Plant Mol Biol (2015) 87:143–156
DOI 10.1007/s11103-014-0266-x

Comparative assessments of CRISPR-Cas nucleases' cleavage efficiency *in planta*

Ross A. Johnson · Vyacheslav Gurevich · Shdema Filler · Aviva Samach · Avraham A. Levy

Received: 7 September 2014 / Accepted: 12 November 2014 / Published online: 18 November 2014

Abstract Custom-designed nucleases can enable precise plant genome editing by catalyzing DNA-breakage at specific targets to stimulate targeted mutagenesis or gene replacement. The CRISPR-Cas system, with its target-specifying RNA molecule to direct the Cas9 nuclease, is a recent addition to existing nucleases that bind and cleave the target through linked protein domains (e.g. TALENs and zinc-finger nucleases). We have conducted a comparative study of these different types of custom-designed nucleases and we have assessed various components of the CRISPR-Cas system. For this purpose, we have adapted our previously reported assay for cleavage-dependent luciferase gene correction in *Nicotiana benthamiana* leaves (Johnson et al. in Plant Mol Biol 82(3):207–221, 2013). We found that cleavage by CRISPR-Cas was more efficient than cleavage of the same target by TALENs. We also compared the cleavage efficiency of the *Streptococcus pyogenes* Cas9 protein based on expression using three different Cas9 gene variants. We found significant differences in cleavage efficiency between these variants, with human and *Arabidopsis thaliana* codon-optimized genes having the highest cleavage efficiencies. We compared the activity of 12 *de novo*-designed single synthetic guide RNA (sgRNA) constructs, and found their cleavage efficiency varied drastically when using the same Cas9 nuclease. Finally, we show that, for one of the targets tested with our assay, we could induce a germinally-transmitted deletion in a repeat array in *A. thaliana*. This work emphasizes the efficiency of the CRISPR-Cas system in plants. It also shows that further work is needed to be able to predict the optimal design of sgRNAs or Cas9 variants.

Keywords CRISPR-Cas · DNA repair · Genome engineering · Non-homologous end-joining · TALENs · Targeted mutagenesis

Introduction

The means to catalyze target sequence-specific DNA breakages is critically important for precise genome editing in higher plant species, both in basic research as well as for agriculture and plant biotechnology (e.g. biopharming) purposes. Until recently, the site-specific endonuclease cleavage of genomes has relied on engineered-meganucleases, as well as protein-chimeras of DNA-binding domains linked to a *FokI* endonuclease DNA-cleavage domain; such nucleases include transcription activator-like effector nuclease (TALEN) enzymes and Zinc-Finger Nuclease (ZFN) enzymes (Voytas 2013; Puchta and Fauser 2014). A new system for target-specific DNA cleavage has emerged based on the *Streptococcus pyogenes* SF390 type II clustered regularly interspaced short palindromic repeats (CRISPR) locus and CRISPR associated factors, termed CRISPR-Cas (Cong et al. 2013; Mali et al. 2013b). The

Electronic supplementary material The online version of this article (doi:10.1007/s11103-014-0266-x) contains supplementary material, which is available to authorized users.

R. A. Johnson · V. Gurevich · S. Filler · A. Samach · A. A. Levy
Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot, Israel

R. A. Johnson (✉)
Department of Plant Biology, University of California, Davis, Davis, CA 95616, USA
e-mail: DrRJohnson@UCDavis.edu

V. Gurevich
Faculty of Medicine in the Galilee, Bar-Ilan University, Safed, Israel

CRISPR-Cas cleavage technology was derived from its use as an adaptive defense against foreign DNA in some bacteria and archaea; there it can convert small portions (protospacers) of invading DNA molecules into genomic arrays to encode RNA molecules, which direct a cleavage defense apparatus against subsequent invasions of DNA that contain these protospacer target sequences (Jinek et al. 2012). CRISPR-Cas is readily site-tailored as a cleavage technology because the specificity of the DNA-cleaving Cas9 enzyme is governed by a separately-encoded single synthetic guide RNA (sgRNA) molecule. To be tailored to a specific site, CRISPR-Cas simply requires the sgRNA's targeting region to be customized by cloning in a complementary-annealed oligonucleotide primer-pair (Cong et al. 2013; Mali et al. 2013b), whereas TALENs and ZFNs, on the other hand, require more-extensive alteration of the nuclease genes to modify their encoded target-specifying protein structure. To find a suitable target site, ZFNs require software-assistance, TALENs have no such limitations (Voytas 2013), whereas CRISPR-Cas requires a simple target sequence, which is $GN_{19}NGG$ when using the archetypal *S. pyogenes*-based technology with an U6 promoter-driven sgRNA (Belhaj et al. 2013). This CRISPR-Cas target site includes a 5′ guide RNA-binding region and a 3′ protospacer-adjacent motif (PAM), which is an NGG (Cong et al. 2013; Mali et al. 2013b) or NAG sequence (Mali et al. 2013a; Jiang et al. 2013a), and affords abundant targets (protospacers) for cleavage.

Targeted mutagenesis by nucleases is a valuable tool in reverse genetics, and has some advantages, in terms of causing permanent and complete loss-of-function, over RNA interference (Puchta and Fauser 2014). Such mutagenesis relies on the catalysis of a double-strand DNA break (DSB), which can trigger repair by error-prone non-homologous end-joining (NHEJ) leading to the deletion or insertion of small sequences adjacent to the cleavage site, sometimes causing a translational reading frame-shifting mutation in a target gene (Feng et al. 2014; Fauser et al. 2014). CRISPR-Cas technology was recently shown to catalyze the cleavage of targets in a variety of plant (as well as metazoan) genomes (Sander and Joung 2014), including *Arabidopsis thaliana*, *Nicotiana benthamiana*, rice, sorghum, wheat (Jiang et al. 2013b; Belhaj et al. 2013), sweet orange (Jia and Wang 2014), tomato (Ron et al. 2014; Brooks et al. 2014), tobacco (Gao et al. 2014) and the liverwort, *Marchantia polymorpha* L. (Sugano et al. 2014). It is difficult to compare efficiencies between the different studies on custom-designed nucleases as different targets were tested, in different species and the quantification method varied between the studies (single cells, transient or stable events, with or without pre-selection, etc.) Nevertheless, overall, it seems that CRISPR-Cas is a promising system with there being reports of plants containing targeted mutations that were recovered at frequencies ranging from 2.5 to 92 % (Feng et al. 2013; Fauser et al. 2014; Miao et al. 2013; Shan et al. 2013b). With CRISPR-Cas, the transmission of the mutation to subsequent generations has sometimes occurred at reduced frequencies (Feng et al. 2014; Fauser et al. 2014; Jiang et al. 2014). Such mutant plants were recovered at frequencies of 1.5–63 % using TALENs (Christian et al. 2013; Li et al. 2012b; Shan et al. 2013a; Zhang et al. 2013) and at 33–69 % using ZFNs (Zhang et al. 2010). CRISPR-Cas has been predicted to have a lower specificity than TALENs (Mali et al. 2013a), with cases of off-target cleavage reported in plants (Xie and Yang 2013; Shan et al. 2013b; Zhang et al. 2014), but these off-target effects can potentially be mitigated by designing an sgRNA with a target-binding region that either has a 5′-truncated end (Sander and Joung 2014), or a unique match to its target in its 3′ half (Sander and Joung 2014), as can be assisted using sgRNA design tools (Montague et al. 2014; Lei et al. 2014; Xie et al. 2014). The ability to carry out multiplex cleavage is a critical advantage of CRISPR-Cas over TALENs and ZFNs, as multiple (up to five, as shown in both mouse and zebrafish cells) genes can be disabled in parallel, as simultaneously expressed sgRNA molecules can each form a target-cleaving complex with a Cas9 protein (Sander and Joung 2014). Moreover, the deletion of a genomic fragment is possible by mediating cleavage in two neighboring targets as was shown by Li et al. (2013) in *Arabidopsis* mesophyll protoplasts who found deletions of up to 48 bp using a transient, sequence-based analysis. Subsequently, in *Arabidopsis* plants, genomic fragment deletions of approximately 160 and 230 bp were germinally-induced within the *AP1* and *TT4* genes, respectively (Feng et al. 2014). Similarly, genomic fragment deletions have been made upon sequences of up to 140 bp in tomato (Brooks et al. 2014) and 1.8 kb in tobacco (Gao et al. 2014). Large chromosomal fragments including several whole genes have been deleted in the rice genome, with fragment sizes ranging from 115 to 245 kb (Zhou et al. 2014). Nucleases specialized to catalyze a single-stranded DNA break only, called nickases, have shown promising potential to increase HR without also triggering NHEJ (Fauser et al. 2014); paired CRISPR-Cas nickases can be used to enhance the specificity of a single nuclease, by separately inducing two adjacent single-strand DNA breaks (forming a DSB, with single-stranded overhangs), in order to stimulate targeted gene replacement with less risk of off-target cleavage (Schiml et al. 2014).

