Metabolic engineering for provitamin D3 biosynthesis in tomato

Sunghwa Choe (shchoe@snu.ac.kr)
Seoul National University https://orcid.org/0000-0001-5541-484X
Sunmee Choi
G+FLAS Life Sciences
Jeongmo Kim
G+FLAS Life Sciences
Jongjin Park
Naturegenic Inc.
Jinhwa Kim
G+FLAS Life Sciences

Article

Keywords: CRISPR-Cas9, Non-GMO, Vegetarian, Vitamin D deficiency, COVID-19

Posted Date: March 28th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1403571/v1

Metabolic engineering for provitamin D3 biosynthesis in tomato

Sunmee Choi[1], Jinhwa Kim[1], Jeongmo Kim[1], Jongjin Park[1,2] & Sunghwa Choe[1,2,3,*]

[1] G+FLAS Life Sciences, CRISPR PLUS Lab, 38 Nakseongdae-ro, Gwanak-Gu, Seoul 08790, Korea

[2] Naturegenic Inc., 1281 Win Hentschel Boulevard, Kurz Purdue Technology Center, West Lafayette, IN 47906, USA

[3] School of Biological Sciences, College of Natural Sciences, Seoul National University, Gwanak-gu, Seoul 08826, Korea

* Corresponding author

E-mail: shchoe@snu.ac.kr

Short title: Vitamin D3 tomato created by CRISPR-Cas9

Keywords: CRISPR-Cas9, Non-GMO, Vegetarian, Vitamin D deficiency, COVID-19

Providing adequate vitamin D3 for the human diet is a long-standing goal of crop breeding because most plants have relatively low levels of this nutrient naturally. Using CRISPR-Cas9 editing of the tomato genome, we increased provitamin D3 levels in tomato fruits to up to 1 mg per 100 grams dry weight. Use of this provitamin D3 tomato could ameliorate widespread vitamin D deficiencies around the world.

In humans, provitamin D3 (ProvitD3) plays dual roles as a protectant against ultraviolet (UV) irradiation in the skin and as a precursor for the biosynthesis of active vitamin D3, an important human steroid hormone. Studies suggest that circulatory vitamin D3 (VitD3) levels in humans around the world are well below desirable levels, especially in those not frequently exposed to substantial sunlight[1]. Dietary supply offers a means to ameliorate this shortage. However, only a limited number of animal-derived foods, including fish, egg yolks, and beef liver, contain VitD[2]; vegetables and fruit are not significant sources of VitD because the biosynthetic activity of plants in the direction of ProvitD3 production is low[3]. We hypothesized that metabolic engineering of tomato plants, a crop widely cultivated and eaten worldwide, to accumulate ProvitD3 may help ameliorate endemic global VitD deficiencies, especially among people whose options for dietary changes are relatively limited, such as the elderly, vegetarians, and patients with obesity or living in nursing homes and hospitals.

ProvitD3 is 7-dehydrocholesterol (7DHC), the penultimate intermediate in the cholesterol biosynthetic pathway. 7DHC is converted to cholesterol by 7DHC reductase [4]. In an alternative branch of this pathway, exposure to UV-B light (290–315 nm) cleaves the bond between the C9 and C10 carbons of 7DHC to yield previtamin D3 (PrevitD3; Fig. 1a). Further enzymatic modification by liver CYP2R1 (Cytochrome P450 Family 2 Subfamily R Member 1) and kidney CYP27B1 solubilizes PrevitD3 to form the active vitamin 1,25-dihydroxy vitamin D (1,25(OH)2D)[5], which plays diverse roles in human metabolism, including calcium and phosphorus absorption.

We hypothesized that knocking out a crop gene homologous to Arabidopsis (*Arabidopsis thaliana*) *DWARF5* (*DWF5*), which encodes sterol 7-reductase, would result in ProvitD3 accumulation in leaves or fruits[6]. An ortholog of Arabidopsis *DWF5* has been identified in humans[4,7], and loss-of-function mutations in this gene cause severe growth defects both in plants and in humans[8]. Like Arabidopsis, lettuce (*Lactuca sativa*) has a single copy of this gene (*LsDWF5*) and a gene organization comprising 13 exons (Fig. 1b). To metabolically engineer lettuce for ProvitD accumulation, we selected appropriate single guide RNAs (sgRNAs) and introduced *LsDWF5* loss-of-function mutations into lettuce through clustered regularly interspaced short palindromic repeats (CRISPR)-CRISPR-associated nuclease (Cas9)-mediated genome editing (Extended Data Fig. 1a). We obtained 13 lines with characteristic insertion or deletion (InDel) mutations (Fig. 1c). Comparing mutant #13-17 (T_1 generation descendant of lettuce T_0 mutant #13) to wild-type plants indicated that, like its Arabidopsis counterpart, the lettuce *Lsdwf5* mutant showed dwarfism with small leaves (Fig. 1d; Extended Data Fig. 2 and 3).

To increase ProvitD3 levels while maintaining overall plant architecture and without compromising growth, an agriculturally desirable trait, we looked for a crop plant with at least two functionally redundant *DWF5* copies. Tomato (*Solanum lycopersicum*) has two copies of the gene, *SlDWF5-1* and *SlDWF5-2*, which are both structurally similar to the Arabidopsis and lettuce genes (Fig. 2a,b) and also known for active biosynthesis of cholesterol as well as campesterol[9]. We established that *SlDWF5-1* is preferentially expressed

in fruits, whereas *SlDWF5-2* was relatively weakly expressed (Fig. 2c). We hypothesized that knocking out *SlDWF5-1* could thus be a good strategy to increase ProvitD3 levels with relatively minimal effects on plant architecture. When we aligned the DNA sequences of exon 6 in the two *SlDWF5* genes, we were able to identify two potential sgRNAs targeting only the *SlDWF5-1* gene (Fig. 2d). We placed these two closely located sgRNAs sequences 31 bases apart under the control of the promoter from the Arabidopsis small nucleolar RNA *U6* and then transformed the resulting Cas9 constructs into tomato hypocotyl tissue derived from 6-day-old tomato seedlings (Extended Data Fig. 1b and 3e through h). Sixteen transgenic seedlings regenerated after transformation (T_0 generation); our examination of the editing events in these seedlings showed that 14 out of 16 carried InDel mutations in at least one of the two sgRNA sites on on-target (Fig. 2d; Extended Data Fig. 3 and 4; Supplementary Table 1). Further targeted deep-se analysis with the DNA prepared from T_2 lines revealed homozygosity for the InDel mutations (Supplementary Table 2). Frequency of InDel mutations at off-target sites was not significant except sgRNA_Sl2-OT1 site in #3-14 line (Supplementary Table 3), which is located at *SlDWF5-2*.

