nature biotechnology

Domestication of wild tomato is accelerated by genome editing

Tingdong Li[1,2,5], Xinping Yang[3–5], Yuan Yu[2–5], Xiaomin Si[1,2], Xiawan Zhai[3,4], Huawei Zhang[1], Wenxia Dong[3,4], Caixia Gao[1,2] & Cao Xu[2–4]

Crop improvement by inbreeding often results in fitness penalties and loss of genetic diversity. We introduced desirable traits into four stress-tolerant wild-tomato accessions by using multiplex CRISPR–Cas9 editing of coding sequences, *cis*-regulatory regions or upstream open reading frames of genes associated with morphology, flower and fruit production, and ascorbic acid synthesis. Cas9-free progeny of edited plants had domesticated phenotypes yet retained parental disease resistance and salt tolerance.

Conventional breeding to domesticate wild plants increases productivity but is often accompanied by decreased fitness and genetic diversity, thus hampering growth in challenging environmental conditions[1]. Introgression of stress-tolerance genes from wild relatives into cultivated species has been used to improve crops, but conventional breeding is slow[2]. *De novo* domestication of wild species has been proposed as an alternative strategy[2,3].

Tomato (*Solanum lycopersicum*; *SI*) is a high-value crop with numerous extant wild relatives. The ancestral line *Solanum pimpinellifolium* is remarkably stress tolerant and can thrive in the Ecuadorian tropics and the Peruvian desert[4]. Furthermore, *S. pimpinellifolium* accession LA1589 is highly resistant to bacterial spot disease caused by race T3 strains (predominately *Xanthomonas perforans*), which can decrease yields of commercial tomato crops by up to 60% (ref. 5), whereas accession LA1357 is salt tolerant[6]. Domestication of wild tomato species for commercial cultivation would require numerous phenotypes to be changed, including flowering and day-length sensitivity, fruit setting and size, ripening synchrony and nutrient content. Fortunately, many domestication phenotypes are monogenic[7–9] and are amenable to CRISPR–Cas9 editing of genes or regulatory regions[10,11]. In short, *de novo* crop domestication is now technically feasible. Here, we report our efforts to domesticate *S. pimpinellifolium*. We used a multiplex CRISPR–Cas9 strategy to edit genes related to day-length sensitivity, shoot architecture, flower and fruit production, and nutrient content to accelerate domestication of *S. pimpinellifolium* (**Fig. 1a** and **Supplementary Table 1**).

Loss of the flowering repressor SELF-PRUNING 5G (SP5G, a member of the CETS protein family) in tomato confers day neutrality and thus could potentially extend the geographical cultivation range[12]. Mutations in the *SP* (*SELF PRUNING*) gene, another CETS family flowering repressor, change tomato indeterminate shoot architecture into determinate vines. This change enables high-density growth and mechanical harvesting, because *sp* variants are compact plants with intense inflorescence and almost synchronous fruit ripening[12,13]. Moreover, simultaneous mutation of *SP5G* and *SP* in tomato varieties can produce 'double determinate' plants that flower early on both primary and sympodial shoots, and permit early harvest[12]. Precise editing of *SP* and *SP5G* in wild tomato species might serve as a first step toward generating commercially cultivable lines. To that end, we designed one guide RNA (gRNA) for the first exon of each gene to induce mutations in *SP* and *SP5G* through genome editing (**Supplementary Fig. 1a,b**).

Fruit-size enlargement due to increased numbers of seed compartments (locules) is a crucial feature of domesticated tomato[14], but *S. pimpinellifolium* has only tiny bilocular fruits. Two quantitative trait loci, *fasciated* (*fas*)[15] and *locule number* (*lc*)[16], are responsible for increasing locule number and fruit size, and are correlated with mutations in the small-peptide-encoding gene *CLV3* (*CLAVATA3*) and the homeobox-encoding gene *WUS* (*WUSCHEL*), respectively. *CLV3* null mutations in tomato and *WUS* loss-of-function mutations in *Arabidopsis* result in developmental defects that have limited their applications in crop improvement[16,17]. However, naturally occurring alleles or those edited with CRISPR–Cas9 in their *cis*-regulatory regions increase locule number and result in fewer developmental defects in tomato[17,18]; moreover, mutations in the *SlCLV3* promoter region adjacent to the translational start site (ATG) often produce strong fasciation that largely mimics that in null mutants[11]. To create weak alleles that increase fruit locule number, we designed two gRNAs spanning the *SlCLV3* promoter from 1.2 kb to 1.8 kb upstream of the ATG (**Supplementary Fig. 1c**). To recreate the effect of *lc*, we designed one gRNA within the predicted 15-bp CArG transcription-repressor element downstream of *SlWUS*[18] and another gRNA targeting the nearby flanking sequence of the CArG element (**Supplementary Fig. 1d**).

