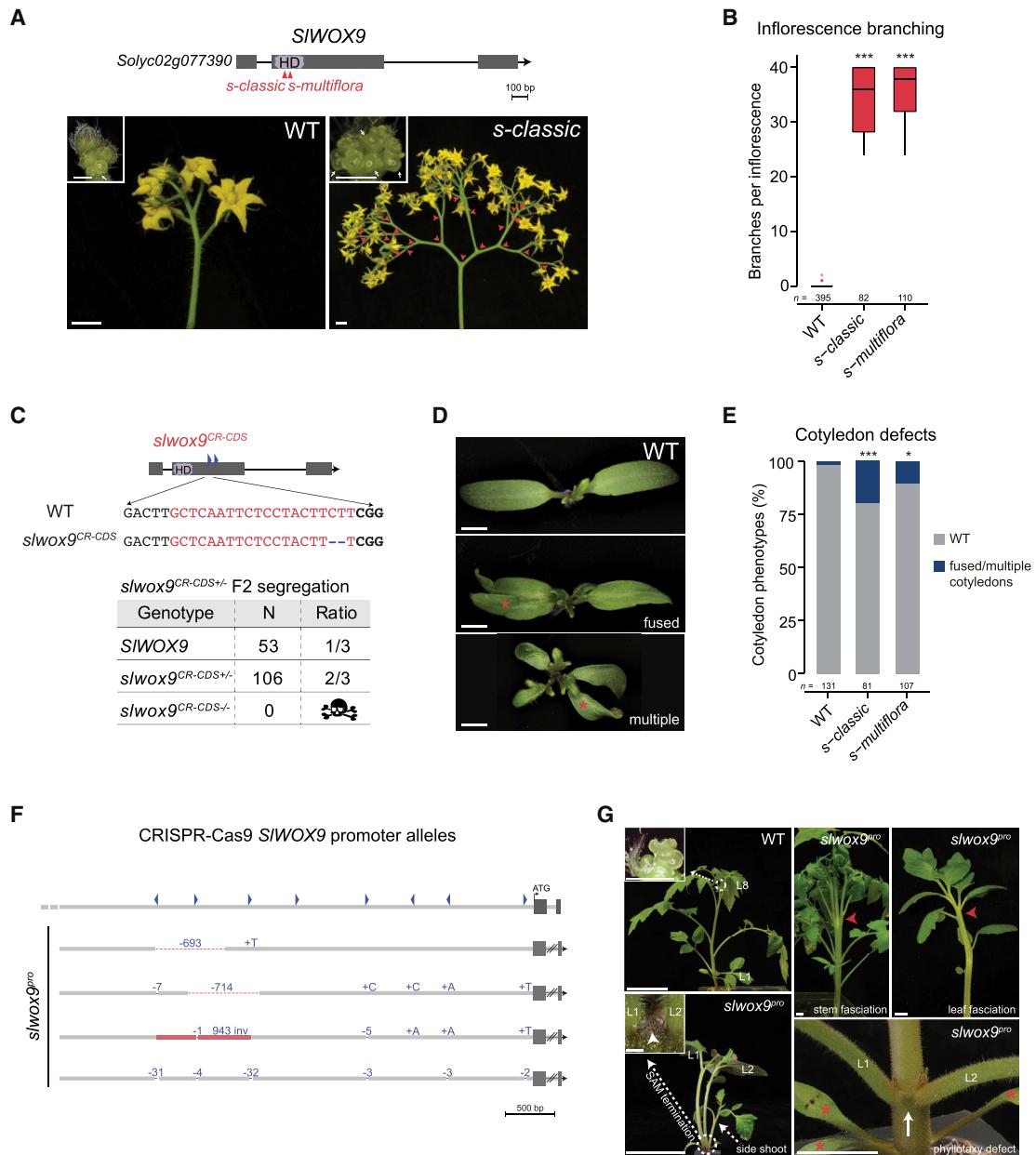


Figure 1. Mutations of the tomato *S/WOX9* and *Arabidopsis WOX9* genes result in embryo defects

(A) Gene model of tomato *S/WOX9* showing positions of coding sequence missense mutations in the mutant alleles *s-classic* and *s-multiflora*. Representative inflorescences of wild-type (WT) and the highly branched *s-classic* mutant. Branching points are marked by red arrowheads. Early development of WT and *s-classic* inflorescences (insets) show additional inflorescence meristems (white arrowheads) in *s-classic* compared with WT. Scale bars represent 1 cm and 500 μ m (insets).

(B) Quantification of inflorescence branching in tomato WT and *s-classic* and *s-multiflora* mutants. Branching was not quantified beyond 40 events. n, inflorescence number.

(C) Null coding sequence mutants of tomato *s/wox9^{CR}* generated by CRISPR-Cas9. Gene diagram is shown. Gray bars are exons, lines are introns, and HD refers to homeodomain. Blue arrowheads are gRNAs. A 2 bp frameshift deletion allele causes embryonic lethality (skull). Summary of F2 segregation is shown.

(D) Defects in tomato *s/wox9* embryo development shown by additional cotyledons (red asterisks) in *s* mutant seedlings. Scale bars represent 1 cm.

(E) Quantification of cotyledon defects in WT, *s-classic*, and *s-multiflora*. n, seedling number.

(F) Schematic of the *S/WOX9* promoter region targeted by CRISPR-Cas9. Blue arrowheads represent gRNAs. Schematics of the four *s/wox9* promoter (*s/wox9^{pro}*) alleles are shown (see also Figure S2 and STAR methods).

(G) Representative images of WT and *s/wox9^{pro}* mutants showing vegetative defects. Insets show the apex of a WT plant after the transition to reproductive growth, in comparison to the terminated SAM (white arrowhead) of *s/wox9^{pro}* mutants. Axillary (side) shoots (dashed arrow) allow continued growth. Red

(legend continued on next page)

et al., 2020). Genetic studies of *WOX* homologs in other plants have shown conservation of meristem function for several family members (Honda et al., 2018; Stuurman et al., 2002). However, cases of functional divergence have also emerged (Costanzo et al., 2014; Meng et al., 2019; Tanaka et al., 2015). One of the most striking examples involve the redundant paralogs *WOX8* and *WOX9* in *Arabidopsis* and their homologs in other species. In *Arabidopsis*, *wox8* *wox9* double mutants and *wox9* null mutants have embryo development defects that can be rescued by transformation with either *WOX8* or *WOX9* (Breuninger et al., 2008). Hypomorphic *wox9* alleles result in weakly penetrant defects in embryo development and meristem maintenance, but mutants otherwise develop normally (Breuninger et al., 2008; Wu et al., 2005, 2007). In contrast, in the Solanaceae plant *Solanum lycopersicum* (tomato), where there is a single *WOX8/9* homolog (Wu et al., 2019), multiple alleles from nature and classical mutagenesis cause overproliferation of reproductive meristems due to a defect in meristem maturation, resulting in excessive branching of flowering shoots known as inflorescences (Lippman et al., 2008). In *Capsicum annuum* (pepper) and *Petunia hybrida* (petunia), two relatives of tomato in the Solanaceae, *wox8/9* mutant plants proliferate what appear to be shoot meristems and rarely produce flowers (Cohen et al., 2014; Rebocho et al., 2008). In *Oryza sativa* (rice), a *wox8/9* mutation not only primarily proliferates axillary shoot meristems but also results in shorter internodes and a subtle effect on inflorescence branching (Wang et al., 2014). Collectively, these cross-species genetic analyses suggested that *WOX8/9* function has diverged across both short and long evolutionary timescales. However, a lack of allelic diversity, uncharacterized *WOX8/9* paralogs (e.g., petunia, rice), and species-specific growth programs that can influence interpretations of mutant phenotypes may all be confounding this conclusion.

Divergence in gene function may be caused by mutations in coding regions, resulting in amino acid changes that modify protein function, or in *cis*-regulatory regions that interact in complex ways to determine domains and levels of expression in different tissues throughout development. Evolutionary innovations often result from changes in *cis*-regulatory regions, as the large number of sequences in these regions and their complex interactions provide a large mutational space to modify expression and phenotype in both qualitative and quantitative ways (Galli et al., 2020; Wittkopp and Kalay, 2011). Thus, interrogating these regions is a powerful approach to expose many facets of gene function and allow differentiation between divergent and conserved gene functions.

Here, we take advantage of our CRISPR-Cas9 multiplex mutagenesis drive system (Rodríguez-Leal et al., 2017) to engineer a large collection of *WOX8/9* promoter alleles in tomato, which exposed multiple pleiotropic roles for this gene that were previously hidden by limited allelic diversity. We show that null mutants of tomato *WOX8/9* are embryonic lethal as in *Arabidopsis*; however, most promoter alleles bypassed this lethality, resulting

in plants showing a spectrum of vegetative and reproductive meristem defects. Notably, different pleiotropic functions were associated with specific regions of overlapping accessible chromatin and *cis*-regulatory sequence conservation. Targeted mutagenesis of the corresponding conserved sequences in the Solanaceae plant shoot apical meristem showed that the pleiotropic meristem functions of groundcherry *WOX8/9* as well as their *cis*-regulatory control are conserved with tomato. Significantly, we showed *WOX8/9* pleiotropy extends to *Arabidopsis*, a species in the Brassicaceae family, despite little conservation of *cis*-regulatory element organization. Our findings suggest that deeply conserved hidden pleiotropy, uncovered through *in vivo* functional dissections of *cis*-regulatory DNA, may be widespread in plant evolution.

RESULTS

Tomato *WOX9* has conserved roles in embryo development

Variation in *WOX8/9* family size and function within and between species has been shaped by gene duplication and loss (Li et al., 2019; Lian et al., 2014; Wu et al., 2019). In our phylogenetic analyses, we found that *Arabidopsis WOX8* and *WOX9* were in a well-supported clade with the single *WOX8/9* gene from tomato that was later independently duplicated in petunia (Figure S1A). However, consistent with the known difficulty of phylogenetic analysis in the *WOX* family, neither our tree nor others (Wu et al., 2019) could resolve the ancestral relationships between these genes. Synteny analysis suggests that the duplication that led to *WOX8* and *WOX9* in *Arabidopsis* was ancient, and one of these ancient paralogs was lost, leading to a single *WOX8/9* homolog in tomato (Figure S1B). Thus, while both *Arabidopsis WOX8* and *WOX9* are homologous to tomato *WOX8/9*, we cannot conclusively determine which is orthologous to the single *WOX8/9* homolog in tomato. For simplicity, we refer to Solanaceae *WOX8/9* homologs as *WOX9* genes throughout. Despite their close relationship, mutational analyses have suggested that *WOX9* homologs have distinct roles in embryo versus inflorescence development in *Arabidopsis* versus tomato (Breuninger et al., 2008; Lippman et al., 2008; Wu et al., 2007).

In tomato, two natural *wox9* alleles (*s-classic* and *s-multiflora*) have changes in highly conserved amino acids in the homeodomain, and these mutations cause overproliferation of reproductive meristems that results in excessive inflorescence branching (Figures 1A and 1B; Table S1B). Since two additional mutations due to structural rearrangements that disrupt the downstream protein-coding region of *S/WOX9* have identical phenotypes, all existing alleles were interpreted as null mutations, and *S/WOX9* in tomato was exclusively linked to regulation of inflorescence branching, in contrast to its essential role in embryogenesis in *Arabidopsis* (Breuninger et al., 2008; Lippman et al., 2008; Wu et al., 2005). To test whether this represents the entire phenotypic

arrowheads mark fasciation, white arrow marks a phyllotaxy defect, and red asterisks mark extra cotyledons. L, leaf number. Scale bars represent 1 cm and 1 mm (insets).

In (B) and (E), * $p < 0.05$ and *** $p < 0.001$, based on two-tailed Dunnett's test.

Raw quantification data and statistical analyses with p values can be found in Table S1.

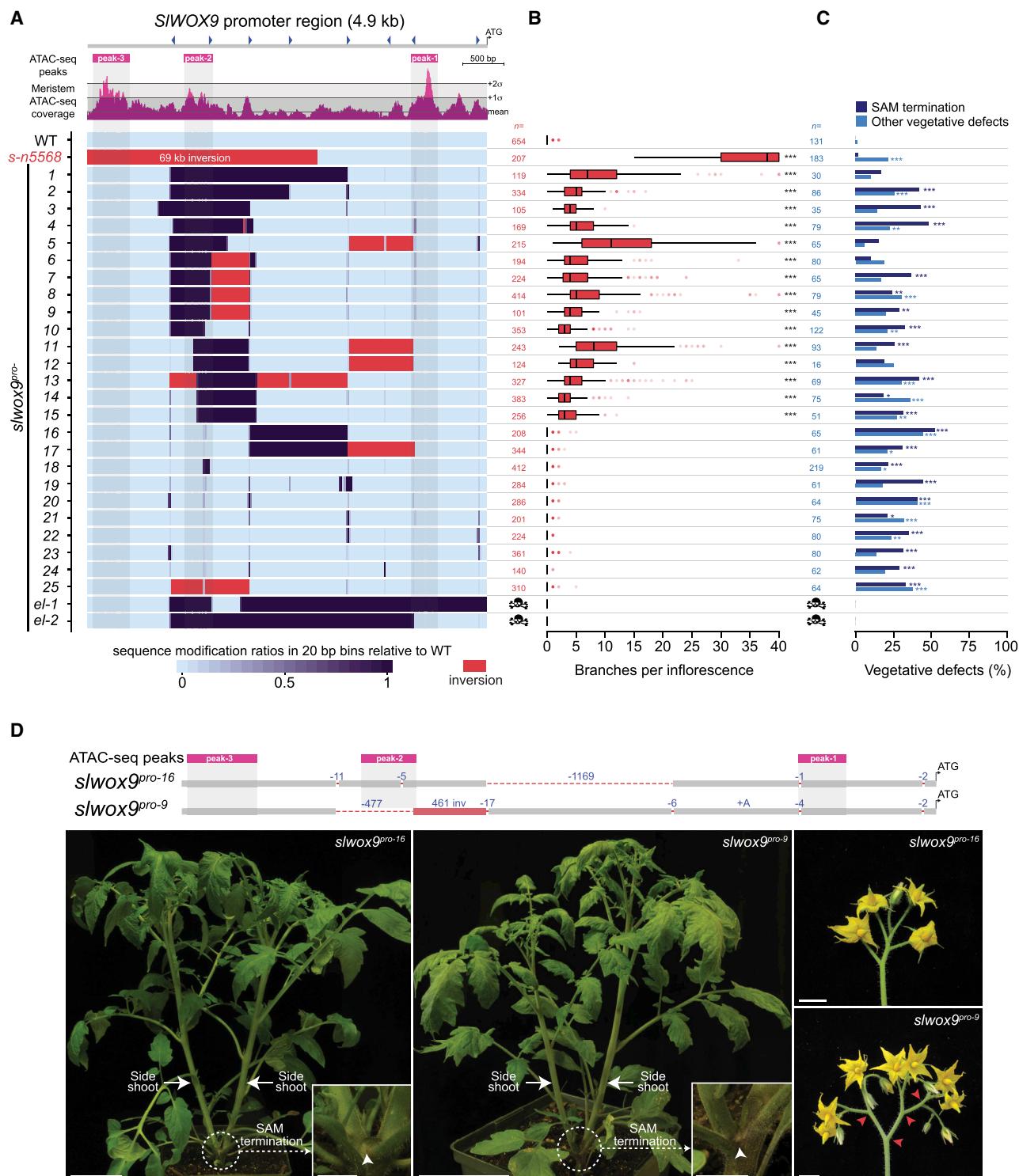


Figure 2. CRISPR-Cas9 promoter alleles of tomato S/IWOX9 reveal pleiotropic roles during vegetative growth and inflorescence development

(A) Meristem ATAC-seq coverage (pink) along with peaks called by Genrich (pink rectangles). Gray lines and color intensity mark the mean and standard deviation. ATAC-seq peaks are numbered sequentially from ATG (see also Figure S2). Heatmap representation of 27 *slwox9^{pro}* promoter alleles, including a 69 kb inversion allele (s-n5568) (see also Figure S2 and STAR methods), and two *slwox9^{pro}* large deletions causing embryonic lethality (skull). The “sequence modification ratios”

(legend continued on next page)

range for tomato *wox9* mutations, we engineered a CRISPR-Cas9 deletion mutation that disrupted the *S/WOX9* reading frame. Surprisingly, this mutation (*s/wox9^{CR-CDS}*) resulted in complete embryonic lethality (Figure 1C), suggesting that *S/WOX9* also functions in embryonic development, a role that was not captured by natural alleles or mutagenesis screens, which only revealed alleles that retained some level of *S/WOX9* embryonic activity. Indeed, reanalysis of large numbers of *s-classic* and *s-multiflora* uncovered low penetrance embryonic defects, such as extra embryonic leaves (i.e., the cotyledons) (Figures 1D and 1E; Table S1C). We also thoroughly phenotyped four previously generated CRISPR-Cas9 *s/wox9* promoter (*s/wox9^{pro}*) alleles causing varying degrees of inflorescence branching (Rodríguez-Leal et al., 2017) and found multiple partially penetrant vegetative defects, including additional cotyledons, early termination of the shoot apical meristem (SAM), organ fusions (stems and leaves), and altered leaf phyllotaxy (Figures 1F and 1G). This newly identified range of *s/wox9* mutant phenotypes was similar to other species, for example, in *Arabidopsis* where *wox9* hypomorphic alleles cause partially penetrant SAM termination (Wu et al., 2005), while stem fusions (i.e., fasciation) and phyllotaxy defects appear in *wox9* mutants of pepper and petunia (Cohen et al., 2014; Rebocho et al., 2008; Schorderet et al., 2018). Our results suggest that the predominant vegetative function of *S/WOX9* is to maintain the embryonic SAM, and later vegetative phenotypes of the *s/wox9^{pro}* alleles may be due to defects in SAM organization (Figure 1G). Together, these observations indicate tomato *S/WOX9* functions in both vegetative and reproductive growth, and its early role in embryo development is shared with *Arabidopsis*.