Next, we collected fruits and harvested T_1 seeds from seven of the 14 mutant plants. To obtain lines that are transgene free but carry the desired mutations, we plated 28–30 T_1 seeds for each line and tested whether the selection marker, conferring hygromycin resistance, remained in the genome or had segregated out. Two plants derived from #3-14-2 and #7-1-24 lines were free of the transgene and displayed vigor indistinguishable from that of wild-type plants (Fig. 2e). We examined the ProvitD3 levels in the offspring plants of #3-14 and #7-1 segregating for mutations in *SlDWF5-1* by liquid chromatography–tandem mass spectrometry (LC-MS/MS; Fig. 2f). Whereas ProvitD3 was undetectable in wild-type plants, the mature fruits from the two mutant lines showed levels ranging from 117 to 1,026 mcg per 100 g dry weight, although these were much lower in leaves and stems (Fig. 2g). When converted to fresh weight equivalents, this would equate to up to 10–100 mcg ProvitD3 per medium-sized tomato fruit of 140 g fresh weight, about half the fruit of beefsteak tomatoes. Thus, we have created a version of tomato—already recognized as a 'superfood' for its high content of other vitamins and antioxidants—further fortified to contain sufficient ProvitD3 to meaningfully improve human vitamin D intake.

The US National Institute of Health (NIH) recommends maintaining serum 25-hydroxyvitamin D [25(OH)D] concentrations of 50 nmol/L, but the majority of people have lower levels. Thus, taking daily supplements containing 10–20 mcg vitamin D is recommended (https://ods.od.nih.gov/factsheets/VitaminD-HealthProfessional/). Our data suggest that this recommendation could be met by eating one of our tomato fruits each day. Thus, through CRISPR-Cas9 genome editing of a widely used crop plant, we have created a new variety with the potential to improve human dietary ProvitD3 intake via ingestion of a sustainable, plant-derived food that is already widely consumed globally, and thereby alleviate VitD deficiency worldwide.

Methods

Vector construction

A construct used for Agrobacterium (*Agrobacterium tumefaciens*)–mediated genetic transformation was created harboring an antibiotic selection cassette, the *Cas9* gene, and tandem polycistronic tRNA-

gRNA repeats for positive selection on both kanamycin and hygromycin and genome editing. The human-codon-optimized *Cas9* gene originated from *Streptococcus pyogenes* (SpyCas9) was cloned into the pCAMBIA1300 plasmid (#44183; Addgene) and placed under the control of the Arabidopsis *Ubiquitin 10* promoter. To facilitate nuclear localization of the Cas9 protein in tomato cells, simian vacuolating virus 40 (SV40 NLS) and bipartite nuclear localization signals (BPNLSs) were added at the N and C termini of Cas9. Using the BsaI restriction enzyme, two sgRNAs were inserted into the pCAMBIA-Cas9 backbone and driven by the promoter of the Arabidopsis small nucleolar *U6* RNA.

Lettuce transformation and regeneration

One-week-old cotyledons of lettuce (*Lactuca sativa* cv. Hanbat Cheongchima) were cut into ~1-cm^2 pieces and co-cultured with Agrobacterium strain GV3101 (OD_{600} = 0.6) harboring the *Cas9* construct for 2 days at 25℃ in the dark. The co-cultured cotyledons were transferred to a callus-inducing medium, consisting of Murashige and Skoog (MS) medium (M0221, Duchefa Farma B.V) with 0.1 mg/L β-benzylaminopurine (BAP, D130, Phytotech Labs), 0.1 mg/L α-naphthalene acetic acid (NAA, N600, Phytotech Labs), 7% (w/v) plant agar (P1001, Duchefa), and 30 g/L sucrose (S0809, Duchefa), and grown at 25℃ in the light for 2 days. To select transformed tissues, the MS-grown cotyledons were transferred to MS medium containing 5 mg/L DL-phosphinothricin (BASTA, MB-P4691, MBcell) and 200 mg/L ticarcillin disodium (T1090, Duchefa) and grown at 25℃ in the light for a further 6–8 weeks, with the growth plates replaced by fresh plates at 2-week intervals until shoot generation occurred. The emerging shoots were transferred to full-strength MS medium and grown into rooted plantlets in containers. The plantlets were transferred to soil-filled pots and maintained until seed harvest.

Tomato transformation and regeneration

Two-week-old hypocotyls of tomato (*Solanum lycopersicum* cv. Seogwang) were cut into approximately 1-cm^2 pieces and co-cultured with Agrobacterium strain LBA4404 (OD_{600} = 0.5) harboring the *Cas9* construct for 2 days at 25℃ in the dark. The co-cultured hypocotyls were transferred to callus-inducing medium (MST) consisting of full-strength MS medium (M0221, Duchefa Farma B.V.) containing 0.5 mg/L nicotinic acid (1414-0130, Showa), 100 mg/L myoinositol (MB-I4715, MB cell), 0.5 mg/L pyridoxine HCl (P-8666, Sigma), 0.1 mg/L thiamine HCl (T0614, Duchefa), 30 g/L sucrose (S0809, Duchefa Farma B.V), 0.25% (w/v) Gelrite (71015-52-1, Duchefa)] with 0.1 mg/L NAA (N600, Phytotech Labs), and 1 mg/L BAP (D130, Phytotech Labs) and grown at 25℃ in the light for 2 days.

To select transformed tissues, the MST-grown hypocotyls were transferred to full-strength MST medium containing 2 mg/L zeatin (Z860, Phytotechlab), 0.2 mg/L indole-3-acetic acid (IAA, I0901, Duchefa Farma B.V.), 25 mg/L hygromycin B (LPS solution, HYB01), and 200 mg/L ticarcillin disodium (T1090; Duchefa Farma B.V.) and grown at 25℃ in the light for 8–12 weeks. The growth plates were replaced by fresh plates at 2-week intervals until shoot generation occurred. The emerging shoots were transferred to half-strength MST medium with 0.2 mg/L IAA and grown into rooted plantlets in containers. The plantlets were transferred to soil-filled pots and maintained until seed harvest.

Reverse-transcription quantitative PCR (RT-qPCR)

Total RNA was isolated from plant tissues using the RNeasy® plant mini kit (#74904, Qiagen, Hilden, Germany). Reverse transcription was performed with a RevertAid RT Reverse Transcription Kit (K1691; Thermo Fisher Scientific, Waltham, MA, USA) using 3 μg of total RNA and was followed by qPCR analysis of the resulting first-strand cDNA. The primers used for qPCR are listed in

Supplementary Table 5. RT-qPCR was performed in 96-well blocks with a Real-Time PCR System (4379216; Applied Biosystems, Foster City, CA, USA) using the KAPA SYBR® FAST qPCR Master Mix Kit (KK4601; KAPA Biosystems, Wilmington, MA, USA) in a volume of 20 μL. The reactions were performed in three technical repeats per run, and three biological replicates were included. Absolute quantification was performed using standard curves generated by amplification of a diluted series of cDNA containing individual transcripts. The transcript levels of each gene in different samples were normalized to an internal control, *ACTIN* (Solyc11g005330.2), using the $2^{-\Delta\Delta CT}$ method.

T7E1 assay

Genomic DNA was isolated from transgenic plants using the DNeasy Plant Mini Kit (#69104, Qiagen). The target DNA region was amplified and subjected to the T7E1 assay as described previously (Woo et al., 2015). For heteroduplex formation, PCR products were denatured at 95°C for 10 min and subjected to an annealing program of 95°C to 85°C (–2°C/sec) followed by a 85°C to 25°C (–0.1°C/sec) slow cooldown using a thermal cycler. Annealed PCR products were incubated with T7 endonuclease I (#m0302, NEB) at 37°C for 20 min and analyzed via agarose gel electrophoresis.

