Cas9-gRNA directed genome editing in maize

Corresponding Author: A. Mark Cigan, DuPont Pioneer, 7300 NW 62nd Ave. Johnston, IA 50131, USA
Telephone: 1-(515)-535-3904
E-mail: Mark.Cigan@Pioneer.com

Research Area: Breakthrough Technologies
Targeted Mutagenesis, Precise Gene Editing and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA

Sergei Svitasev¹, Joshua K. Young¹, Christine Schwartz, Huirong Gao, S. Carl Falco, and A. Mark Cigan*

Trait Enabling Technologies
DuPont Pioneer, 7300 NW 62nd Avenue, Johnston, Iowa 50131, USA

Summary

Advances in double-strand break technologies modernize genome editing in maize.
Footnotes:

1 These authors contributed equally to the article.
* Address correspondence to Mark.Cigan@Pioneer.com

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the Journal policy described in the Instructions for Authors (http://www.plantphysiol.org) is Mark Cigan (Mark.Cigan@Pioneer.com).

A.M.C and S.C.F. conceived the project; S.S., J.K.Y., H.G. and A.M.C. designed the experiments; S.S., J.K.Y., and C.S. performed the experiments; S.S., J.K.Y., C.S., and A.M.C. analyzed the data; S.S. and A.M.C. wrote the article.
ABSTRACT

Targeted mutagenesis, editing of endogenous maize genes, and site-specific insertion of a trait gene using Cas9-guide RNA technology is reported in *Zea mays*. DNA vectors expressing maize codon-optimized *Streptococcus pyogenes* Cas9 endonuclease and single guide RNAs were co-introduced with or without DNA repair templates into maize immature embryos by biolistic transformation targeting five different genomic regions: upstream of the *liguleless-1* gene (LIG), male fertility genes (MS26 and MS45) and acetolactate synthase genes (ALS1 and ALS2). Mutations were subsequently identified at all sites targeted and plants containing biallelic multiplex mutations at LIG, MS26 and MS45 were recovered. Biolistic delivery of guide RNAs (as RNA molecules) directly into immature embryo cells containing pre-integrated Cas9 also resulted in targeted mutations. Editing the ALS2 gene using either single-stranded oligonucleotides or double-stranded DNA vectors as repair templates yielded chlorsulfuron resistant plants. Double-strand breaks generated by RNA guided Cas9 endonuclease also stimulated insertion of a trait gene at a site near *liguleless-1* by homology-directed repair. Progeny demonstrated expected Mendelian segregation of mutations, edits, and targeted gene insertions. The examples reported in this study demonstrate the utility of Cas9-guide RNA technology as a plant genome editing tool to enhance plant breeding and crop research needed to meet growing agriculture demands of the future.
INTRODUCTION

The evolution of plant breeding, one of the first scientific activities known in the history of human kind, has allowed rapid progress of our civilization and continues to be important for the improvement of crops and sustainability of agriculture. Traditional breeding, however, is limited by the starting genetic diversity. With the advent of modern molecular biology, breeder’s options have been substantially broadened allowing the purposeful transfer of genetic information from either related or diverse organisms into different plant species, thus, specifically improving agricultural performance. The most common methods of DNA delivery, particle bombardment and Agrobacterium infection, lead to random gene insertions throughout the genome. Integration of trait genes into a specific location is considerably more attractive and can be accomplished through homology-directed repair (HDR). While a few publications before 2009 have reported successful direct application of HDR to modify endogenous plant genes, this occurs at impractically low frequencies (Halfter et al., 1992; Offringa et al., 1993; Miao and Lam, 1995; Kempin et al., 1997; Hanin et al., 2001; Terada et al., 2002; Terada et al., 2007). Studies in different organisms, including plants, have demonstrated that coincident generation of targeted DNA double-strand breaks (DSBs) is central to improving the frequency (up to 1000-fold or more) of gene editing and site-specific gene insertion via HDR (Puchta et al., 1993; Choulika et al., 1995; Smih et al., 1995; Puchta et al., 1996). DSBs in eukaryotic cells can be repaired using two different pathways – non-homologous end-joining (NHEJ) and HDR (Roth and Wilson, 1986; Moore and Haber, 1996; Puchta et al., 1996; Jasin and Rothstein, 2013). The NHEJ pathway is most common in somatic cells and is prone to imperfect repair, which results in mutations such as deletions, insertions, inversions, or translocations. HDR, although less frequent, can be directed by the addition of DNA molecules homologous to the DSB region. By using repair DNA templates that contain sequence variation or promote the insertion of expression cassettes, gene editing and site-specific gene integration can be achieved.

During the past decade, different nucleases capable of generating targeted DSBs were developed and have become important tools to improve gene editing and site-
specific integration in different species, including plants (Puchta and Fauser, 2014; Voytas and Gao, 2014; Kumar and Jain, 2015). While Zinc-finger nucleases (ZFNs), customized homing endonucleases (meganucleases), and transcription activator-like effector nucleases (TALENs) advanced genome editing, the recently discovered Cas-guide RNA system has revolutionized the field (Doudna and Charpentier, 2014; Hsu et al., 2014). This technology emerged from a bacterial adaptive defense system which includes clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated (Cas) endonuclease which protects the host from plasmid and viral infection (Barrangou et al., 2007). Three CRISPR-Cas systems (Type I, II, and III) with different mechanisms have been described in bacteria and archea (Makarova et al., 2011; Makarova et al., 2011). Type II system is the most simple and requires a single RNA-guided Cas9 protein for recognition and cleavage of invading DNA (Gasiunas et al., 2012; Jinek et al., 2012). To generate site-specific DSB, the system minimally requires Cas9 protein and a duplex of CRISPR RNA (crRNA), a trans-encoded CRISPR RNA (tracrRNA), and the presence of a protospacer adjacent motif (PAM), which flanks the 3' end of the crRNA-targeted sequence (Gasiunas et al., 2012; Cong et al., 2013). It has been demonstrated that fusion of the crRNA and tracrRNA into a single guide RNA (gRNA) molecule not only simplified the system consisting now of two instead of three-components, but also significantly improved the frequency of genomic DNA cleavage (Jinek et al., 2012; Mali et al., 2013). Thus, in contrast to ZFNs, meganucleases, and TALENs which require sophisticated protein engineering to establish specific target site recognition (Smith et al., 2000; Arnould et al., 2006; Smith et al., 2006; Maeder et al., 2008; Boch et al., 2009; Weber et al., 2011), Cas9-gRNA reagents are simple to design. Targeting of genomic sequences by Cas9-gRNA relies on binding of PAM nucleotides (NGG for *S. pyogenes*) by Cas9 and base-pairing of the gRNA to the adjacent target site sequence.

The Cas9-gRNA system has been shown to function in many different organisms, including plants. To date, it has been used to generate targeted gene modifications in multiple plant species: *Arabidopsis thaliana*, *Citrus sinensis*, *Nicotiana benthamiana*, *Nicotiana tabaccum*, *Oryza sativa*, *Solanum lycopersicum*, *Sorghum bicolor*, *Triticum aestivum* (for review see Kumar and Jain, 2015), *Zea mays* (Liang et
Cas9-gRNA directed genome editing in maize

al., 2014) and Glycine max (Jacobs et al., 2015). The majority of these studies, however, report genome modifications in plants via NHEJ to generate mutations and gene knock-outs. While knock-outs are very valuable and allow gene function analysis, their application is limited. HDR brings other opportunities in genome editing, but has been more challenging. As a result, reports describing successful gene editing or site-specific trait gene integration through HDR in plants are limited (Cai et al., 2009; Shukla et al., 2009; Townsend et al., 2009; Ainley et al., 2013; D'Halluin et al., 2013; Li et al., 2013; Shan et al., 2013; Zhang et al., 2013).

