Original Paper Open Access

Characterization of paired Cas9 nickases induced mutations in maize mesophyll protoplasts

Felix Wolter[1,3], Susanne Edelmann[2], Anan Kadri[1], Stefan Scholten[1,2]*

[1]Institute for Plant Breeding, Seed Science and Population Genetics, University of Hohenheim, 70599 Stuttgart, Germany
[2]Biocenter Klein Flottbek, University of Hamburg, 22609 Hamburg, Germany
[3]Current address: Botanisches Institut II, Molekularbiologie und Biotechnologie, KIT, 76131 Karlsruhe, Germany
*Corresponding author: E-mail: stefan.scholten@uni-hamburg.de; s.scholten@uni-hohenheim.de

Abstract

Targeted genome modifications are important for both fundamental and applied research. The CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats / CRISPR-associated protein 9) technology has been successfully used in various plant species with high efficiency. Approaches with paired Cas9 nickase enhance the specificity of the CRISPR/Cas9 system by using guide RNA pairs to create two staggered single strand breaks on complementary DNA strands. Here we used maize mesophyll protoplasts as a transient test system and demonstrated the mutagenic potential of Cas9 nickases. Although we found activity for all the three different guide RNA pairs tested, their efficiency varied considerably. Characterization of the modification events revealed a high ratio of large deletions as well as insertions of donor DNA fragments. By the use of the *maternally expressed in embryo 1 gene* (*mee1*) as model target sequence, we could demonstrate that transcriptionally inactive and methylated genomic loci are practical targets of Cas9 nickase. The high specificity of Cas9 nickase approaches might provide advantage for genome modifications of certain loci in the complex and highly repetitive maize genome.

Keywords: genome editing, sequence specific nucleases, protoplasts, transient transformation

Introduction

In plant breeding mutagenesis is traditionally achieved by treatment with chemicals like EMS or by irradiation like X-rays. Since conventional mutation breeding of this kind is inevitably non-specific as mutations happen at random positions and usually more changes than the desired one are introduced, subsequent laborious screening of the mutants for desired traits and segregation analyses are required (D'Halluin and Ruiter, 2013; Voytas, 2013; Hartung and Schiemann, 2014).

In recent years, more sophisticated and precise methods for engineering DNA were enabled by the development of sequence specific nucleases (SSNs), enabling the formation of double strand breaks (DSBs) at a defined specific sites of the target genome. The cells own DSB repair mechanisms, non-homologous end joining (NHEJ) and homologous recombination (HR), are then harnessed for genome editing purposes (Voytas, 2013). NHEJ is non-template directed, and involves direct re-ligation of the exposed DNA ends. This is an error prone process and frequently causes small insertions or deletions (indels), which may lead to gene knockout by frameshift mutations (Puchta, 2004). This is most widely exploited for gene function analyses in basic research (D'Halluin and Ruiter, 2013). Precise genome editing is possible through HR-mediated repair of the DSB. This repair mechanism requires the availability of a template with regions of homology to the sequence surrounding the DSB (Puchta, 2004). Naturally, this is the sister chromatid or a homologous chromosome. However, in a technique known as gene targeting, a donor DNA molecule containing a desired change surrounded by sequences homologous to the desired insertion region serves as artificial repair template for the cell. The parts between the homologous sequences are then exchanged by homologous recombination, leading to stable integration of the desired sequence at the desired locus. This process can be exploited for both insertion of additional sequences or precise alterations of the existing sequence (Puchta, 2004; Baltes and Voytas, 2015). The advent of efficient SSNs was the breakthrough for gene targeting in plants (Voytas, 2013). Inducing more multiple DSB simultaneously enables even more sophisticated genome editing, e.g. deletion of undesired or inhibitory DNA (Zhou et al, 2014). Sequence inversions or sequence exchanges between chromosomes should also be possible (Puchta and Fauser, 2013).

The first customizable SSNs developed were Zinc Finger Nucleases (ZFN) (Kim et al, 1996). However, ZFN exhibit several drawbacks, first among them that their design is complicated and time consuming and their DNA binding properties are unpredictable and need experimental verification (Puchta and Fauser, 2013; Hartung and Schiemann, 2014; Sprink et al, 2015). More recently, TALENs (transcription activa-

 RECEIVED 12/30/2016

tion like effector nuclease) were developed (Boch et al, 2009), based on TAL effector proteins naturally secreted by pathogenic bacteria of the genus Xanthomonas, providing more flexibility than ZFN. However, due to the highly repetitive sequences in the DNA-binding domain, their construction is quite laborious as well (Fauser et al, 2014).

Most recently, the CRISPR/Cas system emerged as efficient and very easily programmable SSNs. It is based on an adaptive immune system found in prokaryotes. CRISPR (clustered regularly interspaced short palindromic repeats) arrays are segments of DNA containing repetitions in alternation with «spacer DNA». The spacers are homologous to viral and plasmid DNA. They provide protection against invading DNA when combined with Cas (CRISPR associated) genes. CRISPR/Cas systems are highly diverse, the newest classification (Makarova et al, 2015) divides CRISPR/Cas systems into 2 classes, 5 types and 16 subtypes. The typeII system is most commonly used for biotechnological purposes. In its natural form, invading DNA can be specifically cleaved by the nuclease Cas9. Cas9 is guided by a short crRNA with 20 bases conferring homology to the target sequence. A tracrRNA stabilizes the complex and activates Cas9. An NGG PAM (protospacer adjacent motive) is required for efficient cleavage. Cas9 exhibits two different cleavage domains: A RuvC-like nuclease domain near the N terminus that cleaves the non-target-strand and an HNH (McrA-like) nuclease domain in the center of Cas9 that cleaves the target strand (Jinek et al, 2012; Kennedy and Cullen, 2015). The properties of the crRNA (specificity) and the tracrRNA (structural stability) can be combined in a single chimeric guide RNA (gRNA). This is achieved by fusing the 3'end of the crRNA to the 5'end of the tracrRNA with a GAAA tetraloop. At the same time, all regions of the crRNA and tracrRNA not required for guiding Cas9-mediated DNA cleavage can be truncated without losing efficiency (Jinek et al, 2012). Accordingly, in order to use the typeII CRISPR/Cas system for genome editing, all that needs to be done is to transform target cells with vectors that lead to expression of Cas9 and a specific gRNA targeting the gene of interest.

