Engineering quantitative trait variation for crop improvement by genome editing

Daniel Rodríguez-Leal[1], Zachary H. Lemmon[1], Jarrett Man[2], Madelaine E. Bartlett[2], and Zachary B. Lippman[1,3*]

[1] Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA.

[2] Biology Department, University of Massachusetts Amherst, Amherst, Massachusetts 01003, USA.

[3] Lead Contact

*Correspondence:

lippman@cshl.edu (Z.B.L.)

SUMMARY

Major advances in crop yields are needed in the coming decades. However, plant breeding is currently limited by incremental improvements in quantitative traits that often rely on laborious selection of rare naturally occurring mutations in gene regulatory regions. Here we demonstrate that CRISPR/Cas9 genome editing of promoters generates diverse *cis*-regulatory alleles that provide beneficial quantitative variation for breeding. We devised a simple genetic scheme, which exploits *trans*-generational heritability of *Cas9* activity in heterozygous loss-of-function mutant backgrounds, to rapidly evaluate the phenotypic impact of numerous promoter variants for genes regulating three major productivity traits in tomato: fruit size, inflorescence branching, and plant architecture. Our approach allows immediate selection and fixation of novel alleles in transgene-free plants and fine manipulation of yield components. Beyond a platform to enhance variation for diverse agricultural traits, our findings provide a foundation for dissecting complex relationships between gene regulatory changes and control of quantitative traits.

INTRODUCTION

Present crop yield increases will not meet future food demands. There is therefore an urgent need to develop innovative approaches to accelerate crop improvement and make its outcomes more predictable (Council for Agricultural Science and Technology (CAST), 2017). Significant obstacles in plant breeding are limited sources of genetic variation underlying quantitative traits and the time-consuming and labor-intensive phenotypic and molecular evaluation of breeding germplasm required to select plants with improved performance. Enhancing genetic and phenotypic variation in crops has relied on intercrossing with wild relatives to introduce "exotic" allelic diversity, creating novel alleles by random mutagenesis, and genetic engineering (Council for Agricultural Science and Technology (CAST), 2017; Lundqvist et al., 2012; Wang et al., 2017; Zamir, 2001). However, these approaches are inefficient, particularly for providing variants that cause subtle changes in quantitative traits that are most desired by breeders.

Numerous quantitative trait loci (QTL) and genome-wide association studies (GWAS) in both plants and animals have revealed many of the genetic changes driving evolution, domestication, and breeding occurred in *cis*-regulatory regions (Meyer and Purugganan, 2013; Olsen and Wendel, 2013; Wang et al., 2014; Wittkopp and Kalay, 2011). Compared to mutations in coding sequences that alter protein structure, *cis*-regulatory variants are frequently less pleiotropic and often cause subtle phenotypic change by modifying the timing, pattern, or level of gene expression (Wittkopp and Kalay, 2011). A major explanation for this is the complexity of transcriptional control, which includes redundancy and modular organization of the many *cis*-regulatory elements (CREs) in promoters and other regulatory regions, the majority of which remain poorly characterized (Cameron and Davidson, 2009; Priest et al., 2009; Schwarzer and

Spitz, 2014). Adding to this complexity is CRE spacing, chromosomal interactions, epistasis, and compensation between modules (Baxter et al., 2012; Priest et al., 2009; Schwarzer and Spitz, 2014). While these parameters provide flexibility for evolutionary change (Carroll, 2008), they can also complicate predicting phenotypic consequences from mutations in *cis*-regulatory regions (Wittkopp and Kalay, 2011).

Though widely favored in plant and animal evolution and domestication, *cis*-regulatory variants are far from saturated, and thus represent an untapped resource for expanding allelic diversity for breeding. The limited pool of *cis*-regulatory alleles has also precluded a deeper understanding of how regulatory changes impact quantitative traits. For example, a longstanding question is whether alterations in gene regulatory landscapes result in linear or non-linear relationships between transcriptional and phenotypic change, and how such responses vary for different genes (Birchler and Veitia, 2012; Birchler et al., 2016). Thus, expanding *cis*-regulatory variation not only holds promise for crop improvement, but also for elucidating principles underlying the control of quantitative traits.

A powerful approach to create novel allelic variation is through genome editing (Doudna and Charpentier, 2014; Hsu et al., 2014). In plants, this technology has primarily been used to engineer mutations in coding sequences, with the goal of creating null alleles for functional studies (Belhaj et al., 2015). However, based on previous work, we hypothesized that multiple elements of CRISPR/Cas9 technology could be integrated to engineer diverse types and strengths of *cis*-regulatory mutations (Cermak et al., 2017; Soyk et al., 2017; Swinnen et al., 2016). Here, we designed a genetic "drive" system that exploits heritability of CRISPR/Cas9 transgenes carrying multiple gRNAs in "sensitized"

F1 populations to rapidly and efficiently generate dozens of novel *cis*-regulatory alleles for three genes that regulate fruit size, inflorescence architecture, and plant growth habit in tomato. By segregating away the transgene in the following generation, we recovered a wide range of stabilized promoter alleles that provided a continuum of variation for all three traits. For one of these genes, we found that transcriptional change was a poor predictor of phenotypic effect, revealing unexplored complexity in how regulatory variation impacts quantitative traits.

RESULTS

Recreating a fruit size QTL by CRISPR/Cas9 mutagenesis of a *cis*-regulatory element

The major feature of tomato domestication was a dramatic increase in fruit size, caused in large part by an increase in the number of carpels in flowers, and thus seed compartments (locules) in fruits. QTL influencing tomato locule number include genes involved in the classical *CLAVATA-WUSCHEL* stem cell circuit (*CLV-WUS*), which controls meristem size (Somssich et al., 2016) (**Figure 1A**). Mutations in *CLV-WUS,* such as in the signaling peptide gene *CLV3,* can cause meristems to enlarge due to stem cell overproliferation, leading to developmental defects that include additional organs in flowers and fruits (Somssich et al., 2016; Xu et al., 2015). The ancestor of tomato (*S. pimpinellifolium*, *S.pim*) produces small bilocular fruits, and the *fasciated* (*fas*) and *locule number* (*lc*) QTL were major contributors to increased locule number, and thus fruit size, in domesticated tomato (*Solanum lycopersicum, S.lyc*) (**Figure 1B**) (van der Knaap et al., 2014). *fas* is a partial loss-of-function caused by an inversion that disrupts the promoter

of tomato *CLV3* (*SlCLV3*), resulting in a moderate effect on locule number (Huang and van der Knaap, 2011; Xu et al., 2015). In contrast, *lc* is a weak gain-of-function allele previously shown to be associated with two single nucleotide polymorphisms (SNPs) in a predicted 15 bp repressor element downstream of tomato *WUS* (*SlWUS*), a conserved homeobox gene that promotes stem cell proliferation (Somssich et al., 2016). While not functionally validated, this CRE shares similarity with the CArG element of *Arabidopsis* that is bound by the MADS box transcription factor *AGAMOUS* at the end of flower development to downregulate *WUS* and terminate meristem activity (Liu et al., 2011; Muños et al., 2011) (**Figures 1C** and **1D**).

To determine if induced mutations in known CREs can generate predictable quantitative variation, we used CRISPR/Cas9 to target the putative *SlWUS* CArG element (**Figure 1E**, see **STAR methods**). The effect of *lc* is subtle (Muños et al., 2011), with 11% of fruits producing 3-4 locules in *S.pim* near isogenic lines (*S.pim-lc*^{NIL}). Consistent with this, *lc* does not cause detectable changes in *SlWUS* expression, suggesting *lc* weakly affects the level, timing, or pattern of expression (Muños et al., 2011). Notably, 10% of fruits from *S.pim* plants carrying a CRISPR/Cas9-induced 4 bp deletion in the CArG element developed 3 locules (*S.pim-lc*^{CR}), nearly matching the weak effect of *lc* (**Figure 1F**). We previously showed that combining *fas* with *lc* synergistically increases locule number due to epistasis in the *CLV-WUS* circuit (van der Knaap et al., 2014; Lippman and Tanksley, 2001; Xu et al., 2015). We generated *S.pim-lc*^{CR} *fas*^{NIL} plants and found locule number exceeded *fas* alone (76% vs. 43% fruits ≥3 locules) and was similar to *S.pim-lc*^{NIL} *fas*^{NIL} plants, confirming that the CArG deletion allele mimics *lc* (**Figure 1F** and **Table S1**). Importantly, we validated these effects in a domesticated tomato

variety (*S. lyc.* cv. M82) whose fruits develop 2 (60%) or 3 (40%) locules (**Figure 1G**). We found that 70% of fruits from *S.lyc-lc*^{CR} plants carrying a 5 bp deletion in the repressor motif developed 3 or more locules, and this effect was enhanced in *S.lyc-lc*^{CR} *fas*^{NIL} double mutant plants (**Figure 1G** and **Table S1**). These results prove *lc* is caused by mutations in the *SlWUS* CArG element, and demonstrate that QTL can be engineered by mutating CREs with known functions.

CRISPR/Cas9 mutagenesis of the *SlCLV3* promoter generates novel *cis*-regulatory alleles

Recreating the effect of *lc* showed that CRISPR/Cas9 targeting of previously characterized *cis*-regulatory regions can create new alleles of existing QTL. Yet, the precise causative variants underlying the many QTL that map to regulatory regions are rarely known, particularly for the majority of cases where multiple SNPs and/or structural variants (SVs) are in the QTL interval (Meyer and Purugganan, 2013; Olsen and Wendel, 2013). Moreover, the modular organization and inherent redundancy among CREs make it extremely challenging to define useful targets, especially for generating specific desired modifications for a quantitative trait (Priest et al., 2009; Weber et al., 2016). However, we hypothesized these properties could be exploited to create a series of *cis*-regulatory alleles with a range of quantitative transcriptional and phenotypic changes by targeting gene promoter regions with many guide RNAs (gRNAs) (**Figure 2A**). Through synchronous and asynchronous Cas9-gRNA directed cleavage and imprecise repair at each target site, an array of mutation types could be induced, including deletions of various sizes and small indels at target sites (Brooks et al., 2014; Soyk et al., 2017; Xu et

al., 2015). The resulting alleles, having mutations that might impact multiple CREs, *cis*-regulatory modules, or their spacing, could then be evaluated for phenotypic changes by generating stable homozygous mutants in subsequent generations.

We tested this concept by designing a CRISPR/Cas9 construct with 8 gRNAs designed to target the 2 Kbp promoter region immediately upstream of the M82 *SlCLV3* coding sequence (see **STAR Methods**), without considering any predicted CREs (**Figure 2B**). Six first-generation transgenic plants (T_0) were generated as previously described (Brooks et al., 2014), and PCR genotyping revealed four of them carried deletions of various sizes in the target region (**Figure 2C**). Notably, flowers from these plants produced more organs than WT, with $T_{0\text{-}2}$ having organ numbers between *fas* and CRISPR/Cas9-generated *clv3* coding sequence null mutants (*slclv3*^{CR}) (**Figures 2D** and **2E**) (Xu et al., 2015). PCR and sequencing suggested $T_{0\text{-}2}$ was homozygous for an allele with small indels at the first four targets and a ~1 Kbp deletion beginning at target 5 and extending beyond target 8 (**Figure 2F**). In contrast, $T_{0\text{-}1}$ appeared homozygous for a large deletion (1.6 Kbp) that encompassed all targets, yet showed little change in organ number (**Figure 2F**). Two of the four remaining T_0 plants displayed weaker effects and were chimeric for at least three alleles, including one allele in $T_{0\text{-}4}$ having the same deletion as $T_{0\text{-}1}$ (**Figures 2E** and **2F**).

Obtaining homozygous mutants from CRISPR/Cas9 in first-generation transgenics is rare (Brooks et al., 2014; Svitashev et al., 2015; Zhang et al., 2014), and the weak $T_{0\text{-}1}$ phenotype was surprising considering the 1.6 Kbp deletion encompassed most of the $T_{0\text{-}2}$ deletion. To test heritability of these alleles and validate their phenotypic effects we genotyped T_1 progeny generated by self-fertilization. Surprisingly, progeny

from both $T_{0\text{-}1}$ and $T_{0\text{-}2}$ inherited an allele that could not be amplified by PCR (**Figure 2G**). The near 1/4 segregation of homozygosity for these "hidden" alleles in each T_1 family (8/24, 33% and 5/24, 21%; Chi-square: $P = 0.6$ and $P = 0.78$, respectively) indicated both T_0 plants were biallelic, with the second alleles potentially having a larger structural change that prevented PCR amplification. To better characterize these alleles, we sequenced the genomes of homozygous T2 progeny (see **STAR Methods**), which revealed a second allele in $T_{0\text{-}1}$ with a complex rearrangement (designated *SlCLV3*$^{CR\text{-}pro1\text{-}2}$) and a large 7.3 Kbp deletion allele in $T_{0\text{-}2}$ that spanned the *SlCLV3* coding sequence (*SlCLV3*$^{CR\text{-}pro2\text{-}2}$) (**Figure 2H**). We further used these data to show there were no detectable off-target mutations, supporting the high specificity of CRISPR/Cas9 in plants (**Table S2**) (Peterson et al., 2016). Quantitative phenotyping showed that the increase in floral organs for *SlCLV3*$^{CR\text{-}pro2\text{-}2}$ homozygotes matched *slclv3*^{CR} mutants, confirming *SlCLV3*$^{CR\text{-}pro2\text{-}2}$ is a null allele. In contrast, plants homozygous for the original $T_{0\text{-}1}$ and $T_{0\text{-}2}$ alleles (*SlCLV3*$^{CR\text{-}pro1\text{-}1}$ and *SlCLV3CR*$^{\text{-}pro2\text{-}1}$, respectively) showed slightly weaker effects than *slclv3*^{CR} plants, indicating hypomorphic alleles. Finally, we found that *SlCLV3*$^{CR\text{-}pro1\text{-}2}$ homozygotes resembled WT, explaining the weak phenotype of the original biallelic $T_{0\text{-}1}$ plant (**Figure 2I** and **Table S3**). These results demonstrate that CRISPR/Cas9 transgenes carrying diverse gRNAs targeting various regions of a promoter can effectively create novel *cis*-regulatory mutations and alleles with phenotypic effects, including unexpected lesions within and beyond the target region.

