Creation of Early Flowering Germplasm of Soybean by CRISPR/Cas9 Technology

Jinan Han¹, Bingfu Guo², Yong Guo¹, Bo Zhang³, Xiaobo Wang⁴* and Li-Juan Qiu¹*

¹ The National Key Facility for Crop Gene Resources and Genetic Improvement (NFCRI) and MOA Key Labs of Crop Germplasm and Soybean Biology (Beijing), Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing, China, ² Crop Research Institute, Jiangxi Academy of Agricultural Science, Jiangxi, China, ³ School of Plant and Environmental Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA, United States, ⁴ College of Agronomy, Anhui Agricultural University, Hefei, China

Soybean is an important economic crop and a typical short-day crop, sensitive to photoperiod, and has narrow geographical adaptative region, which limit the creation of transgenic materials and reduce the breeding efficiency of new varieties. In addition, the genetic transformation efficiency of soybean is lower than that of many other crops, and the available receptor genotypes are limited. In this study, Agrobacterium-mediated transformation were used to introduce the CRISPR/Cas9 expression vector into soybean cultivar Jack and generated targeted mutants of E1 gene controlling soybean flowering. We obtained two novel types of mutations, 11 bp and 40 bp deletion at E1 coding region, respectively, and frameshift mutations produced premature translation termination codons and truncated E1 proteins, causing obvious early flowering under long day condition. In addition, no off-target effects were observed by predicting and analyzing the potential off-target sites of E1 targets. Significant decreased E1 gene expression of two novel mutants showed that the truncated E1 protein disinhibited GmFT2a/5a and increasing GmFT2a/5a gene expressions resulted obvious early flowering. Homozygous trans-clean mutants without T-DNA elements were also obtained and showed early flowering under long day condition. The photo-insensitive soybean transformation receptor we created laid a foundation for breeding excellent transgenic receptors suitable for high latitudes.

Keywords: soybean, CRISPR/Cas9, gene editing, E1, Agrobacterium-mediated transformation

INTRODUCTION

Soybean is rich of protein and oil and has high economic value. With the increasing demand for soybean globally, it is urgent to clarify gene function, and accelerate functional gene research and breeding speed for increasing yield and improving quality. In recent years, CRISPR (Clustered regular interspaced short palindromic repeat)/Cas9 (CRISPR-associated) provides an effective method for targeted genome editing and gene function research, and supplies a new idea for reverse genetics research. CRISPR/Cas9 system includes gene knock out, knock in, multiple genes and sites editing, large fragment deletion and replacement (Gratz et al., 2013; Shan et al., 2013; Feng et al., 2014; Gratz et al., 2014; Zhou et al., 2014). Feng et al. (2013) and Mao et al. (2013) firstly proved that CRISPR/Cas9 system could be used to target genome editing in crop by introducing site-directed mutations in specific genes in Arabidopsis and rice. Subsequently, CRISPR/Cas9 technology has
been widely used in several species including rice (Shan et al., 2013), wheat (Upadhyay et al., 2013), cotton (Gao et al., 2017), maize (Chen et al., 2018), Arabidopsis (Li et al., 2013), tobacco (Gao et al., 2015) and barley (Kapusi et al., 2017). Jacobs et al. (2015) first used CRISPR/Cas9 technology to knock out the green fluorescent protein gene (GFP), and produced targeted editing for nine soybean endogenous genes. Then CRISPR/Cas9 technology began to be widely applied in soybean (Cai et al., 2015; Jacobs et al., 2015; Li et al., 2015; Michno et al., 2015; Sun et al., 2015; Du et al., 2016; Tang et al., 2016; Cai et al., 2018; Li et al., 2019). Therefore, CRISPR/Cas9 is an effective tool for soybean targeted genome editing, providing a theoretical and technical basis for further research on soybean genome, as well as improving the breeding efficiency and accelerating the breeding process.

CRISPR/Cas9 system relies on transformation technology. However, soybean transformation efficiency is lower than other crops (Du et al., 2016), and receptor genotype dependence is a major limiting factor. The transformation efficiency of soybean is usually low, and the varieties suitable for soybean genetic transformation are very few (Donaldson and Simmonds, 2000; Guo et al., 2015). At the same time, soybean is a typical short-day crop and sensitive to photoperiod, which limits the geographical cultivated region (Xu et al., 2013; Wang et al., 2016), excellent varieties creation and breeding efficiency.

