the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2023) doi: 10.1111/tpj.16586

TECHNICAL ADVANCE

Increasing deletion sizes and the efficiency of CRISPR/Cas9-mediated mutagenesis by SunTag-mediated TREX1 recruitment

Niklas Capdeville, Patrick Schindele and Holger Puchta*

Department of Molecular Biology, Karlsruhe Institute of Technology (KIT), Joseph Gottlieb Kölreuter Institute for Plant Sciences, Fritz-Haber-Weg 4, 76131 Karlsruhe, Germany

Received 2 October 2023; revised 27 November 2023; accepted 30 November 2023.
*For correspondence (e-mail holger.puchta@kit.edu).

SUMMARY

Previously, it has been shown that mutagenesis frequencies can be improved by directly fusing the human exonuclease TREX2 to Cas9, resulting in a strong increase in the frequency of smaller deletions at the cut site. Here, we demonstrate that, by using the SunTag system for recruitment of TREX2, the mutagenesis efficiency can be doubled in comparison to the direct fusion in *Arabidopsis thaliana*. Therefore, we also tested the efficiency of the system for targeted deletion formation by recruiting two other 3′-5′ exonucleases, namely the human TREX1 and *E. coli* ExoI. It turns out that SunTag-mediated recruitment of TREX1 not only improved the general mutation induction efficiency slightly in comparison to TREX2, but that, more importantly, the mean size of the induced deletions was also enhanced, mainly via an increase of deletions of 25 bp or more. EcExoI also yielded a higher amount of larger deletions. However, only in the case of TREX1 and TREX2, the effect was predominately SunTag-dependent, indicating efficient target-specific recruitment. Using SunTag-mediated TREX1 recruitment at other genomic sites, we were able to obtain similar deletion patterns. Thus, we were able to develop an attractive novel editing tool that is especially useful for obtaining deletions in the range from 20 to 40 bp around the cut site. Such sizes are often required for the manipulation of cis-regulatory elements. This feature is closing an existing gap as previous approaches, based on single nucleases or paired nickases or nucleases, resulted in either shorter or longer deletions, respectively.

Keywords: deletion, double strand break repair, exonuclease, gene editing, genome engineering.

INTRODUCTION

In view of growing challenges in the cultivation of crops, the demands for breeding new cultivars by optimizing existing plant lines or domesticating wild varieties are increasing. A central prerequisite for achieving this goal is the accessibility of a wide range of individual traits and varying trait expression, which relies on a high number of different genotypes. However, thousands of years of domestication and selective breeding have impoverished the genetic diversity of modern crops. Thus, different approaches were applied to increase the frequency of DNA damage which, upon defective repair, can result in beneficial mutations. In mutagenesis breeding, plants are treated with chemical or physical genotoxins and subsequently selected for the desired phenotype or genotype. Since this approach results in undirected mutations throughout the whole genome, unwanted changes must be compensated for by tedious rounds of backcrossing. From the mid-1990s onward, the development of sequence-specific nucleases enabled the targeted induction of DNA double-strand breaks (DSB) which accelerated the development of new cultivars (Pacher & Puchta, 2017). In somatic plant cells, DSB is mainly repaired via classical non-homologous end-joining (cNHEJ) (Puchta, 2005). Here, due to its ubiquitous occurrence, the KU70/KU80 heterodimer binds rapidly to the break ends, tethering them together and protecting them from further processing. Nevertheless, since cNHEJ-mediated repair requires the presence of blunt ends for religation, an addition or deletion of single nucleotides can be necessary. Since CRISPR/Cas9 induces DSB with up to three nucleotide 5′-overhangs, their cNHEJ-mediated repair leads to either perfect restoration of the original sequence or incorporation of small insertions or deletions (Jinek et al., 2012;

Lemos et al., 2018; Yang et al., 2017). Therefore, CRISPR/Cas9 can efficiently be used to disrupt open reading frames and, thus, to generate knock-out mutants. However, in the last years, regulatory sequences have come into the focus of plant breeders as many genetic changes underlying the phenotypic diversity of crop plants reside within cis-regulatory elements of unmodified coding regions (Swinnen et al., 2019). While single nucleotide changes can disable the interaction between those regulatory elements (RE) and the respective transcription factors (TF), a protospacer adjacent motif (PAM) in a suitable distance is needed to edit critical bases, greatly impacting the practicability of such approaches. To increase the size of induced mutations, nucleases assigned to the CRISPR/Cas12a systems could be used as they generate DSB with 5 nt 5′-overhangs. However, their main editing outcome is deletions of about 5–10 bp, which might still be insufficient to disrupt RE, which have no suitable cutting sites in close proximity, or for editing of RE clusters. However, the induction of two adjacent breaks, using paired nickases or nucleases, which resulted in the deletion of the sequence in between was successfully used to edit a promoter region in tomato, leading to an increased locule number (Wang et al., 2021). A drawback of this approach is that it necessitates two suitable target sites flanking the targeted region. Furthermore, as both breaks have to be induced simultaneously, steric hindrance between the nucleases requires a distance of at least 20 bp between the breaks, which often results in deletions exceeding 40 bp. Thus, new tools are needed to enable the induction of deletions between 10 and 40 bp, preferably using a single target site.

Previously, it has been shown that mutation size and frequency can be increased by coupling site-specific endonucleases with exonucleases to promote end resection (Certo et al., 2012; Weiss et al., 2020). However, these approaches relied on a general overexpression of the exonuclease, which led to unspecific effects throughout the whole genome. Later, a fusion of Cas9 with the human 3′-5′ THREE PRIME REPAIR EXONUCLEASE 2 (HsTREX2) was used to alternate the outcome of CRISPR/Cas-mediated mutagenesis in human cells (Allen et al., 2018). While the overall mutation frequency could be increased, the observed deletions were mostly between 10 and 15 bp in length. A limiting factor could have been that only one copy of the exonuclease was present at the break site as it is usually active as a homodimer (Perrino et al., 2005). We therefore speculated that increasing the copy number of recruited HsTREX2 might result in even higher mutation frequencies and the generation of larger deletions.

RESULTS

SunTag-mediated recruitment of HsTREX2 increases mutagenesis frequency and alters repair pattern

In order to increase the amount of end-processing enzymes at the target site, we used the SunTag system which is based on multiple antibody epitopes (GCN4) that are recognized by a single chain variable fragment (scFv) (Tanenbaum et al., 2014). By fusing the epitope chain to Cas9 from *Streptococcus pyogenes* (SpCas9) and the antibody to HsTREX2, up to 10 copies of the exonuclease can be recruited to the target sequence located in the *ALCOHOL DEHYDROGENASE 1* (*ADH1*) gene (spacer ADH1-1). To reduce protein agglomeration, a sfGFP-encoding sequence as well as a GB1 solubility tag were added to the scFv exonuclease construct (Tanenbaum et al., 2014). The expression of both constructs was driven by the constitutive Ubiquitin4-2 promoter from *Petroselinum crispum* (PcUbi). In the case of the Cas9 construct, the promoter was followed by an Omega translation enhancer (Papikian et al., 2019). In order to compare our approach with the beforementioned direct fusion, we adapted it according to Allen et al. (2018) (Figure 1A). *Arabidopsis thaliana* plants were transformed with the constructs via Agrobacterium-mediated transformation and genomic DNA was extracted from batches of 30 transgenic T1 plants. Using amplicon deep sequencing, we analyzed the mutagenesis frequency as well as the repair pattern. Indeed, using the SunTag system we could observe a mutagenesis frequency of 81%, corresponding to a 2.4-fold increase compared to Cas9 alone and a twofold increase compared to the direct fusion (39%) (Figure 1B).