The CRISPR/Cas9 system has the great advantage of RNA-based sequence specificity, enabling a very fast and simple design process. Only 20 nucleotides of the gRNA sequence have to be adapted, whereas the Cas9 protein does not require any re-engineering (Puchta and Fauser, 2013). However, in some cases severe off-target issues have been reported for the CRISPR/Cas9 system: Up to five mismatches are tolerated in the gRNA sequence (Fu et al, 2013; Shen, 2014), which translates into a substantial amount of potential off-target sites in most genomes. (Mali et al, 2013) report that mismatches near the 5' end of the gRNA target sequence are more easily tolerated, whereas much higher sensitivity is observed for mismatches in the 8-10 bases near the 3' end.

To counteract the off-target issue, specificity can be enhanced by the use of a double nickase (Mali et al, 2013; Ran et al, 2013; Cho et al, 2014). In this approach, two closely binding gRNAs are used, combined with a mutant version of Cas9, which lacks activity in one of the two DNA cleavage domains. A Cas9 carrying the mutation D10A is used, where a catalytic amino acid in the RuvC-like nuclease domain is mutated (Asp^{10}→Ala^{10}). This disables cleavage of the RuvC-like nuclease domain, converting Cas9 activity from nuclease (catalyzing double strand breaks) to nickase (catalyzing single strand breaks) (Jinek et al, 2012). Isolated nicks, which might occur at off-target sites are repaired with very high fidelity by the base-excision repair pathway, whereas two very closely located nicks on opposite DNA strands at the target region can lead to DSBs and subsequent NHEJ (Liu et al, 2007; Schiml et al, 2014). This way, off-target activity can be decreased 50-1,500 fold at maintained on target efficiency as shown for human cell lines (Ran et al, 2013). The properties «offset» (distance between gRNA targets) as well as the kind of overhang produced and gRNA orientation influence efficiency of double nickases. Several studies are in accordance that highest efficiencies are achieved with short offsets of around -4 to +20 bp and for gRNA pairs generating 5' overhangs and oriented «tail to tail» with their PAM sequences facing outward (Mali et al, 2013; Ran et al, 2013; Cho et al, 2014; Shen et al, 2014).

The efficiency of different SSN constructs varies widely (see e.g. Shan et al, 2013; Li et al, 2013; Liang et al, 2014; Xie et al, 2014). Since the process of creating transgenic plants in crops, such as maize, is time consuming and costly (Ishida et al, 2007), a system, which enables simple, reliable, and fast analysis of the efficiency of SSNs before engaging in the process of transgenic plant production, should be very useful. Here we present an optimized test system for CRISPR/Cas9 nickase constructs in maize. It is based on transient expression in protoplasts and subsequent mutation analysis. We confirmed the functionality of paired CRISPR/Cas9 nickases in maize and characterized the induced mutation events. The methylation status of the *mee1* target sequences in mesophyll cells enabled us to conclude that methylated and transcriptionally inactive genomic loci are practical targets of Cas9 nickases.

Materials and Methods

Methylation analysis

Genomic DNA of mesophyll protoplasts was obtained from 5 independent protoplast isolations. PCR products of *Mee1* served as non-methylated control. For each sample, 350 ng gDNA (or PCR-Product, respectively) was used for bisulfite conversion. Bisulfite conversion was done using EZ DNA Methylation-Gold Kit (Zymo Research). PCR amplification of the bisulfite

treated DNA was performed using EPIK Amplification Kit (Bioline). Binding sites containing no CG and CHG were chosen and the primers were designed to bind to completely converted sequences. Only the coding strand was amplified. The resulting PCR products were cloned into pJET1.2/blunt (Thermo Scientific). Plasmid DNA was isolated from positively tested clones and sequenced using either the Value Read service from Eurofins Genomics or the EZ-seq V2.0 service from Macrogen Europe. The sequences were analyzed using Geneious version 8.05 (http://www.geneious.com; Kearse et al, 2012).

DNA constructs

The Cas9 nickase constructs were cloned with the plasmids B325pUC19-U6Os4, which contains a gRNA scaffold, followed by the rice U6 promoter and B349p7i-Ucas-U6Os, which codes for a Ubiquitin promoter driven codon-optimized version of type II Cas9 D10A from Streptococcus pyogenes, flanked by 2 nuclear localization signals and a rice U6 promoter followed by 2 BsmBI sites for insertion of the gRNA scaffolds and the second rice U6 promotor.

The gRNA pairs for the CRISPR/Cas9 nickase constructs were designed using the design tool from DNA2.0 (https://www.dna20.com). Inserts were amplified from B325pUC19-U6Os4 with primers containing the desired gRNA spacers as well as BsmBI sites and cloned into B349p7i-Ucas-U6Os via BsmBI, resulting in pBin-C34, pBin-C70, and pBin-C140 termed depending on the position of the first SSB. For comparison a TALEN construct (A834p7iU-35sT2AB) was received from DNA Cloning Service (Hamburg). All constructs are binary vectors with pVS1 backbone.

Since the plasmid size might affect protoplast transformation efficiency (Bart et al, 2006), the TALEN and the CRISPR/Cas9 constructs were transferred from the large binary vector into the small vector pBlue (2958 bp) via SfiI, resulting in pBlue-A834 with a size of 10,135 bp, pBlue-C34, pBlue-C70, and pBlue-C140 each with a size of 10,164 bp. For transformation optimization and plasmid size comparisons the GFP expression cassette of the 5.3 kb plasmid pMON30049 carrying a synthetic version of GFP (Pang et al, 1996), was transferred into pBlue-C34 via NotI, resulting in pBlue-C34-GFP that exhibits a size of 12.85kb. Plasmid DNA for protoplast transformation was isolated from 50ml cultures using ZymoPure Plasmid Midiprep Kit (Zymo Research) and adjusted to 1 µg μl^{-1} by ethanol precipitation.

Protoplast isolation and transformation

Maize protoplast transformation was performed by a modified protocol based on (Yoo et al, 2007; Shan et al, 2013) with line B73. Briefly, 10-16 days after planting, the middle part of the youngest but expanded leafs (second or third leaf) from 3 to 5 plants were harvested and transferred into a petri dish containing 15ml of enzyme solution (0.6 M mannitol, 10 mM MES pH 5.7, 1.5% w/v CellulaseR10, 0.75% w/v MacerozymeR10, 0.1% Pectolyase Y23, 10 mM $CaCl_2$, 0.1% v/v BSA) and cut in a right angle from the midrib to the leaf margin as fine as possible. The cut leaves were vacuum infiltrated at -500 mbar for 30 min in the dark at room temperature and subsequently incubated at 26°C with gentle shaking (35 rpm) for 4 to 6 hours.

