HAL
open science

Single and multiple gene knockouts by CRISPR–Cas9 in maize

Nicolas M Doll, Laurine Gilles, Marie-France Gerentes, Christelle Richard, Jérémy Just, Yannick Fierlej, Virginia Borrelli, Ghislaine Gendrot, Gwyneth Ingram, Peter Rogowsky, et al.

To cite this version:

Nicolas M Doll, Laurine Gilles, Marie-France Gerentes, Christelle Richard, Jérémy Just, et al.. Single and multiple gene knockouts by CRISPR–Cas9 in maize. Plant Cell Reports, 2019, 38 (4), pp.487-501. 10.1007/s00299-019-02378-1 . hal-02349639

HAL Id: hal-02349639
https://hal.science/hal-02349639v1

Submitted on 1 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Single and multiple gene knockout by CRISPR-Cas9 in maize

Nicolas M. Doll[1], Laurine M. Gilles[1,2], Marie-France Gérentes[1], Christelle Richard[1], Jeremy Just[1], Yannick Fierlej[1,3], Virginia M. G. Borrelli[4], Ghislaine Gendrot[1], Gwyneth C. Ingram[1], Peter M. Rogowsky[1] and Thomas Widiez[1]*

1 Laboratoire Reproduction et Développement des Plantes, Univ Lyon, ENS de Lyon, UCB Lyon 1, CNRS, INRA, F-69342, Lyon, France

2 Limagrain Europe SAS, Research Centre, F-63720 Chappes, France

3 MAS Seeds, Route de Saint-Sever, F-40280 Haut-Mauco, France

4 Department of Sustainable Crop Production, Università Cattolica del Sacro Cuore, Piacenza, Italy

* Corresponding author: Thomas Widiez (thomas.widiez@ens-lyon.fr).

- PMID: **30684023**
- DOI: 10.1007/s00299-019-02378-1

Plant Cell Reports
https://doi.org/10.1007/s00299-019-02378-1

ORIGINAL ARTICLE

Single and multiple gene knockouts by CRISPR–Cas9 in maize

Nicolas M. Doll[1] · Laurine M. Gilles[1,2] · Marie-France Gérentes[1] · Christelle Richard[1] · Jeremy Just[1] · Yannick Fierlej[1,3] · Virginia M. G. Borrelli[4] · Ghislaine Gendrot[1] · Gwyneth C. Ingram[1] · Peter M. Rogowsky[1] · Thomas Widiez[1]

Received: 31 August 2018 / Accepted: 7 January 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

KEY MESSAGE

The analysis of 93 mutant alleles in 18 genes demonstrated that CRISPR-Cas9 is a robust tool for targeted mutagenesis in maize, permitting efficient generation of single and multiple knockouts.

ABSTRACT

CRISPR-Cas9 technology is a simple and efficient tool for targeted mutagenesis of the genome. It has been implemented in many plant species, including crops such as maize. Here we report single and multiple gene mutagenesis via stably transformed maize plants. Two different CRISPR-Cas9 vectors were used allowing the expression of multiple guide RNAs and different strategies to knock out either independent or paralogous genes. A total of 12 plasmids, representing 28 different single guide RNAs (sgRNAs), were generated in order to target 20 genes. For 18 of these genes, at least one mutant allele was obtained, while two genes were recalcitrant to sequence editing. 19% (16/83) of mutant plants showed biallelic mutations. Small insertions or deletions of less than 10 nucleotides were most frequently observed, regardless of whether the gene was targeted by one or more sgRNAs. Deletions of defined regions located between the target sites of two guide RNAs were also reported although the exact deletion size was variable. Double and triple mutants were created in a single step, which is especially valuable for functional analysis of genes with strong genetic linkage. Off target effects were theoretically limited due to rigorous sgRNA design and random experimental checks at three potential off target sites did not reveal any editing. Sanger chromatograms allowed to unambiguously class the primary transformants; the majority (85%) were fully edited plants transmitting systematically all detected mutations to the next generation, generally following Mendelian segregation.

KEYWORDS (4 to 6 keywords)

CRISPR, gene editing, maize, SDN1, mutagenesis, *Zea mays*

Abbreviations

bp: Base pairs
Cas9: CRISPR associated protein 9
CRISPR: Clustered regularly interspaced short palindromic repeats
DSB: Double strand break
EMS: Ethyl methanesulfonate
ESR: Embryo surrounding region
HR: Homologous recombination
MMEJ: Microhomology mediated end joining
NHEJ: Non homologous end joining
PAM: Protospacer adjacent motif
SDN1: Site directed nuclease 1
sgRNA: Single guide RNA
shRNA: Short hairpin RNA, also referring to scaffold RNA
T-DNA: Transfer DNA

Introduction

Maize has a dual role, being a major crop species and a model species in genetics. Genome edited *waxy* maize characterized by modified starch composed entirely of amylopectin was one of the first crops edited using CRISPR-Cas9 technology that obtained clearance to be cultivated and sold without GM-type oversight by the US Department of Agriculture (Waltz 2016). This example illustrates the intense interest in the potential of CRISPR-Cas9 technology for both applied and fundamental research. The starch industry has appreciated *waxy* maize for decades, because the absence of amylose makes starch easier to process. Although the *waxy* trait is not novel, CRISPR-Cas9 technology allowed the direct creation of *waxy* deletions in elite lines over one or two generations, avoiding time consuming backcrosses and genetic drag experienced with conventional introgression (Cigan et al. 2017).

In fundamental research, understanding the contribution of genes to phenotypic traits in maize has been a challenge for many decades. By comparing a standard (wild-type) to a mutant, the contribution of a genetic sequence to a biological process can be assessed. Although a large number of natural maize mutants exist, increasing their diversity through mutagenesis has been a long term goal (Candela and Hake 2008). For example, the original *waxy* mutation discovered in 1909 (Collins 1909) has been joined over time by hundreds of additional alleles reflecting emerging mutagenesis tools. In the sixties, mutagenesis induced by the chemical agent ethyl methanesulfonate (EMS) was popular and allowed the generation of an allelic series at the *Waxy* locus with different levels of residual amylose (Briggs et al. 1965). A few years later irradiation mutagenesis, which preferentially creates deletions rather than point mutations helped to generate true loss-of-function mutants that are more informative for functional genetics (Amano 1968). With the advent of molecular biology, transposon mutagenesis was developed, since transposon insertions could be easily localized in the genome. The cloning of the *Waxy* gene by transposon tagging was a prime example for the success of this strategy (Shure et al. 1983). Lately, over the past few years CRISPR-Cas9 technology has emerged as an appreciated alternative to sequenced indexed mutant collections (Settles et al. 2007; Vollbrecht et al. 2010), mainly because these collections do not saturate the maize genome, and because CRISPR-Cas9 technology can be targeted.

In contrast to random mutagenesis tools, which require the molecular screening of large mutagenized populations to find a mutation in a given gene, targeted mutagenesis of a gene gives ready access to a specific mutant and its phenotype, but has been a major challenge. It has been made possible by the development of techniques inducing double strand breaks (DSB) of genomic DNA at a predetermined site (Puchta and Fauser 2014). DSB are then repaired by one of several cellular repair mechanisms that can be non-conservative and can therefore lead to mutations at the desired location. The most frequent repair process in plants is Non Homologous End Joining (NHEJ), during which a DNA ligase joins the damaged strands, and which can be classified as either classical NHEJ or alternative NHEJ, also known as Microhomology Mediated End Joining (MMEJ) (Lieber 2010). Classical NHEJ primarily induces the insertion or deletion of a low number of nucleotides, whereas MMEJ generally leads to larger deletions (McVey and Lee 2008; Puchta and Fauser 2014). DSB can also be repaired by homologous recombination (HR), which can also be classified into two classes: conservative and non-conservative. Non-conservative HR, called single strand annealing, occurs if a repeated sequence of more than 30 nucleotides is present upstream and downstream of the DSB (McVey and Lee 2008). The presence of these repeated sequences renders single strand annealing very efficient for mutagenesis, as it repairs up to 1/3 of the DSB and can generate large deletions (Siebert and Puchta 2002; Steinert et al. 2016).

Conservative HR is of particular interest because it can be used for the replacement or the insertion of a sequence of interest, present on an extra chromosome, at the desired genomic locus. However, although the presence of DSB increases the efficiency of conservative HR, this remains two orders of magnitude lower than that of NHEJ (Steinert et al. 2016).

Inducing a DSB at a predetermined site in the genome requires both the recognition of the target sequence and the cleavage of the DNA, hitherto achieved using endonucleases. Several technologies have been developed to direct endonucleases to sequences of interest, either by engineering the DNA binding domains of naturally occurring meganucleases (Choulika et al. 1994) or by linking modular DNA-binding domains such as zinc finger (Bibikova et al. 2003; Porteus and Carroll 2005; Shukla et al. 2009) or Transcription Activator-Like Effector (TALE) domains (Christian et al. 2010) to endonuclease domains such as FokI. All three technologies have been successfully implemented in maize (Bibikova et al. 2003; Porteus and Carroll 2005; Shukla et al. 2009; Gao et al. 2010; Liang et al. 2014; Char et al. 2015). In 2013, the adaptation of the bacterial immune system CRISPR-Cas9 of *Streptococcus pyogenes* offered a novel type of technology in which the recognition of the DNA was not due to a protein domain but to a short guide RNA (sgRNA) that forms an active complex with the Cas9 protein (Jinek et al. 2012; Cong et al. 2013; Nekrasov et al. 2013; Shan et al. 2013). The sgRNA is composed of 20 nucleotides which are homologous to the genomic region targeted, followed by a short hairpin RNA (*shRNA*), also referring to scaffold RNA. Within the genome the 20 targeted nucleotides should be followed by a protospacer adjacent motif (PAM) composed of the nucleotides NGG. DSB induced by Cas9 are generally located three base pairs upstream of the PAM site. The ease of design and low cost explain the rapid success of this user friendly and efficient technology in a wide range of organisms including plants.

