Plant Biotechnology Journal
Open Access
aab Association of Applied Biologists
SEB

Plant Biotechnology Journal (2023), pp. 1–12 doi: 10.1111/pbi.14194

Optimizing *Er*Cas12a for efficient gene editing in *Arabidopsis thaliana*

Janine Pietralla, Niklas Capdeville, Patrick Schindele and Holger Puchta*

Karlsruhe Institute of Technology (KIT), Joseph Gottlieb Kölreuter Institute for Plant Sciences (JKIP), Department of Molecular Biology, Karlrsruhe, Germany

Received 26 May 2023;
revised 24 July 2023;
accepted 23 September 2023.
**Correspondence* (Tel +49 721 608-48894; fax +49 721 608-44874; email holger.puchta@kit.edu)

Keywords: CRISPR/Cas, Cas12, MAD7, gene editing, genome engineering.

Summary

The *Er*Cas12a nuclease, also known as MAD7, is part of a CRISPR/Cas system from *Eubacterium rectale* and distantly related to Cas12a nucleases. As it shares only 31% sequence homology with the commonly used *As*Cas12a, its intellectual property may not be covered by the granted patent rights for Cas12a nucleases. Thus, *Er*Cas12a became an attractive alternative for practical applications. However, the editing efficiency of *Er*Cas12a is strongly target sequence- and temperature-dependent. Therefore, optimization of the enzyme activity through protein engineering is especially attractive for its application in plants, as they are cultivated at lower temperatures. Based on the knowledge obtained from the optimization of Cas12a nucleases, we opted to improve the gene editing efficiency of *Er*Cas12a by introducing analogous amino acid exchanges. Interestingly, neither of these mutations analogous to those in the enhanced or Ultra versions of *As*Cas12a resulted in significant editing enhancement of *Er*Cas12a in *Arabidopsis thaliana*. However, two different mutations, V156R and K172R, in putative alpha helical structures of the enzyme showed a detectable improvement in editing. By combining these two mutations, we obtained an improved *Er*Cas12a (im*Er*Cas12a) variant, showing several-fold increase in activity in comparison to the wild-type enzyme in Arabidopsis. This variant yields strong editing efficiencies at 22 °C which could be further increased by raising the cultivation temperature to 28 °C and even enabled editing of formerly inaccessible targets. Additionally, no enhanced off-site activity was detected. Thus, im*Er*Cas12a is an economically attractive and efficient alternative to other CRISPR/Cas systems for plant genome engineering.

Introduction

The importance of genome engineering in the establishment of new crop varieties is increasing due to climate change and population growth. It allows engineered crops to respond quickly and specifically to the new conditions in order to continuously provide the population with sufficient high-quality food. In many laboratories involved in plant research, genome modification using the RNA-guided target sequence-specific CRISPR (clustered regularly interspaced short palindromic repeats)/Cas (CRISPR associated) nucleases has become the method of choice. Shortly after the first successful CRISPR/Cas-based genome manipulations in eukaryotic cells had been conducted (Cong *et al*., 2013; Mali *et al*., 2013), plant cells were edited for the first time (Li *et al*., 2013; Shan *et al*., 2013). Eventually, CRISPR/Cas systems of all types were adapted for the application in plants (Wada *et al*., 2022). Modification of specific traits was achieved in major food crops, including rice, maize and wheat as well as other important crops, such as tomato, potato and soybean. These have already been described in several reviews, for example Kumlehn *et al*., 2018; Scheben *et al*., 2017; Schindele *et al*., 2018. Besides classical mutagenesis, a wide variety of tools, based on modified Cas enzymes, was developed and enables the induction of predefined mutations (Capdeville *et al*., 2021; Zhu *et al*., 2020). For example, mutations could be introduced which either increased yield or resistance to negative biotic and abiotic influences (Bandyopadhyay *et al*., 2020). This is of high interest for breeders especially in the case of elite crops which are characterized by a limited gene pool due to centuries of inbreeding and, as a consequence of phenotypic selection for preferred traits, have often lost their natural resistance to biotic and abiotic stresses. One approach to overcome this issue is to use CRISPR/Cas to combine the preferred traits of cultivars with the beneficial traits of wild forms (Østerberg *et al*., 2017; Schindele *et al*., 2020).

One decisive factor for the efficient use of CRISPR/Cas in plant breeding is the editing efficiency of the nucleases: the efficiency of the various CRISPR/Cas systems differs strongly among different organisms. Particularly, the Cas12a orthologues show strong variations in their editing efficiency when used in different organisms. On the one hand, these differences might be due to the chosen target sequences. Depending on their GC content or chromatin structure, the enzymes might bind less efficiently to the target DNA sequence and, thus, induce breaks at a lower frequency. On the other hand, direct comparison of orthologues showed differences in efficiencies at identical target sequences (Malzahn *et al*., 2019; Moreno-Mateos *et al*., 2017; Zetsche *et al*., 2015; Zhang *et al*., 2021b). In human cell lines, *As*Cas12a and *Lb*Cas12a showed similarly high editing efficiencies (Zetsche *et al*., 2015), whereas in plants, *As*Cas12a initially showed a very low mutation rate (Bernabé-Orts *et al*., 2019). Upon temperature elevation, editing efficiencies increased significantly in case of the

Please cite this article as: Pietralla, J., Capdeville, N., Schindele, P. and Puchta, H. (2023) Optimizing *Er*Cas12a for efficient gene editing in *Arabidopsis thaliana*. *Plant Biotechnol. J*., https://doi.org/10.1111/pbi.14194.

tested Cas12a orthologues. In contrast to *As*Cas12a, *Lb*Cas12a showed solid editing efficiencies in protoplasts, even at lower temperatures. In Arabidopsis plants, this was the case only upon temperature increase. *Sp*Cas9, which was tested for comparison, has a temperature optimum of 32 °C in plants whereas the Cas12a orthologues *As*Cas12a, *Lb*Cas12a and *Francisella novicida* Cas12a achieved the highest mutagenesis rates at 28 °C in plants (Malzahn *et al*., 2019). Below 28 °C, however, Cas12a nucleases showed a strongly reduced activity. The ongoing search for new nucleases has yielded a number of new Cas12a systems for gene editing in plants, such as Mb2Cas12a, which showed a tolerance to lower temperatures in plants (Zhang *et al*., 2021b). Possible approaches to circumvent this temperature sensitivity are the introduction of mutations within the Cas12a protein sequence. By replacing certain amino acids that interact with the target DNA in the PAM-proximal region, it was possible to generate *As*Cas12a nucleases that exhibited lower temperature sensitivity and, accordingly, showed higher activity in human cell lines at lower temperatures (Kleinstiver *et al*., 2019). Analogous to these enhanced *As*Cas12a (en*As*Cas12a) variants, the corresponding amino acid modifications were also inserted in *Lb*Cas12a, which was already exhibiting a higher editing efficiency in plants (Schindele and Puchta, 2020). The temperature-tolerant *Lb*Cas12a (tt*Lb*Cas12a) variant created in this way showed high editing activity in Arabidopsis, even at low temperatures. This effect could be further enhanced at higher temperatures (Schindele and Puchta, 2020). Furthermore, a variant, called *As*Cas12a-Ultra, could be generated which not only showed increased activity at targets with canonical 5′-TTTV-3′ PAM, but additionally enabled efficient editing of targets with a 5′-TTTT-3′ PAM (Zhang *et al*., 2021a).

In addition to the CRISPR/Cas systems commonly used to date, a relatively unknown CRISPR/Cas system is of particular interest for gene editing-based plant breeding. *Er*Cas12a is a class 2 type V CRISPR/Cas system from *Eubacterium rectale*, and originates from a bacterial strain isolated in Madagascar which, presumably, shows a greater protein sequence divergence from its closest relatives due to long-term separate evolution. Thus, *Er*Cas12a shows only a 31% sequence homology with the commonly used *As*Cas12a. Despite the notable differences from previous Cas12a systems, *Er*Cas12a has already been successfully applied in mammalian cells, zebrafish, bacteria and also in plants (Lin *et al*., 2021; Liu *et al*., 2020; Price *et al*., 2020; Wierson *et al*., 2019; Zhang *et al*., 2021b). *Er*Cas12a showed stable editing in protoplasts, which could be improved upon temperature elevation (Zhang *et al*., 2021b) The required PAM sequence was initially designated as 5′-YTTN-3′. However, studies showed that PAMs with 5′-TTTN-3′ result in a higher editing efficiency (Lin *et al*., 2021). In this, they share a commonality with the better-known Cas12a nucleases which also require a T-rich PAM sequence.

Based on the knowledge obtained from the optimization of other Cas12a nucleases (Kleinstiver *et al*., 2019; Schindele and Puchta, 2020), we opted to improve the gene editing efficiency of *Er*Cas12a at lower temperatures and, thus, to improve its usability in plants.

Results

Similarly to the already well-characterized Cas12a orthologues, *Er*Cas12a recognizes T-rich PAM sequences. It shares the highest similarity with *As*Cas12a and uses an identical crRNA (Maksimova *et al*., 2019; Zetsche *et al*., 2015). In numerous organisms, including mammalian cells, zebrafish as well as rice and wheat, *Er*Cas12a was shown to provide robust cleavage activity (Lin *et al*., 2021; Liu *et al*., 2020; Wierson *et al*., 2019; Zhang *et al*., 2021b). However, a strong target sequence- and temperature-dependent variance of the efficiency was observed. Partially, this could be overcome by increasing the cultivation temperature or by applying brief heat shocks. However, not all plants can be grown under such conditions. Thus, the engineering of nucleases to achieve stable editing efficiencies is of utmost interest for breeders.

