nature plants

LETTERS

https://doi.org/10.1038/s41477-020-00827-4

Check for updates

PAM-less plant genome editing using a CRISPR–SpRY toolbox

Qiurong Ren[1,7], Simon Sretenovic[2,7], Shishi Liu[1,7], Xu Tang[1], Lan Huang[1], Yao He[1], Li Liu[1], Yachong Guo[1], Zhaohui Zhong[1], Guanqing Liu[3], Yanhao Cheng[2], Xuelian Zheng[1], Changtian Pan[2], Desuo Yin[2], Yingxiao Zhang[2], Wanfeng Li[4], Liwang Qi[4], Chenghao Li[5], Yiping Qi[2,6]✉ and Yong Zhang[1]✉

The rapid development of the CRISPR-Cas9, -Cas12a and -Cas12b genome editing systems has greatly fuelled basic and translational plant research[1–6]. DNA targeting by these Cas nucleases is restricted by their preferred protospacer adjacent motifs (PAMs). The PAM requirement for the most popular *Streptococcus pyogenes* Cas9 (SpCas9) is NGG (N = A, T, C, G)[7], limiting its targeting scope to GC-rich regions. Here, we demonstrate genome editing at relaxed PAM sites in rice (a monocot) and the Dahurian larch (a coniferous tree), using an engineered SpRY Cas9 variant[8]. Highly efficient targeted mutagenesis can be readily achieved by SpRY at relaxed PAM sites in the Dahurian larch protoplasts and in rice transgenic lines through non-homologous end joining (NHEJ). Furthermore, an SpRY-based cytosine base editor was developed and demonstrated by directed evolution of new herbicide resistant *OsALS* alleles in rice. Similarly, a highly active SpRY adenine base editor was developed based on ABE8e (ref. [9]) and SpRY-ABE8e was able to target relaxed PAM sites in rice plants, achieving up to 79% editing efficiency with high product purity. Thus, the SpRY toolbox breaks a PAM restriction barrier in plant genome engineering by enabling DNA editing in a PAM-less fashion. Evidence was also provided for secondary off-target effects by de novo generated single guide RNAs (sgRNAs) due to SpRY-mediated transfer DNA self-editing, which calls for more sophisticated programmes for designing highly specific sgRNAs when implementing the SpRY genome editing toolbox.

The targeting scope of SpCas9 (hereafter Cas9) can be broadened with Cas9 orthologs and engineered Cas9 variants. Cas9 orthologs with different PAM requirements, such as StCas9 and SaCas9, have been demonstrated in plants[10,11]. Multiple engineered Cas9 variants have been adopted for plant genome editing at altered PAM sites, including SpCas9-VQR for NGAN or NGNG PAMs[12,13], Cas9-NG for NG PAMs[14–17] and iSpyMacCas9 for NAAR PAMs[18]. Despite this progress, a Cas9 variant without any PAM restrictions had remained elusive until very recently. The recently engineered SpRY Cas9 variant confers nearly PAM-less genome editing in human cells[8]. Compared to wild-type Cas9, SpRY contains 11 amino acid changes (Supplementary Fig. 1). In human cells, SpRY can edit NR (R = A, G) PAM sites more efficiently than NY (Y = C, T) PAM sites[8]. SpRY is poised to further revolutionize genome editing in many organisms.

To assess SpRY in plants, we first targeted a total of 59 NNN PAM sites in rice, a monocot and major crop. Genome editing at these target sites was conducted in rice protoplasts and quantified by next-generation sequencing (NGS) of PCR amplicons. The data indicated SpRY editing, reflected as NHEJ mutations, at nearly all the target sites, albeit with variable efficiencies (Fig. 1a–d). By contrast, Cas9 only showed editing at the canonical NGG PAM sites (Fig. 1b) and to some extent at NAG PAM sites (Fig. 1a), which are known to be editable by Cas9 (ref. [19]). At NAN PAM sites, SpRY showed higher editing efficiency (roughly 20% for the median) at NAG and NAT PAM sites than at NAA and NAC PAM sites (roughly 5–10% for the median) (Fig. 1a and Supplementary Fig. 2). Notably, SpRY appeared to be more robust at editing NAG PAM sites than Cas9 (Fig. 1a). At NGN PAM sites, SpRY resulted in 5–20% median editing efficiency and displayed better editing at NGG PAM sites than at NGA, NGC and NGT PAM sites (Fig. 1b and Supplementary Fig. 3). At NYN PAM sites with NCN and NTN PAMs, SpRY generated roughly 5–10% median editing efficiency (Fig. 1c,d and Supplementary Figs. 4 and 5). Overall, Cas9 slightly outperformed SpRY (roughly 30% as opposed to 20% median editing efficiency) at the canonical NGG PAM sites (Fig. 1b), underperformed compared to SpRY at the non-canonical NAG PAM sites (Fig. 1a) and exhibited very low to no editing activity at relaxed non-canonical NNN PAM sites (Fig. 1a–d). By contrast, SpRY could edit all NNN PAM groups and showed relatively higher editing efficiency at NRN PAM sites than NYN PAM sites (Fig. 1a–d and Supplementary Figs. 2–5), which is consistent with the data from human cells[8]. There also appears to be no strong discrimination at the third nucleotide in the PAMs (Fig. 1a–d), supporting the near PAM-less nature of SpRY.