In the last five years, CRISPR-Cas9 technology has been successfully adapted to maize. For the introduction of the CRISPR-Cas9 machinery, direct DNA transfer to protoplasts (Liang et al. 2014; Xing et al. 2014), particle bombardment of immature embryos (Xing et al. 2014; Svitashev et al. 2015; Feng et al. 2016; Zhu et al. 2016) and *Agrobacterium*-mediated transformation of immature embryos (Xing et al. 2014; Svitashev et al. 2015; Feng et al. 2016; Zhu et al. 2016; Char et al. 2017) have been used. Protoplast experiments serve mainly for the evaluation of the efficiency of different sgRNA designs, since there is presently no protocol for the regeneration of maize plants from protoplasts. Biolistics avoid the use of *Agrobacterium*, which is regulated in certain countries since it is a plant pathogen. *Agrobacterium*-based stable transformation and subsequent elimination of the CRISPR-Cas9 casette by backcross nonetheless remains the most widely used method. The transfer is almost exclusively based on DNA molecules encoding *Cas9* and the sgRNA but the bombardment of Cas9 expressing plants with sgRNA (Xing et al. 2014; Svitashev et al. 2015; Feng et al. 2016; Zhu et al. 2016) and of wild-type plants with pre-assembled Cas9–sgRNA ribonucleoproteins (RNP), have also been reported (Svitashev et al. 2016). Multiplexing with more than one guide RNA in a single construct is of particular interest in maize due to the lengthy and not very efficient transformation protocol. Two techniques have been developed: one based on a multi-guide RNA activated by a single promoter and processed by tRNA motif-mediated self-cleavage into several sgRNAs, and another based on tandem repeats of different U3 and U6 promoters each controlling one guide RNA (Qi et al. 2016; Char et al. 2017). As expected, the mutations resulting from targeted mutagenesis were mainly deletions or insertions of a few nucleotides probably due to classical NHEJ. Larger deletions of more than 10 bases, potentially resulting from an MMEJ repair, have also been reported but are less frequent (Xing et al. 2014; Svitashev et al. 2015; Feng et al. 2016; Zhu et al. 2016). Furthermore, true genome editing, *i.e.* the predetermined modification of an allele based on a repair matrix carrying the desired

mutation by HR (Xing et al. 2014; Svitashev et al. 2015; Feng et al. 2016; Zhu et al. 2016) and the replacement of a promoter has also been achieved in maize (Svitashev et al. 2015; Shi et al. 2017).

Here we describe the CRISPR-Cas9-based mutagenesis of 20 maize genes selected for their putative implication in maize kernel development. The mutagenesis efficiency, the type of mutations obtained, the simultaneous knockout of tightly genetically linked genes and the rate of transmission to the next generation will be addressed.

Results

Multi-sgRNA plasmids for single and multiple gene editing

In order to carry out single or multiple gene mutagenesis using CRISPR-Cas9 technology in maize, two types of vectors were used. The first type was designed in-house and will be named RDP vectors hereafter (**Fig. 1**). The final construct typically contains two guide RNAs and is built by combining derivatives of the initial plasmids L1609 and L1611 (**Fig. 1**) by restriction and ligation. L1609 is a binary vector containing a T-DNA suitable for *Agrobacterium*-mediated maize transformation, which encompasses a plant selection marker conferring resistance to the Basta® herbicide and a *Cas9* coding sequence driven by the maize ubiquitin promoter, which is active in most plant tissues (Christensen and Quail 1996). The specific 20-nt sequence that will hybridize with the target site in the genome and thus guide the Cas9 complex to the gene(s) of interest, is inserted between the *Oryza sativa* U3 (*OsU3*) promoter and the *shRNA* (**Fig. 1**). The other initial plasmid L1611 allows the cloning of a second 20-nt targeting sequence between the *TaU6* (*Triticum aestivum U6*) promoter and a *shRNA*. The sub-cloning of this *TaU6::sgRNA* cassette into the modified L1609 plasmid leads to the generation of the final RDP vector with two sgRNAs (**Fig. 1**). The second type of CRISPR-Cas9 vector used was derived from the Gateway® compatible plasmid pGW-CAS9 developed by Iowa State University (Char et al. 2017) and will be referred to as Iowa vectors hereafter. Two to four sgRNA cassettes flanked by attL sites were entirely synthetized prior to recombination into pGW-CAS9.

Both the RDP and Iowa vectors used in this study contain multi guide RNAs, allowing the targeting of several genes with a single construct. We also designed multi-target 20 nt sequences, targeting up to 10 loci in the genome with one sgRNA, allowing for example to target paralogous genes (**Online Resource 1**). In our functional genetics approaches, we targeted the coding sequence to increase the likelihood of generating loss-of-function mutations. Four main strategies for sgRNA design were employed to achieve different types of gene knock-out(s) (KO) (**Fig. 2** and **Table 2**): (1) targeting two unique, non-related genes with a single guide RNA each, (2) targeting a unique gene with two guide RNAs, (3) targeting paralogs with a single or multiple guide RNAs and (4) targeting a unique gene with four guide RNAs.

Different types of mutations are created using multi guide RNAs strategies

A total of 20 genes were targeted with different RDP or Iowa vectors (**Table 1**). After stable transformation of maize immature embryos, DNA was extracted from young leaves of

transgenic T0 plants to assess the type and frequency of the mutations generated. Based on PCR amplification of the target site, and subsequent Sanger sequencing, at least one mutant allele was obtained for 18 of the 20 genes. All edited alleles are summarized in **Table 1**.

For genes targeted by a single guide RNA (strategies 1 and 3 in **Fig. 2**), a total of 56 mutations were generated in 13 genes (top section of **Table 1**). With the exception of one guide RNA targeting GRMZM2G352274, all other guide RNAs gave rise to new alleles, ranging from one to 12 different alleles (in the case of GRMZM2G089517). In this context it should be noted that the number of alleles does not reflect mutation efficiency, since transformation rates varied over time and not all transformation events were carried to the plantlet stage, and also because identical mutations could be generated independently in different plants. The mutations generated were predominantly (82%, 46/56) small indels, defined as short (<10 bp) insertions or deletions or mixtures of both (**Table 1**). As expected the vast majority of these indels occurred 3 bp upstream of the PAM sequence, the position where the Cas9 nuclease cleaves double stranded DNA (Zuo and Liu 2016). Less frequently (14%, 8/56), larger deletions (>10 bp) were observed, the largest one observed reaching 136 bp (**Table 1**). Interestingly, the majority of these larger deletions concerned a single gene, GRMZM2G089517, in which 6 of the 8 larger deletions were found. In addition, two substantial insertions (10 bp and 11 bp) of unrelated DNA occurred in this gene, both accompanied by the deletion of a few nucleotides, as well as three classical indels (**Table 1**). This atypical example suggests that a specific gene context may influence the type of mutations generated, possibly by favouring a particular repair mechanism. However, in the case of GRMZM2G089517 it was not possible to implicate a specific mechanism with certainty, since the start and end points were not shared between the large deletions and since a search for repeated nucleotides did not detect obvious micro-homologies in proximity to the cutting site. Lastly, two other types of mutations were observed only once. The first, which again concerns the atypical GRMZM2G089517 gene, consists of a substitution of two nucleotides on either side of the PAM site, for which it is difficult to provide a mechanistic explanation (**Table 1**). The second atypical mutation was found in GRMZM2G046086, in which 35 bp next to the putative cutting site were substituted by an insertion of 62 bp (**Fig. 3 and Table 1**). This insertion comprises an adenine nucleotide plus 61 nucleotides corresponding to a stretch of intergenic DNA region found 602 bp downstream of the putative cutting site (**Fig. 3**). Interestingly, this 61 bp intergenic sequence is still present at the original location in the two alleles of the T0 plant, indicating that it was duplicated to create this atypical mutation.

We next analysed the 6 genes that had been targeted by two guide RNAs concomitantly (strategy 2 in **Fig. 2**). The rationale behind this strategy was to increase the probability of success with a single construct, since a mutation at either target site would be sufficient for loss-of-function. In the ideal case the two guide RNAs, spaced between 40 bp and 100 bp apart, would induce deletions of a predictable size that could be easily detected by simple PCR in agarose gels and avoid the Sanger sequencing step to detect and follow the mutant allele. A total of 27 mutations were generated in 4 of the 6 genes, whereas neither deletions between the two cleavage sites nor other mutations were obtained for GRMZM2G040095 and GRMZM2G035701. There were no obvious reasons for the two failures, since the sgRNA design followed the same rules as for the 4 successful constructs. More intriguingly, the two sgRNA for GRMZM2G035701 (failure) were actually present on the same construct and in the same plants as the two sgRNA for GRMZM2G149940 (success). The large majority of the mutations identified (78%, 21/27) did not involve a deletion between the two guide RNA targets, but were caused by indels or larger deletions at one (74%, 20/27) or both target sites (4%, 1/27) (**Table 1**). Clear preferences for one of the two target sites were noted in all four cases and likely reflect differences in mutation efficiency or target site

accessibility. Only 22% (6/27) of mutations harboured deletions of the region located between the two target sites (**Table 1**). In only one case (GRMZM2G049141) the 100 bp deletion corresponded exactly to the zone between the two putative cleavage sites. Regarding the other five deletions, small indels at one or both target sites either caused deletions that were slightly smaller (GRMZM2G039538 and GRMZM2G363552) or slightly larger (GRMZM2G049141) than the expected size (**Table 1**). In summary, it was possible to generate deletions in regions between two guide RNAs. However, the exact size of the deletion was variable and deletions between two target sites were less frequent than indels generated by the action of an individual guide RNA.

Lastly, a vector with four guide RNAs was designed to target a unique gene (strategy 4 in **Fig. 2**). Three guide RNAs gave rise to mutations in GRMZM2G471240, which were all of the indel type (**Table 1**). No deletions between the four target sites were observed.

Mutation efficiency

In total 28 guide RNAs were expressed in plants, 20 using RDP vectors, and 8 using Iowa vectors. Three guides targeted two genes in conserved regions. Among them, 22 resulted in at least one mutation and 6 did not induce any sequence change in the analysed plants (**Table 2**). For RDP vectors, 17 guide RNAs induced at least one mutation and three did not generate a mutation. For Iowa vectors, the proportion of unsuccessful guides was higher (3/8) but this result should be interpreted with caution because considerably fewer transformation events were obtained when using Iowa vectors in our conditions. This was certainly due to the non-optimal combination of our *Agrobacterium* strain LBA4404 (pSB1) and the binary vector, and more precisely an incompatibility between the origins of replication of pSB1 and pGW-Cas9 (Char et al. 2017).

Bi-allelic mutations, meaning that alleles on both the maternal and paternal chromosomes carried mutations, were detected in 19% (16/83) of the mutated plants, and more precisely in 18% (13/74) of the mutants obtained with RDP vectors and in 33% (3/9) of the mutants generated with Iowa vectors (**Table 2**).