Sanger sequencing of target regions

The sgRNA target regions were amplified from genomic DNA using Q5 Polymerase (#M0491; New England Biolabs) in a 25-μL reaction volume. Then, the PCR products were cloned into a 3′-end T-tailed vector using a PCR Cloning Kit (VT201-020; Biofact Pharma Ltd., Kildare, Ireland). Finally, 20 clones for each sample were individually sequenced. The primers used for on-target site mutation analysis are listed in Supplementary Table 4 (lettuce) and 6 (tomato).

Targeted deep sequencing

To compare editing events in the target loci, on-target sequencing was performed for the 23 lines of the T_1 generation, 21 lines of the T2 generation, and the sequences compared to that of the wild type (Supplementary Table 1 and 2). The targeted primers were designed from genomic DNA (1st PCR, Supplementary Table 6). Sequencing adaptors were added for the amplicon PCR (Supplementary Table 6). High-throughput sequencing was performed using the MiniSeq System (SY-420-1001; Illumina, Inc., San Diego, CA, USA) and analyzed using methods available online[10].

Off-target deep sequencing

The comparison target for off-target sequencing is seedling of T1 generation #3-14 and #7-1 were used and compared with wild type. Potential off-target sites were identified in the S. lycopersicum genome using the Cas-OFFinder (http://www.rgenome.net/cas-offinder) algorithm. The Sol Genomics Network (https://solgenomics.net) was used as the reference genome to identify homologous sequences that differed from the on-target sequences by up to 4 nt. Of a total of 32 sites, 11 were selected for targeted deep sequencing. The primers for the on-target and potential off-target sites were designed from genomic DNA (1st PCR, supplementary Table 7). Sequencing adaptors were added for the 2nd PCR (supplementary Table 7). High-throughput sequencing was performed using the MiniSeq System (SY-420-1001; Illumina, Inc., San Diego, CA, USA).

Authentic chemical standards.

7-Dehydrocholesterol (ProvitD3), a 5,7-conjugated diene sterol, is a biosynthetic precursor of cholesterol. Approximately 5 mg of 7-dehydrocholesterol (purity ≥ 95%, #30800, Sigma-Aldrich, Germany) was weighed, and 50 mL of methanol was added to result in a concentration of 100 mcg/mL. The sterol was then dissolved by sonication. The stock solution was diluted appropriately with methanol to produce standard solutions at concentrations of 70, 140, 280, 560, and 1120 ng/mL. The standard solutions were run under the same conditions as the test samples (see below), and the resulting data were used to prepare standard curves and a coherent calculation formula. The results for the test samples were then extrapolated from the standard curve to determine relative concentrations of 7-DHC in the fruit samples.

Sample preparation

Samples from the T_2 generations of #3-14 (four individual lines) and #7-1 (11 individual lines) as well as wild-type (WT) plants were compared. To prepare analytic samples, 15 fruits were picked and weighed to obtain an average 140 g fresh weight per fruit. When the fresh fruits were freeze dried, about 7% of the original fresh weight remained. To approximately 1 g (dry weight) of each fruit sample placed in a flask, 30 mL of ethanol and 1 mL of 10% (v/v) ethanolic pyrogallol were added. Then, 3 mL of 90% (w/v) potassium hydroxide (#484016, Sigma-Aldrich) was added, and each mixture was incubated for 30 min with repeated vortexing every 10 min in an 80°C water bath. After cooling down to room temperature, 30 mL each of distilled water and *n*-hexane were added and the mixture was shaken for 1 h. The upper layer (hexane) was transferred to a separatory funnel, and another 30 mL of *n*-hexane was added to the water layer and mixed by vortexing. The upper layer was again pooled into the separatory funnel. 100 mL of distilled water was used twice for washing the hexane layer. Then, the hexane layer was dehydrated with anhydrous sodium sulfate, concentrated, dissolved in 5 mL methanol, and the filtered through a polytetrafluoroethylene (PTFE) syringe filter (CLS431227, Merck) to prepare test solutions.

LC-MS/MS analysis of ProvitD3

The prepared samples were analyzed by LC-MS/MS (using a UHPLC TSQ Quantis™, Thermo Scientific) at the Korea Health Supplements Institute (http://www.khsi.re.kr). Sample analysis was carried out using two isocratic solvents. Liquid chromatography was performed with Waters HSS T3(2.1 × 100 mm with 1.8 μm of particle size) column. The mobile phase consisted of two solvents (A) 0.1% formic acid in distilled water, and (B) 0.1% formic acid in methanol, and the separation was performed using the following gradient of A–B: 0 min (90:10, v/v), 1 min (0:100, v/v), 10 min (0:100, v/v), 10.1 min (90:10, v/v), and 14 min (90:10, v/v) at a flow-rate of 0.2 ml/min. The injection volume was 20 μL. MS/MS data for 7-DHC were acquired in positive-ion mode using atmospheric pressure chemical ionization (APCI) and selected reaction monitoring (SRM); a parent ion mass of 385.275 and product ion masses of 367.437, 201.315, and 241.325 were used for limit of detection (LOD) determination. MS parameters were optimized for 7-DHC and were as follows: Auxiliary gas pressure was set at 1/5 pressure level of the sheath gas, and nitrogen was utilized for both. The positive-ion discharge current was set to 4 μA, and the ion transfer tube temperature was set at 325°C.

Data availability

Acknowledgments

This project was supported by the CRISPR PLUS R&D program (2019GFRND-GFC102) supported by G+FLAS Life Sciences, Inc.

Author contributions

S.Choe designed and supervised the research, and wrote the article. S.Choi, Jinhwa K., Jeongmo K., and J.P. performed the research and prepared the figures and tables.

Competing financial interests

S.Choe is a founder of the two biotech companies, G+FLAS Life Sciences, Inc. and Naturegenic Inc. S.Choi, Ji.K., and Je.K. are employees of G+FLAS Life Sciences. J.P. is an employee of Naturegenic Inc. Authors are inventors on a patent application covering the genome editing method for plant-based vitamin D3 production described in this manuscript.

Additional information

Supplementary information is available for this paper at https://.... . Correspondence and requests for materials should be addressed to S.C.

References

1 Cashman, K. D. *et al.* Vitamin D deficiency in Europe: pandemic? *Am J Clin Nutr* **103**, 1033-1044, doi:10.3945/ajcn.115.120873 (2016).

2 Schmid, A. & Walther, B. Natural vitamin D content in animal products. *Adv Nutr* **4**, 453-462, doi:10.3945/an.113.003780 (2013).

3 Japelt, R. B. & Jakobsen, J. Vitamin D in plants: a review of occurrence, analysis, and biosynthesis. *Front Plant Sci* **4**, 136, doi:10.3389/fpls.2013.00136 (2013).

4 Wassif, C. A. *et al.* Mutations in the human sterol delta7-reductase gene at 11q12-13 cause Smith-Lemli-Opitz syndrome. *Am J Hum Genet* **63**, 55-62, doi:10.1086/301936 (1998).

5 Saponaro, F., Saba, A. & Zucchi, R. An Update on Vitamin D Metabolism. *Int J Mol Sci* **21**, doi:10.3390/ijms21186573 (2020).