Soybean varieties are adapted to different latitudes and different photoperiods, so they need to have a series of photoinensitivity. Photoperiod is the key meteorological factor that determines flower bud differentiation and adaptation to different ecological regions (Camara et al., 1997). The wide adaptability to different latitudes in soybean is controlled by some major genes and QTLs (Watanabe et al., 2012). At present, 11 genes have been identified to be related to soybean growth period (E1-E10, J) (Bernard, 1971; Buzzell, 1980; Buzzell and Voldeng, 1980; Mcblain and Bernard, 1987; Mcblain et al., 1987; Ray et al., 1995; Cober et al., 1996; Bonato and Antonio, 1999; Cober et al., 2010; Kong et al., 2014; Lu et al., 2017; Yue et al., 2017). Among these genes, E1 has the greatest influence on soybean growth period with the most inhibitory of flowering and is considered as the most important gene controlling soybean growth period (Bernard, 1971; Cober et al., 1996; Watanabe et al., 2012; Xia et al., 2012) and is also a main selection locus in soybean breeding (Xia, 2017). E1, located at centromere of chromosome 6, is a unique transcription to legumes. It has bilateral nuclear location signal and DNA binding site, and is a flowering inhibitor related to the B3 domain (Xia et al., 2012). Seven alleles have been identified for E1. Its amino acids changes in the nuclear location signal region resulted in its protein distribution changing. E1 protein distributes in the nucleus, while e1 protein distributes in the nucleus and cytoplasm (Xia, 2013). E1 gene expression is closely related to the length of day light. Under long day condition (LD), E1 delays flowering by negative regulation of GmFT2a/5a, while its recessive alleles have early flowering time by disinhibiting GmFT2a/5a expression (Zhang et al., 2016). E1 gene expression was significantly inhibited under short day condition which is the main factor of photoperiod sensitivity in soybean (Lü, 2015).

In this study, we used CRISPR/Cas9 system and Agrobacterium-mediated transformation to introduce the CRISPR/Cas9 expression vector into soybean cultivar Jack, knocked out E1 gene and analyzed the effect of novel E1 mutants to soybean flowering. It provides materials for breeding early flowering receptors, promotes the development of soybean genetic improvement, provides a basis for efficient soybean genetic transformation, and establishes an important guide for soybean gene function research, molecular breeding and variety layout.

**MATERIALS AND METHODS**

**Plant Materials and Growth Conditions**

Soybean cultivar Jack was used for genetic transformation receptor. Jack (wild type as a control) and all seeds harvested from T0 plants obtained through CRISPR/Cas9 and Agrobacterium-mediated transformation were planted on June 22, 2017 at the Shunyi Experimental Station of the Institute of Crop Sciences, Chinese Academy of Agricultural Sciences. T1 homozygous mutants were planted in greenhouse under long day conditions (LD, 16 h/8 h, light/dark) and short-day conditions (SD, 12 h/12 h, light/dark) at 28, 70% relative humidity. The flowering time of each plant from seedling emergence to R1 stage (the first flower appeared on any node of the main stem) was recorded according to Fehr et al. (1971). Data analysis was performed by using Microsoft Excel, then adopted one-way analysis of variance. P < 0.01 was considered statistically significant.

**CRISPR/Cas9 Expression Vector Construction**

Cas9 sequence was optimized for the codon-optimized for dicotyledons and connected with CaMV 35S promoter at downstream, then assembled with sgRNA driven by Arabidopsis U6 promoter to construct a plasmid vector containing both sgRNA and Cas9. The bar gene as a selective marker was driven by CaMV 35S promoter. The CRISPR/Cas9 expression cassette was shown in Figure 1. The sequences were synthesized by Shanghai Sangon Biotech (Shanghai, China). The soybean endogenous gene E1 (Glyma.06G207800) sequence and its information were downloaded from the Phytozome database (www.phytozome.net/). The optimal sgRNA sequences (20 bp) with the G as the first base were designed using online tool (http://www.genome.arizona.edu/crispr/CRISPRsearch.html). The base G was added at the 5′ end artificially if the first base was not G (Ran et al., 2013). The online website (http://www.rgenome.net/cas-offinder/) was used to evaluate off-target effects. Two sgRNAs for the E1 gene we selected were named as E1-SP1 (5′-CCCTTCAGATGAAAAGGAGCAG-3′) and E1-SP2 (5′-CCACCATATGGAAAGCCTTAA-3′) respectively. Primers containing either of two sgRNAs were synthesized by Shanghai Sangon Biotech (Supplementary Table 1). Using overlapping polymerase chain reaction (PCR), two sgRNAs were cloned into pPCB plasmid. PCR products containing two sgRNAs were digested and inserted into the pHSE401 plasmid vector to construct the CRISPR/Cas9 expression vector containing Cas9.
and sgRNA (Supplementary Figure 1). And then the CRISPR/Cas9 expression vector was transformed into E. coli Trans1 T1 (TransGen Biotech) used for soybean genetic transformation.