A CRISPR-Cas9 *cis*-regulatory allelic series of tomato *S/WOX9* associates pleiotropic roles with specific regions of accessible chromatin

The existing *s/wox9^{pro}* alleles exposed *S/WOX9* vegetative roles, but were insufficient to determine whether specific *cis*-regulatory regions control *S/WOX9* pleiotropy. Our CRISPR-Cas9 multiplex mutagenesis drive system can be used to rapidly generate large collections of diverse promoter alleles (Rodríguez-Leal et al., 2017). Using transgenic plants carrying a CRISPR-Cas9 transgene that targets a 3.7 kb proximal region of the *S/WOX9* promoter with eight guide RNAs (gRNAs) (Rodríguez-Leal et al., 2017), we isolated an additional 23 *s/wox9^{pro}* alleles having a diversity of large deletions, inversions, and small insertion-deletion (indel) mutations distributed throughout the target region (Figures 2A and S2D; see STAR methods). We also determined by genome sequencing that

another *s/wox9* allele from fast-neutron mutagenesis (*s-n5568*), which branches like *s/wox9* coding sequence mutations and exhibits a substantial reduction in *S/WOX9* expression levels in reproductive meristems (Figures S2E and S2F) (Lippman et al., 2008; Park et al., 2012), is a 69 kb promoter inversion with a breakpoint 2 kb upstream of the ATG translation start site, in the middle of our target region (Figures 2A, S2A, and S2D; see STAR methods). To better understand *S/WOX9* promoter regulation and relate functional consequences of our collection of *s/wox9^{pro}* alleles, we also assessed chromatin accessibility genome-wide using ATAC-seq (assay for transposase-accessible chromatin using sequencing) of meristem-enriched tissue (Figure S3; see STAR methods). Global analysis of ATAC-seq peaks revealed that more than 50% of accessible chromatin regions were located within 5 kb upstream of transcription start sites (TSSs) of genes, with the majority being within 1 kb (Figures S3D and S3E; see STAR methods). In the promoter of *S/WOX9*, there were four ATAC-seq peaks, including two within the CRISPR-Cas9 target region (Figures 2A and S2).

Of the 23 *s/wox9^{pro}* lines, 15 developed branched inflorescences. Strikingly, this phenotype was tightly associated with deletions in the second ATAC-seq peak region (peak-2, 350 bp) in the distal portion of the promoter (Figures 2A, 2B, and S2; Table S2A). Deletions and inversions that occurred outside peak-2 appeared not to cause branching on their own, but may contribute a mild effect on the range of inflorescence branching (Figure 2B; Table S2A). Thus, peak-2, and also additional regions, likely including those within the large *s-n5568* inversion, are associated with *S/WOX9* inflorescence function.

Two *s/wox9^{pro}* alleles could not be recovered as homozygotes, and embryonic lethality (*el*) was confirmed in segregating families (Figures S2B–S2D). One of these alleles (*s/wox9^{pro-el-1}*) removed the first exon and is therefore a coding sequence null mutation like *s/wox9^{CR-CDS}* (Figures 2A and S2D). The second allele (*s/wox9^{pro-el-2}*) harbored a 3 kb deletion that removes the sequences underlying peak-2 and part of the peak-1 sequence, suggesting peak-1 regulatory sequence (300 bp) may be important for *s/wox9* vegetative phenotypes (Figures 2A and S2D). Indeed, all promoter alleles having a lesion in peak-1 exhibited variable penetrance of SAM termination, while the *s-n5568* inversion, which removed most of the distal part of the promoter, leaving peak-1 intact, had no effect on SAM termination (Figures 2A and 2C; Tables S2A and S2B). Our allelic diversity allowed us to associate vegetative functions with one or more small regions; seven *s/wox9^{pro}* alleles with vegetative phenotypes carry only small indels distributed throughout the target region, resulting

of each allele were calculated in 20 bp bins and are reflected in the color range from light cyan to dark purple (see STAR methods). Red bins mark inversions. Blue arrowheads represent gRNAs. See also Figure S2.

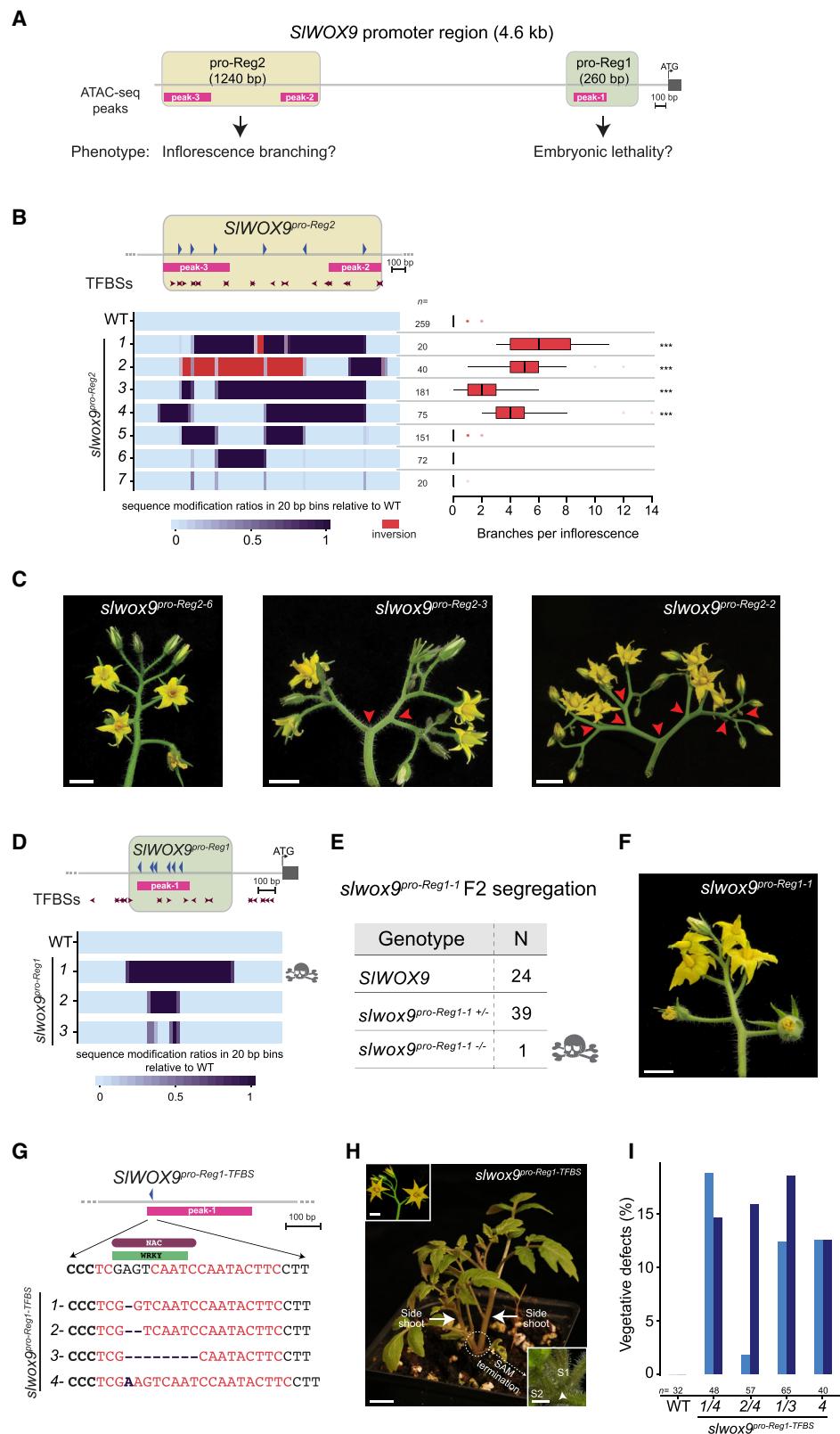
(B) Quantification of inflorescence branching in WT and *s/wox9^{pro}* mutants. n, number of inflorescences. For each allele, at least 10 plants were sampled, with at least 10 inflorescences from each plant.

(C) Quantification of vegetative defects. Traits related to shoot organization and organ boundary defects were classified together as “other vegetative defects.” n, number of seedlings. For each allele, at least 16 seedlings were counted.

(D) Schematics and images comparing the *s/wox9^{pro-16}* and *s/wox9^{pro-9}* alleles in relation to regions of accessible chromatin. Shown are young plants with two side shoots (white arrows). Examination of the arrested main shoot revealed SAM termination (dashed white circle, arrow, insets). *s/wox9^{pro-16}* plants produce unbranched inflorescences like WT (top right), whereas *s/wox9^{pro-9}* inflorescences are branched (bottom right). Branching is associated with the deletion of the accessible chromatin region 2 (peak-2). Red arrowheads mark branching events. Scale bars represents 1 cm.

In (B) and (C), *p < 0.05, **p < 0.01, and ***p < 0.001, based on two-tailed Dunnett’s test.

Raw quantification data and statistical analyses with p values can be found in Table S2.



(legend on next page)

in the concurrent loss of at least four predicted transcription factor binding sites (TFBSs) (Figure S2D; see STAR methods). With the continued improvement of precise genome editing tools in plants (e.g., base editing) (Chen et al., 2019; Mishra et al., 2020), it should be possible to further tease apart the individual and combined roles of these TFBSs, as well as other sequences that are disrupted and potentially involved (e.g., in peak-2) (Figure S2). In contrast to *Arabidopsis* hypomorphic *wox9* plants, which cease growth due to SAM termination (Wu et al., 2005), tomato *slwox9^{pro}* plants showing precocious SAM termination frequently produced axillary shoot meristems that developed normally. Notably, the resulting shoots gave rise to inflorescences that branched only if the sequence underlying peak-2 was removed, validating the critical role of this region in *SIWOX9* reproductive function (Figure 2D).

Results from our broad promoter mutagenesis suggested that *SIWOX9* pleiotropic functions are the result of different activities controlled in part by *cis*-regulatory regions located in two distinct accessible chromatin regions. However, our promoter alleles all contained a range of mutations outside these peak regions. To directly test this hypothesis and further define the *cis*-regulatory regions controlling these different functions, we specifically mutagenized each region of accessible chromatin individually. Since the *s-n5568* allele suggested sequences beyond peak-2 also contribute to the control of inflorescence branching, we first targeted a 1.25 kb distal region spanning the adjacent accessible regions peak-2 and peak-3 (pro-Reg2) (Figures 3A and 3B). Characterization of seven alleles showed that four caused inflorescence branching similar to the 15 *slwox9^{pro}* branching alleles, and also weakly penetrant vegetative phenotypes, suggesting this region has pleiotropic effects (Figures 2B, 3B, and 3C; Table S3A). One of these alleles (*slwox9^{pro-Reg2-1}*) carried a large deletion (1 kb) that spanned most of pro-Reg2, suggesting peak-3 is not a major contributor to inflorescence function (Figure 3B). In support, none of homozygous mutant plants having the alleles that removed portions of peak-3 alone (e.g., *slwox9^{pro-Reg2-5}*) or in combination with peak-2 (e.g., *slwox9^{pro-Reg2-4}*) were

more severely branched than plants with alleles that partially or completely removed peak-2 (Figures 2C and 3B and 3C). Although alleles that eliminated both peaks were not isolated, these data support that the sequence underlying peak-2 is the major contributor to *SIWOX9* inflorescence function in this region and confirm that additional uncharacterized *cis*-regulatory regions also contribute to *SIWOX9* function in reproductive development.

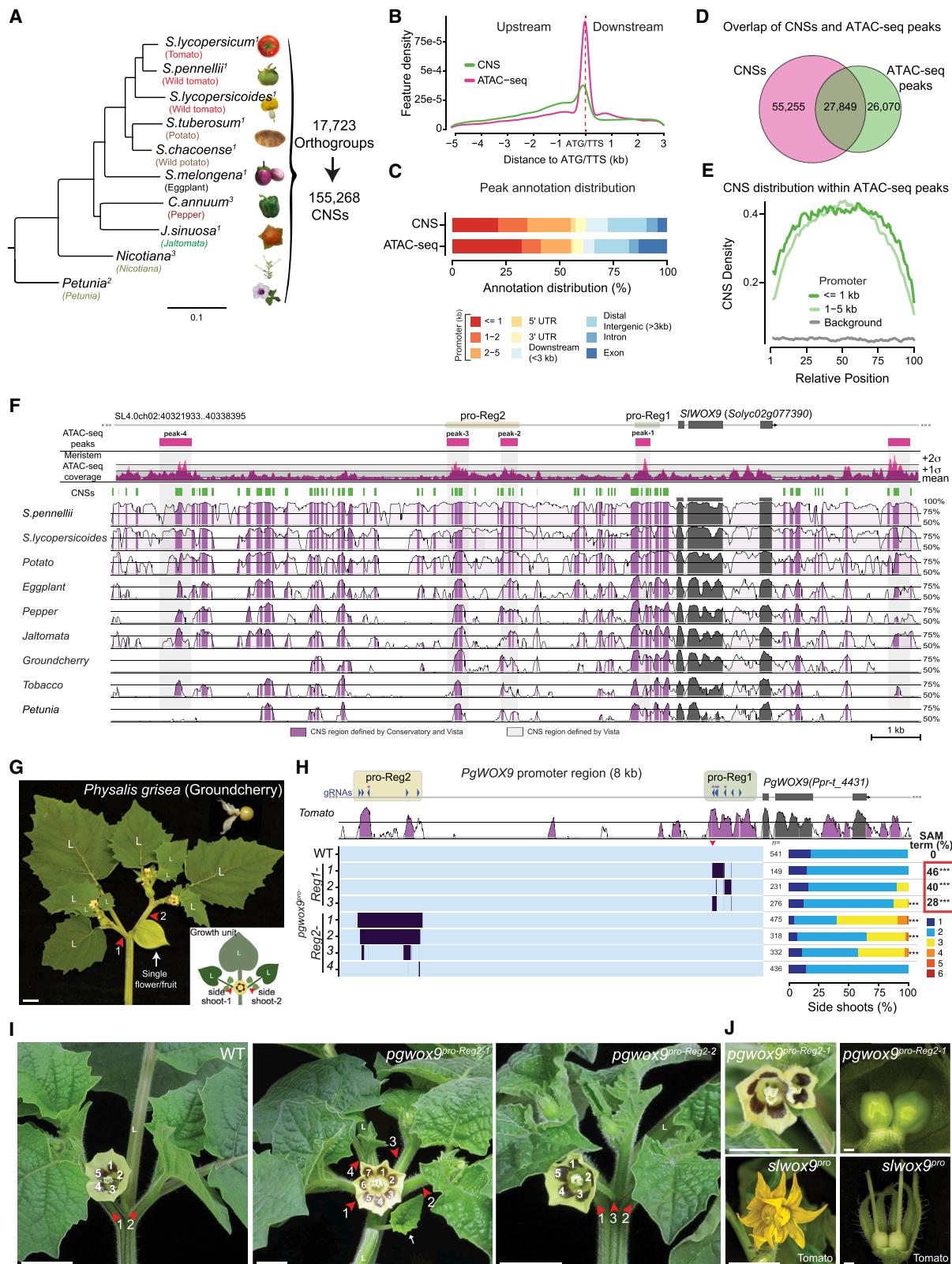
To characterize the function of the sequence underlying peak-1, we targeted this region (pro-Reg1) with a CRISPR-Cas9 construct having six gRNAs. Notably, a 590 bp deletion removing all of pro-Reg1 and some flanking sequences resulted in a high penetrance of embryonic lethality (*slwox9^{pro-Reg1-1}*) (Figures 3D and 3E). A single homozygous plant of *slwox9^{pro-Reg1-1}* developed normally, including unbranched inflorescences, supporting a specific role for pro-Reg1 in embryo development (Figure 3F). Two other alleles having smaller deletions produced normal plants, suggesting these alleles missed a critical sequence or there is redundancy within this region (Figure 3D). We tested this by targeting an 8 bp predicted TFBS that is recognized by the NAC (NAM, ATAF, CUC) and WRKY transcription factor families (Chow et al., 2019; Guo et al., 2015; Rushton et al., 2010; Weirauch et al., 2013) and that is disrupted in all *slwox9^{pro}* alleles (Figures 2A and 3G). Four different indel mutations resulted in partially penetrant vegetative defects, including meristem termination (Figures 3G–3I). These results support that multiple sequences within pro-Reg1 specifically determine *SIWOX9* embryonic function.

Conserved *cis*-regulatory sequences control shared pleiotropy in distantly related Solanaceae species

Our combined promoter dissection by chromatin accessibility and CRISPR-Cas9 *in vivo* analysis showed that in tomato, distinct *SIWOX9* functions controlling early SAM growth or inflorescence branching could be mapped to specific proximal and distal promoter regions. To determine the evolutionary significance of these regions, we first generated a repository of conserved

Figure 3. CRISPR-Cas9 mutagenesis of specific *SIWOX9* *cis*-regulatory regions separates embryonic and inflorescence functions

- (A) Schematic of the *SIWOX9* promoter displaying two regions of accessible chromatin, mutations that are hypothesized to cause embryonic lethality (pro-Reg1) or inflorescence branching (pro-Reg2).
 - (B) Schematic of *SIWOX9^{pro-Reg2}* region targeted by CRISPR-Cas9. ATAC-seq peaks and transcription factor binding sites (TFBSs) (q-value > 0.01) are shown. Encoding for seven *slwox9^{pro-Reg2}* alleles is shown along with quantification of inflorescence branching. Note that mutations in pro-Reg2 alone are associated with increased branching. See also Figure S2.
 - (C) Representative images showing branching (red arrowheads) in two of the *slwox9^{pro-Reg2}* alleles with a deletion that removes accessible chromatin ATAC-seq peak-2, compared with the normal inflorescences of *slwox9^{pro-Reg2-6}* plants with a deletion that removes half of accessible chromatin region ATAC-seq peak-3. n, inflorescence number. Scale bars represent 1 cm.
 - (D) Schematic of the *SIWOX9^{pro-Reg1}* region targeted by CRISPR-Cas9. Blue arrowheads represent gRNAs. ATAC-seq peaks and TFBSs with a q-value > 0.01 are shown (see STAR methods). Encoding for three *slwox9^{pro-Reg1}* alleles is shown. *slwox9^{pro-Reg1-1}* showed high penetrance of embryonic lethality, indicated by a gray skull. See also Figure S2.
 - (E) Summary of *slwox9^{pro-Reg1-1}* F2 segregation.
 - p values in (B) represent <0.001 (**) and are based on two-tailed Dunnett's test.
 - (F) Representative inflorescence from the homozygous *slwox9^{pro-Reg1-1}* mutant plant. Scale bar represents 1 cm.
 - (G) Schematic of the *SIWOX9^{pro-Reg1}* region showing the predicted NAC/WRKY TFBS targeted by CRISPR-Cas9 using a single gRNA (*SIWOX9^{pro-Reg1-TFBS}*). Purple font indicate nucleotides disrupted by gRNA-2 in all *slwox9^{pro}* alleles (Figure 2A). Four *slwox9^{pro-Reg1-TFBS}* indel mutations are shown.
 - (H) Image showing a *slwox9^{pro-Reg1-TFBS}* mutant plant exhibiting SAM termination (dashed white circle, arrow, bottom inset). *slwox9^{pro-Reg1-TFBS}* plants produce unbranched inflorescences like WT (top inset). Scale bar represents 1 cm. Scale bars: 1 mm in bottom inset and 1 cm in top inset.
 - (I) Quantification of vegetative defects in mutant plants homozygous and biallelic for *slwox9^{pro-Reg1-TFBS}* indel alleles. Data from homozygous and biallelic mutant plants are shown. Light blue bars, other vegetative defects. Dark blue bars, SAM termination.
- Raw quantification data and statistical analyses with p values can be found in Table S3.