A *trans*-acting CRISPR/Cas9-driven mutagenesis screen allows rapid generation and evaluation of many *SlCLV3* promoter alleles for quantitative variation

The *SlCLV3* promoter alleles from the T_0 plants showed CRISPR/Cas9 targeting of regulatory sequences could create novel genetic and phenotypic variation. However, each T_0 plant provided only a few alleles with either strong or weak effects. Early on, we expected dozens of alleles or more would be needed to obtain a collection of alleles encompassing a full range of quantitative variation. However, standard transgenic methods would be time-consuming and costly. To address these limitations, we devised a simple genetic scheme that exploits *trans*-generational inheritance of the CRISPR/Cas9 transgene to induce new mutations upon outcrossing to WT plants (**Figure 3A**). With this approach, all F1 plants that inherit the single copy transgene from a T_0 "hemizygous" individual would have the potential to generate one or more new alleles by targeting *in trans* the WT promoter introduced from the cross. However, determining which specific F1 individuals harbor new alleles that result in phenotypic change can be difficult. A telling example is the complex rearrangement of the $SlCLV3^{CR\text{-}pro1\text{-}2}$ allele, which had no effect on floral organ number, and thus complemented and masked the effect of the strong loss-of-function large deletion allele in the original biallelic $T_{0\text{-}1}$ plant (**Figure 2I**). To simultaneously maximize allele creation and efficiently identify those with phenotypes, we outcrossed only T_0 plants with strong loss-of-function alleles to produce a sensitized population of heterozygous F1 plants. In this way, hundreds of F1 progeny carrying a CRISPR/Cas9 transgene, each also having inherited a stable loss-of-function allele, could easily be generated and screened for new loss-of-function alleles, including those causing subtle phenotypes that would otherwise be difficult to detect.

To test this approach, we crossed $T_{0\text{-}2}$ to WT and generated 1,152 F1 plants that were heterozygous for either *SlCLV3CR*$^{-pro2\text{-}1}$ or *SlCLV3CR*$^{-pro2\text{-}2}$ and a WT *SlCLV3* promoter (see **STAR Methods**). PCR genotyping revealed nearly half of the population (42%) inherited the CRISPR/Cas9 transgene (hemizygous *Cas9*$^{-/+}$) (**Figures 3A** and **3B**), and phenotyping these 479 plants revealed 116 individuals (24%) with more floral organs than WT. While most of these plants (80%) showed weak effects, 24 were similar to *fas* or stronger (**Figures 3C**, **S1A-S1B** and **Table S4**). These findings demonstrate the power of combining meiotically heritable Cas9-gRNA activity with a sensitized background to efficiently engineer numerous *cis*-regulatory alleles with readily observable phenotypic consequences.

Novel *cis*-regulatory alleles can immediately be fixed in transgene-gene free plants to achieve a range of variation for fruit locule number

The many F1 plants with increased locule number was a promising indication that the sensitized screen succeeded in generating a collection of loss-of-function alleles that would translate to a continuum of fruit locule number variation. However, given that F1 plants already carried a strong loss-of-function *SlCLV3* promoter allele from $T_{0\text{-}2}$, we expected F1 phenotypes would appear more severe than plants homozygous for newly induced alleles. Thus, to obtain a range of quantitative effects, we focused on isolating new alleles from the 24 F1s with strong and moderate phenotypes. PCR genotyping revealed all of these plants were biallelic or chimeric, with most having novel deletion alleles (**Figure 4A**). Conveniently, an inherent advantage of our genetic scheme is that newly generated alleles in F1 plants can be immediately fixed in a transgene-free

background at a 1/16 ratio in F2 progeny from biallelic plants (**Figure 4B**), and a lower ratio for chimeric plants. We confirmed this segregation using progeny from a moderate biallelic F1 plant (**Figure 4C**).

To enrich for *SlCLV3* promoter alleles covering a full range of quantitative variation, we characterized F2 progeny from a subset of 14 F1 plants with strong and moderate phenotypes that captured the spectrum of allelic diversity and locule number variation (**Figures 4A** and **S1B**). The promoters from homozygous F2 transgene-free mutants were sequenced and F3 progeny evaluated for effects on locule number (**Figures 4D** and **4E**). We found all 14 new alleles were distinct and displayed a variety of mutation types, including large deletions, inversions, small indels and a point mutation throughout the target region (**Figure 4F**). Most significant, homozygous mutants for these alleles displayed a continuum of locule number variation. This included one allele (*SlCLV3*$^{CR\text{-}pro\text{-}s3}$) with a subtle increase in locule number similar to the weak gain-of-function effect of *S.lyc lc*^{CR} (**Figure 1G**), and two alleles (*SlCLV3*$^{CR\text{-}pro\text{-}m6}$ and *SlCLV3*$^{CR\text{-}pro\text{-}m8}$) that phenocopied *fas* (**Figure 4E** and **Table S5**). These results confirm that alleles with weaker quantitative effects can be recovered from moderate and strong F1 plants, and show that both existing and novel QTL variation can be engineered using our approach.

Mutations in conserved *cis*-regulatory regions in the *SlCLV3* promoter and their effects on transcription are poor predictors of phenotypic change.

We took advantage of *SlCLV3* promoter allele collection to address how specific *cis*-regulatory mutations influence locule number variation. Though the resolution

provided by our alleles was insufficient to define functions for specific CREs, sequence analysis revealed many predicted CREs across the *SlCLV3* promoter were differentially deleted among the 14 alleles **(Figure S2A)**. Pairs of alleles, such as *SlCLV3*$^{CR\text{-}pro\text{-}m13}$/*SlCLV3*$^{CR\text{-}pro\text{-}m14}$ and *SlCLV3*$^{CR\text{-}pro\text{-}m7}$/*SlCLV3*$^{CR\text{-}pro\text{-}m11}$, shared overlapping deletions and similar phenotypic effects, pointing to putative transcriptional control regions. To further understand the significance of these and similar observations, we compared the targeted promoter region of *S.lyc* with three *Solanaceae* species, beginning with the wild tomato species *S. pennellii* (*S.pen*) and extending to the more distantly related potato (*S. tuberosum, S.tub*) and pepper (*C. annuum, C.ann*) (**Figure S2B**, and **STAR methods)**. Pairwise sequence alignments revealed four regions of high conservation between tomato and potato ranging in size from ~100-250 bp, with three conserved regions (CR) also found in pepper (**Figure 4E**).

Although some trends were evident, the magnitude of phenotypic effect could not readily be predicted from disrupting these conserved regions. Overall, the largest deletion alleles showed the greatest increase in locule number, whereas alleles with smaller indels had the weakest impact. For example, two of the three strongest alleles that mimicked coding sequence null mutations (*SlCLV3*$^{CR\text{-}pro\text{-}m4}$ and *SlCLV3*$^{CR\text{-}pro\text{-}s6}$) shared large overlapping lesions that disrupted all four CRs. Yet, *SlCLV3*$^{CR\text{-}pro\text{-}m12}$ was an equally strong allele whose large 1042 bp deletion eliminated only CR4 and flanking DNA. Contrary to this was the more moderate effect of *SlCLV3*$^{CR\text{-}pro\text{-}m8}$, an allele with a 1408 bp deletion that overlaps with *SlCLV3*$^{CR\text{-}pro\text{-}m12}$ and removes all of CR3 and CR4 and a part of CR2 (**Figure 4D, Figure S2B**).

One likely explanation for these and other observed poor associations is complex functional interrelationships between the CREs within the conserved regions, including possible redundancy, epistasis, and compensation (Carroll, 2008). In addition, CREs were predicted outside of the 4 conserved regions (**Figure 2SA** and **2SB**), and proper organization and spacing of CREs and *cis*-regulatory modules are also important for promoter function (Baxter et al., 2012; Wittkopp and Kalay, 2011). Though additional alleles will be needed to dissect the function of individual CREs and their relationships, our collection of promoter alleles illustrates how multiple mutations in *cis*-regulatory regions can cause unexpected, unpredictable phenotypic changes. However, it remained possible that different alleles caused similar changes in *SlCLV3* expression, and thus comparable phenotypic effects. In the conserved *CLV-WUS* circuit, CLV3 peptide binds to cell surface leucine-rich repeat (LRR) receptor complexes to initiate a signaling cascade that restricts *WUS* expression and prevents stem cell overproliferation. Through negative feedback, WUS promotes *CLV3* expression to limit its own activity (Somssich et al., 2016). Given this feedback mechanism, we evaluated expression of both *SlCLV3* and *SlWUS* in reproductive transition meristems of the 14 promoter alleles. Remarkably, we found little correlation between locule number and modified transcriptional balance between these genes (**Figure 4F**). For example while two of the strongest alleles showed a dramatic reduction in *SlCLV3* expression, the strong $SlCLV3^{CR\text{-}pro\text{-}s6}$ allele was unchanged for *SlCLV3,* but *SlWUS* was upregulated in all three. Even more, the phenotypically similar *fas* and $SlCLV3^{CR\text{-}pro\text{-}m6}$ showed *SlCLV3:SlWUS* expression ratios of 22 and 1.5, respectively. These and other examples likely reflect the complex spatial, temporal, and biochemical regulatory mechanisms underlying the *CLV-WUS* negative

feedback circuit, one or more of which might be altered in each allele (Perales et al., 2016; Reddy and Meyerowitz, 2005; Rodriguez et al., 2016; Somssich et al., 2016; Xu et al., 2015). Taken together, these findings demonstrate there is not a simple linear relationship between transcriptional and phenotypic change for *SlCLV3* (**Figure 2A**), highlighting the advantage of our approach for isolating novel *cis*-regulatory alleles with a range of quantitative effects.

CRISPR/Cas9 *cis*-regulatory mutagenesis can fine-tune diverse traits

Our ability to engineer a continuum of quantitative variation for locule number relied on a surprising level of dosage-sensitivity for *SlCLV3*, the only previous evidence of which came from *fas* (Xu et al., 2015). In general, developmental genes are more likely to exhibit dosage-sensitivities compared to genes encoding metabolic enzymes, for example (Birchler and Veitia, 2012), suggesting our approach could be used to engineer quantitative variation for diverse traits.

We tested this by targeting the promoters of the inflorescence architecture gene *COMPOUND INFLORESCENCE* (*S,* homolog of *Arabidopsis WUSCHEL-RELATED HOMEOBOX 9*, *WOX9*) (Lippman et al., 2008) and the plant architecture gene *SELF PRUNING* (*SP,* homolog of *Arabidopsis TERMINAL FLOWER 1, TFL1*) (Pnueli et al., 1998), which regulate two major productivity traits. *S* controls tomato inflorescence development by promoting meristem maturation, and coding sequence mutations result in excessively branched inflorescences with hundreds of flowers (Lippman et al., 2008; Park et al., 2012). *SP* encodes a flowering repressor in the florigen gene family that counterbalances the flowering hormone florigen (encoded by *SINGLE FLOWER TRUSS,*

SFT) to ensure continuous shoot and inflorescence production through 'indeterminate' growth (Lifschitz et al., 2006, 2014; Pnueli et al., 1998). A classical coding sequence mutation in *SP* provided the compact bushy 'determinate' growth habit that was critical for large-scale field production (Pnueli et al., 1998).

Limited allelic variation for both genes has made it difficult to improve these traits. Homozygous *s* mutants have poor fertility due to flower abortion, but heterozygosity provides moderate branching and improved yield due to dosage-sensitivity (Soyk et al., 2017). However, this genotype reflects only one point along a possible continuum of inflorescence branching and flower production. The situation is similar for *sp-classic*, the only known loss-of-function allele of *SP*. We recently demonstrated that dosage relationships among genes controlling both developmental programs could be exploited to create a quantitative range of inflorescence and plant architectures that translated to improved productivity (Park et al., 2014; Soyk et al., 2017). However, the need to create specific higher order homozygous and heterozygous mutant combinations for multiple genes by traditional breeding remains a drawback of this approach (Park et al., 2014; Soyk et al., 2017). We reasoned a similar outcome could be achieved much more efficiently by engineering *cis*-regulatory mutations in the promoters of *S* and *SP*.

We used our multiplexed *CRISPR/Cas9* promoter targeting approach with *S* and obtained three T_0 plants that carried deletions of various sizes in the target region. $T_{0\text{-}2}$ exhibited inflorescence branching similar to *s* mutants (**Figure 5A**), and we outcrossed this individual to establish and screen a population of 326 sensitized F1 plants, of which 91 (28%) showed branching ranging from weak to strong. Interestingly, we identified

five chimeric plants that produced sectors with cauliflower-like inflorescences due to massive overproliferation of meristems, suggesting all known *s* mutants might not be null alleles (**Figure 5B**) (Lippman et al., 2008; Soyk et al., 2017). From a subset of 10 strong and moderate F1 plants, we derived a collection of 6 distinct promoter alleles ($S^{CR\text{-}pro}$), which translated to a range of inflorescence branching (**Figure 5C, 5D** and **Table S6**). Notably, this included one allele ($S^{CR\text{-}pro\text{-}6}$) that matched *s*/+ heterozygotes and another allele ($S^{CR\text{-}pro\text{-}5}$) that was more consistently weakly branched, demonstrating a high level of dosage-sensitivity for *S* like *SlCLV3* (**Figure 5D**) (Soyk et al., 2017).