Since the editing efficiency of *Er*Cas12a has not yet been tested in *Arabidopsis thaliana*, we selected five different target sites in the *ECA3* gene for which extensive data are already available following mutagenesis with different *Lb*Cas12a variants (Schindele and Puchta, 2020). The target sites in the *ECA3* gene were accompanied by the PAM sequences 5′-TTTA-3′ and 5′-TTTC-3′ (Figure 1a). Transgenic plants were cultivated for 2 weeks on appropriate selection medium at either 22 °C or 28 °C. Editing efficiency analysis was performed using quantitative TIDE analysis (Brinkman *et al*., 2014) after sequencing of the affected gene segment (Figure 1c,d).

At 22 °C, at Target 2, Target 3 and Target 4, efficiencies of 10.3%, 9.3% and 13.0% were measured on average across 40 plants. At Target 1 and Target 5, not a single plant showed significant editing. Increasing the temperature to 28 °C resulted in editing of Target 1 and 5 with an efficiency of 13.1% and 8.6%, respectively. The three targets that already showed editing at 22 °C showed an editing increase by 1.6-, 2.1- and 1.4-fold at 28 °C, respectively. When the data of individual plants are considered, it is noticeable that a majority of the individual plants were not edited, especially when cultivated at 22 °C. In fact, the average values were raised by individual plants whose editing efficiencies reached up to 86%. At 28 °C, such high values occurred more frequently at the population level, but some also showed efficiencies in the average range. Strikingly, Target 1 and Target 5 demonstrated stable resistance to editing by *Er*Cas12a at 22 °C, whereas a slight increase of the cultivation temperature to 28 °C allowed at least minor editing. Regarding the distribution of the individual values in the boxplots, it appears that, despite the low average efficiency, especially at 28 °C, in some individual plants an editing efficiency of more than 50% could be achieved. In a few individual plants, an editing efficiency of more than 75% could be detected. By and large, however, our results demonstrate that the average efficiency obtained with *Er*Cas12a is rather low in Arabidopsis.

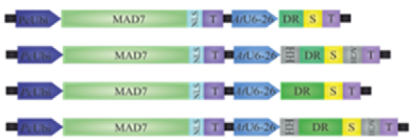
Evaluation of the most efficient crRNA set-up

Two different versions of a compatible crRNA length have been described for *Er*Cas12a: A longer version of the pre-crRNA with a length of 56 nt, which splits into a 35 nt direct repeat and a 21 nt spacer, or a shorter, mature version of the crRNA, consisting of 42 nt, starting with a 21 nt direct repeat, followed by a 21 nt spacer sequence. In human cells, both versions of the crRNA were tested and did not reveal significant differences in editing efficiency (Maksimova *et al*., 2019). A possible disadvantage of the 56 nt version could be the presence of five consecutive thymidine bases in the DNA sequence encoding the crRNA, which represent a possible termination signal for polymerase III. This could result in the generation of truncated and incomplete crRNAs which would consequently have a negative impact on crRNA efficiency. Furthermore, according to

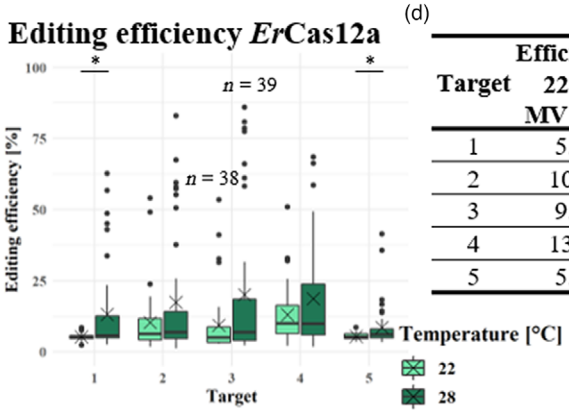
(a)

Target	PAM	Sequence
1	**TTTC**	CTAGCTTATCGTGATAGATCT
2	**TTTA**	AGCAATTATACTATCAGATAT
3	**TTTA**	ATTGATGCAGGAACGCCATTC
4	**TTTA**	CTATTCTATCACAGGTCATTC
5	**TTTC**	TGATTCTCAGGTATGAACCTG

(b)



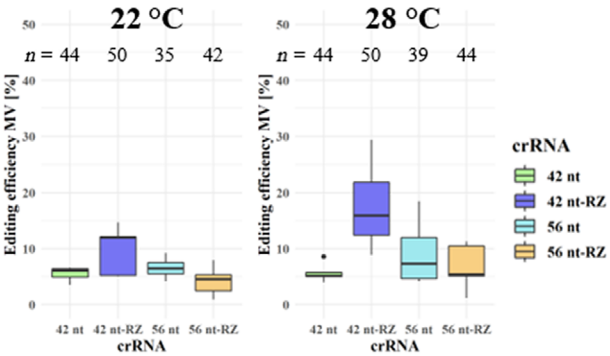
(c)



(d)

Target	Efficiency 22 °C MV [%]	Efficiency 28 °C MV [%]
1	5.2	13.1
2	10.3	17.4
3	9.3	20.1
4	13.0	18.6
5	5.4	8.6

(e)



(f)

crRNA	Efficiency 22°C MV [%]	Efficiency 28°C MV [%]
42 nt	5.5	5.7
42 nt-RZ	9.8	17.7
56 nt	6.5	9.3
56 nt-RZ	4.2	6.7

Figure 1 Gene editing in plants with *Er*Cas12a. (a) Overview of the five target sites in the ECA3 gene. (b) Overview of the expression cassettes used to evaluate the four different approaches to crRNA construction. Nuclease expression was controlled by the constitutive promoter PcUbi4-2 and the pea3A terminator. Expression of crRNAs was controlled by the polymerase III promoter AtU6-26 and a polyT sequence as a termination signal. The crRNAs were composed of either 21 nt direct repeats (DR) and 21 nt spacer sequence (S) or 35 nt direct repeat and 21 nt spacer sequence. Both versions, the 42 nt-long and the 56 nt-long were tested in separate approaches, either with flanking ribozymes (HH and HDV) or without ribozyme flanking. (c) Boxplots show the editing efficiencies in transgenic T1 plants at the tested target sites (1–5) at temperatures of 22 °C and 28 °C. The boxes represent the upper and lower quartiles, and the median is depicted as a solid black line in the boxes. In addition, the mean value of the data is shown as a cross. The values that are still within 1.5 times the interquartile range are indicated by the antennas. Outliers that fall outside this value are represented by black dots. $n = 40$ if not specified otherwise, *P*-values were calculated using the Mann–Whitney-*U* test: $*P < 0.05$. (d) Summary of relative editing efficiencies. The average values of about 40 plants which were tested for each target and the two different temperatures are shown. (e) Boxplots of the editing efficiency of *Er*Cas12a in combination with differently modified crRNA constructs at 22 °C and 28 °C. (f) Overview of relative editing efficiencies, given as average values of ten tested plants per target and for each crRNA set-up at both different temperatures.

sequence comparisons, the crRNA in the 56 nt version shows a completely identical sequence to the non-processed crRNA of *As*Cas12a. However, the *As*Cas12a is standardly used with the 42 nt-long crRNA (Zetsche *et al.*, 2015). Therefore, we tested both versions in Arabidopsis. In addition, to enhance the activity of other Cas12a orthologues, crRNAs are often flanked by the self-cleaving hammerhead (HH) ribozyme and the ribozyme from the hepatitis-delta-virus (HDV) to support crRNA processing (Gao *et al.*, 2018; Tang *et al.*, 2017). In case of the non-processed crRNA of *As*Cas12a, which has an identical sequence to the 56 nt crRNA, flanking with ribozymes led to an increase in crRNA expression and, thus, enhanced activity of *As*Cas12a (Gao *et al.*, 2018). Therefore, approaches using both crRNA versions, 56 nt or 42 nt, flanked with ribozymes, were included in our test. The structure of the different constructs is schematically shown in Figure 1b.

To determine which crRNA set-up was the most efficient for editing using *Er*Cas12a, four constructs were created which differed in crRNA construction but were otherwise identical. As described above, the crRNAs consisted of either 42 nt or 56 nt and both versions were also tested flanked by ribozymes in the HH-crRNA-HDV structure. The boxplots in Figure 1 show the distribution of the individual values of each crRNA, summarizing the values at the five different targets (Figure 1e,f).

The average values show that at both temperatures tested, the strongest editing efficiency for the individual target sequences was achieved using the short crRNA flanked by the two ribozymes HH and HDV, which is why we used this constellation for all further analyses in this study.

Protein engineering of *Er*Cas12a for efficient editing in plants

It has been previously shown in other systems that an increase in temperature is crucial for better editing efficiency by Cas12a orthologues (Moreno-Mateos *et al.*, 2017; Zhang *et al.*, 2021b). At lower temperatures, especially below 28 °C, Cas12a nucleases show a significant reduction in their activities. Based on the modifications in the en*As*Cas12a (Kleinstiver *et al.*, 2019) or

Table 1 Conservation level of amino acids in Cas12a orthologues

	Temperature-tolerant	Enhanced		Ultra	
AsCas12a	E174	S542	K548	M537	F870
LbCas12a	D156	G532	K538	N527	E795
*Er*Cas12a	K172	D532	K538	I527	F843
Conservation	Minor	Minor	Completely	Mostly	Completely

The table shows the amino acids substituted by arginine for the temperature-tolerant and enhanced variants of *As*Cas12a and *Lb*Cas12a, respectively, as well as the amino acids that were changed in the Ultra variant. The homologous amino acids in *Er*Cas12a are marked in bold and the conservation level is indicated depending on their biochemical properties (minor, mostly or completely).