6 Silvestro, D. *et al.* Vitamin D5 in Arabidopsis thaliana. *Sci Rep* **8**, 16348, doi:10.1038/s41598-018-34775-z (2018).

7 Lecain, E., Chenivesse, X., Spagnoli, R. & Pompon, D. Cloning by metabolic interference in yeast and enzymatic characterization of Arabidopsis thaliana sterol delta 7-reductase. *J Biol Chem* **271**, 10866-10873, doi:10.1074/jbc.271.18.10866 (1996).

8 Choe, S. *et al.* Lesions in the sterol delta reductase gene of Arabidopsis cause dwarfism due to a block in brassinosteroid biosynthesis. *Plant J* **21**, 431-443, doi:10.1046/j.1365-313x.2000.00693.x (2000).

9 Kim, T. W. *et al.* Novel biosynthetic pathway of castasterone from cholesterol in tomato. *Plant Physiol* **135**, 1231-1242, doi:10.1104/pp.104.043588 (2004).

10 Park, J., Lim, K., Kim, J. S. & Bae, S. Cas-analyzer: an online tool for assessing genome editing results using NGS data. *Bioinformatics* **33**, 286-288, doi:10.1093/bioinformatics/btw561 (2017).

Figure legends

Figure 1 | Provitamin D3 pathway and CRISPR-Cas9 genome engineering of lettuce for VitD production. a, VitD biosynthetic pathway. The final step in sterol biosynthetic pathways is mediated by plant DWF5 and human DHCR7. Different 7-dehydrosterols varying at the R group lead to vitamin D3 from 7-dehydrocholesterol, D5 from 7-dehydrositosterol, and D7 from 7-dehydrocampesterol. VitD5 and VitD7 are more commonly produced in plants than VitD3. **b**, Gene organization of *DWF5* from *Arabidopsis thaliana* (At1g50430) and *Lactuca sativa* (LOC111910394). Open rectangles, untranslated sequences; filled rectangles, exons; connecting lines, introns; red arrows, locations of single guide RNAs (sgRNAs). **c**, Mutations introduced by CRISPR-Cas9 genome editing: targeted Sanger sequencing results of *LsDWF5* in genome-edited mutant plants at the T_0 generation. DNA prepared from the pooled tissues of T_0 plants were sequenced to find InDel mutations. Underlined bold letters, PAM sites; red dashed lines, deleted nucleotides; red letters, insertions. Numbers with plus/minus signs indicate deletions (–) and insertions (+). Numbers in brackets show frequency of the specific alleles out of 69 T_0 plants for the sgRNA-Ls1 site and 74 T_0 plants for the sgRNA-Ls2LsDWF5-2 site. **d**, Phenotypes of a wild-type and an *Lsdwf5* mutant (#13-17) plant, the latter showing a dwarf phenotype. #13 at T_0 showed heterozygous mutation of single-base insertion at the sgRNA-Ls2 site, and the offspring #13-17 has a homozygous mutation.

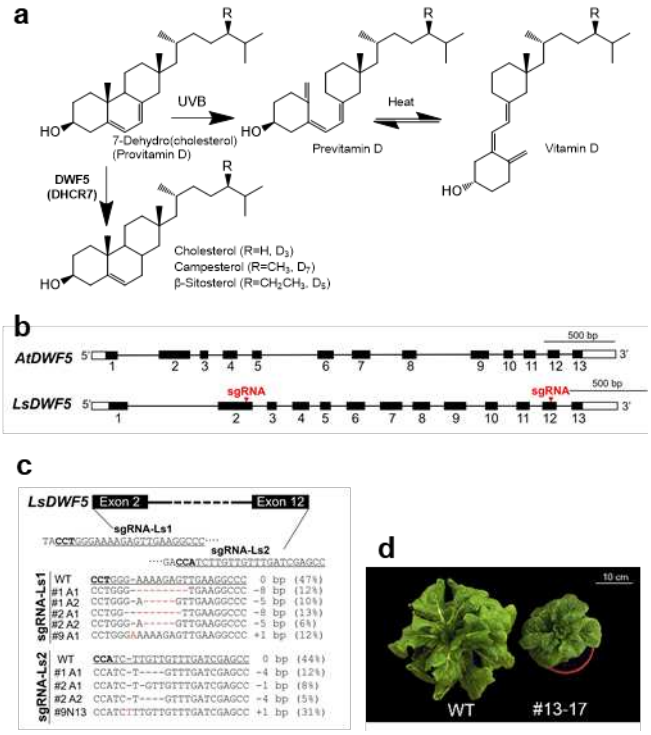
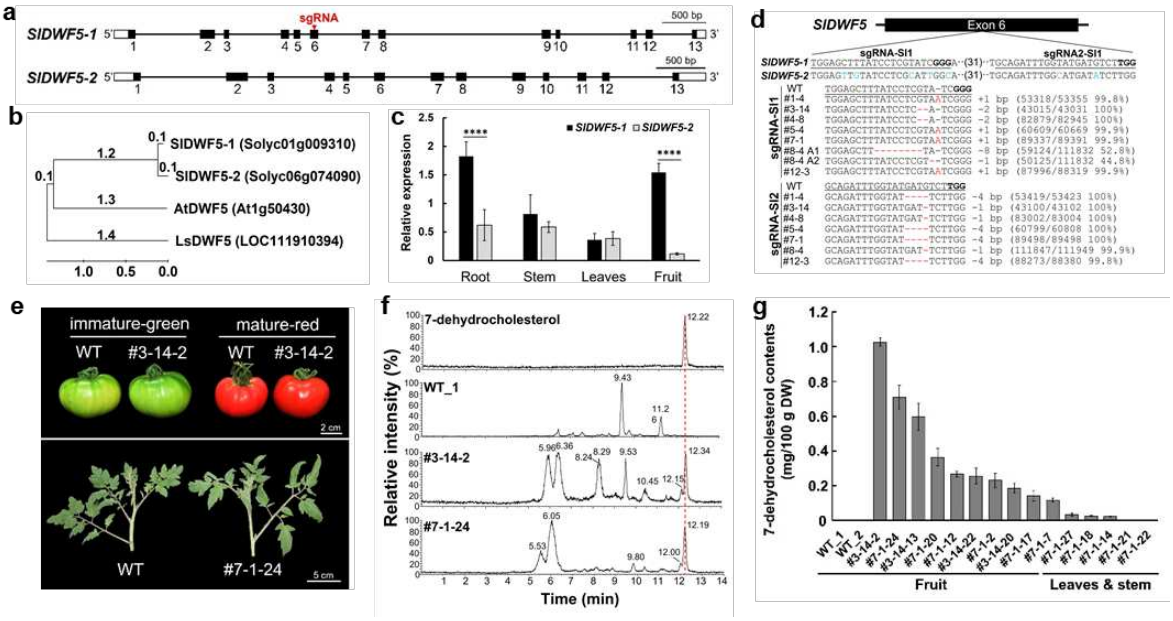
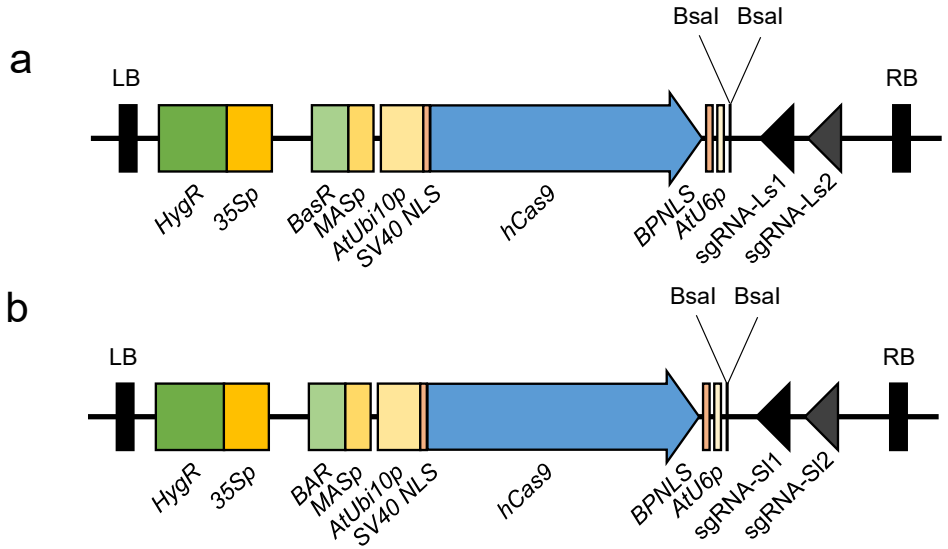