**Agrobacterium-Mediated Soybean Transformation**

The CRISPR/Cas9 expression vector was transferred into *Agrobacterium* strain EHA105 by electroporation. Soybean cultivar Jack was used for tissue culture and soybean transformation. Soybean genetic transformation procedure was referred to the described method in our laboratory (Guo et al., 2015) and made appropriate modification. The healthy soybean seeds were sterilized with chlorine for 16 h and germinated 5 d to prepare explants. Explants were submerged in agrobacterium suspension adding 0.02% surfactant (Silwet L-77) and wounded by ultrasonic treatment. After infection for 30–40 min, explants were co-cultured for 3 d. After the tissue culture (resistant bud induction, shoot elongation and rooting), transgenic plants were regenerated from explants. At the stage of resistant bud induction and shoot elongation, we added 10 mg L⁻¹ and 6 mg L⁻¹ glufosinate (Sigma-Aldrich, USA) respectively into culture medium to screen positive transformed cells.

**RESULTS**

**CRISPR/Cas9-Mediated Mutations**

Two targets for the E1 gene (named E1-SP1 and E1-SP2, respectively) were designed (Figure 2), and the CRISPR/Cas9 expression vector was transferred into the soybean cultivar Jack by *Agrobacterium*-mediated genetic transformation to knock out soybean endogenous gene E1. The whole genome DNAs of transformed plants were used for PCR and sequence analysis. Combining with PAT/Bar test strip detection, 16 T-DNA positive plants were obtained (Figure 3), of which 12 plants had heterozygous mutations at target sites (Supplementary Figure 2). Then, all seeds derived from 9 heterozygous T₀ generation were planted in the Shunyi Station of the Chinese Academy of Agricultural Sciences under LD conditions (3 of 12 T₀ plants had no seed) and the types of mutation at target sites were determined. Six out of nine lines had three types of mutations: homozygous mutation, heterozygous mutation and non-mutation (Table 1). The rest three T₁ lines didn’t show mutation, and site-directed mutagenesis of T₁ generation didn’t inherit to its progeny.

**Real-Time Quantitative PCR Analysis of Gene Expression**

Expression levels of E1 and GmFT2a/5a in wild type plants and T₁ homozygous mutants were analyzed under LD and SD conditions, respectively. Every 5-day interval after 10 days after emergence (DAE), at 10 am (4 h after light), the trifoliate leaves were sampled from plants with different genotypes under LD and SD conditions. Total RNAs were extracted using TransZol Up Plus RNA Kit (TransGen Biotech). For reverse transcription, the first-strand cDNA synthesis was performed using the TransScript First-strand cDNA Synthesis SuperMix Kit (TransGen, China). For qRT-PCR, gene expressions were examined using cDNA templates on an Applied Biosystems 7300 Real-Time PCR System. The relative gene expression levels followed the method (Pfaffl, 2001). The mRNA level of GmActin (Glyma18g52780) was used as a reference for normalization. Specific primers we used in this study were list in Supplementary Table 2. Three biological replicates were used for each gene.

**Supplementary Table 2**

| Target site | E1-SP1 | E1-SP2 |
|------------|--------|--------|
| Bar        | yes    | yes    |
| Kan        | yes    | no     |
| RepA       | yes    | yes    |
| StbA       | yes    | yes    |
| Cas9       | yes    | yes    |
| gRNA       | yes    | yes    |
| U6 promotor| yes    | yes    |
| 35S promotor| yes   | yes    |

**Supplementary Figure 1**

The vector pBSE401 used for CRISPR/Cas9-mediated genome editing. ATU6, Arabidopsis U6 promotor; gRNA, guide RNA; 3SS promotor, CaMV 3SS promotor; Cas9, codon-optimized Cas9; NLS, nuclear location signal; bar, selective marker gene; KanR, Kanamycin resistance gene; pVS1-RepA, pVS1 replication origin; pVS1-StbA, pVS1 stability function.

**Supplementary Table 1**

| T₀ seed | E1-SP1 Bar | E1-SP1 Kan | E1-SP2 Bar | E1-SP2 Kan |
|---------|------------|------------|------------|------------|
| 1st     | yes        | no         | yes        | yes        |
| 2nd     | yes        | no         | yes        | yes        |
| 3rd     | yes        | no         | yes        | yes        |

**Supplementary Table 2**

| Target site | E1-SP1 | E1-SP2 |
|------------|--------|--------|
| Bar        | yes    | yes    |
| Kan        | yes    | no     |
| RepA       | yes    | yes    |
| StbA       | yes    | yes    |
| Cas9       | yes    | yes    |
| gRNA       | yes    | yes    |
| U6 promotor| yes    | yes    |
| 35S promotor| yes   | yes    |

**Supplementary Table 2**

| T₀ seed | E1-SP1 Bar | E1-SP1 Kan | E1-SP2 Bar | E1-SP2 Kan |
|---------|------------|------------|------------|------------|
| 1st     | yes        | no         | yes        | yes        |
| 2nd     | yes        | no         | yes        | yes        |
| 3rd     | yes        | no         | yes        | yes        |

**Supplementary Figure 2**

The vector pBSE401 used for CRISPR/Cas9-mediated genome editing. ATU6, Arabidopsis U6 promotor; gRNA, guide RNA; 3SS promotor, CaMV 3SS promotor; Cas9, codon-optimized Cas9; NLS, nuclear location signal; bar, selective marker gene; KanR, Kanamycin resistance gene; pVS1-RepA, pVS1 replication origin; pVS1-StbA, pVS1 stability function.