Taking a closer look at the mutation types, we found that DSB induction with Cas9 alone mainly resulted in small insertions (75%), consistent with previous reports (Fauser et al., 2014; Feng et al., 2014; Zhang et al., 2016). In contrast, mutations resulting from SunTag-mediated and direct recruitment of HsTREX2 were dominated by deletions making up 83% and 95% of all mutations, respectively. Additionally, we found that recruiting HsTREX2 not only yielded more deletions but also increased their size. While mutations induced by Cas9 alone showed a 2 bp deletion on average, the average mutation size resulting from DSB induction with the direct fusion of HsTREX2 were deletions of around 7 bp and SunTag-mediated HsTREX2 recruitment even resulted in the deletion of 13 bp on average (Figure 1C). In both cases, this increase was mainly due to a frequency of deletions, ranging from 1 to 25 bp.

Analysis of the position of deleted bases within a 40 bp window around the DSB showed that the PAM-distal end of the DSB seems to be more prone for exonuclease-mediated processing than the PAM-proximal end. While DSB induction using Cas9 alone resulted in similar deletion frequencies in both break ends, in the approaches including HsTREX2, bases in the PAM-distal end were disproportionally more often deleted than bases in the PAM-proximal break end (Figure 1D). Namely, 78 and 66% of deleted bases in this range were in the PAM-distal end in the case of the direct fusion and the SunTag-mediated recruitment of HsTREX2, respectively.

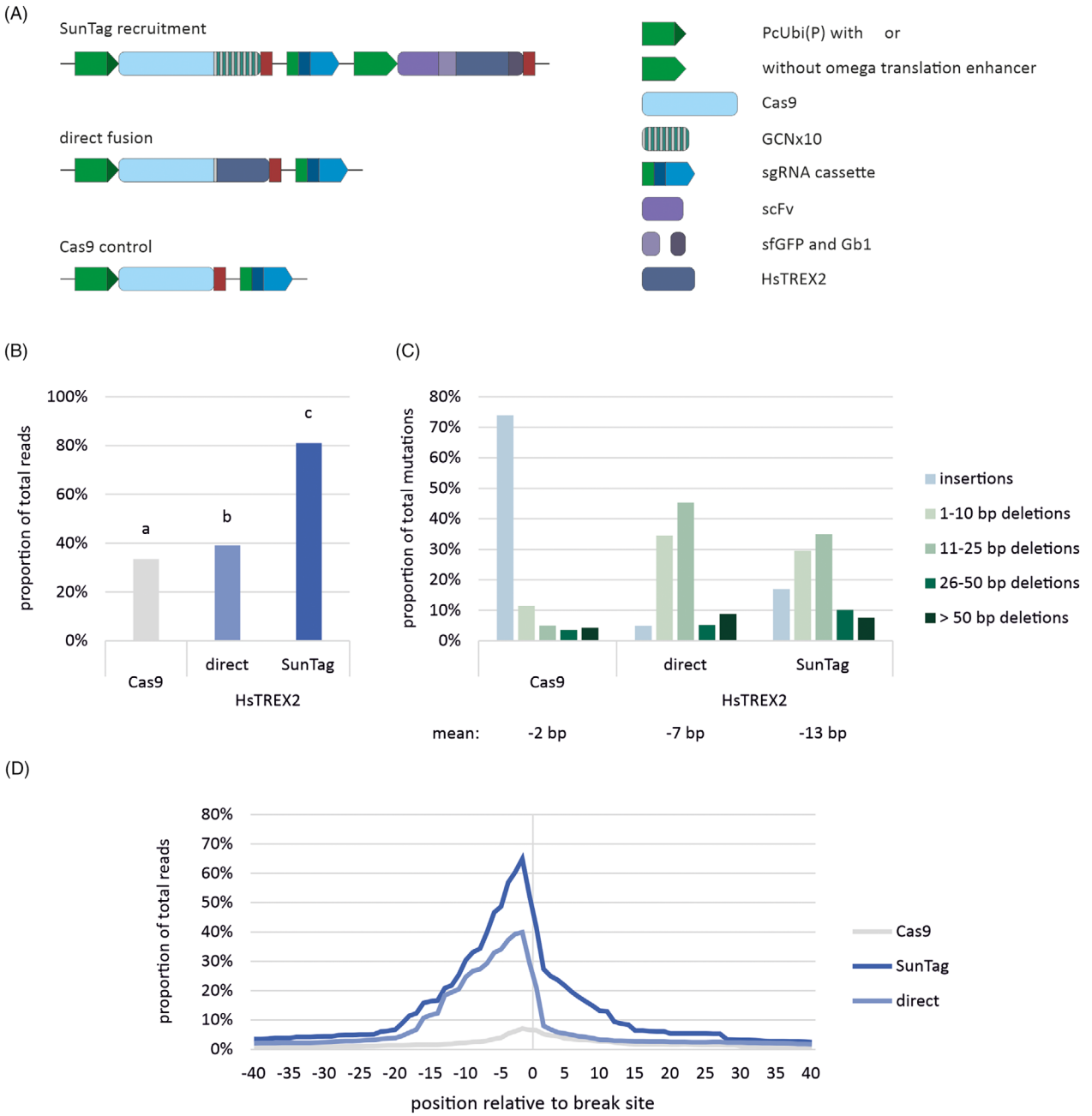

Figure 1. Construct design and repair pattern analysis.

(A) Design of constructs for SunTag-mediated HsTREX2 recruitment and direct fusion. SpCas9 was fused with a chain of 10 GCN4 epitopes, separated by a 22-amino acid (AA) GS linker. Expression was driven by the constitutive Ubiquitin4-2 promoter from *Petroselinum crispum* (PcUbi), followed by the Omega translation enhancer. HsTREX2 was fused with a single chain variable fragment (scFv). sfGFP and GB1 solubility tag were added for reduced protein agglomeration. Direct fusion of HsTREX2 was adapted from the previously described approach by Allen et al. (2018). A construct expressing only SpCas9 was used as a control.

(B) Overall frequency of altered reads in the different approaches. In the presence of HsTREX2, mutagenesis efficiency was increased. However, SunTag-mediated recruitment significantly outperformed the approach based on direct fusion. Statistical differences were calculated by Fisher's exact test with adjustment of *P* values via the Benjamini-Hochberg procedure. a $\neq$ b if $P < 0.05$, with $P < 0.0001$ in all cases. N_{Cas9} = 119 636; N_{direct} = 104 878; N_{SunTag} = 102 037.

(C) Relative frequency of mutation types in altered reads. Mutagenesis using Cas9 alone resulted mainly in small insertions while approaches including HsTREX2 predominantly yielded deletions. Additionally, deletions obtained with Cas9 alone were mostly small whereas recruitment of HsTREX2 shifted the distribution towards larger deletions. The mean size of mutations found in a sample was calculated by adding up the sizes of found mutations and dividing the sum by the number of reads with mutations. N_{Cas9} = 61 266; N_{direct} = 51 386; N_{SunTag} = 114 253.

(D) Position of deleted bases relative to the break site. With Cas9 alone, end processing was approximately symmetrical at both ends of the DSB. In the presence of HsTREX2, the balance shifted towards the PAM-distal end. The graph shows the amount of reads lacking a base at a specific position within a 40 bp window around the cutting site, relative to the total amount of reads. The break site is defined as zero with positive numbers indicating a position at the PAM-proximal break end and negative numbers indicating positions at the PAM-distal end. N_{Cas9} = 177 159; N_{direct} = 161 115; N_{SunTag} = 169 033.

