Synthetic evolution of herbicide resistance using a T7 RNAP–based random DNA base editor

Haroon Butt$, Jose Luis Moreno Ramirez$, Magdy Mahfouz$

Synthetic directed evolution via localized sequence diversification and the simultaneous application of selection pressure is a promising method for producing new, beneficial alleles that affect traits of interest in diverse species; however, this technique has rarely been applied in plants. Here, we designed, built, and tested a chimeric fusion of T7 RNA Polymerase (RNAP) and deaminase to enable the localized sequence diversification of a target sequence of interest. We tested our T7 RNAP–DNA base editor in Nicotiana benthamiana transient assays to target a transgene expressing GFP under the control of the T7 promoter and observed C-to-T conversions. We then targeted the T7 promoter-driven acetolactate synthase sequence that had been stably integrated in the rice genome and generated C-to-T and G-to-A transitions. We used herbicide treatment as selection pressure for the evolution of the acetolactate synthase sequence, resulting in the enrichment of herbicide-responsive residues. We then validated these herbicide-responsive regions in the transgenic rice plants. Thus, our system could be used for the continuous evolution of gene functions to produce variants with improved herbicide resistance.

DOI 10.26508/lsa.202201538 | Received 30 May 2022 | Revised 16 September 2022 | Accepted 19 September 2022 | Published online 28 September 2022

Introduction

In crop plants, natural evolution and artificial selection (i.e., during domestication) relied on naturally occurring genetic variation. Increasing genetic variation in crop plants provides a basis for breeding and developing new traits of value. Genetic variation generated through random mutagenesis provides a basis for creating protein variants and useful alleles for breeding, and for evolution to develop new traits (Prohens, 2011; Hickey et al, 2019). Random mutagenesis has been used to improve crop traits; for example, radiation mutagenesis led to the generation of new traits that made the Green Revolution possible (Khush, 2001). However, radiation and chemical mutagenesis methods are not gene-specific and produce many mutations that are deleterious to the host plant.

Synthetic directed evolution relies on generating targeted or random genetic variability, followed by screening under selection pressure to identify beneficial mutations within a gene or pathway of interest (Simon et al, 2019; Morrison et al, 2020; Castle et al, 2021; Rao et al, 2021). Conventional tools for synthetic directed evolution mainly rely on in vitro diversification of a target gene, followed by screening for variants with improved functions (Smith, 1985; Chen & Arnold, 1991). These methods have been improved to generate different modifications. In single-site recombineering, exogenous oligonucleotides containing the desired sequence with flanking regions of homology are introduced to edit the target gene (Sharan et al, 2009). In the multiplexed automated genome engineering system, a library of oligonucleotides is introduced into an organism to target multiple regions of the genome, producing organisms with combinatorial diversity at the targeted regions (Wang et al, 2009). Similarly, directed evolution with random genomic mutations can be used to evolve multiple loci in a single targeted gene and its promoter regions by iteratively introducing oligonucleotides into an organism by electroporation (Nyerges et al, 2018). Although these approaches have broad applicability, they have several disadvantages, including low mutation rates and the need for labor-intensive, discontinuous steps.

These drawbacks were recently overcome by the development of in vivo mutagenesis systems that simultaneously perform gene diversification and selection to enable continuous directed evolution. The random mutation of plasmid-based genes is achieved by continuous diversification in host organisms with elevated global mutation rates, such as Escherichia coli strain XL1-red (Greener et al, 1997). A similar concept is used to engineer yeast via the OrthoRep system. Using this system, targeted yeast genes can be continuously diversified separately from the rest of the host genome sequence by placing these genes on orthogonal, non-genomic plasmids that replicate using error-prone DNA polymerases (Ravikumar et al, 2014, 2018).