Figure 2 | Provitamin D3 accumulation in genome-edited tomato plants. a, Schematic diagram of the *Solanum lycopersicum DWF5* genes *SlDWF5-1* (Solyc01g009310) and *SlDWF5-2* (Solyc06g074090). Open rectangles, untranslated sequences; filled rectangles, translated exons; connecting lines, introns; red arrow, position of sgRNA. **b**, Phylogenetic tree showing the relationship among Arabidopsis DWF5 and its tomato and lettuce homologs. Scale indicates relative rate of amino acid substitutions per residue. **c**, Spatial expression patterns of the two *SlDWF5* genes determined by reverse-transcription quantitative PCR (RT-qPCR). All tissues were collected from 8-week-old plants, except for the mature fruits. Relative expression levels are normalized to those of the tomato actin gene (*ACT1*, TC194780a). Error bars, s.d. ($n = 3$ biological samples). ****$P < 0.0001$ based on Student's *t*-test. **d**, Results of targeted deep sequencing of T_1 plants at *SlDWF5-1*. Genotypes of *SlDWF5-1* mutants were confirmed by Illumina Mini-Seq. Underlined boldface, PAM site; red dashed lines, deleted nucleotides; red letters, inserted bases. Numbers with plus/minus signs indicate deletions (–) and insertions (+). Numbers in parentheses shows frequency of the specific mutations after targeted deep sequencing reads. **e**, Little phenotypic difference was found between *Sldwf5-1* mutants and the wild type (WT). Immature green and mature red fruits of the WT and *Sldwf5-1* mutants (#3-14-2) after 6 months of growth showed no significant difference in shape or size. The stem node of one *Sldwf5-1* mutant (#7-1-24) grew by one fewer node than the WT. **f**, LC-MS/MS spectra showing ProvitD3 peaks relative to the ProvitD3 (7-dehydrocholesterol) standard, in WT fruits (WT_1), and *Sldwf5-1* mutant fruits (#3-14-2, #7-1-24). Retention time is represented by a red dashed line. **g**, Concentrations of ProvitD3 in tissues of T_2 plants of *Sldwf5-1* mutant lines. Error bars, s.d. ($n = 3$ biological samples).



Extended Data Figures and Supplementary Tables

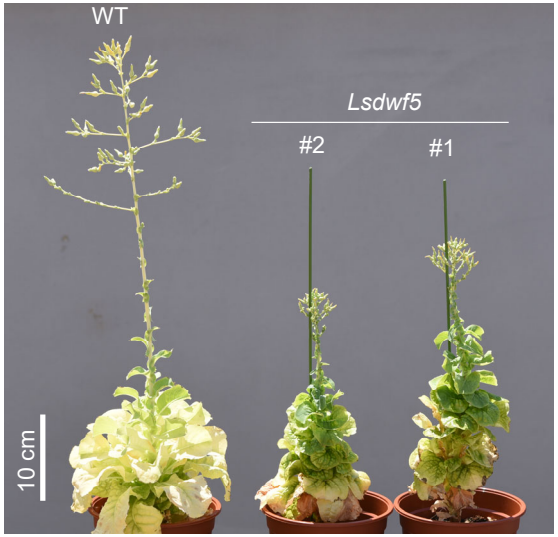


Extended Data Fig. 1. Schematic of the genome-editing binary vector containing two sgRNAs with a gRNA-tRNA linker system. Two sgRNAs were constructed with tandemly arrayed tRNA-spacer 20 bp-sgRNA scaffold systems. Two tandem repeats were placed under the control of a U6 promoter. LB is for the left border sequence of *Agrobacterium* T-DNA, HygR for the coding sequence of the hygromycin B phosphotransferase gene, 35Sp for the cauliflower 35S promoter, BAR for the coding sequence of the BASTA resistance gene phosphinothricin acetyltransferase, MASp for the mannopine synthase (MAS) promoter, AtUbi10p for the *Arabidopsis* ubiquitin 10 gene promoter, SV40 NLS for the nuclear localization signal sequence from simian vacuolating virus 40, hCas9 for the human codon-optimized Cas9 coding sequence, BPNLS for the bipartite nuclear localization sequence, AtU6p for the *Arabidopsis* U6 gene promoter, two sgRNA for the spacer sequences, the gRNA scaffold for the scaffold sequence for sgRNA, tRNA for the tRNA scaffold for RNase P processing, and RB for the right border.