(legend on next page)

non-coding sequences (CNSs) in the Solanaceae family, which includes species with extreme variation in inflorescence architectures, from single flowers to highly branched structures with dozens of flowers (Park et al., 2014). Ten Solanaceae species have genome sequences, but quality and annotation completeness vary greatly (Table S4; see STAR methods). To account for these limitations, we developed Conservatory, an algorithm that combines information from multiple species to improve genome annotation and fill in missing or low-quality regions. This algorithm first groups genes and gene duplications from multiple species into orthogroups. Regulatory sequences of these orthogroups are then aligned to the tomato reference genome, followed by conservation analysis using phyloP (Pollard et al., 2010) to identify single-base-pair resolution conservation in both upstream and downstream sequences (Figures S4A–S4F; see STAR methods). Nucleotides with significant phyloP conservation score within 25 bp windows were combined to form distinct CNSs (see STAR methods). Overall, this approach identified 155,268 CNSs associated with 17,723 orthogroups (Figure 4A). Consistent with previous conservation analyses (Lu et al., 2019; Van de Velde et al., 2016), the Solanaceae CNSs were enriched within 1 kb of the transcription start and stop sites, but 47% of CNSs were distal, located between 3 and 50 kb upstream of a gene (Figures 4B, 4C, and S4C). Importantly, more than half of the accessible chromatin regions we identified in meristems overlapped with at least one CNS, and the highest density of conservation from these CNSs were enriched at the centers of accessible chromatin (Figures 4D and 4E). These data support that the CNSs captured by our Conservatory algorithm represent functional sequences.

Conservatory analysis showed that the *S/WOX9* promoter had a relatively high level of conservation, with 19% of its sequence conserved within the Solanaceae (median promoter conservation

across all genes was 5.2%) (Figure S4E). Many of the underlying CNSs were embedded in larger blocks of conservation defined by traditional pairwise sequence alignment tools that assess *cis*-regulatory sequence similarity (e.g., mVISTA) (Frazer et al., 2004) (Figure 4F). Consistent with CNSs frequently overlapping with regions of accessible chromatin (Figures 4D and 4E) (Lu et al., 2019), all four *S/WOX9* promoter ATAC-seq peaks contained CNSs (Figure 4F). Both pro-Reg1 and pro-Reg2 exhibited high levels of sequence conservation. Conservation of pro-Reg1 sequences might be expected, as embryonic development is a core developmental program with little variation between eudicot species (Radoeva et al., 2019). However, conservation in pro-Reg2, which prevents branching in the multiflowered inflorescences of tomato, was surprising, as it was found even in Solanaceae species that produce unbranched single-flower inflorescences such as pepper and petunia.

To test the roles of these conserved sequences in a single-flower species, we used groundcherry, an emerging diploid Solanaceae system (Figure 4G) (Lemon et al., 2018). Groundcherry has a single *WOX9* gene (*PgWOX9*), and targeting its coding region resulted in embryonic lethality similar to tomato *slwox9^{CR-CDS}* (Figures S4G and S4H). Most of the promoter conservation with tomato is contained within the pro-Reg1 and pro-Reg2 *cis*-regulatory regions (Figure 4H). Following the approach used in tomato, we generated multiple alleles in groundcherry pro-Reg1 and pro-Reg2 by CRISPR-Cas9 mutagenesis (Figures 4H and S4I). Consistent with the role of this *cis*-regulatory region in tomato, all three pro-Reg1 alleles (*pgwox9^{pro-Reg1}*) caused high frequencies of SAM termination (Figure 4H). Notably, all were mutated in the NAC/WRKY TFBS. By contrast, pro-Reg2 alleles (*pgwox9^{pro-Reg2}*) exhibited multiple phenotypes, such as formation of additional branches (three instead of two), flowers with extra petals, and inflorescences comprised of fused flowers (Figures

Figure 4. Solanaceae CNSs largely coincide with accessible chromatin and define conserved *WOX9* pleiotropy in the tomato relative *Physalis grisea* (groundcherry)

- (A) Phylogenetic tree of Solanaceae species used to define orthogroups and conserved non-coding sequences (CNSs) (see STAR methods). Fifteen genomes from 10 species were used. Number of genomes for each species is in superscript.
- (B) Peak density comparison between CNSs and ATAC-seq peaks in orthogroup promoters (5 kb upstream of ATG) and downstream regions (3 kb downstream to the transcription termination site [TTS]). Dashed red line marks the ATG and TTS.
- (C) Comparison of annotated distributions of CNSs and ATAC-seq peaks in the indicated regions.
- (D) Venn diagram showing overlap of CNSs and ATAC-seq peaks from the same regions in (C).
- (E) CNS position distribution relative to overlapping ATAC-seq peaks in the indicated promoter regions. Randomized positions were used to define background. Note that the summit of conservation (CNS density) resides at ATAC-seq peak midpoints.
- (F) Conservation analysis of a 15 kb window surrounding Solanaceae *WOX9*. ATAC-seq coverage and peaks are shown. Green boxes define CNSs. Below is an mVISTA plot of a 20 kb region surrounding *WOX9* orthologs in the indicated species (see STAR methods). Regions identified as CNSs are dark purple. Regions of high similarity defined by mVISTA, but not CNSs, are light purple. Exons are dark gray. CNSs were not called in the gene body. See also Figure S4.
- (G) Representative image of the Solanaceae plant groundcherry. Inset shows diagram of modular growth (shoot units) of groundcherry plants. Compared with tomato multiflowered inflorescences, groundcherry develops single-flower inflorescences that bear fruits with a papery husk. In addition, groundcherry develops two side shoots after inflorescence formation compared with a single side shoot in tomato. Scale bars represent 1 cm.
- (H) mVISTA plot of 15 kb surrounding the tomato *S/WOX9* gene aligned to the corresponding region of *PgWOX9*. Note the two highly conserved regulatory regions (pro-Reg-1 and pro-Reg-2) that were defined in tomato. The gRNAs (blue arrows) used to target the two *cis*-regulatory regions *PgWOX9^{pro-Reg1}* and *PgWOX9^{pro-Reg2}* are shown. Purple asterisks mark identical gRNA sequences used for CRISPR-Cas9 mutagenesis in tomato. Encoding of both *pgwox9^{pro-Reg1}* and *pgwox9^{pro-Reg2}* alleles are shown. Red arrowhead indicates the position of the conserved NAC/WRKY TFBS, which is mutated in all three *pgwox9^{pro-Reg1}* alleles. Quantification of branching is shown (right). SAM termination (%) is shown for *pgwox9^{pro-Reg1}* mutants. See also Figure S3.
- (I) Representative images of WT and the *pgwox9^{pro-Reg2}* shoot units. *pgwox9^{pro-Reg2}* plants show stem and flower fasciation and increased side shoot branching. Numbered red arrowheads mark side shoots. Note the additional petals and side shoots in the *pgwox9^{pro-Reg2}* mutants. Scale bars represent 1 mm.
- (J) *pgwox9^{pro-Reg2}* and *slwox9^{pro}* mutants develop rare twin flowers. Scale bars represent 1 mm.

In (H), ***p < 0.001, based on two-tailed Dunnett's test.

Raw quantification data and statistical analyses with p values can be found in Table S4.

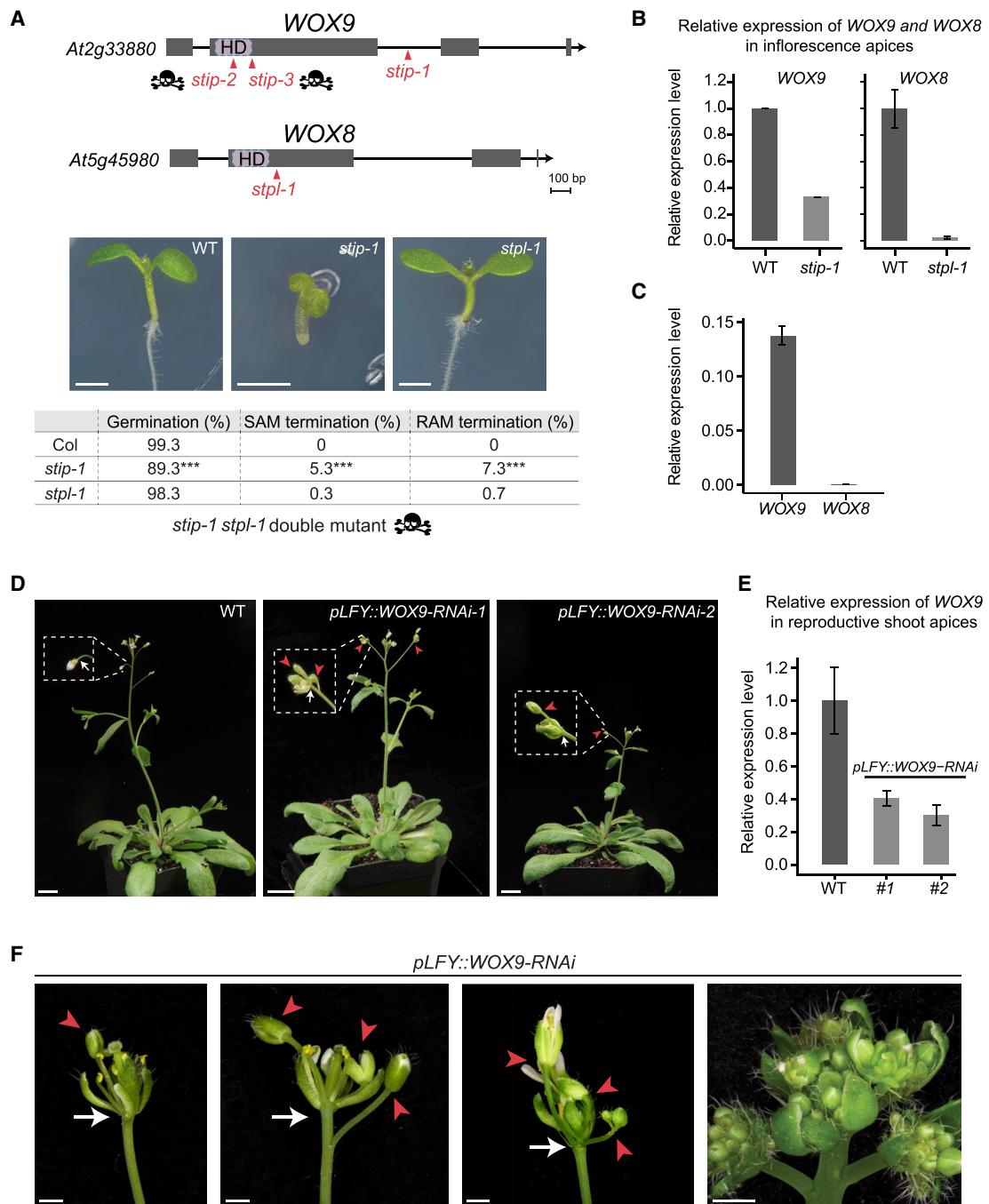


Figure 5. The role of *WOX9* in inflorescence development is conserved in *Arabidopsis*

(A) Gene models of *Arabidopsis* *WOX8* and *WOX9*. Three mutant alleles of *wox9* are shown: *stip-1* (T-DNA insertion, hypomorphic), *stip-2* (T-DNA insertion, null), and *stip-3* (nonsense mutation, null). One mutant allele of *wox8* is shown: *stpl-1* (T-DNA insertion, null). *stip-2* and *stip-3* are both embryonic lethal (skull). Images showing seedling development of WT, *stip-1*, and *stpl-1*. Table below shows low frequencies of seedling shoot and root growth (meristem) termination in *stip-1* and *stpl-1*. Double mutants of *stip-1 stpl-1* are embryonic lethal. Scale bars represent 1 mm.

(B) Quantitative reverse transcriptase PCR (qRT-PCR) expression analysis of *WOX8* and *WOX9* in *stip-1* and *stpl-1* mutants relative to WT.

(C) Expression analysis (qRT-PCR) of *WOX8* and *WOX9* in inflorescences of WT.

(D) Representative images of WT and *pLFY::WOX9-RNAi* plants and inflorescences showing the flower-in-flower phenotype (insets) due to reduction of *WOX9* expression in reproductive apices (STAR methods). Ectopic flowers (red arrowheads) develop where individual flowers should form (white arrows). Scale bars represent 1 cm.

(E) Expression analysis (qRT-PCR) showing reduction of *WOX9* transcripts in reproductive apices of two *pLFY::WOX9-RNAi* transgenic lines compared with WT.

(legend continued on next page)

4H, 4I, S4J, and S4K). These diverse phenotypes commonly result from overproliferation of meristem cells and delayed meristem maturation (Barton, 2010; Lemmon et al., 2016). Importantly, we also noted these floral phenotypes, albeit rarely, in many tomato *sIwox9^{pro}* branching alleles (Figure 4J). The groundcherry branching phenotypes also resembled, in a weaker manner, the *wox9* fasciation and branching phenotypes in pepper and petunia resulting from meristem overproliferation (Cohen et al., 2014; Rebocho et al., 2008; Schorderet et al., 2018). Taken together, our results suggest that, rather than undergoing divergence, both *WOX9* pleiotropic vegetative and reproductive roles and the *cis*-regulatory regions that control them are conserved in the Solanaceae.

The pleiotropic role of *WOX9* in inflorescence development is conserved in *Arabidopsis*

Despite differences in inflorescence architecture, the conserved pleiotropic functions of *WOX9* within the Solanaceae prompted us to ask whether this conservation extended to *Arabidopsis*, a distantly related plant in the Brassicaceae family. Similar to previously published tomato *wox9* alleles, which revealed weakly penetrant effects on embryo development (Figure 1), reproductive phenotypes may have been missed in previously characterized *Arabidopsis wox9* and *wox8* mutants. We confirmed that *Arabidopsis wox9* null alleles as well as double mutants of a hypomorphic *wox9* allele with a null *wox8* allele are embryonic lethal. Single hypomorphic *wox9* mutants exhibit a weakly penetrant early meristem termination phenotype, and all null *wox8* mutants appeared normal (Figure 5A; Table S5). We grew large numbers of homozygous plants, but did not find reproductive branching or inflorescence phenotypes in *wox9* hypomorphic or *wox8* null mutant plants, despite reduced expression of *WOX9* and complete loss of *WOX8* expression in inflorescence tissue, respectively (Figure 5B).

However, *WOX9* is abundantly expressed in inflorescence tissue, whereas *WOX8* is barely detectable (Figure 5C), suggesting *WOX9* does have a potential function in reproductive development. Limited *wox9* allelic diversity in *Arabidopsis* may explain the failure to find reproductive phenotypes, but could be revealed through *cis*-regulatory mutagenesis. Conservatory analysis using eight Brassicaceae species revealed that the promoter of *Arabidopsis WOX9* is relatively highly conserved (21% conserved sequence), similar to tomato *S/WOX9*, whereas *WOX8* promoter conservation is much lower (11%; median promoter conservation across all genes was 13%) (Figures S5A–S5E). Sequence divergence precluded informative mVISTA alignments of *Arabidopsis WOX8* and *WOX9* promoters against the promoter of *S/WOX9*. To address this problem, we extracted the consensus sequences for each family's CNSs, and using our Conservatory algorithm, compared the consensus CNSs from the Solanaceae and Brassicaceae families. This reduction in search space al-

lowed us to identify and deconvolute 24 deeply conserved short (10- to 30-bp) sequences between the promoters of tomato *S/WOX9* and *Arabidopsis WOX9*, including one element having the NAC/WRKY TFBS (Figures 3G and S5F). Only four of these sequences were shared between tomato *S/WOX9* and *Arabidopsis WOX8*, supporting a closer *cis*-regulatory functional relationship between the promoters of *S/WOX9* and *Arabidopsis WOX9*. Importantly, more than half of these sequences are clustered within the *S/WOX9* pro-Reg1 and pro-Reg2 *cis*-regulatory regions, and no conserved sequences from pro-Reg2 were found in the promoter of *WOX8*, consistent with low *WOX8* expression in inflorescences. However, grouping, order, and orientation of these conserved sequences was not maintained between the promoters of tomato *S/WOX9* and *Arabidopsis WOX9* (Figure S5F).

Our Conservatory analysis suggested that *Arabidopsis WOX9* might function in inflorescence development, so we mutagenized its promoter using two multiplex CRISPR-Cas9 constructs that targeted two regions of deep conservation (Figure S5G) (see STAR methods). Screening more than 300 independent transgenic lines resulted in many chimeric plants but only two heritable alleles, likely due to low CRISPR-Cas9 efficiency in *Arabidopsis* (LeBlanc et al., 2018). Neither allele affected the deeply conserved sequences and no vegetative or reproductive phenotypes were observed (Figure S5G). To bypass this technical limitation, we mimicked the effects of tomato *sIwox9^{pro-Reg2}* *cis*-regulatory alleles by expressing an inverted repeat targeting *WOX9* using the *LEAFY* (*LFY*) promoter, which is active predominantly in reproductive meristems (*pLFY::WOX9-RNAi*) (Blázquez et al., 1997) (see STAR methods). Remarkably, two independent transgenic lines and their progeny showing more than 50% reduction in *WOX9* expression converted flowers into inflorescence-like structures (Figures 5D–5F), ranging from a few additional flowers to massive overproliferation of floral meristems (Figure 5F). These results show that *Arabidopsis WOX9* shares vegetative and reproductive functions with *WOX9* in the Solanaceae and indicate that pleiotropic control of development by these plant homeobox genes is deeply conserved.