The pioneering method of phage-assisted continuous evolution is used for the continuous evolution of proteins. Using this system, the mutational throughput in E. coli can be increased in vivo via a continuous process using phages, enabling the rapid evolution of gene-encoded molecules that can be linked to protein production (Eswell et al, 2011; Simon et al, 2019; Morrison et al, 2020). However,
this approach is best suited for phenotypes that can be linked to phage growth and cannot be directly applied to eukaryotes. These tools for directed evolution require extensive effort, long time scales, and ample resources to select for sought-after traits and are limited by the frequency and nature of mutagenesis.

To date, synthetic directed evolution has mainly been performed in single-celled organisms. Developing systems that induce localized sequence diversification at high efficiency will expand our ability to evolve traits of interest in multicellular eukaryotes (Simon et al., 2019; Hendel & Shoulders, 2021). Eukaryotic genes that have evolved in prokaryotic systems do not seem to exhibit the functions expected in single-celled organisms. Developing systems that induce localized mutagenesis within native cellular context is essential for evolving gene functions that affect traits of interest in multicellular eukaryotes such as crop plants.

DNA base editors, including cytidine base editors (CBEs) and adenosine base editors (ABEs), are powerful tools for introducing point mutations for genome engineering and synthetic evolution. Various naturally occurring cytosine deaminases, such as activation-induced cytidine deaminase (AID), RAPOBEC1, and pmCDAI, convert cytosine to uracil (Eid et al., 2018; Porto et al., 2020; Sakata et al., 2020). The U-G mismatch can be misread, resulting in C-G to T-A inversions. The base excision repair enzyme uracil DNA glycosylase reduces deamination activity by catalyzing the removal of uracil to initiate the base excision repair pathway (Krokan et al., 2002; Hegde et al., 2008). The uracil DNA glycosylase activity of CBES is inactivated by fusing the CBE with a uracil N-glycosylase inhibitor (UGI) (Cortazar et al., 2007; Rees & Liu, 2018). ABEs do not exist in nature; instead, they were developed using a directed evolution approach. ABEs can introduce A→T or G→C point mutations into living cells (Gaudelli et al., 2017). The recently evolved ABE8e has much faster catalytic activity and offers improved editing efficiency compared with the earlier ABE10.7 variant (Lapinaite et al., 2020; Richter et al., 2020).

Targeting the base editors to the desired sequence requires programmable systems that specifically recognize the sequence of interest. CRISPR/Cas has recently been applied to diversify gene sequences of interest. For example, CRISPR-X technology has been successfully used for protein engineering in mammalian cells. In this system, catalytically inactive dCas9 recruits variants of the cytidine deaminase base editor AID with MS2-modified sgRNAs to mutate endogenous targets (Hess et al., 2016). Using a similar strategy, Cas9 fused with base editors and sgRNAs targeting the coding sequence of OsALS were used to produce variants tolerant to the herbicide bispyribac sodium (BS) (Kuang et al., 2020). However, this system can only target short segments of DNA and not in a continuous fashion, which may be desired for the synthetic evolution of traits. In the EvolVR system, which has thus far been used only in bacteria, CRISPR-guided nCas9 fused with nick-translating error-prone DNA polymerase nick the target locus and performs mutagenesis at the target site (Halperin et al., 2018). Using CRISPR-X and EvolVR, high rates of mutagenesis at ~50-bp regions adjacent to the sgRNA target site have been achieved, but the mutation rate dropped with increasing distance from the sgRNA target window.

CRISPR-directed evolution (CDE) was recently performed by transforming rice (Oryza sativa) callus with CRISPR/Cas9 along with a pool of sgRNAs targeting the splicing factor locus OsSF3B1. Mutants were produced due to non-homologous end joining (NHEJ) under selection pressure imposed by the splicing inhibitor GEX1A. The recovered SF3B1 variants showed different levels of tolerance of GEX1A (Butt et al., 2019a, 2020b, 2021). Therefore, efficient mutagenesis requires dozens of sgRNAs to diversify the target gene. In addition, these CRISPR/Cas-based mutagenesis platforms have limited utility because of PAM sequence restrictions, the change of PAM sites by mutation, the change of sgRNA-binding sites by mutation, and the narrow genomic window adjacent to the sgRNA binding site, thus exhibiting an overall lack of efficiency for self-recurring continuous mutagenesis.