The protoplasts were passed though a 45 µm sieve, collected by centrifugation at 100 g, and re-suspended in a total of 5 ml W5. The protoplasts were further purified on a «sucrose cushion»: the protoplast solution was layered on top of CPW16S (Banks and Evans, 1976) and centrifuged at 217 g for 8 min. The protoplast band was transferred into tubes containing W5, washed by centrifugation, re-suspended in W5, and counted. Total protoplast yields were usually in the range of 3-4 x 10^6.

After 30 min incubation on ice, the protoplasts were pelleted by 1 min centrifugation at 94 g. For transformation, the protoplasts were re-suspended to a density of 2,5 x 10^6 cells ml^{-1} in MMG (0.4 M mannitol, 15 mM $MgCl_2$, 4 mM MES pH 5.7). 200µl of this suspension (500,000 cells) were added to the DNA to be transformed in 2 ml Eppendorf tubes. For routine transformations, the DNA mixture was made up of 5 µg pMON30049 (as transformation control) and 40 µg SSN-construct (for approximately 13kb pBlue constructs). Next, 220 µl freshly prepared PEG solution (40% w/v PEG4000, 0.2 M mannitol, 0.1 M $CaCl_2$) was added, mixed gently but thoroughly and incubated for 15 min at room temperature in the dark. In order to wash, 850 µl W5 solution was added to each tube followed by centrifugation at 100 g for 1.5 min. The pellet was re-suspended in 1 ml W5 and the suspension transferred to a 35 mm petri dish. Incubation was performed at 26°C for 48 h.

Before sampling the protoplasts were analyzed using the microscope Zeiss Axiovert 135. The transformation rate was determined by dividing the number of cells expressing GFP by the total number of viable cells, genomic DNA was extracted using Quick gDNA MicroPrep Kit (Zymo Research).

Mutation detection and quantification

PCR amplification of the target region was performed with Phusion DNA Polymerase (Thermo Scientific). The PCR products were used for restriction assays without purification. The Surveyor assay (CELII) was performed using the kit «Surveyor mutation detection kit for standard gel electrophoresis» (Integrated DNA Technologies). Per reaction, 15 µl of unpurified PCR product were mixed with 1.5 µl 0.15 M $MgCl_2$, 1 µl Enhancer S and 1 µl Nuclease S. The duplex formation was performed following manufacturer's recommendations. The digestion reaction was incubated 45 min at 42°C. For the T7E1 assay 300 ng column-purified PCR-product was used for duplex formation in NEBuffer 2 (NEB). The cycling program for duplex formation was as follows: (95°C - 5 min; 85°C to 25°C in 5°C steps for 48s each with ramping of -2°C s^{-1})

After duplex formation, 0.5 µl T7E1 (NEB) was added to the reaction, followed by 30 min incubation at 37°C. To stop the reaction, 1.5 µl 0.25 M EDTA was added.
All assays were analyzed by agarose gel electrophoresis and quantified by capillary gel electrophoresis. For the latter, the reactions were purified using Agencourt Ampure XP Magnetic Beads (Beckman Coulter) and analyzed using High Sensitivity NGS Fragment Analysis Kit 1 - 6,000bp (Advanced Analytical).
For restriction assays a smear analysis from 400 -1,000 bp was performed and this value was used as uncleaved DNA amount for the calculation of the uncleaved fraction. The same smear analysis was performed with untransformed controls to quantify background. In case of the T7E1 assays, the cleaved bands were quantified by smear analysis from 150 - 350 bp. The obtained DNA amount was subtracted from the total DNA amount of cleaved and uncleaved bands to obtain the cleaved fraction. The same procedure was performed with untransformed controls to quantify background.

Sequencing of mutated fragments

Uncleaved fragments from restriction assays at around 480 bp were gel purified, cloned into pJET1.2/blunt and transformed into DH5α and sequenced. In order to obtain sequences of insertions, fragments larger than the main uncut band (size range of around 600 - 900 bp) were purified from the restriction assay of C34 transformations. In case of the TALEN samples, the gel purified DNA was concentrated by ethanol precipitation using glycogen as carrier. Sequences were analyzed using Geneious version 8.05.

Statistical analysis

Statistical analysis was performed in R (version 0.98.484). Student's t-tests were performed for restriction assays and T7E1 assays to test for significant differences between SSN transformations and GFP-only transformed controls. In case of restriction assays, the uncleaved fractions of 4 controls from BfoI-assays were tested against the uncleaved fraction of the SSN transformations. In case of T7E1 assays, the cleaved fractions of 4 controls were tested against the cleaved fraction of the transformations. The Shapiro-Wilk test was performed to test for normal distribution of the samples. In case of a significant result (null hypothesis of normally distributed sample is rejected), a log transformation was performed. This way normality could be achieved. Student's t-test further requires equal variances between the two tested samples. Thus, F-tests were performed between untransformed controls and samples from transformation experiments to confirm equality of variance.

Results

Effects of plasmid size and concentration on protoplast transformation efficiency

Since the plasmid size can affect protoplast transformation efficiency (Bart et al, 2006), we addressed the question to what degree the transformation efficiency is affected by construct size and concentration in our experimental system with maize mesophyll protoplasts. Typical test transformation results with expression of GFP in protoplasts are shown in Figure 1 A,B. Transformation experiments were performed with two plasmids in increasing molar amounts both containing the same GFP expression cassette but varying in size (Figure 1C). With comparable absolute DNA amounts, the larger 12.9 kb GFP-plasmid (pBlue-C34-GFP) showed generally lower transformation rates than the 5.3 kb GFP-plasmid (pMON30049). With 20 µg pMON30049 per 200 µl protoplasts suspension, the transformation rate reached a plateau, whereas DNA amounts higher than 20 µg of the larger plasmid still led to an increased transformation rate. An amount of 48 µg of the 12.9 kb GFP-plasmid showed a transformation rate comparable to 20 µg of the 5.3 kb plasmid. These results indicate that the molar concentration determines transformation efficiency. When equimolar amounts are compared, transformation rates did not differ significantly (Figure 1C, Student's t-test). Accordingly, using higher amounts of DNA compensate for larger plasmids in PEG-mediated maize mesophyll protoplast transformation.

