Application of the CRISPR–Cas System for Efficient Genome Engineering in Plants

Dear Editor,

Recently, engineered endonucleases, such as Zinc-Finger Nucleases (ZFNs) (Carroll, 2011), Transcription Activator-Like Effector Nucleases (TALENs) (Mahfouz et al., 2011; Li et al., 2012), and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated (Cas) systems (Cong et al., 2013) have been successfully used for gene editing in a variety of species. These systems generate double-strand breaks (DSBs) at target loci to drive site-specific DNA sequence modifications. The modifications include sequence insertion and deletion and other mutations in the host genomes via the error-prone non-homologous end joining (NHEJ) pathway or sequence correction or replacement through the error-free homologous recombination (HR) pathway (Symington and Gautier, 2011). Here, we show that the CRISPR–Cas system can be applied to generate targeted gene mutations and gene corrections in plants, and the system can also be readily engineered to achieve deletion of large DNA fragments and for multiplex gene editing in plants.

Both ZFNs and TALENs have tandem repeats in their DNA-binding domains that can be engineered to recognize specific DNA sequences; the resulting chimeric nucleases can thus be guided to the desired target sequences in the genome to generate DSBs. For each target site, a new ZFN or TALEN chimeric protein needs to be engineered to recognize the target. This has been a major hurdle in the wide use of these two gene-editing systems because engineering a new protein is no trivial task. In comparison, the newly developed CRISPR–Cas system uses a short single guide RNA (sgRNA) to direct the Cas9 endonuclease to complementary target DNA (Gaj et al., 2013), so only a new sgRNA is needed for a new target site. This system thus greatly simplifies the gene-editing process and widens target-site selection. An additional requirement for the Cas9 nuclease activity is the presence of the protospacer-associated motif (PAM) NGG downstream of the target site. This requirement is an important consideration in target-site selection (Jinek et al., 2012).

Several expression vectors were constructed for CRISPR–Cas-based gene editing in *Arabidopsis* and rice, in which the designed sgRNAs and optimized SpCas9 (Cong et al., 2013) were driven by *AtU6* or *OsU6* and *AtUBQ*, *OsUBQ*, or *CaMV 35S* promoters, respectively (Supplemental Figure 1). A transient expression system was developed in *Arabidopsis* protoplasts according to an earlier study (Zhang et al., 2013) to assess the activity of the CRISPR–Cas construct psgR–Cas9–At. The yellow fluorescent protein (YFP)-based reporter contained two partially overlapped YFP fragments that were interrupted by a multiple recognition site (MRS). Three different nucleases, I-SceI, gdTALEN, and CRISPR–Cas, were designed to target the MRS region (Supplemental Figure 2). When the MRS sequence is recognized and cleaved, a functional copy of the YFP gene could be restored through the HR pathway in the cells, so that the efficiency of these endonucleases can be estimated by counting the number of cells emitting yellow fluorescent light using flow cytometry. In an optimized experiment, 11.0% sgR–MRS/YFFP co-transfected protoplasts showed fluorescence—a frequency lower than the 18.8% for gdTALEN but comparable to the 12.5% for I-SceI (Supplemental Figure 2). These results suggested that the CRISPR–Cas system was functional in generating DSBs and triggering gene correction in plant cells.

To test the importance of the PAM sequence for target recognition in plants, an improper sgRNA (sgR-MRS*) with a shifted PAM sequence (from GGG to GGA) was used to target the MRS in the YFFP reporter (Supplemental Table 1 and Supplemental Figure 2). The proportion of YFP florescent cells in the sgR-MRS*-targeted protoplasts was 5.4%, compared to the 11% for the sgR-MRS target with a correct PAM. Thus, the altered PAM sequence greatly reduced but did not abolish the activity of CRISPR–Cas9, suggesting that, although PAM is important, it is not absolutely required for the function of CRISPR–Cas in plant cells.