DISCUSSION

We discovered hidden conserved pleiotropy in homologous plant homeobox genes through *in vivo* functional dissection of *cis*-regulatory DNA. By generating genome-edited promoter alleles in the tomato *S/WOX9* gene, we uncovered pleiotropic roles in vegetative and reproductive development and further showed that specific *cis*-regulatory regions control this pleiotropy. Despite differences in reproductive meristem activity that result in a simpler inflorescence architecture in the distantly related groundcherry compared with tomato, *WOX9* pleiotropy and its *cis*-regulatory control were conserved. Our findings show that

(F) Representative flowers from *pLFY::WOX9-RNAi* inflorescences arranged by severity of floral meristem overproliferation. On the left, an example of weak floral meristem overproliferation, where a single flower is transformed into an inflorescence developing two flowers. On the right, an example of massive overproliferation of inflorescence and floral meristems, and flower buds developing in individual flowers. Ectopic flowers (red arrowheads) develop where individual flowers normally form (white arrows). Scales bars represent 1 mm.

In (A), ***p < 0.001, based on two-tailed Dunnett's test.

Raw quantification data and statistical analyses with p values can be found in Table S5. See also Figure S5.

s/wox9 branching phenotypes in single-flower (groundcherry, pepper, petunia) versus multiflower (tomato) Solanaceae species is likely not due to differences in *WOX9* function, but rather is the result of species-specific differences in developmental programs (Cohen et al., 2014; Park et al., 2014; Rebocho et al., 2008). We found similar defects in floral meristem maturation and overproliferation when we reduced *WOX9* expression in *Arabidopsis*, similar to weak inflorescence branching in a *wox9* mutant in rice (Wang et al., 2014). Notably, the growth program that determines inflorescence architecture in *Arabidopsis* and rice is distinct from that in the Solanaceae (Kyozuka et al., 2014; Park et al., 2014). Our approach of engineering *cis*-regulatory allelic diversity in tomato and groundcherry was critical to expose the spectrum of *WOX9* pleiotropy, suggesting that embryonic lethality and SAM termination in the *wox9* mutants of pepper and petunia could emerge from creating similar allelic diversity. In the case of petunia, potential redundancy with a *WOX9* paralog may also reveal missing embryonic phenotypes. Rice has two additional *WOX9* genes (Figure S1A), and it seems likely that a full spectrum of vegetative and reproductive pleiotropic roles will emerge from engineering and combining protein coding and *cis*-regulatory allelic diversity in all three paralogs. We propose that *WOX9* genes have preserved ancient functions in vegetative and reproductive development in flowering plants (Zhu et al., 2014).

There remains much debate about the extent of pleiotropy and how pleiotropic genes evolve to shape phenotypes (Hodgkin, 1998; Paaby and Rockman, 2013; Stearns, 2010; Wagner and Zhang, 2011). Global assessments of pleiotropy have thus far been based on quantitative genetic studies and analyses of large-scale mutant populations (Paaby and Rockman, 2013; Rockman, 2012; Solovieff et al., 2013; Tyler et al., 2016; Visscher and Yang, 2016). Natural alleles are largely hypomorphic, but as their number is limited, such variation typically reveals only a subset of the possible gene functions (Paaby and Rockman, 2013; Stern, 2000), as found for tomato *s/wox9* alleles (Lippman et al., 2008). On the other hand, laboratory-generated knockout mutants can reveal more pleiotropy; however, they may also mask functions due to developmental epistasis, like *Arabidopsis wox9* embryonic lethal mutants (Breuninger et al., 2008; Wu et al., 2005). We demonstrated that *cis*-regulatory dissection by genome editing is a powerful approach to expose hidden pleiotropy. However, it remains challenging to identify the specific *cis*-regulatory sequences that will reveal pleiotropic gene functions. Conservation of NCS is an excellent proxy for *cis*-regulatory function (Lu et al., 2019), as we demonstrated by applying Conservatory in the Solanaceae family to identify and characterize corresponding functional *cis*-regulatory sequences in groundcherry. However, identifying conservation over large evolutionary distances is more challenging due to the rapid divergence of these sequences (Fuqua et al., 2020; Nelson and Wardle, 2013; Rebeiz and Tsiantis, 2017; Wong et al., 2020). Our use of Conservatory to compare CNS blocks between the distantly related Solanaceae and Brassicaceae families allowed us to identify short CNSs in the tomato *S/WOX9* and *Arabidopsis WOX9* promoters, despite substantial rearrangements in their clustering, order, and orientation (Figure S5F). This suggests that functional characterization of CNSs can be used first to

expose pleiotropic roles of a gene in one species and then specific regions harboring short, deeply conserved sequences, many of which encode TFBSs, provide the most promising targets to reveal conserved pleiotropy across diverse species. Indeed, similar patterns of functional and regulatory element conservation were recently observed for an ancient animal enhancer whose role was maintained between sponges and mammals (Wong et al., 2020).

Our *cis*-regulatory dissection of Solanaceae *WOX9* pleiotropy also revealed variable penetrance and quantitative effects in our collection of alleles, including from disrupting a specific TFBS, demonstrating that multiple CNSs contribute to vegetative and reproductive functions (Figures 2A–2C and 4H). This is similar to findings in animals showing that complex, often redundant, relationships within and among short *cis*-regulatory sequences determine precise expression patterns and levels of gene expression (Cameron and Davidson, 2009; Fuqua et al., 2020; Priest et al., 2009; Schwarzer and Spitz, 2014; Wittkopp and Kailay, 2011). Overall, our work lays the foundation to broadly explore the prevalence of gene functional divergence and conserved pleiotropy in plant and animal development and study how the evolution of *cis*-regulatory complexity shapes these trajectories.

Finally, our findings have important implications for the deployment of genome editing in agriculture. Mutations in *cis*-regulatory regions are not only critical in evolution because of their weak phenotypic effects but also because such variation can have fewer pleiotropic consequences than mutations in protein-coding regions (Galli et al., 2020; Stern, 2000). Indeed, evidence that crop domestication and improvement depend disproportionately on mutations in *cis*-regulatory regions continues to grow (Eshed and Lippman, 2019; Meyer and Purugganan, 2013; Swinnen et al., 2016); however, this large mutational space has not been fully explored or exploited (Scheben and Edwards, 2018; Springer et al., 2019). Genome editing has the potential to rapidly improve agricultural traits in many crops, and we previously showed that quantitative trait variation can be engineered through CRISPR-Cas9 mutagenesis of promoters (Rodríguez-Leal et al., 2017). As shown here, by prioritizing conserved NCSs, this approach can simultaneously define, partition, and quantitatively tune the functions of pleiotropic genes. Taming pleiotropy through precise genome editing of *cis*-regulatory regions may open the door to a vast pool of new genes and variants that could be exploited in agriculture.

Limitations of study

In this study, we aimed to expose the entire functional spectrum of the *WOX9* gene in an evolutionary context. We generated dozens of tomato *S/WOX9* *cis*-regulatory mutant alleles, which exposed its pleiotropic roles and identified *cis*-regulatory regions associated with different pleiotropic functions, also in the tomato relative groundcherry. However, expression analyses of our deep and diverse collections of *cis*-regulatory alleles on vegetative and reproductive meristems in both species could reveal the extent that expression changes are associated with partitioning of pleiotropic functions. Our preliminary analysis of the severe *s/wox9* allele *s-n5568* suggests that reduced expression levels, rather than altered patterns, are the relevant change for the

quantitative phenotypic outputs, but a comprehensive analysis of all of our alleles is necessary to address this hypothesis. Apart from this, while we were able to show conservation of function from specific *cis*-regulatory regions within the Solanaceae, technical limitations in generating *Arabidopsis* promoter CRISPR alleles prevented us from demonstrating that the identified conserved short *cis*-regulatory elements between Solanaceae and Brassicaceae also share functional roles. If indeed such deep functional conservation exists between these elements in *WOX9*, its characterization will be a major step forward in the approaches needed to identify and study conserved *cis*-regulatory elements in many genes across flowering plants. The rapid introduction of high-efficiency genome-editing techniques in many species, together with advanced sequence analysis tools will allow deeper exploration of this fundamental biological question in the near future.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
 - Lead contact
 - Materials availability
 - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
 - Plant materials and growth conditions
- **METHOD DETAILS**
 - Plant phenotyping
 - Genome sequencing
 - Phylogenetic analyses and sequence analyses
 - CRISPR-Cas9 mutagenesis, plant transformation, and selection of mutant alleles
 - Characterization of *cis*-regulatory mutant alleles
 - Assay for Transposase-Accessible Chromatin using sequencing (ATAC-Seq) libraries
 - Identification of conserved non-coding sequences (CNSs), Conservatory
 - Tissue collection and RNA extraction
 - *S/**WOX9* expression analysis
 - *In situ* hybridization
- **QUANTIFICATION AND STATISTICAL ANALYSES**
- **ADDITIONAL RESOURCES**

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.cell.2021.02.001>.

ACKNOWLEDGMENTS

We thank members of the Lippman laboratory for valuable comments and discussions. We thank J. Kim and A. Krainer for technical support. We thank L. Randall, A. Horowitz Doyle, P. Keen, K. Swartwood, M. Tjahjadi, and J. Van Eck for performing tomato and groundcherry transformations. We thank T. Mulligan, K. Schlecht, A. Krainer, and S. Qiao for assistance with plant care. We thank M. Karaca from Inari Agriculture for assistance with ATAC-seq (prior to September 2018). We thank B. Usadel, S. Strickler, and L. Mueller for

sharing the *S. lycopersicoides* genome assembly and annotation sequences. We thank D. Jackson for comments on the manuscript. This research was supported by the Howard Hughes Medical Institute and grant support from the BARD (United States-Israel Binational Agricultural Research and Development Fund, IS-5120-18C) to Y.E. and Z.B.L. and the Vaadia-BARD Postdoctoral Fellowship (award no. FL-542-16) to A.H., the Howard Hughes Medical Institute International Scholar Grant (55008730) and Israel Science Foundation (ISF966/17) to I.E., and the National Science Foundation Plant Genome Research Program (grant no. IOS-1546837) to M.E.B. and Z.B.L.

AUTHOR CONTRIBUTIONS

Z.B.L. conceived the project. A.H., S.Z., D.R.-L., X.W., I.E., and Z.B.L. designed and planned experiments. A.H., S.Z., D.R.-L., G.R., X.W., L.T., P.W., Y.E., I.E., and Z.B.L. performed experiments and collected data. A.H., S.Z., N.D., X.W., J.K., P.W., M.E.B., I.E., and Z.B.L. analyzed data and A.H., I.E., and Z.B.L. wrote the manuscript with input from M.E.B. and Y.E. All authors read, edited, and approved the manuscript.

DECLARATION OF INTERESTS

Z.B.L. is a consultant for and a member of the Scientific Strategy Board of Inari Agriculture, and he is also a named inventor on a number of patents and patent applications directed to related technology that have been exclusively licensed from CSHL to Inari Agriculture.

Received: November 17, 2020

Revised: January 3, 2021

Accepted: February 1, 2021

Published: March 4, 2021

REFERENCES

- Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37, W202–W208.
- Barton, M.K. (2010). Twenty years on: the inner workings of the shoot apical meristem, a developmental dynamo. *Dev. Biol.* 341, 95–113.
- Bäurle, I., and Laux, T. (2005). Regulation of WUSCHEL transcription in the stem cell niche of the *Arabidopsis* shoot meristem. *Plant Cell* 17, 2271–2280.
- Belhaj, K., Chaparro-Garcia, A., Kamoun, S., and Nekrasov, V. (2013). Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods* 9, 39.
- Blanchette, M., Kent, W.J., Riemer, C., Elnitski, L., Smit, A.F.A., Roskin, K.M., Baertsch, R., Rosenblom, K., Clawson, H., Green, E.D., et al. (2004). Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.* 14, 708–715.
- Blázquez, M.A.F., Soowal, L.N., Lee, I., and Weigel, D. (1997). LEAFY expression and flower initiation in *Arabidopsis*. *Development* 124, 3835–3844.
- Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.
- Bray, N., Dubchak, I., and Pachter, L. (2003). AVID: A global alignment program. *Genome Res.* 13, 97–102.
- Breuninger, H., Rikirsch, E., Hermann, M., Ueda, M., and Laux, T. (2008). Differential expression of *WOX* genes mediates apical-basal axis formation in the *Arabidopsis* embryo. *Dev. Cell* 14, 867–876.
- Brooks, C., Nekrasov, V., Lippman, Z.B., and Van Eck, J. (2014). Efficient gene editing in tomato in the first generation using the clustered regularly interspaced short palindromic repeats/CRISPR-associated9 system. *Plant Physiol.* 166, 1292–1297.
- Brudno, M., Malde, S., Poliakov, A., Do, C.B., Couronne, O., Dubchak, I., and Batzoglou, S. (2003). Glocal alignment: finding rearrangements during alignment. *Bioinformatics* 19 (Suppl 1), i54–i62.

- Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr. Protoc. Mol. Biol.* 109, 21.29.1–21.29.9.
- Cameron, R.A., and Davidson, E.H. (2009). Flexibility of transcription factor target site position in conserved cis-regulatory modules. *Dev. Biol.* 336, 122–135.
- Chen, K., Wang, Y., Zhang, R., Zhang, H., and Gao, C. (2019). CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture. *Annu. Rev. Plant Biol.* 70, 667–697.
- Chow, C.-N., Lee, T.-Y., Hung, Y.-C., Li, G.-Z., Tseng, K.-C., Liu, Y.-H., Kuo, P.-L., Zheng, H.-Q., and Chang, W.-C. (2019). PlantPAN3.0: a new and updated resource for reconstructing transcriptional regulatory networks from ChIP-seq experiments in plants. *Nucleic Acids Res.* 47 (D1), D1155–D1163.
- Cohen, O., Borovsky, Y., David-Schwartz, R., and Paran, I. (2014). Capsicum annuum S (CaS) promotes reproductive transition and is required for flower formation in pepper (*Capsicum annuum*). *New Phytol.* 202, 1014–1023.
- Costanzo, E., Trehin, C., and Vandebussche, M. (2014). The role of WOX genes in flower development. *Ann. Bot.* 114, 1545–1553.
- Deveaux, Y., Toffano-Nioche, C., Claisse, G., Thureau, V., Morin, H., Laufs, P., Moreau, H., Kreis, M., and Lecharny, A. (2008). Genes of the most conserved WOX clade in plants affect root and flower development in *Arabidopsis*. *BMC Evol. Biol.* 8, 291.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.
- Dolzblasz, A., Nardmann, J., Clerici, E., Causier, B., van der Graaff, E., Chen, J., Davies, B., Werr, W., and Laux, T. (2016). Stem Cell Regulation by *Arabidopsis* WOX Genes. *Mol. Plant* 9, 1028–1039.
- Dutheil, J.Y., Gaillard, S., and Stukenbrock, E.H. (2014). MafFilter: a highly flexible and extensible multiple genome alignment files processor. *BMC Genomics* 15, 53.
- Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30, 207–210.
- Eshed, Y., and Lippman, Z.B. (2019). Revolutions in agriculture chart a course for targeted breeding of old and new crops. *Science* 366, eaax0025.
- Fornes, O., Castro-Mondragon, J.A., Khan, A., van der Lee, R., Zhang, X., Richmond, P.A., Modi, B.P., Corread, S., Gheorghe, M., Baranasić, D., et al. (2020). JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 48 (D1), D87–D92.
- Frazer, K.A., Pachter, L., Poliakov, A., Rubin, E.M., and Dubchak, I. (2004). VISTA: computational tools for comparative genomics. *Nucleic Acids Res.* 32, W273–W279.
- Fuqua, T., Jordan, J., van Breugel, M.E., Halavaty, A., Tischer, C., Polidor, P., Abe, N., Tsai, A., Mann, R.S., Stern, D.L., and Crocker, J. (2020). Dense and pleiotropic regulatory information in a developmental enhancer. *Nature* 587, 235–239.
- Galli, M., Feng, F., and Gallavotti, A. (2020). Mapping Regulatory Determinants in Plants. *Front. Genet.* 11, 591194.
- Gilchrist, C.L.M., and Chooi, Y.-H. (2020). clinker & clustermap.js: Automatic generation of gene cluster comparison figures. *Bioinformatics*, btab007.
- Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27, 1017–1018.
- Guo, D., Zhang, J., Wang, X., Han, X., Wei, B., Wang, J., Li, B., Yu, H., Huang, Q., Gu, H., et al. (2015). The WRKY transcription factor WRKY71/EXB1 controls shoot branching by transcriptionally regulating RAX genes in *Arabidopsis*. *Plant Cell* 27, 3112–3127.
- Haecker, A., Gross-Hardt, R., Geiges, B., Sarkar, A., Breuninger, H., Herrmann, M., and Laux, T. (2004). Expression dynamics of WOX genes mark cell fate decisions during early embryonic patterning in *Arabidopsis thaliana*. *Development* 131, 657–668.
- Hodgkin, J. (1998). Seven types of pleiotropy. *Int. J. Dev. Biol.* 42, 501–505.
- Holland, P.W.H., Marlétaz, F., Maeso, I., Dunwell, T.L., and Paps, J. (2017). New genes from old: asymmetric divergence of gene duplicates and the evolution of development. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 372, 20150480.
- Honda, E., Yew, C.-L., Yoshikawa, T., Sato, Y., Hibara, K.-I., and Itoh, J.-I. (2018). LEAF LATERAL SYMMETRY1, a Member of the WUSCHEL-RELATED HOMEOBOX3 Gene Family, Regulates Lateral Organ Development Differentially from Other Paralogs, NARROW LEAF2 and NARROW LEAF3 in Rice. *Plant Cell Physiol.* 59, 376–391.
- Hosmani, P.S., Flores-Gonzalez, M., van de Geest, H., Maumus, F., Bakker, L.V., Schijlen, E., van Haarst, J., Cordewener, J., Sanchez-Perez, G., Peters, S., et al. (2019). An improved de novo assembly and annotation of the tomato reference genome using single-molecule sequencing, Hi-C proximity ligation and optical maps. *BioRxiv*, 767764.
- Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780.
- Kyozuka, J., Tokunaga, H., and Yoshida, A. (2014). Control of grass inflorescence form by the fine-tuning of meristem phase change. *Current Opinion in Plant Biology* 17, 110–115. <https://doi.org/10.1016/j.pbi.2013.11.010>.
- Lanfear, R., Hua, X., and Warren, D.L. (2016). Estimating the Effective Sample Size of Tree Topologies from Bayesian Phylogenetic Analyses. *Genome Biol. Evol.* 8, 2319–2332.
- LeBlanc, C., Zhang, F., Mendez, J., Lozano, Y., Chatpar, K., Irish, V.F., and Jacob, Y. (2018). Increased efficiency of targeted mutagenesis by CRISPR/Cas9 in plants using heat stress. *Plant J.* 93, 377–386.
- Leisner, C.P., Hamilton, J.P., Crisovan, E., Manrique-Carpintero, N.C., Marand, A.P., Newton, L., Pham, G.M., Jiang, J., Douches, D.S., Jansky, S.H., and Buell, C.R. (2018). Genome sequence of M6, a diploid inbred clone of the high-glycoalkaloid-producing tuber-bearing potato species *Solanum chacoense*, reveals residual heterozygosity. *Plant J.* 94, 562–570.
- Lemmon, Z.H., Park, S.J., Jiang, K., Van Eck, J., Schatz, M.C., and Lippman, Z.B. (2016). The evolution of inflorescence diversity in the nightshades and heterochrony during meristem maturation. *Genome Res.* 26, 1676–1686.
- Lemmon, Z.H., Reem, N.T., Dalrymple, J., Soyk, S., Swartwood, K.E., Rodriguez-Leal, D., Van Eck, J., and Lippman, Z.B. (2018). Rapid improvement of domestication traits in an orphan crop by genome editing. *Nat. Plants* 4, 766–770.
- Lemoine, F., Domelevo Entfellner, J.-B., Wilkinson, E., Correia, D., Dávila Felipe, M., De Oliveira, T., and Gascuel, O. (2018). Renewing Felsenstein's phylogenetic bootstrap in the era of big data. *Nature* 556, 452–456.
- Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* 1303.3997. <https://arxiv.org/abs/1303.3997>.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.
- Li, M., Wang, R., Liu, Z., Wu, X., and Wang, J. (2019). Genome-wide identification and analysis of the WUSCHEL-related homeobox (WOX) gene family in allotetraploid *Brassica napus* reveals changes in WOX genes during polyploidization. *BMC Genomics* 20, 317.
- Lian, G., Ding, Z., Wang, Q., Zhang, D., and Xu, J. (2014). Origins and evolution of WUSCHEL-related homeobox protein family in plant kingdom. *ScientificWorldJournal* 2014, 534140.
- Lin, H., Niu, L., McHale, N.A., Ohme-Takagi, M., Mysore, K.S., and Tadege, M. (2013). Evolutionarily conserved repressive activity of WOX proteins mediates leaf blade outgrowth and floral organ development in plants. *Proc. Natl. Acad. Sci. USA* 110, 366–371.
- Lippman, Z.B., Cohen, O., Alvarez, J.P., Abu-Abied, M., Pekker, I., Paran, I., Eshed, Y., and Zamir, D. (2008). The making of a compound inflorescence in tomato and related nightshades. *PLoS Biol.* 6, e288.
- Lu, Z., Marand, A.P., Ricci, W.A., Ethridge, C.L., Zhang, X., and Schmitz, R.J. (2019). The prevalence, evolution and chromatin signatures of plant regulatory elements. *Nat. Plants* 5, 1250–1259.