Mutation efficiency was calculated as the number of transformation events harbouring at least one mutation as a proportion of all transformation events obtained for a given guide RNA (**Table 2**). Although this number may be somewhat influenced by differences in the accessibility of certain targets, for example due to chromatin differences between centromeric and telomeric chromosome regions, or by competition between guide RNAs in the plants that produced more than one guide, it was clear that mutation efficiency was very variable despite similar rules for guide RNA design (**Table 2)**. Concerning the promoter used to drive guide RNA expression in the RDP vectors, mutations were obtained using both the *OsU3* and the *TaU6* promoter. Averaging the percentages for each promoter, a higher overall mutation efficiency was observed with the *TaU6* promoter (65%) as compared to the *OsU3* promoter (39%) (**Table 2**). Using the same approach, a slightly higher efficiency was noted when the 20-nt target and the NGG were chosen on the coding (+) strand (58%) compared to the non-coding (-) strand (48%) (**Table 2**). Finally, the mutation efficiency was not strongly correlated to the overall GC content of the 20 nt targeted sequence (r=0.31, **Table 2**).

Although the sample number (3 cases) in which 1 guide RNA was used to target two paralogous genes precluded a quantitative analysis, the mutation efficiency seemed to be in the same range for both target genes with 63%/88% for the first (GRMZM2G039538/GRMZM2G363552), 25%/25% for the second (GRMZM2G039538/GRMZM2G363552) and 75%/50% for the third guide RNA GRMZM2G140302/GRMZM2G046086) (**Table 2**). Our results on the first two case suggests

that the difference in mutagenesis efficiency between two guide RNAs targeting the same gene was more important than the difference between the mutagenesis efficiency for a single guide RNA targeted the two paralogs.

Transmission of edited genes to the next generation

Mutations must be present in germinal cells in order to be passed on to the next generation. We therefore tested whether mutations detected in leaf material of T0 plants fulfill this criterion. During the detection by PCR amplification and Sanger sequencing of leaf material of T0 transformation events, two categories of chromatograms indicative of editing were observed (**Online Resource 2)**: (1) the first and most common category (85%) was characterised by a switch from a homogenous chromatogram to two overlapping sequences with similar peak height (**Online Resource 2a)**, indicating two alleles present in approximately the same proportion in the extracted DNA; (2) the second, less frequent category (15%) presented a main signal and a very weak overlapping signal in the chromatograms (**Online Resource 2b)**, suggesting that the proportion of mutated DNA is very low compared to wild-type DNA. Without any exception, all mutations of the first category were systematically detected in the next generation. Transmission from the T0 to the T1 generation generally followed Mendelian segregation rules, suggesting that the edited alleles had been fixed and been present in all leaf cells and that the mutations had probably occurred early in the maize transformation process, likely during the callus formation step. It should be noted that all alleles presented in **Table 1**, including the alleles with multiple edits, were of this type. With regard to the second category, we never observed any transmission of the mutations to the next generation suggesting that the mutations were present only in few leaf cells and that the mutations had probably occurred during leaf development. These data indicate that, although chimeras may exist in maize, fully edited T0 plants are predominant and that the distinction between chimeric and fully edited T0 plants can be made on the basis of the Sanger chromatograms.

Limitation of off-target effects

During targeted mutagenesis of a gene of interest, the presence of additional mutations elsewhere in the genome should ideally be minimised. To reduce unintended off-target effects at sites with no homology to the target site, primary transformants were backcrossed to the parent line A188 and only Cas9-free mutant T1 plants, in which the T-DNA had been segregated away based on a negative PCR assay for the *BAR* gene (**Online Resource 3),** were used for subsequent molecular analyses. Furthermore, the use of at least two independent mutations (transformation events) and of several plants per mutation will allow to establish a clear link between the mutated gene and observed phenotypes.

To minimize off target effects at sites with substantial homology to the target, a rigorous design of the 20 nt recognition sequence was put in place (see Material and Methods). Since the objective of obtaining knockouts left a lot of freedom to the exact position of the targets in the coding sequence, targets with high similarities to sites elsewhere in the genome (less than 3 mismatches) were excluded from the design whenever possible. In fact, early works on the specificity of CRISPR-Cas9 had established that cleavage at target sites with more than 2 mismatches to the sgRNA were generally extremely rare (<0.01%), although exceptions existed (Mali et al. 2013; Pattanayak et al. 2013). While the design of the present study did not consider bulges (Lin et al. 2014), targets with similarity to other sites followed by NAG rather

than NGG PAM sites were also excluded. In the case of strategy 3 (**Fig. 2**) targeting two or more paralogous genes with a single sgRNA, these rules were not applied to the paralog(s) but maintained for the rest of the genome. As a quality control of the selected 20 nt recognition sequences used in this study, a customized BLAST search was performed (see Material and Methods section). For each of the 28 targets the most likely off target was selected based on the BLAST score and the position of the mismatches relative to the 7 to 12 bp seed sequence close to the PAM (**Online Resource 4**). Three of the putative off-target sites with 3 mismatches were chosen for experimental analysis of off-target mutations by specific PCR amplification followed by Sanger sequencing on T0 plants. No editing was detected at any of the three sites.

Creation of multiple mutants

One of the advantages of the CRISPR-Cas9 technology is that it allows the creation of multiple mutants in a single step, thereby avoiding time consuming crosses and/or backcrosses. With regard to unlinked genes located on different chromosomes, three double mutants were produced in the T0 generation using a construct with two guide RNAs, one for each gene (strategy 1 **in Fig. 2**). They concerned members of the same gene family in the case of GRMZM2G157313/GRMZM2G014499 (two double mutants in four transformation events, **Table 2**) and GRMZM2G059165/GRMZM2G120085 (1/3) and true paralogs in the case of GRMZM2G039538/GRMZM2G363552 (5/8). More importantly, multiple mutants were also obtained in genes that were tightly linked on the same chromosome, and for which the production of a double knockout mutant would have been difficult to achieve. Double mutants were identified for GRMZM2G089517/GRMZM2G352274 (separated by 75 kb on chromosome 5 (1 mutant found out of 14 transformation evens) and GRMZM2G145466/GRMZM2G573952 (located within 53 kb on chromosome 7 (1 out of 2). Finally, we successfully managed to knock out three small (<600 bp) paralogous genes that are situated in the same region of chromosome 1. These genetically strongly linked genes are *ZmEsr1*, *ZmEsr2* and *ZmEsr3* (GRMZM2G046086, GRMZM2G315601, GRMZM2G140302) (Opsahl-Ferstad et al. 1997). Since *ZmEsr2* and *ZmEsr1* are separated by only 29 kb, and *ZmEsr1* and *ZmEsr3* by only 13 kb, the production of a triple knockout mutant underlines the power of CRISPR-Cas9 technology. Using the CRISPR-Cas9 strategy 3 illustrated in **Fig. 2**, a plant with a frame-shift mutation in each of the three *ZmEsr* genes was obtained. By a simple self-pollination, we have been able to generate T1 plants homozygous for the three mutated *ZmEsr* genes that are now available for functional analysis. However, no large deletions between the cleavage sites in the linked genes were found, despite specific PCR reactions designed to detect them.

Discussion

The present study examined CRISPR-Cas9-mediated targeted mutagenesis in maize aimed at routine use for functional genetics studies. Analysing mutations in 20 genes in genome edited maize plants, it was conducted at a larger scale than previous studies in maize, which either simply demonstrated the feasibility for a single gene or addressed a maximum of 5 genes (Liang et al. 2014; Xing et al. 2014). It also focused on regenerated plants rather than protoplasts or calli, systematically analysed offspring and is the first study to use the inbred line A188. The results indicate that CRISPR-Cas9 is a robust technology for gene knockout in

maize, and can be used to generate various types of mutations with a high frequency of success. Furthermore it allows the production of double and triple mutants in tightly linked genes.

Three types of mutations were observed in the 93 mutant maize plants analysed: indels, larger deletions and local chromosome rearrangements. The occurrence of larger chromosome rearrangements, such as those reported recently for mouse embryonic stem cells (Kosicki et al. 2018), cannot be excluded but would not be detected with our method. Indeed the detection method, based on PCR amplification and subsequent Sanger sequencing, can only detect mutations in which the two primer binding sites on either side of the putative cleavage site are conserved in head to head orientation and remain at a distance allowing standard PCR amplification. Small indels as produced in the case of classical NHEJ repair (Ma et al. 2016; Bortesi et al. 2016) were, as expected, the most frequent outcome (80%, 74/93) and were documented for each of the 18 genes that were successfully mutagenized. They were generally located at, or close to, the putative cleavage site 3 bp upstream of the PAM. Larger deletions (>10 bp) ranging from 11 bp to 136 bp were considerably less frequent (11%, 10/93) and concerned only 4/18 genes. Thought to be generated by the MMEJ repair mechanism, short (2 bp to 4 bp) microhomologies were indeed present on both sides of the putative cleavage site in the wild-type sequence of GRMZM2G120085 (GC), GRMZM2G149940 (CCG) and GRMZM2G049141 (GACT) and the large deletions tended to correspond more or less precisely to recombination products between these direct repeats. In contrast, for GRMZM2G089517 larger deletions were more frequent than indels, the start and end points of the deletions were not conserved between events, and two other atypical mutations were obtained: a combination of a 7 bp deletion with an 11 bp insertion, and two point mutations flanking the PAM. The mechanism generating these atypical mutations remains unclear, although it is known that strand resection and random DNA synthesis can lead to unpredictable outcomes during MMEJ repair (Wang and Xu 2017; Sinha et al. 2017).

An unexpected allele was also detected for GRMZM2G046086 alongside 5 other classical indels. This allele consists of a 35 bp deletion accompanied by the insertion of a 61 bp DNA fragment copied from the intergenic region downstream of the gene (**Fig. 3**).

Importantly, defined deletions (6%, 6/93) of predetermined size and position were successfully provoked by the simultaneous action of two guide RNAs on target sites separated by between 44 bp and 102 bp in a given gene. The precision of these deletion events was not perfect, since only one deletion was precisely of the expected size, whereas the other five contained indels of 1 bp or 2 bp at least at one end of the deletion. In addition, this approach worked only for 3/7 targeted genes and in the three successful cases indels at only one of the target sites were more frequent than the deletion. Since variations in target accessibility over such short distances in coding regions are unlikely, this suggests that very similar efficiency of the two guide RNAs is crucial for the successful generation of defined deletions. Differences between guide RNA efficiency may also partially explain why larger deletions involving target sites distant between 13 kb and 75 kb in genetically linked paralogs were not detected, despite the fact that CRISPR-Cas9-mediated deletions of up to 120 kb have been documented in plants (Gantner et al. 2018).

