ABSTRACT

Title of Dissertation:	DEVELOPING AND EXPANDING CRISPR-CAS PLANT GENOME ENGINEERING SYSTEMS
	Aimee Alyssa Malzahn, Doctor of Philosophy in Plant Science, 2022
Dissertation directed by:	Dr. Yiping Qi, Department of Plant Science and Landscape Architecture

In order to advance plant biology and speed up crop breeding, researchers have used genome engineering tools in their research. Genome engineering with CRISPR has revolutionized agriculture by providing an easy, fast, and accessible tool to induce desirable mutations.

This thesis works on addressing problems in the application of CRISPR for plant genome engineering. CRISPR systems are adopted from bacterial immune systems and consists of a Cas endonuclease and a guide RNA (or crRNA). Cas variants have different characteristics and exploring natural variants can provide systems with enhanced or new applications. The first aim in this thesis is to demonstrate novel LbCas12a for genome editing in Arabidopsis. To overcome the temperature sensitivity of LbCas12a, a heat treatment regime was developed. In order to expand LbCas12a use beyond genome editing, a transcription repression system was developed and used successfully for multiplexed repression of two homologs of

EDS1. Two crRNA processing systems were compared, and results suggest that either can be used successfully in Arabidopsis.

The second aim is to improve Cas9 and Cas12a editing outcomes by creating Cas-effector fusions. Cas9 and Cas12a were fused with six different exonucleases and compared at three targets in rice protoplasts. Several Cas-exonuclease fusions resulted in an increase in editing efficiency and the production of larger deletions. The Cas-exonuclease fusions' editing efficiency differed between Cas9 and Cas12a, along with the deletion profile. Additionally, chromatin modulating peptides were fused to Cas9, which resulted in higher editing efficiency without altering deletion profiles. These engineered Cas proteins can be used to create unique editing outcomes, and paired with an increased editing efficiency, could be used to target difficult-to-edit target sites for gene knockout and cis-regulatory elements for fine-tuning gene expression.

In summary, this work explored new Cas variant LbCas12a, developed multiplex gene repression systems, and compared engineered fusion Cas9 and Cas12a proteins for increased editing efficiency and larger deletions. The developed and improved CRISPR systems expand the number of available targets, improve efficiency, demonstrate novel editing outcomes, and enable multiplexed transcriptional regulation in plants.

DEVELOPING AND EXPANDING CRISPR-CAS PLANT GENOME ENGINEERING SYSTEMS

by

Aimee Alyssa Malzahn

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:
Professor Yiping Qi, Chair
Professor Shirley Micallef
Professor Shunyuan Xiao
Professor James Culver
Professor Stephen Mount

Dedication

Dedicated to Andrew T. Goffin

I would need more words than are in this thesis to describe how much you mean to me.
Thank you for all that you've done and all that you are.

Acknowledgements

There are many people that contributed to my Ph.D. journey and to this thesis. Thank you to everyone who contributed to this work and to my personal and professional growth.

A huge thank you:

To my advisor, Dr. Yiping Qi, for providing advice, research opportunities, funding, and mentorship.

To my thesis committee, Drs. Shirley Micaleff, James Culver, Shunyan Xiao, and Stephen Mount, for their feedback and guidance.

To the Qi lab members past and present, for their research support and technical advice.

Contributions to this dissertation

Chapter 3:

Simon Sretenovic completed the analysis of the qRT-PCR data shown in figures 5,6 and 7. Aimee A. Malzahn produced the raw qRT-PCR data and constructed the figures.

Dr. Yingxiao Zhang completed the statistical analysis of the multiplexed repression qRT-PCR data and constructed figure 8. Aimee A. Malzahn produced the raw qRT-PCR data and completed the analysis.

Chapter 4:

Dr. Xiaoyu Wang completed the Cas entry vector construction of Cas12a-exonuclease fusions, the T-DNA construction, protoplast transformation, and the data analysis.

Table of Contents

Chapter 1: Introduction

Contains the published work of “Plant Genome Editing with TALEN and CRISPR”:

Malzahn A., Lowder L., Qi Y. (2017) Plant Genome Editing with TALEN and CRISPR. Cell & Bioscience. 7:21 doi: 10.1186/s13578-017-0148-4

Introduction

The field of genome editing is experiencing rapid growth as new methods and technologies continue to emerge. Using genome editing to boost agriculture productivity is needed as the world population is expected to grow to 9.6 billion by 2050 while the amount of arable land decreases [1]. Besides potential for boosting crop yields, genome editing is now one of the best tools for carrying out reverse genetics and is emerging as an especially versatile tool for studying basic biology.

Genome edited plants are differentiated from conventional transgenic plants as they may not incorporate foreign DNA. Although genome editing can be used to introduce foreign DNA into the genome, it may simply involve changes of a few base pairs in the plant's own DNA. This distinction makes genome editing a novel and powerful breeding tool that has promising applications in agriculture, especially when genome edited crops are not regulated as genetically modified (GM) [2].

Genome editing relies on DNA repair

DNA damage occurs naturally in all cells either due to exogenous factors, such as UV radiation, or endogenous agents such as metabolic by-products and free radicals. A Double-strand break (DSB) is the most lethal type of DNA damage and must be repaired before DNA replication, which has led to the evolution of two major DNA repair pathways in eukaryotes: non-homologous end-joining and homology-directed repair [3–6] (Figure 1).

Non-homologous end-joining (NHEJ) is an error-prone repair pathway. When a DSB occurs, NHEJ can quickly, although often imprecisely, be used in two ways to repair the break. In classical NHEJ (Figure 1a), several different proteins (e.g. Ku70 and Ku80) bind to broken DNA ends and are joined together by a ligase that can result in the insertion or deletion (indel) of nucleotides. In microhomology-based alternative NHEJ (Figure 1b), 5' ends are cut until 3' overhangs with homology are created. DNA strands then bind at their complementary sequence, and flaps of non-homologous DNA are excised. This typically results in deletions as DNA between homologous sections are removed. NHEJ often leads to frameshift mutations which can result in premature stop codons, rendering genes non-functional (Figure 1a and b). This is helpful for creating knockout plants useful for reverse genetic studies, but can also create desirable agricultural traits. For example, a powdery mildew resistant wheat line was created by knocking out three redundant *MLO* genes [7].

The second DNA repair pathway is homology directed repair (HDR) which relies on template DNA. Homologous recombination is an important process that occurs in somatic cells to repair DSBs and in meiotically dividing cells to exchange genetic material between parental chromosomes. The most common conservative HDR mechanism in plants, which repairs almost all DSBs in somatic cells, is the synthesis-dependent strand annealing (SDSA) pathway [5, 8] (Figure 1c). As a DSB occurs, 3' overhangs are extended from the break site. A 5' end invades the homologous strand forming a D-loop. Synthesis fills in the gaps using homologous DNA as a template, and the 3' end reanneals with the second 3' end without crossover. The result is a precisely integrated template or "donor" DNA strand. In nature, template DNA in the

form of a sister chromatid or homologous chromosome is not always available, which may hinder HDR. However, synthetic template DNA can be provided exogenously and used for gene insertion, replacement, or epitope/florescent tagging. There are many exciting applications in basic and applied science using HDR. For example, HDR was used to engineer an herbicide resistant trait in tobacco plants [9].

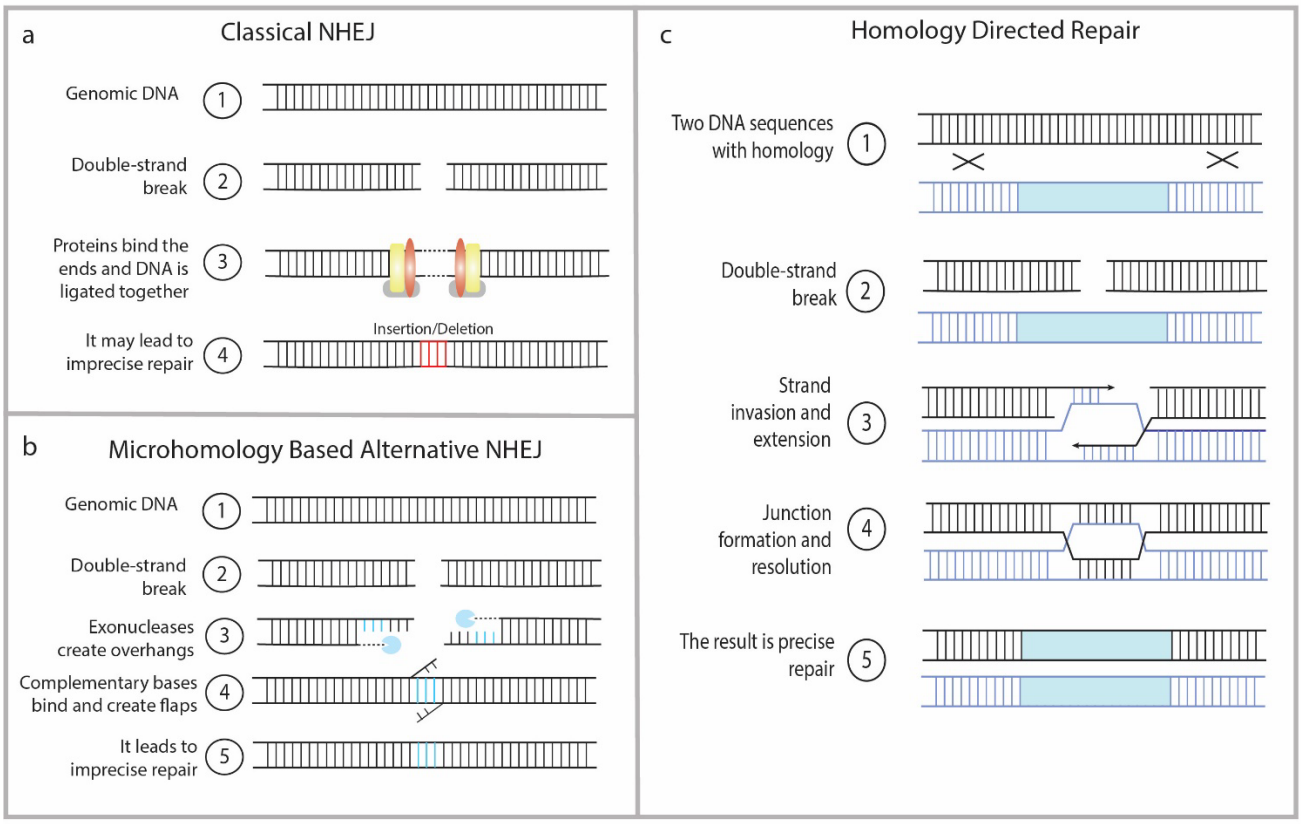


Figure 1. Major DNA repair pathways in plants.
Non-homologous end joining (NHEJ) and homology directed repair (HDR) are two main repair pathways. Classical NHEJ may lead to insertions or deletions, while Microhomology based alternative NHEJ always results in deletions. Homology directed repair is less efficient but can result in precise integration of a donor DNA template into the genome.

Rapid evolution of sequence specific nucleases (SSNs) for plant genome editing

Meganucleases, or homing endonucleases, are site specific endonucleases found in eukaryotes, archaea, and bacteria which recognize DNA sequences over 12 bp long [10]. Several hundred meganucleases have been discovered and they can be divided into four families: LAGLIDADG, His-Cys box, GIY-YIG, and the HNH family [10]. The LAGLIDADG family consists of popular meganucleases I-CreI and I-SceI. Originally, meganucleases were only able to target a single sequence and thus were not capable of targeting endogenous genes. After it was discovered that only a few amino acid residues make direct contact with nucleotides, the binding specificity was successfully altered for targeting endogenous genes. For example, targeted mutagenesis was successfully achieved in maize with de novo-engineered meganucleases [11]. However, DNA binding properties of meganucleases cannot be completely separated from their nuclease activity, making them

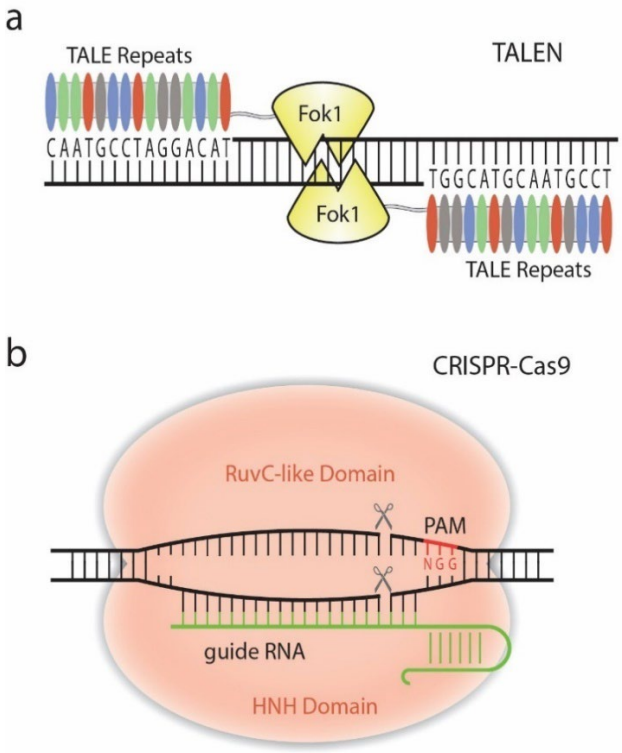


Figure 2. TALEN and CRISPR-Cas9.
a) A TALEN is composed of two monomers with each containing a TALE DNA binding domain and a FokI nuclease domain. Fok1 dimerizes to create a double-strand break. b) CRISPR-Cas9 is a two-component system composed of Cas9 and a gRNA. Once Cas9 finds a PAM site, if the gRNA binds to the DNA, a double break occurs three base pairs upstream the PAM.

difficult to engineer and use in research.

Zinc finger nucleases (ZFNs) function as dimers and each monomer is a fusion protein of a zinc finger DNA binding domain and a non-specific FokI nuclease domain [12, 13]. A zinc finger is formed by repeated groupings of cysteine and histidine residues and recognize 3 nucleotides (nt). Each ZFN monomer is typically composed of 3 or 4 zinc fingers, recognizing 9 or 12 nt DNA. The zinc fingers are thought to be modular, making it possible to recognize a long stretch of DNA by putting multiple zinc fingers together [14, 15]. However, ZFNs based on modular assembly typically have poor activity and high toxicity [16, 17], suggesting there is context dependency among neighboring fingers. This context dependency in ZFN engineering has been largely addressed by a proprietary platform developed by Sangamo Bioscience [18] and by academically developed platforms such as "OPEN" [19] and "CoDA" [20]. "OPEN" or "CoDA" generated ZFNs were later used for generating mutants and studying DNA repair mechanisms in the model plant Arabidopsis [21–23].

The possibility of engineering transcription activator-like (TAL) effectors for DNA targeting was realized in 2009 when their DNA binding mechanism was discovered [24, 25]. TAL effectors in nature are introduced into plant host cells by the bacterium *Xanthomonas* via the type III secretion system, where they alter host gene expression to meet the bacteria's needs. In the nucleus, TAL effectors bind target genes' promoters within 60 base pairs of start codons and activate transcription [24]. The DNA binding central repeat domain of each TAL effector is composed of a few to 33.5 repeats which are typically made of 34 amino acids [26]. Using a β-

Glucuronidase (GUS) reporter in tobacco, Boch *et al.* discovered repeat variable diresidue (RVD) at positions 12 and 13 of each repeat determines nucleotide binding specificity [25]. This breakthrough quickly led to the creation of a new kind of SSN called TAL effector nuclease (TALEN), which is based on the fusion of a Fok1 nuclease domain to the DNA binding TALE repeats [27–30] (Figure 2a). There are benefits to choosing TALENs over ZFNs. First, TALEs are less toxic and secondly, they are easier to engineer because recognizing each DNA nucleotide simply relies on using a TALE repeat with the corresponding RVD. However, the repetitive sequence of TALE makes them difficult to construct via polymerase chain reaction (PCR). This was addressed with the development of multiple assembly methods mostly based on Golden gate cloning (e.g. [31–33]), which furthered rapid adoption of TALEN technology for genome editing in many organisms including plants.

Just two years after the realization of TALENs, another genome editing tool was introduced. Clustered regularly interspaced palindromic repeats (CRISPR) had been found to function as an adaptive immune system in bacteria and archaea against invading viruses, phages and plasmids [34–36]. The bacteria can protect themselves using a series of CRISPR associated (Cas) proteins that cleave viral DNA, insert pieces of viral DNA into their own genomes, and then use certain Cas9 protein(s) paired with RNA transcribed from the viral DNA library to make targeted double-strand breaks in invading viral DNA. Class 2 CRISPR-Cas systems utilize single-protein effectors, such as Cas9, for DNA targeting [37]. Cas9 is composed of two endonuclease domains, HNH and a RuvC-like domain that each cut one strand of DNA (Figure 2b). It was demonstrated in 2012 that Cas9 of *Streptococcus pyogenes*

could be paired with a synthetic single guide RNA (gRNA) to create a targeted DNA DSB *in vitro* and in *E.coli* [38]. Shortly after, CRISPR-Cas9 was demonstrated as a powerful RNA-guided SSN for genome editing in human cells [3, 39]. Although off target effects have been a concern, the simple design and ease of vector construction has dramatically increased the number of genome editing studies using CRISPR-Cas9 in plants [40, 41] (Table 1).

Both TALEN and CRISPR-Cas9 have been used extensively for genome editing and each have their own unique disadvantages and advantages (Table 1), that will be further explored in this review. Both systems will continue to be useful as molecular scissors for a wide variety of applications.

Table 1. Comparison of TALEN and CRISPR-Cas9 Systems

TALEN	
Advantages	**Disadvantages**
• ~30 bp target requirement results in less off-target effects • No PAM requirement; can target any sequence	• Difficult protein engineering potentially increases time and financial investment • Efficiency varies for each construct • Cannot target methylated DNA • Difficult to engineer nickase
CRISPR-Cas9	
Advantages	**Disadvantages**
• Able to multiplex • Easy to engineer • Can target methylated DNA • Easy to create a nickase	• Higher potential for off-target effects • PAM requirement limits target

Structure of CRISPR-Cas9 and Cas12a

Since the first use of Cas9 for genome editing, many Cas proteins have been described and used for genome editing. The most efficient of these novel endonucleases is Cas12a (formerly Cpf1). Cas12a has a different classification and

characteristics, which makes it a valuable expansion to the CRISPR toolkit. CRISPR systems are divided into two main classes which are then further classified into different types based on the way the system processes their gRNAs. The two main endonucleases used for genome editing, Cas9 and Cas12a are Type II and Type V systems, respectively. The Cas9 endonuclease is made up of two lobes containing several different domains (Figure 1a) [42]. The REC lobe contains the bridge helix, Rec1, Rec2, and Rec3 domains. The NUC lobe contains the RuvC and HNH for cleavage, and the PAM-interacting domain. The Cas9 endonuclease touches down onto a DNA strand and if it does not find a PAM, it releases from the DNA. This process was shown to take 17 (+/- 4) ms in bacteria [43]. If Cas9 binds to an appropriate PAM next to a sequence complementary to the sgRNA, the DNA will unwind (starting at the PAM) and bind to the RNA [44]. This interaction between RNA and double stranded DNA forms an R-loop, and it is the rate-limiting step for the reaction [45]. R- loop formation triggers conformational changes in the HNH domain, which cleaves the target stand, and in the RuvC domain which cleaves the non-target strand. Cleavage results in a blunt double-stranded break 3bp upstream of the PAM. Once Cas9 cleaves, it stays bound to the DNA [46].

In contrast to Cas9's structure which relies on two catalytic domains, Cas12a relies only on RuvC to make double-stranded breaks in DNA. The Nuc domain contributes to binding and mutations in the Nuc domain affect target strand cleavage, but do not result in complete loss of cleavage [47, 48]. Although the complete cleavage mechanism has not been elucidated, the most recent hypothesis is that RuvC cleaves the target strand, which in turn triggers conformational changes in the REC and NUC

lobes to place the non-target strand in position to be cleaved by the RuvC domain [47]. This results in a double stranded break 18 bp downstream from the PAM on the target strand and 23 bp away on the non-target strand, leaving a 5' overhang at each end. This overhang results in a different deletion profile from Cas9 DSB. Cas12a has been shown to commonly result in 6-15 bp deletions, while Cas9 often produces smaller deletions and 1 bp insertions [49, 50]. Because the DSB is distal to the PAM, Cas12a can theoretically target the same site multiple times, contrary to Cas9 where a single mutation next to PAM would prevent further Cas9 sgRNA binding Although it has not been achieved in plants, the overhangs produced by Cas12a have been viewed as a potential strategy for NHEJ gene insertion. If the insert had complementary overhangs, nucleotide binding would facilitate insertion instead of the tempospatial array of proteins involved in HDR.

NHEJ based genome editing by TALEN

Over 50 genes have been targeted for mutations using TALEN in plants, including Arabidopsis, Barley, Brachypodium, Maize, Tobacco, Rice, Soybean, Tomato and Wheat (Table 2). Many of these have been proof-of-concept studies. TALEN scaffolds were optimized for high activity in plants [51]. The optimized TALEN scaffold was then demonstrated by targeted mutagenesis in Arabidopsis [52], tomato [53], Brachypodium [54] and wheat [7]. More recently, TALEN was shown to induce a variety of heritable mutations in rice [55], demonstrating its usefulness in plant genome editing.

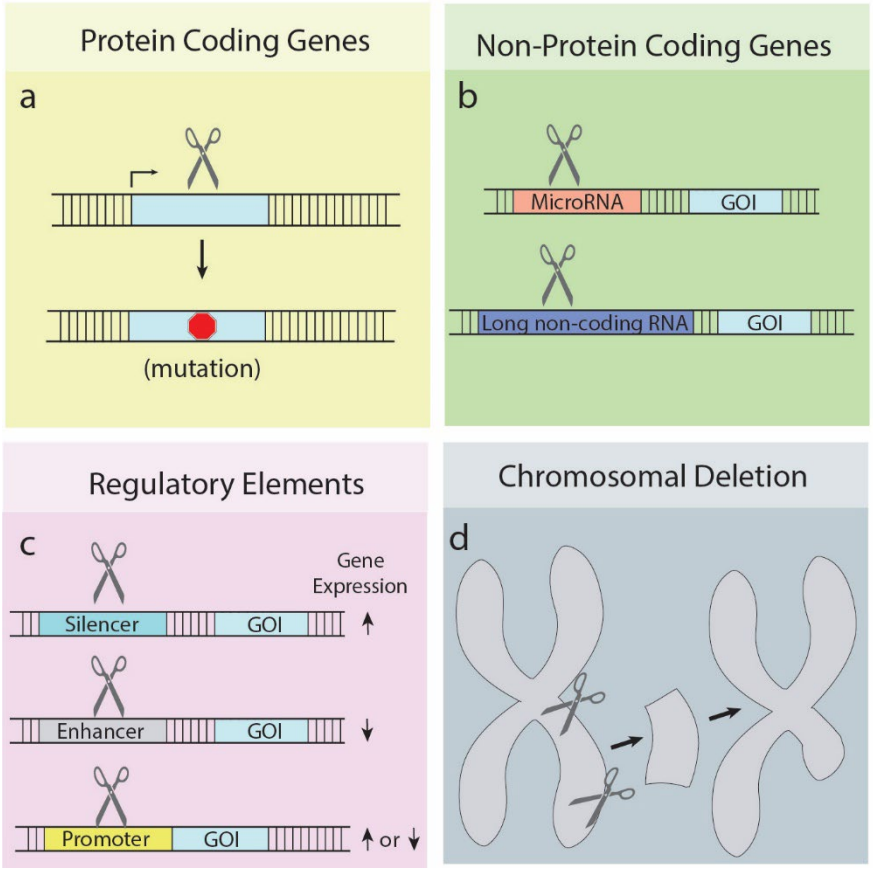


Figure 3. NHEJ based genome editing applications.

a) NHEJ repair of an SSN induced break can create a premature stop codon. A stop codon is indicated by a red octagon. GOI is an acronym for Gene of Interest. b) Non-protein coding genes such as microRNA and long non-coding RNA can be rendered non-functional through targeted mutations by SSNs. c) Regulatory elements involved in the activation or repression of genes can be disrupted by SSNs. d) Pieces of chromosomes that may involve regulatory networks or related genes can be deleted by SSNs.

As an effective genome editing tool, TALEN has been applied to generate useful traits in crops. In an elegant study, TALEN was used to engineer disease resistance in *Xanthomonas oryzae* pv. *oryzae* by destroying the target sequence of TALE effectors in rice [55]. In soybean, the *FAD2* gene was targeted for improved oil quality [56]. In wheat, three homologs of *MLO* were successfully targeted for simultaneous knockout,

conferring heritable disease resistance to powdery mildew [7]. Improved rice seeds have been engineered with TALEN, creating traits such as fragrance [57] and storage tolerance [58]. Improved cold storage and processing traits have also been engineered in potato [59]. Most of these studies targeted protein coding genes for mutagenesis (Figure 3a). Other types of NHEJ based editing can also be achieved by TALEN, such as targeted mutagenesis of non-protein coding genes (Figure 3b) and regulatory elements [60] (Figure 3c), and generating large chromosomal deletions [52] (Figure 3d).

NHEJ based genome editing by CRISPR-Cas9

Due to ease of engineering, CRISPR-Cas9 has been widely adopted for genome editing in plants (Table 3). At the time of this review, CRISPR-Cas9 edited plants include Arabidopsis, barley, *B. oleracea*, cotton, dandelion, flax, lettuce, liverwort, corn, petunia, populus, rice, sorghum, soybean, sweet orange, tomato, wheat, and several tobacco varieties (Table 3). CRISPR-Cas9 quickly moved beyond proof-of-concept; promoting a reverse genetics revolution in plant research and creating many desirable traits in major crops. Using rice as an example, multiple yield-related genes have been targeted in rice [61]. CRISPR-Cas9 has been widely used for functional study on rice genes (Table 3). In addition, environment-induced male sterility has been engineered to facilitate hybrid-based breeding [62, 63]. Disease resistance traits have been developed by knocking out host genes in rice [64] and Arabidopsis [65].

The intrinsic property of CRISPR-Cas9 for targeting viral DNA for cleavage makes it a great tool to increase plant immunity against DNA viruses. For example,

such immunity has been shown in tobacco by stably expressing Cas9 and introducing gRNAs that target geminiviruses [66]. Many similar studies have targeted geminiviruses because they must maintain circular structure for replication, thus one DSB will destroy the virus [67]. Tobacco with resistance to the geminiviruses beet severe curly top virus, bean yellow dwarf virus, and tomato yellow leaf curl virus have been created [66, 68, 69]. These findings were also replicated in Arabidopsis [69]. Because Cas9 can complex with any compatible and programmable gRNAs, it may offer a robust protection strategy against double stranded DNA viruses. Single stranded viruses can also be potentially targeted by NMCas9 which exhibit DNase H activity [70].

CRISPR-Cas9 is a valuable reverse genetic tool in plant science research. Large chromosomal deletion in Arabidopsis was used to demonstrate redundant functionality of tandem arrayed *CBF* genes in cold acclimation [71] (Figure 3d). CRISPR-Cas9 based reverse genetics was even made possible in poplar [72], a woody tree that has traditionally proven difficult for genetic manipulation. Despite challenges with editing polyploidy plants, both hexaploid bread wheat and tetraploid durum wheat were effectively edited by CRISPR-Cas9 [7, 73, 74]. Editing of the tetraploid cotton genome was also recently reported [75, 76]. *Camelia sativa* is a hexaploid relative to Arabidopsis and editing three copies of the *FAD2* gene was demonstrated when screen was carried to T3 generation [77, 78]. Using CRISPR-Cas9, two recent studies disproved conclusions made by earlier work using traditional genetic techniques, further demonstrating that CRISPR-Cas9 is a great addition to existing genetic tools. In one study, knockout alleles of *ABP1* were generated in

Arabidopsis and it was discovered this gene is not required for auxin signaling or development as originally thought [79]. In another study [80], *Rj4* was found to control nodulation specificity in soybean and the identity of this gene confirmed by CRISPR-Cas9 corrected earlier reports.

CRISPR-Cas9 will also further reverse genetic studies on non-protein coding genes (Figure 3b) and regulatory elements (Figure 3c). MicroRNAs are short RNAs that can repress translation, but mostly cleave mRNA transcripts [81]. Both mechanisms silence protein expression. Long non-coding RNAs are diverse groups of non-coding transcripts longer than 200 nucleotides whose function is poorly understood in plants [82]. Small indel mutations in non-protein coding genes may not alter or destroy their function, making them more challenging targets with CRISPR-Cas9 [83]. CRISPR-Cas9 mediated targeted chromosomal deletion is very efficient in rice [84] and this approach was recently applied for deleting microRNA genes in Arabidopsis [85]. Moreover, CRISPR-Cas9 was used to target a non-coding regulatory element of *OsRAV2* in rice to confirm its function in response to salt treatment [86].

Multiplex CRISPR-Cas9 systems

One distinct advantage of CRISPR-Cas9 over TALEN is the ability to multiplex (Table 1). By expressing multiple gRNAs that independently pair with Cas9, multiple target sites can be mutated in a single cell. This multiplexing property of CRISPR-Cas9 has enabled targeted deletion of large chromosomal segments containing multiple genes in rice [84] and in Arabidopsis [71]. Simultaneous targeting

of multiple genes can result in more than one improved trait in crops and can also be used in basic research to deduce the role of each gene in a complex network.

The first toolkit to demonstrate multiplexing knockout of three Arabidopsis genes was released in 2014 [87]. Since then, several toolkits have been developed. A second toolkit was released in 2015 by Ma et al. [88], that constructed vectors using PCR and Golden Gate cloning. These constructs were validated in both monocots and dicots. A third toolkit was released in that same year by Lowder et al. [89]. This kit contains vectors that could be used for genome editing and transcriptional regulation without the need for PCR, ensuring that no mutations occur during assembly. Other multiplex systems were also developed that, while more time consuming, allowed for targeting of up to six target sites or theoretically unlimited target sites respectively [90, 91].

Paired CRISPR-Cas9 nickase for improving editing specificity

TALEN works in pairs to recognize 30bp or even longer DNA sequences and presumably has higher targeting specificity than CRISPR-Cas9 which recognizes ~20bp DNA sequence. However, the targeting specificity of CRISPR can be improved by using a paired nickase strategy (Figure 4a). One of the Cas9 endonuclease domains, either HNH or RuvC-like, is inactivated to produce a Cas9 nickase that can only cut one DNA strand. By pairing two nickases and their gRNAs, the target sequence grows from ~20bp to ~40bp and specificity is drastically increased. It was shown this increase in specificity results in a 20-1,500- fold reduction in off-target effects without a decrease in cleavage efficiency in human cells [92]. There are several examples of successful genome editing using nickases in

plants [93–95]. A single transcript unit (STU) was effectively shown to express Cas9 nickase and a gRNA pair [96], in which Cas9 and two gRNAs flanked by hammerhead ribozyme sequences were expressed under a single Polymerase II promoter. The ribozyme successfully processed the single transcript, demonstrating a system for simultaneous, inducible expression of both Cas9 and gRNAs.