The data presented in this report demonstrate that the Cas9-gRNA system can efficiently facilitate simultaneous multiple gene knockouts, native gene editing, and site-specific gene integration in maize. Analysis of T1 plants established that the gene knockouts, edits and insertions were heritable and exhibited expected Mendelian segregation. This is the first report of Cas9-gRNA technology application in maize demonstrating gene mutagenesis upon direct delivery of gRNA in the form of RNA molecules, native gene editing, and site-specific gene integration.
RESULTS

Targeted Mutations in Maize

To test whether maize-optimized RNA-guided Cas9 endonuclease can generate DSBs and alter maize sequences through error-prone NHEJ repair, 12 different sites at 5 genomic locations were targeted for cleavage (Fig. 1 and Supplemental Table S1). These target sequences were selected first by identifying the NGG PAM sequence required for *S. pyogenes* Cas9 and then capturing the 17-24 nucleotides immediately upstream of the PAM sequence for use as the spacer in the gRNA. To ensure good U6 polymerase III expression and not introduce a mismatch within the gRNA spacer, all target sequences contained a G at their 5' end. Two of these locations, one near *liguleless-1* and another within the 5th exon of the maize fertility gene MS26, were previously mutated in corn by custom designed homing endonucleases named LIG3:4 (Gao et al., 2010) and Ems26 (Djukanovic et al., 2013), respectively. To indirectly compare the efficiency of generating DSBs by Cas9-gRNA to these homing endonucleases, maize immature embryos were bombarded with DNA vectors containing constitutively expressed meganucleases (LIG3:4, Ems26) or Cas9 components (*S. pyogenes* Cas9, ZmUbi:Cas9, and gRNAs, ZmPolIII:gRNA) and helper genes (ODP2 and Wus) then analyzed by amplicon sequencing for the presence of mutations at the target sites. Embryos bombarded with either Cas9 or gRNA DNA expression cassettes served as controls. Embryos were harvested 7 days after bombardment, total genomic DNA extracted, and fragments surrounding the targeted sequences were amplified by PCR. Barcoded PCR amplicons were sequenced to an approximate depth of 400,000 to 800,000 reads, aligned to their reference sequence and analyzed. It is important to emphasize that during bombardment only a small percentage of embryo cells receive vectors with Cas9 and gRNA expression cassettes while the majority of the cells were not transformed and would be expected to contain unmodified DNA sequences. As shown in Table 1, in contrast to controls, Cas9-gRNA treated embryos yielded a high frequency of mutations. With the exception of one gRNA (MS45-CR1), the majority of the gRNAs tested yielded mutation frequencies greater than 1.3%. Notably, embryos transformed with gRNAs that targeted the same
sequences as LIG3:4 and Ems26 endonucleases yielded 10- to 20-fold higher mutation frequencies (3.9% vs. 0.2% and 1.75% vs. 0.13%, respectively). Ten of the most prevalent mutations generated by either the LIG-CR3 gRNA or the LIG3:4 homing endonuclease are shown in Supplemental Figure S1. In contrast to the predominance of various size deletions present in the sequences derived from the LIG3:4 endonuclease treated embryos and in agreement with previous reports (Feng et al., 2014), the most common types of mutations promoted by LIG-CR3 guided Cas9 were single nucleotide insertions and single nucleotide deletions. Similar results were observed for all other gRNAs used in this study (data not shown).

Maize embryos, stably transformed by co-bombardment with Cas9, gRNA and selectable and visible marker gene (MoPAT-DsRED), were resistant to treatment with bialaphos due to the presence of the phosphotriokinase acetyl transferase gene, MoPAT, and displayed red fluorescence. These traits allowed selection and growth of transformed calli, regeneration of fertile plants, and evaluation of inheritance of the
Cas9-gRNA induced mutations describe above in subsequent generations. As shown in Table 2, mutations were recovered at all targeted loci when multiple gRNA expression cassettes were introduced, either in duplex or triplex, with mutant read frequencies similar to those when gRNAs were delivered individually. These results demonstrate that the maize codon-optimized RNA-guided Cas9 can simultaneously introduce mutations at multiple loci in a single transformation experiment. The frequency of stable event recovery has been reported to be an indicator of cytotoxicity due to off-target recognition for ZFNs in human cell lines and plants (Cornu et al., 2007; Szczepak et al., 2007; Townsend et al., 2009). In this multiplex experiment, the frequency of establishing callus sectors growing on selection was very similar (more than 1 event per embryo when helper genes are co-bombarded) to the frequency in control experiments (no gRNA delivery). Normal event recovery suggests that, for the gRNAs used in this experiment, off-target activity was either absent or below a level sufficient to confer detectable cytotoxicity.

Red fluorescing callus, indicative of a stably integrated DsRED gene, were used to regenerate plants from these multiplex gRNA experiments. Thirty T0 plants (25 duplex and 5 triplex) were selected for mutation detection at the target site by copy-number determination using qPCR and by direct sequencing of the targeted alleles. qPCR analysis revealed that all 30 plants were mutated at the corresponding target sites with a high percentage of plants containing bi-allelic mutations (LIGCas-3 – 91%, MS26Cas-2 – 77%, MS45Cas-2 – 100%; Supplemental Table S2). To corroborate the results generated by qPCR, DNA fragments spanning these three target sites were amplified by PCR, cloned, and sequenced from selected plants. Sequencing data were in agreement with the qPCR results, confirming the high frequency of mutations at all three sites (Table 3), and as expected, T0 plants containing bi-allelic mutations at either Ms26 or Ms45 were male sterile (Cigan et al., 2001; Djukanovic et al., 2013) (data not shown). Plants were also screened by Cas9 and guideRNA-specific PCR (Supplemental Table S3) for the presence of stably integrated copies of Cas9 and gRNA DNA expression cassettes. As shown in Table 3, plant 5 did not contain detectable Cas9 and gRNA expression cassettes suggesting that transient delivery of DNA vectors was sufficient to confer biallelic mutations in this experiment. Due to this
observation, this plant was selected for transmission studies. This plant was fertilized using pollen from wild-type Hi-II plants to produce T1 progeny. Examination of these T1 plants by qPCR and DNA sequencing, demonstrated sexual transmission of parental mutations with expected segregation ratios for all three sites tested (LIGCas-3 – 16:15 [Wt:+T], MS26Cas-2 – 14:12 [Wt/+A:Wt/+G], MS45Cas-2 – 14:13 [Wt/+T:Wt/del]). Additional plants, pollinated with wild-type pollen and containing multiple biallelic gene mutations in the first generation, were observed to transmit these mutant alleles to progeny in an expected Mendelian fashion (data not shown). These results demonstrate that RNA-guided Cas9 can simultaneously cleave multiple chromosomal loci and sexually transmit these mutations to progeny.

**Transient gRNA Delivery into Cells with Pre-Integrated Cas9 Generates Mutations in Maize**

As described above, co-delivery of DNA vectors containing Cas9 and gRNA expression cassettes into maize embryos yielded heritable mutations. However, it has been shown in Arabidopsis (Feng et al., 2014) and wheat (Wang et al., 2014) that progeny plants often do not follow expected Mendelian (1:1) segregation. In the present study, segregation distortion was also observed in progeny of two T0 plants (data not shown). One possibility is that stable integration and constitutive expression of gRNAs and Cas9 endonuclease in T0 and T1 plants leads to somatic mutations and, consequently, to chimeric plants. To overcome the potential problems associated with stable integration of DNA vectors containing Cas9-gRNA components, gRNA in the form of *in vitro* synthesized RNA molecules was co-bombarded with the Cas9 DNA vector and analyzed for mutations by amplicon sequencing 7 days post-transformation. As shown in Table 4, in contrast to the delivery of both Cas9 and gRNA expressed from DNA vectors, the mutation frequencies observed when LIG-CR3 gRNA delivered as RNA were approximately 100-fold lower (2.6% vs. 0.02%, respectively).

One possible explanation for this difference may be the requirement for coincident function of Cas9 and gRNA was not met when gRNA was delivered as RNA and Cas9 as a DNA vector. To test this idea, two *Agrobacterium* vectors (Supplemental Fig. S2) containing maize-codon optimized Cas9 under the transcriptional control of a
constitutive (maize UBI) or a temperature regulated (maize MDH) promoter were introduced into Hi-II embryo cells to establish lines containing pre-integrated genomic copies of Cas9 endonuclease. These Agrobacterium T-DNAs also contained an embryo-preferred END2 promoter regulating the expression of a blue-fluorescence gene (AmCYAN) as a visible marker and an interrupted copy of the DsRED gene transcriptionally regulated by a maize Histone 2B promoter. Part of the DsRED sequence was duplicated in a direct orientation (369 bp fragment) and consisted of two fragments of the DsRED (RF-FP) gene which were separated by a 347 bp spacer that could be targeted by gRNAs (Supplemental data). DSBs within the spacer region promote intramolecular recombination restoring function to the disrupted DsRED gene which results in red fluorescing cells. Maize plants with single-copy T-DNA inserts containing either UBI:Cas9 or MDH:Cas9 were used as a source of immature embryos for delivery of gRNAs as DNA expression cassettes or, as described later, as in vitro transcribed RNA. Blue-fluorescing embryos containing pre-integrated Cas9 were excised and incubated at 28°C (UBI:Cas9) or at 37°C (MDH:Cas9) for 24 hours. Post-bombardment, embryos with MDH:Cas9 were incubated at 37°C for 24 hours and then moved to 28°C. As shown in Supplemental Figure S3, in contrast to control (no gRNA), UBI:Cas9 and MDH:Cas9 containing embryos bombarded with two DNA-expressed gRNAs that targeted sequences within the 347 bp spacer, readily produced red fluorescing foci after 5 days demonstrating expression of functional Cas9 protein in these plants.