tt*Lb*Cas12a variants (Schindele and Puchta, 2020), which mediated temperature tolerance, we searched for the corresponding amino acids in *Er*Cas12a. These amino acid substitutions supposedly lead to an increased efficiency due to a stronger interaction between the nuclease, the crRNA and the target DNA (Kleinstiver *et al*., 2019; Schindele and Puchta, 2020). However, *Er*Cas12a exhibits an evolutionary distance from the other Cas12a proteins which made the identification of homologous amino acids challenging. To identify corresponding amino acids that can form interactions with the target DNA, the amino acid sequence of *Er*Cas12a was compared with that of *As*Cas12a and *Lb*Cas12a (Table 1). As a result, many highly conserved stretches between the proteins, as well as some areas of lower conservation, could be identified. High conservation exists at the N-terminal region, as well as in the REC1 domain where the interaction with the crRNA/DNA heteroduplex occurs (Yamano *et al*., 2016). In addition, there is high conservation in the WED II/PI domain and especially in the region where the RuvC I and RuvC II single motifs and the bridging helix are present. The amino acids whose substitutions were responsible for the enhanced efficiency of the tt*Lb*Cas12a variant, D156 (Schindele and Puchta, 2020), and had the main influence on the function of en*As*Cas12a, E174 (Kleinstiver *et al*., 2019), are located within the REC1 domain where the interaction with the crRNA/DNA heteroduplex occurs. Both amino acids are acidic, negatively charged amino acids and, thus, exhibit structural similarity. No conservation, on the other hand, is evident concerning these amino acids between *Er*Cas12a and the other orthologous Cas12a nucleases, since lysine, a basic, positively charged amino acid, is present at this position in *Er*Cas12a. The strong activity of en*As*Cas12a and tt*Lb*Cas12a is likely mainly due to the change in charge, based on a substitution of acidic amino acids by the large, basic amino acid arginine. In this way, new contacts to the target DNA can be established. In the case of *Er*Cas12a, however, lysine, a basic, positively charged amino acid, is already present at this position. The other amino acid substitutions of S542R/K548R in en*As*Cas12a and G532R/K538R in en*Lb*Cas12a, took place in the WED-II/PI domain. In *Er*Cas12a, D532 and K538 are located at these sites, and lysine is highly conserved at this position between the examined Cas12a nucleases.

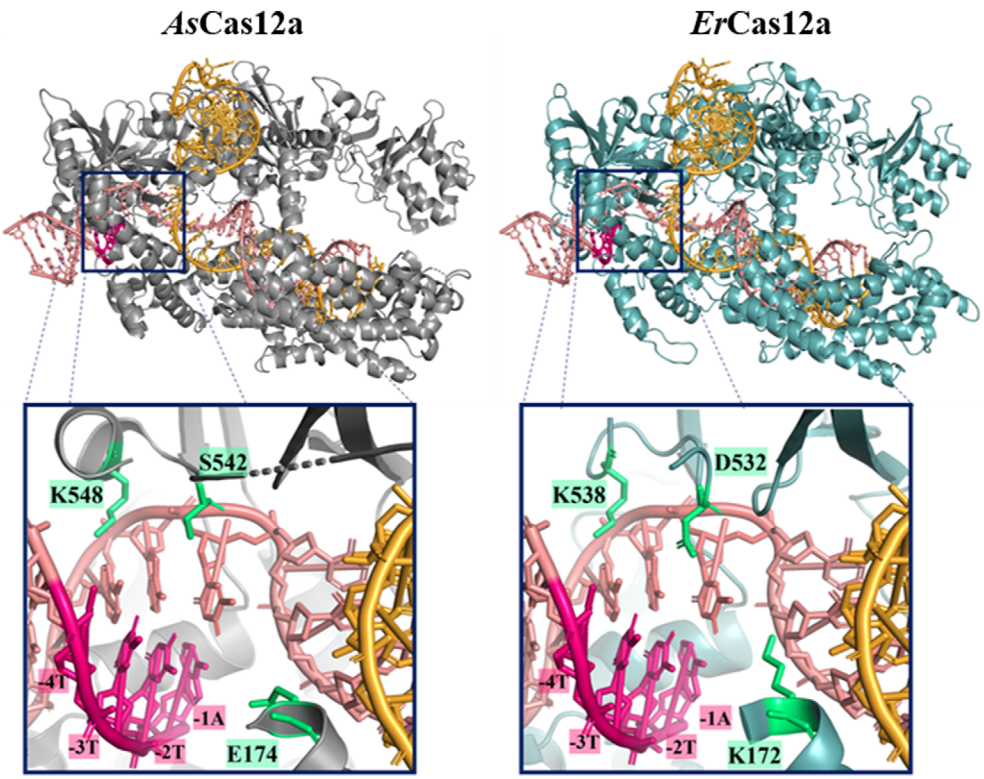
Furthermore, using existing crystal structures of *As*Cas12a (Yamano *et al*., 2016) and *Lb*Cas12a (Yamano *et al*., 2017), the bioinformatic calculation of a theoretical model of the *Er*Cas12a protein structure was performed using the bioinformatics online tool SWISS-MODEL (Bertoni *et al*., 2017). The comparison of the crystal structures with the calculated model of *Er*Cas12a showed a high structural similarity between the orthologous Cas12a proteins and *Er*Cas12a. The majority of the two structures could be visually overlapped using PyMOL for visualization since most of the secondary structures occupy almost identical positions (Schrödinger, 2015). The analogous amino acids (Table 1), which were determined by the alignment of the protein sequences, are also localized at comparable positions in the calculated model (Figure 2a). Although the structure of *Er*Cas12a is a theoretically calculated model, it can nevertheless provide first indications that the homologous amino acids might be located at conserved positions in both proteins. Accordingly, substitutions comparable to the ones described for the generation of improved *As*Cas12a and *Lb*Cas12a variants, might lead to comparable effects on the editing efficiency.

Analysis of the en*Er*Cas12a and *Er*Cas12a-Ultra variants

An en*Er*Cas12a variant was generated analogously to en*As*-Cas12a and en*Lb*Cas12a, with amino acid substitution K172R as an analogue to D156R in en*Lb*Cas12a and E174R in en*As*Cas12a. Because of the earlier mentioned charge differences between the respective amino acids, it was uncertain if this lysine in *Er*Cas12a takes on the same function and if the effect of the substitution by arginine would lead to an effect comparable to the one described for en*As*Cas12a and en*Lb*Cas12a. Unlike lysine, arginine contains a hydrophilic, basic-reacting guanidino group that acts as a very strong base and provides more opportunities for interactions than the amino group in lysine (Berg *et al*., 2018). Furthermore, analogously to the efficient *As*Cas12a-Ultra variant, a *Er*Cas12a-Ultra variant was generated that contained the two mutations I527R and F843L. The *As*Cas12a-Ultra variant was generated in another attempt to reduce the temperature sensitivity of *As*Cas12a and achieve constant activity even at lower

Figure 2 Comparison of *Er*Cas12a and *Er*Cas12a variants. (a) Shown is a model of the experimentally confirmed structure of the *As*Cas12a-crRNA/DNA complex (NCBI code 5B43; top left), and a theoretical model of *Er*Cas12a calculated based on the model of the *As*Cas12a-crRNA/DNA complex, with crRNA/DNA duplex optically inserted (top right). *As*Cas12a is shown in grey, *Er*Cas12a in turquoise and the corresponding crRNA and DNA in orange and pink, respectively. In the close-up of the PAM-proximal region, the PAM sequence is highlighted in pink ([−4T], [−3T], [−2T], [−1A]), and the three amino acids altered in the variants of *As*Cas12a and *Er*Cas12a are shown in green. (b) Overview of relative editing efficiencies, shown as averages of 20–40 plants tested for each target and the two different temperatures. (c) Plot of editing efficiency of en*Er*Cas12a and *Er*Cas12a Ultra variants compared with *Er*Cas12a at 22 °C and 28 °C. Boxplots show the distribution of editing efficiency of individual plants for each tested target (1–5) at temperatures 22 °C and 28 °C, respectively. The boxes represent the upper and lower quartiles and the median is drawn as a solid black line in the boxes. In addition, a cross represents the respective mean value of the data. The values that are still within 1.5 times the interquartile range are indicated by the antennas. Outliers that fall beyond this value are represented by black dots. The number of analysed plants for each variant and target is specified in the table, *P*-values were calculated using the one-way ANOVA test: $*P < 0.05$; $**P < 0.01$; $***P < 0.001$.

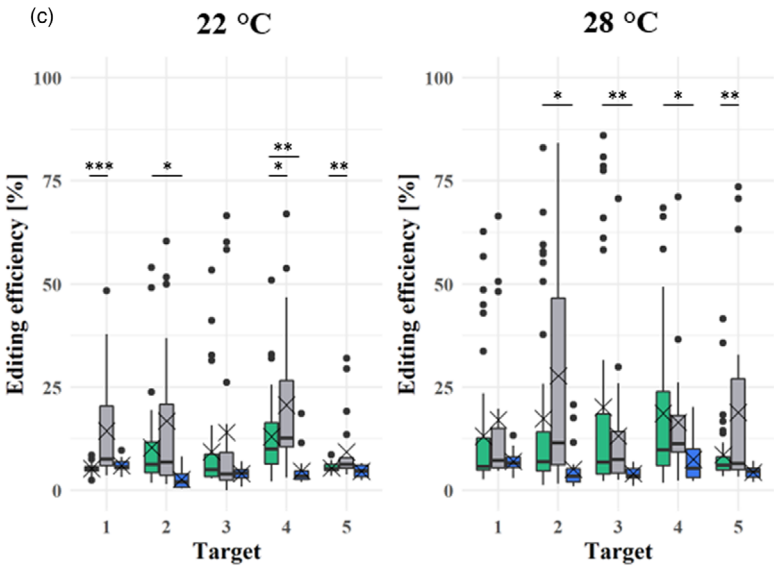
(a)



(b)

Target	Variant	Efficiency 22 °C MV [%]	Efficiency 28 °C MV [%]
1	*Er*Cas12a	5.2	13.1
	en*Er*Cas12a	14.4	17.1
	Ultra	6.0	7.0
2	*Er*Cas12a	10.3	17.4
	en*Er*Cas12a	16.8	27.7
	Ultra	2.6	5.1
3	*Er*Cas12a	9.3	20.1
	en*Er*Cas12a	14.0	13.2
	Ultra	4.0	4.0
4	*Er*Cas12a	13.0	18.6
	en*Er*Cas12a	20.6	16.4
	Ultra	4.6	7.5
5	*Er*Cas12a	5.4	8.6
	en*Er*Cas12a	9.3	18.9
	Ultra	4.6	4.3