The overall mutation efficiency (averaging the percentages for each guide RNA) of 53% was in the global range (2% to 100%) of previous reports on targeted mutagenesis in maize, as was the 19% rate of biallelic mutations obtained (Liang et al. 2014; Xing et al. 2014; Lee et al. 2018). Since higher rates have been achieved in maize with the same basic elements (maize ubiquitin promoter, codon optimized Cas9, cereal U3 or U6 promoters), the specific

choices made during vector design, such as the choice of different versions of the *ZmUbi* promoter, the choice of the terminator, the position of promoter-*Cas9* and *Cas9*-terminator junctions, as well as the presence of an NLS domain, of tags for immuno-detection or of introns in the *Cas9* coding sequence, are possible parameters for optimisation. However, this suboptimal rate of biallelic mutations also has advantages in the context of functional genetics studies of genes involved in maize kernel development, since mutations could be lethal for the embryo and/or seedling in the homozygous state (Neuffer and Sheridan 1980; Doll et al. 2017). It is therefore preferable to generate heterozygous plants and to assess the (lethal) phenotype after self-pollination in segregating ears.

More importantly, the mutation efficiency was very variable at different levels. Firstly, two of the 20 genes could not be mutated at all, despite the use of two guide RNAs per gene and the generation of 8 and two transformation events, respectively. Secondly, among the 18 genes successfully mutated, not all transformation events caused mutations. For example, in the case of GRMZM2G352274 only one of the 16 transformation events yielded a mutation. Thirdly, in transformation events carrying novel mutations, not all guide RNAs present in the same plant induced mutations. The reasons for failure are likely linked either to the intrinsic quality of the sgRNA design or to the accessibility of the target sequence. Although the design of all sgRNAs followed the same rationale, the online and in house tools used only ensure a relatively high minimum quality standard, but they do not exclude quality differences between the possible designs. The GC content of the binding site (Ren et al. 2014; Labuhn et al. 2018), the secondary structure of the sgRNA, and its capacity not only to guide but also to activate the nuclease activity of Cas9 are known to be important parameters (Liu et al. 2016). In this context it is noteworthy that the GC content of both target sites in GRMZM2G035701 (failure) was relatively low (45%), whereas the GC content of the two sites in GRMZM2G149940 (targeted with success by the same construct in the same plants) was considerably higher (60% and 65%). The criteria for target site accessibility are less clear. Although Cas9 cleavage activity is not thought to be strongly affected by DNA CpG methylation (Hsu et al. 2013), it is generally accepted that the chromatin status of the target region influences the efficiency of CRISPR-Cas9 approaches, that DNase I hypersensitivity (DHS) is a good indicator for Cas9 binding (Wu et al. 2014) and that heterochromatin may be less accessible (Jensen et al. 2017). On the other hand, the accessibility of genes located in globally heterochromatic, centromeric regions of maize chromosomes to Cas9-mediated targeted mutagenesis has been demonstrated in protoplasts (Feng et al. 2016). In our study, the two recalcitrant genes GRMZM2G035701 and GRMZM2G040095 are located in gene-rich regions on the long arm of chromosome 8 and close to the end of chromosome 2, respectively. These regions do not present any obvious features explaining failure.

Differences in mutation efficiency between transformation events are expected, since the genomic environment is known to influence the expression level of transgenes, in the present case of the *Cas9* and *sgRNA* genes. However, very low success rates, such as the single edit for GRMZM2G352274 in 16 transformation events, are difficult to explain by insufficient expression, in particular since the second guide RNA present in the same plants caused mutations in 14/16 events. In this as in other cases, the competition of guide RNAs of unequal quality, or differences in target gene accessibility, are more likely explanations for differences in successful mutagenesis than positional effects on transgene expression. Our study suggests that other parameters with a minor impact on mutation efficiency were the choice of the type III promoter with a preference for the *TaU6* over the *OsU3* promoter, and the choice of the DNA strand with mutagenesis improved by binding of the sgRNA to the template rather than non-template strand. This last observation is likely caused by a quicker

release of the Cas9 from the template strand due to displacement by RNA polymerase II and faster repair of the DSB by the cellular machinery (Clarke et al. 2018). Overall, these results can be translated into 5 recommendations for gene-knockout in maize: 1) The use of at least two guide RNAs per gene. 2) The generation of at least 5 transformation events. 3) The retargeting of recalcitrant genes with constructs targeting a single gene. 4) The use of maize or wheat *U6* promoters. 5) The preferential use of target sequences on the coding strand.

Chimerism is an important issue in CRISPR-Cas9-mediated mutagenesis, since in stably transformed plants constitutively expressing *Cas9* and *sgRNA* genes, genome editing can occur at any time and in any number of cells during the life cycle of the plant, raising the question of whether the mutations detected in the leaves or other organs of primary transformants will be present in germinal cells and thus transmitted to the offspring. Our results indicate that chimerism does occur, but that in the majority of events detected in leaf material are fully edited and that sequencing chromatograms with overlapping sequencing peaks of equal height are predictive for transmission to the next generation. This is in agreement with earlier reports in maize (Liang et al. 2014; Xing et al. 2014) and seems to indicate that the majority of editing events occur very early on during the transformation of immature maize embryos, likely at the callus stage.

The ease of multiplexing is frequently cited as one of the major advantages of CRISPR-Cas9 technology over the use of other site-directed nucleases such as meganucleases, zinc finger nucleases or TALENs, and CRISPR-Cas9 constructs harbouring as many as 14 guide RNAs have been used successfully in Arabidopsis (Peterson et al. 2016). Three double mutants in gene family members residing on different chromosomes, two double mutants in paralogues separated by 53 kb or 75 kb, and a triple mutant in paralogues separated by 13 kb or 29 kb were generated in our study. These examples underline the power of CRISPR-Cas9 technology since the production of double or triple knockout mutants in tightly linked genes would have been nearly impossible to achieve by crossing of single mutants, and would have required the analysis of thousands of recombinants. Multiplexing is of particular interest in maize, which is an ancient tetraploid known to contain numerous functionally redundant paralogues, hampering functional analysis. As a result the production of multiple mutants by CRISPR-Cas9 will almost certainly become a prime tool for functional genomics studies in this species.

Acknowledgement

We thank Bing Yang (Iowa State University) for sharing the Iowa vectors prior to publication, Jean-Philippe Pichon (Biogemma SA) for providing unpublished genomic sequences of genotype A188, Justin Berger, Patrice Bolland and Alexis Lacroix for maize culture, Isabelle Desbouchages and Hervé Leyral for buffers and media preparation, as well as Sandrine Chaignon, Jérôme Laplaige and Edwige Delahaye for technical assistance. We acknowledge support by the Investissement d'Avenir program of the French National Agency of Research for the project GENIUS (ANR-11-BTBR-0001_GENIUS) to PMR and by the INRA Plant Science and Breeding Division for the project SeedCom to TW. NMD is funded by a PhD fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche. LMG and YF are supported by CIFRE fellowships of the ANRT (grants 2015/0777 and 2018/0480). VMGB is supported by the Doctoral School on the Agro-Food System (Agrisystem) of Università Cattolica del Sacro Cuore (Italy).

Author contribution statement

NMD, PMR and TW conceived and designed research; NMD, LMG, MFG, CR, YF, VMG, GG, and TW conducted experiments. MFG, CR and GG performed maize transformation. NMD prepared tables and figures. JJ performed bioinformatics analyses to produce **Online Resource 1** and **Online Resource 4**. NMD, GI, PMR and TW wrote the manuscript. PMR and TW were involved in project management and obtained funding.

Data availability statement

The two plasmids generated during the current study are available at GeneBank under the accession number MH662439 for L1609 and MH662440 for L1611.

Conflict of Interest

LMG is employed by Limagrain Europe. YF is employed by MAS seed. PMR is part of the GIS-BV ("Groupement d'Intérêt Scientifique Biotechnologies Vertes").

Material and Methods

Plant Material and Growth Conditions

The maize (*Zea mays*) inbred line A188 (Gerdes and Tracy 1993) and derived transgenic or edited plants were grown in growth chambers that fulfill the French S2 safety standards for the culture of transgenic plants. In the 15 m^2 growth chambers the plants were illuminated by a mixture of 10 LED spots of 500 W (Neptune LED, Ste Anne sur Gervonde) set at 60% intensity and 8 high-pressure sodium lamps of 400 W, resulting in the spectrum presented in **Online Resource 5** and a photosynthetic photon flux density (PPF) of about 300-400 µmol.s-1 at plant height. The photoperiod consisted of 16 h light and 8 h darkness in a 24 h diurnal cycle. Temperature was set to 24°C/17°C (day/night) during the first 84 days after sowing (DAS) and then to 26°C/28°C for the remaining 30 days of the life cycle. The relative humidity was controlled at 55% (day) and 65% (night). Seeds were germinated in 0.2 L of Favorit MP Godets substrate (Eriterre, Saint-André-de-Corcy) and were transferred at between 12 and 20 DAS to 8 L of Favorit Argile TM + 20% perlite substrate (Eriterre, Saint-André-de-Corcy) supplemented with 50 ml of Osmocote Exact Hi.End 5-6M (15-9-12+2MgO+TE) fertilizer (Scotts, Écully). All plants were propagated by hand pollination.

Vector cloning

The integrative plasmid L1609 (**Fig. 1**) is based on the backbone of pSB11 (Ishida et al., 1996), from which a *Sap*I site was removed. It contains between the T-DNA borders a rice codon optimized *Cas9* (Miao et al., 2013) driven by a synthetic maize ubiquitin promoter lacking several restriction sites, a rice U3 promoter separated from a sgRNA scaffold (Shan et al., 2013) by two adjacent but otherwise unique *Sap*I sites, unique *EcoR*V and I-*Ceu*I sites and a Basta® resistance cassette. The small plasmid L1611 (**Fig. 1**) contains a wheat U6 promoter followed by two adjacent *Sap*I sites and a shRNA (Shan et al., 2013), the entire cassette being

flanked by unique *Eco*RV and I-*Ceu*I sites. Annealed oligonucleotides with *Sap*I compatible overhangs and corresponding to 20 nt targeted sequences containing at their 5' end an A in the case of the U3 promoter or a G in the case of the U6 promoter were cloned in L1609 and L1611 respectively. The U6 driven target cassette present in L1611 was subsequently excised with *Eco*RV and I-*Ceu*I and cloned into the L1609 derivative downstream of the U3 driven target cassette. The resulting plasmid was transferred to *Agrobacterium tumefaciens* strain LBA4404 (pSB1) and used for maize transformation. Alternatively, Gateway compatible assemblies of two to four cassettes consisting each of a long or short maize U6 promoter, followed by a 20 nt target site starting with a G and a shRNA (Char et al. 2017) were entirely synthesized (GENEWIZ, New Jersey) and recombined into plasmid pGW-Cas9 (Char et al. 2017) containing between T-DNA borders a maize codon optimized Cas9 driven by maize ubiquitin promoter and a Basta® resistance cassette conferring glufosinate-ammonium herbicide resistance.