To measure mutation frequencies at the LIG and MS26 endogenous target sites, LIG-CR3 and MS26-CR2 gRNAs as DNA vectors (25 ng/shot) or as RNA molecules (100 ng/shot) were delivered into MDH:Cas9 and UBI:Cas9 containing embryo cells with temperature treatments described above. In these experiments, embryos were harvested two days post-bombardment and analyzed by amplicon sequencing. As opposed to the results observed in the previous experiment, nearly similar frequencies were detected for gRNAs delivered as DNA vectors and as RNA molecules, particularly in the case of Cas9 regulated by the Ubiquitin promoter (Table 4). Together these data demonstrate that delivery of gRNA in the form of RNA directly into maize cells...
containing a pre-integrated Cas9 is a viable alternative to DNA delivery for the
generation of mutations in plant cells.

Editing Maize ALS2

Sulfonylurea herbicides prevent branched amino acid biosynthesis in plants due
to the inhibition of the enzyme acetolactate synthase (ALS) (Haughn et al., 1988; Lee et
al., 1988; Yu and Powles, 2014). Resistance to one of these herbicides, chlorsulfuron,
has been described as a result of single amino acid changes in the ALS protein at
position 197 in Arabidopsis (Pro to Ser) (Haughn et al., 1988) and tobacco (Pro to Gln
or Ala) (Lee et al., 1988) and at a corresponding location, position 178, in soybean (Pro
to Ser) (Walter et al., 2014).

ALS was chosen for gene editing to demonstrate that RNA-guided Cas9 can
facilitate specific DNA sequence changes in a native maize gene. There are two ALS
genes in maize, ALS1 and ALS2, located on chromosomes 4 and 5, respectively (Burr
and Burr, 1991), with 94% sequence identity at the DNA level. In the first experiments,
gRNA-expressing DNA vectors targeting three different sites within both ALS1 and
ALS2 genes were tested (Fig. 1D and Supplemental Table S1). As these gRNAs were
not gene-specific, both ALS1 and ALS2 were mutated with approximately the same
efficiency (data not shown) and resulted in low stable event recovery. As acetolactate
synthase is a critical enzyme for cell function in plants, callus events containing
simultaneous bi-allelic knockouts of ALS1 and ALS2 genes would not be expected to
survive. It is likely that the rare events recovered in this experiment did not receive
gRNA or Cas9 necessary to generate mutations, and therefore contained one or more
functional ALS alleles. Alternatively, alleles may have been mutated but still retained
wild-type function. To overcome this problem, an ALS2 specific ALS-CR4 gRNA (Fig.
2A) was designed based on the polymorphisms between ALS1 and ALS2 nucleotide
sequences, and tested. As shown in Figure 2B, amplicon sequence analysis revealed
that ALS2 was highly preferred over ALS1 when ALS-CR4 gRNA was used. One
important difference contributing to this result was that ALS1 lacks an NGG PAM
sequence required for DNA cleavage by the S. pyogenes Cas9 endonuclease.
To generate ALS2 edited alleles, a 794 bp fragment of homology was cloned into a plasmid vector and two 127 nt single-stranded DNA oligos were tested as repair templates (Fig. 2C). The 794 bp fragment had the same sequence modifications as Oligo1. Repair templates contained several nucleotide changes in comparison to the native sequence (Fig. 2C and Supplemental data). Single-stranded Oligo1 and the 794 bp repair templates included a single nucleotide change which would direct editing of DNA sequences corresponding to the proline at amino acid position 165 to a serine (P165S), as well as three additional changes within the ALS-CR4 target site and PAM sequence. Modification of the PAM sequence within the repair template altered the methionine codon (AUG) to isoleucine (AUU), which naturally occurs in the ALS1 gene (Fig. 2 A and C). A second 127 nt single-stranded oligo repair template (Oligo2) was also tested which preserved the methionine at position 157 but contained three
additional single nucleotide changes in the sequence which would influence base pairing with the ALS-CR4 gRNA (Fig. 2C).

Approximately 1,000 immature embryos per treatment were bombarded with the two oligo or single plasmid repair templates, Cas9, ALS-CR4 gRNA, and MoPAT-DsRED in DNA expression cassettes and placed on media to select for bialaphos resistance conferred by PAT. Five weeks post-transformation, two hundred (per treatment) randomly selected independent young callus sectors growing on selective media were separated from the embryos and transferred to fresh bialaphos plates. The remaining embryos (>800 per treatment) with developing callus events were transferred to the plates containing 100 ppm of chlorosulfuron as direct selection for an edited ALS2 gene. A month later, a total of 384 randomly picked callus sectors growing on bialaphos (approximately 130 events for each repair template) and 7 callus sectors that continued growing on media with chlorosulfuron were analyzed by PCR amplification and sequencing. Edited ALS2 alleles were detected in nine callus sectors: two derived from the callus sectors growing on bialaphos and generated using the 794 bp repair DNA template, and the remaining 7 derived from chlorosulfuron resistant callus sectors edited using the 127 nt single-stranded oligos, three by Oligo1 and four by Oligo2 (Fig. 2D; and Supplemental Table S4). The second ALS2 allele in these callus sectors was mutated as a result of NHEJ repair (data not shown). Analysis of the ALS1 gene revealed only wild-type sequence confirming high specificity of the ALS-CR4 gRNA.

Plants were regenerated from 7 out of 9 callus sectors containing edited ALS2 alleles for additional molecular analysis and progeny testing. DNA sequence analysis of ALS2 alleles confirmed the presence of the P165S modifications as well as the other nucleotide changes associated with the respective repair templates. T1 and T2 progeny of two T0 plants generated from different callus events (794 bp repair DNA and Oligo2) were analyzed to evaluate the inheritance of the edited ALS2 alleles. Progeny plants derived from crosses using pollen from wild type Hi-II plants were analyzed by sequencing and demonstrated sexual transmission of the edited alleles observed in the parent plant with expected 1:1 segregation ratio (57:56 and 47:49, respectively). To test whether the edited ALS sequence confers herbicide resistance, selected four-week old segregating T1 plants with edited and wild-type ALS2 alleles were sprayed with four
different concentrations of chlorsulfuron (50, 100 (1x), 200, and 400 mg/liter). Three weeks after treatment, plants with an edited allele showed normal phenotype, while plants with only wild-type alleles demonstrated strong signs of senescence (Fig. 3A). In addition, embryos isolated from seed that had been pollinated with wild-type HI-II pollen, were germinated on media with 100 ppm of chlorsulfuron. Fourteen days after germination, plants with edited alleles showed normal height and a well-developed root system, while plants with wild-type alleles were short and did not develop roots (Fig. 3B).

These experiments demonstrate that Cas9-gRNA can stimulate HDR-dependent targeted sequence modifications in maize resulting in plants with an edited endogenous gene which properly transmits to subsequent generations.

**HDR-Mediated Gene Insertion**
In contrast to Agrobacterium- or biolistic-mediated delivery of trait genes which can often result in random insertions and potential disruption of native gene function, double-strand breaks using site-directed nucleases would allow for precise gene integration at a desired genomic location. In this experiment, given the differences in mutation frequencies observed between RNA-guided Cas9 and meganucleases described above, the ability of these two DSB technologies to promote site-specific gene insertion was compared in maize. As LIG-CR3-guided Cas9 cleaves genomic DNA at the same position as LIG3:4 meganuclease, these DSB reagents were selected for this study. DNA donor repair template contained the constitutively expressed PAT gene (UBI:MoPAT) flanked by approximately 1.0 kb of DNA fragments homologous to genomic sequences immediately adjacent to the LIG cleavage site (Supplemental Fig. S1, S4 and Supplemental data). This donor vector was co-bombarded with Cas9 and LIG-CR3 gRNA expression cassettes or LIG3:4 homing endonuclease vector. In a separate experiment, Cas9, gRNA and a modified donor DNA template containing UBI regulated MoPAT-DsRED fusion flanked with the same regions of homology were delivered on the same vector (Supplemental Fig. S4). To compare different delivery systems, this vector was moved into Agrobacterium strain LBA4044 and used in an Agrobacterium-mediated transformation experiment.