To edit multiple domestication genes simultaneously and stack the resulting allelic variants, we assembled the set of six gRNAs to edit four genes into one construct, by using the Csy4 multi-gRNA system[9]. We then transformed the construct pDIRECT_22C_6gR_A into four wild-tomato accessions, all of which are resistant to bacterial spot disease, and two of which (LA1357 and LA1547) are salt tolerant (**Supplementary Figs. 2–4**). We generated 140 independent first-generation transgenic lines (T0). Genotyping revealed that all

[1]State Key Laboratory of Plant Cell and Chromosome Engineering, Center for Genome Editing, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China. [2]University of Chinese Academy of Sciences, Beijing, China. [3]State Key Laboratory of Plant Genomics, National Center for Plant Gene Research (Beijing), Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China. [4]CAS-JIC Centre of Excellence for Plant and Microbial Science (CEPAMS), Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China. [5]These authors contributed equally to this work. Correspondence should be addressed to C.X. (caoxu@genetics.ac.cn) or C.G. (cxgao@genetics.ac.cn).

Received 20 July; accepted 12 September; published online 1 October 2018; doi:10.1038/nbt.4273

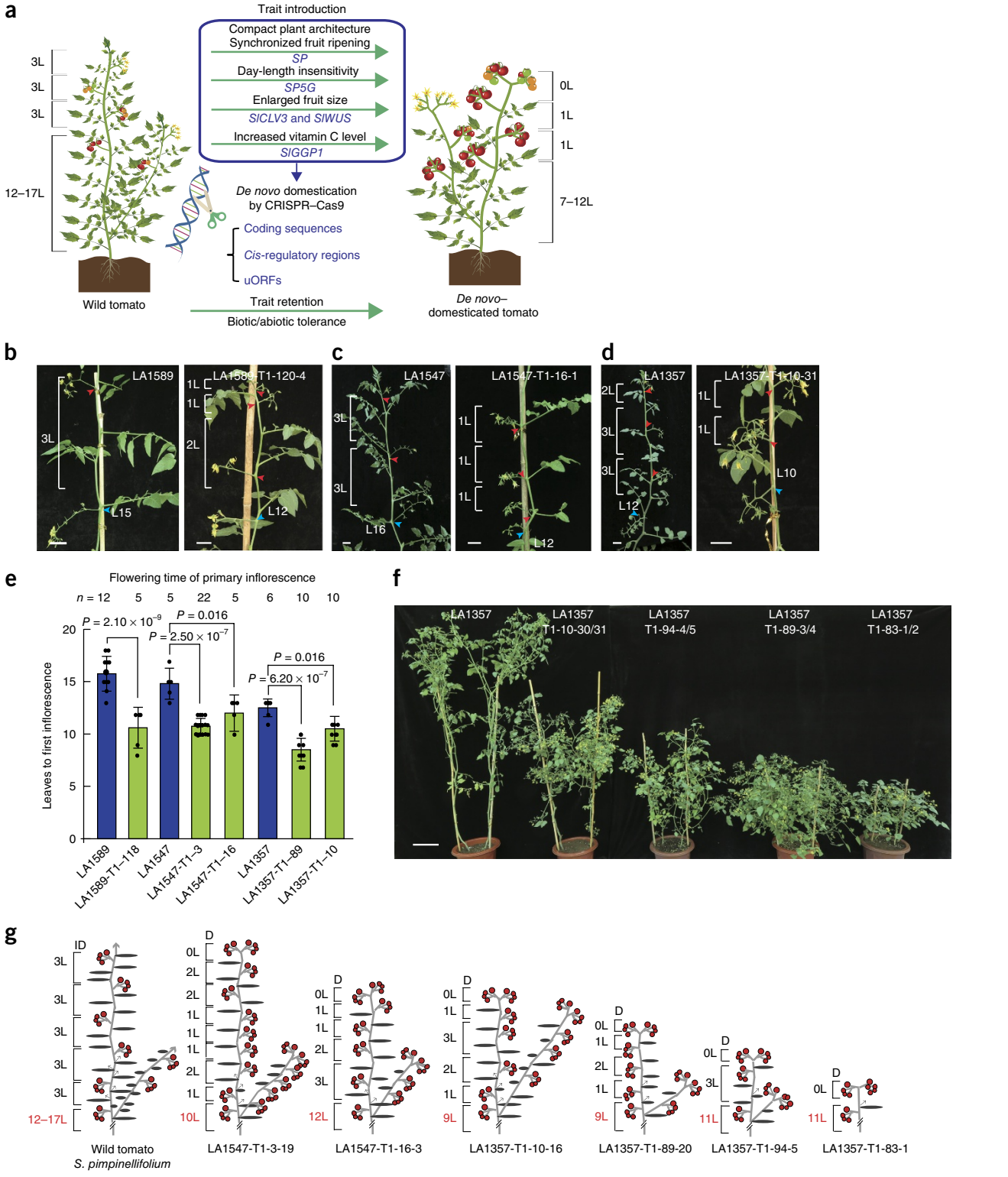

Figure 1 Genome editing of coding and regulatory regions of *SP*, *SP5G*, *SlCLV3*, *SlWUS* and phenotypes of T1 progeny. (**a**) Schematic of accelerating domestication of wild tomatoes through CRISPR–Cas9 technology. L, leaf. (**b**–**d**) Representative vines of LA1589 (**b**), LA1547 (**c**), LA1357 (**d**) and the corresponding T1 CRISPR mutant plants. At least five plants from each wild accession and T1 CRISPR mutants with identical mutation forms of SP and SP5G were observed and are represented. Brackets and numbers indicate leaf number within a sympodial cycle. Red and blue arrowheads indicate sympodial and primary inflorescences, respectively. L, leaf. Scale bars, 2 cm. (**e**) Quantification of flowering time of wild-type and T1 CRISPR plants. Tukey *post hoc* tests for multiple comparisons were performed. Data are means ± s.d.; the number of plants per line (*n*) is indicated above each bar. (**f**) Representative whole plants showing compact plant architecture and early flowering of T1 Cas9-free progeny with mutation of *SP* and *SP5G*. Scale bars, 10 cm. At least six plants from LA1357 and each CRISPR T1 mutant line were observed and are represented. (**g**) Schematic showing a continuum of plant-architecture and flower-production traits created by stacking *sp* and *sp5g* allelic variants in T1 CRISPR plants. L, leaf; red color highlights leaf number for the primary inflorescence; ID, indeterminate; D, determinate. Mutation allele information of CRISPR lines for **b**–**g** is in listed **Supplementary Figure 1** and **Supplementary Table 2**.