We reason that the differences in mutagenesis frequencies reported within studies could relate to differences in the cleavage efficiency of the sgRNA. The sgRNA efficiency values can vary greatly even between sgRNAs recognizing sites that are adjacent in the chromosome, and so would be expected to have similar accessibility with respect

to the chromatin (Wang et al. 2014a). In addition, we reason that the general differences in mutagenesis frequencies between studies could relate to differences in the form of Cas9 gene used, for example, their codon optimization (Belhaj et al. 2013). Studies of CRISPR-Cas in plants have either used a human codon-optimized Cas9 gene (Nekrasov et al. 2013; Xie and Yang 2013; Jia and Wang 2014; Sugano et al. 2014; Feng et al. 2013; Zhang et al. 2014), a general plant codon-optimized Cas9 gene (Shan et al. 2013b; Li et al. 2013), a rice codon-optimized Cas9 gene (Miao et al. 2013; Jiang et al. 2013b), as well as Cas9 genes optimized to *Chlamydomonas reinhardtii* and monocotyledonous plants (Jiang et al. 2013b), in addition to the wild type *S. pyogenes* Cas9 gene (Jiang et al. 2013b). A previous study found a comparable cleavage efficiency between the native *S. pyogenes* Cas9 and one construct that was codon-optimized towards rice (Jiang et al. 2013b). Their continued work found that their rice codon-optimized Cas9 gene variant was more efficient for mutagenesis in rice protoplasts than genes with bacterial, human or *C. reinhardtii* codon-optimizations (Zhou et al. 2014), some of which had different usage of nuclear localization and FLAG-tag sequences. Another study, also in rice, found that a plant codon-optimized Cas9 construct had a greater activity than a human one (Xu et al. 2014). The ranging success of data reported for CRISPR-Cas *in planta* shows that the optimization of CRISPR-Cas components is warranted in order to obtain mutagenesis at predictable frequencies or the stimulation of inherently more rare events, such as biallelic mutagenesis (Miao et al. 2013), genomic fragment deletion (Li et al. 2013; Feng et al. 2014) and inversion (Gao et al. 2014), NHEJ-mediated knock-in (Wang et al. 2014b), and targeted homologous recombination (Shan et al. 2013b; Li et al. 2013; Fauser et al. 2014; Feng et al. 2014) or gene replacement, at routinely detectable levels. As previously stated by Belhaj et al. (2013), side-by-side comparisons are needed to determine optimal CRISPR-Cas components, such as the form of Cas9 gene used, for cleavage *in planta*.

Here we present a comparative study of the CRISPR-Cas system versus other custom-designed nucleases. We also show a comparison of the cleavage activities of different *S. pyogenes* Cas9 gene variants and twelve different sgRNA sequences. In addition to these studies of Cas9 variants, we have adapted the transient assay recently published by Johnson et al. (2013) to perform a rapid assessment of *de novo*-designed sgRNAs prior to instigating long-term studies in transgenic plants. We conducted preliminary testing of CRISPR-Cas nucleases that were *de novo*-developed to cleave target sequences in the tomato (*Solanum lycopersicum*) phytoene synthase-encoding gene 1, (*SlPSY1*) and the carotene cis–trans-isomerase-encoding gene (*SlCrtISO*) targets (Kachanovsky et al. 2012; Isaacson et al. 2002). Finally, we also show the germinally-transmitted genomic fragment deletion in part of the *Arabidopsis thaliana CRUCIFERIN 3* (*AtCRU3*) gene. The naturally repeated target sites of our sgRNA allowed us to assay for the deletion of a genomic fragment, albeit at a low frequency.

Materials and methods

Reporter constructs

To compare the cleavage efficiencies of each cleavage technology, targets previously found to be cleaved by the QQR ZFN as well as TALENs recognizing the *A. thaliana CRUCIFERIN3* (*AtCRU3*) gene (Johnson et al. 2013; Shaked et al. 2005; Even-Faitelson et al. 2011), were used to develop CRISPR-Cas nucleases. The pGreenII Tfs-494::*LUC* vector, which was previously-described to report cleavage by the *AtCRU3* T494 TALENs, allowed CRISPR-Cas nucleases to be designed in its target sequence. *LUC*-based reporter targets for the 'QQR' ZFN and previously reported T852 TALENs (Johnson et al. 2013) were expanded to allow recognition by the CRISPR-Cas nucleases. These new Tfs-qqr-t and Tfs-852-t sites have 'fs' referring to extra base-pairs creating a 1 bp background signal-cancelling frameshift mutation (Johnson et al. 2013) and 't' referring to an additional triplet(s)-some reporter constructs necessitated the inclusion of additional base-pairs, and/or the use of reverse complementary targets, to shift additional start and stop codons to appropriate positions in the translational reading frame. These reporter constructs were made by cloning a T4 Polynucleotide Kinase-treated (New England BioLabs, Ipswich, MA) complementary-annealed oligonucleotide primer-pair (see Online Resource 1 for all primer sequences) into the *Xba*I and *Xho*I sites of pGreenII Tqqr-sc::*LUC* (Johnson et al. 2013). All restriction endonucleases were sourced from New England BioLabs, Ipswich, MA. Additional sites for the Tfs-qqr-t, Tfs-852-t, Tfs-psy-1, Tfs-psy-2, Tfs-psy-4, Tfs-psy-5, Tfs-crtiso-3, Tfs-crtiso-4, Tfs-crtiso-5 and Tfs-3m targets were prepared in the same way. These vectors were detected with PCR colony screening, using RAJ-198 and the reverse primer of the target site. These reporter constructs were partially-sequenced using RAJ-365 and RAJ-417. The pGreenII 0000 (No *LUC*) and pGreenII 0579-1 (35S::*LUC*) vectors were used to quantify background LUC activity and constitutive LUC activity, respectively, in *Agro*-infiltrated leaves (Johnson et al. 2013). These plant transformation vectors were delivered into *Agrobacterium tumefaciens* strain *LBA4404* cells along with the pSoup 0800 (35S::*REN*) vector (Johnson et al. 2013) using electroporation. The resulting strains were verified by extraction of their plasmid DNA, and propagation in *Escherichia coli*. Verification of pGreenII plasmid DNA used sequencing with RAJ-198 and -365, and pSoup 0800 (35S::*REN*) used digestion with *Xho*I and *Stu*I.

CRISPR-Cas constructs

An *S. pyogenes* CRISPR-associated (Cas9) gene was gratefully obtained from the laboratory of Prof. Holger Puchta at Karlsruhe Institute of Technology (Fauser et al. 2014). This '*At*eCas9' construct had been codon-optimized for expression in *A. thaliana* plants and was amplified with RAJ-629 and -630 using Phusion DNA polymerase (New England BioLabs, Ipswich, MA). The *At*eCas9 gene was then cloned into pDONR 221 using Gateway® technology (Life Technologies, Inc; Carlsbad, CA). This Cas9 gene variant (and others that we refer to subsequently) was used as an intron-less construct (Nekrasov et al. 2013; Xie and Yang 2013; Feng et al. 2013; Zhang et al. 2014; Miao et al. 2013; Jiang et al. 2013b; Shan et al. 2013b). The resulting entry clone was then sequenced using RAJ-639, -640, -641, -642, and -643 as well as Cas9R01. The *At*eCas9 gene was then sub-cloned into pHEX2 using Gateway® technology (Life Technologies, Inc; Carlsbad, CA). The resulting clone was checked with a *Sac*I-*Hpa*I digest, and sequenced again using the same aforementioned primers. An sgRNA expression construct was synthesized from Biomatik (Cambridge, ON) with a promoter from the U6-26 small nuclear RNA gene (Fauser et al. 2014). The site for inserting the protospacer was made using *Bbs*I sites (Cong et al. 2013), with the remainder of the sequence being derived from Mali et al. (2013b). The sequence had *Avr*II and *Sac*I sites positioned outside it. The scaffold construct for sgRNA expression, pBMH U626 RNA Chimera, was ordered as a synthesized construct from Biomatik (Cambridge, ON). Target-binding sequences were cloned into the *Bbs*I sites of this vector as a T4 Polynucleotide Kinase-treated (New England BioLabs, Ipswich, MA) complementary-annealed oligonucleotide primer-pair for the CRISPR-Cas target-binding sequences in the Tqqr (#1 sgRNA; RAJ-648 and -649; #3 sgRNA, RAJ-652 and -653), '3m' (3m-S and -AS), T494 (#1 sgRNA, RAJ-633 and -634, and; #2 sgRNA, RAJ-635 and -636), T852 (RAJ-637 and -638), *PSY1* #1 sgRNA (PSY1 TARGET1 F and R), *PSY1* #2 sgRNA (PSY1 TARGET2 F and R), *PSY1* #4 sgRNA (PSY1 TARGET4 F and R), *PSY1* #5 sgRNA (*PSY*1 target 5, RAJ-658 and -659), *CrtISO* #3 sgRNA (CRTISO TARGET3 F and R), and *CrtISO* #5 sgRNA (RAJ-662 and -663). We could not clone additional sgRNAs designed for Tfs-qqr-t targets 2 and 4, along with several other sgRNAs, and have referred to them in Online Resource 2. Target-binding sequences for the Tqqrt #1 and #2, 3m, T494 #1 and #2, T852, *PSY1* #1, #2, #4, #5, *CrtISO* #3 and #5 sgRNAs were cloned into the *Bbs*I sites of this vector as a kinase-treated complementary-annealed oligonucleotide primer-pair. A further 6 sgRNAs were developed but proved problematic to clone; they are referred to in Online Resource 2. The sgRNA constructs were verified by sequencing, generally using RAJ-319 and RAJ-197, but in the case of the '3m' sgRNA, sequencing used the M13-reverse primer. The sgRNA expression constructs were excised from the *Sac*I sites of these pBMH vectors, and cloned into the same sites of pHEX2 *At*eCas9 and/or their respective pGreenII reporter vectors. For cloning sgRNAs into pHex2 *At*eCas9, PCR screening employed the reverse primer used to create the target-binding sequence in the sgRNA, as well as RAJ-374. Follow up PCR screening was done with Cas9R01 and the forward primer that created the sgRNA's target-binding sequence. The resulting clones were verified by sequencing using RAJ-374, and in some cases, RAJ-731. This DNA was then used to transform *Agrobacterium tumefaciens* strain LBA4404, with the resulting strains being verified by sequencing extracted DNA, after it was propagated in *E. coli*, with RAJ-374. As an alternative to delivering sgRNAs with pHEX2 *At*eCas9, sgRNAs were also cloned into the pGreenII *LUC*-based cleavage reporter vector, as facilitated by the unique *Sac*I site present on this vector. The resulting clones were verified by PCR colony screening with RAJ-198, and the reverse primer that was used to create the target-binding sequence in the sgRNA. The clones were verified by diagnostic *Sac*I restriction enzyme digestion, and by sequencing using RAJ-198 and -365.