For continuous in vivo mutagenesis of target sequences, several kilobases long, base editors can be fused with bacteriophage T7 RNA polymerase (T7 RNAP) (Moore et al., 2018; Chen et al., 2020; Park & Kim, 2021). The target sequences are then placed under the control of the T7 RNAP promoter (pT7). Mutations are introduced by fusion proteins of the base editor with T7 RNAP, which recognizes the pT7 and reads through the target DNA. Chimeric T7 RNAP deaminase enzymes can perform continuous mutagenesis of several kilobases of DNA and have been used for directed evolution, including Mutat7 in bacteria (Moore et al., 2018; Álvarez et al., 2020; Park & Kim, 2021), TRIDENT in yeast (Cravens et al., 2021), and TRACE in mammalian cells (Chen et al., 2020). However, the use of this T7 RNAP deaminase editing system to introduce localized sequence diversification, leading to synthetic evolution, has not been demonstrated in plants.

Here, we established a T7 RNAP deaminase editor system for targeted mutagenesis for the first time in plant cells. In transient experiments, we targeted the GFP sequence and achieved localized sequence diversification at high efficiency. Most of the transitions were C-to-T substitutions. Single-vector and two-vector approaches for the evolution of OsALS allowed us to recover gene variants conferring herbicide resistance, a trait of interest in rice. We identified herbicide-responsive residues in the resulting acetolactate synthase (ALS) protein, and generated ALS variants conferring herbicide resistance. This technique opens up myriad possibilities for synthetic evolution in plants, including the development of crops with increased resistance to changing climate conditions, resistance to pests and pathogens, and improved productivity.

Results

Design, construction, and testing of the T7 RNAP–DNA base editor in Nicotiana benthamiana transient assays

Bacteriophage T7 RNAP transcribes DNA sequences under the control of the T7 promoter. A fusion protein of T7 RNAP with a cytidine deaminase (Editor) could continuously edit the DNA bases downstream of the T7 promoter (Target) (Fig. 1A). Once the cytidine deaminase (AID) converts C > U, the uracil–guanine (U-G) mismatch can be misread, resulting in a C > T or G > A transition. Alternatively,
an error-prone polymerase could be recruited through the mismatch-repair pathway, generating transitions and transversions near the lesion (Odegard & Schatz, 2006) (Fig 1A).

To test the ability of the T7 RNAP deaminase fusion to induce mutations within a target region, we constructed the AID-T7 RNAP-UGI fragment with hyperactivating mutations (P266L, G645A, and G645V).

**Figure 1: A cytidine deaminase T7 RNA polymerase fusion mediates editing in Nicotiana benthamiana transient assays.**

(A) Schematic diagram of the targeted mutagenesis system. The hyperactive cytidine deaminase AID is fused to the N-terminus of T7 RNA polymerase (Editor) and recognizes the T7 promoter inserted upstream of the target gene (Target). Once the T7 polymerase transcribes the target gene, ssDNA is available for AID deaminase activity, which converts cytosine to uracil. The U-G mismatch can be misread, resulting in C > T or G > A substitutions. Alternatively, error-prone polymerase can be recruited through the mismatch-repair pathway, generating transitions and transversions near the lesion. (B) Design of the pEditor and pTarget plasmids used for transient expression in N. benthamiana leaves. pEditor without AID was used as a control for editing. (C) The leaves were Agro-infiltrated with pEditor and pTarget plasmids. At 3 dpi, images were taken under UV light to detect GFP. (D) Sanger sequencing analysis of p35S-GFP targeted by AID-T7 RNAP showing C > T substitutions. (E) Amplicon deep sequencing analysis of GFP targeted by p35S-AID-T7 RNAP; p35S-GFP co-infiltrated with p35S-T7 RNAP was used as the mock control. Two independently infiltrated leaf samples were analyzed. The PCR products GFP-1 and GFP-2 were processed using an Illumina kit, and samples were sequenced on the NovaSeq platform. The p35S-T7 RNAP data were used as a control and subtracted during analysis to detect the true substitution rate during base editing. (F) Mutation percent frequency represents the distribution of all types of base edits found in GFP samples treated with p35S-AID-T7 RNAP. The data were analyzed using Geneious Prime. Source data are available for this figure.
OsALS in transgenic rice plants