As delivery of the donor DNA would result in both random and targeted insertions, approximately two month old callus sectors growing on bialaphos-containing media were screened for mutations and gene insertions at the LIG target site. In contrast to sectors derived from biolistic-transformed LIG3:4 meganuclease, qPCR copy-number analysis revealed that Cas9-gRNA yielded a higher target site mutation frequency (>83% vs 6% mutant reads; Table 5). Notably, for the Cas9-gRNA system, the mutation frequency was independent of the delivery method (bombardment or Agrobacterium-mediated transformation). Callus events that contained site-specific gene insertions were identified by junction PCR using primer pairs shown in Figure 4A. Amplification products from positive integration events were verified by subcloning and sequencing. As shown in Table 5, two of the 288 callus events in the meganuclease experiment yielded amplicons across both junctions (0.7% HDR frequency), while biolistic delivery of Cas9-gRNA and donor DNA resulted in frequencies of 2.5% and 4%,
Cas9-gRNA directed genome editing in maize

when used either separately or as a single vector, respectively. Screening of the 192 callus events from Agrobacterium-mediated delivery did not produce amplicons indicative of HDR-mediated gene insertions (Table 5), suggesting that biolistic transformation is the preferred delivery method to generate gene insertions at the LIG site in maize.

T0 plants regenerated from 7 independent HDR positive events generated in the experiment when donor DNA, Cas9, and gRNA were delivered separately, were used for additional molecular characterization. DNA blot hybridization analysis of regenerated plants was conducted using genomic- and donor DNA-specific probes shown in Figure 4A. These analyses, in addition to sequencing across the entire inserted fragment (Supplemental data), demonstrated that plants regenerated from two events (Events 1 and 2) contained a single copy of an intact MoPAT gene indicative of homology-directed gene insertion at the LIG target site (Fig. 4). Plants regenerated
from the remaining 5 events contained extra, rearranged and randomly integrated
copies of MoPAT. It should be noted that upon further molecular examination of Events
1 and 2 by PCR analysis, the Cas9 and gRNA DNA vectors were not detected in plants
suggesting that transient expression of these components was sufficient to stimulate
gene integration.

To examine transmission and segregation of MoPAT, T0 plants from Events 1
and 2 were fertilized with wild-type pollen to produce T1 seeds. Ninety-six T1 progeny
plants from Event 1 and 64 plants from Event 2, were analyzed by junction PCR and gel
electrophoresis, and demonstrated expected Mendelian segregation of the gene
integrated at the LIG target site (48:48 and 31:33, respectively). DNA hybridization
analysis performed on 20 random, junction PCR positive T1 plants from both events,
using DNA probes described above, confirmed nuclear stability and integrity of MoPAT
in progeny plants (data not shown).

Together these genetic and molecular data demonstrate that maize optimized
RNA-guided Cas9 can be used to stimulate homology-directed repair for the insertion of
genes at specific locations in maize chromosomes. Moreover, in these experiments the
frequency of site-specific trait integration depended upon the endonuclease target site
cleavage efficiency and repair template delivery method.
DISCUSSION

During the past decade, several technologies, ZFNs, meganucleases, TALENs and, recently, Cas9-gRNA system were developed to generate targeted DSB that can be used for genome editing by NHEJ or HDR. DSB technologies have three major applications in plant biotechnology: gene mutagenesis resulting in gene knock-outs, gene editing allowing modification of a gene product or metabolic pathway in a desired way, and site-specific trait insertion which would allow gene stacking in plant breeding programs. Here we reported successful application of the Cas9-gRNA technology to generate multiple gene knock-outs, ALS2 gene editing resulting in plants resistant to chlorsulfuron, and site-specific gene integration in an important crop species, maize.

Guide RNA Evaluation by Amplicon Sequencing

Maize transformation is a time consuming process, often taking three to four months from DNA delivery into immature embryo cells to plant regeneration. Thus, testing the DSB reagents is an important step in developing proper tools for genome editing experiments. The rapid method based on amplicon sequencing described here allows for evaluation of mutation frequencies at any chromosomal site targeted by different gRNAs within two weeks after transformation. Analysis of multiple gRNAs (including ones not presented in this report) indicated that >90% of the guides tested led to a high frequency of mutations at the corresponding target sites. Additional studies will be needed to determine if the inability of certain guide RNAs that did not promote high frequencies of mutations in planta can be attributed to such factors as target site accessibility, secondary structure of gRNA molecules, and/or high frequency of flawless DSB repair at some target sites.

NHEJ-Mediated Gene Mutagenesis

Selected gRNAs with high mutation frequencies were used in multiplexed (duplexed and triplexed) targeted mutagenesis experiments. All genes were targeted with 100% success rate, and more than 80% of the regenerated T0 plants contained biallelic mutations. These results, along with previous publications, illustrate that the Cas9-gRNA technology can be used to efficiently promote the alteration of individual or
multiple gene sequences (for review see Kumar and Jain, 2015), gene families or homologous genes in polyploid species (Wang et al., 2014), and delete chromosomal segments of various sizes (Brooks et al., 2014; Zhou et al., 2014; Gao et al., 2015). In this report, analysis of T0 plants and T1 segregating populations indicated that stably integrated Cas9 and gRNA expression cassettes remain active, often resulting in chimeric T0 plants and segregation distortions. Similar observations have been previously reported in several plant species (Brooks et al., 2014; Feng et al., 2014; Jiang et al., 2014; Wang et al., 2014; Zhang et al., 2014). At the same time, T0 plants regenerated from callus events without integrated Cas9-gRNA components were not chimeric and exhibited proper sexual transmission of the mutated alleles to the next generation. These results suggest that transient delivery of Cas9 and gRNA is sufficient to mutagenize multiple alleles and can speed the discovery and breeding process. In mammalian systems, both Cas9 and gRNA can be delivered transiently in the form of RNA-protein complex providing high mutation frequencies (Kim et al., 2014; Lin et al., 2014; Zuris et al., 2015). However, this approach has limited application among plant species due to the currently available transformation methods. Therefore, one method to limit the Cas9-gRNA system to transient activity is to deliver gRNA in the form of RNA molecules rather than DNA expression cassettes. In this report, simultaneous delivery of gRNA in the form of RNA and a DNA vector containing a Cas9 expression cassette, yielded about 100-fold lower frequencies of mutations in comparison to the control in which both components were delivered as DNA vectors. Because Cas9-gRNA is a two component system, both gRNA and Cas9 protein must be present in a cell at the same time. In the experiment in which gRNA and a Cas9 gene were delivered simultaneously, this condition likely was not met due to the predicted rapid turnover of RNA molecules and the delayed and, therefore, non-coincidental expression of functional Cas9 protein encoded on a DNA vector. To overcome this challenge, maize lines containing constitutively or conditionally expressed pre-integrated Cas9 were generated. Amplicon sequence analysis performed on embryos two days after transformation demonstrated that transient gRNA delivery into the maize embryo cells with pre-integrated Cas9 resulted in mutation frequencies comparable to delivery of Cas9 and gRNA as DNA vectors. It is likely that the frequency of mutations would be
reduced in comparison to the experiments when DNA vectors were used to deliver
gRNAs because expression cassettes often stably integrate into the genome and, as
the result, the Cas9-gRNA system continues to be active. Nevertheless, delivery of
gRNA as RNA may not limit successful gene editing and site-specific gene integration,
as gene insertion events were recovered without stable integration of either Cas9 or
gRNA expression vectors. Therefore, gRNA delivery in the form of RNA can be a viable
alternative to gRNA delivery in DNA vectors for plant gene and genome editing.

HDR-Mediated Gene Editing and Targeted Gene Insertion

While gene mutagenesis can be the result of the error-prone process of DNA
DSB repair through NHEJ, gene editing and targeted gene integration rely on the HDR
repair pathway. Cas9-gRNA has been well documented in different plant species for
targeted gene mutagenesis, both single and multiplexed (for review see Kumar and
Jain, 2015), however examples of HDR-mediated gene editing and site-specific gene
insertion have been limited to model systems (Li et al., 2013; Shan et al., 2013). Until
now, delivery of the gene of interest and its targeted insertion via HDR in plants remains
a serious challenge (Feng et al., 2013; Voytas and Gao, 2014).

