

Redesigning the tomato fruit shape for mechanized production

Received: 5 April 2023

Accepted: 23 August 2023

Published online: 18 September 2023

 Check for updates

Qiang Zhu , Lei Deng , Jie Chen , Gustavo R. Rodríguez , Chuanlong Sun , Zeqian Chang , Tianxia Yang , Huawei Zhai , Hongling Jiang , Yasin Topcu , David Francis , Samuel Hutton , Liang Sun , Chang-Bao Li , Esther van der Knaap & Chuanyou Li ✉

Crop breeding for mechanized harvesting has driven modern agriculture. In tomato, machine harvesting for industrial processing varieties became the norm in the 1970s. However, fresh-market varieties whose fruits are suitable for mechanical harvesting are difficult to breed because of associated reduction in flavour and nutritional qualities. Here we report the cloning and functional characterization of *fs8.1*, which controls the elongated fruit shape and crush resistance of machine-harvestable processing tomatoes. *FS8.1* encodes a non-canonical GT-2 factor that activates the expression of cell-cycle inhibitor genes through the formation of a transcriptional module with the canonical GT-2 factor SIGT-16. The *fs8.1* mutation results in a lower inhibitory effect on the cell proliferation of the ovary wall, leading to elongated fruits with enhanced compression resistance. Our study provides a potential route for introducing the beneficial allele into fresh-market tomatoes without reducing quality, thereby facilitating mechanical harvesting.

The use of mechanized machines in crop harvesting significantly saves labour and time costs, and puts forward new requirements for breeding. In tomato (*Solanum lycopersicum*), one of the most consumed vegetable and fruit crops worldwide¹, fruit shape is an important horticultural trait that determines the main market class and culinary purpose of a particular variety: elongated and blocky fruits are preferred for processing, while round and attractive fruits are ideal for the fresh market and are used for slicing^{2,3}. Notably, while machine harvesting for elongate-fruited processing tomatoes became the norm in the 1970s, typical round-fruited fresh-market tomatoes are usually too soft to endure machine harvesting^{4–6},

suggesting an association between fruit shape and suitability for machine harvesting.

In tomato, two categories of quantitative trait loci (QTLs) that control fruit shape have been identified: *locule number* (*lc*) and *fasciated* (*fas*) QTLs control locule number and flat shape, while *sun*, *ovate* and *fruit shape chr 8.1* (*fs8.1*) control elongated shape^{7–10}. In the early 1960s, the processing-tomato industry in California was threatened due to labour shortage in harvesting the crops. In response, efforts began in earnest to develop tomato varieties suited to machine harvesting. Selection of a tomato suitable for mechanical harvesting involved a suite of traits including compact plant stature, concentrated fruit set

¹State Key Laboratory of Plant Genomics, National Center for Plant Gene Research (Beijing), Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China. ²CAS Center for Excellence in Biotic Interactions, University of Chinese Academy of Sciences, Beijing, China.

³College of Life Sciences, Shandong Agricultural University, Tai'an, China. ⁴College of Horticulture, China Agricultural University, Beijing, China. ⁵Instituto de Investigaciones en Ciencias Agrarias de Rosario (IICAR-CONICET-UNR), Rosario, Argentina. ⁶College of Horticulture Science and Engineering, Shandong Agricultural University, Tai'an, China. ⁷Institute of Plant Breeding, Department of Horticulture, University of Georgia, Athens, GA, USA.

⁸Bati Akdeniz Agricultural Research Institute, Antalya, Turkey. ⁹Department of Horticulture and Crop Science, The Ohio State University, Columbus, OH, USA. ¹⁰Gulf Coast Research and Education Center, University of Florida, Gainesville, FL, USA. ¹¹Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (North China), Ministry of Agriculture, Beijing Vegetable Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China. ¹²These authors contributed equally: Qiang Zhu, Lei Deng and Jie Chen. ✉e-mail: cyl@genetics.ac.cn

and, perhaps most importantly, enhanced fruit firmness to withstand the rigours of machine harvesting and bulk handling. Thus, fruit firmness was intentionally selected to avoid fruit breakage during mechanized production. These efforts eventually led to the so-called ‘square tomatoes’ that could withstand machine harvesting^{4–6}. Thus, the breeding of ‘square tomatoes’ has played a significant part in streamlining the mechanization of the processing-tomato industry^{4,5}.

One of the noticeable changes in fruit shape of these machine-harvestable processing tomatoes is that they are no longer round but rather slightly elongated^{5,11}. Later genetic studies revealed that the elongated fruit shape is controlled by the major QTL *fs8.1* (refs. 7,12–15). However, the gene underlying *fs8.1* has not been identified due to reduced recombination rates around this locus^{13,14}.

Here we identified the *FS8.1* gene and report on its mechanistic function in regulating fruit shape. *FS8.1* encodes a GT-2-like protein that lacks the featured trihelix DNA-binding domain^{16,17}. *FS8.1* interacts with and enhances the transcriptional activity of SIGHT-16, a canonical GT-2 factor that activates the expression of cell-cycle inhibitor genes through promoter binding. Notably, *fs8.1*-mediated shape change leads to enhanced fruit resistance to squeezing force. Distribution analyses of a diverse collection of tomato accessions indicated that the *fs8.1* mutation allele is present in modern processing lines but largely absent in fresh-market varieties. The cloning of *FS8.1* and the CRISPR/Cas9 gene-editing technology probably lead to the possibility of a convenient trait-stacking strategy to breed new varieties that combine the advantages of fresh-market tomatoes and processing tomatoes.

Results

fs8.1-mediated shape change leads to enhanced fruit compression resistance (CR)

Compared with fresh-market tomato plants that produce round fruits, machine-harvestable processing-tomato fruits are usually elongated in shape and have a higher fruit shape index (the ratio of fruit length to width) (Extended Data Fig. 1a,b). To determine whether the elongated fruit shape is associated with increased fruit endurance to rigours, we used a texture analyser to assess CR, which is defined as the pressure at which point a compressed fruit is broken¹⁸, reflecting the fruit’s resistance to squeezing force. We compared the CR of multiple typical fresh-market and processing-tomato fruits and found that the processing-tomato fruits had a significantly increased resistance to squeezing (Extended Data Fig. 1c).

We then aimed to understand whether *fs8.1*-mediated shape change contributes to the increased CR of processing-tomato fruits. To achieve this goal, we first phenotyped the well-established introgression line (IL) population derived from a cross between the wild species *Solanum pennellii* (whose fruits are round) and the processing cultivar M82 (whose fruits are square shaped), which has been shown to carry a mutant allele of *fs8.1* (refs. 15,19,20). We found that IL8-1-1, harbouring a marker-defined segment of *S. pennellii* chromosome 8, bears round fruits and ovaries (Fig. 1a–c and Extended Data Fig. 1d), indicating that IL8-1-1 carries the wild-type (WT) allele of the *FS8.1* gene. Thus, IL8-1-1 and its recurrent parent M82 were used as near-isogenic lines (NILs) with respect to the *fs8.1* locus. As expected, compared with IL8-1-1, M82 had a higher fruit/ovary shape index (Fig. 1b,c) and enhanced CR (Fig. 1d,e). Consistent with this observation, we analysed the *fs8.1* NILs developed between the round-fruited wild *S. pimpinellifolium* accession LA1589 and Rio Grande, a cultivated *S. lycopersicum* accession bearing elongated fruits (Extended Data Fig. 1e,f)¹⁴. Results showed that compared with the WT allele, the *fs8.1* mutant allele had higher fruit shape index and afforded increased CR (Extended Data Fig. 1g–i). Collectively, these results indicated that the *fs8.1*-mediated shape change is probably associated with enhanced CR of tomato fruits.

To gain insight into the cellular mechanism by which *fs8.1* changes fruit shape, we compared the cellular parameters of the M82 ovaries with those of the IL8-1-1 ovaries at the tissue level (Fig. 1a,c). The ovary

wall length and columella length of M82 were significantly larger than those of IL8-1-1 (Fig. 1f). Closer histological observations revealed that the *fs8.1* mutation led to significantly increased cell number in the proximal–distal direction of both the ovary wall and the columella (Fig. 1g). In contrast, *fs8.1* exerted a negligible effect on the cell size of both tissue layers (Fig. 1h). Notably, the *fs8.1*-mediated increase in cell number was significantly greater in the ovary wall than in the columella (Fig. 1f,g), suggesting that the *fs8.1* mutation results in a stronger promoting effect on cell proliferation of the ovary wall than of the columella.

FS8.1 encodes a non-canonical GT-2 factor

We previously mapped the *fs8.1* locus to a 3.03-Mb interval on chromosome 8 (Fig. 2a)¹⁴. On the basis of the overlap between the *fs8.1* mapping interval and the introgressed fragment defined by IL8-1-1 (refs. 19,20), the *fs8.1* locus was narrowed to a region flanked by markers 11EP239 and CT92 (Fig. 2a).

To further fine-map the locus, we screened approximately 14,000 seedlings in a population that carried a chromosome 8 segment of LA1589 and Rio Grande¹⁴. From this screen, we identified five recombinant plants in the interval (Fig. 2b). Progeny testing of these five recombinants showed that the locus was mapped to a ~270-kb interval, carrying three genes (Fig. 2b). Among them, *Solyc08g061910* carried two non-synonymous variations (T52C and A857T) in Rio Grande compared with LA1589 (Fig. 2b). The coding regions of the other two genes did not carry consequential single nucleotide polymorphisms (SNPs) between the alleles and were not differentially expressed¹⁴, rendering them less likely to encode FS8.1. Thus, *Solyc08g061910* was considered to be the candidate gene encoding FS8.1.

The *Solyc08g061910^{A857T}* mutation was predicted to result in a premature stop codon of the predicted protein (Fig. 2b–d). Sequence analyses indicated that this premature stop codon mutation is present in all of the analysed processing tomatoes but not in fresh-market tomatoes (Extended Data Fig. 1j). To verify that *Solyc08g061910* is the causative gene for the *fs8.1* phenotype, we generated null alleles of this gene (Extended Data Fig. 2a) in the IL8-1-1 background by CRISPR/Cas9-mediated gene editing^{21–23}. Fruits and ovaries of two of the resulting mutants displayed a shape change from round to elongated (Fig. 2e–g) and had higher CR than IL8-1-1 fruits (Fig. 2h). Histological observations indicated that these *fs8.1* mutants displayed an increased ovary wall length and decreased ovary wall thickness as compared with WT (Fig. 1a and Extended Data Fig. 2b–f). The fruit weight of edited mutants was comparable to that of WT (Extended Data Fig. 2g). In addition, the WT *Solyc08g061910* genomic DNA with its native promoter was introduced into the M82 background (Extended Data Fig. 2h). Two of the resulting complementation lines (*Comp-1* and *Comp-2*) rescued the M82 phenotype in both fruit shape and ovary development (Fig. 2i–k and Extended Data Fig. 2i–m). Collectively, these results demonstrated that *Solyc08g061910* underlies the *fs8.1* locus.

The predicted FS8.1 protein is homologous to members of the GT-2 clade of trihelix transcription factors (Extended Data Fig. 3a,b). The members of this clade of GT factors are characterized by the presence of duplicate trihelical DNA-binding domains (TD1 and TD2) and a middle α-helix domain (HD) that is involved in protein–protein interactions^{16,17}. Compared with SIGHT-16, a canonical GT-2 factor that is most closely related to FS8.1, the FS8.1 protein contains a middle HD but lacks the featured TDs that are important for DNA binding (Fig. 2d and Extended Data Fig. 3a–c). Sequence comparison between IL8-1-1 and M82 revealed that the *fs8.1* mutation resulted in a truncated protein with an incomplete α-helical domain (Fig. 2d and Extended Data Fig. 3a,c). Among the sequenced genomes of the Solanaceae family, the FS8.1 orthologue could only be identified from pepper (Extended Data Fig. 4a), whereas SIGHT-16 orthologues were identified from potato, eggplant and pepper (Extended Data Fig. 4b). Further phylogenetic analyses indicated that while FS8.1 orthologues are only present in

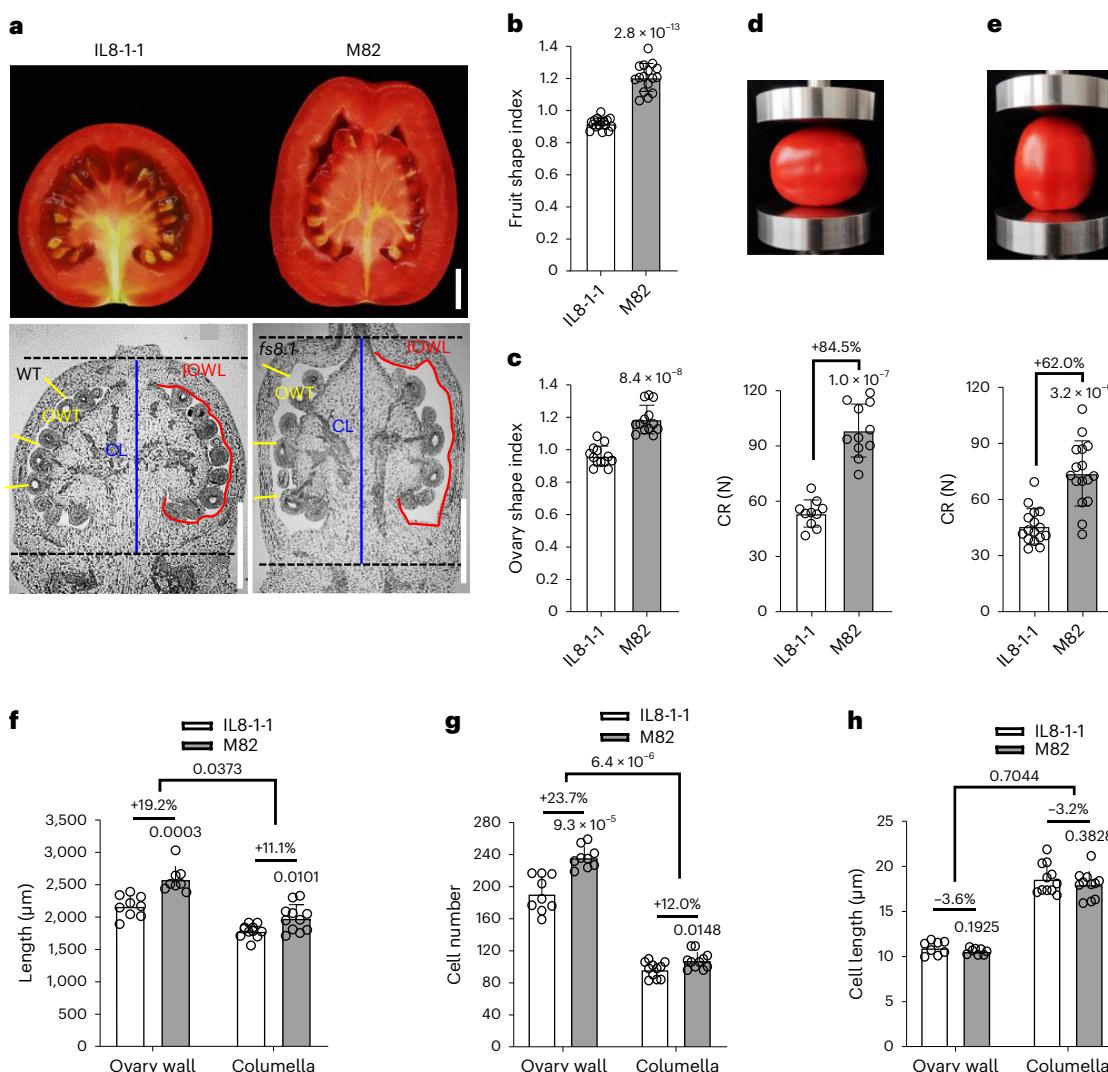


Fig. 1 | *fs8.1*-mediated shape change confers enhanced fruit resistance to squeezing. **a**, Ripe fruits (top) and ovaries at anthesis (bottom) of IL8-1-1 and M82. The red and blue lines indicate the inner ovary walls and columella, respectively. IOWL, inner ovary wall length; CL, columella length; OWT, ovary wall thickness. Scale bars, 1 cm (top), 0.5 mm (bottom). **b**, Fruit shape index of IL8-1-1 and M82 ($n=16$ fruits). **c**, Ovary shape index of IL8-1-1 and M82 ($n=13$ ovaries). **d**, **e**, CR of IL8-1-1 and M82 in the abaxial–adaxial (**d**) and proximal–distal (**e**) directions. The fold changes (FCs) between M82 and IL8-1-1 are given.

In **d**, $n=10$ (for IL8-1-1) and 11 (for M82) fruits. In **e**, $n=16$ fruits. **f**–**h**, Histological and cellular observations of the anthesis ovaries of IL8-1-1 and M82. Cell number and cell size were measured along the red and blue lines indicated in **a**. In **f**, $n=9, 8, 11$, 11 ovaries from left to right. In **g**, $n=9, 9, 11, 11$ ovaries from left to right. In **h**, $n=8, 8, 11, 11$ ovaries from left to right. Bars represent means \pm s.d. The significance of differences was evaluated using two-tailed Student's *t*-tests, with exact *P* values indicated on the graph.

some dicot plants (Extended Data Fig. 4a), SIGT-16 orthologues (that is, canonical GT-2 factors) are broadly conserved in both dicot and monocot plants (Extended Data Fig. 4b).

FS8.1 regulates fruit shape by activating *SIKRP2* expression

To investigate how FS8.1 regulates fruit shape at the transcriptional level, we performed an RNA sequencing (RNA-seq) assay to identify genes regulated by FS8.1 during the early stages of ovary development. Consistent with the observation that the *fs8.1* mutation exerts its effects on fruit shape early in carpel development at least 6 d before anthesis (DBA)¹³, we noted that *fs8.1* led to obvious ovary shape change at 9 DBA (Fig. 3a,b and Extended Data Fig. 5a,b). Thus, we compared the transcriptome profiles between M82 and *Comp-2* ovaries at 9 DBA (Fig. 3a,b and Extended Data Fig. 5a,b). These analyses led to the identification of 1,061 genes that were differentially expressed (fold change (FC) > 1.5 , false discovery rate (FDR)-adjusted *P* < 0.05) between the two genotypes. These genes were defined as FS8.1-regulated genes

(Fig. 3c and Supplementary Table 1). Among them, 655 genes (61.7%) were upregulated by FS8.1, while 406 genes (38.3%) were downregulated by FS8.1 (Fig. 3c). Gene Ontology (GO) analysis revealed that the FS8.1-regulated genes were significantly enriched in several pathways, including those related to the cell cycle and cell differentiation (Fig. 3d).

Notably, many positive regulators of cell cycle progression were downregulated by FS8.1. This list includes genes encoding cyclins, cyclin-dependent kinases (CDKs) and microtubule-associated proteins (Extended Data Fig. 5c and Supplementary Table 2). In contrast, many negative regulators of the cell cycle were upregulated by FS8.1 (Extended Data Fig. 5c and Supplementary Table 2). For example, many genes encoding the well-characterized CDK inhibitors of the KIP-RELATED PROTEIN (KRP) family^{24–26} and the SIAMESE-RELATED (SMR) family^{27,28} were upregulated by FS8.1 (Extended Data Fig. 5c and Supplementary Table 2). Taken together, these results revealed that FS8.1 downregulates positive regulators of the cell cycle while upregulating negative regulators of the cell cycle.