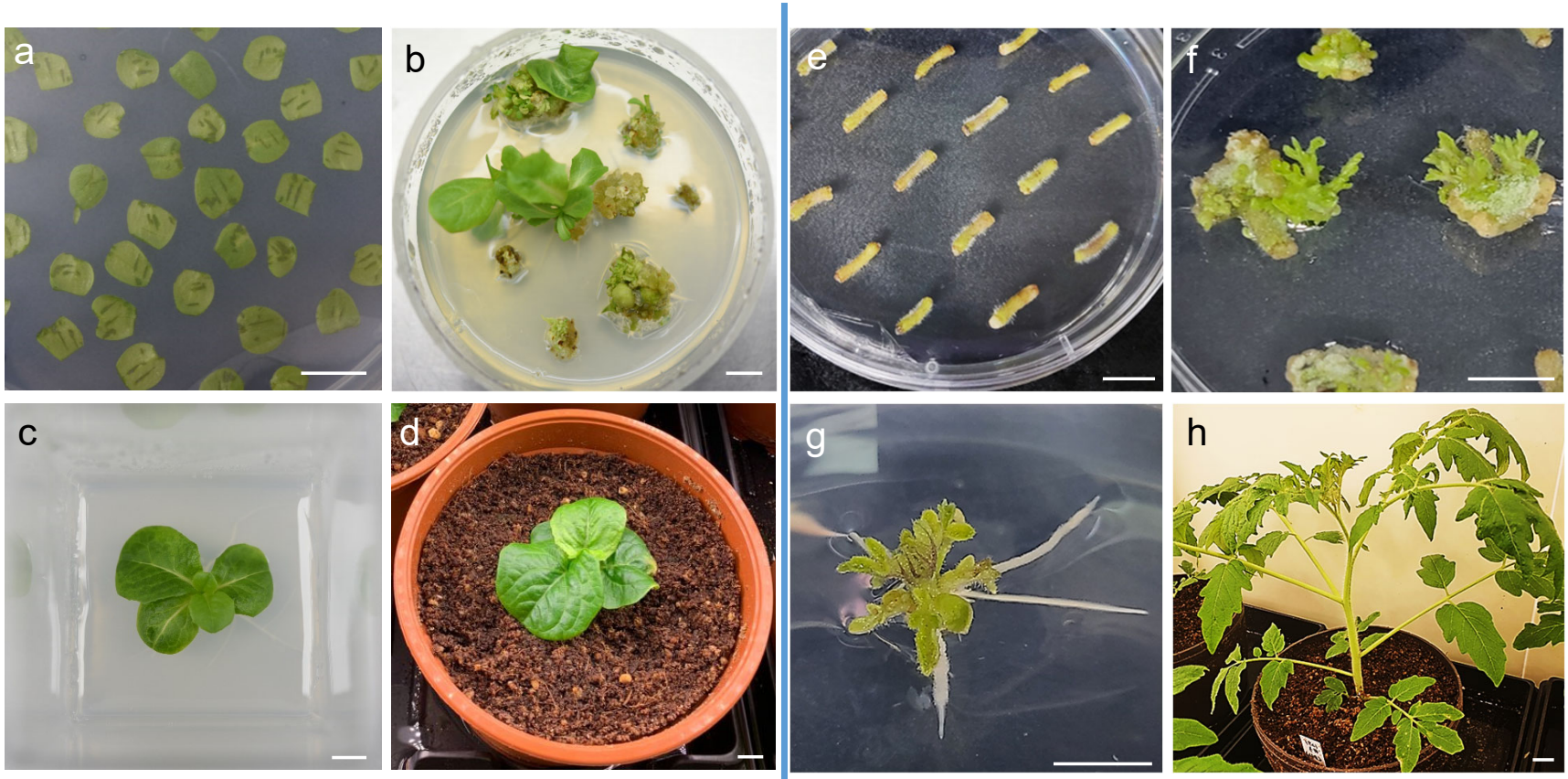
a

Generation	Line	Indel	note
T0	**#1**	(-)4/(-)5/(-)8/WT	small amount of seeds
	#2	(-)1/(-)4/(-)5/(-)8/WT	small amount of seeds
	#3	WT	
	#4	(-)5/(-)10/(-)38	died
	#5	(-)8/WT	sterile
	#6	(-)5/(-)8/WT	sterile
	#7	(-)8/(-)195/WT	sterile
	#8	WT	
	#9	+1/+1	sterile
	#10	(-)2/(-)14/WT	small amount of seeds
	#11	WT	
	#12	(-)2/WT	small amount of seeds
	#13	+1/WT	normal amount seeds
	#14	(-)2/(-)5/(-)14/WT	sterile

b

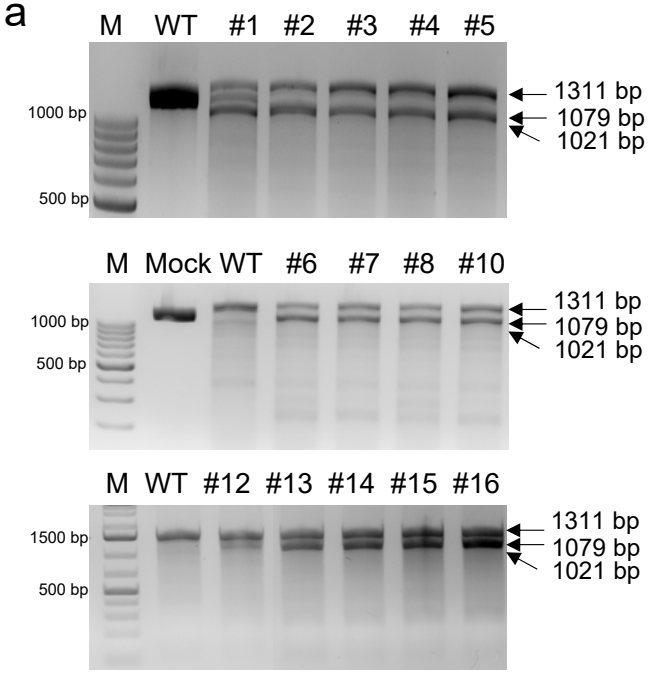


Extended Data Fig. 2. Detection of gene editing events in transgenic lettuce plants. **a,** Gene editing results by sanger sequencing. Gene editing was confirmed in 11 out of 14 transgenic plants. #4 transgenic plants died during growth. #5, #6, #7, #9 and #14 lines were sterile. **b,** *Lsdwf5* showed dwarfism, short-round leaves and reduced seed setting, or a few seeds, if any.



Extended Data Fig. 3. Plants regenerated after genome editing in *Lactuca sativa* (a-d) and *Solanum lycopersicum* (e-h). **a,** Cotyledons after cocultivation with Agrobacterium (GV3101). **b,** Shoots from eight-week-old explants under 5 mg/L BASTA . **c,** Roots from four-week-old explants. **d,** A transgenic lettuce plant. **e,** Hypocotyls after co-cultivation with Agrobacterium (LBA4404). **f,** Shoots from ten-week-old explants under 25 mg/L hygromycin. **g,** Roots from four-week-old explants. **h,** A transgenic tomato plant in soil. White scale bars, 1 cm.

a



b

Generation	Line	Indel	note
T0	#1	(-)159/(-)11/(-)4/+1	
	#2	(-)8/(-)5/(-)4/(-)1/+1	
	#3	(-)4/(-)2/(-)1/+1	
	#4	(-)4/(-)2/(-)1/+1	
	#5	(-)6/(-)4/(-)1/+1	
	#6	(-)33/(-)11/(-)5/(-)4/(-)1/+1	
	#7	(-)17/(-)4/(-)1/+1	
	#8	(-)4/(-)1/+1	
	#9		died
	#10	(-)5/(-)4/(-)1/+1	
	#11		dead
	#12	(-)12/(-)5/(-)4/(-)1/+1	
	#13	(-)8/(-)4/(-)1/+1	
	#14	(-)5/(-)4/(-)1/+1	
	#15	(-)5/(-)4/(-)1/+1	
	#16	(-)8/(-)4/(-)1/+1	

Extended Data Fig. 4. Confirmation of gene editing events in tomato transgenic plants by T7E1 assay (**a**) and Sanger sequencing (**b**). **a,** Results of T7E1 analysis on 16 transgenic plants (T0). Confirm that gene editing has occurred in all individual plant. M; Size maker, Mock; no treatment T7Endonuclease 1, WT; Wild type. Treatment of PCR product with 1,311 bp length with T7E1 is cut into 1,079 and 232 for sgRNA_Ls1, and 1,021 and 290 for sgRNA_Ls2. **b,** Gene editing results by Sanger sequencing. Gene editing was confirmed in 14 out of 16 transgenic plants. #9 and #11 transgenic plants died during growth.