The sgRNA and Cas9 expression cassettes were cloned into a binary vector that contains a nonfunctional GUUS reporter (Figure 1A and Supplemental Figure 1) for *Agrobacterium*-mediated transformation. Among 44 T1 transgenic *Arabidopsis* plants tested, five showed a GUS signal in their cotyledons (Figure 1B). We did not see any plant organ with a uniform GUS signal, and observed GUS expression in one guard cell but not in the other one in the same stoma (Figure 1B), indicating that the CRISPR–Cas induced cleavage and HR repair events happened in individual cells. Using a SURVEYOR assay, we found that 35 of the 44 plants, including three of the five GUS-positive lines, had mutations at the target site (Figure 1C and Supplemental Table 2). These

doi:10.1093/mp/sst121, Advance Access publication 20 August 2013

Received 9 August 2013; accepted 12 August 2013

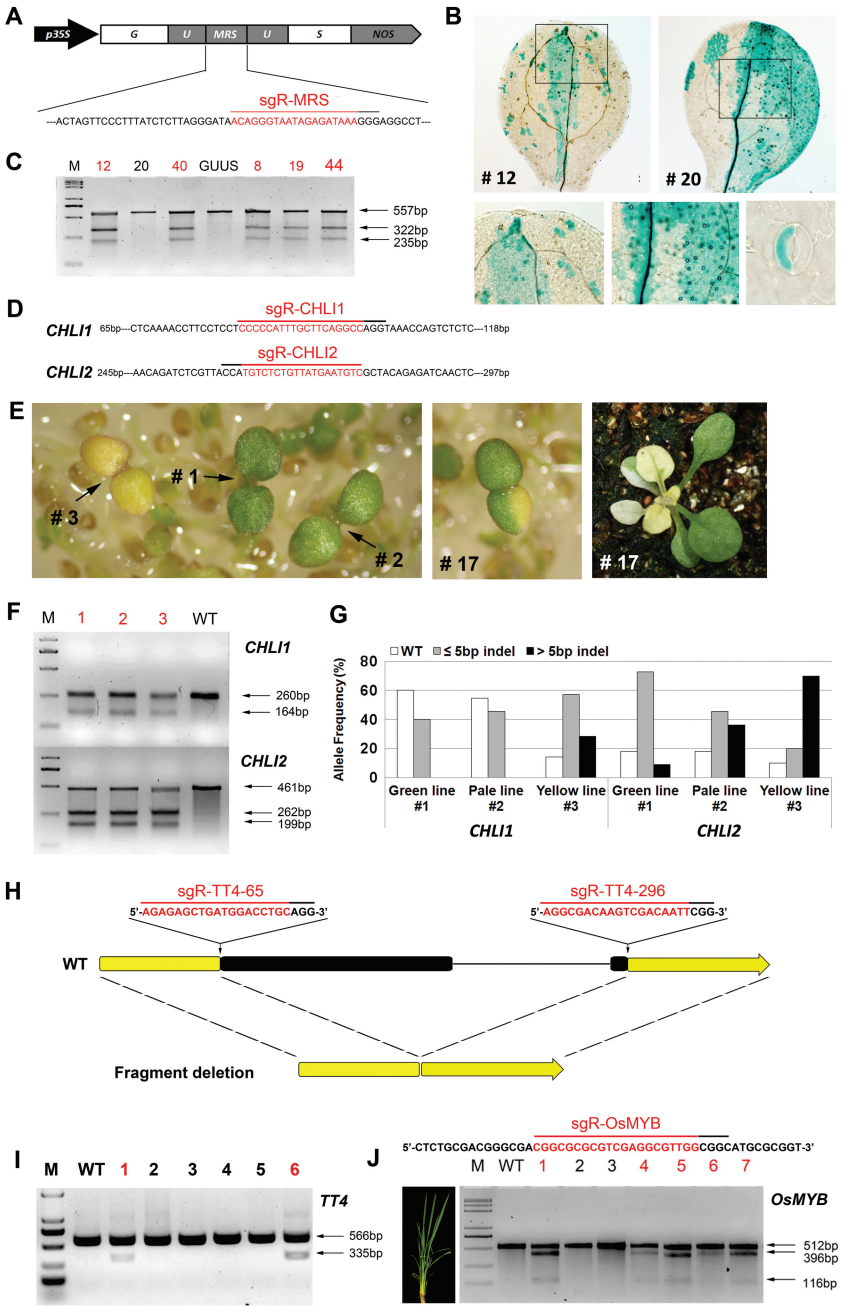

Figure 1. The CRISPR–Cas9 System Induces Efficient Targeted Gene Editing and Correction in Plants.