20-nt target sequence choice

For the design of sgRNAs targeting specifically a single gene in the maize genome, the online tools CRISPR-P (http://crispr.hzau.edu.cn/CRISPR/) (Lei et al. 2014) and CRISPOR (http://crispor.tefor.net/) (Haeussler et al. 2016) were interrogated and targets at convenient positions with high scores in both tools were chosen. Since these tools are not readily suited to target several members of a gene family with a single sgRNA, we wrote custom Perl scripts to design sgRNAs directed against up to 10 genes each. All candidate CRISPR-Cas9 targets were identified in the B73 maize reference genome sequence v3.26 (Schnable et al. 2009) using the following criteria: 23-mers ending with NGG, not containing more than 4 Ts in a row, and with no variant of the last 12 nt ending in NAG existing in the genome. Using Jellyfish v2.2.0 (Marçais and Kingsford 2011), we counted the number of occurrences in the genome of the last 15 nt of each candidate (excluding NGG), and kept only those occurring at most 10 times. The resulting database (**Online Resource 1** and https://flower.ens-lyon.fr/maize/crispr/) contained 15,715,633 23-nt sequences, targeting 19,024,477 loci. We queried it to identify targets in the genes we wanted to edit. In both cases the design was only retained, if the sequence of the reference genome of genotype B73 v3.26 (Schnable et al. 2009) available in the design tools did not show any polymorphism in the 20-nt target sequence and the PAM with the sequence of genotype A188 used for transformation.

Identification of off-target loci

To identify potential off-target loci in the maize genome, the 23 nt sequence of each sgRNA was used as query in a WU-Blast search of the B73 maize reference genome sequence v3.26 with the very relaxed parameters “W=1 M=4 N=-5 Q=8 R=7 gapX=100 E=1e7 V=50 B=1e6 filter=none kap pingpong”. Alignments were subsequently filtered with *ad-hoc* scripts to keep those covering the whole sgRNA length, with at most three mismatches in the last 15 nt, if the NGG had a perfect match; in case the NGG was not conserved, we kept only alignments with a perfect match on the last 15 nt. Alignments were sorted by decreasing match quality, favoring those with the longest match on the 3’ region of the sgRNA, and manually examined (**Online Resource 4** and https://flower.ens-lyon.fr/maize/crispr/). For experimental validation, off target-specific PCR primers were designed (**Online Resource 3**) and used for PCR amplification and Sanger sequencing.

Maize transformation and screen for Cas9-free edited plants

Immature embryos of maize inbred line A188 were transformed with *A. tumefaciens* strain LBA4404 harboring pSB1 and the construct of interest according to a standard protocol (Ishida et al. 1996, 2007). T-DNA integrity was checked as described elsewhere (Gilles et al. 2017). Genome editing was evaluated on leaves of T0 plants, individually for each targeted gene by specific PCR amplification of the targeted region (see **Online Resource 3** for primer sequences) followed by Sanger sequencing. Segregation of T-DNA in T1 plants was evaluated by PCR amplification on the *Bar* gene, checking the presence and quality of genomic DNA by PCR amplification of the GRMZM2G136559 control gene (see **Online Resource 3** for primer sequences).

Figures

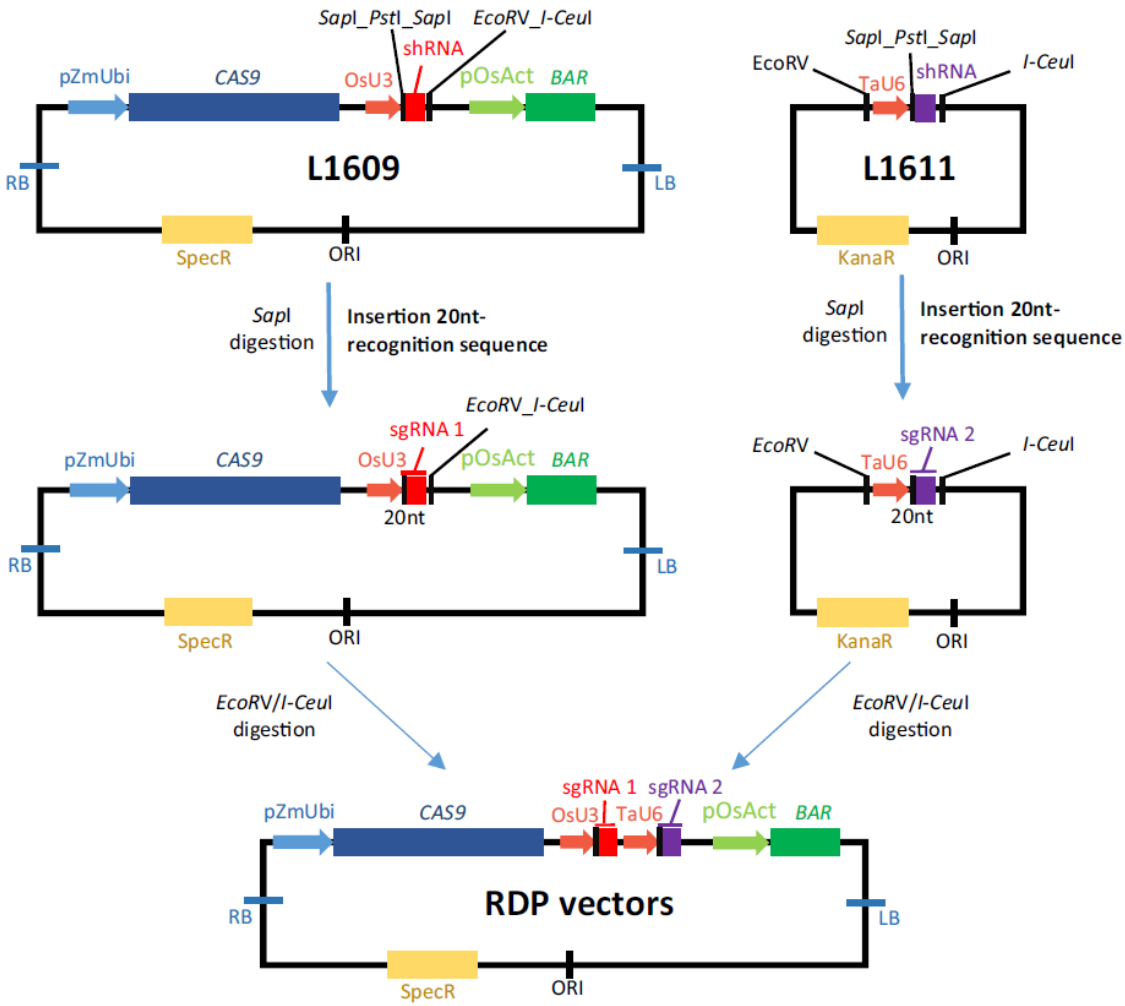

Fig. 1 CRISPR-Cas9 cloning vectors.

Cloning strategy for RDP vectors. The final RDP plasmids contain two small guide RNAs (*sgRNA1* and *sgRNA2*) and are generated by assembly of the two initial plasmids L1609 and L1611. First the 20 nt corresponding to the recognition sequences are synthesized as oligonucleotides with *SapI* compatible ends and inserted between the U3 or U6 promoter and

the scaffold RNA (*shRNA*) after *Sap*I digestion in both plasmids, forming *sgRNA1* and *sgRNA2*. Then the *TaU6::sgRNA2* cassette is transferred by *Eco*RV/I-*Ceu*I digestion into the plasmid already containing the *OsU3::sgRNA1* cassette. *BAR* = Basta® resistance gene, *Cas9* = rice codon optimized *Cas9* gene, LB = T-DNA left border, *OsU3* = rice U3 promoter, *pActUbi* = maize ubiquitin promoter, *pOsAct* = rice actin promoter, RB = T-DNA right border, *shRNA* = short hairpin RNA, *sgRNA* = small guide RNA, *TaU6* = wheat (*Triticum aestivum*) U6 promoter, 20 nt = recognition sequence of 20 nucleotides inserted before the *shRNA*

Fig. 2

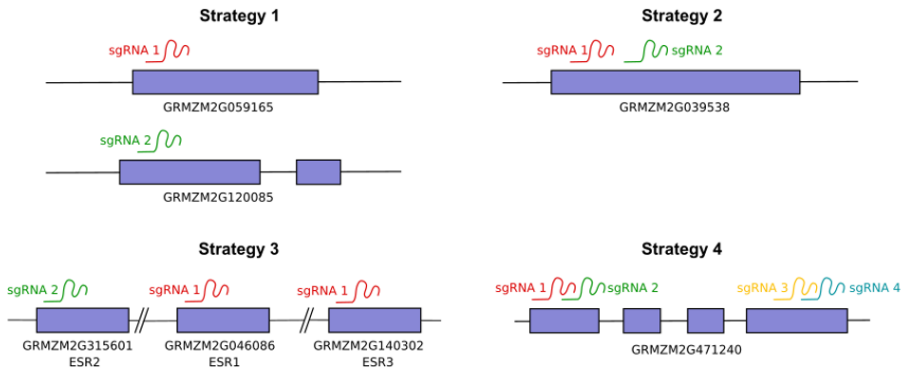

Fig 2 Different approaches to generate single and multiple gene knockout in maize.
Scheme illustrating different types of strategies for single or multiple gene knockout. The first strategy consists of targeting two distinct genes with specific guide RNAs for each gene, the second of targeting a single gene with two guide RNAs, the third of targeting several paralogous genes with one or several guide RNAs, and the fourth of targeting a single gene with four guide RNAs