Hundreds of F1 plants were employed in our *SlCLV3* and *S* sensitized screens, but for both cases a smaller subset of individuals was sufficient to isolate alleles that provided a continuum of trait variation. We used *SP* to test whether alleles derived from T_0 plants combined with a smaller screen would yield similar results. Four T_0 plants were obtained and PCR genotyping revealed all were biallelic or chimeric and produced deletions of various sizes (**Figure 5E**). Sequencing of eight alleles showed all were distinct; however, half of the alleles shared a ~500 bp deletion between target site 2 and 4, suggesting these sites were more efficiently targeted (**Figure 5F**). We evaluated the phenotypic effects of three *SP* promoter alleles ($SP^{CR\text{-}pro}$) derived from $T_{0\text{-}3}$ and $T_{0\text{-}4}$ by recovering stable homozygous transgene-free plants. Importantly, these $SP^{CR\text{-}pro}$ variants displayed desirable modified shoot architectures, representing several points along the continuum of plant architectures generated previously by combining rare coding sequence mutations for multiple genes in the florigen pathway (**Figures 5G**, **5H** and **Table S7**) (Park et al., 2014). To expand the collection of *SP cis*-regulatory alleles, we initiated a smaller screen with a population of 81 F1 plants, of which 25 (30%) showed a

range of modified architectures. Progeny from these individuals should carry additional alleles to expand variation for this major productivity trait. Collectively, these results demonstrate that our approach can be widely applied to efficiently engineer desirable variation for diverse genes and traits.

DISCUSSION

CRISPR/Cas9 genome editing has been widely adopted in plants as a tool for understanding fundamental biological processes (Belhaj et al., 2015; Voytas and Gao, 2014). Attention is now shifting towards the promise of genome editing for agricultural applications. In this study, we have taken a major step forward on both fronts by integrating multiple virtues of CRISPR/Cas9 to rapidly and efficiently engineer quantitative trait variation by mutagenizing *cis*-regulatory regions. The collections of *cis*-regulatory alleles we created demonstrate that a wide range of quantitative variation can be achieved from altering the expression of individual genes. This remarkable level of dosage sensitivity bodes well for modifying diverse traits by targeting the promoters of other developmental regulators, and possibly other classes of genes. Our characterization of the *SlCLV3* promoter alleles in particular provided fundamental insights into the complex architecture of *cis*-regulatory regions, transcriptional regulation, and the control of quantitative traits. For example, and most strikingly, we discovered that changes in locule number are not predicted by changes in *SlWUS* or *SlCLV3* expression levels. This lack of predictability lends empirical support to recent models suggesting that non-linear relationships may be widespread for dose-sensitive genes, particularly developmental genes encoding transcriptional regulators and components of signal transduction

pathways that often function in complex regulatory networks (Birchler et al., 2016). Our approach can now be used to study if and to what extent similar complexity exists for other genes, pathways, and phenotypes. With the flexibility and expandability of CRISPR/Cas9 (Cermak et al., 2017), our approach can produce hundreds of regulatory mutations to systematically assess the association of *cis*-regulatory regions with phenotypic variation by dissecting the functions of specific CREs, the modules in which they function, and their spatial organization.

Though of great value for understanding the control of complex traits, the most immediate impact of our findings will be for enhancing breeding, where the precise function of an individual CRE is less important than phenotypic outcomes. Such streamlined trait improvement is evident for all three genes and traits we targeted. The phenotypic variation we achieved by engineering novel regulatory alleles for a single gene previously required stacking multiple natural and induced mutations for several genes (Park et al., 2014; Soyk et al., 2017). There is also potential for engineering gain-of-function alleles, of which there are several examples in evolution and domestication (Meyer and Purugganan, 2013). For example, the *cis*-regulatory region downstream of *SlWUS* could have repressor elements beyond *lc,* presenting a promising proof-of-principle target to engineer a range of dominant or semi-dominant effects. Finally, breeders expend great time and effort to adapt beneficial allelic variants to diverse breeding germplasm, and our approach can help bypass this constraint by directly generating and selecting for the most desirable regulatory variant in the context of modifier loci and epistatic environments of specific genetic backgrounds (Shen et al., 2016). With the remarkable pace that genome editing and plant transformation

technologies are advancing (Barrangou and Doudna, 2016; Cermak et al., 2017; Lowe et al., 2016), expansive libraries of regulatory alleles could soon be created in both plants and animals (Van Eenennaam, 2017). Beyond enhancing and customizing diverse trait variation in elite breeding germplasm, our approach opens the door to improve orphan crops and engineer domestication in wild plants with agricultural potential.

AUTHOR CONTRIBUTIONS

D.R.-L and Z.B.L designed the experiments; D.R.-L., Z.H.L, Z.B.L performed experiments and data analyses. J.M. and M.E.B. performed CRE predictions and comparative sequence analyses. D.R.-L and Z.B.L wrote the paper with input from Z.H.L. All authors have read, edited and approved the content of the manuscript.

ACKNOWLEDGEMENTS

We thank members of the Lippman lab and D. Jackson, Z. Nimchuck, and E. Vollbrecht for discussions and comments on the manuscript. We thank C. Brooks, A. Krainer, and J. Dalrymple for technical support, Y. Eshed for assistance with screening, E. van der Knaap for providing *S.pim. fas* and *lc* NILs, M. Tjahjadi and J. Van Eck for tomato transformation. We thank T. Mulligan, S. Vermylen, S. Qiao, N. Springsteen, E. Schoenfeld and B. Berube for assistance with plant care and data collection. This research was supported by a PEW Latin American Fellowship (29661) to D.R.-L., a National Science Foundation Postdoctoral Research Fellowship in Biology grant (IOS-1523423) to Z.H.L., a National Science Foundation Plant Genome Research Program grant (IOS-1732253) to Z.B.L and a National Science Foundation Plant Genome Research Program grant (IOS-1546837) to M.E.B and Z.B.L.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

REFERENCES

100 Tomato Genome Sequencing Consortium, Aflitos, S., Schijlen, E., de Jong, H., de Ridder, D., Smit, S., Finkers, R., Wang, J., Zhang, G., Li, N., et al. (2014). Exploring genetic variation in the tomato (Solanum section Lycopersicon) clade by whole-genome sequencing. Plant J. *80*, 136–148.

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*, W202-8.

Barrangou, R., and Doudna, J.A. (2016). Applications of CRISPR technologies in research and beyond. Nat. Biotechnol. *34*, 933–941.

Baxter, L., Jironkin, A., Hickman, R., Moore, J., Barrington, C., Krusche, P., Dyer, N.P., Buchanan-Wollaston, V., Tiskin, A., Beynon, J., et al. (2012). Conserved noncoding sequences highlight shared components of regulatory networks in dicotyledonous plants. Plant Cell *24*, 3949–3965.

Belhaj, K., Chaparro-Garcia, A., Kamoun, S., Patron, N.J., and Nekrasov, V. (2015). Editing plant genomes with CRISPR/Cas9. Curr. Opin. Biotechnol. *32*, 76–84.

Birchler, J.A., and Veitia, R.A. (2012). Gene balance hypothesis: connecting issues of dosage sensitivity across biological disciplines. Proc. Natl. Acad. Sci. U. S. A. *109*, 14746–14753.

Birchler, J.A., Johnson, A.F., and Veitia, R.A. (2016). Kinetics genetics: Incorporating the concept of genomic balance into an understanding of quantitative traits. Plant Sci. *245*, 128–134.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y., and Buckler, E.S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics *23*, 2633–2635.

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.

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.

Carroll, S.B. (2008). Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution. Cell *134*, 25–36.

Cermak, T., Curtin, S.J., Gil-Humanes, J., Čegan, R., Kono, T.J.Y., Konečná, E., Belanto,

J.J., Starker, C.G., Mathre, J.W., Greenstein, R.L., et al. (2017). A multi-purpose toolkit to enable advanced genome engineering in plants. Plant Cell tpc.00922.2016.

Council for Agricultural Science and Technology (CAST) (2017). Plant Breeding and Genetics—A paper in the series on The Need for Agricultural Innovation to Sustainably Feed the World by 2050. CAST.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158.

Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science *346*, 1258096.

Van Eenennaam, A.L. (2017). Genetic modification of food animals. Curr. Opin. Biotechnol. *44*, 27–34.

Frazer, K.A., Pachter, L., Poliakov, A., Rubin, E.M., and Dubchak, I. (2004). VISTA: computational tools for comparative genomics. Nucleic Acids Res. *32*, W273-9.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017–1018.

Gupta, S., and Van Eck, J. (2016). Modification of plant regeneration medium decreases the time for recovery of Solanum lycopersicum cultivar M82 stable transgenic lines. Plant Cell, Tissue Organ Cult. *127*, 417–423.

Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. Cell *157*, 1262–1278.

Huang, Z., and van der Knaap, E. (2011). Tomato fruit weight 11.3 maps close to fasciated on the bottom of chromosome 11. Theor. Appl. Genet. *123*, 465–474. van der Knaap, E., Chakrabarti, M., Chu, Y.H., Clevenger, J.P., Illa-Berenguer, E.,

Huang, Z., Keyhaninejad, N., Mu, Q., Sun, L., Wang, Y., et al. (2014). What lies beyond the eye: the molecular mechanisms regulating tomato fruit weight and shape. Front. Plant Sci. *5*, 227.

Lei, Y., Lu, L., Liu, H.-Y., Li, S., Xing, F., and Chen, L.-L. (2014). CRISPR-P: a web tool for synthetic single-guide RNA design of CRISPR-system in plants. Mol. Plant *7*, 1494–1496.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

Lifschitz, E., Eviatar, T., Rozman, A., Shalit, A., Goldshmidt, A., Amsellem, Z., Alvarez, J.P., and Eshed, Y. (2006). The tomato FT ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli. Proc. Natl. Acad.

Sci. U. S. A. *103*, 6398–6403.

Lifschitz, E., Ayre, B.G., and Eshed, Y. (2014). Florigen and anti-florigen - a systemic mechanism for coordinating growth and termination in flowering plants. Front. Plant Sci. *5*, 465.

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

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.

Liu, X., Kim, Y.J., Muller, R., Yumul, R.E., Liu, C., Pan, Y., Cao, X., Goodrich, J., and Chen, X. (2011). AGAMOUS Terminates Floral Stem Cell Maintenance in Arabidopsis by Directly Repressing WUSCHEL through Recruitment of Polycomb Group Proteins. Plant Cell *23*, 3654–3670.

Lowe, K., Wu, E., Wang, N., Hoerster, G., Hastings, C., Cho, M.-J., Scelonge, C., Lenderts, B., Chamberlin, M., Cushatt, J., et al. (2016). Morphogenic Regulators *Baby boom* and *Wuschel* Improve Monocot Transformation. Plant Cell *28*, 1998–2015.

Lundqvist, U., Franckowiak, J.D., and Forster, B.P. (2012). Plant mutation breeding and biotechnology (Wallingford: CABI).

Lyons, E., and Freeling, M. (2008). How to usefully compare homologous plant genes and chromosomes as DNA sequences. Plant J. *53*, 661–673.

Mathelier, A., Zhao, X., Zhang, A.W., Parcy, F., Worsley-Hunt, R., Arenillas, D.J., Buchman, S., Chen, C., Chou, A., Ienasescu, H., et al. (2014). JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic Acids Res. *42*, D142-7.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.

Meyer, R.S., and Purugganan, M.D. (2013). Evolution of crop species: genetics of domestication and diversification. Nat. Rev. Genet. *14*, 840–852.

Muños, S., Ranc, N., Botton, E., Bérard, A., Rolland, S., Duffé, P., Carretero, Y., Le Paslier, M.-C., Delalande, C., Bouzayen, M., et al. (2011). Increase in Tomato Locule Number Is Controlled by Two Single-Nucleotide Polymorphisms Located Near WUSCHEL. Plant Physiol. *156*, 2244–2254.

Narzisi, G., O'Rawe, J.A., Iossifov, I., Fang, H., Lee, Y., Wang, Z., Wu, Y., Lyon, G.J.,

Wigler, M., and Schatz, M.C. (2014). Accurate de novo and transmitted indel detection in exome-capture data using microassembly. Nat. Methods *11*, 1033–1036.

O'Malley, R.C., Huang, S.-S.C., Song, L., Lewsey, M.G., Bartlett, A., Nery, J.R., Galli, M., Gallavotti, A., and Ecker, J.R. (2016). Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. Cell *165*, 1280–1292.

Olsen, K.M., and Wendel, J.F. (2013). A bountiful harvest: genomic insights into crop domestication phenotypes. Annu. Rev. Plant Biol. *64*, 47–70.

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. *109*, 639–644.

Park, S.J., Jiang, K., Tal, L., Yichie, Y., Gar, O., Zamir, D., Eshed, Y., and Lippman, Z.B. (2014). Optimization of crop productivity in tomato using induced mutations in the florigen pathway. Nat. Genet. *46*, 1337–1342.

Perales, M., Rodriguez, K., Snipes, S., Yadav, R.K., Diaz-Mendoza, M., and Reddy, G.V. (2016). Threshold-dependent transcriptional discrimination underlies stem cell homeostasis. Proc. Natl. Acad. Sci. *113*, E6298–E6306.