The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd.,
The Plant Journal, (2023), doi: 10.1111/tpj.16586

Recruiting EcExol and HsTREX1

Encouraged by those findings, we opted to test other 3′-5′ exonucleases. We speculated that, due to higher processivity or different affinities to DNA repair intermediates, we might be able to obtain a different deletion spectrum. Therefore, we decided to test *E. coli* Exonuclease I (EcExol) as well as human TREX1 (HsTREX1). EcExol is a processive exonuclease with preference for single-strands that was shown to stimulate RecA-mediated DNA strand exchange between linear duplex and circular single-strand molecules by progressively digesting the displaced 3′-strand from the linear duplex (Bedale et al., 1993). HsTREX1 is closely related to HsTREX2 but was shown to have a higher substrate affinity with a dissociation constant 23-fold lower than that of HsTREX2 for double-stranded DNA (Mazur & Perrino, 2001). As its C-terminal domain contains an ER localization signal, we used a truncated version ($HsTREX1_{1\text{-}293}$, hereafter referred to as HsTREX1) with a reportedly higher catalytic activity (Lindahl et al., 2009). In the case of both exonucleases, we were able to observe a more than twofold increase in overall mutagenesis frequencies, with the recruitment of HsTREX1 yielding the highest amount of altered reads at 88%, which corresponds to 2.6-fold more mutations than when using Cas9 alone and, thus, even outperforming HsTREX2 recruitment. Recruitment of EcExol resulted in 74% mutagenesis (2.2-fold increase) (Figure 2A). Similar to the approach using HsTREX2, the relative amount of deletions was more than tripled in both cases, making up for 80% of all mutations induced with HsTREX1 and 77% when EcExol had been recruited. Furthermore, both nucleases outperformed HsTREX2 with respect to the size of the induced deletions. Mutations induced by SunTag-mediated recruitment of EcExol and HsTREX1 showed deletions of 15 and 20 bp on average, respectively. This is mainly due to the fact that, in contrast to the approach using HsTREX2, only the amount of small deletions ranging from 1 to 10 bp was decreased while the amount of larger deletions was generally increased (Figure 2B). Remarkably, with both, EcExol and HsTREX1, the unequal resection of the break ends was less pronounced than when recruiting HsTREX2, with 57% of the deleted bases within a 40 bp window around the break site located in the PAM-distal break end in both cases (Figure 2C). In contrast, the recruitment of exonucleases did not alter the size of the induced insertions. Similarly to the DSB induction with Cas9 alone, 1 bp insertions made up for over 90% of detected insertions in all approaches (Figure S1).

Analysis of the repair pathways

The main repair pathway for DSB in somatic plant cells is NHEJ. While cNHEJ plays a major role and accommodates for most small insertions and deletions, larger deletions are often a result of another NHEJ pathway, microhomology-mediated end-joining (MMEJ). Thus, we analyzed the frequency of deletions which had resulted from MMEJ using MH of 2 nt or more. MH were defined as the highest sequence identity between the end of the deleted sequence and the sequence of the newly built junction. Indeed, 39% of deletions found in the approach using Cas9 alone were caused by a repair using microhomologies. Nevertheless, end resection in the approaches including the exonucleases did not remarkably increase microhomology usage, which in all cases ranged between 32 and 41% (Figure 3). However, the frequency of deletions which had resulted from MMEJ using microhomologies exceeding 5 bp in length was slightly increased in the HsTREX1-based approach. On the contrary, recruitment of HsTREX2 and EcExol did not remarkably change the distribution of the length of used microhomologies.

Effects of SunTag-mediated exonuclease recruitment are target-specific

Since the SunTag system is based on an indirect fusion, the high abundance of exonucleases might result in unspecific effects on naturally occurring DSB. Thus, in order to verify the specificity of our approach, we targeted the same sequence with a Cas9 enzyme lacking the GCN4×10 epitope chain, resulting in the overexpression of non-recruitable scFv exonuclease molecules. Indeed, without SunTag-mediated recruitment of the exonuclease, in the TREX1- as well as TREX2-based approaches, mutagenesis frequencies were clearly reduced. Only in the case of EcExol, the difference in the relative frequency of the induced mutations with and without SunTag-mediated recruitment was similar (Figure 4A). However, even in this case, analysis of the repair pattern (Figure 4B) showed that the obtained mutations were on average deletions of 7 bp in length, about half the size of the ones obtained when recruiting this exonuclease. In the case of HsTREX1, the difference is comparable as the average mutation size obtained in the overexpression control amounted to about one third of the deletion size obtained when recruiting it via SunTag. Similarly, overexpression of HsTREX2 yielded deletions of only 2 bp on average, which is identical to Cas9-mediated break induction without recruitment of end-processing enzymes. This indicates that, in the case of HsTREX1 and HsTREX2, although non-recruited exonuclease constructs can act on free DNA ends, SunTag-mediated recruitment is necessary for the efficient induction of larger deletions. Thus, the SunTag approach is not only very effective in recruiting the respective enzymes to the target site but also enables the active processing of the broken ends in a target-specific manner.

Effects of SunTag-mediated HsTREX1 recruitment are target-independent

Taken together, SunTag-mediated recruitment of HsTREX1 enables a target-specific induction of larger deletions. However, the efficiency of CRISPR/Cas-mediated mutagenesis