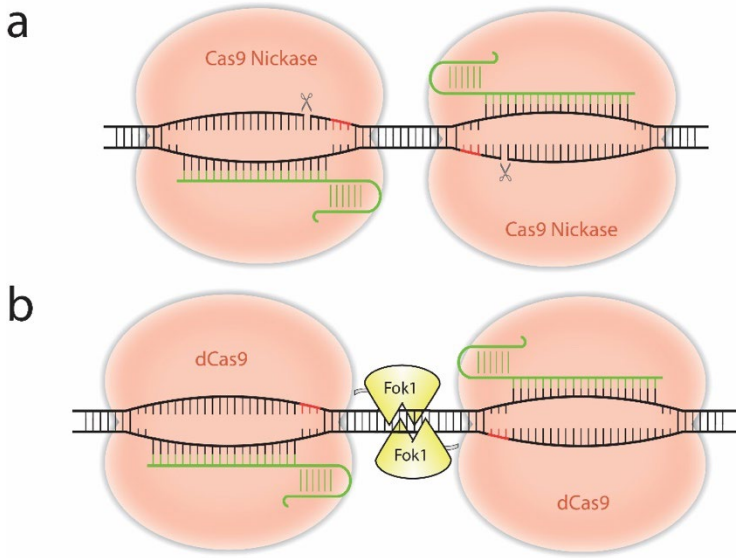


Figure 4. Paired Cas9 nickase and FokI-dCas9 systems. Alternative Cas9 proteins can decrease off-target effects. a) Two nickases are required to make a double-strand break, increasing the gRNA requirement and length of target sequence. b) A catalytically dead Cas9 is paired to a Fok1 nuclease, also resulting in an increased length of target sequence for enhanced targeting specificity.

Alternatively, FokI-dCas9 can be engineered to work in pairs [97, 98], which relies on fusing a catalytically dead Cas9 (dCas9) with a FokI nuclease domain (Figure 4b). When the two Fok1-dCas9s are carefully positioned on both DNA strands, the gRNAs lead dCas9 to the target sites and FokI nuclease domains dimerize resulting in DNA cleavage. As with the paired nickase strategy, the requirement of two gRNAs should decrease off-target effects. This takes advantage of the simple design of gRNAs and avoids the protein engineering required for TALEN. However, the editing frequency for both techniques will need to be improved for wide-scale adoption.

HDR based genome editing with TALEN and CRISPR-Cas9

There are many powerful applications for HDR based genome editing using both TALEN and CRISPR-Cas9. The applications include, but are not limited to, gene replacement (Figure 5a), epitope tagging (Figure 5b) or florescent protein tagging (Figure 5c) of endogenous genes, and gene insertion which can be used for trait stacking (Figure 5d).

Gene replacement with HDR was first accomplished using TALENs in human cells in 2011 [99], but it wasn't until 2013 that HDR initiated by TALEN was demonstrated in plants [51] (Table 2). Barley was the first monocot to demonstrate HDR with TALEN. A green fluorescent protein (GFP) was converted into yellow fluorescent protein (YFP) by one amino acid change with a 3% efficiency in protoplasts, demonstrating an effective system for optimizing TALENs [100]. Replacing *ALS* with an herbicide resistant gene was successful in tobacco protoplasts and rice with TALEN [51, 101]. In the tobacco protoplasts, about 30% of transformed cells had NHEJ mutations and 14% showed targeted insertion due to HDR [51]. For this study, transient expression of TALEN was efficient enough to get edited plants without selection. In rice, it was reported that between 1.4% and 6.3% of transformants had one or both alleles edited [101]. In tomato, targeted insertion of a strong promoter ahead of the *ANT1* gene led to ectopic accumulation of anthocyanin, producing purple tomatoes [102]. The study utilized a geminivirus replicon system that has the advantage of amplifying the genome editing reagents in plant cells [103].

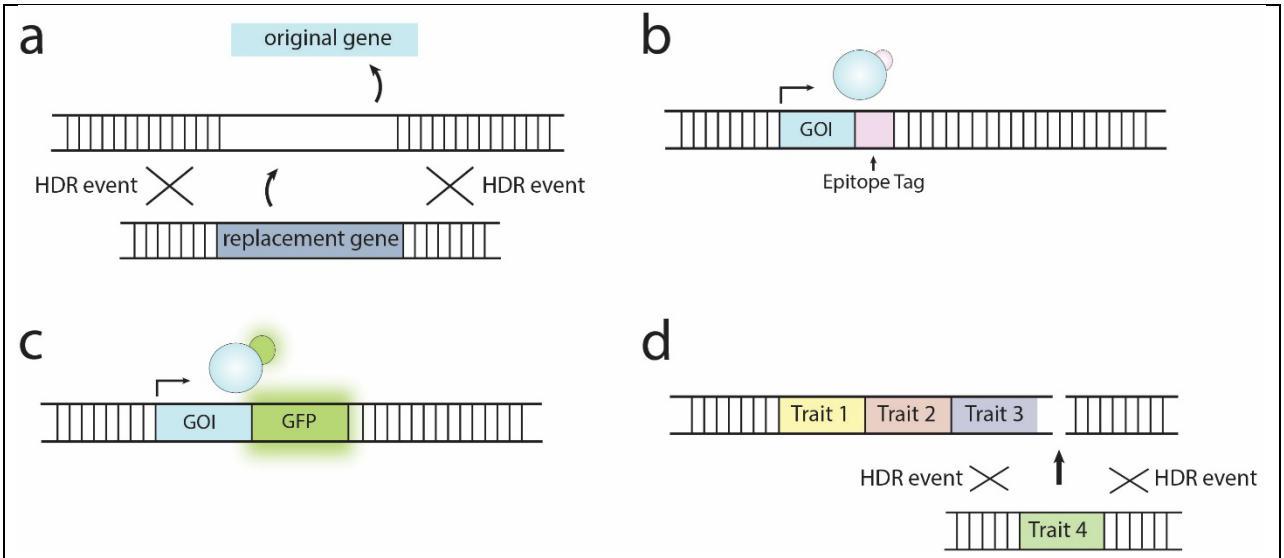


Figure 5. HDR based genome editing applications.
a) Gene replacement is applicable for basic research and agriculture. b) HDR can add a tag to a protein for easy purification and study. c) Fluorescent proteins such as green fluorescent protein (GFP) can be fused to a gene of interest for *in vivo* study. d) Gene stacking is useful for placing genes physically close together on a chromosome. This is accomplished by creating a target site for HDR at the end of each gene, which allows for modular addition of genes.

HDR utilizing CRISPR-Cas9 was first demonstrated in 2013 [104] (Table 3). A plant codon-optimized Cas9 and gRNAs were transiently expressed in Arabidopsis and tobacco protoplasts for targeting respective *PDS* genes. A much higher mutagenesis frequency was observed in the tobacco protoplasts compared to Arabidopsis. HDR was accomplished at 9% frequency with a donor template harboring an *Avr*II digestion site, a 533 bp left homology arm, and a 114 bp right homology arm. This proof-of-concept study demonstrated that it is possible to replace a wild-type gene with an altered one using CRISPR-Cas9 in plant cells. A year later, germline editing of the *ADH1* gene was demonstrated in Arabidopsis [94]. CRISPR-Cas9 has also been used to alter *ALS* in rice to confer herbicide resistance [105, 106]

and both studies explored different strategies to enhance HDR in rice. In one study, plants with a *lig4* mutation were shown to have between a 0.147% and 1% gene targeting efficiency and these contained biallelic mutations [106]. Lig4 is involved in the classic NHEJ pathway (Figure 1a) and *Lig4* mutants have been shown to undergo increased rates of HDR and microhomology-based alternative NHEJ in Arabidopsis [22]. In the second study, the authors observed high frequency HDR when using two gRNAs for cutting off the target gene and liberating donors that were provided in the form of both plasmids and free double-stranded DNAs [105].

For all HDR applications, efficiency will need to be improved. Increasing the efficiency of SSN delivery will greatly help genome editing, including HDR applications. If a higher percentage of plants or plant cells can receive SSNs, then more of them will have the potential to undergo HDR without increasing sample size. Although easy to use, agrobacterium-mediated delivery is not as efficient as ballistic bombardment because the latter can introduce multiple copies of donor DNA [101, 106]. One of the potential methods that may solve issues with difficult delivery, as well donor copy number, is geminivirus delivery. In tomatoes, geminiviruses replicons were found to create mutations at a tenfold higher frequency when compared to agrobacterium mediated transfer [102]. Recently, geminivirus systems were successfully used for CRISPR-Cas9 mediated HDR in rice [107] and wheat [108]. Alternatively, donor DNA may be liberated from integrated chromosome regions with an *in-planta* gene targeting strategy [94, 109]. The second issue to address is low occurrence of HDR in cells, especially in non-dividing cells. If all cells in culture or *in planta* were synchronized, then SSN and donor DNA could be

introduced during replication which will boost HDR events. Cas9 nickases, with their ability to create single stand breaks (SSBs), have been utilized for HDR in Arabidopsis at high efficiencies and the authors have speculated the mechanism of HDR initiated by SSBs could be different from that of DSBs [93]. The mechanism of SSB based HDR, if discovered, should be useful for enhancing HDR. There are many exciting possibilities for HDR based genome editing, and innovative ideas will continue to further this area.

TAL effector and CRISPR-Cas9 for transcriptional regulation

Either a TAL effector or a deactivated Cas9 (dCas9) can be fused to an activator such as VP64 [110] or a repressor such as SRDX [111] for transcriptional regulation in plants (Figure 6). There may be some differences intrinsic to TAL effector and Cas9 that make one more suitable for activating or repressing gene expression than the other. To date, no study has been carried out to make an accurate comparison of both systems in plants.

TAL effectors are natural transcriptional activators in plants [112, 113]. This property was cleverly used for decoding the DNA recognition code of TAL repeats [25]. Although the endogenous transcriptional activation domain of a TAL effector seems potent for activation, it could be swapped with VP64 to make smaller proteins (Figure 6a). TAL repeats, when fused to SRDX, repressed gene expression by more than two-fold in Arabidopsis [114]. Interestingly, it was recently reported in Arabidopsis that binding of TAL proteins to the sense strand of a gene of interest is enough to result in gene repression [115], which is likely due to TAL proteins

blocking transcription. Despite proven concept, there is almost no report on utilizing de novo-engineered TAL activators or repressors in plant research. This could be due to the difficulty of engineering of TAL proteins and multiplexing them in plant cells.

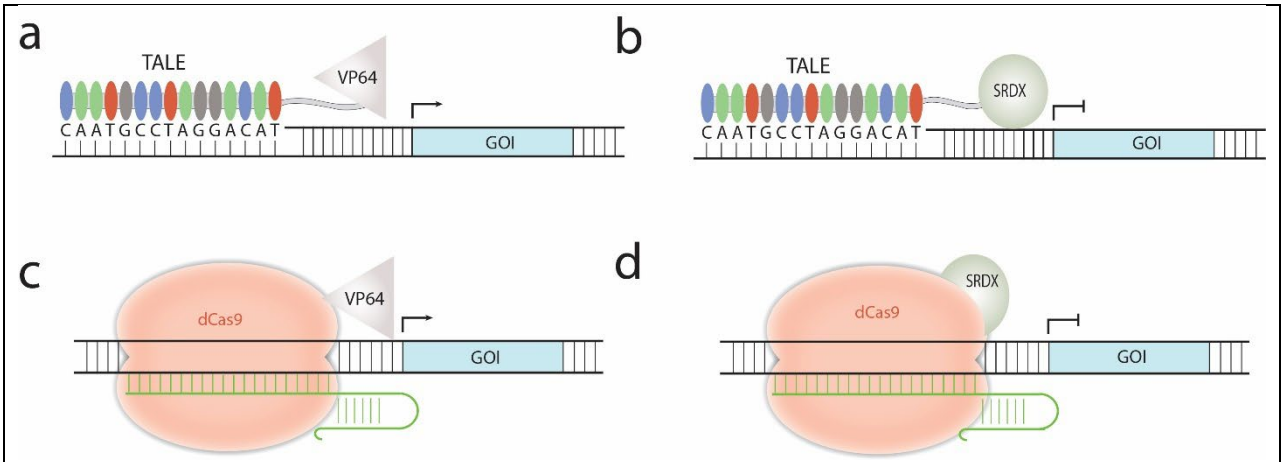


Figure 6. TALE and CRISPR-Cas9 based transcriptome modulation systems. a) The activator VP64 is fused to TALE for gene activation. b) The repressor SRDX is fused to TALE for gene repression. c) The activator VP64 is fused to dCas9 for gene activation. d) The repressor SRDX is fused to dCas99 for gene repression.

CRISPR-Cas9 may be more suitable for developing transcriptional regulation tools due to facile engineering and multiplexing. CRISPR-dCas9 based activators and repressors were demonstrated in transiently transformed tobacco [116] and in stably transformed Arabidopsis [89]. In the latter study, a tool kit was developed for easy assembly of a final T-DNA construct for simultaneous transcriptional modulation at multiple genetic loci in plants [89]. By targeting dCas9-VP64 to a highly-methylated promoter region, a 400-fold increase in mRNA expression of the imprinted gene, *AtFIS2*, occurred in Arabidopsis rosette leaves. The result demonstrated that methylated DNA, difficult to target with TAL proteins [117], is targetable by

CRISPR-Cas9 (Table 1). Although these results are exciting, they merely represent the first generation of such activators and repressors. Further improvement of CRISPR-dCas9 based transcriptional regulation systems for high efficiency in plants is anticipated.

Current Perspectives

CRISPR-Cas9 has been widely adopted for basic and applied research and as efficiency improves will continue as a popular tool. Currently, gene targets are somewhat limited by the NGG PAM site required by SpCas9 [38] (Table 1). However, target ranges will broaden as more systems are further explored. Orthogonal Cas9s have garnered attention for their unique PAM sites and gRNA structure, creating the possibility of expressing multiple Cas9s and gRNAs in a cell without interference. These orthogonal Cas9 variants differ in size and specificity as well as PAM sequences. Some of the most promising are NmCas9, StCas9 and SaCas9, all of which have been demonstrated in human cells [118] and the latter two in plants [119–122]. A CRISPR-Cpf1 system was reported in 2015 and it differs from the Cas9 system on several key parameters [123]. Cpf1 requires only a crRNA, making the gRNA 42 nt instead of ~100 nt for Cas9. The Cpf1 PAM is TTTN and cleavage results in 5' overhangs distall from protospacer elements. A shorter gRNA is easier to synthesize and an overhang may improve efficiency for NHEJ based gene insertion if the insert is designed with a complementary overhang. Lastly, the location of the DSB means that any indels will likely fail to disrupt the PAM site, leaving the possibility for multiple Cpf1 targeting events and allowing a second chance for gene

insertion should the first attempt fail. Reports of Cpf1 in plants have also been published recently [124–130]. The CRISPR-Cpf1 system developed by Tang et.al achieved 100% mutagenesis frequency at all target sites in rice [126], demonstrating promising applications of Cpf1 in plants.

DNA independent delivery of SSNs for plant genome editing is another trend. Development of such methods are likely motivated for use in crop improvement in regards to regulation [2]. Nucleic-acid free delivery of TALEN has been successfully accomplished [130]. This study demonstrated that delivery of pure nuclease protein into protoplasts was possible albeit at a low frequency. DNA-free delivery of Cas9 was accomplished by incubating Arabidopsis, rice, tobacco, and lettuce protoplasts with Cas9/gRNA ribonucleoprotein complexes [131]. Bread wheat was shown to be amenable to genome editing based on mRNA or ribonucleoprotein delivery of CRISPR-Cas9 [74, 132]. More recently, ribonucleoprotein delivery of CRISPR-Cpf1 was also demonstrated in soybean and wild tobacco protoplasts [127].

Genome editing may be achieved without introducing DNA DSBs. DNA base editing tools based on fusing cytidine deaminase to Cas9n or dCas9 were first demonstrated in human cells [133, 134]. Encouragingly, this technology was recently shown to work in rice [135–139], Arabidopsis [140], tomato [139], maize and wheat [137]. Without question, first generation base editing tools will be further expanded, improved and applied in many other plant species soon. Finally, as genome editing moves ahead into many crop plants, improving transformation and tissue culture methods will be critical for success. A recent report of using *Baby boom* and *Wuschel*

genes to improve transformation efficiency in recalcitrant monocot plants set an exciting example of this endeavor [141].

Challenges of using CRISPR for plant genome engineering

CRISPR has rapidly been adopted for use in research and product development. Compared to earlier genome engineering systems, CRISPR is a much faster and easier to use molecular tool. However, there is much room for improvement ten years after CRISPR's debut and increased demand will likely drive innovation of new applications for CRISPR. For plants, editing efficiency is an especially important consideration. Many important crops such as strawberry, wheat, potato, and others are polyploid and increases in editing efficiency are important for targeting all copies of a gene. It is also important to be able to target different PAM sites. Promoters are AT-rich meaning that many targets are inaccessible with Cas9, which limits promoter engineering for making changes in the genome without creating changes in protein coding genes.

In order to increase editing efficiency and expand applications, natural variants of Cas proteins are identified from prokaryotes and tested for unique characteristics. Cas proteins can also be engineered. Mutations can create Cas nickases for applications in base editing and prime editing, as well as nucleases capable of targeting new PAM sequences [142–147]. Recently, the addition of effector proteins such as activators and exonucleases have been demonstrated to increase editing efficiency [148–151].

In addition to changes to the endonuclease, guide RNA engineering has also been pursued for improving CRISPR-Cas mediated genome editing efficiency.

Addition of a 5' cap and poly A tail have been shown to greatly increase the efficiency of editing in human cell lines by decreasing sgRNA degradation [152]. Highly precise gRNA processing that prevents the addition of bases can increase efficiency through production of accurate gRNA sequences that avoid mismatch, especially for Cas12a [126]. Expression of CRISPR systems in the plant affects the efficiency. Mimicking what occurs in the plant nucleus, the first studies used a Pol II promoter to drive protein expression and a Pol III for sgRNA expression. The substitution of endogenous RNA promoters for Pol II protein promoters can increase sgRNA production, which suggests that low expression of gRNAs can result in a low efficiency CRISPR system [153]. The problem of low expression becomes even more detrimental in polyploid plants and in multiplexed editing, one of the highlights of the CRISPR systems over previous SSNs.

Table 2. TALEN mediated genome editing in plants

Plant species	Target gene	Modification	Reference
Arabidopsis	*ADH1, TT4, MAPKKK1, DSK2B, NATA2, GLL22a, GLL22b*	NHEJ	[31, 44]
Arabidopsis	*CLV3*	NHEJ	[134]
Arabidopsis	*CRU3*	NHEJ	[135]
Barley	*HvPAPhy_a*	NHEJ	[136]
Barley	*GFP (transgene)*	NHEJ	[137]
Barley	*GFP (transgene)*	HDR	[92]
Brachypodium	*ABA1, CKX2, SMC6, SPL, SBP, COII, RHT, HTA1*	NHEJ	[46]
Maize	*GL2*	NHEJ	[138]
Maize	*IPK1A, IPK, MRP4*	NHEJ	[139]
N. benthamiana	*FucT, XylT*	NHEJ	[140]
N. tabacum	*ALS*	NHEJ, HDR	[43]
Potato	*Vlnv*	NHEJ	[52]
Potato	*ALS*	NHEJ	[141, 142]
Rice	*11N3*	NHEJ	[48]
Rice	*DEP1, BADH2, CKX2, SD1*	NHEJ	[46, 50]
Rice	*EPSPS*	NHEJ	[143]

Rice	*MST7, MST8, PMS3, CSA, DERF1*	NHEJ	[47]
Rice	*LOX3*	NHEJ	[51]
Rice	*ALS*	HDR	[93]
Rice	*SWEET14*	NHEJ	[144]
Rice	*WAXY*	NHEJ	[145]
Soybean	*FAD2-1A, FAD2-1B, FAD3A*	NHEJ	[49, 146]
Soybean	*PDS11, PDS18*	NHEJ	[147]
Sugarcane	*COMT*	NHEJ	[148]
Tomato	*PROCERA*	NHEJ	[45]
Tomato	*ANT1*	HDR	[94]
Wheat	*MLO*	NHEJ	[7]

Table 3. CRISPR-Cas9 mediated genome editing in plants

Plant species	Target gene	Modification	Reference
Arabidopsis	*PDS3, FLS2, RACK1b, RACK1c*	NHEJ	[96]
Arabidopsis	*BRI1, GAI, JAZ1*	NHEJ	[149, 150]
Arabidopsis	*CHLI1, CHLI2, TT4, AP1, GL2*	NHEJ	[150-152]
Arabidopsis	*GFP (transgene)*	NHEJ	[153, 154]
Arabidopsis	*ADH1, TT4, RTEL, GUS (transgene)*	NHEJ, HDR	[85, 86]
Arabidopsis	*FT, SPL4*	NHEJ	[155]
Arabidopsis	*ABP1*	NHEJ	[71]
Arabidopsis	*Cru3*	NHEJ	[156]
Arabidopsis	*TRY, CPC, ETC2, CHIL1, CHIL2*	NHEJ	[79, 157]
Arabidopsis	*1g03180, 1g16210, 1g56650, 5g55580*	NHEJ	[80]
Arabidopsis	*05g55580, 1g56650, 1g03180, 1g16210*	NHEJ	[80]
Arabidopsis	*PHYB, BRI1*	NHEJ	[123]
Arabidopsis	*BRI1, PDS3*	NHEJ	[158]
Arabidopsis	*PYR1, PYL1, PYL2, PYL4, PYL5, PYL8*	NHEJ	[82]
Arabidopsis	*SH3P3*	NHEJ	[159]
Arabidopsis	*eIF(iso)4E*	NHEJ	[57]
Arabidopsis	*CBF1, CBF2, CBF3*	NHEJ	[63, 160]
Arabidopsis	*DM2*	NHEJ	[161]
Arabidopsis	*UGT79B2, UGT79B3*	NHEJ	[162]
Arabidopsis	*CWIN1*	NHEJ	[163]
Arabidopsis	*MIR169a, MIR827a, TFL1*	NHEJ, HDR	[77]
Arabidopsis	*TTG1*	NHEJ	[164]
Barley	*HvPM19*	NHEJ	[165]
Cabbage	*BoIC.GA4.a*	NHEJ	[165]

Camelina	*FAD2*	NHEJ	[69, 70]
C. reinhardtii	*CpFTSY, ZEP*	NHEJ	[166]
Cotton	*GFP (transgene)*	NHEJ	[167]
Cotton	*MYB25-like A, MYB25-like D*	NHEJ	[67]
Cotton	*CLA1, VP*	NHEJ	[68]
Dandelion	*1-FFT*	NHEJ	[168]
Flax	*EPSPS, BFP (transgene)*	NHEJ, HDR	[169]
Grape	*IdnDH*	NHEJ	[170]
Lettuce	*BIN2*	NHEJ	[123]
Liverwort	*ARF1*	NHEJ	[171]
Lotus japonicus	*SYMRK, LjLb1, LjLb2, LjLb3*	NHEJ	[172]
Maize	*IPK*	NHEJ	[139]
Maize	*LIG1, Ms26, Ms45, ALS1, ALS2*	NHEJ, HDR	[173]
Maize	*PSY1, and other 90 loci*	NHEJ	[174]
Maize	*ZB7, 2g332562, 2g080129, 2g099580, 2g170586, 2g438243,*	NHEJ	[175]
Maize	*ARGOS8*	NHEJ	[176]
Maize	*AGO18a, Ago18b, a1, a4*	NHEJ	[177]
Moss	*PpAPT*	NHEJ, HDR	[178]
Moss	*PpKAI2L, PpAP2/ERF*	NHEJ	[179]
N. oceanica	*NR*	NHEJ	[180]
N. attenuata	*AOC*	NHEJ	[123, 159]
N. benthamiana	*PDS3*	NHEJ, HDR	[96]
N. benthamiana	*PDS*	NHEJ	[181-183]
N. benthamiana	*PCNA, PDS*	NHEJ	[60]
N. benthamiana	*FLS2, BAK1*	NHEJ	[81]
N. benthamiana	*PDS, blspH*	NHEJ	[184]
N. benthamiana	*XT1, XT2*	NHEJ	[185]
N. benthamiana	*EDS1a, PAD4*	NHEJ	[161]
N. tabacum	*GFP (transgene)*	NHEJ	[153]
N. tabacum	*PDS, PDR6*	NHEJ	[186]
N. tabacum	*mCherry (transgene)*	NHEJ	[187]
Petunia	*PDS*	NHEJ	[188]
Petunia	*NR*	NHEJ	[189]
Populus	*4CL1, 4CL2, 4CL5*	NHEJ	[64]
Populus	*PDS*	NHEJ	[190, 191]
Potato	*IAA2*	NHEJ	[192]
Potato	*ALS*	NHEJ	[142, 193]
Potato	*GBSS*	NHEJ	[194]
Potato	*MYB44*	NHEJ	[195]
Rice	*PDS, BADH2, MPK2, 02g23823*	NHEJ	[65, 83]
Rice	*MPK5*	NHEJ	[196]
Rice	*ROC5, SPP, YSA*	NHEJ	[149, 197]

Rice	*MYB1*	NHEJ	[151, 197]
Rice	*DERF1, EPSPS, MSH1, PDS, PMS3*	NHEJ	[197]
Rice	*SWEET11*	NHEJ	[198]
Rice	*SWEET11, SWEET14*	NHEJ	[153]
Rice	*CAO1, LAZY1*	NHEJ	[199]
Rice	*BEL*	NHEJ	[200]
Rice	*SWEET11, SWEET13, SWEET1a, SWEET1b, CPS4, CYP99A2, CYP76M5, CYP76M6, KO1, KOL5*	NHEJ	[76]
Rice	*CDKA2, CDKB1, CDKB2*	NHEJ	[201]
Rice	*MPK1, MPK2, MPK5, MPK6, PDS*	NHEJ	[202]
Rice	*ALS*	HDR	[97, 98]
Rice	*GSTU, MRP15, ANP, WAXY, 7 FTL genes, and 21 other genes*	NHEJ	[80]
Rice	*AOX1a, AOX1b, AOX1c, BEL*	NHEJ	[203]
Rice	*DsRed (transgene), YSA, PDS, DL*	NHEJ	[204, 205]
Rice	*P450, DWD1*	NHEJ	[123]
Rice	*RAV2*	NHEJ	[78]
Rice	*DMC1A, DMC1B*	NHEJ	[87]
Rice	*NAL1, LPA1, LG1, GL1-1*	NHEJ	[206]
Rice	*DEP1, ROC5*	NHEJ	[207]
Rice	*Gn1a, DEP1, GS3, IPA1*	NHEJ	[53]
Rice	*ERF922*	NHEJ	[56]
Rice	*OST2*	NHEJ	[208]
Rice	*CSA*	NHEJ	[54]
Rice	*RUPO*	NHEJ	[209]
Rice	*EPSPS*	NHEJ, HDR	[210]
Rice	*TMS5*	NHEJ	[55]
Rice	*PMR*	NHEJ	[211]
Rice	*MEGs, PEGs*	NHEJ	[212]
Rice	*Hd2, Hd4, Hd5*	NHEJ	[213]
Rice	*SBEI, SBEIIB*	NHEJ	[214]
Rice	*ACT, GST*	HDR	[99]
Rice	*RBOHH*	NHEJ	[215]
Rice	*EPFL9*	NHEJ	[121]
Salvia miltiorrhiza	*CPS1*	NHEJ	[216]
Sorghum	*DsRED2 (transgene)*	NHEJ	[153]

Soybean	*GFP (transgene), 07g14530, 01g38150, 11g07220, miR1514, miR1509*	NHEJ	[217]
Soybean	*06g14180, 08g02290, 09g00490, 12g37050*	NHEJ	[218]
Soybean	*PDS11, PDS18*	NHEJ	[147]
Soybean	*DD20, DD43, ALS*	NHEJ, HDR	[219]
Soybean	*FEI1, FEI2, SHR, bar (transgene)*	NHEJ	[220]
Soybean	*Rj4*	NHEJ	[72]
Sweet orange	*PDS*	NHEJ	[221]
Sweet orange	*LOB1*	NHEJ	[222]
Tomato	*SHR, GFP (transgene)*	NHEJ	[223]
Tomato	*AGO, 08g041770, 07g021170, 12g044760*	NHEJ	[224]
Tomato	*RIN*	NHEJ	[225]
Tomato	*PDS, PIF4*	NHEJ	[226]
Tomato	*SIAGL6*	NHEJ	[227]
Tomato	*SP5G*	NHEJ	[228]
Tomato	*SIBOP*	NHEJ	[229]
Tomato	*SIIAA9*	NHEJ	[230]
Tomato	*MLO*	NHEJ	[231]
Wheat (Common)	*MLO*	NHEJ	[7, 65]
Wheat (Common)	*INOX*	NHEJ	[183]
Wheat (Common)	*GASR7, GW2, DEP1, NAC2, PIN1, LOX2,*	NHEJ	[66]
Wheat (Common)	*Ubi, MLO*	HDR	[100]
Wheat (Durum)	*GASR7*	NHEJ	[66]

Research Objectives

This thesis work covers expansion of LbCas12a for genome editing and multiplexed transcriptional repression in Arabidopsis and engineering of Cas9 and Cas12a-effector fusion proteins to influence NHEJ outcomes. This work is further divided into several subaims.