Fig. 3

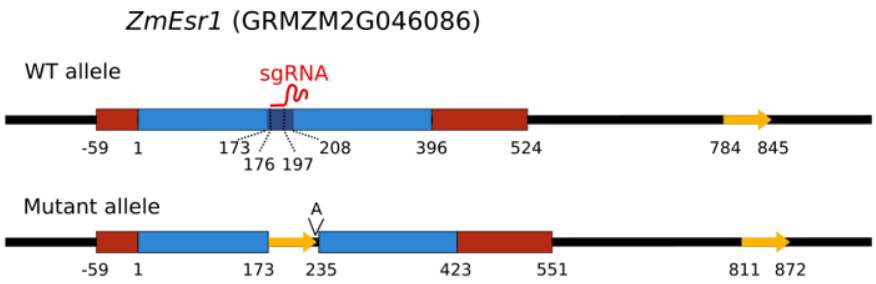

Fig 3 Scheme of an atypical mutant allele of *ZmEsr1* (GRMZM2G046086). The intronless *ZmEsr1* gene is represented by a square box with the open reading frame in blue and the UTRs in red. Numbering starts at the first nucleotide of ATG start codon. The duplicated intergenic sequence is depicted in yellow. The 35 bp segment deleted in the mutant allele is indicated in dark blue

Table 1 CRISPR-Cas9 alleles generated in 20 maize genes.

a: Sequence of the coding strand of the targeted genes around the recognition site (underlined) and the PAM (in blue)
b: Position of the mutation relative to the putative cleavage site (3 bp upstream of the PAM)
c: Number of independent transformation events with the same mutation

Table 1 CRISPR–Cas9 alleles generated in 20 maize genes

Gene targeted (ID Maize Genome V3)	Allele	Sequence[a]	Position from the putative cutting site[b]	Occurrence[c]
Gene targeted by one guide with RDP vectors				
GRMZM2G157313	WT A188	GACCGGAACGCGACATG-GTACGGGAGCACA		
	ins G	GACCGGAACGCGACATGGGTACGGAGCACA	0	X2
GRMZM2G014499	WT A188	ACAGGTCTACATAGTGT-ACCTGGGCGAGC		
	del TGTACCT	ACAGGTCTACATAG--------GGGCGAGC	-3	
	ins T	ACAGGTCTACATAGTGTTACCTGGGCGAGC	0	X2
GRMZM2G059165	WT A188	AGCATTCTGCACCAGGGTATCGGACCCAT		
	del T	AGCATTCTGCACCAGGG-ATCGGACCCAT	1	
GRMZM2G120085	WT A188	CGCACTCCTCGC-AAGCCCCTCGCTTCCCC		
	del CTCCTCGCAAG	CGCA----------CCCCTCGCTTCCCC	-8	
	del AA	CGCACTCCTCGC---GCCCCTCGCTTCCCC	0	
	del A	CGCACTCCTCGC--AGCCCCTCGCTTCCCC	0	X2
	ins A	CGCACTCCTCGCAAAGCCCCTCGCTTCCCC	0	
GRMZM2G145466	WT A188	GGAGAAGCACACAAAGC-GCGTGGACGCAC		
	del CG	GGAGAAGCACACAAAG---CGTGGACGCAC	0 or 1	
	ins A	GGAGAAGCACACAAAGCAGCGTGGACGCAC	0	
GRMZM2G573952	WT A188	AAAAGGTGCTACTGCTG-CGTTGGGGGTAG		
	ins A	AAAAGGTGCTACTGCTGACGTTGGGGGTAG	0	
	ins G	AAAAGGTGCTACTGCTGGCGTTGGGGGTAG	0	
GRMZM2G046086	WT A188	TATTGGCCATCG--TCCAAGACTTGCATCTT		
	del 35pb ins 62pb	see Figure 3	-6	
	del TCCA ins G	TATTGGCCATCG-G----AGACTTGCATCTT	0	
	del CG ins A	TATTGGCCAT---ATCCAAGACTTGCATCTT	-2	
	ins C	TATTGGCCATCG-CTCCAAGACTTGCATCTT	0	
	ins A	TATTGGCCATCG-ATCCAAGACTTGCATCTT	0	
	ins TT	TATTGGCCATCGTTTCCAAGACTTGCATCTT	0	
GRMZM2G140302	WT A188	TATTGGCCATCG-TCCAAGACTTGCATCTT		
	del TCCAAG	TATTGGCCATCG-------ACTTGCATCTT	0	
	del T	TATTGGCCATCG--CCAAGACTTGCATCTT	0	x2
	ins C	TATTGGCCATCGCTCCAAGACTTGCATCTT	0	
	ins A	TATTGGCCATCGATCCAAGACTTGCATCTT	0	
GRMZM2G315601	WT A188	AAAAGACCTGTG-CCTAGCGGGCCAGACCC		
	del CCT	AAAAGACCTGTG----AGCGGGCCAGACCC	0	
	del C	AAAAGACCTGTG--CTAGCGGGCCAGACCC	0	X3
	ins T	AAAAGACCTGTGTCCTAGCGGGCCAGACCC	0	
	ins A	AAAAGACCTGTGACCTAGCGGGCCAGACCC	0	
GRMZM2G134341	WT A188	TCCTCCCCAGGT-TGCCGGGTGCGACCTGT		
	del CCGGG	TCCTCCCCAGGT-TG-----TGCGACCTGT	2	
	del TG	TCCTCCCCAGGT---CCGGGTGCGACCTGT	0	
	ins A	TCCTCCCCAGGTATGCCGGGTGCGACCTGT	0	
AC208201.3_FG003	WT A188	GCAGACGTGCGACCTGT-ACCGGGGCAGCT		
	ins T	GCAGACGTGCGACCTGTTACCGGGGCAGCT	0	X2
	del ACCGGG	GCAGACGTGCGACCTGT-------GCAGCT	0	
	del A	GCAGACGTGCGACCTGT--CCGGGGCAGCT	1	
	del GT	GCAGACGTGCGACCT---ACCGGGGCAGCT	-2	
	ins A	GCAGACGTGCGACCTGTAACCGGGGCAGCT	0	X2
GRMZM2G089517	WT A188	CTTGAAGTGAGGACTGCAAGAA----GGCCGGCTCTCGCCGACATCTCTG		
	del 136pb	CTTGAAG----------15	-15	
	del 43pb	----------TCTG	-31	
	del 32pb	----------CCGGCTCTCGCCGACATCTCTG	-30	
	del 17pb	CTTGAAGTGAGGACTGCAA----------GACATCTCTG	-3	
	del 8pb	CTTGAAGTGAGGACTGCAAGAA----------CTCGCCGACATCTCTG	0	
	del 28pb, ins GT	----------GTCTCTCGCCGACATCTCTG	-21	
	del 16pb, ins 10pb	CTTGAAGTGAGGACT----------ACCAGAGATGTCGCCGACATCTCTG	-7	
	del 7pb ins 11pb	CTTGAAGTGAGGACTGCAAGATGCTCTTTGGTCTCTCGCCGACATCTCTG	-1	
	del AAG	CTTGAAGTGAGGACTGCAAG-------GCCGGCTCTCGCCGACATCTCTG	-2	
	ins T	CTTGAAGTGAGGACTGCAAGAAT---GGCCGGCTCTCGCCGACATCTCTG	0	
	ins A	CTTGAAGTGAGGACTGCAAGAAA---GGCCGGCTCTCGCCGACATCTCTG	0	
	2 mutations	CTTGAAGTGAGGACTGCAAGAA----GGTCGGATCTCGCCGACATCTCTG	3 and 7	
GRMZM2G089517	WT A188	GAATGGTGCTGTCAAGC-GGCCGGCTCGGC		
	ins T	GAATGGTGCTGTCAAGCTGGCCGGCTCGGC	0	
GRMZM2G352274	WT A188	CGG-CAACACATCCAATCGAATGAAGATTCTTCAC		
	ins T	CGGTCAACACATCCAATCGAATGAAGATTCTTCAC	14	
GRMZM2G352274	WT A188	AACGGACTGCTCCTTGCAGGCGGCTCCAT		

Table 1 (continued)

