RESEARCH ARTICLE

Optimization of T-DNA architecture for Cas9-mediated mutagenesis in Arabidopsis

Baptiste Castel[1], Laurence Tomlinson[1], Federica Locci[1,2], Ying Yang[1,3], Jonathan D. G. Jones[1]*

1 The Sainsbury Laboratory, Norwich Research Park, Norwich, United Kingdom, 2 Department of Biology and Biotechnology Charles Darwin, Sapienza University of Rome, Rome, Italy, 3 Center for Plant Science Innovation, Beadle Center, University of Lincoln-Nebraska, Lincoln, Nebraska, United States of America

* jonathan.jones@tsl.ac.uk

Abstract

Bacterial CRISPR systems have been widely adopted to create operator-specified site-specific nucleases. Such nuclease action commonly results in loss-of-function alleles, facilitating functional analysis of genes and gene families We conducted a systematic comparison of components and T-DNA architectures for CRISPR-mediated gene editing in Arabidopsis, testing multiple promoters, terminators, sgRNA backbones and *Cas9* alleles. We identified a T-DNA architecture that usually results in stable (*i.e.* homozygous) mutations in the first generation after transformation. Notably, the transcription of sgRNA and Cas9 in head-to-head divergent orientation usually resulted in highly active lines. Our Arabidopsis data may prove useful for optimization of CRISPR methods in other plants.

Introduction

CRISPR (clustered regularly interspaced short palindromic repeat)-Cas (CRISPR associated) site-specific nucleases evolved as components of prokaryotic immunity against viruses, and are widely deployed as tools to impose operator-specified nucleotide sequence changes in genomes of interest [1–4]. During infection by bacteriophages, Cas1 and Cas2 can integrate phage DNA sequences into 'spacer' regions of tandem CRISPR loci in the bacterial genome. The crRNA (CRISPR-RNA) transcription product of the spacer associates with nucleases from the Cas family to form ribonucleoproteins that can cleave nucleic acid sequences homologous to the spacer. This enables elimination of viral nucleic acid upon subsequent infection. CRISPR systems are divided in two classes [5,6]. Class 1 systems comprise multi-subunit complexes whereas Class 2 systems function with single ribonucleoproteins. Within Class 2, Type-II and Type-V cleave dsDNA (double-stranded DNA) via Cas9 and Cas12/Cpf1 respectively, while Type-VI cleaves ssRNA (single-stranded RNA) via Cas13/C2c2.

Cas9, Cas12 and Cas13-based systems function in heterologous organisms enabling applications such as targeted mutagenesis, dynamic imaging of genomic loci, transcriptional regulation, pathogen detection and RNA quantification [7–9]. Expression of Cas9 with its associated sgRNA (single-guide RNA, an artificial fusion of the dual endogenous crisprRNA/trans-acting-crisprRNA), results in targeted DNA mutations in animals and plants [3,10,11].

Cas9-sgRNA ribonucleoprotein cleaves genomic DNA at loci homologous to the sgRNA spacer sequence. Cleaved DNA strands can be religated by the endogenous Non-Homologous End Joining (NHEJ) system, which can result in insertions or deletions (indels) at the repaired site. Indels in the CDS (coding DNA sequence) can cause a codon reading frame shift resulting in loss-of-function alleles.

Arabidopsis thaliana (Arabidopsis) is widely used for plant molecular genetics. Expression of CRISPR-Cas9 components can result in loss-of-function alleles of targeted genes in Arabidopsis, with variable efficiency [12–14]. To improve induced mutation rates in Arabidopsis, several groups have evaluated various promoters to drive *Cas9* expression. [15–17].

We set out to optimize mutation rates in Arabidopsis, and report here an extensive comparison of promoters, *Cas9* alleles, terminator, sgRNA and construct architecture. Cas9-sgRNA ribonucleoprotein can be directly delivered by protoplast transformation or particle bombardment into plant cells [18,19], but these methods require regeneration via tissue culture. To avoid this process, we delivered *Cas9* and the *sgRNA* in transgenic Arabidopsis. This method requires three steps: (i) DNA assembly of a binary vector with selectable marker, a *Cas9* and a *sgRNA* expression cassettes in the T-DNA, (ii) *Agrobacterium tumefaciens*-mediated stable transformation of the plasmid via the floral dip method [20] and (iii) identification of mutants among the transformed lines. Multiple T-DNA architectures were tested for their ability to trigger homozygous mutations in the *ADH1* gene, including presence or absence of an "*overdrive*" sequence to promote T-DNA transfer [21]. ADH1 converts allyl alcohol into lethal allyl aldehyde, so *adh1* mutant lines resist allyl-alcohol treatment, enabling facile measurement of CRISPR-induced mutation rates [13,16]. We defined combinations of CRISPR components that enable high efficiency recovery of stable homozygous mutants in one generation.

Results

Golden Gate cloning enables facile assembly of diverse Cas9 T-DNA architectures

In Golden Gate modular cloning, the promoter, reading frame and 3' end modules at 'Level 0', are assembled using Type IIS restriction enzymes to 'Level 1' complete genes, that can then be easily combined into T-DNAs carrying multiple genes at 'Level 2'. This enables facile assembly of diverse T-DNA conformations [22,23]. Level 0 acceptor vectors are designed to clone promoter, coding sequence (CDS) or terminator fragments (see Materials and methods). For our purpose, we used three Level 1 vectors: a glufosinate plant selectable marker in position 1 (pICSL11017, cloned into pICH47732), a *Cas9* expression cassette in position 2 (cloned into pICH47742) and a *sgRNA* expression cassette in position 3 (cloned into pICH47751) (Fig 1). Some *Cas9* expression cassettes were cloned into a Level 1 position 2 variant: pICH47811. This vector can be assembled in Level 2 in the same fashion as pICH47742, but it enables *Cas9* transcription in the opposite direction as compared to the other Level 1 modules. We assembled 25 different Level 1 *Cas9* constructs and four *sgRNA* expression cassettes. The sequence targeted by the sgRNA was CGTATCTTCGGCCATGAAGC (*NGG*) (Protospacer Adjacent Motif indicated in italics) which targets specifically *ADH1* in Col-0, enabling pre-selection of CRISPR-induced *adh1* mutants by selecting with allyl alcohol [13]. Assembly of these Level 1 modules resulted in 39 Level 2 T-DNA vectors (S1 Table). More details of the assembly protocols can be found in the 'Materials and Methods' section.
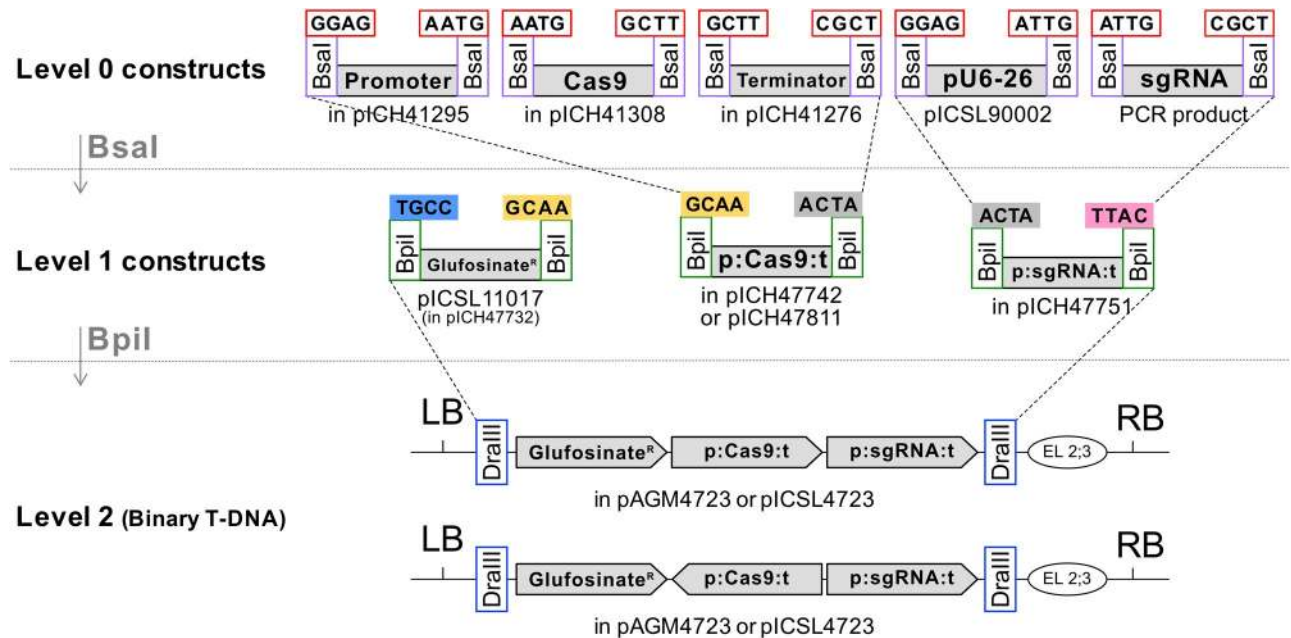
Fig 1. Golden Gate cloning method enables assembly of CRISPR modules in various combinations. *Cas9* alleles, promoters and terminators were cloned into the indicated Level 0 acceptor vectors as described in *Materials and Methods* and were assembled in Level 1 acceptor vector pICH47742. sgRNAs targeting *AtADH1* were amplified by PCR and assembled with the *U6-26* promoter vector pICSL90002 in the same manner. Both Cas9 and sgRNA expression units were assembled in Level 2 acceptors pAGM4723 (not containing an *overdrive* sequence) or pICSL4723 (containing an *overdrive*) along with a Glufosinate resistance plant selectable marker. An end-linker pICH41766 (EL2;3) was used to link the sgRNA expression unit to the Level 2 acceptor vector. For a "head-to-head orientation" of the sgRNA and Cas9 expression cassettes, Cas9 allele, promoter and terminator were assembled into pICH47811 instead of pICH47742.

https://doi.org/10.1371/journal.pone.0204778.g001

CRISPR-induced Arabidopsis mutations can be selected using allyl-alcohol

The 39 Level 2 plasmids were transformed in *A. tumefaciens* strain GV3101 and used to generate Arabidopsis Col-0 transgenic lines. 'T1' refers to independent primary transformants selected from the seeds of the dipped plant; 'T2' refers to the T1 progeny. For each of the 39 constructs, about 100 T2 progenies from six independent T1 lines were screened for allyl alcohol resistance ([Fig 2](#)). T2 seeds were selected with 30 mM allyl-alcohol for two hours. Six survivors (or all survivors if there were less than six) were screened by PCR amplification and capillary sequencing to confirm the mutation in *ADH1* at the expected target site. This genotyping step enabled us to estimate the percentage of non-mutated plants that escape the allyl-alcohol selection. We indeed identified some lines surviving the allyl-alcohol screen that are heterozygous (*ADH1/adh1)*. CRISPR activity is expressed as [(number of allyl-alcohol surviving plants) x (% of homozygous or biallelic mutants confirmed by sequencing among the surviving plants tested) / (number of seeds sown)]. It was measured for six independent T2 families, for each of 39 constructs. When more than 75% of the lines survived the allyl-alcohol treatment and all the lines genotyped are knock-out (KO) alleles with the exact same mutation within one T2 family, we assumed that the T1 parent was a homozygous mutant. Such T2 families are indicated in red.