Aim 1 - LbCas12a genome engineering in Arabidopsis: temperature sensitive editing and multiplexed transcriptional repression

Sub-aim 1- Comparing LbCas12a genome editing efficiency at different temperatures

Sub-aim 2- Exploring temperature sensitive mechanism through gene repression

Sub-aim 3- Demonstrating transcriptional repression with multiplexed dLbCas12a-SRDX systems

Aim 2 – Improving NHEJ outcomes in rice with Cas9 and Cas12a fusions

Sub-aim 1- Improving genome editing with chromatin modulating peptides

Sub-aim 2- Improving genome editing with exonucleases

Chapter 2: LbCas12a genome engineering in Arabidopsis: temperature sensitive editing and multiplexed transcriptional repression

The work presented in this chapter has been published and can be found in the following sources:

Malzahn AA, Tang X, Lee K, Ren Q, Sretenovic S, Zhang Y, Chen H, Kang M, Bao Y, Zheng X, Deng K, Zhang T, Salcedo V, Wang K, Zhang Y, Qi Y (2019) Application of CRISPR-Cas12a temperature sensitivity for improved genome editing in rice, maize, and Arabidopsis. BMC Biology 17:9. doi: 10.1186/s12915-019-0629-5

Lowder LG, Malzahn A, Qi Y (2018) Plant Gene Regulation Using Multiplex CRISPR-dCas9 Artificial Transcription Factors. In: Maize: methods and protocols. Springer Science+Business Media, New York, NY, pp 197–214

Zhang Y, Ren Q, Tang X, Liu S, Malzahn AA, Zhou J, Wang J, Yin D, Pan C, Yuan M, Huang L, Yang H, Zhao Y, Fang Q, Zheng X, Tian L, Cheng Y, Le Y, McCoy B, Franklin L, Selengut JD, Mount SM, Que Q, Zhang Y, Qi Y (2021) Expanding the scope of plant genome engineering with Cas12a orthologs and highly multiplexable editing systems. Nat Communications 12:1944. doi: 10.1038/s41467-021-22330-w

2.1 Introduction

CRISPR endonucleases are diverse with unique features

CRISPR-based genome engineering was pioneered with the *Streptococcus pyogenes* endonuclease CRISPR Associated 9 (SpCas9) as the "molecular scissor" inducing double stranded DNA breaks or changes in transcription. However, there are a range of different Cas endonucleases with unique characteristics. Databases such as CasPDB have identified endonucleases from sequence databases, which can then be tested *in vivo* to quantify features such as editing efficiency, deletion profile, and PAM specificity [154]. One endonuclease, Cas12a, formerly designated Cpf1, has emerged as an important expansion to the CRISPR toolbox.

Cas12a expands the CRISPR toolbox

Genome engineering has focused on effective Cas12a variants that have been identified in three species of bacteria: *Francisella novicida* (FnCas12a), *Lachnospiraceae bacterium* (LbCas12a), and *Acidaminococcus* sp. *BV3L6* (AsCas12a) [42]. These variants have the same general structure, but have differences in features such as editing efficiency, which has been demonstrated with SpCas9 and other Cas9 variants [42]. Cas12a is structurally different from Cas9 (Figure 1) [42, 123] and has several unique characteristics: endogenous crRNA processing, a T-rich PAM, and formation of distal DSBs with nucleotide overhangs.

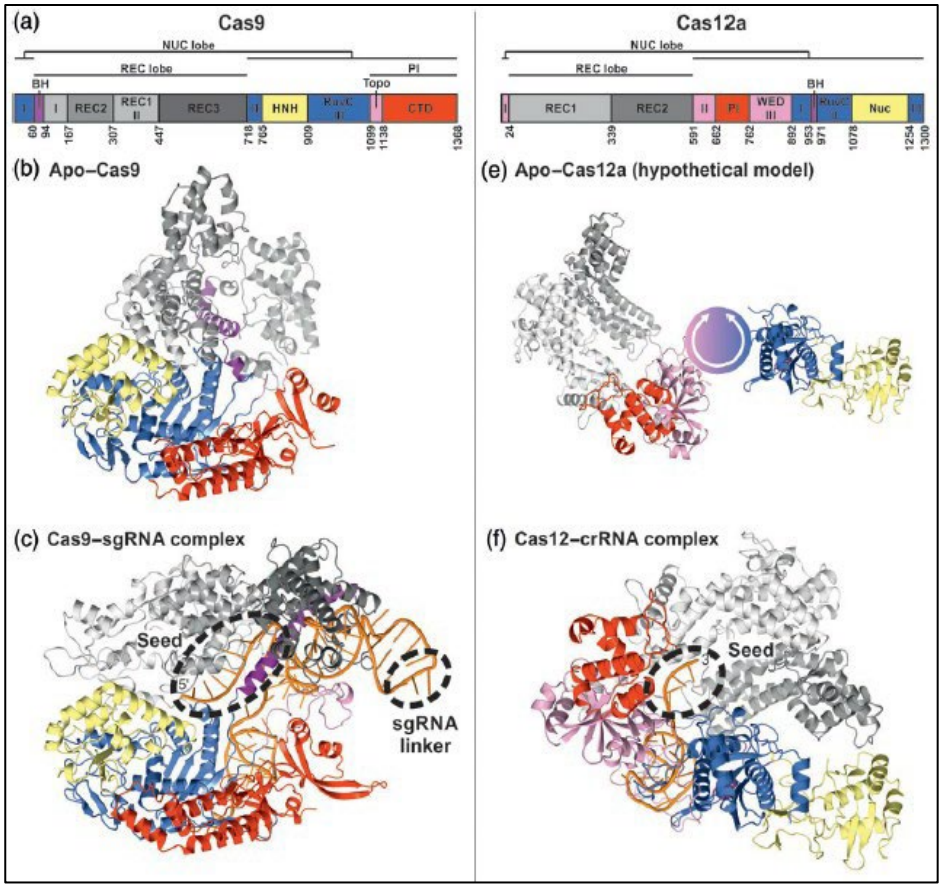


Figure 1. **Architecture of SpCas9 and FnCas12a**. (A) Schematic of Cas9 and Cas12a genes with labeled domains. Structure of Cas9 (B) and Cas12a (E) proteins not bound to a sgRNA. Cas9 (C) and Cas12a (F) in complex with sgRNA or crRNA. Figure from Swarts and Jinek [4].

For the first unique Cas12a characteristic, the wedge (WED) domain controls crRNA processing, an endogenous mechanism not found in Cas9 (Figure 1) [123]. This means Cas12a is uniquely capable of processing its own pre-crRNA into mature crRNA. This feature negates the need for processing systems such as ribozymes or tRNA and provides space in the delivery vector, which may be vital to increased multiplexing systems or viral vectors that have a limited size.

The second unique characteristic, the T-rich PAM, is caused by the binding channel formed by the PI, RecI and WED domains collectively. Hydrogen bonds and Van der Waals interactions in these three domains facilitate binding between Cas12a and the minor groove of the T-rich PAM 5'-TTTV-3', where V= A,G, or C. The binding channel can be slightly opened during PAM binding which results in an optimal and suboptimal PAM for Cas12a endonucleases [155]. This expands the number of available targets editable by Cas12a, which is useful even at lower efficiencies.

For cleavage, Cas12a relies only on RuvC to make double-stranded breaks in DNA, although the Nuc domain contributes to binding. Mutations in the Nuc domain affect target strand cleavage, but do not result incomplete loss of cleavage [47, 48]. The most recent hypothesis is that RuvC cleaves the target strand, which triggers conformational changes in the REC and NUC lobes and possibly requires DNA unwindingpast the crRNA-DNA duplex. The non-target strand is then in place to be cleaved by the RuvC domain [47]. This results in an overhanging double stranded break 18 bp downstream from the PAM on the target strand and 23 bp away on the non-target strand. This overhang results in a different deletion profile from Cas9 DSB. Cas12a has been shown to commonly result in 6-15 bp deletions [49, 50]. Because the DSB is distal to the PAM, Cas12a can theoretically target the same site multiple times, contrary to Cas9 where a single mutation next to PAM would prevent further Cas9 sgRNA binding Although it has not been achieved in plants, the overhangs produced by Cas12a have been viewed as a potential

strategy for NHEJ gene insertion. If the insert had complementary overhangs, nucleotide binding would facilitate insertion instead of the tempospatial array of proteins involved in HDR. Cas12a's unique features (T-rich PAM, distal DSB overhangs, short crRNA) make it an important expansion to the CRISPR toolbox.

While Cas12a's desirable characteristics are interesting, there is the problem of editing efficiency. High efficiency editing has been achieved in rice, but not in other plants [126]. Gene repression has also been achieved in plants, but not with multiplexed systems. In order to capitalize on Cas12a's unique features, we tested LbCas12a in Arabidopsis and developed methodology to increase editing efficiency [156]. Our results indicate that LbCas12a is temperature sensitive but editing efficiency can be rescued by heat-treating Arabidopsis in controlled conditions. In addition to editing, transcriptional repression was demonstrated at three different temperatures, indicating that Cas12a binding is unaffected by temperature and another mechanism(s) is responsible for low efficiency editing at ambient temperatures. Two multiplexed crRNA processing systems were compared to further explore LbCas12a transcriptional repression. No significant difference was observed between the crRNA processing systems, indicating that simpler strategies for crRNA processing can be used effectively.

In summary, this work presents a significant expansion of the CRISPR toolkit by providing methodology for Cas12a-based editing and gene repression in Arabidopsis, a well-known plant model organism. This provides

plant researchers with a new multiplexing endonuclease to target T-rich PAMs untargetable by Cas9, and produce unique editing outcomes. Researchers can obtain our vectors for use in their own work via Addgene [157] and step-by-step instructions for toolkit assembly can be found in book chapters in the Methods in Molecular Biology series [158].

2.2 Results

2.2.1 Comparing LbCas12a genome editing efficiency at different temperatures

SpCas9 (*Streptococcus pyogenes* Cas9) can target sequences with an NGG PAM, which limits the ability to target T-rich sequences. For example, in rice, the addition of Cas12a to the CRISPR toolkit increased the number of target sites available for genome editing by 14 million TTTV PAM sites (Lb and FnCas12a) and 19 million VTTV PAM sites (FnCas12a only) [155]. The application of Cas12a endonucleases with T- rich PAMs expands the available target sites. Three variants have been tested in plants (FnCas12a, LbCas12a, and AsCas12a), but studies remain limited [126].

Rice, an important crop and monocot model organism, was successfully edited with Cas12a but no genome editing studies had been conducted in Arabidopsis, which is an important dicot model organism [124–126]. In order to test LbCas12a in Arabidopsis, wild type Arabidopsis thaliana (var. Columbia) was transformed via floral dip with two separate T-DNA vectors that contained a unique crRNA. A dual Pol II promoter system (a ZmUbi promoter driving LbCas12a and a second ZmUbi

promoter driving crRNA expression) was used to drive one crRNA. The dual Pol II promoter system previously produced high efficiency LbCas12a editing in rice [126]. Two target genes were tested: *Gl2 (GLABRA 2)* and *TT4 (TRANSPARENT TESTA 4)*. Given our previous success in rice, we were surprised to find that none of the Arabidopsis T1 plants harbored mutations from LbCas12a (data not shown). We hypothesized that LbCas12a's DNA cleavage ability was temperature dependent, which would explain why rice, which is grown at 32 °C, was mutated but not Arabidopsis which is grown at 22 °C.

In order to determine whether LbCas12a was temperature sensitive, select T2 lines underwent temperature treatment at 29 °C with control plants grown at 22 °C for comparison. These T2 lines were grown on MS-hygromycin selection medium at 22 °C and 29 °C for 2 weeks, and then individual plants were sacrificed for DNA extraction, followed by Restriction Fragment Length Polymorphism (RFLP)-based mutation analysis (Fig. 2). RFLP relies on an overlap between the area targeted for deletion and a restriction enzyme site. Altered sequences will not be cut during digestion, unlike a wild-type sequence, providing a method for estimating editing efficiency on a gel via a software such as Bio-Rad's Image Lab [159]. GL2 and TT4 T2 lines all revealed detectable mutation frequencies at 29 °C, but not at 22 °C (Fig. 2). The highest mutation percentage of 35% was produced at 29 °C in GL2 T2 line #8, followed by GL2 #3 at 12%, GL2 #9 at 10%, and GL2 #2 at 7% (Fig. 2b). Similarly, while two examined T2 lines (#2 and #7) for TT4 showed mutation frequencies of about 14% and 15% at 29 °C, no mutation could be detected at 22 °C (Fig. 2 c,d).

To further increase the editing efficiency, plants from untested T2 lines GL2 #7 and TT4 #9 underwent a longer heat treatment at 29 °C for 29 days before being moved to 22 °C for recovery. T3 plants from these T2 parents were grown in a greenhouse at ~ 22 °C to demonstrate stable germline editing and evaluate mutation frequencies. Leaf tissues were collected from the three lines (GL2 #7-4, GL2 #7-7, and TT4 #9-7) for RFLP and Sanger sequencing (Fig. 3a, b). Thirty-three percent of the plants from GL2 #7-4, 21% of plants from GL2 #7-7, and 5.8% from TT4 #9-7 had targeted mutations (Fig. 3a).

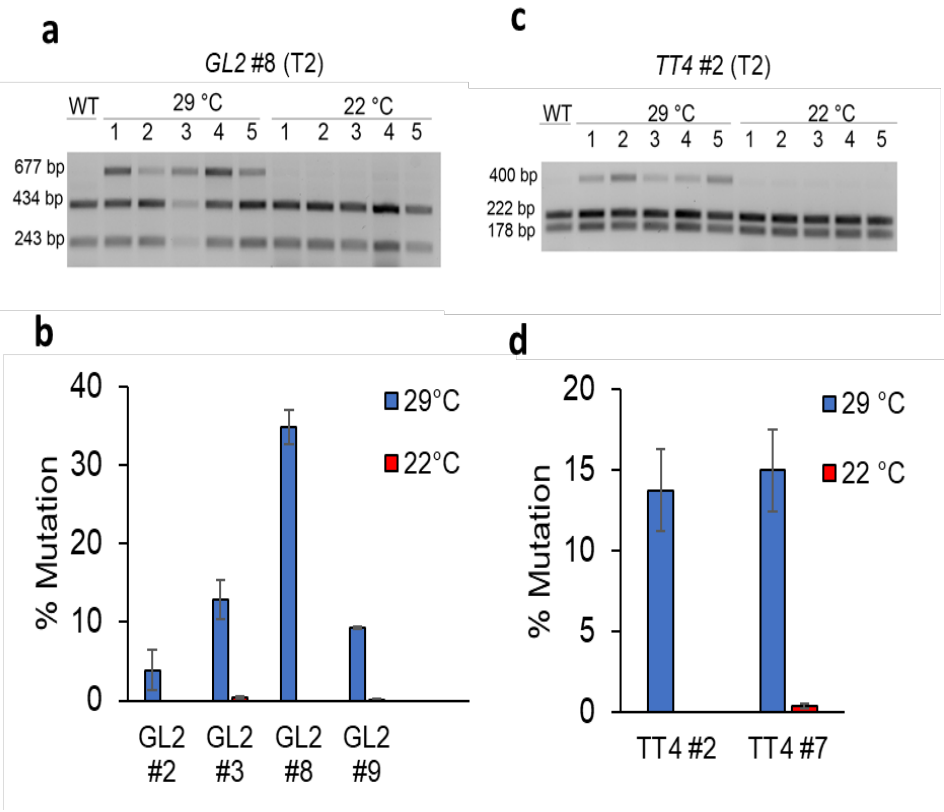


Figure 2. LbCas12a is temperature sensitive. A) Gel image comparing a wild-type sample to five individual samples from LbCas12a- GL2 plants, progeny from line #8, grown in 29 °C and 22 °C temperature regimes. B) Summary of estimated mutation percentages of four LbCas12a-GL2 lines. C) Gel image comparing a wild-type sample to five individual samples from LbCas12a-TT4 plants, from line #2, grown in 29 °C and 22 °C temperature regimes. D) Summary of estimated mutation percentages of two LbCas12a-TT4 lines. Figure from Malzahn et al. [44].

Among T3 lines from GL2 #7-4, the same mutation (a deletion of 5 bp) was found in all the homozygous and heterozygous plants, suggesting germline transmission from the parent (Fig. 3b). All homozygous GL2 #7-4 plants had a trichome-less (gl2 loss-of-function) phenotype (Fig. 3c). For GL2 #7-7 T3 plants, only GL2 #7-7-1 showed a loss-of-function phenotype. Sequencing analysis showed that the other homozygous or biallelic plants (e.g., GL2 #7-7-14 and GL2 #7-7-17) had the same allele of 3-bp deletion, suggesting the mutated protein, albeit missing one amino acid, is still functional. LbCas12a seems less efficient at the TT4 locus since only two heterozygous mutants were identified out of 34 TT4 #9-7 T3 lines (Fig. 3a, b). Together, we demonstrated LbCas12a-mediated germline editing in Arabidopsis with a high-temperature treatment.

A

Target gene + line (T3 population)	Mutation Rates	Genotype
GL2 #7-4	33.3% (7 out of 21)	6 Ho/1 He/13 W
GL2 #7-7	21.0% (4 out of 19)	2 Ho/1 He/1 B/14 W
TT4 #9-7	5.8% (2 out of 34)	2 He/ 32 W

B

GL2 #7-4-7 WT

C

```
T3 generation GL2 #7-4
WT (GL2): AAGACATGTCGACGGCCATTGACATACAAATCC
GL2 #7-4-7
Allele 1: AAGACA-----ACGGCCATTGACATACAAATCC -5 bp
Allele 2: AAGACA-----ACGGCCATTGACATACAAATCC -5 bp
GL2 #7-4-19
Allele 1: AAGACATGTCGACGGCCATTGACATACAAATCC WT
Allele 2: AAGACA-----ACGGCCATTGACATACAAATCC -5 bp
GL2 #7-4-23
Allele 1: AAGACA-----ACGGCCATTGACATACAAATCC -5 bp
Allele 2: AAGACA-----ACGGCCATTGACATACAAATCC -5 bp
GL2 #7-4-24
Allele 1: AAGACA-----ACGGCCATTGACATACAAATCC -5 bp
Allele 2: AAGACA-----ACGGCCATTGACATACAAATCC -5 bp
GL2 #7-4-25
Allele 1: AAGACA-----ACGGCCATTGACATACAAATCC -5 bp
Allele 2: AAGACA-----ACGGCCATTGACATACAAATCC -5 bp
GL2 #7-4-26
Allele 1: AAGACA-----ACGGCCATTGACATACAAATCC -5 bp
Allele 2: AAGACA-----ACGGCCATTGACATACAAATCC -5 bp
GL2 #7-4-28
Allele 1: AAGACA-----ACGGCCATTGACATACAAATCC -5 bp
Allele 2: AAGACA-----ACGGCCATTGACATACAAATCC -5 bp

T3 generation GL2 #7-7
GL2 #7-7-1
Allele 1: AA--------------------------TCC -28 bp
Allele 2: AA--------------------------TCC -28 bp
GL2 #7-7-11
Allele 1: AAGACATGTCGACGGCCATTGACATACAAATCC WT
Allele 2: AA--------------------------TCC -28 bp
GL2 #7-7-14
Allele 1: AAGACATG---ACGGCCATTGACATACAAATCC -3 bp
Allele 2: AAGACATG---ACGGCCATTGACATACAAATCC -3 bp
GL2 #7-7-17
Allele 1: AAG-------------------------TCC -28 bp
Allele 2: AAGACATG---ACGGCCATTGACATACAAATCC -3 bp

T3 generation TT4 #9-7
WT (TT4): GTATTTACTATTCACAGGCGACAAGTCGACAAT
TT4 #9-7-3
Allele 1: GTATTTACTATTCACAGGCGACAAGTCGACAAT WT
Allele 2: GTATTTACTATTCACAGGCG-----TCGACAAT -5 bp
TT4 #9-7-31
Allele 1: GTATTTACTATTCACAGGCGACAAGTCGACAAT WT
Allele 2: GTATTTACTATTCACAGGCG-----TCGACAAT -5 bp
```

Figure 3. Temperature treatment rescues LbCas12a efficiency for germline editing. A) A summary of genotyping results from GL2 and TT4 T3 generation. B) Image of wild-type and gl2 mutant plants. Mutant plants lack trichomes. C) Sequences of mutants from the T3 generation. Figure from Malzahn et al. [3].

We also tested a late treatment of transgenic plants at 29 °C which was less effective in inducing LbCas12a-mediated mutations at both target sites. The treatment consisted of germination at 22 °C on MS media. After 6 days, the plants were kept at 29 °C for 8 days. After 24 days of recovery at 22 °C, plants were kept at 29 °C again for 14 days. Leaf tissue was collected and mutation percentages were estimated via

RFLP analysis. No mutations were observed in the TT4 lines tested (Fig. 4b). Mutations were observed in both GL2 lines, but at very low efficiencies (Fig. 4a). Since low mutation percentages were observed in the late heat treatment compared to previous heat treatment during the first month of the plant's life cycle, the data suggests that there is an early window to effectively induce mutations. Early heat treatment could induce mutations in young cells that divide into many other cells, meaning that even a few mutated cells could have a significant impact.

This work not only demonstrates the temperature sensitivity of LbCas12a in Arabidopsis, but also highlights that heat treatment can rescue the efficiency and produce germline mutations, making LbCas12a an important expansion to the plant CRISPR toolkit. Following a heat treatment regime, researchers can successfully use LbCas12a for genome editing. This CRISPR- Cas12a toolkit is publicly available at Addgene for other's use [157]. In addition to temperature, timing impacts editing efficiency. Timing heat treatments earlier in the life cycle may be more effective than late heat treatments. Further research on Arabidopsis genome editing at higher temperatures and with different cycles of heating and recovery can be expected to produce better methodology that results in high efficiency editing without sacrificing plant health and seed yield.

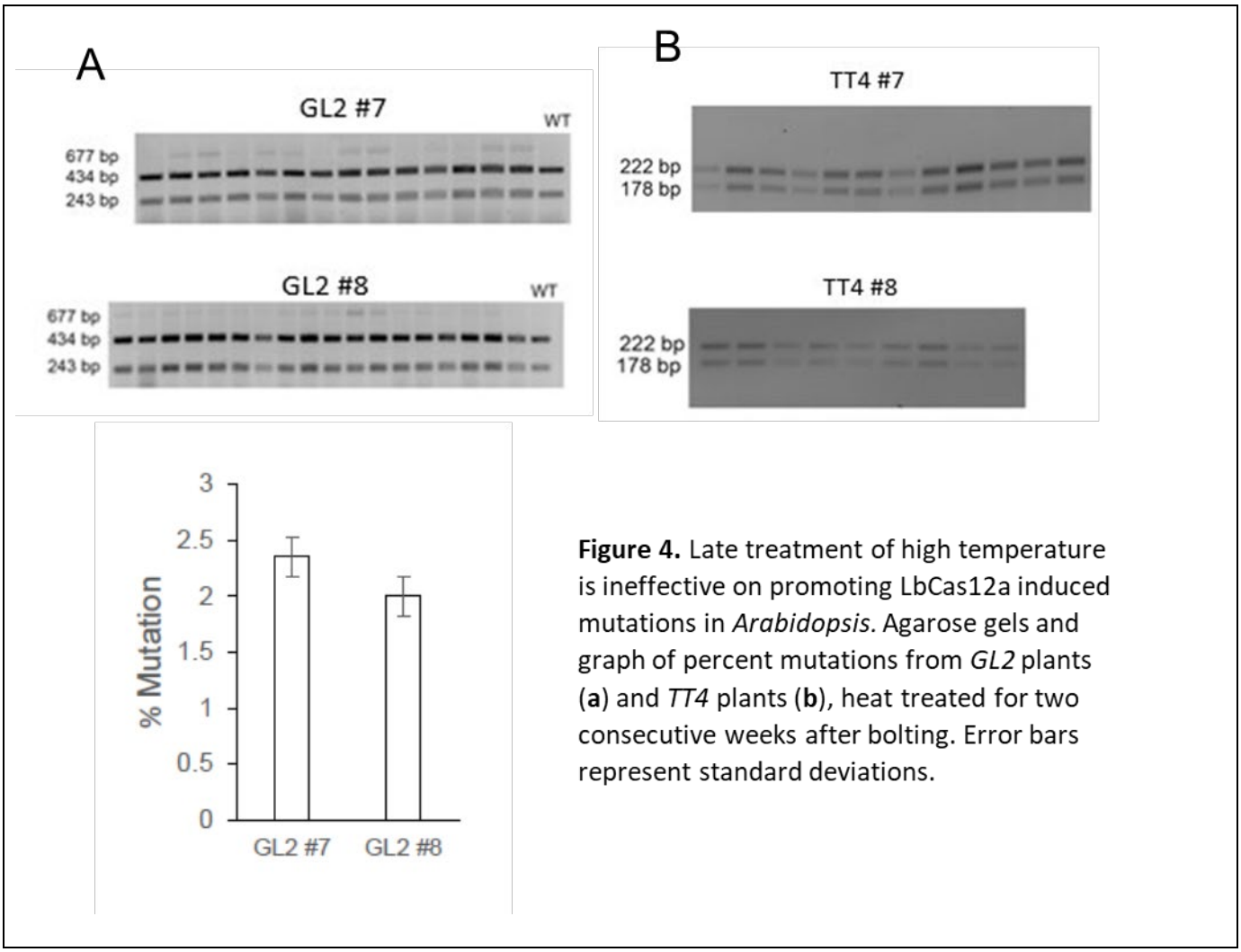


Figure 4. Late treatment of high temperature is ineffective on promoting LbCas12a induced mutations in *Arabidopsis*. Agarose gels and graph of percent mutations from *GL2* plants (**a**) and *TT4* plants (**b**), heat treated for two consecutive weeks after bolting. Error bars represent standard deviations.

2.2.2 Exploring LbCas12a temperature sensitivity through gene repression

While heat treatment can rescue LbCas12a efficiency, there are some plant species that may not be able to tolerate heat treatment. Previously, we successfully applied LbCas12a for repression at a low temperature in Arabidopsis, therefore, some functions of LbCas12a were not wholly dependent on temperature [126]. Determining what mechanism of LbCas12a cleavage is dependent on high temperature would pave the way for protein

engineering of the endonuclease. An engineered enzyme would not only be widely applicable to plants but to zebrafish, Drosophila, Xenopus, and other model organisms that grow best at ambient temperatures.

In order to explore the mechanism behind LbCas12a temperature sensitivity, catalytically dead dLbCas12a-SRDX was targeted to PAP1 (PRODUCTION OF ANTHOCYANIN PIGMENT 1) for repression with Arabidopsis plants grown for a week at 16 °C, 22 °C, and 29 °C. Repression was achieved at all three temperatures, suggesting that Cas12a binding was not the mechanism behind our temperature sensitive genome editing results (Fig. 5) [156]. This has been supported by experiments in zebrafish that suggest binding and cleavage are not temperature sensitive; potentially chromatin blockage or another uncharacterized mechanism is [160]. Before Cas12a can cleave DNA, it must scan the genome, find a PAM and crRNA complimentary sequence, unwind the DNA, and make conformational changes that activate nuclease activity [161, 162]. Temperature may affect any one, or all, of these steps. Although testing all these in vivo will be challenging, we reasoned that we could test whether Cas12a in planta binding to DNA is affected by temperature. Given the contrasting differences of LbCas12a editing activities in Arabidopsis at 22 °C and 29 °C, we decided to test this in Arabidopsis by using a dLbCas12a-SRDX repressor that we had previously developed [48]. We constructed a dLbCas12a-SRDX vector to target the promoter of PAP1 (PRODUCTION OF

ANTHOCYANIN PIGMENT 1) with a single crRNA (Fig. 7a). Successful transcriptional repression would indicate successful binding of deactivated LbCas12a (dLbCas12a) to the target DNA. We selected two independent T1 lines for obtaining T2

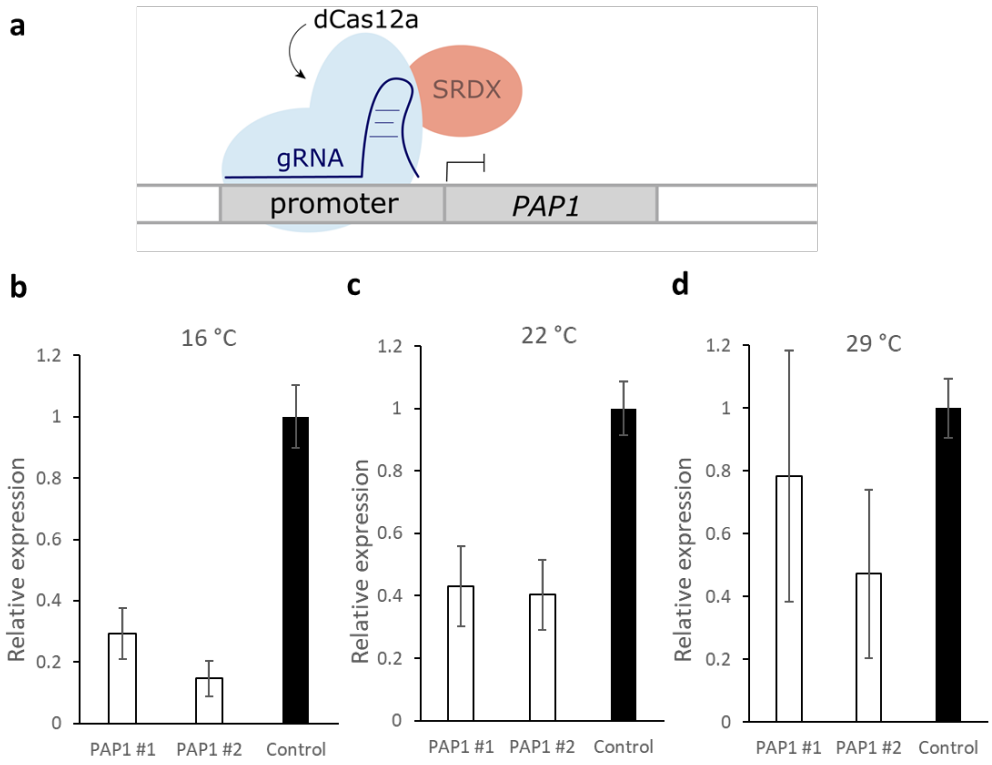


Figure 5. Gene repression with LbCas12a occurred at low temperatures. A) Schematic of dLbCas12a- SRDX targeting the PAP1 promoter for repression. Relative expression of PAP1 mRNA normalized to EF1α for lines grown at B) 16 °C, C) 22 °C and D) 29 °C. Figure from Malzahn et al. [3].

generation seeds. T2 plants were germinated in MS-hygromycin selection medium for a week and then transplanted to new MS medium for 1-week treatment at 16 °C, 22 °C, and 29 °C, before analysis for gene expression by quantitative real-time (qRT)-PCR. The highest levels of repression were seen at 16 °C, where the expression of PAP1 was reduced to 30% and 20% of WT expression in the two transgenic lines, respectively (Fig. 7b). At 22 °C, expression of PAP1 was reduced to 40% of the WT in both lines (Fig. 5c). At 29 °C, PAP1 was also repressed, but more variation was seen between the two lines (Fig. 5d). Unsurprisingly, a comparative analysis of PAP1 in control plants shows that expression was higher at 29 °C, likely as a response to

heat stress (Fig. 6). This could explain some variation between the lines and larger error bars for the repression results at 29 °C. Because repression was accomplished at lower temperatures, these results suggest that unlike genome editing, binding of LbCas12a to the target DNA is not abolished at lower temperatures. Further analysis of dLbCas12a-SRDX expression in transgenic lines indicated that its mRNA level was not elevated at a higher temperature (e.g., 29 °C) (Fig. 7). Rather, PAP1 #2 line showed higher expression of dLb-cas12a-SRDX at 16 °C (Fig. 7), which is consistent with stronger transcriptional repression on the target gene (PAP1) observed in this line at 16 °C (Fig. 5b).