Gene targeted (ID Maize Genome V3)	Allele	Sequence[a]	Position from the putative cutting site[b]	Occurrence[c]
Gene targeted by two guides with RDP vectors				
GRMZM2G039538	WT A188	CCTCTTCCACTC-GGGCGGCGAGCTCCAGCAGCATCCTGTACACCTACGACACCGTCAT-GCACGGCTTCGC		
	del 45pb	CCTCTTCCACTC----------T-GCACGGCTTCGC	0 and -1	
	del 44pb	CCTCTTCCACTC-G----------T-GCACGGCTTCGC	1 and -1	
	del GGGCG ins C ; ins T	CCTCTTCCACTC-C----GCGAGCTCCAGCAGCATCCTGTACACCTACGACACCGTCATTGCACGGCTTCGC	0 and 0	
	del GGGC	CCTCTTCCACTC-----GGCGAGCTCCAGCAGCATCCTGTACACCTACGACACCGTCAT-GCACGGCTTCGC	0	
	del G	CCTCTTCCACTC--GGCGGCGAGCTCCAGCAGCATCCTGTACACCTACGACACCGTCAT-GCACGGCTTCGC	0	x2
	del CTCGGGC ins T	CCTCTTCCA----T---GGCGAGCTCCAGCAGCATCCTGTACACCTACGACACCGTCAT-GCACGGCTTCGC	-3	
	ins T	CCTCTTCCACTCTGGGCGGCGAGCTCCAGCAGCATCCTGTACACCTACGACACCGTCAT-GCACGGCTTCGC	0	X2
GRMZM2G363552	WT A188	CCTCTTCCACTC-GGGCGGCGAGCTCCAGCAGCATCCTGTACACCTACGACACCGTCAT-GCACGGCTTCGC		
	del 44pb	CCTCTTCCACTC-G----------T-GCACGGCTTCGC	1 and -1	X2
	del GGG ins TT	CCTCTTCCACTC-TT-CGGCGAGCTCCAGCAGCATCCTGTACACCTACGACACCGTCAT-GCACGGCTTCGC	0	
	del GG	CCTCTTCCACTC---GCGGCGAGCTCCAGCAGCATCCTGTACACCTACGACACCGTCAT-GCACGGCTTCGC	0	
	del G	CCTCTTCCACTC--GGCGGCGAGCTCCAGCAGCATCCTGTACACCTACGACACCGTCAT-GCACGGCTTCGC	0	
	ins G	CCTCTTCCACTCGGGGCGGCGAGCTCCAGCAGCATCCTGTACACCTACGACACCGTCAT-GCACGGCTTCGC	0	
	ins T	CCTCTTCCACTCTGGGCGGCGAGCTCCAGCAGCATCCTGTACACCTACGACACCGTCAT-GCACGGCTTCGC	0	
	ins A	CCTCTTCCACTCAGGGCGGCGAGCTCCAGCAGCATCCTGTACACCTACGACACCGTCAT-GCACGGCTTCGC	0	
GRMZM2G049141	WT A188	ACAGTGCCGACGATGGCAGTATATCGTCCAGGCCAGC-(60nt)-GGCCCCAGCCCCGAACCCTGAATCTTCATCAT		
	del 100pb	ACAGTGCCGACG----------(60nt)----------CCCTGAATCTTCATCAT	0 and 0	
	del 102pb	ACAGTGCCGACG----------(60nt)----------CTGAATCTTCATCAT	0 and 2	
	del TGCCGACG	ACAG--------ATGGCAGTATATCGTCCAGGCCAGC-(60nt)-GGCCCCAGCCCCGAACCCTGAATCTTCATCAT	-8	
	del ATG	ACAGTGCCGACG---GCAGTATATCGTCCAGGCCAGC-(60nt)-GGCCCCAGCCCCGAACCCTGAATCTTCATCAT	0	
	del 17 pb	A----------GTATATCGTCCAGGCCAGC-(60nt)-GGCCCCAGCCCCGAACCCTGAATCTTCATCAT	-11	
	del 15 pb	A----------CAGTATATCGTCCAGGCCAGC-(60nt)-GGCCCCAGCCCCGAACCCTGAATCTTCATCAT	-11	
	del CCTG	ACAGTGCCGACGATGGCAGTATATCGTCCAGGCCAGC-(60nt)-GGCCCCAGCCCCGAAC----AATCTTCATCAT	1	
GRMZM2G040095	WT A188	CGAGCGCCCTGCTCAAGTACCGGGAGGACGAGCTCGGCG-(40nt)-GCCGTGGGACCCCGTGTACGACTACGCGCTGT		
Gene targeted by 2 guides with Iowa vectors				
GRMZM2G149940	WT A188	CTTCCACCAATA-CCCCGCCGGCTTGATCCCAGCGCCGGTGGCACTGCCGGTTCACGCACCCGGTGTCGTCACAGACGTCTCC		
	del 54pb	CTTCCACCAATA-CCCCGCCGGCTTGATCCCAGCGCCGGTGGCACTGCCGGTTCACGCACCCGGT----------	-1	
	del G	CTTCCACCAATA-CCCCGCCGGCTTGATCCCAGCGCCGGTGGCACTGCCGGTTCACGCACCCGGT-TCGTCACAGACGTCTCC	-1	
	ins A	CTTCCACCAATAACCCCGCCGGCTTGATCCCAGCGCCGGTGGCACTGCCGGTTCACGCACCCGGTGTCGTCACAGACGTCTCC	0	
GRMZM2G035701	WT A188	GAATACAGCTGCTCTTGATCCGGATCATT-(30nt)-GGAGCTCCGAAATAGCGATGTAAGCCAGC		
Gene targeted by 4 guides with Iowa vectors				
GRMZM2G471240	WT A188	GCAATACCTGTAGCACGAAGGCGATGGCC		
GRMZM2G471240	WT A188	GCGGCCTCTCTACGCTG-CCAAGGACATCA		
	del G	GCGGCCTCTCTACGCT--CCAAGGACATCA	0	
	ins G	GCGGCCTCTCTACGCTGGCCAAGGACATCA	0	
	ins T	GCGGCCTCTCTACGCTGTCCAAGGACATCA	0	
GRMZM2G471240	WT A188	GAGGGTGTCCAGGGTCAACGTGGAGACAG		
	del CAA	GAGGGTGTCCAGGGT---CGTGGAGACAG	-2	
	del A	GAGGGTGTCCAGGGTC-ACGTGGAGACAG	0	
GRMZM2G471240	WT A188	GGAGACAGGGAGGTACG-AACCGGTGACTG		
	del GA	GGAGACAGGGAGGTAC--AACCGGTGACTG	0	
	del GA ins C	GGAGACAGGGAGGTAC-C-ACCGGTGACTG	0	
	del G	GGAGACAGGGAGGTAC--AACCGGTGACTG	0	
	ins T	GGAGACAGGGAGGTACGTAACCGGTGACTG	0	
	ins A	GGAGACAGGGAGGTACGAAACCGGTGACTG	0	

Table 2 Guide RNAs used and relationship with plant transformation events.
a Refers to strategies described in Fig. 2
b Mutation efficiency is defined as percentage of transformation events leading to at least one mutation within the targeted gene(s)
c Chromosome carrying the targeted gene
d DNA strand targeted by sgRNA, relative to gene orientation. "−" refers to the non-coding strand and "+" to the coding strand

20nt target (including PAM site NGG)	Gene targeted (ID Maize Genome V3)	Strategy used [a]	Number of transfor-mation events	Number of events with at least one mutation	Mutat-ion efficie-ncy [b]	Number of transformati-on events with bi-allelic mutation	Promot-er driving the sgRNA	%GC witin the 20nt targeted	Chromos-ome [c]	DNA strand [d]
RDP vectors										
GCTGGAGCTCGCCGCCCGAGTGG	GRMZM2G039538	Strategy 2 applied twice to two paralogous genes	8	5	63%	0	TaU6	80	2	-
	GRMZM2G363552			7	88%	0	TaU6	80	7	-
ACCTACGACACCGTCATGCACGG	GRMZM2G039538			2	25%	0	OsU3	55	2	+
	GRMZM2G363552			2	25%	0	OsU3	55	7	+
GACCGGAACGCGACATGGTACGG	GRMZM2G157313	Strategy 1	4	3	75%	0	TaU6	60	10	+
ACAGGTCTACATAGTGTACCTGG	GRMZM2G014499			2	50%	0	OsU3	45	3	+
AGCATTCTGCACCAGGGTATCGG	GRMZM2G059165	Strategy 1	3	1	33%	0	OsU3	50	7	+
GGGGAAGCGAGGGGCTTGCGAGG	GRMZM2G120085			3	100%	1	TaU6	75	1	-
GGAGAAGCACACAAAGCGCGTGG	GRMZM2G145466	Strategy 1	2	2	100%	0	TaU6	60	7	+
AAAAGGTGCTACTGCTGCGTTGG	GRMZM2G573952			1	50%	0	OsU3	50	7	+
ACAGGTCGCACCCGGCAACCTGG	GRMZM2G134341	Strategy 1	3	1	33%	1	OsU3	70	7	-
GCAGACGTGCGACCTGTACCGGG	AC208201.3_FG003			2	67%	2	TaU6	65	1	+
AGTGAGGACTGCAAGAAGGCCGG	GRMZM2G089517	Strategy 1	16	14	88%	0	OsU3	55	5	+
GTGAAGAATCTTCATTCGATTGG	GRMZM2G352274			1	6%	0	TaU6	35	5	-
GAATGGTGCTGTCAAGCGGCCGG	GRMZM2G089517	Strategy 1	5	3	60%	0	TaU6	60	5	+
AACGGACTGCTCCTTGCAGGTGG	GRMZM2G352274			0	0%	0	OsU3	60	5	+
ATGATGAAGATTCAGGGTTCGGG	GRMZM2G049141	Strategy 1	7	3	43%	1	OsU3	40	2	-
GGACGATATACTGCCATCGTCGG	GRMZM2G049141			6	86%	1	TaU6	50	2	-
AAGATGCAAGTCTTGGACGATGG	GRMZM2G140302	Strategy 3	8	6	75%	2	OsU3	45	1	-
	GRMZM2G046086			4	50%	3	OsU3	45	1	-
GGGTCTGGCCCGCTAGGCACAGG	GRMZM2G315601			6	75%	2	TaU6	75	1	-
ACAGCGCGTAGTCGTACACGCGG	GRMZM2G040095	Strategy 1	8	0	0%	0	OsU3	65	2	-
GTCCTCCCGGTACTTGAGCAGGG	GRMZM2G040095			0	0%	0	TaU6	65	2	-
Iowa vectors										
GAATACAGCTGCTCTTGATCCGG	GRMZM2G035701	Strategy 4	2	0	0%	0	ZmU6	45	8	+
GCTGGCTTACATCGCTATTTCGG	GRMZM2G035701	Strategy 4	2	0	0%	0	ZmU6	45	8	-
GGATCAAGCCGGCGGGGTATTGG	GRMZM2G149940			1	50%	0	ZmU6	65	2	-
GGAGACGTCTGTGACGACACCGG	GRMZM2G149940			2	100%	2	ZmU6	60	2	-
GGCCATCGCCTTCGTGCTACAGG	GRMZM2G471240	Strategy 4	2	0	0%	0	ZmU6	65	1	-
GCGGCCTCTCTACGCTGCCAAGG	GRMZM2G471240			2	100%	0	ZmU6	70	1	+
GAGGGTGTCCAGGGTCAACGTGG	GRMZM2G471240			2	100%	0	ZmU6	65	1	+
GGAGACAGGGAGGTACGAACCGG	GRMZM2G471240			2	100%	1	ZmU6	60	1	+

Electronic Supplementary Material

Online Resource 1 List of all CRISPR-Cas9 targets identified in maize B73 genome (v3.26), presented as 23-mers ending with NGG
Online Resource 2 Detection of genome editing in T0 plants
Online Resource 3 List of the primers used in this study
Online Resource 4 List of the most probable off-target for each sgRNA
Online Resource 5 Emission spectrum of the light used for maize culture in growth chambers

Bibliography

Amano E (1968) Comparison of ethyl methanesulfonate- and radiation-induced waxy mutants in maize. Mutat Res Mol Mech Mutagen 5:41–46. doi: 10.1016/0027-5107(68)90079-1

Bibikova M, Beumer K, Trautman JK, Carroll D (2003) Enhancing gene targeting with designed zinc finger nucleases. Science 300:764. doi: 10.1126/science.1079512

Bortesi L, Zhu C, Zischewski J, et al (2016) Patterns of CRISPR/Cas9 activity in plants, animals and microbes. Plant Biotechnol J 14:2203–2216. doi: 10.1111/pbi.12634

Briggs RW, Amano E, Smith HH (1965) Genetic Recombination with Ethyl-methanesulphonate-induced Waxy Mutants in-Maize. Nature 207:890–891. doi: 10.1038/207890a0

Candela H, Hake S (2008) The art and design of genetic screens: maize. Nat Rev Genet 9:192–203. doi: 10.1038/nrg2291

Char SN, Neelakandan AK, Nahampun H, et al (2017) An Agrobacterium-delivered CRISPR/Cas9 system for high-frequency targeted mutagenesis in maize. Plant Biotechnol J 15:257–268. doi: 10.1111/pbi.12611

Char SN, Unger-Wallace E, Frame B, et al (2015) Heritable site-specific mutagenesis using TALENs in maize. Plant Biotechnol J 13:1002–1010. doi: 10.1111/pbi.12344