The above results support the hypothesis that FS8.1 regulates fruit shape via upregulation of *KRP* genes that inhibit cell cycle progression. To test this hypothesis, we aimed to verify the FS8.1-mediated regulation of the expression of *SlKRP2*, a tomato homologue of the *Arabidopsis KRP2* gene (Extended Data Fig. 5d). It was previously reported that *AtKRP2* plays a critical role in the inhibition of cell cycle progression during leaf development²⁵. We found that *SlKRP2* expression levels were markedly decreased in the *fs8.1* mutant compared with the WT (that is, IL8-1-1) ovaries (Fig. 3e). Consistently, *SlKRP2* expression levels were significantly increased in the *Comp-1* and *Comp-2* ovaries compared with the *fs8.1* mutant (that is, M82) ovaries (Fig. 3f). Together, these results supported the idea that FS8.1 activates *SlKRP2* expression.

To test whether and how *SlKRP2* regulates fruit shape, we generated its null mutants in IL8-1-1 and found that the resulting *slkrp2* single mutants did not affect fruit shape (Extended Data Fig. 5e,f). We then generated CRISPR/Cas9-mediated mutants in which *SlKRP2* and its paralogous genes, *SlKRP1* and *SlKRP4*, were knocked out (Extended Data Fig. 5d,e). Compared with the round-shaped IL8-1-1 fruits, the *slkrp1 slkrp2 slkrp4* triple mutant fruits were elongated and had an increased fruit shape index (Fig. 3g,h), indicating that *SlKRP2* acts redundantly with *SlKRP1* and *SlKRP4* to regulate fruit shape.

To determine the genetic relationship between FS8.1 and *SlKRP2* in regulating fruit shape, we generated *SlKRP2*-overexpression (*SlKRP2-OE*) plants (Extended Data Fig. 5g) in the background of M82, which harbours a loss-of-function allele of *FS8.1* (Extended Data Fig. 1j). Compared with the elongated M82 fruits, the *SlKRP2-OE* fruits were round and had a decreased fruit shape index (Fig. 3i,j), indicating that *SlKRP2* could complement the *fs8.1*-mediated fruit shape changes. These results implied that *SlKRP2* acts genetically downstream of *FS8.1* in regulating fruit shape. Collectively, these results supported the notion that FS8.1 regulates fruit shape via activation of the expression of *SlKRP* genes.

FS8.1 activates *SlKRP2* expression by interacting with SIGT-16

We then explored how FS8.1 activates *SlKRP2* expression. Chromatin immunoprecipitation-quantitative polymerase chain reaction (ChIP-qPCR) assays using *FS8.1-GFP* plants (Extended Data Fig. 2h) and anti-GFP antibody indicated that *FS8.1-GFP* was enriched in the *SlKRP2* promoter (Fig. 4a,b). However, electrophoretic mobility shift assays (EMSA) failed to detect direct binding of maltose-binding protein (MBP)-tagged FS8.1 (MBP-FS8.1) to the GT element (GGTAATT) of the *SlKRP2* promoter (Fig. 4c,d), suggesting the possibility that FS8.1 is recruited to the *SlKRP2* promoter by interacting proteins that can directly bind to the GT element.

Given that GT-2 factors usually act as dimers to regulate gene expression^{17,29}, we hypothesized that FS8.1 could form a complex with GT-2 factors that are able to bind GT elements. To test this hypothesis, we conducted firefly luciferase (LUC) complementation imaging (LCI) assays to test whether FS8.1 interacts with five tomato GT-2 factors that are phylogenetically closest related³⁰ (Extended Data Fig. 3b). The results showed that FS8.1 strongly interacted with SIGT-16 and SIGT-34 (Extended Data Fig. 6a). Consistently, both MBP-SIGT-16

and MBP-SIGT-34 were able to pull down glutathione S-transferase (GST)-tagged FS8.1 (GST-FS8.1) (Fig. 4e,f), indicating that FS8.1 interacts with SIGT-16 and SIGT-34 in vitro.

Considering the high sequence similarity between SIGT-16 and SIGT-34 (Extended Data Fig. 3a,b), we focused on the former for further analysis. The results of our LCI assays indicated that the truncated protein encoded by the *fs8.1* mutant allele failed to interact with SIGT-16 (Fig. 4g,h). Additional LCI assays (Fig. 4g,h) and pull-down assays (Fig. 4i,j) revealed that SIGT-16 interacted with itself but that FS8.1 did not. Moreover, the FS8.1–SIGT-16 interaction was significantly stronger than the SIGT-16–SIGT-16 interaction (Fig. 4g,h). Collectively, these results support the idea that FS8.1 might regulate *SlKRP2* expression through complexation with the canonical GT-2 factor SIGT-16.

We then investigated whether SIGT-16 regulates *SlKRP2* expression through promoter binding. ChIP-qPCR analysis using *SIGT-16-GFP* plants (Extended Data Fig. 6b) and anti-GFP antibody indicated that SIGT-16-GFP was enriched in the *SlKRP2* promoter (Fig. 4k). Moreover, EMSAs indicated that SIGT-16 directly binds the GT element of the *SlKRP2* promoter (Fig. 4c) and that this binding depends on the N-terminal trihelix domain (TD1) of SIGT-16 (Fig. 4d). These results, together with the observation that the expression pattern of *SlKRP2* was similar to that of *SIGT-16* in developing ovaries (Extended Data Fig. 5a), support the idea that SIGT-16 activates *SlKRP2* expression through promoter binding.

To examine the role of *SIGT-16* and *SIGT-34* in regulating fruit shape, we generated CRISPR/Cas9-mediated mutants of these genes in the genetic backgrounds of the round-fruited Ailsa Craig (AC) and IL8-1-1 (Extended Data Fig. 6c). Whereas the single mutants did not affect fruit shape, the *slgt-16 slgt-34* double mutants in both backgrounds displayed elongated fruits (Fig. 4l,m and Extended Data Fig. 6d,e). Consistently, *SlKRP2* expression levels in the ovaries of these double mutants were significantly reduced compared with those in the WT ovaries (Fig. 4n and Extended Data Fig. 6f). These results implied that SIGT-16 acts redundantly with SIGT-34 to regulate fruit shape formation through activation of *SlKRP2* expression.

Notably, *FS8.1-GFP* enrichment in the *SlKRP2* promoter was largely reduced in the *slgt-16 slgt-34* double mutant compared with the WT (Fig. 4b), supporting the notion that FS8.1 recruitment to the *SlKRP2* promoter depends on SIGT-16 and SIGT-34. Consistently, overexpression of *SIGT-16* rescued the fruit shape of M82 from elongated to round (Fig. 4o,p and Extended Data Fig. 6b). These results collectively suggest that SIGT-16 acts genetically downstream of FS8.1 in regulating fruit shape.

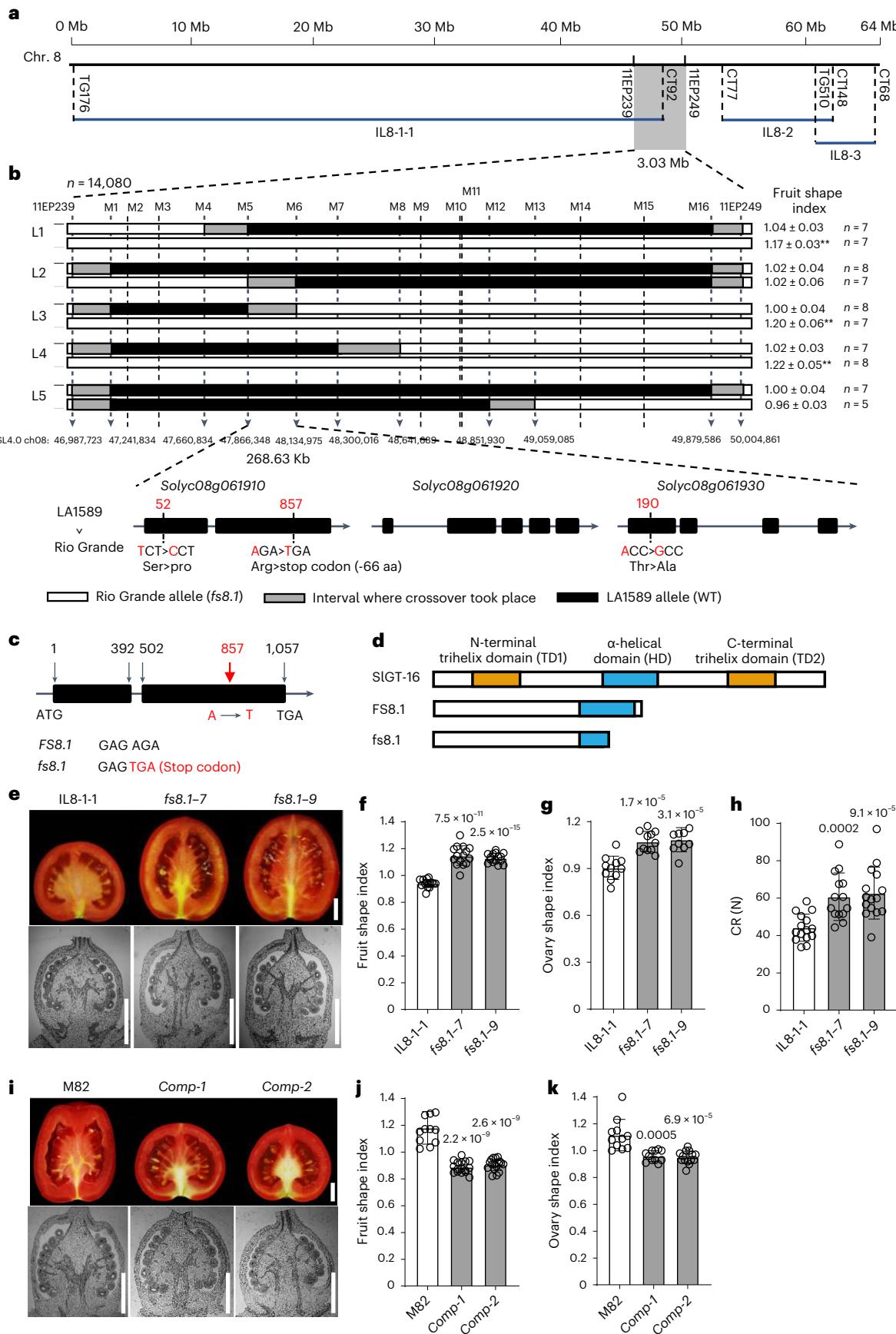
Taken together, our results support the idea that FS8.1 regulates fruit shape via interaction with canonical GT-2 factors such as SIGT-16 and possibly others to activate *KRP* expression.

FS8.1 enhances the transcriptional activity of SIGT-16

Next, we aimed to understand the mechanism and functional significance of the FS8.1–SIGT-16 interaction. The results of domain mapping experiments with pull-down assays revealed that the HD of SIGT-16 is involved in its interaction with FS8.1 (Fig. 5a), whereas both the

Fig. 2 | FS8.1 encodes a GT-2-like protein that lacks the duplicate trihelix domains. **a**, Schematic representation of the introgressed regions of IL8-1-1, IL8-2 and IL8-3. The blue lines represent the introgressed chromosome fragments from *S. pennellii* (LA0716). The vertical dashed lines indicate the positions of molecular markers. **b**, Fine-mapping of *FS8.1*. Top: high-resolution mapping narrowed *FS8.1* to the DNA segment between markers M5 and M6. *n* = number of plants. Fruit shape indexes of recombinant progenies (L1–L5) are on the right. Data are means ± s.d. Asterisks indicate significant differences between WT and *fs8.1* progenies (***P* < 0.01; Student's *t*-tests). Bottom: candidate genes in the *FS8.1* mapping region. Three non-synonymous SNPs between LA1589 and Rio Grande are labelled on the gene sketch. **c**, Schematic diagram of *Solanum lycopersicum* O8g061910. The black boxes represent exons. The SNP (A857T) causing early termination

(R250Stop) is highlighted in red. **d**, Schematic domain architecture of the indicated proteins. **e**, Ripe fruits (top) and ovaries at anthesis (bottom) of IL8-1-1 and *fs8.1* mutants. Scale bars, 1 cm (top), 0.5 mm (bottom). **f,g**, Fruit shape index (**f**) and ovary shape index (**g**) of IL8-1-1 and *fs8.1* mutants. In **f**, *n* = 15 fruits. In **g**, *n* = 11, 11, 10 ovaries from left to right. **h**, CR of IL8-1-1 and *fs8.1* ripe fruits (*n* = 15, 14, 15 fruits from left to right). **i**, Ripe fruits (top) and anthesis ovaries (bottom) of M82 and complementation (*Comp*) lines. Scale bars, 1 cm (top), 0.5 mm (bottom). **j,k**, Fruit shape index (**j**) and ovary shape index (**k**) of M82 and *Comp* lines. In **j**, *n* = 12, 16, 16 fruits from left to right. In **k**, *n* = 11, 11, 14 ovaries from left to right. Bars represent means ± s.d. The significance of differences was evaluated using two-tailed Student's *t*-tests, with exact *P* values indicated on the graph.



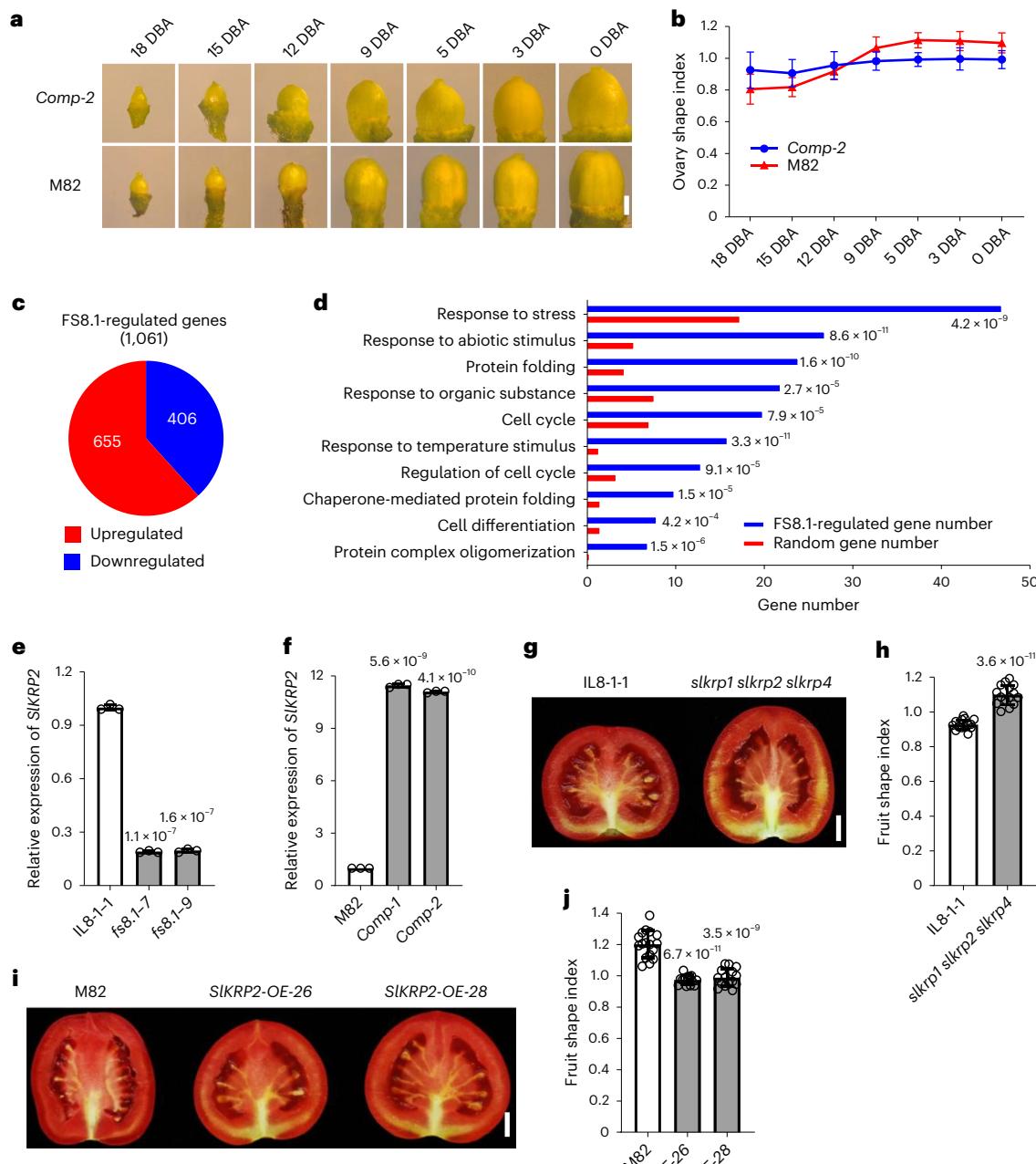
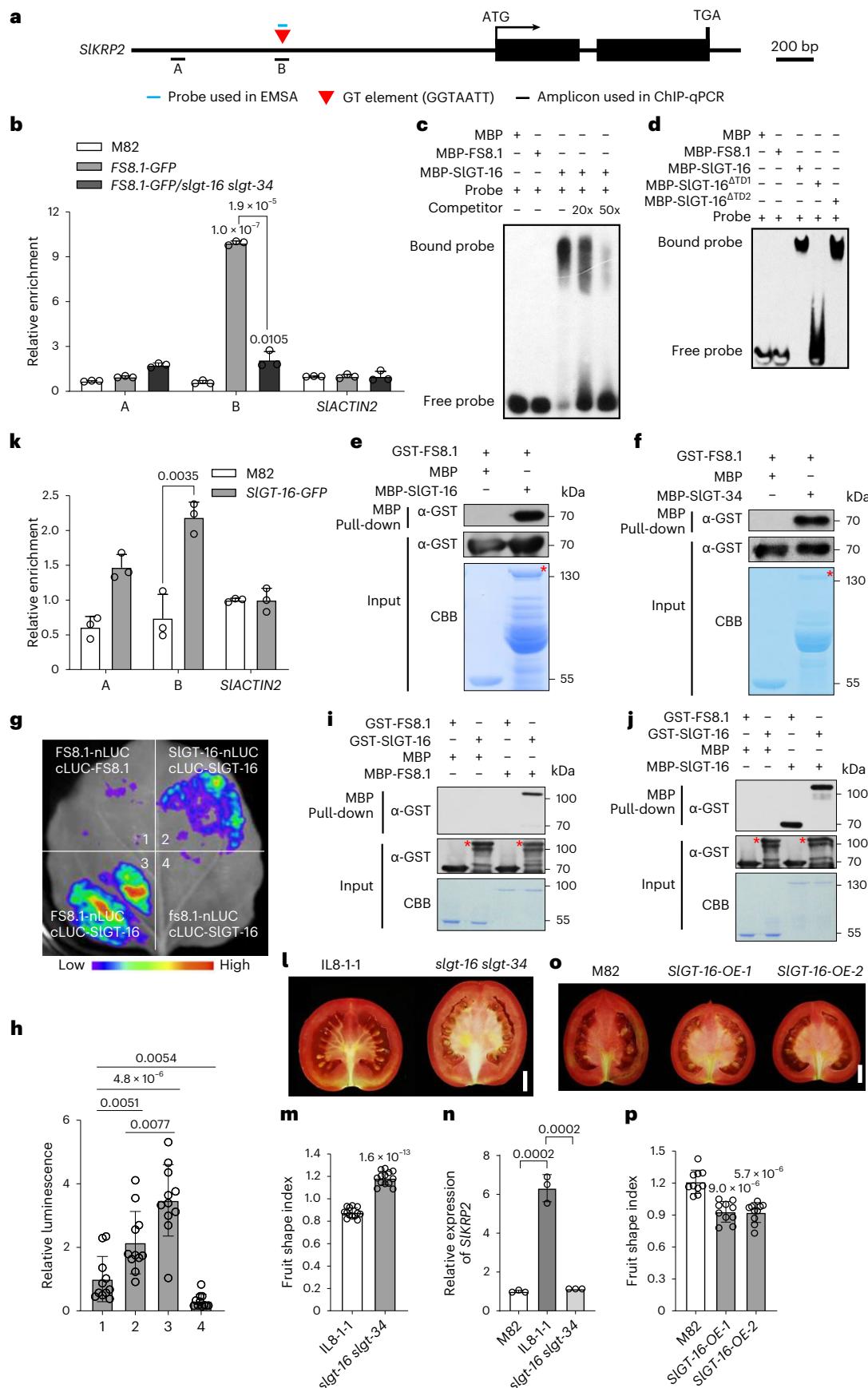