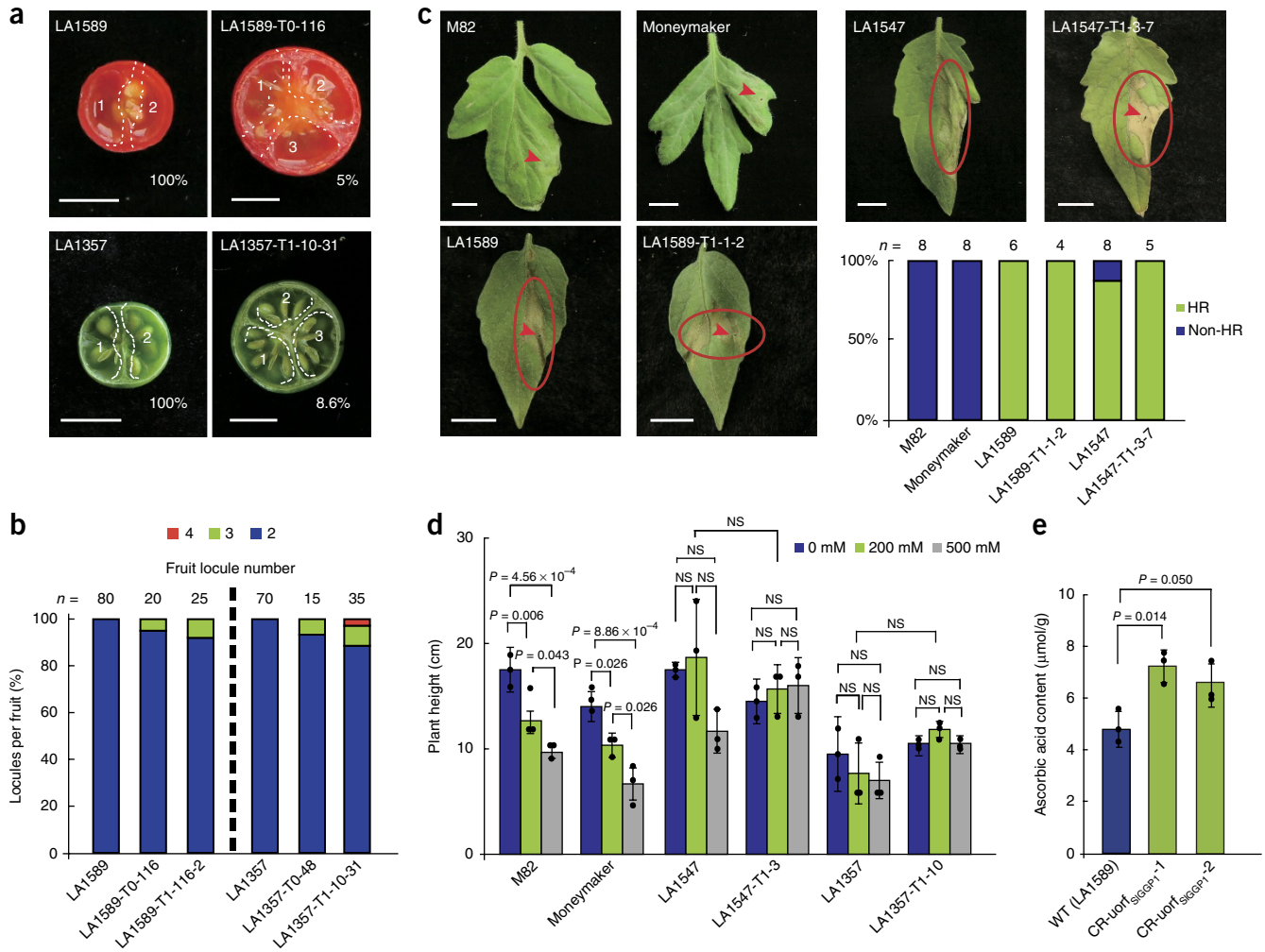

Figure 2 Improvement of fruit size, nutrient content and retention of stress tolerance in CRISPR plants. (**a**,**b**) Representative images (**a**) and quantification (**b**) of fruit locules of T0 and T1 CRISPR plants. The percentages of fruit locule numbers are indicated in **a**. The fruits of T0-48 and T0-116 were observed with three locules; the fruits of T1-10 and T1-116 were observed with three or four locules. (**c**) HR in wild accessions and the corresponding Cas9-free CRISPR T1 plants after bacterial inoculation. Red arrowheads indicate inoculation sites, and red circles indicate HR area. The bar graphs show percentages of plants with HR. The number of plants used for the HR test is indicated above each bar. Scale bars, 0.5 cm. (**d**) Comparison of plant height between wild type and the corresponding Cas9-free CRISPR T1 plants under salt treatment. (**e**) Foliar ascorbic acid content of LA1589 and the T0 CRISPR mutants. One-way analysis of variance (ANOVA) (**d**) and Tukey *post hoc* tests (**d**,**e**) for multiple comparisons were performed. Data are means ± s.d., $n = 3$ individual plants. NS, nonsignificant difference; WT, wild type. Mutation information of selected CRISPR lines for **a**–**e** is provided in **Supplementary Figures 1** and **8**.

six gRNAs produced a variety of mutations in target genes. We categorized the 140 lines of T0 plants into the following four groups on the basis of the mutation efficiency of four target genes: at least one gene mutated (80 lines); at least two genes mutated (77 lines); at least three genes mutated (68 lines); and four genes mutated