Genes and genomes in many important crop species are often present in
multiple copies due to rearrangements, duplications and/or polyploidization. In some
instances, all copies of a gene may need to be mutagenized to result in complete
inactivation. However, in some cases, only one copy of a gene family would require
editing to produce a desired phenotype. In the present study, the acetolactate synthase
(ALS) gene was chosen to test the ability of the Cas9-gRNA system to facilitate
endogenous gene editing via HDR in maize immature embryo cells transformed by
particle bombardment. ALS is a well characterized gene family consisting of two genes,
ALS1 and ALS2, which can be edited to confer plant resistance to sulfonylurea class
herbicides. Several gRNAs targeting both genes simultaneously, or ALS2 uniquely,
were tested. Delivery of a guideRNA targeting both ALS1 and ALS2 lead to low
recovery of stable events likely due to the high frequency of biallelic knock-out
mutations and the important nature of the gene. However, when an ALS2-specific
gRNA was tested, typical frequencies of event recovery were observed. To edit ALS2,
three different repair templates (double-stranded vector DNA and two single-stranded oligos) were tested. All three experimental designs resulted in editing of ALS2, indicating that small single-stranded DNA oligonucleotides were sufficient for gene editing experiments in maize. Although all regenerated plants contained integrated and, presumably, functional copies of Cas9 and gRNA expression cassettes, only wild-type alleles of ALS1 were observed, suggesting high targeting specificity by the ALS2-CR4 gRNA. This experiment demonstrated that by carefully choosing target sites and corresponding gRNAs, multiple or individual genes within a family may be targeted if sufficient nucleotide polymorphisms are present between the family members.

Targeted gene insertion allows clustering of multiple genes in a genomic location which would allow for segregation as a single locus, significantly simplifying the breeding process. Previously, gene insertion at a specific target site was demonstrated using ZFNs in tobacco (Cai et al., 2009) and maize (Shukla et al., 2009; Ainley et al., 2013), and meganucleases in cotton (D'Halluin et al., 2013). The Cas9-gRNA system was tested for its ability to facilitate targeted gene insertion in maize immature embryo cells in comparison to a meganuclease targeting the same genomic location. Two different delivery systems, biolistic and Agrobacterium-mediated transformation, were tested. Only particle bombardment experiments resulted in events with a gene inserted into the target site through an HDR process. Delivery of the donor DNA on the same plasmid with Cas9-gRNA components yielded approximately two times higher frequency of integration events in comparison to experiments in which components were delivered as separate vectors. In this experiment, Cas9-gRNA yielded approximately 5 times more integration events than LIG3:4 meganuclease. This observed difference was likely due to the higher activity of Cas9-LIG-CR-3 gRNA as determined by the mutation frequencies measured at the target site. The inability to recovery integration events in the Agrobacterium-mediated experiment may be explained by the low number of copies of the T-DNA delivered into the plant cell compared with particle bombardment, which delivers multiple copies of all DNA molecules. However, this does not necessarily mean that Agrobacterium transformation cannot result in HDR-mediated gene integration, but likely indicates lower frequency of such events (Svitashev, unpublished observations). Therefore, selection of a delivery
method is very important and should be considered based on the goal of individual experiments.

The efficiency of gene editing and targeted gene insertions has typically been associated with the activity of a nuclease and, therefore, frequency of mutations at a given target site. Although the ability of a nuclease to generate DSB is the most important condition, this is not the only requirement and can sometimes be misleading in predicting frequency of HDR at the target site. Factors such as efficiency of DSB repair, target site accessibility, presence of repetitive sequences close to the target site, and yet to be discovered factors may significantly affect frequencies of gene editing and targeted gene insertions. Our unpublished data indicate that although similar mutation rates can be observed at different target sites, gene editing and gene insertion frequencies can vary from less than 1% to as high as 15%.

CONCLUSIONS

Only three years after the first publications introduced the Cas-gRNA system, it has become the DSB technology of choice for many laboratories working in the field of genome editing. The advantages of the system lay not only in its simplicity and activity, but also in its versatility to more fully address problems associated with DSB technology applications in plant biology. These challenges include the inability to generate DSBs at a specific genomic location and lack of complete specificity that can result in off-target DNA cleavage. In the not too distant future, increasing the density of target sites may be achieved by the discovery and utilization of new Cas9 proteins with different PAM recognition sequences. PAM diversity would be particularly useful in plant species with high AT content, while longer PAM sequences may increase the proportion of unique and highly specific target sites. Testing of S. pyogenes Cas9-gRNA in a variety of plant species has resulted in extremely low frequency of non-specific DNA cleavage and, therefore, high specificity of the system (Feng et al., 2013; Shan et al., 2013; Brooks et al., 2014; Zhang et al., 2014). While targeting various sites may yield different results, incorporation of the Cas9 D10A nickase (Gasiunas et al., 2012) may further improve gene editing options for Cas9 as the nickase version has been successfully applied to
improve specificity (Mali et al., 2013; Cho et al., 2014; Fauser et al., 2014).
Furthermore, transient delivery of purified Cas9 protein and gRNA (as a
ribonucleoprotein complex) into human culture cells has demonstrated high frequencies
of targeted mutagenesis and gene editing while reducing off-target mutations (Kim et
al., 2014; Lin et al., 2014). These combined approaches may further improve the
precision and accuracy of genome modification in plants. Thus, targeted mutagenesis
or editing of specific genes or gene families to yield new allelic variants will no longer be
a limiting step for the discovery of gene function and establishment of their relationship
to complex biological pathways. Adoption of robust genome modification methods to
introduce genetic variation aided by inexpensive sequencing and new plant
transformation methods will accelerate the continued development of new breeding
techniques to address challenges of modern agriculture.

MATERIALS AND METHODS

Plasmids and Reagents Used for Plant Transformation

Standard DNA techniques as described in Sambrook et al (Sambrook et al.
1898) were used for vector construction. The Cas9 sequence from Streptococcus
pyogenes M1 GAS (SF370) was maize codon-optimized, and designed to contain the
potato ST-LS1 intron (Accession number X04753, (nucleotides 2837-2892) at
nucleotide position (+402-591). To facilitate nuclear localization of the Cas9 protein in
maize cells, the SV40 (MAPKKKRKV) and the A. tumefaciens VirD2
(KRPRDRHDGEGLGKRAR) nuclear localization signals were added at the amino and
carboxyl-termini of the Cas9 open reading frame, respectively (Supplemental data).
The chemically synthesized Cas9 gene, contained on a 4.335 kb Ncol-Hpal DNA
fragment, was subcloned into PHP17720 (Sabelli et al., 2009) digested with Ncol-Hpal
which links Cas9 to a maize constitutive Ubiquitin-1 promoter, including the first intron (–
899 to+1092, Christensen et al., 1992) while transcription is terminated by the addition
of the 3’ sequences from the potato proteinase inhibitor II gene (PinII) (nucleotides 2–
310, (An et al., 1989) to generate UBI:Cas9 vector. Single guide RNAs (gRNAs) were
designed using the methods described by Mali et al., 2013. A maize U6 polymerase III promoter and terminator residing on chromosome 8 (nt 165535024-165536023 Maize (B73) Public Genome Assembly, (AGP_v3.8)) were isolated and used to direct initiation and termination of gRNAs, respectively. Two BbsI restriction endonuclease sites were introduced in an inverted tandem orientation with cleavage orientated in an outward direction as described in Cong et al., 2013 to facilitate the rapid introduction of maize genomic DNA target sequences into the gRNA expression constructs (Supplemental data). Only target sequences starting with a G nucleotide were used to promote favorable polymerase III expression of the gRNA. The guideRNA expression cassettes contained on a 1.1 kb XhoI-EcoRI DNA fragment were subcloned into Bluescript SK vectors and used for bombardment and subsequent vector construction.