Peterson, B.A., Haak, D.C., Nishimura, M.T., Teixeira, P.J.P.L., James, S.R., Dangl, J.L., and Nimchuk, Z.L. (2016). Genome-Wide Assessment of Efficiency and Specificity in CRISPR/Cas9 Mediated Multiple Site Targeting in Arabidopsis. PLoS One *11*, e0162169.

Pnueli, L., Carmel-Goren, L., Hareven, D., Gutfinger, T., Alvarez, J., Ganal, M., Zamir, D., and Lifschitz, E. (1998). The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1. Development *125*, 1979–1989.

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.

Qiu, Z., Li, R., Zhang, S., Wang, K., Xu, M., Li, J., Du, Y., Yu, H., and Cui, X. (2016). Identification of Regulatory DNA Elements Using Genome-wide Mapping of DNase I Hypersensitive Sites during Tomato Fruit Development. Mol. Plant *9*, 1168–1182.

Reddy, G.V., and Meyerowitz, E.M. (2005). Stem-cell homeostasis and growth dynamics can be uncoupled in the Arabidopsis shoot apex. Science *310*, 663–667.

Rodriguez, K., Perales, M., Snipes, S., Yadav, R.K., Diaz-Mendoza, M., and Reddy, G.V. (2016). DNA-dependent homodimerization, sub-cellular partitioning, and protein destabilization control WUSCHEL levels and spatial patterning. Proc. Natl. Acad. Sci. *113*, E6307–E6315.

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.

Shen, L., Wang, C., Fu, Y., Wang, J., Liu, Q., Zhang, X., Yan, C., Qian, Q., Wang, K. (2016). QTL editing confers opposing yield performance in different rice varieties. *Journal of integrative plant biology*, 22, pp. 1–4. doi: 10.1111/jipb.12501.

Somssich, M., Je, B. Il, Simon, R., and Jackson, D. (2016). CLAVATA-WUSCHEL signaling in the shoot meristem. Development *143*, 3238–3248.

Soyk, S., Lemmon, Z.H., Oved, M., Fisher, J., Liberatore, K.L., Park, S.J., Goren, 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.

Svitashev, S., Young, J.K., Schwartz, C., Gao, H., Falco, S.C., and Cigan, A.M. (2015). Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA. Plant Physiol. *169*, 931–945.

Swinnen, G., Goossens, A., and Pauwels, L. (2016). Lessons from Domestication: Targeting Cis-Regulatory Elements for Crop Improvement. Trends Plant Sci. *21*, 506–515.

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

Voytas, D.F., and Gao, C. (2014). Precision genome engineering and agriculture: opportunities and regulatory challenges. PLoS Biol. *12*, e1001877.

Wang, C., Hu, S., Gardner, C., and Lübberstedt, T. (2017). Emerging Avenues for Utilization of Exotic Germplasm. Trends Plant Sci. *22*, 624–637.

Wang, G.-D., Xie, H.-B., Peng, M.-S., Irwin, D., and Zhang, Y.-P. (2014). Domestication genomics: evidence from animals. Annu. Rev. Anim. Biosci. *2*, 65–84.

Weber, B., Zicola, J., Oka, R., and Stam, M. (2016). Plant Enhancers: A Call for Discovery. Trends Plant Sci. *21*, 974–987.

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. Bioengineered *3*, 38–43.

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

Xu, C., Liberatore, K.L., MacAlister, C.A., Huang, Z., Chu, Y.-H., Jiang, K., Brooks, C., Ogawa-Ohnishi, M., Xiong, G., Pauly, M., et al. (2015). A cascade of arabinosyltransferases controls shoot meristem size in tomato. Nat. Genet. *47*, 784–792.

Zamir, D. (2001). Improving plant breeding with exotic genetic libraries. Nat. Rev. Genet. *2*, 983–989.

Zhang, H., Zhang, J., Wei, P., Zhang, B., Gou, F., Feng, Z., Mao, Y., Yang, L., Zhang, H., Xu, N., et al. (2014). The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. Plant Biotechnol. J. *12*, 797–807.

FIGURE LEGENDS

Figure 1. Recreating a known fruit size QTL in tomato

(A) The conserved *CLV3*-*WUS* negative feedback circuit controls meristem size. LP, leaf primordia.

(B) The *fas* and *lc* fruit size QTL increased locule number (arrowheads) during domestication. Yellow arrowheads, locules.

(C) *fas* is caused by an inversion with a breakpoint 1 Kbp upstream of *SlCLV3*.

(D) The *lc* QTL (red rectangle) is associated with two SNPs (in bold) in a putative repressor motif (CArG, blue-dashed square) 1.7 Kbp downstream of *SlWUS*.

(E) CRISPR/Cas9-induced deletions in the CArG repressor motif (blue-dashed square) of *S.pim* and *S.lyc*. The gRNA target sequence is highlighted in red and the PAM site underlined.

(F) *S.pim*-lc^{CR} plants produce fruits with more than two locules. *S.pim*-fas^{NIL} *S.pim*-lc^{CR} double mutants synergistically increase locule number.

(G) Locule number is increased in *S.lyc*-lc^{CR} lines, and double mutants with *S.lyc*-fas^{NIL} are enhanced. N, plant number; n, fruit number. *P*: two-tailed, two-sample *t*-test.

Data in (F and G) are presented as percentage of fruits per locule number category. N = plants per genotype; n = fruit number. See also **Table S1**. *P*: two-tailed, two-sample *t*-test. Scale bars = 100 μm (A) and 1 cm (B, F and G).

Figure 2. Inducing mutations in the *SlCLV3* promoter using CRISPR/Cas9

(A) Model showing how an allelic series of *SlCLV3* transcriptional alleles could provide a range of quantitative effects on floral organ number according to a simple linear relationship of reduced expression resulting in increased phenotypic severity. WT, *fas*, and *clv3*^{CR} are shown as reference points in this hypothesized continuous relationship.

(B) Schematic of *SlCLV3* promoter targeted by eight gRNAs (numbered blue arrowheads). Blue arrows, PCR primers.

(C) PCR showing multiple deletion alleles in four T_0 plants. Amplicons were obtained using primers spanning the entire target region.

(D) Weak and strong effects on flower morphology and fruit size were observed among T_0 lines. Number of floral organs and locules are indicated.

(E) Quantification of floral organ number (mean ± SD; n≥10) in T_0, WT, *fas*, and *slclv3*^{CR} plants.

(F) Sequencing of *SlCLV3* promoter alleles for all T_0 plants. Deletions (-) and insertions (+) indicated by numbers or letters. $T_{0\text{-}5}$ and $T_{0\text{-}6}$ contained only WT alleles (not shown). Blue arrowheads, gRNAs; a, allele.

(G) PCR genotyping of T_1 progeny from $T_{0\text{-}1}$ and $T_{0\text{-}2}$. *UBIQUITIN* (*UBI*) served as an internal control. Absence of amplification for the target region of *SlCLV3* indicated homozygous plants for hidden alleles in both $T_{0\text{-}1}$ and $T_{0\text{-}2}$.

(H) Genome sequencing of $T_{0\text{-}1}$ and $T_{0\text{-}2}$ offspring homozygous for non-amplifiable alleles. Vertical dashed lines show target region. Neighboring genes, transposable elements and repeats upstream of the target region are shown. See also **Table S2**.

(I) Floral organ quantification (mean ± SD; N≥5) from homozygous plants for each of the four $T_{0\text{-}1}$ and $T_{0\text{-}2}$ alleles. Black arrowheads indicate WT values. See also **Table S3.**

Scale bars = 100 μm and 1cm (A), 1 cm (D).

Figure 3. A CRISPR/Cas9 driven genetic screen to rapidly generate and evaluate many *SlCLV3* promoter alleles for quantitative variation

(A) Crossing scheme for generating a sensitized F1 population heterozygous for a $T_{0\text{-}2}$ inherited allele and segregating for a CRISPR/Cas9 transgene (blue-dashed square). Expected segregation frequencies are indicated (%).

(B) Model showing how Cas9 activity in $Cas9^{+/-}$ hemizygous plants creates new mutant alleles (colored boxes) by targeting the WT *SlCLV3* promoter ($SlCLV3^{pro}$) introduced from the cross. Alleles derived from $T_{0\text{-}2}$ are shown as black or dark gray boxes. The transgene containing the CRISPR/Cas9 cassette (*Cas9*) is shown (red box).

(C) Locule number for WT, *fas*, and F1 plants grouped into three phenotypic categories: strong, moderate, weak. Data are presented as percentage of fruits per locule number category. N = plants per category. See also **Figures S1A, S1B** and **Table S4**.

Gray arrow in (A) points to the number of plants obtained for the forward genetics screen.

Figure 4. A collection of 14 engineered *SlCLV3* promoter alleles provides a continuum of locule number variation

(A) PCR showing new *SlCLV3* alleles from F1 plants with strong (s) and moderate (m) effects. Red asterisks, chimeric plants. Blue arrowheads, selected F1 plants for recovering new alleles (See also **Figure S1B**). Green arrow, inherited $SlCLV3^{CR\text{-}pro2\text{-}1}$ allele. Lower panel, $SlCLV3^{CR\text{-}pro2\text{-}2}$ genotyping (see **STAR Methods**).

(B) Non-transgenic F2 individuals homozygous for a new allele (blue asterisk and yellow box) are expected to segregate 1/16 from biallelic F1s (highlighted in dark gray).

(C) Segregation for a new allele and *Cas9* from $SlCLV3^{CR\text{-}pro2\text{-}1/m1}$ F2 population. *m1*, new allele from F1 moderate-1. Blue arrowheads indicate non-transgenic homozygotes. Lower panel, *Cas9*. Absence of band indicates transgene-free individuals.

(D) Sequences of 14 new $SlCLV3^{pro}$ alleles. Deletions (-) and insertions (+) indicated as numbers or letters. gRNAs, blue arrowheads. Parental F1s marked at right (See also **Figure S1B**). Connected arrows indicate similar phenotypes for *fas* and the ($SlCLV3^{CR\text{-}pro}$) m6-derived allele. Quantification of locule number (percent of fruits; N≥5 plants; mean ± SD shown) from homozygous F3 families is shown next to each allele sequence. *fas* and WT are references (See also **Table S5**).

(E) mVISTA plots of the *SlCLV3* promoter across four *Solanaceae* species (see **STAR METHODS**). Four conserved regions (CR) are indicated by blue shading. Dark blue regions indicate high sequence similarity (>70%) over at least 100 bp, as compared to *S.lyc*. CR1, 3, and 4 are conserved between *S.lyc, S.pen, S.tub,* and *C.ann*. CR2 is conserved only between *S.lyc, S.pen,* and *S.tub*.

(F) qRT-PCR of *SlCLV3* and *SlWUS* from reproductive meristems (mean ± SEM; two biological and three technical replicates) for WT, *fas*, and each $SlCLV3^{CR\text{-}pro}$ allele, normalized to *UBI*. Dashed lines mark WT levels for *SlCLV3* (red) and *SlWUS* (blue).

Figure 5. CRISPR/Cas9 targeting of the *S* and *SP* promoters results in a range of modified inflorescence and plant architectures

(A) The *S* promoter was targeted by eight gRNAs (blue arrowheads). PCR genotyping showed distinct deletion alleles in all three T_0 plants. Blue arrows, PCR and cloning primers. $T_{0\text{-}2}$ showed inflorescence branching similar to *s* mutants (red arrowheads).

(B) Representative inflorescence from F1 sectored individual showing excessive meristem overproliferation.

(C) Sequences of six new $S^{CR\text{-}pro}$ alleles and associated phenotypes. Deletions (-) and insertions (+) indicated as numbers or letters. gRNAs, blue arrowheads. Quantification of percent of inflorescences with indicated branch number for homozygous mutants (N≥5 plants) from segregating F2 families. WT, *s-classic* and heterozygotes (*s*/+) are references (See also **Table S6**). Red asterisks indicate *s-classic* coding sequence mutation (Lippman et al., 2008). Data for *s*/+ from Soyk et al., 2017.

(D) Representative images showing the range of inflorescence branching (red arrowheads) for three $S^{CR\text{-}pro}$ alleles.

(E) The *SP* promoter was targeted by eight gRNAs (blue arrowheads). PCR genotyping showed distinct deletion alleles in all four T_0 plants, which appeared biallelic or chimeric. Blue arrows, PCR and cloning primers.

(F) Sequences of *SP*$^{CR\text{-}pro}$ alleles from the four T_0 plants. Deletions (-) and insertions (+) indicated as numbers or letters. gRNAs, blue arrowheads.

(G) Representative main shoots from WT, *sp-classic* and three *SP*$^{CR\text{-}pro}$ alleles. Brackets and numbers indicate number of leaves between inflorescences in successive shoots. Red arrowheads, inflorescences. D, determinate; ID, indeterminate; SD, semideterminate.

(H) Average number of leaves in sympodial shoots from WT, *sp-classic* and three *SP*$^{CR\text{-}pro}$ alleles. See also **Table S7**.

Data in (H) are shown as mean ± SEM. N = number of individuals. Scale bars = 1 cm (A, B, D); 100 μm (B); 5 cm (G).

SUPPLEMENTARY FIGURE LEGENDS

Figure S1. A library of transgenic F1 plants exhibiting a range of quantitative variation for fruit locule number. Related to Figure 3.

(A) Locule number quantification (percent of fruits) from 92 F1 plants comprising the weak (w) phenotypic category.

(B) 24 F1 plants showed moderate (m) to strong (s) increases in locule number (percent of fruits), from which 14 plants were selected for isolating new *SlCLV3*^{pro} alleles (black arrowheads).