This work explores the mechanism behind temperature-sensitive genome editing with LbCas12a. Elucidating what mechanism is the most temperature sensitive is important for protein engineering of Cas12a. Additionally, this study demonstrates that LbCas12a can be used for gene repression in Arabidopsis without altering the normal growth temperature; an important consideration for other plant researchers applying this technology. These vectors are publicly available for other's use for gene repression [157].

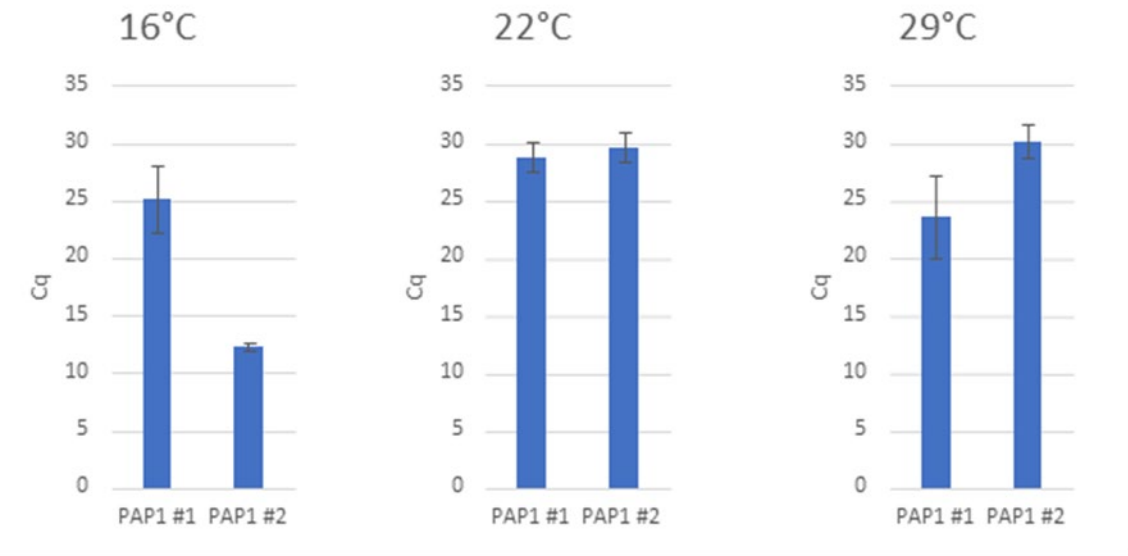


Figure 6. PAP1 expression at three different temperatures in the GUS control plants (WT). PAP1 expression was normalized to EF1α at 16°C, 22 °C, and 29 °C. Error bars represent standard errors of three biological replicates except for WT at 16 °C for which four biological replicates were used.

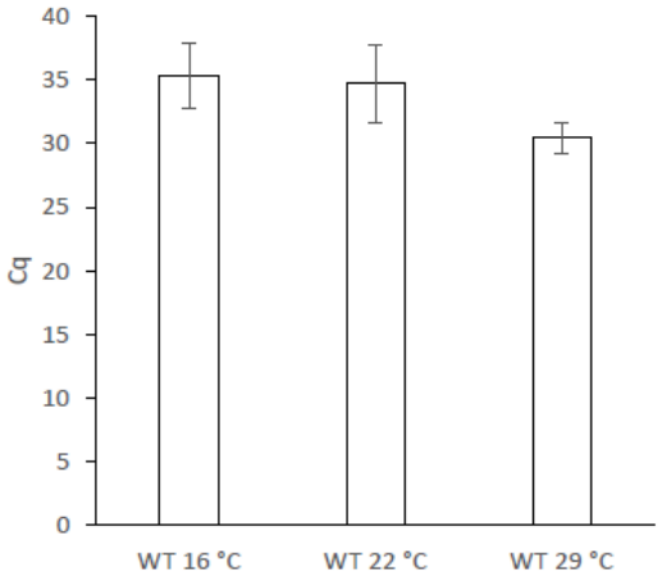


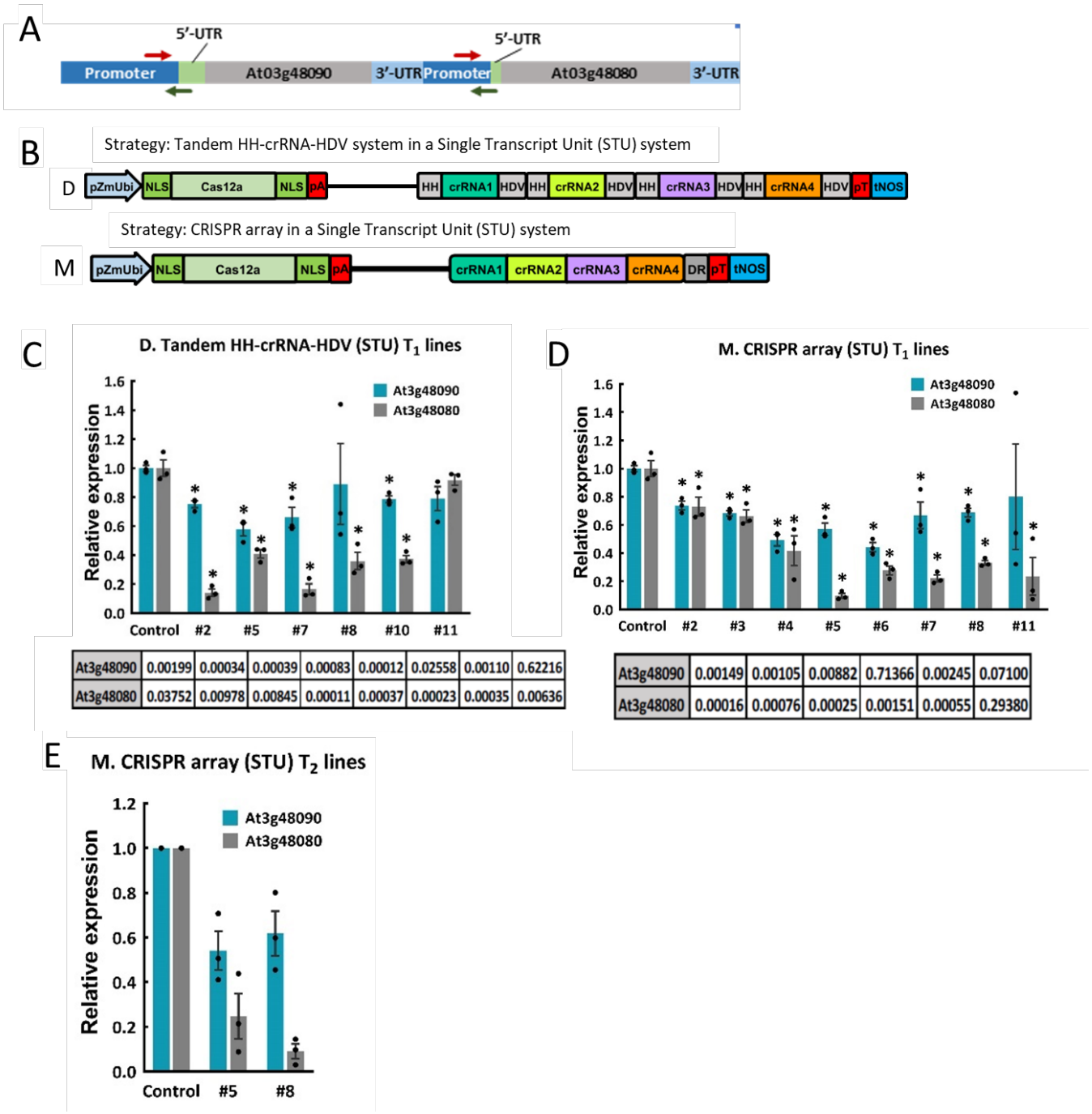
Figure 7. *dLbCas12a-SRDX* expression at three different temperatures in transgenic lines. *dLbCas12a-SRDX* expression was normalized to *EF1α* at 16°C, 22 °C, and 29 °C. Error bars represent standard errors of four biological replicates except for PAP1 #2 at 29 °C for which only three biological replicates were used.

2.2.3 Demonstrating transcriptional repression with multiplexed dLb12a-SRDX systems

A mutation (D832A) in the RuvC-like domain of LbCas12a results in a catalytically dead endonuclease (dLbCas12a) that is capable of binding to the target site but not DNA cleavage, creating a targeted transporter for effector proteins such as activators and repressors [123, 163]. Transcriptional regulators are important for gene function studies [111]. There are still knowledge gaps in studying gene regulation, even well-known effectors such as repressor SRDX are not fully understood [114]. In order to address this need, a multiplexed repressor was developed [164]. Two crRNA expression systems were compared: HH-HDV and Single Transcript Unit (STU). Processing crRNAs is an important step. Extra bases could result in crRNA-DNA mismatches or decreased activity [126]. Cas12a can uniquely process its own crRNA, which may eliminate the need for ribozyme processing systems, such as Hammerhead (HH) and HDV, that take up limited space in the T-DNA vector.

Two *Enhanced Disease Susceptibility 1* (*EDS1*) homologs were selected for targeting (Fig. 1a). EDS1 is part of R-gene mediated disease response and Arabidopsis mutants show enhanced disease susceptibility [165]. Because EDS1 mutants and methods for measuring disease susceptibility are established, this target can be used to measure the impact of the dLbCas12a-SRDX system on phenotype. Previously, we demonstrated transcriptional repression of single genes in plants with dCas12a-SRDX [123, 166]. The STU systems represent the most compact expression

systems for simultaneous and coordinated repression of multiple genes, which can potentially reduce the difficulties in cloning and transformation due to large constructs, as well as avoid gene silencing. Therefore, we chose two top performing STU systems, D and M (Figure 8b), for multiplexed transcriptional repression based on dLbCas12a-SRDX. We further tested the systems in Arabidopsis by targeting two tandemly arrayed genes, At3g48090 and At3g48080, which encode two EDS1 homologs (Figure 8a). Simultaneous transcriptional repression of both genes was observed in multiple T1 lines with either the D or M STU system (Figure 8c,d), and the repression effects were transmitted into the T2 generation (Figure 8e).



At3g48090	0.00199	0.00034	0.00039	0.00083	0.00012	0.02558	0.00110	0.62216
At3g48080	0.03752	0.00978	0.00845	0.00011	0.00037	0.00023	0.00035	0.00636

At3g48090	0.00149	0.00105	0.00882	0.71366	0.00245	0.07100			
At3g48080	0.00016	0.00076	0.00025	0.00151	0.00055	0.29380			

Figure 8. Multiplexed transcriptional repression with two compacted STU dCas12a-SRDX systems. (a) Schematics of the Arabidopsis target genes and crRNAs. Red arrows indicate crRNAs targeting the antisense strand of DNA. Green arrows indicate crRNAs targeting the sense strand of DNA. **(b)** Schematics of multiplexing strategies. Tandem HH-crRNA-HDV system (denoted strategy "M") is driven by a ZmUbi promoter driving a crRNA array with HH-HDV processing. The CRISPR array system is driven by a ZmUbi promoter driving a crRNA array processed by endogenous pre-crRNA processing **(c)** Simultaneous transcriptional repression of two tandemly arrayed genes in Arabidopsis T1 lines by the STU system D. **(d)** Simultaneous transcriptional repression in Arabidopsis T1 lines by the STU system M. **(e)** Targeted transcriptional repression is inherited to the T2 generation. T2 lines are the progenies of T1 lines with the same line number. Transcription levels of target genes were quantified by

qRT-PCR. Data are presented as mean values ± SEM. n=3 technical replicates in "**c**" and "**d**" and n=3 biological replicates in "**e**". Asterisks indicate significant differences ($p<0.05$) between CRISPR lines and the control line using two-sided Student's t-test. *P* values are shown under each line in "**c**" and "**d**".

2.3 Materials and Methods

2.3.1 T-DNA vector construction

Vector Construction for Arabidopsis temperature sensitivity experiments (see Results 2.2.1)

T-DNA vectors for CRISPR-Cas12a were constructed based on the protocols described previously (Figure 9) [89, 126, 167, 168]. Forward and reverse oligos for AtGL2-crRNA1, and AtTT4-crRNA1 (Table 1) were phosphorylated, annealed, and cloned into the Esp3I sites of pYPQ141-ZmUbi-RZ-Lb (Addgene #86197). The resulting crRNA expression vectors were mixed with pYPQ230 (LbCas12a, Addgene #86210) to generate the final T-DNA binary vectors using multi-site LR reactions (1-5-2) [47, 48].

Oligo Name	Sequence	Purpose
AtGL2-crRNA1-F	TAGATTATGTCAATGGCCGTCGACATGT	Target GL2 for editing
AtGL2-crRNA1-R	GGCCACATGTCGACGGCCATTGACATAA	Target GL2 for editing
AtTT4-crRNA1-F	TAGATCTATTCACAGGCGACAAGTCGAC	Target TT4 for editing

AtTT4-crRNA1-R	GGCCGTCGACTTGTCGCCTGTGAATAGA	Target TT4 for editing

Table 1. List of oligos used to construct the crRNAs for editing.

Vector Construction for Arabidopsis single gene repression experiments (see Results 2.2.2)

The T-DNA vector for transcriptional repression in Arabidopsis was constructed similarly (Figure 9) [89, 126, 167, 168]. The protospacer of AtPAP1-crRNA1 (Table 2) was cloned into pYPQ141-ZmUbi-RZ-Lb at Esp3I sites in the form of phosphorylated and annealed oligos. Then, a multi-site LR reaction using pYPQ141-ZmUbi-RZ-Lb-AtPAP1-crRNA1, pYPQ233 (dLbCas12a-SRDX, Addgene #86211), and pYPQ202 (Addgene #86198) was conducted to generate the final T-DNA vector.

Oligo Name	Sequence	Purpose
AtPAP1-crRNA1-F	TAGATTGTATTAGCTGTCGTGCTTAATT	Target PAP1 for repression
AtPAP1-crRNA1-R	GGCCAATTAAGCACGACAGCTAATACAA	Target PAP1 for repression

Table 2. List of oligos used to construct the crRNAs for repression.

Vector Construction for Arabidopsis multiplexed gene repression experiments (see Results 2.2.3)

New destination vectors and crRNA entry vectors were generated as part of this experiment. To generate pYPQ233-STU (dLbCas12a-SRDX-STU, Addgene

#138111) to express Cas12a and its crRNAs as a single transcript unit (STU), the NOS terminator in pYPQ230 (LbCas12a, Addgene #86210) was replaced by a polyA signal at the AatII and BspEI sites.

To express four crRNAs in attL5-attL2 entry vectors, the cloning method for each multiplexing system is as follows: (B) To make Golden Gate recipient vector pYPQ144-ZmUbi-pT (Addgene #138108) an intermediate vector was constructed: pYPQ144-pT. To make pYPQ144-pT, a polyT signal was inserted into pYPQ144 (Addgene #69296) at the SpeI and EcoRI sites. To make Golden Gate recipient vector pYPQ144-ZmUbi-pT, pZmUbi was cut off from pYPQ141-ZmUbi-RZ-Lb (Addgene #86197) with AflII and BamHI and cloned into pYPQ144-pT. The crRNA1–4 were cloned into pYPQ131-STU-Lb (Addgene #138096), pYPQ132-STU-Lb (Addgene #138099), pYPQ133-STU-Lb (Addgene #138102), and pYPQ134-STU-Lb (Addgene #138105). These four vectors were assembled with recipient vector pYPQ144-ZmUbi-pT using Golden Gate reactions for multiplexing system B. (M) The CRISPR array without promoter but with an extra crRNA direct repeat, and a polyT signal at the end was synthesized and ligated into pYPQ144-ZmUbi-pT at the EcoRI and BamHI sites.

The attL1-attR5 entry vectors and attL5-attL2 entry vectors were further assembled with the destination vector pYPQ203 (Addgene #86207) through LR reactions (Figure 9). The final T-DNA vectors used for gene transcriptional regulation were constructed based on the previous protocols (Figure 9) [89, 126, 167, 168].

Oligo Name	Sequence	Purpose
EDS-A-crRNA1-F	TAGATATCCGCTCCTACTCTGTTAATTT	Target EDS for repression
EDS-A-crRNA1-R	GGCCAAATTAACAGAGTAGGAGCGGATA	Target EDS for repression
EDS-A-crRNA2-F	TAGATCATTGAAATGGTCTCATGATGGG	Target EDS for repression
EDS-A-crRNA2-R	GGCCCCCATCATGAGACCATTTCAATGA	Target EDS for repression
EDS-B-crRNA1-F	TAGATAGTCGTCTTGATTCTTAGTCCTC	Target EDS for repression
EDS-B-crRNA1-R	GGCCGAGGACTAAGAATCAAGACGACTA	Target EDS for repression
EDS-B-crRNA2-F	TAGATGGTAATTACTCCTGTCTACATCT	Target EDS for repression
EDS-B-crRNA2-R	GGCCAGATGTAGACAGGAGTAATTACCA	Target EDS for repression

Table 3. List of oligos used to construct the multiplexed repression constructs.

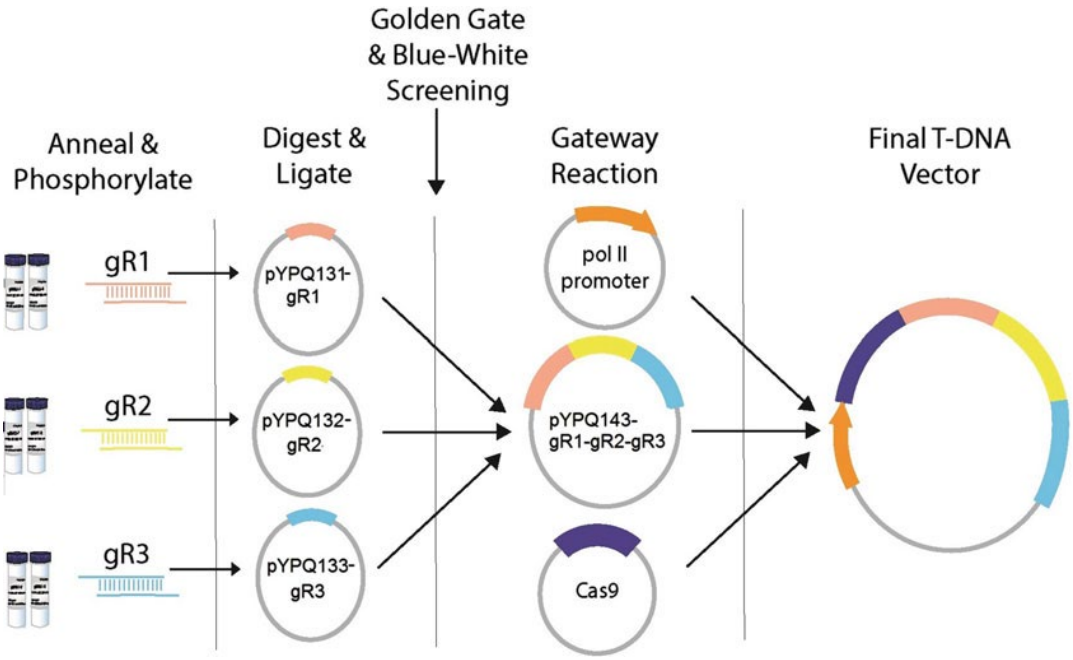


Figure 9. General diagram showing the steps of T-DNA vector construction.

Left: Assembly starts by ordering forward and reverse oligonucleotides (tube pairs with blue caps) for each functional gRNA sequence. Forward and reverse oligonucleotides are phosphorylated and annealed to form double-stranded DNA sequences. gRNA DNA oligomers are ligated into digested gRNA entry clones (pYPQ130 series plasmids). gRNA entry vectors are confirmed using sequencing and are then combined into a single expression array using Golden Gate cloning. Blue–white screening helps streamline positive clone selection after Golden Gate reactions and transformation. Successful Golden Gate assemblies are confirmed using double restriction analysis and then combined with Gateway compatible entry vectors. Available Gateway compatible vectors harbor T-DNA expression backbones with various Polymerase II promoters allowing for flexible heterologous Cas9 transcription (orange arrow). Other compatible vectors harbor Cas9 or Cas9

derivatives such as dCas9-VP64 (artificial transcriptional activator; from pYPQ152) or dCas9-SRDX (artificial transcriptional repressor; from pYPQ153) (Purple bar on plasmid). Invitrogen Multisite LR Gateway cloning is used to assemble desiredCas9 variant (e.g., dCas9 fused with a transcriptional activator or repressor), Pol II promoter with T-DNA expression elements and selective markers and multiplex gRNA expression unit. Figure from [169].

2.3.2 Arabidopsis stable transformation

Arabidopsis thaliana variety Columbia wild type plants were transformed with Agrobacterium tumefaciens strain GV3101 using the floral dip method [170]. Plants were trimmed about 4-5 days before transformation to promote growth of unopened buds. A transformed Agrobacterium colony was scraped from solid LB media (supplemented with kanamycin and gentamycin) and cultured overnight at 28 °C in 2 mL of liquid LB Kan+ Gen+ media. One mL of culture was added to 200 mL of liquid LB Kan+ Gen+ media and cultured overnight at 28 °C. Culture was transferred to a clean centrifuge tube and spun at 3700g for 20 minutes at 4 °C. Cells were re-suspended in fresh buffer consisting of 400 mL of milliq water, 20 g of sucrose and 120 ul silwet L77. Plants were carefully dipped into the buffer and held for a minute before being laid sideways into a standard planting tray. A second tray was used to cover the plants and left overnight to increase the humidity and block the light. Plants were then placed back into a growth chamber.

2.3.3 Temperature treatment and mutation analysis

Seeds for T1, T2, and T3 generations were sterilized using 50% bleach and 0.05% Tween, vernalized at 4 °C for 3 days, then plated on ½ MS media with 15 mg l−1 hygromycin. After a week, transgenic plants were transferred to MS clean plates for a week of recovery before soil transplantation. To test a variety of T2 lines, five individual plants were sacrificed after 2 weeks of heat treatment at 29 °C on MS plates. Leaf tissue was used for DNA extraction using a modified CTAB method [50]. The rest of the plants were transferred to soil and kept at 29 °C for a total of 29 days before recovery at 22 °C. A second batch of plants was heat treated to test an alternative method of late treatment at 29 °C. For this, 6 days after plating, the plants were kept at 29 °C for 8 days. After 24 days of recovery at 22 °C, plants were kept at 29 °C again for 14 days. For mutation analysis, a ~ 677 bp fragment covering the GL2 target site was amplified using the primers GL1-F1 5-GATGGCTGCCAATGCTGTAGCTGG-3 and GL2-R1 5-CGTCAACTACTCTTCTGCCCAGG-3, and a ~400-bp fragment covering the TT4 target site was amplified using the primers TT4-F2 5'-AGGCATCTTGGCTATTGGCACTG-3' and TT4-gR3-top 5'-gattGGGCTGGCCCCACTCCTTGA-3'. GL2 and TT4 PCR products underwent RFLP analysis through direct digestion with restriction enzyme SalI and analysis on 1.5% and 2% agarose gels, respectively. The mutation percentages for each plant were estimated from RFLP gels with Image Lab (BioRad) [159]. PCR products were cleaned using Exonuclease I and Antarctic Phosphatase and sent for Sanger

sequencing. Results were aligned in Snapgene and decoded with CRISP-ID (GSL Biotech LLC) [54].

2.3.4 Single and multiplexed gene transcriptional repression and analysis

For the single-gene repression experiments, transgenic Arabidopsis dLbCas12a-PAP1 T2 plants were grown on ½ MS media with 15 mg/L hygromycin at 22 °C for a week. They were then transferred to MS medium without hygromycin, allowed to recover at 22 °C for 2 days, and then grown at 16 °C, 22 °C, and 29 °C for a week. Individual plants were treated as single transgenic lines. Arabidopsis leaf tissue was collected from these T2 seedlings.

For the multiplexed-gene repression experiments, transgenic Arabidopsis T1 plants were grown on ½ MS media with 15 mg/L hygromycin at 22 °C for a week. They were then transferred to MS medium without hygromycin and allowed to recover for a week before leaf tissue was collected. Plants transformed with a GUS gene and a hygromycin resistance gene were used as controls. Individual plants were treated as single transgenic lines. Seeds collected from these lines were treated using the same method to obtain the T2 generation plants. Three plants of each line were used for qRT-PCR analysis.

The qRT-PCR analysis was carried out following previously described protocols [47] with minor modifications. GUS transformed plants, also expressing a hygromycin resistance gene, were used as controls. Total RNA was extracted using TRIzol™ Reagent (Invitrogen) following the manufacturer's instructions with the exception that samples were extracted twice using chloroform and washed twice by 75% ethanol. RNA was treated with DNase I (New England BioLabs) to remove

DNA contamination. Complementary DNA (cDNA) was synthesized using SuperScript III First-Strand Synthesis System (Invitrogen) with Oligo dT. The qRT-PCR was set up using Applied Biosystems SYBR Green Master Mix (Invitrogen) and ran on the CFX96 Touch™ Real-Time PCR Detection System.

The following primers were used to analyze gene repression at different temperatures: PAP1-F 5-AGTATGGAGAAGGCAAATGGC-3' and PAP1-R 5-CACCTATTCCCTAGAAGCCTATG-3', LbCas12a-RT-F1 5-TTCGTTCAACGGATTCACAA-3, and LbCas12a-RT-R1 5-GCTTGTCAAAAATTGCGTCA-3'. The following primers were used to analyze multiplexed gene repression: EDS1A-F1 5'-CTGGTACAGTCGATGGGAAAG-3', EDS1A-R1 5'-TCCTCTAATGCAGCTTGAACG-3', EDS1A-F2 5'-CCGTGTTCAGTTTCCTTGTATG-3', EDS1A-R2 5'-TGTTGCTAAGATTGCAGTTGC-3', EDS1B-F1 5'-GATCTGCATAATCCGCTTTTGG-3', EDS1B-R1 5'-TTCCACTCGTTGCTTCTCAG-3', EDS1B-F2 5'-GAACCTCATTTCATGCTTCTGTG-3', EDS1B-R2 5'-CGTCACGTATGAAGTATGTCTCC-3'. Elongation factor1 α (EF1α) was used as the internal control and amplified with the following primers: EF-1α-F 5'-TGAGCACGCTCTTCTTGCTTTCA-3' and EF-1α-R 5'-GGTGGTGGCATCCATCTTGTTACA-3'.

The average of three technical replicates was used for data analysis of each biological replicate. Relative expression to controls was calculated using the comparative threshold cycle method. Student's t-test was used for pairwise

comparison. One asterisk ($p < 0.05$) and two asterisks ($p < 0.01$) indicate significant differences between two treatments. Tukey's Honest Significant Difference (HSD) test was used for multiple comparisons. Treatments with the same letter are not significantly different when $\alpha = 0.05$.

2.4 Discussion

Cas12a endonucleases have several unique features: a T-rich PAM, unique DSB cleavage, and endogenous crRNA processing, to name a few. Cas12a genome engineering toolkits greatly expand the number of available targets for mutation or transcriptional regulation. Arabidopsis is a well-known model organism that grows best between 22 °C - 25 °C and is a dicot in the Brassicaea. While LbCas12a worked well in rice, which is a monocot grown between 28 °C - 32 °C, low editing efficiency was observed in Arabidopsis. LbCas12a efficiency could be rescued through heat treatment. This highlights the importance of testing genome engineering systems in different conditions, not only to ensure the system works, but because different conditions can reveal new information about CRISPR endonucleases and how they function.

In this study, we investigated temperature sensitivity of LbCas12a in Arabidopsis. While LbCas12a showed reasonable nuclease activities in rice protoplasts and detectable activity in Arabidopsis protoplasts at 22 °C, it barely worked in Arabidopsis cells of stable transgenic lines at the same temperature. This could be due to the procedures of delivering CRISPR-Cas12a reagents, in vitro cell culture for the transient protoplast assay, and floral dip for Arabidopsis. Also, the length of Cas12a treatment and the tissue sources for evaluating mutagenesis are

different among transient and stable transgenesis. It is possible that any of these factors had contributed to the drastic difference between rice and Arabidopsis on editing efficiencies we observed. It is likely that chromatin structure plays a role as it has been shown to impact genome editing. For example, the natural cycle of nucleosome breathing and ATP-driven chromatin remodelers are essential for Cas9 binding at target sites [37, 38], and the same can be true for Cas12a. Recent comparative studies have revealed distinct chromatin packing in rice and Arabidopsis [39, 40]. It will be interesting to investigate how chromatin states in different plant species impact Cas12a and other CRISPR-Cas systems on genome editing. Nevertheless, we were able to rescue LbCas12a activity at a higher temperature and demonstrated LbCas12a-based germline editing in Arabidopsis. In zebrafish, AsCas12a had poor activity at 28 °C and its activity was drastically improved by elevating the temperature to 34 °C [29]. By contrast, we found Cas12a nucleases seem to reach optimal activities in plants at around 28–29 °C, which is more feasible given most plants grow in temperatures around 22–29 °C. For example, we have grown rice and maize constantly at 28 °C and have treated Arabidopsis at 29 °C for up to ~ 4 weeks continuously. However, lengthy and continuous treatment at 29 °C significantly impedes Arabidopsis growth. Further exploration of heat treatment regimens will probably result in more robust genome editing in plants. While Cas12a nucleases are temperature sensitive, it should not be a barrier that prevents adoption of them for genome editing in many other plant species.

Currently, it is unclear why Cas12a-based genome editing is temperature sensitive. With the analysis of NHEJ mutations by all three Cas12a nucleases across

four different temperatures, we ruled out the possible involvement of DNA repair pathways in this difference. Opposite effects of high temperatures on the activities of ZFN and TALEN versus CRISPR-Cas9 were reported in mammalian cells [43], which also suggested the effects were unlikely due to DNA repair machinery. Using a dLbCas12a-SRDX repressor, we further demonstrated that the DNA binding property of LbCas12a at lower temperatures is as good as, if not better than, higher temperatures. This is consistent with our previous observation that the dAsCas12a-SRDX repressor could work robustly in mediating targeted transcriptional repression at room temperature in Arabidopsis [17]. While DNA binding is not significantly affected under these temperatures, it is still possible that chromatin structure is affected by temperature in a way that impacts the necessary conformation change of the Cas12a/crRNA ribonucleoprotein complex that is required for activation of nuclease activity, as supported by the data in zebrafish [29]. Our data collectively points to a working hypothesis that Cas12a nuclease activity is affected by temperature. Upon activation, Cas12a proteins also unleash single-stranded DNase activities [44,45]. We predict such non-specific DNase activities are likely also temperature sensitive. Finally, given that we have narrowed down the main cause of temperature sensitivity to Cas12 nuclease activities, it will be highly valuable and should also be possible to engineer Cas12a variants that are more active at lower temperatures, similar to engineering Cas12a variants with altered PAM specificities [25], which has been recently demonstrated in plants [18, 26].

