Multiplex CRISPR/Cas9-mediated metabolic engineering increases soya bean isoflavone content and resistance to soya bean mosaic virus

Peipei Zhang1,†, Hongyang Du1,2,†, Jiao Wang1,†, Yixiang Pu1, Changyun Yang1, Rujuan Yan1, Hui Yang1,3, Hao Cheng1,* and Deyue Yu1,3,*

1National Center for Soybean Improvement, National Key Laboratory of Crop Genetics and Germplasm Enhancement, Jiangsu Collaborative Innovation Center for Modern Crop Production, Nanjing Agricultural University, Nanjing, China
2Key Laboratory of Rice Genetic Breeding of Anhui Province, Rice Research Institute, Anhui Academy of Agricultural Science, Hefei, China
3School of Life Sciences, Guangzhou University, Guangzhou, China

Received 17 February 2019; accepted 18 November 2019.
*Correspondence (Tel 86-25-84399530; fax 86-25-84396410; email jenny21star@njau.edu.cn) and (Tel 86-25-84396410; fax 86-25-84396410; email dyu@njau.edu.cn)
†These authors contribute equally to this work.

Summary
Isoflavonoids, which include a variety of secondary metabolites, are derived from the phenylpropanoid pathway and are distributed predominantly in leguminous plants. These compounds play a critical role in plant–environment interactions and are beneficial to human health. Isoflavone synthase (IFS) is a key enzyme in isoflavonoid synthesis and shares a common substrate with flavanone-3-hydroxylase (F3H) and flavone synthase II (FNS II). In this study, CRISPR/Cas9-mediated multiplex gene-editing technology was employed to simultaneously target GmF3H1, GmF3H2 and GmFNSI-1 in soya bean hairy roots and plants. Various mutation types and frequencies were observed in hairy roots. Higher mutation efficiencies were found in the T0 transgenic plants, with a triple gene mutation efficiency of 44.44%, and these results of targeted mutagenesis were stably inherited in the progeny. Metabolomic analysis of T0 triple-mutants leaves revealed significant improvement in isoflavone content. Compared with the wild type, the T-g generation homozygous triple mutants had approximately twice the leaf isoflavone content, and the soya bean mosaic virus (SMV) coat protein content was significantly reduced by one-third after infection with strain SC7, suggesting that increased isoflavone content enhanced the leaf resistance to SMV. The isoflavone content in the seeds of T0 triple mutants was also significantly increased. This study provides not only materials for the improvement of soya bean isoflavone content and resistance to SMV but also a simple system to generate multiplex mutations in soya bean, which may be beneficial for further breeding and metabolic engineering.

Keywords: Soya bean, Isoflavonoid, Multiplex gene editing, Metabolic engineering, Soya bean mosaic virus.

Introduction
Soya bean [Glycine max (L.) Merr.] is an important legume crop that is grown widely in various agroecological environments worldwide. It provides not only high-quality vegetable oils and proteins but also a variety of physiologically active substances that are beneficial to humans. Isoflavonoids are important group of secondary metabolites synthesized predominantly in leguminous plants. Notably, the content of isoflavonoids in soya bean is approximately 100 times that in other legumes (Yu et al., 2003). These compounds provide a wide range of functions; for instance, as common constituents of the human diet, they have vital functions in human health, such as reducing the incidence of specific cancers (Budryn et al., 2018; Malloy et al., 2018) and improving cardiovascular disease risk markers (Sathyalapan et al., 2018). Furthermore, they play important roles in plant disease resistance (Cheng et al., 2015; Graham et al., 2007; Kant et al., 2019; Liu et al., 2017) and as signal molecules for the induction of nod genes (Algar et al., 2014; Subramanian et al., 2006). Soya bean mosaic virus (SMV) disease is one of the most harmful diseases of soya beans; it seriously harms soya bean yield and reduces soya bean quality (Hill and Whitham, 2014). Previous studies have shown that polymorphisms of the soya bean GmIFS1 are associated with SMV resistance (Cheng et al., 2010); however, the relationship between isoflavones and SMV resistance needs further confirmation.

The biosynthesis of isoflavonoids is a complex process that is derived from the phenylpropanoid pathway. To date, more than 12 isoflavones have been identified in soya bean; the main components are daidzein, genistein and glycitein (Li and Zhang, 2017). The key step at the beginning of the isoflavonoid metabolic pathway is the hydroxylation of flavanone (naringenin, etc.) at the C-2 position, catalysed by a CYP450 enzyme isoflavone synthase (IFS). Two IFS (GmIFS1 and GmIFS2) have been isolated from soya bean (Jung et al., 2000). IFS competes for the naringenin substrate with flavone synthase (FNS) and flavanone-3-hydroxylase (F3H) (Figure S1) (Liu et al., 2002; Yu et al., 2000). In addition, FNS includes two fundamentally different enzymatic systems, FNS I and FNS II (Martens and Mithofer, 2005). FNS I is limited to a few umbelliferous species, such as celery and parsley (Gebhardt et al., 2005; Martens et al., 2001). FNS II is widely found in plants (Gebhardt et al., 2005). Two FNS II (GmFNSII-1 and GmFNSII-2) have been identified in soya bean (Jiang et al., 2010; Yan et al., 2014). F3H belongs to...
the 2-ketopamyl dioxygenase (2-ODD)-dependent family (Holton, 1995), and four copies exist in soya bean, two of which (GmF3H1 and GmF3H2) have been characterized (Zabala and Vodkin, 2005).

Metabolic engineering methods have been used for isoflavone content improvement by either expressing native/heterologous key regulators of the biosynthesis pathway or suppressing the expression of enzymes in competing metabolic pathways. The coexpression of Arabidopsis AtMYB12 and soya bean GmMy51 in tobacco led to enhanced isoflavone biosynthesis (Pandey et al. 2014). Studies have shown that there is a competitive relationship between IFS and branch pathway enzymes such as F3H, FNS and dihydroflavonol reductase (DFR) in the phenylpropanoid pathway (Jiang et al., 2010; Liu et al., 2002; Yu et al., 2000). In tobacco, when F3H expression was suppressed by RNAi while soya bean IFS was overexpressed, the yield of genistein was approximately 11 times higher than that of plants overexpressing only IFS (Liu et al., 2007). Overexpression of soya bean IFS in the Arabidopsis F3H and DFR double-mutant tt6/tt3 resulted in an increase in approximately 5- to 30-fold in genistein content (Liu et al., 2002). The coexpression of Arabidopsis AtMYB12 and Lotus japonicus LJF5 in tomato f3h mutant also led to enhanced biosynthesis of isoflavones (Zhang et al., 2015). When GmF3H-1 and GmFNSI-2 were silenced separately or simultaneously in soya bean hairy roots by RNAi, the isoflavone content was strongly increased (Jiang et al., 2014; Jiang et al., 2010).

Compared with suppressing gene expression by RNAi, targeted genome editing could be more precise and efficient. Both TALENs and the CRISPR/Cas9 system, especially the latter, have been widely used to modify target genes in recent years (Bak et al., 2018). Early-stage genome editing technologies have mainly been used to edit single genes, but single mutations cannot meet the needs of all studies, and the generation of multiplex mutations is required for investigating relationships among several related genes and the functions of gene families. To date, multiplex CRISPR/Cas9 systems have been developed in a variety of organisms (Minkenberg et al., 2017). These systems use different strategies to integrate multiple sgRNAs and Cas9 expression cassettes into a single binary vector. Previously, we developed an efficient single-target editing platform in soya bean using the CRISPR/Cas9 system (Du et al., 2016). To study traits such as isoflavone content that are controlled by multiple pathways and genes, a multiplex gene-editing platform is needed.