Supplementary Table

Table S1. Targeted deep-sequencing to confirm homozygosity for the InDel mutations. DNAs prepared from the select T1 generation of *Sldwf5-1* mutants were subject to Illumina Mini-Seq.

T0 plant No.	T1 plant No.	Total reads (raw fastq)	sgRNA_SI1			Total reads (raw fastq)	sgRNA_SI2		
			Insertions	Deletions	Indel frequency(%)		Insertions	Deletions	Indel frequency(%)
#1	#1-1	86818	86370	340	86710 (99.8%)	86957	0	86913	86913 (99.9%)
	#1-2	54277	53495	29	53524 (99.9%)	53645	0	53629	53629 (100.0%)
	#1-4	53415	55318	51	53369 (99.8%)	53423	0	53419	53419 (100.0%)
	#1-5	68709	68498	167	68665 (99.9%)	68855	0	68821	68821 (100.0%)
#3	#3-5	50317	44	50256	50300 (99.9%)	50401	0	50376	50376 (100.0%)
	#3-6	70729	70528	65	70593 (99.8%)	70792	0	70775	70775 (99.9%)
	#3-11	54564	17	54531	54548 (100.0%)	54615	0	54615	54615 (100.0%)
	#3-14	43035	16	43015	43031 (100.0%)	43102	0	43100	43100 (100.0%)
	#3-15	53598	53473	70	53543 (99.9%)	53663	0	53663	53663 (100.0%)
	#3-18	52997	49	52868	52917 (99.8%)	52978	0	52978	52978 (100.0%)
#4	#4-4	83111	386	82710	83096 (100.0%)	82938	0	82938	82938 (100.0%)
	#4-8	82901	65	82814	82879 (100.0%)	83004	0	83002	83002 (100.0%)
	#4-9	78655	108	78516	78624 (100.0%)	78519	0	78519	78519 (100.0%)
#5	#5-4	60669	60609	60	60669 (99.9%)	60808	0	60799	60799 (100.0%)
	#5-9	70656	70473	134	70607 (99.9%)	70750	0	70699	70699 (99.9%)
#7	#7-1	89457	89337	54	89391(99.9%)	89541	0	89498	89498 (100.0%)
	#7-5	98801	58	98701	98759 (100.0%)	98838	0	98803	98803 (100.0%)
#8	#8-1	71913	37	39	76 (0.1%)	71691	0	71658	71658 (100.0%)
	#8-3	75821	75656	81	75737 (99.9%)	75901	0	75867	75867 (100.0%)
	#8-4	111832	156	111676	111832 (100.0%)	111849	0	111847	111847 (100.0%)
#12	#12-3	88259	87996	163	88159 (99.9%)	88452	0	88273	88273 (99.8%)
	#12-7	109897	109131	623	109754 (99.9%)	109974	0	109946	109946 (100.0%)
	#12-8	7402	7377	13	7390 (99.8%)	7421	0	7412	7412 (99.9%)

Supplementary Table

Table S2. Targeted deep-sequencing to confirm homozygosity for the InDel mutations. DNAs prepared from the select T2 generation of *Sldwf5-1* mutants were subject to Illumina Mini-Seq.

T2 plant No.	Total reads	sgRNA_SI1			Total reads	sgRNA_SI2		
	(raw fastq)	Insertions	Deletions	Indel frequency(%)	(raw fastq)	Insertions	Deletions	Indel frequency(%)
#3-14-1	15202	428	14774	15202 (100.0%)	15179	0	15179	15179 (100.0%)
#3-14-2	9996	15	9977	9992 (100.0%)	10446	0	10446	10446 (100.0%)
#3-14-6	15009	758	14233	14991 (99.9%)	15014	0	14995	14995 (99.9%)
#3-14-13	6601	15	6586	6601 (100.0%)	6735	0	6735	6735 (100.0%)
#3-14-20	3319	10	3309	3319 (100.0%)	3905	0	3905	3905 (100.0%)
#3-14-22	1973	7	1966	1973 (100.0%)	2406	0	2406	2406 (100.0%)
#3-14-29	9791	351	9440	9791 (100.0%)	9795	0	9795	9795 (100.0%)
#7-1-1	16711	16654	37	16691 (99.9%)	16839	0	16828	16828 (99.9%)
#7-1-2	6741	6734	5	6739 (100.0%)	7096	0	7096	7096 (100.0%)
#7-1-3	14058	14033	15	14048 (99.9%)	14057	0	14057	14057 (100.0%)
#7-1-7	7950	7940	2	7942 (99.9%)	8335	0	8335	8335 (100.0%)
#7-1-12	8018	8014	2	8016 (100.0%)	8295	0	8295	8295 (100.0%)
#7-1-14	23130	23117	8	23125 (100.0%)	23204	0	23204	23204 (100.0%)
#7-1-15	20284	20238	26	20264 (99.9%)	20403	0	20403	20403 (100.0%)
#7-1-17	11392	11379	4	11383 (99.9%)	12154	0	12154	12154 (100.0%)
#7-1-18	9984	9979	5	9984 (100.0%)	10080	0	10080	10080 (100.0%)
#7-1-20	7658	7652	0	7652 (99.9%)	8136	0	8136	8136 (100.0%)
#7-1-21	5348	5348	0	5348 (100.0%)	5670	0	5670	5670 (100.0%)
#7-1-22	7719	7711	2	7713 (99.9%)	8139	0	8136	8136 (100.0%)
#7-1-24	6645	6625	18	6643 (100.0%)	6728	0	6728	6728 (100.0%)
#7-1-27	9320	9311	7	9318 (100.0%)	9385	0	9385	9385 (100.0%)

Supplementary Table

Table S3. Targeted deep-sequencing to detect InDel mutations at off-target sitesusing Illumina Mini-Seq. Mutations were scanned for the three lines including wild type, #3-14, and #7-1. Off-target effect was negligible except OT1 of SlDWF5-1 sgRNA_Sl2. This OT1 site is located on *SlDWF5-2*.

sgRNA			WT				#3-14				#7-1		
			Total reads (raw fastq)	Insertions	Deletions	Indel frequency(%)	Total reads (raw fastq)	Insertions	Deletions	Indel frequency (%)	Total reads (raw fastq)	Insertions	Deletions	Indel frequency (%)
SlDWF5-1 sgRNA_Sl1	OT1	gGGAGCTTTAaCCTCtcATCTGG	55188	0	10	10 (0.0%)	48078	0	0	0 (0.0%)	54495	0	0	0 (0.0%)
	OT2	catAGCTTTATCCTCaTATCTGG	40820	0	0	0 (0.0%)	63347	0	0	0 (0.0%)	22263	0	0	0 (0.0%)
	OT3	TGGAGCTTcATCCTCagATtTGG	45434	2	2	4 (0.0%)	48846	0	4	4 (0.0%)	47858	0	2	2 (0.0%)
	OT4	TGtAGCTaaATCCTCGTATaAGG	41279	0	7	7 (0.0%)	35503	0	5	5 (0.0%)	41664	0	0	0 (0.0%)
	OT5	TGGAGgcTTATtCTCGaATCAGG	26174	0	0	0 (0.0%)	15815	0	2	2 (0.0%)	15122	0	2	2 (0.0%)
SlDWF5-1 sgRNA_Sl2	OT1	GCAGATTTGGcATGATaTCTTGG	32736	0	0	0 (0.0%)	48131	0	173	173 (0.4%)	21145	0	2	2 (0.0%)
	OT2	GCAGATTTGGTtTGATtTtTGGG	45194	3	5	8 (0.0%)	50614	2	8	10 (0.0%)	42789	3	10	13 (0.0%)
	OT3	GaAGATTcGGTAgGATtTCTAGG	31961	0	0	0 (0.0%)	55929	0	2	2 (0.0%)	31906	0	6	6 (0.0%)
	OT4	GCAtATTTGtTATcATtTCTAGG	61779	0	6	6 (0.0%)	52921	0	3	3 (0.0%)	53101	0	3	3 (0.0%)
	OT5	GaAaATTTGcTAaGATGTCTCGG	48830	0	13	13 (0.0%)	30867	0	12	12 (0.0%)	41365	0	12	12 (0.0%)
	OT6	GtAGAcTTGGTAgGATGTCcTGG	48871	0	0	0 (0.0%)	37297	0	0	0 (0.0%)	56254	0	0	0 (0.0%)