To examine the characteristics of SpRY editing, we compared NHEJ editing profiles of SpRY and Cas9 at NAG PAM sites and NGG PAM sites using NGS data from rice protoplast amplicons. The analysis of deletion positions showed a large overlap between SpRY and Cas9 (Fig. 1e). However, a closer look at deletion sizes revealed a higher frequency of larger deletions by SpRY (Fig. 1f). For example, SpRY resulted in a higher frequency of 5 bp deletions at both target sites when compared to Cas9 (Fig. 1f). Further analysis of SpRY editing profiles at additional 14 NAN, NGN, NCN and

[1]Department of Biotechnology, School of Life Science and Technology, Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu, China. [2]Department of Plant Science and Landscape Architecture, University of Maryland, College Park, MD, USA. [3]Jiangsu Key Laboratory of Crop Genetics and Physiology, Key Laboratory of Plant Functional Genomics of the Ministry of Education, Jiangsu Key Laboratory of Crop Genomics and Molecular Breeding, Agricultural College of Yangzhou University, Yangzhou, China. [4]State Key Laboratory of Tree Genetics and Breeding, Research Institute of Forestry, Chinese Academy of Forestry, Beijing, China. [5]State Key Laboratory of Forest Genetics and Tree Breeding, Northeast Forestry University, Harbin, China. [6]Institute for Bioscience and Biotechnology Research, University of Maryland, Rockville, MD, USA. [7]These authors contributed equally: Qiurong Ren, Simon Sretenovic, Shishi Liu. ✉e-mail: yiping@umd.edu; zhangyong916@uestc.edu.cn

NTN PAM sites consistently revealed a high frequency of 3–7 bp deletions (Supplementary Figs. 6–9), which are larger than the 1–3 bp deletions typical of Cas9 and similar to Cas12a with 5–10 bp deletions[2,20]. This feature indicates that SpRY is probably more suitable than Cas9 for knocking out microRNA genes[21] or engineering quantitative trait variation through promoter editing[22]. We also compared SpRY to Cas9-NG and xCas9 (ref. [23]) at NGN PAM sites in rice protoplasts. On average, SpRY demonstrated editing efficiency comparable to Cas9-NG, but yielded much higher editing efficiency than xCas9 at the NGN sites in rice protoplasts (Supplementary Fig. 10). Taken together, SpRY appears to be PAM-less for genome editing in rice. With SpRY's PAM-less feature, we would expect vector self-editing. In fact, the larger deletions by SpRY could be attributed to editing guided by newly generated sgRNAs after vector self-editing, which warranted further investigation in stable transgenic plants. Nevertheless, the ability to generate larger deletions by SpRY may aid gene knockout and *cis* regulatory element editing applications.

CRISPR–Cas genome editing systems have been widely applied in angiosperms, including monocots and dicots[1,4,5]. Genome editing has, however, not been demonstrated in gymnosperms, which contain many plant species of evolutionary and ecological importance. We hence sought to apply SpRY for genome editing in a gymnosperm species. To this end, we targeted three genomic sites in the Dahurian larch (*Larix gmelinii*), a coniferous tree (Fig. 1g). An average editing efficiency of 72.5% was achieved by SpRY at a canonical GGG PAM site in the Dahurian larch protoplasts (Fig. 1h). Robust editing efficiency was detected at two relaxed GAG and TGT PAM sites, with average editing efficiencies of roughly 17.0 and 12.0%, respectively (Fig. 1h and Supplementary Fig. 11). Genome editing by SpRY in the Dahurian larch protoplasts was further confirmed by Sanger sequencing (Fig. 1i). These data indicate the promising applications of SpRY for PAM-less genome editing in diverse plant species.

To see whether SpRY can readily generate edited plants, we transformed 12 SpRY transfer DNA (T-DNA) vectors and two Cas9 control T-DNA vectors in rice by *Agrobacterium*-mediated transformation. SpRY resulted in 62.5% editing efficiency at the OsPDS-AGG-02 site, which was slightly lower than Cas9's editing efficiency (79.0%) at this canonical NGG PAM site (Fig. 2a and Supplementary Fig. 12). At a NAG PAM site (OsPDS-CAG-01), SpRY displayed higher editing efficiency than Cas9 (23.5% as opposed to 13.6%) (Fig. 2a). Editing efficiencies in T_0 lines by SpRY at ten relaxed PAM sites (TAA, GAA, GAT, CAC, TGC, GGT, TTG, CTG, ACT and TCA) were 93.8, 13.3, 100, 63.2, 95.7, 40.9, 46.7, 5.3, 8.7 and 12.5%, respectively (Fig. 2a). The genotypes of targeted mutations these T_0 rice plants were revealed by Sanger sequencing, which showed high-frequency biallelic editing in many lines (Fig. 2a and Supplementary Figs. 13–16). Albino T_0 plants due to biallelic editing of *OsPDS* by SpRY at these PAM sites were readily recovered (Fig. 2b). These data indicate SpRY is robust for targeted mutagenesis at relaxed PAM sites in rice T_0 lines, consistent with the protoplast data (Fig. 1a). Altogether, we demonstrated robust targeted mutagenesis by SpRY in stable rice plants in a PAM-less fashion.

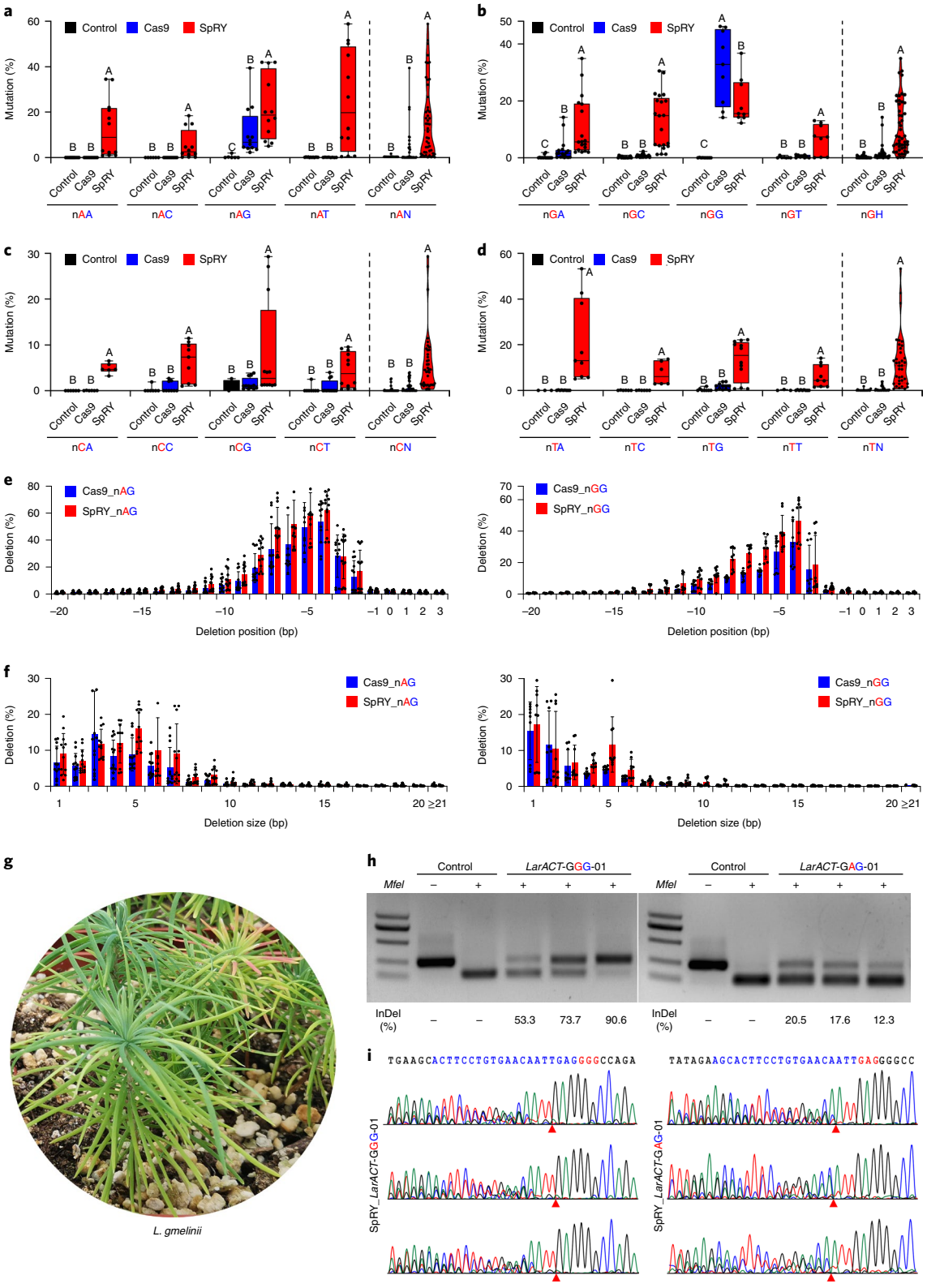
However, the PAM-less nature of SpRY makes the system vulnerable to vector self-editing. The canonical sgRNA scaffold contains the GTT trinucleotide immediately after the protospacer sequence. In rice protoplasts, SpRY resulted in roughly 5% median editing efficiency at NTT PAM sites (Fig. 1d), indicating that the GTT trinucleotide PAM is targetable by SpRY, which is further supported by SpRY induced larger deletions at the target sites (Fig. 1f). By genotyping the T_0 lines resulting from the 12 SpRY T-DNA constructs, we indeed found evidence of T-DNA self-editing for 11 constructs (Fig. 2a), including the one targeting an AGG PAM site. By contrast, Cas9 did not result in T-DNA self-editing at the two target sites (Fig. 2a). Both single-site and multi-site self-editing events were found in T_0 lines, and in some cases the T_0 lines contained both on-target editing and T-DNA self-editing (Fig. 2c–e and Supplementary Fig. 17). On-target editing efficiencies at NRN PAM sites (for example, AGG and GGT PAMs) are generally higher than T-DNA self-editing efficiencies at the vector's own GTT PAM site, while T-DNA self-editing efficiencies at this GTT PAM site seem to be higher than on-target editing efficiencies at NYN PAM sites (Fig. 2a). Hence, these data not only reveal SpRY's high tendency for self-editing, but also further confirm SpRY's general preference for editing NRN PAMs over NYN PAMs. At all four NYN PAM sites tested (TTG, CTG, ACT and TCA), much higher frequencies of T-DNA self-editing were observed (80.0, 36.8, 52.2 and 79.2%) than target site editing (46.7, 5.3, 8.7 and 12.5%). These data supported the idea that the GTT PAM on the T-DNA was favoured by SpRY over many NYN PAMs, which may partly contribute to lower on-target editing at NYN PAM sites. Since we have successfully obtained targeted mutants for all NNN SpRY vectors in the T_0 generation. we conclude that SpRY is a potent nuclease for editing PAM-less sites, despite self-editing.

The PAM-less nature of SpRY would potentially increase the chance of off-targeting on two levels. The first level of off-targeting is solely based on sequence similarity to the target sites. To assess this type of off-targeting, we selected five constructs for editing GAA, TTG, GGT, AGG and ACT PAMs. Between 15 and 23 rice T_0 transgenic lines for each construct were genotyped by Sanger sequencing at all top off-target sites with 1–3 mismatches, identified by Cas-OFFinder[24]. We have shown previously that Sanger sequencing is sensitive enough to capture off-target mutations identified by whole genome sequencing[6]. No off-target mutations were detected at these sites (Supplementary Table 1). The second level of off-targeting may result from the de novo generated sgRNAs due to T-DNA self-editing. To assess this, we chose eight edited T_0 plants that carried small self-editing deletions resulting from four different SpRY constructs (nos. 3136, 3139, 3144 and

Fig. 1 | PAM-less gene editing by SpRY in the protoplasts of rice and the Dahurian larch. a–**d**, Comparison of Cas9 and SpRY at editing NAN (16 sites) (**a**), NGN (19 sites) (**b**), NCN (12 sites) (**c**) and NTN (12 sites) (**d**) PAM sites in rice cells. Editing efficiency was quantified by NGS of PCR amplicons. Left panel, editing at PAM subgroups. Right panel, editing at the whole PAM group. Each dot represents a biological replicate. Each target contains three biological replicates. Different capital letters indicate significant differences ($P < 0.05$; one-way analysis of variance, Bonferroni post hoc test). Samples with the same uppercase letters have no significant difference. The median, interquartile range (IQR) and 1.5× IQR are shown. The maxima, centre and minima of each box refer to the upper quartile (Q3), median (Q2) and lower quartile (Q1). The maxima and minima of whiskers refer to Q3 + 1.5× IQR, Q1 − 1.5× IQR. **e**, Comparison of deletion positions between Cas9 and SpRY at three NGG and four NAG PAM sites in rice cells. Each dot represents a biological replicate. Each target contains three biological replicates. Data are presented as mean values ± s.d. ($n = 12$ for NAG PAM sites, $n = 9$ for NGG PAM sites). **f**, Comparison of deletion sizes between Cas9 and SpRY at three NGG and four NAG PAM sites in rice cells. Each dot represents a biological replicate. Each target contains three biological replicates. Data are presented as mean values ± s.d. ($n = 12$ for NAG PAM sites, $n = 9$ for NGG PAM sites). **g**, A photograph of the Dahurian larch (*L. gmelinii*) seedlings sourced for the protoplast assay. **h**, RFLP analysis of editing efficiency of SpRY at *LarACT*-GGG-01 and *LarACT*-GAG-01 sites in the *L. gmelinii* protoplasts. The upper bands uncut by *MfeI* denote edited mutations. Editing results of three biological replicates are shown. The sizes of the DNA marker from the top to the bottom are 1,000, 750, 500, 250 and 100 bp. **i**, Detection of targeted mutations by SpRY in *L. gmelinii* protoplasts by Sanger sequencing. The red arrowheads indicate the cleavage positions of SpRY on the target sequences.

3146) (Supplementary Table 2). These lines carried newly generated sgRNAs with altered protospacers (for example, 17 and 21 nt) due to T-DNA self-editing. We assessed 42 top off-target sites preferred by these new sgRNAs and found one T_0 line (no. 3139-3-1) contained an off-target mutation caused by a de novo generated sgRNA that has 17 nucleotides perfect match to this off site (Supplementary Table 2 and Supplementary Fig. 14b). No mutations were found at the remaining 41 putative off-target sites in the seven T_0 lines, even though some sites only have one nucleotide mismatch to the protospacers (Supplementary Table 2). Thus, despite self-editing, we did not find strong and concerning evidence for off-target effects of SpRY.

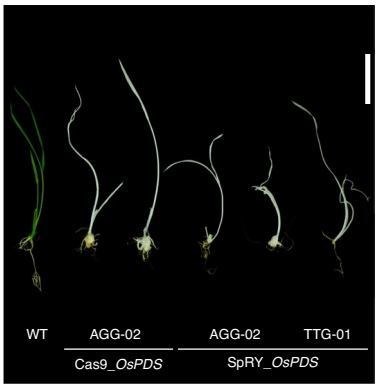
The PAM-less feature of SpRY renders many previously inaccessible bases amenable to base editing. We fused PmCDA1-UGI to the C terminus of SpRY-D10A nickase, generating a PAM-less cytosine base editor termed SpRY-PmCDA1 (Supplementary Fig. 1). Testing