(c)



analysed plants	22 °C T1	T2	T3	T4	T5	28 °C T1	T2	T3	T4	T5
*Er*Cas12a	40	40	38	40	40	40	40	39	40	40
en*Er*Cas12a	20	20	20	20	20	19	15	20	20	20
Ultra	20	21	20	20	23	19	20	20	20	20

temperatures (Zhang *et al*., 2021a). In the case of *As*Cas12a-Ultra, substitution of the amino acids M537R and F870L resulted in a significant increase in editing efficiencies, even at lower temperatures (Zhang *et al*., 2021a). The editing efficiency of the different variants was tested just like the editing efficiency of wild-type *Er*Cas12a: at the five *ECA3* target sites and by comparing the editing efficiency with that of wild-type *Er*Cas12a. The average values showed a slightly higher editing efficiency of en*Er*Cas12a compared to *Er*Cas12a (Figure 2b,c). For both *Er*Cas12a and en*Er*Cas12a, the average editing efficiency at 28 °C was higher than at 22 °C. A 1.2- to 1.7-fold stronger editing efficiency of en*Er*Cas12a was detected when the

temperature was increased to 28 °C. Comparing the median of these two variants, no clear difference could be detected between *Er*Cas12a and en*Er*Cas12a. On the other hand, no editing activity of *Er*Cas12a-Ultra could be detected, neither at 22 °C nor at 28 °C. In contrast to the other two lines, no individual plants displaying higher editing efficiencies could be observed in this case.

Analyses of the *Er*Cas12a α-helix1 variants

Analogously to the generation of the highly efficient tt*Lb*Cas12a, the sole substitution of the amino acid K172 was of particular interest to improve *Er*Cas12a. Additionally, we searched for further amino acids that could play key roles in the protein–DNA interaction within the REC1 domain and whose exchange could lead to enhanced interactions within the protein. In proximity of K172 at position 169 lies a negatively charged aspartic acid that could possibly represent the actual equivalent to D156 from *Lb*Cas12a or E174 from *As*Cas12a. Furthermore, we searched for amino acids in the same region that either had a negative charge or were small and nonpolar or uncharged. In these cases, a substitution by the large, positively charged arginine, would likely have the greatest effect and might enable additional interactions with the DNA, which could have implications for editing efficiency. Thus, the three amino acids D169, K172 and A175 were selected due to being located in the α-helix structure which is likely analogous to the α-helices in which the amino acids D156 and E174 are positioned in *Lb*Cas12a and *As*Cas12a, respectively. In the further course, this α-helix will be referred to as α-helix1.

The D169R and A175R variants showed on average editing efficiencies which were comparable to that of *Er*Cas12a (Figure 3a,b). For variant K172R, successful edits were observed at both tested temperatures. The average values indicate slightly increased editing efficiencies for all tested targets compared to *Er*Cas12a and a stronger activity at 28 °C. Compared to *Er*Cas12a, editing at 22 °C was increased between 1.5- and 1.7-fold at three of the five tested targets. At 28 °C, only two out of five targets, target 3 and 4, showed an increase in efficiency, by 1.5- and 1.6-fold compared to *Er*Cas12a with an editing efficiency of approximately 30%.

Analyses of the *Er*Cas12a α-helix2 variants

In addition to the before-mentioned variants, substitutions of three amino acids, E152, K153 and V156, located in another adjacent α-helix structure, were also tested. This structure is from now on referred to as α-helix2. Depending on the folding of the protein, mutations in this area could lead to further interactions in the protein/crRNA/DNA complex or have effects on the position of other α-helices, which could indirectly lead to further interactions. In the case of the E152R and K153R variants, successful edits were detected in individual plants (Figure 3c,d). The average values indicate similar editing efficiencies to *Er*Cas12a at all five targets tested. An exception to this finding was the occurrence of a high editing efficiency of K153R at 22 °C which can be traced back to a few individuals showing very high editing efficiency. However, this effect could not be reproduced at 28 °C. In the case of the variant V156R, already at 22 °C, a slight increase in editing was observed at two of the tested targets compared to *Er*Cas12a. At 28 °C, at three targets, a strikingly higher editing efficiency, with averages between 30% and 40%, could be achieved.

Generation and validation of an improved *Er*Cas12a variant

Due to the already notable increase in editing efficiencies, especially at 28 °C, which were achieved with the two variants V156R and K172R, we generated a new variant, called improved*Er*Cas12a (im*Er*Cas12a), which combines both substitutions. By combining these two variants, a dramatic increase in editing efficiency was achieved (Figure 4). Already at 22 °C, strong average editing efficiencies between 30% and 50% were achieved at the targets 2, 3 and 4. This number increased to over 50% at 28 °C. Importantly, im*Er*Cas12a also enabled editing of the formerly inaccessible targets 1 and 5 at 28 °C with efficiencies of 41.4% and 27.2%, respectively. A closer look at the graphs shows that, already at 22 °C, editing efficiencies of more than 50% could be achieved at target 2 in half of the tested plants. At targets 2, 3 and 4, individual editing efficiencies of more than 75% were observed. At 28 °C, a median of over 50% editing efficiency could be achieved at three targets. While other variants failed to edit targets 1 and 5, we achieved robust editing at both targets by using im*Er*Cas12a. Remarkably, despite the mediocre average editing efficiency of 27.2% at target 5, single plants showed efficiencies of up to 70%. This demonstrates the clear superiority of the im*Er*Cas12a variant over the previously described variants and the wild-type *Er*Cas12a. However, no change in the repair pattern of the induced breaks could be detected, indicating that the protein modifications only alter the frequency of break induction and not the break pattern itself (Data S1–S4).

Off-target analysis of the improved *Er*Cas12a variant

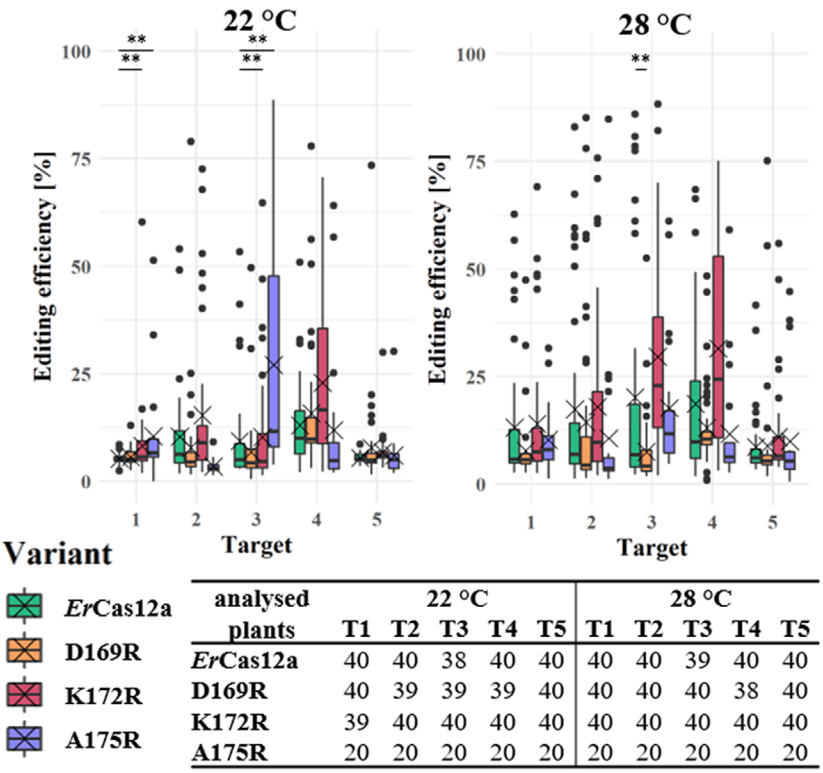
To date, the specificity of gene editing using *Er*Cas12a has only been investigated in mammalian cells *in vivo* (Liu *et al*., 2020) and in regenerated rice plants (Lin *et al*., 2021).In both cases, no editing at non-predetermined sites has been described. For Arabidopsis, the specificity of *Er*Cas12a has not yet been described. Furthermore, the question arose whether the significant increase in activity of im*Er*Cas12a compared to *Er*Cas12a could also lead to lower specificity and, accordingly, higher off-target activity. In general, Cas12a nucleases show high target specificity (Kim *et al*., 2016, 2017; Kleinstiver *et al*., 2016). However, for en*As*Cas12a with increased editing efficiency, increased editing at non-matching targets could also be observed, most notably at similar target sites accompanied by different PAM

Figure 3 Quantitative determination of the editing efficiency of the *Er*Cas12a variants. Overview of relative editing efficiencies of the α-helix1 variants (a) and the α-helix2 variants (c), given as average values of 20–40 plants tested for each target and the two different temperatures. (b) and (d) Plot of editing efficiencies of tested variants compared with *Er*Cas12a at 22 °C and 28 °C. Boxplots show the distribution of editing efficiencies of individual plants for each tested target (1–5) at temperatures of 22 °C and 28 °C. The boxes represent the upper and lower quartiles, and the median is drawn as a solid black line in the boxes. In addition, a cross represents the respective mean value of the data. The values that are still within 1.5 times the interquartile range are indicated by the antennas. Outliers that fall beyond this value are represented by black dots. The number of analysed plants for each variant and target is specified in the tables, *P*-values were calculated using the one-way ANOVA test: $*P < 0.05$; $**P < 0.01$; $***P < 0.001$.