In this study, a multiplex gene-editing system was established in soya bean hairy root and then in soya bean transgenic plants by simultaneously targeting the GmF3H1, GmF3H2 and GmFNSI-1. Genotype analysis of the T0 generation plants and their descendents revealed the stable inheritance of the CRISPR-edited genes. Metabolomics analysis was employed to detect changes in isoflavonoids and other flavonoid profiles in multiplex CRISPR-edited soya bean plants. Compared with the wild type, the isoflavone content in the seeds of T2 generation triple-mutants transgenic plants and leaves of T2 generation homozygous triple-mutants transgenic plants was significantly increased, and the increasing isoflavone content in leaves enhanced soya bean resistance to SMV disease. This study provided not only new breeding materials for soya bean isoflavone content improvement and soya bean resistance to SMV enhancement but also a useful multiplex gene editing system for the study of traits controlled by multiple pathways or genes in soya bean.

Results

Design of CRISPR/Cas9 multiplex gene constructs to knock out GmF3H1, GmF3H2 and GmFNSI-1 simultaneously in soya bean

To increase the flow of the common substrate naringenin into the isoflavonoid metabolic pathway by blocking the branch pathways that compete with IFS, we set out to simultaneously knock out GmF3H1, GmF3H2 and GmFNSI-1 in soya bean. To generate Cas9-induced mutations in GmF3H1 and GmF3H2, three sgRNAs, that is sgRNA1 (sg1), sgRNA2 (sg2) and sgRNA3 (sg3), were designed to target the first exon (Figure 1a). sg1 precisely matched both F3H copies, whereas sg2 and sg3 matched well with GmF3H1 and GmF3H2, respectively (Figure 1a). Three sgRNAs, that is sgRNA4 (sg4), sgRNA5 (sg5) and sgRNA6 (sg6), that targeted the first exon of GmFNSI-1 were also designed (Figure 1a). All sgRNAs were expressed under the GmU3 or GmU6 promoter. Three binary constructs carrying four sgRNAs with dpcas9 driven by the GmUbi3 promoter were generated. Vector Tri-KO1 contained sg1, two sg4 and sg5; vector Tri-KO2 contained sg1, two sg4 and sg6; vector Tri-KO3 contained sg2, sg3 and two sg4 (Figure 1b).

Variety and efficiency of GmF3H1, GmF3H2 and GmFNSI-1 mutations in soya bean hairy roots

To estimate the efficiency of the multiplex gene-editing system, the three multiplex CRISPR/Cas9 vectors, Tri-KO1, Tri-KO2 and Tri-KO3, were transformed into soya bean hairy roots by the A. rhizogenes-mediated method. The mutation frequency was calculated as the number of hairy roots showing mutations divided by the total number of positive hairy roots. The transformation of the 144 hairy roots (48 for each vector) was confirmed by bar-specific primers. To test the mutation efficiency, all positive samples were sequenced after amplification with target-specific primers. For each target, the mutation frequency ranged from 10.42% to 31.25% (Table 1). For each hairy root, the frequency of single-gene mutations ranged from 4.17% to 12.50%, the frequency of double-gene mutations ranged from 8.33% to 16.67%, and the frequency of triple gene mutations ranged from 6.25% to 12.50%. To determine whether mutations in the three target genes were correlated or independent, we analysed the theoretical mutation rate and compared it with the observed value. The results showed that the observed number of hairy roots with co-mutation of all three genes deviated significantly from the expected number ($X^2_{12} = 113.83, P < 0.01$) by the Chi-square test (Table S4), suggesting that mutations at different targets tended to occur simultaneously.

To examine the variety of mutations in hairy roots, the amplified products of the mutant target genes, identified by sequencing chromatograms, were ligated into the pClone007 vector and sequenced. Various mutations, including insertions and deletions, were observed in the sgRNA targets. Combining the results of all targets, 13.64% of the mutations were insertions, 68.18% were deletions, and 18.18% were simultaneous deletions and insertions. Of the insertion mutations, most were 1 bp insertions, with a marked preference for A (53.85%, 7/13) over T (15.38%, 2/13), C (7.69%, 1/13) and G (23.08%, 3/13) nucleotides. Among the mutations, 73.86% (65/88) were missing less than 10 bp, while the remaining 26.14% (23/88) were longer deletions of more than 10 bp (Figure 54 and Table S5). The product with the largest deletion (294 bp) was due...
to simultaneous cleavage at both the sg5 and sg4 targets. The deleted chromosomal region encompassed the sg5 and sg4 target sequences and the sequence between them, representing a perfect ligation after Cas9 cleavage.

Variety and frequency of \( \text{GmF3H1} \), \( \text{GmF3H2} \), and \( \text{GmFNSII-1} \) mutations in T0 transgenic plants

To further confirm that the multiplex gene-editing system was suitable for soya bean, the Tri-KO1 vector was transformed into soya bean using \( \text{Agrobacterium} \)-mediated cot-node transformation. A total of 27 independent transgenic T0 lines were generated according to the \( \text{bar} \) strip test. The DNA of the 27 lines was isolated and amplified with target-specific primers and sequenced. The mutation frequency was calculated as the number of plants showing mutations divided by the total number of T0 plants. We found that the frequencies of the \( \text{GmF3H1} \), \( \text{GmF3H2} \) and \( \text{GmFNSII-1} \) mutations were 55.56%, 48.15% and 81.48%, respectively (Table 1). The frequencies of single, double and triple mutations were 37.03%, 7.41% and 44.44%, respectively (Table 1). These efficiencies were higher than those in hairy roots. Both \( \text{GmF3H1} \) and \( \text{GmF3H2} \) were targeted by sg1, and 13 T0 lines showed mutations in both, while two lines carried

Table 1 Modification efficiency of transformation events induced by the multiplex CRISPR/Cas9 system

| Transformation type | Vector ID | Total number of transformations | Target gene | Mutation frequency* | Single mutation frequency* | Double mutation frequency* | Triple mutation frequency* |
|--------------------|-----------|--------------------------------|-------------|---------------------|---------------------------|---------------------------|---------------------------|
| Hairy root         | Tri-KO1   | 48                            | \( \text{GmF3H1} \) | 20.83% (10)         | 4.17% (2)                 | 8.33% (4)                 | 12.50% (6)                |
|                    | Tri-KO2   | 48                            | \( \text{GmF3H1} \) | 20.83% (10)         | 12.50% (6)                | 16.67% (8)                | 6.25% (3)                 |
|                    | Tri-KO3   | 48                            | \( \text{GmF3H1} \) | 25.00% (12)         | 8.33% (4)                 | 12.50% (6)                | 8.33% (4)                 |
| Cotyledon†         | Tri-KO1   | 27                            | \( \text{GmF3H1} \) | 55.56% (15)         | 37.03% (10)               | 7.41% (2)                 | 44.44% (12)               |

*The number of mutations is presented in parentheses.
†Cotyledon represents the stably transformed T0 plants.
mutations only in GmF3H1, implying that sg1 tended to target both genes simultaneously. Similarly, the Chi-square test ($\chi^2 = 22.90, P < 0.01$) suggested that mutations at the three different genes tended to occur simultaneously, which was consistent with the result in hairy roots (Table S4).

To assess the variety of mutations in T0 transgenic plants, four T0 transgenic plants (T0-3-1, T0-10-1, T0-16-1 and T0-17-2) whose three genes were simultaneously mutated were used for PCR amplification of the targets and sequenced. Similar to the results in hairy roots, mutations including both insertions and deletions were observed (Fig. 2 and Table S5). Combining the results at all gene loci, 8.51% of the mutations were insertions, 87.23% were deletions, and 4.26% were simultaneous deletions and insertions. Of the insertion mutations, 100% were 1 bp insertions. Most mutations (82.98%, 39/47) were missing less than 10 bp in length, while the remaining 17.02% (8/47) were longer deletions of more than 10 bp.