We next expanded the applicability of the T7 RNAP deaminase editor system to mutate the genes of interest in stable plants using the rice gene acetolactate synthase (OsALS, LOC_Os02g30630, 1,935 bp). To this end, we cloned OsALS under the control of the T7 promoter in various rice expression vectors as a pTarget (Fig 2A) using two approaches. In the first approach, the single-vector system, we generated a vector harboring pEditor and pTarget. To avoid continuous targeting of pTarget in the progeny and to inhibit any possible activities in Agrobacteria, the two-vector system was designed such that pEditor and pTarget were in two different vectors (Fig 2A).

For the single-vector system, the AID-T7 RNAP-U7I fragment was cloned via Gateway recombination reaction under the control of the Ubiquitin promoter and inserted OsALS into this vector using unique restriction enzymes to generate the full-length vector p35S-pT7-ALS/pUBL-AID-T7pol (Fig 2A). The AID-T7pol fragment was removed from this vector and p35S-pT7-ALS only was used as the pTarget.

For the two-vector system, we also used the N. benthamiana expression vector p35S-AID-T7pol in rice. Rice callus was co-transformed with p35S-pT7-ALS and p35S-AID-T7pol via Agrobacterium-mediated transformation and a combination of hygromycin and G418 was used for callus selection and shoot regeneration; for the single-vector system, only hygromycin was used for selection and regeneration. Once transgenic plantlets were established in soil, we extracted the DNA and amplified the 2.2-kb fragment covering the complete OsALS sequence along with vector sequences. The amplified fragments were cloned into the pJET1.2 cloning vector and 10 reads were analyzed for mutagenesis assessment via Sanger sequencing (Fig 2B and C).

In the single-vector experiment, six of eight (75%) plants showed at least one read with a base substitution, whereas in the two-vector experiment, three of eight (37.5%) plants contained at least one read with base substitutions. For the single-vector approach, 2 of 10 reads in plant #1 contained C > T substitutions. These substitutions convert glycine 133 to aspartic acid (G to A) and histidine 415 to tyrosine (CAC to TAC) (Fig 2B). Plant #2 contained G > A and A > G substitutions, which convert glycine 133 to aspartic acid (GGC to GAC). The other mutation in plant #2 was a silent mutation at proline 248 (CCG to CCG) (Fig 2B). For the two-vector system, plant #1 contained two reads with G > A edits. These edits convert alanine 467 to threonine (GCA to ACA) and glycine 485 to aspartic acid (GGG to ACG) (Fig 2C). Plant #2 of the two-vector system contained two silent mutations in the same read: glutamine 102 (CAG to CAA) and serine 132 (TCC to TCT) (Fig 2C).

To perform detailed analysis of the substitution rate and to calculate the mutation percent frequency, we performed amplicon deep sequencing using the PacBio platform. We analyzed the data using the offline version of CRISPR-sub and subtracted p35S-pT7-ALS reads to calculate the substitution efficiency. Similar to the Sanger sequencing results, the substitution frequency was 1.06% and 0.53% for the single- and two-vector systems, respectively (Fig 2D). We also analyzed the frequency of each type of substitution using Geneious prime (Fig 2E). Interestingly, 36.1% of the mutations were G > A substitutions, 32.3% were C > T substitutions, 17.2% were T > N substitutions, 6.2% were G > Y substitutions, 6.1% were C > R
substitutions, and 2.0% were A > N substitutions (Fig 2E). Overall, these results indicate that T7 RNAP–mediated targeted mutagenesis can be performed in transgenic rice plants, albeit at low but sufficient frequencies.