(a)

Target	Variant	Efficiency 22 °C MV [%]	Efficiency 28 °C MV [%]
1	*Er*Cas12a	5.2	13.1
	D169R	5.6	7.7
	K172R	8.2	13.9
	A175R	10.4	10.2
2	*Er*Cas12a	10.3	17.4
	D169R	7.8	14.3
	K172R	15.3	18.0
	A175R	3.4	10.6
3	*Er*Cas12a	9.3	20.1
	D169R	6.8	7.4
	K172R	10.2	29.6
	A175R	27.0	17.6
4	*Er*Cas12a	13.0	18.6
	D169R	15.8	13.1
	K172R	22.9	31.5
	A175R	11.9	11.6
5	*Er*Cas12a	5.4	8.6
	D169R	7.6	8.9
	K172R	6.8	10.9
	A175R	6.0	9.9

(b)

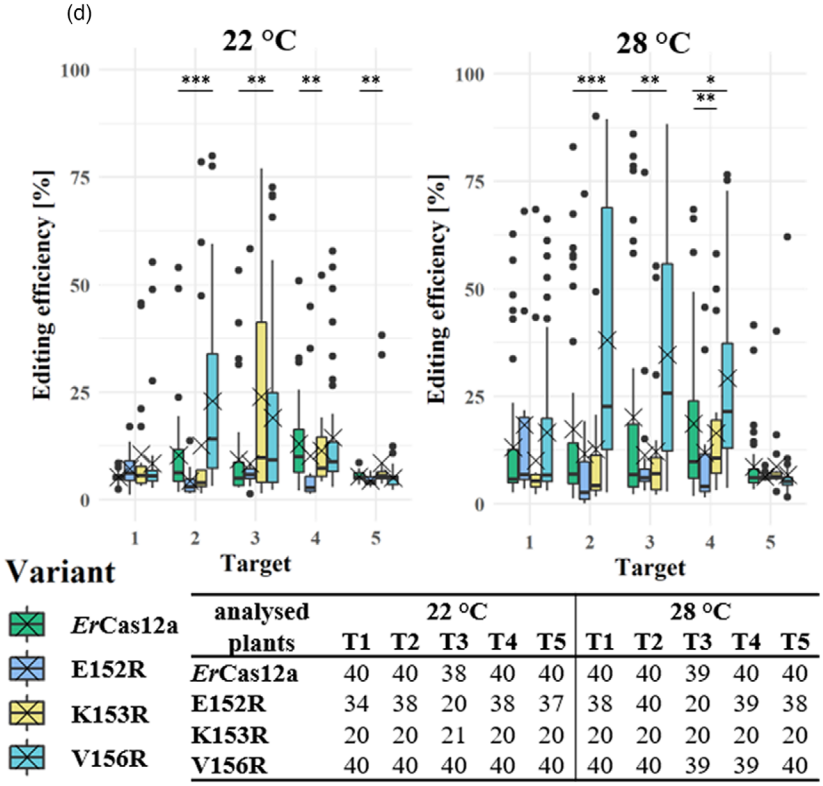


analysed plants	22 °C T1	T2	T3	T4	T5	28 °C T1	T2	T3	T4	T5
*Er*Cas12a	40	40	38	40	40	40	40	39	40	40
D169R	40	39	39	39	40	40	40	40	38	40
K172R	39	40	40	40	40	40	40	40	40	40
A175R	20	20	20	20	20	20	20	20	20	20

(c)

Target	Variant	Efficiency 22 °C MV [%]	Efficiency 28 °C MV [%]
1	*Er*Cas12a	5.2	13.1
	E152R	7.2	18.4
	K153R	10.5	10.1
	V156R	8.4	16.6
2	*Er*Cas12a	10.3	17.4
	E152R	4.3	11.6
	K153R	12.6	12.7
	V156R	22.9	38.1
3	*Er*Cas12a	9.3	20.1
	E152R	8.4	11.3
	K153R	24.0	12.2
	V156R	19.1	34.7
4	*Er*Cas12a	13.0	18.6
	E152R	10.1	11.8
	K153R	11.4	16.5
	V156R	14.4	28.5
5	*Er*Cas12a	5.4	8.6
	E152R	4.4	6.3
	K153R	8.5	8.6
	V156R	5.1	6.7

(d)



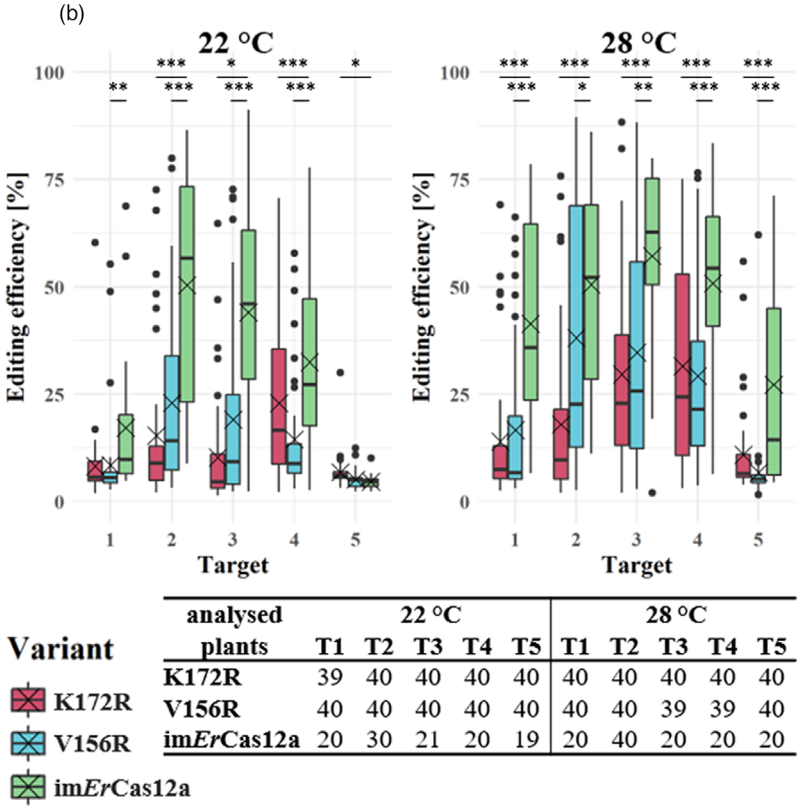
analysed plants	22 °C T1	T2	T3	T4	T5	28 °C T1	T2	T3	T4	T5
*Er*Cas12a	40	40	38	40	40	40	40	39	40	40
E152R	34	38	20	38	37	38	40	20	39	38
K153R	20	20	21	20	20	20	20	20	20	20
V156R	40	40	40	40	40	40	40	39	39	40

sequences (Kleinstiver *et al*., 2019). To determine the off-target activity of *Er*Cas12a and im*Er*Cas12a in Arabidopsis we chose a target in the *SUC1* gene which is part of the highly conserved *SUCROSE CARRIER* (*SUC*) gene family that consists of nine homologous genes (*SUC1*–*SUC9*). Within those genes, we found five potential off-targets with one to three mismatches compared to the on-target sequence (Figure 5a). Analysis of the editing efficiencies at the on- and off-target sites was performed using

(a)

Target	Variant	Efficiency 22 °C MV [%]	Efficiency 28 °C MV [%]
1	K172R	8.2	13.9
	V156R	8.4	16.6
	im*Er*Cas12a	17.1	41.4
2	K172R	15.3	18.0
	V156R	22.9	38.1
	im*Er*Cas12a	50.4	50.5
3	K172R	10.2	29.6
	V156R	19.1	34.7
	im*Er*Cas12a	43.4	57.2
4	K172R	22.9	31.5
	V156R	14.4	28.5
	im*Er*Cas12a	32.5	43.2
5	K172R	6.8	10.9
	V156R	5.1	6.7
	im*Er*Cas12a	4.6	27.2

(b)



analysed plants	22 °C T1	T2	T3	T4	T5	28 °C T1	T2	T3	T4	T5
K172R	39	40	40	40	40	40	40	40	40	40
V156R	40	40	40	40	40	40	40	39	39	40
im*Er*Cas12a	20	30	21	20	19	20	40	20	20	20

Figure 4 Quantitative determination of the editing efficiency of engineered *Er*Cas12a variants. Overview of the relative editing efficiencies of engineered *Er*Cas12a variants (a), given as average values of 20–40 plants tested for each target and the two different temperatures. (b) Plot of editing efficiencies of the tested variants at 22 °C and 28 °C. Boxplots show the distribution of editing efficiencies of individual plants for each tested target (1–5) at temperatures of 22 °C and 28 °C. The boxes represent the upper and lower quartiles, and the median is drawn as a solid black line in the boxes. In addition, a cross represents the respective mean value of the data including the outliers. The values that are still within 1.5 times the interquartile range are indicated by the antennas. Outliers that fall beyond this value are represented by black dots. The number of analysed plants for each variant and target is specified in the table, *P*-values were calculated using the one-way ANOVA test: *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

TIDE analysis as described before. The average values show that im*Er*Cas12a can only edit the on-targets with noteworthy efficiency, at both 22 °C and 28 °C. *Er*Cas12a could neither efficiently edit the on-targets at 22 °C or 28 °C, nor could it efficiently edit any of the five possible off-targets. Im*Er*Cas12a also showed a low editing efficiency at 22 °C. The on-target editing of *Er*Cas12a could not be increased, even when the temperature was increased to 28 °C. Im*Er*Cas12a already showed an increase in editing at 22 °C, which could be further increased at the on-target *SUC1* at 28 °C. In the off-target analysis, neither efficient editing could be detected for im*Er*Cas12a, nor were there any individual outliers in the off-targets. All average values lie in the single-digit range. Based on the detected values of the editing efficiencies at the off-target sequences, a high specificity of im*Er*Cas12a can be confirmed.