Fig. 3 | FS8.1 regulates fruit shape via activation of SlkRP2 expression.
a, Representative images of *Comp-2* and *M82* ovaries at the indicated developmental stages. Scale bar, 0.5 mm. **b**, Ovary shape indexes of *Comp-2* and *M82* at the indicated developmental stages. For *Comp-2*, $n = 7, 7, 10, 10, 10, 10, 12$ ovaries from left to right; for *M82*, $n = 4, 6, 6, 5, 8, 7, 6$ ovaries from left to right. **c**, Distribution of genes differentially expressed between *Comp-2* and *M82* ovaries at 9 DBA. **d**, GO enrichment analysis of FS8.1-regulated genes. The results of classification within the biological process categories with P values from the statistical overrepresentation test (<https://pantherdb.org/>) are shown. **e**, RT-qPCR assays showing *SlkRP2* expression in 9-DBA ovaries of

the indicated genotypes ($n = 3$ independent biological replicates). **f**, RT-qPCR assays showing *SlkRP2* expression in 9-DBA ovaries of the indicated genotypes ($n = 3$ independent biological replicates). **g**, Ripe fruits of IL8-1-1 and the *slkRP1 slkRP2 slkRP4* mutant. Scale bar, 1 cm. **h**, Fruit shape indexes of IL8-1-1 and the *slkRP1 slkRP2 slkRP4* mutant ($n = 15$ fruits). **i**, Ripe fruits of *M82* and *SlkRP2-OE* lines. Scale bar, 1 cm. **j**, Fruit shape indexes of *M82* and *SlkRP2-OE* lines ($n = 17, 15, 15$ fruits from left to right). Bars represent means \pm s.d. The significance of differences was evaluated using two-tailed Student's *t*-tests, with exact P values indicated on the graph.

HD and the C-terminal TD2 are important for SIGT-16 homodimerization (Fig. 5b). Because GT-2 factors usually undergo dimerization through the HD^{17,29}, we were interested in determining whether and how FS8.1 affects SIGT-16 dimerization. For this purpose, we performed an *in vitro* pull-down assay in which the amounts of MBP-SIGT-16 and GST-SIGT-16 were kept constant in each sample, whereas the amount of

GST-FS8.1 was increased through a gradient. In this assay, recombinant MBP-SIGT-16 was used to pull down GST-SIGT-16 and GST-FS8.1. The results showed that increasing concentrations of GST-FS8.1 had a minor (if any) effect on the ability of MBP-SIGT-16 to pull down GST-SIGT-16 (Fig. 5c), suggesting that FS8.1 tends to complex with SIGT-16 homodimers rather than impair the formation of SIGT-16 homodimers.



We then employed a dual-luciferase reporter system³¹ to study the effect of FS8.1 on SIGT-16-mediated transcriptional activation of *SIKRP2* expression. For this purpose, we cloned a 2,053-bp *SIKRP2* promoter

sequence and inserted it into a dual-LUC reporter construct to generate a *proSIKRP2::LUC* reporter construct (Fig. 5d). Co-expression of *35S::SIGT-16* with *proSIKRP2::LUC* in *Nicotiana benthamiana* leaves

Fig. 4 | FS8.1 activates SIKRP2 expression by interacting with SIGT-16 and SIGT-34. **a**, Schematic representation of *SIKRP2* showing the amplicons and probe used for ChIP-qPCR and EMSAs, respectively. **b**, ChIP-qPCR results showing the enrichment of FS8.1-GFP in the *SIKRP2* promoter ($n = 3$ independent biological replicates). The amplicons detected by qPCR are indicated in **a**. *SIACtIN2* was used as a non-specific target. **c**, EMSA results showing that SIGT-16, but not FS8.1, directly binds to the *SIKRP2* promoter. **d**, EMSA results showing that SIGT-16 directly binds the *SIKRP2* promoter through its TD1. SIGT-16^{ATD1} and SIGT-16^{ATD2} are SIGT-16 derivatives lacking TD1 and TD2, respectively. **e,f**, Pull-down assay results showing that FS8.1 interacts with SIGT-16 (**e**) and SIGT-34 (**f**). CBB, Coomassie brilliant blue staining. **g**, LCI assay results showing the association between FS8.1 and SIGT-16. **h**, Quantitative analysis of luminescence intensity in **g** ($n = 11$ independent biological replicates). **i**, Pull-down assay

results showing that FS8.1 interacts with SIGT-16 but not with itself. **j**, Pull-down assay results showing that SIGT-16 interacts with FS8.1 and itself. **k**, ChIP-qPCR results showing the enrichment of SIGT-16-GFP in the *SIKRP2* promoter ($n = 3$ independent biological replicates). The amplicons detected by qPCR are indicated in **a**. *SIACtIN2* was used as a non-specific target. **l**, Ripe fruits of IL8-1-1 and *slgt-16 slgt-34* mutants. Scale bar, 1 cm. **m**, Fruit shape indexes of IL8-1-1 and *slgt-16 slgt-34* mutants ($n = 13$ fruits). **n**, RT-qPCR assay results showing *SIKRP2* expression in ovaries at anthesis of the indicated genotypes ($n = 3$ independent biological replicates). **o**, Ripe fruits of M82 and *SIGT-16-OE* lines. Scale bar, 1 cm. **p**, Fruit shape indexes of M82 and *SIGT-16-OE* lines ($n = 10$ fruits). Bars represent means \pm s.d. The significance of differences was evaluated using two-tailed Student's *t*-tests, with exact *P* values indicated on the graph. Experiments in **c-f, i and j** were repeated independently at least three times with similar results.

led to increased LUC activity, suggesting that SIGT-16 activates *proSIKRP2::LUC* expression (Fig. 5d,e). When FS8.1 was co-expressed with SIGT-16 and the *proSIKRP2::LUC* reporter, the SIGT-16-dependent activation of LUC activity was further enhanced (Fig. 5d,e). In parallel control experiments, the *fs8.1* mutant protein showed a negligible effect on SIGT-16-dependent activation of LUC activity (Fig. 5d,e). These results suggest that FS8.1 functions as a co-activator of SIGT-16.

Taken together, our results support a scenario in which FS8.1 forms a transcriptional complex together with SIGT-16 and thereby enhances its transcriptional activation for the expression of cell cycle inhibitory genes.

In summary, our results suggest that the FS8.1-SIGT-16 transcriptional module exerts a stronger cell proliferation-inhibitory effect on the ovary wall than on the columella. If this were true, we anticipated that *FS8.1* and *SIGT-16* may be differentially expressed in the ovary wall and the columella of developing ovaries. To test this hypothesis, we measured the tissue-specific expression of *FS8.1* and *SIGT-16* by using transgenic plants harbouring their promoter-driven β-glucuronidase (*GUS*) fusion constructs. The results showed that the *proFS8.1::GUS* and *proSIGT-16::GUS* activities were much stronger in the ovary wall than in the columella (Fig. 5f). Consistently, reverse transcription-quantitative polymerase chain reaction (RT-qPCR) assays showed that the expression levels of both *FS8.1* and *SIGT-16* were higher in the ovary wall than in the columella (Fig. 5g,h). As expected, the expression levels of *SIKRP2*, which serves as a readout of the transcriptional output of the FS8.1-SIGT-16 module, were also higher in the ovary wall than in the columella (Fig. 5g,i). Given that SIKRP2 functions as an inhibitor of cell cycle progression, our results support a model where, in the WT, the FS8.1-SIGT-16 module exerts a stronger inhibitory effect on the cell cycle progression of the ovary wall than on the columella, leading to round fruits (Fig. 5j).

We then examined how the *fs8.1* mutation alters the differential expression pattern of relevant genes between the ovary wall and the columella. The results showed that in the *fs8.1* mutant ovaries, the *SIGT-16* expression levels in both tissue layers were largely unchanged (Fig. 5h), but the relative expression levels of *SIKRP2* in the ovary walls were lower than those in the columella (Fig. 5i). Thus, disruption of *FS8.1* function resulted in a weaker inhibitory effect on the cell cycle

progression of the ovary walls than of the columella, leading to elongated fruits (Fig. 5j).

fs8.1 is selected in machine-harvestable processing tomatoes

We sought to determine how common the *fs8.1* allele is in modern lines and when the mutation arose in tomato domestication and selection. The analyses showed that the mutation arose relatively late as it was not found in the Varitome collection, a group of 168 accessions consisting of the closest wild relatives to cultivated tomato, *S. pimpinellifolium*, and many accessions of *S. lycopersicum* var. *cerasiforme*, an intermediate type between the fully wild and cultivated accessions³² (Fig. 6a and Supplementary Tables 3–7). The *fs8.1* allele was found at low levels in Regional Latin American (8%) and Regional Italian accessions (14%) including the well-known San Marzano type cultivar (Fig. 6a and Supplementary Tables 3–5), suggesting that the allele arose during initial selections of cultivated types in Latin America or Europe well after domestication of the crop. The distribution of the *fs8.1* allele was significantly increased (60%) in transitional/early processing accessions (Fig. 6a and Supplementary Table 6). Strikingly, a survey of modern processing and fresh-market accessions from China and the United States showed that the mutant allele is present in current-day processing lines but absent in fresh-market varieties (Fig. 6a and Supplementary Tables 6 and 7). The cultivar Red Top VF was released in 1952 and carried the *fs8.1* mutation (Supplementary Table 6). The release dates for San Marzano and Roma VF were not known but they also carried the *fs8.1* mutation (Supplementary Tables 5 and 6). These were key parental genotypes in initial crosses to develop cultivars with the characteristics demanded by the processing industry. The historical pedigree information and the year of release for the modern cultivars (Supplementary Tables 5 and 6) suggested that selection of the *fs8.1* allele in processing-tomato breeding occurred during the 1960s–1970s. Thus, in line with the history of redesigning the tomato for mechanized production^{6,33,34}, the *fs8.1* mutation was highly selected in the breeding of modern processing cultivars.

Editing of FS8.1 improves CR without compromising quality

Next, we were interested in determining whether *FS8.1* can be used to improve the CR of fresh-market tomato fruits. To accomplish this

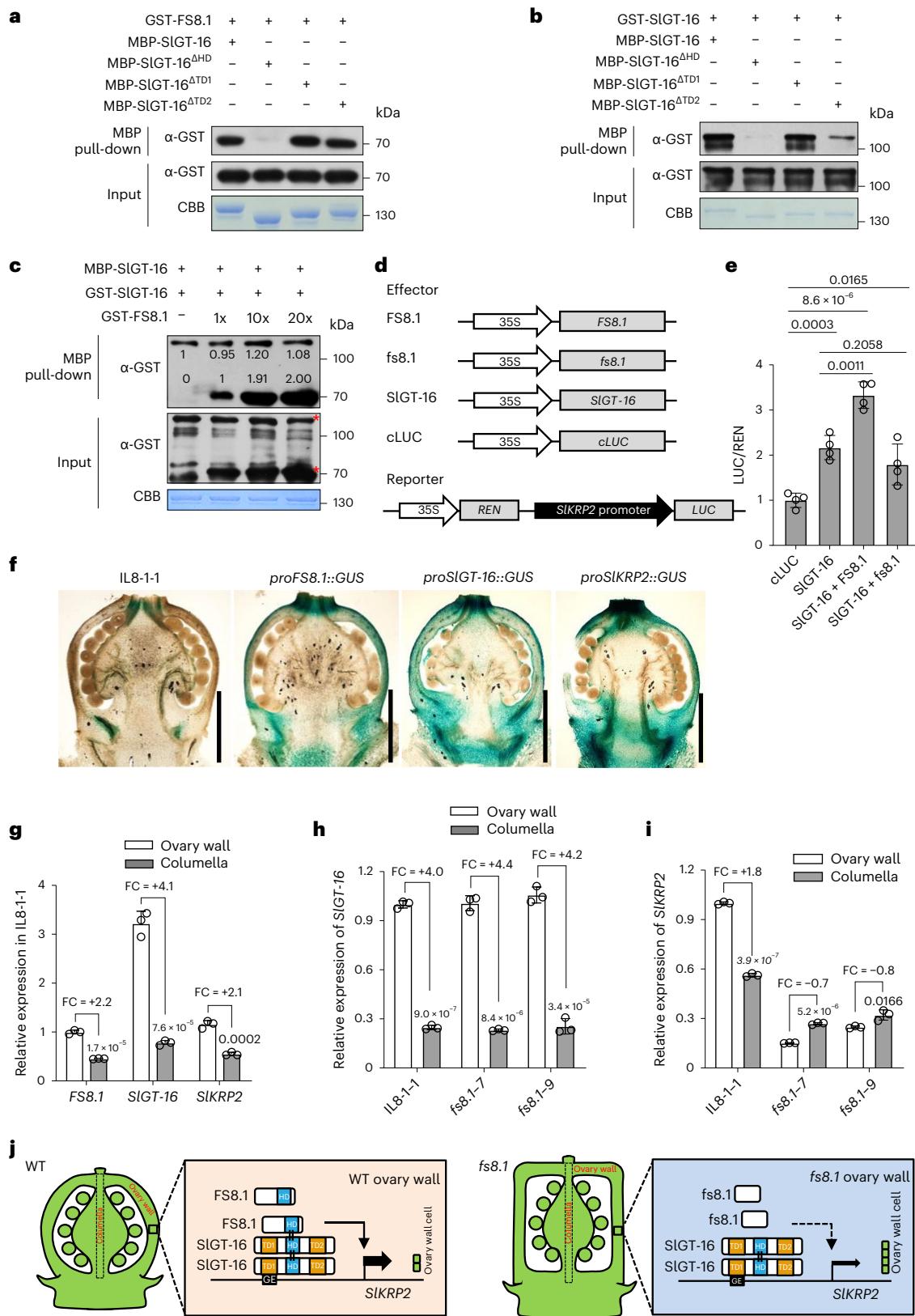
Fig. 5 | FS8.1 increases the transcriptional activity of SIGT-16. **a**, Pull-down assay results showing that SIGT-16 interacts with FS8.1 through HD. **b**, Pull-down assay results showing that both HD and TD2 are important for SIGT-16 homodimerization. **c**, Pull-down assay results showing that FS8.1 tends to form a complex with SIGT-16 homodimers rather than impair the formation of SIGT-16 homodimers. The bands were quantified using ImageJ software. **d**, Schematic diagram showing the constructs used in the transient expression assays of **e**. **e**, Transient expression assays showing that FS8.1 functions as a co-activator of SIGT-16-mediated gene transcription ($n = 4$ independent biological replicates). The *proSIKRP2::LUC* reporter was cotransfected with the indicated effector constructs. The LUC:REN ratio represents the *proSIKRP2::LUC* activity relative to that of the internal control (*REN* driven by the 35S promoter). **f**, Longitudinal section of IL8-1-1 (WT), *proFS8.1::GUS*, *proSIGT-16::GUS* and *proSIKRP2::GUS*

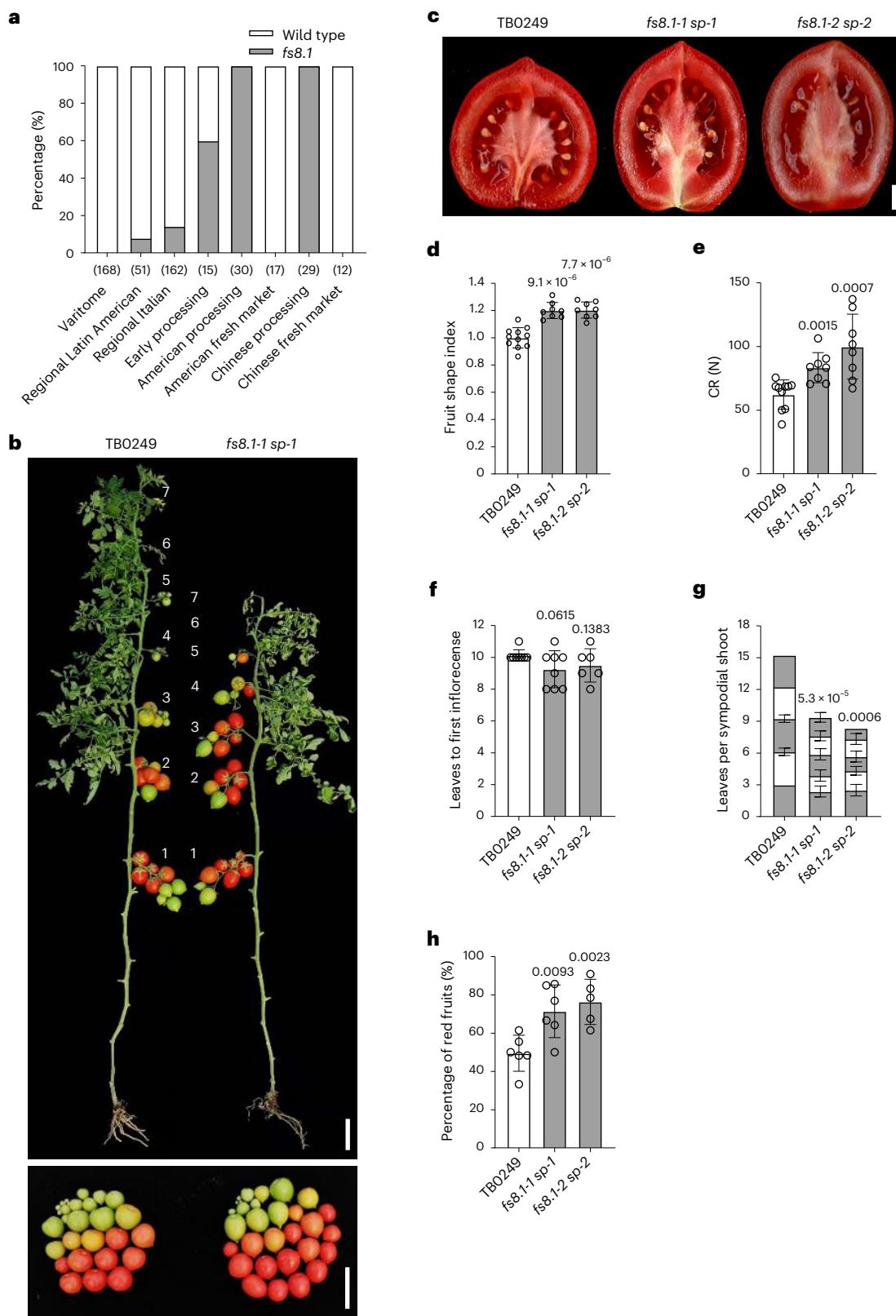
ovaries showing the tissue-specific expression patterns of these genes. **g**, RT-qPCR results showing the relative expression levels of the indicated genes in the ovary walls and ovary columella of IL8-1-1 anthesis ovaries ($n = 3$ independent biological replicates). The expression FCs between the ovary walls and columella are given. **h,i**, RT-qPCR results showing the relative expression levels of *SIGT-16* (**h**) and *SIKRP2* (**i**) in the ovary walls of ovaries at anthesis of the indicated genotypes ($n = 3$ independent biological replicates). The expression FCs between the ovary walls and columella are given. **j**, Proposed model for FS8.1-mediated fruit shape formation. GE, GT element. Bars represent means \pm s.d. The significance of differences was evaluated using two-tailed Student's *t*-tests, with exact *P* values indicated on the graph. Experiments in **a–c** were repeated independently at least three times with similar results.

task, we used the CRISPR/Cas9 gene-editing system to mutate *FS8.1* in the background of the typical fresh-market cultivar AC (Extended Data Fig. 7a). As expected, the resulting mutants produced elongated fruits that had significantly increased CR (Extended Data Fig. 7b–d). Notably, fruit weight (Extended Data Fig. 7e) and the levels of total soluble solids (Extended Data Fig. 7f), sugars (Extended Data Fig. 7g,h),

acids (Extended Data Fig. 7i,j) and lycopene (Extended Data Fig. 7k) of mutant fruits were comparable to those of their parental fruits, indicating that manipulation of *FS8.1* does not lead to unfavourable effects on fruit weight and quality.