Data in (A and B) are presented as percentage of fruits per locule number category (See also **Table S4**).

Figure S2. Comparative sequence analysis of the *SlCLV3* promoter in the *Solanaceae*. Related to Figure 4.

(A) Predicted conserved regions and cis-regulatory elements in the *SlCLV3* promoter. Transcription factor binding sites (TFBS) were predicted using either the Cistrome or JASPAR at relative profile score thresholds of 95% (low stringency) and 99% (high stringency) (see **STAR METHODS**). DNase hypersensitivity sites (DHS) were from the break stage of tomato fruit development. Regions of low genetic diversity (low SNP diversity and Tajima's D hits) indicate regions of the *SlCLV3* promoter potentially under selection (calculated as described in the **STAR METHODS**). gRNA target sites are specified as blue arrowheads. Conserved regions (CRE) were defined using pairwise sequence alignments, as implemented in mVISTA. The pattern of deletions for each

allele at every gRNA target site is shown and color-coded according to their phenotypic strength. The allele name and deletion sizes are indicated.

(B) Full VISTA plot showing conserved regions (at least 70% similarity, 100bp window) of the *SlCLV3* promoter (as compared to *S. lyc*) in *S. pen*, *S. tub*, and *C. ann*.

STAR METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

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).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Plant materials and growth conditions

Seeds of *S. pimpinellifolium* (LA1589), *S. lycopersicum* cv. M82 (LA3475), fas^{NIL} in both *S.pimpinellifolium* and M82, lc^{NIL} and $fas^{NIL}lc^{NIL}$ (*S. pimpinellifolium*) and the mutants $slcv3^{CR}$, *s* and *sp-classic* in M82 background were our own stocks and from stocks kindly provided by E. van der Knaap. Seeds were directly sown in soil in 96-cell plastic flats. Plants were grown in a greenhouse under long-day conditions (16-h light/8-h dark) supplemented with artificial light from high-pressure sodium bulbs (~250 μmol m^{-2} s^{-1}).

METHOD DETAILS

Plant phenotyping

To quantity floral organ and locule number, flowers or fruits from multiple inflorescences were dissected and each organ quantified separately. Shoot determinacy was assessed by counting leaves from five successive sympodial shoots from greenhouse

grown plants. Inflorescence branching was determined by counting the number of branch points in at least 5 inflorescences per plant.

CRISPR/Cas9 constructs with two or eight gRNAs

A binary vector containing a *CRISPR* cassette with a functional Cas9 under a constitutive promoter (*CaMV 35s*) and either two or eight gRNA (**Table S8**) was made using standard Golden Gate assembly (Brooks et al., 2014; Werner et al., 2012). To check for specificity, BLAST analyses for each gRNA target site including the PAM site (NGG) were performed against the tomato genome (SL2.50) (Tomato Genome Consortium, 2012). For the CRISPR/Cas9 construct targeting the *LC* QTL region, two gRNA target sites were selected manually, based on the previously annotated *LC* SNPs and the predicted repressor motif (CArG). To produce each gRNA, a PCR reaction was carried out with a primer containing the gRNA sequence (**Table S8**), using the plasmid pICH86966::AtU6p::gRNA_PDS (Addgene plasmid 46966) as template. Each gRNA was cloned individually into the level 1 vectors *pICH47732* (gRNA1) and *pICH47742* (gRNA2). Level 1 construct *pICH47732-NOSpro::NPTII* (selection maker), *pICH47742-35S:Cas9* and the gRNAs were then assembled in the binary Level 2 vector *pAGM4723*.

For the CRISPR/Cas9 construct carrying eight gRNAs targeting the promoters of *SlCLV3, S* and *SP*, eight potential 20 bp sites were selected for gRNA design within a region of 2 Kbp (*SlCLV3* and *SP*) and 4 Kbp (*S*) upstream of the start codon (ATG) of each gene using the CRISPR-P tool (Lei et al., 2014). To minimize recovering *SlCLV3* alleles with strong effects similar to coding sequence null mutations, the first gRNA targeted 130 bp upstream of the translational start site (ATG) to avoid the transcription

start site. Target sites were mainly selected based on BLAST results that gave little or no potential off-targets and were spaced apart between 100, 400 bp or more (**Table S8**). The *SP* promoter was derived from an introgression of *S.pen* to create indeterminate plants and 3 gRNAs had target site mismatches. Each gRNA was cloned into level 1 vectors *pICH47732* (gRNA1 or gRNA8), *pICH47742* (gRNA2), *pICH47751* (gRNA3), *pICH47761* (gRNA4), *pICH47772* (gRNA5), *pICH47781* (gRNA6), *pICH47791* (gRNA7). gRNAs were then assembled into two groups in an intermediate cloning step, using level M vectors *pAGM8055* and *pAGM8093*. Level 1 construct *pICH47732-NOSpro::NPTII* (selection maker), *pICH47742-35S:Cas9* and level M vectors containing the assembled gRNAs cassettes were then assembled in the binary Level 2 vector *pAGM4723*. All restriction-ligation Golden Gate reactions were carried out in a volume of 15 µL in a thermal cycler (3 min at 37 °C and 4 min at 16° for 20 cycles; 5 min at 50 °C, 5 min at 80 °C, and final storage at 4 °C).

The final binary vectors were introduced into either *S. pim* or *S. lyc* M82 by *Agrobacterium tumefaciens*–mediated transformation as previously described (Brooks et al., 2014; Gupta and Van Eck, 2016). First-generation (T_0) transgenic plants were transplanted in soil and grown under standard greenhouse conditions. CRISPR/Cas9-generated mutations were genotyped by PCR amplification of the target region in DNA extracted from pooled main and axillary shoots. Primers were designed to bind 250 and 400 bp away from the outermost gRNAs (**Table S8**). PCR products were analyzed by gel electrophoresis and cloned into *pSC-A-amp/kan* vector (Agilent) following manufacturer's instructions. At least 3 clones per sample were sequenced using sequencing primers spanning the target region (**Table S8**). Sequence assembly was

carried out using SeqMan software from Lasergene 8 suite, using standard parameters. Due to the presence of big deletions or inversions, manual assembly edition was performed to ensure proper alignment and reconstruction of the sequenced alleles.

Recovery of homozygous progeny from T_0 plants

S.pim or *S.lyc* lc^{CR} T_0 lines were crossed to fas^{NIL} and F2 populations were genotyped for *Cas9* and lc^{CR}. Genotyping for fas^{NIL} was carried out as described (Xu et al., 2015). Genotyping for lc^{CR} was performed by sequencing of cloned PCR products (**Table S8**). At least 3 clones per sample were sequenced and reads were assembled using SeqMan software from Lasergene 8 suite, using standard parameters. Locule number was quantified in stable non-transgenic single and double mutants.

To isolate plants homozygous for new alleles derived from $SlCLV3^{pro}$ $T_{0\text{-}1}$ and $T_{0\text{-}2}$, DNA was extracted from T_1 and T_2 individuals and genotyping was performed using primers that amplified the target region (**Table S8**). All T_1 plants showing consistent lack of PCR amplification produced T_2 progeny also lacking the PCR products of the target region, and floral organ and fruit locule number quantification was performed for at least 6 replicate plants, using multiple inflorescences and flowers (≥80) from each replicate (**Table S4**). Offspring of independent T_1 plants showing positive PCR reactions were analyzed for segregation of unsuccessful PCR amplification of the target region. T_1 progeny that did not segregate for failed PCR reactions were selected for floral organ quantification. *UBI* gene was used as internal control of quality of DNA and PCR (**Table S8**).

To obtain homozygotes for the alleles from *SP*^{pro} T_{0-3} and T_{0-4} plants, T_1 progeny was genotyped using PCR primers amplifying the target region (**Table S8**). The progeny from both T_0 plants only inherited alleles that were successfully amplified by PCR. Phenotyping for shoot determinacy was done as previously described (Park et al., 2014) on T_2 homozygous offspring for the recovered alleles from T_{0-3} and T_{0-4} plants (**Table S7**).

Genome sequencing

One T_2 line carrying *SlCLV3*$^{CR\text{-}pro1\text{-}2}$ (isolated from T_{0-1}) and three T_2 lines carrying *SlCLV3*$^{CR\text{-}pro2\text{-}2}$ (isolated from T_{0-2}), were sequenced to elucidate the nature of the CRISPR/Cas9-induced lesions and analyze potential off-target mutations. DNA was extracted from one plant for each line and libraries were produced from 2 µg of genomic DNA sheered to 550 bp with the Illumina TruSeq DNA PCR-free prep kit. All four libraries were sequenced on a single lane of the Illumina NextSeq 500 platform at the Cold Spring Harbor Laboratory Genome Center (Woodbury, NY). For the *SlCLV3*$^{CR\text{-}pro1\text{-}2}$ line, we obtained 36,032,000 paired end 151 bp reads. A total of 34,775,494, 44,539,109, and 30,820,245 paired end 151 bp reads were collected for the three *SlCLV3*$^{CR\text{-}pro2\text{-}2}$ lines (pedigrees: 16-2377-1, 16-2378-1, and 16-2378-2), respectively. In total 146,166,848 paired end 151 bp reads were obtained for the *SlCLV3*$^{CR\text{-}pro}$ mutants. Additionally, 287,344,857 million Illumina Genome Analyzer IIx paired end 80 and 100 bp reads were obtained from the European Nucleotide Archive (ENA) at the European Bioinformatics Institute (EBI) for wild type *S. lycopersicum* cv. M82.

Genomic DNA reads were trimmed by quality using Trimmomatic v0.32 (parameters: ILLUMINACLIP:TruSeq3-PE-2.fa:2:40:15:1:FALSE LEADING:30 TRAILING:30 MINLEN:75) (Bolger et al., 2014) and paired reads mapped to the tomato reference genome (SL2.50) (Tomato Genome Consortium, 2012) using BWA-MEM v0.7.10-r789 (parameters: -M) (Li and Durbin, 2009). Alignments were then sorted with samtools and duplicates marked with PicardTools v1.126 (parameters: VALIDATION_STRINGENCY=LENIENT) (Li and Durbin, 2009), http://broadinstitute.github.io/picard). Alignments were analyzed using Scalpel v0.5.3 (parameters: scalpel-discovery --single --pathlimit 100000) to detect small to moderately sized indels (Narzisi et al., 2014), the typical lesion induced by Cas9. A custom Perl script was used to scan the genome for putative off target sites including the protospacer adjacent motif (PAM) NGG, with at most 3 mismatches in the 12 bp proximal to the PAM. Of the 94,069 identified potential off target sites, only 33 putative Scalpel indels were located ±5 bp. There were no indels found within genes. Read alignments at these sites were hand inspected and after removing repeated scalpel calls, the 24 remaining sites represented either tandem repeats within transposable elements (TE) or in intergenic regions surrounded by repetitive sequences (**Table S3**).

Raw sequencing reads generated in this study have been deposited at the Sequence Read Archive (http://ncbi.nlm.nih.gov/sra) under accession number SRP107576. Wild type *S. lycopersicum* cv. M82 sequencing reads were obtained from the European Bioinformatics Institute under project number PRJEB6302.

Genetic scheme to generate *SlCLV3 and S* promoter alleles

F1 populations were obtained by crossing the original *SlCLV3*$^{CR\text{-}pro}$ $T_{0\text{-}2}$ or *S*$^{CR\text{-}pro}$ $T_{0\text{-}2}$ plant as a pollen donor to emasculated wild type M82 flowers. F1 seeds were extracted and germinated in 96-wells flats and genotyped for the *Cas9* coding sequence (**Table S8**). Confirmed F1 individuals carrying the inherited CRISPR/Cas9 transgene in both outcrosses were transplanted in the field at the Uplands Farm of Cold Spring Harbor Laboratory, New York. Plants were grown under standard drip irrigation and fertilizer regimes. Six weeks after planting, the *SlCLV3*$^{CR\text{-}pro}$ F1 plants were individually inspected for increased sepals and petals in flowers from the first inflorescences, and plants showing phenotypes were marked for later analysis of locule number. Wild-type plants as well as those with multiple phenotypic sectors were removed to allow better growth of the remaining F1 plants. Following fruit set, locule number was quantified and plants were grouped into three phenotypic categories for effect on increased locule number: "weak", "moderate", and "strong". DNA was extracted from moderate and strong classes and genotyped by PCR amplification of the target region and for the two alleles inherited from $T_{0\text{-}2}$ (**Figure 4A** and **Table S8**). F2 progeny seed was collected for each F1 plant from the three phenotypic categories. For *S*$^{CR\text{-}pro}$, F1 plants were individually inspected and those exhibiting inflorescence branching were tagged as moderate in effect if most inflorescences showed four or less branching events, and tagged as strong if more than four. Seeds were collected for all F1 plants showing inflorescence branching.

For *SlCLV3*$^{CR\text{-}pro}$, the 24 F2 families from moderate and strong categories were grown under greenhouse conditions and genotyped for both *Cas9* and the *SlCLV3* promoter to identify non-transgenic biallelic or homozygous plants carrying new alleles.

Sequencing of new alleles was performed for at least three cloned products per sample in homozygous F3 progeny from a subset of 14 F2 families spanning the phenotypic range of locule number (**Figure S1B**). The effect of new alleles on locule number was quantified in five or more homozygous F3 progeny (**Table S6**). For $S^{CR\text{-}pro}$, ten families were selected and progeny was genotyped for *Cas9* and the targeted region of the promoter. Sequencing and promoter allele assembly was done as described for $SlCLV3^{CR\text{-}pro}$. The effect of new alleles on inflorescence branching was quantified in at least 5 homozygous F2 individuals growing under field conditions.

Sequence assembly was carried out using SeqMan software from Lasergene 8 suite, using standard parameters. Due to the presence of big deletions or inversions, manual assembly edition was performed to ensure proper alignments and reconstruction of the sequenced alleles.

Comparative sequence analysis of the *CLV3* promoter in the *Solanaceae*

Conserved regions and putative *cis*-regulatory elements in the *CLV3* promoter were identified using phylogenetic shadowing, searches for transcription factor binding sites, searches for regions of open chromatin (as assessed by other investigators (Qiu et al., 2016), analyses of SNP diversity, and tests for selection. For phylogenetic footprinting, orthologous *SlCLV3* genome regions in *S. pennellii, S. tuberosum*, and *C. annuum* were identified using CoGE GEvo tool (Lyons and Freeling, 2008). The regions ~2,000bp upstream of the *CLV3* CDS were scored for sequence homology to *S. lycopersicum CLV3* using mVISTA (Frazer et al., 2004). Alignment windows of 100bp at a similarity threshold of 70%. JASPAR Core Plantae (Mathelier et al., 2014) was used to

find potential TFBSs. All plant TFs in the curated database were included, using relative profile score thresholds of 95% and 99% (shown in different tracks in **Figure S2A**). TFBSs in the noncoding regions flanking *Arabidopsis thaliana CLV3* were identified from Cistrome (O'Malley et al., 2016), and mapped onto the *SlCLV3* promoter. The MEME (Bailey et al., 2009) global binding profile matrixes for these TFs were used to interrogate the *SlCLV3* promoter region for matches using FIMO (Grant et al., 2011). Tomato DNAse hypersensitivity data from break stage fruit, indicating regions of open chromatin, had only one peak near *SlCLV3* (Qiu et al, 2016). Tomato genome re-sequencing data (100 Tomato Genome Sequencing Consortium et al., 2014) was used to assess SNP-density in the *SlCLV3* promoter at the population level, and to test for selection. Variable call files containing SNPs relative to the reference genome for 76 tomato accessions were downloaded from solgenomics.org and merged using the GATK toolkit CombineVariants function (McKenna et al., 2010) . We calculated SNP density in 20bp bins using the VCFtools SNPdensity function (Danecek et al., 2011). Regions where average SNP density was lower than one standard deviation below the mean SNP density for the *SlCLV3* exons were identified and mapped to the tomato genome using R. Estimates of Tajima's D were calculated using Tassel software (Bradbury et al., 2007), using a sliding window analysis with 20bp steps and a 200bp window size. Sub-regions with a Tajima's D score $\leq$ 1 standard deviation below the mean (-2.0059) Tajima's D for the entire *CLV3* genomic region (-2500bp and +500bp) were interpreted to be evolving non-neutrally. Sub-regions meeting this criterion were then mapped to the tomato genome using R.

RNA extraction and Quantitative RT-PCR (qPCR)

Seeds were germinated in Petri dishes with moistened paper towels and then transferred to 96-well flats. Mersitems at the transition stage (11-13 days after germination) were collected from shoot apices and fixed in 100% acetone as previously described (Xu et al., 2015).

Total RNA from meristems was then extracted using the PicoPure RNA Extraction kit (Thermo Fisher). 200 ng of total RNA was used for cDNA synthesis with SuperScript III reverse-transcriptase kit (Invitrogen). qPCR was performed with gene-specific primers using the iQ SYBR Green SuperMix (Bio-Rad) reaction system on the CFX96 Real-Time system (Bio-Rad), following manufacturer's instructions. *UBI* gene was used as internal control (Table S8).

QUANTIFICATION AND STATISTICAL ANALYSES

For quantitative analysis in floral organ and locule number in **Figures 1F-G**, **2I**, **3C**, **4E, S1A-B**, at least 3 primary or secondary inflorescences from ≥3 individuals per genotype were analyzed. For quantitative analysis in **Figure 2E**, at least 10 flowers from each T_0 plant (N = 1), *fas* (N = 3), and $slclv3^{CR}$ (N = 3) were analyzed. Quantification of sympodial shoot flowering in **Figure 5D** was performed by counting the number of leaves in the first 5 inflorescences of the main shoot. Number of individuals (N) are presented in the figures and shown in **Tables S1-S8**. Statistical calculations were performed using Microsoft Excel using raw numbers. Mean values for each measured parameter were compared using two-tailed, two-samples Student's *t*-test.

DATA AND SOFTWARE AVAILABILITY

Raw sequencing reads generated in this study have been deposited at the Sequence Read Archive (http://ncbi.nlm.nih.gov/sra) under BioProject SRP107576.

ADDITIONAL RESOURCES

CRISPR design: http://cbi.hzau.edu.cn/cgi-bin/CRISPR

Data deposition: http://ncbi.nlm.nih.gov/sra

SUPPLEMENTARY ITEMS

Table S1. Quantification of locule number from *S.pim and S.lyc* lc^{CR} fruits. Related to Figure 1.

Table S2. Predicted off target sites from $SlCLV3^{pro}$ gRNAs. Related to Figure 2.

Table S3. Floral organ number quantification from plants homozygous for each of four alleles recovered from $SlCLV3^{CR\text{-}pro}$ $T_{0\text{-}1}$ and $T_{0\text{-}2}$ plants. Related to Figure 2.

Table S4. Quantification of locule number in $SlCLV3^{CR\text{-}pro}$ F1 plants. Related to Figure 3.

Table S5. Quantification of locule number from F3 homozygous plants for each of 14 alleles derived from $SlCLV3^{CR\text{-}pro}$ F2 populations. Related to Figure 4.

Table S6. Quantification of inflorescence branching on $S^{CR\text{-}pro}$ alleles. Related to Figure 5.

Table S7. Quantification of flowering time measured by leaf number in successive sympodial shoots derived from the primary shoots of WT, *sp-classic*, and three $SP^{CR\text{-}pro}$ alleles. Related to Figure 5.

Table S8. List of all gRNA and primer sequences used in this study for CRISPR/Cas9 goldengate design, PCR detection, cloning and qPCR analysis. Related to STAR Methods.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
N/A		
Bacterial and Virus Strains		
N/A		
Biological Samples		
DNA and leaf tissue from tomato wild species and cultivar M82 .	See **Table S2**	N/A
Chemicals, Peptides, and Recombinant Proteins		
CTAB	Sigma Aldrich	Cat# H6269-500G
Agarose	VWR	Cat# 97062-250
BsaI	NEB	Cat# R0535L
BpiI	Thermo Fisher	Cat# ER1012
T4 DNA Ligase	NEB	Cat# M0202L
Acetone	Fisher Scientific	Cat# A928-4
Taq DNA Polymerase with Standard Taq Buffer	NEB	Cat# M0273L
KOD Xtreme Hot Start DNA Polymerase	Millipore	Cat# 71975
iQ SYBR Green Supermix	Bio-Rad	Cat# 17-8882
Critical Commercial Assays		
TruSeq DNA PCR-Free HT Library Preparation Kit	Illumina	Cat#FC-121-3003
RNase Free DNase Set	Qiagen	Cat# 79254
QIAprep Spin Miniprep Kit	Qiagen	Cat# 27106
QIAquick PCR Purification Kit	Qiagen	Cat# 28106
StrataClone Blunt PCR Cloning Kit	Stratagene	Cat# 240207
SuperScript III First-Strand Synthesis System	Invitrogen	Cat# 18080051
RNeasy Plant Mini Kit	Qiagen	Cat# 74904
ARCTURUS PicoPure RNA Isolation Kit	Thermo Fisher	Cat# KIT0204
Deposited Data		
Whole-genome sequencing data	This study	SRP107576
Experimental Models: Cell Lines		
N/A		
Experimental Models: Organisms/Strains		
Tomato wild species (*Solanum pimpinellifolium*)	See **STAR Methods**	N/A
Tomato cultivars (M82)	See **STAR Methods**	N/A
Oligonucleotides		
Guide RNA (gRNA) sequences, see **Table S8**	This study	N/A
Primer sequences for cloning, see **Table S9**	This study	N/A

Primer sequences for genotyping, see **Table S9**	This study	N/A
Primer sequences for sequencing, see **Table S9**	This study	N/A
Primer sequences for qRT-PCR, see **Table S9**	This study	N/A
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
Software and Algorithms		
Trimmomatic	(Bolger et al., 2014b)	http://www.usadellab.org/cms/?page=trimmomatic
BWA-MEM	(Li, 2013; Li and Durbin, 2009)	http://bio-bwa.sourceforge.net/
PicardTools	N/A	http://broadinstitute.github.io/picard
Scalpel	(Narzisi et al., 2014)	http://scalpel.sourceforge.net/
Custom Perl script for off-target sites search	This study	NA
SeqMan	DNASTAR Lasergene	http://www.dnastar.com
CoGe Gevo tool	(Lyons and freeling, 2008)	https://www.genomevolution.org/coge/
mVISTA	(Mathelier et al., 2014)	http://genome.lbl.gov/vista/mvista/about.shtml
JASPAR Core PLANTAE	(Mathelier et al., 2014)	http://jaspar.genereg.net/
Cistrome	(O'Malley et al., 2016)	http://www.cistrome.org/Cistrome/Cistrome_Project.html
MEME-suite	(Bailey et al., 2009)	http://meme-suite.org/
FIMO	(Grant et al., 2011)	http://meme-suite.org/
GATK toolkit	(McKenna et al., 2010)	https://software.broadinstitute.org/gatk/
VFCtools	(Danecek et al., 2011)	https://vcftools.github.io/man_latest.html
TASSEL	(Bradbury et al., 2007)	http://www.maizegenetics.net/tassel
Other		
N/A		

Figure 1

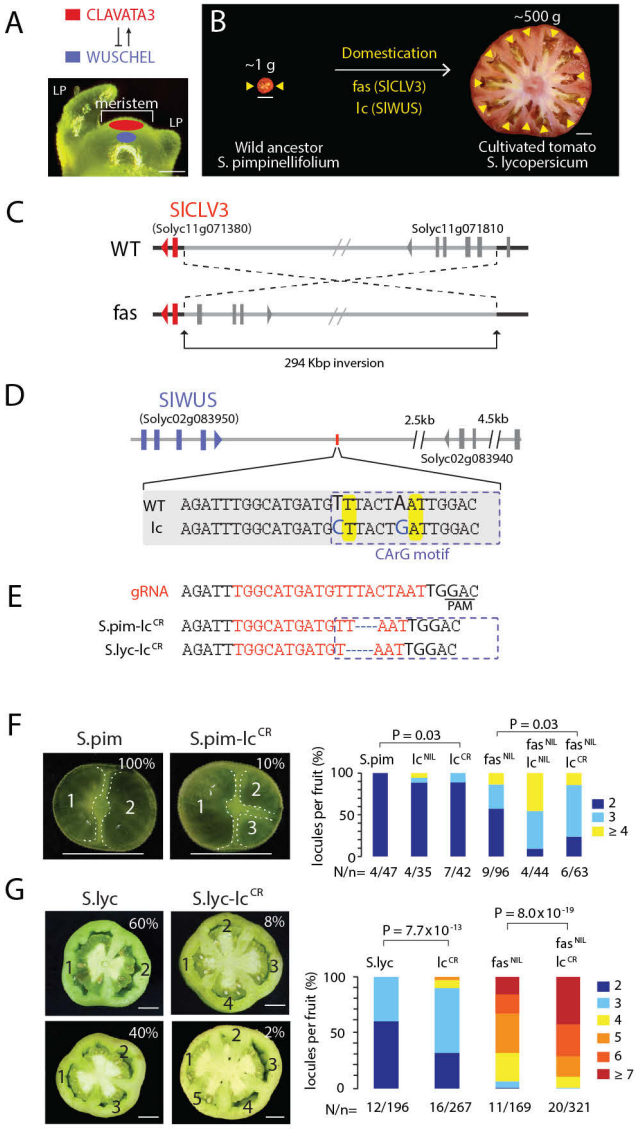

A
CLAVATA3
WUSCHEL
LP
meristem
LP
B
~1 g
Domestication
fas (SlCLV3)
lc (SlWUS)
~500 g
Wild ancestor
S. pimpinellifolium
Cultivated tomato
S. lycopersicum
C
SlCLV3
(Solyc11g071380)
Solyc11g071810
WT
fas
294 Kbp inversion
D
SlWUS
(Solyc02g083950)
2.5kb
4.5kb
Solyc02g083940
WT AGATTTGGCATGATGTTTACTAATTGGAC
lc AGATTTGGCATGATGCTTACTGATTGGAC
CArG motif
E
gRNA AGATTTGGCATGATGTTTACTAATTGGAC
PAM
S.pim-lc^CR AGATTTGGCATGATGTT----AATTGGAC
S.lyc-lc^CR AGATTTGGCATGATGT-----AATTGGAC
F
S.pim
S.pim-lc^CR
100%
10%
P = 0.03
P = 0.03
S.pim
lc^NIL
lc^CR
fas^NIL
fas^NIL lc^NIL
fas^NIL lc^CR
locules per fruit (%)
100
50
0
2
3
≥ 4
N/n= 4/47 4/35 7/42 9/96 4/44 6/63
G
S.lyc
S.lyc-lc^CR
60%
8%
40%
2%
P = 7.7x10^-13
P = 8.0x10^-19
S.lyc
lc^CR
fas^NIL
fas^NIL lc^CR
locules per fruit (%)
100
50
0
2
3
4
5
6
≥ 7
N/n= 12/196 16/267 11/169 20/321