Mutations by the CRISPR/Cas9 multiplex gene editing system can be efficiently inherited

Of the 27 transgenic T0 plants, 24 plants contained mutations, and the remaining three had no editing at any target. To test the heritability of the mutations, a total of 231 T1 plants derived from 17 T0 lines were examined (Table 2). Wild-type plants and homozygotes could be clearly identified by sequencing chromatograms, while bialleles, heterozygotes and chimeras could not be distinguished and were treated as heterozygotes for the sake of convenience. In the T1 generation, some plants carried mutant genes that were not detected in the T0 lines, and triple gene editing was found in the descendants of T0 lines with double (T0-19-4) and single mutations (T0-15-1). This result indicated that Cas9 can continue to function in descendants. The segregation patterns of the T1 generation were diverse and unpredictable. A total of 134 T1 plants contained mutant genes, including single,
average expression of GmF3H1 mutant T0 and wild-type plants. The results showed that the GmF3H1 expression, qRT-PCR was used to detect the expression levels of isoflavone content in the phenylpropanoid pathway.

Metabolomic analysis revealed significant increase of inherited by descendants.

Table 2 Positive identification of T1 transgenic plants

| T0 genotype | T0 plant no. | AABbCC | AABbCC/AABbcA | AaBbCc | AaBbCc/AABbcA | AaBbCc/AaBbCc/AaBbCc/AaBbcc/AABBCC | AABbCc/AaBbCc/AABBCC | Total |
|-------------|--------------|--------|----------------|--------|----------------|--------------------------------------|------------------------|-------|
| AABbCc      | T0-3-1       | 6      | 5/0            | 2      | 1              | 2/0                                  | 0                      | 23    |
|             | T0-10-1      | 4      | 0              | 0      | 0              | 1/0                                  | 0                      | 18    |
|             | T0-16-1      | 1      | 0              | 0      | 1              | 2/0                                  | 6/0/0/0                | 2     |
|             | T0-17-2      | 3      | 2/0            | 0      | 2              | 0                                    | 4/0/0/0                | 5     |
|             | T0-17-3      | 0      | 0              | 0      | 0              | 0/2                                  | 6/4/1/1                | 0     |
|             | T0-18-1      | 0      | 0              | 0      | 0              | 0                                    | 0/0/1/0                | 0     |
|             | T0-22-1      | 0      | 0              | 0      | 0              | 1                                    | 2/1/0/0                | 0     |
|             | T0-17-1      | 8      | 0              | 0      | 0              | 0                                    | 0                      | 20    |
|             | T0-14-2      | 0      | 0/3            | 0      | 1              | 0                                    | 7/1/2/0/1             | 0     |
|             | T0-15-3      | 4      | 0              | 0      | 0              | 0                                    | 0                      | 9     |
| AABbCC      | T0-19-4      | 0      | 1/1            | 0      | 3              | 0                                    | 4/0/1/0                | 0     |
|             | T0-22-2      | 0      | 0              | 0      | 1              | 0                                    | 0                      | 1     |
| AABbCc      | T0-22-3      | 0      | 0              | 0      | 0              | 1/0                                  | 0                      | 1     |
|             | T0-15-4      | 0      | 5/2            | 0      | 1              | 2/4                                  | 0                      | 0     |
|             | T0-19-1      | 4      | 1/0            | 0      | 0              | 1                                    | 0                      | 10    |
|             | T0-15-5      | 1      | 0              | 0      | 1              | 1/0                                  | 1/0/0/0                | 9     |
|             | T0-15-2      | 0      | 7/3            | 0      | 0              | 2/2                                  | 0                      | 0     |
| Total       | 31            | 30     | 2              | 12     | 17             | 42                                   | 97                     | 231   |

a, b and c represent GmF3H1-, GmF3H2- and GmFNSII-1-edited alleles, respectively. #/#/#/# represents the numbers of various genotypes corresponding to the headers in the same columns.

double and triple mutations (Table 2). Fourteen T1 plants contained homozygous GmFNSII-1 mutations, six T1 plants contained homozygous GmF3H2 mutations, and one T1 plant contained homozygous GmF3H1 and GmFNSII-1 double mutations. Among these mutant, 68 'transgene-clean' plants containing no transgenic elements were detected (Table S6). These 'transgene-clean' plants contained single, double and triple mutations, making them ideal materials for the further study of gene function and the molecular mechanism of isoflavone biosynthesis.

To verify that the edited genes could be inherited, we cloned the mutated genes of eight T1 generation plants derived from T0-17-2 and T0-16-1 (each had four descendants). The descendants had the same types of mutations as the previous generation, and some new types also appeared (Figure 3 and Table S8). Different T1 lines from the same T0 generation contained the same new mutations, implying that they could have originated in the T0 generation, although we did not detect them then. To summa- 

Metabolomic analysis revealed significant increase of isoflavone content in the phenylpropanoid pathway

To verify whether the gene mutations caused changes in gene expression, qRT-PCR was used to detect the expression levels of GmF3H1, GmF3H2 and GmFNSII-1 in the leaves of four triple-mutant T0 and wild-type plants. The results showed that the average expression of GmF3H1, GmF3H2 and GmFNSII-1 in the transgenic plants was significantly decreased compared to that in the wild type (Figure S5). To detect metabolic changes in leaves, flavonoid metabolites were profiled using a widely targeted metabolic method, and a total of 217 flavonoids were detected (Tables S9 and S10). Based on partial least squares discriminant analysis, 64 differential metabolites were detected, among which 34 were significantly decreased and 30 were significantly increased in the triple mutants. The content of the direct synthetic products of F3H, such as dihydrokaempferol, dihydroquercetin and dihydromyricetin, did not change significantly in the triple mutants (Figure 4 and Table S10), but some of their flavonol derivatives were significantly decreased, such as astragalin, kaempferin and myricitrin. In addition, the levels of six anthocyanins, including delphinidin, were significantly decreased. FNS catalyses flavonanes to generate flavones. The content of apigenin did not change significantly, while its derivatives rhoifolin and isorhoifolin were significantly decreased (Figure 4 and Table S10). Moreover, FNS catalyses eriodictyol to produce luteolin. Although eriodictyol was increased significantly by 128.06-fold, the content of luteolin did not change significantly. Flavone C-glycosides represent a group of oxygen heterocyclic compounds in which the sugar group and the aglycone are linked by a C-glycosidic bond. Among the detected flavone C-glyco- 

As expected, compared with that in wild-type plants, naringenin was significantly increased by 12.07-fold in the triple mutants (Figure 4 and Table S10). Furthermore, the naringenin precursor naringenin chalcone was also significantly increased by 11.99-fold. The increase in naringenin could cause increases in the downstream products of the competing branches of F3H and FNS. For example, the flavones prunin, hesperetin and eri- 

© 2019 The Authors. Plant Biotechnology Journal published by Society for Experimental Biology and The Association of Applied Biologists and John Wiley & Sons Ltd., 18, 1384–1395
genistin, prunetin, biochanin A and sissotrin increased by 4.74-, 9.17-, 7.79- and 7.38-fold, respectively. These results suggested that the isoflavone content in the leaves of T0 plants were significantly increased after multiple branch pathways were knocked out by the multiplex gene-editing system.

Increased isoflavone content enhanced the leaf resistance to SMV

Isoflavonoids play important roles in plant disease resistance (Graham et al., 2007), and polymorphisms of soya bean GmIFS1 were associated with SMV resistance (Cheng et al., 2010). To explore whether the increased isoflavone content could enhance leaf resistance to SMV, we detected the isoflavone content and SMV coat protein (cp) content of leaves that were infected with strain SC7, a virulent strain prevalent in the Huang-Huai-Hai valleys in China (Cai et al., 2014). Two homozygous triple-mutants lines, T2-17-3 and T2-22-1, were isolated in the T2 generation (Table S7). In their offspring, seedlings of T3 generation were inoculated with SC7 inoculum (SC7) at the primary leaves at the VC-cotyledon stage, and the control groups were either inoculated with phosphate buffer (Mock) or noninoculated (NI). In the NI group, the expression levels of GmF3H1, GmF3H2 and GmFNSII-1 were significantly decreased in the triple mutants compared with the wild type (Figure S7a). GmIFS1, GmIFS2 and GmCHIA did not change significantly, except that the expression levels of GmIFS2 and GmCHIA were increased in T3-22-1. In the SC7 group, the expression levels of GmF3H1, GmF3H2 and GmFNSII-1 were also significantly decreased in mutants compared with the wild type. The GmIFS1, GmIFS2 and GmCHIA were significantly increased, except that the expression levels of GmIFS2 did not change significantly in the T3-17-3 line.