Additional Cas9 gene variants analyzed

A translational enhancer-less version of the aforementioned *At*eCas9 gene (Fauser et al. 2014) was PCR amplified using RAJ-702 and -630, and cloned into pDONR 221 using Gateway® technology (Life Technologies, Inc; Carlsbad, CA) with putative clones verified using PCR colony screening using RAJ-643 and -320, and sequencing using RAJ-319, -320, -639, -640, -641, -642, and -643. This '*At*Cas9' gene was then sub-cloned into pHEX2 using Gateway® technology (Life Technologies, Inc; Carlsbad, CA), with putative clones verified by PCR colony screening with RAJ-643 and -246, as well as sequencing using RAJ-246, -417, -639, -640, -641, -642, and -643.

The 'hCas9' gene was cloned from pK7WGF2::hCas9 (Nekrasov et al. 2013) into pDONR 221 using Gateway® technology (Life Technologies, Inc; Carlsbad, CA). Putative clones of the pENTR hCas9 vector were verified by PCR colony screening using RAJ-707 and -320, followed by sequencing using RAJ-319, -320, -705, 706, and -707. The gene was then sub-cloned into the pHEX2 vector, with verification using PCR colony screening with RAJ-707 and -246, then sequencing using RAJ-246, 417, -705, 706, and -707.

A Cas9 gene was kindly provided from the laboratory of Prof. Daniel Voytas at the University of Minnesota. This 'DPCas9' gene was amplified using RAJ-645 and -646, and cloned into pDONR 221 using Gateway® technology (Life

Technologies, Inc; Carlsbad, CA). Putative clones were verified by PCR colony screening using primers RAJ-698 and -699, and sequencing using RAJ-319, -320, -644, 647, -698, and -699. The gene was then sub-cloned into pHEX2 with putative clones being verified by PCR colony screening using RAJ-417 and Cas9R01, then sequencing using RAJ-246, -417, -644, -647, -698, and -699.

Assaying for nuclease cleavage efficiencies

Assays for transient dual-luciferase® (Promega, Madison, WI) reporter expression activity were conducted in *N. benthamiana* leaves after they were infiltrated with *A. tumefaciens* ('*Agro*-Infiltrated'), according to a previous study (Johnson et al. 2013). Some incremental changes were made, including using *N. benthamiana* plants of a larger size, assaying for luciferase activity in a more concentrated buffered leaf extract, and making sure the infiltration buffer was at pH 5.25. We did not carry out Cas9 only (no-sgRNA) controls due to studies reporting no mutagenesis when Cas9 was expressed alone (Li et al. 2013; Nekrasov et al. 2013). This experimental work and the other research that we report were done in plants without impacting on human and animal rights. While these transient assays involved the delivery of three T-DNA constructs, in order for the cleavage-based restoration of *LUC* gene expression to occur, only two of these T-DNA constructs (those based on Fig. 1A, C) were required to enter the same *N. benthamiana* leaf cell. The other T-DNA construct, which served as the transformation standard (Fig. 1B), could normalize the aforementioned signal by expression within a *N. benthamiana* host cell even if it entered in the absence of the other T-DNAs.

Mutagenesis testing in stably transgenic *A. thaliana* plants

The pFZ19 vector (Zhang et al. 2010) was used to clone the *At*eCas9 nuclease for its estradiol-inducible expression. This cloning used the aforementioned entry clone of the *At*eCas9 gene construct as part of Gateway® technology (Life Technologies, Inc; Carlsbad, CA). As both recombination substrates were kanamycin-resistant plasmids, in order to obtain the expression vector product, this cloning used the *At*eCas9 gene entry clone as a *Pvu*II-linearized fragment. The '3m' sgRNA expression construct was then cloned as a second expression cassette into the pFZ19 *At*eCas9 plasmid using the steps described here: firstly, the '3m' sgRNA construct was cloned using *Sac*I sites into the pBluescript II KS vector; Secondly, to flank the sgRNA construct with *Kpn*I sites, this construct was amplified using the oligonucleotide primers, gRNA-KpnI-F and -R, and Phusion DNA polymerase (New England BioLabs, Ipswich, MA), and; Thirdly, this amplification product was sub-cloned into the pFZ19 *At*eCas9 vector using *Kpn*I sites. Clones were verified by PCR colony screening using RAJ-643 and 3m-AS, and predictive *Nco*I, *Bst*XI, *Xba*I, and *Spe*I + *Xmn*I restriction enzyme digestion. The resulting *At*eCas9 + '3m' sgRNA vector was used to transform *A. tumefaciens* strain GV3101, with the correct resulting strain being verified by predictive restriction enzyme digestion (using *Xba*I and *Nco*I, respectively) of its extracted DNA, after it was propagated in *E. coli*. This strain was then used to stably transform wild-type *A. thaliana* (Columbia) plants using the floral-dip technique (Clough and Bent 1998). Transformed seeds were sown on Murashige-Skoog plates containing 25 mg/L hygromycin. To directly verify the specific estradiol-induction of the *At*eCas9 nuclease, in addition to '3m' sgRNA (constitutive) expression, a sub-set of seeds were sown with the additional inclusion of 10 μM β-estradiol (Christian et al. 2013), and compared with a estradiol-free control group. A PCR-based analysis of '3m' sgRNA/*At*eCas9-mediated cleavage in T1 plants was performed 10 days after germination on both medias. This analysis used PCR across the '3m' sites with Phusion DNA polymerase (New England BioLabs, Ipswich, MA), and the oligonucleotide primers, Cru-for probe-F and 852mutR. This PCR analysis, and the other such analyses done, used the DNA polymerase enzyme manufacturer's instructions with 30 reaction cycles. The template for these reactions was DNA extracted from 5 whole 10-day-old T1 seedlings (pooled together) using a previously reported protocol (Shaked et al. 2005). Such sampling was used to compare between induced and not-induced nuclease expression (Fig. 7B). One of the amplified fragments was sequenced (Fig. 7E), after prior cloning into pGEM® T-easy (Promega, Madison, WI). To analyze for the induction of genomic fragment deletions in the *A. thaliana* germline after estradiol-induction during T1 germination, we took 7 independent induced T1 transformants to analyze for transmission of the mutation to the T2 generation (which was not exposed to estradiol itself). Firstly, a preliminary screening analysis for the presence of mutant fragments was done using enrichment PCR. Samples of approximately 1000 T2 seeds from each of these seven T1 plant lines were taken for DNA isolation, using the method described herein. The above seed samples were homogenized in liquid nitrogen using a micro-centrifuge tube and pestle. An extraction buffer, comprising 120 μl of 1 M Tris (adjusted, using HCl, to pH = 9.5; Bio-Lab Ltd; Jerusalem, Israel), 45 μl of 4 M NaCl (Bio-Lab Ltd; Jerusalem, Israel), 240 μl of N-Lauroylsarcosine sodium salt (5 %, Sigma-Aldrich, St. Louis, MO), 6 μl of 2-mercaptoethanol (0.5 %, Sigma-Aldrich, St. Louis, MO), and double-distilled water up to 1.2 ml, was then added per sample, and mixed for 5 min. The samples were then pelleted by centrifugation at 11,000*g* for 5 min at room temperature.