UBI10, *YAO* and *RPS5a* promoter-controlled *Cas9* expression enhance mutation rates

CRISPR-mediated DNA sequence changes are only inherited if they occur in the germline. The Cauliflower Mosaic Virus 35S promoter and ubiquitin promoters are strongly expressed
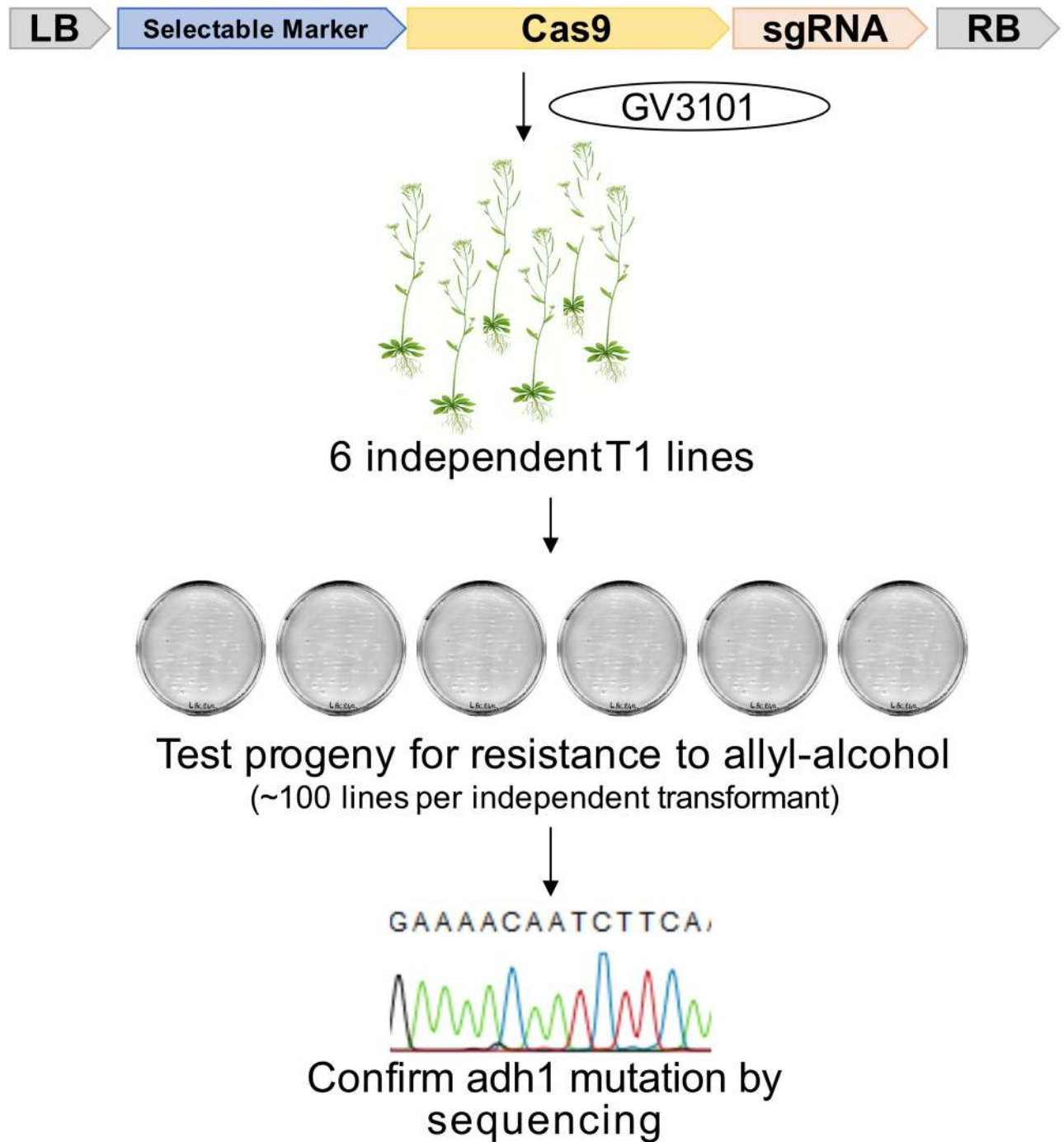
Fig 2. Evaluation of mutation rates. Constructs were transformed into Arabidopsis accession Col-0 via *Agrobacterium tumefaciens* strain GV3101. Six independent transformants (T1) were selected using Glufosinate. About 100 progeny (T2) of each transformant were selected for allyl-alcohol resistance. For each independent T2 family, up to six allyl-alcohol resistant plants were genotyped at the *ADH1* locus. For each T2 family, the mutation rate was calculated as [(number of allyl-alcohol surviving plants) x (% of homozygous or biallelic mutants confirmed by sequencing among the surviving plants tested) / (number of seeds sown)].

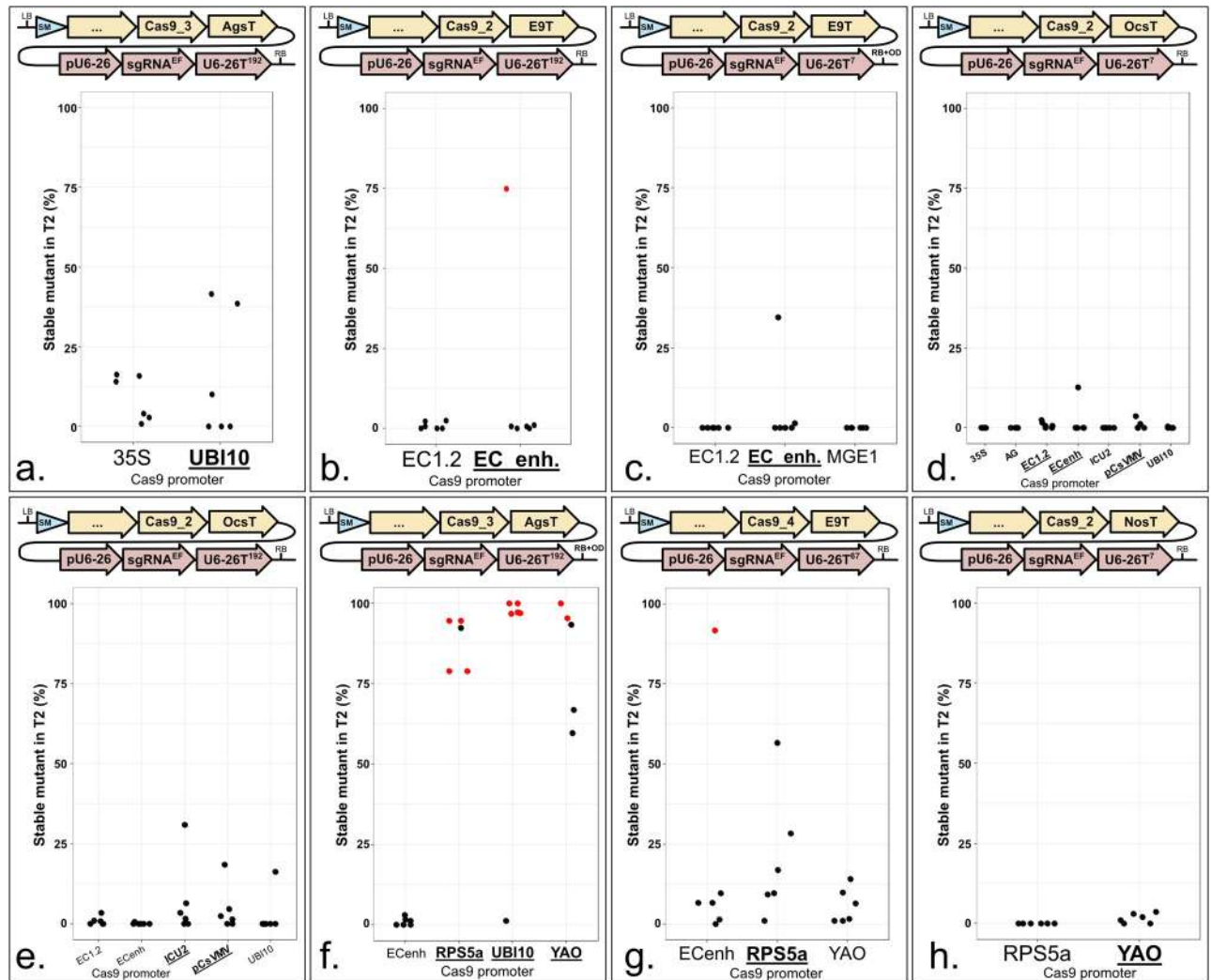**Fig 3. *UBI10*, *YAO* and *RPS5a* promoter-regulated *Cas9* expression enhances mutation rates. a. to h.** Each panel represents a promoter comparison in the same T-DNA context. Promoters can be compared within each panel, but not from one panel to another. The modules were assembled into pICSL4723 (RB+OD, with an *overdrive*) or pAGM4723 (RB, without an *overdrive*) and transformed into Col-0 via *Agrobacterium tumefaciens* strain GV3101. LB: Left Border. SM: Selectable Marker (*Glufosinate resistance* gene). 35S: 426 bp of the 35S promoter from *Cauliflower Mosaic Virus*. UBI10: 1327 bp of the At4g05320 promoter. EC1.2: 1014 bp of the At2g21740 promoter. EC_enh.: 752 bp of the At2g21740 promoter fused to 548 bp of the At1g76750 promoter. MGE1: 1554 bp of the At5g55200 promoter. AG: 3101 bp of the At4g18960 promoter. ICU2: 625 bp of the At5g67100 promoter. CsVMV: 517 bp of a promoter from *Cassava Vein Mosaic* virus. RPS5a: 1688 bp of the At3g11940 promoter. YAO: 596 bp of the At4g05410 promoter. Cas9_1: Mali *et al.*, 2013 [3]. Cas9_2: Fauser *et al.*, 2014 [13]. Cas9_3: Li *et al.*, 2013 [25]. Cas9_4: Le Cong *et al.*, 2013 [10]. E9T: 631 bp of the *Pisum sativum rbcS* E9 terminator. OcsT: 714 bp of the *Agrobacterium tumefaciens octopine synthase* terminator. AgsT: 410 bp of the *Agrobacterium tumefaciens agropine synthase* terminator. NosT: 267 bp of the *Agrobacterium tumefaciens nopaline synthase* terminator. pU6-26: 205 bp of the At3g13855 promoter. sgRNA[EF]: "extension-flip" sgRNA. U6-26T: 7, 67 or 192 bp of the At3g13855 terminator. RB: Right Border. **d.** Five lines were tested for UBI10 and ICU2 and four lines for AG instead of six. **F.** Five lines were tested for YAO and RPS5a instead of six. The sgRNA targets *ADH1*. CRISPR activity measured in % of homozygous or biallelic stable mutants in the second generation after transformation (T2). Each dot represents an independent T2 family. Red dot: All the T2 lines from this family carry the same mutation, indicating a mutation more likely inherited from the T1 parent rather than being *de novo* from the T2 line. Bold and underlined: Most active construct(s) for each panel.

in most tissues [24]. We compared the *35S* and Arabidopsis *UBI10* promoters. More mutants were recovered using the *UBI10* promoter, suggesting it is more active than *35S* in the germ-line (Fig 3A). Following this observation, we tested other germline-expressed promoters.