Choulika A, Perrin A, Dujon B, Nicolas JF (1994) The yeast I-Sce I meganuclease induces site-directed chromosomal recombination in mammalian cells. C R Acad Sci III 317:1013–1019

Christensen AH, Quail PH (1996) Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. Transgenic Res 5:213–218

Christian M, Cermak T, Doyle EL, et al (2010) Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186:757–761. doi: 10.1534/genetics.110.120717

Cigan A., Gadlage MJ, Gao H, et al (2017) Waxy corn. p p.1-61

Clarke R, Heler R, MacDougall MS, et al (2018) Enhanced Bacterial Immunity and Mammalian Genome Editing via RNA-Polymerase-Mediated Dislodging of Cas9 from Double-Strand DNA Breaks. Mol Cell 71:42-55.e8. doi: 10.1016/j.molcel.2018.06.005

Collins GN (1909) A new type of Indian corn from China. US Dept Agric Cur Plant Indust Bull 1–30

Cong L, Ran FA, Cox D, et al (2013) Multiplex Genome Engineering Using CRISPR/Cas Systems. Science 339:819–823. doi: 10.1126/science.1231143

Doll NM, Depège-Fargeix N, Rogowsky PM, Widiez T (2017) Signaling in Early Maize Kernel Development. Mol Plant 10:375–388. doi: 10.1016/j.molp.2017.01.008

Feng C, Yuan J, Wang R, et al (2016) Efficient Targeted Genome Modification in Maize Using CRISPR/Cas9 System. J Genet Genomics Yi Chuan Xue Bao 43:37–43. doi: 10.1016/j.jgg.2015.10.002

Gantner J, Ordon J, Ilse T, et al (2018) Peripheral infrastructure vectors and an extended set of plant parts for the Modular Cloning system. PloS One 13:e0197185. doi: 10.1371/journal.pone.0197185

Gao H, Smith J, Yang M, et al (2010) Heritable targeted mutagenesis in maize using a designed endonuclease. Plant J 61:176–187. doi: 10.1111/j.1365-313X.2009.04041.x

Gerdes JT, Tracy WF (1993) Pedigree Diversity within the Lancaster Surecrop Heterotic Group of Maize. Crop Sci 33:334–337. doi: 10.2135/cropsci1993.0011183X003300020025x

Gilles LM, Khaled A, Laffaire J-B, et al (2017) Loss of pollen-specific phospholipase NOT LIKE DAD triggers gynogenesis in maize. EMBO J 36:707–717. doi: 10.15252/embj.201796603

Haeussler M, Schönig K, Eckert H, et al (2016) Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome Biol 17:148. doi: 10.1186/s13059-016-1012-2

Hsu PD, Scott DA, Weinstein JA, et al (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol 31:827–832. doi: 10.1038/nbt.2647

Ishida Y, Hiei Y, Komari T (2007) *Agrobacterium*-mediated transformation of maize. Nat Protoc 2:1614–1621. doi: 10.1038/nprot.2007.241

Ishida Y, Saito H, Ohta S, et al (1996) High efficiency transformation of maize (Zea mays L.) mediated by Agrobacterium tumefaciens. Nat Biotechnol 14:745–750. doi: 10.1038/nbt0696-745
Jensen KT, Fløe L, Petersen TS, et al (2017) Chromatin accessibility and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency. FEBS Lett 591:1892–1901. doi: 10.1002/1873-3468.12707
Jinek M, Chylinski K, Fonfara I, et al (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337:816–821. doi: 10.1126/science.1225829
Kosicki M, Tomberg K, Bradley A (2018) Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. Nat Biotechnol 36:765–771. doi: 10.1038/nbt.4192
Labuhn M, Adams FF, Ng M, et al (2018) Refined sgRNA efficacy prediction improves large- and small-scale CRISPR-Cas9 applications. Nucleic Acids Res 46:1375–1385. doi: 10.1093/nar/gkx1268
Lee K, Zhang Y, Kleinstiver BP, et al (2018) Activities and specificities of CRISPR/Cas9 and Cas12a nucleases for targeted mutagenesis in maize. Plant Biotechnol J. doi: 10.1111/pbi.12982
Lei Y, Lu L, Liu H-Y, et al (2014) CRISPR-P: A Web Tool for Synthetic Single-Guide RNA Design of CRISPR-System in Plants. Mol Plant 7:1494–1496. doi: 10.1093/mp/ssu044
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. doi: 10.1016/j.jgg.2013.12.001
Lieber MR (2010) The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem 79:181–211. doi: 10.1146/annurev.biochem.052308.093131
Lin Y, Cradick TJ, Brown MT, et al (2014) CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. Nucleic Acids Res 42:7473–7485. doi: 10.1093/nar/gku402
Liu X, Homma A, Sayadi J, et al (2016) Sequence features associated with the cleavage efficiency of CRISPR/Cas9 system. Sci Rep 6:19675. doi: 10.1038/srep19675
Ma X, Zhu Q, Chen Y, Liu Y-G (2016) CRISPR/Cas9 Platforms for Genome Editing in Plants: Developments and Applications. Mol Plant 9:961–974. doi: 10.1016/j.molp.2016.04.009
Mali P, Aach J, Stranges PB, et al (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol 31:833–838. doi: 10.1038/nbt.2675
Marçais G, Kingsford C (2011) A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinforma Oxf Engl 27:764–770. doi: 10.1093/bioinformatics/btr011
McVey M, Lee SE (2008) MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. Trends Genet TIG 24:529–538. doi: 10.1016/j.tig.2008.08.007
Nekrasov V, Staskawicz B, Weigel D, et al (2013) Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. Nat Biotechnol 31:691–693. doi: 10.1038/nbt.2655
Neuffer MG, Sheridan WF (1980) Defective kernel mutants of maize. I. Genetic and lethality studies. Genetics 95:929–944
Opsahl-Ferstad H-G, Deunff EL, Dumas C, Rogowsky PM (1997) ZmEsr, a novel endosperm-specific gene expressed in a restricted region around the maize embryo. Plant J 12:235–246. doi: 10.1046/j.1365-313X.1997.12010235.x
Pattanayak V, Lin S, Guilinger JP, et al (2013) High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat Biotechnol 31:839–843. doi: 10.1038/nbt.2673
Peterson BA, Haak DC, Nishimura MT, et al (2016) Genome-Wide Assessment of Efficiency and Specificity in CRISPR/Cas9 Mediated Multiple Site Targeting in Arabidopsis. PloS One 11:e0162169. doi: 10.1371/journal.pone.0162169
Porteus MH, Carroll D (2005) Gene targeting using zinc finger nucleases. Nat Biotechnol 23:967–973. doi: 10.1038/nbt1125
Puchta H, Fauser F (2014) Synthetic nucleases for genome engineering in plants: prospects for a bright future. Plant J 78:727–741. doi: 10.1111/tpj.12338

Qi W, Zhu T, Tian Z, et al (2016) High-efficiency CRISPR/Cas9 multiplex gene editing using the glycine tRNA-processing system-based strategy in maize. BMC Biotechnol 16:58. doi: 10.1186/s12896-016-0289-2
Ren X, Yang Z, Xu J, et al (2014) Enhanced specificity and efficiency of the CRISPR/Cas9 system with optimized sgRNA parameters in Drosophila. Cell Rep 9:1151–1162. doi: 10.1016/j.celrep.2014.09.044
Schnable PS, Ware D, Fulton RS, et al (2009) The B73 maize genome: complexity, diversity, and dynamics. Science 326:1112–1115. doi: 10.1126/science.1178534
Settles AM, Holding DR, Tan BC, et al (2007) Sequence-indexed mutations in maize using the UniformMu transposon-tagging population. BMC Genomics 8:116. doi: 10.1186/1471-2164-8-116
Shan Q, Wang Y, Li J, et al (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotechnol 31:686–688. doi: 10.1038/nbt.2650
Shi J, Gao H, Wang H, et al (2017) ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. Plant Biotechnol J 15:207–216. doi: 10.1111/pbi.12603
Shukla VK, Doyon Y, Miller JC, et al (2009) Precise genome modification in the crop species Zea mays using zinc-finger nucleases. Nature 459:437–441. doi: 10.1038/nature07992
Shure M, Wessler S, Fedoroff N (1983) Molecular identification and isolation of the Waxy locus in maize. Cell 35:225–233
Siebert R, Puchta H (2002) Efficient repair of genomic double-strand breaks by homologous recombination between directly repeated sequences in the plant genome. Plant Cell 14:1121–1131
Sinha S, Li F, Villarreal D, et al (2017) Microhomology-mediated end joining induces hypermutagenesis at breakpoint junctions. PLoS Genet 13:e1006714. doi: 10.1371/journal.pgen.1006714
Steinert J, Schiml S, Puchta H (2016) Homology-based double-strand break-induced genome engineering in plants. Plant Cell Rep 35:1429–1438. doi: 10.1007/s00299-016-1981-3
Svitashev S, Schwartz C, Lenderts B, et al (2016) Genome editing in maize directed by CRISPR-Cas9 ribonucleoprotein complexes. Nat Commun 7:13274. doi: 10.1038/ncomms13274
Svitashev S, Young JK, Schwartz C, et al (2015) Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA. Plant Physiol 169:931–945. doi: 10.1104/pp.15.00793
Vollbrecht E, Duvick J, Schares JP, et al (2010) Genome-wide distribution of transposed Dissociation elements in maize. Plant Cell 22:1667–1685. doi: 10.1105/tpc.109.073452
Waltz E (2016) CRISPR-edited crops free to enter market, skip regulation. Nat Biotechnol 34:582. doi: 10.1038/nbt0616-582
Wang H, Xu X (2017) Microhomology-mediated end joining: new players join the team. Cell Biosci 7:6. doi: 10.1186/s13578-017-0136-8
Wu X, Scott DA, Kriz AJ, et al (2014) Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat Biotechnol 32:670–676. doi: 10.1038/nbt.2889
Xing H-L, Dong L, Wang Z-P, et al (2014) A CRISPR/Cas9 toolkit for multiplex genome editing in plants. BMC Plant Biol 14:. doi: 10.1186/s12870-014-0327-y
Zhu J, Song N, Sun S, et al (2016) Efficiency and Inheritance of Targeted Mutagenesis in Maize Using CRISPR-Cas9. J Genet Genomics Yi Chuan Xue Bao 43:25–36. doi: 10.1016/j.jgg.2015.10.006
Zuo Z, Liu J (2016) Cas9-catalyzed DNA Cleavage Generates Staggered Ends: Evidence from Molecular Dynamics Simulations. Sci Rep 5:37584. doi: 10.1038/srep37584