- Meng, Y., Liu, H., Wang, H., Liu, Y., Zhu, B., Wang, Z., Hou, Y., Zhang, P., Wen, J., Yang, H., et al. (2019). HEADLESS, a WUSCHEL homolog, uncovers novel aspects of shoot meristem regulation and leaf blade development in *Medicago truncatula*. *J. Exp. Bot.* **70**, 149–163.
- Meyer, R.S., and Purugganan, M.D. (2013). Evolution of crop species: genetics of domestication and diversification. *Nat. Rev. Genet.* **14**, 840–852.
- Miller, M.A., Pfeiffer, W., and Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop (GCE), 1–8. <https://doi.org/10.1109/GCE.2010.5676129>.
- Mishra, R., Joshi, R.K., and Zhao, K. (2020). Base editing in crops: current advances, limitations and future implications. *Plant Biotechnol. J.* **18**, 20–31.
- Nelson, A.C., and Wardle, F.C. (2013). Conserved non-coding elements and cis regulation: actions speak louder than words. *Development* **140**, 1385–1395.
- Paaby, A.B., and Rockman, M.V. (2013). The many faces of pleiotropy. *Trends Genet.* **29**, 66–73.
- Panchy, N., Lehti-Shiu, M., and Shiu, S.-H. (2016). Evolution of Gene Duplication in Plants. *Plant Physiol.* **171**, 2294–2316.
- Paradis, E., and Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528.
- Park, S.J., Jiang, K., Schatz, M.C., and Lippman, Z.B. (2012). Rate of meristem maturation determines inflorescence architecture in tomato. *Proc. Natl. Acad. Sci. USA* **109**, 639–644.
- Park, S.J., Eshed, Y., and Lippman, Z.B. (2014). Meristem maturation and inflorescence architecture—lessons from the Solanaceae. *Curr. Opin. Plant Biol.* **17**, 70–77.
- Pollard, K.S., Hubisz, M.J., Rosenbloom, K.R., and Siepel, A. (2010). Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* **20**, 110–121.
- Potok, M.E., Wang, Y., Xu, L., Zhong, Z., Liu, W., Feng, S., Naranbaatar, B., Rayatpisheh, S., Wang, Z., Wohlschlegel, J.A., et al. (2019). Arabidopsis SWR1-associated protein methyl-CpG-binding domain 9 is required for histone H2A.Z deposition. *Nat. Commun.* **10**, 3352.
- Powell, A.F., Courtney, L.E., Schmidt, M.H.-W., Feder, A., Vogel, A., Xu, Y., Lyon, D.A., Dumschat, K., McHale, M., Sulpice, R., et al. (2020). A *Solanum lycopersicoides* reference genome facilitates biological discovery in tomato. *BiorXiv*. <https://doi.org/10.1101/2020.04.16.039636>.
- Priest, H.D., Filichkin, S.A., and Mockler, T.C. (2009). Cis-regulatory elements in plant cell signaling. *Curr. Opin. Plant Biol.* **12**, 643–649.
- Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842.
- Radoeva, T., Vaddepalli, P., Zhang, Z., and Weijers, D. (2019). Evolution, Initiation, and Diversity in Early Plant Embryogenesis. *Dev. Cell* **50**, 533–543.
- Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44** (W1), W160–W165.
- R Core Team (2020). R: A language and environment for statistical computing (R Foundation for Statistical Computing).
- Rebeiz, M., and Tsiantis, M. (2017). Enhancer evolution and the origins of morphological novelty. *Curr. Opin. Genet. Dev.* **45**, 115–123.
- Rebocho, A.B., Bliek, M., Kusters, E., Castel, R., Procissi, A., Roobeek, I., Souer, E., and Koes, R. (2008). Role of EVERGREEN in the development of the cymose petunia inflorescence. *Dev. Cell* **15**, 437–447.
- Rockman, M.V. (2012). The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* **66**, 1–17.
- Rodríguez-Leal, D., Lemmon, Z.H., Man, J., Bartlett, M.E., and Lippman, Z.B. (2017). Engineering Quantitative Trait Variation for Crop Improvement by Genome Editing. *Cell* **171**, 470–480.e8.
- Rushton, P.J., Somssich, I.E., Ringler, P., and Shen, Q.J. (2010). WRKY transcription factors. *Trends Plant Sci.* **15**, 247–258.
- Scheben, A., and Edwards, D. (2018). Towards a more predictable plant breeding pipeline with CRISPR/Cas-induced allelic series to optimize quantitative and qualitative traits. *Curr. Opin. Plant Biol.* **45** (Pt B), 218–225.
- Schorderet, M., Duvvuru Muni, R.R., Fiebig, A., and Reinhardt, D. (2018). Dereulation of MADS-box transcription factor genes in a mutant defective in the WUSCHEL-LIKE HOMEOBOX gene EVERGREEN of Petunia hybrida. *Plant Signal. Behav.* **13**, e1471299.
- Schwarzer, W., and Spitz, F. (2014). The architecture of gene expression: integrating dispersed cis-regulatory modules into coherent regulatory domains. *Curr. Opin. Genet. Dev.* **27**, 74–82.
- Shubin, N., Tabin, C., and Carroll, S. (2009). Deep homology and the origins of evolutionary novelty. *Nature* **457**, 818–823.
- Solovieff, N., Cotsapas, C., Lee, P.H., Purcell, S.M., and Smoller, J.W. (2013). Pleiotropy in complex traits: challenges and strategies. *Nat. Rev. Genet.* **14**, 483–495.
- Soyk, S., Lemmon, Z.H., Oved, M., Fisher, J., Liberatore, K.L., Park, S.J., Gorén, A., Jiang, K., Ramos, A., van der Knaap, E., et al. (2017). Bypassing Negative Epistasis on Yield in Tomato Imposed by a Domestication Gene. *Cell* **169**, 1142–1155.e12.
- Springer, N., de León, N., and Grotewold, E. (2019). Challenges of Translating Gene Regulatory Information into Agronomic Improvements. *Trends Plant Sci.* **24**, 1075–1082.
- Stamatakis, A., Ludwig, T., and Meier, H. (2005). RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics* **21**, 456–463.
- Stearns, F.W. (2010). One hundred years of pleiotropy: a retrospective. *Genetics* **186**, 767–773.
- Stern, D.L. (2000). Evolutionary developmental biology and the problem of variation. *Evolution* **54**, 1079–1091.
- Stuurman, J., Jäaggi, F., and Kuhlemeier, C. (2002). Shoot meristem maintenance is controlled by a GRAS-gene mediated signal from differentiating cells. *Genes Dev.* **16**, 2213–2218.
- Swartwood, K., and Van Eck, J. (2019). Development of plant regeneration and Agrobacterium tumefaciens-mediated transformation methodology for *Physalis pruinosa*. *Plant Cell Tissue Organ Cult.* **137**, 465–472.
- Swinnen, G., Goossens, A., and Pauwels, L. (2016). Lessons from Domestication: Targeting Cis-Regulatory Elements for Crop Improvement. *Trends Plant Sci.* **21**, 506–515.
- Tanaka, W., Ohmori, Y., Ushijima, T., Matsusaka, H., Matsushita, T., Kumanomaru, T., Kawano, S., and Hirano, H.-Y. (2015). Axillary Meristem Formation in Rice Requires the WUSCHEL Ortholog TILLERS ABSENT1. *Plant Cell* **27**, 1173–1184.
- Tyler, A.L., Crawford, D.C., and Pendergrass, S.A. (2016). The detection and characterization of pleiotropy: discovery, progress, and promise. *Brief. Bioinform.* **17**, 13–22.
- Van de Velde, J., Van Bel, M., Vaneechoutte, D., and Vandepoele, K. (2016). A Collection of Conserved Noncoding Sequences to Study Gene Regulation in Flowering Plants. *Plant Physiol.* **171**, 2586–2598.
- Van Eck, J., Keen, P., and Tjahjadi, M. (2019). Agrobacterium tumefaciens-Mediated Transformation of Tomato. In *Transgenic Plants: Methods and Protocols*, S. Kumar, P. Barone, and M. Smith, eds. (Humana Press), pp. 171–182.
- Visscher, P.M., and Yang, J. (2016). A plethora of pleiotropy across complex traits. *Nat. Genet.* **48**, 707–708.
- Wagner, G.P., and Zhang, J. (2011). The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. *Nat. Rev. Genet.* **12**, 204–213.
- Wang, Z., Liao, B.Y., and Zhang, J. (2010). Genomic patterns of pleiotropy and the evolution of complexity. *Proc. Natl. Acad. Sci. USA* **107**, 18034–18039.
- Wang, W., Li, G., Zhao, J., Chu, H., Lin, W., Zhang, D., Wang, Z., and Liang, W. (2014). Dwarf Tiller1, a Wuschel-related homeobox transcription factor, is required for tiller growth in rice. *PLoS Genet.* **10**, e1004154.

Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011). A modular cloning system for standardized assembly of multigene constructs. *PLoS ONE* 6, e16765.

Weirauch, M.T., Cote, A., Norel, R., Annala, M., Zhao, Y., Riley, T.R., Saez-Rodríguez, J., Cokelaer, T., Vedenko, A., Talukder, S., et al.; DREAM5 Consortium (2013). Evaluation of methods for modeling transcription factor sequence specificity. *Nat. Biotechnol.* 31, 126–134.

Werner, S., Engler, C., Weber, E., Gruetzner, R., and Marillonnet, S. (2012). Fast track assembly of multigene constructs using Golden Gate cloning and the MoClo system. *Bioeng. Bugs* 3, 38–43.

Wesley, S.V., Helliwell, C.A., Smith, N.A., Wang, M.B., Rouse, D.T., Liu, Q., Gooding, P.S., Singh, S.P., Abbott, D., Stoutjesdijk, P.A., et al. (2001). Construct design for efficient, effective and high-throughput gene silencing in plants. *Plant J.* 27, 581–590.

Wittkopp, P.J., and Kalay, G. (2011). Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.* 13, 59–69.

Wong, E.S., Zheng, D., Tan, S.Z., Bower, N.L., Garside, V., Vanalleghem, G., Gaiti, F., Scott, E., Hogan, B.M., Kikuchi, K., et al. (2020). Deep conservation of the enhancer regulatory code in animals. *Science* 370, eaax8137.

Wu, X., Dabi, T., and Weigel, D. (2005). Requirement of homeobox gene STIMPY/WOX9 for *Arabidopsis* meristem growth and maintenance. *Curr. Biol.* 15, 436–440.

Wu, X., Chory, J., and Weigel, D. (2007). Combinations of WOX activities regulate tissue proliferation during *Arabidopsis* embryonic development. *Dev. Biol.* 309, 306–316.

Wu, C.-C., Li, F.-W., and Kramer, E.M. (2019). Large-scale phylogenomic analysis suggests three ancient superclades of the WUSCHEL-RELATED HO-MEOBOX transcription factor family in plants. *PLoS ONE* 14, e0223521.

Zhang, Z., Runions, A., Mentink, R.A., Kierzkowski, D., Karady, M., Hashemi, B., Huijser, P., Strauss, S., Gan, X., Ljung, K., and Tsiantis, M. (2020). A WOX/Auxin Biosynthesis Module Controls Growth to Shape Leaf Form. *Curr. Biol.* 30, 4857–4868.e6.

Zhu, T., Moschou, P.N., Alvarez, J.M., Sohlberg, J.J., and von Arnold, S. (2014). Wuschel-related homeobox 8/9 is important for proper embryo patterning in the gymnosperm Norway spruce. *J. Exp. Bot.* 65, 6543–6552.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological Samples		
DNA and RNA from various plant tissues from tomato, groundcherry, and <i>Arabidopsis</i> .	See STAR methods	N/A
Chemicals, peptides, and recombinant proteins		
CTAB	Sigma Aldrich	Cat# H6269-500G
Agarose	VWR	Cat# 97062-250
Bsal	NEB	Cat# R0535L
BpI	Thermo Fisher	Cat# ER1012
T4 DNA Ligase	NEB	Cat# M0202L
Taq DNA Polymerase with Standard Taq Buffer	NEB	Cat# M0273L
KOD OneTM PCR Master Mix	TOYOB0	Cat# KMM-101
Fast SYBR Green Master Mix	Applied Biosystems	Cat# 4385612
Critical commercial assays		
TruSeq DNA PCR-Free HT Library Preparation Kit	Illumina	Cat#FC-121-3003
QIAprep Spin Miniprep Kit	QIAGEN	Cat# 27106
QIAquick PCR Purification Kit	QIAGEN	Cat# 28106
StrataClone Blunt PCR Cloning Kit	Stratagene	Cat# 240207
SuperScript IV VILO Master Mix	Invitrogen	Cat# 11756500
TRIzol	Invitrogen	Cat# 15596018
DIG RNA Labeling Kit (SP6/T7)	Roche	Cat# 11175025910
Deposited data		
ATAC-seq sequencing data	This study	GSE164297
Experimental models: cell lines		
N/A	N/A	N/A
Experimental models: organisms/strains		
Tomato (M82)	See STAR methods	N/A
Groundcherry (<i>Physalis grisea</i>)	See STAR methods	N/A
<i>Arabidopsis thaliana</i> (Columbia-0)	See STAR methods	N/A
Oligonucleotides		
Guide RNA (gRNA) sequences, see Table S5C	This study	N/A
Primer sequences for cloning, see Table S5D	This study	N/A
Primer sequences for genotyping, see Table S5D	This study	N/A
Primer sequences for <i>in situ</i> probe, see Table S5D	This study	N/A
Primer sequences for qRT-PCR, see Table S5D	This study	N/A

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Recombinant DNA		
MoClo Toolkit	Weber et al., 2011	Addgene #1000000044
pICH86966::AtU6p::sgRNA_PDS	Belhaj et al., 2013	Addgene #46966
pICH47732::NOSp::NPTII	Belhaj et al., 2013	Addgene #51144
pICH47742::35S::Cas9	Belhaj et al., 2013	Addgene #49771
pHannibal	Wesley et al., 2001	N/A
Software and algorithms		
Trimmomatic	Bolger et al., 2014	http://www.usadellab.org/cms/?page=trimmomatic
BWA-MEM	Li, 2013	http://bio-bwa.sourceforge.net/
PicardTools	N/A	http://broadinstitute.github.io/picard
PartitionFinder2	Lanfear et al., 2016	http://www.robertlanfear.com/partitionfinder/
BOOSTER	Lemoine et al., 2018	https://booster.pasteur.fr/
R	R Core Team, 2020	https://www.r-project.org/
phyloP	Pollard et al., 2010	http://compgen.cshl.edu/phast/phyloFit-tutorial.php
Geneious Prime® 2020.1.1	N/A	http://www.geneious.com/
MAFFT v7.402	Katoh and Standley, 2013	https://mafft.cbrc.jp/alignment/software/
CIPRES	Miller et al., 2010	http://www.phylo.org/
RAxML v8.2.12	Stamatakis et al., 2005	https://cme.h-its.org/exelixis/web/software/raxml/
Bedtools	Quinlan and Hall, 2010	https://bedtools.readthedocs.io/en/latest/index.html
Samtools	Li et al., 2009	http://www.htslib.org
Clinker v0.0.12	Gilchrist and Chooi, 2020	https://github.com/gamcil/clinker
mVISTA	Frazer et al., 2004	http://genome.lbl.gov/vista/mvista/about.shtml
Shuffle-LAGAN	Brudno et al., 2003	http://lagan.stanford.edu/lagan_web/index.shtml
JASPAR 2020 Core Plants collections	Fornes et al., 2020	http://jaspar.genereg.net/
FIMO	Grant et al., 2011	https://meme-suite.org/
MEME-suite	Bailey et al., 2009	https://meme-suite.org/
lastz 1.04.01		http://www.bx.psu.edu/~rsharris/lastz/README.lastz-1.04.03.html#history_recent
Genrich V0.6	N/A	https://github.com/jsh58/Genrich
AVID 2.1	Bray et al., 2003	http://toucan.aertslab.org/software/help/WebServices/vista.htm
Multiz	Blanchette et al., 2004	https://hgdownload-test.gi.ucsc.edu/goldenPath/mm9/multiz30way/multiz30way.html
deepTools	Ramírez et al., 2016	https://deeptools.readthedocs.io/en/develop/index.html
STAR-2.6.1.d	Dobin et al., 2013	https://github.com/alexdobin/STAR/tree/2.6.1
Conservatory	This study	https://github.com/idanefroni/Conservatory
maffilter 1.3.1	Dutheil et al., 2014	https://jydu.github.io/maffilter/Manual/index.html

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Zachary B. Lippman (lippman@cshl.edu).