simultaneously (45 lines) (**Supplementary Table 2**). To evaluate the domestication outcomes in these plants, we examined the flowering time of primary inflorescences in T1 progeny. In the wild species, 12–17 leaves formed before the first inflorescence, but the T1 progeny with *SP5G* and *SP* mutations flowered early by developing the first inflorescence after 7–12 leaves on primary shoots under long-day conditions (**Fig. 1b–e**), comparably to modern tomato cultivars. Sympodial cycling (alternating developed leaves and inflorescences on sympodial shoots) is an important trait determining growth density, flower and fruit production, and harvesting feasibility[12,13]. Wild tomato plants have an indeterminate vine architecture characterized by endless sympodial cycling with three leaves per inflorescence (**Fig. 1b–d**). Simultaneous mutation of *SP* and *SP5G* converted the indeterminate vine architecture of wild tomatoes into determinate growth with early termination of sympodial cycling, thus resulting in compact tomato plants with intensive and almost synchronously ripening fruits (**Fig. 1b–d,f** and **Supplementary Fig. 5**). Although *SP* and *SP5G* are crucial for improving the harvest index, the limited allelic variation has hampered efforts to optimize this trait. Sequencing suggested that we had created at least 128 and 113 mutated alleles of *SP* and *SP5G*, respectively (**Supplementary Table 3**). Stacking this allelic variation allowed us to produce a continuum of plant architectures, and flower- and fruit-production traits in just one generation (**Fig. 1f,g**, **Supplementary Fig. 1a,b** and **Supplementary Table 4**).