Discussion

Influence of the crRNA system on *Er*Cas12a editing

At the beginning of this study, we tested the optimal structure of the crRNA of *Er*Cas12a for application in plants. The coding DNA sequence of the pre-crRNA (Maksimova *et al.*, 2019) contains five consecutive thymidine bases in the 35 nt-long direct repeat which can also be recognized as a premature termination sequence when expressed under the control of a RNA polymerase III promoter (Gao *et al.*, 2017). Using the mature 42 nt crRNA, it is possible to bypass the consecutive thymidine bases of the coding DNA sequence and, thus, the premature termination of transcription. On the other hand, studies using *Lb*Cas12a in rice showed that using the non-processed pre-crRNA can work better than the short, processed crRNAs (Xu *et al.*, 2017). In this study, we compared the non-processed crRNA with the processed crRNA, and both variants were also tested with flanking ribozymes. In previous work in rice and wheat protoplasts, no difference was found between the 56 nt and the 42 nt ribozyme-flanked crRNA (Lin *et al.*, 2021), whereas we found that the 56 nt crRNA, flanked by ribozymes, performed considerably worse at 22 °C than its shorter counterpart. In addition, its efficiency could not be enhanced by increasing the temperature to 28 °C. In agreement with the results of Xu *et al.*, 2017 using *Lb*Cas12a in rice, the 42 nt crRNA without ribozymes showed low activity. Most studies on this topic found that the use of flanking ribozymes with the mature crRNA achieves robust editing by Cas12a nucleases in plants (Gao *et al.*, 2018; Lin *et al.*, 2021; Schindele and Puchta, 2020; Tang *et al.*, 2017; Zhang *et al.*, 2021b). In this study, we could confirm that *Er*Cas12a has comparable crRNA expression requirements for gene editing approaches as other Cas12a orthologues.

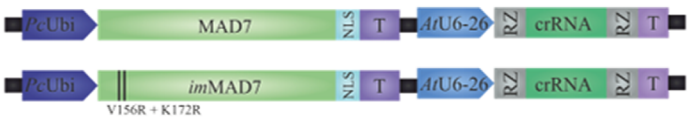
Establishment of optimized *Er*Cas12a variants

*Er*Cas12a has not yet been used in Arabidopsis but experiments in rice, wheat and maize showed partly comparable activity with

(a)

Target	PAM	Sequence
SUC1	**TTTG**	TTCGATACTGATTGGATGGGT
SUC6	**TTTG**	T**A**CGATACTGATTGGATGGGT
SUC7	**TTTA**	T**A**CGATACTGATTGGATGGGT
SUC8	**TTTA**	T**A**CGATAC**C**GATTGGATGGGT
SUC9	**TTTA**	T**A**CGATAC**C**GATTGGATGGGT
SUC5	**TTTG**	T**AT**GATAC**C**GATTGGATGGGT

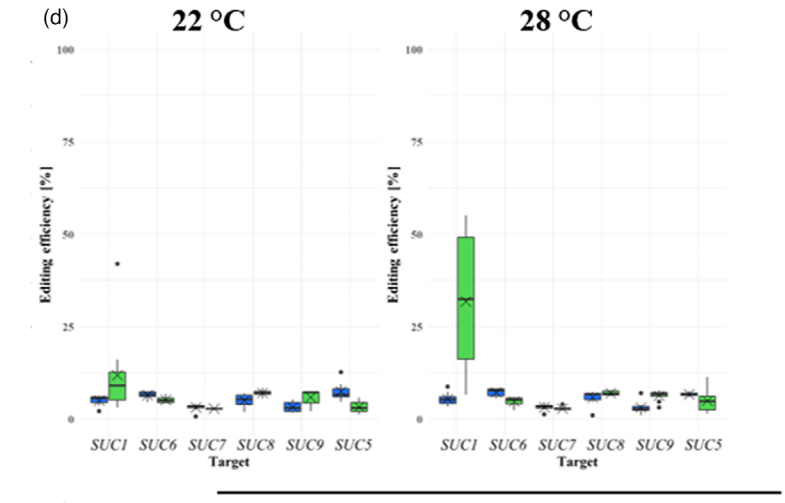
(b)



(c)

Target	Variant	Efficiency 22 °C MV [%]	Efficiency 28 °C MV [%]
SUC1	*Er*Cas12a	5.0	5.5
	im*Er*Cas12a	11.9	31.8
SUC6	*Er*Cas12a	6.5	7.2
	im*Er*Cas12a	5.2	4.8
SUC7	*Er*Cas12a	3.2	3.2
	im*Er*Cas12a	2.8	2.9
SUC8	*Er*Cas12a	5.0	5.8
	im*Er*Cas12a	7.1	7.0
SUC9	*Er*Cas12a	3.3	3.1
	im*Er*Cas12a	5.9	6.3
SUC5	*Er*Cas12a	7.5	6.7
	im*Er*Cas12a	3.3	5.0

(d)



analysed plants	*SUC1*	*SUC6*	*SUC7*	*SUC8*	*SUC9*	*SUC5*
			22 °C			
*Er*Cas12a	9	6	10	10	10	7
im*Er*Cas12a	10	9	9	10	9	10
			28 °C			
*Er*Cas12a	10	6	10	10	8	8
im*Er*Cas12a	10	10	10	10	8	10

Figure 5 Quantitative determination of off-target editing efficiency of im*Er*Cas12a. (a) Overview of the on-target site in the *SUC1* gene, as well as the five off-target sites *SUC5* to *SUC9*. The nucleotides highlighted in red indicate the mismatches in each sequence. (b) Schematic structure of the *Er*Cas12a expression cassettes, which are analogous to the expression cassettes used to determine editing efficiencies in the ECA3 gene. The guide was designed to address the target in *SUC1*. (c) Overview of the relative editing efficiencies of *Er*Cas12a and im*Er*Cas12a, given as average values of ten tested plants for the on-target, and the examined possible off-target sequences. The editing efficiency was determined at the two different temperatures, 22 °C and 28 °C. (d) Boxplots show the distribution of editing efficiencies of individual plants at the respective targets (*SUC1*/*SUC5*–*SUC9*) at temperatures of 22 °C and 28 °C. Boxes represent the upper and lower quartiles, and the median is drawn as a solid black line in the boxes. In addition, the mean value of the data is shown as a cross. The values that are still within 1.5 times the interquartile range are indicated by the antennas. Outliers that lie outside this value are represented by black dots.

*Lb*Cas12a in protoplasts (Lin *et al.*, 2021; Zhang *et al.*, 2021b). In this study, we observed stronger target sequence-dependent variations in editing efficiencies of *Er*Cas12a compared to *Lb*Cas12a as well as a temperature sensitivity comparable to other Cas12a systems (Malzahn *et al.*, 2019; Schindele and Puchta, 2020; Zhang *et al.*, 2021b). Editing by *Er*Cas12a resulted in low efficiencies at all five tested targets at 22 °C; two targets were completely inaccessible for *Er*Cas12a-mediated gene editing. While increasing the temperature to 28 °C resulted in enhanced editing efficiencies, the average remained below 20% at all targets.

Based on knowledge obtained from former Cas12a nuclease improvements, we designed several *Er*Cas12a variants, of which two, K172R and V156R, showed higher efficiencies at some targets. By combining these variants, we obtained an improved *Er*Cas12a which enabled robust editing at all targets (Figure 4), including those that were previously inaccessible to every other tested *Er*Cas12a variant and en*Lb*Cas12a (Schindele and Puchta, 2020). At target 1, im*Er*Cas12a even outperformed the highly efficient tt*Lb*Cas12a variant at 22 °C (Schindele and Puchta, 2020).

For both the amino acid substitution K172R as well as V156R, it can be seen that arginine can form more interactions with neighbouring structures in the predicted structure (Figure S1). In the case of K172R, additional contacts with the backbone of the DNA in the PAM-proximal region as well as additional hydrogen

bonds with the base at the 3′ end of the PAM might form. The resulting enhanced editing efficiency is in line with results from other studies, which showed that non-negatively charged amino acids within DNA-interacting protein domains increase editing activities of different Cas12a orthologues, probably due to improved DNA binding (Zhang *et al*., 2021b, 2023). However, it should be kept in mind that this structure only reflects a probable model of *Er*Cas12a and that the crRNA/DNA duplex was only visually inserted into the model, based on the analogous *As*Cas12a. Thus, this model does not necessarily correspond to the actual structure. However, at this point, the use of the calculated model of *Er*Cas12a and inserted crRNA/DNA duplex offers a possible explanation for the slight increase in editing due to the replacement of L172R. Polar interactions between amino acids and the DNA backbone may lead to stabilization of the protein/DNA complex (Luscombe *et al*., 2001). In the case of V156R, it is likely that steric repulsion may occur at this site between the two helices, in which case the α-helix2 would distance itself from the other helix. Because of the short random coil connection of three amino acids between α-helix1 and α-helix2, the altered positioning of α-helix2 is likely to affect α-helix1 as well. In turn, the altered positioning of α-helix1 could also strengthen the binding between the nuclease and the crRNA/DNA duplex in conjunction with lysine 172. This could also lead to an increase in editing efficiency of the V156R variant. The presumably stronger binding of the nuclease to DNA, due to the single mutation K172R, only led to a minor increase in efficiency. Similarly, the possible structural change of the protein, due to the substitution of V156R in a region where PAM-proximal binding of DNA occurs, suggests a slight increase in efficiency. Only the combination of the two substitutions V156R and K172R led to a substantial increase in editing efficiencies. The respective variant, im*Er*Cas12a, exhibited much stronger editing at all targets compared to the two variants with only one of the two mutations.