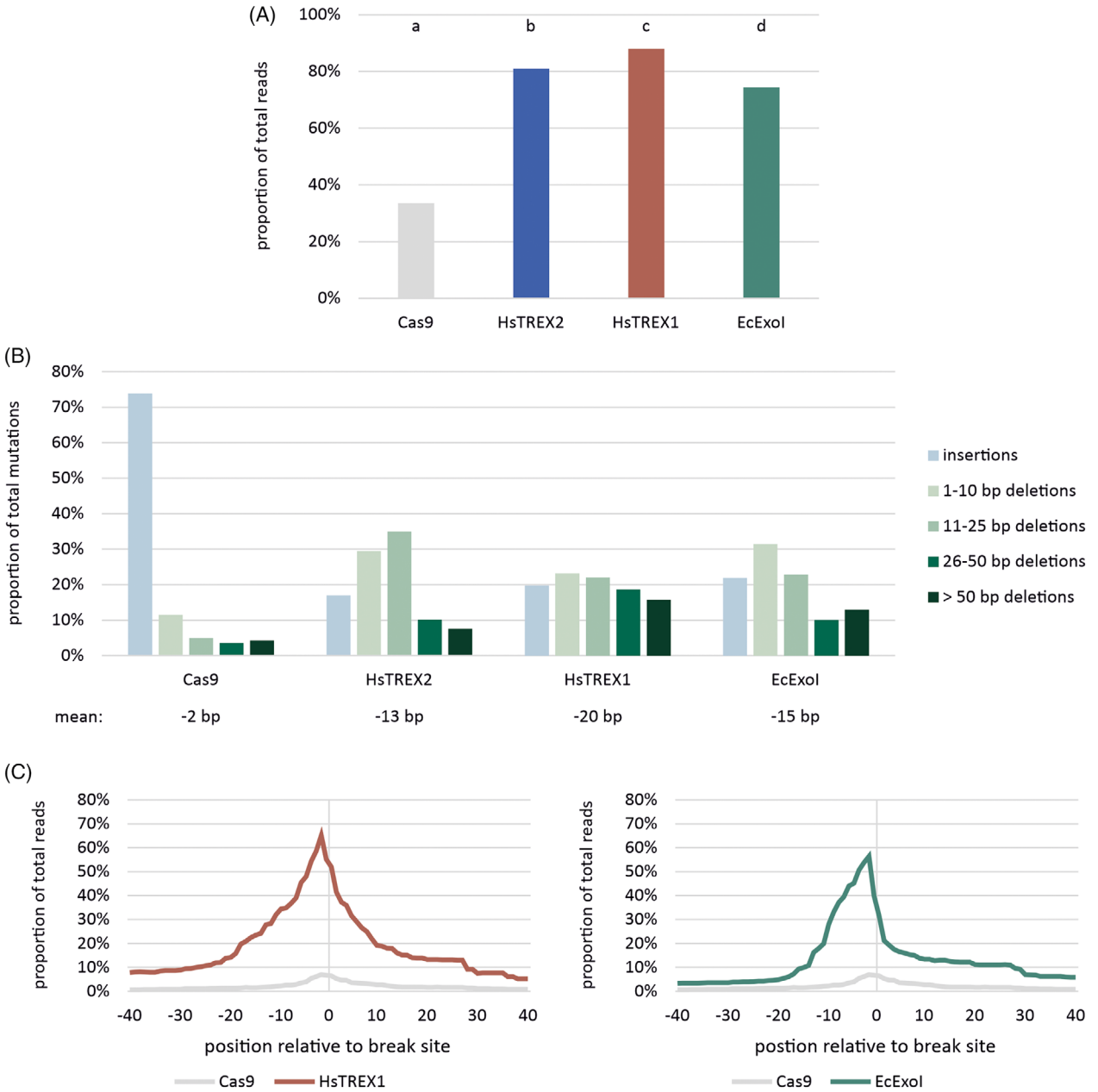

Figure 2. Repair pattern after SunTag-mediated recruitment of HsTREX1 and EcExoI.

(A) Overall frequency of altered reads. SunTag-mediated recruitment of both exonucleases not only significantly increased overall mutation frequencies compared to Cas9 alone. Recruitment of HsTREX1 even outperformed HstREX2 recruitment. but also changed the main mutation type from insertions to deletions. Recruitment of HsTREX1 even outperformed HsTREX2 in regards to mutagenesis efficiency. Statistical differences were calculated by Fisher's exact test with adjustment of P values via the Benjamini-Hochberg procedure. a $\neq$ b if $P < 0.05$, with $P < 0.0001$ in all cases. N_{Cas9} = 119 636; $N_{HsTREX2}$ = 102 037; $N_{HsTREX1}$ = 101 753; N_{EcExoI} = 113 758.

(B) Relative frequency of mutation types in altered reads. While with Cas9 alone mostly short deletions were obtained, recruiting exonucleases shifted the distribution towards larger deletions. Recruitment of HsTREX1 showed the greatest impact. The mean size of mutations found in a sample was calculated by adding up the sizes of found mutations and dividing the sum by the number of reads with mutations. N_{Cas9} = 40 111; $N_{HsTREX2}$ = 82 646; $N_{HsTREX1}$ = 89 586; N_{EcExoI} = 84 588.

(C) Position of deleted bases relative to the break site. With Cas9 alone, end processing appears to be approximately symmetrical at both ends of the DSB. In the presence of exonucleases, this balance shifted slightly towards the PAM-distal end. However, both break ends can be processed by EcExoI and HsTREX1. The graphs show the number of reads lacking a base at a specific position within a 40 bp window around the cutting site, relative to the total amount of reads. The break site is defined as zero with positive numbers indicating a position on the PAM-proximal break end and negative numbers indicating positions on the PAM-distal end. N_{Cas9} = 177 159; $N_{HsTREX1}$ = 163 316; N_{EcExoI} = 178 329.

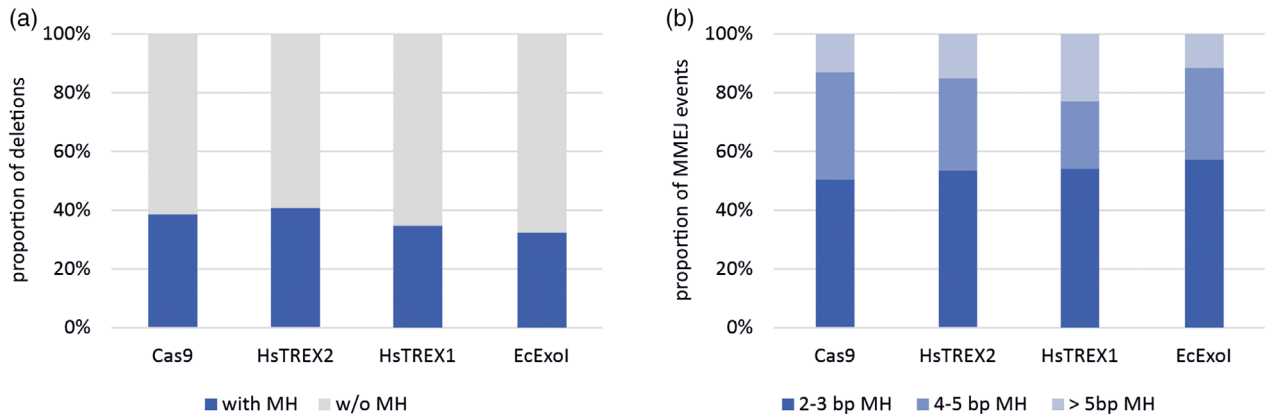

Figure 3. Microhomology-mediated repair.
(a) Amount of deletions showing repair patterns that indicate usage of microhomologies relative to the total amount of reads with deletions. No major changes in the frequency of microhomology-mediated end-joining (MMEJ) could be found in the approaches including exonucleases. N_{Cas9} = 9427; $N_{HsTREX2}$ = 66 703; $N_{HsTREX1}$ = 68 903; N_{EcExoI} = 66 010.
(b) Size of microhomologies (MH) used for repair of the induced DSB. In all cases, microhomology-mediated repair (MMEJ) was mainly based on small MH from 1 to 3 bp between the break ends. N_{Cas9} = 5808; $N_{HsTREX2}$ = 45 151; $N_{HsTREX1}$ = 44 838; N_{EcExoI} = 44 270.

was shown to be highly target-dependent. While some factors, such as the chromatin status and sequence identity of target sites, are known to affect the mutagenesis frequency, no reliable prediction of the cutting efficiencies for specific targets can be made yet (Liang et al., 2016; Liu et al., 2016; Verkuijl & Rots, 2019). Thus, we chose to test two more targets, one located in the *ADH1* gene (spacer ADH1-2) and one located in the *TRANSPARENT TESTA 4* (*TT4*) gene (spacer TT4), to confirm that the outcome of SunTag-mediated HsTREX1 recruitment to CRISPR/Cas-induced DSB is independent of the target sequence. At both targets, recruiting of HsTREX1 yielded mutagenesis efficiencies of about 95% which represents a significant increase of 1.5-fold at the *ADH1* target and 2.7-fold at the *TT4* target compared to the respective control using Cas9 alone (Figure 5a). Furthermore, similar to the first tested target, the repair pattern was changed, with the size of induced mutations averaging in deletions of 32 and 52 bp at the *ADH1* and *TT4* target, respectively. In contrast, DSB induction using Cas9 alone led to the deletion of 8 and 12 bp on average, respectively (Figure 5b). As described before, these deletions resulted from approximately equal end resection at both, the PAM-distal and -proximal end of the DSB, with 54% and 59% of all deletions being located on the PAM-distal end of the DSB in the case of the *ADH1* and the *TT4* target, respectively (Figure 5c).