In addition to presenting enhanced fruit CR, cultivars suitable for machine harvesting should have determinate growth and uniform





fruit ripening. Indeed, *fs8.1* (refs. 5,7,13) and *self-pruning* (*sp*)^{35–37} natural mutations have facilitated mechanical harvesting of processing tomatoes and thus revolutionized the tomato industry. In this context, we used CRISPR/Cas9 to engineer both *FS8.1* and *SP* in the background of TB0249, an elite fresh-market cultivar with an indeterminate growth habit and that produces round fruits (Extended Data Fig. 7a). As expected, the resulting *fs8.1 sp* double mutants displayed determinate

growth habits and produced elongated fruits with enhanced CR (Fig. 6b–e). Furthermore, the double mutants displayed typical *sp* determinate growth habits that provide progressively faster sympodial shoot flowering and synchronized fruit ripening, leading to early yields (Fig. 6b,f–h). Thus, our work probably exemplifies a simple one-step approach for rapidly improving fruit CR and generating early-yielding, machine-harvestable fresh-market tomatoes.

Fig. 6 | Editing of *FS8.1* improves the CR of fresh-market tomato fruits without compromising quality. **a**, Population of *FS8.1* alleles in different tomato germplasm classes. Varitome is an ancestral population of tomato composed of *Solanum pimpinellifolium* (SP) accessions from Ecuador and Peru; *Solanum lycopersicum* var. *cerasiforme* (SLC) accessions from Ecuador, Peru and Central America; and *Solanum lycopersicum* var. *lycopersicum* (SLL) from Mexico. The number of accessions in each category is indicated in parentheses. **b**, Representative images of TB0249 and *fs8.1* sp double mutant plants (top) and their fruits (bottom). The images were taken at 167 d after germination. The inflorescences are indicated by numbers. Scale bar, 10 cm. **c**, Ripe fruits of the

TB0249 and *fs8.1* sp double mutant. Scale bar, 1 cm. **d,e**, Fruit shape indexes (**d**) and CR in the abaxial–adaxial direction (**e**) of TB0249 and *fs8.1* sp double mutant fruits. In **d**, $n = 11$, 8, 8 fruits from left to right. In **e**, $n = 10$, 8, 8 fruits from left to right. **f,g**, Flowering time for primary (**f**) and sympodial (**g**) shoots of TB0249 and *fs8.1* sp plants ($n = 8$, 8, 6 plants from left to right). The sum of the leaves from five sympodial shoots was used to determine statistical significance in **g**. **h**, Percentage of red fruits from TB0249 and *fs8.1* sp plants ($n = 6$, 6, 5 plants from left to right). In **d–h**, bars represent means \pm s.d. The significance of differences was evaluated using two-tailed Student's *t*-tests, with exact *P* values indicated on the graph.

Discussion

We have identified the long-sought-after gene underlying the elongated fruit shape of machine-harvestable processing tomatoes and elucidated its mode of action at the cellular and molecular levels. *FS8.1* encodes a non-canonical GT-2 factor and complexes with canonical GT-2 factors such as *SLGT-16* to activate cell cycle inhibitory genes. *FS8.1* itself and *SLGT-16* are differentially expressed in the ovary wall and the columella, thereby exerting distinct transcriptional output on these tissue layers. Our results support a scenario in which *FS8.1* determines fruit shape by differentially inhibiting the cell proliferation rate of the ovary wall and the columella.

Notably, the wild progenitors of cultivated tomatoes have tremendous genetic diversity at the whole-genome level but produce fruits that are almost invariably round. In contrast, cultivated tomatoes have little genetic variation in their genomes but exhibit tremendous diversity in terms of fruit shape^{1,5,9,38,39}. Presumably, with respect to fruits, roundness provides several advantages over other shapes for seed dispersal in the wild, and domestication and selection have presumably led to dramatic changes in the fruit shape of tomato cultivars^{5,9,11,32,40}. Indeed, our analysis of the distribution of *FS8.1* alleles demonstrates that the *fs8.1* mutation arose recently and was highly selected in the breeding of machine-harvestable processing tomatoes. These findings are in line with the notion that selection for firmer-fruited varieties resulted in a shape change from round to elongated or torpedo-shaped tomatoes^{4,5,11}.

Our study raised the intriguing question of how the *fs8.1* mutation leads to enhanced CR of tomato fruits. It is conceivable that elongated fruits have several aspects of advantage over round fruits to withstand mechanized harvesting and postharvest processing. First, they are much easier to transport by a conveyor belt. Second, they are more resistant to crushing because the contact areas among elongated fruits are much larger than those of round fruits^{4–6}. In addition to shape change, the enhanced CR of *fs8.1* mutant fruits could also be due to additional changes such as cell wall structure. However, several lines of evidence do not support this scenario. First, the *FS8.1* gene is temporally expressed during the early stages of ovary development and determines the pattern of ovary shape before anthesis¹³. Second, the ovary wall thickness of the *FS8.1* knockout mutants was actually decreased due to reduced cell layers (Extended Data Fig. 2e,f). Third, the *fs8.1* mutation does not affect fruit ripening. Thus, it is not likely that the *fs8.1* mutation enhances fruit CR by changing cell wall remodelling. It will be interesting to further explore the cellular, molecular and biochemical mechanisms underlying *fs8.1*-mediated effects on fruit CR in future studies.

Compared with processing-tomato fruits, typical fresh-market tomato fruits are of higher quality in terms of nutrition and flavour. However, fresh-market tomato fruits can experience severe soft fruit-associated deterioration in terms of harvest, transportation and storage. Thus, enhancing fruit firmness without compromising other quality values is an earnest target for modern breeding of fresh-market tomatoes. Given that fruit firmness is determined by the rate of softening during ripening, a process in which fruit colour, texture and flavour traits are formed, breeding firm fruits with desirable qualities has

proven to be challenging. For example, the natural mutations *ripening inhibitor* (*rin*) and *non-ripening* (*nor*) have been used to improve fruit firmness^{41,42}. However, these ripening-related mutations often adversely affect colour, flavour and nutritional qualities^{41,43–45}. Thus, an alternate strategy to improve fruit firmness without affecting the ripening process would ideally achieve both goals of long shelf-life and high quality.

Although the *fs8.1* allele has revolutionized the breeding of processing tomatoes in the 1970s, the application of this allele in fresh-market tomato breeding has proven to be challenging. Considering that the *fs8.1* locus is located in a heterochromatic region¹⁴, it is likely that the extremely low recombination rate around the locus and unfavourable linkage drag have prevented the introduction of this beneficial allele into fresh-market tomatoes. From this perspective, our identification of the *FS8.1* gene makes it possible to breed this beneficial allele into fresh-market tomatoes using gene editing, a powerful breeding technique that can effectively avoid the effects of linkage drag^{46–48}. Indeed, CRISPR/Cas9-mediated precise mutation of *FS8.1* led to enhanced CR of elite fresh-market tomatoes without loss of quality. Previous studies have shown that *PECTATE LYASE*⁴⁹, the GA2-oxidase-encoding gene *FIRMSKIN1* (ref. 18), the LATERAL ORGAN BOUNDARIES transcription factor gene *SILLOB1* (ref. 50), β -galactosidase genes *TBG4* and *TBG6* (ref. 51,52), the MADS-box transcription factor gene *SIMBP3* (ref. 53,54), and several polygalacturonase genes and expansin genes^{55,56} are useful targets to enhance fruit firmness by targeted control of fruit softening. Given that *FS8.1* is not involved in fruit ripening, our work provides an alternate strategy to improve fruit CR without affecting softening.

Furthermore, simultaneous mutation of *FS8.1* and *SP* converted indeterminate fresh-market tomato plants that produce round fruits into determinate plants that produce square fruits suitable for machine harvesting. In this regard, our study probably underscores a potential strategy to breed fresh-market tomatoes suitable for mechanized production.

It is also possible that the consumer preference and culinary utility of round fruits have precluded the breeding of the *fs8.1* allele into fresh-market tomatoes. In this regard, it will be interesting to test whether the *FS8.1* gene can be used to develop more flavourful processing tomatoes from fresh-market varieties.

Methods

Plant material and growth conditions

Seeds of M82 (LA3475), IL8-1-1 (LA4072), IL8-2 (LA4074), IL8-3 (LA4076), AC (LA2838A), LA0716, LA1706, LA3528, E6203 (LA4024) and LA3144 were obtained from the C.M. Rick Tomato Genetics Resource Center (TGRC; <https://tgrc.ucdavis.edu>) at the University of California, Davis. Seeds of Springo, T46567 and Fafnir were purchased from Syngenta. Seeds of Provence and SV4424TH were purchased from Monsanto. Seeds of TB0249 were obtained from our own stocks. IL8-1-1 was used as the WT for *fs8.1-7, fs8.1-9, slgt-16 slgt-34, slkrp2, slkrp1slkrp2slkrp4, proFS8.1::GUS, proSIIGT-16::GUS* and *proSIKRp2::GUS*. M82 was used as the WT for *proFS8.1::FS8.1-GFP* (*Comp*), *SIKRp2-OE* and *SIIGT-16-OE*. AC was used as the WT for *fs8.1-3, fs8.1-16, slgt-16-1, slgt-16-2, slgt-34-1*,

slgt-34-2, *slgt-16-3* *slgt-34-3* and *slgt-16-4* *slgt-34-4*. TB0249 was used as the WT for *fs8.1-1 sp-1* and *fs8.1-2 sp-2*. Seeds of *N. benthamiana* were obtained from our own stocks.

The tomato seeds were germinated on moistened filter paper at room temperature and then sown in 32-cell plastic flats. The seedlings were grown in a growth chamber and maintained at 60% relative humidity under 16 h of light ($200 \mu\text{E m}^{-2} \text{s}^{-1}$) at 26°C and 8 h of darkness at 18°C . Thirty-day-old seedlings were transplanted to the field and grown under standard water and fertilizer management. *N. benthamiana* seeds were germinated in soil in 8-cm diameter pots and then transplanted to 32-cell plastic flats. The *N. benthamiana* seedlings were grown under the same conditions as those of the tomato seedlings.

To collect the developing ovaries more precisely at pre-anthesis stages, the timing of ovary development was determined as described previously¹³. Briefly, flower buds were tagged and the size of the flower was recorded every day until anthesis. Then, flower sizes were plotted against the relative number of days pre-anthesis to estimate the relationship between flower size and the timing of ovary development (that is, DBA). Next, ovaries were collected from flowers estimated to be at 18, 15, 12, 9, 5, 3 and 0 DBA (on the basis of flower size) and used for analysis.

Phenotypic evaluation

For fruit shape analysis, full-size ripe fruits were cut longitudinally, scanned at 300 dpi and analysed using Tomato Analyser (v.4.0)⁵⁷. The fruit shape index was calculated as the ratio of the maximum height/length to the maximum width of a fruit.

For ovary shape analysis, ovaries at anthesis were cut longitudinally and digitized using a Lecia DFC490 digital camera. The maximum length and width of the ovaries were measured using ImageJ software. The ovary shape index was calculated as the ratio of the maximum length to the maximum width of an ovary.

Fruit CR was measured using an Instron 5542 texture analyser as previously described¹⁸. More than eight red ripe fruits of each genotype were used for measurements.

Histological analysis

Ovaries at anthesis were fixed in formalin-acetic acid-alcohol (10% formaldehyde, 5% glacial acetic acid, 50% ethanol) at 4°C overnight. The samples were dehydrated on ice with an ethanol-double distilled H₂O series (50, 70, 85, 95, 100% \times 2) and then embedded in a resin-ethanol series (50, 100%) at room temperature. The embedded ovaries were subsequently sectioned using a Leica RM2265 semithin slicer, followed by staining with 0.5% toluidine blue in a 0.1% sodium carbonate solution. The sections were imaged using an Olympus BX51 microscope, and histological and cellular observations were made using ImageJ. The red, blue and yellow lines in Fig. 1a indicate the inner ovary wall, the columella boundary and ovary wall thickness, respectively. The cell number and size were measured along these lines.

Fine-mapping

FS8.1 has been mapped to a 3.03-Mb interval between two markers, 11EP239 and 11EP249, on the long arm of chromosome 8 by using recombinant plants derived from hybrid progenies of *S. lycopersicum* cv. Rio Grande and a wild species, *S. pimpinellifolium* (accession no. LA1589, <https://tgrc.ucdavis.edu/>) (ref. 14). On the basis of previous mapping, we used a BC₄F₅ plant (13S140) to self-pollinate twice and obtained five recombinants from 14,080 progenies, which narrowed the region to 268.63 kb. There are three open reading frames in this region. Primers used for genotyping are listed in Supplementary Table 8.

Plasmid construction and plant transformation

For genetic complementation of M82, an *FS8.1* genomic fragment containing its 2,854-bp promoter and coding sequence from AC was cloned into a pCAMBIA1300-GFP vector⁵⁸ to generate a *proFS8.1::FS8.1-GFP* construct. For overexpression of *SIKRP2*, the coding DNA sequence

(CDS) of *SIKRP2* was cloned into a pENTR vector (Invitrogen), which was then recombined with the binary vector pGWB18 (ref. 59) to generate a *pro35S::SIKRP2-myc* construct. For overexpression of *SIGT-16*, the *SIGT-16* CDS was cloned into pENTR (Invitrogen), which was then recombined with the binary vector pK7FWG2.0 (ref. 60) to generate a *pro35S::SIGT-16-GFP* construct. All the constructs were transformed into M82 by *Agrobacterium tumefaciens* (GV3101)-mediated cotyledon explant transformation as described previously²³. The resulting transformants were selected on the basis of their resistance to hygromycin B or kanamycin. Homozygous T₂ transgenic plants were used for further experiments. All primers used for plasmid construction are listed in Supplementary Table 8.

CRISPR/Cas9-mediated mutations

Null mutations of *FS8.1*, *SIKRP2*, *SIGT-16* and *SIGT-34* were generated by the tomato U6 promoter-controlled CRISPR/Cas9 system as previously described²³. Briefly, two sets of primers containing guide RNA (gRNA) sequences of *FS8.1* were used in PCR to generate a tomato *U6-26-FS8.1-gRNA* cassette. The resulting *U6-26-FS8.1-gRNA* cassette was then cloned into the binary vector pTX041 (ref. 23) to form a *pTX041-FS8.1* construct. *pTX041-SIKRP2*, *pTX041-SIGT-16* and *pTX041-SIGT-34* constructs were generated following the same protocol. The final binary vectors were subsequently transformed into IL8-1-1, AC or TB0249 by *A. tumefaciens* (GV3101)-mediated cotyledon explant transformation. The *slgt-16 slgt-34* double mutant line was generated by a CRISPR/Cas9 multiplex editing system on the basis of the endogenous transfer RNA (tRNA)-processing system^{61,62}. Briefly, a pGTR plasmid containing the gRNA scaffold with a tRNA sequence was used as a PCR template to generate gRNA-tRNA units. The resulting units were cloned into pTX041 to form a *pTX041-SIGT-16&SIGT-34* construct. The final construct was then transformed into IL8-1-1 and AC. An *slkrp1 slkrp2 slkrp4* triple mutant line and *fs8.1 sp* double mutant lines were generated following the same protocol. The CRISPR/Cas9-mediated mutations were genotyped by PCR amplification and DNA sequencing. Homozygous lines without Cas9 were identified for further experiments. The primers used for plasmid construction and genotyping are listed in Supplementary Table 8.

RNA extraction and gene expression analysis

Total RNA was extracted from each sample using TRIzol reagent (Invitrogen, 15596018) according to manufacturer instructions. The quality of the total RNA was determined using a NanoDrop spectrophotometer (Thermo Fisher). Each sample (2 μg) of total RNA was used to synthesize first-strand complementary DNA with a PrimeScript RT kit with gDNA eraser (RR0447A, Takara). RT-qPCR was performed using a Roche LightCycler 480 system with a SYBR Fast qPCR kit (KK4601, KAPA Biosystems). The expression levels of the target genes were normalized against that of *SIACTIN2*. The error bars represent the s.d.s of 3 biological replicates. Each replicate was collected from more than 10 pooled ovaries or other organs at the same stage. The primers used to quantify the gene expression levels are listed in Supplementary Table 8.

Phylogenetic tree construction and protein structure prediction

Phylogenetic trees were constructed with MEGA (v.7.0)⁶³ using the neighbour-joining method. The branches were compared with bootstrap support values from 500 replicates per node. Protein structures of *SIGT-16*, *FS8.1* and *fs8.1* were predicted by the artificial intelligence (AI) system, AlphaFold^{64,65}.