Next, we identified small indels and large insertions in the targeted regulatory regions of *SlCLV3* and *SlWUS* in 140 T0 and their T1 mutant plants (**Supplementary Fig. 1c,d**). Quantification of the fruit locule number of T0 and T1 plants suggested that LA1589-T0-116, T1-116-2 and LA1357-T0-48 and T1-10-31 plants show increased

fruit size, producing approximately 5–12% of fruits with three locules (**Fig. 2a,b**), similarly to the reported subtle effect of *lc* mutation in *S. pimpinellifolium*[11]. Sequencing suggested that the *SlCLV3* promoter region in LA1589-T0-116 and LA1357-T0-48 contained biallelic (1- and 2-bp deletions on each strand) and heterozygous (1-bp insertion) mutations, respectively (**Supplementary Fig. 1c**), thus suggesting that *SlCLV3* mutation has an almost negligible contribution to the increase in fruit locule number. Locule number was not increased in those T0 and T1 plants with large insertions and inversions in the targeted *SlCLV3* promoter region (**Supplementary Fig. 1c**). One explanation for this finding is that the targeted region of the *SlCLV3* promoter may not be essential for regulating *SlCLV3* transcription. Alternatively, disruption of regions (gRNA-5) flanking the CArG element downstream of *SlWUS* may have decreased its transcription and counteracted the effects of mutation of *SlCLV3*, owing to a negative feedback loop of CLV3–WUS in controlling stem cell proliferation[16]. Although we recreated the fruit locule effect of *lc*, this was possible only because of previous characterization of the CArG functional element[11,18].

To test whether T1 progeny retained disease and environmental-stress tolerance, we inoculated T1 lines with bacterial spot race T3 and examined their salt tolerance (**Supplementary Fig. 6** and **Supplementary Tables 5** and **6**). We found that all four tested Cas9-free progeny had the same strong hypersensitivity reaction (HR) to race T3 as that of LA1589 (**Fig. 2c** and **Supplementary Fig. 4**). The seedling growth of two tested lines on medium supplemented with NaCl suggested that the T1 progeny were as salt tolerant as LA1357 (**Fig. 2d**). One concern about CRISPR–Cas9 is the occurrence of off-target mutations, which was excluded by examination of potential off-target sites of six targets in transgene-free T1 plants with T7E1 assays and Sanger sequencing (**Supplementary Table 7**).

Improving the nutritional benefits of wild plants is desirable. Vitamin C content was previously increased in *Arabidopsis* and lettuce[19,20]. We designed a gRNA targeting the upstream open reading frame (uORF) of *SlGGP1* (encoding a vitamin C–biosynthetic enzyme) in accession LA1589. *SlGGP1* uORF mutations in two independent CRISPR lines were associated with increased foliar ascorbic acid content, thus demonstrating that editing of uORFs by CRISPR–Cas9 can substantially increase ascorbic acid levels in wild tomatoes (**Fig. 2e** and **Supplementary Figs. 7** and **8**).

In summary, we report that manipulation of wild tomato can begin to recapitulate breeding without causing an associated drag on other useful traits. Genome engineering could thus be applied for *de novo* domestication of wild species to create climate-smart crops.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank W. Yang and N. Li (China Agricultural University) for providing the bacterial spot race T3 strain and for assistance with inoculation, and D. Bartlem (KWS) for critical reading and editing of an earlier draft of this manuscript. We also thank Y. Wang (Institute of Genetics and Developmental Biology, CAS) for help with Figure 1a. We thank L. Yan and Y. Li for assistance with tissue culture. This work was supported by grants from the National Key Research and Development Program of China (2016YFD0101804), the National Science Foundation of China (31788103 and 31420103912) and the Chinese Academy of Sciences (QYZDY-SSW-SMC030 and GJHZ1602) to C.G., and the Thousand Talents Plan to C.X.