The supernatant was then isolated, then had 0.5 ml of chloroform added, and then was vortex-mixed for 2 min. A 0.5 ml volume of Tris buffer saturated phenol (pH 8, Sigma-Aldrich, St. Louis, MO) was added, and the sample was vortex-mixed for another 2 min. The sample was then pelleted by centrifugation at the above speed and temperature for 15 min. In a fresh tube, the supernatant was then mixed with 90 µl of NaAc (3 M, pH 5.2) and 600 µl of isopropanol, then incubated at room temperature for at least 10 min. The samples were then pelleted by centrifugation at the above speed and temperature for 10 min. The pellet was then washed twice, each time by adding 1 ml of 75 % ethanol, vortex mixing, and then pelleting by centrifugation at 11,000*g* for 5 min at room temperature. The pellet was then allowed to dry, dissolved in 100 µl of double-distilled water containing 10 µg/ml RNase A (Sigma-Aldrich, St. Louis, MO), and incubated at 37 °C for 30 min. Ten µl of NaAc (3 M, pH 5.2, Sigma-Aldrich, St. Louis, MO) and 250 µl of ethanol were then added, and the sample was then incubated at room temperature for 10 min. The samples were then pelleted by centrifugation at the above speed and temperature for 15 min. The pellet was then was then washed using the same ethanol-based procedure described above. The DNA pellet was then dried and dissolved in double-distilled water in preparation for PCR analysis of the *AtCRU3* '3m' locus. In order to increase the PCR detection of rare mutant alleles potentially among a potentially much larger background of wild-type (WT) alleles, the DNA was pre-treated with *Bbs*I, a restriction enzyme capable of cleaving the WT locus template in the PCR amplification. We used *Bbs*I to cleave our DNA as part of a 20 µl volume reaction in Buffer 2.1 (supplied by the manufacturer) overnight at 37 °C. One µl of the cleavage reaction was added as a template for PCR with REDtaq® DNA polymerase (Sigma-Aldrich, St. Louis, MO) according to the manufacturer's instructions for 30 cycles, in a total reaction volume of 25 µl. We detected germinal genomic fragment deletion events (apparently in a high ratio relative to WT alleles) in 6/7 lines, with approximately one thousand T2 seeds analyzed from each T1 line (Fig. 7C). A control reaction was carried out using the DNA extracted from a pool of induced T1 seedlings (see Fig. 7B) to serve as a marker for the sizes of deletion products. For a more accurate analysis of the mutation frequency, individual genotyping of seeds was carried out on a pool of seeds, comprised from 20 mg of seeds from each of these PCR-positive pools (Fig. 7C, lanes 1–6, where only part of each group was analyzed): the T2 progeny plants were individually genotyped as 10-day-old seedlings, with the PCR being conducted using REDtaq® DNA polymerase (Sigma-Aldrich, St. Louis, MO) as described before (including the T1 control), except with template DNA provided in the form of <1 mm^2 leaf discs supplied directly into the PCR reaction. One plant was found with a mutant allele, in addition to a wild-type *AtCRU3* allele (the plant was a heterozygote), in a background of 253 WT plants (Fig. 7D). The sequence-based analysis of this mutation (Fig. 7E) was performed according to the procedure described previously. The T3 progeny of this T2 mutant plant was analyzed for segregation (Fig. 7F) according to the same criteria described above, which used the DNA extraction procedure described in part b. DNA previously extracted from the T2 parental plant was used as a template for a control amplification to show the correct amplification product sizes of both alleles.

Results

CRISPR-Cas outperformed TALENs in cleavage efficiency assays *in planta*

CRISPR-Cas nucleases were *de novo*-developed and comparatively assayed *in planta* against other nucleases using shared candidate target sites, which were co-delivered as part of the constructs shown in Fig. 1.

The first component (Fig. 1A) is used to report the relative level of cleavage by each system through the activation of luciferase (*LUC*) gene expression from a constitutively-expressed reporter containing a candidate target sequence and a deliberate frameshift mutation (in reporter names, 'T' stands for target, and 'fs' stands for frameshift). The repair of the cleaved target is typically conducted by the error-prone mechanism of NHEJ, which can lead to a restoration of the *LUC* open reading frame for some of these repaired Transfer-DNA (T-DNA) reporter molecules (Johnson et al. 2013). The second component (Fig. 1B) is a constitutively-expressed *Renilla reniformis*-derived luciferase (*REN*) gene, which can be used to normalize the variable transformation frequency in measurements of the cleavage reporter (Johnson et al. 2013); these LUC:REN measurements are made using the dual-luciferase® reporter assay system (Promega, Madison, WI) and allow more accurate quantification than cleavage analyses with single reporter constructs (Jiang et al. 2013b, 2014). The vector is contained in the same *Agrobacterium* strain as the construct described in Fig. 1A, with both T-DNA molecules being capable of separate T-DNA delivery. The third component (Fig. 1C) can express genes encoding a nuclease. Such nucleases include a TALEN or ZFN (Johnson et al. 2013), or, alternatively, a CRISPR-Cas nuclease. A Cas9 nuclease can be delivered, along with a target-specific sgRNA molecule (Fig. 1D), as two separate expression cassettes in the pHEX2 vector. In order to keep these components separate from their DNA target, this vector is delivered as part of a separate *Agrobacterium* strain.

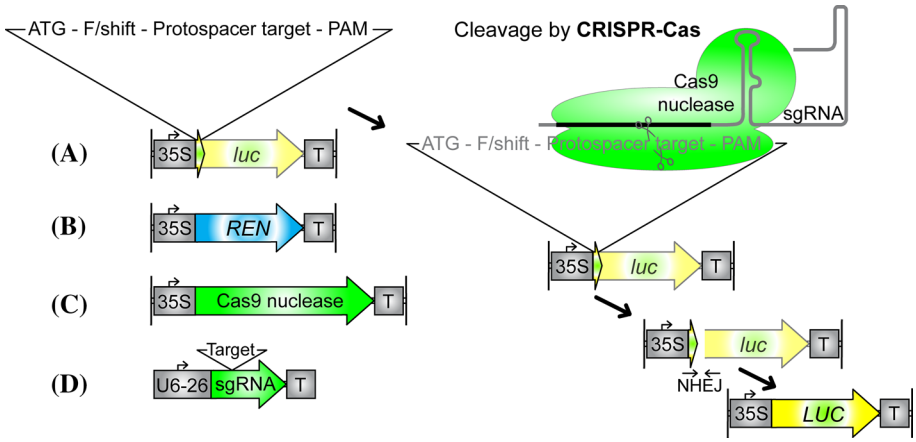

Fig. 1 Constructs used to test CRISPR-Cas cleavage *in planta*, based on the activation of transient luciferase expression. **A** A firefly-derived luciferase (*LUC*)-based reporter vector for nuclease cleavage-induced non-homologous end-joining (NHEJ). The sequence immediately downstream of the ATG start codon is expanded above the construct to show the presence of a frame-shift mutation (F/shift) that prematurely disrupts *LUC*, as well as the CRISPR-Cas target (Protospacer target—PAM). The annealing of the sgRNA construct (shown as a *curved grey line*) based on target sequence-complementarity (shown as a *black line* within the sgRNA) can elicit cleavage of the target by the Cas9 nuclease protein (shown in *green with scissor symbols*), leading to restoration of the *LUC* translational reading-frame as part of error-prone repair by NHEJ; **B** a co-delivered *Renilla reniformis*-derived luciferase (*REN*)-based vector for normalizing the firefly LUC activity; **C** represents a vector for expressing the Cas9 nuclease, or alternative nucleases, and; **D** an sgRNA expression construct, which can be included either as part of the aforementioned vector, or, alternatively, as part of the *LUC*-based reporter vector. '*U6-26*' refers to a U6 (loci 26) small nuclear RNA gene promoter that is trans-activated by a RNA polymerase III, '35S' refers to a constitutively-active cauliflower-mosaic virus promoter sequence and 'T.' refers to a terminator/polyadenylation sequence

To compare the relative cleavage efficiencies of these nucleases, the target sequences of our previously-reported TALENs (Johnson et al. 2013), and the well-established 'QQR' zinc-finger nuclease control (Even-Faitelson et al. 2011; Johnson et al. 2013), were used to *de novo*-develop CRISPR-Cas nucleases, as shown in Fig. 2.

The findings from testing newly-developed CRISPR-Cas components, against the same target sites for previously-described TALENs, as part of the same reporter system, are shown in Fig. 3.

In both test cases, our CRISPR-Cas components performed better than the corresponding TALENs, 'T494' and 'T852', which we previously found to catalyze cleavage *in planta* (Johnson et al. 2013). The T494 #1 and #2 sgRNAs resulted in 83-fold ($p = 1.3 \times 10^{-12}$) and 79-fold ($p = 1.8 \times 10^{-8}$) average 'cleavage efficiency' values. These 'cleavage efficiency' values were determined by the LUC:REN signal measured in the presence of the nuclease divided by that signal obtained for the corresponding negative control made with a non-recombinogenic *GUS* gene instead of a nuclease. This means of calculating the cleavage efficiency was chosen because we observed that the base-line level of expression tended to vary between reported constructs (e.g. Fig. 3B vs. C), presumably due to varying amounts of ribosomal frameshifting/slippage, which would not be accounted for if we were to normalize to the

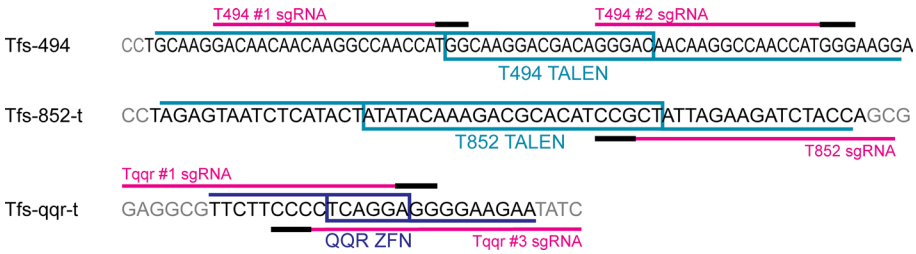

Fig. 2 To compare each cleavage technology, CRISPR-Cas protospacer target sites were located/introduced into the targets of the previously-reported T494 and T852 TALENs, as well as the 'gold standard' QQR ZFN. The aforementioned TALENs were previously reported by us to cleave sequences in the *Arabidopsis thaliana CRUCIFERIN 3* (*AtCRU3*), with their names reflecting their distance downstream of the start codon. The sequence (5′–3′) shown in *black* is required for TALEN or ZFN cleavage, whereas sequence shown in *grey* represents bases inserted to allow cleavage by CRISPR-Cas and/or to appropriately position start and stop codons in the reporter. *Thin magenta lines* represent protospacer targets and *thick black lines* represent PAM sequences