Figure 2

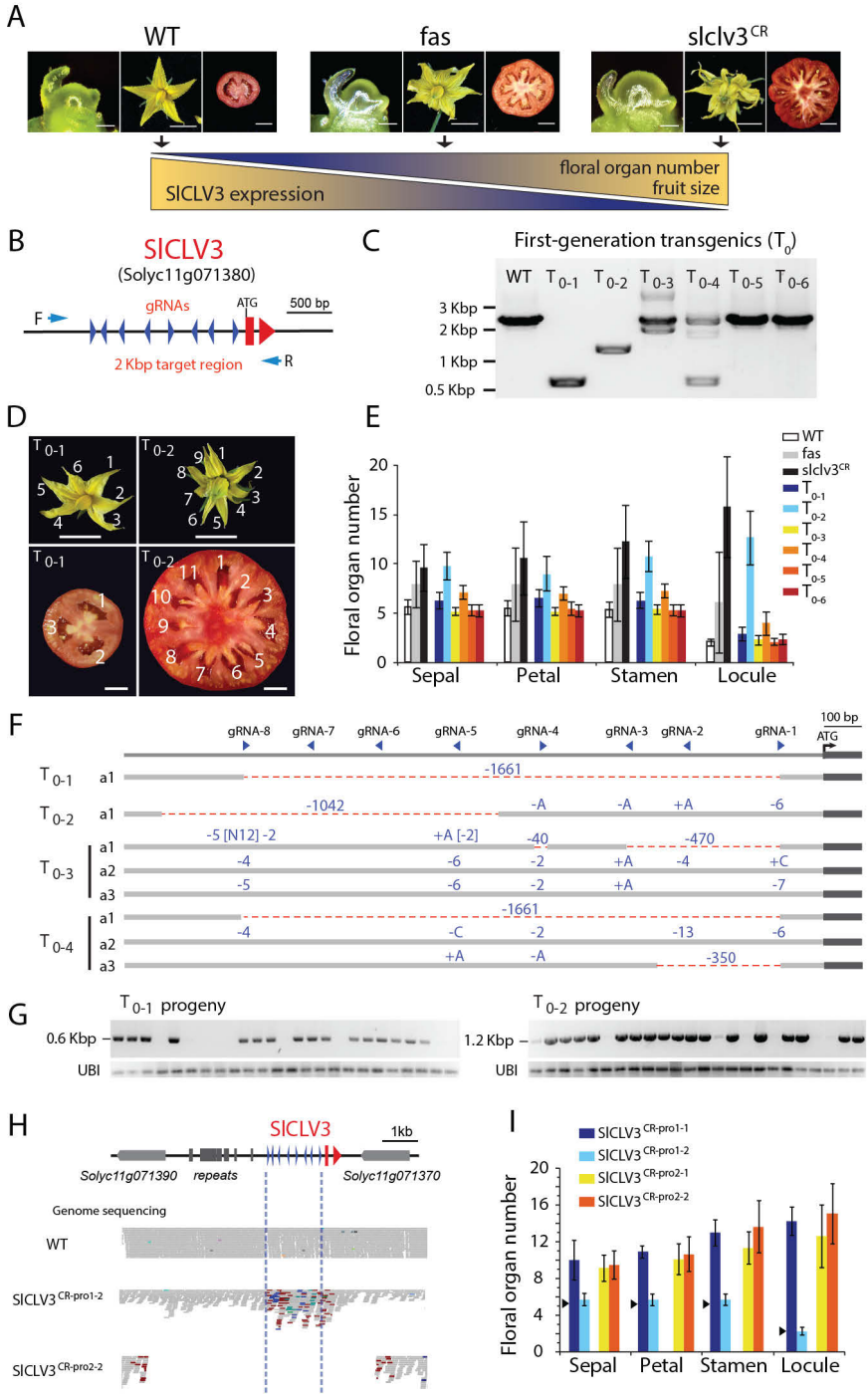

A
WT
fas
slclv3 CR
floral organ number
fruit size
SlCLV3 expression
B
SlCLV3
(Solyc11g071380)
gRNAs
ATG
500 bp
F
R
2 Kbp target region
C
First-generation transgenics (T0)
WT
T0-1
T0-2
T0-3
T0-4
T0-5
T0-6
3 Kbp
2 Kbp
1 Kbp
0.5 Kbp
D
E
Floral organ number
Sepal
Petal
Stamen
Locule
F
gRNA-8
gRNA-7
gRNA-6
gRNA-5
gRNA-4
gRNA-3
gRNA-2
gRNA-1
100 bp
ATG
-1661
-1042
-A
+A
-6
-5 [N12] -2
+A [-2]
-40
-470
-4
-2
-7
-C
-13
-350
G
T0-1 progeny
T0-2 progeny
0.6 Kbp
1.2 Kbp
UBI
H
SlCLV3
1kb
Solyc11g071390
repeats
Solyc11g071370
Genome sequencing
WT
SlCLV3 CR-pro1-2
SlCLV3 CR-pro2-2
I
SlCLV3 CR-pro1-1
SlCLV3 CR-pro1-2
SlCLV3 CR-pro2-1
SlCLV3 CR-pro2-2
Floral organ number
Sepal
Petal
Stamen
Locule

Figure 3

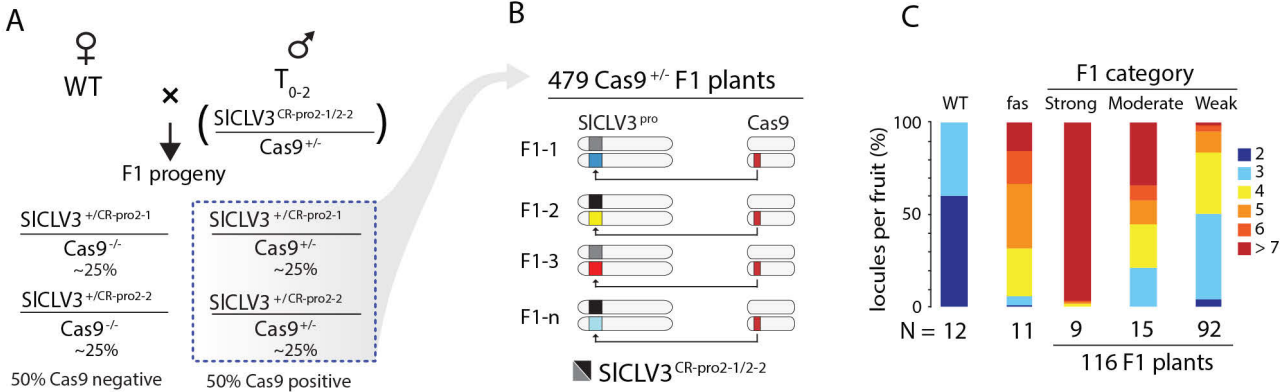
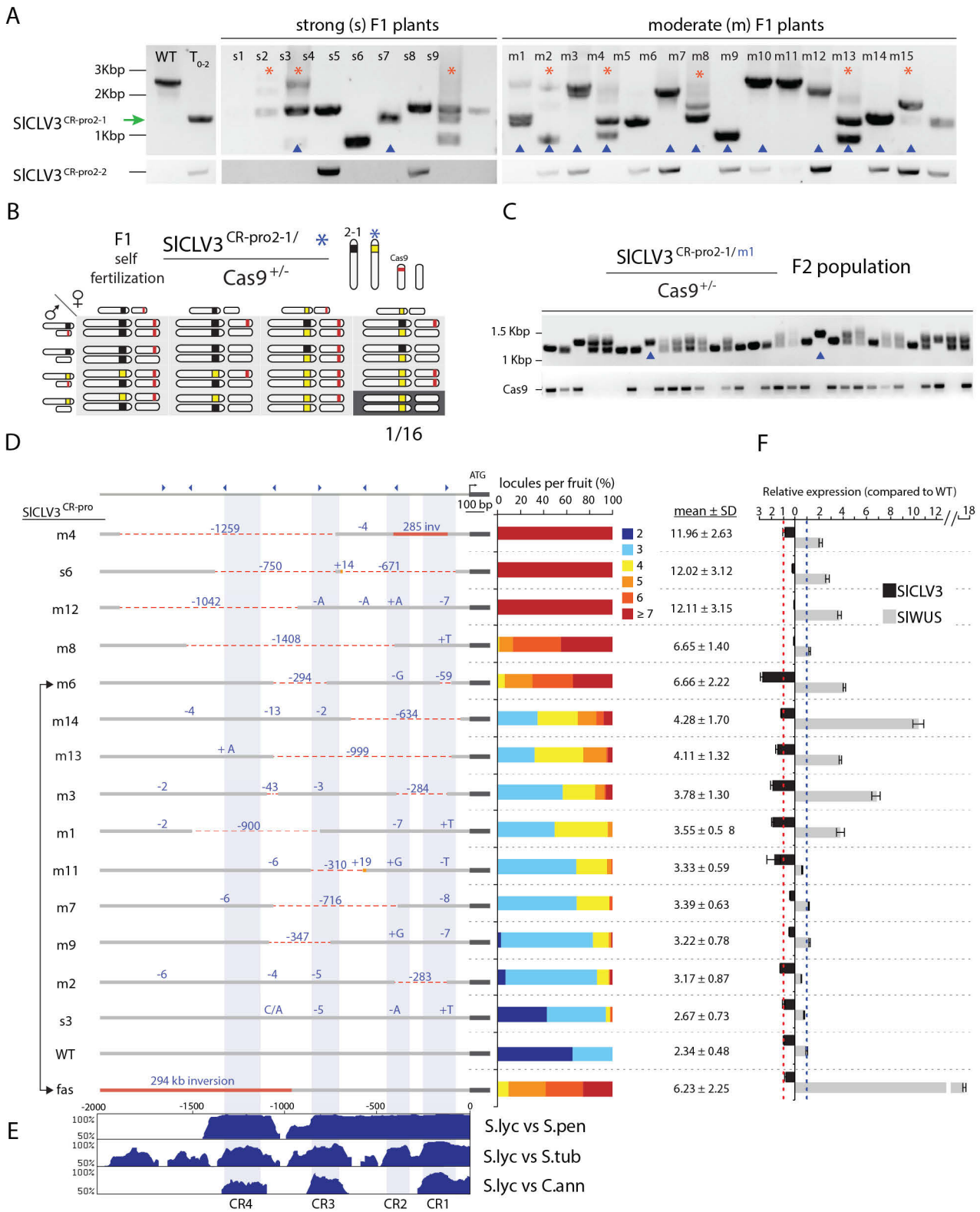

A
♀
WT
×
♂
T0-2
(SlCLV3 CR-pro2-1/2-2 / Cas9 +/-)
F1 progeny
SlCLV3 +/CR-pro2-1 / Cas9 -/- ~25%
SlCLV3 +/CR-pro2-2 / Cas9 -/- ~25%
SlCLV3 +/CR-pro2-1 / Cas9 +/- ~25%
SlCLV3 +/CR-pro2-2 / Cas9 +/- ~25%
50% Cas9 negative
50% Cas9 positive
B
479 Cas9 +/- F1 plants
SlCLV3 pro
Cas9
F1-1
F1-2
F1-3
F1-n
SlCLV3 CR-pro2-1/2-2
C
F1 category
WT
fas
Strong
Moderate
Weak
locules per fruit (%)
100
50
0
2
3
4
5
6
> 7
N = 12
11
9
15
92
116 F1 plants

Figure 4


A
strong (s) F1 plants
moderate (m) F1 plants
WT
T0-2
s1 s2 s3 s4 s5 s6 s7 s8 s9
m1 m2 m3 m4 m5 m6 m7 m8 m9 m10 m11 m12 m13 m14 m15
3Kbp
2Kbp
SlCLV3 CR-pro2-1
1Kbp
SlCLV3 CR-pro2-2
B
F1
self
fertilization
SlCLV3 CR-pro2-1/
Cas9 +/-
2-1
Cas9
1/16
C
SlCLV3 CR-pro2-1/m1
Cas9 +/-
F2 population
1.5 Kbp
1 Kbp
Cas9
D
SlCLV3 CR-pro
ATG
100 bp
locules per fruit (%)
0 20 40 60 80 100
2
3
4
5
6
≥ 7
mean ± SD
m4 -1259 -4 285 inv 11.96 ± 2.63
s6 -750 +14 -671 12.02 ± 3.12
m12 -1042 -A -A +A -7 12.11 ± 3.15
m8 -1408 +T 6.65 ± 1.40
m6 -294 -G -59 6.66 ± 2.22
m14 -4 -13 -2 -634 4.28 ± 1.70
m13 + A -999 4.11 ± 1.32
m3 -2 -43 -3 -284 3.78 ± 1.30
m1 -2 -900 -7 +T 3.55 ± 0.5 8
m11 -6 -310 +19 +G -T 3.33 ± 0.59
m7 -6 -716 -8 3.39 ± 0.63
m9 -347 +G -7 3.22 ± 0.78
m2 -6 -4 -5 -283 3.17 ± 0.87
s3 C/A -5 -A +T 2.67 ± 0.73
WT 2.34 ± 0.48
fas 294 kb inversion 6.23 ± 2.25
F
Relative expression (compared to WT)
3 2 1 0 2 4 6 8 10 12 18
SlCLV3
SlWUS
E
-2000 -1500 -1000 -500 0
100%
50%
CR4 CR3 CR2 CR1
S.lyc vs S.pen
S.lyc vs S.tub
S.lyc vs C.ann