Compared with the wild type, the triple mutants had significantly increased isoflavone content, regardless of whether the NI group was inoculated with phosphate buffer or virus (Figure 5a). In the NI group, the average isoflavone content of the wild type was 1393.60 \( \mu \)g/g, while the average content of the triple mutants was 2859.38 \( \mu \)g/g. In the mock group, the average isoflavone content of the wild type was 1325.13 \( \mu \)g/g, while the average content of the triple mutants was 3302.55 \( \mu \)g/g. In the SC7 group, the average isoflavone content of the wild type was 1600.26 \( \mu \)g/g, while the average content of the triple mutants was 3094.32 \( \mu \)g/g. Compared with NI, there was no significant change in the isoflavone content after inoculation, suggesting that the content of isoflavones was not affected by inoculation with phosphate buffer or virus after 21 days (Figure 5a). We further examined the SMV coat protein (cp) content after virus inoculation. Compared with the wild type, the SMV cp content of triple mutants decreased significantly (Figure 5b). The average SMV cp content of T3-17-3 and T3-22-1 was 25.49 \( \mu \)g/mg and 24.31 \( \mu \)g/mg, respectively, while the average SMV cp content of

---

**Figure 3**  Mutation types of T0 plants (T0-17-2) and their descendants (PAM in green). T1-17-2-1, T1-17-2-2, T1-17-2-3 and T1-17-2-4 represent four individual descendants from T0-17-2, respectively. Bold text represents mutations consistent with the T0 generation, ‘-’ represents base pairs (bp) deleted from the target site, and red font represents inserted base pairs.
the wild-type plants was 36.59 µg/mg, suggesting that the proliferation or spread of the virus in the mutants was inhibited (Figure 5b). These results implied that the increased isoflavone content enhanced the leaf resistance to SMV.

Figure 4 Schematic diagram of isoflavones and their competing metabolic pathways. The green box represents flavonol and anthocyanin metabolic pathway compounds. The yellow box represents flavone metabolic pathway compounds, and the purple box represents isoflavone metabolic pathway compounds. Blue text represents a significant decrease in the substance, and red text represents a significant increase in the substance. Gray text indicates that the substance was not detected. Multiple arrows in sequence represent omitted metabolites.

Figure 5 Isoflavone content and soya bean mosaic virus coat protein content in the leaves of T3 homozygous triple mutants and wild type. (a) Isoflavone content in leaves of the non-inoculated group (NI), group inoculated with phosphate buffer (Mock) group inoculated with virus (SC7). (b) soya bean mosaic virus coat protein content in inoculated leaves with SC7 inoculum after 21 days. (Student’s t-test; **P < 0.001; two-tailed t-test).
Isoflavone content in the seeds of T2 generation triple mutants was increased

To verify the phenotypes in the transgenic lines, analyses were made of isoflavone content in the beans of T2 generation triple mutants. Compared with the wild type, five triple mutants had significantly increased seed isoflavone contents, with the exception of T2-17-2 (Figure S6). We further measured the 100-grain weight, protein and oil content, and most transgenic lines showed no significant change compared with the wild type (Figure S6). The exception was a significant increase in the protein content and decrease in the oil levels of T2-18-1.

We measured the expression levels of several genes in the seeds at three different developmental stages of T3 homozygous triple mutants (Figure S8). The expression levels of GmF3H1 and GmF3H2 gradually decreased with the development and maturity of the seeds, while the expression levels of GmIFS1, GmIFS2 and GmCHI gradually increased. Compared with the wild type, the expression levels of GmF3H1, GmF3H2 and GmIFSII-1 were significantly decreased in the seeds at different developmental stages except in the seeds of T3-17-3, which showed no significant changes in the seeds with lengths of 7 and 10 mm. Compared with the wild-type levels, GmIFS1 did not change significantly during different developmental stages of seeds, while the expression of GmIFS2 decreased to varying degrees. The expression levels of GmCHIA did not change significantly at different stages of seed development except in the seeds of T3-17-3, which showed a significant decrease in GmCHIA expression in 10 mm seeds.

Discussion

Efficient genome editing in soya bean using the CRISPR/Cas9 multiplex gene editing system

Targeted gene editing has been one of the most sought-after technologies for fundamental research and biotechnology. Since the advent of the CRISPR/Cas9 system, gene editing can be easily achieved and is widely used in plants (Ma et al., 2016; Yin et al., 2017). Previous work in our laboratory has demonstrated that targeted mutagenesis can be successfully achieved through both TALENs and the CRISPR/Cas9 system in soya bean (Du et al., 2016). Multiplex gene editing has become a research hotspot in recent years. Six PYLS of Arabidopsis were simultaneously targeted by a multiplex CRISPR/Cas9 platform (Zhang et al., 2016). In rice, multiplex gene editing has been reported several times (Lan et al., 2017; Shen et al., 2018; Wang et al., 2017; Wang et al., 2018b; Wang et al., 2015; Xu et al., 2016; Zhou et al., 2018). In soya bean, several QTLs for agronomic traits, including growth habit, flowering time and salinity tolerance, have been positionedally cloned. The multiplex CRISPR/Cas9 platform provides an opportunity to simultaneously manipulate different agronomic traits. Previously, a CRISPR/Cas9 system that allowed two sgRNA targeting sites was used to create a biallelic double mutant for the two soya bean paralogous double-stranded RNA-binding2 (GmDrb2a and GmDrb2b) genes (Curtin et al., 2018). Recently, the CRISPR/Cas9-mediated targeted mutagenesis of GmSPL9, which were expressed using Arabidopsis U3 and U6 promoters, altered plant architecture in soya bean (Bao et al., 2019).

Here, we developed a multiplex CRISPR/Cas9 platform that allows the application of four or more sgRNAs with a single binary vector, with two different soya bean Pol III-dependent promoters (GmU3 and GmU6) driving the expression of the individual sgRNAs. To evaluate the gene-editing efficiency of the multiplex CRISPR/Cas9 system, GmF3H1, GmF3H2 and GmIFSII-1, which encode key enzymes in the competing metabolic pathways of isoflavone biosynthesis, were selected as targets for mutagenesis. The mutation frequency of each target ranged from 10.42% to 31.25% in hairy roots (Table 1). In each edited hairy root, various mutation types were found, including single, double and triple mutations (Table 1). We found that the observed frequency of the triple mutation deviated significantly from the expected frequency, suggesting that mutations at the three different targets tended to occur simultaneously. This result implies that these synchronous mutations were not independent of each other but were correlated with common factors, such as the activity of the Cas9 protein. The mutation frequencies of each target in the T0 transgenic plants ranged from 48.15% to 81.48% (Table 1), significantly higher than those of hairy roots, perhaps due to the difference in culture time and inoculated tissue. The frequency of triple mutations in plants was 44.44% (Table 1), which was also significantly higher than that in hairy roots. In this work, most detected mutations were small deletions, and a few were insertions (Figure 2 and Figure S4), which is consistent with the results of our previous study and others (Du et al., 2016; Wang et al., 2018a).