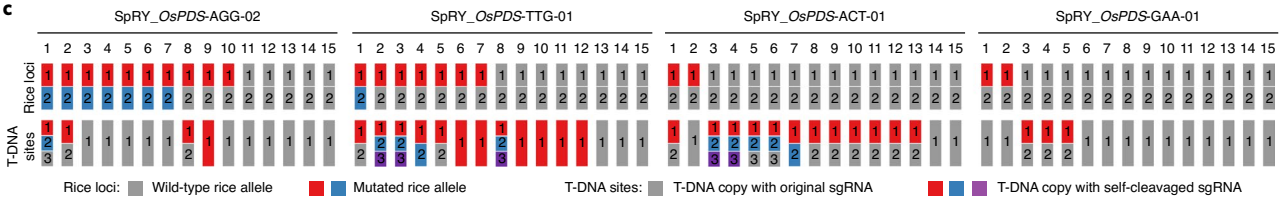
a

Target gene	Reagents	sgRNA + PAM	Tested T_0 lines	Mutated T_0 lines (number; ratio)	Biallelic T_0 lines (number; ratio)	Self-cleavage T_0 lines (number; ratio)
OsPDS	Cas9_*OsPDS*-AGG-02	GTTGGTCTTTGCTCCTGCAGAGG	19	15; 79.0%	15; 79.0%	0; 0.0%
OsPDS	SpRY_*OsPDS*-AGG-02	GTTGGTCTTTGCTCCTGCAGAGG	16	10; 62.5%	7; 43.8%	4; 25.0%
OsPDS	Cas9_*OsPDS*-CAG-01	TTATGTTCATAGATGACAGGCAG	22	3; 13.6%	1; 4.6%	0; 0.0%
OsPDS	SpRY_*OsPDS*-CAG-01	TTATGTTCATAGATGACAGGCAG	17	4; 23.5%	2; 11.8%	5; 29.4%
OsPDS	SpRY_*OsPDS*-TAA-03	ACCTCCACTAGAAAACACAATAA	16	15; 93.8%	13; 81.3%	12; 75.0%
OsPDS	SpRY_*OsPDS*-GAA-01	GGCATTTCTACCTTATCGATGAA	15	2; 13.3%	0; 0.0%	3; 20.0%
OsDEP1	SpRY_*OsDEP1*-GAT-02	GACCTGTGCGGCCGGCGGCGGAT	20	20; 100.0%	20; 100.0%	9; 45.0%
OsPDS	SpRY_*OsPDS*-CAC-01	CGCAAGTAGCAGCATCCAAGCAC	19	12; 63.2%	8; 42.1%	9; 47.4%
OsPDS	SpRY_*OsPDS*-TGC-02	ACATGCTGACTACTTTTCAGTGC	23	22; 95.7%	13; 56.5%	19; 82.6%
OsDEP1	SpRY_*OsDEP1*-GGT-01	CCGCCGGCCGCACAGGTCCGGGT	22	9; 40.9%	8; 36.4%	0; 0.0%
OsPDS	SpRY_*OsPDS*-TTG-01	GCAATAATAGAGTACAAAGGTTG	15	7; 46.7%	1; 6.7%	12; 80.0%
OsPDS	SpRY_*OsPDS*-CTG-02	GGAAAGTCCTGGCAAACAACCTG	19	1; 5.3%	1; 5.3%	7; 36.8%
OsPDS	SpRY_*OsPDS*-ACT-01	TCTGCAGGTTGTTTGCCAGGACT	23	2; 8.7%	0; 0.0%	12; 52.2%
OsPDS	SpRY_*OsPDS*-TCA-01	ACAACCTGCAGAAAAGCCAATCA	24	3; 12.5%	0; 0.0%	19; 79.2%

b



c



d

Genome editing: SpRY_*OsPDS*-AGG-02

OsPDS (WT): CTGGAGTTGGTCTTTGCTCCTG-CAGAGGAATGG

no.3,157–1–1
Allele 1: CTGGAGTTGGTCTTTGCTCCTGtCAGAGGAATGG +1 bp
Allele 2: CTGGAGTTGGTCTTTGCTCCTG-CAGAGGAATGG WT
no.3,157–1–2
Allele 1: CTGGAGTTGGTCTTTGCTCCTGtCAGAGGAATGG +1 bp
Allele 2: CTGGAGTTGGTCTTTGCTCCTGtCAGAGGAATGG +1 bp
no.3,157–2–1
Allele 1: CTGGAGTTGGTCTTTGCTCCTGtCAGAGGAATGG +1 bp
Allele 2: CTGGAGTTGGTCTTTGCTCCTGtCAGAGGAATGG +1 bp
no.3,157–2–2
Allele 1: CTGGAGTTGGTCTTTGCTCCTGtCAGAGGAATGG +1 bp
Allele 2: CTGGAGTTGGTCTTTGCTCCTG-CAGAGGAATGG WT
no.3,157–4–2
Allele 1: CTGGAGTTGGTCTTTGCTCCTGgCAGAGGAATGG +1 bp
Allele 2: CTGGAGTTGGTCTTTGCTCCTGtCAGAGGAATGG +1 bp
no.3,157–5–1
Allele 1: CTGGAGTTGGTC----------AGAGGAATGG –11 bp
Allele 2: CTGGAGTTGGTCTTTGCTCCTGgCAGAGGAATGG +1 bp
no.3,157–5–2
Allele 1: CTGGAGTTGGTCTTTGCTCCTGcCAGAGGAATGG +1 bp
Allele 2: CTGGAGTTGGTCTTTGCTCCTG-CAGAGGAATGG WT
no.3,157–7–1
Allele 1: CTGGAGTTGGTCTTTGCTCCTGaCAGAGGAATGG +1 bp
Allele 2: CTGGAGTTGGTCTTTGCTCCTGtCAGAGGAATGG +1 bp
no.3,157–8–1
Allele 1: CTGGAGTTGGTCTTTGCTCCTG-tCAGAGGAATG +1 bp
Allele 2: CTGGAGTTGGTCTTTGCTCCTGaaCAGAGGAATG +2 bp
no.3,157–9–1
Allele 1: CTGGAGTTGGTCTTTGCTCCTGcCAGAGGAATGG +1 bp
Allele 2: CTGGAGTTGGTCTTTGCTCCTGtCAGAGGAATGG +1 bp