AUTHOR CONTRIBUTIONS

C.X. and C.G. designed the experiments; T.L., X.Y., Y.Y. and X.S. performed most of the experiments; X.Z. and W.D. generated transgenic plants. T.L., X.Y., Y.Y., X.S. and H.Z. analyzed the results; C.X. and C.G. supervised the project; C.X. and C.G. wrote the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

1. Meyer, R.S. & Purugganan, M.D. *Nat. Rev. Genet.* **14**, 840–852 (2013).
2. Østerberg, J.T. *et al. Trends Plant Sci.* **22**, 373–384 (2017).
3. Zsögön, A., Cermak, T., Voytas, D. & Peres, L.E. *Plant Sci.* **256**, 120–130 (2017).
4. Zuriaga, E. *et al. Genet. Resour. Crop Evol.* **56**, 39–51 (2008).
5. Robbins, M.D., Darrigues, A., Sim, S.C., Masud, M.A. & Francis, D.M. *Phytopathology* **99**, 1037–1044 (2009).
6. Rao, E.S. *et al. Euphytica* **190**, 215–228 (2013).
7. Doebley, J.F., Gaut, B.S. & Smith, B.D. *Cell* **127**, 1309–1321 (2006).
8. Huang, X. *et al. Nature* **537**, 629–633 (2016).
9. Lin, T. *et al. Nat. Genet.* **46**, 1220–1226 (2014).
10. Yin, K., Gao, C. & Qiu, J.L. *Nat. Plants* **3**, 17107 (2017).
11. Rodríguez-Leal, D., Lemmon, Z.H., Man, J., Bartlett, M.E. & Lippman, Z.B. *Cell* **171**, 470–480 e8 (2017).
12. Soyk, S. *et al. Nat. Genet.* **49**, 162–168 (2017).
13. Park, S.J. *et al. Nat. Genet.* **46**, 1337–1342 (2014).
14. Lippman, Z. & Tanksley, S.D. *Genetics* **158**, 413–422 (2001).
15. van der Knaap, E. *et al. Front. Plant Sci.* **5**, 227 (2014).
16. Xu, C. *et al. Nat. Genet.* **47**, 784–792 (2015).
17. Endrizzi, K., Moussian, B., Haecker, A., Levin, J.Z. & Laux, T. *Plant J.* **10**, 967–979 (1996).
18. Somssich, M., Je, B.I., Simon, R. & Jackson, D. *Development* **143**, 3238–3248 (2016).
19. Čermák, T. *et al. Plant Cell* **29**, 1196–1217 (2017).
20. Zhang, H.W. *et al. Nat. Biotechnol.* **36**, 894–898 (2018).

ONLINE METHODS

Guide RNA design and CRISPR–Cas9 constructs. Vector pDIRECT_22C (**Supplementary Fig. 6a**) was used to express the Csy4-multi-gRNA system[19] to modify several targets simultaneously. Csy4 (187 amino acids) and P2A fused with *Arabidopsis* codon-optimized SpCas9 were driven by the cauliflower mosaic virus (CaMV) *35S* promoter. Six gRNAs were separated by 20-bp Csy4-binding sites and introduced simultaneously in one step into pDIRECT_22C through the Golden Gate cloning method with a previously reported protocol[19]. gRNAs were designed to target regions of the coding sequences, the promoter, 3′ regulatory motif and uORF as follows: gRNA-1 and gRNA-2 targeting the first exon in *SP* and *SP5G*, respectively; gRNA-3 and gRNA-4 targeting the promoter region (1.2–1.8 kb upstream of ATG) of *SlCLV3*; gRNA-5 and gRNA-6 targeting the nearby flanking region and 15-bp core region of the CArG transcription-repressor element downstream of *SlWUS*. The gRNA array in construct pDIRECT_22C_6gR_A was arranged in the following order: gRNA-1–gRNA-2–gRNA-3–gRNA-4–gRNA-5–gRNA-6 (**Supplementary Table 1**). The gRNA array was driven by the *Cestrum* yellow leaf curling virus (*CmYLCV*) promoter.

To induce mutations in the uORF of *SlGGP1*, the construct pKSE401-*SlGGP1uORF*-gRNA-7 was generated as previously described[21]. gRNA-7 was introduced into pKSE401 by BsaI digestion. To produce the constructs for the dual-luciferase reporter assays, the *35S* promoter followed by wild-type and mutated forms of the 5′ leader sequence of *SlGGP1* were cloned into the pGreenII0800-LUC vector with a ClonExpress II One Step Cloning Kit (Vazyme Biotech), thus yielding the constructs $uORF_{SlGGP1}$, $uorf_{SlGGP1}$-1(-3), $uorf_{SlGGP1}$-2(-6) and $uorf_{SlGGP1}$-2(-12).

Gene sequences of *S. lycopersicum* and *S. pimpinellifolium* were obtained from the Sol Genomics Network (SGN) database (https://solgenomics.net/). The relative sequences of the five target genes of accession LA1589 used in this study are listed in **Supplementary Sequences.** All the primers used in this study are listed in **Supplementary Table 8**.