Interestingly, the other *Er*Cas12a tested variants did not show an increase in editing efficiencies (Figures 2 and 3) and partially even resulted in a loss of nuclease activity (Figure 2). Especially for the en*Er*Cas12a and *Er*Cas12a-Ultra variants, these results are not in line with expectations. In the case of en*Er*Cas12a, the third substituted amino acid lysine 538 was conserved between the *Er*Cas12a and *As*Cas12a. The substitution of lysine 172 to arginine, which is also part of the en*Er*Cas12a variant, individually resulted in a slight increase in editing activity (Figure 3). However, the two additional mutations of aspartic acid 532 and lysine 538 to arginine, respectively, completely abolished the slight positive effect of the first mutation. Although this observation was also made in the case of *Lb*Cas12a with the comparison between the two variants en*Lb*Cas12a and tt*Lb*Cas12a (Schindele and Puchta, 2020), the closer similarity of *Er*Cas12a to *As*Cas12a led to the assumption that the effect that was observed for en*As*Cas12a was likely to be reproducible in the case of en*Er*Cas12a (Kleinstiver *et al*., 2019). Of the two amino acids aspartic acid 532 and lysine 538, lysine 538 showed the highest conservation between the orthologous Cas12a nucleases, whereas aspartic acid 532 was not conserved between the orthologues. Accordingly, it can be assumed that the conserved lysine 538 performs a comparable function in all orthologues. However, mutation of this amino acid did not cause a reduction in nuclease activity in en*As*Cas12a, so it seems more likely that substitution of aspartic acid 532 by arginine caused the reduction in editing efficiencies. A similar effect was likely achieved by substitution of glycine 532 by arginine in *Lb*Cas12a (Schindele and Puchta, 2020).

Despite the high degree of conservation, essential differences can nevertheless exist between the nucleases, as was observed when analysing the *Er*Cas12a-Ultra variant. Here, the corresponding amino acids methionine 537 and phenylalanine 870 in *As*Cas12a and isoleucine 527 and phenylalanine 843 in *Er*Cas12a, respectively, were almost completely conserved between *As*Cas12a and *Er*Cas12a, since isoleucine and methionine, as non-polar, hydrophobic amino acids, are both likely to have similar properties in the protein structure. Therefore, an increased editing efficiency was expected when the amino acids were exchanged, analogously to the effect observed in the *As*Cas12a-Ultra variant (Zhang *et al*., 2021a). However, the opposite was true for *Er*Cas12a. In fact, no significant editing by the *Er*Cas12a-Ultra variant was detected at any of the targets, even at increased temperatures. Phenylalanine 843 in *Er*Cas12a, and at position 870 in *As*Cas12a, is highly conserved between the two orthologues and is located in the region that, at least in *As*Cas12a, interacts with the stem loop of the crRNA, possibly strengthening the stability of the Cas12a/crRNA complex. Using the two examples of en*Er*Cas12a and *Er*Cas12a-Ultra, it is again clear that, despite seemingly similar structures, orthologous nucleases nevertheless differ from each other and that essential differences in properties exist.

Off-target activity of *Er*Cas12a and im*Er*Cas12a

The nucleases en*As*Cas12a, which showed much more efficient editing at on-target sequences (Kleinstiver *et al*., 2019), as well as tt*Lb*Cas12a, with exceedingly efficient editing at the respective target sites, showed at least partially lower specificity (Kleinstiver *et al*., 2019; Schindele and Puchta, 2020). The possibility of increased editing at off-targets by increased editing activity also exists in the case of im*Er*Cas12a. In general, Cas12a nucleases are more specific than Cas9 nucleases and, accordingly, show less off-target effects (Xu *et al*., 2017). On the one hand, due to the increased editing efficiency, there is some likelihood that im*Er*Cas12a might also show higher off-target activity. To date, no off-target activities of *Er*Cas12a have been detected in mice and rats (Liu *et al*., 2020), nor have any off-target events been detected in plants (Lin *et al*., 2021). Im*Er*Cas12a also shows no editing of off-target sites, neither at 22 °C nor at 28 °C, confirming its high specificity in plants. The enhancement of the interaction of im*Er*Cas12a with the target DNA in the PAM-proximal region does not seem to affect the specificity of im*Er*Cas12a, contrary to was has been observed in the case of tt*Lb*Cas12a. Taken together, we added a new nuclease variant to the toolbox for efficient and target-specific gene editing in plants.

Future applications of im*Er*Cas12a

Can the cutting efficiency of imErCas12a be further improved? This will most probably be the case as the example of LbCas12a demonstrates: Recently, besides the already known activity enhancing mutation D156R (Schindele and Puchta, 2020), two other mutation could be identified by saturation mutagenesis, that in combination resulted in the variant LbCas12a-RRV showing a further increase of editing efficiency (Zhang *et al*., 2023). On the other hand, the integration of introns in the ORF of ttLbCas12a lead to drastic enhancement of gene editing and gene targeting frequencies (Schindele *et al*., 2023). Thus, it might well be that by applying or even combining these

strategies new versions of imErCas12a might be able to compete with, or even outcompete LbCas12 variants in the long run.

Besides its use as an efficient nuclease for the knock-out of genes in plants, there are a number of further applications that can be envisaged using imErCas12a: There is a direct correlation between the efficiency of DSB induction and the frequency of homologous recombination-based gene targeting in plants (Huang *et al*., 2021; Schindele *et al*., 2023; Wolter *et al*., 2018). Moreover, efficient DSB induction is a critical prerequisite for successful plant chromosome engineering (Rönspies *et al*., 2022). DNA-binding modules based on Cas12a-like enzymes have also been used for base editing (BE) (Gaillochet *et al*., 2023; Kleinstiver *et al*., 2019) and transcriptional control (Tang *et al*., 2017) before. Here, especially Cas12a-mediated BE was shown to rely on the enhanced protein–DNA interaction of engineered variants as BE based on wild-type Cas12a orthologues showed only little activity in plants (Cheng *et al*., 2023; Gaillochet *et al*., 2023). Moreover, due to its similarity to other Cas12a systems, established and highly efficient multiplexing approaches should be adoptable for imErCas12a-mediated targeting of multiple loci (Zhang *et al*., 2021b). Last but not least, imErCas12a might also be an attractive tool for application in other organisms than plants, especially if they require lower cultivation temperatures.

Methods

T-DNA constructs

Coding sequence of *Er*Cas12a was obtained codon optimized for Arabidopsis from biocat GmbH (Heidelberg, Germany) (Data S1).

T-DNA constructs used in this study are based on the previously described pDe-*Lb*Cas12a plasmid (Wolter and Puchta, 2019). Exchange of the Cas12 coding sequence was performed via AscI digestion and subsequent ligation resulting in pDe-*Er*Cas12a. Variants were generated via PCR-based site-directed mutagenesis, inducing codon changes listed in Data S2.

Gateway compatible entry-vectors containing expression cassettes for crRNAs are based on pEn-RZ-*As*Cas12a (Tang *et al*., 2017). Sequences of the 56 nt and 42 nt crRNAs as well as of the employed ribozymes are listed in Data S3. crRNA programming of the entry-vector was achieved using oligos which were inserted into the entry-vector as annealed oligos. The programmed crRNAs were transferred from the entry-vector to the destination vectors via Gateway LR reaction.

Plant transformation and cultivation

Arabidopsis plants were transformed via floral dip using the Agrobacterium strain GV3101, as previously described (Clough and Bent, 1998). T1 plants were grown on GM with cefotaxime (500 mg/L) and an appropriate selection marker for 2 weeks at either 22 °C or 28 °C with 16 h light per day. Genomic DNA was extracted using the method of Edwards *et al*., 1991.

Analysis of gene editing efficiency

Amplification of regions of interest was performed from genomic DNA using the proof-reading polymerase Q5® High-Fidelity DNA Polymerase (New England Biolabs). Sanger sequencing was performed by Eurofins Genomics GmbH. TIDE-analysis (Brinkman *et al*., 2014) was performed using a decomposition window from position 150–620 and an indel size range of 35 bp. Low-level editing efficiencies of <10% were considered noise and not true editing. Gene editing efficiencies of every plant are listed in Data S4. Statistical differences were calculated using the Mann–-Whitney-*U* test or the one-way ANOVA test.

Acknowledgements

We thank Carina Jülch, Alisa Röder, Clara Pfunder and Theresa Kraus for their excellent technical assistance. We thank Michelle Rönspies for critical reading of the manuscript. The im*Er*Cas12a construct is freely available on request.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

H.P. conceived the research. J.P. and P.S. designed the experiments. J.P. and N.C. executed the experiments. J.P., N.C. and H.P. wrote the article.

References

Bandyopadhyay, A., Kancharla, N., Javalkote, V.S., Dasgupta, S. and Brutnell, T.P. (2020) CRISPR-Cas12a (Cpf1): a versatile tool in the plant genome editing tool box for agricultural advancement. *Front. Plant Sci.* **11**, 584151.

Berg, J.M., Tymoczko, J.L., Gatto, G.J. and Stryer, L. (2018) *Stryer Biochemie*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Bernabé-Orts, J.M., Casas-Rodrigo, I., Minguet, E.G., Landolfi, V., Garcia-Carpintero, V., Gianoglio, S., Vázquez-Vilar, M. *et al*. (2019) Assessment of Cas12a-mediated gene editing efficiency in plants. *Plant Biotechnol. J.* **17**, 1971–1984.

Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L. and Schwede, T. (2017) Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci. Rep.* **7**, 10480.

Brinkman, E.K., Chen, T., Amendola, M. and van Steensel, B. (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* **42**, e168.

Capdeville, N., Merker, L., Schindele, P. and Puchta, H. (2021) Sophisticated CRISPR/Cas tools for fine-tuning plant performance. *J. Plant Physiol.* **257**, 153332.

Cheng, Y., Zhang, Y., Li, G., Fang, H., Sretenovic, S., Fan, A., Li, J. *et al*. (2023) CRISPR-Cas12a base editors confer efficient multiplexed genome editing in rice. *Plant Commun.* **4**, 100601.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J. Cell Mol. Biol.* **16**, 735–743.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D. *et al*. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science (New York, N.Y.)*, **339**, 819–823.

Edwards, K., Johnstone, C. and Thompson, C. (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res.* **19**, 1349.

Gaillochet, C., Peña Fernández, A., Goossens, V., D'Halluin, K., Drozdzecki, A., Shafie, M., van Duyse, J. *et al*. (2023) Systematic optimization of Cas12a base editors in wheat and maize using the ITER platform. *Genome Biol.* **24**, 6.

Gao, Z., Herrera-Carrillo, E. and Berkhout, B. (2017) Delineation of the exact transcription termination signal for type 3 polymerase III. *Mol. Ther. Nucleic Acids*, **10**, 36–44.