Directed evolution of OsALS to identify herbicide-resistant mutations

To use and exploit the T7 RNAP deaminase editing system for trait engineering in plants, we use the rice gene OsALS for evolution of herbicide resistance. For this purpose, we performed Agrobacterium-mediated stable transformation of rice callus using our single-vector and two-vector systems harboring OsALS as the target locus. After co-cultivation, the callus was transferred to selection medium supplemented with 0.5 or 0.75 μM BS to inhibit the growth of wild-type callus and exert selection pressure to generate OsALS variants. For the mock control, the callus was grown on selection medium without BS. After 2 wk of selection, five independent actively growing callus pieces were pooled and used to amplify the OsALS sequence in the transgene (Fig 3A). The full-length amplicon was subjected to deep sequencing using PacBio sequencing technology and the data analyzed using the online version of CRISPR-Sub. To identify the BS-responsive regions of OsALS, we compared the reads from BS-treated samples with data from mock-treated sample reads. (E) Mutation percent frequency represents the distribution of base edits of all types found in the ALS samples with p35S-AID-T7 RNAP. The data are the means of the results for the single- and two-vector systems. Source data are available for this figure.
treated samples. The mutations that were common among BS-treated and mock-treated reads were subtracted during analysis.

Many more substitution mutations were observed at this locus in BS-versus mock-treated samples (Fig 3B). We considered these to be BS-responsive substitutions. Interestingly, these substitutions were enriched in the C- and N-terminal regions of the ALS protein. In samples transformed with the single-vector system and selected on 0.5 μM BS, we identified the enrichment of G > C causing the G629R substitution, G > A causing the A39T substitution, and A > G causing the K591E substitution. By contrast, in single-vector samples selected on 0.75 μM BS, we identified the enrichment of G > T causing the W548C substitution, and T > G causing F575C substitution (Fig 3B).

For samples transformed with the two-vector system and selected on 0.5 μM BS, we identified the enrichment of C > G causing the P70R substitution, G > C causing the G94R substitution, C > G causing the P70R substitution, A > G causing the K591E substitution, and G > A causing the G629S substitution. By contrast, for two-vector samples selected on 0.75 μM BS, we identified the enrichment of C > T causing the P93S substitution and G > C causing the G629R substitution (Fig 3B). Interestingly, G629 appeared to be an important residue in our analysis, as substitutions at this site were enriched in samples transformed with the single-vector system selected on 0.5 μM BS and in samples transformed with the two-vector system on both concentration of BS (Fig 3B). These results highlight the important residues in the OsALS sequence that are enriched in response to BS treatment.

ALS mutant variants confer variable levels of herbicide resistance

To further engineer the BS-responsive regions at the ALS sequence and to generate herbicide-resistant ALS variants, we mutate the OsALS of pTarget, p35S:ALS, via site-directed mutagenesis (Fig 4A). We selected nine different herbicide-responsive point mutations (Fig 3) that cause substitutions A39T, P70R, P93S, G94R, W548C, F575C, K591E, G629R, and G629S (Fig 4A). The vectors with these point mutations were used to transform embryonic rice callus to generate transgenic rice plants harboring ALS mutant variants.