Plant materials and growth conditions. Seeds of *S. pimpinellifolium* accessions LA1589, LA1547, LA1357, LA1606, M82 and Moneymaker were obtained from the Tomato Genetics Resource Center at the University of California, Davis (http://tgrc.ucdavis.edu/). The seeds were directly sown in soil in 72-cell plastic flats covered with vermiculite under standard greenhouse conditions. Four-week-old seedlings were transplanted into 6-L pots (for phenotypic observation and harvesting) or 8 × 8 × 8 cm^3 (width × length × height) pots (for bacterial inoculation). Analyses of shoot architecture and flowering time were conducted on plants grown in a greenhouse under natural light with supplementation from high-pressure sodium bulbs (50 $mM/m^2/s$) on a 16-h light/8-h dark photoperiod. The day and night temperatures were 26–30 °C and 18–20 °C, respectively.

Generation and genotyping of the CRISPR plants. Transformations mediated by *Agrobacterium tumefaciens* strain LBA4404 were performed as previously described[22]. Four *S. pimpinellifolium* accessions (LA1589, LA1357, LA1547 and LA1606) were used as recipients for the transformations. Transplantation and acclimation of tissue-culture-generated CRISPR-seedlings were performed as previously described[23].

For genotyping of the T0 transgenic lines, three different leaf samples were collected from each plant to capture all possible induced mutations due to sectoring, and genomic DNA was extracted with the cetyltrimethylammonium bromide (CTAB) method. T7E1, PCR restriction-enzyme digestion (PCR–RE) assays and Sanger sequencing were performed to identify mutations in target genes of T0 and T1 CRISPR plants, as described previously[24].

Salt-stress treatment. Salt treatments were performed as previously reported with some modifications[6]. Briefly, treatments of wild accessions and derived Cas9-free T1 CRISPR plants were divided into three groups on the basis of NaCl concentration (0 mM, 200 mM and 500 mM) supplied in the watering solution. The NaCl solution for watering was prepared by dissolving NaCl in double-distilled water. Four-week-old seedlings with four to five true leaves were grown in 8 × 8 × 8 cm (width × length × height) pots filled with potting soil. Nine plants for each genotype were used for salt treatment under different NaCl concentrations with three biological replicates per concentration. The pots with different plants were randomly placed to avoid possible location effects. For the treatment, each plant was watered with 10 mL of the appropriate NaCl solution every 2 d and grown under standard greenhouse conditions. Plant height was measured and calculated after 14 d of salt treatment. No fertilizer was applied to avoid interference with the salt treatment.

Inoculation of bacteria. The *X. perforans* race T3 strain Xv829 used in this study was provided by W. Yang from China Agricultural University. The bacteria were grown on yeast, dextrose and calcium carbonate (YDC) agar medium at 28 °C for 48–72 h. A bacterial suspension was prepared by washing the bacteria with sterilized double-distilled water and adjusting the concentration to approximately 3×10^8 CFU/ml (A_{600} = 0.20). Plants with four to five fully expanded true leaves were used for inoculation. The inoculation was performed as previously described[25], by infiltration of the bacterial suspension with a 5-ml syringe without a needle through the abaxial side of leaflets, to a diameter of approximately 1.5 cm. Four to eight plants per biological replicate and three leaves per plant were subjected to inoculation (n = 4–8). The number of individuals (n) is presented in the figures. The inoculated plants were then kept at 22–28 °C in the greenhouse and misted with water regularly to maintain humidity.

Transient expression in *S. pimpinellifolium* protoplasts. *S. pimpinellifolium* protoplasts were isolated from well-expanded leaves from 14- to 20-d-old plants grown on MS medium. The protoplast isolation and transient expression were performed as previously reported[16]. 15–20 μg of plasmid and approximately 5×10^5 protoplast cells were used for each PEG-mediated transfection. The cells were harvested 48 h after transfection for further analysis. For dual luciferase reporter assays, luciferase/*Renilla* luciferase activity was measured with a Dual-Luciferase Reporter Assay System (Promega).

Preparation of RNA and qRT–PCR. Total RNA was extracted with TRIzol reagent (Life Technologies) from transfected *S. pimpinellifolium* protoplasts. M-MLV Reverse Transcriptase (Promega) was used for reverse transcription. qRT–PCR was performed with a SsoFast EvaGreen Supermix kit (Bio-Rad) according to the instruction manual.