DISCUSSION

The CRISPR/Cas system has revolutionized molecular biology by enabling the targeted induction of mutations. Multiple tools have been developed in the last years, enabling the induction of a wide variety of changes in the plant genome (Capdeville et al., 2023). While the repair of Cas9-induced DSB predominantly results in small insertions, Cas12a-mediated mutagenesis often leads to the generation of small deletions of about 5–10 bp (Zhang et al., 2021). Larger deletions can be obtained by simultaneously inducing two adjacent double- or single-strand breaks (Schiml et al., 2014, 2016). However, this approach is limited by PAM requirements and the requirement of a minimal distance between the breaks to avoid steric hindrance between the Cas proteins. Thus, deletions ranging from 10 to 40 bp are difficult to obtain using existing tools. Previously, it could be shown that overexpressing or fusing exonucleases to Cas endonucleases increases the size of the induced deletions (Čermák et al., 2017; Clements et al., 2017; Weiss et al., 2020; Zhang et al., 2020). However, the efficiency of the fusion-based approach might have still been limited by the fact that only one copy of the exonuclease was present at the break site. This could especially be true for HsTREX2 which is usually active as a homodimer (Perrino et al., 2005). Using the SunTag system, we were able to increase the copy number of recruited HsTREX2 and achieved higher mutation frequencies compared to a single copy recruitment via a direct fusion. Based on these findings, we set up the SunTag system to be able to recruit two other 3′-5′ exonucleases: EcExoI and HsTREX1. While EcExoI was slightly less efficient than HsTREX2 in overall mutation induction, its recruitment resulted in an enhanced frequency of larger deletions. However, our analysis also showed that, even without recruitment, EcExoI overexpression doubled mutation frequencies at a Cas9-induced DSB site. This indicates that naturally arising DSB might also be resected by EcExoI which could enhance the danger of off-site mutations at sites with no homology to the used sgRNA. In contrast, recruitment of HsTREX1 outperformed HsTREX2 recruitment, not only in the overall mutagenesis efficiency but also concerning the size of the induced

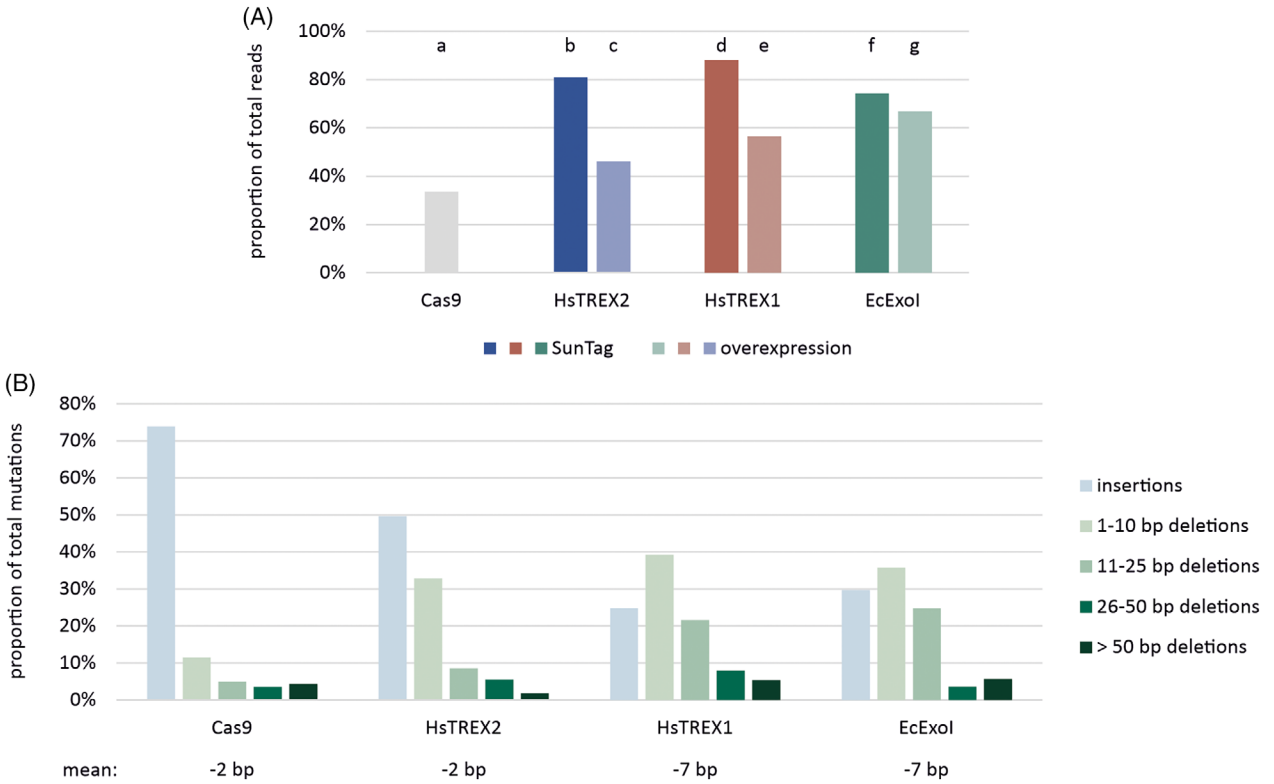

Figure 4. Repair pattern in cells overexpressing exonucleases.
Overall frequency of altered reads (A) and main mutation types (B).
(A) While increased mutagenesis frequencies could be observed when exonucleases were expressed in the cell, the effect was significantly less pronounced compared to approaches with SunTag-mediated recruitment. Only in the case of EcExoI, mutagenesis frequencies were similar when recruited or overexpressed. Statistical differences were calculated by Fisher's exact test with adjustment of P values via the Benjamini-Hochberg procedure. a $\neq$ b if $P < 0.05$. N_{Cas9} = 119 636; HsTREX2: N_{SunTag} = 102 037; $N_{Overexpression}$ = 117 835; HsTREX1: N_{SunTag} = 101 753; $N_{Overexpression}$ = 55 351; EcExoI: N_{SunTag} = 113 758; $N_{Overexpression}$ = 69 311.
(B) However, in all cases the repair pattern showed only a slight shift towards small deletions compared to break induction with Cas9 alone. The mean size of mutations found in a sample was calculated by adding up the sizes of found mutations and dividing the sum by the number of reads with mutations. N_{Cas9} = 40 111; $N_{HsTREX2}$ = 54 292; $N_{HsTREX1}$ = 31 245; N_{EcExoI} = 46 283.

deletions. Furthermore, we could show that these effects are target-independent and limited to DSB to which the exonucleases are recruited.

Interestingly, we found that, in the case of HsTREX2, the PAM-distal end is more prone to exonuclease-mediated processing than the PAM-proximal break end. This could be due to the characteristic of SpCas9 to stay bound to the DNA for several hours after break induction, possibly protecting the PAM-proximal end from end-processing enzymes (Ma et al., 2016; Zhang et al., 2019). In contrast, both, EcExoI and HsTREX1, were able to also resect the PAM-proximal end of the DSB which, in the case of HsTREX1, could be due to an increased substrate affinity compared to HsTREX2, thus displacing Cas9 from the DNA (Mazur & Perrino, 2001). Similarly, EcExoI could remove Cas9 as it encloses the DNA to achieve processive degradation (Breyer & Matthews, 2000).