RNA-seq and data analysis

M82 and *Comp-2* ovaries were collected at 9 DBA for RNA isolation. Total RNA was extracted using TRIzol reagent (15596018, Invitrogen) and then treated with DNase I (M0303S, New England Biolabs). The quality of the total RNA was assessed using a NanoDrop spectrophotometer

and Agilent 2100 Bioanalyzer. For each sample, 1 µg of total RNA was used to construct Illumina sequencing libraries according to manufacturer instructions. The libraries were sequenced by staff at Biomarker Technologies via an Illumina HiSeq X-ten platform and preliminarily analysed using the BMKCloud platform (<http://www.biocloud.net/>). Low-quality reads and reads containing adapter sequences and poly-N sequences were removed from the raw data, thus resulting in more than 7 Gb of clean reads per library. The clean reads were mapped onto the tomato reference genome (SL2.5), which was downloaded from the Solanaceae (SOL) Genomics Network (<http://solgenomics.net/>) via TopHat2 (ref. 66) with default parameters. The gene expression levels were calculated as fragments per kilobase of transcript per million fragments mapped (FPKM) using StringTie⁶⁷. Genes that were differentially expressed between two groups were identified using the DEGseq R package⁶⁸, with an FC ≥ 1.5 and FDR-adjusted $P < 0.05$. GO enrichment analysis was implemented using Protein Analysis Through Evolutionary Relationships (PANTHER; <http://geneontology.org/>). GO term enrichment is shown using the most specific subclass in the enrichment analysis.

ChIP-qPCR assays

ChIP assays were performed as previously described^{69,70}, with slight modifications. Briefly, ovaries of *proFS8.1::FS8.1-GFP-2* or *SIGT-16-OE-1* plants at anthesis were collected and ground to a powder in liquid nitrogen. Four grams of powder of each sample was crosslinked with 1% (v/v) formaldehyde on ice for 10 min, followed by neutralization with 0.125 M glycine. The chromatin complex was isolated, resuspended in lysis buffer (50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, pH 7.5), 150 mM NaCl, 1 mM EDTA, 1% (w/v) SDS, 1% (v/v) Triton X-100, 0.1% (w/v) sodium deoxycholate and 1 mM phenylmethylsulfonyl fluoride) containing 1× protease inhibitor cocktail (05056489001, Roche) and sonicated to reduce the average DNA fragment size to ~500 bp. Then, 50 µl of sheared chromatin was saved as an input control. Anti-GFP antibody (AB290, Abcam, 1:750 dilution) was incubated together with Dynabeads Protein G (10004D, Invitrogen) at 4 °C for at least 6 h and then added to the remaining chromatin, which was subsequently incubated overnight at 4 °C. The immunoprecipitated chromatin–protein complex was sequentially washed in low-salt buffer (20 mM Tris-HCl (pH 8.0), 150 mM NaCl, 2 mM EDTA, 0.2% (w/v) SDS and 0.5% (v/v) Triton X-100), high-salt buffer (20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 2 mM EDTA, 0.2% (w/v) SDS and 0.5% (v/v) Triton X-100), LiCl buffer (10 mM Tris-HCl (pH 8.0), 25 mM LiCl, 1 mM EDTA, 0.5% (w/v) Nonidet P-40 and 0.5% (w/v) sodium deoxycholate) and TE buffer (10 mM Tris-HCl (pH 8.0) and 1 mM EDTA). After washing, the immunoprecipitated chromatin was eluted with elution buffer (1% SDS and 100 mM NaHCO₃). The protein–DNA crosslinks were reversed by incubating the immunoprecipitated complexes in 20 µl 5 M NaCl at 65 °C overnight. The DNA was recovered using a QIAquick PCR Purification kit (28106, Qiagen) and analysed by qPCR. The ChIP signals were displayed as the percentage of precipitated DNA relative to that of input DNA. The fold enrichment in selected regions was normalized against the non-specific binding region of the *SIACTIN2* promoter. The error bars represent the s.d.s of 3 biological replicates. Each replicate was collected from more than 100 pooled ovaries at the same stage. The primers used for qPCR are listed in Supplementary Table 8.

EMSA

EMSA were performed as described previously^{71,72}. Briefly, *FS8.1* CDS, *SIGT-16* CDS and *SIGT-16* derivatives were cloned into a pMAL-c2X vector (N8076S, New England Biolabs). The recombinant MBP-*FS8.1*, MBP-*SIGT-16*, MBP-*SIGT-16*^{ΔTD1} and MBP-*SIGT-16*^{ΔTD2} fusion proteins were expressed in *Escherichia coli* Rosetta (DE3) cells and affinity-purified with amylose resin (E8021L, New England Biolabs) according to manufacturer instructions. Oligonucleotide probes were synthesized and

labelled with biotin at the 5' end by Invitrogen. EMAs were performed using a chemiluminescent EMSA kit (GS009, Beyotime). Biotin-labelled probes were incubated together with MBP-fusion proteins at room temperature for 20 min, and free and bound probes were separated in an acrylamide gel. A labelled probe incubated with MBP was used as a negative control. Unlabelled probes were used as competitors. The probes used for EMSA are listed in Supplementary Table 8.

LCI assays

LCI assays were performed as described previously⁷³. The CDSs of *FS8.1*, *fs8.1* and *SIGT-16* were cloned into a pCAMBIA-35S-nLuc vector, and the CDSs of *FS8.1*, *SIGT-16*, *SIGT-34*, *SIGT-30*, *SIGT-26* and *SIGT-26L* were cloned into a pCAMBIA-35S-cLuc vector. The primers used for plasmid construction are listed in Supplementary Table 8. *A. tumefaciens* GV3101 transformed with different constructs were incubated, collected and resuspended in a 10 mM MgCl₂ solution including 0.2 mM acetosyringone to a final concentration of optical density (OD)₆₀₀ = 1.5. After 2 h of incubation at room temperature, equal volumes of *Agrobacterium* suspensions carrying the indicated constructs were mixed and co-infiltrated into *N. benthamiana* leaves with a needleless syringe. The infiltrated plants were incubated at 23 °C for 72 h under a 16-h light/8-h dark photoperiod before the LUC activity was measured. A low-light-cooled charge-coupled device (CCD) imaging apparatus (NightOWL II LB983 with Indigo software) was used to capture the LUC image. The leaves were sprayed with a 0.5 mM luciferin solution and placed in darkness for 3 min before luminescence detection. Data from at least 9 biological replicates were collected.

In vitro pull-down assays

To produce MBP-*FS8.1*, MBP-*SIGT-16*, MBP-*SIGT-34*, MBP-*SIGT-16*^{ΔTD1}, MBP-*SIGT-16*^{ΔTD2} and MBP-*SIGT-16*^{ΔHD} fusion proteins, the CDSs of *FS8.1*, *SIGT-16* and *SIGT-34* and the derivatives of *SIGT-16* were cloned into a pMAL-c2X vector (N8076S New England Biolabs). To produce GST-*FS8.1* and GST-*SIGT-16* fusion proteins, the CDSs of *FS8.1* and *SIGT-16* were cloned into a pGEX-4T-3 vector (27-4583-01, GE Healthcare). The primers used for plasmid construction are listed in Supplementary Table 8. The resulting recombinant vectors were transformed into *E. coli* Rosetta (DE3) cells, and the fusion proteins were purified using amylose resin (E8021, New England Biolabs,) or GST Bind Resin (17-0756-01, GE Healthcare) according to manufacturer instructions.

To detect *FS8.1-SIGT-16* and *FS8.1-SIGT-34* interactions, 10 µl of amylose resin bound to 1 µg of MBP-*SIGT-16* or MBP-*SIGT-34* was incubated together with 1 µg of GST-*FS8.1* in 1 ml of reaction buffer (25 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM dithiothreitol and Roche protease inhibitor cocktail) at 4 °C for 1 h. Subsequently, the beads were collected and washed three times in washing buffer (25 mM Tris-HCl (pH 7.5), 150 mM NaCl and 1 mM dithiothreitol). Then, the samples were denatured using sodium dodecyl sulfate (SDS) protein loading buffer and separated using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The GST fusion proteins were detected by immunoblotting with anti-GST antibody (M20007, Abmart, 1:3,000 dilution); purified MBP was used as a negative control. One microgram each of MBP and MBP-fusion proteins were separated by SDS-PAGE, and the staining of polyacrylamide gels with Coomassie Brilliant Blue R 250 was used as a loading control. To investigate whether *FS8.1* and *SIGT-16* form homodimers or heterodimers with each other and to determine which domains are involved in dimerization, subsequent manipulations similar to those described above were performed.

To determine whether *FS8.1* affects *SIGT-16-SIGT-16* homodimer formation, 1 µg each of purified MBP-*SIGT-16* and GST-*SIGT-16* proteins was added to each sample. Purified GST-*FS8.1* protein was added in accordance with a concentration gradient. Amylose resin was used to pull down the proteins. Subsequent procedures performed and buffers used were the same as those described above.

Transient expression assays

Transient expression assays were performed on *N. benthamiana* leaves as previously described⁵⁸. The 2,053-bp promoter sequence of *SIKRP2* was amplified from genomic DNA and cloned into a pGreenII 0800-LUC vector³¹ for use as a reporter. The *Renilla* (*REN*) LUC gene under the control of the cauliflower 35S promoter in a pGreenII 0800-LUC vector was used as the internal control. The CDSs of *FS8.1*, *fs8.1* and *SIGT-16* under the control of the 35S promoter were cloned into a pCAMBIA-35S-cLuc vector, and the resulting vectors were used as effectors. The pCAMBIA-35S-cLuc vector was used as a negative control. The primers used for plasmid construction are listed in Supplementary Table 8. *A. tumefaciens* GV3101 transformed with the different constructs were incubated, collected and resuspended in a 10 mM MgCl₂ solution including 0.2 mM acetosyringone to a final concentration of OD₆₀₀ = 1.5. After 2 h of incubation at room temperature, equal volumes of *Agrobacterium* suspensions carrying the indicated constructs were mixed and co-infiltrated into *N. benthamiana* leaves with a needleless syringe. The infiltrated plants were incubated at 23 °C for 48 h under a 16-h light/8-h dark photoperiod before luminescence detection. Firefly LUC and REN activities were measured using a Dual-LUC Reporter Assay system (Promega) following manufacturer instructions, and LUC:REN ratios were calculated and presented. Data from 4 independent biological replicates were collected, with error bars used to represent the s.d.s.

Genotyping of tomato germplasm

A derived cleaved amplified polymorphic sequence (dCAPS) marker, FS8.1-dCAPS, was developed to detect the *fs8.1* mutation. After amplification, the products were digested with *Hind*III and separated on a 3% (w/v) agarose gel. The null allele of *FS8.1* yielded a fragment of 174 bp, whereas the WT allele yielded a fragment of 154 bp. The primers used for genotyping are listed in Supplementary Table 8.

Content analysis of fruit quality-related metabolites

Total soluble solids (°Brix) were measured with a digital Brix refractometer (ATAGO PAL-BX/ACID3). In the corresponding figures, error bars are used to represent the s.d.s of 4 biological replicates. For each replicate, the fruit juice of 3 red ripe fruits was pooled and used for measurements.

Sugar and acid analyses were performed as described previously¹⁸, with modifications. More than 3 red ripe fruits were collected from each genotype, and the pericarps of the mixed fruits were ground in liquid nitrogen. For saccharide analysis, 800 mg of ground powder was diluted in 500 µl of water. After sonication and centrifugation, the samples were filtered through a 0.22 µm polyethersulfone ultrafiltration membrane. The saccharide contents were measured using a high-performance liquid chromatograph (LC-20AT, Shimadzu). A Waters BEH Amide 5 µm column was used as the analytical column (250 × 4.6 mm). The mobile phase consisted of acetonitrile as solvent A and 1 mg ml⁻¹ ammonium hydroxide as solvent B. The temperatures of the column and autosampler were 40 °C and 4 °C, respectively. Each saccharide was separated under isocratic elution (A:B = 7:3) under a flow rate of 0.8 ml min⁻¹. In the corresponding figures, the error bars represent the s.d.s of 3 or 4 biological replicates. For the acid analysis, 500 mg of ground powder was diluted in 500 µl 0.2 mol l⁻¹ metaphosphoric acid. After sonication and centrifugation, the samples were filtered through a 0.22 µm polyethersulfone ultrafiltration membrane. Acid contents were measured using a high-performance liquid chromatograph (LC-20AT, Shimadzu). A CNW Athena C18-WP 5 µm column was used as the analytical column (250 × 4.6 mm). The mobile phase consisted of water with 1 mg ml⁻¹ phosphoric acid as solvent A and methanol as solvent B. The temperatures of the column and autosampler were 30 °C and 4 °C, respectively. Each acid was separated by increasing the solvent B concentration from 2% to 80% during 3.5 min after the first 7.5 min of the run at 2% under a flow rate of 0.7 ml min⁻¹, followed by washing with 2% solvent B for 8 min.

In the corresponding figures, the error bars represent the s.d.s of 4 biological replicates.

Lycopene analysis was performed as previously described⁷⁴, with modifications. Briefly, 500 mg of ground powder was diluted in 15 ml of methanol and ethyl acetate (methanol:ethyl acetate = 1:9 (v/v)) including 0.1% butylated hydroxytoluene. After sonication and centrifugation, the samples were dried under low temperature and dissolved in 1 ml of ethyl acetate. Then, the samples were filtered through a 0.22 µm polyethersulfone ultrafiltration membrane. The lycopene contents were measured using a high-performance liquid chromatograph (LC-20AT, Shimadzu). A C18 5 µm column was used as the analytical column (250 × 4.6 mm). The mobile phase consisted of methanol as solvent A and ethyl acetate as solvent B. The temperatures of the column and autosampler were 20 °C and 4 °C, respectively. Each acid was separated by increasing the solvent B concentration from 6% to 90% during 4 min after the first 1 min of the run at 6% under a flow rate of 1 ml min⁻¹, followed by washing with 6% solvent B for 7 min. In the corresponding figures, the error bars represent the s.d.s of 4 biological replicates.

Quantification and statistical analysis

The significance of differences was evaluated using two-tailed Student's *t*-tests. For protein quantification, band intensities were quantified using ImageJ software.

Reporting summary

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

Data availability

The RNA-seq data have been deposited in the Genome Sequence Archive (GSA; <https://ngdc.cncb.ac.cn/gsa/>) at the Beijing Institute of Genomics (BIG) Data Center, Chinese Academy of Sciences, under accession number CRA008400. The sequence data of the following genes (and their accession numbers) in this article can be found in the Sol Genomics Network (SGN): *FS8.1* (Solyc08g061910), *SIGT-16* (Solyc04g071360), *SIGT-34* (Solyc12g056510), *SIGT-30* (Solyc11g005380), *SIGT-26* (Solyc09g009250), *SIGT-26L* (Solyc10g083567), *SIKRP1* (Solyc03g044480), *SIKRP2* (Solyc02g090680), *SIKRP4* (Solyc12g098310) and *SP* (Solyc06g074350). Materials used in this study are available upon request. Source data are provided with this paper.

References

1. The Tomato Genome Consortium. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* **485**, 635–641 (2012).
2. van der Knaap, E. & Tanksley, S. D. The making of a bell pepper-shaped tomato fruit: identification of loci controlling fruit morphology in Yellow Stuffer tomato. *Theor. Appl. Genet.* **107**, 139–147 (2003).
3. García-Valverde, V., Navarro-González, I., García-Alonso, J. & Periago, M. J. Antioxidant bioactive compounds in selected industrial processing and fresh consumption tomato cultivars. *Food Bioproc. Tech.* **6**, 391–402 (2013).
4. Lukyanenko, A. N. in *Genetic Improvement of Tomato* (ed. Kalloo, G.) 213–230 (Springer, 1991).
5. Tanksley, S. D. The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. *Plant Cell* **16**, S181–S189 (2004).
6. Webb, R. E. & Bruce, W. M. in *Yearbook of Agriculture, 1968: Science for Better Living* 103–107 (U. S. Department of Agriculture, 1968).
7. Grandillo, S., Ku, H.-M. & Tanksley, S. D. Characterization of *fs8.1*, a major QTL influencing fruit shape in tomato. *Mol. Breed.* **2**, 251–260 (1996).