Two Agrobacterium-transformation vectors were designed and introduced into maize to generate stably integrated Cas9 expression cassettes. To generate constitutively expressed Cas9, the UBI:Cas9 vector was digested with HindIII-EcoRI to release a 6.6 kb HindIII-EcoRI fragment (UBI:Cas9) and this fragment was subcloned into HindIII-EcoRI digested plasmid pSB11 (Ishida et al., 1996). To generate the UBI:Cas9 T-DNA vector, the pSB11-Ubi:Cas9 vector was digested with NotI and a 5.8 kb NotI fragment containing the blue-fluorescence gene (CFP: AmCYAN1, Clontech Laboratories Inc., CA, USA) regulated by the maize END2 promoter (Chromosome 8, nt 171120325-171119383 Maize (B73) Public Genome Assembly, (AGP_v3.8)) a visible marker for the identification of transformed embryos), and a red fluorescence gene (pDsRED Express Product number 632412; Clontech Laboratories Inc., CA, USA) regulated by the maize Histone 2B promoter (Chromosome 3, nt 16569805–16571259 Maize (B73) Public Genome Assembly, (AGP_v3.8)) was subcloned downstream of UBI:Cas9. A portion of DsRED gene in this vector was duplicated (nt 148-516) in a direct orientation with 369 bp of overlap. The duplicated fragments were separated by a 347-bp spacer contained sequences compatible for recognition and targeting by gRNAs and the LIG3:4 meganuclease (Supplemental Fig. S2). To generate a conditionally expressed Cas9 (MDH:Cas9), a 1048 bp fragment of temperature regulated promoter from the maize mannitol dehydrogenase gene (MDH; Chromosome 2, nt 8783228-8784275 Maize (B73) Public Genome Assembly, (AGP_v3.2)) was modified to contain
an RcaI restriction site at the 3’ end of the sequence (Cigan and Unger-Wallace, 2013) to accommodate translational fusion to the maize optimized Cas9 to generate MDH:Cas9 expression cassette. The 6.6 kb HindIII-EcoRI DNA fragment in UBI:Cas9 T-DNA vector was replaced with the MDH:Cas9 DNA contained on a 5.7 kb HindIII-EcoRI DNA fragment to generate the MDH:Cas9 T-DNA vector. These plasmids were introduced into Agrobacterium strain LBA4404 (Komari et al., 1996) by electroporation using a Gene Pulser II (Bio-Rad, CA, USA) as described by Gao (Gao et al 2010).

To generate gRNA in the form of RNA molecules, the maize-optimized U6 polymerase III gRNA expression cassettes were amplified by PCR using a 5’ oligonucleotide primer that also contained the sequence of the T7 polymerase promoter and transcriptional initiation signal just 5’ of the spacer to gene. T7 in vitro transcription was carried out with the AmpliScribe™ T7-Flash™ kit (Epicentre, WI, USA) according to the manufacturer’s recommendations and products were purified using NucAway™ Spin Columns (Invitrogen, Life Technologies Inc., USA) followed by ethanol precipitation.

The LIG3:4 and the Ems26 endonucleases vectors were previously described (Gao et al., 2010; Djukanovic et al., 2013, respectively). Plasmids containing cell division promoting polypeptides, ZmODP2 and ZmWus, and selectable and visible markers, MoPAT-DsRED (a translational fusion of the bialophos resistance gene, phosphinothricin-N-acetyl-transferase, and the red fluorescent protein, DsRED), were previously described in Ananiev et al., 2009.

For ALS2 gene editing, single-stranded 127 nt long oligos (sense and antisense strands) homologous to the fragment spanning the ALSCas-4 target site with 4 and 7 single nucleotide changes (Fig. 2C) were synthesized (Integrated DNA Technologies, IA, USA). SNPs within gRNA target site and PAM sequence were introduced to prevent the donor DNA from cleavage by Cas9 protein and SNPs, a C to T, Oligo1, and a C to T and a G to C replacements, downstream from the cleavage site, to change proline to serine at position 165 in the ALS2 gene. These oligos (sense and antisense strands) were annealed to generate double-stranded DNA, digested with BcoDI and ligated to PCR amplified genomic DNA fragments upstream and downstream from the oligo sequence with primers 5’-ACAGCCGCGCAACCATGGCCA-3’ (forward) and 5’-
ACCATGGGGACGGAATCGAGCA-3' (reverse) and 5'-
ATAAGAGTAGTTGTGACCGGGA-3' (forward) and 5'-
CTGCTCAAGCAACTCAGTCGCAGGG-3' (reverse), respectively, also digested with
BcoDI to extend the total homology region to 794 bp, and cloned into the pUC19 vector.

Vectors for HDR-mediated targeted genome integration were generated by
flanking a UBI:MoPAT or UBI:MoPAT-DsRED fusion with 1099 bp and 1035 bp maize
genomic DNA fragments homologous to the sequences immediately upstream and
downstream from the cleavage site of the of LigCas-3 target, HR1 and HR2,
respectively (Fig. 4A). Vector containing the MoPAT-DsRED fusion was retrofitted
either with UBI:Cas9 and U6:gRNA or LIG3:4 meganuclease expression cassettes
(Supplemental Fig. S4B). For Agrobacterium-mediated transformation, vector
containing expression cassettes shown on Supplemental Figure S4B was moved into
the A. tumefaciens strain LBA4404 by electroporation using a Gene Pulser II (Bio-Red,
CA, USA).

Plant Material

Publicly available maize (Zea mays L.) hybrid high type II (Hi-II) line (Armstrong
and Green, 1985) was obtained from internal Pioneer sources.

To generate plants with either constitutive or regulated expression of Cas9
protein, constructs described above and shown in Supplemental Figure S2 were used to
transform Hi-II maize embryos using Agrobacterium-mediated transformation.
Regenerated T0 plants were screened by qPCR and DNA blot hybridization analysis for
single-copy T-DNA insertions. Independent UBI:Cas9 (4) and DH:Cas9 (11) T-DNA
containing plants were identified; T0 plants were fertilized with wild-type Hi-II pollen and
blue fluorescing embryos were used for further biolistic transformation experiments.

Maize Transformation

Biolistic-mediated transformation of maize immature embryos was performed as
described in Ananiev et al (Ananiev et al., 2009) with slight modification. Briefly, the
DNA was precipitated onto 0.6 µm (average diameter) gold microprojectiles using a
water-soluble cationic lipid TransIT®-2020 (Mirus, WI, USA) as follows: 50 µl of gold
particles (water solution, 10 mg/ml) and 1 µl of TransIT-2020 water solution were added to the premixed DNA constructs (DNA and RNA amounts used in each experiment are shown in Supplemental Table S5), mixed gently and incubated on ice for 10 min. DNA-coated gold particles were then centrifuged at 10,000 rpm for 1 minute. The pellet was rinsed with 100 µl of absolute alcohol and resuspended by a brief sonication. Immediately after sonication, DNA-coated gold particles were loaded onto the center of a macrocarrier (10 µl each) and allowed to air dry. Immature maize embryos 8-10 days after pollination were bombarded using a PDS-1000/He Gun (Bio-Rad, CA, USA) with a rupture pressure of 425 PSI. Post-bombardment culture, selection, and plant regeneration were performed as previously described (Gordon-Kamm et al., 2002). Agrobacterium-mediated transformation followed the detailed protocol previously described in Gao et al., 2010. Regenerated plantlets were moved to soil where they were sampled and grown to maturity in greenhouse conditions.

**PCR, qPCR Analysis and Amplicon Sequencing**

For PCR analysis, DNA was extracted from a small amount (0.5 cm in diameter) of callus tissue or two leaf punches as described in Gao et al., 2010. PCR was performed using REDExtract-N-Amp™ PCR ReadyMix (Sigma, USA) or Phusion® High Fidelity PCR Master Mix (NEB, MA, USA) according to the manufacturer’s recommendations. PCR amplified fragments were cloned using the pCR2.1-TOPO cloning vector (Life Technologies Inc., USA). The primers used in PCR and qPCR reactions are shown in Supplemental Table 3. qPCR analysis was performed as in (Wu et al., 2014). As mutations within a population of individual callus sectors may differ from simple (single base pair insertions or deletions) to complex (deletions and additions) sequence modifications, target site copy-number, as determined by qPCR analysis, was used to as a semi-qualitative method to determine mutation frequencies. In general, mutation frequency was measured based on target site allele copy-number: both alleles are defined to be intact if target site copy-number is 2, and the event would be called wild type. Events for which qPCR revealed target site copy-number reduced to one would be scored as one mutated allele and one wild type allele. Finally, if both
alleles of the target site are modified, the copy-number is zero, and event would be called “null”.