Two Single Transcript Unit (STU) multiplexed transcriptional repression systems were tested to expand on the dLbCas12a-SRDX transcriptional repression

work. Several gene editing multiplex systems were tested in rice, and the top two performing systems, labeled “D” and “M” were utilized for transcriptional repression in Arabidopsis. STU-D used flanking HH and HDV ribozymes to process four crRNAs from a single transcript, while STU-M relied on endogenous crRNA processing from LbCas12a. Neither system achieved consistently high rates of repression as a great deal of variance in the level of repression was observed between individual lines. These results demonstrate that the compact STU-M system, with endogenous crRNA processing, can be used successfully. Future work will likely test different repressors in comparison with SRDX, which while commonly used, is not well understood. Exploring the ability of natural Cas variants to repress gene expression may uncover Cas proteins with characteristics that are helpful for repression, such as tight protein-DNA binding. Engineering Cas proteins could also produce a more effective repressor. Outside of the Cas-repressor proteins, future work may focus on expanding multiplexing systems which will be useful for targeting the gene(s) of interest and simultaneously targeting endogenous controls of gene expression.

Chapter 3: Improving NHEJ outcomes in rice with Cas9 and 12a fusions

4.1 Introduction

CRISPR systems can be improved through strategic engineering

While Cas9 and Cas12a have both been used effectively in plants, their efficiency in all species and under different conditions leaves room for improvement. One strategy to improve CRISPR systems is to explore natural variants for different characteristics and editing efficiency. The earliest genome editing studies use the Cas9 endonuclease. Now, there are several working Cas endonucleases with different properties including Cas12a and Cas12b, both targeting T-rich PAMs (protospacer adjacent motifs) [171]. The mutation efficiency for these new endonucleases remains relatively low in plants.

Previously, engineering Cas with effector proteins has been used to create new editing properties and applications. Some examples include base editing, which uses a deaminase fused to a Cas nickase to edit nucleotides from a C to T or A to G within a targeting window [144], and the transformation of CRISPR from a genome editor into a transcriptome editor via fusion proteins of Cas9-VP64 and Cas9/Cas12a-SRDX [171–173]. The success of fusion proteins in novel applications highlight the versatility of the CRISPR system. The use of new and creative designs will yield improvements in efficiency as well as new editing outcomes. In the following experiments, the strategy to improve CRISPR systems centers on engineering Cas9 and Cas12a with effector proteins fusions.

Because CRISPR systems evolved in bacteria and archaea for defense against viruses, these endonucleases did not evolve to target eukaryotic DNA. Unlike in bacteria, plants grow at ambient temperatures, have complex regulation and genomic

structure, and are often polyploid. These features can be roadblocks to efficient application of CRISPR and thus consideration of the microenvironment within the plant cells should inform the design of engineered CRISPR systems. In this chapter, the impact of NHEJ repair and chromatin remodeling on CRISPR-Cas9 and Cas12a is explored through the design and comparison of Cas-fusion proteins.

Chromatin may hinder target DNA accessibility

One of the distinguishing features of eukaryotic DNA is the way it is tightly packaged into chromatin. Accessibility of CRISPR in chromatin is an important consideration for genome editing and transcriptional regulation in eukaryotes. Heterochromatin has been shown to impede CRISPR efficiency, although it remains unclear what the precise cause is [33–35]. There are several regulatory and chromatin remodeling proteins found in plants and animals. High Mobility Group Nucleosome-binding (HMGN) proteins are a family of proteins that bind to the nucleosome and regulate chromatin. While the exact mechanism is unknown, HMGNs decondense chromatin through competition for chromatin binding sites, or through interaction with the histone and histone tails [51]. Three HMGNs have been identified and demonstrated to increase editing efficiency in human cell lines [174]. High Mobility Group Protein B1 binds and bends to broken DNA and enhances chromatin formation [52]. Human histone H1 central globular domain is a part of the linker histone H1. H1 binds to the DNA at the beginning and end of the nucleosome and stabilizes chromatin. By fusing portions of these proteins to Cas9, increased mutation efficiency was observed in human cells [56]. Because these chromatin modulators are highly

conserved, they can likely also be used in plants to increase editing efficiency.

NHEJ repair is triggered by double-strand breaks

After CRISPR-Cas causes a double-stranded break, that break must be repaired for cell survival. There are multiple DNA repair pathways in plant cells and the DNA repair outcomes vary between pathways, target sequences, and cell cycle stages. DNA repair pathways are split into two major groups, Non-Homologous End Joining (NHEJ) and Homology Directed Repair (HDR) [3, 4, 175]. NHEJ is an imprecise DNA repair pathway, mostly occurring in somatic cells during the G0 phase. The result is an insertion/deletion (indel) of a few base pairs. An indel may result in an early stop codon and thus render a gene knockout, which is often the primary end goal for researchers using CRISPR in plants.

Cas12a and Cas9 produce different double-stranded breaks that often result in deletions. Cas9 generates blunt ends and deletions typically range from 1 to 3 bp, while Cas12a generates staggered ends and deletions are usually between 6 and 15 bp [40, 176]. These deletions do not guarantee a knockout, which is often the goal during genome editing. Indeed, we observed this firsthand when Cas12a produced homozygous mutations of 3 bp that did not cause a mutant phenotype [156]. In order to increase the size of the deletions, a strategy has been developed that deploys two gRNAs to flank the target DNA [84, 177, 178]. However, these systems rely on the availability of two close target sites and similar editing efficiency at the target sites. To circumvent these issues, Cas-exonucleases fusions can degrade the ends of DNA

breaks generated by Cas9 or Cas12a [148]. While exonucleases have been utilized for genome editing in other organisms, only one exonuclease, T5, has been tested in plants [151]. New Cas-fusions are likely to create larger deletions and/or increase editing efficiency.

4.2 Results

4.2.1 Comparison of Cas9-chromatin modulating peptides (Cas9-CMPs) in rice protoplasts

Two varieties of Cas9-CMP were created and tested in a rice protoplast assay. The chromatin modulating peptide HN1 was fused to the N-terminus of Cas9 with a 5xGS linker (Figure 1). Each variety had an N-terminus HN1 and a different C-terminus CMP fused to Cas9 (Figure 1), creating HN1-Cas9-HB1 and HN1-Cas9-H1G.

Two target genes were selected *LOC_Os04g22730* and *GRAIN WIDTH 2 (GW2)* (Figure 2). GW2 is involved in regulating the grain width in rice, and therefore is a highly relevant gene to rice breeders. Indeed, GW2 has previously been targeted for CRISPR editing and specific alleles have been studied for their impact on rice yield [179, 180]. Alternatively, LOC_Os04g22730 was chosen because this target sites may be more difficult to edit. Previous work

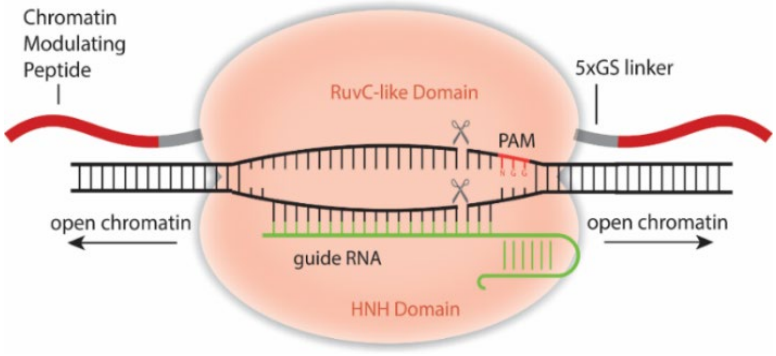


Figure 1. Schematic diagram for the fusion of Cas9 and a chromatin modulating peptide.

targeted these sites because they are in heterochromatic regions, which can impede Cas9 activity [150]. The sites were targeted with a Cas9-activator fusion (Cas9-TV), which increased editing efficiency compared to canonical Cas9. Based on the results, it is theorized that this increase was caused by the VP64 activator opening up the tightly

wound DNA and creating a more accessible target for Cas9 editing.

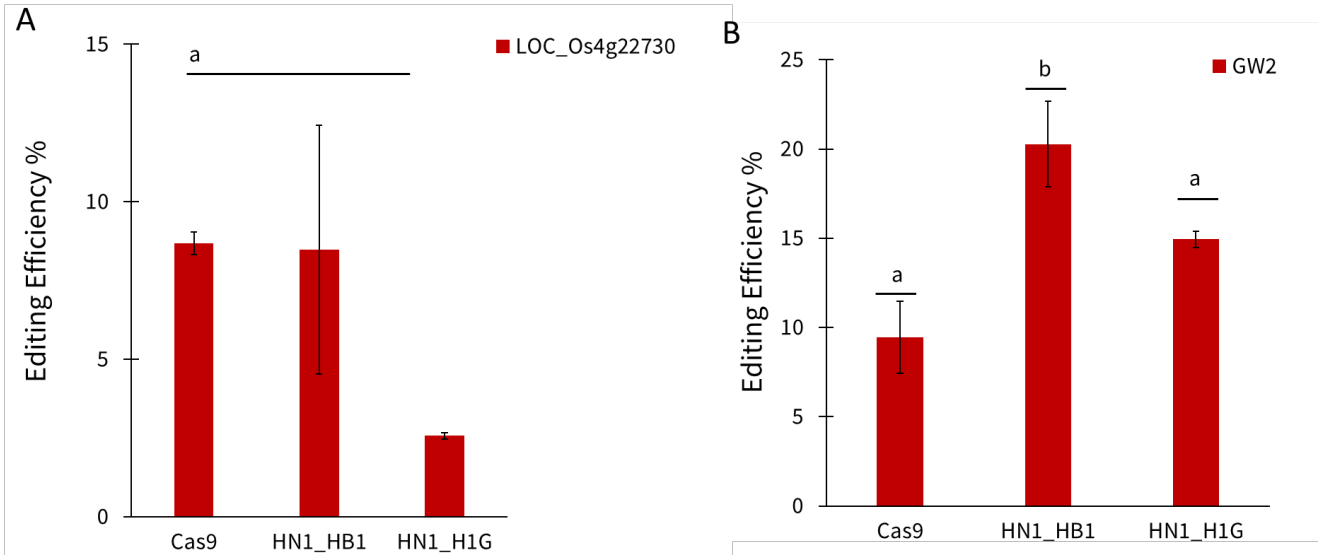


Figure 2: Comparing editing efficiency of Cas9-CMP fusions. Each Cas9-CMP is compared to Cas9 and the other Cas9 fusions. A) Cas9-CMPs targeting LOC_Os04g22730 B) Cas9-CMPs targeting GW2. Matching "a" and "b" labels indicate there no significant increase in editing efficiency between treatments when $\alpha = 0.05$ as determined by a pairwise T-test. Data presented as the mean +/- SEM (standard error mean). Three technical repeats were analyzed for all samples.

At the LOC_Os04g22730 target site, the average efficiency of HN1-Cas9-HB1 was 8.5%, HN1-Cas9-H1G was 2.6%, and Cas9 was 8.7% (Figure 2a). At GW2, the average editing efficiency of HN1-Cas9-HB1 was 20.3%, HN1-Cas9-H1G was 14.9%, and Cas9 was 9.4% (Figure 2b). At both targets, HN1-Cas9-H1G

produced a lower editing efficiency than HN1-Cas9-HB1. HN1-Cas9-HB1 increased editing efficiency significantly compared to Cas9 at GW2 (Figure 2b). These results demonstrate that Cas9-CMP fusions can successfully increase editing efficiency in rice protoplasts, and HN1-Cas9-HB1 is a more effective Cas9-CMP fusion than HN1-Cas9-HB1.

The position of deletions produced by Cas9 and the Cas9-CMPs was measured around target site LOC_Os04g22730 (Figure 3). The percentage of aligned reads with deletions at each nucleotide was calculated by CRISPResso2 [181]. Cas9 and Cas9-CMPs had a similar overall deletion profile, with the highest percentage of deleted bases occurring between the 4th and 10th base pair from the PAM sequence. For Cas9, there is a higher percentage of deletions occurring 5 to 9 bps from the end of the PAM sequence when compared to the Cas9-CMPs, while in contrast HN1-Cas9-HB1 had a higher percentage of deleted bases upstream and downstream of the target site compared to compared to Cas9.

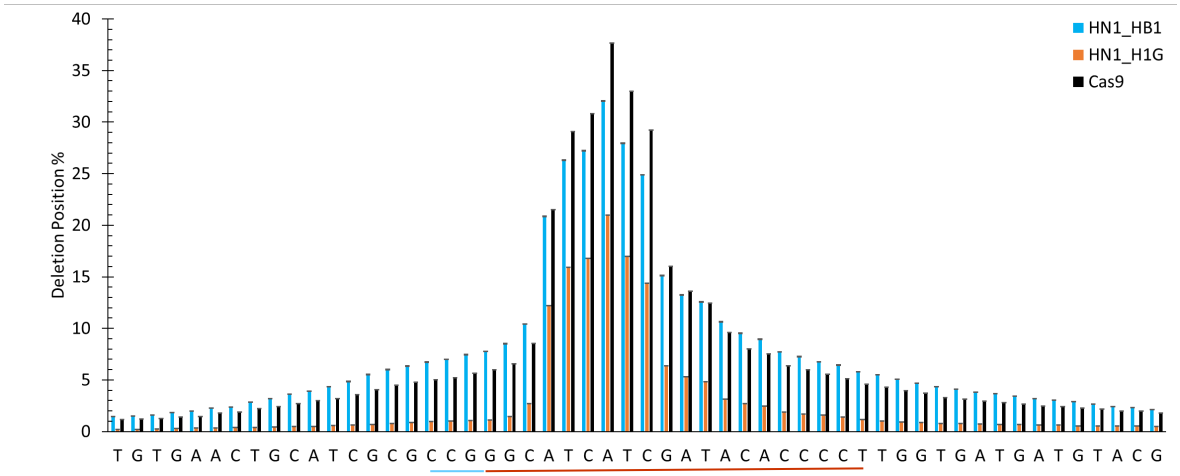


Figure 3. Deletion position percentages of Cas9 and Cas9-CMPs at LOC_Os04g22730. Bars indicate the percentage of reads that showed a deletion at that nucleotide. The red line labels the sgRNA target, and the blue lines is the PAM site. Data presented as the mean +/- SEM (standard error mean). Three technical repeats were analyzed for all samples.

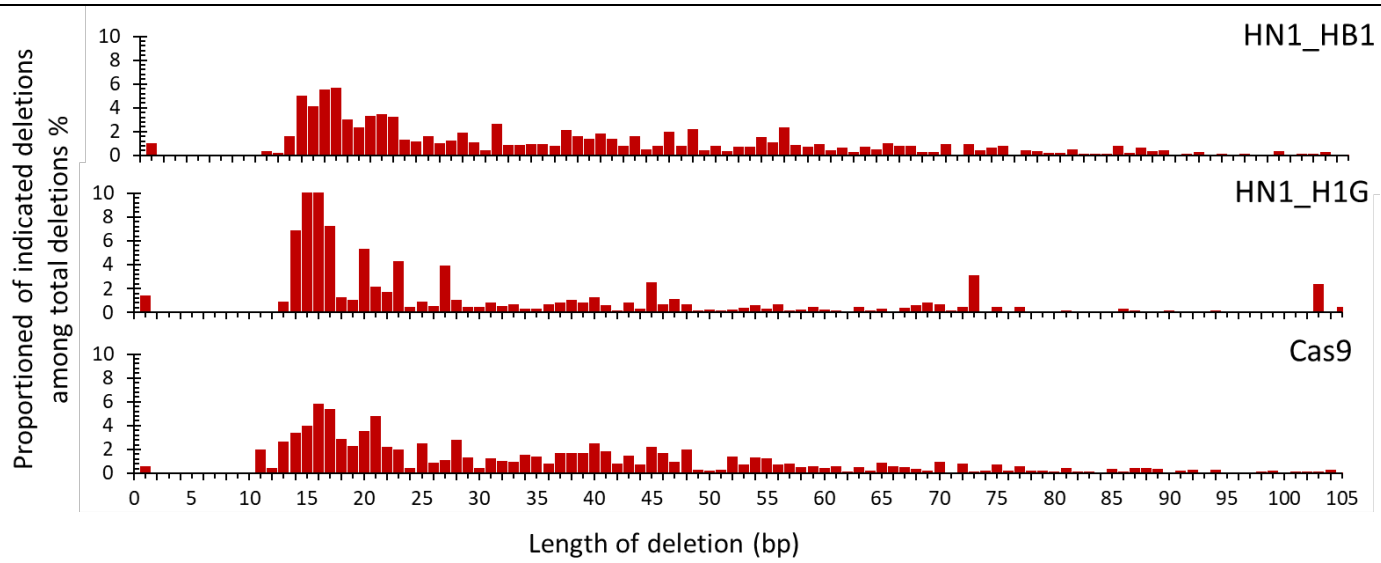


Figure 4. Deletion position percentages of Cas9 and Cas9-CMPs at LOC_Os04g22730. Bars indicate the percentage of reads that showed a deletion at that nucleotide. The red line labels the sgRNA target, and the blue lines is the PAM site. Data presented as the mean +/- SEM (standard error mean). Three technical repeats were analyzed for all samples.

The length of deletions were calculated at LOC_Os04g22730 for Cas9 and the Cas9-CMPs (Figure 4). A small percentage of 1 bp deletions were produced by each system, which is a result often seen with Cas9 editing. For all three systems, the highest percentage of deletion lengths were between 15-17 bps as well as between 20-21 bps. HN1-Cas9-H1G produced the highest percentage of larger deletions as 3% of deletions were 73 bp long and 2.4% of deletions were 103 bp long. While higher percentages of deletion lengths from Cas9-CMPs match the deletion lengths produced by Cas9, some longer deletions were observed with HN1-Cas9-H1G.

4.2.2 Comparison of Cas9-exonuclease fusions in rice protoplasts

Cas9-exonuclease fusions were designed with two configurations: N and C terminus (Figure 4). Cas9 and the exonuclease are fused with an XTEN linker either at the N or C terminus of Cas9. Comparison of the N and C terminus Cas9-fusions were compared to determine whether any difference in editing efficiency could be observed that may indicate interference between Cas9 and the exonucleases.

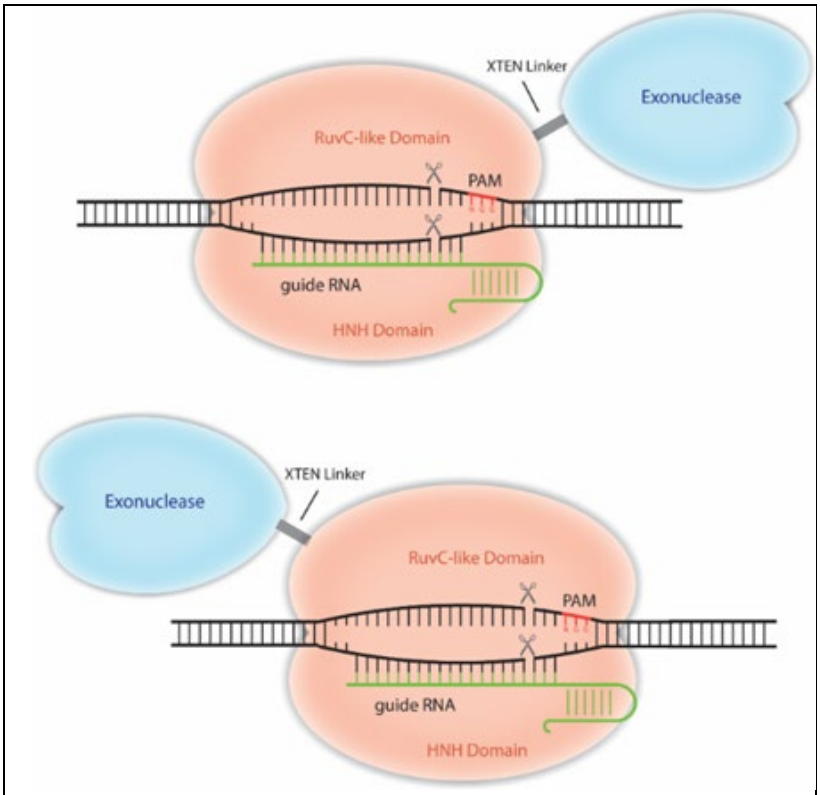


Figure 4. Cas9-exonuclease fusions in C and N terminus configurations.

Three exonucleases were used for comparison: AtDPD1, TREX2, and AtExo1b. These exonuclease represent a sampling of exonuclease from different sources, and importantly, different functions. 1) Defective in Pollen organelle DNA Degradation 1 (DPD1) is found only in angiosperms and degrades organelle DNA during pollen development and leaf senescence [182–184]. AtDPD1 is a Mg^{2+} dependent 3' to 5' exonuclease. It can degrade double-stranded (ds) DNA, but higher activity is observed on single-stranded (ss) DNA [183]. 2) Three prime repair exonuclease 2 (TREX2) is also a

Mg^{2+} dependent 3’ to 5’ exonuclease, but can be found in both plants and mammals [185]. Trex2 is involved in DNA repair and is active on ssDNA. 3) Exonuclease 1b (AtExo1b) is a Mg^{2+} dependent 5’ to 3’ exonuclease involved in DNA repair. It can act on the 5’ of duplexed dsDNA and degrades the 5’ to 3’ strand, leaving behind a tail of 3’ to 5’ ssDNA. AtExo1b can also degrade ssDNA 5’ to 3’.

Two targets were selected: LOC_Os04g22730, and *GW2* (Figure 5). At LOC_Os04g22730, N-AtDPD1 and N-TREX2 had a higher average editing efficiency than their C-terminus counterparts. At GW2, all three N-terminus versions (AtDPD1, TREX2, and AtExo1b) had a higher average editing efficiency than their C-terminus versions, notably with N-TREX2 significantly outperforming C-TREX2.

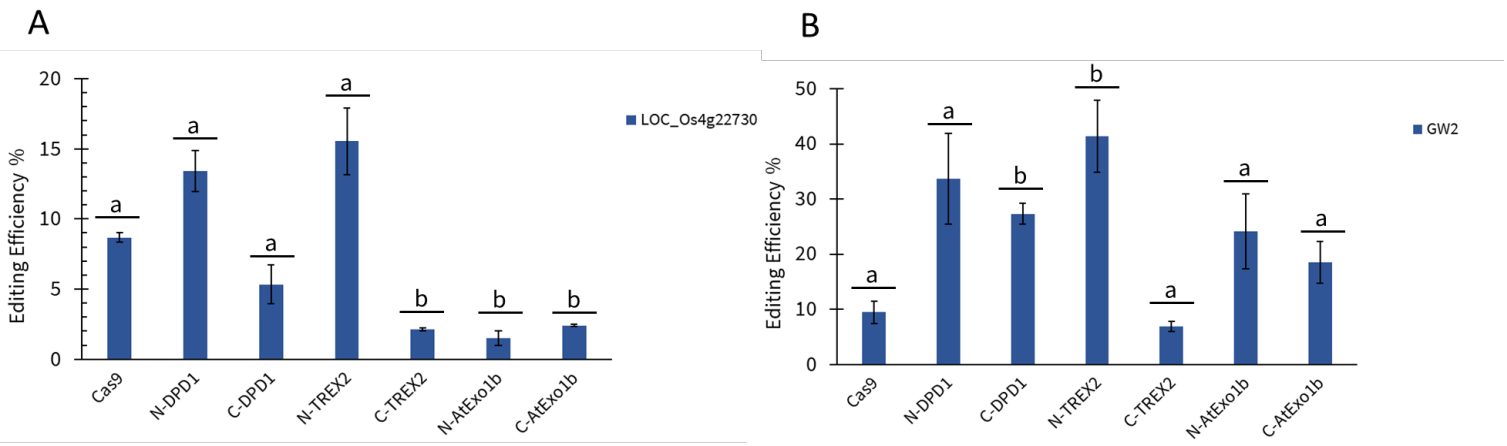


Figure 5: Comparing editing efficiency of Cas9 C and N-terminus fusions. Each Cas9-C and N terminus fusion is compared to Cas9 and the other Cas9 fusions. A) Cas9-fusions targeting LOC_Os04g22730 B) Cas9-CMPs targeting GW2. Matching “a” and “b” labels indicate there no significant increase in editing efficiency between treatments when $\alpha = 0.05$ as determined by a pairwise T-test. Data presented as the mean +/- SEM (standard error mean). Three technical repeats were analyzed for all samples.

In addition to editing efficiency, a deletion profile was constructed for LOC_Os04g22730 (Figure 6). Interestingly, larger deletions were observed with Cas9-AtDPD1 and Cas9-TREX2 when compared to Cas, but Cas9-AtExo1b fusions produced the same deletion profile as Cas9. N and C-terminus versions of the same exonuclease produced the same deletion profile.

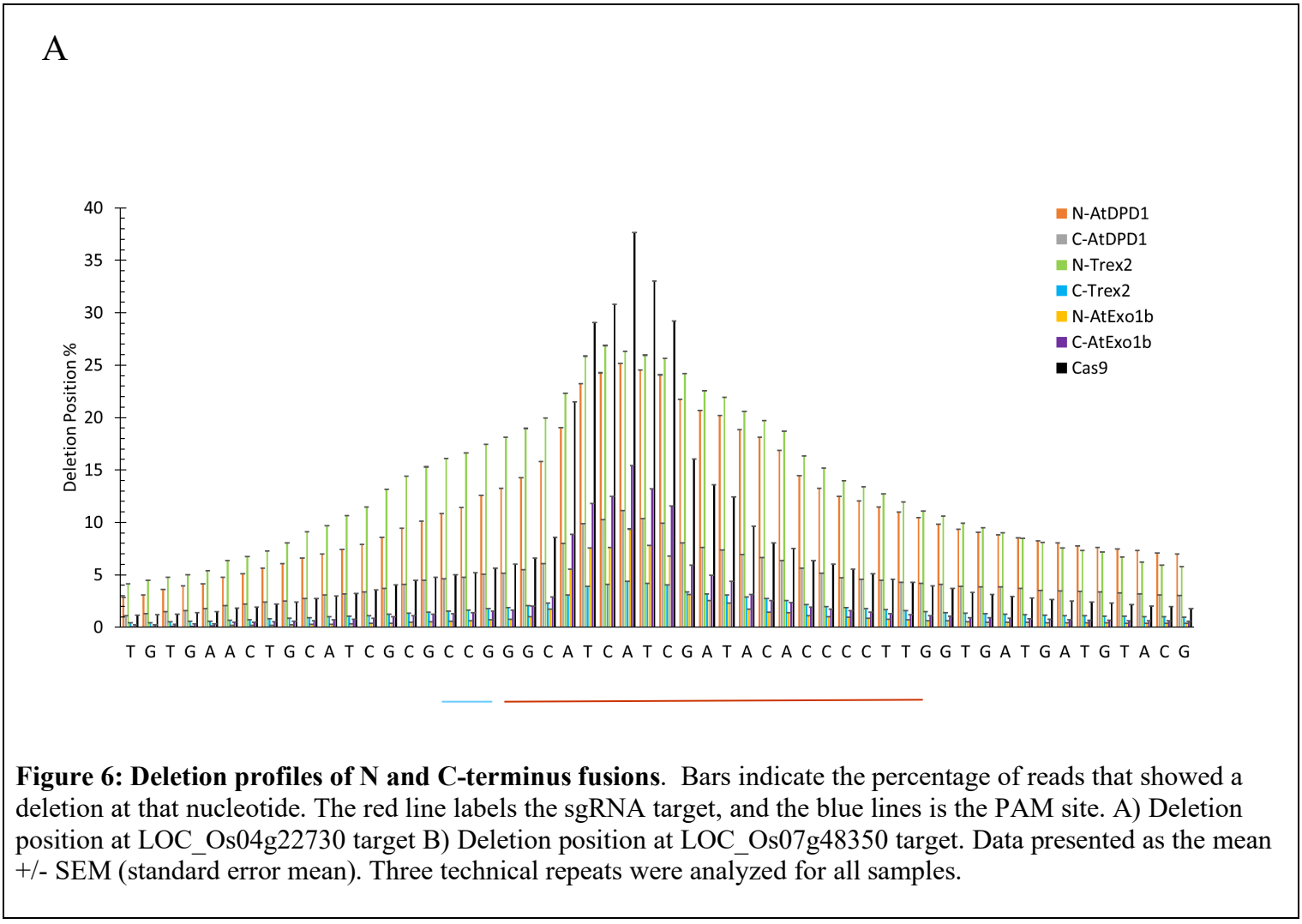


Figure 6: Deletion profiles of N and C-terminus fusions. Bars indicate the percentage of reads that showed a deletion at that nucleotide. The red line labels the sgRNA target, and the blue lines is the PAM site. A) Deletion position at LOC_Os04g22730 target B) Deletion position at LOC_Os07g48350 target. Data presented as the mean +/- SEM (standard error mean). Three technical repeats were analyzed for all samples.

Based on the previous results, Cas-exonucleases were constructed only as an N-terminus fusion. In addition to AtDPD1, Trex2, and AtExo1b, three additional exonucleases were tested: sbcB, T7, and T5. 1) sbcB is an E. coli exonuclease that degrades ssDNA 3’ to 5’ [186]. 2) T7 is sourced from a bacteriophage and recognizes duplexed dsDNA to degrade DNA 5' to 3'. T7 also is active on 5’ ssDNA flaps, but shows higher activity on dsDNA [187]. 3) T5 is sourced from a bacteriophage and degrades DNA 5’ to 3’. T5 has the highest activity at 5’ ssDNA flaps, but can degrade DNA-RNA hybrids and dsDNA, albeit at a slower rate than ssDNA [188].