The predictable inheritance and segregation of genome modifications in later generations is highly desirable for basic research as well as molecular breeding. In the observed descendants of seventeen T2 plants (Table 2), more than half (134/231) carried different numbers of mutated genes. By detecting the mutation types of two T0 plants (T0-17-2, T0-16-1) and their descendants, we found that the edited genes of T0 plants can be passed to the next generation. In addition, some mutations that were not detected in the T0 generation were found in T1 plants. One possible reason is that the T0 plants were chimeric, and the WT copies of the target genes could have continued to mutate in either the T0 or T1 generation. The existence of transgenic elements greatly increases the risk of potential off-target effects and impedes the commercial application of CRISPR-edited crops. In this study, the examined descendants of seventeen T0 plants (Table 2) included more than two-thirds (162/231) that were bar negative (Table S6), implying that transgenic elements might readily be lost from the T0 to T1 generation. Three plants of the 42 heterozygous triple mutants in the T1 generation lacked any transgenic elements. These transgene-clean T1 plants yielded homozygous triple mutant (T2-17-3) in the T2 generation.

Multipathway metabolic engineering is effective for improving soya bean isoflavone content and resistance to SMV

Metabolic engineering is aimed at changing the metabolic composition of the cell, and when applied to plants, food, feed quality and resistance improvements are the main goals. Metabolic engineering strategies change the metabolic composition by introducing biosynthetic genes, using transcription factors and improving metabolic flux (Fu et al., 2018). Overexpressing key regulators in biosynthetic pathways is an effective strategy to enhance the yield of desired products. Another strategy is to suppress competing metabolic pathways. Using a multiplex pYLCRISPR/Cas9 system that targeted five key genes, the γ-amino butyric acid (GABA) shunt was converted in tomato (Li et al., 2018a), and lycopene was enriched in tomato fruit by
knocking down five genes associated with the carotenoid metabolic pathway (Li et al., 2018c).

Here, we developed a multiplex CRISPR/Cas9 platform to target the competing metabolic pathways of isoflavones. A metabolomic assay of the leaves from four transgenic plants revealed changes in the content of substances in the upstream, downstream, and competing pathways, resulting in manipulated isoflavone biosynthesis (Figure 4). Isoflavones, including glycinein, daidzein and genistein, were synthesized by the catalytic action of IFS to liquiritigenin and naringenin. This study improves the flow of genistein by knocking out GmF3H1, GmF3H2 and GmFNSII-1 of the genistein competing pathway. Our greatest concern is the increase in isoflavones in soya bean seeds and the impact of multiplex gene editing on other traits, such as resistance. The content of isoflavones in the seeds of five T$_2$ generation triple mutants was significantly increased, except for T$_2$-17-2, which may be due to a chimeric or heterozygous genotype. The isoflavone competition pathway and the synthetic pathway genes showed opposite trends with seed development. Studies have shown that the content of isoflavones gradually increases with the development and maturity of the seeds (Chu et al., 2017), and this trend was consistent with our gene expression levels. These results showed that the method of knocking out the isoflavone competition pathway genes is effective for increasing the content.

Soya bean mosaic virus seriously reduces soya bean yield and quality, and in pursuit of improving the resistance of soya bean to SMV disease, strain-specific resistance genes such as Rsx1, Rsx3 and Rsx5 have been discovered (Tran et al., 2018; Wu et al., 2019). However, soya bean resistance to SMV is a quantitative trait controlled by multiple genes; most disease resistance genes are trivial, and the evolution of different physiological races makes the resistant materials lose resistance. Therefore, it seems more effective to use the metabolites of the plants themselves, such as lignin and isoflavones, to resist invasion by pathogenic bacteria. Previous studies have found that lower isoflavone content leads to a decrease in soya bean resistance to phytophthora (Graham et al., 2007). Our results showed that increasing isoflavones can improve soya bean resistance to SMV.

In conclusion, a CRISPR/Cas9-mediated multiplex gene-editing system was established for soya bean, and a high mutation efficiency was found in the T$_0$ plants. The edited genes could be inherited from the T$_0$ to the T$_1$ and later generations. Three genes in the isoflavone competition pathway were knocked out by this system, resulting in significantly increased isoflavone content in the leaves and seeds of transgenic plants, and higher isoflavone content enhanced soya bean resistance to SMV. All the data in this study indicated that the multiplex CRISPR/Cas9 system is a powerful tool for soya bean improvement through targeting multiple genes and the study of metabolic engineering by simultaneously regulating multiple metabolic pathways.

**Experimental procedures**

**Plant material and growth conditions**

Soya bean cotyledons of the cultivar ‘Jack’ were used as the transformation recipient in this study. All seeds were provided by the National Center for soya bean Improvement (Nanjing, China). T$_0$, T$_1$ and T$_2$ plants were grown in a medium in which the proportion of vermiculite and nutrient soil was 1:1 under a 16 h light/8 h dark cycle at 25 °C. To avoid the impact of soil-borne diseases, T$_3$ plants that were used for the SMV inoculation experiments were grown in pots containing sand in the net room.

**Vector construction for multiplex gene editing**

To construct multiplex sgRNA modules for gene editing in soya bean, we used a step-by-step assembly strategy based on a vector in our laboratory (Du et al., 2016) that contained the GmU6-16g-1 promoter. Here, we added a GmU3 (GmU3-19g-1) promoter to regulate the expression of additional sgRNAs. The promoter sequences of GmUBi3 and GmU3-19g-1 were amplified from the soya bean cultivar ‘Williams 82’ (Sequence). We developed a binary vector that contains a dicotyledon codon-optimized dpCas9 regulated by the GmUBi3 promoter (Hernandez-Garcia et al., 2010), referred to as pGmUBi-Cas9 (Figure S2). Two scaffold vectors that contain the sgRNA module were developed: pU6-sgR and pU3-sgR (Figure S2). Using these three vectors (pU6-sgR, pU3-sgR and pGmUBi-Cas9) and a step-by-step strategy, up to four sgRNAs can be assembled in a single carrier plasmid (Figure S2).

Four F3H copies are present in the soya bean genome and are hereafter called GmF3H1 (Glyma.02G0484000), GmF3H2 (Glyma.02G0486000), GmF3H3 (Glyma.16G1287000) and GmF3H4 (Glyma.01G1662000). GmF3H1 and GmF3H2 were selected as the targets. Of GmFNSII-1 (Glyma.12G0670000) and GmFNSII-2 (Glyma.12G0671000), only GmFNSII-1 was selected as a target. The target sequences of the sgRNA were identified by the web-based tools CRISPR-P (Lei et al., 2014) and CRISPR-PLANT (Xie et al., 2014). To construct a pGmUBi-Cas9-4XsgR vector targeting GmF3H1, GmF3H2 and GmFNSII-1, complementary 23-bp oligonucleotides with a 20-bp target sequence were synthesized (Table S1). These oligo pairs were annealed to generate double-strand DNA with 3-bp overhangs on both ends and cloned into the BsmBI sites of the vector pU6-sgR or pU3-sgR. The GmU6pro-gRNA and GmU3pro-gRNA modules were amplified from their host vectors by PCR using specific primers (Table S2, pU3-F/R and pU6-F/R) and then simultaneously digested and ligated into the pGmUBi-Cas9 vector (digested with NcoI and XbaI) to obtain pGmUBi-Cas9-2XsgR (Figure S2). To generate the final quadruple CRISPR/Cas9 construct, the two tandem sgRNA modules in one pGmUBi-Cas9-2XsgR vector were amplified by PCR using specific primers (Table S2, F/R), and the amplification product was recombinant ligated into another pGmUBi-Cas9-2XsgR vector (digested with BstEI). Finally, the pGmUBi-Cas9-4XsgR vector for quadruple gene editing was obtained (Figure S3).

**Soya bean hairy root transformation**

Each pGmUBi-Cas9-4XsgR vector was independently transformed into Agrobacterium rhizogenes strain K599 for hairy root transformation. ‘Jack’ seeds were inoculated with the transformed K599 with few modifications from the protocol as previously described (Du et al., 2016). Briefly, intact seeds were surface sterilized in Cl$_2$ gas and then germinated on SG4 medium for 6 days at 25 °C (16 h light/8 h dark). After germination, the cotyledons were separated, and 1-mm-deep cuts were made on the abaxial surfaces. The Agrobacterium rhizogenes solution was cultured to an OD 600 of 0.8-1.0, centrifuged and suspended in 10 mM MgCl$_2$ buffer (containing 50 μM acetosyringone) to an OD 600 of 0.5-0.6. A drop of the suspension was applied to the wounded side of each cotyledon, and the cotyledons were cocultivated with Agrobacterium on White medium at 25 °C in the dark. After two weeks, cotyledons with hairy roots were
transferred to a new White medium plate, followed by the collection of regenerated roots two weeks later.