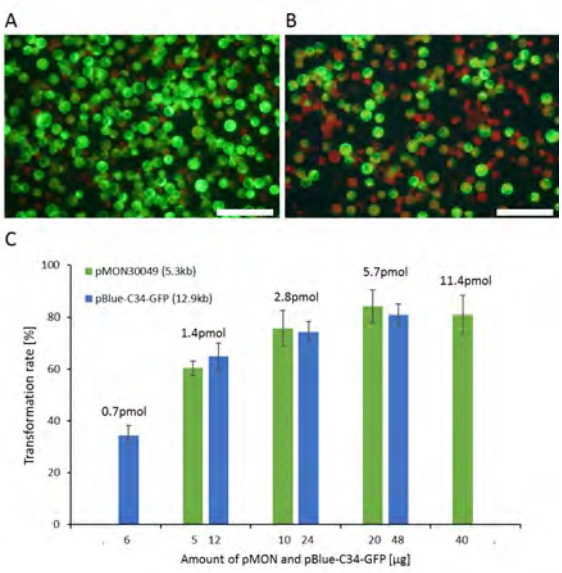

Figure 1 - Protoplast transformation rate as a function of plasmid amount and size. Exemplar fluorescence micrographs of protoplast 24 hours after transformation with 20 µg (**A**) or 5 µg (**B**) pMON30049 showing 88% or 57% transformation rate, respectively. Bars in (**A**) and (**B**) = 100 µm. Protoplast transformations with identical GFP expression cassette but two plasmids of varying size revealed the transformation rate to be dependent on the plasmid concentration and size (C). Numbers below the bars indicate the amount of transformed plasmid in µg. Bars of equimolar amounts of the two plasmid sizes are grouped together. The actual plasmid concentration during transformation is indicated above the bars.

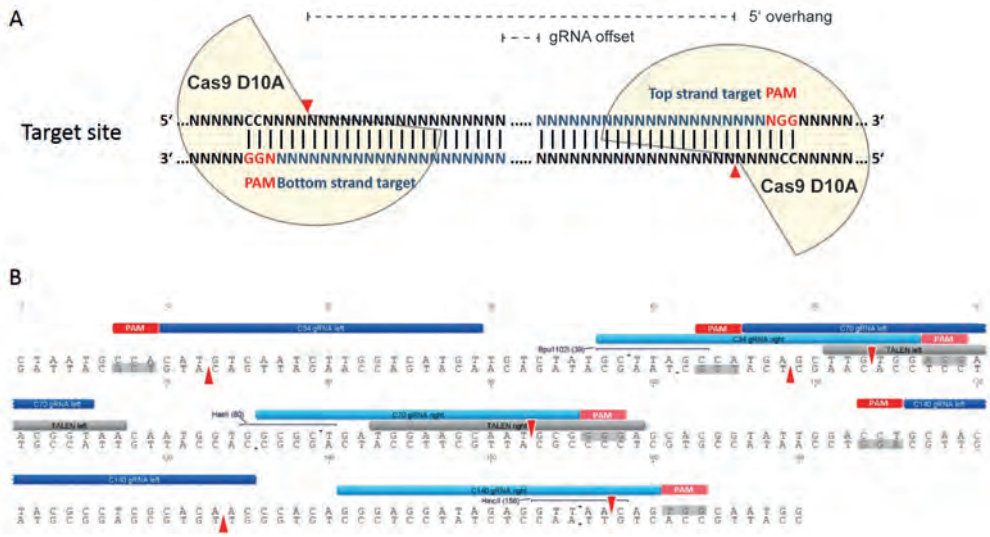

Figure 2 – Sequence specific nuclease (SSN) constructs. **A**) Schematic illustration of DNA double-stranded break induction by a pair of gRNAs guiding Cas9 D10A nickases. The D10A mutation causes Cas9 to cleave only the strand complementary to the gRNA; a pair of gRNA-Cas9 D10A complexes can nick both strands simultaneously. gRNA offset is the distance between the 5' ends of the guide sequence of a given gRNA pair. Positive offset implies the gRNA complementary to the top strand (sequence in blue on the left) to be 5' of the gRNA complementary to the bottom strand (sequence in blue on the right) and creates a 5' overhang. **B**) Partial sequence of the Mee1 ORF with DNA binding and cleavage sites of the SSNs used in this study. Red arrows indicate SSBs induced by the Cas9 nickases. TALEN induced cleavage occurs in the spacer between the two binding regions. The restriction sites used for the restriction assays of the respective SSNs are indicated.

Features of the SSN constructs

The paired gRNA design principles we applied for double nicking is illustrated in Figure 2A. The maize ubiquitin promoter drives expression of the Cas9 D10A nickase, which is flanked by eukaryotic NLS on both sides. The two gRNAs are both driven by the U6 promoter from rice. For PEG-mediated protoplast transformation the original constructions in binary vectors were subcloned resulting in a size of 10,164 bp for the functional plasmids each. All target sequences of the gRNAs in the *mee1* gene (Jahnke and Scholten, 2009) with the respective sites of single stranded nicks are shown in Figure 2B. In order to distinguish between the 3 CRISPR/Cas9 nickase constructs, they were termed according to the position of the first SSB within the target sequence as C34, C70, and C140. The offsets between their gRNAs are 7, 10 and 5 bp, creating 41, 44, and 39 bp 5' overhangs upon cleavage, respectively. Additionally, a TALEN construct targeting the same locus was used. Both TALEN proteins contain an N-terminal NLS and are both driven by the CaMV 35S promoter. Please refer to Figure 2B for the relation of SSN DSB or nick sites with the restriction enzymes (HaeII, HincII, and Bpu1102I) cleavage sites, which were used for the analyses of mutagenic events.

Quantification of mutagenesis efficiency

First, restriction assays were used to quantify the mutagenesis efficiency. In this kind of assay, the fraction of fragments that is resistant to cleavage by the restriction enzyme indicates the mutagenesis efficiency (Figure 3A). Mutagenesis activity was clearly detectable for all 3 CRISPR/Cas9 nickase constructs. Quantification of the fragments by capillary gel electrophoresis showed highly significant differences to the background with Student's t-test p-values of 0.0024 (C70), 0.00016 (C34), and 0.0020 (C140). Background subtraction revealed mutagenesis efficiencies of 1.7%, 5.4%, and 2.1% for C70, C34, and C140, respectively. The restriction assays revealed no detectable mutagenesis events for the TALEN construct (Figure 3B).