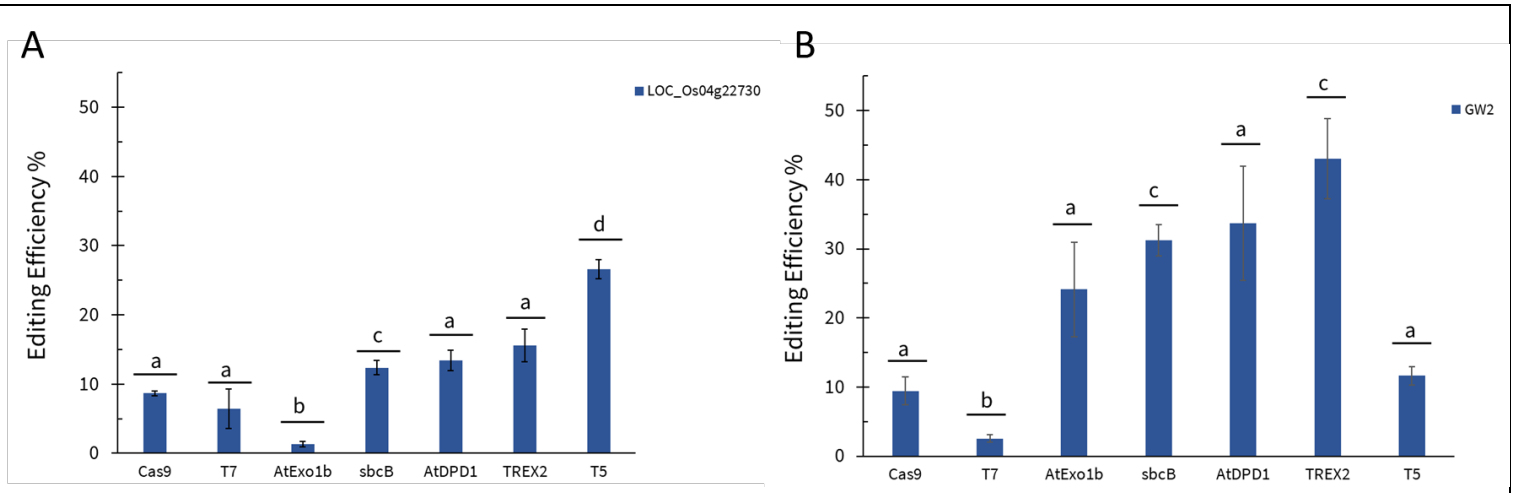


Figure 7: Comparing editing efficiency of Cas9-exonuclease fusions. Each Cas9-exonuclease is compared to Cas9 and the other Cas9 fusions. A) Cas9-exonucleases targeting LOC_Os07g48350 B) Cas9-exonucleases targeting LOC_Os04g22730 C) Cas9-exonucleases targeting GW2. Matching “a” and “b” labels indicate there no significant increase in editing efficiency between treatments when $\alpha = 0.05$ as determined by a pairwise T-test. Data presented as the mean +/- SEM (standard error mean). Three technical repeats were analyzed for all samples.

The six Cas9-exonuclease fusions showed similar patterns in the average editing efficiency between the targets, although statistically significant differences

were not observed for all targets (Figure 7). T7 and AtExo1b had the lowest average editing efficiency, and editing efficiency increased with sbcB, AtDPD1, and TREX2. Although T5 had high editing efficiency at the low-efficiency target LOC_Os04g22730 (Figure 7a), lower editing efficiency was observed at *GW2* (Figure 7b). These results demonstrate that Cas9-exonuclease fusions can increase genome editing efficiency in comparison to Cas9.

In addition to editing efficiency, a deletion profile was created for LOC_Os04g22730 (Figure 8). Many of the Cas9-exonuclease fusions produced deletions that extended past the 5' and 3' sgRNA target sites, which contrasted with the deletions produced by Cas9 where the largest percentage of nucleotides were deleted between 4 to 15 bp from the PAM. Different deletion lengths were produced by different Cas9-exonuclease fusions (Figure 9). AtExo1b, TREX2, and AtDPD1 produced a larger percentage of deletions in the 70-80 bp range than the other Cas9-exonuclease fusions and Cas9. T7, sbcB, and AtExo1b produced a high percentage of deletions in the 13-18 bp range. While the other exonucleases produced deletions in this range, there is a more even spread of deletions across sizes. AtDPD1 and T5 did not produce high percentages of deletions within a specific range of lengths, rather the deletions are spread out between 11 to 85 bp deletions.

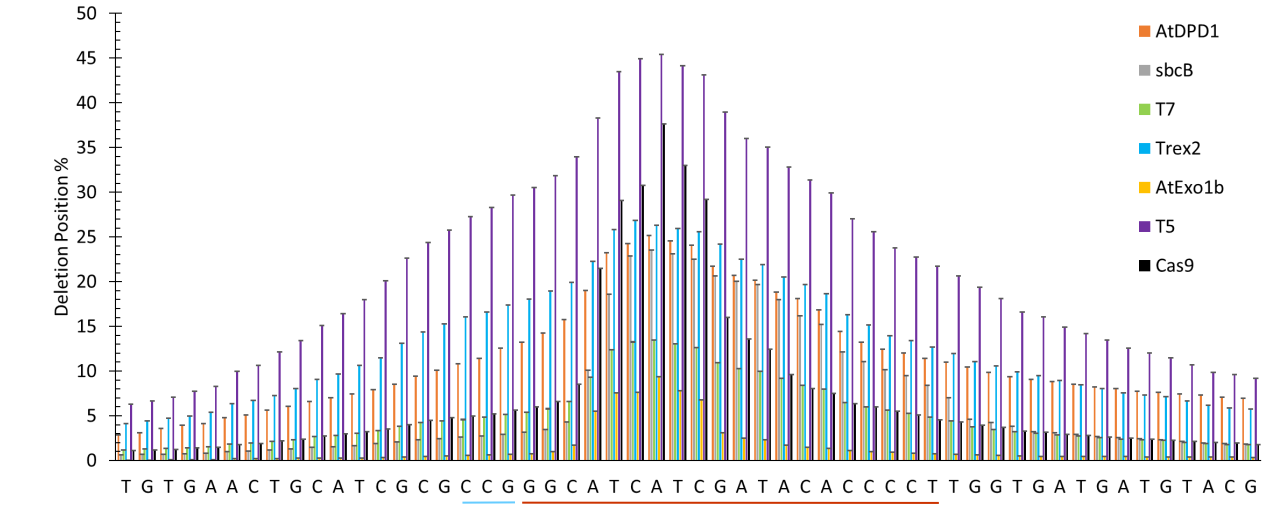


Figure 8: Deletion profiles of Cas9 N-terminus fusions at LOC_Os04g22730. Bars indicate the percentage of reads that showed a deletion at that nucleotide. The red line labels the sgRNA target, and the blue lines is the PAM site. Deletion position target B) Deletion position at LOC_Os07g48350 target. Data presented as the mean +/- SEM (standard error mean). Three technical repeats were analyzed for all samples.

4.2.3 Comparison of Cas12a-exonuclease fusions in rice protoplasts

The six previously tested exonucleases were fused to the N-terminus of LbCas12a and tested at four different target sites (Figure 9). Because many of these exonucleases robustly degrade ssDNA, there was interest to see if any exonucleases would produce different editing outcomes when presented with 5' overhangs by Cas12a in comparison to the blunt ends generated by Cas9. Indeed, Cas12a-exonucleases showed a different editing pattern than Cas9-exonuclease. T7, AtExo1b, AtDPD1, and T5 produced similar average editing efficiencies which were lower than Cas12a at the four target sites. sbcB and TREX2 produced higher average editing efficiencies, although only sbcB produced statistically significant higher editing than Cas12a at the AA1 target site (Figure 10a).

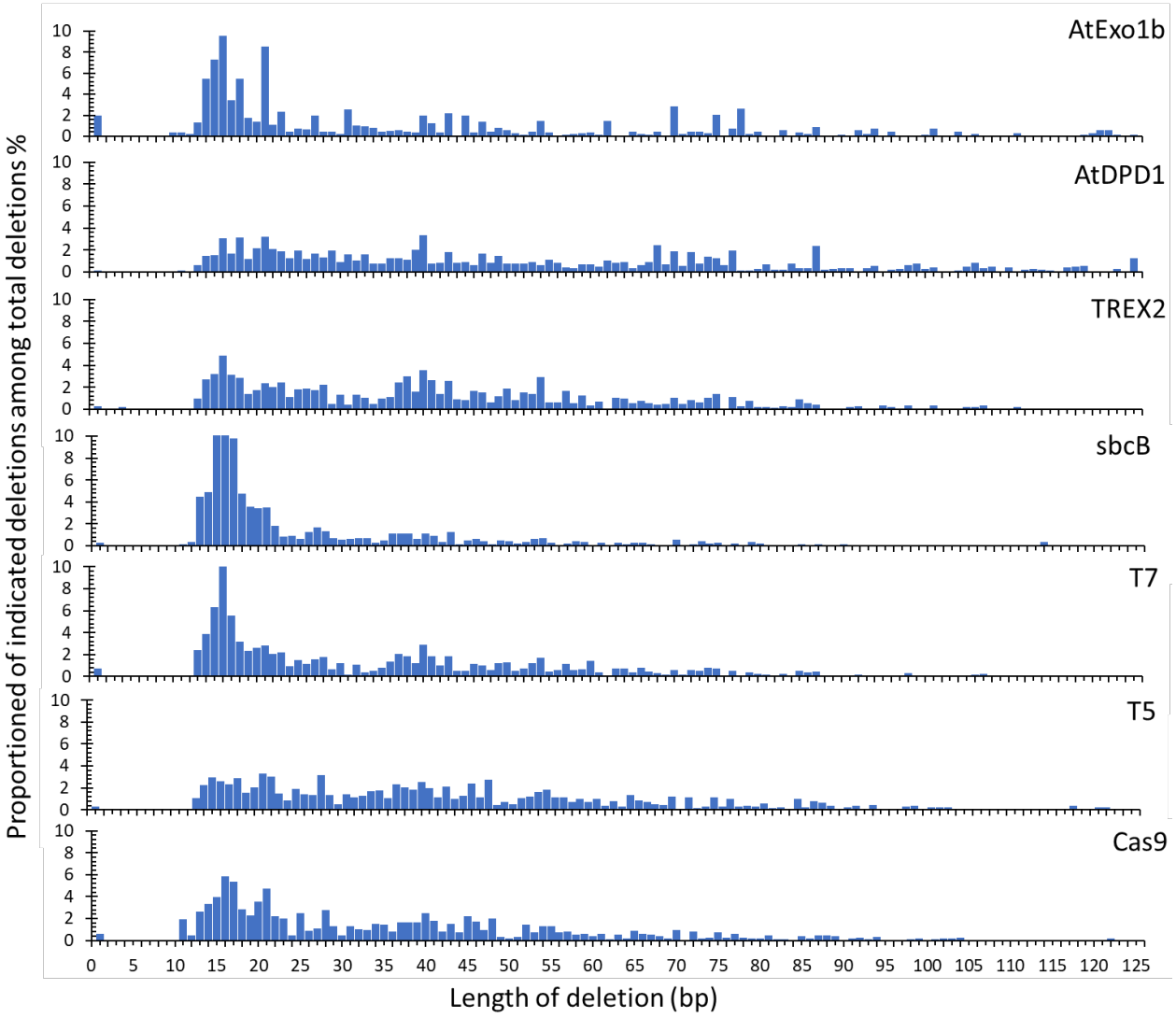


Figure 9: Percentage of deletion lengths by Cas9-exonuclease fusions at LOC_Os04g22730. Bars indicate the percentage of reads that showed a deletion at that nucleotide. Data presented as the mean +/- SEM (standard error mean). Three technical repeats were analyzed for all samples.

In addition to editing efficiency, deletion profiles were created for each target site (Figure 11). Deletion percentages peaked between the 13th and 23rd base pairs from the end of the PAM for Cas12a. T5, T7, and AtExo1b showed slightly larger

deletions, but sbcB and TREX2 produced the largest deletions at all target sites. AtDPD1 also produced larger deletions, most notably at the GC1 target (Figure 11). Interestingly, Cas12a-sbcB produced more deletions downstream from the crRNA target site, which contrasts the edits seen from Cas9-exonucleases which flank the sgRNA target site. The length of the deletions produced by each Cas12a-exonuclease at the targets was calculated as a percentage of deletions categorized by base-pair length (Figure 12). Categorizing the deletions by length instead of position allows for analysis beyond the crRNA target. All six Cas12a-exonucleases had a higher percentage of deletions over 20 bp than Cas12a. Cas12a deletions were mostly between 8 to 11 bp long, with a spike of deletions over 20 bp. SbcB had the highest percentages of deletions over 20 bp at all target sites, which correlates with the large percentage of deletions observed at the crRNA target site (Figure 11). These results suggest that Cas12a-exonucleases are capable of producing long deletions, and that different deletion profiles can be observed between Cas9 and Cas12a fusions.

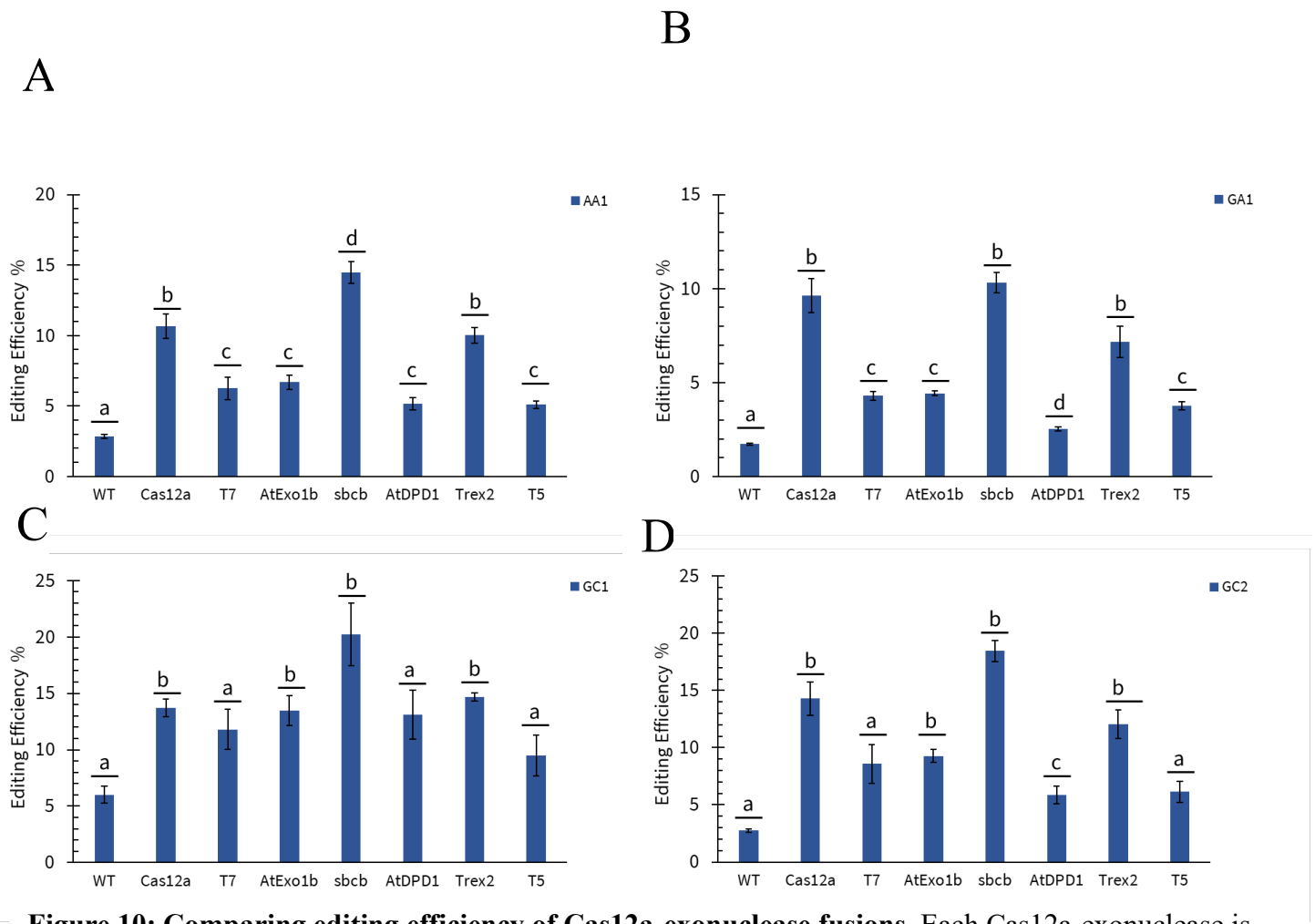
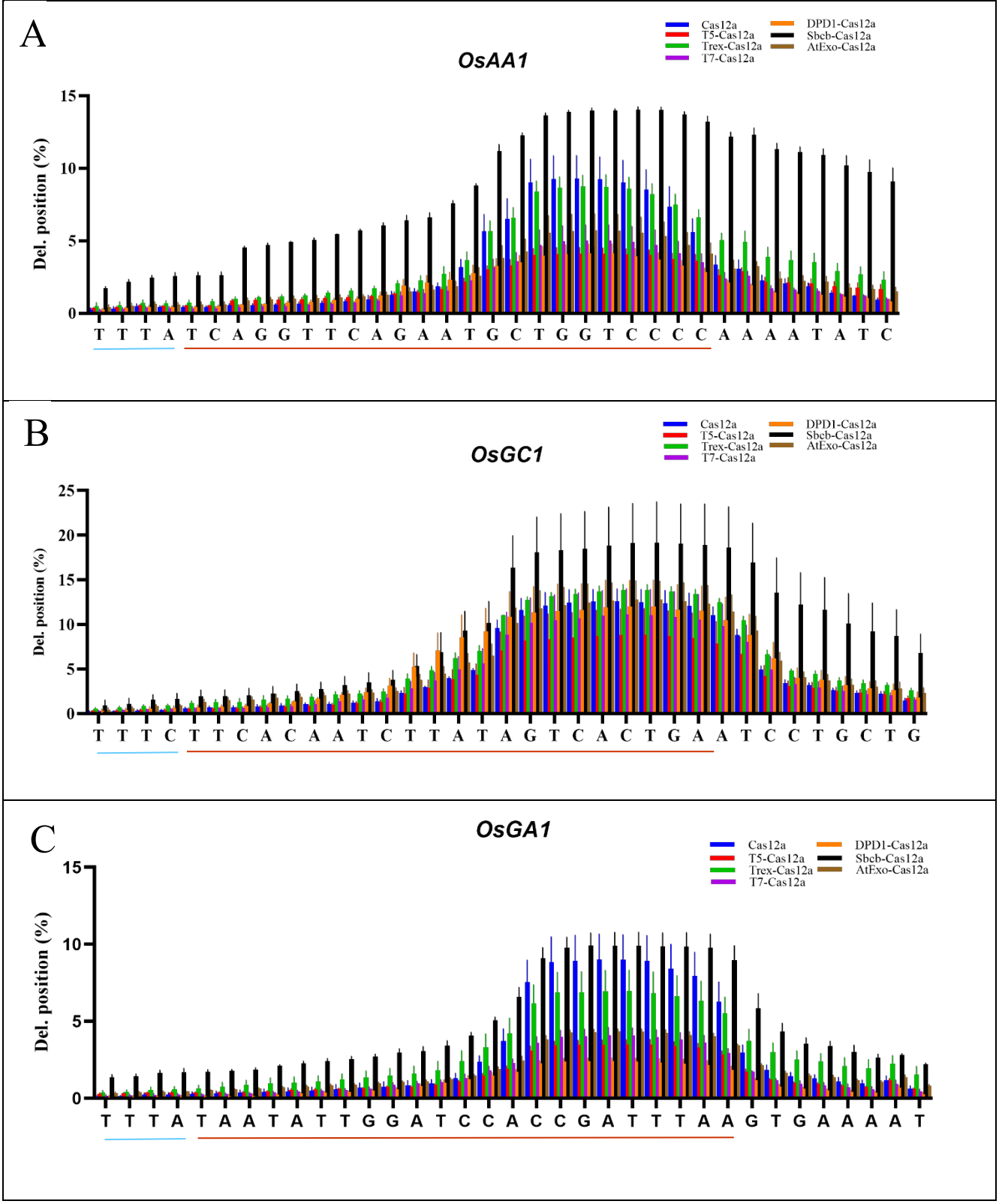


Figure 10: Comparing editing efficiency of Cas12a-exonuclease fusions. Each Cas12a-exonuclease is compared to Cas12a, a wild-type sample, and the other Cas12a fusions. A) Cas12a-exonucleases targeting AA1 B) Cas12a-exonucleases targeting GA1 C) Cas12a-exonucleases targeting GC1. D) Cas12a-exonucleases targeting GC2. Matching "a" and "b" labels indicate there no significant increase in editing efficiency between treatments when $\alpha = 0.05$ as determined by a pairwise T-test. Data presented as the mean +/- SEM (standard error mean). Three technical repeats were analyzed for all samples.


A
OsAA1
Cas12a
T5-Cas12a
Trex-Cas12a
T7-Cas12a
DPD1-Cas12a
Sbcb-Cas12a
AtExo-Cas12a
Del. position (%)
15
10
5
0
T T T A T C A G G T T C A G A A T G C T G G T C C C C A A A A T A T C
B
OsGC1
Cas12a
T5-Cas12a
Trex-Cas12a
T7-Cas12a
DPD1-Cas12a
Sbcb-Cas12a
AtExo-Cas12a
Del. position (%)
25
20
15
10
5
0
T T T C T T C A C A A T C T T A T A G T C A C T G A A T C C T G C T G
C
OsGA1
Cas12a
T5-Cas12a
Trex-Cas12a
T7-Cas12a
DPD1-Cas12a
Sbcb-Cas12a
AtExo-Cas12a
Del. position (%)
15
10
5
0
T T T A T A A T A T T G G A T C C A C C G A T T T A A G T G A A A A T

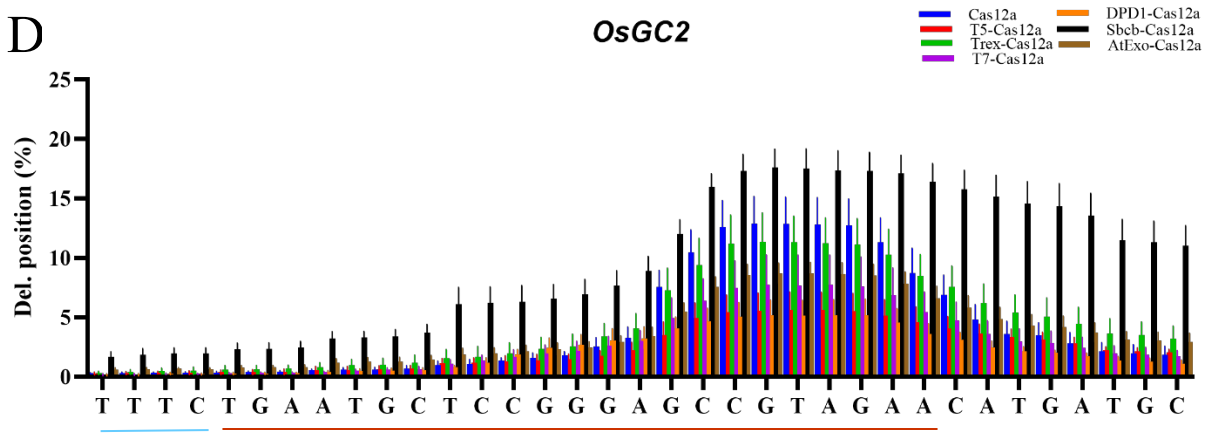
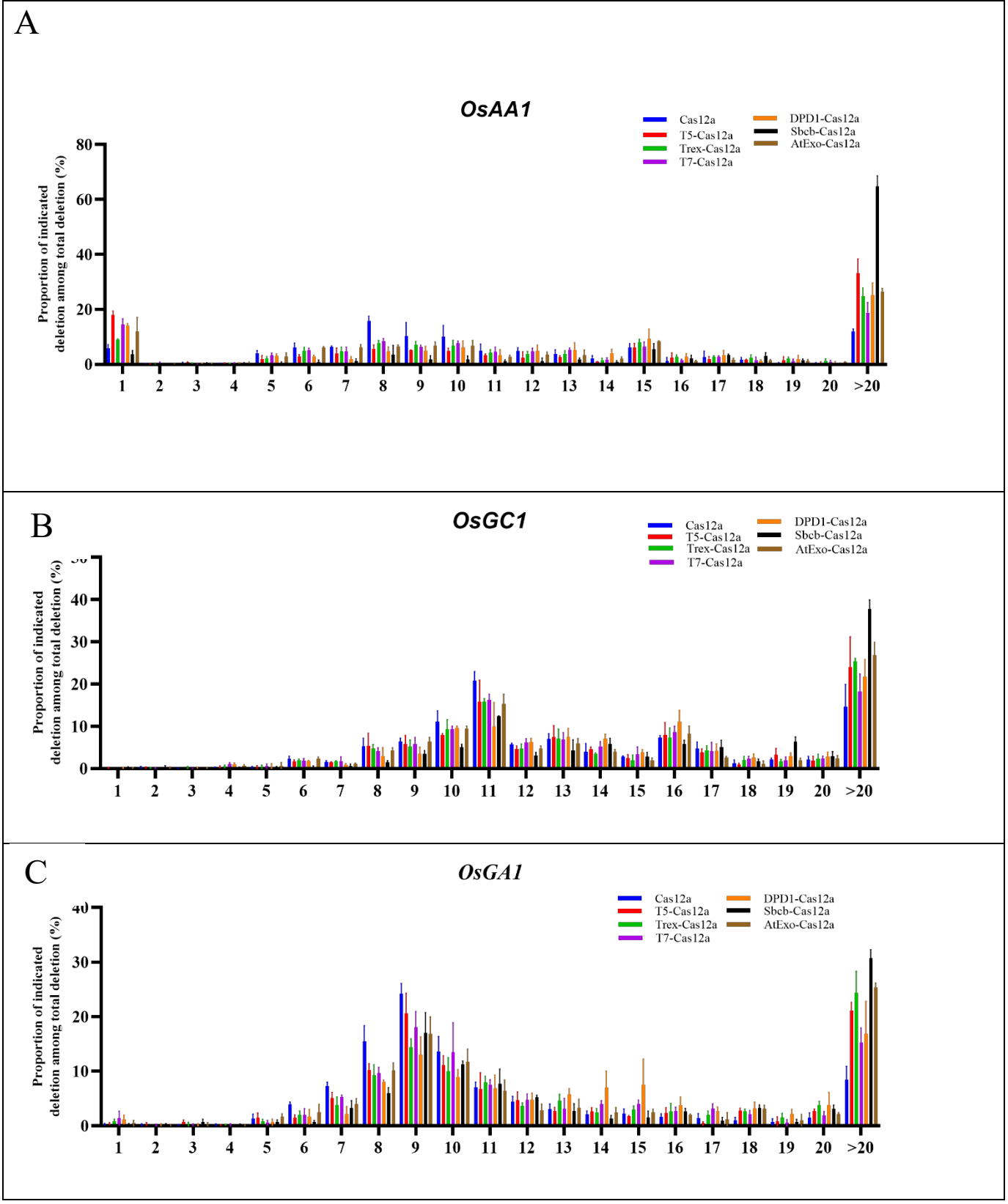


Figure 10: Deletion profiles of Cas9 N-terminus fusions. Bars indicate the percentage of reads that showed a deletion at that nucleotide. A) Cas12a-exonucleases targeting AA1 B) Cas12a-exonucleases targeting GA1 C) Cas12a-exonucleases targeting GC1. D) Cas12a-exonucleases targeting GC2. Data presented as the mean +/- SEM (standard error mean). Three technical repeats were analyzed for all samples.


A
OsAA1
Cas12a
T5-Cas12a
Trex-Cas12a
T7-Cas12a
DPD1-Cas12a
Sbcb-Cas12a
AtExo-Cas12a
Proportion of indicated deletion among total deletion (%)
0
20
40
60
80
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 >20
B
OsGC1
Cas12a
T5-Cas12a
Trex-Cas12a
T7-Cas12a
DPD1-Cas12a
Sbcb-Cas12a
AtExo-Cas12a
Proportion of indicated deletion among total deletion (%)
0
10
20
30
40
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 >20
C
OsGA1
Cas12a
T5-Cas12a
Trex-Cas12a
T7-Cas12a
DPD1-Cas12a
Sbcb-Cas12a
AtExo-Cas12a
Proportion of indicated deletion among total deletion (%)
0
10
20
30
40
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 >20

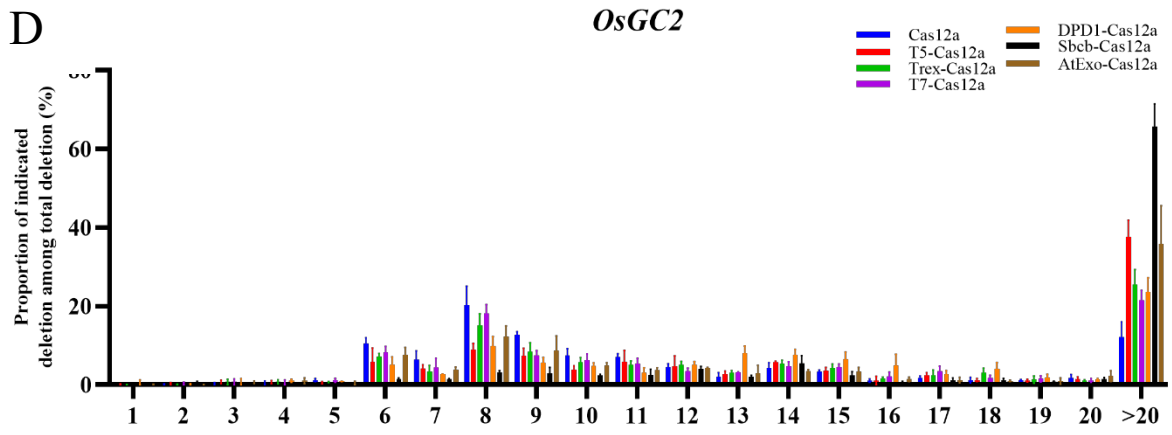


Figure 11: Percentage of deletion lengths by Cas12a-exonuclease fusions. Bars indicate the percentage of reads that showed a deletion at that nucleotide. A) Cas12a-exonucleases targeting AA1 B) Cas12a-exonucleases targeting GA1 C) Cas12a-exonucleases targeting GC1. D) Cas12a-exonucleases targeting GC2. Data presented as the mean +/- SEM (standard error mean). Three technical repeats were analyzed for all samples.

4.3 Discussion

Several studies suggest that chromatin structure affects Cas9 editing, but the level of impact, and underlying mechanism, is unclear [161, 174, 189]. Many variables including timing it takes to induce DSBs, NHEJ outcomes, and level of gene activation could interplay with chromatin and impact the editing efficiency. Cas9-CMPs could provide interesting insights into chromatin structure and regulation and interactions between Cas9 and chromatin. Chromatin regulation is thought to be highly conserved, so binding domains from humans should be effective in other species, although future work will likely include different domains sourced from

various species. Potentially, Cas9-CMPs could shed insight onto the process and mechanisms of chromatin remodeling. Beyond insights into CRISPR mechanism, Cas-CMPs could possibly be used to improve other applications of CRISPR, such as gene regulation. Catalytically dead Cas proteins carrying an activator or repressor remain bound to the target site but can be removed from the DNA by other proteins. It will be interesting to see whether dCas fusions with CMPs or other DNA binding domains, can improve gene regulation outcomes through tighter DNA binding.

Exonucleases have been used in genome editing since TALEN to improve genome editing efficiency and produce larger deletions [148, 149, 190, 191]. However, more research on using exonucleases with CRISPR systems and creating Cas-fusion proteins needs to be carried out, especially in plants where co-transformation of Cas and a separate exonuclease is impractical. Increased editing efficiency was observed with both Cas9 and Cas12a-fusions and the majority of fusions produced longer deletions. Variability in the results may be reduced with further analysis in stable rice plants rather than protoplasts. Illumina NGS sequencing produces a large amount of reads, but is size limited, which could allow larger deletions to be excluded from analysis.