Supplementary Tables

Table S4. Oligonucleotide sequences used for amplification of the DNA flanking the target region in lettuce.

Gene	Name	Sequence	Tm (°C)
LsDWF5	LsDWF5_F1	CAGATGGTATACAATGGTACATCTCGATGGTTC	61
	LsDWF5_R1	GAAAATGAGTGCGGAGTATATTTCTCCCAAATGATC	61
	LsDWF5_F4	GTGGGGTTTATCTCGTCATTTCCACTATGTACC	62
	LsDWF5_R5	TCAGTAGATCCCTGGTATGATCCTATATGGAACC	62

Table S5. Oligonucleotide sequences used for realtime PCR analysis in tomato.

Gene	Name	Sequence	Tm (°C)
ACTIN	SlACT_qRT-PCR_F	GAGCGTGGTTACTCGTTCA	58
	SlACT_qRT-PCR_R	CTAATATCCACGTCACATTTCAT	58
SlDWF5-1	SlDWF5-1_qRT-PCR_F	CGGAGACTTCTGTAGTTCCA	55
	SlDWF5-1_qRT-PCR_R	GGTGGTGTGAAAGATAGAAGAC	55
SlDWF5-2	SlDWF5-2_qRT-PCR_F	AGGTGCAAGTCAAAGTATGG	55
	SlDWF5-2_qRT-PCR_R	CTACTACAATACTAGAAGCAGCACT	55

Table S6. Oligonucleotide sequences used for amplification of the DNA flanking the target region in tomato.

Gene	Name	Sequence	Tm (°C)
SlDWF5-1	SlDWF5-1_F	GGCTCATCAGGGAACATAATAGTTGACTTCTATTGG	60
	SlDWF5-1_R2	CTCCATGCCAAGTCAAAATTGG	60
	Amplicon_SlDWF5-1_F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGTACTTGCA TCATCTCTGG	68
	Amplicon_SlDWF5-1_R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCTCACTATC AAGGACTTCAT	68

Supplementary Table

Table S7. Oligonucleotide sequences used for amplification of the DNA flanking the off-target sites for NGS analysis.

sgRNA	Name	Sequence	Tm (°C)
SlDWF5-1 sgRNA_Sl1	OT-1_F	CGTTGCACTCCACCCA	60
	OT-1_R	CCCCGTTTATTCTCTTGAGCT	60
	OT-2_F	GGACAGTTGAAGGAGAAGTTAGTG	60
	OT-2_R	CAACTGGGGCCATAAAGTCAATG	60
	OT-3_F	GTCGCGGCTCATAATCGAGA	60
	OT-3_R	TCATAGCAAAGTACATGGATATAGCATCAA	60
	OT-4_F	AGTCATGGTAGTTAAGTGCACCG	60
	OT-4_R	TGTCTGCGACGTCTGAGAG	60
	OT-5_F	GAGGGGAGAAAGACCCCT	60
	OT-5_R	GCTCAATGAGGCCACGAC	60
SlDWF5-1 sgRNA_Sl2	OT-1_F	AGGCCTGATTTGAATCTTGTACTTG	60
	OT-1_R	GGTGTTCCAGTAACCTGCT	60
	OT-2_F	CGGGTCTGTTTGGATCAGTTATT	60
	OT-2_R	CTCTATAGGTTTAGCTAGCACGGA	60
	OT-3_F	GCGCCTGCTGAAATGACA	58
	OT-3_R	GTCCAGCCTAACTACCCCG	58
	OT-4_F	GGTCTCCTAAACGGAATGAGC	58
	OT-4_R	GAGTTAGGGAATGCGACGT	58
	OT-5_F	AGGGTCCTTAGAAAATACATGCAC	58
	OT-5_R	GGGAGAACTAGGGAGATGAACATTC	58
	OT-6_F	GGGCGGGTTTCTAGAAGCA	60
	OT-6_R	CGGACTTGAGCTGCTCC	60
SlDWF5-1 sgRNA_Sl1	Amplicon_OT-1_F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCACACGCTCAGTAGGTTAATCATAG	58
	Amplicon_OT-1_R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCCCTGTGGGGCTACTTTTAGA	58
	Amplicon_OT-2_F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGGTACAAGTGTTGCAAGAGT	58
	Amplicon_OT-2_R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTGCGAAATCGGAGAACCATG	58
	Amplicon_OT-3_F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGACTATTGCTTTTAATGTGCAAAATCGATTG	58
	Amplicon_OT-3_R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTCATAGCAAAGTACATGGATATAGCATCAA	58
	Amplicon_OT-4_F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCCAGCTGGGTTGACAGA	58
	Amplicon_OT-4_R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGGACCCCCTCCTAAGGC	58
	Amplicon_OT-5_F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCATGTCCGAGAAGCCAAAA	58
	Amplicon_OT-5_R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCCCAACCAGGTGGATATATCCTG	58
SlDWF5-1 sgRNA_Sl2	Amplicon_OT-1_F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGGGATGGAGTTGTATCC TCG	58
	Amplicon_OT-1_R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCCTTCATTCTCACTTTCGAGCAC	58
	Amplicon_OT-2_F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCATAACCAAATCAAATACTATCGGTTTTCAAAAA	58
	Amplicon_OT-2_R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGAATTCAAAGATGAGTAACAGTTTCTTGGG	58
	Amplicon_OT-3_F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGTGGAAGAATTACAATGTTTATTTCTTCTCTAC	68
	Amplicon_OT-3_R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGTAGTGAGTTTCTAACGAAGAAATTTGAATTC	68
	Amplicon_OT-4_F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGAAGCAACACTCGACGAC	68
	Amplicon_OT-4_R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACCATGTACATGATGGATCATTTTGATC	68
	Amplicon_OT-5_F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCATCCTTGTACCTAACCTAGGATGG	68
	Amplicon_OT-5_R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCCAAAACGTCCCTCCTTAATTGATTAAAAA	68
	Amplicon_OT-6_F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCGACACCAGATTTCATGTAGATGACA	68
	Amplicon_OT-6_R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCAACACACGCGCATGCATTC	68