JIPB Journal of Integrative Plant Biology

Research Article
https://doi.org/10.1111/jipb.13543

Artificial evolution of *OsEPSPS* through an improved dual cytosine and adenine base editor generated a novel allele conferring rice glyphosate tolerance[OO]

Chen Zhang[1], Xue Zhong[1,2], Shaoya Li[1,2,3], Lei Yan[1], Jingying Li[1,2,3], Yubing He[1,2,3], Yong Lin[4], Yangjun Zhang[4] and Lanqin Xia[1,2,3]*

1. Institute of Crop Sciences (ICS), Chinese Academy of Agricultural Sciences (CAAS), Beijing 100081, China
2. CAAS/Hainan Yazhou Bay Seed Laboratory, National Nanfan Research Institute (Sanya), Sanya 572024, China
3. Key Laboratory of Gene Editing Technologies (Hainan), Ministry of Agricultural and Rural Affairs, Sanya 572024, China
4. Beijing Dabeinong Technology Group Co., Ltd, Beijing 10080, China

*Correspondence: Lanqin Xia (xialanqin@caas.cn)

Chen Zhang

Lanqin Xia

ABSTRACT

Exploiting novel endogenous glyphosate-tolerant alleles is highly desirable and has promising potential for weed control in rice breeding. Here, through fusions of different effective cytosine and adenine deaminases with nCas9-NG, we engineered an effective surrogate two-component composite base editing system, STCBE-2, with improved C-to-T and A-to-G base editing efficiency and expanded the editing window. Furthermore, we targeted a rice endogenous *OsEPSPS* gene for artificial evolution through STCBE-2-mediated near-saturated mutagenesis. After hygromycin and glyphosate selection, we identified a novel *OsEPSPS* allele with an Asp-213-Asn (D213N) mutation (*OsEPSPS*-D213N) in the predicted glyphosate-binding domain, which conferred rice plants reliable glyphosate tolerance and had not been reported or applied in rice breeding. Collectively, we developed a novel dual base editor which will be valuable for artificial evolution of important genes in crops. And the novel glyphosate-tolerant rice germplasm generated in this study will benefit weeds management in rice paddy fields.

Keywords: artificial evolution, dual cytosine and adenine base editor, nCas9-NG, rice (*Oryza sativa* L), near-saturated mutagenesis

Zhang, C., Zhong, X., Li, S., Yan, L., Li, J., He, Y., Lin, Y., Zhang, Y., and Xia, L. (2023). Artificial evolution of *OsEPSPS* through an improved dual cytosine and adenine base editor generated a novel allele conferring rice glyphosate tolerance. J. Integr. Plant Biol. **00:** 1–10.

INTRODUCTION

Rice is one of the most important food crops consumed by nearly half of the world population. Improving rice yield and quality will be of great significance to alleviate the food crisis and ensure global food security. In the process of rice production, farmland weeds often compete with rice plants for sunlight, water, nutrients, and growth. Meanwhile, they are intermediate hosts of pathogens and pests, which severely affect the yield and quality of rice. Glyphosate (*N*-(phosphonomethyl)-glycine) is one of the most extensively used broad-spectrum herbicides for weeds control in crop production (Steinrücken and Amrhein, 1980). The primary action mode of glyphosate is though competitive binding and inhibition of the enzyme 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), which catalyzes one important step in the shikimate pathway by using the phosphoenolpyruvate (PEP) as substrate (Steinrücken and Amrhein, 1980). In plants, as much as 20% of all fixed carbon flows through the shikimate pathway which results in the synthesis of the aromatic amino

acids Tyr, Phe, and Trp, as well as tetrahydrofolate, ubiquinone, vitamins K and E (Baerson et al., 2002). Following the aging of the population, decreasing labor resources, and mechanized crop production, cultivating glyphosate-tolerant rice varieties suitable for direct seeding and mechanized production, which is more labor-saving and cost-effective, has becoming a major priority in rice breeding programs. However, due to the fact that nearly all the currently cultivated rice varieties are sensitive to glyphosate treatment, application of glyphosate in the rice paddy field for weed control frequently leads to yield losses (Jin et al., 2022). Therefore, it is essential to develop glyphosate-tolerant rice varieties in order to secure rice production and quality.

At present, the commonly used strategy to generate glyphosate-tolerant crop plants is to introduce exogenous glyphosate-resistant genes into crops through transgenesis. For example, an adenosine triphosphate-binding cassette (ABC) transporter (EcABCC8) was identified from *Echinochloa colona*; overexpression of this gene could confer rice plants glyphosate tolerance (Pan et al., 2021). However, due to the long biosafety assessment process and public concerns about genetically modified food crops, so far, no glyphosate-tolerant transgenic rice varieties have yet been released for commercialization. Recently, glyphosate-tolerant rice plants were recovered by prime editing of endogenous rice *EPSPS* gene (*OsEPSPS*) to generate alleles with either PS (P177S) or TAP-IVS (T173I, A174V, and P177S) mutations in three independent laboratories, respectively (Perotti et al., 2019; Jiang et al., 2022; Qiao et al., 2022). Furthermore, a novel glyphosate-tolerant rice allele, m*OsEPSPS*-GATIPS (G172A, T173I, and P177S), was created through homology donor repair (HDR) strategy by using a repair template harboring the desired mutations (Sony et al., 2023). These results indicated that more novel alleles of *OsEPSPS* gene could be exploited for non-transgenic glyphosate-tolerant rice breeding.

Single nucleotide polymorphism (SNP) is the genetic basis for the improvement of agriculturally important traits in crop plants. Base editing enables a single nucleotide substitution in a target gene in a sequence-specific manner to generate either loss-of-function or gain-of-function mutations, thus greatly accelerating functional annotation, crop improvement, *de novo* domestication, or directed evolution of target genes in crop plants (Bharat et al., 2020; Kuang et al., 2020; Li et al., 2020; Zeng et al., 2020; Xu et al., 2021; Yan et al., 2021; Tan et al., 2022). Three base editors are currently in use: cytosine base editors (CBEs) for C:G to T:A transition (Komor et al., 2016), adenine base editors (ABEs) for A:T to G:C transition (Gaudelli et al., 2017), and C-to-G base editors (CGBEs) for C:G to G:C transversion (Kurt et al., 2021; Zhao et al., 2021). So far, impressive progresses have been achieved in improving the performance of CBEs and ABEs (Li et al., 2023). For example, a base editor, PhieCBE, consisting of a codon-optimized evolved cytidine deaminase, evoFERNY, which is derived from a wild-type deaminase Anc656, exhibited the highest editing efficiency (86.3%) in an editing window of C_3–C_{10} in the respective target genes (Zeng et al., 2020). And another cytidine deaminase with wide editing window, a human APOBEC3A (hA3A), when fused with nCas9 (D10A), was reported to improve the efficiency of base editing in wheat, rice, and potato with a 17-nucleotide editing window at all examined sites (Zong et al., 2018). Subsequently, a BEACON base editing system composed of a dCas12a and a hA3A with W98Y/W104A mutations (mhA3A) significantly improved C-to-T editing efficiency in comparison with the original hA3A (Wang et al., 2020). For ABEs, monomer TadA8e, which has eight amino acid mutations in a deoxyadenosine deaminase TadA, significantly improved the efficiency of A-to-G base editing (Richter et al., 2020). In addition, the efficiencies of CBEs or ABEs editors have been further improved by using a surrogate system to enrich the precisely edited events (Xu et al., 2020), optimizing the lengths of linkers between fusion proteins, increasing the numbers of uracil DNA glycosylase inhibitor (UGI) and the nuclear localization sequences (NLS) (Zeng et al., 2020), respectively.