CRISPR tools have been rapidly developed since the first CRISPR-Cas9 paper ten years ago and are becoming commonplace in labs. Demand continues to grow for highly efficient CRISPR systems, as well as for systems that can produce different editing outcomes for a variety of applications. Further research will likely expand Cas-exonuclease fusions and pair exonucleases with systems such as nickases and base editing.

4.4 Materials and Methods

4.4.1 Vector Construction

Vectors were constructed to create the Cas-fusion genes, both an N-terminus and C-terminus version. A vector (pYPQ166) containing a zCas9 was digested and a gene block of AtExo1b was inserted via restriction enzyme cloning to create pYPQ166-N-AtExo1b. "N" indicates this is the N-terminus version. To create the C-terminus version, pYPQ166 was digested and ligated with gene block AtExo1b to create pYPQ166- C-AtExo1b. The rest of the exonuclease vectors were made by removing the AtExo1b gene with restriction enzymes and ligating the individual exonucleases into the vector at the XTEN linker and NLS sequence. The Cas12a-exonuclease vectors were created with NEB HiFi cloning by amplifying the exonuclease from the corresponding Cas9-exonuclease vector with high fidelity polymerase Q5 and inserting the exonuclease gene into LbCas12a vector pYPQ230. Entry vectors for the sgRNAs and crRNAs, as well as the creation of the final T-DNA vector, were constructed with the methods described in Chapter 2.

4.4.2 Rice Protoplast transformation

Rice protoplast creation and PEG transformation followed a previously published protocol, with minor modifications listed in the following description [192]. Protoplasts were made from the stems of Kitaake *(japonica* family) rice plants that were grown on 1/2 MS media for three weeks in the dark at 25 °C. After digestion, the cells underwent PEG transformation with 25 µg of plasmid. Plasmids

were eluted in water and concentrations of 1,000 ng/µl were obtained via midiprep with Qiagen Plasmid Midi Kit. Transformed cells were incubated for 48 hours before being collected.

4.4.3 Stable rice transformation

Rice calli induction, transformation, and regeneration from Kitaake seeds followed a previously published protocol, with minor modifications listed in the following description [193]. Calli were grown on calli induction media in a tissue culture chamber at 28 °C with a long-day cycle (16 hours light/8 hours dark). The agrobacterium strain GV3101 was used to transform the calli. Regeneration was carried out according to the protocol, and calli were transferred to the final regeneration media after shoots and roots appeared. Care was taken to not split the calli.

4.4.4 Amplification of target genes and Next-Gen Sequencing

Target genes were amplified with the ThermoFisher Phire Kit using barcoded primers for NGS with analysis via Hi-TOM [194]. An Illumina adaptor sequence was added to sequence-specific primers and used with a second set of barcoding primers according to the Hi-TOM protocol. Pooled amplicons were sequenced by Genewiz on an Illumina machine as part of their Amplicon-EZ service.

Chapter 5: Conclusions and Future Directions

Development of Cas12a for genome engineering

The goal of this work was to develop novel CRISPR genome engineering systems in order to provide more efficient tools to the plant science community and expand genome engineering applications. In summary, this was accomplished by exploring natural Cas variants and engineering Cas-effector protein fusions. First, novel LbCas12a was tested in Arabidopsis, with single and multiplexed targeting, for genome editing and gene repression. Second, Cas9 and Cas12a was fused to effector proteins which increased editing efficiency and produced longer deletions. The following paragraphs provide further detail as well as possibilities for future CRISPR research.

In Chapter 2, novel endonuclease LbCas12a (formerly Cpf1) was tested in Arabidopsis. LbCas12a is an important expansion of the CRISPR toolkit because it can target T-rich PAMs, and produce staggered DSB overhangs with a compact crRNA. While previous work had demonstrated high efficiency editing in rice, Arabidopsis editing with LbCas12a was unsuccessful. Our results show that LbCas12a is temperature sensitive and editing efficiency could be achieved by heat treating Arabidopsis plants at 29 °C. Further exploration of the LbCas12a mechanism revealed that gene repression could be achieved at lower temperatures, and a multiplexed LbCas12a gene repression system was developed to target four genes. Because many promoters are T-rich, LbCas12a-mediated gene repression is highly desirable. Although there was variation between targets and plant lines, gene repression was successfully demonstrated, making this the first multiplexed LbCas12a gene repression system demonstrated in Arabidopsis.

Further research into LbCas12a will address temperature sensitivity, possibly through several strategies. Engineered Cas nucleases have been developed through mutating the protein domains as well as with continuous evolution systems. For example, LbCas12aD156R, also known as ttLbCas12a, was an engineered LbCas12a variant that is more tolerant than the WT LbCas12a at lower temperature [195, 196]. For targeted mutations in the different domains, more insight into the temperature sensitive mechanism(s) will be required. Because gene repression can be achieved at Arabidopsis's normal growth condition at 22 °C, gene repression systems can be improved with the canonical LbCas12a. Possibly novel repressors will create more efficient repression systems, and when paired with induced expression systems or tissue-specific promoters, will allow researchers to finely control gene expression.

Engineering Cas9 and Cas12a-effector fusions

In Chapter 3, effector proteins were added to Cas9 and Cas12a to increase editing efficiency and produce large deletions. First, chromatin-modulating peptides (CMPs) were added to Cas9 to create Cas9-CMP fusions. These fusions resulted in higher editing efficiency with a similar deletion profile to Cas9. Second, six exonucleases were fused to both Cas9 and Cas12a. Of the six exonucleases, only one (T7) did not result in a higher average editing efficiency at both Cas9 target sites. In contrast, only sbcB resulted in a higher average editing efficiency with Cas12a. This suggest that the staggered overhang after Cas12a cleavage may already promote end

processing by plant endogenous exonucleases. Several Cas9 exonuclease fusions produced larger deletions and the deleted base pairs spanned upstream and downstream from the gRNA target sequence. Cas12a-sbcB fusion created a higher percentage of large deletions compared to Cas12a, and Cas12a-sbcB was also more likely to delete base pairs upstream and downstream from the crRNA target sequence than Cas12a, although Cas12a-sbcB deleted more bases downstream than upstream. These results demonstrate novel Cas-fusions in rice protoplasts for increased efficiency and larger deletions.

Future research will test a variety of chromatin-modulating peptides from diverse sources along with other DNA-binding domains to increase editing efficiency. In gene regulation, the Cas protein stays bound to the target DNA, and it will be interesting to see if CMPs facilitate DNA binding for enhanced gene regulation. Additionally, CMPs can also be added to low efficiency Cas variants that have unique characteristics, such as PAM recognition, but haven't been widely adopted because the mutation efficiency is so low. Future research should also test for any off-target effects that may occur as well as determine whether CMPs affect Cas binding kinetics. Cas-exonuclease fusions will likely continue to improve editing efficiency, and a diverse array of exonucleases can be tested. As with Cas-CMPs, exonucleases can possibly be paired with low efficiency Cas systems, as well as with nickases that may trigger exonuclease activity. Furthermore, some of the 5' to 3' exonuclease discovered in this study may be used in the prime editing system, an exciting precise genome editing technology, to boost the prime editing efficiency [197].

Regulation and policy governing CRISPR crops

CRISPR has moved beyond basic research into use for crop breeding to create genetically engineered foods for human consumption. The USDA's Sustainable, Ecological, Consistent, Uniform, Responsible, and Efficient (SECURE) rule has updated agricultural policy to regulate gene-edited crops produced with tools such as CRISPR [198, 199]. The US has product-driven regulation, meaning that the final product determines regulation, not whether it was produced with conventional or traditional methods. The SECURE rule was established in 2020 and exempts plants from regulation if the genome contains a single base pair substitution, a deletion that results from endogenous DNA repair, or a modification that has been previously regulated or can be found in nature. Notably, only one modification can be introduced. Crops with multiple edits can be exempt from regulation, as long as the edits are introduced individually and combined with conventional breeding. The regulatory process is carried out by the Animal and Plant Health Inspection Service (APHIS). Developers submit a permit request to APHIS, which determines whether the plant could potentially become a pest. Plants that pose low risk would then be analyzed to determine if they are regulated as genetically modified. Newly developed crops that are exempt from regulation can be submitted voluntarily to the USDA for approval.

In contrast to US policy, European regulation focuses on the method of production [200]. Plants with single base pair changes would be regulated with the same strict standards as genetically modified plants that harbor a transgene. This regulation has limited the planting of genome edited crops within the European Union

(EU). Imports of genetically modified and genome edited crops are allowed. Differing regulation between countries can create problems for farmers and stakeholders in agricultural trade. For example, African countries trade predominantly with the EU. Many regions would benefit from crop breeding efforts on nutrition, drought tolerance, and disease resistance, but adopt the strict genome editing policies of EU trading partners. Many countries with national biosafety committees will soon need to update their policies in regard to trade, academic research, field trials of genetically engineered crops, receiving foreign aid, and industry crop breeding, to name a few examples. Countries without national biosafety committees will need support to establish the committee to write agricultural policy. As CRISPR systems continue to rapidly evolve and become more commonplace in plant science, biotech scientists will serve as technical experts and help develop science-based policies.

Bibliography

1. Ray DK, Mueller ND, West PC, Foley JA (2013) Yield Trends Are Insufficient to Double Global Crop Production by 2050. PLOS ONE 8:e66428. https://doi.org/10.1371/journal.pone.0066428

2. Wolt JD, Wang K, Yang B (2016) The Regulatory Status of Genome-edited Crops. Plant Biotechnol J 14:510–518. https://doi.org/10.1111/pbi.12444

3. Hartlerode AJ, Scully R (2009) Mechanisms of double-strand break repair in somatic mammalian cells. Biochemical Journal 423:157–168. https://doi.org/10.1042/BJ20090942

4. Kanaar R, Hoeijmakers JHJ (1998) Molecular mechanisms of DNA double-strand break repair. trends in CELL BIOLOGY 8:7

5. Steinert J, Schiml S, Puchta H (2016) Homology-based double-strand break-induced genome engineering in plants. Plant Cell Rep 35:1429–1438. https://doi.org/10.1007/s00299-016-1981-3

6. Pastwa E, Błasiak J (2003) Non-homologous DNA end joining. Acta Biochim Pol 50:891–908. https://doi.org/035004891

7. Wang Y, Cheng X, Shan Q, Zhang Y, Liu J, Gao C, Qiu J-L (2014) Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. Nature Biotechnology 32:947–951. https://doi.org/10.1038/nbt.2969

8. Puchta H (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. Journal of Experimental Botany 56:1–14. https://doi.org/10.1093/jxb/eri025

9. Townsend JA, Wright DA, Winfrey RJ, Fu F, Maeder ML, Joung JK, Voytas DF (2009) High-frequency modification of plant genes using engineered zinc-finger nucleases. Nature 459:442–445. https://doi.org/10.1038/nature07845

10. Pâques F, Duchateau P Meganucleases and DNA Double-Strand Break-Induced Recombination: Perspectives for Gene Therapy. 18

11. Gao H, Smith J, Yang M, Jones S, Djukanovic V, Nicholson MG, West A, Bidney D, Falco SC, Jantz D, Lyznik LA (2010) Heritable targeted mutagenesis in maize using a designed endonuclease. Plant J 61:176–187. https://doi.org/10.1111/j.1365-313X.2009.04041.x

12. Kim YG, Cha J, Chandrasegaran S (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A 93:1156–1160

13. Bibikova M, Carroll D, Segal DJ, Trautman JK, Smith J, Kim YG, Chandrasegaran S (2001) Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. Mol Cell Biol 21:289–297. https://doi.org/10.1128/MCB.21.1.289-297.2001

14. Engineering polydactyl zinc-finger transcription factors | Nature Biotechnology. https://www.nature.com/articles/nbt0202-135. Accessed 11 Nov 2022

15. Segal DJ, Beerli RR, Blancafort P, Dreier B, Effertz K, Huber A, Koksch B, Lund CV, Magnenat L, Valente D, Barbas CF (2003) Evaluation of a modular strategy for the construction of novel polydactyl zinc finger DNA-binding proteins. Biochemistry 42:2137–2148. https://doi.org/10.1021/bi026806o

16. Ramirez CL, Foley JE, Wright DA, Müller-Lerch F, Rahman SH, Cornu TI, Winfrey RJ, Sander JD, Fu F, Townsend JA, Cathomen T, Voytas DF, Joung JK (2008) Unexpected failure rates for modular assembly of engineered zinc fingers. Nat Methods 5:374–375. https://doi.org/10.1038/nmeth0508-374

17. Cornu TI, Thibodeau-Beganny S, Guhl E, Alwin S, Eichtinger M, Joung JK, Joung JK, Cathomen T (2008) DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. Mol Ther 16:352–358. https://doi.org/10.1038/sj.mt.6300357

18. Doyon Y, McCammon JM, Miller JC, Faraji F, Ngo C, Katibah GE, Amora R, Hocking TD, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Amacher SL (2008) Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. Nat Biotechnol 26:702–708. https://doi.org/10.1038/nbt1409

19. Maeder ML, Thibodeau-Beganny S, Osiak A, Wright DA, Anthony RM, Eichtinger M, Jiang T, Foley JE, Winfrey RJ, Townsend JA, Unger-Wallace E, Sander JD, Müller-Lerch F, Fu F, Pearlberg J, Göbel C, Dassie JP, Pruett-Miller SM, Porteus MH, Sgroi DC, Iafrate AJ, Dobbs D, McCray PB, Cathomen T, Voytas DF, Joung JK (2008) Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol Cell 31:294–301. https://doi.org/10.1016/j.molcel.2008.06.016

20. Sander JD, Dahlborg EJ, Goodwin MJ, Cade L, Zhang F, Cifuentes D, Curtin SJ, Blackburn JS, Thibodeau-Beganny S, Qi Y, Pierick CJ, Hoffman E, Maeder ML, Khayter C, Reyon D, Dobbs D, Langenau DM, Stupar RM, Giraldez AJ, Voytas DF, Peterson RT, Yeh J-RJ, Joung JK (2011) Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). Nat Methods 8:67–69. https://doi.org/10.1038/nmeth.1542

21. 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. Proceedings of the National Academy of Sciences 107:12028–12033. https://doi.org/10.1073/pnas.0914991107

22. Qi Y, Zhang Y, Zhang F, Baller JA, Cleland SC, Ryu Y, Starker CG, Voytas DF (2013) Increasing frequencies of site-specific mutagenesis and gene targeting in Arabidopsis by manipulating DNA repair pathways. Genome Res 23:547–554. https://doi.org/10.1101/gr.145557.112

23. Qi Y, Zhang Y, Baller JA, Voytas DF (2016) Histone H2AX and the small RNA pathway modulate both non-homologous end-joining and homologous recombination in plants. Mutat Res 783:9–14. https://doi.org/10.1016/j.mrfmmm.2015.12.002

24. Moscou MJ, Bogdanove AJ (2009) A Simple Cipher Governs DNA Recognition by TAL Effectors. Science 326:1501–1501. https://doi.org/10.1126/science.1178817

25. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326:1509–1512. https://doi.org/10.1126/science.1178811

26. Boch J, Bonas U (2010) Xanthomonas AvrBs3 family-type III effectors: discovery and function. Annu Rev Phytopathol 48:419–436. https://doi.org/10.1146/annurev-phyto-080508-081936

27. Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, Bogdanove AJ, Voytas DF (2010) Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186:757–761. https://doi.org/10.1534/genetics.110.120717

28. Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DP, Yang B (2011) TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. Nucleic Acids Res 39:359–372. https://doi.org/10.1093/nar/gkq704

29. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, Dulay GP, Hua KL, Ankoudinova I, Cost GJ, Urnov FD, Zhang HS, Holmes MC, Zhang L, Gregory PD, Rebar EJ (2011) A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29:143–148. https://doi.org/10.1038/nbt.1755

30. Mahfouz MM, Li L, Shamimuzzaman M, Wibowo A, Fang X, Zhu J-K (2011) De novo-engineered transcription activator-like effector (TALE) hybrid

nuclease with novel DNA binding specificity creates double-strand breaks. Proc Natl Acad Sci U S A 108:2623–2628. https://doi.org/10.1073/pnas.1019533108

31. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Research 39:e82–e82. https://doi.org/10.1093/nar/gkr218

32. Li T, Huang S, Zhao X, Wright DA, Carpenter S, Spalding MH, Weeks DP, Yang B (2011) Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. Nucleic Acids Res 39:6315–6325. https://doi.org/10.1093/nar/gkr188

33. Morbitzer R, Elsaesser J, Hausner J, Lahaye T (2011) Assembly of custom TALE-type DNA binding domains by modular cloning. Nucleic Acids Res 39:5790–5799. https://doi.org/10.1093/nar/gkr151

34. Wiedenheft B, Sternberg SH, Doudna JA (2012) RNA-guided genetic silencing systems in bacteria and archaea. Nature 482:331–338. https://doi.org/10.1038/nature10886

35. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P (2007) CRISPR provides acquired resistance against viruses in prokaryotes. Science 315:1709–1712. https://doi.org/10.1126/science.1138140

36. Marraffini LA, Sontheimer EJ (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science 322:1843–1845. https://doi.org/10.1126/science.1165771

37. Makarova KS, Zhang F, Koonin EV (2017) SnapShot: Class 2 CRISPR-Cas Systems. Cell 168:328-328.e1. https://doi.org/10.1016/j.cell.2016.12.038

38. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science 337:816–821. https://doi.org/10.1126/science.1225829

39. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM (2013) RNA-guided human genome engineering via Cas9. Science 339:823–826. https://doi.org/10.1126/science.1232033

40. Paul JW, Qi Y (2016) CRISPR/Cas9 for plant genome editing: accomplishments, problems and prospects. Plant Cell Reports 35:1417–1427. https://doi.org/10.1007/s00299-016-1985-z

41. Lowder L, Malzahn A, Qi Y (2016) Rapid evolution of manifold CRISPR systems for plant genome editing. Front Plant Sci 7:1683. https://doi.org/10.3389/fpls.2016.01683

42. Swarts DC, Jinek M Cas9 versus Cas12a/Cpf1: Structure–function comparisons and implications for genome editing. Wiley Interdisciplinary Reviews: RNA 0:e1481. https://doi.org/10.1002/wrna.1481

43. Martens KJA, van Beljouw SPB, van der Els S, Vink JNA, Baas S, Vogelaar GA, Brouns SJJ, van Baarlen P, Kleerebezem M, Hohlbein J (2019) Visualisation of dCas9 target search in vivo using an open-microscopy framework. Nat Commun 10:3552. https://doi.org/10.1038/s41467-019-11514-0

44. Rutkauskas M, Sinkunas T, Songailiene I, Tikhomirova MS, Siksnys V, Seidel R (2015) Directional R-Loop Formation by the CRISPR-Cas Surveillance Complex Cascade Provides Efficient Off-Target Site Rejection. Cell Reports 10:1534–1543. https://doi.org/10.1016/j.celrep.2015.01.067

45. Gong S, Yu HH, Johnson KA, Taylor DW (2018) DNA Unwinding Is the Primary Determinant of CRISPR-Cas9 Activity. Cell Reports 22:359–371. https://doi.org/10.1016/j.celrep.2017.12.041

46. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9 | Nature. Nature 62–67

47. Swarts DC, van der Oost J, Jinek M (2017) Structural Basis for Guide RNA Processing and Seed-Dependent DNA Targeting by CRISPR-Cas12a. Molecular Cell 66:221–233. https://doi.org/10.1016/j.molcel.2017.03.016

48. Yamano T, Nishimasu H, Zetsche B, Hirano H, Slaymaker IM, Li Y, Fedorova I, Nakane T, Makarova KS, Koonin EV, Ishitani R, Zhang F, Nureki O (2016) Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA. Cell 165:949–962. https://doi.org/10.1016/j.cell.2016.04.003

49. Banakar R, Schubert M, Collingwood M, Vakulskas C, Eggenberger AL, Wang K (2020) Comparison of CRISPR-Cas9/Cas12a Ribonucleoprotein Complexes for Genome Editing Efficiency in the Rice Phytoene Desaturase (OsPDS) Gene. Rice (N Y) 13:4. https://doi.org/10.1186/s12284-019-0365-z

50. Lee K, Zhang Y, Kleinstiver BP, Guo JA, Aryee MJ, Miller J, Malzahn A, Zarecor S, Lawrence-Dill CJ, Joung JK, Qi Y, Wang K (2019) Activities and specificities of CRISPR/Cas9 and Cas12a nucleases for targeted mutagenesis in maize. Plant Biotechnol J 17:362–372. https://doi.org/10.1111/pbi.12982

51. 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 Engineering1[W][OA]. Plant Physiol 161:20–27. https://doi.org/10.1104/pp.112.205179

52. Christian M, Qi Y, Zhang Y, Voytas DF (2013) Targeted Mutagenesis of Arabidopsis thaliana Using Engineered TAL Effector Nucleases. G3 Genes|Genomes|Genetics 3:1697–1705. https://doi.org/10.1534/g3.113.007104

53. Lor VS, Starker CG, Voytas DF, Weiss D, Olszewski NE (2014) Targeted Mutagenesis of the Tomato PROCERA Gene Using Transcription Activator-Like Effector Nucleases1[W]. Plant Physiol 166:1288–1291. https://doi.org/10.1104/pp.114.247593

54. Shan Q, Wang Y, Chen K, Liang Z, Li J, Zhang Y, Zhang K, Liu J, Voytas DF, Zheng X, Zhang Y, Gao C (2013) Rapid and Efficient Gene Modification in Rice and Brachypodium Using TALENs. Mol Plant 6:1365–1368. https://doi.org/10.1093/mp/sss162

55. Zhang H, Gou F, Zhang J, Liu W, Li Q, Mao Y, Botella JR, Zhu J-K (2016) TALEN-mediated targeted mutagenesis produces a large variety of heritable mutations in rice. Plant Biotechnol J 14:186–194. https://doi.org/10.1111/pbi.12372

56. Haun W, Coffman A, Clasen BM, Demorest ZL, Lowy A, Ray E, Retterath A, Stoddard T, Juillerat A, Cedrone F, Mathis L, Voytas DF, Zhang F (2014) Improved soybean oil quality by targeted mutagenesis of the fatty acid desaturase 2 gene family. Plant Biotechnol J 12:934–940. https://doi.org/10.1111/pbi.12201

57. Shan Q, Zhang Y, Chen K, Zhang K, Gao C (2015) Creation of fragrant rice by targeted knockout of the OsBADH2 gene using TALEN technology. Plant Biotechnology Journal 13:791–800. https://doi.org/10.1111/pbi.12312

58. Ma L, Zhu F, Li Z, Zhang J, Li X, Dong J, Wang T (2015) TALEN-Based Mutagenesis of Lipoxygenase LOX3 Enhances the Storage Tolerance of Rice (Oryza sativa) Seeds. PLoS One 10:e0143877. https://doi.org/10.1371/journal.pone.0143877

59. Clasen BM, Stoddard TJ, Luo S, Demorest ZL, Li J, Cedrone F, Tibebu R, Davison S, Ray EE, Daulhac A, Coffman A, Yabandith A, Retterath A, Haun W, Baltes NJ, Mathis L, Voytas DF, Zhang F (2016) Improving cold storage and processing traits in potato through targeted gene knockout. Plant Biotechnology Journal 14:169–176. https://doi.org/10.1111/pbi.12370

60. Li T, Liu B, Spalding MH, Weeks DP, Yang B (2012) High-efficiency TALEN-based gene editing produces disease-resistant rice. Nat Biotechnol 30:390–392. https://doi.org/10.1038/nbt.2199

61. Li M, Li X, Zhou Z, Wu P, Fang M, Pan X, Lin Q, Luo W, Wu G, Li H (2016) Reassessment of the Four Yield-related Genes Gn1a, DEP1, GS3, and IPA1 in Rice Using a CRISPR/Cas9 System. Frontiers in Plant Science 7:. https://doi.org/10.3389/fpls.2016.00377

62. Li Q, Zhang D, Chen M, Liang W, Wei J, Qi Y, Yuan Z (2016) Development of japonica Photo-Sensitive Genic Male Sterile Rice Lines by Editing Carbon Starved Anther Using CRISPR/Cas9. J Genet Genomics 43:415–419. https://doi.org/10.1016/j.jgg.2016.04.011

63. Zhou H, He M, Li J, Chen L, Huang Z, Zheng S, Zhu L, Ni E, Jiang D, Zhao B, Zhuang C (2016) Development of Commercial Thermo-sensitive Genic Male Sterile Rice Accelerates Hybrid Rice Breeding Using the CRISPR/Cas9-mediated TMS5 Editing System. Sci Rep 6:37395. https://doi.org/10.1038/srep37395

64. Enhanced Rice Blast Resistance by CRISPR/Cas9-Targeted Mutagenesis of the ERF Transcription Factor Gene OsERF922 | PLOS ONE. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0154027. Accessed 11 Nov 2022

65. Pyott DE, Sheehan E, Molnar A (2016) Engineering of CRISPR/Cas9-mediated potyvirus resistance in transgene-free Arabidopsis plants. Molecular Plant Pathology 17:1276–1288. https://doi.org/10.1111/mpp.12417

66. Baltes NJ, Hummel AW, Konecna E, Cegan R, Bruns AN, Bisaro DM, Voytas DF (2015) Conferring resistance to geminiviruses with the CRISPR–Cas prokaryotic immune system. Nature Plants 1:1–4. https://doi.org/10.1038/nplants.2015.145

67. Gutierrez C (1999) Geminivirus DNA replication. Cell Mol Life Sci 56:313–329. https://doi.org/10.1007/s000180050433

68. Ali Z, Abul-faraj A, Li L, Ghosh N, Piatek M, Mahjoub A, Aouida M, Piatek A, Baltes NJ, Voytas DF, Dinesh-Kumar S, Mahfouz MM (2015) Efficient Virus-Mediated Genome Editing in Plants Using the CRISPR/Cas9 System. Mol Plant 8:1288–1291. https://doi.org/10.1016/j.molp.2015.02.011

69. Ji X, Zhang H, Zhang Y, Wang Y, Gao C (2015) Establishing a CRISPR–Cas-like immune system conferring DNA virus resistance in plants. Nature Plants 1:1–4. https://doi.org/10.1038/nplants.2015.144

70. Zhang Y, Rajan R, Seifert HS, Mondragón A, Sontheimer EJ (2015) DNase H Activity of Neisseria meningitidis Cas9. Mol Cell 60:242–255. https://doi.org/10.1016/j.molcel.2015.09.020

71. Zhao C, Zhang Z, Xie S, Si T, Li Y, Zhu J-K (2016) Mutational Evidence for the Critical Role of CBF Transcription Factors in Cold Acclimation in Arabidopsis. Plant Physiol 171:2744–2759. https://doi.org/10.1104/pp.16.00533

72. Zhou X, Jacobs TB, Xue L-J, Harding SA, Tsai C-J (2015) Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial Populus reveals 4-coumarate:CoA ligase specificity and redundancy. New Phytol 208:298–301. https://doi.org/10.1111/nph.13470

73. Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu J-L, Gao C (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotechnol 31:686–688. https://doi.org/10.1038/nbt.2650

74. Zhang Y, Liang Z, Zong Y, Wang Y, Liu J, Chen K, Qiu J-L, Gao C (2016) Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA. Nat Commun 7:12617. https://doi.org/10.1038/ncomms12617

75. Li C, Unver T, Zhang B (2017) A high-efficiency CRISPR/Cas9 system for targeted mutagenesis in Cotton (Gossypium hirsutum L.). Sci Rep 7:43902. https://doi.org/10.1038/srep43902

76. Chen X, Lu X, Shu N, Wang S, Wang J, Wang D, Guo L, Ye W (2017) Targeted mutagenesis in cotton (Gossypium hirsutum L.) using the CRISPR/Cas9 system. Sci Rep 7:44304. https://doi.org/10.1038/srep44304

77. Morineau C, Bellec Y, Tellier F, Gissot L, Kelemen Z, Nogué F, Faure J-D (2017) Selective gene dosage by CRISPR-Cas9 genome editing in hexaploid *Camelina sativa*. Plant Biotechnology Journal 15:729–739. https://doi.org/10.1111/pbi.12671

78. Jiang WZ, Henry IM, Lynagh PG, Comai L, Cahoon EB, Weeks DP (2017) Significant enhancement of fatty acid composition in seeds of the allohexaploid, *Camelina sativa* , using CRISPR/Cas9 gene editing. Plant Biotechnology Journal 15:648–657. https://doi.org/10.1111/pbi.12663

79. Gao Y, Zhang Y, Zhang D, Dai X, Estelle M, Zhao Y (2015) Auxin binding protein 1 (ABP1) is not required for either auxin signaling or Arabidopsis development. Proceedings of the National Academy of Sciences 112:2275–2280. https://doi.org/10.1073/pnas.1500365112

80. Tang F, Yang S, Liu J, Zhu H (2016) Rj4, a Gene Controlling Nodulation Specificity in Soybeans, Encodes a Thaumatin-Like Protein But Not the One Previously Reported. Plant Physiol 170:26–32. https://doi.org/10.1104/pp.15.01661

81. Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M, Weigel D (2005) Specific effects of microRNAs on the plant transcriptome. Dev Cell 8:517–527. https://doi.org/10.1016/j.devcel.2005.01.018

82. Liu X, Hao L, Li D, Zhu L, Hu S (2015) Long non-coding RNAs and their biological roles in plants. Genomics Proteomics Bioinformatics 13:137–147. https://doi.org/10.1016/j.gpb.2015.02.003

83. Basak J, Nithin C (2015) Targeting Non-Coding RNAs in Plants with the CRISPR-Cas Technology is a Challenge yet Worth Accepting. Frontiers in Plant Science 6:. https://doi.org/10.3389/fpls.2015.01001

84. Zhou H, Liu B, Weeks DP, Spalding MH, Yang B (2014) Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. Nucleic Acids Research 42:10903–10914. https://doi.org/10.1093/nar/gku806

85. Zhao Y, Zhang C, Liu W, Gao W, Liu C, Song G, Li W-X, Mao L, Chen B, Xu Y, Li X, Xie C (2016) An alternative strategy for targeted gene replacement in plants using a dual-sgRNA/Cas9 design. Scientific Reports 6:. https://doi.org/10.1038/srep23890

86. Duan Y-B, Li J, Qin R-Y, Xu R-F, Li H, Yang Y-C, Ma H, Li L, Wei P-C, Yang J-B (2016) Identification of a regulatory element responsible for salt induction of rice OsRAV2 through ex situ and in situ promoter analysis. Plant Mol Biol 90:49–62. https://doi.org/10.1007/s11103-015-0393-z

