Brief Communication

Target base editing in soybean using a modified CRISPR/Cas9 system

Yupeng Cai1,2,†, Li Chen1,†, Yan Zhang1,†, Shan Yuan1,†, Qiang Su1, Shi Sun1, Cunxiang Wu1, Weiwei Yao1, Tianfu Han1 and Wensheng Hou1,†

1National Center for Transgenic Research in Plants, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China
2State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, South China Agricultural University, Guangzhou, China

Received 20 December 2019; revised 26 February 2020; accepted 4 April 2020.
*Correspondence (Tel +86 10 82105172; fax +86 10 82108784; e-mail houwensheng@caas.cn)
†These authors contributed equally to this work.

Keywords: Soybean, CRISPR/Cas9, base editing, GmFT2a, GmFT4, flowering time.

In recent years, the CRISPR (clustered regularly interspaced short palindromic repeat)/Cas9 (CRISPR-associated 9) system has revolutionized functional genomic research and crop improvement, owing to its advantages of simplicity, efficiency, cost-effectiveness and versatility (Chen et al., 2019). There have been abundant reports in plants showing that a small amount of base insertions/deletions (indels) at the intended target site through non-homologous end joining (NHEJ) are typically used to disrupt gene function by frameshift mutations (Chen et al., 2019).

Recently, a new and powerful strategy called ‘base editing’ has been developed from modifications of the CRISPR/Cas9 system, which enables single base substitution into another through a RNA-programmed manner, without requiring DNA double-strand breaks or a donor template (Komor et al., 2016). Cas9 can be modified to serve as a nickase enzyme (Cas9n) by inactivating either of its two endonuclease domains. It has been reported that the fusion of a nicked Cas9 (D10A mutation) with a cytidine deaminase enzyme and an uracil glycosylase inhibitor (UGI) typically allows C to T (or G to A) substitution within a small window (from positions 4 to 8, counting the end distal to the PAM (protospacer adjacent motif) as position 1) of the target sequence (Komor et al., 2016; Li et al., 2017).

Many agriculturally important traits are associated with single nucleotide polymorphism (SNP) variation. Utilization of functional SNPs is also an important means to improve agronomic characters of crops. Therefore, generation of point mutations at specific sites associated with diverse important agronomic traits is of great value in molecular breeding (Mishra et al., 2020). For instance, using base editing in rice, C to T substitution was achieved at a frequency of 1.6%–3.9%. Moreover, the SLR1 mutants with C to T substitution (S97L) display obvious semi-dwarf phenotypes (Lu and Zhu, 2017). In addition, base editing events in two important crops, maize and wheat, have also been reported (Zong et al., 2017). However, it has not been systematically explored in soybean (Glycine max (L.) Merr.), an important gene function by frameshift mutations (Chen et al., 2019).

In this study, we developed a CRISPR/Cas9-mediated base editing tool to specifically induce single base substitution in soybean. We combined the Cas9n (D10A) nickase, rat cytosine deaminase (APOBEC1) and uracil glycosylase inhibitor (UGI) as the base editor (BE), and then cloned these elements into the pTF101.1 vector to generate pTF101.1-BE. BE was driven by a 2X CaMV 35S promoter. Expression of the sgRNA expression cassettes was driven by the Arabidopsis U6 promoter within the pUC57 vector, thus generating the puUC57-sgRNA. Once a desired target sequence is selected, only the DNA sequence encoding the sgRNA needs to be cloned. The expression cassettes containing desired sgRNA sequences were cut from puUC57-sgRNA and inserted into pTF101.1-BE to generate the pTF101.1-sgRNA-BE vector, which would be used to attempt single base substitution.