**Supplementary Table 1**

| T₀ seed | E1-SP1 Bar | E1-SP1 Kan | E1-SP2 Bar | E1-SP2 Kan |
|---------|------------|------------|------------|------------|
| 1st     | yes        | no         | yes        | yes        |
| 2nd     | yes        | no         | yes        | yes        |
| 3rd     | yes        | no         | yes        | yes        |

**Supplementary Figure 2**

The vector pBSE401 used for CRISPR/Cas9-mediated genome editing. ATU6, Arabidopsis U6 promotor; gRNA, guide RNA; 3SS promotor, CaMV 3SS promotor; Cas9, codon-optimized Cas9; NLS, nuclear location signal; bar, selective marker gene; KanR, Kanamycin resistance gene; pVS1-RepA, pVS1 replication origin; pVS1-StbA, pVS1 stability function.
Two types of homozygous mutations were detected at the target sites (named \( e1-1 \) and \( e1-2 \), 11 bp deletion and 40 bp deletion, respectively) in three lines including \( L7 \) (\( e1-1 \)), \( L9 \) (\( e1-2 \)) and \( L16 \) (\( e1-1 \)). All of them were frameshift mutations, resulting in premature translation termination codons (Supplementary Figure 3). The 11 bp deletion of \( e1-1 \) created a truncated protein encoding 79 amino acids, and caused the absence of all the B3 domains while keeping part of nuclear location signal. The 40 bp deletion of \( e1-2 \) created a truncated protein encoding 88 amino acids, and caused the absence of neither B3 domain or nuclear location signal (Supplementary Figure 4).

Potential Off-Target Analysis

In order to determine whether the CRISPR/Cas9 expression vector we used to have variations at potential off-targets and avoid the possibility of potential off-targets affecting on phenotype, by using online website (http://cbi.hzau.edu.cn/crispr/), four most potential off-target sites at the two targets (\( E1\text{-SP1} \) and \( E1\text{-SP2} \)) were selected. Every potential off-target site mismatched 2-4 bases with the \( E1 \) target sequences (Table 2). Using specific primers of potential off-target sites (Supplementary Table 2) and genome DNAs of 35 T\(_1\) mutants as templates for PCR, we didn't detect variation at four potential off-targets. The sequence comparison analysis shown in Supplementary Figure 5 indicated that CRISPR/Cas9 expression vector had specific edits in \( E1\text{-SP1} \) and \( E1\text{-SP2} \) targets.

Inheritance Analysis and Phenotype Identification

To identify whether the mutations at the target of the homozygous T\(_1\) mutants could inherited to T\(_2\) generation stably, T\(_2\) seeds derived from homozygous mutants of T\(_1\) lines (\( L7 \), \( L9 \) and \( L16 \)) were planted under LD and SD conditions, respectively (Table 3). Sequence analyses of 28 T\(_2\) individuals indicated that CRISPR/Cas9-mediated mutagenesis of \( E1 \) gene could be stably inherited from T\(_1\) generation to T\(_2\) generation and maintained the same type of variation. To analyze the flowering time accurately, R1 period of all above T\(_2\) plants were recorded under LD and SD conditions respectively. Under LD condition, when T\(_2\) homozygous mutants were flowering, the wild type plants had...
creased to the correlation between the expression of E1 and GmFT2a/5a under LD or SD condition, but different expression levels. Under LD condition, E1 had the highest expression level at 15 DAE both in mutants and wild type plants, but significantly lower E1 expression level in mutants was observed at 15 and 20 DAE (P < 0.01). GmFT2a/5a exhibited two expression peaks at 15 DAE and 25 DAE both in mutants and wild type plants, but there were significantly higher expression levels in mutants at 15 and 25 DAE (P < 0.01). Under SD condition, all the three genes (E1/GmFT2a/5a) showed bimodal expression patterns and the expression peaks appeared at 15 DAE and 25 DAE, respectively. However, E1 gene expression was significantly lower and almost no expression than that under LD condition.

Trans-Clean Mutants Without T-DNA Elements
To obtain novel soybean germplasm with homozygous mutation but without T-DNA elements, PAT/Bar test strip was used to identify the selective marker gene bar firstly and PCR strategy was used to exam sgRNA/Cas9 on T-DNA by using specific primers (Supplementary Table 2). Among three T1 lines, only L7 didn’t show T-DNA elements, and its 50 progenies were all free of T-DNA. Only 11 out of 211 homozygous mutants were free of T-DNA in T2 generation derived from L9 and L16 (Table 4, Figure 6).