In order to repair these generated 5′-overhangs, either the truncated 3′-ends need to be elongated by a polymerase or the 5′-ends have to be resected in order to generate blunt ends. As in the presence of all three exonucleases, the repair pattern is shifted massively towards deletions, repair of DSB with 5′-overhangs via NHEJ is much more probable to happen by resection of the overhangs rather than by elongation of the shortened 3′-ends. However, as no proof of a major increase in microhomology-mediated repair was found, the formation of longer 3′-overhangs is unlikely to happen as their presence would enable hybridization via microhomologies which are larger or located further away from the break site. Instead, it is probable that once the elimination of the overhang yields a blunt-ended DSB, it is repaired via cNHEJ-mediated ligation. The fact that, in the case of HsTREX1, MMEJ as well as the size of the used MHs was slightly enhanced might be simply due to its more pronounced end-resection which might make longer microhomologies, located further from the break site, more accessible for hybridization.

Taken together, we have developed a new tool for a target-specific induction of larger deletions using a single guide RNA. We could show that SunTag-mediated

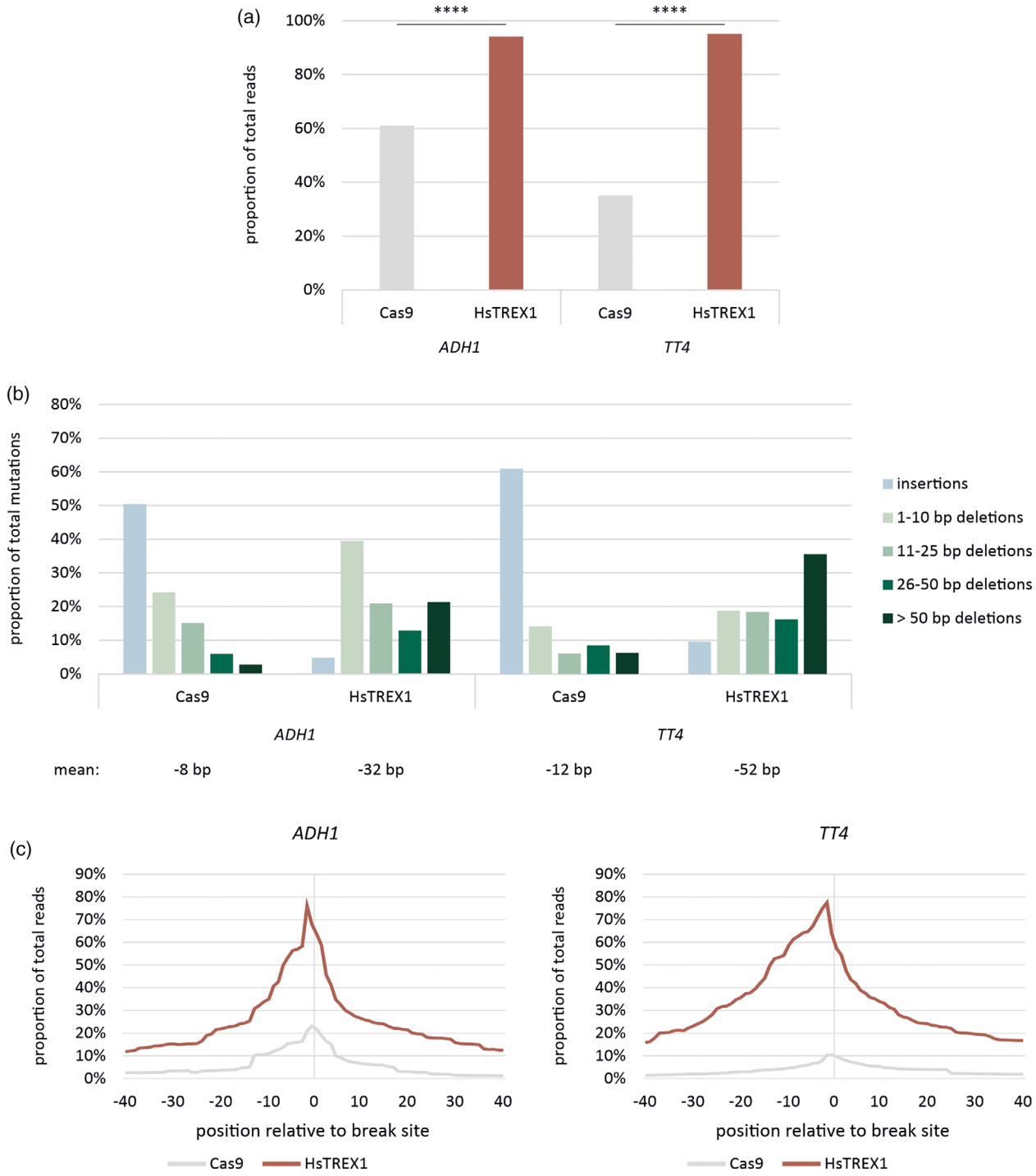

Figure 5. Validation of target-independent effects of SunTag-mediated HsTREX1 recruitment.
(a) Overall frequency of altered reads. SunTag-mediated recruitment of HsTREX1 significantly increased the frequency of altered reads compared to the respective Cas9 control at both targets, yielding around 95% mutagenesis efficiency in both cases. Statistical differences were calculated by Fisher's exact test. ****: $P < 0.0001$. ADH1: N_{Cas9} = 22 849; $N_{HsTREX1}$ = 24 882; TT4: N_{Cas9} = 29 535; $N_{HsTREX1}$ = 46 792.
(b) Relative frequency of mutation types in altered reads. While with Cas9 alone, mostly insertions and short deletions were obtained, recruiting HsTREX1 shifted the distribution towards larger deletions. The mean size of mutations found in a sample was calculated by adding up the sizes of found mutations and dividing the sum by the number of reads with mutations. ADH1: N_{Cas9} = 13 946; $N_{HsTREX1}$ = 23 427; TT4: N_{Cas9} = 10 355; $N_{HsTREX1}$ = 44 530.
(c) Position of deleted bases relative to the break site. Both break ends can be processed approximately to the same degree by HsTREX1. The graphs show the number of reads lacking a base at a specific position within a 40 bp window around the break site, relative to the total amount of reads with deletions. The break site is defined as zero with positive numbers indicating a position at the PAM-proximal break end and negative numbers indicating positions at the PAM-distal end. ADH1: N_{Cas9} = 53 798; $N_{HsTREX1}$ = 62 890; TT4: N_{Cas9} = 73 502; $N_{HsTREX1}$ = 97 001.

recruitment of up to 10 copies of the 3′-5′ exonuclease HsTREX1 outperforms previous attempts based on a direct fusion of exonucleases with no considerable effects on DSB to which it is not recruited. Even more importantly, we could show that, by SunTag-mediated recruitment of HsTREX1, deletions in the range of 10–40 bp can be obtained at different genomic sites with high efficiency.

Thus, this system might be especially suitable for targeting non-coding DNA, such as cis-regulatory elements (CRE). These sequences are key factors in the complex patterns of gene expression regulation and, thus, of utmost interest for trait improvement in crops (Rodríguez-Leal et al., 2017; Schmitz et al., 2022; Wang et al., 2021). Small InDels generated by Cas9 alone show very limited effects in these regions. On the contrary, the simultaneous induction of two DSB necessitates two target sites, limiting the size of inducible deletions to the distance between the identified PAM sequences which often leads to the generation of much larger deletions. In contrast, our system provides the possibility to precisely induce a set of different alleles containing medium-sized deletions which could enable an inactivation of individual cis-regulatory motifs that often don't exceed 50 bp in length (Korkuc et al., 2014).