To validate the herbicide resistance of ALS mutant variants, we tested the progeny of these transgenic plants. We conducted germination and root inhibition assays at several BS concentrations (Figs S2 and S3). The seeds of wild-type and progeny of p35S:ALS without ALS mutation were used as controls. We observed that the germination of WT seeds was severely inhibited at 0.75 μM BS, whereas the germination of p35S:ALS seed with WT-ALS sequence was affected at 0.75 μM BS and severely inhibited at 1 μM BS (Figs 4B and 5A).
and S2). We considered the 1 μM BS for screening of ALS mutant variants. We observed that the germination of ALS-W548C is insensitive to 1 μM BS and slightly affected at 2 μM BS (Fig 4B). The G629S and G629R are insensitive to 0.75 μM BS and showing resistance at 1 μM BS (Fig 4B). The F575C and K591E are resistant to BS at 0.75 and 1 μM concentration, however at lower levels than W548C, G629S, and G629R (Fig 4B). The other ALS mutant variants A39T, P70R, P93S, and G94R did not show resistance to BS like the p35S:ALS-WT control (Fig S2). We also tested the response of the seedlings of ALS mutant variants to different concentrations of BS (Fig S3). The BS resistance of seedlings is similar to our germination analysis, where W548C was showing the highest levels of resistance at 1.5 and 2 μM BS concentrations (Fig S3). The G629S and G629R are resistant to 1.5 μM BS and slightly affected at 2 μM BS (Fig S3). The F575C and K591E are showing slight resistant to BS at 1.5 μM and affected 2 μM concentration (Fig S3). All the other ALS mutant variants A39T, P70R, P93S, and G94R, including WT and p35S:ALS-WT control are sensitive to at 1.5 μM BS (Fig S3). Some of these ALS sites and the nearby sites G95, W548, and G629 (Endo et al, 2007; Okuzaki et al, 2007; Butt et al, 2017; Zhang et al, 2020) are known to confer resistance to BS. However, we identified the novel ALS BS-resistant substitutions W548C and G629R, and residues F575 and K591, which were not reported before.

In conclusion, we successfully used the T7 RNAP deaminase editing system to identify the BS-responsive regions in the ALS sequence and validate for herbicide resistance trait engineering in rice.

Discussion

Synthetic directed evolution involves targeted gene diversification, selection, and replication to obtain new functional biomolecules for use in research, biotechnology, and medicine (Zeymer & Hilvert, 2018; Simon et al, 2019; Morrison et al, 2020). The ideal synthetic directed evolution system must be performed in vivo in a continuous manner, preferably in the organism in which new variants are
Evolution of herbicide resistance using T7 RNAP base editor

Butt et al.

Materials and Methods

Plant materials

O. sativa L. ssp. japonica cv. Nipponbarest was used for the rice experiments. Wild-type N. benthamiana plants were used for the
N. benthamiana experiments. All wild-type and transgenic plants were grown in a greenhouse at 28°C/25°C day/night temperatures under a natural light/dark cycle.

Vector construction and cloning strategy

The pEditor fragment of AID_T7_UGI was amplified from the vector pcDNA3.1 (+)-Hyperactive AID-T7 RNA Polymerase-U2G-2A-td Tomato (138610; Addgene; [Chen et al, 2020]) using primers SphI_AID_T7_F and NsiI_XbaI_AID_T7_R, digested with SphI and NsiI, and cloned into the pK2GW7 vector using HindIII and PstI to generate the 35S promoter. The 35S promoter was cloned upstream of the T7 promoter in the rice pEditor vector. The terminator_CaMV poly(A) signal sequence were annealed and ligated upstream of the T7 promoter and RBS sequences were annealed with SbfI and XbaI for expression in N. benthamiana. oligos, with the T7 promoter and RBS sequences were annealed with SbfI and XbaI overhangs and cloned into plpLSLGFpSaci by replacing the upstream RIR. To clone the T7 promoter upstream of GFP for expression in N. benthamiana, oligos with the T7 promoter and RBS sequences were annealed with SbfI and XbaI overhangs and cloned into plpLSLGFpSaci by replacing the upstream RIR. To clone the T7 promoter upstream of the T7 promoter, oligos (35sP_HindF and 35sP_PstI_HindR) were designed to amplify the 35S promoter. The 35S promoter was cloned upstream of the T7 promoter using HindIII and PstI to generate the final plasmid 35S_pT7_pGFpNOSTerm.