DISCUSSION
Soybean is an important source of protein and oil for human. Therefore, gene function exploration and utilization are of great significance to increase yield and improve seed quality. However, soybean is a short-day crop, sensitive to photoperiod and had narrow geographical cultivated regions. In Agrobacterium-mediated genetic transformation, soybean transformation efficiency is lower than many other crops, and depends on genotypes used for transformation. At present, the soybean receptors with higher transformation efficiency are limited. Therefore, its sensitivity to photoperiod and difficulty in developing transgenic offspring have significantly restricted the improvement of breeding efficiency of new varieties and germplasm. Therefore, it is important to study soybean growth-related genes, breed and create new soybean germplasm with wide adaptability to different latitudes. Among soybean growth-related genes, E1 has the greatest impact on growth period and has the most obvious photoperiod response, so it is considered as a major gene for controlling soybean flowering (Bernard, 1971; Cober et al., 1996; Watanabe et al., 2012; Xia et al., 2012).

In this study, we screened and obtained early flowering homozygous mutants without transgenic elements of soybean endogenous E1 gene by CRISPR/Cas9 system. The random integration of foreign genes in plant chromosome may lead
Creation of Early Flowering Soybean by CRISPR/Cas9
Han et al.

To destruction of plant endogenous gene, gene silencing and other undesirable phenomena, which caused a great controversy about the biosecurity of genetically modified organisms (Napoli et al., 1990). Therefore, the promotion and application of transgenic crops are greatly restricted (Zhang et al., 2011). CRISPR/Cas9 system affords an effective method to eliminate undesirable phenomena in the transformation by editing the target gene accurately. At the same time, the transgenic elements in CRISPR/Cas9 system such as Cas9 and other selective markers can be lost by progeny separation or selling and obtain trans-clean mutants (Cai et al., 2018; Chen et al., 2018). Besides, off-target effects could be reduced by selecting target sequences specifically (Xie et al., 2014; Xu et al., 2015). As an effective genome editing tool, CRISPR/Cas9 technology has been widely used in many crops and there are also many applications in soybean. Jacobs et al. (2015) reported the site-directed mutagenesis in soybean by using CRISPR/Cas9 technology firstly, which laid a good foundation for soybean genome editing. Many studies have also successfully obtained mutants without transgenic elements in many crops by CRISPR/Cas9. Haun et al. (2014) generated a high oleic acid content soybean variety without transgenic components and improved the quality of soybean. Cai et al. (2018) obtained trans-clean soybean homozygous mutants with late flowering time by using CRISPR/Cas9 technology to knock out GmFT2a.

| T_{1} generation | Mutation type | T_{2} generation |
|-------------------|---------------|------------------|
|                   | No. of plants in LD | No. of plants in SD |
| L7                | 40 bp delete     | 3                | 3                |
| L9                | 11 bp delete     | 6                | 6                |
| L16               | 40 bp delete     | 19               | 17               |

**TABLE 3** | T_{2} homozygous mutants under LD and SD conditions.

**TABLE 3** | T_{2} homozygous mutants under LD and SD conditions.

**FIGURE 4** | CRISPR/Cas9-induced E1 mutants flowering time under both LD and SD conditions. (A) Phenotypes of wild type (WT, Jack) and homozygous T_{2} mutant under LD condition, respectively. Top panel, WT did not have floral buds when T_{2} mutant was flowering. Bottom panel, T_{2} mutant produced the pods when WT was flowering. Red box, magnified view. (B) Flowering time of WT and homozygous T_{2} mutants under LD condition. (C) Phenotypes of wild type (WT, Jack) and homozygous T_{2} mutant under SD condition, respectively. (D) Flowering time of WT and homozygous T_{2} mutants under SD condition. n, exact numbers of individual plants identified. **, homozygous T_{2} mutants exhibit significant early flowering time (P < 0.01). The flowering time is shown as the mean values ± standard deviation.
**E1** is a unique transcription factor to legumes and its B3 domain and nuclear location signal are important modules controlling soybean flowering (Xia et al., 2012). **E1** gene includes at least 7 allelic natural variations with **E1** and **e1-as** as the two basic genotypes. **e1-as** has a single missense mutation at the region of nuclear location signal. This one amino acid change led to the cell localization change and e1 protein distribution in the nucleus and cytoplasm at the same time. However, **e1-as** is a leaky allele and has partially function of delaying flowering (Xia, 2013). The other three nonfunctional alleles are **e1-as-fs**, **e1-as-nl** and **e1-as-b3a**. **e1-as** has 1 bp deletion in the B3 domain, and this frameshift mutation resulted in a truncated protein encoding 41 amino acids. **e1-nl** is a null allele and all the **E1** gene is deleted. **e1-b3a** allele has 3 SNPs and 2 bp deletions in the B3 domain resulting in frameshift mutation (Zhai et al., 2015). **e1-re** and **e1-p** have variations at 5′UTR, but the flowering mechanism of these two alleles is unclear (Tsubokura et al., 2014). The two novel types of homozygous germplasm we obtained had 11 bp deletion and 40 bp deletion in the coding region, respectively, resulting in pre-terminate codons and truncated E1 proteins. **e1-1** coded a 79-aa and deleted all B3 domain. **e1-2** coded an 88-aa and completely deleted the nuclear location signals and B3 domain.