CBEs and ABEs enable the artificial evolution of agriculturally important genes in crops to generate novel gene/allele resources and germplasm. For example, a CBE- and/or ABE-mediated gene evolution (BEMGE) strategy was developed to obtain novel allelic variants in *OsALS* and *OsACC* in rice (Kuang et al., 2020; Liu et al., 2020; Wang et al., 2022). In addition, prime editing was also employed for artificial evolution through saturated mutagenesis (Xu et al., 2021). Furthermore, dual base editors for saturated targeted endogenous mutagenesis editors (STEME or STEME-NG), were developed with the capability of simultaneously performing C-to-T and A-to-G transitions by using one single-guide RNA (sgRNA) in rice (Li et al., 2020). Development of an effective dual cytosine and adenine base editor is highly desirable in saturated mutagenesis of targeted functional domains or regulatory regions of important genes in crops. However, due to the strict requirement of protospacer adjacent motif (PAM) of canonical NGG for nCas9(D10A) and the relatively low efficiency of dual base editor with nCas9(D10A)-NG (STEME-NG) (Li et al., 2020), the wide application of the dual base editing system in artificial evolution for crop improvement still remains challenging.

In this study, in order to improve the efficiency of the dual cytosine and adenine base editor and expand its editing window, we developed several surrogate two-component composite base editing systems (STCBEs) through fusions of different effective cytosine and adenine deaminases with nCas9(D10A)-NG, respectively. We first evaluated the activities of these STCBEs in rice protoplasts by using six endogenous targets. Then, we targeted a rice endogenous *OsEPSPS* gene for near-saturated mutagenesis by the more effective dual cytosine and adenine base editor, STCBE-2, developed in our study. Through *Agrobacterium*-mediated transformation and both hygromycin and glyphosate selection, we successfully generated a novel *OsEPSPS* allele which conferred rice plants higher level of glyphosate

tolerance. We truly believe the developed novel dual cytosine and adenine base editor will also have significant potentials for artificial evolution of other important traits in crops. And the generated novel glyphosate-tolerant rice germplasm, which has not been reported or applied in rice breeding, will be of great significance for weed management in rice paddy fields to facilitate sustainable agricultural development.

RESULTS

Development of a STCBE with improved performance

We first attempted to fuse different cytidine deaminases and adenosine deaminases to establish the novel STCBEs to enable simultaneously performing C-to-T and A-to-G transitions with improved editing efficiency by using a single sgRNA. We fused either mhA3A (hA3A with W98Y/W104A mutations) (Zong et al., 2018; Wang et al., 2020) or evo-FERNY (Zeng et al., 2020), and in combination with a monomer TadA8e (Wei et al., 2021) to the N terminus of nicked SpCas9-NG (nCas9-NG) (Ren et al., 2019; Zhong et al., 2019), followed by two copies of UGI at the C terminus of nCas9-NG, to generate STCBE-1 and STCBE-2, respectively (Figure 1A). Alternatively, we fused either mhA3A or evo-FERNY and a TadA-TadA8e dimer, to the N terminus of nCas9-NG, followed by one copy of UGI at the C terminus of nCas9-NG, to generate STCBE-3 and STCBE-4, respectively (Figure 1A). The broad PAM compatibility of nCas9-NG makes it feasible to generate the high-density mutations in a specific target gene/sequence (Ren et al., 2019; Zhong et al., 2019). Furthermore, we used a surrogate system to restore the defective *hptII* gene into the functional one which could confer hygromycin resistance and enrich the precisely edited plants (Xu et al., 2020; Li et al., 2022a). Thus, the gene

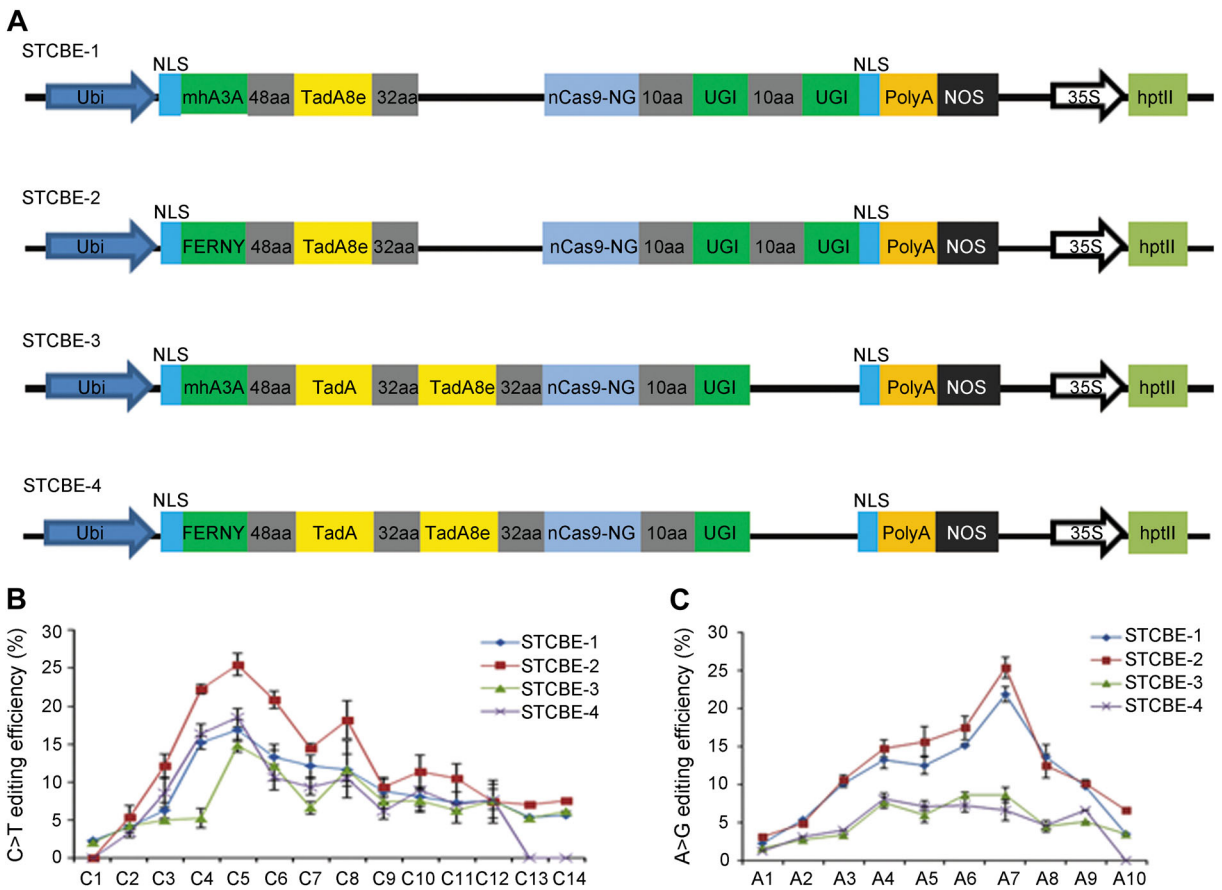


Figure 1. Development of an efficient surrogate two-component composite base editing system (STCBE) for simultaneous cytidine and adenine base editing

(A) A schematic showing the architectures of STCBE-1, STCBE-2, STCBE-3, and STCBE-4, respectively. From left to right, Ubi, ubiquitin promoter; NLS, nuclear localization signal; mhA3A, a mutated human deaminase APOBEC3A; FERNY, evoFERNY derived from a wild-type deaminase Anc656; TadA8e, a deaminase variant containing eight amino acid mutations in deoxyadenosine deaminase TadA7.10; TadA, a natural adenine deaminase; nCas9-NG, a Cas9 nickase nCas9-NG(D10A) recognizing NG protospacer adjacent motif; UGI, a uracil DNA glycosylase inhibitor; PolyA, polyadenylation signal; NOS, NOS terminator; 35S, cauliflower mosaic virus 35S promoter; *hptII*, a defective *HptII* gene. The boxes in gray indicate different lengths of amino acid linkers. **(B)** The editing efficiencies and editing windows of different STCBEs at C1–C14 sites of the tested six endogenous targets. Error bars indicate the standard error of three independent biological replicates. **(C)** The editing efficiencies and editing windows of different STCBEs at A1–A10 sites of the tested six endogenous targets. Error bars indicate the standard error of three independent biological replicates.

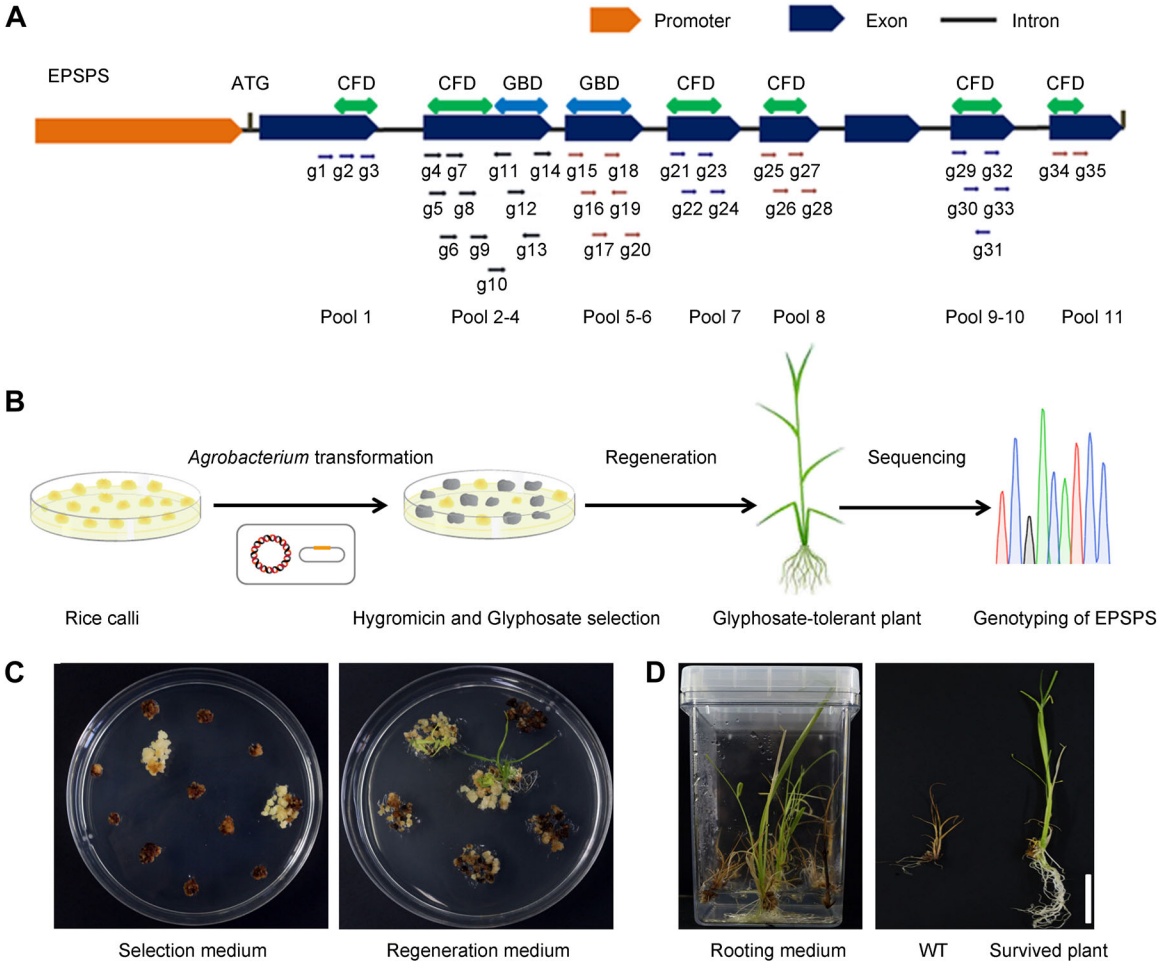
encoding hygromycin (*HptII*) in all the four STCBEs was in a defective form of *hptII* (Figure 1A).

To examine the base editing activities of these four STCBEs in rice, six sgRNAs targeting different genes/encoding regions including *OsFTL2*, *OsGRF4*, *OsSD1-1, OsSD1-2, OsEPSPS-1*, and *OsEPSPS-2* were designed (Table S1) and cloned into OsU3-sgRNA cassette in each of the four STCBEs, respectively, and then transiently expressed in rice protoplasts. The targeted fragments of the selected target genes/encoding regions were amplified by polymerase chain reaction (PCR), then for Hi-tom high-throughput sequencing (Liu et al., 2019). Our sequencing results indicated that STCBE-2 produced C-to-T transitions more efficiently than the other three STCBEs (Figure 1B). The C-to-T base editing efficiency of STCBE-2 could reach up to 25.5%, while the highest editing efficiencies of STCBE-1, STCBE-3, and STCBE-4 were 16.85%, 14.80%, and 18.50%, respectively. All the four STCBEs had a most efficient C-to-T editing efficiency at target site *OsGRF4*, while the editing efficiencies at *OsFTL2*, *OsSD1-1*, *OsSD1-2*, *OsEPSPS-1*, and *OsEPSPS-2* target sites were relatively lower (Figure S1A), consistent with previous reports that the editing efficiency was dependent on the innate nature of target genes/sequences (Li et al., 2022b). Further, the cytidine deaminase, evoFENRY, in STCBE-2 and STCBE-4 exhibited higher editing efficiency of C-to-T than mhA3A in STCBE-1 and STCBE-3, respectively (Figure 1B). Furthermore, among the four STCBEs developed in this study, STCBE-2 also had the highest A-to-G base editing efficiency (25.30%), while the highest editing efficiencies of STCBE-1, STCBE-3, and STCBE-4 were 21.90%, 8.70%, and 7.30%, respectively (Figure 1C). In addition, all the four STCBEs exhibited a more efficient A-to-G editing at *OsFTL2*, whereas the editing efficiencies of other target sites, Os*GRF4*, *OsSD1-1, OsSD1-2, OsEPSPS-1*, and *OsEPSPS-2*, were relatively lower (Figure S1B). In particular, the A-to-G base editing efficiencies of STCBE-1 and STCBE-2 with monomeric TadA8e were significantly improved in comparison to STCBE-3 and STCBE-4 which were composed of a dimer TadA-TadA8e (Figure 1C), consistent with a previous report (Wei et al., 2021). As for the editing window, in comparison with STCBE-4 and the previously reported STEME-NG (Li et al., 2020), STCBE-2 exhibited a wider editing window which encompassed C_1–C_{14} (counting the end distal to the PAM as position 1) of the protospacer sequences for C-to-T editing, and A_1 to A_{10} for A-to-G editing, respectively (Figures 1B, C, S2). Furthermore, we evaluated the performance of STCBE-2 by using the same set of 16 targets in four endogenous genes as reported for testing the activity of STEME-NG (Table S1) in rice protoplasts. Our result indicated that STCBE-2 could enable the C-to-T and A-to-G base editing efficiencies reaching up to 23.1% and 24.2%, respectively. In comparison to STEME-NG, STCBE-2 increased the efficiencies of C-to-T and A-to-G base editing by 2.9-(23.1%/7.92%) and 13.2-fold (24.2%/1.84%), respectively, further indicating that STCBE-2 outperformed STEME-NG in conducting simultaneous C-to-T and A-to-G base editing (Figure S2). Collectively, these results indicate that STCBE-2 is a more effective dual cytosine and adenine base editor, and can be used for saturated mutagenesis of the agriculturally important genes in rice for artificial evolution (Figure 1).

Near-saturated mutagenesis of *OsEPSPS* by STCBE-2 generated a novel allele conferring rice glyphosate tolerance

Next, to further prove the effectiveness of STCBE-2, which is capable of performing C-to-T and A-to-G transition simultaneously by using a single sgRNA, we chose an endogenous *OsEPSPS* gene for saturated mutagenesis in rice stable lines. Combined with the protein structure prediction of OsEPSPS and bioinformatics analysis, we designed a total of 35 sgRNAs, tiled on both the forward and reverse DNA strands, to target the conserved functional domains of OsEPSPS and the predicted glyphosate-binding domains (Figure 2A; Tables S2, 3). The sgRNA library was divided into 11 pools; each was composed of two to four sgRNAs, depending on the length of the targeted region (Figure 2A). The sgRNAs in each pool were cloned into the STCBE-2 vector individually, and then introduced into *Agrobacterium* cells. The *Agrobacterium* cells harboring these vectors in each pool were mixed in equal amounts and introduced into the calli of the rice cultivar *cv* Zhonghua 11 through *Agrobacterium*-mediated transformation. Regenerated plants recovered from both hygromycin and glyphosate selection were genotyped by amplicon sequencing of the endogenous *OsEPSPS* gene to identify the evolved *OsEPSPS* alleles (Figure 2B).

In total, approximately 2,600 independent calli were transformed with 11 pools. These calli were then screened on medium added with both hygromycin (50 mg/L) and glyphosate (10 mg/L) at callus induction and regeneration stages. Hygromycin and glyphosate selection could potentially inhibit the growth of wild-type rice calli (Figure 2C, D). After 4–6 weeks of selection, whereas no regenerated plants were obtained from other pools, we recovered nine independent lines derived from the transformation of Pool 6. All these lines had the precisely edited *HptII* gene which restored its function and conferred rice plants hygromycin resistance (Figure S3). Among these lines, two independent lines (line 1: P6-D213-1 including three individual plants; line 2: P6-D213-2, including five individual plants) were identified as lines carrying the precisely edited *OsEPSPS* by amplicon sequencing (Figures 2E, 3A), whereas the other lines were wild-type plants. The escape of wild-type plants under the strict selection regime of both hygromycin and glyphosate might be due to the growth of these plants along with the edited *OsEPSPS* plants. Furthermore, our sequencing results also indicated that all the regenerated *OsEPSPS* edited plants carried the same sgRNA-g19, and the C-to-T transition in *OsEPSPS* was at a heterozygous status (Figures 2E, 3A; Table S2). The C_{10}–T_{10} transition at the sgRNA-g19 targeting site in these *OsEPSPS* edited lines resulted in a substitution of an amino acid aspartic acid (Asp; D) with an asparagine (Asn; N) at



E

Target site	No. of transformed calli	No. of independent transgenic plants	No. of independent edited plants	Base-editing type	Base-editing efficiency	Amino acid variation type
EPSPS-D213	120	9	2	C-to-T	22.2%(2/9)	D213N

Figure 2. Near-saturated mutagenesis of *Oryza sativa* 5-enolpyruvylshikimate-3-phosphate synthase (OsEPSPS) by surrogate two-component composite base editing system 2 (STCBE-2) generates glyphosate-tolerant rice plants

(A) Schematic diagram of OsEPSPS gene structure and a tiled single-guide RNA (sgRNA) library designed to target the conserved domains (CFD; green bidirectional arrow) of OsEPSPS and the predicted glyphosate-binding domains (GBD; blue bidirectional arrow). A total of 35 sgRNAs were designed and divided into 11 pools. **(B)** A transformation process for screening potential glyphosate-tolerant rice mutant. **(C)** The induction and regeneration of calli after *Agrobacterium*-mediated rice transformation. Calli on selection medium containing 10 mg/L glyphosate and 50 mg/L hygromycin and glyphosate-resistant seedlings on regeneration medium containing 10 mg/L glyphosate after 6 weeks of culture. **(D)** Glyphosate-resistant plants cultured in rooting medium containing 10 mg/L glyphosate for 2 weeks. Bar = 5 cm. **(E)** Summary of the precise base editing efficiency at the target site of *OsEPSPS* in T0 plants in rice. The editing frequency is estimated by dividing the number of precisely edited independent lines by the number of independent transgenic lines detected.

position 213 amino acid (OsEPSPS-D213N), a key amino acid located in between the α-helix and β-sheet in the predicted glyphosate-binding domain (Figures 2A, 3B). The change of acidic Asp with negative charge to neutral Asn might decrease the glyphosate-binding activity of OsEPSPS, and thus conferred rice plants with glyphosate tolerance.

Evaluation of the glyphosate tolerance of rice plants with the novel *OsEPSPS* allele

Following segregation, we obtained the *OsEPSPS*-D213N transgene-free homozygous lines in the T_1 generation. To determine if the transgene sequences of *Cas9*, *sgRNA* cassette and *hptII*, were present in the T_1 lines, we performed PCR amplification using the primer sets designed to specifically amplify *nCas9-NG*, *sgRNA* cassette and *hptII* sequences, respectively (Figure S4). We successfully recovered *nCas9-NG*, *sgRNA* cassette and *hptII* transgene-free homozygous lines with *OsEPSPS*-D213N from the T_1 progenies (Figure S4). Subsequently, the potential sgRNA-dependent off-target sites for sgRNA19 were predicted using the web tool CRISPR-GE (Clustered Regularly Interspaced Palindromic Repeats – Genome Editing). The two

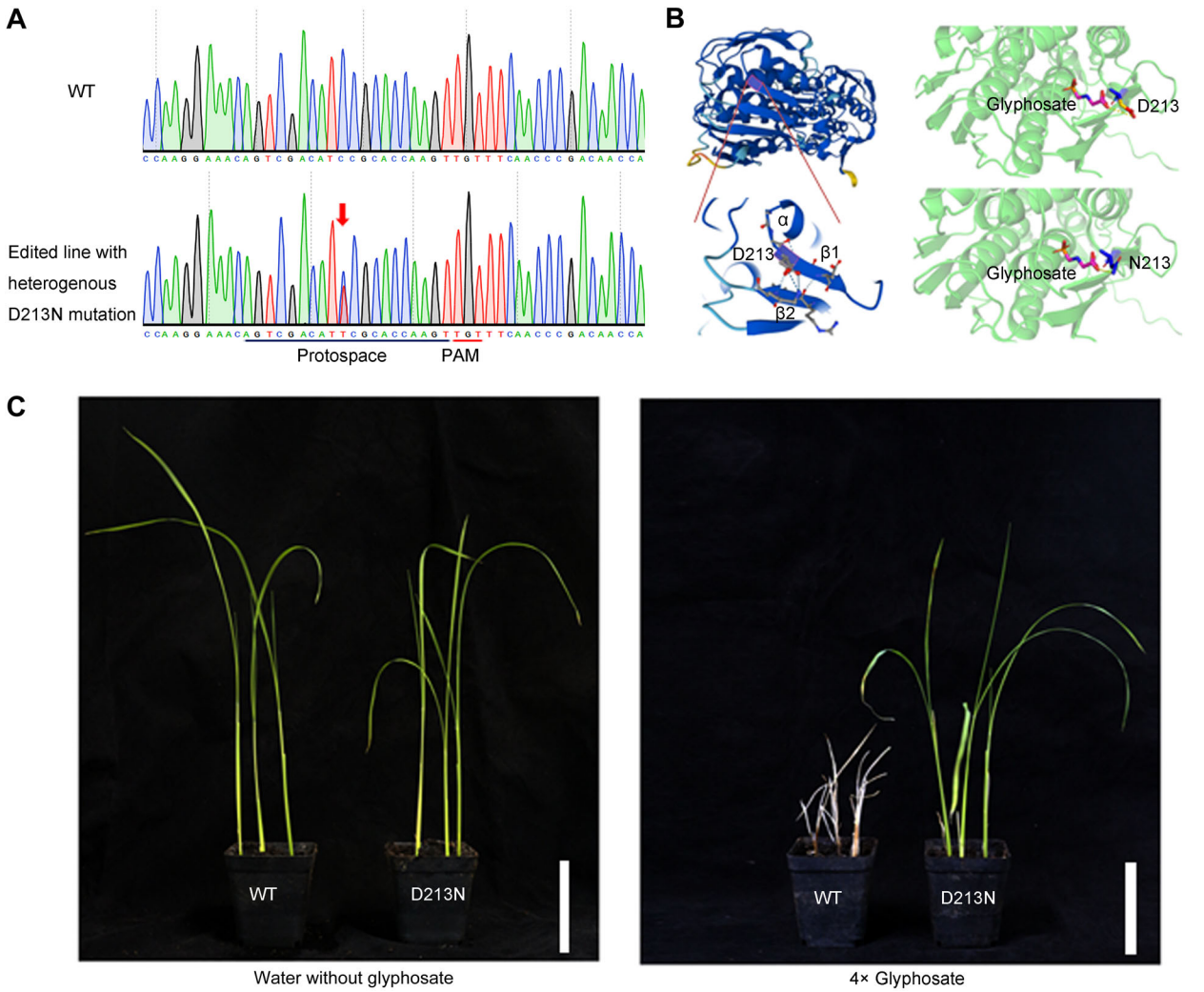


Figure 3. Evaluation of the glyphosate tolerance of rice plants with the novel *Oryza sativa* 5-enolpyruvylshikimate-3-phosphate synthase (OsEPSPS) allele

(A) Sanger sequencing of the survived plant in the T^0 generation. The mutant site is marked by a red arrow. **(B)** The structural model of the OsEPSPS protein and location of aspartic acid (Asp; D) at position 213 amino acid (D213) in the predicted binding domain of glyphosate. The residue (D213/N213) and glyphosate are labeled as colored sticks. α, α-helix; β, β-sheet. **(C)** Glyphosate resistance of the transgene-free homozygous lines with OsEPSPS-D213N allele. Water without glyphosate: the growth of the seedlings of wild-type and transgene-free homozygous lines with OsEPSPS-D213N allele at 14 d after spraying with water. 4× Glyphosate, the growth of the seedlings of wild-type and transgene-free homozygous lines with OsEPSPS-D213N allele at 14 d after spraying with glyphosate at four times field-recommended application dose (5,400 g a.i./ha). Bar = 7 cm.

potential off-target sites with 4-nt mismatches regarding the sgRNA19 were examined, and no off-target effect was detected in the tested *OsEPSPS*-D213N lines (Table S4).

To further validate the glyphosate tolerance of the rice lines with *OsEPSPS*-D213N allele, the transgene-free homozygous plants at the three-leaf stage were sprayed with glyphosate at a dose of 4× field-recommended dose (5,400 g a.i./ha). After 2 weeks, the *OsEPSPS*-D213N seedlings exhibited normal growth, whereas the wild-type seedlings died (Figure 3C). However, whereas the homozygous lines displayed growth defects at seedling stage (Figure S5) and sterility with no seed setting at all, the heterozygous *OsEPSPS*-D213N plants exhibited a slightly delayed flowering compared with the wild-type when grown in the rice paddy field (Figure S6). This result indicated that the homozygous *OsEPSPS*-D213N allele might affect its activity in the shikimate pathway and thus leading to growth defects and sterility. The potential mechanism underlying this phenomenon remains to be exploited in the future. Nevertheless, our data suggest that STCBE-2 enables generation of a novel *OsEPSPS*-D213N allele which confers rice a higher level of glyphosate tolerance through artificial evolution, and could be employed as an effective tool in directed evolution of other agriculturally important genes in rice and other crops for genetic improvement.

DISCUSSION

Development an efficient dual cytosine and adenine base editor enabling simultaneous C-to-T and A-to-G transitions by using a single sgRNA would be highly desirable for artificial evolution of agriculturally important genes in crop improvement. In this study, through fusions of different effective cytosine and adenine deaminases with nCas9-NG, we developed a series of STCBEs for simultaneous C-to-T and A-to-G base editing. Testing the base editing activity of four STCBEs in rice protoplast by using six endogenous targets demonstrated that STCBE-2, which is composed of evoFERNY and TadA8e fused to the N-termini of nCas9-NG, and two copies of UGI at the C-termini of nCas9-NG, exhibited significantly improved efficiencies of 25.5% for C-to-T and 25.3% for A-to-G base editing, respectively. In comparison the previously reported dual cytosine and adenine base editor, STEME-NG, which was composed of hA3A (APOBEC3A), TadA7.10, and nCas9-NG, and exhibited editing activity of 7.92% for C-to-T and 1.84% for A-to-G in rice protoplast (Li et al., 2020), STCBE-2 substantially increased the simultaneous C-to-T and A-to-G base editing efficiencies by 2.9- (23.1%/7.92%) and 13.2-fold (24.2%/1.84%), respectively (Figure S2). Furthermore, the editing byproducts of STCBE-2, which mainly consisted of C-to-G and C-to-A, occurred at relatively lower frequencies of 1.5% and 1.7% (Figure S1C), respectively, in comparison with those of 13.73% and 4.41% for STEMEs, respectively (Li et al., 2020). In addition, no indels were detected in different target sites of STCBE-2 and STCBE-4 in rice protoplasts, while the indel frequencies of STCBE-1 and STCBE-3 were 1.10% and 3.50%, respectively, demonstrating the combination of FERNY and TadA8e increased the purity of editing products (Figure S1C). Moreover, STCBE-2 extended the A-to-G editing window of A_1–A_{10}, compared to the STEMEs in a window of A_4–A_8 (Li et al., 2020). Finally, it is worth mentioning that both STCBEs and the previously reported STEME-NG (Li et al., 2020) displayed a compromised activity in comparison with STEMEs, indicating that while nCas9-NG could expand the editing scope of base editing, it might decrease the editing activity. Given that the presence of limited target sites with a canonical NGG PAM within the encoding region/functional domains of a target gene when nCas9 were used, we proposed that STCBE-2, which is capable of dual base editing at target with NG PAM, would be more favorable for artificial evolution of important protein-coding genes to generate new alleles or alternative functional activities through saturated mutagenesis.

In addition, a novel glyphosate-tolerant allele, *OsEPSPS*-D213N, was successfully exploited through STCBE-2-mediated artificial evolution of endogenous *OsEPSPS* gene in rice in this study. To understand the molecular basis of glyphosate tolerance mutation, the structural model of the OsEPSPS protein was predicted using the web tools AlphaFold Protein Structure Database (www.alphafold.com). Aspartic acid (D) is an acidic amino acid with negative charge, while Asn (N) is neutral and has no charge at all. We speculate that the mutation D213N might alter the molecular structure of β-sheet in OsEPSPS and cause the conformational changes of the glyphosate-binding domain (Figure 3B). In addition, the change of the hydrogen bond in the interaction between OsEPSPS and glyphosate might decrease the glyphosate-binding activity. Therefore, through near-saturated mutagenesis of *OsEPSPS* by STCBE-2, we exploited a functional allele *OsEPSPS*-D213N conferring rice glyphosate tolerance, demonstrating the potential of STCBE-2 in artificial evolution of agriculturally important genes for crop improvement. Moreover, the novel transgene-free rice germplasm harboring *OsEPSPS*-D213N allele exhibited a substantially higher level of glyphosate tolerance (Figure 3C) and could potentially be applied for weed control in rice paddy fields.

In conclusion, we successfully engineered an efficient dual cytosine and adenine base editing system, STCBE-2, for artificial evolution of agriculturally important genes in rice. The established dual cytosine and adenine base editing system may also be extended to the artificial evolution of important traits in other food crops through saturated mutagenesis. Furthermore, through STCBE-2-mediated saturated mutagenesis, we exploited a novel glyphosate-tolerant *OsEPSPS*-D213N allele which has not been documented in existing natural resources or applied in rice breeding. The novel transgene-free glyphosate rice germplasm generated in our study will have potential in weed management in rice paddy fields, thus facilitating the direct seeding and mechanized production of hybrid rice.

MATERIALS AND METHODS

Construction of the base editing vectors

To establish the hygromycin-based surrogate system, we first mutated the *hptII* gene at the start codon (ATG) to ACG and cloned it into our basic vector pCXUN-Ubi-NLS-nCas9(D10A)-NG-PolyA-Nos, to generate a pCXUN-Ubi-NLS-nCas9(D10A)-NG-PolyA-Nos-35S-*hptII* vector (hereafter referred to as the BE vector-Original). For STCBE-1, the cytidine deaminase (hA3A with W98Y/W104A mutations), adenosine deaminases (TadA8e), and two copies of UGI were amplified from BEACON (Wang et al., 2020), pCXUN-Ubi-TadA-Tad8e-nCas9 (D10A) preserved in our laboratory, and PevoFERNY-NG (Zeng et al., 2020), respectively. For STCBE-2, the cytidine deaminase (FERNY) and two copies of UGI, and adenosine deaminases (TadA8e), were amplified from PevoFERNY-NG, and pCXUN-Ubi-TadA-Tad8e-nCas9 (D10A) preserved in our laboratory, respectively. For STCBE-3, the cytidine deaminase (hA3A with W98Y/W104A mutations), and adenosine deaminases (a dimer composed of TadA and TadA8e), and one copy of UGI were amplified from BEACON, pCXUN-Ubi-TadA-Tad8e-nCas9 (D10A) preserved in our laboratory, and PevoFERNY-NG, respectively. For STCBE-4, the cytidine deaminase (FERNY) and

one copy of UGI, and adenosine deaminases (a dimer composed of TadA and TadA8e), were amplified from PevoFERNY-NG, and pCXUN-Ubi-TadA-Tad8e-nCas9 (D10A) preserved in our laboratory, respectively. Polymerase chain reaction was performed using high-fidelity DNA polymerase Phusion (New England BioLabs). These PCR products were assembled into the BE vector-Original by One Step Cloning (ClonExpress II One Step Cloning Kit; Vazyme), respectively. Each OsU3-sgRNA expression cassette was then cloned into the STCBEs backbone, respectively. The spacers were inserted into *Pme*I (NEB) digested STCBEs-sgRNA. The primer sets used for construction of these vectors are listed in Table S5.

Protoplast transfection

We used the *Japonica* rice cultivar (*Japonica* cv Zhonghua11) to prepare protoplasts. Protoplast isolation and transformation were performed as previously described (Ren et al., 2019). Rice protoplasts were isolated from 14-d-old seedlings grown on 1/2 Murashige and Skoog (MS) medium. They were transfected by STCBEs-sgRNA (*OsGRF4*), STCBEs-sgRNA (*OsFTL2*), STCBEs-sgRNA (*OsSD1-1*), STCBEs-sgRNA (*OsSD1-2*), STCBEs-sgRNA (*OsEPSPS-1*), STCBEs-sgRNA (*OsEPSPS-2*), and STCBEs without sgRNA as a control. Further, STCBE-2 constructs containing the same set of 16 targets as reported for testing the activity of STEME-NG (Li et al., 2020) were also generated in order to evaluate the performance of STCBE-2 in comparison with STEME-NG in rice protoplasts. These targets included four targets for each *OsAAT*, *OsCDC48*, *OsDEP1*, and *OsOD* gene with PAMs of NGA, NGT, NGC, and NGG, respectively (Table S1). In each transfection, 20 μg of STCBEs-sgRNA plasmid DNA was introduced into approximately 1×10^6 protoplasts by polyethylene glycol (PEG)-mediated transfection. The transfected protoplasts were incubated at 25°C, and at 36–48 h post-transfection they were collected and genomic DNA was amplified by PCR, then for Hi-Tom high-throughput sequencing (Liu et al., 2019).

sgRNA design and sgRNA library assembly for evolving *OsEPSPS*

The genomic DNA sequence of *OsEPSPS* (LOC_Os06g04280) was PCR amplified from the *Japonica* rice cultivar Zhonghua11 with high-fidelity DNA polymerase Phusion (NEB) and determined by Sanger sequencing. The conserved functional domains of OsEPSPS and the glyphosate-binding domains were predicted using the web tools EnsemblPlants (www.plants.ensembl.org) and NCBI (www.ncbi.nlm.nih.gov). To cover most of the sections of functional domains of OsEPSPS, we used the NG PAM on both the sense and anti-sense strands. In total, 35 spacers were designed and divided into 11 pools based on the sgRNA-target region of *OsEPSPS* (Table S2).

Rice *Agrobacterium*-mediated transformation

For rice transformation of aforementioned vectors, the *Japonica* rice cultivar Zhonghua11 was used as donor material in this study. The binary vectors in each group were pooled in equimolar ratios and transformed into *Agrobacterium tumefaciens* EHA-105 by electroporation and used to transform about 240 rice calli. *Agrobacterium*-mediated transformation was conducted following the procedure as previously reported (Kuang et al., 2020). Afterwards, the calli were subjected to hygromycin (50 mg/L) and glyphosate (10 mg/L) selection in the induction and selection media for 4–6 weeks at 28°C, respectively. Finally, the well-grown calli were transferred to regeneration medium to recover plants.

Molecular characterization of the regenerated plants

The genomic DNA of the regenerated plant was extracted using the DNA quick Plant System (Tiangen Biotech, Beijing, China). Polymerase chain reaction amplification and detailed genotyping were performed as previously described (Li et al., 2022a). The Sanger sequencing chromatograms at target sites were analyzed using SnapGene software. To investigate off-target effects, we selected potential off-target sites, based on the prediction of the CRISPR-GE (skl.scau.edu.cn) for the target of sgRNA19 (Table S4). The site-specific genomic PCR and Sanger sequencing were used to determine the off-target effects. The primer sets used are listed in Table S5.

Glyphosate test in rice plants with D213N mutation

Rice seeds were germinated on 1/2 MS medium in a growth chamber (28°C, 16 h light/8 h dark). After 14 d, the plantlets with *OsEPSPS*-D213N allele and the wild-type plants were transferred to the soil and grown in the greenhouse. The rice plants at the three-leaf stage were then sprayed with glyphosate at a dose of 4× field-recommended dose (5,400 g a.i./ha) using a cabinet spray chamber at 0.3 MPa pressure; the phenotypes were then investigated after 14 d of glyphosate applications, with wild-type plants included as the control.

ACKNOWLEDGEMENTS

This work is partly funded by the National Natural Science Foundation of China (32188102 to LX), Hainan Yazhou Bay Seed Lab (B21HJ0215 to LX), the Central Public-Interest Scientific Institution-Based Research Fund (S2023ZD03 to LX), and Key Laboratory of Gene Editing Technologies (Hainan), China as well as National Engineering Research Center of Crop Molecular Breeding.

CONFLICT OF INTEREST

The authors declare they have no conflicts of interest associated with this work. A patent application has been filed on the novel OsEPSPS-D213N allele conferring rice glyphosate tolerance.

AUTHOR CONTRIBUTIONS

L.X. and C.Z. conceived the study. C.Z. and X.Z. contributed to the vector constructions. C.Z. and X.Z., S.L., Y.H., Y.L.,

and Y.Z. contributed to the rice transformations, molecular analysis, herbicide treatments and phenotyping. C.Z. wrote the manuscript. L.X. revised the manuscript. All the authors read and approved the final version of this manuscript.

Edited by: Jian-Kang Zhu, Southern University of Science and Technology, China

Received Mar. 27, 2023; **Accepted** Jun. 29, 2023; **Published** Jul. 4, 2023

OO: OnlineOpen

REFERENCES

Baerson, S.R., Rodriguez, D.J., Tran, M., Feng, Y., Biest, N.A., and Dill, G.M. (2002). Glyphosate-resistant goosegrass. Identification of a mutation in the target enzyme 5-enolpyruvylshikimate-3-phosphate synthase. Plant Physiol. **129:** 1265–1275.

Bharat, S.S., Li, S.Y., Li, J.Y., Yan, L., and Xia, L.Q. (2020). Base editing in plants: Current status and challenges. Crop J. **8:** 384–395.

Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. Nature **551:** 464–471.

Jiang, Y., Chai, Y., Qiao, D., Wang, J., Xin, C., Sun, W., Cao, Z., Zhang, Y., Zhou, Y., Wang, X.C., et al. (2022). Optimized prime editing efficiently generates glyphosate-resistant rice plants carrying homozygous TAP-IVS mutation in EPSPS. Mol. Plant **15:** 1646–1649.

Jin, M., Chen, L., Deng, X.W., and Tang, X. (2022). Development of herbicide resistance genes and their application in rice. Crop J. **10:** 26–35.

Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature **533:** 420–424.

Kuang, Y., Li, S., Ren, B., Yan, F., Spetz, C., Li, X., Zhou, X., and Zhou, H.B. (2020). Base-editing-mediated artificial evolution of OsALS1 in planta to develop novel herbicide-tolerant rice germplasms. Mol. Plant **13:** 565–572.

Kurt, I.C., Zhou, R., Iyer, S., Garcia, S.P., Miller, B.R., Langner, L.M., Grunewald, J., and Joung, J.K. (2021). CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. Nat. Biotechnol. **39:** 41–46.

Li, C., Zhang, R., Meng, X., Chen, S., Zong, Y., Lu, C., Qiu, J.L., Chen, Y.H., Li, J., and Gao, C.X. (2020). Targeted, random mutagenesis of plant genes with dual cytosine and adenine base editors. Nat. Biotechnol. **38:** 875–882.

Li, H.Y., Zhu, Z.W., Li, S.Y., Li, J.Y., Yan, L., Zhang, C., Ma, Y.Z., and Xia, L.Q. (2022a). Multiplex precision gene editing by a surrogate prime editor in rice. Mol. Plant **15:** 1077–1080.

Li, J., Chen, L., Liang, J., Xu, R., Jiang, Y., Li, Y., Ding, J., Li, M., Qin, R., and Wei, P.C. (2022b). Development of a highly efficient prime editor 2 system in plants. Genome Biol. **23:** 161.

Li, J.Y., Zhang, C., He, Y.B., Li, S.Y., Yan, L., Li, Y.C., Zhu, Z.W., and Xia, L.Q. (2023). Plant base editing and prime editing: The current status and future perspectives. J. Integr. Plant Biol. **65:** 444–467.

Liu, Q., Wang, C., Jiao, X., Zhang, H., Song, L., Li, Y., Gao, C., and Wang, K.J. (2019). Hi-TOM: A platform for high-throughput tracking of mutations induced by CRISPR/Cas systems. Sci. China: Life Sci. **62:** 1–7.

Liu, X., Qin, R., Li, J., Liao, S., Shan, T., Xu, R., Wu, D., and Wei, P.C. (2020). A CRISPR-Cas9-mediated domain-specific base-editing screen enables functional assessment of ACCase variants in rice. Plant Biotechnol. J. **18:** 1845–1847.