Ten FLOWERING LOCUS T (FT) homologs have been identified in soybean (Kong et al., 2010). To verify the base editing vector in soybean for its feasibility and efficacy, we selected GmFT2a (Glyma.16G150700) and GmFT4 (Glyma.08G363100) as target genes. The target sequences were located in the first exon of GmFT2a and fourth exon of GmFT4, respectively (Figure 1a). The corresponding pTF101.1-sgRNA-BE vectors were transformed into the soybean cultivar Jack via Agrobacterium tumefaciens-mediated transformation. We then obtained 22 and 34 corresponding T-DNA-positive T0 plants of GmFT2a and GmFT4, respectively. Genomic DNA was extracted from their leaves, and PCR amplification was performed with specific primers designed to specifically amplify the fragment containing the target sites. As shown in Figure 1b, Sanger sequencing results showed that four T0 plants (T0-rt2a-BE-#1, #2, #17, #18; 18.2% of T0 plants) had chimeric nucleotide changes at the target site of GmFT2a. There were two types of resulting base substitution mutations. There were C to T mutations and C to G mutations, both types occurred at position 7 of the target sequence, counting the end distal to the PAM as position 1. Similarly, as shown in Figure 1c, two T0 plants (T0-f4-BE-#27, #34; 6.0% of T0 plants) had chimeric nucleotide changes at the target site of GmFT4. Both the two plants had C to G changes at position 6 of the target sequence. These results indicated the feasibility of this base editing system in soybean.

To detect the heritability of the base editing mutations, the genotypes of the T1 plants of all four T0 lines for GmFT2a and two T0 lines for GmFT4 were examined. The base editing mutations found in T0-rt2a-BE-#1 and T0-rt2a-BE-#17 were transmitted to the T1 generation, but the mutations in other lines were not inherited, maybe because they occurred in somatic cells of the T0 plants. From 59 independent plants of T1-rt2a-BE-#1, we identified 15 plants with a homozygous C to G change, 9 plants with no mutation, and 35 plants with heterozygous mutations containing C to G or C to T changes. From 39
Figure 1  Base editing of GmFT2a and GmFT4 in soybean. (a) Gene structures of GmFT2a and GmFT4 with target sites for base editing. Black stripe, exon. Black line, intron. Grey stripe, untranslated regions. Nucleotides in blue represent the target sequences. Nucleotides in red represent the PAM (protospacer adjacent motif). (b, c) Sequences and peaks of representative mutation types of base editing of GmFT2a and GmFT4 in the T0 lines, respectively. The red arrowheads and underlines indicate the positions of these base editing mutations. (d) Base editing mutation types of GmFT2a in the T1 generation. (e) Sequence and peak of the homozygous ft2a mutant with C to G change. (f) and (g) Flowering time of WT, T2-ft2a-C;G-BE plants and ft2a+1A-Cas9 plants under SD conditions. Red box, magnified view. n, exact numbers of individual plants identified. **, P < 0.01. DAE, days after emergence. The flowering time is shown as the mean values ± standard deviation. (h) and (i) Flowering time of WT, T2-ft2a-C;G-BE plants and ft2a+1A-Cas9 plants under LD conditions. Red box, magnified view. n, exact numbers of individual plants identified. **, P < 0.01. DAE, days after emergence. The flowering time is shown as the mean values ± standard deviation.
independent plants of T1-rt2a-BE-#17, we identified 37 plants with a homozygous C to G change and 2 plants with heterozygous targeted mutations (Figure 1d). The plants with homozygous C to G change (named as rt2a-C_{G-BE}) (Figure 1e) were used for further phenotypic analysis.

In a previous study, we employed the CRISPR/Cas9 system to specifically knockout GmFT2a in the Jack variety. Site-directed mutations in the first exon of GmFT2a generated frameshift mutations and thereby interfered with gene function. The rt2a mutants (1-bp insertion at target site GmFT2a-SP2), frameshift mutation) which exhibit a late-flowering phenotype (Cai et al., 2018) were used as a control and named as rt2a+1A-Cas9 in this study. The GmFT2a of rt2a-C_{G-BE} plants did not generate a frameshift mutation, but the proline at position 8 of its amino acid was changed to alanine. The T2 progeny of rt2a-C_{G-BE}, wild-type plants (WT, Jack) and rt2a+1A-Cas9 plants were grown under LD (16 h light/8 h dark) and SD (12 h light/12 h dark) photoperiodic conditions, and the flowering time was compared by one-way ANOVA with pairwise comparisons. Compared to WT plants under SD conditions (32.0 ± 1.1 days after emergence, DAE), the T2-rt2a-C_{G-BE} plants showed significantly later flowering (34.3 ± 1.7 DAE), although not to the extent of rt2a+1A-Cas9 plants (36.6 ± 1.3 DAE) (Figure 1f, g). Compared to WT plants under LD conditions (43.5 ± 1.5 DAE), the rt2a+1A-Cas9 plants exhibited later flowering by about 4 days (47.5 ± 2.6 DAE), while the T2-rt2a-C_{G-BE} plants showed slightly later flowering (44.9 ± 1.8 DAE) (Figure 1h,i). The relatively weaker late-flowering phenotype of T2-rt2a-C_{G-BE} plants under SD and LD conditions may be due to the influence on a functional domain or an interaction site with other proteins, rather than disruption of protein function through a frameshift mutation as in rt2a+1A-Cas9 plants.