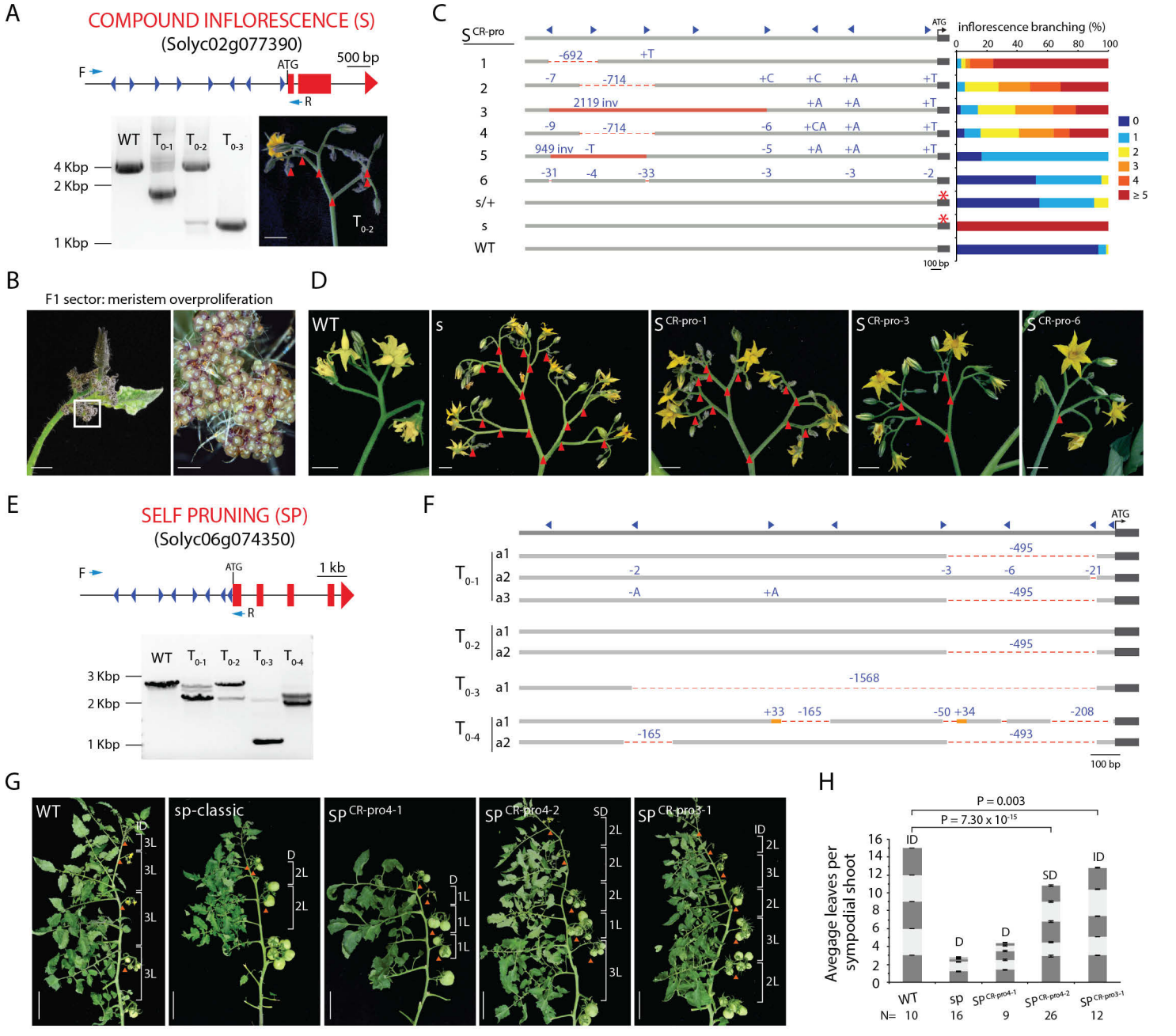

A
COMPOUND INFLORESCENCE (S)
(Solyc02g077390)
F
ATG
500 bp
R
WT
T0-1
T0-2
T0-3
4 Kbp
2 Kbp
1 Kbp
T0-2
C
S CR-pro
ATG
inflorescence branching (%)
0
20
40
60
80
100
1
2
3
4
5
6
s/+
s
WT
-692
+T
-7
-714
+C
+C
+A
+T
2119 inv
+A
+A
+T
-9
-714
-6
+CA
+A
+T
949 inv
-T
-5
+A
+A
+T
-31
-4
-33
-3
-3
-2
100 bp
0
1
2
3
4
≥ 5
B
F1 sector: meristem overproliferation
D
WT
s
S CR-pro-1
S CR-pro-3
S CR-pro-6
E
SELF PRUNING (SP)
(Solyc06g074350)
F
ATG
1 kb
R
WT
T0-1
T0-2
T0-3
T0-4
3 Kbp
2 Kbp
1 Kbp
F
ATG
T0-1
a1
a2
a3
-495
-2
-3
-6
-21
-A
+A
-495
T0-2
a1
a2
-495
T0-3
a1
-1568
T0-4
a1
a2
+33
-165
-50
+34
-208
-165
-493
100 bp
G
WT
sp-classic
SP CR-pro4-1
SP CR-pro4-2
SP CR-pro3-1
ID
D
SD
3L
2L
1L
H
P = 0.003
P = 7.30 x 10-15
Avegage leaves per sympodial shoot
0
2
4
6
8
10
12
14
16
ID
D
D
SD
ID
WT
sp
SP CR-pro4-1
SP CR-pro4-2
SP CR-pro3-1
N=
10
16
9
26
12

Figure S1

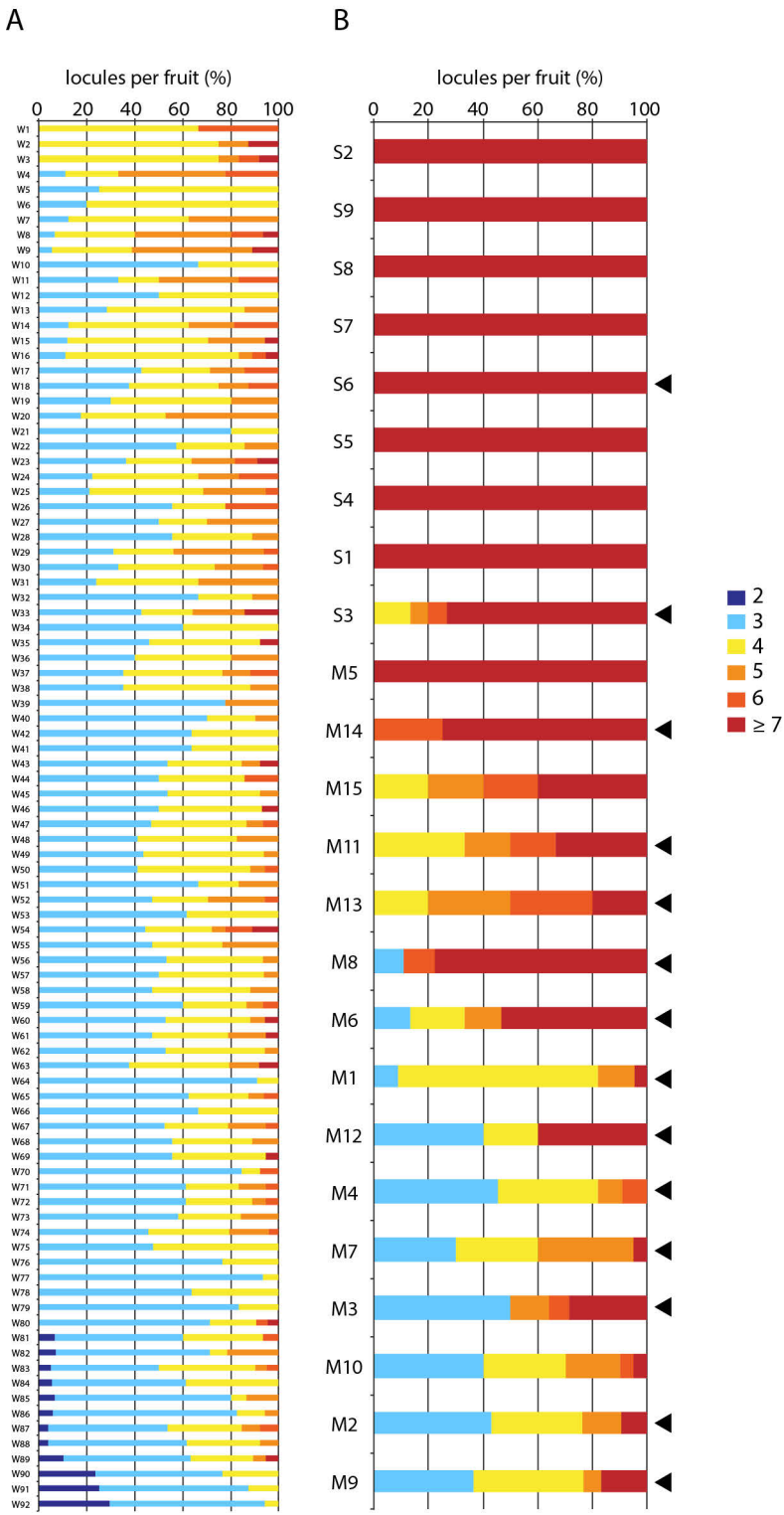
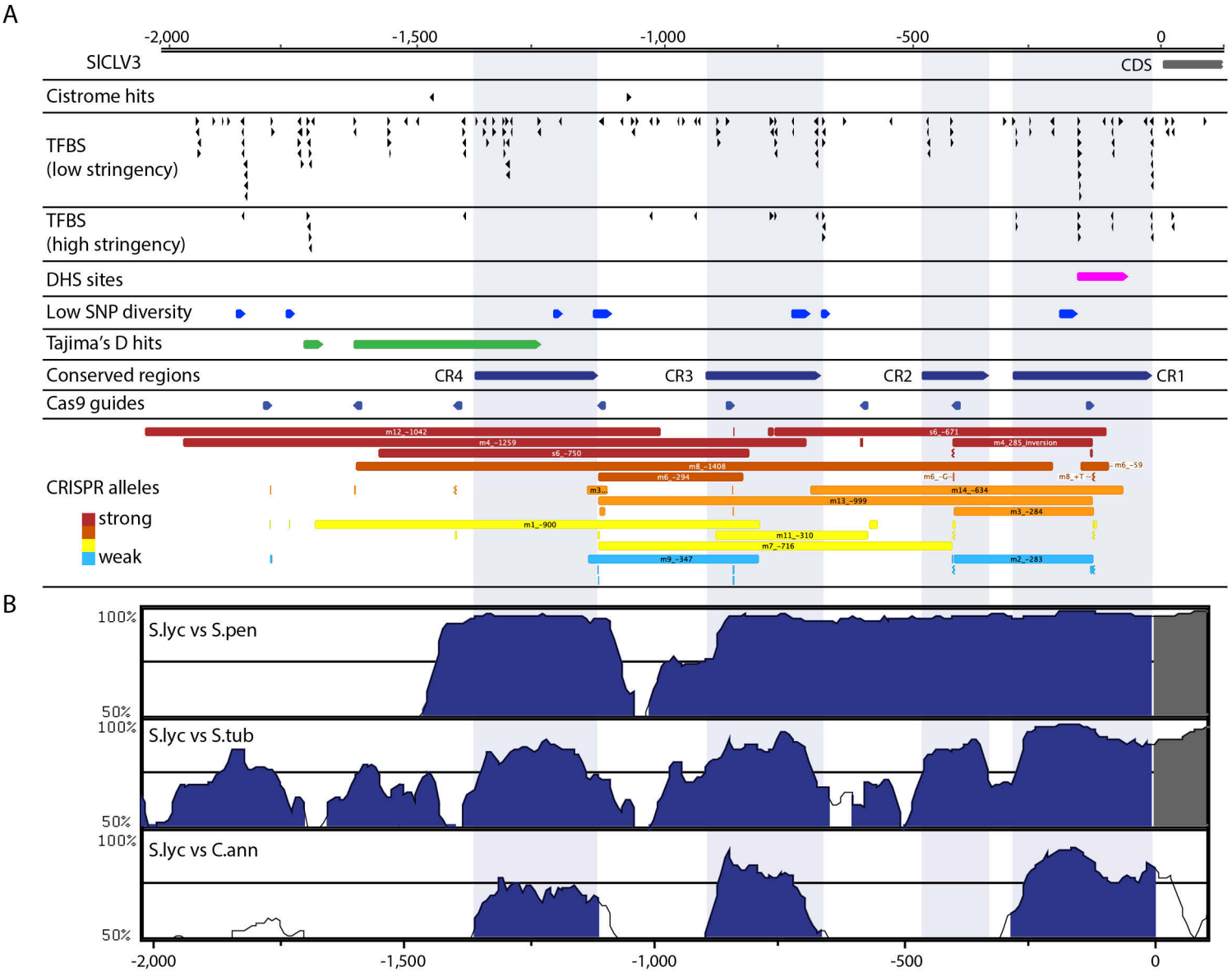

A
B
locules per fruit (%)
0
20
40
60
80
100
S2
S9
S8
S7
S6
S5
S4
S1
S3
M5
M14
M15
M11
M13
M8
M6
M1
M12
M4
M7
M3
M10
M2
M9
2
3
4
5
6
≥ 7

Figure S2


A
-2,000
-1,500
-1,000
-500
0
SlCLV3
CDS
Cistrome hits
TFBS
(low stringency)
TFBS
(high stringency)
DHS sites
Low SNP diversity
Tajima's D hits
Conserved regions
CR4
CR3
CR2
CR1
Cas9 guides
CRISPR alleles
strong
weak
B
100%
50%
S.lyc vs S.pen
S.lyc vs S.tub
S.lyc vs C.ann
-2,000
-1,500
-1,000
-500
0