(A) Schematic diagram of the GUUS reporter and sgR-MRS target site.

(B) GUS staining of cotyledons of two T1 10-day-old seedlings.

(C) SUVEYOR assay for three of the GUS-positive lines (#12, #20, and #40) and three GUS-negative lines (#8, #9, and #44). The 557-bp band corresponds to the wild-type allele while the 322-bp and 235-bp bands correspond to the mutated alleles.

(D) The target sites in the *Arabidopsis CHLI1* and *CHLI2* genes that were simultaneously targeted for mutagenesis.

(E) Leaf color phenotypes of the p2xsgR-CHLI1&2 T1 transformants. Left panel: three seedlings were numbered according to the severity of their cotyledon color phenotype. Middle panel: a T1 seedling showing a mosaic cotyledon phenotype. This plant later developed albino leaves at the rosette stage (right panel).

(F) SUVEYOR assay of the three seedlings in (E) showing mutations at both loci.

(G) Allele frequency of the three seedlings in (E).

(H) Diagram of the two target sites in the *AtTT4* gene and resulting fragment deletion induced by CRISPR–Cas.

(I) AFLP assay detected events of fragment deletion in some T1 seedlings (red-colored numbers).

(J) CRISPR–Cas-based editing of the rice *OsMYB* gene. Shown are the target site of sgRNA, a T0 transgenic rice seedling, and the SURVEYOR assay of seven T0 seedlings.

data suggest that CRISPR–Cas-generated DSB can be repaired via both NHEJ and HR (Supplemental Figure 3), but NHEJ is the dominant DNA repair pathway in plants. Cloning and sequencing of the PCR products from three of the lines with mutations in the target site revealed that deletions were more abundant and longer in size (6–25 bp) than insertions (1–2 bp) (Supplemental Figure 4).

We then tried to simultaneously target two sites in the *Arabidopsis* genome using a CRISPR–Cas9 construct that contained two sgRNA expression cassettes. The *magnesium-chelatase subunit I* (CHLI) genes, *CHLI1* (At4g18480) and *CHLI2* (At5g45930), were selected for the test. The *chli1 chli2* double mutant was albino, while mutations in either gene alone resulted in pale-green plants (Huang and Li, 2009). Two sgRNAs, each targeting one of the *CHLI* genes, were placed into the p2xsgR–Cas-At vector for *Arabidopsis* transformation (Figure 1D and Supplemental Table 2). Out of 60 T1 transformants, 23 could not survive to have true leaves because of their severe albino phenotype. The rest grew slowly and most of them exhibited a mosaic leaf color phenotype (Figure 1E). Three transgenic lines with different leaf colors were selected to test for mutations by the SURVEYOR assay and sequencing (Figure 1E). All three lines were found to harbor mutations in both genes (Figure 1G). In the dark-green and pale-green lines, 60% of the *CHLI1* PCR products were wild-type, while, in the albino line, the proportion of wild-type *CHLI1* was reduced to 14% (Supplemental Figure 5). In contrast, the proportions of wild-type *CHLI2* in all three lines were similar, between 10% and 20% (Supplemental Figure 6). The proportion of small insertion and deletion (indels) (<5 bp) in *CHLI1* was about 50% in all three lines but larger deletions (>5 bp) were only found in the albino line. The rate of long deletions in *CHLI2* increased with the severity of the albino phenotype. In addition, longer deletions were more abundant in *CHLI2* than in *CHLI1*, and more prevailing in the albino line than in the green lines (Figure 1F). Transgenic line #17 had one cotyledon green while the other one had half green and half yellow (Figure 1E). Later, new true leaves emerged as albino, suggesting that double gene mutation events occurred in the shoot apical meristem. The data showed that CRISPR–Cas could be used for multiplex gene editing in plants and suggested that the efficiency of CRISPR–Cas varied at different target sites.