Pan, L., Yu, Q., Wang, J., Han, H., Mao, L., Nyporko, A., Maguza, A., Fan, L., Bai, L., and Powles, S. (2021). An ABCC-type transporter endowing glyphosate resistance in plants. Proc. Natl. Acad. Sci. U.S.A. **118:** e2100136118.

Perotti, V.E., Larran, A.S., Palmieri, V.E., Martinatto, A.K., Alvarez, C.E., Tuesca, D., and Permingeat, H.R. (2019). A novel triple amino acid substitution in the EPSPS found in a high-level glyphosate-resistant *Amaranthus hybridus* population from Argentina. Pest. Manag. Sci. **75:** 1242–1251.

Qiao, D., Wang, J., Lu, M.H., Xin, C., Chai, Y., Jiang, Y., Sun, W., Cao, Z., Guo, S., Wang, X.C., et al. (2022). Optimized prime editing efficiently generates heritable mutations in maize. J. Integr. Plant Biol. **65:** 900–906.

Ren, B., Liu, L., Li, S., Kuang, Y., Wang, J., Zhang, D., Zhou, X., Lin, H., and Zhou, H.B. (2019). Cas9-NG greatly expands the targeting scope of the genome-editing toolkit by recognizing NG and other atypical PAMs in rice. Mol. Plant **12:** 1015–1026.

Richter, M.F., Zhao, K.T., Eton, E., Lapinaite, A., Newby, G.A., Thuronyi, B.W., Wilson, C., Koblan, L.W., Zeng, J., Bauer, D.E., et al. (2020). Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. Nat. Biotechnol. **38:** 883–891.

Sony, S.K., Kaul, T., Motelb, K.F.A., Thangaraj, A., Bharti, J., Kaul, R., Verma, R., and Nehra, M. (2023). CRISPR/Cas9-mediated homology donor repair base editing confers glyphosate resistance to rice (*Oryza sativa* L.). Front. Plant Sci. **14:** 1122926.

Steinrücken, H.C., and Amrhein, N. (1980). The herbicide glyphosate is a potent inhibitor of 5-enolpyruvylshikimic acid-3-phosphate synthase. Biochem. Biophys. Res. Commun. **94:** 1207–1212.

Tan, J., Zeng, D., Zhao, Y., Wang, Y., Liu, T., Li, S., Xue, Y., Luo, Y., Xie, X., Chen, L., et al. (2022). PhieABEs: A PAM-less/free high-efficiency adenine base editor toolbox with wide target scope in plants. Plant Biotechnol. J. **20:** 934–943.

Wang, H., He, Y., Wang, Y., Li, Z., Hao, J., Song, Y., Wang, M., and Zhu, J.K. (2022). Base editing-mediated targeted evolution of ACCase for herbicide-resistant rice mutants. J. Integr. Plant Biol. **64:** 2029–2032.

Wang, X., Ding, C., Yu, W., Wang, Y., He, S., Yang, B., Xiong, Y.C., Wei, J., Li, J., Liang, J., et al. (2020). Cas12a base editors induce efficient and specific editing with low DNA damage response. Cell Rep. **31:** 107723.

Wei, C., Wang, C., Jia, M., Guo, H.X., Luo, P.Y., Wang, M.G., Zhu, J.K., and Zhang, H. (2021). Efficient generation of homozygous substitutions in rice in one generation utilizing an rABE8e base editor. J. Integr. Plant Biol. **63:** 1595–1599.

Xu, R., Liu, X., Li, J., Qin, R., and Wei, P.C. (2021). Identification of herbicide resistance OsACC1 mutations via in planta prime-editing-library screening in rice. Nat. Plants **7:** 888–892.

Xu, W., Yang, Y., Liu, Y., Kang, G., Wang, F., Li, L., Lv, X., Zhao, S., Yuan, S., Song, J., et al. (2020). Discriminated sgRNAs-based surrogate system greatly enhances the screening efficiency of plant base-edited cells. Mol. Plant **13:** 169–180.

Yan, D., Ren, B., Liu, L., Yan, F., Li, S., Wang, G., Sun, W., Zhou, X., and Zhou, H.B. (2021). High-efficiency and multiplex adenine base editing in plants using new TadA variants. Mol. Plant **14:** 722–731.

Zeng, D., Liu, T., Tan, J., Zhang, Y., Zheng, Z., Wang, B., Zhou, D., Xie, X., Guo, M., Liu, Y.G., et al. (2020). PhieCBEs: Plant high-efficiency cytidine base editors with expanded target range. Mol. Plant **13:** 1666–1669.

Zhao, D., Li, J., Li, S., Xin, X., Hu, M., Price, M.A., Rosser, S.J., Bi, C., and Zhang, X. (2021). Glycosylase base editors enable C-to-A and C-to-G base changes. Nat. Biotechnol. **39:** 35–40.

Zhong, Z., Sretenovic, S., Ren, Q., Yang, L., Bao, Y., Qi, C., Yuan, M., He, Y., Liu, S., Liu, X., et al. (2019). Improving plant genome editing with high-fidelity xCas9 and non-canonical PAM-targeting Cas9-NG. Mol. Plant **12:** 1027–1036.

Zong, Y., Song, Q., Li, C., Jin, S., Zhang, D., Wang, Y., Qiu, J.L., and Gao, C.X. (2018). Efficient C-to-T base editing in plants using a fusion of nCas9 and human APOBEC3A. Nat. Biotechnol. **36:** 950–954.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article: http://onlinelibrary.wiley.com/doi/10.1111/jipb.13543/suppinfo

Figure S1. Base editing efficiencies and product purity of different surrogate two-component composite base editing systems (STCBEs) in rice protoplasts

Figure S2. Base editing efficiencies of surrogate two-component composite base editing system 2 (STCBE-2) at the tested 16 targets of four endogenous genes, *OsAAT*, *OsCDC48*, *OsDEP1*, and *OsOD*, in rice protoplasts

Figure S3. The genotypes of *HptII* in the surviving edited plants in T_0 generation

Figure S4. Identification of transgene-free edited T_1 plants by polymerase chain reaction (PCR) amplification of the *OsEPSPS, sgRNA, nCas9-NG*, and *HptII* genes with gene-specific primers

Figure S5. The phenotypes of wild-type, heterozygous and homozygous *OsEPSPS-D213N* plants at seedling stage

Figure S6. The phenotypes of wild-type and representative heterozygous *OsEPSPS-D213N* plants at seed setting stage

Table S1. The sequences of guide RNAs (gRNAs) targeting different rice genes selected for testing the editing activities of surrogate two-component composite base editing systems (STCBEs) in protoplasts in this study

Table S2. The sequences of guide RNAs (gRNAs) used for saturated mutagenesis of the *OsEPSPS*

Table S3. The locations of conserved functional domains and glyphosate-binding domains in *Oryza sativa* 5-enolpyruvylshikimate-3-phosphate synthase (OsEPSPS) and the respective corresponding single-guide RNAs (sgRNAs)

Table S4. Potential off-target effects of guide RNAs (gRNA) used for D213N mutation in the rice genome

Table S5. Primers used in this study

Scan using WeChat with your smartphone to view JIPB online

Scan with iPhone or iPad to view JIPB online