In addition, we compared the flowering time (R1) of **E1** mutants and wild-type. Under LD condition, the two types of homozygous mutants showed about 20 days earlier flowering time than wild type plants. The flowering time of the wild type plants

![Graphs showing expression analyses of E1/GmFT2a/5a in WT plants and mutants under LD and SD conditions.](https://example.com/graphs.png)
Creation of Early Flowering Soybean by CRISPR/Cas9

Han et al.

was about 57 days and the average flowering time of mutants was only about 38 days. Studies have shown that the flowering time of cultivars carrying E1, e1-as, e1-fs and e1-nl were 70 d, 50 d, 30 d and 30 d under LD condition, respectively (Xia et al., 2012). In our study, the average flowering time of novel germplasm was about 38 days which was like the natural alleles e1-fs and e1-nl, indicating that the new E1 mutants we obtained by CRISPR/Cas9 system may have the same mechanism for flowering as natural alleles. Studies have shown that nonsense mediated mRNA decay (NMD) reduces mRNAs with premature translation termination codons (PTCs) by down-regulating gene expression, and reduces its encoding truncated protein production (Baker and Parker, 2004; Maquat, 2004; Conti and Izaurralde, 2005; Maquat, 2005). The two mutants obtained in our study had bases deletion, leading to PTCs and truncated E1 protein. We speculated that NMD resulted in decreased E1 gene expression in mutants compared with wild type plants.

Soybean is a typical short-day crop and its flowering time is closely related to the length of day light. The expression of E1 was negatively correlated with GmFT (GmFT2a/5a) and controlled the GmFT2a/5a that functionally coordinated each other (Xia et al., 2012). In this study, compared with wild type plants, E1 gene sequence of mutants caused not only truncated E1 protein but also significantly decreased gene expressions. Meanwhile, GmFT2a/5a expressions of mutants were significantly increased due to decreased E1 gene and appeared earlier flowering time. In order to identify whether the decreased E1 gene expression had influence on its homologous gene E1-L (Glyma04g24640/Glyma18g22670) and avoid its potential effect on the flowering time, we analyzed the E1-L gene expression. Result showed that there was no obvious difference on E1-L gene expression between mutants and wild type plants (Supplementary Figure 7).

E1 is the most important gene controlling flowering and is also the major determinant to short-day crops. Photosensitivity in soybean reduced the flowering time, leading to early maturation and low yield in low latitude areas. Under SD condition, long juvenile trait can ensure enough growth period in soybean (Hartwig and Kiihl, 1979). J (Glyma.04G050200) is a transcriptional suppressor of E1 gene and its recessive allele, j, delaying flowering time by disinhibit E1. The amplitude to delay flowering time is determined by different E1 alleles (Lu et al., 2017). Furthermore, studies have shown that long-juvenile trait in soybean may be controlled by other genes (Ray et al., 1995, Valéria et al., 2000, Yue et al., 2017). Therefore, mutants we obtained not only provide soybean new receptor for high latitudes, and can also change the function of J at low latitudes, which may provide soybean high yield potential. In addition, it provides materials and theoretical basis for identifying other genes controlling long-juvenile trait and studying further the flowering regulation pathways. Compared with cross breeding,

### Table 4: Transgene-clean homozygous mutants from T1 and T2 generations.

| T1 generation | T-DNA | T2 generation | No. of plants identified | No. of trans-clean plants |
|---------------|-------|---------------|--------------------------|--------------------------|
| L7            | free  |               | 50                       | 50                       |
| L9            | positive |             | 5                        | 2                        |
| L16           | positive |             | 206                      | 9                        |
| Total         |        |               | 261                      | 61                       |

FIGURE 6 | Identifying of trans-clean mutants. (A) Detection of the selectable marker gene bar by PAT/Bar test strip. The red arrowhead indicates that bar gene is positive. (B) Gel image of PCR products for T-DNA elements. Cas9, part of the Cas9 coding sequence. sgRNA, region from the U6 promoter to the downstream vector sequence spanning the sgRNA. GmActin was used as a normalization control. V: plasmid of the vector in transformation. WT, wild type. Labels 1-20, individual mutant lines.
using CRISPR/Cas9 technology to create early flowering soybean material has the advantages significantly short breeding period with high efficiency.

**CONCLUSION**

Our study provides materials support for breeding early-maturing transgenic receptors suitable for high latitudes and contributes to the soybean introduction. The photo-insensitive soybean transformed receptors could improve the soybean genetic development and contribute to efficient soybean genetic transformation. It offers important guidance for molecular breeding, soybean gene function research and variety development.

**DATA AVAILABILITY STATEMENT**

The datasets generated for this study are available on request to the corresponding author.