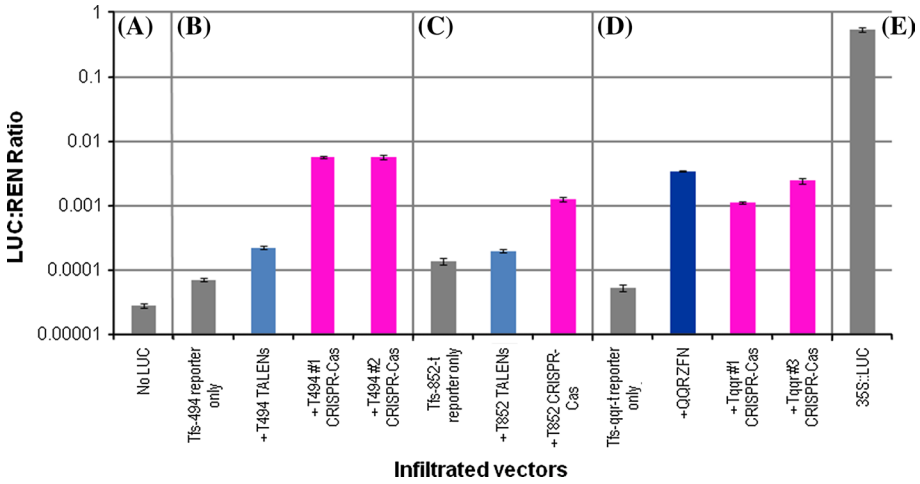

Fig. 3 Comparing the cleavage efficiency of CRISPR-Cas technology with that of TALENs and the QQR ZFN *in planta*. Average normalized LUC activity (LUC:REN ratio) assayed from a series of constructs designed to compare cleavage by CRISPR-Cas components (*pink bars*), TALENs (*light blue bars)* and the QQR ZFN reference (*dark blue bars*), following NHEJ DSB repair in *Agro*-infiltrated *N. benthamiana* leaves. Negative controls are shown as *grey bars* for each target site. The following reporter constructs were used: **A** a negative control (No *LUC*); Nuclease cleavage reporter constructs embedded with target sites, including **B** Tfs-494, **C** Tfs-852-t, and **D** Tfs-qqr-t sites, which were delivered either without or together with their respective TALEN, CRISPR-Cas nuclease, or ZFN constructs, as well as; **E** A positive control (35S::*LUC*) reporter construct. All data shown has been combined together from two separate experiments. There were seven replicas in each experiment. *Error bars* corresponding to the standard error of two experiments, are not always visible due to their small size

no *LUC* or 35S::*LUC* controls. The cleavage efficiency of these CRISPR-Cas components was clearly higher than the T494 TALEN pair, which had a 3.4-fold ($p = 1.0 \times 10^{-10}$) average cleavage efficiency (Fig. 3B). Similarly, the T852 sgRNA resulted in an 11-fold ($p = 1.9 \times 10^{-8}$) average cleavage efficiency, which was again better than the T852 TALEN pair, which had a 1.7-fold ($p = 2.8 \times 10^{-3}$) average cleavage efficiency (Fig. 3C). While the cleavage efficiency of the T494 #1 and #2 sgRNAs was higher than the QQR ZFN positive control with a 66-fold ($p = 3.1 \times 10^{-15}$) cleavage efficiency, the QQR ZFN outperformed our Tqqr #1 and #3 CRISPR-Cas components, which had average cleavage efficiencies of 21-fold ($p = 1.2 \times 10^{-12}$) and 48-fold, respectively (Fig. 3D).

The Cas9 gene variants differed in their cleavage efficiencies

We wished to comparatively test three previously-described Cas9 gene variants, which had been codon-optimized to *Arabidopsis thaliana* (*At*Cas9) by Fauser et al. (2014), humans (hCas9) by Mali et al. (2013b), and dicotyledonous plants (DPCas9) by Baltes et al. (2014). A derivative of *At*Cas9 that contained a translational enhancer (*Ate*Cas9) was also tested (Fauser et al. 2014). Excluding the case of the aforementioned enhancer sequence, as well as a 42 amino-acid sequence tag and three additional amino acids located in DPCas9 (in the 5′ and 3′ extremities, respectively), these Cas9 gene variants had the same protein sequence. In order to facilitate a comparison between different Cas9 gene variants, the sgRNA was delivered as part of the reporter construct (Fig. 1, construct D delivered as part of construct A), as opposed to delivery with the Cas9 gene as in Fig. 3, in order to avoid needing to clone the sgRNA with each Cas9 gene variant (this sgRNA configuration and related forms were verified beforehand as shown in Online Resources 3 and 4). The sgRNA that showed the most consistently high signal in assessments made until this point, T494 #1 sgRNA (Fig. 3B) was chosen for comparing between the different Cas9 gene variants. The results of this Cas9 gene testing are shown below in Fig. 4.

The cleavage efficiencies of the three Cas9 gene variants, *At*Cas9, *Ate*Cas9 and hCas9, were comparable and all three statistically significant, while DPCas9 was significantly less efficient than the other three. The highest cleavage-mediated activation of LUC:REN was obtained with hCas9 (Nekrasov et al. 2013; Mali et al. 2013b), which was 113-fold greater than the negative control (Standard Error, SE 2.6, p value 4.3×10^{-10}). *At*Cas9 (Fauser et al. 2014) showed the next highest signal, which was 105-fold over the negative control (SE 4.0, p value 1.5×10^{-9}) and the same gene with a translational enhancer, *Ate*Cas9, had an 89-fold higher signal than the negative control (SE 8.0, p value 3.0×10^{-10}). Similarly to the data presented in Fig. 3, this *Ate*Cas9/T494 #1 combination had a higher cleavage activity than the QQR ZFN positive control, which was measured at 62-fold (SE 16, p value 3.1×10^{-16}). We found that DPCas9 (Baltes et al. 2014)

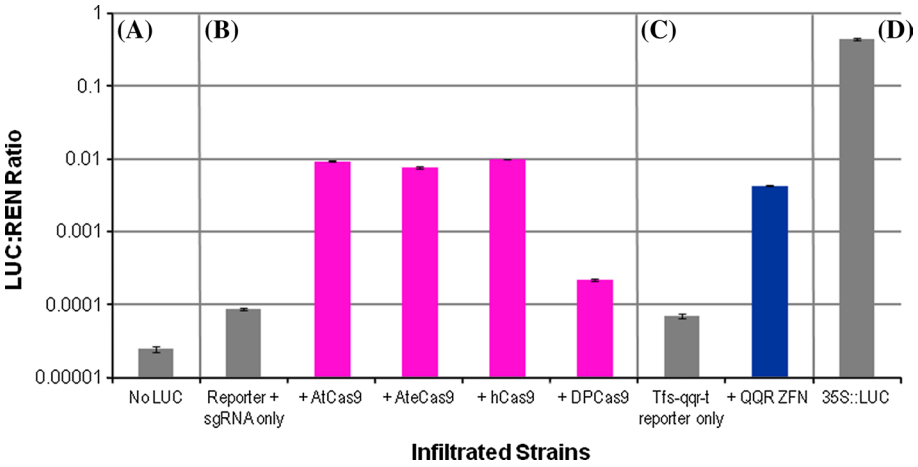

Fig. 4 Comparing the cleavage efficiencies of different *S. pyogenes* Cas9 gene variants. Average normalized LUC activity (LUC:REN ratio) assayed from a series of constructs designed to test cleavage by CRISPR-Cas using a number of Cas9 gene variants following NHEJ DSB repair in *Agro*-infiltrated *N. benthamiana* leaves. The following reporter constructs were used: **A** a negative control (No *LUC*); **B** An Tfs-494 + sgRNA 1 reporter (and sgRNA) construct alone and with a variety of Cas9 genes delivered in a separate vector; **C** Tfs-qqr-t site, which was delivered either without or together with its respective nuclease, and; **D** A positive control (35S::*LUC*) reporter construct. All data shown has been combined together from two separate experiments. *At*Cas9, hCas9 and DPCas9 refer to *Arabidopsis*, human and dicotyledonous plant codon-optimizations, respectively, for these gene variants. *Ate*Cas9 refers to the former gene variant with a translational enhancer. There were seven replicas in each experiment. *Error bars* corresponding to the standard error of two experiments, are not always visible due to their small size

had a 3-fold (SE 0.78, p value 2.5×10^{-13}) cleavage efficiency in our analyses, over an order of magnitude lower than the other Cas9 constructs tested. Another *S. pyogenes* Cas9 gene that was synthesized for our research purposes with codon-optimization to *Solanum lycopersicon* was attempted for inclusion in this comparison; however, this gene could not be routinely propagated in both *E. coli* and *A. tumefaciens* cells. In order to compare the cleavage efficiency of various sgRNAs (including those whose testing was shown in Fig. 3) using a consistent Cas9 gene, subsequent experiments were conducted with this *Ate*Cas9 gene construct.