For deep sequencing analysis, 60-90 Hi-II maize immature embryos were co-bombarded with the maize optimized Cas9 endonuclease and gRNA expression cassettes or I-CreI modified endonuclease vectors and screenable marker UBI:MoPAT-DsRED, and cell division promoting genes, UBI:ZmODP2 and IN2:ZmWUS2. Seven days after bombardment, 20-30 of the most uniformly transformed embryos (based on transient expression of DsRED) from each treatment were pooled and total genomic DNA was extracted. The DNA region surrounding the intended target site was amplified by PCR using Phusion® High Fidelity PCR Master Mix (NEB, MA, USA) adding the amplicon-specific barcodes and Illumina sequencing through two rounds of PCR. The primers used in the primary PCR reaction are shown in Supplemental Table 3 and the primers used in the secondary PCR reaction were 5’-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACAG-3’ (forward) and 5’-CAAGCAGAAGACGGCATA-3’ (reverse). The resulting PCR amplifications were purified with a Qiagen PCR purification spin column (Qiagen, Germany), concentration measured with a Hoechst dye-based fluorometric assay, combined in an equimolar ratio, and single read 100 nucleotide-length amplicon sequencing was performed on Illumina’s MiSeq Personal Sequencer with a 30-40% (v/v) spike of PhiX control v3 (Illumina, FC-110-3001) to off-set sequence bias. Only those reads with a ≥1 nucleotide INDEL arising within the 10 nt window centered over the expected site of cleavage and not found in the negative controls were classified as NHEJ mutations. NHEJ mutant reads with the same mutation were counted and collapsed into a single read and the top 10 most prevalent mutations were visually confirmed as arising within the expected site of cleavage. The total numbers of visually confirmed mutations were then used to calculate the percentage of mutant reads based on the total number of reads of an appropriate length containing a perfect match to the barcode and forward primer.

DNA Isolation and Blot Hybridization Analysis

For DNA isolation and blot hybridization analyses, maize leaf tissue (approximately 2.5 g fresh weight), from transformed and control plants was freeze-dried and ground. Total
genomic DNA was extracted as described in Cigan et al., 2001. Five micrograms of DNA from each sample were digested overnight with \textit{Hind}III and the DNA fragments were separated on a 0.8% agarose gel run at 23 V overnight. DNA was transferred to a nylon membrane (Roche, Germany) using standard protocol (Sambrook et al., 1989), and fixed to the membrane by UV irradiation of 150 mJ in a Gene Linker (Bio-Rad, CA, USA). Membranes were pre-hybridized for 1 hour in DIG Easy Hyb (Roche, Germany) at 50°C. Hybridization probes LIG-5' (459 bp), LIG-3' (456 bp), and Mo-PAT (446 bp) were labeled by PCR with primers 5'-AGCTTTATCCATCCATCGC-3' (forward) and 5'-AATTCTCGTACGTACAGCACGCG-3' (reverse), 5'-ACCCCTACAGCCACTTAGTCTCCG-3' (forward) and GCCAGAGAGATCGAGATCGATGGA-3' (reverse), and 5'-GTGTGCACATCGTAACCACT-3' (forward) and 5'-TCGAAGTCGCGCTGCCAGAA-3' (reverse), respectively, using the DIG Probe PCR Synthesis Kit (Roche, Germany) according to manufacturer's recommendations. Hybridization was performed overnight at 50°C. Membranes were washed twice in 2× SSC/0.1% SDS at room temperature, followed by three washes in 0.1× SSC/0.1% SDS for HR probe or 0.05× SSC/0.05% SDS for Mo-PAT probe at 63°C. Membranes were prepared for detection by washing, at room temperature, in 1× Wash Buffer (DIG Wash and Block Buffer Set) for 10 minutes, then 1× Blocking Buffer (DIG Wash and Block Buffer Set) for 60 minutes, followed by 1× Antibody solution (Anti-Digoxigenin-AP Fab Fragments (Roche, Germany) in 1× Blocking Buffer) for 60 minutes, and finally three times in 1× Wash Buffer for 8 minutes per wash. Detection was done by placing membranes in 1× Detection Buffer (DIG Wash and Block Buffer Set) for 5 minutes, removing to a clean tray, and pipetting CDP-Star, ready-to-use (Roche, Germany) onto the membranes then exposed to X-ray film (Kodak, USA).

\textbf{ACKNOWLEDGMENTS}

Authors are thankful to Brian Lenderts, Meizhu Yang, Shan Zhao, Wally Marsh, Susan Wagner, and Min Zeng for technical and maize transformation support, DuPont Pioneer Controlled Environment group for managing plants, and Genomics group for DNA sequencing. The authors extend special thanks to Brooke Peterson-Burch for
bioinformatics support, Wayne Crismani, Jeff Sander, Doane Chilcoat, Philippe Horvath
and the anonymous *Plant Physiology* reviewers for thoughtful comments and
suggestions.
SUPPLEMENTAL MATERIAL

Table S1. Maize genomic sites targeted by Cas9-gRNA system.

Table S2. T0 plant analysis by qPCR for mutations at maize target loci of 30 T0 plants produced in the multiplexed Cas9-gRNA experiments.

Table S3. List of primers used in the study.

Table S4. Edited ALS2 events recovered using different templates and selection.

Table S5. DNA and RNA amounts used in bombardment experiments (single shot).

Figure S1. Ten of the most prevalent mutations generated by either the Cas9-gRNA system or the LIG3:4 homing endonuclease identified by deep sequencing of the PCR amplicons across the corresponding target site.

Figure S2. Agro vector for stable integration of the ZmUBI:Cas9 into the maize genome.

Figure S3. Repair of the RF-FP gene via intramolecular homologous recombination.

Figure S4. Constructs used for HDR-mediated targeted gene insertion at the liguleless-1 region using Cas-gRNA system and LIG3:4 meganuclease.

Nucleotide sequence of maize optimized Cas9

Nucleotide sequence of LIG-CR3 gRNA expression cassette

Nucleotide sequence and description of the donor DNA in the integration experiment

Nucleotide sequence and description of the RF-FP repair cassette
Table 1. Percentage of mutant reads at 5 different target sites 7 days post-transformation.

<table>
<thead>
<tr>
<th>Target</th>
<th>DSB Reagents</th>
<th>Total Number of Reads</th>
<th>Number of Mutant Target Gene Reads</th>
<th>Percentage of Mutant Reads in Target Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIG (Chr. 2)</td>
<td>LIG3:4 meganuclease</td>
<td>616,536</td>
<td>1,211</td>
<td>0.20%</td>
</tr>
<tr>
<td></td>
<td>LIG-CR1 gRNA+Cas9</td>
<td>716,854</td>
<td>33,050</td>
<td>4.61%</td>
</tr>
<tr>
<td></td>
<td>LIG-CR2 gRNA+Cas9</td>
<td>711,047</td>
<td>16,675</td>
<td>2.35%</td>
</tr>
<tr>
<td></td>
<td>LIG-CR3 gRNA+Cas9</td>
<td>713,183</td>
<td>27,959</td>
<td>3.92%</td>
</tr>
<tr>
<td>MS26 (Chr. 1)</td>
<td>Ems26 meganuclease</td>
<td>512,784</td>
<td>642</td>
<td>0.13%</td>
</tr>
<tr>
<td></td>
<td>MS26-CR1 gRNA+Cas9</td>
<td>575,671</td>
<td>10,073</td>
<td>1.75%</td>
</tr>
<tr>
<td></td>
<td>MS26-CR2 gRNA+Cas9</td>
<td>543,856</td>
<td>16,930</td>
<td>3.11%</td>
</tr>
<tr>
<td></td>
<td>MS26-CR3 gRNA+Cas9</td>
<td>538,141</td>
<td>13,879</td>
<td>2.58%</td>
</tr>
<tr>
<td>MS45 (Chr. 9)</td>
<td>MS45-CR1 gRNA+Cas9</td>
<td>812,644</td>
<td>3,795</td>
<td>0.47%</td>
</tr>
<tr>
<td></td>
<td>MS45-CR2 gRNA+Cas9</td>
<td>785,183</td>
<td>14,704</td>
<td>1.87%</td>
</tr>
<tr>
<td></td>
<td>MS45-CR3 gRNA+Cas9</td>
<td>728,023</td>
<td>9,203</td>
<td>1.26%</td>
</tr>
<tr>
<td>ALS1 (Chr. 4) and ALS2 (Chr. 5)</td>
<td>ALS-CR1 gRNA+Cas9</td>
<td>434,452</td>
<td>9,669</td>
<td>2.23%</td>
</tr>
<tr>
<td></td>
<td>ALS-CR2 gRNA+Cas9</td>
<td>472,351</td>
<td>6,352</td>
<td>1.35%</td>
</tr>
<tr>
<td></td>
<td>ALS-CR3 gRNA+Cas9</td>
<td>497,786</td>
<td>8,535</td>
<td>1.72%</td>
</tr>
<tr>
<td>Controls</td>
<td>Cas9 only</td>
<td>640,063</td>
<td>1</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>LIG-CR1 gRNA only</td>
<td>646,774</td>
<td>1</td>
<td>0.00%</td>
</tr>
</tbody>
</table>
**Table 2.** Read counts and percentage of mutant reads at maize target sites in multiplexed Cas9-gRNA experiments 7 days post-transformation.