**REFERENCES**

Baker, K. E., and Parker, R. (2004). Nonsense-mediated mRNA decay: terminating erroneous gene expression. *Curr. Opin. Cell Biol.* 16 (3), 293–299. doi: 10.1016/j.ceb.2004.03.003

Bernard, R. L. (1971). Two major genes for flowering and maturity in soybeans. *Crop Sci.* 11, 242–244. doi: 10.2135/cropsci1971.0111183X001100020022x

Bonato, E. R., and Antonio, V. N. (1999). Soybean introduction. The photo-insensitive soybean transformed receptors suitable for high latitudes and contributes to the soybean introduction. The photo-insensitive soybean transformed receptors could improve the soybean genetic development and contribute to efficient soybean genetic transformation. It offers important guidance for molecular breeding, soybean gene function research and variety development.

**AUTHOR CONTRIBUTIONS**

JH performed the experiments, analyzed the data, and wrote the manuscript. BG, YG, and XW provided advice on experimental implementation. XW, BZ, and L-JQ supervised the project, reviewed and revised the manuscript.

**ACKNOWLEDGMENTS**

This work was supported by the National Science and Technological innovation Program of China.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.01446/full#supplementary-material

Fehr, W. R., Caviness, C. E., Burmood, D. T., and Pennington, J. S. (1971). Stage of development descriptions for soybeans, *Glycine max* (L). *Merrill. Crop Sci.* 11 (6), 929–931. doi: 10.2135/cropsci1971.0111183X001100060051x

Feng, Z. Y., Mao, Y. F., Xu, N. F., Zhang, B. T., Wei, P. L., Yang, D., et al. (2014). Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 111 (12), 4632. doi: 10.1073/pnas.1400822111

Feng, Z. Y., Zhang, B. T., Dong, W. N., Liu, X. D., Yang, D. L., Wei, P. L., et al. (2013). Efficient genome editing in plants using a CRISPR/Cas system. *Cell Res.* 23 (10), 1229–1232. doi: 10.1038/cr.2013.114

Gao, J. P., Wang, G. H., Ma, S. Y., Xie, X. D., Wu, X. W., Zhang, X. T., et al. (2015). CRISPR/Cas9-mediated targeted mutagenesis in *Nicotiana tabacum*. *Plant Mol. Biol.* 87 (2), 99–110. doi: 10.1007/s11103-014-0263-0

Gao, W., Lu, L., Tian, X. Q., Xu, F. C., Liu, J., Singh, P. K., et al. (2017). Genome editing in cotton with the CRISPR/Cas9 system. *Front. Plant Sci.* 8, 1364. doi: 10.3389/fpls.2017.01364

Gratz, S. J., Cummings, A. M., Nguyen, J. N., Hamm, D. C., Donohue, L. K., Harrison, M. M., et al. (2013). Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nucleases. *Genetics* 194 (4), 1029. doi: 10.1534/genetics.113.152710

Gratz, S. J., Ukken, F. P., C Dustin, R., Gene, T., Donohue, L. K., Cummings, A. M., et al. (2014). Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in *Drosophila*. *Genetics* 196 (4), 961–971. doi: 10.1534/genetics.113.160713

Guo, B. F., Guo, Y., Hong, H. L., Jin, L. G., Zhang, L. J., Chang, R. Z., et al. (2015). Co-treatment with surfactant and sonication significantly improves *Agrobacterium*-mediated resistant bud formation and transient expression efficiency in soybean. *J. Integr. Agric.* 14 (7), 1242–1250. doi: 10.1016/S2095-3119(14)60907-2

Hartwig, E. E., and Kühn, R. A. S. (1979). Identification and utilization of a delayed flowering character in soybeans for short-day conditions. *Field Crops Res.* 2 (2), 145–151. doi: 10.1016/0378-4290(79)90017-0

Hauen, W., Coffman, A., and Clasen, B. M. (2014). Improved soybean oil quality by targeted mutagenesis of the fatty acid desaturase 2 gene family. *Plant Biotechnol. J.* 12 (7), 934–940. doi: 10.1111/pbi.12201

Jacobs, T. B., Lafayette, P. R., Schmitz, R. J., and Parrott, W. A. (2015). Targeted genome modifications in soybean with CRISPR/Cas9. *BMBC Biotechnol.* 15 (1), 1–10. doi: 10.1016/s12896-015-0131-2

Kapusi, E., Corduera-Gómez, M., Melnik, S., and Stoger, E. (2017). Heritable genomic fragment deletions and small indels in the putative ENGase gene induced by CRISPR/Cas9 in barley. *Front. Plant Sci.* 8, 540–. doi: 10.3389/fpls.2017.00540