87. Xing H-L, Dong L, Wang Z-P, Zhang H-Y, Han C-Y, Liu B, Wang X-C, Chen Q-J (2014) A CRISPR/Cas9 toolkit for multiplex genome editing in plants. BMC Plant Biology 14:. https://doi.org/10.1186/s12870-014-0327-y

88. Ma X, Zhang Q, Zhu Q, Liu W, Chen Y, Qiu R, Wang B, Yang Z, Li H, Lin Y, Xie Y, Shen R, Chen S, Wang Z, Chen Y, Guo J, Chen L, Zhao X, Dong Z, Liu Y-G (2015) A Robust CRISPR/Cas9 System for Convenient, High-Efficiency Multiplex Genome Editing in Monocot and Dicot Plants. Molecular Plant 8:1274–1284. https://doi.org/10.1016/j.molp.2015.04.007

89. Lowder LG, Zhang D, Baltes NJ, Paul JW, Tang X, Zheng X, Voytas DF, Hsieh T-F, Zhang Y, Qi Y (2015) A CRISPR/Cas9 Toolbox for Multiplexed Plant Genome Editing and Transcriptional Regulation. Plant Physiology 169:971–985. https://doi.org/10.1104/pp.15.00636

90. Zhang Z, Mao Y, Ha S, Liu W, Botella JR, Zhu J-K (2016) A multiplex CRISPR/Cas9 platform for fast and efficient editing of multiple genes in Arabidopsis. Plant Cell Rep 35:1519–1533. https://doi.org/10.1007/s00299-015-1900-z

91. Wang C, Shen L, Fu Y, Yan C, Wang K (2015) A Simple CRISPR/Cas9 System for Multiplex Genome Editing in Rice. J Genet Genomics 42:703–706. https://doi.org/10.1016/j.jgg.2015.09.011

92. Hao Y, Wang Q, Li J, Yang S, Ma L, Zheng Y, Peng W (2021) Double nicking by RNA-directed Cascade-nCas3 for high-efficiency large-scale genome engineering. bioRxiv 2021.07.12.451994. https://doi.org/10.1101/2021.07.12.451994

93. Fauser F, Schiml S, Puchta H (2014) Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. The Plant Journal 79:348–359. https://doi.org/10.1111/tpj.12554

94. 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. The Plant Journal 80:1139–1150. https://doi.org/10.1111/tpj.12704

95. Mikami M, Toki S, Endo M (2016) Precision Targeted Mutagenesis via Cas9 Paired Nickases in Rice. Plant Cell Physiol 57:1058–1068. https://doi.org/10.1093/pcp/pcw049

96. Tang X, Zheng X, Qi Y, Zhang D, Cheng Y, Tang A, Voytas DF, Zhang Y (2016) A Single Transcript CRISPR-Cas9 System for Efficient Genome Editing in Plants. Molecular Plant 9:1088–1091. https://doi.org/10.1016/j.molp.2016.05.001

97. Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, Goodwin MJ, Aryee MJ, Joung JK (2014) Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. Nat Biotechnol 32:569–576. https://doi.org/10.1038/nbt.2908

98. Guilinger JP, Thompson DB, Liu DR (2014) Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. Nat Biotechnol 32:577–582. https://doi.org/10.1038/nbt.2909

99. Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, Cost GJ, Zhang L, Santiago Y, Miller JC, Zeitler B, Cherone JM, Meng X, Hinkley SJ, Rebar EJ, Gregory PD, Urnov FD, Jaenisch R (2011) Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol 29:731–734. https://doi.org/10.1038/nbt.1927

100. Budhagatapalli N, Rutten T, Gurushidze M, Kumlehn J, Hensel G (2015) Targeted Modification of Gene Function Exploiting Homology-Directed Repair of TALEN-Mediated Double-Strand Breaks in Barley. G3 (Bethesda) 5:1857–1863. https://doi.org/10.1534/g3.115.018762

101. Li T, Liu B, Chen CY, Yang B (2016) TALEN-Mediated Homologous Recombination Produces Site-Directed DNA Base Change and Herbicide-Resistant Rice. J Genet Genomics 43:297–305. https://doi.org/10.1016/j.jgg.2016.03.005

102. High-frequency, precise modification of the tomato genome | Genome Biology | Full Text. https://genomebiology.biomedcentral.com/articles/10.1186/s13059-015-0796-9. Accessed 11 Nov 2022

103. Baltes NJ, Gil-Humanes J, Cermak T, Atkins PA, Voytas DF (2014) DNA Replicons for Plant Genome Engineering. The Plant Cell 26:151–163. https://doi.org/10.1105/tpc.113.119792

104. Li J-F, Norville JE, Aach J, McCormack M, Zhang D, Bush J, Church GM, Sheen J (2013) Multiplex and homologous recombination–mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. Nature Biotechnology 31:688–691. https://doi.org/10.1038/nbt.2654

105. Sun Y, Zhang X, Wu C, He Y, Ma Y, Hou H, Guo X, Du W, Zhao Y, Xia L (2016) Engineering herbicide-resistant rice plants through CRISPR/Cas9-mediated homologous recombination of Acetolactate Synthase. Molecular Plant 9:628–631. https://doi.org/10.1016/j.molp.2016.01.001

106. Endo M, Mikami M, Toki S (2016) Biallelic gene targeting in rice. Plant Physiol 170:667–677. https://doi.org/10.1104/pp.15.01663

107. Wang M, Lu Y, Botella JR, Mao Y, Hua K, Zhu J (2017) Gene targeting by homology-directed repair in rice using a geminivirus-based CRISPR/Cas9 system. Molecular Plant 10:1007–1010. https://doi.org/10.1016/j.molp.2017.03.002

108. Gil-Humanes J, Wang Y, Liang Z, Shan Q, Ozuna CV, Sánchez-León S, Baltes NJ, Starker C, Barro F, Gao C, Voytas DF (2017) High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. The Plant Journal 89:1251–1262. https://doi.org/10.1111/tpj.13446

109. Fauser F, Roth N, Pacher M, Ilg G, Sánchez-Fernández R, Biesgen C, Puchta H (2012) In planta gene targeting. PNAS 109:7535–7540. https://doi.org/10.1073/pnas.1202191109

110. Beerli RR, Segal DJ, Dreier B, Barbas CF (1998) Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. Proceedings of the National Academy of Sciences 95:14628–14633. https://doi.org/10.1073/pnas.95.25.14628

111. Hiratsu K, Matsui K, Koyama T, Ohme-Takagi M (2003) Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in Arabidopsis. The Plant Journal 34:733–739. https://doi.org/10.1046/j.1365-313X.2003.01759.x

112. Kay S, Hahn S, Marois E, Hause G, Bonas U (2007) A bacterial effector acts as a plant transcription factor and induces a cell size regulator. Science 318:648–651. https://doi.org/10.1126/science.1144956

113. Römer P, Hahn S, Jordan T, Strauss T, Bonas U, Lahaye T (2007) Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. Science 318:645–648. https://doi.org/10.1126/science.1144958

114. Mahfouz MM, Li L, Piatek M, Fang X, Mansour H, Bangarusamy DK, Zhu J-K (2012) Targeted transcriptional repression using a chimeric TALE-SRDX repressor protein. Plant Molecular Biology 78:311–321. https://doi.org/10.1007/s11103-011-9866-x

115. Lin S, Zhao Y, Zhu Y, Gosney M, Deng X, Wang X, Lin J (2016) An Effective and Inducible System of TAL Effector-Mediated Transcriptional Repression in Arabidopsis. Mol Plant 9:1546–1549. https://doi.org/10.1016/j.molp.2016.09.003

116. Piatek A, Ali Z, Baazim H, Li L, Abulfaraj A, Al-Shareef S, Aouida M, Mahfouz MM (2015) RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. Plant Biotechnol J 13:578–589. https://doi.org/10.1111/pbi.12284

117. Valton J, Dupuy A, Daboussi F, Thomas S, Maréchal A, Macmaster R, Melliand K, Juillerat A, Duchateau P (2012) Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. J Biol Chem 287:38427–38432. https://doi.org/10.1074/jbc.C112.408864

118. Esvelt KM, Mali P, Braff JL, Moosburner M, Yaung SJ, Church GM (2013) Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nature Methods 10:1116–1121. https://doi.org/10.1038/nmeth.2681

119. Steinert J, Schiml S, Fauser F, Puchta H (2015) Highly efficient heritable plant genome engineering using Cas9 orthologues from *Streptococcus thermophilus* and *Staphylococcus aureus*. The Plant Journal 84:1295–1305. https://doi.org/10.1111/tpj.13078

120. Kaya H, Mikami M, Endo A, Endo M, Toki S (2016) Highly specific targeted mutagenesis in plants using *Staphylococcus aureus* Cas9. Scientific Reports 6:26871. https://doi.org/10.1038/srep26871

121. Zhang H-Y, Wang X-H, Dong L, Wang Z-P, Liu B, Lv J, Xing H-L, Han C-Y, Wang X-C, Chen Q-J (2017) MISSA 2.0: an updated synthetic biology toolbox for assembly of orthogonal CRISPR/Cas systems. Sci Rep 7:41993. https://doi.org/10.1038/srep41993

122. Kaya H, Ishibashi K, Toki S (2017) A Split Staphylococcus aureus Cas9 as a Compact Genome-Editing Tool in Plants. Plant Cell Physiol 58:643–649. https://doi.org/10.1093/pcp/pcx034

123. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, Koonin EV, Zhang F (2015) Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. Cell 163:759–771. https://doi.org/10.1016/j.cell.2015.09.038

124. Xu R, Qin R, Li H, Li D, Li L, Wei P, Yang J (2017) Generation of targeted mutant rice using a CRISPR-Cpf1 system. Plant Biotechnology Journal 15:713–717. https://doi.org/10.1111/pbi.12669

125. Endo A, Masafumi M, Kaya H, Toki S (2016) Efficient targeted mutagenesis of rice and tobacco genomes using Cpf1 from Francisella novicida. Sci Rep 6:. https://doi.org/10.1038/srep38169

126. Tang X, Lowder LG, Zhang T, Malzahn AA, Zheng X, Voytas DF, Zhong Z, Chen Y, Ren Q, Li Q, Kirkland ER, Zhang Y, Qi Y (2017) A CRISPR–Cpf1 system for efficient genome editing and transcriptional repression in plants. Nature Plants 3:17018. https://doi.org/10.1038/nplants.2017.18

127. Kim H, Kim S-T, Ryu J, Kang B-C, Kim J-S, Kim S-G (2017) CRISPR/Cpf1-mediated DNA-free plant genome editing. Nature Communications 8:14406. https://doi.org/10.1038/ncomms14406

128. Wang M, Mao Y, Lu Y, Tao X, Zhu J (2017) Multiplex Gene Editing in Rice Using the CRISPR-Cpf1 System. Molecular Plant 10:1011–1013. https://doi.org/10.1016/j.molp.2017.03.001

129. Yin X, Biswal AK, Dionora J, Perdigon KM, Balahadia CP, Mazumdar S, Chater C, Lin H-C, Coe RA, Kretzschmar T, Gray JE, Quick PW, Bandyopadhyay A (2017) CRISPR-Cas9 and CRISPR-Cpf1 mediated targeting of a stomatal developmental gene *EPFL9* in rice. Plant Cell Rep 36:745–757. https://doi.org/10.1007/s00299-017-2118-z

130. Luo S, Li J, Stoddard TJ, Baltes NJ, Demorest ZL, Clasen BM, Coffman A, Retterath A, Mathis L, Voytas DF, Zhang F (2015) Non-transgenic Plant Genome Editing Using Purified Sequence-Specific Nucleases. Molecular Plant 8:1425–1427. https://doi.org/10.1016/j.molp.2015.05.012

131. DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins | Nature Biotechnology. https://www.nature.com/articles/nbt.3389/. Accessed 25 Sep 2018

132. Liang Z, Chen K, Li T, Zhang Y, Wang Y, Zhao Q, Liu J, Zhang H, Liu C, Ran Y, Gao C (2017) Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. Nature Communications 8:14261. https://doi.org/10.1038/ncomms14261

133. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 533:420–424. https://doi.org/10.1038/nature17946

134. Nishida K, Arazoe T, Yachie N, Banno S, Kakimoto M, Tabata M, Mochizuki M, Miyabe A, Araki M, Hara KY, Shimatani Z, Kondo A (2016) Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. Science 353:aaf8729. https://doi.org/10.1126/science.aaf8729

135. Li J, Sun Y, Du J, Zhao Y, Xia L (2017) Generation of targeted point mutations in rice by a modified CRISPR/Cas9 system. Molecular Plant 10:526–529. https://doi.org/10.1016/j.molp.2016.12.001

136. Lu Y, Zhu J-K (2017) Precise editing of a target base in the rice genome using a modified CRISPR/Cas9 system. Molecular Plant 10:523–525. https://doi.org/10.1016/j.molp.2016.11.013

137. Zong Y, Wang Y, Li C, Zhang R, Chen K, Ran Y, Qiu J-L, Wang D, Gao C (2017) Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. Nature Biotechnology 35:438–440. https://doi.org/10.1038/nbt.3811

138. Ren B, Yan F, Kuang Y, Li N, Zhang D, Lin H, Zhou H (2017) A CRISPR/Cas9 toolkit for efficient targeted base editing to induce genetic variations in rice. Science China Life Sciences 60:516–519. https://doi.org/10.1007/s11427-016-0406-x

139. Shimatani Z, Kashojiya S, Takayama M, Terada R, Arazoe T, Ishii H, Teramura H, Yamamoto T, Komatsu H, Miura K, Ezura H, Nishida K, Ariizumi T, Kondo A (2017) Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. Nature Biotechnology 35:441–443. https://doi.org/10.1038/nbt.3833

140. Chen Y, Wang Z, Ni H, Xu Y, Chen Q, Jiang L (2017) CRISPR/Cas9-mediated base-editing system efficiently generates gain-of-function mutations in Arabidopsis. Sci China Life Sci 60:520–523. https://doi.org/10.1007/s11427-017-9021-5

141. Lowe K, Wu E, Wang N, Hoerster G, Hastings C, Cho M-J, Scelonge C, Lenderts B, Chamberlin M, Cushatt J, Wang L, Ryan L, Khan T, Chow-Yiu J, Hua W, Yu M, Banh J, Bao Z, Brink K, Igo E, Rudrappa B, Shamseer PM, Bruce W, Newman L, Shen B, Zheng P, Bidney D, Falco C, Register J, Zhao Z-Y, Xu D, Jones T, Gordon-Kamm W (2016) Morphogenic Regulators Baby boom and Wuschel Improve Monocot Transformation. The Plant Cell 28:1998–2015. https://doi.org/10.1105/tpc.16.00124

142. Fu BXH, Smith JD, Fuchs RT, Mabuchi M, Curcuru J, Robb GB, Fire AZ (2019) Target-dependent nickase activities of the CRISPR–Cas nucleases Cpf1 and Cas9. Nature Microbiology 4:888–897. https://doi.org/10.1038/s41564-019-0382-0

143. Gao L, Cox DBT, Yan WX, Manteiga JC, Schneider MW, Yamano T, Nishimasu H, Nureki O, Crosetto N, Zhang F (2017) Engineered Cpf1 variants with altered PAM specificities. Nat Biotechnol 35:789–792. https://doi.org/10.1038/nbt.3900

144. Gurel F, Zhang Y, Sretenovic S, Qi Y (2020) CRISPR-Cas nucleases and base editors for plant genome editing | SpringerLink. aBIOTECH 1:74–87

145. Hua K, Tao X, Han P, Wang R, Zhu J-K (2019) Genome engineering in rice using Cas9 variants that recognize NG PAM sequences. Molecular Plant. https://doi.org/10.1016/j.molp.2019.03.009

146. Nelson JW, Randolph PB, Shen SP, Everette KA, Chen PJ, Anzalone AV, An M, Newby GA, Chen JC, Hsu A, Liu DR (2021) Engineered pegRNAs improve prime editing efficiency. Nat Biotechnol. https://doi.org/10.1038/s41587-021-01039-7

147. Sretenovic S, Yin D, Levav A, Selengut JD, Mount SM, Qi Y (2020) Expanding Plant Genome Editing Scope by An Engineered iSpyMacCas9 System Targeting the A-rich PAM Sequences. Plant Comm 0: https://doi.org/10.1016/j.xplc.2020.100101

148. Clements TP, Tandon B, Lintel HA, McCarty JH, Wagner DS (2017) RICE CRISPR: Rapidly increased cut ends by an exonuclease Cas9 fusion in zebrafish. genesis 55:e23044. https://doi.org/10.1002/dvg.23044

149. Lainšček D, Forstnerič V, Mikolič V, Malenšek Š, Pečan P, Benčina M, Sever M, Podgornik H, Jerala R (2022) Coiled-coil heterodimer-based recruitment of an exonuclease to CRISPR/Cas for enhanced gene editing. Nat Commun 13:3604. https://doi.org/10.1038/s41467-022-31386-1

150. Liu G, Yin K, Zhang Q, Gao C, Qiu J-L (2019) Modulating chromatin accessibility by transactivation and targeting proximal dsgRNAs enhances Cas9

editing efficiency in vivo. Genome Biol 20:145. https://doi.org/10.1186/s13059-019-1762-8

151. 5'-exonuclease Increases Gene Editing Efficiency of Plants - 20160020 - University of Minnesota Office for Technology Commercialization

152. Mu W, Zhang Y, Xue X, Liu L, Wei X, Wang H (2019) 5′ capped and 3′ polyA-tailed sgRNAs enhance the efficiency of CRISPR-Cas9 system. Protein & Cell 10:223–228. https://doi.org/10.1007/s13238-018-0552-5

153. Zhou J, Yang Y, Wang X, Yu F, Yu C, Chen J, Cheng Y, Yan C, Chen J (2013) Enhanced transgene expression in rice following selection controlled by weak promoters. BMC Biotechnology 13:29. https://doi.org/10.1186/1472-6750-13-29

154. Tang Z, Chen S, Chen A, He B, Zhou Y, Chai G, Guo F, Huang J (2019) CasPDB: an integrated and annotated database for Cas proteins from bacteria and archaea. Database (Oxford) 2019:baz093. https://doi.org/10.1093/database/baz093

155. Zhong Z, Zhang Y, You Q, Tang X, Ren Q, Liu S, Yang L, Wang Y, Liu X, Liu B, Zhang T, Zheng X, Le Y, Zhang Y, Qi Y (2018) Plant Genome Editing Using FnCpf1 and LbCpf1 Nucleases at Redefined and Altered PAM Sites. Molecular Plant 11:999–1002. https://doi.org/10.1016/j.molp.2018.03.008

156. Malzahn AA, Tang X, Lee K, Ren Q, Sretenovic S, Zhang Y, Chen H, Kang M, Bao Y, Zheng X, Deng K, Zhang T, Salcedo V, Wang K, Zhang Y, Qi Y (2019) Application of CRISPR-Cas12a temperature sensitivity for improved genome editing in rice, maize, and Arabidopsis. BMC Biology 17:9. https://doi.org/10.1186/s12915-019-0629-5

157. Addgene (2020) Yiping Qi Lab Plasmids. https://www.addgene.org/Yiping_Qi/

158. Zhang Y, Zhang Y, Qi Y (2019) Plant Gene Knockout and Knockdown by CRISPR-Cpf1 (Cas12a) Systems. In: Qi Y (ed) Plant Genome Editing with CRISPR Systems: Methods and Protocols. Springer, New York, NY, pp 245–256

159. Image Lab Software. In: Bio-Rad Laboratories. https://www.bio-rad.com/en-us/product/image-lab-software?ID=KRE6P5E8Z. Accessed 8 Jun 2022

160. Liu P, Luk K, Shin M, Idrizi F, Kwok S, Roscoe B, Mintzer E, Suresh S, Morrison K, Frazão JB, Bolukbasi MF, Ponnienselvan K, Luban J, Zhu LJ, Lawson ND, Wolfe SA (2019) Enhanced Cas12a editing in mammalian cells and zebrafish. Nucleic Acids Res 47:4169–4180. https://doi.org/10.1093/nar/gkz184

161. Chen Y, Zeng S, Hu R, Wang X, Huang W, Liu J, Wang L, Liu G, Cao Y, Zhang Y (2017) Using local chromatin structure to improve CRISPR/Cas9 efficiency in zebrafish. PLOS ONE 12:e0182528. https://doi.org/10.1371/journal.pone.0182528

162. Shan Q, Baltes NJ, Atkins P, Kirkland ER, Zhang Y, Baller JA, Lowder LG, Malzahn AA, Haugner JC, Seelig B, Voytas DF, Qi Y (2018) ZFN, TALEN and CRISPR-Cas9 mediated homology directed gene insertion in Arabidopsis: a disconnect between somatic and germinal cells. Journal of Genetics and Genomics. https://doi.org/10.1016/j.jgg.2018.07.011

163. Leenay RT, Maksimchuk KR, Slotkowski RA, Agrawal RN, Gomaa AA, Briner AE, Barrangou R, Beisel CL (2016) Identifying and Visualizing Functional PAM Diversity across CRISPR-Cas Systems. Molecular Cell 62:137–147. https://doi.org/10.1016/j.molcel.2016.02.031

164. Zhang Y, Ren Q, Tang X, Liu S, Malzahn AA, Zhou J, Wang J, Yin D, Pan C, Yuan M, Huang L, Yang H, Zhao Y, Fang Q, Zheng X, Tian L, Cheng Y, Le Y, McCoy B, Franklin L, Selengut JD, Mount SM, Que Q, Zhang Y, Qi Y (2021) Expanding the scope of plant genome engineering with Cas12a orthologs and highly multiplexable editing systems. Nat Commun 12:1944. https://doi.org/10.1038/s41467-021-22330-w

165. Rogers EE, Ausubel FM (1997) Arabidopsis enhanced disease susceptibility mutants exhibit enhanced susceptibility to several bacterial pathogens and alterations in PR-1 gene expression. Plant Cell 9:305–316. https://doi.org/10.1105/tpc.9.3.305

166. Wang Z-P, Xing H-L, Dong L, Zhang H-Y, Han C-Y, Wang X-C, Chen Q-J (2015) Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. Genome Biology 16:. https://doi.org/10.1186/s13059-015-0715-0

167. Tang X, Ren Q, Yang L, Bao Y, Zhong Z, He Y, Liu S, Qi C, Liu B, Wang Y, Sretenovic S, Zhang Y, Zheng X, Zhang T, Qi Y, Zhang Y (2018) Single transcript unit CRISPR 2.0 systems for robust Cas9 and Cas12a mediated plant genome editing. Plant Biotechnology Journal 0: https://doi.org/10.1111/pbi.13068

168. Zhang Y, Zhang Y, Qi Y (2019) Plant Gene Knockout and Knockdown by CRISPR-Cpf1 (Cas12a) Systems. In: Qi Y (ed) Plant Genome Editing with CRISPR Systems: Methods and Protocols. Springer, New York, NY, pp 245–256

169. Lowder LG, Malzahn A, Qi Y (2018) Plant Gene Regulation Using Multiplex CRISPR-dCas9 Artificial Transcription Factors. In: Maize: methods and protocols. Springer Science+Business Media, New York, NY, pp 197–214

170. Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana: Floral dip transformation of Arabidopsis. The Plant Journal 16:735–743. https://doi.org/10.1046/j.1365-313x.1998.00343.x

171. Zhang Y, Malzahn AA, Sretenovic S, Qi Y (2019) The emerging and uncultivated potential of CRISPR technology in plant science. Nat Plants 5:778–794. https://doi.org/10.1038/s41477-019-0461-5

172. Piatek A, Mahfouz MM (2017) Targeted genome regulation via synthetic programmable transcriptional regulators. Critical Reviews in Biotechnology 37:429–440. https://doi.org/10.3109/07388551.2016.1165180

173. Schindele P, Wolter F, Puchta H Transforming plant biology and breeding with CRISPR/Cas9, Cas12 and Cas13. FEBS Letters 592:1954–1967. https://doi.org/10.1002/1873-3468.13073

174. Ding X, Seebeck T, Feng Y, Jiang Y, Davis GD, Chen F (2019) Improving CRISPR-Cas9 efficiency with chromatin modulating peptides. The CRISPR Journal 2:

175. Steinert J, Schiml S, Puchta H (2016) Homology-based double-strand break-induced genome engineering in plants. Plant Cell Rep 35:1429–1438. https://doi.org/10.1007/s00299-016-1981-3

176. Chen JS, Ma E, Harrington LB, Costa MD, Tian X, Palefsky JM, Doudna JA (2018) CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. Science 360:436–439. https://doi.org/10.1126/science.aar6245

177. Ordon J, Gantner J, Kemna J, Schwalgun L, Reschke M, Streubel J, Boch J, Stuttmann J (2017) Generation of chromosomal deletions in dicotyledonous plants employing a user-friendly genome editing toolkit. The Plant Journal 89:155–168. https://doi.org/10.1111/tpj.13319

178. Wu R, Lucke M, Jang Y, Zhu W, Symeonidi E, Wang C, Fitz J, Xi W, Schwab R, Weigel D (2018) An efficient CRISPR vector toolbox for engineering large deletions in Arabidopsis thaliana. Plant Methods 14:. https://doi.org/10.1186/s13007-018-0330-7

179. Pan C, Li G, Malzahn AA, Cheng Y, Leyson B, Sretenovic S, Gurel F, Coleman GD, Qi Y (2022) Boosting plant genome editing with a versatile CRISPR-Combo system. Nat Plants 8:513–525. https://doi.org/10.1038/s41477-022-01151-9

180. Yamaguchi K, Yamamoto T, Segami S, Horikawa M, Chaya G, Kitano H, Iwasaki Y, Miura K (2020) gw2 mutation increases grain width and culm

thickness in rice (Oryza sativa L.). Breed Sci 70:456–461. https://doi.org/10.1270/jsbbs.20018

181. Clement K, Rees H, Canver MC, Gehrke JM, Farouni R, Hsu JY, Cole MA, Liu DR, Joung JK, Bauer DE, Pinello L (2019) CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nature Biotechnology 37:224–226. https://doi.org/10.1038/s41587-019-0032-3

182. Sakamoto W, Takami T (2014) Nucleases in higher plants and their possible involvement in DNA degradation during leaf senescence. Journal of Experimental Botany 65:3835–3843. https://doi.org/10.1093/jxb/eru091

183. Takami T, Ohnishi N, Kurita Y, Iwamura S, Ohnishi M, Kusaba M, Mimura T, Sakamoto W (2018) Organelle DNA degradation contributes to the efficient use of phosphate in seed plants. Nature Plants 4:1044–1055. https://doi.org/10.1038/s41477-018-0291-x

184. Tang LY, Sakamoto W (2011) Tissue-specific organelle DNA degradation mediated by DPD1 exonuclease. Plant Signal Behav 6:1391–1393. https://doi.org/10.4161/psb.6.9.16595

185. Cheng H-L, Lin C-T, Huang K-W, Wang S, Lin Y-T, Toh S-I, Hsiao Y-Y (2018) Structural insights into the duplex DNA processing of TREX2. Nucleic Acids Res 46:12166–12176. https://doi.org/10.1093/nar/gky970

186. Phillips GJ, Prasher DC, Kushner SR (1988) Physical and biochemical characterization of cloned sbcB and xonA mutations from Escherichia coli K-12. J Bacteriol; (United States) 170:5:

187. Mitsunobu H, Zhu B, Lee S-J, Tabor S, Richardson CC (2014) Flap Endonuclease Activity of Gene 6 Exonuclease of Bacteriophage T7. J Biol Chem 289:5860–5875. https://doi.org/10.1074/jbc.M113.538611

188. Pickering TJ, Garforth SJ, Thorpe SJ, Sayers JR, Grasby JA (1999) A single cleavage assay for T5 5'-->3' exonuclease: determination of the catalytic parameters forwild-type and mutant proteins. Nucleic Acids Res 27:730–735. https://doi.org/10.1093/nar/27.3.730

189. Jensen KT, Fløe L, Petersen TS, Huang J, Xu F, Bolund L, Luo Y, Lin L (2017) Chromatin accessibility and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency. FEBS Letters 591:1892–1901. https://doi.org/10.1002/1873-3468.12707

190. Mashimo T, Kaneko T, Sakuma T, Kobayashi J, Kunihiro Y, Voigt B, Yamamoto T, Serikawa T (2013) Efficient gene targeting by TAL effector nucleases coinjected with exonucleases in zygotes. Scientific Reports 3:1253. https://doi.org/10.1038/srep01253

191. Park J, Yoon J, Kwon D, Han M-J, Choi S, Park S, Lee J, Lee K, Lee J, Lee S, Kang K-S, Choe S (2021) Enhanced genome editing efficiency of CRISPR PLUS: Cas9 chimeric fusion proteins. Sci Rep 11:16199. https://doi.org/10.1038/s41598-021-95406-8

192. Chen S, Tao L, Zeng L, Vega-Sanchez ME, Umemura K, Wang G-L (2006) A highly efficient transient protoplast system for analyzing defence gene expression and protein–protein interactions in rice. Molecular Plant Pathology 7:417–427. https://doi.org/10.1111/j.1364-3703.2006.00346.x

193. Main M, Frame B, Wang K (2015) Rice, Japonica (Oryza sativa L.). In: Wang K (ed) Agrobacterium Protocols: Volume 1. Springer, New York, NY, pp 169–180

194. Liu Q, Wang C, Jiao X, Zhang H, Song L, Li Y, Gao C, Wang K (2019) Hi-TOM: a platform for high-throughput tracking of mutations induced by CRISPR/Cas systems. Sci China Life Sci 62:1–7. https://doi.org/10.1007/s11427-018-9402-9

195. Kleinstiver BP, Sousa AA, Walton RT, Tak YE, Hsu JY, Clement K, Welch MM, Horng JE, Malagon-Lopez J, Scarfò I, Maus MV, Pinello L, Aryee MJ, Joung JK (2019) Engineered CRISPR-Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. Nat Biotechnol 37:276–282. https://doi.org/10.1038/s41587-018-0011-0

196. Schindele P, Puchta H (2020) Engineering CRISPR/LbCas12a for highly efficient, temperature-tolerant plant gene editing. Plant Biotechnol J 18:1118–1120. https://doi.org/10.1111/pbi.13275

197. Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, Chen PJ, Wilson C, Newby GA, Raguram A, Liu DR (2019) Search-and-replace genome editing without double-strand breaks or donor DNA. Nature 1–1. https://doi.org/10.1038/s41586-019-1711-4

198. Revisions to USDA biotechnology regulations: The SECURE rule. https://www.pnas.org/doi/10.1073/pnas.2004841118. Accessed 31 May 2022

199. USDA APHIS | About the SECURE Rule. https://www.aphis.usda.gov/aphis/ourfocus/biotechnology/biotech-rule-revision. Accessed 25 Feb 2021

200. gene-editing (2019) European Union: Crops / Food. In: Global Gene Editing Regulation Tracker. https://crispr-gene-editing-regs-tracker.geneticliteracyproject.org/european-union-crops-food/. Accessed 9 Nov 2022