**Agrobacterium-mediated transformation of soya bean cotyledons**

Stable transformation was performed using an optimized cotyledon transformation protocol (Du et al., 2016; Zeng et al., 2004). *Agrobacterium tumefaciens* strain EHA105 was chosen to transform the plasmid. Briefly, *Agrobacterium rhizogenes* solution was cultured to an OD 600 of 0.8–1.0, centrifuged and suspended in cocultivation medium (CCM (liquid)) to an OD 600 of 0.4–0.5. ‘Jack’ seeds were surface sterilized and germinated as previously described. The root that was approximately 10 mm below the cotyledonary node was removed, and the cotyledonary nodes were wounded and inoculated in the cocultivation suspension for 30 min at 28 °C and then transferred onto solid cocultivation medium CCM for cocultivation for 5 days at 25 °C under dark conditions. After cocultivation, the infected explants were washed in Washing Medium and subsequently cultured on SI medium for 2 weeks at 25 °C with a 16 h photoperiod. The explants were then transferred to fresh SI medium with 6 mg/L glufosinate for bar screening. After four weeks of culture on SI medium, the explants were transferred to SE medium to initiate shoot elongation. Selection during the shoot elongation stage was performed with gradients of 4 and 2.5 mg/L every two weeks. To initiate rooting, the bottom of each individual shoot was cut and then transferred to Rooting Medium. After two weeks, the rooted plantlets were transplanted to soil.

**Genomic DNA extraction and PCR/sequencing-based genotyping**

Genomic DNA was isolated from approximately 100 mg of hairy roots, T₀, T₁, and T₂ plant leaves using the DNAsecure Plant Kit (Tiangen Biotechnology (Beijing) Co., Ltd.; Catalog No. DP320; Beijing, China). The presence of bar was assessed by PCR using specific primers (Table S1). To identify mutations in the target genes, PCR was performed to amplify the genomic region surrounding the CRISPR target sites using specific primers (Table S1). The PCR products were directly sequenced or cloned into the pClone007 Blunt Simple Vector (Beijing TsingKe Biotech Co., Ltd.; Catalog No. TSV-007B; Nanjing, Jiangsu Province, China) and subsequently sequenced. A total of 50–200 ng of DNA and PhantaMax Super-Fidelity DNA Polymerase (Vazyme Biotechnology (Nanjing) Co., Ltd.; Catalog No. P505; Nanjing, Jiangsu Province, China) were used in the PCR system.

**Determination of expression levels by quantitative real-time PCR (qRT-PCR)**

RNA was isolated from approximately 100 mg of leaves or seeds using the RNAprep Pure Plant Kit (Tiangen Biotechnology (Beijing) Co., Ltd.; Catalog No. DP320; Beijing, China), and cDNA was synthesized using a HisScript® II 1st Strand cDNA Synthesis Kit (Vazyme Biotechnology (Nanjing) Co., Ltd.; Catalog No. R212-01/02). Gene expression was determined by qRT-PCR assays using an Applied Biosystems 7500 Real-Time PCR System (Life Technologies, USA) with ChamQ® SYBR® qPCR Master Mix (Vazyme Biotechnology (Nanjing) Co., Ltd.; Catalog No. Q331). The housekeeping gene *tubulin* (GenBank: AY907703.1) was used as a control. Three replicates were performed for each reaction. The normalized expression, reported as the fold change, was calculated for each sample with the 2−ΔΔCT method (Livak and Schmittgen, 2001). The qRT-PCR primers used are listed in Table S3.

**Metabolomic measurement**

The sample used in the metabolomic analysis was the same as the sample used for T₀ qRT-PCR. Sample extraction for metabolomic analysis was carried out as in previous studies (Chen et al., 2014). The sample extracts were analysed using a liquid chromatography-electrospray ionization-tandem mass spectrometry (LC-ESI-MS/MS) system (HPLC, Shim-pack UFLC Shimadzu CBM30A system; MS, Applied Biosystems 6500 Q TRAP; Wuhan, Hubei province, China). In brief, 5 μL extracts were injected into the high-performance liquid chromatography (HPLC) system. The effluent was connected to the ESI-triple quadrupole-linear ion trap (Q TRAP)-MS system. The conditions and operation parameters were set as in previous studies (Chen et al., 2014). Metabolite identification was based on the primary and secondary spectral data annotated against the self-compiled database MWDB (MetWare Biological Science and Technology Co., Ltd., Wuhan, China) and publicly available metabolite databases, including MassBank (http://www.massbank.jp/), KNAPSAck (http://kana.ya.naist.jp/KNAPSAck/), HMDB (http://www.hmdb.ca/), MoToDB (http://www.ab.wur.nl/moto/) and METLIN (http://metlin.scripps.edu/index.php). Metabolite quantification was carried out using multiple-reaction monitoring (MRM) mode. Partial least squares discriminant analysis (PLS-DA) was used to study the identified metabolites. Metabolites with significant differences in content were set with thresholds of variable importance in projection (VIP) ≥ 1 and fold change ≥ 2 or ≤ 0.5.

**Extraction and determination of isoflavone content**

Isoflavone content was determined with few modifications from the protocol as previously described (Chu et al., 2017). First, 0.0400 g of leaf powder or 0.0200 g of seed powder was weighed into 2.0 mL centrifuge tubes containing 800 or 1000 μL of 80% chromatographic methanol (solid–liquid ratio: 1:20 or 1:50). The mixture was spun for 1 min, and 50 °C ultrasound (frequency: 40 kHz, power: 300 W) was used to assist extraction for 1 h. After centrifugation at 21 499 g for 10 min at 4 °C, the supernatant was passed through a 0.22-μm organic phase needle filter and injected into a water autoinjection vial (2 mL), which was stored at −20 °C before use. Samples were analysed with an ultra-performance liquid chromatography (UPLC) system (National Key Laboratory of Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University. Specification model: Hclass, Column: Waters ACQUITY UPLC HSS T3 1.8 μm, 2.1 mm×100 mm column, Part No. 186003539, Serial No. 012631029257) under the following conditions: solvent A was 0.5% aqueous acetic acid, and solvent B was 100% acetonitrile; the solvent system was 0–16 min 15%–26% B (v/v), 16–16.1 min 26%–15% B, 16.1–18 min 15% B. The solvent flow rate was 4 mL/min, and the UV absorption was measured at 254 nm (DAD). The column temperature was set at 40 °C, and the injection volume was 2 μL. Twelve isoflavones were determined, including daidzein, genistein, glycitein, daidzin, genistin, glycitin, 6’-O-acetylglycitein, 6’-O-acetylgenistin, 6’-O-acetylglycitin, 6’-O-acetyl-daidzin, 6’-O-malonyldaidzin, 6’-O-malonylgenistin and 6’-O-malonylglycitin.

Seed protein and oil were determined using a FOSS near-infrared grain quality analyzer (Specification model: Infratec™ 1241).
Resistance evaluation of T₃ generation transgenic plant leaves

For isoflavone content and resistance evaluation, 165 soybean seedlings were planted in round plastic pots (diameter × depth: 31 cm × 24 cm) filled with sand in an aphid-free net room. Fifteen pots were planted for each transgenic line and wild type. Control groups were inoculated with phosphate buffer or noninoculated. Two seeds were planted per pot, one of which was infected with phosphate buffer, and the other was not infected at the VC-cotyledon stage. Treatment groups were inoculated with SC7 virus-containing phosphate buffer at the VC-cotyledon stage. In the third week after infection, the first compound leaves, the second compound leaves and the third compound leaves were mixed and divided into three parts, which were used for RNA extraction, isoflavone and cp content determination. Virus coat protein and total protein content were determined using SMV and total protein enzyme-linked immunosorbent assay kit (ELISA), Co., Ltd.; Catalog No. KTEP0013-96T and Co., Ltd.; Catalog No. KTEP0012-96T.