The cleavage efficiency of preliminarily tested, *de novo*-developed sgRNAs is variable and difficult to predict

The ability of CRISPR-Cas to cleave user-selected target sites within the tomato (*S. lycopersicum*) genome was tested against two genes involved in carotenoid biosynthesis. Targets were chosen as described in Online Resource 5 for the *de novo*-development of sgRNAs for use with CRISPR-Cas nuclease cleavage. The cleavage efficiency of the CRISPR-Cas components using sgRNAs designed to cleave these two tomato (*S. lycopersicum*) carotenoid genes were tested as shown in Fig. 5. Further sgRNAs were tested as shown in Fig. 6, which, to avoid cloning constraints, was done with the separate delivery of Cas9 as opposed to using the co-delivery of Cas9 (Fig. 5), with either basis offering similar cleavage efficiencies (Fig. 5B vs. 6B). For the sgRNAs that we developed for these tomato gene targets, six out of six tested were found to have positive cleavage activities *in planta*. We found that the cleavage efficiencies of these sgRNAs varied markedly (Online Resource 6). Their cleavage efficiencies ranged from 2.2-fold (Fig. 5D, SE 0.40, p value 8.4×10^{-5}) to 120-fold (Fig. 6C, SE 58, p value 1.3×10^{-7}).

Our preliminary nuclease testing could forecast the feasibility of generating a germinally-transmissible genomic fragment deletion in *Arabidopsis* plants

The naturally repeated '3m' target sequences in the *A. thaliana CRU3* gene (Fig. 7A) allowed us to design an effective sgRNA (it had a 30-fold cleavage efficiency, p value 1.4×10^{-11}, Online Resource 7); these targets were then used to assay for a genomic fragment deletion in *A. thaliana* plants. The '3m' sgRNA (constitutive) expression cassette and the *Ate*Cas9 nuclease gene were delivered into plant cells using the pFZ19 vector (Zhang et al. 2010); this vector was used to allow for the controlled timing of *Ate*Cas9 gene expression in transgenic *A. thaliana* plants owing to its estrogen-inducible expression system. Following the induction of our CRISPR-Cas nuclease's expression during germination of the primary transformant (T1) seeds, we found preliminary PCR-based evidence of cleavage within those T1 individuals as nuclease-induced plants (Fig. 7B), and their T2 seed progeny (Fig. 7C). The latter analysis was conducted on pooled seed lines, which were

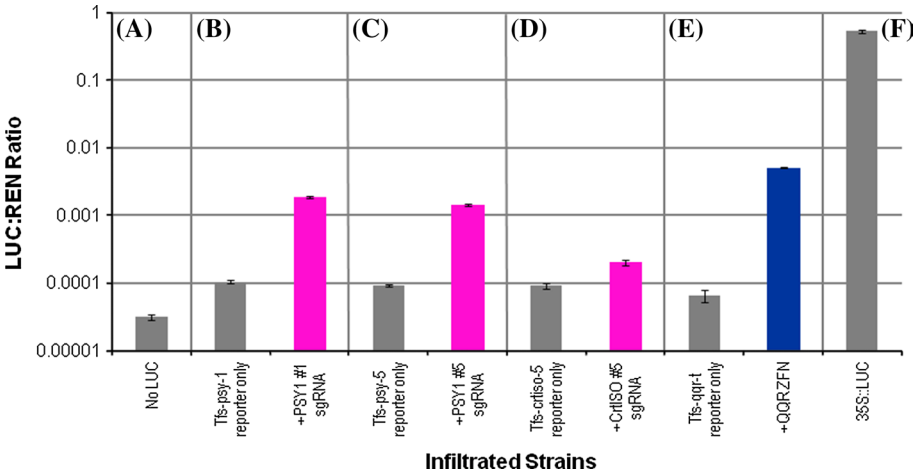

Fig. 5 The cleavage efficiency of CRISPR-Cas sgRNAs designed to cleave target sequences in tomato *PSY1* and *CrtISO* genes. Average normalized LUC activity (LUC:REN ratio) assayed from a series of constructs designed to test cleavage by CRISPR-Cas components (delivered in the same vector) following NHEJ DSB repair in *Agro*-infiltrated *N. benthamiana* leaves. The following reporter constructs were used: **A** a negative control (No *LUC*); Nuclease cleavage reporter constructs embedded with target sites, including **B** Tfs-psy-1, **C** Tfs-psy-5, **D** Tfs-crtiso-5, **E** Tfs-qqr-t sites, which were delivered either without or together with their respective nuclease components, as well as; **F** A positive control (35S::*LUC*) reporter construct. All data shown has been combined together from two separate experiments. There were seven replicas in each experiment. *Error bars* corresponding to the standard error of two experiments, are not always visible due to their small size

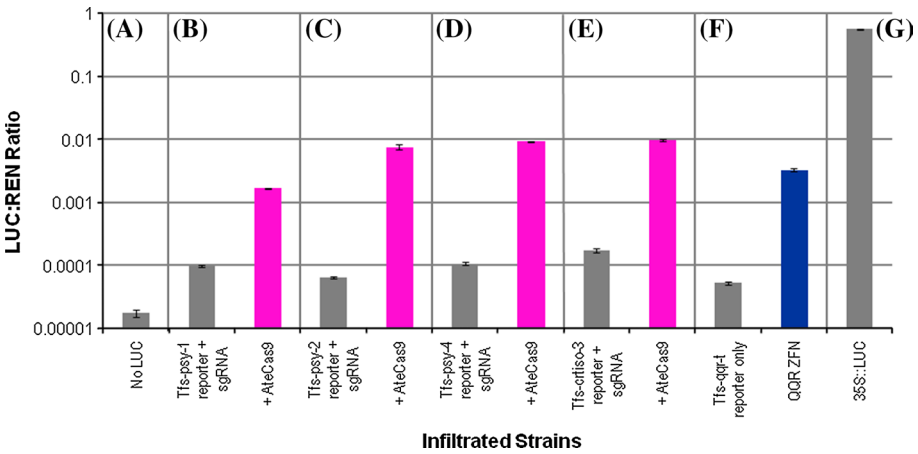

Fig. 6 The cleavage efficiency of more CRISPR-Cas sgRNAs designed to cleave target sequences in tomato *PSY1* and *CrtISO* genes. Average normalized LUC activity (LUC:REN ratio) assayed from a series of constructs designed to test cleavage by CRISPR-Cas components (delivered in different vectors) following NHEJ DSB repair in *Agro*-infiltrated *N. benthamiana* leaves. The following reporter constructs were used: **A** a negative control (No *LUC*); nuclease cleavage reporter constructs embedded with target sites and their corresponding sgRNA construct, including **B** Tfs-psy-1 + sgRNA, **C** Tfs-psy-2 + sgRNA, **D** Tfs-psy-4 + sgRNA, and **E** Tfs-crtiso-3 + sgRNA, which were delivered either without or together with the *Ate*Cas9 gene; **F** the Tfs-qqr-t site, which was delivered without and together with the QQR ZFN, as well as; **G** A positive control (35S::*LUC*) reporter construct. All data shown has been combined together from two separate experiments. There were seven replicas in each experiment. *Error bars* corresponding to the standard error of two experiments, are not always visible due to their small size

not nuclease-induced themselves, and used *Bbs*I restriction enzyme-mediated enrichment PCR to help identify lines for further analysis. By individually genotyping T2 progeny plants (from the seed lines showing preliminary evidence of mutagenesis), we detected a 48 bp genomic fragment deletion that occurred in one out of 253 plants (Fig. 7D). This deletion occurred between the first and second targets of the three *AtCRU3* '3m' repeats (Fig. 7E), and was stably-inherited, as shown by the segregation of this mutation in the T3 generation (Fig. 7F).