8. Gonzalo, M. J. & van der Knaap, E. A comparative analysis into the genetic bases of morphology in tomato varieties exhibiting elongated fruit shape. *Theor. Appl. Genet.* **116**, 647–656 (2008).
9. Rodríguez, G. R. et al. Distribution of SUN, OVATE, LC, and FAS in the tomato germplasm and the relationship to fruit shape diversity. *Plant Physiol.* **156**, 275–285 (2011).
10. van der Knaap, E. et al. What lies beyond the eye: the molecular mechanisms regulating tomato fruit weight and shape. *Front. Plant Sci.* **5**, 227 (2014).
11. Paran, I. & van der Knaap, E. Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. *J. Exp. Bot.* **58**, 3841–3852 (2007).
12. Grandillo, S. & Tanksley, S. D. QTL analysis of horticultural traits differentiating the cultivated tomato from the closely related species *Lycopersicon pimpinellifolium*. *Theor. Appl. Genet.* **92**, 935–951 (1996).
13. Ku, H. M., Grandillo, S. & Tanksley, S. D. *fs8.1*, a major QTL, sets the pattern of tomato carpel shape well before anthesis. *Theor. Appl. Genet.* **101**, 873–878 (2000).
14. Sun, L. et al. Candidate gene selection and detailed morphological evaluations of *fs8.1*, a quantitative trait locus controlling tomato fruit shape. *J. Exp. Bot.* **66**, 6471–6482 (2015).
15. Frary, A., Fulton, T. M., Zamir, D. & Tanksley, S. D. Advanced backcross QTL analysis of a *Lycopersicon esculentum* × *L. pennellii* cross and identification of possible orthologs in the Solanaceae. *Theor. Appl. Genet.* **108**, 485–496 (2004).
16. Zhou, D. X. Regulatory mechanism of plant gene transcription by GT-elements and GT-factors. *Trends Plant Sci.* **4**, 210–214 (1999).
17. Kaplan-Levy, R. N., Brewer, P. B., Quon, T. & Smyth, D. R. The trihelix family of transcription factors—light, stress and development. *Trends Plant Sci.* **17**, 163–171 (2012).
18. Li, R. et al. *FIS1* encodes a GA2-oxidase that regulates fruit firmness in tomato. *Nat. Commun.* **11**, 5844 (2020).
19. Eshed, Y. & Zamir, D. An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. *Genetics* **141**, 1147–1162 (1995).
20. Lippman, Z. B., Semel, Y. & Zamir, D. An integrated view of quantitative trait variation using tomato interspecific introgression lines. *Curr. Opin. Genet. Dev.* **17**, 545–552 (2007).
21. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
22. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
23. Deng, L. et al. Efficient generation of pink-fruited tomatoes using CRISPR/Cas9 system. *J. Genet. Genomics* **45**, 51–54 (2018).
24. De Veylder, L. et al. Functional analysis of cyclin-dependent kinase inhibitors of *Arabidopsis*. *Plant Cell* **13**, 1653–1668 (2001).
25. Verkest, A. et al. The cyclin-dependent kinase inhibitor KRP2 controls the onset of the endoreduplication cycle during *Arabidopsis* leaf development through inhibition of mitotic CDKA;1 kinase complexes. *Plant Cell* **17**, 1723–1736 (2005).
26. Zhao, X. et al. A general G1/S-phase cell-cycle control module in the flowering plant *Arabidopsis thaliana*. *PLoS Genet.* **8**, e1002847 (2012).
27. Komaki, S. & Sugimoto, K. Control of the plant cell cycle by developmental and environmental cues. *Plant Cell Physiol.* **53**, 953–964 (2012).
28. Kumar, N. et al. Functional conservation in the SIAMESE-RELATED family of cyclin-dependent kinase inhibitors in land plants. *Plant Cell* **27**, 3065–3080 (2015).
29. Hiratsuka, K., Wu, X. D., Fukuzawa, H. & Chua, N. H. Molecular dissection of Gt-1 from *Arabidopsis*. *Plant Cell* **6**, 1805–1813 (1994).
30. Yu, C., Cai, X., Ye, Z. & Li, H. Genome-wide identification and expression profiling analysis of trihelix gene family in tomato. *Biochem. Biophys. Res. Commun.* **468**, 653–659 (2015).
31. Hellens, R. P. et al. Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. *Plant Methods* **1**, 13 (2005).
32. Razifard, H. et al. Genomic evidence for complex domestication history of the cultivated tomato in Latin America. *Mol. Biol. Evol.* **37**, 1118–1132 (2020).
33. Brandt, J. A. & French, B. C. Mechanical harvesting and the California tomato industry: a simulation analysis. *Am. J. Agric. Econ.* **65**, 265–272 (1983).
34. Scott, J. W. in *Report of the Tomato Genetics Cooperative* (ed. Scott, J. W.) 5–13 (University of Florida, 2014).
35. Yeager, A. F. Determinate growth in the tomato. *J. Hered.* **18**, 263–265 (1927).
36. Pnueli, L. et al. 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 (1998).
37. MacArthur, J. W. Inherited characters in the tomato: I. The self-pruning habit. *J. Hered.* **23**, 395–396 (1932).
38. Rodriguez, G. R., Kim, H. J. & van der Knaap, E. Mapping of two suppressors of OVATE (sov) loci in tomato. *Heredity* **111**, 256–264 (2013).
39. Gao, L. et al. The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nat. Genet.* **51**, 1044–1051 (2019).
40. Lin, T. et al. Genomic analyses provide insights into the history of tomato breeding. *Nat. Genet.* **46**, 1220–1226 (2014).
41. Garg, N., Cheema, D. S. & Pathak, D. Heterosis breeding in tomato involving rin, nor and alc alleles: a review of literature. *Adv. Hort. Sci.* **22**, 54–62 (2008).
42. Kitagawa, M. et al. Characterization of tomato fruit ripening and analysis of gene expression in F1 hybrids of the *ripening inhibitor* (*rin*) mutant. *Physiol. Plant.* **123**, 331–338 (2005).
43. Osorio, S. et al. Genetic and metabolic effects of ripening mutations and vine detachment on tomato fruit quality. *Plant Biotechnol. J.* **18**, 106–118 (2020).
44. Wang, R. et al. The *rin*, *nor* and *Cnr* spontaneous mutations inhibit tomato fruit ripening in additive and epistatic manners. *Plant Sci.* **294**, 110436 (2020).
45. Adaskaveg, J. A., Silva, C. J., Huang, P. & Blanco-Ulalte, B. Single and double mutations in tomato ripening transcription factors have distinct effects on fruit development and quality traits. *Front. Plant Sci.* **12**, 647035 (2021).
46. Chen, K., Wang, Y., Zhang, R., Zhang, H. & Gao, C. CRISPR/Cas genome editing and precision plant breeding in agriculture. *Annu. Rev. Plant Biol.* **70**, 667–697 (2019).
47. Yin, K., Gao, C. & Qiu, J. L. Progress and prospects in plant genome editing. *Nat. Plants* **3**, 17107 (2017).
48. Kwon, C. T. et al. Rapid customization of Solanaceae fruit crops for urban agriculture. *Nat. Biotechnol.* **38**, 182–188 (2020).
49. Uluisik, S. et al. Genetic improvement of tomato by targeted control of fruit softening. *Nat. Biotechnol.* **34**, 950–952 (2016).
50. Shi, Y. et al. A tomato LATERAL ORGAN BOUNDARIES transcription factor, SILOB1, predominantly regulates cell wall and softening components of ripening. *Proc. Natl Acad. Sci. USA* **118**, e2102486118 (2021).
51. Smith, D. L., Abbott, J. A. & Gross, K. C. Down-regulation of tomato beta-galactosidase 4 results in decreased fruit softening. *Plant Physiol.* **129**, 1755–1762 (2002).
52. Moctezuma, E., Smith, D. L. & Gross, K. C. Antisense suppression of a beta-galactosidase gene (TB G6) in tomato increases fruit cracking. *J. Exp. Bot.* **54**, 2025–2033 (2003).

53. Huang, B. et al. Interaction of two MADS-box genes leads to growth phenotype divergence of all-flesh type of tomatoes. *Nat. Commun.* **12**, 6892 (2021).
54. Zhang, J. et al. An AGAMOUS MADS-box protein, SLMBP3, regulates the speed of placenta liquefaction and controls seed formation in tomato. *J. Exp. Bot.* **70**, 909–924 (2019).
55. Cantu, D. et al. The intersection between cell wall disassembly, ripening, and fruit susceptibility to *Botrytis cinerea*. *Proc. Natl Acad. Sci. USA* **105**, 859–864 (2008).
56. Jiang, F. et al. Disassembly of the fruit cell wall by the ripening-associated polygalacturonase and expansin influences tomato cracking. *Hortic. Res.* **6**, 17 (2019).
57. Rodríguez, G. R. et al. Tomato Analyzer: a useful software application to collect accurate and detailed morphological and colorimetric data from two-dimensional objects. *J. Vis. Exp.* **16**, 1856 (2010).
58. You, Y., Zhai, Q., An, C. & Li, C. LEUNIG_HOMOLOG mediates MYC2-dependent transcriptional activation in cooperation with the coactivators HAC1 and MED25. *Plant Cell* **31**, 2187–2205 (2019).
59. Nakagawa, T. et al. Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *J. Biosci. Bioeng.* **104**, 34–41 (2007).
60. Karimi, M., Inze, D. & Depicker, A. GATEWAY vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci.* **7**, 193–195 (2002).
61. Xie, K., Minkenberg, B. & Yang, Y. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc. Natl Acad. Sci. USA* **112**, 3570–3575 (2015).
62. Yang, T. et al. Recoloring tomato fruit by CRISPR/Cas9-mediated multiplex gene editing. *Hortic. Res.* **10**, uhac214 (2023).
63. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).
64. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
65. Varadi, M. et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444 (2022).
66. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
67. Pertea, M. et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
68. Wang, L., Feng, Z., Wang, X., Wang, X. & Zhang, X. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* **26**, 136–138 (2010).
69. Sun, C. et al. A transcriptional network promotes anthocyanin biosynthesis in tomato flesh. *Mol. Plant* **13**, 42–58 (2020).
70. Liu, Y. et al. MYC2 regulates the termination of jasmonate signaling via an autoregulatory negative feedback loop. *Plant Cell* **31**, 106–127 (2019).
71. Du, M. et al. MYC2 orchestrates a hierarchical transcriptional cascade that regulates jasmonate-mediated plant immunity in tomato. *Plant Cell* **29**, 1883–1906 (2017).
72. Deng, L. et al. Tomato MED25 regulates fruit ripening by interacting with EIN3-like transcription factors. *Plant Cell* **35**, koac349 (2022).
73. Chen, H. et al. Firefly luciferase complementation imaging assay for protein–protein interactions in plants. *Plant Physiol.* **146**, 368–376 (2008).
74. Powell, A. L. et al. Uniform ripening encodes a Golden 2-like transcription factor regulating tomato fruit chloroplast development. *Science* **336**, 1711–1715 (2012).

Acknowledgements

We thank J. Li for assistance with the fruit CR measurements. This work was supported by the National Key Research and Development Program of China (2022YFF1003000 and 2021YFF1000103 to L.D.), the National Natural Science Foundation of China (31991183 to C.L., 32072582 to L.D., U22A20459 to C.L. and 31772319 to L.S.) and the Beijing Joint Research Program for Germplasm Innovation and New Variety Breeding (G20220628003 to C.L.).

Author contributions

C.L. conceived and supervised the project. Q.Z. and L.D. performed most of the experiments. J.C., G.R.R., Z.C., T.Y., H.Z., Y.T., S.H. and L.S. performed the genotyping of the tomato germplasm. Q.Z., C.S., H.Z. and H.J. performed the plant transformations. T.Y. and C.-B.L. helped grow the plants. D.F. and E.v.d.K. reviewed and edited the manuscript. With input from all authors, C.L., L.D. and Q.Z. wrote the manuscript. All authors read and approved the content of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41477-023-01522-w>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41477-023-01522-w>.

Correspondence and requests for materials should be addressed to Chuanyou Li.

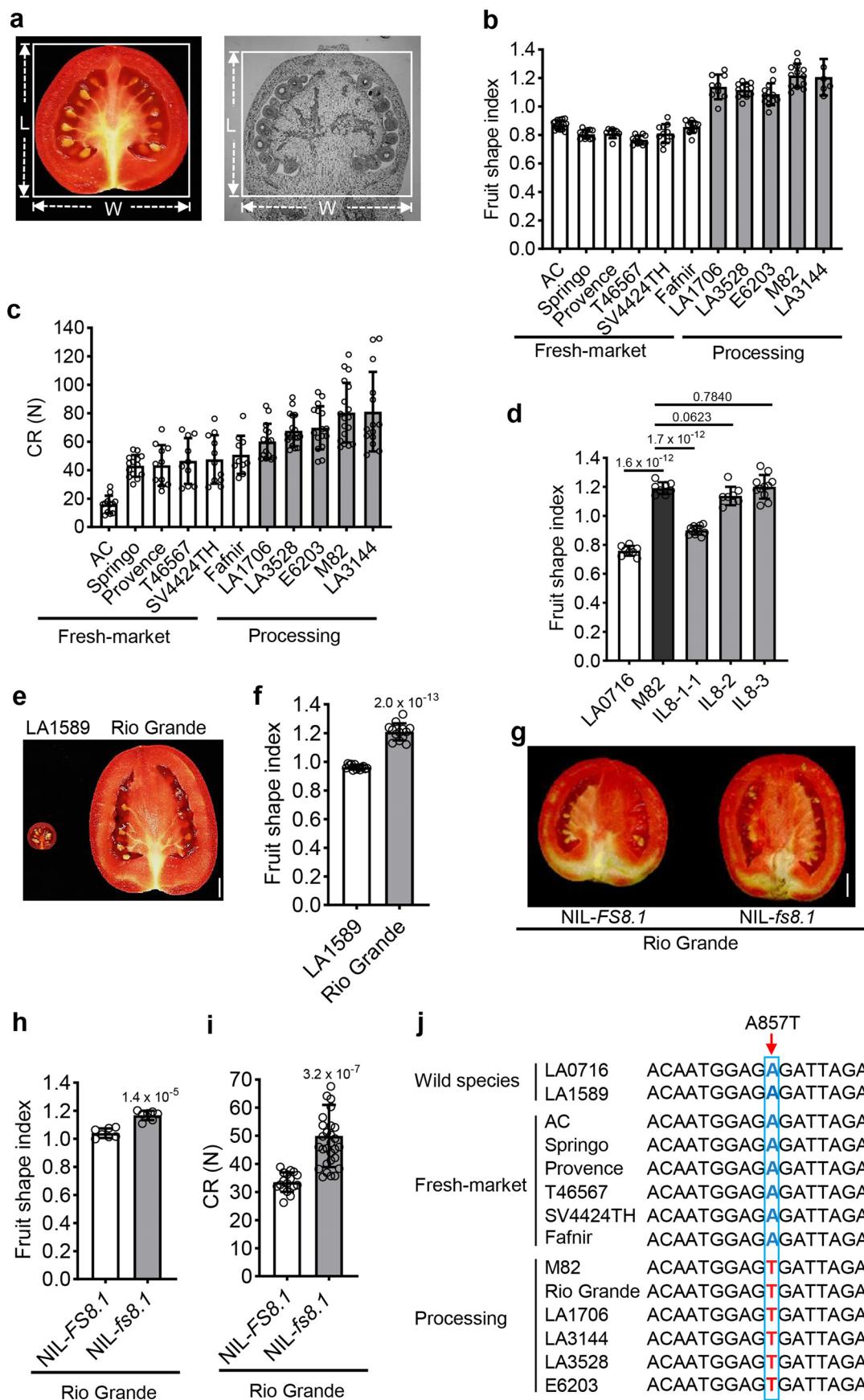
Peer review information *Nature Plants* thanks James Giovannoni and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

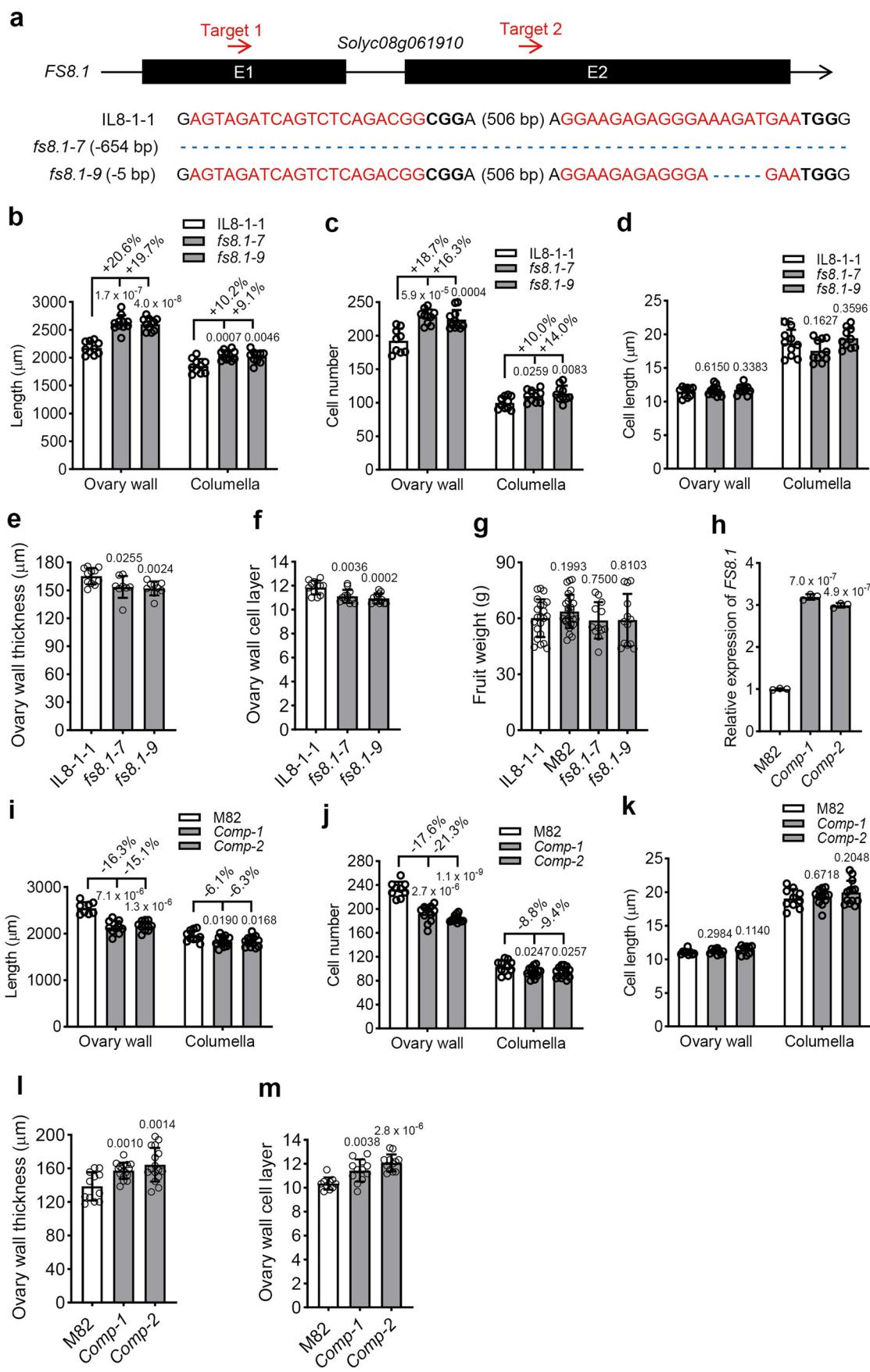
© The Author(s), under exclusive licence to Springer Nature Limited 2023, corrected publication 2024



Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | *fs8.1*-mediated shape change leads to increased fruit firmness. a, Fruit and ovary shape index measurements via Tomato Analyzer version 4.0 and ImageJ (NIH). The fruit and ovary shape indexes are defined as the ratios of the maximum height length (L) to maximum width (W) of a fruit or an ovary. b,c, Fruit shape indexes (b) and CR (c) of different fresh-market and processing tomatoes. In b, $n = 16, 12, 9, 11, 11, 12, 9, 13, 13, 13, 6$ fruits from left to right. In c, $n = 12, 14, 11, 10, 11, 10, 15, 18, 16, 18, 14$ fruits from left to right. d, Fruit shape indexes of LA0716, M82 and NILs ($n = 8, 8, 11, 8, 11$ fruits from left to right). e, Ripe fruits of LA1589 and Rio Grande. Bar = 1 cm. f, Fruit shape index of LA1589

and Rio Grande ($n = 13$ fruits). g, Ripe fruits of *fs8.1* NILs in the background of Rio Grande. Bar = 1 cm. h, Fruit shape indexes of *fs8.1* NILs in the background of Rio Grande ($n = 7$ fruits). i, CR of *fs8.1* NILs in the background of Rio Grande. $n = 17$ (for NIL-*fs8.1*) and 31 (for NIL-*fs8.1*) fruits. j, Sequence analysis showing the *SolyC08g061910*^{A857T} mutation in different fresh-market or processing tomato accessions. Bars represent the means \pm SD. The significance of the difference was evaluated by two-tailed Student's t tests, the exact *P* value is indicated on the graph.



Extended Data Fig. 2 | See next page for caption.