In the combinations we tested, we detected low CRISPR activity using the meiosis I-specific promoter *MGE1* [26] (Fig 3C), the homeotic gene promoter *AG* [27] (Fig 3D) and the DNA

polymerase subunit-encoding gene promoter *ICU2* [28] (Fig 3D). They were tested with constructs inducing an overall low activity and we do not exclude that they can perform efficiently in other conditions. In one context specifically, *ICU2* promoter resulted in moderate activity in four of the six T2 families tested, while only one T2 family showed activity with the *UBI10* promoter (Fig 3E).

EC1.2 and an *EC1.2*::*EC1.1* fusion (referred as '*EC enhanced*' or '*ECenh*') are specifically expressed in the egg cell and were reported to trigger elevated mutation rates with CRISPR in Arabidopsis [17]. In our Golden Gate compatible system, only *ECenh* induced homozygous mutants in T1 and at low frequency (Fig 3B and 3G). In one comparison, *EC1.2* and *ECenh* performed slightly better than *pUBI10* (Fig 3D), but in another, they induced lower activity (Fig 3E).

A promoter from Cassava Vein Mosaic Virus (*CsVMV*) was reported to mediate CRISPR activity in *Brassica oleracea* [29]. We found that it induced more CRISPR activity than *pUBI10* in two combinations tested (Fig 3D and 3E).

We also tested the *YAO* and *RPS5a* promoters. Both of them were reported to boost CRISPR activity in Arabidopsis [15,16]. Both triggered elevated mutation rates compared with the *UBI10* promoter (Fig 3F). In one comparison, *pRPS5a* performed slightly better (Fig 3G), but in another, *pYAO* performed better (Fig 3H).

As have others, we conclude that the promoter driving *Cas9* expression influences CRISPR-mediated mutation rates [15–17,26]. We observed the best mutation rates using *RPS5a*, *YAO* and *UBI10* promoters.

Codon optimization of *Cas9* and presence of an intron elevate mutation rates

The activity of different constructs with the same promoter can be very different. For instance, *pRPS5a*:*Ca9* and *pYAO*:*Ca9* lines were recovered that displayed either high or low activity (Fig 3F and 3H). The most active constructs carried *Cas9_3* or *Cas9_4* alleles. We thus compared four *Cas9* alleles side-by-side (Fig 4). *Cas9_1* is a human codon-optimized version with a single C-terminal Nuclear Localization Signal (NLS) [3]. *Cas9_2* is an Arabidopsis codon-optimized version with a single C-terminal NLS [13]. *Cas9_3* is a plant codon-optimized version with both N- and C-terminal NLSs, an N-terminal FLAG tag and a potato intron *IV* [25]. *Cas9_4* is a human codon-optimized version with both N- and C-terminal NLSs and an N-terminal FLAG tag [10].

We found that in comparable constructs, *Cas9_2* performs better than *Cas9_1* (Fig 4E to 4H), consistent with the fact that *Cas9_2* was designed for Arabidopsis codon usage. However, human codon-optimized *Cas9_4* induced more mutants than Arabidopsis optimized *Cas9_2* in one experiment (Fig 4B). Cas9_4 has an extra N-terminal NLS compared to Cas9_2, which may explain this difference. In this comparison specifically, *Cas9_3* was less efficient than *Cas9_4*. However, by comparing *Cas9_3* and *Cas9_4* in combination with *YAO* or *RPS5a* promoters, we found that *Cas9_3* resulted in high mutation rates (Fig 4C and 4D). *Cas9_3* efficiency can be explained by the plant codon optimization, the presence of two NLSs and the inclusion of a plant intron. This intron was originally added to avoid expression in bacteria during cloning and, as side effect, can also increase expression *in planta* [30]. We recommend the use of *Cas9_3* for gene editing in Arabidopsis.

A modified sgRNA triggers CRISPR-induced mutations more efficiently

In the endogenous CRISPR immune system, Cas9 binds a CRISPR RNA (crRNA) and a trans-acting CRISPR RNA (tracrRNA) [31]. A fusion of both, called single guide RNA (sgRNA), is
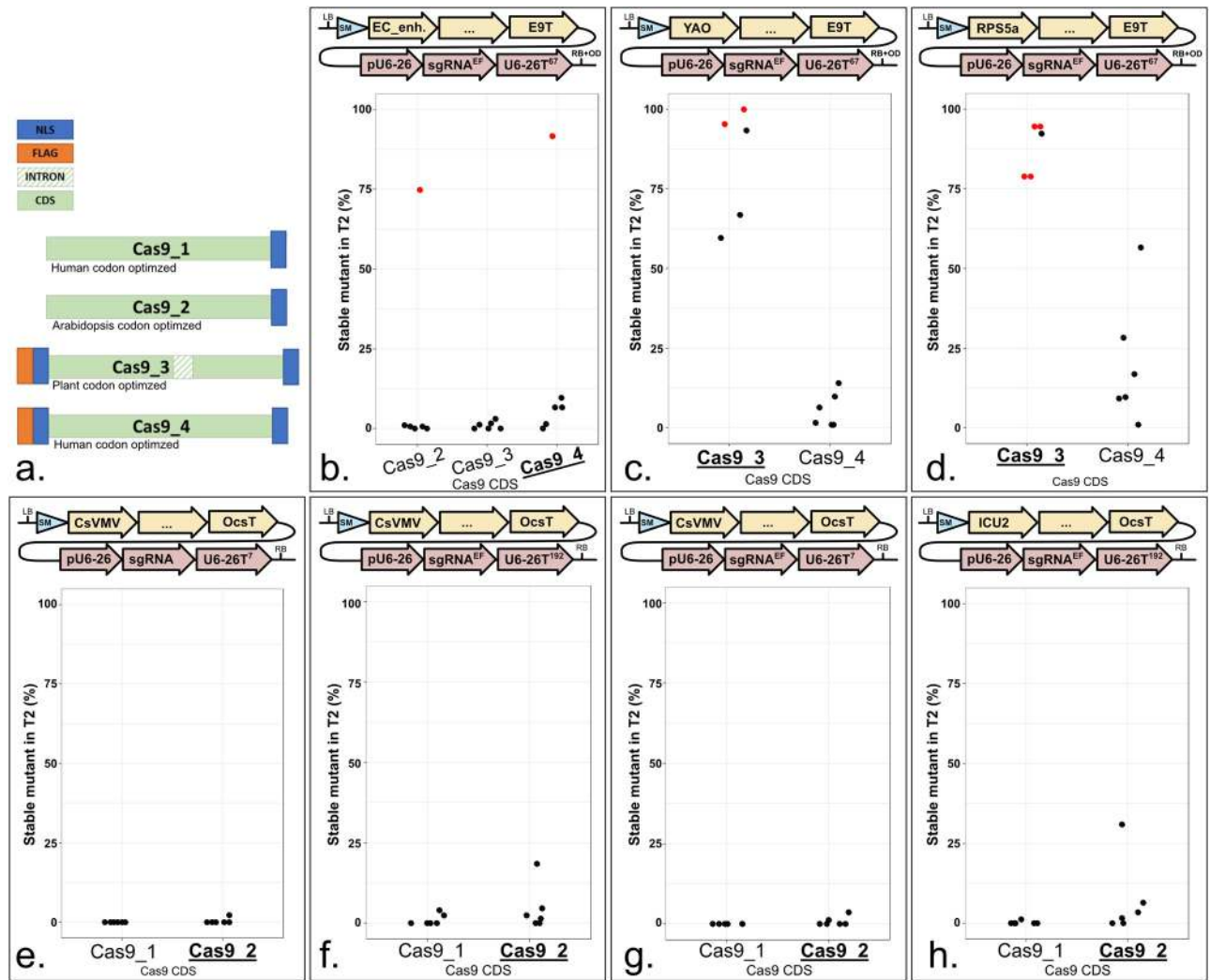
Fig 4. An intron-containing allele of *Cas9* triggers elevated mutation rates. a. Cas9_1: Mali *et al*., 2013 [3]. Cas9_2: Fauser *et al*., 2014 [13]. Cas9_3: Li *et al*., 2013 [25]. Cas9_4: Le Cong *et al*., 2013 [10]. NLS: Nuclear Localization Signal. FLAG: DYKDDDDK peptide. Apart from the FLAG and NLS, the amino acid sequences are identical. The nucleotide sequence (codon optimization) are different. Bars are not in scale. **b. to h.** Each panel represents a CDS comparison in the same context. CDSs can be compared within each panel, not from one panel to another. The modules were assembled into pICSL4723 (RB+OD, with an *overdrive*) or pAGM4723 (RB, without an *overdrive*) and transformed into Col-0 via *Agrobacterium tumefaciens* strain GV3101. LB: Left Border. SM: Selectable Marker (*Glufosinate resistance* gene). pEC_enh.: 752 bp of the At2g21740 promoter fused to 548 bp of the At1g76750 promoter. pYAO: 596 bp of the At4g05410 promoter. pRPS5a: 1688 bp of the At3g11940 promoter. CsVMV: 517 bp of a promoter from *Cassava Vein Mosaic* virus. pICU2: 625 bp of the At5g67100 promoter. E9T: 631 bp of the *Pisum sativum rbcS E9* terminator. OcsT: 714 bp of the *Agrobacterium tumefaciens octopine synthase* terminator. U6-26p: 205 bp of the At3g13855 promoter. sgRNA^{EF}: "extension-flip" sgRNA. U6-26T: 7, 67 or 192 bp of the At3g13855 terminator. RB: Right Border. **b.** Cas9_2 is in pAGM4723 (*i.e.* RB) in combination with U6-26^{192}; Cas9_3 and Cas9_4 are in pICSL4723 (*i.e.* RB+OD) in combination with U6-26^{67}. **c. and d.** Five lines were tested for Cas9_3 instead of six. The sgRNA targets *ADH1*. CRISPR activity measured in % of homozygous or biallelic stable mutants in the second generation after transformation (T2). Each dot represents an independent T2 family. Red dot: All the T2 lines from this family carry the same mutation, indicating a mutation more likely inherited from the T1 parent rather than being *de novo* from the T2 line. Bold and underlined: Most active construct(s) for each panel.

sufficient for CRISPR-mediated genome editing [32]. sgRNA stability was suggested to be a limiting factor in CRISPR system [33]. Chen *et al*. proposed an improved sgRNA to tackle this issue [8]. It carries an A-T transversion to remove a TTTT potential termination signal, and an extended Cas9-binding hairpin structure (Fig 5A). We compared side-by-side the 'Extended' and 'Flipped' sgRNA (sgRNA^{EF}) with the classic sgRNA (Fig 5B and 5C). In two independent

Fig 5. A modified sgRNA is slightly more efficient to trigger mutations. a. Original sgRNA proposed by Mali *et al.*, 2013 [3]. Extension-Flip (EF) sgRNA proposed by Chen *et al.*, 2013 [8]. **b. and c.** Each panel represents a sgRNA backbone comparison in the same context. sgRNA backbones can be compared within each panel but not from one panel to another. The modules were assembled into pICSL4723 (RB+OD, with an *overdrive*) or pAGM4723 (RB, without an *overdrive*) and transformed into Col-0 via *Agrobacterium tumefaciens* strain GV3101. LB: Left Border. SM: Selectable Marker (*Glufosinate resistance* gene). CsVMV: 517 bp of a promoter from *Cassava Vein Mosaic* virus. UBI10: 1327 bp of the At4g05320 promoter. Cas9_2: Fauser *et al.*, 2014 [13]. OcsT: 714 bp of the *Agrobacterium tumefaciens octopine synthase* terminator. U6-26p: 205 bp of the

At3g13855 promoter. U6-26T: 7 bp of the At3g13855 terminator. RB: Right Border. **c.** Five lines were tested for sgRNA[EF] and four lines for sgRNA instead of six. The sgRNA targets *ADH1*. CRISPR activity measured in % of homozygous or biallelic stable mutants in the second generation after transformation (T2). Each dot represents an independent T2 family. Bold and underlined: Most active construct(s) for each panel.

comparisons, the efficiency was higher with sgRNA[EF]. The improvement was not dramatic but sufficient to lead us to recommend use of 'EF'-modified guide RNAs for genome editing in Arabidopsis.