For the single-vector system in rice, the 3xFLAG_AID_T7_UGI fragment was cloned into pK2GW7 via LR reaction under the control of the 35S promoter to generate pEditor p35S_3xFLAG_AID_T7_UGI_pK2GW7. To prepare p35S_3xFLAG_AID_T7_UGI_pK2GW7 without AID, the 35S promoter was cloned upstream of the T7 promoter in the rice pEditor vector. The terminator_CaMV poly(A) signal sequence were annealed and ligated with HindIII-SbfI overhangs and cloned into the rice pEditor vector. The cloning was performed with oligos ApaI_AvrII_ALS_F and XbaI_KpnI_ALS_R (1.963 Kb) for the single-vector system in rice, the 3xFLAG_AID_T7_UGI fragment was cloned into pK2GW7 via LR reaction under the control of the OsUbiquitin promoter to generate pEditor pUbI_3xFLAG_AID_T7_UGI_pK2GW7. Oligos for the T7-promoter_RBS_MCS_T7-terminator_CaMV poly(A) signal sequence were annealed and ligated with HindIII-SbfI overhangs and cloned into the rice pEditor vector. The 35S promoter was cloned upstream of the T7 promoter in the rice pEditor vector under the control of the OsALS sequence (LOC_Os02g30630) was amplified with oligos ApaI_Avrl_ALS_F and XbaI_Kpnl_ALS_R (1.963 Kb) and cloned under the control of pT7 or p35S-pT7 in rice pEditor vectors via digestion with ApaI and XbaI.

In rice for the two-vector system, pEditor and pTarget were transformed into rice as separate vectors. N. benthamiana pEditor p35S_3xFLAG_AID_T7_UGI_pK2GW7 was used as pEditor for rice. The AID_T7pol_UGI fragment was removed from p35S_pT7_ALS_AID_T7pol_UGI_pRGEB32 by BsrGI digestion to prepare p35S_pT7_ALS_pRGEB32 as pTarget. All the primers sequences used in this study are given in Table S1.

Agro-infiltration of N. benthamiana leaves and GFP imaging

Constructs harboring pEditor (p35S_3xFLAG_AID_T7_UGI_pK2GW7), pEditor without AID (p35S_3xFLAG_AID_T7_UGI_pK2GW7), pTarget (p35S_pT7_F + GEP_F + GEP_R2), and pTarget (p35S_pT7_ALS_AID_T7pol_UGI_pRGEB32 by BsrGI digestion to prepare p35S_pT7_ALS_pRGEB32 as pTarget. All the primers sequences used in this study are given in Table S1.

Rice transformation and mutant screening

Rice transformation plasmids were introduced into A. tumefaciens strain EHA105. Agrobacterium-mediated rice transformation was performed as described previously (Hiei & Komari, 2008; Butt et al., 2019b). Calli transformed with pRGEB32 were selected on medium containing 50 mg/l hygromycin B. calli transformed with pK2GW7 were selected on medium containing 150 mg/l G418 and regenerated on 100 mg/l G418. To select resistant variants of pTarget OsALS, BS was used at a concentration of 0.5 or 0.75 μM. After 1 wk of growth, when plants were established in soil, DNA was extracted from a leaf sample. To genotype pTarget OsALS, the HindIII_T7_PT_F and Sbf_T7_PT_R oligos were used to amplify the OsALS sequence from T-DNA (2.28 Kb) and cloned using a CloneJET PCR Cloning Kit (K1231). hiuqjpr The same fragment was sequenced with three to four overlapping primers to cover the entire coding sequence. At least 10 colonies were subjected to Sanger sequencing to analyze the mutation.