Next, the mutagenesis efficiency was quantified by mismatch cleavage assays. Here, mutagenesis efficiency is estimated by the cleaved fragment fraction, which is determined by dividing the DNA amount of the 2 cleaved fragments by the amount of all 3 fragments. The same procedure was performed with untransformed controls to determine background signal, which was generally higher and showed larger variation than in the restriction assays. Two different enzymes are commonly used for mismatch cleavage assays: CELII (Surveyor) and T7 Endonuclease 1 (T7E1). First, the sensitivity of these 2 enzymes regarding the kinds of mutations obtained from paired nickases in protoplasts was tested. T7E1 showed a better signal to noise ratio (Figure 4A) and was used for subsequent efficiency determination of the SSNs. Cleaved bands were not sharp and were determined by smear analysis of fragments ranging from 150 to 350 bp. Only C34 and C140 showed a signal discernable from background with efficiencies of 4.2% and 2.3%, respectively, after background subtrac-

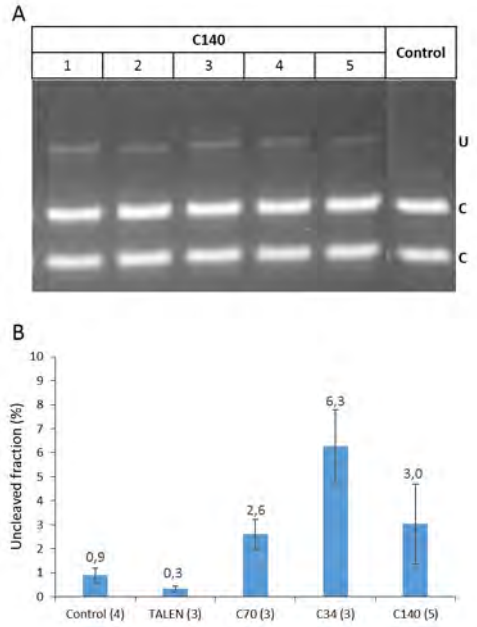

Figure 3 – Mutagenesis efficiencies determined by restriction assays. **A**) Agarose gel electrophoresis of exemplary restriction assays for gRNA pair C140 with HincII. Numbers below construct label indicate replicates. One control transformation without SSN plasmid is shown. «U» indicates uncut fragments. «C» indicates cut fragments. **B**) Mutagenesis efficiencies quantified by capillary electrophoresis of the restriction assays for all SSN constructs. Restrictions enzymes used for each contruct: BfoI: TALEN and C70, Bpu1102I: C34, HincII: C140. Number of replicate experiments are given in brackets. The mean of the uncleaved (mutated) fractions for each construct is given above the bars.

tion (Figure 4B). However, due to the large variation, the difference against background was not statistically significant (student's t-test p-values 0.149 and 0.115).

DNA methylation state at the mee1 target sequence in protoplasts

The SSN induced mutagenesis events occurred on a transcriptional inactive genomic sequence, since the *mee1* gene is not expressed in leafs (Jahnke and Scholten, 2009). To test whether this inactivity is associated with DNA methylation we analyzed the DNA methylation state of the locus at single nucleotide resolution by bisulfite sequencing. We used genomic DNA isolated from protoplasts to ensure that we analyze the target cells of our SSN test system. On the basis of 23 clones we found 95%, 94%, and 2% of the cytosines in the CG, CHG, and CHH context, respectively, methylated. Three clones of DNA directly isolated from leaf tissue confirmed these results. How many of the predominantly methylated cytosines in CG and CHG context each of the TALEN or CRISPR/Cas9 target sequences contain is summarized in Table 1.

Table 1 - Number of methylated cytosines in SSN binding sequences.

SSN	TALEN	C34	C70	C140
CG and CHG	2	0	4	3

Mutation pattern by paired nickases

Restriction assays after protoplast transformation of all three paired nickase constructs indicated that the mutation events created mainly either relatively long insertions or relatively long deletions. Fragments of near wild-type size (483 bp), related to InDels of one or a few nucleotides that impede cleavage, were hardly detectable (Figure 5). The size and intensity distribution of cleavage-resistant fragments indicated that deletions as well as insertions largely ranged in size and that the potential insertions prevail. In the example shown in Figure 5, the smear related to insertions with fragment sizes from approx. 500 - 700 bp contained roughly 4 times as much DNA as the smear related to deletions with fragment sizes of approx. 400 - 470 bp.

Uncleaved fragments from restriction assays were purified from agarose gels, cloned and sequenced. Mutated fragments could be recovered for the TALEN construct (Figure 6A). This confirms the presence of low mutagenesis activity despite negative assay re-

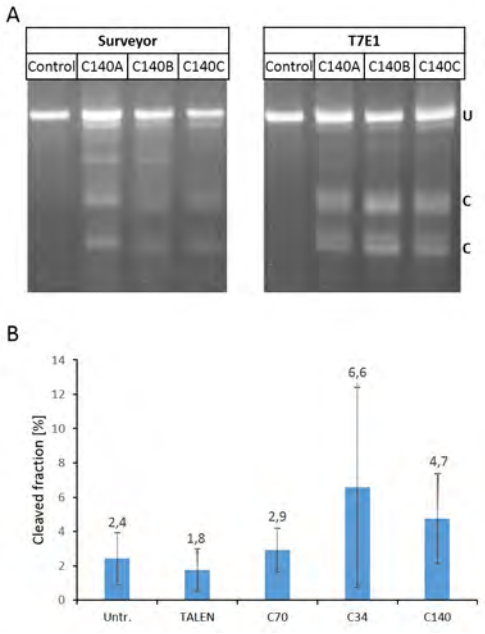

Figure 4 - Mutagenesis efficiencies determined by mismatch cleavage assays. **A**) Agarose gel electrophoresis of exemplary mismatch cleavage assays for gRNA pair C140 with Surveyor (left) and T7E1 (right) nucleases. Letters in brackets next to labels indicate replicates. For each nuclease one control transformation without SSN plasmid is shown. «U» indicates uncut fragments. «C» indicates cut fragments. **B**) Mutagenesis efficiencies quantified by capillary electrophoresis of the mismatch cleavage assays with T7E1 nuclease for all SSN constructs. Three replicate experiments were performed for all constructs. The mean of the cleaved (mutated) fractions for each construct is given above the bars.