The 3' regulatory sequences of Cas9 and the sgRNA influence the overall activity

To avoid post-transcriptional modifications such as capping and polyadenylation, sgRNA must be transcribed by RNA polymerase III (Pol III). Several approaches involving ribozymes, Csy4 ribonuclease or tRNA-processing systems have been proposed but were not tested here [34–36]. *U6-26* is a Pol III-transcribed gene in Arabidopsis [37]. We used 205 bp of the 5' upstream region of *U6-26* as promoter and we compared a synthetic polyT sequence (seven thymidines) and 192 bp of the 3' downstream region as terminator. A T-rich stretch has been reported to function as a termination signal for Pol III [38].

In seven out of nine side-by-side comparisons, the authentic 192 bp of *U6-26* terminator directed a higher efficiency of the construct, as compared to a synthetic polyT termination sequence (Fig 6 and S2 Fig). We speculate that a stronger terminator increases the stability of the sgRNA. For multiplex genome editing, the use of 192 bp per sgRNA will result in longer T-DNAs and increase the risk of recombination and instability. We generated constructs with only 67 bp of the *U6-26* 3' downstream sequence. Such constructs were not compared side-by-side with the '192 bp terminator', although they enabled modest to high mutation rates (*e.g.* Fig 3F and 3G). With these results in mind, we recommend using 67 bp of the 3' downstream sequence of U6-26 as terminator for the sgRNA.

Since 3' regulatory sequences can influence sgRNA stability, we tested if the same was true for *Cas9*. We compared the *Pisum sativum rbcS E9* with two *A. tumefaciens* terminators commonly used in Arabidopsis: *Ocs* and *Ags* (Fig 7). We did not observe consistent differences between *E9* and *Ocs* (Fig 7A and 7B). However, in one comparison, *E9* outperformed *Ags* (Fig 7C). This is consistent with previous observations that RNA Polymerase II (Pol II) terminators quantitatively control gene expression and influence CRISPR efficiency in Arabidopsis [17,39]. We propose that a weak terminator after *Cas9* enables Pol II readthrough that could interfere with Pol III transcription of sgRNAs in some T-DNA construct architectures. This limiting factor can be corrected by divergent transcription of *Cas9* and *sgRNAs*.

Divergent transcription of Cas9 and sgRNA expression can elevate mutation rates

The Golden Gate Level 1 acceptor vector collection contains seven 'forward' expression cassettes and seven 'reverse' expression cassettes, which are interchangeable [23]. We assembled '*RPS5a*:*Cas9_4*:*E9*' and '*YAO*:*Cas9_3*:*E9*' in both the Level 1 vector position 2 forward (pICH47742) and reverse (pICH47811) (Figs 1 and 6). In one case, CRISPR activity was moderate when *Cas9* and *sgRNA* are expressed in the same direction and high when they are expressed in opposite direction (Fig 8A). In another case, CRISPR activity was very high in both cases (Fig 8B).

We thus recommend to both use a strong terminator after *Cas9* (*e.g. E9* or *Ocs*) and express *Cas9* and *sgRNA* in opposite directions.
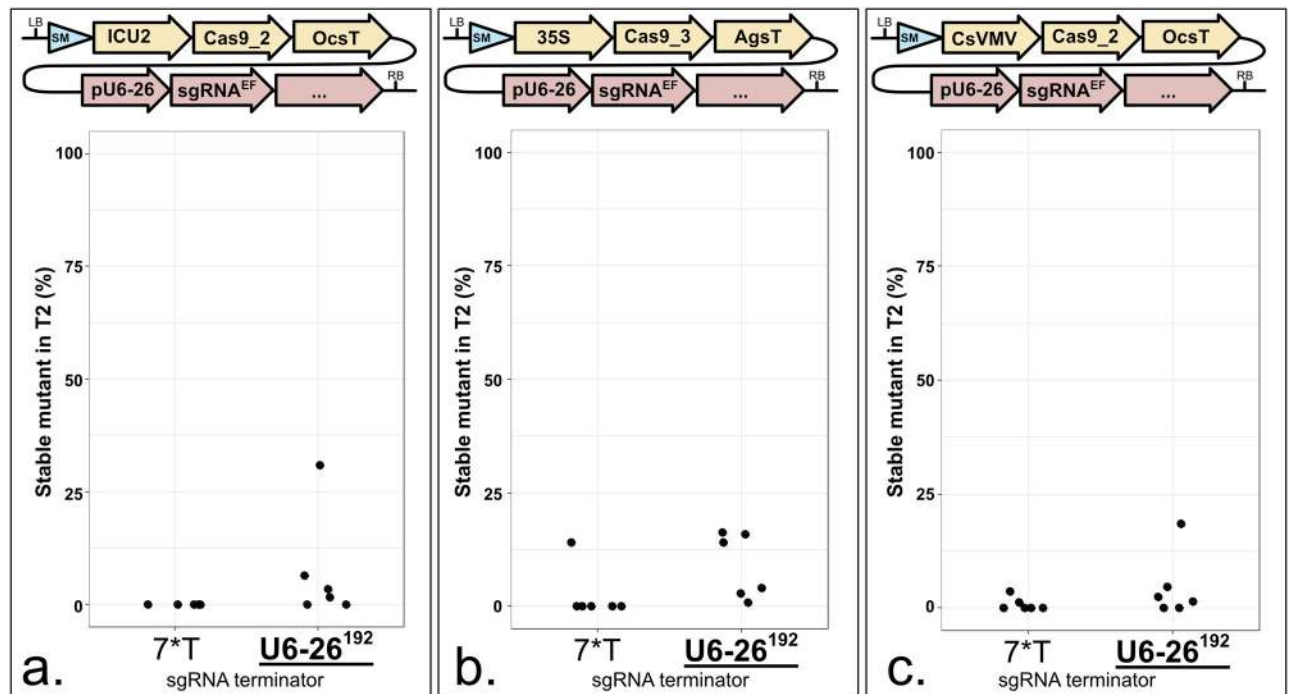
Fig 6. The sgRNA expression regulated by an authentic 3' regulatory sequence of U6-26 produces greater mutation rates. a. to c. Each panel represents a terminator comparison in the same context. Terminators can be compared within each panel, not from one panel to another. The modules were assembled into pAGM4723 and transformed into Col-0 via *Agrobacterium tumefaciens* strain GV3101. LB: Left Border. SM: Selectable Marker (*Glufosinate resistance* gene). ICU2: 625 bp of the At5g67100 promoter. 35S: 426 bp of the 35S promoter from Cauliflower Mosaic Virus. CsVMV: 517 bp of a promoter from *Cassava Vein Mosaic* virus. Cas9_2: Fauser *et al.*, 2014 [13]. Cas9_3: Li *et al.*, 2013 [25]. OcsT: 714 bp of the *Agrobacterium tumefaciens octopine synthase* terminator. AgsT: 410 bp of the *Agrobacterium tumefaciens agropine synthase* terminator. U6-26p: 205 bp of the At3g13855 promoter. sgRNA^{EF}: "extension-flip" sgRNA. U6-26T: 7 or 192 bp of the At3g13855 terminator. RB: Right Border. **a.** Five lines were tested for U6-26^7 instead of six. The sgRNA targets *ADH1*. CRISPR activity measured in % of homozygous or biallelic stable mutants in the second generation after transformation (T2). Each dot represents an independent T2 family. Bold and underlined: Most active construct(s) for each panel.

https://doi.org/10.1371/journal.pone.0204778.g006

Most of the stable double events are homozygous rather than biallelic

From the mutant screen, 315 allyl-alcohol resistance lines were confirmed by capillary sequencing (S5 Table). We classified them in four categories: (i) 59% were homozygous (single sequencing signal, different than *ADH1* WT), (ii) 11% were heterozygous (dual sequencing signal, one matching *ADH1* WT), (iii) 10% were biallelic (dual sequencing signal, none matching *ADH1* WT) and (iv) 20% were difficult to assign (unclear sequencing signals, either biallelic or due to somatic mutations, but clearly different than WT, heterozygous or homozygous genotypes) (Fig 9). The recovery of heterozygous (*ADH1/adh1*) lines indicates that the loss of a single copy of *ADH1* can sometimes enable plants to survive the allyl-alcohol selection.