Ampli-on sequenc-in g

For ampi-on sequenc-in g of rice callus samples, five in-dependent callus pieces per treatment were collected after 2 wk of culture on medium without BS, with 0.5 μM BS, or with 0.75 μM BS. Equal amounts of samples from the same treatment groups were pooled together and used for DNA extraction. The ALS sequence from T-DNA was amplified using the primers HindIII_T7_PT_F and Sbf_T7_PT_R (2.28 bp). The purified PCR products were sequenced using PacBio and the same fragment was sequenced using Illumina TruSeq DNA Nano libraries according to the manufacturer’s instructions. The libraries were analyzed on the NovaSeq platform. The online tool CRISPR-Sub was used for data analysis (Hwang et al., 2020). pTarget-GFP samples co-infiltrated into N. benthamiana leaves with p35S_3xFLAG_T7_UGI_pK2GW7 without AID were used as the control during data analysis. The edited samples were compared with the control samples to calculate the substitution rate. For ampi-on sequenc-in g of rice callus samples, five in-dependent callus pieces per treatment were collected after 2 wk of culture on medium without BS, with 0.5 μM BS, or with 0.75 μM BS. Equal amounts of samples from the same treatment groups were pooled together and used for DNA extraction. The ALS sequence from T-DNA was amplified using the primers HindIII_T7_PT_F and Sbf_T7_PT_R (2.28 bp). The purified PCR products were sequenced using PacBio Sequel I technology. The data were analyzed using the offline version of CRISPR-Sub. To avoid the factor of sequencing errors and somaclonal variations we subtracted the mock-treated sample (without AID) reads from the experimental sample (with AID) reads to calculate editing efficiency. Data from the 0.5 μM BS- and 0.75 μM BS-treated samples were compared with data from untreated samples to identify the BS-responsive substitutions.

Geneious Prime software was used to calculate the mutation percent frequency of each type of base editing for the NovaSeq and PacBio data. The results of the data analysis are available as Additional_File_2 for Mutation rate GFP and ALS, Additional_File_3 for Mutation rate ALS and Additional_File_4 for Mutation rate AID.
for Substitution rate for ALS and Additional_File_4 for BS-responsive regions at the ALS.

**Generation of transgenic ALS mutant variants**

The herbicide-responsive point mutations were introduced into the ALS sequence of the vector pTarget, p35S_pT7_ALS_pRGE32, via site-directed mutagenesis (GenScript). The vectors with ALS mutations were transformed into wild-type rice callus and selected on medium containing 50 mg/l hygromycin B. The seeds of the progeny plants were used for herbicide resistance analysis.

**BS-resistance analysis of ALS mutant variants**

For seedling analysis, the seeds were de-husked, sterilized, and grown vertically on square plates containing ½ MS media (with 50 mg/l hygromycin B). Seeds of wild-type and pTarget, p35S_pT7_ALS_pRGE32, without ALS mutations were used as controls. After 3 d, seedlings with similar root growth were transferred to ½ MS plates supplemented with BS. The root tips were marked to observe growth. The seedlings were grown vertically for another 14-d and then imaged.

For germination analysis, the seeds were de-husked, sterilized, and grown vertically on square plates containing ½ MS media containing different concentrations of BS. Seeds of wild-type and pTarget, p35S_pT7_ALS_pRGE32, without ALS mutations were used as controls. The seeds were grown vertically for 14-d and then imaged.

**Supplementary Information**

Supplementary Information is available at https://doi.org/10.26508/lsa.202201538 vol 5 | no 12 | e202201538.

**Acknowledgements**

We would like to thank members of the genome engineering and synthetic biology laboratory at KAUST for their critical discussion and technical help in this work. This work was funded by KAUST-baseline funding to M Mahfouz.

**Author Contributions**

H Butt: data curation, investigation, visualization, methodology, and writing—original draft.

JLM Ramirez: investigation and methodology.

M Mahfouz: conceptualization, supervision, funding acquisition, investigation, visualization, and writing—original draft, review, and editing.

**Conflict of Interest Statement**

The authors declare that they have no conflict of interest.

**References**


Evolution of herbicide resistance using T7 RNAP base editor

Brett et al.