Materials availability

This study did not generate new unique reagents. Plasmids and transgenic plants generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

Data and code availability

The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus ([Edgar et al., 2002](#)) and are accessible through GEO Series accession number GSE164297 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE164297>). Software is available from <https://github.com/idanefroni/Conservatory>.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Plant materials and growth conditions

Seeds of *Physalis grisea* and *Solanum lycopersicum* cv. M82 (LA3475), and the *s/wox9* (*Solyc02g077390*) mutants (i.e., *s-classic*, *s-multiflora*, *s-n5568*, *s/wox9^{pro}*), and *anantha-e1546* (*Solyc02g081670*) in M82 background were our own stocks. Unless otherwise indicated, seeds were directly sown into soil in 96-cell plastic flats and grown in a greenhouse under long-day conditions (16-hr light/8-hr dark) supplemented with artificial light from high-pressure sodium bulbs (~250 $\mu\text{mol m}^{-2} \text{s}^{-1}$). The temperature ranged between 26–28°C during the day to 18–20°C during the night, with a relative humidity of 40%–60%. One-month old seedlings were either transplanted to the fields at Cold Spring Harbor Laboratory (CSHL), Cold Spring Harbor, NY, or transplanted to a 4L pot and grown in the greenhouse at the same conditions described above. Plants were grown under drip irrigation and standard fertilizer regimens. *Arabidopsis thaliana* Columbia ecotype seeds were used for *Arabidopsis* experiments. Also used were the *Arabidopsis* mutants *stip-1* (AT2G33880), *stpl-1* (AT5G45980), and *stpl-1/stip-1:stip-2/+*, which were previously published ([Wu et al., 2005, 2007](#)). Seeds were bleach sterilized, cold treated for three days, sown onto ½ MS media and grown for five to ten days under long-day conditions (16-hr light/8-hr dark with ~100 $\mu\text{mol m}^{-2} \text{s}^{-1}$ light, 22°C constant temperature). Plants were then transplanted to soil for further growth in a controlled growth chamber under the same long-day conditions and were phenotyped.

METHOD DETAILS

Plant phenotyping

We define vegetative trait phenotypes as shoot apical meristem (SAM) termination, cotyledon defects (multiple cotyledons, lobed cotyledons, and fused cotyledons), leaf fusion, and phyllotaxis defects. Vegetative trait phenotyping was done on young seedlings at various stages of vegetative growth. In order to determine if a mutant allele was embryonic lethal, an F2/T1 population of at least 40 seedlings was screened and genotyped. Tomato inflorescence complexity was measured by counting the number of branching events per inflorescence in F2 derived plants growing in CSHL fields. Stereoscope images of young inflorescences were imaged with a Nikon SMZ-25 stereo zoom microscope (Nikon). Quantification of groundcherry side shoots per growth unit and petal number were conducted on T1 plants growing at CSHL fields and greenhouses. For *pgwox9^{pro-Reg1}* SAM termination analysis, seeds were incubated for three to seven days on a Petri dish with moistened filter paper in the dark at 28°C until germination (marked by root meristem emergence). Germinated seeds were transplanted to soil and two weeks after transplanting, arrested seedlings were counted as having SAM termination. To phenotype known *wox9* and *wox8* mutant vegetative defects (germination, root apical meristem (RAM) and SAM termination), *Arabidopsis* seeds were sown onto ½ MS media with 1% agar (Phyto Technology Laboratories) and grown vertically for five days under long-day conditions before imaging with a Nikon SMZ-25 stereo zoom microscope (Nikon), or ten days before phenotyping and transplanting to soil for further growth in a controlled growth chamber under long-day conditions. Inflorescences were imaged at five weeks old.

Genome sequencing

The *s-n5568*, a *s/wox9* *cis*-regulatory inversion mutation allele ([Lippman et al., 2008](#)), was determined by genomic sequencing ([Soyk et al., 2017](#)). Briefly, reads for the *compound inflorescence* (*s/slw9*) mutant were trimmed by quality using Trimmomatic v0.32 (HiSeq2500 parameters: ILLUMINACLIP: TruSeq3-PE-2.fa:2:40:15:1:FALSE LEADING:30 TRAILING:30 MINLEN:75 TOPHRED33) ([Bolger et al., 2014](#)) and aligned to the tomato reference genome sequence (Heinz, SL4.0) ([Hosmani et al., 2019](#)) using bwa mem -M ([Li, 2013](#)). Breakpoint coordinates were then validated by Sanger sequencing (see primers in [Table S5D](#)).

Phylogenetic analyses and sequence analyses

To reconstruct the evolutionary history of the WOX8/9 lineage, amino acid translations of the 15 WOX genes (WUS + WOX1-14) in *Arabidopsis thaliana* were used as BLAST search queries against NCBI, SOL Genomics, Phytozome and JGI to identify putative WOX genes from other species with fully sequenced genomes. Amino acid gene models for *Solanum chacoense* were obtained from Leisner et al. (2018), and for *S. lycopersicoides* from Powell et al. (2020). Putative WOX orthologs were retained following reciprocal best-BLAST results against the *Arabidopsis* sequences. Sequences were aligned using MAFFT v7.402 (Katoh and Standley, 2013) via CIPRES (Miller et al., 2010), and the best scoring ML tree inferred using default parameters in RAxML v8.2.12 (Stamatakis et al., 2005) via CIPRES. Phylogenetically informative sequences included in the WOX8/9 clade with strong support were then retained for a more focused analysis. For this WOX8/9 gene tree, the best scoring ML tree was inferred using RAxML as before, except the number of bootstraps was increased to 1000 and a partition file (following PartitionFinder2 (Lanfear et al., 2016) via CIPRES) was included. To further assess support, the best scoring and all bootstrap trees were used in BOOSTER to calculate TBE values (Lemoine et al., 2018). Trees were visualized in R package ape (Paradis and Schliep, 2019). Synteny analysis of *Arabidopsis* WOX9 and WOX8 surrounding genomic regions to corresponding genomic regions in tomato and pepper was performed using Clinker v0.0.12 (Gilchrist and Chooi, 2020; Wu et al., 2019).

CRISPR-Cas9 mutagenesis, plant transformation, and selection of mutant alleles

CRISPR-Cas9 mutagenesis and generation of transgenic tomato and groundcherry plants were performed following our standard protocol (Brooks et al., 2014; Van Eck et al., 2019; Lemmon et al., 2018; Swartwood and Van Eck, 2019). Briefly, guide RNAs (gRNAs) were designed using the Geneious Prime software (<https://www.geneious.com/>) (all gRNAs used in this study are listed in Table S5C). The Golden Gate cloning system was used to assemble the binary vector containing the Cas9 and the specific gRNAs (Rodríguez-Leal et al., 2017; Soyk et al., 2017; Werner et al., 2012). Final binary vectors were then transformed into the tomato cultivar M82 or groundcherry by *Agrobacterium tumefaciens*-mediated transformation through tissue culture (Van Eck et al., 2019; Swartwood and Van Eck, 2019). First-generation transgenic plants (T_0) were genotyped with specific primers surrounding the target sites. To generate new alleles and purify them from potential spontaneous mutations or CRISPR-Cas9 off-target effects following plant transformation, all CRISPR-Cas9 T_0 transgenic lines were backcrossed to parental wild-type M82 cultivar (WT) plants. These F1 populations were then screened for plants lacking the Cas9 transgene, and PCR products of the targeted regions were sequenced to confirm inheritance of alleles. Selected F1 plants were self-fertilized to generate F2 populations, and these segregating populations were used to validate the phenotypic effects of each allele by co-segregation. F2 or F3 homozygous mutant plants were then used for quantitative phenotypic analyses. Screening for the embryonic lethal phenotype and validation of expected segregation ratios was performed on F2 families. The targeting of the NAC/WRKY TFBS in the tomato pro-Reg1 cis-regulatory region was performed using *s/wox9^{pro}* gRNA2 (Table S5C). For groundcherry mutant allele characterization, T1 homozygous progeny plants were used. For *Arabidopsis*, we cloned three level 0 constructs; the AtUBQ10 promoter (AT4G05320) into pICH41295, zCas9 into pICH41308, and the rubisco terminator rbsSE9t into pICH41276. The three constructs were assembled into the level 1 vector pICH47742. For selection, we cloned the *CRUCIFERIN 3* (AT4G28510) promoter into pICH41295, Venus into pAGM1287, N7 NLS signal into pAGM1301 and the nos terminator into pICH41276, followed by assembly into the level 1 vector pICH47751. Both level 1 plasmids were assembled into the level 2 vector pAGM4723 together with the kanamycin resistance cassette at position 1. gRNAs were cloned and assembled into the level 2 vector as described above. Fluorescence seed selection was used to isolate transformed plants. For the pLFY::WOX9-RNAi construct, a 410 bp region from the second exon of the WOX9 gene (AT2G33880, Chr2: 14343186-14342777) was cloned in both sense and antisense orientation into the pHANNIBAL vector (Wesley et al., 2001), which then was shuttled into a binary vector containing the previously described 2.3 kb *Arabidopsis LEAFY* (LFY, AT5G61850) promoter (Blázquez et al., 1997). All *Arabidopsis* plants were transformed using agrobacterium-mediated floral dip.

Characterization of cis-regulatory mutant alleles

In order to rapidly generate large collections of diverse promoter alleles, we used our CRISPR-Cas9 mutagenesis drive system (Rodríguez-Leal et al., 2017). Briefly, Cas9 positive F1 *s/wox9^{pro}* plants were grown in the field at the Uplands Farm of Cold Spring Harbor Laboratory, New York. Seeds were collected for all F1 plants showing inflorescence branching, indicating trans-targeting of the WT *S/WOX9* promoter allele inherited from the WT M82 parent. F2 progeny plants were genotyped for both Cas9 and the *S/WOX9* targeted-promoter region to identify non-transgenic, biallelic, and homozygous plants carrying new alleles. This resulted in the isolation of an additional 25 alleles, among which two were embryonic lethal and could not be recovered as homozygotes (Figures 2A and S2B–S2D). Sequencing of new alleles was performed for at least three cloned individuals per family (StrataClone Blunt PCR Cloning Kit, Stratagene). Sequence assembly was carried out using Geneious Prime software, using modified parameters enabling big deletions. Ten T_0 mutant plants for both *s/wox9^{pro-Reg1}* and *s/wox9^{pro-Reg2}* promoter specific targeted regions were crossed to WT plants and selfed to generate F1 and T1 progeny plants respectively. Non-transgenic T1 and F1 plants were then genotyped, cloned, and sequenced to characterize the inherited allele. In the case of *s/wox9^{pro-Reg1-1}*, genotyping of 149 T1 seedlings showed distorted allele segregation of *S/WOX9*: *s/wox9^{pro-Reg1-1/+}*: *s/wox9^{pro-Reg1-1/-}* 61:88:0 (ratio of 1/3 – 2/3 of WT:heterozygous) suggesting embryo lethality, which was then validated in an F2 population (Figures 3D and 3E). The same mutant allele characterization approach was carried out for the groundcherry *pgwox9^{CR-CDS}* and *pgwox9^{pro-Reg1}*, and *pgwox9^{pro-Reg2}* mutant alleles (Figures 4H and S4G–S4I). For *Arabidopsis wox9^{pro-proximal}* and *wox9^{pro-distal}* more than 200 T1 plants and 1000 T2 plants for each targeted region were grown and

genotyped. More than 40 plants were cloned and sequenced; however, we were able to isolate only two individual homozygous plants that were heritable (*wox9^{pro-proximal-1}* and *wox9^{pro-distal-1}*). All other sequenced plants were not inherited, despite having high Cas9 activity which resulted in high chimerism (Figure S5G). Specific primers/CAPS-markers were designed and used for further genotyping each mutant allele (all primers used in this study are listed in Table S5D).

Visualization of the aligned mutant alleles was done in Illustrator (Adobe Illustrator CS6) (Figures 1F, S2D, S4G, and S5G). The encoding allele visualization was done in R and Perl scripts. For each allele, the ratio-change in the WT sequence of a 20 bp sliding window was calculated in Perl, and scored between 0 (no change) and 1 (fully changed). Insertions were considered as a 1 bp deletion and inversions were considered as a 2 bp deletion (1 bp deletion at each breakpoint). The allele ratio-change map was then plotted in R as a heatmap (Figures 2A, 3B, 3D, and 4H). The JASPAR 2020 Core Plants collections (Forbes et al., 2020) was used to predict transcription factor binding sites (TFBSs) in tomato and groundcherry *WOX9* promoter regions by FIMO (Grant et al., 2011), with a p value < 0.00001 and a q-value < 0.01 cutoffs (Figures 3B, 3D, S2D, and S4G).

Assay for Transposase-Accessible Chromatin using sequencing (ATAC-Seq) libraries

ATAC-seq was performed as described in (Potok et al., 2019) with small modifications. Briefly, six biological replicates of 1g of freshly collected *anantha* inflorescences (meristematic enriched tissue) were ground in ice-cold isolation buffer (300 mM sucrose, 20 mM Tris pH 8.0, 5 mM MgCl₂, 5 mM KCl, 5 mM beta-mercaptoethanol, 35% glycerol, 0.2% Triton X-100). After grinding, the suspension was filtered through a series of cell strainers (the smallest being 70 mm). Samples were further disrupted using a Dounce homogenizer, and then were washed four times with ice-cold isolation buffer. The enriched nuclei samples were resuspended in Resuspension Buffer (50 mM Tris pH 8.0, 5 mM MgCl₂, 5 mM beta-mercaptoethanol, 20% glycerol). In order to continue with 100,000 nuclei, a small portion of the nuclei enriched sample was stained with 4,6-Diamidino-2-Phenylindole (DAPI), and counted under a microscope using a hematocytometer. For the transposase reaction 100,000 nuclei were resuspended in 22.5 mL freezing buffer (5 mM MgCl₂, 0.1 mM EDTA, 50 mM TRIS pH 8.0, 40% glycerol), and were mixed with 2X DMF Buffer (66 mM Tris-acetate (pH 7.8), 132 K-Acetate, 20 mM Mg-Acetate, 32% DMF) and 2.5ul Transposase enzyme from Illumina Nextera Kit (Catalog No. FC-121-1031), and incubated at 37°C for 30 min. Tagmented DNA was isolated using NEB Monarch PCR Purification kit (Catalog No. NEB #T1030). Next-generation sequencing (NGS) libraries were amplified in two steps (Buenrostro et al., 2015) using KAPA HiFi HotStart ReadyMix PCR kit (Catalog No. KK2601). NGS libraries were PE50 sequenced on an Illumina MiSeq SY-410-1003 instrument. ATAC-seq reads were trimmed with trimmomatic (Bolger et al., 2014) and aligned to Heinz SL4.0 reference genome (Hosmani et al., 2019) with BWA mem. Reads that mapped to either chloroplast or mitochondria were filtered out together with PCR duplicate reads (picard MarkDuplicates) and multiple position aligned reads. Finally, only high quality properly paired reads were retained for the analysis (samtools view -b -h -f 3 -F 4 -F 8 -F 256 -F 1024 -F 2048 -q 30). For Jbrowse visualization of the ATAC-seq, coverage was calculated with genomeCoverageBed followed by bedGraphToBigWig command. ATAC-seq peak calling was done by Genrich V0.6 (<https://github.com/jsh58/Genrich>) using all six libraries with the following parameters: -j -r -v -q 0.5 -g 30. The deepTools (Ramírez et al., 2016) were used to calculate Pearson correlation analysis for the called ATAC-seq peak regions, and the six biological-replicates (Figure S3B). Further analyses of the ATAC-seq peaks were done with deepTools and R. ATAC-seq peak size, peak distributions surrounding gene bodies, and annotation distributions are shown in Figures S3C-S3E.

Identification of conserved non-coding sequences (CNSs), Conservatory

Identification of conserved cis-regulatory sequences was performed by comparing upstream and downstream sequences among orthologs in different species (Figure S4A). First, genes were assembled into putative orthogroups. To this end, we performed reciprocal BLAST for each protein sequence in the reference genome with the other family genomes (blastp –evalue 10; Table S4A) and up to eight genes with the lowest combined alignment E-value were selected as putative orthologs for every genome. To remove spurious BLAST hits, all putative orthologs with –log(E-value) less than one third of the maximum –log(E-value) for the putative orthogroup were removed.

A database of upstream *cis*-regulatory regions was compiled by extracting the sequences upstream to the translation start site (TSS) either to the nearest gene or up to a length of 50 kb, which ever was shorter. Similarly, downstream regulatory regions were extracted from the stop codon until the next gene, or 5 kb, which ever was shorter.

To identify the orthogroups from the putative ortholog list, the regulatory region of each putative ortholog for each genome was aligned against the regulatory region of each gene from the reference genome using AVID 2.1 (Bray et al., 2003). Gene pairs having a lower quality score than that detailed in Table S4A were discarded. If there were multiple genes from different genomes of the same species, the species genome with the highest quality score was selected. If there were multiple putative orthologs that had regulatory region alignments of sufficient scores from the same genome, they were all included in the orthogroup. After the selection of the orthogroup, every gene regulatory sequence was aligned against the regulatory sequence of the ortholog from the reference genome using lastz 1.04.01 (lastz –format = maf –gap = 400,50 –nochain –seed = match3 –gapped –strand = both –step = 1 –identity 60 –ydrop = 500 –hsptreshold = 2500 –gappedthresh = 2500 –inner = 1800). Resulting gene-gene alignments for the entire orthogroup were combined using Multiz v11.2 (Blanchette et al., 2004) into a single alignment file in MAF format for each orthogroup. Alignments were filtered using maftfilter 1.3.1 (Dutheil et al., 2014) using the commands “Concatenate(), MinBlockLength(min_length=10), Merge(dist_max=4), MinBlockSize(min_size=4), EntropyFilter(), QualFilter()” to produce the final alignment file.