Discussion

CRISPR emerged in 2012 as a useful tool for targeted mutagenesis in many organisms including plants [11,32]. In Arabidopsis, the transgenic expression of CRISPR components can be straightforward, avoiding tedious tissue culture steps. Many strategies to enhance the overall CRISPR-induced mutation rate have been proposed [8,13,15–17,40]. Here we report a systematic comparison of putative limiting factors including promoters, terminators, codon optimization, sgRNA improvement and T-DNA architecture.
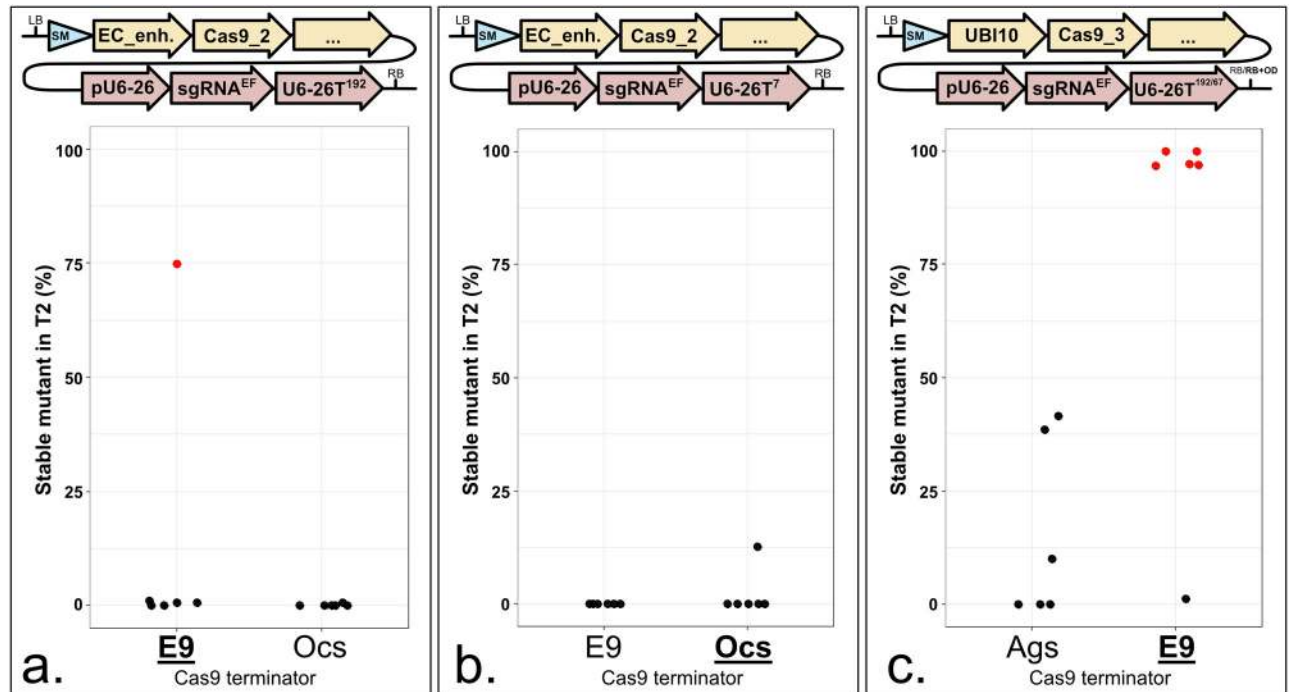
Fig 7. A weak 3' regulatory sequence reduces the CRISPR-induced mutation rate. a. to c. Each panel represents a terminator comparison in the same context. Terminators can be compared within each panel, not from one panel to another. The modules were assembled into pAGM4723 and transformed into Col-0 via *Agrobacterium tumefaciens* strain GV3101. LB: Left Border. SM: Selectable Marker (*Glufosinate resistance* gene). EC_enh.: 752 bp of the At2g21740 promoter fused to 548 bp of the At1g76750 promoter. UBI10: 1327 bp of the At4g05320 promoter. Cas9_2: Fauser *et al.*, 2014 [13]. Cas9_3: Li *et al.*, 2013 [25]. E9T: 631 bp of the *Pisum sativum rbcS E9* terminator. OcsT: 714 bp of the *Agrobacterium tumefaciens octopine synthase* terminator. AgsT: 410 bp of the *Agrobacterium tumefaciens agropine synthase* terminator. U6-26p: 205 bp of the At3g13855 promoter. sgRNA^{EF}: "extension-flip" sgRNA. U6-26T: 7, 67 or 192 bp of the At3g13855 terminator. RB: Right Border. For the comparison using the UBI10 promoter, the AgsT is in combination with U6-26^{192}; OcsT is in combination with U6-26^{67}. The sgRNA targets *ADH1*. CRISPR activity measured in % of homozygous or biallelic stable mutants in the second generation after transformation (T2). Each dot represents an independent T2 family. Red dot: All the T2 lines from this family carry the same mutation, indicating a mutation more likely inherited from the T1 parent rather than being *de novo* from the T2 line. Bold and underlined: Most active construct(s) for each panel.

https://doi.org/10.1371/journal.pone.0204778.g007

We found that the best promoters to control Cas9 expression are *UBI10*, *YAO* and *RPS5a*. The best terminators in our hands were *Ocs* from *A. tumefaciens* and *rbcS E9* from *P. sativum*. A plant codon-optimized, intron-containing *Cas9* allele outperformed the other alleles tested. A modified sgRNA with a hairpin Extension and a nucleotide Flip, called sgRNA^{EF}, triggers slightly elevated mutations rates. The sgRNA transcription regulation by the authentic 3' regulatory sequence of *AtU6-26* results in better CRISPR activity. We get high mutation rates with either 67 bp or 192 bp of terminator and recommend using the shortest (67 bp). We hypothesise that a weak terminator after *Cas9* enables RNA-polymerase II readthrough within the sgRNA expression cassette, preventing optimal expression of the sgRNA. Indeed, we noted an elevated CRISPR-Cas9 efficiency by expressing *Cas9* and *sgRNA* in opposite directions.

Considering the combinations of Cas9 and sgRNA genes tested in this study, we recommend to use a '*YAO:Cas9_3:E9*' and a '*pU6-26:sgRNA^{EF}:U6-26T^{67}*' cassettes in head-to-head orientation. This combination is included in the constructs tested here ([Fig 8B](#)) and enabled us to recover one homozygous mutants in five T1 plants tested. We also obtained useful rates with other constructs (*e.g.* [Fig 3F](#)), indicating that the CRISPR components do not entirely explain the final CRISPR activity. It was recently reported that heat stress increases the efficiency of CRISPR in Arabidopsis [41]. Environmental conditions may explain fluctuation of the CRISPR activity, independently of the T-DNA architecture.
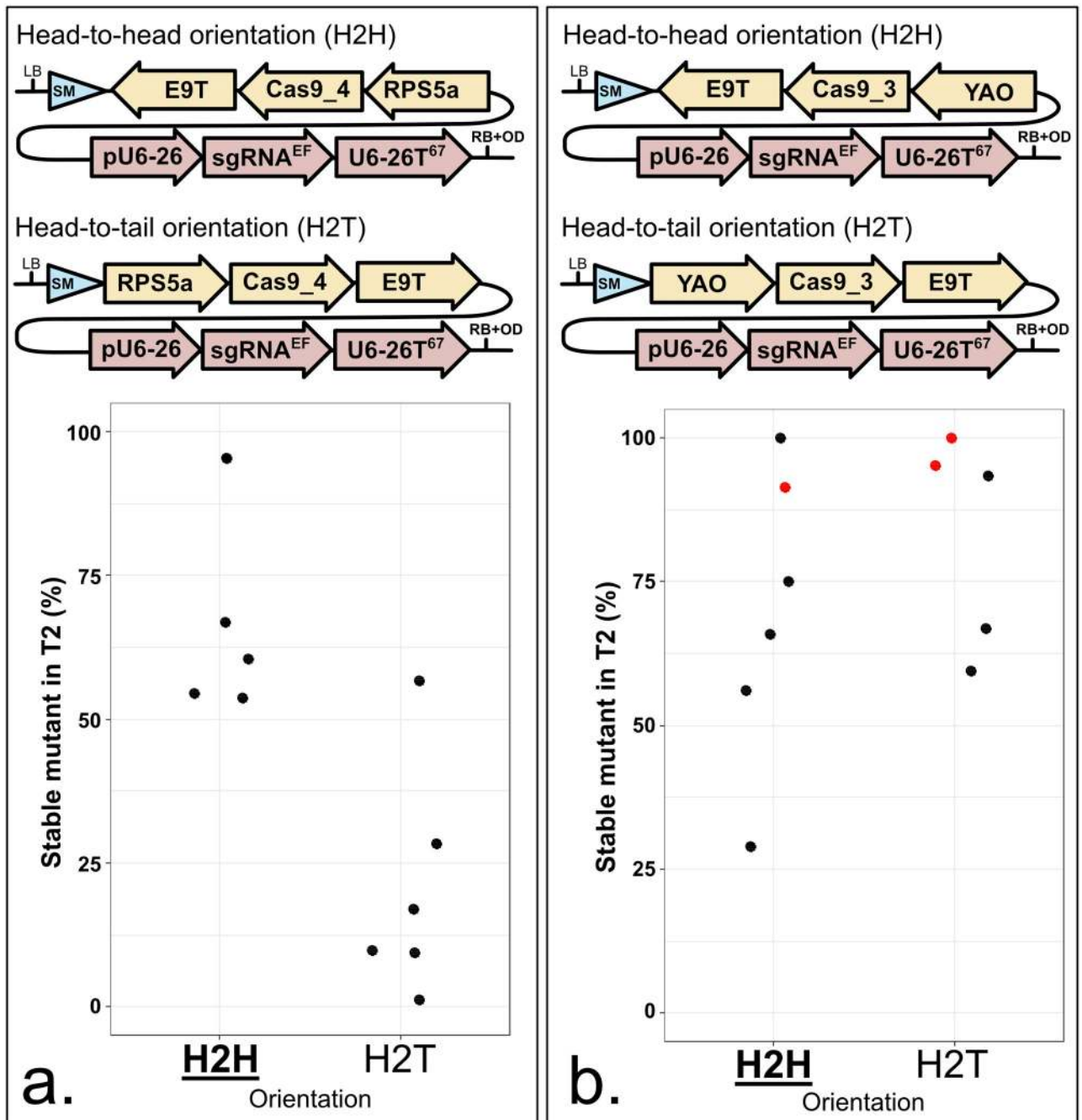
Fig 8. CRISPR activity is similar or higher when the sgRNA and the Cas9 expression cassettes are in a head-to-head orientation. a. and b. Each panel represents an orientation comparison in the same context. Orientations can be compared within each panel, not from one panel to another. The modules have been assembled by Golden Gate into pICSL4723 (RB+OD, with an *overdrive*) and transformed into Col-0 via *Agrobacterium tumefaciens* strain GV3101. LB: Left Border. SM: Selectable Marker (*Glufosinate resistance* gene. RPS5a: 1688 bp of the At3g11940 promoter. YAO: 596 bp of the At4g05410 promoter. Cas9_3: Li *et al.*, 2013 [25]. Cas9_4: Le Cong *et al.*, 2013 [10]. E9T: 631 bp of the *Pisum sativum rbcS E9* terminator. U6-26p: 205 bp of the At3g13855 promoter. sgRNA^EF: "extension-flip" sgRNA. U6-26T: 67 bp of the At3g13855 terminator. RB: Right Border. **a.** Five lines were tested for H2H instead of six. **b.** Five lines were tested for H2T instead of six. The sgRNA targets *ADH1*. CRISPR activity measured in % of homozygous or biallelic stable mutants in the second generation after transformation (T2). Each dot represents an independent T2 family. Red dot: All the T2 lines from this family carry the same mutation, indicating a mutation more likely inherited from the T1 parent rather than being *de novo* from the T2 line. Bold and underlined: Most active construct(s) for each panel.
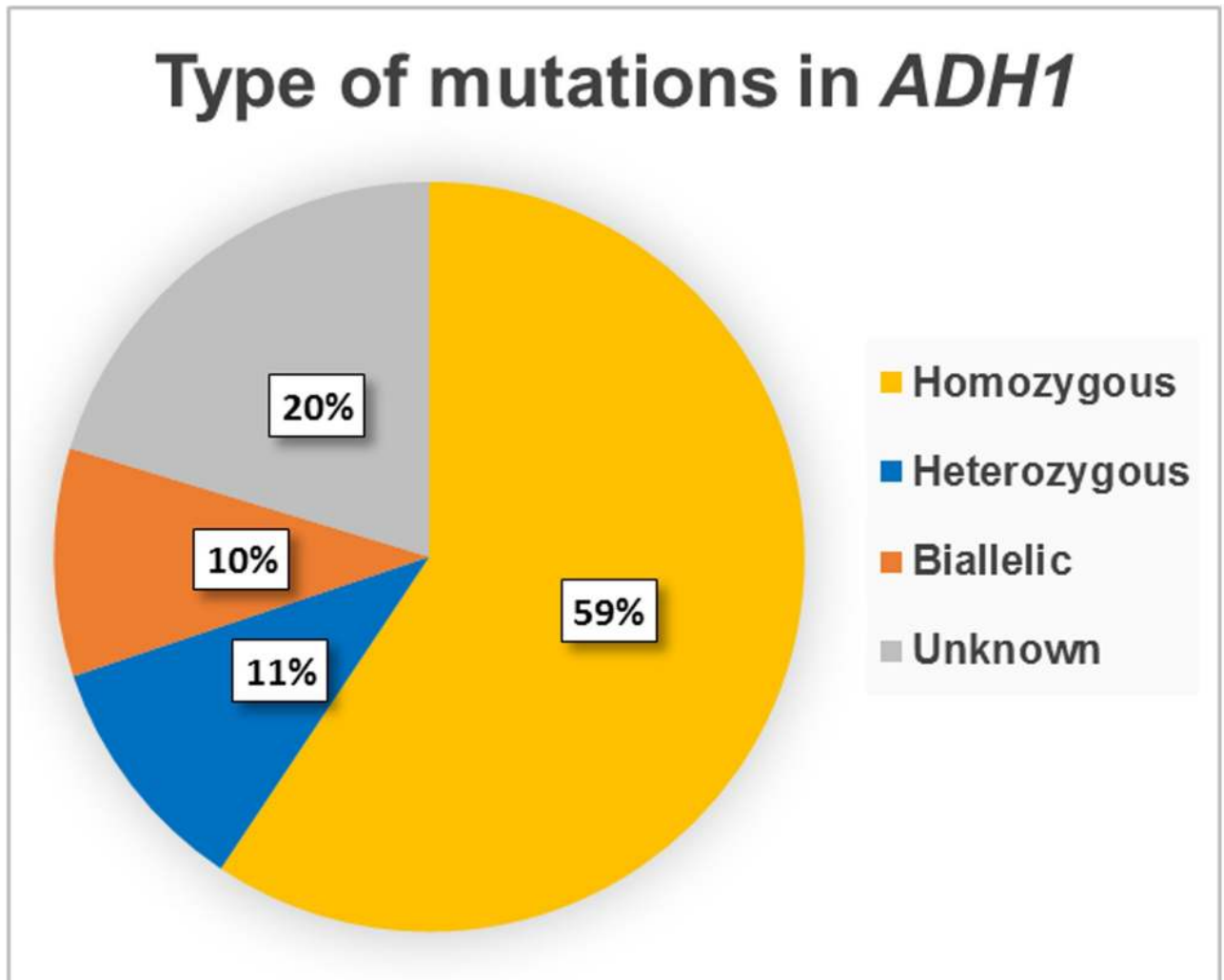
Fig 9. Genotype at *ADH1* locus confirmed by capillary sequencing. For each T2 family tested, up to six allyl-alcohol resistant plants were genotyped by capillary sequencing of an sgRNA target (*ADH1*) PCR amplicon. We retrieved a total of 315 sequences with a mutation. 59% (187) showed a single sequencing signal, different than *ADH1* WT and were classified as "Homozygous". 11% (33) showed an overlap of two sequencing signals, one matching *ADH1* WT and one different; and were classified as "Heterozygous". 10% (31) showed an overlap of two sequencing signals, none matching *ADH1* WT; and were classified as "Biallelic". 64 (20%) showed an overlap of signals different than WT but not clear enough to distinguish; and were classified as "Unknown". The "Unknown" sequences can be biallelic or due to somatic mutations but are different than WT, heterozygous or homozygous genotypes.