Acknowledgements

This work was supported in part by the Ministry of Science and Technology (2016YFD0100504, 2017YFE0111000) and Key Transgenic Breeding Program of China (2016ZX08004-003, 2016ZX08009003-004). We are grateful to Ms Shaoyan Lin for the help of isoflavone content measure.

Conflict of interest

The authors declare no competing interests.

Author contributions

D.Y. and H.C. conceived and supervised the project. P.Z., H.D., H.C. and D.Y. designed the research. P.Z., H.D., J.W., Y.P., C.Y., R.Y. and H.Y. performed the experiments. P.Z., H.D., J.W. and D.Y. conceived and supervised the project. P.Z., H.D., J.W., Y.P., C.Y., D.Y. and H.C. performed the experiments. P.Z., H.D., J.W., Y.P., C.Y., D.Y. and H.C. performed the experiments. P.Z., H.D., J.W., Y.P., C.Y., D.Y. and H.C. conceived and supervised the project.

References

Algar, E., Gutierrez-Mañero, F.J., García-Villaraco, A., García-Seco, D., Lucas, J.A. and Ramos-Solano, B. (2014) The role of isoflavone metabolism in plant protection depends on the rhizobacterial MAMP that triggers systemic resistance against Xanthomonas axonopodis pv. glycines in Glycine max (L.) Merr. cv. Osumi. Plant Physiol. Biochem. 82, 9–16.

Bak, R.O., Gomez-Ospina, N. and Porteus, M.H. (2018) Gene editing on center stage. Trends Genet. 34, 600–611.

Bao, A., Chen, H., Chen, L., Chen, S., Hao, Q., Guo, W., Qiu, D. et al. (2019) CRISPR/Cas9-mediated targeted mutagenesis of GmSPS9 genes alters plant architecture in soybean. BMC Plant Biol. 19, 131.

Budryn, G., Grzelačyz, J. and Perez-Sanchez, H. (2018) Binding of red clover isoflavones to actin as a potential mechanism of anti-metastatic activity restricting the migration of cancer cells. Molecules, 23, 2471.

Cai, C.M., Jiang, X., Zhao, C.M. and Ma, J.X. (2014) Sequence analysis of the coat protein gene of Chinese soybean mosaic virus strain SC7 and comparison with those of SMV strains from the USA. Bing Du Xue Bao, 30, 489–494.

Chen, W., Gao, Y., Xie, W., Gong, L., Lu, K., Wang, W., Li, Y. et al. (2014) Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. Nat. Genet. 46, 714–721.

Cheng, H., Yang, H., Zhang, D., Gai, J. and Yu, D. (2010) Polymorphisms of soybean isoflavone synthase and flavanone 3-hydroxylase genes are associated with soybean mosaic virus resistance. Mol. Breeding, 25, 13–24.

Cheng, Q., Li, N., Dong, L., Zhang, D., Fan, S., Jiang, L., Wang, X. et al. (2015) Overexpression of soybean isoflavone reductase (GmIFR) enhances resistance to phytophthora sojae in soybean. Front. Plant Sci. 6, 1024.

Chu, S., Wang, J., Zhu, Y., Liu, S., Zhou, X., Zhang, H., Wang, C. et al. (2017) An R2R3-type MYB transcription factor, GmMYB29, regulates isoflavone biosynthesis in soybean. PLoS Genet. 13, e1006770.

Curtin, S.J., Xiong, Y., Michino, J.M., Campbell, B.W., Stec, A.O., Cermak, T., Starker, C. et al. (2018) CRISPR/Cas9 and TALENs generate heritable mutations for genes involved in small RNA processing of Glycine max and Medicago truncatula. Plant Biotechnol. J. 16, 1125–1137.

Du, H., Zeng, X., Zhao, M., Cui, X., Wang, Q., Yang, H., Cheng, H. et al. (2016) Efficient targeted mutagenesis in soybean by TALENs and CRISPR/Cas9. J. Biotechnol. 217, 90–97.

Fu, R., Martin, C. and Zhang, Y. (2018) Next-generation plant metabolic engineering, inspired by an ancient chinese irrigation system. Molecular Plant, 11, 47–57.

Gebhardt, W., Witte, S., Forkmann, G., Lukàin, R., Matern, U. and Martens, S. (2005) Molecular evolution of flavonoid dioxygenases in the family Apiaceae. Phytochemistry, 66, 1273–1284.

Graham, T.L., Graham, M.Y., Subramanian, S. and Yu, O. (2007) RNAi silencing of genes for elicitation or biosynthesis of 5-deoxyisoflavonoids suppresses race-specific resistance and hypersensitive cell death in Phytophthora sojae infected tissues. Plant Physiol. 144, 728–740.

Hernandez-Garcia, C.M., Bouchard, R.A., Rushton, P.J., Jones, M.L., Chen, X., Timko, M.P. and Finer, J.J. (2010) High level transgenic expression of soybean (Glycine max) GmERF and Gmubi gene promoters isolated by a novel promoter analysis pipeline. BMC Plant Biol. 10, 237.

Hill, J.H. and Whitham, S.A. (2014) Control of virus diseases in soybeans. Adv. Virus Res. 90, 355–390.

Holton, T.A. (1995) Modification of flower colour via manipulation of P450 gene expression in transgenic plants. Drug Metab. Drug Interact. 12, 359–368.

Jiang, Y.N., Wang, B., Li, H., Yao, L.M. and Wu, T.L. (2010) Flavonoid production is effectively regulated by RNAi interference of two flavone synthase genes from Glycine max. J. Plant Biol. 53, 425–432.

Jiang, Y., Hu, Y., Wang, B. and Wu, T. (2014) Bivalent RNA interference to increase isoflavone biosynthesis in soybean (Glycine max). Braz. Arch. Biol. Technol. 57, 163–170.

Jung, W., Yu, O., Lau, S.M., O’Keefe, D.P., Odell, J., Fader, G. and Mccongulf, B. (2000) Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes. Nat. Biotechnol. 18, 208–212.

Kant, R., Tyagi, K., Ghosh, S. and Jha, G. (2019) Host alternative NADH: Ubiquinone oxidoreductase serves as a susceptibility factor to promote pathogenesis of Rhizoctonia solani in plants. Phytopathology, 109, 1741–1750.

Lan, S., Hua, Y., Fu, Y., Jian, L., Liu, Q., Jiao, X., Xin, G. et al. (2017) Rapid generation of genetic diversity by multiplex CRISPR/Cas9 genome editing in rice. Sci. China Life Sci. 60, 1–10.

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

Li, Y. and Zhang, H. (2017) Soybean isoflavones ameliorate ischemic cardiomyopathy by activating Nrf2-mediated antioxidant responses. Food Funct. 8, 2935–2944.

Li, R., Li, R., Li, X., Fu, D., Zhu, B., Tian, H., Luo, Y. et al. (2018a) Multiplexed CRISPR/Cas9-directed metabolic engineering of gamma-aminobutyric acid levels in Solanum lycopersicum. Plant Biotechnol J. 16, 415–427.

Li, X., Wang, Y., Chen, S., Tian, H., Fu, D., Zhu, B., Luo, Y. et al. (2018c) Lycopene is enriched in tomato fruit by CRISPR/Cas9-mediated multiplex genome editing. Front. Plant Sci. 9, 559.