Measurement of ascorbic acid content. Ascorbic acid concentrations were measured through high-performance liquid chromatography as previously described, with minor modifications[20]. Briefly, tomato leaves were ground to a powder in liquid nitrogen, then vortexed for 30 s in extraction buffer (74.45 mg EDTA, 286.65 mg TCEP and 5 ml of 98% orthophosphoric acid). After incubation on ice for 10 min, the samples were centrifuged at 12,000*g* at 4 °C for 30 min. Ascorbic acid concentrations were measured with a Pursuit XRs C18 A2000250X046 column (Agilent) with ultraviolet-light (244 nm) detection.

Detection of potential off-target mutations. Potential off-target sites in the *S. lycopersicum* genome were predicted with the online tool CRISPR-P v2.0 (ref. 26). For each target, the top six potential off-target sites were examined in four representative Cas9-free T1 plants (T1-1-4, T1-3-28, T1-10-3 and T1-89-2). Amplicons from the potential off-target sites were examined with T7E1 assays and Sanger sequencing (**Supplementary Table 7**).

Statistical analysis. For quantitative analysis of flowering time, luciferase/*Renilla* luciferase activity, relative transcription level and ascorbic acid content, at least three individual plants per genotype or three biological replicates for each experiment were used for statistical analysis. The statistical analysis was performed with Tukey *post hoc* tests for multiple comparisons. In all cases, the alpha level for significant difference was 0.05. Statistical comparisons were made with commercially available software SPSS (IBM SPSS Statistics 20). One-way ANOVA was performed to test the group-to-group differences between LA1547 and LA1357 and their corresponding CRISPR–Cas9 mutants under different salt concentrations in **Figure 2d**. All numerical values are presented as means ± s.d.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability. All data supporting the findings of this study are available in the article or its Supplementary Information files, or are available

doi:10.1038/nbt.4273

from the corresponding author upon reasonable request. Sequence data of *S. pimpinellifolium* (LA1589) were obtained from the draft genome sequence in the Sol Genomics Network (SGN) database (https://solgenomics.net/). The accession numbers of the relevant *S. pimpinellifolium* genes are as follows: *SP*, Solyc06g074350; *SP5G*, Solyc05g053850; *SlCLV3*, Solyc11g071380; *SlWUS*, Solyc02g083950; and *SlGGP1*, Solyc02g091510.

21. Xing, H.L. *et al. BMC Plant Biol.* **14**, 327 (2014).
22. Van Eck, J., Kirk, D.D. & Walmsley, A.M. *Methods Mol. Biol.* **343**, 459–473 (2006).
23. Brooks, C., Nekrasov, V., Lippman, Z.B. & Van Eck, J. *Plant Physiol.* **166**, 1292–1297 (2014).
24. Shan, Q., Wang, Y., Li, J. & Gao, C. *Nat. Protoc.* **9**, 2395–2410 (2014).
25. Yang, W. & Francis, D.M. *J. Am. Soc. Hortic. Sci.* **130**, 716–721 (2005).
26. Liu, H. *et al. Mol. Plant* **10**, 530–532 (2017).

nature research

Corresponding author(s): Caixia Gao, Cao Xu

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Sanger sequencing was used to analyze the mutations of the target genes; DNAMAN 6.0 was used to to align and analyze the sequences.
Data analysis	IBM SPSS Statistics 20 was used to analyze the data; all the genes sequence data were obtained from Sol Genomics Network (SGN) (https://solgenomics.net) database; CRISPR-P v2.0 (http://cbi.hzau.edu.cn/CRISPR2/) was was used to predict the potential off-target sites.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data supporting the findings of this study are available in the article or in Supplementary Information files, or are available from the corresponding author upon

Nature Biotechnology: doi:10.1038/nbt.4273

request. Sequence data of S. pimpinellifolium (LA1589) was obtained from the draft genome sequence on Sol Genomics Network (SGN) database (https://solgenomics.net/). The accession numbers of relevant S. pimpinellifolium genes is: SP (Solyc06g074350), SP5G (Solyc05g053850), SlCLV3 (Solyc11g071380), SlWUS (Solyc02g083950), SlGGP1 (Solyc02g091510).

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. Experiments were performed three times independently unless indicated. In previous studies using related experiments, the sample size has been determined to be sufficient to ensure reproducibility.
Data exclusions	No data exclusion.
Replication	All attempts for replication were successful. A minimum of 3 biological replicates were included.
Randomization	Samples were randomly allocated into groups.
Blinding	Not applicable, as samples were processed identically through standard and in some cases automated procedures (DNA sequencing, transfection, DNA isolation) that should not have bias outcomes.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Nature Biotechnology: doi:10.1038/nbt.4273