We were surprised to recover more homozygous than biallelic events. Stable double mutations are the result of two CRISPR events, on the male and female inherited chromosome respectively. In this scenario, lines can be recovered with two different mutations, resulting in a biallelic (*e.g. adh1-2/adh1-3*) genotype, rather than having the same mutation on both chromosomes (*e.g. adh1-1/adh1-1*). Double-strand break-induced homologous recombination occurs between allelic sequences [42]. It has been reported that double strand breaks caused by CRISPR-Cas9 can increase this phenomenon [43]. Allelic recombination can explain our observation of the same mutation on both copies of *ADH1*. The prevalence of homozygous over biallelic genotypes facilitates the genotyping and is an advantage for targeted mutagenesis using CRISPR-Cas9.
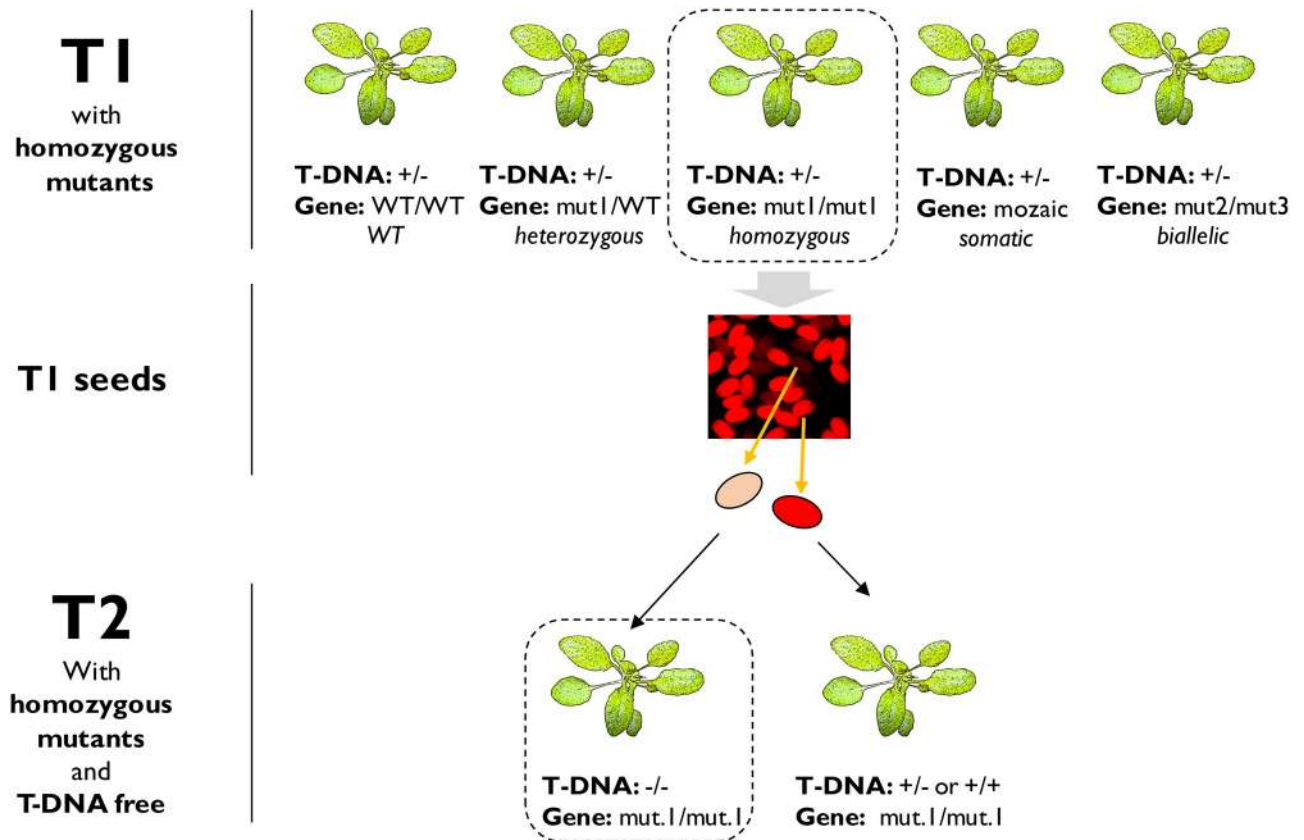
Fig 10. FAST-Red combined with CRISPR to generate T-DNA free mutants. The five T1 lines are independent transformants. They are all hemizygous for the T-DNA. At the sgRNA target site, they can be WT, or display somatic, heterozygous, biallelic of homozygous mutations. All the possibilities are represented here. "Somatic" describes events happening in somatic cells, not inherited in the next generation. As somatic events can happen independently in each cell, they often result in mosaic pattern of mutations across the leaf. One line has homozygous mutation (*mut1/mut1*). It produces seeds segregating for the T-DNA, visible under microscope if using FAST-Red. The seeds will segregate 3:1 (Red: Non-red) if there is one locus insertion, 15:1 (Red: Non-red) if there are two loci insertion, etc. The T2 progeny of (mut1/mut1) is 100% homozygous for the mutation. The non-red seeds are also T-DNA free.

We used a glufosinate resistance selectable marker which enables easy selection of transgenic lines. It can be important to segregate away the T-DNA in the CRISPR mutant line for multiple reasons. For instance, a loss-of-function phenotype must be confirmed by complementation of the CRISPR-induced mutation. A CRISPR construct still present in the mutant can target the complementation transgene and interfere with the resulting phenotypes. Selection of non-transgenic lines is possible but complicated with classic selectable markers such as kanamycin or glufosinate resistance, since a selective treatment kills the non-transgenic plants. FAST-Green and FAST-Red provide a rapid non-destructive selectable marker and involve expression of a GFP- or RFP-tagged protein in the seed [44]. Transgenic and non-transgenic seeds can be distinguished under fluorescence microscopy [16,45,46]. This facilitates recovery of mutant seeds lacking the T-DNA (Fig 10). Homozygous mutants can be identified among the independent T1 lines. Non-fluorescent seeds can be selected from the T1 seeds. The resulting T2 plants are homozygous mutant and non-transgenic.

We report a CRISPR- and Golden Gate-based method to generate stable Arabidopsis mutant lines in one generation. In our efforts to elevate mutation rates in Arabidopsis, we found several limiting factors mostly related to *Cas9* and *sgRNA* transcription. Some of these

findings can be tested for other plant species and for knock-in breeding. The generation of null alleles via CRISPR is today quick and simple, facilitating the investigation of gene function. Improvement of rates of gene 'knock-ins' provides the next challenge. *In vivo* gene tagging or knock-in breeding are theoretically possible and have been reported [47–50]. Improvements in CRISPR-based genome editing techniques will facilitate the study of genes and proteins and be beneficial for both basic and applied plant science.