To derive the conservation p value for each base pair and identify conserved base pairs, a phylogenetic tree of the species in each family was constructed using alignment of common protein sequences, and a substitution model was fitted using phyloFit by using just the 3rd variant base pair of the codons of the aligned proteins. phyloP (Pollard et al., 2010) was then applied using the parameters:–seed 123 –wig-scores –no-prune –mode CON –method SCORE. phyloP scores of less than 2 for Solanaceae or less than 1.2 for Brassicaceae were filtered. The coordinates of the conserved base pairs determined by phyloP were then remapped to the reference genome and concatenated together for all genes into a single BED file. To identify CNSs, regions of conservation within 25 bp of each other were merged (bedtools merge –d 25) and regions less than 10 bp in length were filtered out. Analyses of the upstream and downstream distributions (*cis*-regulatory region size distribution, CNS distribution in *cis*-regulatory regions, CNS length, and the calculated conservation percentage of the *cis*-regulatory regions) were performed in R for both Solanaceae and Brassicaceae (Figures S4B–S4E and S5A–S5D) (see also Table S4A for additional information). To identify conserved CNSs between plant families, we extracted the consensus sequence for all the CNSs in the *WOX9* promoters for each family separately. The CNSs were then aligned using permissive alignment (–threshold = 1000 –ydrop = 500 –gap = 100,40 –seed = match3) to identify deeply conserved regions. All code was written in Perl and will be made available on github following publication. For the mVISTA analysis, the 40 kb sequence surrounding *WOX9* in the indicated species was scored for sequence homology to *S/WOX9* or *PgWOX9* using mVISTA Shuffle-LAGAN aligner with the following parameters: windows of 70 bp at a similarity threshold of 70% (Brudno et al., 2003; Frazer et al., 2004).

Tissue collection and RNA extraction

Reproductive shoot apices (~4 mm apices of axillary inflorescences collected within days of emergence) or inflorescence apices (~2 mm primary inflorescence apices including meristem and small floral buds collected when the inflorescence had at least one silique and eight flowers) were collected from n = 20 plants at seven and six weeks old, respectively. Tissue was frozen, ground with beads, and RNA was extracted with TRIzol (Invitrogen). RNA was quantified by Nanodrop One spectrophotometer (Thermo Scientific), normalized, DNase treated (ezDNase, Invitrogen), and cDNA synthesis was performed with Superscript IV Vilo (Invitrogen). qPCR was performed with Fast SYBR Green Master Mix (Applied Biosystems) on a QuantStudio 6 Real-Time PCR machine (Applied Biosystems). Transcripts were normalized to the internal control *UBC28* (AT1G64230) (all primers used in this study are listed in Table S5D). For *in situ* hybridization, tomato apices containing young inflorescences were dissected and fixed in 4% paraformaldehyde with 0.3% Triton X-100.

S/WOX9 expression analysis

Published RNA-seq data of wild-type M82 and *s-n5568* mutant meristems were downloaded from SRA PRJNA376115, and PRJNA343677 (Lemmon et al., 2016; Park et al., 2012; Soyk et al., 2017). Reads were trimmed by quality using Trimmomatic (ILLUMINA_CLIP:TruSeq2-PE.fa:2:30:10:1:FALSE LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36) and aligned to the cDNA annotation of reference genome sequence of tomato (SL4.0) using STAR v2.6.1.d (Dobin et al., 2013). Normalization and quantification of *S/WOX9* expression was done in R by calculating transcripts per million (TPM).

In situ hybridization

non-radioactive mRNA *in situ* hybridization was done as described previously (Lippman et al., 2008; Lemmon et al., 2016). Antisense *S/WOX9* RNA probe were transcribed with DIG RNA Labeling Kit (SP6/T7) (Roche).

QUANTIFICATION AND STATISTICAL ANALYSES

Detailed methods for CNSs and ATAC-seq peak analyses can be found in “Identification of conserved non-coding sequences (CNSs), Conservatory” and “ATAC-sequencing (ATAC-seq) libraries” respectively. All statistical tests were performed in R, and significance is defined as a p value less than 0.05 shown by the following code: * < 0.05, ** < 0.1, and *** < 0.001. For quantitative analysis of inflorescence complexity in Figures 1B, 2B, and 3B, n = “number of inflorescences quantified” was used for pairwise comparisons using Dunnett’s T3 test for multiple comparisons with unequal variances with default parameters (see Tables S1B, S2A, and S3A for raw data and statistical analysis data). For quantitative analysis of vegetative trait phenotypes, such as cotyledon defects in Figure 1E and other vegetative defects in Figure 2C, each plant was counted once regardless of how many vegetative trait defects it showed (cotyledon defects, leaf or stem fusion, and phyllotaxis defects). For SAM termination phenotypes shown in Figure 2C, seedlings were counted once for SAM termination and once for other vegetative defects. n = “number of seedlings quantified” was used for pairwise comparisons using Dunnett’s T3 test for multiple comparisons with unequal variances with defaults parameters (see Tables S1C, S2B, and S5A for raw data and statistical analysis data). For quantitative analysis of groundcherry side shoots per growth unit, and petal number per flower, in Figures 4H and S4I, n = “number of growth units” and n = “flower number” respectively (see Tables S4B and S4C for raw data and statistical analysis data). For qRT-PCR expression analysis, three biological replicates of pooled reproductive shoot apices, or inflorescence apices, were used for each genotype and three technical replicates were performed for each biological replicate. Mean values of normalized expression were compared using two-tailed, two sample t tests. Means ± s.d. were shown and mean values between groups were compared by two-sample t tests.

ADDITIONAL RESOURCES

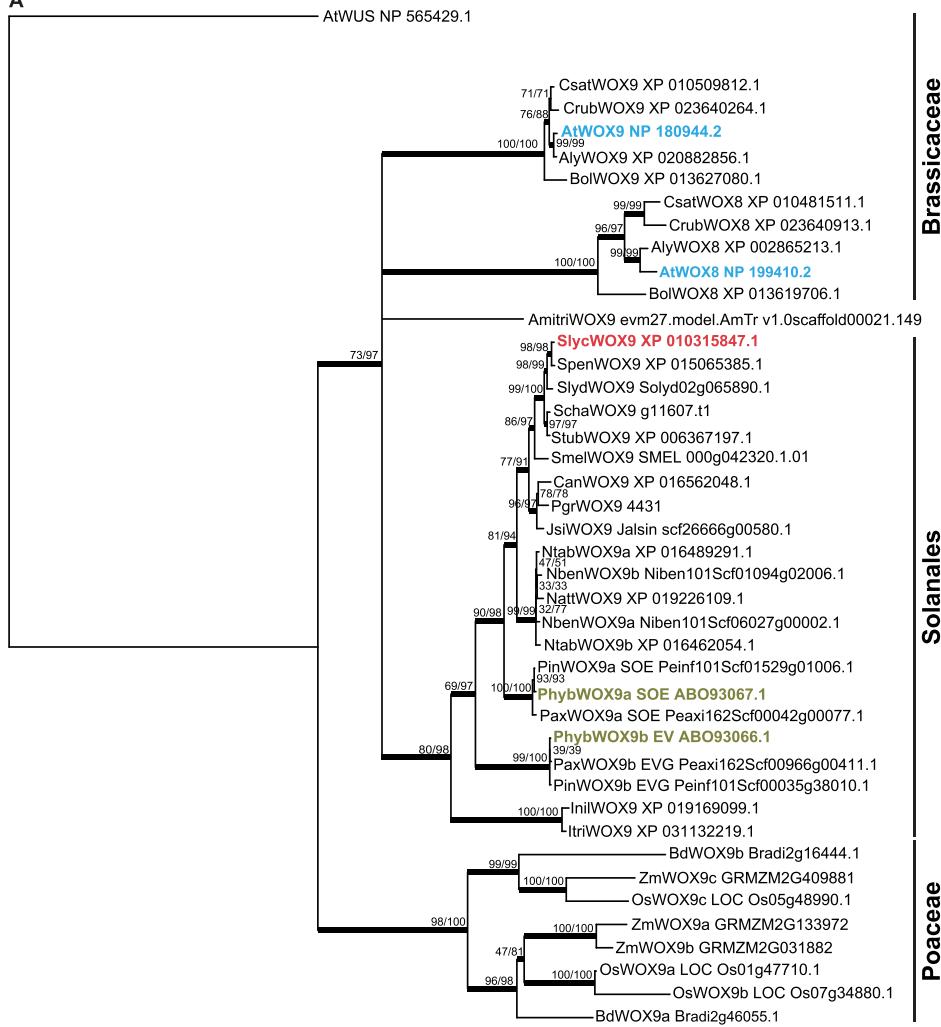
Sequence retrieval: <https://phytozome.jgi.doe.gov/pz/portal.html>

Data deposition: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE164297>

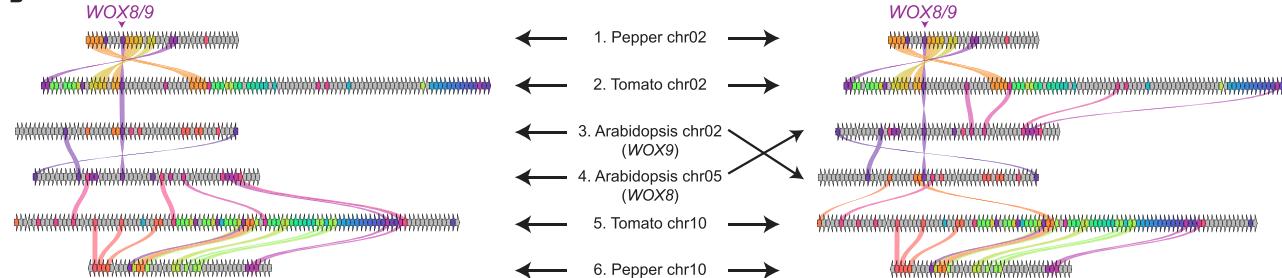
Conservatory: <https://github.com/idanefroni/Conservatory>

Supplemental figures

A



B



1. *Capsicum annuum glabriusculum* (V2.0), Chr02:129542918..129972016
2. *Solanum lycopersicum* 'Heinz 1706' (ITAG4.0), SL4.0Ch02:40291255-40482899
3. *Arabidopsis thaliana* (TAIR10), Chr02:14304893-14395067

4. *Arabidopsis thaliana* (TAIR10), Chr05:18613242-18700774
5. *Solanum lycopersicum* 'Heinz 1706' (ITAG4.0), SL4.0Ch10:2700445-2885864
6. *Capsicum annuum glabriusculum*, Chr10:12645877-15246401

(legend on next page)

Figure S1. Phylogeny of WOX8 and WOX9 proteins in selected plant families and synteny analysis showing loss of the tomato S/WOX8 paralog, related to Figure 1

(A) Phylogenetic tree of WOX8 and WOX9 proteins. Numbers at each node are bootstrap values (%) from 1000 replicates/TBE values (see [Table S1A](#) and see [STAR methods](#)). Thicker branches indicate TBE values higher than 75. Clades with less than 50% bootstrap support were collapsed.

(B) Clinker analyses of WOX8/9 synteny regions from *Arabidopsis*, tomato and pepper. The synteny diagrams to the left and right show reciprocal comparisons of *Arabidopsis* chromosome 2 (*WOX9*) and 5 (*WOX8*) with tomato and pepper chromosomes 2 and 10. Syntenic genes are represented by colored arrows and are connected by lines in adjacent chromosomal segments. Colored genes in non-adjacent chromosome segments are also syntenic, but are not connected by lines. Chromosome 2 of tomato and pepper carry a *WOX8/9* homolog and are syntenic to both *Arabidopsis* chromosome 2 and 5. Chromosome 10 of tomato and pepper are also syntenic to *Arabidopsis* chromosome 2 and 5, but lack a *WOX8/9* gene, indicating loss of one *WOX8/9* homolog in the Solanaceae lineage. Chromosomal coordinates for each species are listed.

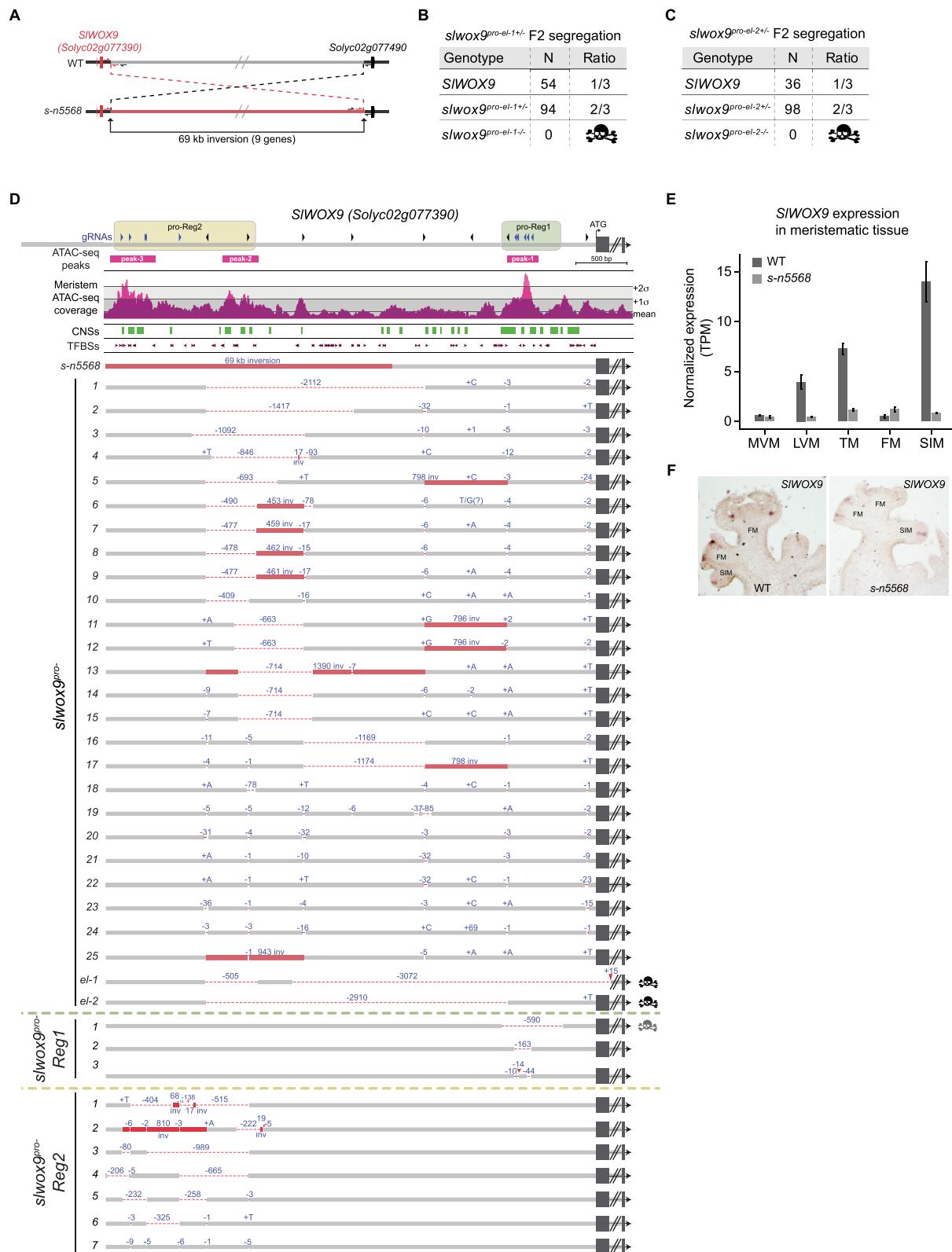


Figure S2. Characterization of *SIWOX9* cis-regulatory mutant alleles, related to Figure 2

(A) The *s-n5568* mutation is a 63 kb inversion, with the first break point proximal to the gene located 2,034 bp upstream to *SIWOX9* (see STAR methods).

(legend continued on next page)

(B-C) Summary of *slwox9^{pro-el-1}* F2 segregation (B) and *slwox9^{pro-el-2}* F2 segregation (C).

(D) Schematic of the *SlWOX9* regulatory region targeted by CRISPR-Cas9. gRNAs (black arrowheads mark the eight gRNAs used in the original *SlWOX9* promoter mutagenesis; blue arrowheads mark the gRNAs used to target the specific *cis*-regulatory regions pro-Reg1 and pro-Reg2). *SlWOX9^{pro-Reg1}* and *SlWOX9^{pro-Reg2}* targeted regions are marked and highlighted with green and yellow colors, respectively. Meristem ATAC-seq coverage is shown, along with peak calls. Gray lines and color intensities mark mean and standard deviations. Peaks are numbered sequentially from ATG (see also [STAR methods](#)). Solanaceae CNS blocks are shown in green (see [Figures 4](#) and [S4](#) and [STAR methods](#)) and TFBSS with a q-value > 0.01 are shown (see [STAR methods](#)). Shown are 27 *slwox9^{pro}* alleles, including the 69 kb inversion allele s-n5568. The three *slwox9^{pro-Reg1}* alleles are shown beneath the green dashed line. Note that the *slwox9^{pro-Reg1-1}* allele causes high penetrance of embryonic lethality (gray skull). The seven *slwox9^{pro-Reg2}* alleles are shown beneath the yellow dashed line. Deletions (-) are indicated by red dashed lines, insertions by plus signs (+), inversions by red boxes.

(E) RNA-seq analysis of *SlWOX9* expression from meristems of WT and the s-n5568 promoter inversion mutant showing unchanged expression in vegetative meristems and ~80% lower expression in reproductive meristems. MVM, Middle Vegetative Meristem; LVM, Late Vegetative Meristem; TM, Transition Meristem; FM, Floral Meristem; SIM, Sympodial Inflorescence Meristem. Data are from [Park et al. \(2012\)](#).

(F) mRNA *in situ* hybridization showing that *SlWOX9* expression domains in reproductive meristems (SIM, FM) of s-n5568 are unchanged compared to WT. Similar *SlWOX9* expression patterns in reproductive meristems were observed in [Lemmon et al. \(2016\)](#).

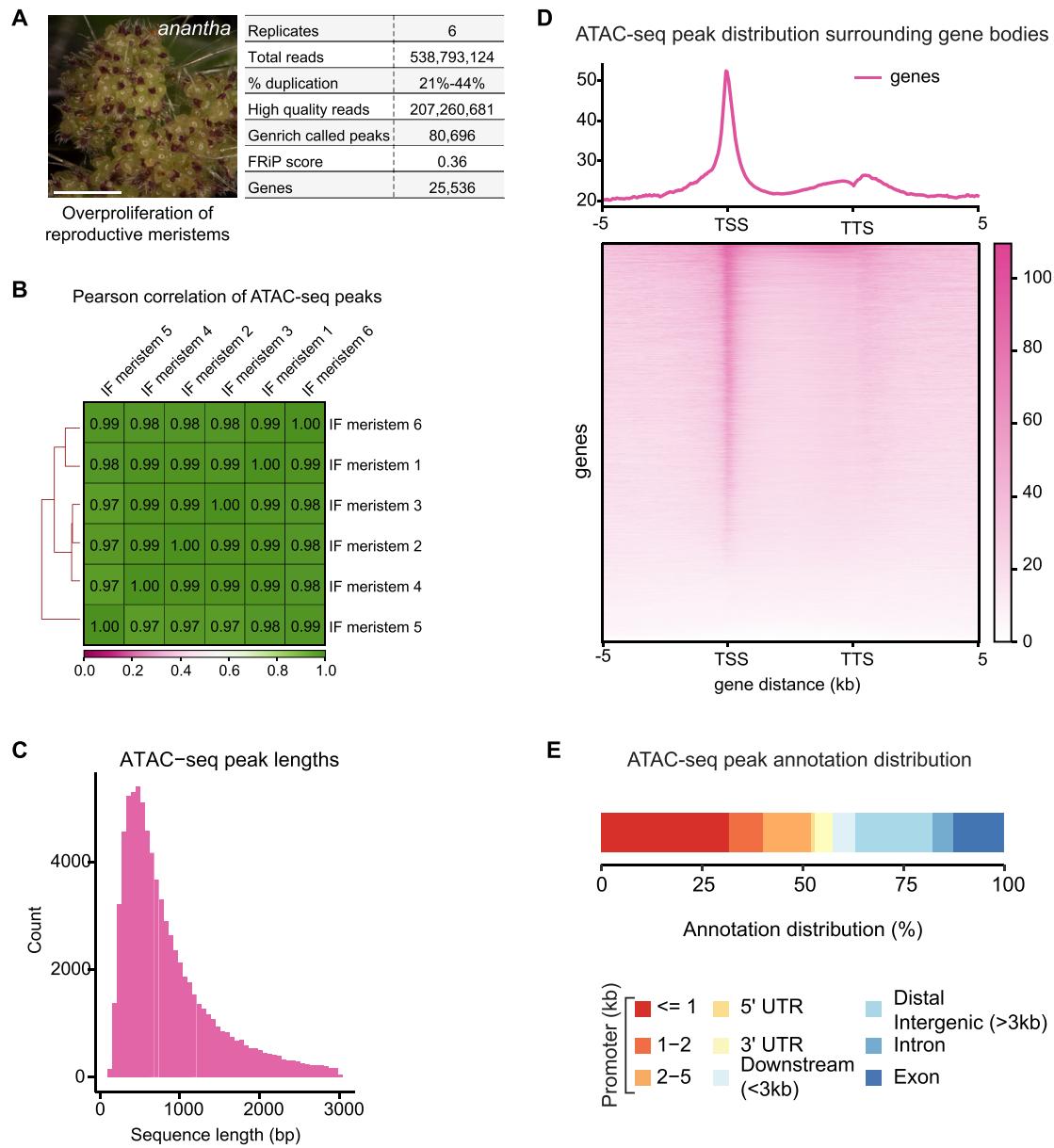
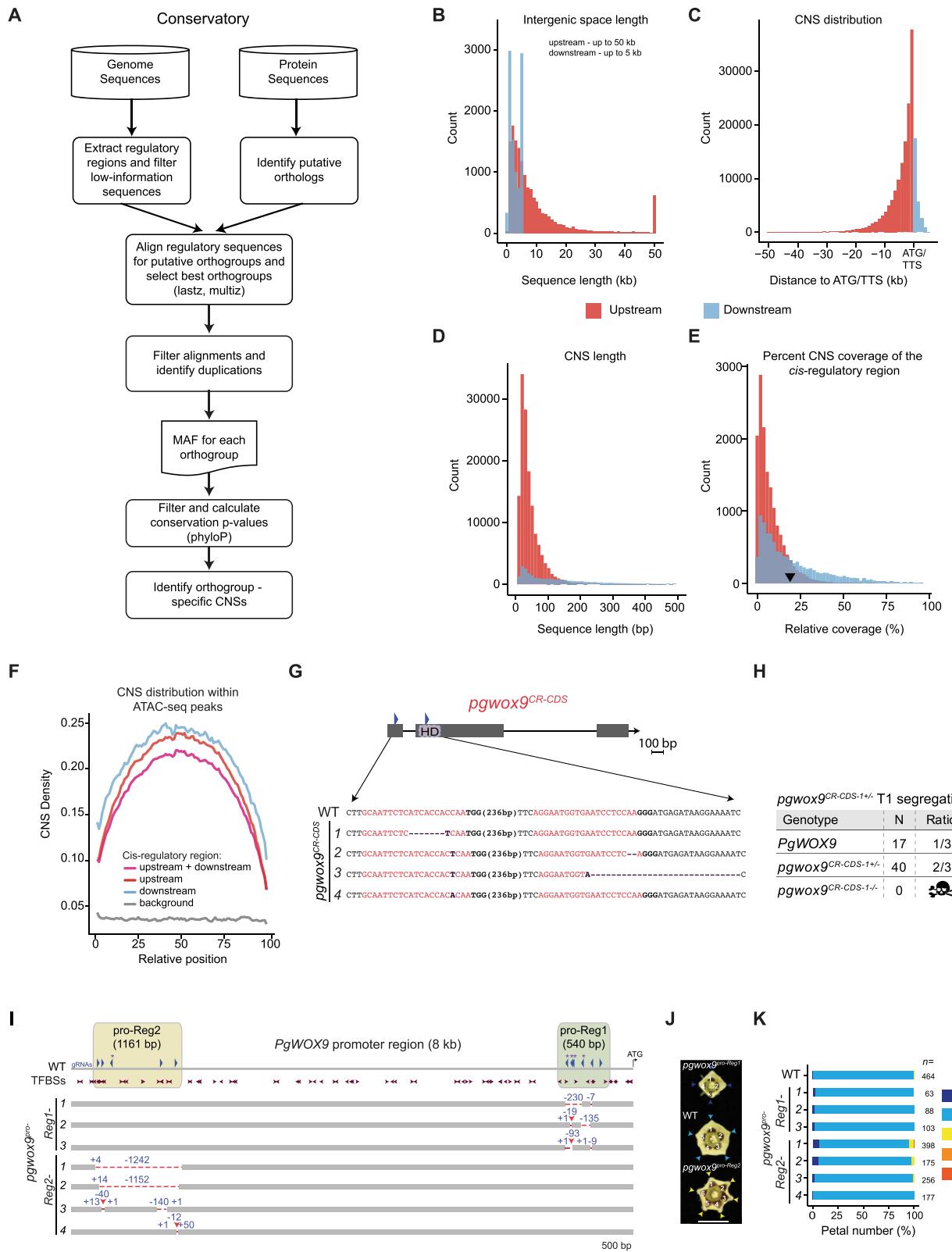


Figure S3. Summary of ATAC-seq analysis on meristem-enriched tissue, related to Figure 3

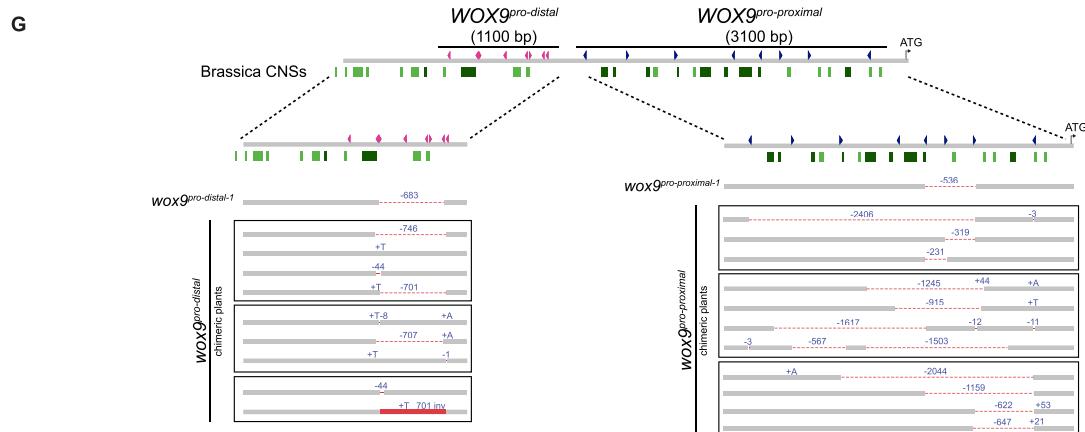
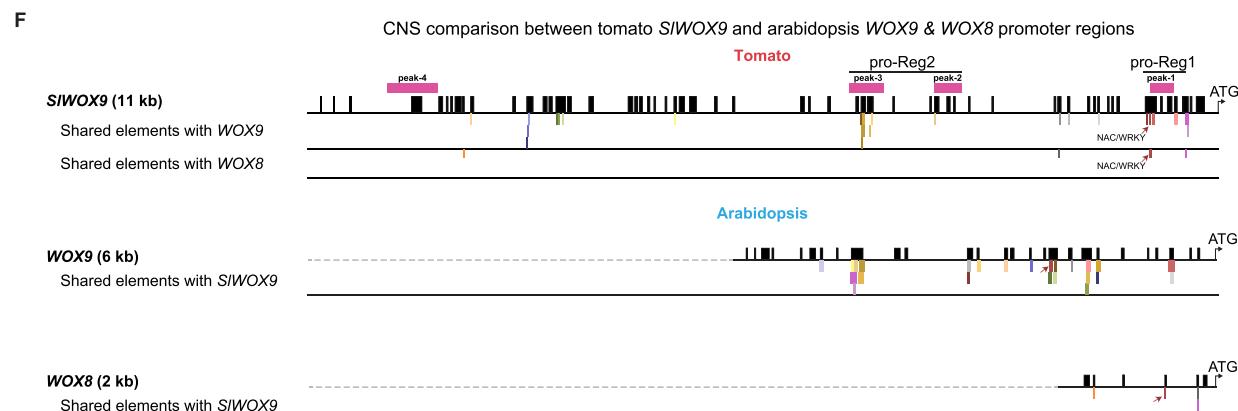
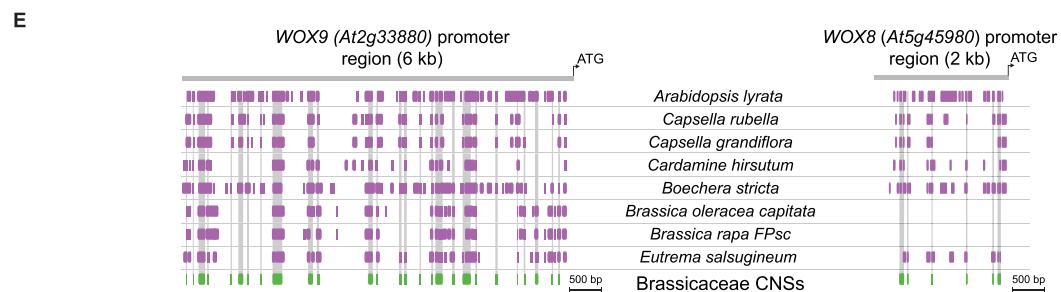
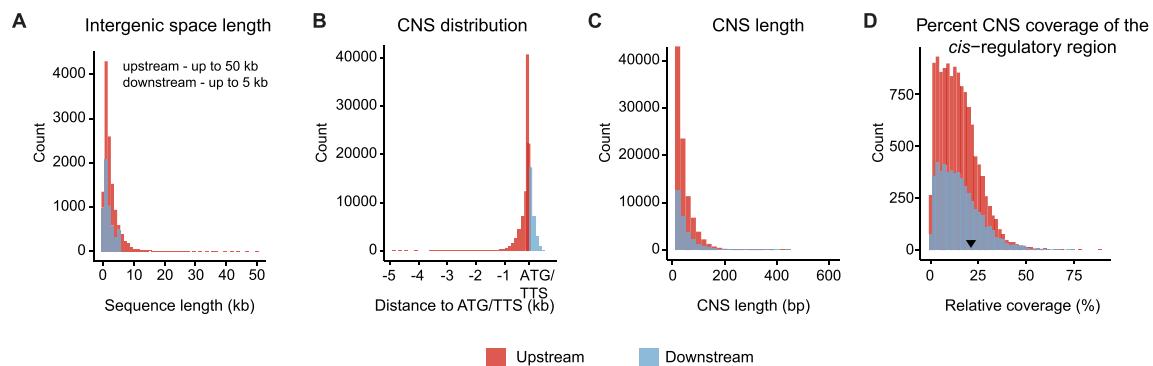
- (A) Representative image of an *antha* mutant inflorescence showing massive overproliferation of reproductive meristems used for generating the meristem ATAC-seq libraries. The table (right) shows the summary of ATAC-seq data (see STAR methods).
- (B) Heatmap showing Pearson correlation coefficients of the ATAC-seq peak regions called by Genrich (see STAR methods).
- (C) ATAC-seq peak size distribution (see STAR methods).
- (D) Enrichment of ATAC-seq coverage surrounding genes' TSSs and TTSs (see STAR methods).
- (E) Annotation of ATAC-seq peaks. Note that promoter region was defined as 5 kb upstream relative to TSSs (see STAR methods).



(legend on next page)

Figure S4. Solanaceae Conservatory analysis and additional characterization of the groundcherry-targeted *cis*-regulatory mutant alleles generated by CRISPR-Cas9, related to Figure 4

- (A) Flow chart of the “Conservatory” analysis used to define conserved non-coding sequences in this study (see STAR methods).
- (B) Distribution of upstream and downstream intergenic regions surrounding the orthogroups. 50 kb and 5 kb cut-offs were defined as upstream and downstream regions, respectively.
- (C) CNS distribution in 50 kb upstream and 5 kb downstream regions relative to the ATG and TTS, respectively (see STAR methods).
- (D) CNS size distributions for upstream and downstream regions (see STAR methods).
- (E) Conservation percentage of the intergenic *cis*-regulatory space surrounding the orthogroups. Black arrowhead marks conservation percentage of the Solanaceae WOX9 promoter region (19%).
- (F) CNS position distribution relative to overlapping ATAC-seq peaks in the indicated regulatory regions. Randomized positions were used to define background. Note that the summit of conservation (CNS density) resides at ATAC-seq peak midpoints regardless of position.
- (G) Null coding sequence mutants of groundcherry *pgwox9^{CR-CDS}* generated by CRISPR-Cas9. Gene diagram is shown. Grey bars are exons, lines are introns, and HD refers to homeodomain. Blue arrowheads are gRNAs. Four different indel mutations that cause frameshifts were generated, and no homozygous mutant plants could not be recovered, indicating embryonic lethality like the tomato *slwox9^{CDS}* mutant (Figure 1C).
- (H) Summary of segregation from progeny derived from a plant heterozygous for *pgwox9^{CR-CDS-1}*.
- (I) Schematic of the *PgWOX9* regulatory region targeted by CRISPR-Cas9. Blue arrowheads mark gRNAs used to target the specific *cis*-regulatory regions pro-Reg1 and pro-Reg2, as defined in the tomato *SlWOX9* promoter (see also Figure 3). Purple asterisks above specific gRNAs mark identical gRNA sequences used for CRISPR-Cas9 mutagenesis in tomato. *PgWOX9^{pro-Reg1}* and *PgWOX9^{pro-Reg2}* targeted regions are marked and highlighted with green and yellow colors, respectively. TFBSs with a q-value > 0.01 are shown (see STAR methods). Shown are three *pgwox9^{pro-Reg1}* alleles followed by four *pgwox9^{pro-Reg2}* alleles. Deletions (-) are indicated by red dashed lines, insertions by plus signs (+), inversions by red boxes.
- (J) Representative images of groundcherry flowers showing petal number variability. Colored arrowheads mark the petals in each flower. Dark blue, light blue and yellow marks four, five, and six petals per flower, respectively. Scale bar represent 1 cm.
- (K) Quantification of petal number. n, flower number. Note the increased frequency of 4 and 6 petal flowers, particularly in *pgwox9^{pro-Reg1}* and *pgwox9^{pro-Reg2}* alleles (see Table S4C).



(legend on next page)

Figure S5. Brassicaceae Conservatory analysis, comparison of tomato *SlWOX9* and *Arabidopsis WOX8* and *WOX9* CNSs, and CRISPR-Cas9 targeting of the *Arabidopsis WOX9* promoter, related to Figure 5

(A) Distribution of upstream and downstream intergenic regions surrounding the orthogroups. 50 kb and 5 kb cut-offs were defined as upstream and downstream regions, respectively.

(B) CNS distribution in 50 kb upstream and 5 kb downstream regions relative to the ATG and TTS, respectively (see [STAR methods](#)).

(C) CNS size distributions for upstream and downstream regions (see [STAR methods](#)).

(D) Conservation percentage of the intergenic *cis*-regulatory space surrounding the orthogroups. Black arrowhead marks conservation percentage of the Brassicaceae *WOX9* promoter region (21%).

(E) Conservatory analysis of Brassicaceae *WOX9* and *WOX8* promoters. Purple boxes define highly similar region of *WOX9* orthologs in the indicated species (see [STAR methods](#)). Green boxes define Brassicaceae CNSs.

(F) Conservatory analysis and comparison of tomato *SlWOX9* and *Arabidopsis WOX9* and *WOX8* promoter regions revealing shuffling of conserved *cis*-regulatory elements (see [STAR methods](#)). Boxes represent the short sequence elements colored by their position relative to the *SlWOX9* promoter. Dark red arrows mark the conserved NAC/WRKY TFBS. Note that only four elements were defined between *SlWOX9* and *WOX8*, compared to 24 elements conserved between *SlWOX9* and *WOX9*, supporting the shared functions of *SlWOX9* and *WOX9* (see [STAR methods](#)).

(G) Schematic of the *Arabidopsis WOX9* promoter regions targeted by CRISPR-Cas9. Blue arrowheads mark gRNAs used to target the proximal 3 kb region (*WOX9^{pro-proximal}*), and pink arrowheads mark gRNAs used to target 1 kb region upstream to *WOX9^{pro-proximal}* (*WOX9^{pro-distal}*). Shown is a *wox9^{pro-proximal}* homozygous allele followed by genotyping data of three *wox9^{pro-proximal}* chimeric plants. Alleles in each rectangle are from individual chimeric plants. WT alleles are not shown. Deletions (-) are indicated by red dashed lines, insertions by plus signs (+), inversions by red boxes.