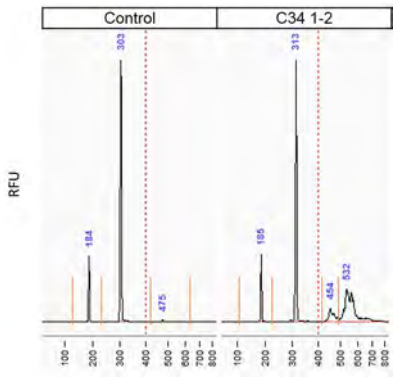

Figure 5 - Fragment analysis of restriction assays by capillary electrophoresis. Exemplary electropherogram of a capillary electrophoresis of a restriction assay for gRNA pair C34 with Bpu1102I. The sizes of the main fragments are indicated above the peaks in bp. RFU: relative fluorescence units.

sults (Figures 3B,4B). The deletions produced by CRISPR/Cas9 nickase constructs variegated in size, ranging from 5 to 72 bp (Figures 6B,D,E). Cloning and sequencing of fragments larger in size than the wild type fragment from transformations with C34 confirmed that the smear larger than wild type size detected by electrophoresis (Figure 5) was caused by insertions. The average insertion size was 165 bp and the insertions are exclusively combined with deletions (Figure 6D). Sequence analysis revealed that the insertions do not originate from the maize genome but that the transformed plasmid construct was exclusively used as template for the insertions. Interestingly, larger insertions proved to be a concatenation of multiple smaller insertions. This is the case for around half the insertion recovered. An alignment of the individual smaller insertions to the map of the CRISPR/Cas9 nickase construct is shown in Supplemental Figure 1. The templates for insertions seem to be randomly distributed across the plasmid. Even in case of concatenated insertions, their multiple templates are not located near to each other, but are randomly distributed. In many cases microhomologies can be found at the junctions, indicating a one-sided invasion of a single stranded overhang into the plasmid template.

Discussion

We established an efficient transient test system for CRISPR/Cas9 nickase constructs in maize. On one hand, such a system requires a biological platform for fast transient expression of the respective constructs. Plant protoplasts isolated from leaf mesophyll are especially suited to serve as an expression platform, because the physiological response from protoplasts is largely comparable to cells of whole plants (Yoo et al, 2007) and the cell autonomous response to DNA brakes facilitate testing the activity of SSNs. Expression of any SSN is easily and efficiently achieved by transient transformation of protoplasts using the PEG-calcium method to get SSN encoding plasmids into the cells (Yoo et al, 2007; Jiang et al, 2013). We found that within the plasmid size range tested, the most important factor for transformation frequency was the molar amount of plasmid. This indicates that in order to simplify the whole procedure of genome editing, large binary plasmids, which are required for Agrobacterium-mediated stable transformation of maize, could be used in the efficiency testing system without prior subcloning of the construct into smaller plasmids. However, high amounts of DNA are needed to compensate for large plasmid sizes and to obtain sufficiently high transformation efficiencies. Our protocols showed that protoplast can be handled with minimal laboratory equipment and tests for mutations are possible within 48 hours after transformation without laborious and time-consuming subculturing.

On the other hand, a SSN efficiency test system requires methods that enable detection and quantification of the mutations induced in genomic DNA (Voytas and Gao, 2014). For this purpose, we amplified the SSN target region from genomic DNA by PCR and subjected the amplicon either to restriction assays or mismatch cleavage assays. Restriction assays followed by agarose gel electrophoresis allowed in principal the detection of mutations events although the determination of the SSN efficiencies was more accurate by capillary gel electrophoresis. In case of the TALEN construct, the mutagenesis signal was not discernable from background. The two enzymes most commonly used for mismatch cleavage assays T7E1 and Surveyor (CELII) were tested for their potential to detect mutations induced by paired nickases in protoplasts. For these mutations, T7E1 showed a better signal to noise ratio than Surveyor, which is in accordance with earlier findings (Vouillot et al, 2015). Therefore, T7E1 assays were performed to determine SSN efficiencies. The comparison between restriction assays and T7E1 assays indicates that under the conditions of transient expression of CRISPR/Cas9 nickases in protoplasts, T7E1 assays were less sensitive than restriction assays due to higher background signal and large variance. We conclude that for information on the functionality and approximate efficiency of SSN constructs restriction assays and T7E1 assays are most appropriate.

Concerning the efficiencies of the different constructs, a comparison of the TALEN and the CRISPR/Cas9 nickase constructs was constricted essentially by the use of different promoters. In maize, the CaMV 35S promoter used for TALEN expression is much weaker than the Ubiquitin promoter, which was used to drive expression of the Cas9 nickase. Among the three nickase constructs we tested, the mutagenesis efficiencies ranged from 1.7% to 5.4% after background subtraction. Several studies explored mutations by transient expression of CRISPR/Cas9 nuclease constructs in plant protoplasts of various species