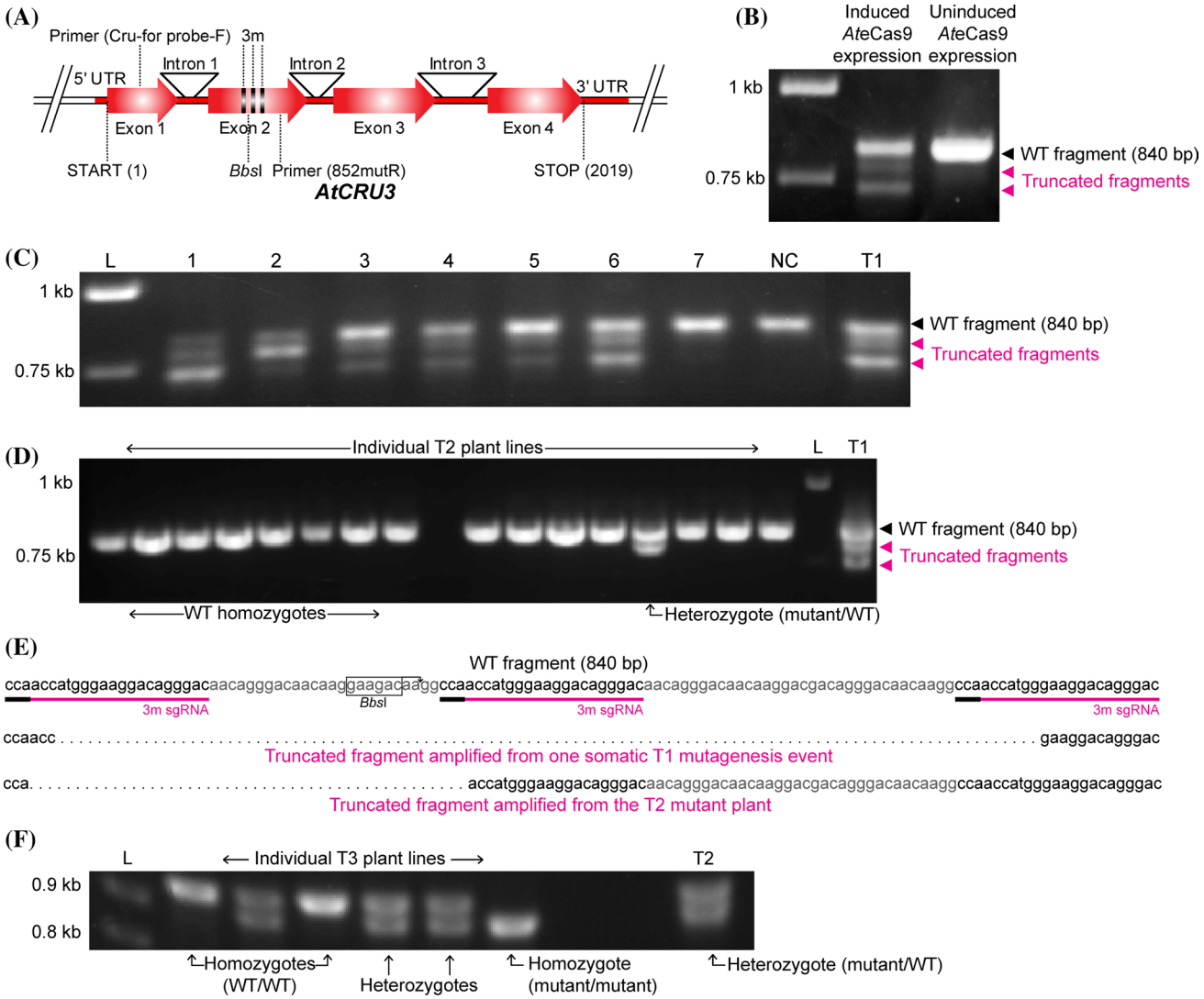

Fig. 7 The stable mutagenesis of the '3m' target sequences that are naturally repeated in the *Arabidopsis thaliana CRU3* gene. **A** The *AtCRU3* gene with the '3m' target sites labeled as black boxes, the exons labeled as *red arrows*, and the introns labeled as *red lines* in between exons. 'Primer (Cru-for probe-F)' and 'Primer (852mutR)' respectively refer to the binding sites of the forward and reverse primers used for the detection of mutagenesis events. **B** The PCR detection of '3m' sgRNA-mediated mutagenesis events in the somatic tissue of T1 seedlings (the expression of *Ate*Cas9 was induced at the T1 seed stage) with five plants pooled together for this analysis. **C** The preliminary detection of germinally-transmitted truncation mutagenesis events in seven T2 seed lines (labels 1–7), on DNA extracted from approximately one thousand T2 seeds, using *Bbs*I restriction enzyme-mediated enrichment PCR. **D** Individual screening for germinally-transmitted truncation mutations in the *AtCRU3* gene within a pool of T2 seeds from the PCR-positive lines (part C, labels 1–6 only); here only 17 of 253 of these genotyping reactions are shown for space reasons. **E** The location of the '3m' sgRNA target sequences and the *Bbs*I site in the wild type (WT) *AtCRU3* gene sequence, as well as the sequences of two truncated fragments, one amplified from a somatic mutagenesis event in T1 plants, the other amplified from a single T2 mutant plant containing an inherited mutagenesis event. **F** The mutant locus was transmitted to the T3 generation, based on our segregation analysis. 'L' stands for a DNA ladder (with relevant band sizes labeled), 'NC' stands for not estradiol-exposed control, and the 'T1' control is a marker for the sizes of deletion products based on the analysis of plants in part B with induced nuclease expression

Discussion

Here we report that CRISPR-Cas out-performed TALENs in a comparative analysis of their cleavage activities *in planta*, for two independent targets (Fig. 3). Moreover, all of the TALENs we have tested were less effective than the gold-standard QQR ZFN (Even-Faitelson et al. 2011; Johnson et al. 2013); while here, several sgRNAs catalyzed cleavage at levels similar and sometimes higher than the QQR ZFN (Figs. 3, 6). Across the variations that we made to CRISPR-Cas nucleases, we found that eleven-of-twelve CRISPR-Cas sgRNAs and two-of-three Cas9 gene variants

had higher cleavage efficiencies than any TALEN we tested (Figs. 4, 5 and 6, vs. 3).

Our findings and previous studies of CRISPR-Cas (Feng et al. 2013; Fauser et al. 2014; Miao et al. 2013; Shan et al. 2013b) and TALENs (Christian et al. 2013; Li et al. 2012b; Shan et al. 2013a; Zhang et al. 2013), suggest that cleavage mediated by CRISPR-Cas might be more generally efficient than cleavage mediated by TALENs in plants. It is important to note though that for other aspects, such as specificity and the lack of target sequence constraints, TALENs might be preferable to CRISPR-Cas (Puchta and Fauser 2014).

We also found that different variants of the *S. pyogenes* Cas9 genes, coding for the same protein, but codon-optimized for different species had very different activities in our analyses. While the human and *Arabidopsis* gene variants both induced high cleavage activity, other Cas9 gene variants were either less efficient or could not be cloned in *E. coli* or *A. tumefaciens*. The reasons for the differences between a number of Cas9 gene variants, as determined by us, Zhou et al. (2014), and by Xu et al. (2014), are not clear. We suggest that further work is needed to optimize Cas9, for aspects such as gene translation efficiency and protein stability, to ensure the highest cleavage efficiency. While previous studies had found that their rice (Zhou et al. 2014) or plant codon-optimized (Xu et al. 2014) Cas9 variants out-performed their humanized counterpart, their hCas9 variant (Cong et al. 2013) was not the same as the one that we tested (Nekrasov et al. 2013; Mali et al. 2013b). We do not have cleavage efficiency data for the hCas9 variant used in these previous studies for comparison, as our analyses involved a different hCas9 variant, which lacked the 5′-localized nuclear localization and FLAG-tag sequences. The lower cleavage activity of Cas9 encoded by the DPCas9 gene variant, relative to the other constructs we tested, might be due to its 5′-located 42 amino-acid sequence tag.

We also report the relative cleavage efficiencies of twelve different *de novo*-developed sgRNAs. These values are compared with features of their sequences that might indicate their performance as suggested in a large-scale study in human cells by Wang et al. (2014a). Previously reported studies conducting stable mutagenesis in plants were also included in this analysis (Online Resource 6). We found that, based on the very limited numbers that we compare to the much larger study of Wang et al. (2014a), there is no clear correlation between the performance of sgRNAs with respect to their performance indicators, both for our work in a transient system, as well as for the work reported by others in stable transgenic systems. Similarly, we did not find that the tool of Doench et al. (2014) could accurately predict sgRNA performance in plant systems. Moreover, we found no correlation between the cleavage activity of our sgRNAs and the strand in which the target site existed as part of the reporter construct, the GC content, or RNA secondary structure. The accurate prediction of sgRNA cleavage performance *in planta* may require analyzing greater numbers of sgRNAs; this analysis might be more readily facilitated by studying their promoter transactivation using artificial transcription factors based on CRISPR-Cas (Farzadfard et al. 2013; Li et al. 2012a), rather than studying their cleavage efficiency. Furthermore, it seems possible that information from the Cas9 crystal structure might be used to assist the prediction of sgRNA performance (Nishimasu et al. 2014).

Here we report the stably-inherited deletion of a genomic fragment by the cleavage of two naturally-repeated target sites, which was mediated by one sgRNA; in contrast, previously-reported genomic fragment deletion experiments (targeting non-repeated sites) have necessitated the use of two different sgRNAs (Li et al. 2013; Feng et al. 2014). The frequency of germinal transmission of our genomic fragment deletion mutation based on '3m' sites, was 1 in 253 plants, which was lower than the 25.8 % frequency of germinally-induced genomic fragment deletion reported in *Arabidopsis* plants by Feng et al. (2014). The reasons for this difference may have been due to Feng et al. (2014) using sgRNAs that were more efficient than our '3m' sgRNA. Additionally, we used an inducible nuclease expression system, rather than the constitutive (2 × 35S) promoter-driven Cas9 construct used by Feng et al. (2014). This difference in frequencies may also relate to the close proximity of our target sites, which were 25 bp apart (with a third site that the nuclease could bind only 34 bp away), whereas the targets were 150–229 bp apart in the work of Feng et al. (2014). Furthermore, due to the aforementioned lack of comparative information, we have to consider that their hCas9 variant, originally derived from Cong et al. (2013), may have allowed more efficient cleavage than the *At*eCas9 gene variant that we employed in this case. Our work shows, for the first time in plants, that CRISPR-Cas nucleases can be used to target multiple sites in a genomic repeat array, inducing a deletion, and reduce the number of repeats.