We also tried to modify the *Arabidopsis TT4* (At5g13930) gene. Two sites separated by 230 bp in the *TT4* gene were selected for gene editing. Among the 58 T1 transgenic seedlings, 89% of them had mutations at the first site, 84% at the second site, and 74% at both sites (Supplemental Table 2). Furthermore, an amplification fragment length polymorphism (AFLP) assay was performed with primers flanking the two target sites (Figure 1I). Fifteen plants (26%) showed a short PCR product of about 335 bp in addition to the 566-bp wild-type product, indicating that DSBs were created at both sites, which resulted in the deletion of the fragment between the two sites (Figure 1H). Sequencing of the PCR products from 11 of the 15 seedlings showed that most of the mutations were small indels (61%) and relatively large deletions (32%), while duplications (2.5%) and inversions (4%) were occasionally found (Supplemental Figures 7 and 8).

To test whether the CRISPR–Cas9 system functions in a monocot, we used the system to target the *OsMYB1* gene (LOC_Os1g12700) of rice. Twenty T0 transgenic rice plants were recovered (Figure 1K). Out of 20 T0 transgenic rice plants, 10 were found to be wild-type for the *OsMYB* gene and 10 contained mutations in the target site (Figure 1J). Sequencing revealed that the plants were mosaics of various deletions and small insertions as well as the wild-type allele (Supplemental Figure 9).

Our results demonstrated that the CRISPR–Cas system was efficient in targeted genome engineering in both monocot and dicot plants. In general, the system generated detectable mutations at a frequency of 50–89% for a single locus and 68–74% for double loci in plants (Supplemental Table 2). Our results suggest that CRISPR–Cas can be used not only for targeted gene mutagenesis, but also for gene correction and deletion of large genomic fragments.

SUPPLEMENTARY DATA

Supplementary Data are available at *Molecular Plant Online*.

FUNDING

The work was supported by the Chinese Academy of Sciences. No conflict of interest declared.

Note Added in Proof: Related work from our lab has been published in Feng et al. 2013. Cell Res. doi:10.1038/cr.2013.114

Yanfei Mao[a], *Hui Zhang*[a], *Nanfei Xu*[a], *Botao Zhang*[a], *Feng Gou*[a,b], and *Jian-Kang Zhu*[a,c,1]

[a] Shanghai Center for Plant Stress Biology, Chinese Academy of Sciences, Shanghai 200032, China
[b] University of Chinese Academy of Sciences (CAS), Shanghai 200032, PR China
[c] Department of Horticulture and Landscape Architecture, Purdue University, West Lafayette, IN 47907, USA
[1] To whom correspondence should be addressed. E-mail jkzhu@purdue.edu.

REFERENCES

Carroll, D. (2011). Genome engineering with zinc-finger nucleases. Genetics. **188**, 773–782.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science **339**, 819–823.

Gaj, T., Gersbach, C.A., and Barbas, C.F., III (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. **31**, 397–405.

Huang, Y.S., and Li, H.M. (2009). *Arabidopsis* CHLI2 can substitute for CHLI1. Plant Physiol. **150**, 636–645.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. **337**, 816–821.

Li, T., Liu, B., Spalding, M.H., Weeks, D.P., and Yang, B. (2012). High-efficiency TALEN-based gene editing produces disease-resistant rice. Nat. Biotechnol. **30**, 390–392.

Mahfouz, M.M., Li, L., Shamimuzzaman, M., Wibowo, A., Fang, X., and Zhu, J.K. (2011). *De novo*-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. Proc. Natl Acad. Sci. U S A. **108**, 2623–2628.

Symington, L.S., and Gautier, J. (2011). Double-strand break end resection and repair pathway choice. Annu. Rev. Genet. **45**, 247–271.

Zhang, Y., Zhang, F., Li, X., Baller, J.A., Qi, Y., Starker, C.G., Bogdanove, A.J., and Voytas, D.F. (2013). Transcription activator-like effector nucleases enable efficient plant genome engineering. Plant Physiol **161**, 20–27.