Materials and methods

CRISPR constructs assembly

The vectors were assembled using the Golden Gate modular cloning method [23]. To generate the Cas9 expression cassettes, the *RPS5a*, *YAO*, *ICU2*, *CsVMV*, *EC1.2*, *EC_enh.*, *UBI10*, *AG*, *MGE1* and *35S* promoters, the *Cas9_1*, *Cas9_2*, *Cas9_3* and *Cas9_4* coding sequences, the *Ocs*, *Nos*, *Ags* and *E9* terminators were amplified using primers flanked with BpiI restriction sites associated with Golden Gate compatible overhangs (S3 Table). 0.02 pmoles of the purified PCR products were mixed with the same molar amount of the corresponding Level 0 vector (S3 Table), 0.5 μl of BpiI enzyme (10U/μl, ThermoFisher), 0.5 μl of T4 ligase (400U/μl, NEB), 1.5 μl of CutSmart Buffer (NEB), 1.5μl of Bovine Serum Albumin (10X) and water in a total reaction volume of 15 μl. The reaction was placed in a thermocycler and the following 'Golden Gate' program was applied: 20 seconds 37˚C, 25 cycles of [3 minutes 37˚C / 4 minutes 16˚C], 5 minutes 50˚C and 5 minutes 80˚C.

Combinations of three Level 0 vectors containing respectively a promoter, a Cas9 coding sequence and a terminator were assembled in Level 1 vector pICH7742 (Position 2) or pICH47811 (Position 2, reverse) by the same 'Golden Gate' protocol but using 0.5 μl of BpiI enzyme (10U/μl, ThermoFisher) instead of 0.5 μl of BsaI-HF.

To generate the sgRNA expression cassettes, DNA fragments containing the classic or the 'EF' backbone with 7, 67 or 192 bp of the *U6-26* terminator were amplified using primers flanked with BsaI restriction sites associated with Golden Gate compatible overhangs (S3 Table). The amplicons were assembled with the U6-26 promoter (pICSL90002) in Level 1 vector pICH7751 (Position 3) by the 'Golden Gate' protocol using the BsaI-HF enzyme. Combinations of three Level 1 vectors containing a glufosinate resistance selectable maker (pICSL11017), a Cas9 expression cassette and a sgRNA expression cassette were assembled in Level 2 pAGM4723 (- *overdrive*) or pICSL4723 (+ *overdrive*) by the 'Golden Gate' protocol using the BpiI enzyme. All the plasmids were prepared using a QIAPREP SPIN MINIPREP KIT on *Escherichia coli* DH10B electrocompetent cells selected with appropriate antibiotics and X-gal.

All the plasmid identification numbers refer to the 'addgene database' (www.addgene.org/).

Plant transformation, growth and selection

Agrobacterium tumefaciens strain GV3101 was transformed with plasmids by electroporation and used for stable transformation of Arabidopsis accession Col-0. Arabidopsis plants were grown in 'short days' conditions (10 hr light/14 hr dark, 21˚C). Transformants were selected by spraying three times 1- to 3-weeks old seedlings with phosphinotrycin at a concentration of 0.375g/l. 4-weeks old resistant plants were transferred in 'long days' conditions (16 hr light/8 hr dark, 21˚C) for flowering. For each genotype, six independent T1 were self-pollinated to obtain six independent T2 families per construct.

Characterisation of CRISPR events

T2 families were tested for resistance to allyl-alcohol. ~100 seeds were sterilized, immersed in water (4˚C, dark, overnight), treated with allyl-alcohol (30mM, room temperature, 2 hours, shaken at 750rpm), rinsed three times with water and sown on $MS^{1/2}$ medium. After two weeks, the number of germinated and non-germinated seeds was monitored. DNA was extracted from up six allyl-alcohol resistant plants (or all the resistant plants if there were less than six) for genotyping. ~$0.5cm^2$ of leaf tissue was printed by mechanical pression onto an FTA filter paper (Whatman Bioscience). 1-mm disks were punched out from FTA filter paper by using a punch and placed in a 200μl PCR tubes. One disc was used per tube. Samples were incubated in 50μl of FTA buffer (1.25ml Tris 1M, 500μl EDTA 0.5M, 12.5μl Tween 20 and water up to a total volume of 125ml) for 2 hours and rinsed with water. PCR was performed on this template using primers flanking the sgRNA target in *ADH1* (S3 Table) and Q5 High-Fidelity DNA Polymerase (NEB, following the manufacturer recommendations). After amplification, the PCR products were resolved by electrophoresis on a 1.5% agarose gel and purified using the QIAquick Gel Extraction Kit (QIAGEN). The purified PCR product was sequenced using the same primer set for amplifications by capillary sequencing (GATC Biotech). Sequencing results were compared to the Col-0 sequence of *ADH1* using CLC Main Workbench 7.7.1. *ADH1* genotypes were reported as WT (identical to Col-0), heterozygous (both Col-0 and single mutation detected), biallelic (two different mutations detected), homozygous (single mutation detected) or somatic (more than two signals detected). The number of confirmed mutants among all the allyl-alcohol resistant lines was used to estimate the total number of real mutants among allyl-alcohol survivors from each plate. For each T2 family, the CRISPR efficiency was defined as the ratio of homozygous and biallelic mutants compared to the total number of seeds sown. Plots presented in this article were made using *ggplot2* in *R* version 3.3.2.

Supporting information

S1 Table. List of the Level 2 T-DNA used in this article.
(XLSX)

S2 Table. List of plasmids used in this article, available through addgene.
(XLSX)

S3 Table. List of primers used in this article. Some vectors were not cloned using a PCR step (*e.g.* synthesised or cloned prior this article), which are indicated in this table.
(XLSX)

S4 Table. Mutation rates table. LBC number indicates a unique independent T1 line. Clone, SLJ number and Genotype refers to the "Level 2 Constructs" table (S1 Table). Vector "pAGM4723" lacks an *overdrive*; Vector "pICH4723" has an *overdrive*. "Same_mutation" indicates whether all the lines carry the same mutation. It is applied only if more than 75% of the seeds germinated. If so, it indicates that the parent was likely a homozygous mutant and the mutation was inherited to all progenies.
(CSV)

S5 Table. List of mutations obtained in *ADH1*. From capillary sequencing data.
(XLSX)

S1 Fig. The presence of the *overdrive* sequence in the T-DNA right border does not affect the CRISPR efficiency. A. Sequence of the right border with (pICSL4723) or without (pAGM4723) the *overdrive* sequence. **B. and C.** Each panel represents a vector comparison in the same context. Vectors can be compared within each panel, not from one panel to another.

The modules have been assembled by Golden Gate into pICSL4723 (OD+, with an *overdrive*) or pAGM4723 (OD-, without an *overdrive*) and transformed into Col-0 via *Agrobacterium tumefaciens* strain GV3101. LB: Left Border. SM: Sel. Marker (Glufosinate *resistance* gene). EC1.2: 1014 bp of the At2g21740 promoter. EC_enh.: 752 bp of the At2g21740 promoter fused to 548 bp of the At1g76750 promoter. Cas9_2: Fauser *et al.*, 2014 [13]. E9T: 631 bp of the *Pisum sativum rbcS E9* terminator. U6-26p: 205 bp of the At3g13855 promoter. sgRNA^{EF}: "extension-flip" sgRNA. U6-26T: 7 bp of the At3g13855 terminator. RB: Right Border. The sgRNA targets *ADH1*. CRISPR activity measured in % of homozygous or biallelic stable mutants in the second generation after transformation (T2). Each dot represents an independent T2 family. Bold and underlined: Most active construct(s) for each panel. The *overdrive* sequence can increase the integration efficiency [21]. In one comparison the presence of the *overdrive* results in slightly better activity (C), but in another one it did not (B). We concluded that the presence of an *overdrive* does not influence the CRISPR efficiency. Thus, we could compare constructs independently of the presence of an *overdrive*.
(TIFF)

S2 Fig. The sgRNA expression regulated by an authentic 3' regulatory sequence of U6-26 produces greater mutation rates. A. to F. Each panel represents a terminator comparison in the same context. Terminators can be compared within each panel, not from one panel to another. The modules were assembled into pICSL4723 (RB+OD, with an *overdrive*) or pAGM4723 (RB, without an *overdrive*) and transformed into Col-0 via *Agrobacterium tumefaciens* strain GV3101. LB: Left Border. SM: Sel. Marker (Glufosinate *resistance* gene). CsVMV: 517 bp of a promoter from Cassava Vein Mosaic virus. UBI10: 1327 bp of the At4g05320 promoter. EC1.2: 1014 bp of the At2g21740 promoter. EC_enh.: 752 bp of the At2g21740 promoter fused to 548 bp of the At1g76750 promoter. Cas9_1: Mali *et al.*, 2013 [3]. Cas9_2: Fauser *et al.*, 2014 [13]. E9T: 631 bp of the *Pisum sativum rbcS E9* terminator. OcsT: 714 bp of the *Agrobacterium tumefaciens octopine synthase* terminator. EF: 205 bp of the At3g13855 promoter controlling the expression of an "extension-flip" sgRNA. U6-26T: 7 or 192 bp of the At3g13855 terminator. **B.** Five lines were tested for 7*T instead of six. The sgRNA targets *ADH1*. CRISPR activity measured in % of homozygous or biallelic stable mutants in the second generation after transformation (T2). Each dot represents an independent T2 family. Red dot: All the T2 lines from this family carry the same mutation, indicating a mutation more likely inherited from the T1 parent rather than being *de novo* from the T2 line. Bold and underlined: Most active construct(s) for each panel.
(TIFF)

S1 File. Plasmid maps. Maps are in genbank format in a ZIP file.
(ZIP)

Acknowledgments

Author Contributions

Conceptualization: Baptiste Castel, Laurence Tomlinson, Ying Yang, Jonathan D. G. Jones.

Data curation: Baptiste Castel.

Formal analysis: Baptiste Castel.

Funding acquisition: Jonathan D. G. Jones.

Investigation: Baptiste Castel, Laurence Tomlinson, Federica Locci.

Methodology: Baptiste Castel, Laurence Tomlinson, Ying Yang, Jonathan D. G. Jones.

Project administration: Jonathan D. G. Jones.

Resources: Baptiste Castel, Laurence Tomlinson, Federica Locci, Ying Yang.

Supervision: Jonathan D. G. Jones.

Visualization: Baptiste Castel.

Writing – original draft: Baptiste Castel.

Writing – review & editing: Baptiste Castel, Laurence Tomlinson, Jonathan D. G. Jones.