Kong, F. J., Nan, H. Y., Cao, D., Li, Y. W., Fu, F. E., Wang, J. L., et al. (2014). A new dominant gene conditions early flowering and maturity in soybean. *Crop Sci.* 54, 2529–2535. doi: 10.2135/cropsci2014.03.0228
Li, C. L., Nguyen, V., Liu, J., Fu, W. J., Chen, C., Yu, K. F., et al. (2019). Mutagenesis of seed storage protein genes in soybean using CRISPR/Cas9. BMC Res. Notes 12 (1), 176. doi: 10.1186/s13104-019-4207-2

Li, J. F., Norville, J. E., Ach, J., Mccormack, M., Zhang, D., Bush, J., et al. (2013). Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. Nat. Biotechnol. 31 (8), 688–691. doi: 10.1038/nbt.2654

Li, Z., Liu, Z. B., Xing, A., Moon, B. P., Koellhofer, J. P., Huang, L., et al. (2015). Cas9-guide RNA directed genome editing in soybean. Plant Physiol. 169 (2), 960–970. doi: 10.1104/pp.15.00773

Lü, S. (2015). Studies on molecular mechanism regulating the expression of E1 in soybean [Glycine Max (L.) Merr.]. [dissertation]. northeast institute of geography and agroecology. Chin. Acad. Sci.

Lu, S., Zhao, X., Hu, Y., Liu, S., Nan, H., Li, X., et al. (2017). Natural variation at the soybean J locus improves adaptation to the tropics and enhances yield. Nat. Genet. 49 (5), 773. doi: 10.1038/ng.3819

Mao, Y. F., Hui, Z., Xu, N. F., Zhang, B. T., Feng, G., and Zhu, J. K. (2013). Application of CRISPR-Cas system for efficient genome engineering in plants. Mol. Plant 6 (6), 2008–2011. doi: 10.1093/mp/ssl121

Maquat, L. E. (2004). Nonsense-mediated mRNA decay: a comparative analysis of different species. Curr. Genomics 5 (3), 75–190. doi: 10.2174/138920243945493

Maquat, L. E. (2005). Nonsense-mediated mRNA decay in mammals. J. Cell Sci. 118 (9), 1773–1776. doi: 10.1242/jcs.01701

Mcblain, B. A., and Bernard, R. L. (1987). A new gene affecting the time of flowering and maturity in soybeans. J. Heredity 78 (3), 160–162. doi: 10.1093/oxfordjournals.jhered.a110349

Mcblain, B. A., Bernard, R. L., Cremeens, C. R., and Korczak, J. F. (1987). A procedure to identify genes affecting maturity using soybean isoline testers. Crop Sci. 27 (6), 1127–1132. doi: 10.2135/cropsres1987.0011183x002700060008x

Michno, J. M., Wang, X. B., Liu, J. Q., Curtin, S. J., Kono, T. J., and Stupar, R. M. (2015). CRISPR/Cas mutagenesis of soybean and Medicago trunculata using a new web-tool and a modified Cas9 enzyme. GM Crops Food 6 (4), 243–252. doi: 10.1080/21645698.2015.1106063

Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a chimeric gene into the actin promoters in trans. Plant Cell 2, 279–289. doi: 10.1105/tpc.2.4.279

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29 (9), e45. doi: 10.1093/nar/29.9.e45

Pan, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. 8, 2281–2306. doi: 10.1038/nprot.2013.143

Ray, J. D., Hinson, K., Mankono, J. E. B., and Malo, M. F. (1995). Functional control of a long juvenile trait in soybean. Curr. Genet. 25 (4), 1001–1006.

Sagarbl-Maroeif, M. A., Sobhan, K. M., Jorgensen, R. A., and Allard, R. W. (1984). Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. Proc. Natl. Acad. Sci. U.S.A. 81 (24), 8014–8018. doi: 10.1073/pnas.81.24.8014

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., et al. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. Nat. Biotechnol. 31 (8), 686–688. doi: 10.1038/nbt.2650

Sun, X. J., Hu, Z., Chen, R., Jiang, Q. Y., Song, G. H., Zhang, H., et al. (2015). Targeted mutagenesis in soybean using the CRISPR-Cas9 system. Sci. Rep. 5, 10342. doi: 10.1038/srep10342

Tang, F., Yang, S., Liu, J., and Zhu, H. (2016). B7a, a gene controlling nodule specificity in soybeans, encodes a thaliana-like protein but not the one previously reported. Plant Physiol. 170 (1), 26–32. doi: 10.1104/pp.15.01661

Tsukobura, Y., Watanabe, S., Xia, Z. J., Kanamori, H., Yamagata, H., Kaga, A., et al. (2014). Natural variation in the genes responsible for maturity loci E1, E2, E3 and E4 in soybean. Ann. Bot. 113 (3), 429–441. doi: 10.1093/aob/mct269

Upadhyay, S. K., Kumar, J., Alok, A., and Tuli, R. (2013). RNA-guided genome editing for target gene mutations in wheat. Genes Genomes Genet. 3 (12), 2233–2238. doi: 10.1534/g3.113.008847-G3.