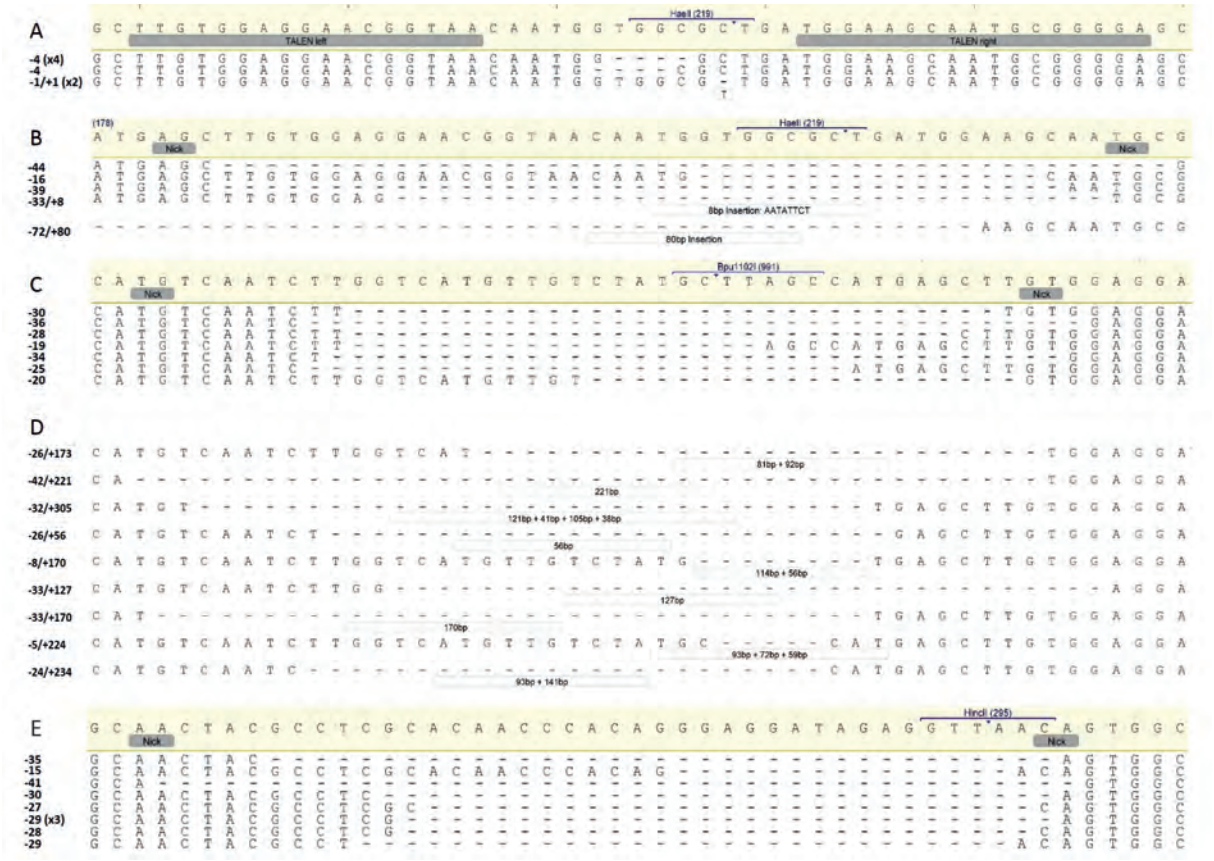

Figure 6 - Sequences of mutated fragments. The respective wild-type sequence of the mee1 locus is shown on top for every construct with binding sites for the TALEN or nick sites for Cas9 constructs indicated below the sequence by grey boxes. Within the mutated fragment deletions are indicated by dashes. Insertions are indicated by boxes below each fragment with numbers indicating insertion size and letters the sequence. Identifiable multiple insertions are given separately. Numbers to the left of the sequences indicate total number of nucleotides deleted/inserted. Fragments recovered multiple (n) times are marked by (xn). SSN Constructs: **A**) TALEN; **B**) C70; **C, D**) C34; **E**) C140; **D**) Fragments with large insertions of transformations with C34.

(Shan et al, 2013; Li et al, 2013; Xie and Yang, 2013; Liang et al, 2014; Gao et al, 2015). The efficiencies determined by restriction assays range from 1% to 38%. Specifically for maize, 16% and 19% mutation efficiency was reported for Cas9 nucleases (Liang et al, 2014).

One possibility for these contrasts might be that paired Cas9 nickases are less efficient in plants than the Cas9 nuclease, although the efficiencies of paired nickases reported for human embryonic kidney cells using T7E1 assays are comparable with the efficiencies of the nuclease. (Mali et al, 2013; Ran et al, 2013). To our knowledge no prior study tested efficiencies of paired nickases in protoplasts. For transgenic Arabidopsis plants with stable expression of paired nickase constructs 42.8% mutation efficiency was shown by deep sequencing of DNA extracted after 2 weeks from T1 plants (Fauser et al, 2014). However, in stable transgenic plants, the SSNs are expressed for a much longer period of time, which is not possible in protoplasts.

On the other hand, it might be possible that the kinds of mutations induced by paired nickases in protoplasts aggravate their quantitative analysis. Capillary electrophoresis and fragment analysis indicated that large deletions and especially large insertions constitute the highest proportion of mutations. The inherent size heterogeneity of these mutated fragments produces smear during electrophoresis, which in turn raises problems with quantitative analysis and background correction. More importantly, very likely our restriction assays did not detect all mutations induced by paired nickases. The sequences obtained by cloning and sequencing do only include mutations that affect the restriction side, very likely many smaller deletions did not span the restriction sites and were thus not detectable by the restriction assay. The problem also applies to the fragments with insertions as they are exclusively coupled with deletions. Concerning the T7E1 assay it is likely that large insertions affect it efficiency. Fluorescence graphs from capillary electrophoresis showed that the large insertions are only cleaved partially, because much of the smear larger than wild type size was still present after T7E1 digestion. It might well be that insertions beyond 150 bp impair heteroduplex formation. Taken

these considerations into account higher actual mutation efficiencies of our nickase constructs are likely.

The use of the *mee1* gene as target demonstrated that the SSN constructs we tested function on transcriptionally inactive sequences. The analysis of the DNA methylation status of the nickase target regions on the *mee1* locus confirmed the presence of methylated cytosines in CG and CHG contexts as detected earlier (Jahnke and Scholten, 2009). All nickase constructs tested in this study, except C34 exhibit methylated cytosines in their binding sequences (Table 1). Both C70 and C140 showed nuclease activity in spite of four and three methylated cytosines within the target sequences, respectively. This proves that the CRISPR/Cas9 system is able to bind and cleave methylated DNA. The question if DNA methylation quantitatively affects gRNA cleavage efficiency cannot be answered by this study. However, the fact that the efficiency of C140 was comparable to C34, which did not have methylated cytosine in its target sequence suggests that DNA methylation do not strongly affect CRISPR/Cas9 activity. C34 did show higher efficiency in the restriction assay than C140, but the restriction assay for C140 might be biased due to unfavorable location of the HincII restriction site available for C140.

The mutation types induced by paired nickases in protoplasts revealed relatively large insertions and deletions. Most insertions fall in the size range of 100 - 200 bp, most deletions are in the range of 15 - 45 bp. Capillary electrophoresis revealed that insertions up to 250 bp actually constitute the largest part of the mutations. This kind of mutation pattern has not been reported in earlier investigations of paired nickases so far. The main mutations induced by transiently expressed paired nickases in human embryonic kidney cells were also deletions (Ran et al, 2013). However, in contrast to plant cells, short deletions (< 5bp) made up a significant fraction of the mutations. In stable transgenic Arabidopsis plants paired nickases with an offset of 18 bp induced mainly deletions, ranging in size from 1 to more than 100 bp, but also large insertions up to more than 80 bp were found (Fauser et al, 2014). However, in their case of stably transformed transgenic plants, the templates for the insertions were provided by the sequences immediately upstream or downstream of the insertion site. None of the above-mentioned studies found a comparable high portion of large insertions to our findings. The high portion of large insertions is also unexpected as no repair donor with homology arms for HR was delivered. Thus the insertions must result from NHEJ. A synthesis-dependent strand annealing (SDSA) like mechanism leading to larger insertions by NHEJ has been described earlier (Salomon and Puchta, 1998). In this model, a single stranded 3'-end resulting from DSB formation and resection invades a double stranded donor molecule via binding to a short microhomology and formation of a D-loop structure. The invading strand is elongated via repair synthesis. The newly synthesized sequence can be religated to the other end of the DSB, resulting in the insertion of the newly synthesized sequence.