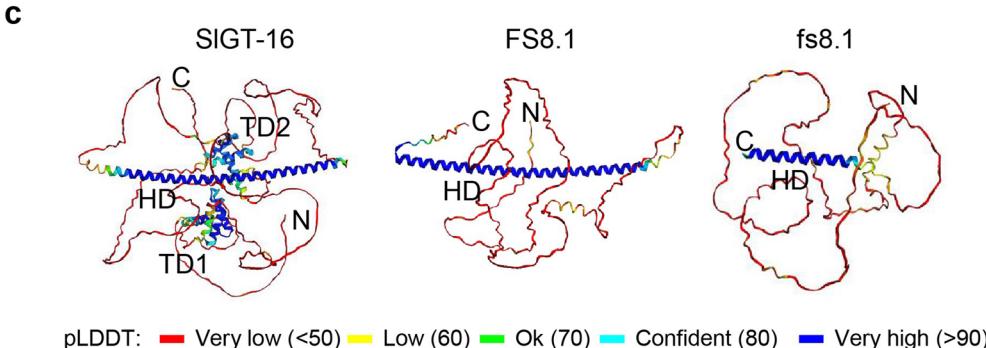
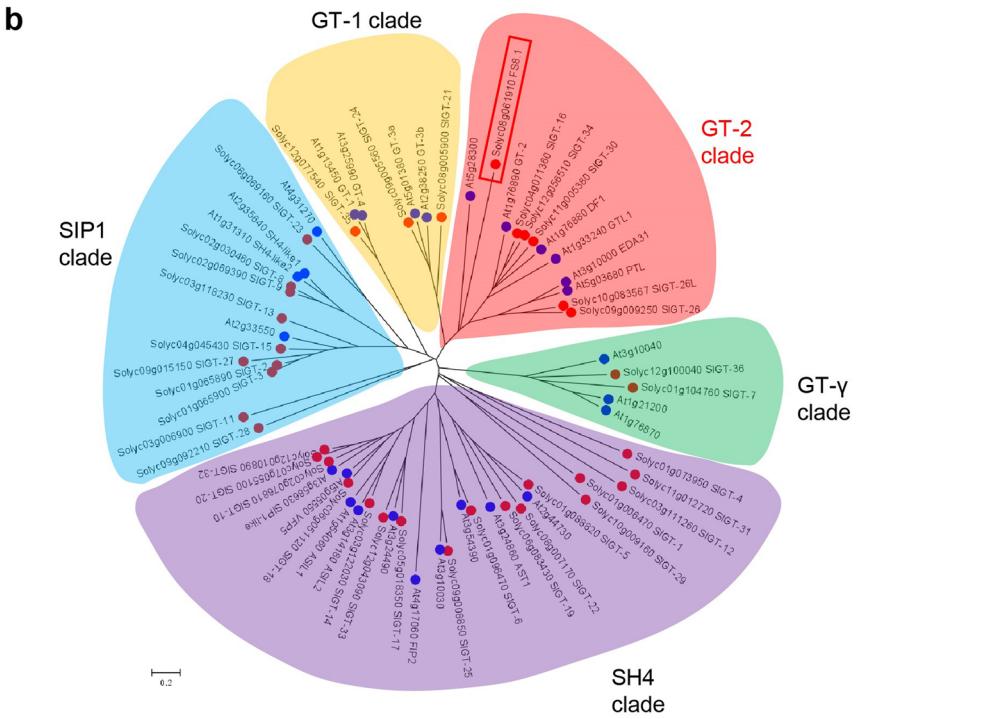
Extended Data Fig. 2 | Genetic validation of the *FS8.1* candidate gene.

a, Generation of *fs8.1* mutants in the IL8-1-1 background. The sgRNA targets and protospacer adjacent motifs (PAMs) are highlighted in red and bold fonts, respectively. The blue dashes indicate deletions, and the numbers indicate the numbers of nucleotides involved. b-f, Histological and cellular observations of anthesis ovaries of IL8-1-1 and *fs8.1* mutants. The cell number and cell size were measured along the red, blue and yellow lines indicated in Fig. 1a. IOWL, inner ovary wall length; OWT, ovary wall thickness; CL, columella length. In b, $n = 10$, 11, 11, 10, 11, 10 ovaries from left to right. In c, $n = 9, 10, 11, 10, 10, 10$ ovaries from left to right. In d, $n = 10, 12, 11, 10, 11, 10$ ovaries from left to right. In e, $n = 11, 8, 9$ ovaries from left to right. In f, $n = 12$ ovaries. g, Fruit weight of IL8-1-1 and *fs8.1*

mutants ($n = 22, 27, 12, 13$ fruits from left to right). h, RT-qPCR results showing *SolyC08g061910* expression in 9 DBA ovaries of the indicated genotypes ($n = 3$ independent biological replicates). i-m, Histological and cellular observations of the anthesis ovaries of M82 and *Comp* lines. The cell number and cell size were measured along the red, blue and yellow lines indicated in Fig. 1a. IOWL, inner ovary wall length; OWT, ovary wall thickness; CL, columella length. In i, $n = 8, 10, 11, 10, 13, 13$ ovaries from left to right. In j, $n = 9, 11, 11, 11, 14, 13$ ovaries from left to right. In k, $n = 9, 10, 9, 10, 13, 13$ ovaries from left to right. In l, $n = 12, 16, 16$ ovaries from left to right. In m, $n = 10, 12, 12$ ovaries from left to right. Bars represent the means \pm SD. The significance of the difference was evaluated by two-tailed Student's t tests, the exact *P* value is indicated on the graph.

a

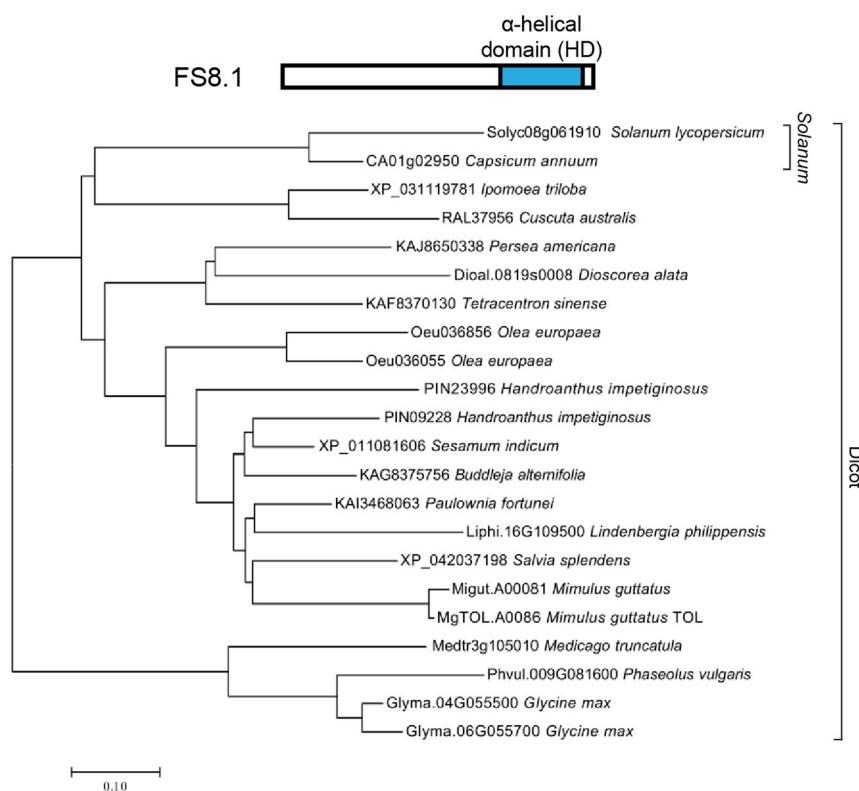
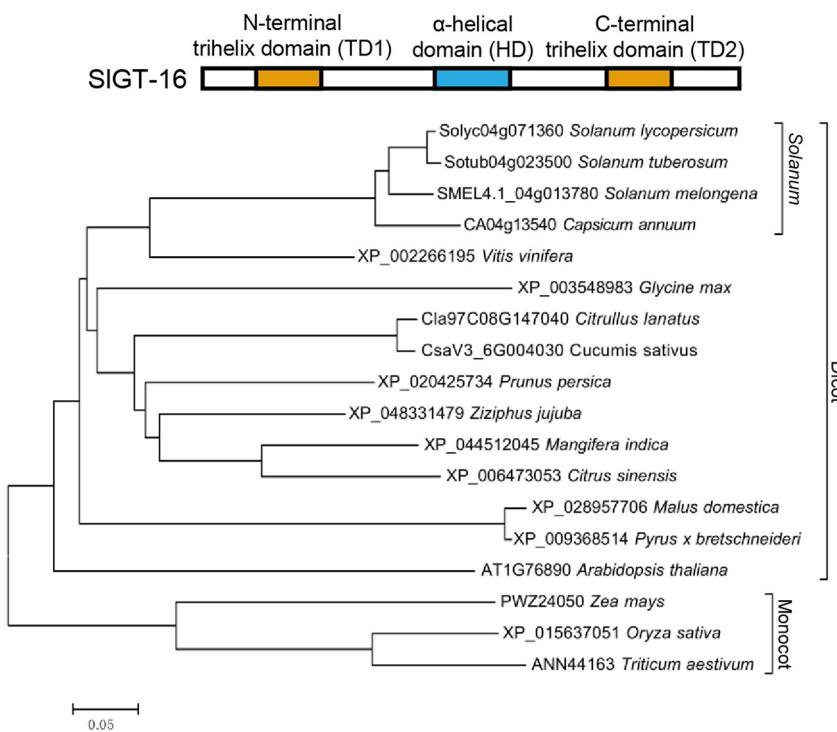
SIGT-16	MLGVSGLVLSSGGGDNPESGGAGSGGGS-----EIGLGGGS-----GGGGGSSGGFMTEDGERNS	G G N R W P R Q E T I	Trihelix																							
SIGT-34	MLGVSSSLIASSNTSITAGAAGDGAAISAAAPSQLAPPQEAPESGGSEGGGGGGDLSIGGEDEGRNS	G G N R W P R Q E T I																								
FS8.1	M Q S D Y E M G	G I H Q C E D E P R	Domain 1																							
fs8.1	M Q S D Y E M G	G I H Q C E D E P R																								
SIGT-16	A L L K I R S E M D V I F R D S S L K G P L W E E V S R K M A D L G F H R S S K K C K E K F E N V Y K Y H K R T K D G R A S K A D G K N Y R F F E Q L E A L E																									
SIGT-34	A L L K I R S E M D V V F K D S S L K G P L W E E V S R K M A D L G F H R S S K K C K E K F E N V Y K Y H K R T K D G R A S K A D G K T Y R F F D Q L Q A L E																									
FS8.1	F M V E N N A S T S S P P Y P N Y H F P P T S H P I L Q Q I H S L P I T O H F F P Y Q H A H -----Y R S M S E E I R V D Q S Q T A E L V A F F A S E I R G																									
fs8.1	F M V E N N A S T S S P P Y P N Y H F P P T S H P I L Q Q I H S L P I T O H F F P Y Q H A H -----Y R S M S E E I R V D Q S Q T A E L V A F F A S E I R G																									
SIGT-16	N I T S H H S L M P V P S S N T R P P P P L E A T P I N M A M P M A S S N V Q V T A S Q G T -----I P H N V T I S S A P P P N S L F A P S H O N A P S S S																									
SIGT-34	N N -----P S S H R S G N P P P L A A T P I T M A M P M R S G N N S A N P P M P T P T P T P Q N H N H F F S V S Q K S V V T G A A Q P A V M T A P																									
FS8.1	G Q E -----D A L I R G S E R Y C T Q P R Q T C V A V W Q N Q E D S A M K Q P V W K G E Y S N G N A I E K N K Q E D E E L Y S L E E T N K R R V V																									
fs8.1	G Q E -----D A L I R G S E R Y C T Q P R Q T C V A V W Q N Q E D S A M K Q P V W K G E Y S N G N A I E K N K Q E D E E L Y S L E E T N K R R V V																									
SIGT-16	P V P L P P P S Q Q P S P Q P A V N P I N N I P O -----Q V N A S A M S Y S T S S S T S S D E D I O R R H K K K R K W K D Y F F E K F T K D V I N K O E E S H R R																									
SIGT-34	A L P L S Q V P I G N N N L N Q M H R P G N T T T K T S F L S N S T S S S D E D I O R R H K K K R K W K D Y F F E K F T K D V I N K O E E S H R R																									
FS8.1	F G E L E A I C I R G I A S E S A D N L P T N H N V T F P E L A L N E A M V N K M D N T M G K F H K R K R G K D E W G R F F K S L V K K L A N H G E D L Q R S																									
fs8.1	F G E L E A I C I R G I A S E S A D N L P T N H N V T F P E L A L N E A M V N K M D N T M G K F H K R K R G K D E W G R F F K S L V K K L A N H G E D L Q R S																									
SIGT-16	F L E K L L E K E H D R M V R E E A W K V E E M A R M N R E H D L L V Q E R A M A A A K D A A V I S F L L Q K I T E Q Q N I Q I P N S I N V G P P S A Q V Q I Q																									
SIGT-34	F L E T L E K R E R D R L M R E E A W R V Q E M A R L N R E H D L L V Q E R S M A A A K D A I I A F L L Q K I T E Q Q N T Q T P N S T N N T S P S P F P I A Q																									
FS8.1	L M E T M E R L D Q E R K E R E E L W R E K E L E K L Q N E E A R A H E R R L A S T R E A L V S C L E K L T G Q K I D F Q T F K I K E D E T																									
fs8.1	L M E T M E R L D Q E R K E R E E L W R E K E L E K L Q N E E A R A H E R R L A S T R E A L V S C L E K L T G Q K I D F Q T F K I K E D E T																									
SIGT-16	L P E N P L S A P V P T Q I Q P T T V T A A A P P Q P A P V P V S L P V T I P A P V P A L I P S L S L P L T P P V P S K N M E L V P K S D N G G D S Y S P A S																									
SIGT-34	I Q L K L S E K P F S T P P Q P Q -----P Q P S A T A V S L P M T I H T P T P A P P Q T L T L P V -----V S S K S L E --P P K S D N G G E N F S P A S																									
FS8.1																										
fs8.1																										
SIGT-16	S S R W P K A E V E A L I K L R T N L D V K Y Q E N G P K G P L W E E I S S G M K K I G Y N R N A K R C K E K W E N I N K Y F K K V K E S N K K R P E D S K T		Trihelix																							
SIGT-34	S S R W P K E E I E A L I S L R T C L D L K Y Q E N G P K G P L W E E I S S G M R K I G Y N R N A K R C K E K W E N I N K Y F K K V K E S N K K R P E D S K T		FS8.1			Domain 2	fs8.1				SIGT-16	C P Y F H Q L D A L Y K E K A K N P E T A S S T S F N P S F A L N P D N N Q M A P I M A R P E Q Q W P L P Q H -----H E S T T R I D H E N E S D N M D E D D H			SIGT-34	C P Y F H Q L E A L Y K E K A K L -----E P V P H N T T F G L T P Q N N P P P P -----P P I M A Q P E Q Q W P I P Q N Q L H Q Q N R D H H D N E S D S M D H D L E			FS8.1				fs8.1			
FS8.1			Domain 2																							
fs8.1																										
SIGT-16	C P Y F H Q L D A L Y K E K A K N P E T A S S T S F N P S F A L N P D N N Q M A P I M A R P E Q Q W P L P Q H -----H E S T T R I D H E N E S D N M D E D D H																									
SIGT-34	C P Y F H Q L E A L Y K E K A K L -----E P V P H N T T F G L T P Q N N P P P P -----P P I M A Q P E Q Q W P I P Q N Q L H Q Q N R D H H D N E S D S M D H D L E																									
FS8.1																										
fs8.1																										



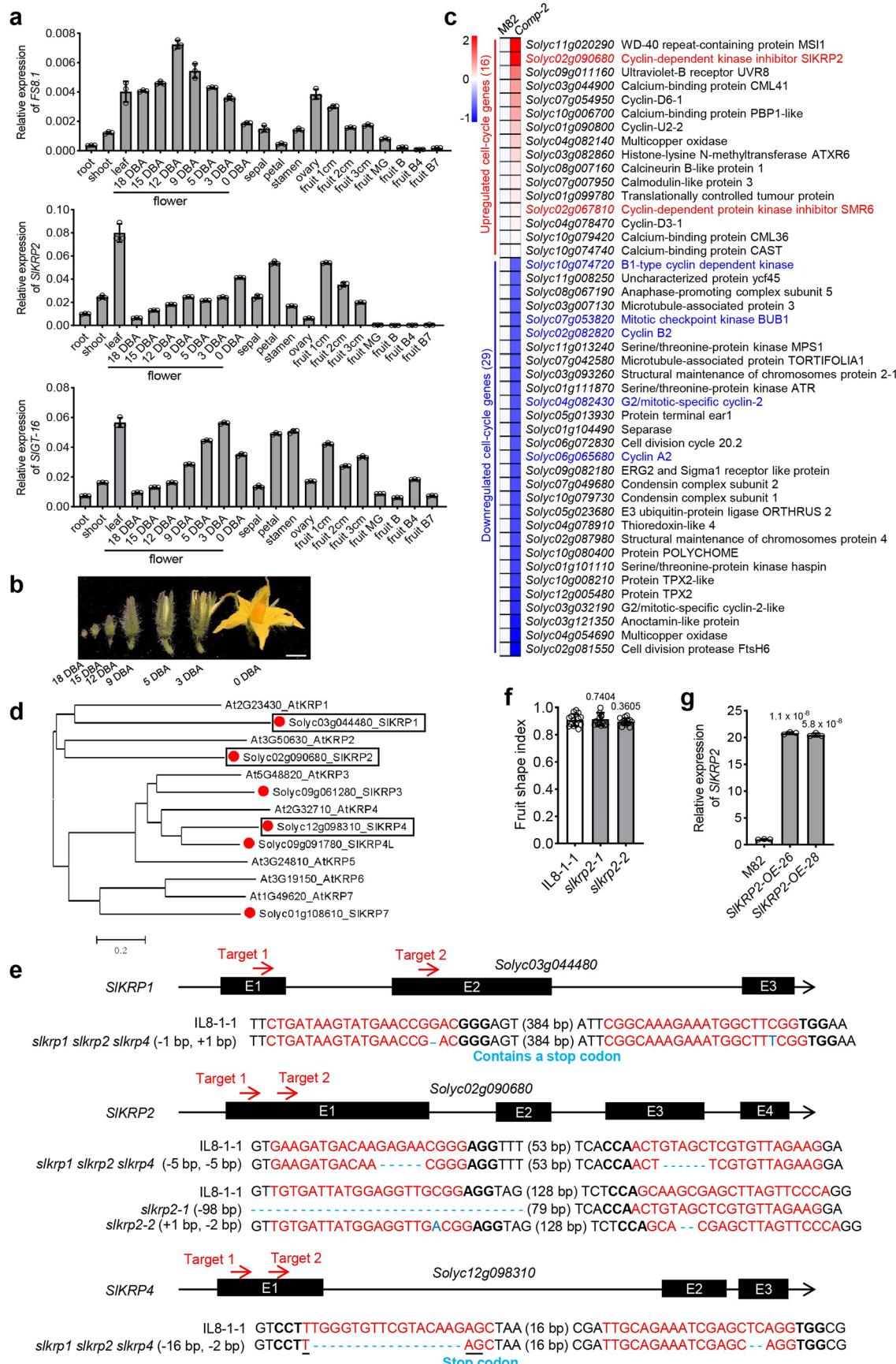
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | FS8.1 encodes a GT-2-like protein that lacks the duplicate trihelix domains. a, Protein sequence alignment of SIGT-16, SIGT-34, FS8.1 and fs8.1. The gray boxes indicate the conserved domains. b, Phylogenetic tree of trihelical transcription factors in tomato and *Arabidopsis*. The phylogenetic tree was constructed by the neighbor-joining method using MEGA version 7.0. The five clades composing the trihelix family are indicated by different colors. The *FS8.1* gene is highlighted in the red box. c, Structures of

SIGT-16, FS8.1 and fs8.1 predicted by AlphaFold. N, N-terminus; C, C-terminus; TD1, trihelix domain 1; TD2, trihelix domain 2; HD, α -helical domain. AlphaFold produces a per-residue confidence score (pLDDT) between 0 and 100. Some regions below 50 pLDDT may be unstructured in isolation. Accession numbers are from the SGN and TAIR database. The phylogenetic tree was constructed by the neighbor-joining method using MEGA version 7.0.

a**b**

Extended Data Fig. 4 | Phylogenetic analyses of FS8.1 and SIGT-16. a,b, Phylogenetic analyses of FS8.1 orthologs (a) and SIGT-16 orthologs (b). Accession numbers are from the SGN, CuGenDB, TAIR, Phytozome and NCBI databases.