Liu, C.J., Blount, J.W., Steele, C.L. and Dixon, R.A. (2002) Bottlenecks for metabolic engineering of isoflavone glycoconjugates in Arabidopsis. Proc. Natl Acad. Sci. USA, 99, 14578–14583.
Liu, R., Hu, Y., Li, J. and Lin, Z. (2007) Production of soybean isoflavone genistein in non-gerume plants via genetically modified secondary metabolism pathway. *Metab. Eng.* 9, 1–7.

Liu, Y., Hassan, S., Kidd, B.N., Garg, G., Mathiesius, U., Singh, K.B. and Anderson, J.P. (2017) Ethylene signaling is important for isoflavonoid-mediated resistance to *rhizoctonia solani* in roots of medicago *truncatula*. *Mol. Plant Microbe Interact.* 30, 691–700.

Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(−ΔΔC(T)) Method. *Methods.* 25, 402–408.

Ma, X., Zhu, Q., Chen, Y. and Liu, Y.G. (2016) CRISPR/Cas9 platforms for genome editing in plants: developments and applications. *Molecular Plant.* 9, 961–974.

Malloy, K.M., Wang, J., Clark, L.H., Fang, Z., Sun, W., Yin, Y., Kong, W. et al. (2018) Novel soy and genistein inhibit endometrial cancer cell proliferation through disruption of the AKT/mTOR and MAPK signaling pathways. *Am J Transl Res.* 10, 784–795.

Martens, S. and Mithofer, A. (2005) Flavones and flavone synthases. *Phytochemistry.* 66, 2399–2407.

Martens, S., Forkmann, G., Matern, U. and Lukacin, R. (2001) Cloning of parsley flavone synthase I. *Phytochemistry.* 58, 43–46.

Minkenberg, B., Wheatley, M. and Yang, Y. (2017) CRISPR/Cas9-Enabled Multiplex Genome Editing and Its Application. *Prog. Mol. Biol. Transl. Sci.* 140, 111.

Pandey, A., Misra, P., Khan, M.P., Swarnik, G., Tewari, M.C., Bhamhani, S., Trivedi, R. et al. (2014) Co-expression of Arabidopsis transcription factor, AIMYB12, and soybean isoflavone synthase, GmIFS1, in tobacco leads to enhanced biosynthesis of isoflavones and flavonols resulting in osteoprotective activity. *Plant Biotechnol. J.* 12, 69–80.

Sathyapalan, T., Aye, M., Rigby, A.S., Thatcher, N.J., Dargham, S.R., Kilpatrick, E.S. and Atkin, S.L. (2018) Soy isoflavones improve cardiovascular disease risk markers in women during the early menopause. *Nutr. Metab. Cardiovasc. Dis.* 28, 691–697.

Shen, L., Wang, C., Yu, F., Wang, J., Liu, Q., Zhang, X., Yan, C. et al. (2018) QTL editing confers opposing yield performance in different rice varieties. *J. Integr. Plant Bio.* 60, 89–93.

Subramanian, S., Stacey, G. and Yu, O. (2006) Endogenous isoflavones are essential for the establishment of symbiosis between soybean and *Bradyrhizobium japonicum*. *Plant J.* 48, 261–273.

Tran, P.T., Widyasari, K., Seo, J.K. and Kim, K.H. (2018) Isolation and validation of a candidate *Rsv3* gene from a soybean genotype that confers strain-specific resistance to soybean mosaic virus. *Virology.* 513, 153–159.

Wang, C., Shen, L., Fu, Y., Yang, J., Liu, Q., Zhang, X., Yan, C. and Wang, K. (2015) A simple CRISPR/Cas9 system for multiplex genome editing in rice. *J. Genet. Genom.* 42, 703–706.

Wang, M., Mao, Y., Lu, Y., Tao, X. and Zhu, J.-K. (2017) Multiplex gene editing in rice using the CRISPR-Cpf1 system. *Molecular Plant.* 10, 1011–1013.

Wang, B., Wu, Z., Li, Z., Zhang, Q., Hu, J., Xiao, Y., Cai, D. et al. (2018a) Dissection of the genetic architecture of three seed-quality traits and consequences for breeding in *Brassica napus*. *Plant J.* 136, 1336–1348.

Wang, M., Mao, Y., Lu, Y., Wang, Z., Tao, X. and Zhu, J.-K. (2018b) Multiplex gene editing in rice with simplified CRISPR-Cpf1 and CRISPR-Cas9 systems. *J. Integr. Plant Biol.* 60, 626–631.

Wu, M., Li, Y.N., Zhang, C., Liu, X.T., Liu, C.C., Guo, R., Niu, K.X. et al. (2019) Molecular mapping of the gene conferring resistance to soybean mosaic virus and bean common mosaic virus in the soybean cultivar Raiden. *Mol. Plant Microbe Interact.* 32, 1323–1334.

Xie, K., Zhang, J. and Yang, Y. (2014) Genome-wide prediction of highly specific guide RNA spacers for CRISPR-Cas9-mediated genome editing in model plants and major crops. *Molecular Plant.* 7, 923–926.

Xu, R., Yang, Y., Qin, R., Li, H., Qiu, C., Li, L., Wei, P. et al. (2016) Rapid improvement of grain weight via highly efficient CRISPR/Cas9-mediated multiplex genome editing in rice. *J. Genet. Genom.* 43, 525–532.

Yan, J., Wang, B., Jiang, Y., Cheng, L. and Wu, T. (2014) GmNFsII-controlled soybean flavone metabolism responds to abiotic stresses and regulates plant salt tolerance. *Plant Cell Physiol.* 55, 74–86.

Yin, K., Gao, C. and Qiu, J.L. (2017) Progress and prospects in plant genome editing. *Nat. Plants.* 3, 17107.

Yu, O., Jung, W.S., Shi, J., Croues, R.A., Fader, G.M., McGonigle, B. and Odell, J.T. (2000) Production of the isoflavones genistein and daidzein in non-gerume dicots and monocots tissues. *Plant Physiol.* 124, 781–793.

Zhang, Z., Ma, Y., Zhu, H., Li, S. and Zou, J.K. (2016) A multiplex CRISPR/Cas9 platform for fast and efficient editing of multiple genes in *Arabidopsis*. *Plant Cell Rep.* 35, 1519–1533.

Zhou, J., Zou, J., Hu, J., Li, Q., Zou, X., Zhong, Z. et al. (2018) Multiplex QTL editing of grain-related genes improves yield in elite rice varieties. *Plant Cell Rep.* 38, 475–485.

### Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Figure S1** Schematic diagram of the isoflavone, flavone, flavonol and anthocyanin synthetic pathways.

**Figure S2** Flowchart for construction of the CRISPR/Cas9 system for multiplex gene editing in soya bean.

**Figure S3** Map of the pGmUbi-Cas9-4XsgR binary vector for multiplex genome editing.

**Figure S4** CRISPR/Cas9-induced mutation types and frequency in soy bean hairy roots.

**Figure S5** Relative expression of GmF3H1, GmF3H2 and GmNFsII-1 in the leaves of wild-type and T3 transgenic plants.

**Figure S6** Isoflavone, protein, oil content and 100-grain weight in beans of T2 generation triple mutants.

**Figure S7** Relative expression levels of selected genes in T3 generation homozygous triple mutants leaves.

**Figure S8** Relative expression levels of selected genes in the seeds of T3 generation homozygous triple mutants at three different developmental stages.

**Table S1** Target sequences and primers used to detect mutations.

**Table S2** Primers used to construct the vector.

**Table S3** Primers used for quantitative real-time PCR.

**Table S4** Examination of the co-mutation frequencies of the CRISPR/Cas9-induced mutations.

**Table S5** Metabolomic analysis of T0 transgenic plants and wild type.

**Table S6** T1 generation genotype and Bar positive status.

**Table S7** Positive identification of T2 transgenic plants.

**Table S8** Mutation types of T0 plants (T0–16–1) and its descendants (PAM in green).

**Table S9** Metabolomic analysis of T0 transgenic plants and wild type.

**Table S10** All substance types determined by the metabolome.