References

1. Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science. 2014; 343: 1247997. https://doi.org/10.1126/science.1247997 PMID: 24505130

2. Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol. 2013; 31: 833–838. https://doi.org/10.1038/nbt.2675 PMID: 23907171

3. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. Science. 2013; 339: 823–8266. https://doi.org/10.1126/science.1232033 PMID: 23287722

4. Wright A V., Nuñez JK, Doudna JA. Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. Cell. 2016; 164: 29–44. https://doi.org/10.1016/j.cell.2015.12.035 PMID: 26771484

5. Makarova KS, Zhang F, Koonin E V. SnapShot: Class 1 CRISPR-Cas Systems. Cell. 2017; 168: 946. https://doi.org/10.1016/j.cell.2017.02.018 PMID: 28235204

6. Makarova KS, Zhang F, Koonin E V. SnapShot: Class 2 CRISPR-Cas Systems. Cell. 2017; 168: 328. https://doi.org/10.1016/j.cell.2016.12.038 PMID: 28086097

7. Barrangou R, Horvath P. A decade of discovery: CRISPR functions and applications. Nat Microbiol. 2017; 2: 17092. https://doi.org/10.1038/nmicrobiol.2017.92 PMID: 28581505

8. Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li G-W, et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell. 2013; 155: 1479–1491. https://doi.org/10.1016/j.cell.2013.12.001 PMID: 24360272

9. Gootenberg JS, Abudayyeh OO, Lee JW, Essletzbichler P, Dy AJ, Joung J, et al. Nucleic acid detection with CRISPR-Cas13a/C2c2. Science. 2017; 356: 438–442. https://doi.org/10.1126/science.aam9321 PMID: 28408723

10. Cong L, Ran F, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. Science. 2013; 339: 819–822. https://doi.org/10.1126/science.1231143 PMID: 23287718

11. Nekrasov V, Staskawicz BJ, Weigel D, Jones JDG, Kamoun S. Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. Nat Biotechnol. 2013; 31: 691–693. https://doi.org/10.1038/nbt.2655 PMID: 23929340

12. Jiang W, Zhou H, Bi H, Fromm M, Yang B, Weeks DP. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. Nucleic Acids Res. 2013; 41: e188. https://doi.org/10.1093/nar/gkt780 PMID: 23999092

13. Fauser F, Schiml S, Puchta H. Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. Plant J. 2014; 79: 348–359. https://doi.org/10.1111/tpj.12554 PMID: 24836556

14. Feng Z, Mao Y, Xu N, Zhang B, Wei P, Yang D-L, et al. Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. Proc Natl Acad Sci. 2014; 111: 4632–4637. https://doi.org/10.1073/pnas.1400822111 PMID: 24550464

15. Yan L, Wei S, Wu Y, Hu R, Li H, Yang W, et al. High efficiency genome editing in Arabidopsis using *YAO* promoter-driven CRISPR/Cas9 system. Mol Plant. 2015; 8: 1820–1823. https://doi.org/10.1016/j.molp.2015.10.004 PMID: 26524930

16. Tsutsui H, Higashiyama T. pKAMA-ITACHI vectors for highly efficient CRISPR/Cas9-mediated gene knockout in *Arabidopsis thaliana*. Plant Cell Physiol. 2017; 58: 46–56. https://doi.org/10.1093/pcp/pcw191 PMID: 27856772

17. Wang Z-P, Xing H-L, Dong L, Zhang H-Y, Han C-Y, Wang X-C, et al. Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. Genome Biol. 2015;16.

18. Liang Z, Chen K, Zhang Y, Liu J, Yin K, Qiu J-L, et al. Genome editing of bread wheat using biolistic delivery of CRISPR/Cas9 in vitro transcripts or ribonucleoproteins. Nat Protoc. 2018; 13: 413–430. https://doi.org/10.1038/nprot.2017.145 PMID: 29388938

19. Woo JW, Kim J, Kwon S Il, Corvalán C, Cho SW, Kim H, et al. DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. Nat Biotechnol. 2015; 10–13.

20. Clough SJ, Bent AF. Floral dip: A simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J. 1998; 16: 735–743. https://doi.org/10.1046/j.1365-313X.1998.00343.x PMID: 10069079

21. Peraltal EG, Hellmiss R, Ream W. *Overdrive*, a T-DNA transmission enhancer on the *A. tumefaciens* tumour-inducing plasmid. EMBO J. 1986; 5: 1137–1142. https://doi.org/10.1002/j.1460-2075.1986.tb04338.x PMID: 15966101

22. Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S. A modular cloning system for standardized assembly of multigene constructs. PLoS One. 2011; 6: e16765. https://doi.org/10.1371/journal.pone.0016765 PMID: 21364738

23. Engler C, Youles M, Gruetzner R, Ehnert TM, Werner S, Jones JDG, et al. A Golden Gate modular cloning toolbox for plants. ACS Synth Biol. 2014; 3: 839–843. https://doi.org/10.1021/sb4001504 PMID: 24933124

24. Sunilkumar G, Mohr L, Lopata-Finch E, Emani C, Rathore KS. Developmental and tissue-specific expression of *CaMV 35S* promoter in cotton as revealed by GFP. Plant Mol Biol. 2002; 50: 463–474. https://doi.org/10.1023/A:1019832123444 PMID: 12369622

25. Li J-F, Norville JE, Aach J, McCormack M, Zhang D, Bush J, et al. Multiplex and homologous recombination–mediated genome editing in Arabidopsis and *Nicotiana benthamiana* using guide RNA and Cas9. Nat Biotechnol. 2013; 31: 686–688.

26. Eid A, Ali Z, Mahfouz MM. High efficiency of targeted mutagenesis in Arabidopsis via meiotic promoter-driven expression of Cas9 endonuclease. Plant Cell Rep. 2016; 35: 1555–1558. https://doi.org/10.1007/s00299-016-2000-4 PMID: 27236699

27. Hong RLRL, Hamaguchi L, Busch MA, Weigel D. Regulatory elements of the floral homeotic gene AGAMOUS identified by phylogenetic footprinting and shadowing. Plant Cell. 2003; 15: 1296–1309. https://doi.org/10.1105/tpc.009548 PMID: 12782724

28. Hyun Y, Kim J, Cho SW, Choi Y, Kim J-S, Coupland G. Site-directed mutagenesis in *Arabidopsis thaliana* using dividing tissue-targeted RGEN of the CRISPR/Cas system to generate heritable null alleles. Planta. 2014; 241: 271–284. https://doi.org/10.1007/s00425-014-2180-5 PMID: 25269397

29. Lawrenson T, Shorinola O, Stacey N, Li C, Østergaard L, Patron NJ, et al. Induction of targeted, heritable mutations in barley and Brassica oleracea using RNA-guided Cas9 nuclease. Genome Biol. 2015; 16: 258. https://doi.org/10.1186/s13059-015-0826-7 PMID: 26616834

30. Callis J, Fromm M, Walbot V. Introns increase gene expression in cultured maize cells. Genes Dev. 1987; 1: 1183–1200. https://doi.org/10.1101/gad.1.10.1183 PMID: 2828168

31. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. 2014; 157: 1262–1278. https://doi.org/10.1016/j.cell.2014.05.010 PMID: 24906146

32. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science. 2012; 337: 816–821. https://doi.org/10.1126/science.1225829 PMID: 22745249

33. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. Elife. 2013; 2013: e00471. https://doi.org/10.7554/eLife.00471 PMID: 23386978

34. Gao Y, Zhao Y. Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. J Integr Plant Biol. 2014; 56: 343–349. https://doi.org/10.1111/jipb.12152 PMID: 24373158

35. Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. Nat Biotechnol. 2014; 32: 569–576. https://doi.org/10.1038/nbt.2908 PMID: 24770325

36. Xie K, Minkenberg B, Yang Y. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proc Natl Acad Sci. 2015; 112: 3570–3575. https://doi.org/10.1073/pnas.1420294112 PMID: 25733849

37. Li X, Jiang DH, Yong K, Zhang DB. Varied transcriptional efficiencies of multiple Arabidopsis U6 small nuclear RNA genes. J Integr Plant Biol. 2007; 49: 222–229. https://doi.org/10.1111/j.1744-7909.2007.00393.x

38. Waibel F, Filipowicz W. U6 snRNA genes of Arabidopsis are transcribed by RNA polymerase III but contain the same two upstream promoter elements as RNA polymerase II-transcribed U-snRNA genes. Nucleic Acids Res. 1990; 18: 3451–3458. https://doi.org/10.1093/nar/18.12.3451 PMID: 2362802

39. Nagaya S, Kawamura K, Shinmyo A, Kato K. The HSP terminator of Arabidopsis thaliana increases gene expression in plant cells. Plant Cell Physiol. 2010; 51: 328–332. https://doi.org/10.1093/pcp/pcp188 PMID: 20040586

40. Peterson BA, Haak DC, Nishimura MT, Teixeira PJPL, James SR, Dangl JL, et al. Genome-Wide Assessment of Efficiency and Specificity in CRISPR/Cas9 Mediated Multiple Site Targeting in Arabidopsis. PLoS One. 2016; 11: e0162169. https://doi.org/10.1371/journal.pone.0162169 PMID: 27622539

41. Le Blanc C, Zhang F, Mendez J, Lozano Y, Chatpar K, Irish V, et al. Increased efficiency of targeted mutagenesis by CRISPR/Cas9 in plants using heat stress. Plant J. 2017; 9: 377–386. https://doi.org/10.1111/tpj.13782 PMID: 29161464

42. Gisler B, Salomon S, Puchta H. The role of double-strand break-induced allelic homologous recombination in somatic plant cells. Plant J. 2002; 32: 277–284. https://doi.org/10.1046/j.1365-313X.2002.01421.x PMID: 12410807

43. Hayut SF, Bessudo CM, Levy AA. Targeted recombination between homologous chromosomes for precise breeding in tomato. Nat Commun. Nature Publishing Group; 2017; 8: 15605.

44. Shimada TL, Shimada T, Hara-Nishimura I. A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of Arabidopsis thaliana. Plant J. 2010; 61: 519–528. https://doi.org/10.1111/j.1365-313X.2009.04060.x PMID: 19891705

45. Wu R, Lucke M, Jang Y, Zhu W, Symeonidi E, Wang C, et al. An efficient CRISPR vector toolbox for engineering large deletions in Arabidopsis thaliana. Plant Methods. BioMed Central; 2018; 14: 65. https://doi.org/10.1186/s13007-018-0330-7 PMID: 30083222

46. Morineau C, Bellec Y, Tellier F, Gissot L, Kelemen Z, Nogué F, et al. Selective gene dosage by CRISPR-Cas9 genome editing in hexaploid Camelina sativa. Plant Biotechnol J. 2017; 15: 729–739. https://doi.org/10.1111/pbi.12671 PMID: 27885771

47. Čermák T, Curtin SJ, Gil-Humanes J, Čegan R, Kono TJY, Konečná E, et al. A multi-purpose toolkit to enable advanced genome engineering in plants. Plant Cell. 2017; 29: 1196–1217. https://doi.org/10.1105/tpc.16.00922 PMID: 28522548

48. Čermák T, Baltes NJ, Čegan R, Zhang Y, Voytas DF. High-frequency, precise modification of the tomato genome. Genome Biol. 2015; 16: 232. https://doi.org/10.1186/s13059-015-0796-9 PMID: 26541286

49. Li Z, Liu Z-B, Xing A, Moon BP, Koellhoffer JP, Huang L, et al. Cas9-Guide RNA Directed Genome Editing in Soybean. Plant Physiol. 2015; 169: 960–970. https://doi.org/10.1104/pp.15.00783 PMID: 26294043

50. Svitashev S, Young JK, Schwartz C, Gao H, Falco SC, Cigan AM. Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA. Plant Physiol. 2015; 169: 931–945. https://doi.org/10.1104/pp.15.00793 PMID: 26269544