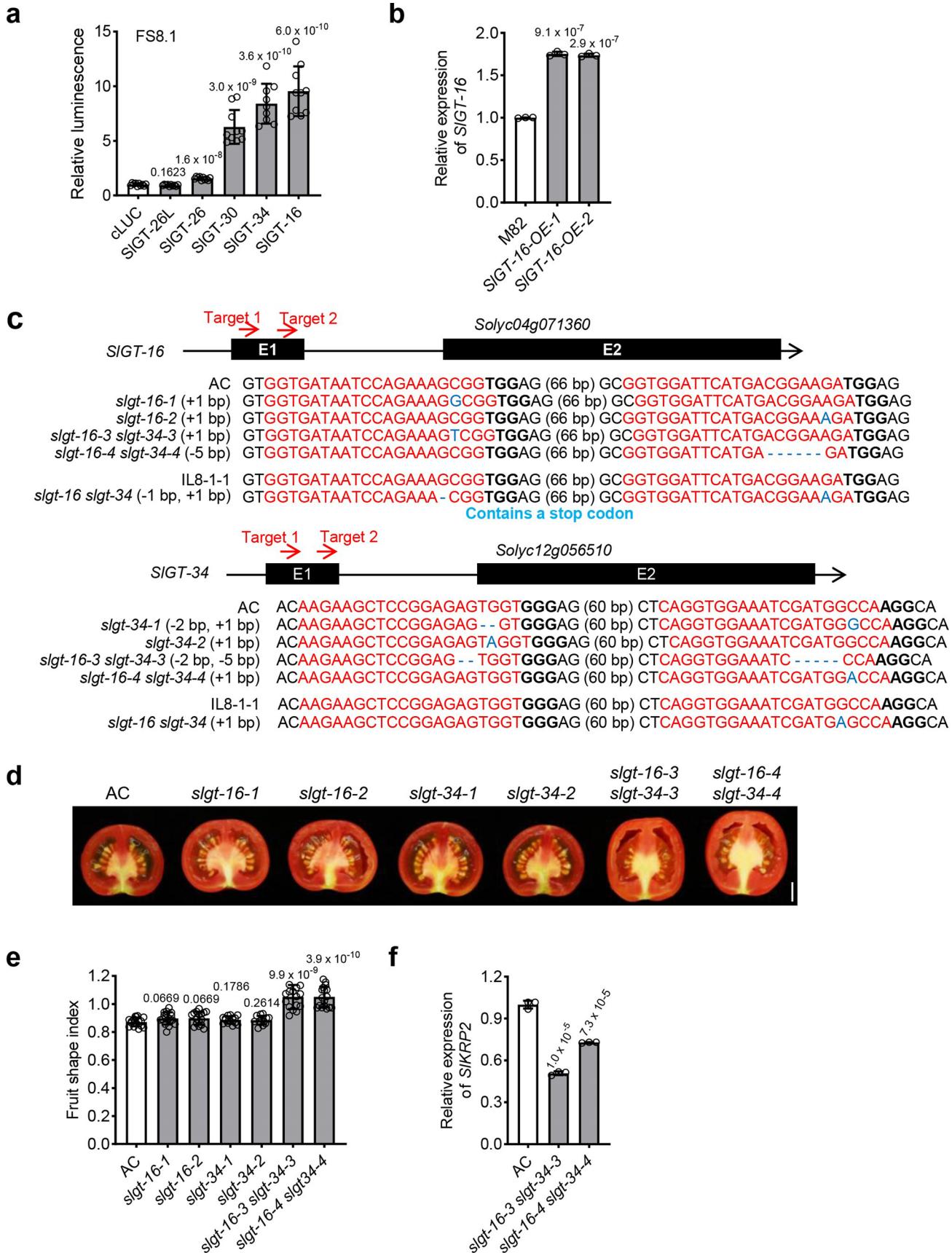


Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | FS8.1 differentially regulates cell cycle-related genes.

a, RT-qPCR results showing the expression levels of *FS8.1*, *SlKRP2* and *SlGT-16* in various tissues of AC plants ($n = 3$ independent biological replicates). MG, mature green; B, breaker; B + 4/7, 4/7 days after the breaker stage. b, Photograph of flowers at the indicated developmental stages. Bar = 5 mm. c, Expression profiles of cell cycle-related genes regulated by FS8.1. The FC (log₁₀ scale) of the average expression of each gene is shown. d, Phylogenetic tree of tomato KRP s and their *Arabidopsis* orthologs based on protein sequences. The phylogenetic tree was constructed by the neighbor-joining method using MEGA version 7.0. The scale bar indicates the average number of amino acid substitutions per site.

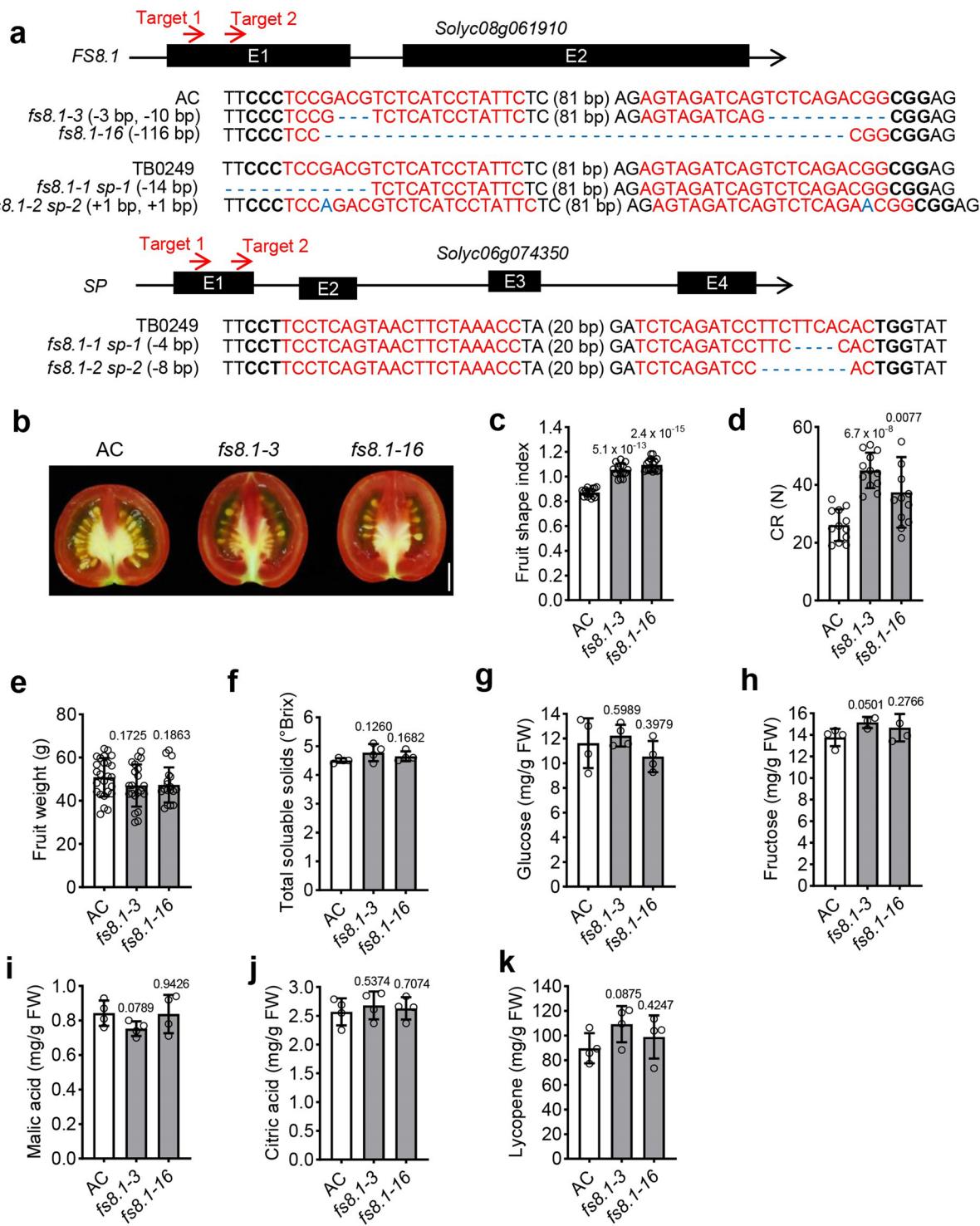
e, Generation of *slkrp1 slkrp2 slkrp4* triple mutants in the IL8-1-1 background. The sgRNA targets and PAM are highlighted in red and bold font, respectively. The blue dashes and letters indicate deletions and insertions, respectively, and the numbers indicate the numbers of nucleotides involved (+, insertion; -, deletion). f, Fruit shape indexes of IL8-1-1, *slkrp2* mutants ($n = 14$ fruits). g, RT-qPCR results showing *SlKRP2* expression in the anthesis ovaries of the indicated genotypes ($n = 3$ independent biological replicates). Bars represent the means \pm SD. The significance of the difference was evaluated by two-tailed Student's t tests, the exact *P* value is indicated on the graph.



Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | SIGT-16 and SIGT-34 redundantly regulate *SIKRP2* expression and fruit shape formation. a, LCI assays designed to determine interactions between FS8.1 and five canonical GT-2 factors ($n = 10, 10, 10, 10, 9, 10$ independent biological replicates from left to right). Tobacco cells coexpressing FS8.1-nLUC and cLUC were used as negative controls. b, RT-qPCR results showing *SIGT-16* expression in the ovaries of the indicated genotypes at anthesis ($n = 3$ independent biological replicates). c, Generation of single and double mutants of *SIGT-16* and *SIGT-34* in different backgrounds. The sgRNA targets and PAMs are highlighted in red and bold font, respectively. The blue dashes and letters

indicate deletions and insertions, respectively, and numbers indicate the numbers of nucleotides involved (+, insertion; -, deletion). d, Ripe fruits of AC, *slgt-16*, *slgt-34* and *slgt-16 slgt-34* mutants. Bar = 1 cm. e, Fruit shape indexes of AC, *slgt-16*, *slgt-34* and *slgt-16 slgt-34* mutants ($n = 16, 17, 16, 14, 13, 16, 17$ fruits from left to right). f, RT-qPCR results showing *SIKRP2* expression in the ovaries of the indicated genotypes at anthesis ($n = 3$ independent biological replicates). Bars represent the means \pm SD. The significance of the difference was evaluated by two-tailed Student's t tests, the exact *P* value is indicated on the graph.



Extended Data Fig. 7 | Editing of *FS8.1* enhances fruit firmness without compromising nutrition. a, Generation of *fs8.1* single mutants and *fs8.1sp* double mutants in different backgrounds. The sgRNA targets and PAM are highlighted in red and bold font, respectively. The blue dashes and letters indicate deletions and insertions, respectively, and the numbers indicate the numbers of nucleotides involved (+, insertion; -, deletion). b, Ripe fruits of AC and *fs8.1* mutants. Bar = 1 cm. c, d, Fruit shape indexes (c) and fruit CR (d)

of AC and *fs8.1* mutants. $n = 16$ (for c) and 12 (for d) fruits. e, Fruit weight of AC and *fs8.1* mutants ($n = 25, 21, 18$ fruits from left to right). f-k, Contents of fruit quality-related metabolites. In f, g, i-k, $n = 4$ independent biological replicates. In h, $n = 3$ independent biological replicates. Bars represent the means \pm SD. The significance of the difference was evaluated by two-tailed Student's t tests, the exact P value is indicated on the graph.

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the 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
- A statement on 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 statistical parameters 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

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection RNA-seq raw data was generated by sequencing using an Illumina HiSeq X-ten platform

Data analysis All software used in our study are publicly available and the corresponding software version were described in the Methods
1. Tomato Analyzer 4.0 <http://vanderknaaplab.uga.edu/index.html>
2. ImageJ <https://imagej.nih.gov/ij/>
3. Prism 8 <https://www.graphpad.com/>
4. MEGA 7 <https://megasoftware.net/>
5. AlphaFold <https://github.com/deepmind/alphafold>

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The RNA-seq data have been deposited in the Genome Sequence Archive (GSA; <https://ngdc.cncb.ac.cn/gsa/>) at the Beijing Institute of Genomics (BIG) Data Center, Chinese Academy of Sciences, under accession number CRA008400. The data have been released.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	This manuscript does not deal with this issue.
Reporting on race, ethnicity, or other socially relevant groupings	This manuscript does not deal with this issue.
Population characteristics	This manuscript does not deal with this issue.
Recruitment	This manuscript does not deal with this issue.
Ethics oversight	This manuscript does not deal with this issue.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is 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/documents/nr-reporting-summary-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 establish sample size. Sample size for each experiment was selected based on recently published studies, which were mentioned in the Methods.
Data exclusions	No data was excluded.
Replication	The number of replication are indicated in the figure legends. All experiments were performed at least three replicates. And all the replicates were successful.
Randomization	All samples were selected randomly for experiments.
Blinding	The investigators were blinded to group allocation during data collection and analysis.

Behavioural & social sciences study design

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

Study description	Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).
Research sample	State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.

Sampling strategy

Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.

Data collection

Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.

Timing

Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.

Data exclusions

If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.

Non-participation

State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.

Randomization

If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.

Ecological, evolutionary & environmental sciences study design

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

Study description

Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.

Research sample

Describe the research sample (e.g. a group of tagged *Passer domesticus*, all *Stenocereus thurberi* within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.

Sampling strategy

Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.

Data collection

Describe the data collection procedure, including who recorded the data and how.

Timing and spatial scale

Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken

Data exclusions

If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.

Reproducibility

Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.

Randomization

Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.

Blinding

Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.

Did the study involve field work? Yes No

Field work, collection and transport

Field conditions

Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).

Location

State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).

Access & import/export

Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).

Disturbance

Describe any disturbance caused by the study and how it was minimized.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input type="checkbox"/>	<input checked="" type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used

Rabbit polyclonal anti-GFP for ChIP (Abcam, cat#ab290, 1:750 dilution), anti-GST antibody (Abmart, cat. # M20007, 1:3000 dilution).

Validation

Information of Rabbit polyclonal anti-GFP for ChIP can be found at the product websit <<https://www.abcam.com/gfp-antibody-ab290.html>>

Information of anti-GST validation can be found at the product websit <<http://www.ab-mart.com.cn/page.aspx?node=%2059%20&id=%20967>>

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)

This manuscript does not deal with this issue.

Authentication

This manuscript does not deal with this issue.

Mycoplasma contamination

This manuscript does not deal with this issue.

Commonly misidentified lines (See [ICLAC](#) register)

This manuscript does not deal with this issue.

Palaeontology and Archaeology

Specimen provenance

Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information). Permits should encompass collection and, where applicable, export.

Specimen deposition

Indicate where the specimens have been deposited to permit free access by other researchers.

Dating methods

If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.

Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight

Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals

For laboratory animals, report species, strain and age OR state that the study did not involve laboratory animals.

Wild animals

Provide details on animals observed in or captured in the field; report species and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.

Reporting on sex

Indicate if findings apply to only one sex; describe whether sex was considered in study design, methods used for assigning sex. Provide data disaggregated for sex where this information has been collected in the source data as appropriate; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex-based analyses where performed, justify reasons for lack of sex-based analysis.

Field-collected samples

For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.

Ethics oversight

Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration

Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.

Study protocol

Note where the full trial protocol can be accessed OR if not available, explain why.

Data collection

Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.

Outcomes

Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.

Dual use research of concern

Policy information about [dual use research of concern](#)

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

- | | |
|--------------------------|---|
| No | Yes |
| <input type="checkbox"/> | <input type="checkbox"/> Public health |
| <input type="checkbox"/> | <input type="checkbox"/> National security |
| <input type="checkbox"/> | <input type="checkbox"/> Crops and/or livestock |
| <input type="checkbox"/> | <input type="checkbox"/> Ecosystems |
| <input type="checkbox"/> | <input type="checkbox"/> Any other significant area |

Experiments of concern

Does the work involve any of these experiments of concern:

- | | |
|--------------------------|--|
| No | Yes |
| <input type="checkbox"/> | <input type="checkbox"/> Demonstrate how to render a vaccine ineffective |
| <input type="checkbox"/> | <input type="checkbox"/> Confer resistance to therapeutically useful antibiotics or antiviral agents |
| <input type="checkbox"/> | <input type="checkbox"/> Enhance the virulence of a pathogen or render a nonpathogen virulent |
| <input type="checkbox"/> | <input type="checkbox"/> Increase transmissibility of a pathogen |
| <input type="checkbox"/> | <input type="checkbox"/> Alter the host range of a pathogen |
| <input type="checkbox"/> | <input type="checkbox"/> Enable evasion of diagnostic/detection modalities |
| <input type="checkbox"/> | <input type="checkbox"/> Enable the weaponization of a biological agent or toxin |
| <input type="checkbox"/> | <input type="checkbox"/> Any other potentially harmful combination of experiments and agents |

Plants

Seed stocks	Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.
Novel plant genotypes	Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.
Authentication	Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.

ChIP-seq

Data deposition

- Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links

May remain private before publication.

For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.

Files in database submission

Provide a list of all files available in the database submission.

Genome browser session (e.g. [UCSC](#))

Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

Methodology

Replicates	Describe the experimental replicates, specifying number, type and replicate agreement.
Sequencing depth	Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.
Antibodies	Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.
Peak calling parameters	Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.
Data quality	Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.
Software	Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.
Instrument	Identify the instrument used for data collection, specifying make and model number.
Software	Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.

Cell population abundance

Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.

Gating strategy

Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type

Indicate task or resting state; event-related or block design.

Design specifications

Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.

Behavioral performance measures

State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).

Acquisition

Imaging type(s)

Specify: functional, structural, diffusion, perfusion.

Field strength

Specify in Tesla

Sequence & imaging parameters

Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.

Area of acquisition

State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.

Diffusion MRI

Used

Not used

Preprocessing

Preprocessing software

Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).

Normalization

If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.

Normalization template

Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.

Noise and artifact removal

Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).

Volume censoring

Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.

Statistical modeling & inference

Model type and settings

Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).

Effect(s) tested

Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.

Specify type of analysis: Whole brain ROI-based Both

Statistic type for inference

Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.

(See [Eklund et al. 2016](#))

Correction

Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).

Models & analysis

n/a Involved in the study

Functional and/or effective connectivity

Graph analysis

Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).

Graph analysis

Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).

Multivariate modeling and predictive analysis

Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.