In summary, while more work is needed to optimize the CRISPR-Cas system in plants, their current performance shows encouraging potential for producing new mutant plant varieties that are regulated as non genetically-modified in the USA and Canada, and potentially in other countries adopting product-based regulatory frameworks for genetically modified plants (Podevin et al. 2012).

Acknowledgments We thank all members of Avraham A. Levy's laboratory at the Weizmann Institute of Science, including Dr. Cathy Melamed-Bessudo for technical advice, and Tal Dahan for her help maintaining *N. benthamiana* plants. We thank Prof. Holger Puchta from Karlsruhe Institute of Technology and Prof. Daniel Voytas at the

University of Minnesota for providing Cas9 gene constructs. Further thanks to Prof. Sophien Kamoun from the Sainsbury Laboratory for providing another Cas9 gene construct. We also thank Dario Breitel and Prof. Asaph Aharoni who developed the *A. thaliana* seed genomic DNA extraction protocol that was used here. This work was funded by an EU-FP7 TRACTAR grant from the European Research Council.

Conflict of interest The authors declare that they have no conflict of interest.

References

Baltes NJ, Gil-Humanes J, Cermak T, Atkins PA, Voytas DF (2014) DNA replicons for plant genome engineering. Plant Cell 26(1):151–163

Belhaj K, Chaparro-Garcia A, Kamoun S, Nekrasov V (2013) Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. Plant Methods 9(1):39

Brooks C, Nekrasov V, Lippman Z, Van Eck J (2014) Efficient gene editing in tomato in the first generation using the CRISPR/Cas9 system. Plant Physiol 166(3):1292–1297

Christian M, Qi Y, Zhang Y, Voytas DF (2013) Targeted mutagenesis of *Arabidopsis thaliana* using engineered TAL effector nucleases. G3 (Bethesda) 3(10):1697–1705

Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J 16(6):735–743

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339(6121):819–823

Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, Sullender M, Ebert BL, Xavier RJ, Root DE (2014) Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol. doi:10.1038/nbt.3026

Even-Faitelson L, Samach A, Melamed-Bessudo C, Avivi-Ragolsky N, Levy AA (2011) Localized egg-cell expression of effector proteins for targeted modification of the *Arabidopsis* genome. Plant J 68(5):929–937

Farzadfard F, Perli SD, Lu TK (2013) Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas. ACS Synth Biol 2(10):604–613

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

Feng Z, Zhang B, Ding W, Liu X, Yang DL, Wei P, Cao F, Zhu S, Zhang F, Mao Y, Zhu JK (2013) Efficient genome editing in plants using a CRISPR/Cas system. Cell Res 23:1229–1232

Feng Z, Mao Y, Xu N, Zhang B, Wei P, Yang D-L, Wang Z, Zhang Z, Zheng R, Yang L, Zeng L, Liu X, Zhu J-K (2014) Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in *Arabidopsis*. Proc Natl Acad Sci 111(12):4632–4637

Gao J, Wang G, Ma S, Xie X, Wu X, Zhang X, Wu Y, Zhao P, Xia Q (2014) CRISPR/Cas9-mediated targeted mutagenesis in *Nicotiana tabacum*. Plant Mol Biol. doi:10.1007/s11103-014-0263-0

Isaacson T, Ronen G, Zamir D, Hirschberg J (2002) Cloning of tangerine from tomato reveals a carotenoid isomerase essential for the production of b-carotene and xanthophylls in plants. Plant Cell Online 14(2):333–342

Jia H, Wang N (2014) Targeted genome editing of sweet orange using Cas9/sgRNA. PLoS One 9(4):e93806

Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA (2013a) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol 31(3):233–239

Jiang W, Zhou H, Bi H, Fromm M, Yang B, Weeks DP (2013b) Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. Nucl Acids Res 41(20):e188

Jiang W, Yang B, Weeks DP (2014) Efficient CRISPR/Cas9-mediated gene editing in *Arabidopsis thaliana* and inheritance of modified genes in the T2 and T3 generations. PLoS One 9(6):e99225

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(6096):816–821

Johnson RA, Gurevich V, Levy AA (2013) A rapid assay to quantify the cleavage efficiency of custom-designed nucleases *in planta*. Plant Mol Biol 82(3):207–221

Kachanovsky DE, Filler S, Isaacson T, Hirschberg J (2012) Epistasis in tomato color mutations involves regulation of phytoene synthase 1 expression by *cis*-carotenoids. Proc Natl Acad Sci 109(46):19021–19026

Lei Y, Lu L, Liu HY, Li S, Xing F, Chen LL (2014) CRISPR-P: a web tool for synthetic single-guide rna design of CRISPR-system in plants. Mol Plant 7(9):1494–1496

Li L, Piatek M, Atef A, Piatek A, Wibowo A, Fang X, Sabir J, Zhu J-K, Mahfouz M (2012a) Rapid and highly efficient construction of TALE-based transcriptional regulators and nucleases for genome modification. Plant Mol Biol 78(4):407–416

Li T, Liu B, Spalding MH, Weeks DP, Yang B (2012b) High-efficiency TALEN-based gene editing produces disease-resistant rice. Nat Biotechnol 30(5):390–392

Li JF, 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(8):688–691

Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, Yang L, Church GM (2013a) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol 31(9):833–838

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM (2013b) RNA-guided human genome engineering via Cas9. Science 339(6121):823–826

Miao J, Guo D, Zhang J, Huang Q, Qin G, Zhang X, Wan J, Gu H, Qu LJ (2013) Targeted mutagenesis in rice using CRISPR-Cas system. Cell Res 23:1233–1236

Montague TG, Cruz JM, Gagnon JA, Church GM, Valen E (2014) CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucl Acids Res 42(W1):W401–W407

Nekrasov V, Staskawicz B, Weigel D, Jones JD, Kamoun S (2013) Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. Nat Biotechnol 31(8):691–693

Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F, Nureki O (2014) Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell 156(5):935–949

Podevin N, Devos Y, Davies HV, Nielsen KM (2012) Transgenic or not? No simple answer! EMBO Rep 13(12):1057–1061

Puchta H, Fauser F (2014) Synthetic nucleases for genome engineering in plants: prospects for a bright future. Plant J 78(5):727–741

Ron M, Kajala K, Pauluzzi G, Wang D, Reynoso MA, Zumstein K, Garcha J, Winte S, Masson H, Inagaki S, Fn Federici, Sinha N, Deal RB, Bailey-Serres J, Brady SM (2014) Hairy Root transformation using *Agrobacterium rhizogenes* as a tool for exploring cell type-specific gene expression and function using tomato as a model. Plant Physiol 166(2):455–469

Sander JD, Joung JK (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol 32(4):347–355

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. doi: 10.1111/tpj.12704

Shaked H, Melamed-Bessudo C, Levy AA (2005) High-frequency gene targeting in *Arabidopsis* plants expressing the yeast *RAD54* gene. PNAS 102(34):12265–12269

Shan Q, Wang Y, Chen K, Liang Z, Li J, Zhang Y, Zhang K, Liu J, Voytas DF, Zheng X, Zhang Y, Gao C (2013a) Rapid and efficient gene modification in rice and brachypodium using TALENs. Mol Plant 6(4):1365–1368

Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu JL, Gao C (2013b) Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotechnol 31(8):686–688

Sugano SS, Shirakawa M, Takagi J, Matsuda Y, Shimada T, Hara-Nishimura I, Kohchi T (2014) CRISPR/Cas9-mediated targeted mutagenesis in the liverwort *Marchantia polymorpha* L. Plant Cell Physiol 55(3):475–481

Voytas DF (2013) Plant genome engineering with sequence-specific nucleases. Annu Rev Plant Biol 64(1):327–350

Wang T, Wei JJ, Sabatini DM, Lander ES (2014a) Genetic screens in human cells using the CRISPR-Cas9 system. Science 343(6166):80–84

Wang Y, Cheng X, Shan Q, Zhang Y, Liu J, Gao C, Qiu JL (2014b) Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. Nat Biotechnol 32(9):947–951

Xie K, Yang Y (2013) RNA-guided genome editing in plants using a CRISPR-Cas system. Mol Plant 6(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(5):923–926

Xu R, Li H, Qin R, Wang L, Li L, Wei P, Yang J (2014) Gene targeting using the *Agrobacterium tumefaciens*-mediated CRISPR-Cas system in rice. Rice (N Y) 7(1):1–4

Zhang F, Maeder ML, Unger-Wallace E, Hoshaw JP, Reyon D, Christian M, Li X, Pierick CJ, Dobbs D, Peterson T, Joung JK, Voytas DF (2010) High frequency targeted mutagenesis in *Arabidopsis thaliana* using zinc finger nucleases. Proc Natl Acad Sci 107(26):12028–12033

Zhang Y, Zhang F, Li X, Baller JA, Qi Y, Starker CG, Bogdanove AJ, Voytas DF (2013) Transcription activator-like effector nucleases enable efficient plant genome engineering. Plant Physiol 161(1):20–27

Zhang H, Zhang J, Wei P, Zhang B, Gou F, Feng Z, Mao Y, Yang L, Xu N, Zhu JK (2014) The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. Plant Biotechnol J 12(6):797–807

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. Nucl Acids Res 42(17):10903–10914