Gao, Z., Herrera-Carrillo, E. and Berkhout, B. (2018) Improvement of the CRISPR-Cpf1 system with ribozyme-processed crRNA. *RNA Biol.* **15**, 1458–1467.

Huang, T.-K., Armstrong, B., Schindele, P. and Puchta, H. (2021) Efficient gene targeting in Nicotiana tabacum using CRISPR/SaCas9 and temperature tolerant LbCas12a. *Plant Biotechnol. J.* **19**, 1314–1324.

Kim, D., Kim, J., Hur, J.K., Been, K.W., Yoon, S.-H. and Kim, J.-S. (2016) Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat. Biotechnol.* **34**, 863–868.

Kim, H.K., Song, M., Lee, J., Menon, A.V., Jung, S., Kang, Y.-M., Choi, J.W. *et al*. (2017) In vivo high-throughput profiling of CRISPR-Cpf1 activity. *Nat. Methods*, **14**, 153–159.

Kleinstiver, B.P., Sousa, A.A., Walton, R.T., Tak, Y.E., Hsu, J.Y., Clement, K., Welch, M.M. *et al*. (2019) Engineered CRISPR-Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. *Nat. Biotechnol.* **37**, 276–282.

Kleinstiver, B.P., Tsai, S.Q., Prew, M.S., Nguyen, N.T., Welch, M.M., Lopez, J.M., McCaw, Z.R. *et al*. (2016) Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat. Biotechnol.* **34**, 869–874.

Kumlehn, J., Pietralla, J., Hensel, G., Pacher, M. and Puchta, H. (2018) The CRISPR/Cas revolution continues: from efficient gene editing for crop breeding to plant synthetic biology. *J. Integr. Plant Biol.* **60**, 1127–1153.

Li, J.-F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M. *et al*. (2013) Multiplex and homologous recombination-mediated genome editing in Arabidopsis and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.* **31**, 688–691.

Lin, Q., Zhu, Z., Liu, G., Sun, C., Lin, D., Xue, C., Li, S. *et al*. (2021) Genome editing in plants with MAD7 nuclease. *J. Genet. Genomics*, **48**, 444–451.

Liu, Z., Schiel, J.A., Maksimova, E., Strezoska, Ž., Zhao, G., Anderson, E.M., Wu, Y. *et al*. (2020) ErCas12a CRISPR-MAD7 for model generation in human cells, mice, and rats. *CRISPR J.* **3**, 97–108.

Luscombe, N.M., Laskowski, R.A. and Thornton, J.M. (2001) Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res.* **29**, 2860–2874.

Maksimova, E., Schiel, J.A., Zhao, G., Machado, H.B., Anderson, E.M., D'Angelo, M., Collin, P. *et al*. (2019) *In vitro and in vivo CRISPR gene editing with MAD7. Poster*.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. *et al*. (2013) RNA-guided human genome engineering via Cas9. *Science (New York, N.Y.)*, **339**, 823–826.

Malzahn, A.A., Tang, X., Lee, K., Ren, Q., Sretenovic, S., Zhang, Y., Chen, H. *et al*. (2019) Application of CRISPR-Cas12a temperature sensitivity for improved genome editing in rice, maize, and Arabidopsis. *BMC Biol.* **17**, 9.

Moreno-Mateos, M.A., Fernandez, J.P., Rouet, R., Vejnar, C.E., Lane, M.A., Mis, E., Khokha, M.K. *et al*. (2017) CRISPR-Cpf1 mediates efficient homology-directed repair and temperature-controlled genome editing. *Nat. Commun.* **8**, 2024.

Østerberg, J.T., Xiang, W., Olsen, L.I., Edenbrandt, A.K., Vedel, S.E., Christiansen, A., Landes, X. *et al*. (2017) Accelerating the domestication of new crops: feasibility and approaches. *Trends Plant Sci.* **22**, 373–384.

Price, M.A., Cruz, R., Bryson, J., Escalettes, F. and Rosser, S.J. (2020) Expanding and understanding the CRISPR toolbox for *Bacillus subtilis* with MAD7 and dMAD7. *Biotechnol. Bioeng.* **117**, 1805–1816.

Rönspies, M., Schindele, P., Wetzel, R. and Puchta, H. (2022) CRISPR-Cas9-mediated chromosome engineering in *Arabidopsis thaliana*. *Nat. Protoc.* **17**, 1332–1358.

Scheben, A., Wolter, F., Batley, J., Puchta, H. and Edwards, D. (2017) Towards CRISPR/Cas crops – bringing together genomics and genome editing. *New Phytol.* **216**, 682–698.

Schindele, A., Dorn, A. and Puchta, H. (2020) CRISPR/Cas brings plant biology and breeding into the fast lane. *Curr. Opin. Biotechnol.* **61**, 7–14.

Schindele, P., Merker, L., Schreiber, T., Prange, A., Tissier, A. and Puchta, H. (2023) Enhancing gene editing and gene targeting efficiencies in Arabidopsis thaliana by using an intron-containing version of ttLbCas12a. *Plant Biotechnol. J.* **21**, 457–459.

Schindele, P. and Puchta, H. (2020) Engineering CRISPR/LbCas12a for highly efficient, temperature-tolerant plant gene editing. *Plant Biotechnol. J.* **18**, 1118–1120.

Schindele, P., Wolter, F. and Puchta, H. (2018) Das CRISPR/Cas-System. *Biol. Unserer Zeit*, **48**, 100–105.

Schrödinger, L.L. (2015) *The PyMOL molecular graphics system, version 1.8*.

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., Zhang, K. *et al*. (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 686–688.

Tang, X., Lowder, L.G., Zhang, T., Malzahn, A.A., Zheng, X., Voytas, D.F., Zhong, Z. *et al*. (2017) A CRISPR-Cpf1 system for efficient genome editing and transcriptional repression in plants. *Nat. Plants*, **3**, 17018.

Wada, N., Osakabe, K. and Osakabe, Y. (2022) Expanding the plant genome editing toolbox with recently developed CRISPR-Cas systems. *Plant Physiol.* **188**, 1825–1837.

Wierson, W.A., Simone, B.W., WareJoncas, Z., Mann, C., Welker, J.M., Kar, B., Emch, M.J. *et al*. (2019) Expanding the CRISPR Toolbox with ErCas12a in zebrafish and human cells. *CRISPR J.* **2**, 417–433.

Wolter, F., Klemm, J. and Puchta, H. (2018) Efficient in planta gene targeting in Arabidopsis using egg cell-specific expression of the Cas9 nuclease of *Staphylococcus aureus*. *Plant J. Cell Mol. Biol.* **94**, 735–746.

Wolter, F. and Puchta, H. (2019) In planta gene targeting can be enhanced by the use of CRISPR/Cas12a. *Plant J. Cell Mol. Biol.* **100**, 1083–1094.

Xu, R., Qin, R., Li, H., Li, D., Li, L., Wei, P. and Yang, J. (2017) Generation of targeted mutant rice using a CRISPR-Cpf1 system. *Plant Biotechnol. J.* **15**, 713–717.

Yamano, T., Nishimasu, H., Zetsche, B., Hirano, H., Slaymaker, I.M., Li, Y., Fedorova, I. *et al*. (2016) Crystal structure of Cpf1 in complex with guide RNA and target DNA. *Cell*, **165**, 949–962.

Yamano, T., Zetsche, B., Ishitani, R., Zhang, F., Nishimasu, H. and Nureki, O. (2017) Structural basis for the canonical and non-canonical PAM recognition by CRISPR-Cpf1. *Mol. Cell*, **67**, 633–645.e3.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E. *et al*. (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, **163**, 759–771.

Zhang, L., Li, G., Zhang, Y., Cheng, Y., Roberts, N., Glenn, S.E., DeZwaan-McCabe, D. *et al*. (2023) Boosting genome editing efficiency in human cells and plants with novel LbCas12a variants. *Genome Biol.* **24**, 102.

Zhang, L., Zuris, J.A., Viswanathan, R., Edelstein, J.N., Turk, R., Thommandru, B., Rube, H.T. *et al*. (2021a) AsCas12a ultra nuclease facilitates the rapid generation of therapeutic cell medicines. *Nat. Commun.* **12**, 3908.

Zhang, Y., Ren, Q., Tang, X., Liu, S., Malzahn, A.A., Zhou, J., Wang, J. *et al*. (2021b) Expanding the scope of plant genome engineering with Cas12a orthologs and highly multiplexable editing systems. *Nat. Commun.* **12**, 1944.

Zhu, H., Li, C. and Gao, C. (2020) Applications of CRISPR-Cas in agriculture and plant biotechnology. *Nat. Rev. Mol. Cell Biol.* **21**, 661–677.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Data S1 *Er*Cas12a coding sequence codon optimized for *Arabidopsis thaliana*.

Data S2 Codon changes induced in order to generate *Er*Cas12a variants.

Data S3 Coding sequences for used *Er*Cas12a crRNAs and ribozymes.

Data S4 Editing efficiency of different ErCas12a variants.

Figure S1 Transformation of the protein structure of *Er*Cas12a in the PAM-proximal region. Shown is a close-up of the PAM-proximal region (top) of the theoretical *Er*Cas12a structure, with crRNA/DNA complex optically inserted and of the α-helix2 region (bottom). Highlighted is the PAM in pink ([−4T], [−3T], [−2T], [−1A]), the *Er*Cas12a structure in turquoise, the DNA in pink and the crRNA in orange. In the top left image, the amino acid lysine is depicted in green at position 172, as it is present at this position in wild-type *Er*Cas12a. The top right image shows potential interactions calculated by PyMOL, resulting from a substitution by arginine. The lower left image shows the amino acid valine in green at position 156, as present at this position in wild-type *Er*Cas12a. The lower right image shows potential interactions calculated by PyMOL, resulting from substitution by arginine. In all four images, dashed lines in yellow indicate the possible polar contacts as well as their distances in angstroms (Å).