In summary, we have successfully introduced single base substitutions at target sites, indicating the feasibility of this base editing tool in soybean, which may expand the scope of application of CRISPR/Cas9 in soybean. However, it should be noted that most of the changes in our study were C to G mutations, not the C to T changes previously reported in many plant species. We suggest that the variety of C to T or C to G substitution frequencies may depend on the sequence context or species. There are abundant haplotypes in varieties of soybean, which has effects on phenotypic diversity and is associated with diverse important agronomic traits. Utilization of base editing technology holds great potential for future customized genetic improvement and breeding. Moreover, it also has been reported that cytosine base editors generate substantial off-target single nucleotide variants in both animal and plant systems (Jin et al., 2019; Zuo et al., 2019); thus, the deleterious effects should be noticed during the further breeding practices.

Acknowledgements

This work was supported by the Major Science and Technology Projects of China (2016ZX08010-004), the Ministry of Science and Technology of China (2016YFD0100504), the National Natural Science Foundation of China (31871644), and the CAAS (Chinese Academy of Agriculture Sciences) Agricultural Science and Technology Innovation Project.

Conflict of interests

The authors declare no conflict of interest.

Author contributions

Y.C., L.C., Y.Z. and S.Y. performed the experiments. Y.C. and L.C. wrote the manuscript. S.Y., Q.S. and W.Y. assisted in soybean transformation. S.S. and C.W. provided soybean varieties. W.H. and T.H. designed and advised on the experiments and revised the manuscript.

References

Cai, Y., Chen, L., Liu, X., Guo, C., Sun, S., Wu, C., Jiang, B. et al. (2018) CRISPR/Cas9-mediated targeted mutagenesis of GmFT2a delays flowering time in soybean. Plant Biotechnol. J. 16, 176–185.

Chen, K., Wang, Y., Zhang, R., Zhang, H. and Gao, C. (2019) CRISPR/Cas genome editing and precision plant breeding in agriculture. Annu. Rev. Plant Biol. 70, 667–697.

Jin, S., Zong, Y., Gao, Q., Zhu, Z., Wang, Y., Qin, P., Liang, C. et al. (2019) Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. Science, 364, 292.

Komor, A.C., Kim, Y.B., Packer, M.S., Zanis, J.A. and Liu, D.R. (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature, 533, 420.

Kong, F., Liu, B., Xia, Z., Sato, S., Kim, B.M., Watanabe, S., Yamada, T. et al. (2010) Two coordinately regulated homologs of FLOWERING LOCUS T are involved in the control of photoperiodic flowering in soybean. Plant Physiol. 154, 1220–1231.

Li, J., Sun, Y., Du, J., Zhao, Y. and Xia, L. (2017) Generation of targeted point mutations in rice by a modified CRISPR/Cas9 system. Mol. Plant. 10, 526–529.

Lu, Y. and Zhu, J.K. (2017) Precise editing of a target base in the rice genome using a modified CRISPR/Cas9 system. Mol. Plant. 10, 523–525.

Mishra, R., Joshi, R.K. and Zhao, K. (2020) Base editing in crops: current advances, limitations and future implications. Plant Biotechnol. J. 18, 20–31.

Zong, Y., Wang, Y., Li, C., Zhang, R., Chen, K., Ran, Y., Qiu, J.L. et al. (2017) Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. Nat. Biotechnol. 35, 438–440.

Zuo, E., Sun, Y., Wei, W., Yuan, T., Ying, W., Sun, H., Yuan, L. et al. (2019) Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. Science, 364, 289.