By combining our SunTag-HsTREX1 approach with other natural or modified Cas enzymes, the number of putative target sites could be massively increased, making deletions of a few dozen bp almost anywhere in eukaryotic genomes achievable in the future.

EXPERIMENTAL PROCEDURES

T-DNA constructs

Coding sequences of HsTREX1, HsTREX2 and EcExoI were obtained in codon-optimized form for Arabidopsis from BioCat GmbH (Heidelberg, Germany) (Supplemental Data S1). Similarly, an entry-vector based on pEn-Sp-Chimera (Fauser et al., 2014) with an additional expression cassette, encoding scFv-sfGFP-GB1 under the control of constitutive Ubiquitin4-2 promoter from *Petroselinum crispum* (PcUbi) (pEn-SpsgRNA+scFv), was synthesized by BioCat GmbH (Heidelberg, Germany) (Supplemental Data S2). Integration of exonuclease coding sequences in between sfGFP- and GB1-encoding sequences was performed via BamHI digestion and subsequent ligation. For the Cas9 reference approach without an exonuclease, the original pEn-Sp-Chimera was used. Spacer integration was performed as described previously (Fauser et al., 2014).

T-DNA constructs used in this study are based on the previously described pDe-SpCas9 plasmid (Fauser et al., 2014). A plasmid containing the Omega enhancer as well as the GCN4 epitope chain sequences was obtained from BioCat GmbH (Heidelberg, Germany). For subsequent Gibson assembly of the final SunTag destination vector (pDe-Omega-SpCas9-GCN4 + PPTR), these sequences were designed with appropriate overhangs and flanked by MlyI sites. First, the PcUbi promoter was excised from pDe-SpCas9 using EcoRI. Together with the Omega enhancer sequence, it was then reintegrated using an amplified version with specific overhangs via Gibson assembly, resulting in pDe-Omega-SpCas9 + PPTR, used for reference and overexpression constructs. For the addition of the GCN4 epitope chain sequence, the SpCas9 coding sequence was excised from pDe-Omega-SpCas9 + PPTR with AscI. Gibson assembly of pDe-Omega-SpCas9-GCN4 + PPTR (Supplemental Data S2) was performed together with the coding sequence of SpCas9 which was amplified from pDe-SpCas9 with primers encoding homologous flanks. Similarly, cloning of the HsTREX2 direct fusion construct (pDe-Omega-SpCas9-HsTREX2 + PPTR, Supplemental data S2) was performed via Gibson assembly using amplified versions of SpCas9 and HsTREX2 coding sequences containing appropriate overhangs. The overhangs further contained the coding sequence for an 8-aa GS linker.

Assembly of the final expression vectors (Supplemental data S3), containing the T-DNA with a Cas9 cassette as well as the sgRNA cassette, the scFv cassette and a phosphinothricin (PPT) resistance cassette as selection marker, was performed via a Gateway LR reaction.

Plant transformation and cultivation

Arabidopsis plants were transformed via floral dip using the Agrobacterium strain GV3101 as previously described (Clough & Bent, 1998). T1 plants were grown on a germination medium with cefotaxime (500 mg/l) and phosphinothricin (6 mg/l) as selection markers. DNA was extracted from batches of 30 plants after 2 weeks of cultivation as described previously (Fulton et al., 1995).

Amplicon deep sequencing and data analysis

Amplicons with universal adapter sequences were generated with 30 ng of genomic DNA using a proofreading polymerase. Sequences of the used oligonucleotides are listed in Table S3. Illumina paired-end sequencing was performed by Eurofins Genomics Germany GmbH. Data analysis was performed using CRISPResso2 (Clement et al., 2019) with default settings, CRISPR RGEN Tools Cas-Analyzer (comparison range (*R*) = use both ends, minimum frequency (*n*) = 1 and no WT marker (*r*)) (Park et al., 2017), custom *R* and Linux shell scripts and Microsoft Office Excel. To identify MH, they were defined as the maximum identity between the sequence of deleted bases and the bases flanking the deletion.

AUTHOR CONTRIBUTIONS

HP conceived the research. NC and PS designed and executed the experiments. NC and HP wrote the article.

ACKNOWLEDGEMENTS

We thank Sebastian Scheiner for his help with the cloning of the constructs and Ricky Koch and Niklas Bucherer for their excellent technical assistance. We thank Michelle Rönspies for the critical reading of the manuscript. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Supplemental Data S1. Sequences of synthesized codon-optimized exonucleases coding sequences.

Supplemental Data S2. Sequences of entry vector and fusion constructs.

Supplemental Data S3. List of used oligonucleotides for NGS amplicon generation.

Figure S1. Relative frequency of insertion sizes after SunTag-mediated recruitment of exonucleases.

The recruitment of exonuclease did not alter the size of induced insertions. Similarly to the DSB induction with Cas9 alone, 1 bp insertions made up for over 90% of found insertions in all approaches. N_{Cas9} = 45 279; $N_{HsTREX2}$ = 19 417; $N_{HsTREX1}$ = 24 601; N_{EcExoI} = 22 620.

OPEN RESEARCH BADGES

This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at Next generation sequencing data have been deposited in the Sequence Read Achive (SRA, https://www.ncbi.ncbi.nlm.nih.gov/sra) under the accession number PRJNA1040112 (https://www.ncbi.ncbi.nlm.nih.gov/bioproject/PRJNA1040112).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request. Next-generation sequencing data that support the findings of this study have been deposited in SRA under the accession number PRJNA1040112.

REFERENCES

Allen, F., **Crepaldi, L.**, **Alsinet, C.**, **Strong, A.J.**, **Kleshchevnikov, V.**, **de Angeli, P.** ***et al.*** (2018) Predicting the mutations generated by repair of Cas9-induced double-strand breaks. *Nature Biotechnology*, **37**, 64–72.

Bedale, W.A., **Inman, R.B.** & **Cox, M.M.** (1993) A reverse DNA strand exchange mediated by recA protein and exonuclease I. The generation of apparent DNA strand breaks by recA protein is explained. *The Journal of Biological Chemistry*, **268**, 15004–15016.

Breyer, W.A. & **Matthews, B.W.** (2000) Structure of Escherichia coli exonuclease I suggests how processivity is achieved. *Nature Structural Biology*, **7**, 1125–1128.

Capdeville, N., **Schindele, P.** & **Puchta, H.** (2023) Getting better all the time — recent progress in the development of CRISPR/Cas-based tools for plant genome engineering. *Current Opinion in Biotechnology*, **79**, 102854.

Čermák, T., **Curtin, S.J.**, **Gil-Humanes, J.**, **Čegan, R.**, **Kono, T.J.Y.**, **Konečná, E.** ***et al.*** (2017) A multipurpose toolkit to enable advanced genome engineering in plants. *The Plant Cell*, **29**, 1196–1217.

Certo, M.T., **Gwiazda, K.S.**, **Kuhar, R.**, **Sather, B.**, **Curinga, G.**, **Mandt, T.** ***et al.*** (2012) Coupling endonucleases with DNA end-processing enzymes to drive gene disruption. *Nature Methods*, **9**, 973–975.

Clement, K., **Rees, H.**, **Canver, M.C.**, **Gehrke, J.M.**, **Farouni, R.**, **Hsu, J.Y.** ***et al.*** (2019) CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nature Biotechnology*, **37**, 224–226.