According to Xie et al (2014), the maize genome is prone to off-target effects when a single gRNA is used in combination with Cas9 nucleases. By bioinformatic analyses they showed that only 30% of the annotated transcript units in maize are targetable by specific gRNA spacers. Hence, the dual nickase approach, which provides enhanced specificity (Cho et al, 2014; Fauser et al, 2014), is probably advantageous for maize genome editing. Despite its potential, to our knowledge, this system has not yet been tested in maize. Thus, our study demonstrated activity of the paired nickase system in maize. Further we can conclude that transient expression of SSNs in maize mesophyll protoplasts followed by restriction and/or mismatch cleavage assays is a helpful tool for fast analysis of SSN function and efficiency. Positively tested constructs can subsequently be used for stable transformation of maize with higher confidence that the intended genome editing will be induced. Finally, we showed that paired nickases could efficiently induce insertions via NHEJ from templates without large homology arms when present in abundance. Together, our findings might promote the use of paired nickases in maize for high specific genome editing.

Acknowledgements

We thank Monika Stanke for excellent technical support.

References

Baltes NJ, Voytas DF, 2015. Enabling plant synthetic biology through genome engineering. Trends Biotechnol 33: 120-131

Banks MS, Evans PK, 1976. A comparison of the isolation and culture of mesophyll protoplasts from several nicotiana species and their hybrids. Plant Science Letters 7: 409-416

Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U, 2009. Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. Science 326: 1509-1512

Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S, Kim J-S, 2014. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Research 24: 132-141

D'Halluin K, Ruiter R, 2013. Directed genome engineering for genome optimization. Int J Dev Biol 57: 621-627

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

Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D,

Joung JK, Sander JD, 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol 31: 822-826

Gao J, Wang G, Ma S, Xie X, Wu X, Zhang X, Wu Y, Zhao P, Xia Q, 2015. CRISPR/Cas9-mediated targeted mutagenesis in *Nicotiana tabacum*. Plant Mol Biol 87: 99-110

Hartung F, Schiemann J, 2014. Precise plant breeding using new genome editing techniques: opportunities, safety and regulation in the EU. Plant J 78: 742-752

Ishida Y, Hiei Y, Komari T, 2007. Agrobacterium-mediated transformation of maize. Nat Protoc 2: 1614-1621

Jahnke S, Scholten S, 2009. Epigenetic resetting of a gene imprinted in plant embryos. Curr Biol 19: 1677-1681

Jiang F, Zhu J, Liu H-L, 2013. Protoplasts: a useful research system for plant cell biology, especially dedifferentiation. Protoplasma 250: 1231-1238

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E, 2012. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science 337: 816-821

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al, 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28: 1647-1649

Kennedy EM, Cullen BR, 2015. Bacterial CRISPR/Cas DNA endonucleases: A revolutionary technology that could dramatically impact viral research and treatment. Virology 479-480: 213-220

Kim YG, Cha J, Chandrasegaran S, 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci USA 93: 1156-1160

Li J-F, Norville JE, Aach J, McCormack M, Zhang D, Bush J, Church GM, Sheen J, 2013. Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. Nat Biotechnol 31: 688-691

Liang Z, Zhang K, Chen K, Gao C, 2014. Targeted Mutagenesis in Zea mays Using TALENs and the CRISPR/Cas System. J Genet Genomics 41: 63-68

Liu Y, Prasad R, Beard WA, Kedar PS, Hou EW, Shock DD, Wilson SH, 2007. Coordination of Steps in Single-nucleotide Base Excision Repair Mediated by Apurinic/Apyrimidinic Endonuclease 1 and DNA Polymerase beta. Journal of Biological Chemistry 282: 13532-13541

Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, Barrangou R, Brouns SJJ, Charpentier E, Haft DH, et al, 2015. An updated evolutionary classification of CRISPR-Cas systems. Nat Rev Microbiol 13: 722-736

Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, Yang L, Church GM, 2013. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol 31: 833-838

Pang SZ, DeBoer DL, Wan Y, Ye G, Layton JG, Neher MK, Armstrong CL, Fry JE, Hinchee MA, Fromm ME, 1996. An improved green fluorescent protein gene as a vital marker in plants. Plant Physiology 112: 893-900

Puchta H, Fauser F, 2013. Gene targeting in plants: 25 years later. Int J Dev Biol 57: 629-637

Puchta H, 2004. The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. J Exp Bot 56: 1-14

Ran FA, Hsu PD, Lin C-Y, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, Zhang F, 2013. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell 154: 1380-1389

Salomon S, Puchta H, 1998. Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. Embo J 17: 6086-6095

Schiml S, Fauser F, Puchta H, 2014. The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. Plant J 80: 1139-1150

Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu J-L, Gao C, 2013. Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotechnol 31: 686-688

Shen B, Zhang W, Zhang J, Zhou J, Wang J, Chen L, Wang L, Hodgkins A, Iyer V, Huang X, Skarnes WC, 2014. Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. Nat Meth 11: 399-402

Sprink T, Metje J, Hartung F, 2015. Plant genome editing by novel tools: TALEN and other sequence specific nucleases. Current Opinion in Biotechnology 32: 47-53

Vouillot L, Thelie A, Pollet N, 2015. Comparison of T7E1 and Surveyor Mismatch Cleavage Assays to Detect Mutations Triggered by Engineered Nucleases. G3 (Bethesda). 5: 407-415

Voytas DF, Gao C, 2014. Precision genome engineering and agriculture: opportunities and regulatory challenges. PLoS Biol 12: e1001877

Voytas DF, 2013. Plant genome engineering with sequence-specific nucleases. Annu Rev Plant Biol 64: 327-350

Xie K, Yang Y, 2013. RNA-guided genome editing in plants using a CRISPR-Cas system. Mol Plant 6: 1975-1983

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

Yoo S-D, Cho Y-H, Sheen J, 2007. Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. Nat Protoc 2: 1565-1572

Zhou H, Liu B, Weeks DP, Spalding MH, Yang B, 2014. Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. Nucleic Acids Research 42: 10903-10914