<table>
<thead>
<tr>
<th>Target Site</th>
<th>Co-Transformed gRNAs</th>
<th>Total Number of Reads</th>
<th>Number of Mutant Target Gene Reads</th>
<th>Percentage of Mutant Reads in Target Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIGCas-3</td>
<td>LIGCas-3</td>
<td>645,107</td>
<td>12,631</td>
<td>1.96%</td>
</tr>
<tr>
<td></td>
<td>LIGCas-3, MS26Cas-2</td>
<td>579,992</td>
<td>10,348</td>
<td>1.78%</td>
</tr>
<tr>
<td></td>
<td>LIGCas-3, MS26Cas-2, MS45Cas-2</td>
<td>648,901</td>
<td>12,094</td>
<td>1.86%</td>
</tr>
<tr>
<td>MS26Cas-2</td>
<td>MS26 Cas 2</td>
<td>699,154</td>
<td>17,247</td>
<td>2.47%</td>
</tr>
<tr>
<td></td>
<td>MS26Cas-2, LIGCas-3</td>
<td>717,158</td>
<td>10,256</td>
<td>1.43%</td>
</tr>
<tr>
<td></td>
<td>MS26Cas-2, MS45Cas-2</td>
<td>613,431</td>
<td>9,931</td>
<td>1.62%</td>
</tr>
<tr>
<td></td>
<td>MS26Cas-2, MS45Cas-2, LIGCas-3</td>
<td>471,890</td>
<td>7,311</td>
<td>1.55%</td>
</tr>
<tr>
<td>MS45Cas-2</td>
<td>MS45Cas-2</td>
<td>503,423</td>
<td>10,034</td>
<td>1.99%</td>
</tr>
<tr>
<td></td>
<td>MS45Cas-2, MS26Cas-2</td>
<td>480,178</td>
<td>8,008</td>
<td>1.67%</td>
</tr>
<tr>
<td></td>
<td>MS45Cas-2, MS26Cas-2, LIGCas-3</td>
<td>416,711</td>
<td>7,190</td>
<td>1.73%</td>
</tr>
</tbody>
</table>
Table 3. Mutation analysis at maize target sites in multiplexed experiments.

<table>
<thead>
<tr>
<th>Target Sites</th>
<th>T0 Plant</th>
<th>qPCR Results</th>
<th>LIG3/4 TS</th>
<th>MS26 TS</th>
<th>MS45 TS</th>
<th>Stable Integration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Allele 1</td>
<td>Allele 2</td>
<td>Allele 1</td>
<td>Allele 2</td>
</tr>
<tr>
<td>LIGCas-3/MS26Cas2</td>
<td>1</td>
<td>null/null</td>
<td>1bp del</td>
<td>2bp del+ 1 bp ins</td>
<td>1bp ins (A)</td>
<td>19bp del</td>
</tr>
<tr>
<td>MS26Cas-2/MS45Cas-2</td>
<td>2</td>
<td>null/null</td>
<td>1bp ins (T)</td>
<td>1bp del</td>
<td>1bp ins (A)</td>
<td>1bp ins (G)</td>
</tr>
<tr>
<td>MS26Cas-2/MS45Cas-2</td>
<td>3</td>
<td>het/null</td>
<td>-</td>
<td>-</td>
<td>1bp ins (C)</td>
<td>WT 2bp del</td>
</tr>
<tr>
<td>LIGCas-3/MS26Cas-2/MS45Cas-2</td>
<td>4</td>
<td>null/null/null</td>
<td>1bp ins (T)</td>
<td>large mut*</td>
<td>1bp ins (T)</td>
<td>1 bp del 15bp del</td>
</tr>
<tr>
<td>LIGCas-3/MS26Cas-2/MS45Cas-2</td>
<td>5</td>
<td>het/null/null</td>
<td>1bp ins (T)</td>
<td>Wild-type</td>
<td>1bp ins (A)</td>
<td>1bp ins (G)</td>
</tr>
</tbody>
</table>

Null indicates that both alleles are mutated; het indicates that one alleles is mutated and the other one is wild-type.

* Due to the size and nature of the mutation (likely the transformation vector insertion) it could not be resolved by the methods used in the analysis.
Table 4. Percentage of mutant reads at maize LIG and MS26 target sites produced in transient gRNA delivery into wild-type embryos and embryos with pre-integrated Cas9 under constitutive (UBI) or regulated (MDH) promoters.

<table>
<thead>
<tr>
<th>Target Site</th>
<th>Hi-II Embryos</th>
<th>Transformation</th>
<th>Percentage of Mutant Reads</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 days post-transformation</td>
</tr>
<tr>
<td>LIG</td>
<td>Wild-type</td>
<td>Cas9 (DNA) + gRNA (DNA)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cas9 (DNA) + gRNA (RNA)</td>
<td>-</td>
</tr>
<tr>
<td>UBI:Cas9</td>
<td>gRNA (DNA)</td>
<td>1.22%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>gRNA (RNA)</td>
<td>1.86%</td>
<td></td>
</tr>
<tr>
<td>MDH:Cas9 event 1</td>
<td>gRNA (DNA)</td>
<td>0.25%</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>gRNA (RNA)</td>
<td>0.12%</td>
<td>-</td>
</tr>
<tr>
<td>MDH:Cas9 event 2</td>
<td>gRNA (DNA)</td>
<td>0.57%</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>gRNA (RNA)</td>
<td>0.26%</td>
<td>-</td>
</tr>
<tr>
<td>MDH:Cas9 event 3</td>
<td>gRNA (DNA)</td>
<td>0.46%</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>gRNA (RNA)</td>
<td>0.35%</td>
<td>-</td>
</tr>
<tr>
<td>MS26</td>
<td>MDH:Cas9 event 2</td>
<td>gRNA (DNA)</td>
<td>0.58%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gRNA (RNA)</td>
<td>0.17%</td>
</tr>
</tbody>
</table>

* In these experiments, analysis was performed on embryos collected 7 days after transformation, while in the experiments with pre-integrated Cas9, analysis was done on embryos collected 2 days after transformation. The additional 5 days allows constitutively expressed Cas9 endonuclease to continue DNA cleavage, thus accounting for the difference in percentage of mutant reads in the DNA delivery experiments.
Table 5. Summary of gene integration at LIG target site.

<table>
<thead>
<tr>
<th>System / Experiment</th>
<th>Number of Events Analyzed</th>
<th>Mutation Frequency LIG target site</th>
<th>2 Junction PCR Positive Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIG3:4 Meganuclease + donor DNA Separately vectors Particle bombardment</td>
<td>288</td>
<td>6%</td>
<td>2 (0.7%)</td>
</tr>
<tr>
<td>LIG-CR3 gRNA + Cas9 + donor DNA 3 separate vectors Particle bombardment</td>
<td>480</td>
<td>83%</td>
<td>11 (2.5%)</td>
</tr>
<tr>
<td>LIG-CR3 gRNA + Cas9 + donor DNA All in a single vector Particle bombardment</td>
<td>336</td>
<td>86%</td>
<td>14 (4.1%)</td>
</tr>
<tr>
<td>LIG-CR3 gRNA + Cas9 + donor DNA Agrobacterium-mediated delivery</td>
<td>192</td>
<td>84%</td>
<td>0</td>
</tr>
</tbody>
</table>


transcript splicing, and promoter activity following transfer to protoplasts by electroporation. Plant Mol. Biol. 18: 675-689


Fauser F, Schiml S, Puchta H (2014) Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. The Plant Journal 79: 348-359


Puchta H, Dujon B, Hohn B (1996) Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. Proceedings of the National Academy of Sciences 93: 5055-5060