Clements, T.P., **Tandon, B.**, **Lintel, H.A.**, **McCarty, J.H.** & **Wagner, D.S.** (2017) RICE CRISPR: Rapidly increased cut ends by an exonuclease Cas9 fusion in zebrafish. *Genesis (New York, N.Y: 2000)*, **55**, Article e23044. Available from: https://doi.org/10.1002/dvg.23044

Clough, S.J. & **Bent, A.F.** (1998) Floral dip: a simplified method for agrobacterium-mediated transformation of Arabidopsis thaliana. *The Plant Journal: For Cell and Molecular Biology*, **16**, 735–743.

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: for Cell and Molecular Biology*, **79**, 348–359.

Feng, Z., **Mao, Y.**, **Xu, N.**, **Zhang, B.**, **Wei, P.**, **Yang, D.-L.** ***et al.*** (2014) Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 4632–4637.

Fulton, T.M., **Chunwongse, J.** & **Tanksley, S.D.** (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Molecular Biology Reporter*, **13**, 207–209.

Jinek, M., **Chylinski, K.**, **Fonfara, I.**, **Hauer, M.**, **Doudna, J.A.** & **Charpentier, E.** (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, N.Y.)*, **337**, 816–821.

Korkuc, P., **Schippers, J.H.M.** & **Walther, D.** (2014) Characterization and identification of cis-regulatory elements in Arabidopsis based on single-nucleotide polymorphism information. *Plant Physiology*, **164**, 181–200.

Lemos, B.R., **Kaplan, A.C.**, **Bae, J.E.**, **Ferrazzoli, A.E.**, **Kuo, J.**, **Anand, R.P.** ***et al.*** (2018) CRISPR/Cas9 cleavages in budding yeast reveal templated insertions and strand-specific insertion/deletion profiles. *Proceedings of the National Academy of Sciences of the United States of America*, **115**, E2040–E2047.

Liang, G., **Zhang, H.**, **Lou, D.** & **Yu, D.** (2016) Selection of highly efficient sgRNAs for CRISPR/Cas9-based plant genome editing. *Scientific Reports*, **6**, 21451.

Lindahl, T., **Barnes, D.E.**, **Yang, Y.-G.** & **Robins, P.** (2009) Biochemical properties of mammalian TREX1 and its association with DNA replication and inherited inflammatory disease. *Biochemical Society Transactions*, **37**, 535–538.

Liu, X., **Homma, A.**, **Sayadi, J.**, **Yang, S.**, **Ohashi, J.** & **Takumi, T.** (2016) Sequence features associated with the cleavage efficiency of CRISPR/Cas9 system. *Scientific Reports*, **6**, 19675.

Ma, H., **Tu, L.-C.**, **Naseri, A.**, **Huisman, M.**, **Zhang, S.**, **Grunwald, D.** ***et al.*** (2016) CRISPR-Cas9 nuclear dynamics and target recognition in living cells. *The Journal of Cell Biology*, **214**, 529–537.

Mazur, D.J. & **Perrino, F.W.** (2001) Excision of 3′ termini by the Trex1 and TREX2 3′→5′ exonucleases. *The Journal of Biological Chemistry*, **276**, 17022–17029.

Pacher, M. & **Puchta, H.** (2017) From classical mutagenesis to nuclease-based breeding - directing natural DNA repair for a natural end-product. *The Plant Journal: For Cell and Molecular Biology*, **90**, 819–833.

Papikian, A., **Liu, W.**, **Gallego-Bartolomé, J.** & **Jacobsen, S.E.** (2019) Site-specific manipulation of Arabidopsis loci using CRISPR-Cas9 SunTag systems. *Nature Communications*, **10**, 729.

Park, J., **Lim, K.**, **Kim, J.-S.** & **Bae, S.** (2017) Cas-analyzer: an online tool for assessing genome editing results using NGS data. *Bioinformatics (Oxford, England)*, **33**, 286–288.

Perrino, F.W., **Harvey, S.**, **McMillin, S.** & **Hollis, T.** (2005) The human TREX2 3′ - 5′-exonuclease structure suggests a mechanism for efficient nonprocessive DNA catalysis. *The Journal of Biological Chemistry*, **280**, 15212–15218.

Puchta, H. (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *Journal of Experimental Botany*, **56**, 1–14.

Rodríguez-Leal, D., **Lemmon, Z.H.**, **Man, J.**, **Bartlett, M.E.** & **Lippman, Z.B.** (2017) Engineering quantitative trait variation for crop improvement by genome editing. *Cell*, **171**, 470–480.e8.

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: For Cell and Molecular Biology*, **80**, 1139–1150.

Schiml, S., **Fauser, F.** & **Puchta, H.** (2016) Repair of adjacent single-strand breaks is often accompanied by the formation of tandem sequence duplications in plant genomes. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, 7266–7271.

Schmitz, R.J., **Grotewold, E.** & **Stam, M.** (2022) Cis-regulatory sequences in plants: their importance, discovery, and future challenges. *The Plant Cell*, **34**, 718–741.

Swinnen, G., **Goossens, A.** & **Pauwels, L.** (2019) Lessons from domestication: targeting cis-regulatory elements for crop improvement. *Trends in Plant Science*, **24**, 1065.

Tanenbaum, M.E., **Gilbert, L.A.**, **Qi, L.S.**, **Weissman, J.S.** & **Vale, R.D.** (2014) A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*, **159**, 635–646.

Verkuijl, S.A. & **Rots, M.G.** (2019) The influence of eukaryotic chromatin state on CRISPR-Cas9 editing efficiencies. *Current Opinion in Biotechnology*, **55**, 68–73.

Wang, X., **Aguirre, L.**, **Rodríguez-Leal, D.**, **Hendelman, A.**, **Benoit, M.** & **Lippman, Z.B.** (2021) Dissecting cis-regulatory control of quantitative trait variation in a plant stem cell circuit. *Nature Plants*, **7**, 419–427.

Weiss, T., **Wang, C.**, **Kang, X.**, **Zhao, H.**, **Elena Gamo, M.**, **Starker, C.G.** ***et al.*** (2020) Optimization of multiplexed CRISPR/Cas9 system for highly efficient genome editing in Setaria viridis. *The Plant Journal: For Cell and Molecular Biology*, **104**, 828–838.

Yang, H., **Wu, J.-J.**, **Tang, T.**, **Liu, K.-D.** & **Dai, C.** (2017) CRISPR/Cas9-mediated genome editing efficiently creates specific mutations at multiple loci using one sgRNA in Brassica napus. *Scientific Reports*, **7**, 7489.

Zhang, B., **Yang, X.**, **Yang, C.**, **Li, M.** & **Guo, Y.** (2016) Exploiting the CRISPR/Cas9 system for targeted genome mutagenesis in petunia. *Scientific Reports*, **6**, 20315.

Zhang, Q., **Wen, F.**, **Zhang, S.**, **Jin, J.**, **Bi, L.**, **Lu, Y.** ***et al.*** (2019) The post-PAM interaction of RNA-guided spCas9 with DNA dictates its target binding and dissociation. *Science Advances*, **5**, eaaw9807.

Zhang, Q., **Yin, K.**, **Liu, G.**, **Li, S.**, **Li, M.** & **Qiu, J.-L.** (2020) Fusing T5 exonuclease with Cas9 and Cas12a increases the frequency and size of deletion at target sites. *Science China. Life Sciences*, **63**, 1918–1927.

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