Self cleavage: SpRY_*OsPDS*-AGG-02

T-DNA_ref: TGTGTGTTGGTCTTTGCTCCTG-CAGGTTTTAGA
no.3,157–1–1
Copy 1: TGTGTGTTGGTCTTTGCTCCTG--AGGTTTTAGA –1 bp
no.3,157–1–2
Copy 1: TGTGTGTTGGTCTTTGCTCCTG------TTTAGA –5 bp
Copy 2: TGTGTGTTGGTCTTTGCTCCTG-------TTAGA –6 bp
Copy 3: TGTGTGTTGGTCTTTGCTCCTG-CAGGTTTTAGA WT
no.3,157–5–2
Copy 1: TGTGTGTTGGTCTTTGC------CAGGTTTTAGA –5 bp
Copy 2: TGTGTGTTGGTCTTTGCTCCTG-CAGGTTTTAGA WT
no.3,157–8–1
Copy 1: TGTGTGTTGGTCTTTGCTCCTGcCAGGTTTTAGA +1 bp
Copy 2: TGTGTGTTGGTCTTTGCTCCTG-CAGGTTTTAGA WT

e

Genome editing: SpRY_*OsPDS*-ACT-01

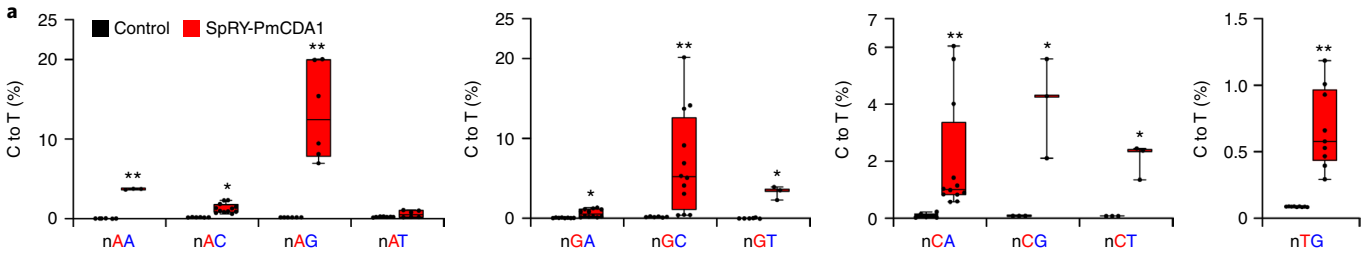
OsPDS (WT): GCTTTTCTGCAGGTTGTTTGCCAGGACTTTCCA

no.3,142–11–2
Allele 1: GCTTTTCTGCAGGTTGTT-----GGACTTTCCA –5 bp
Allele 2: GCTTTTCTGCAGGTTGTTTGCCAGGACTTTCCA WT
no.3,142–14–2
Allele 1: GCTTTTCTGCAGGTTGTT---CAGGACTTTCCA –3 bp
Allele 2: GCTTTTCTGCAGGTTGTTTGCCAGGACTTTCCA WT

Self cleavage: SpRY_*OsPDS*-ACT-01

T-DNA (ref): TGTGTTCTGCAGGTTGTTTGCCAGGGTTTTAGA
no.3,142–2–1
Copy 1: TGTGTTCTGCAGGTTGTT----AGGGTTTTAGA –4 bp
Copy 2: TGTGTTCTGCAGGTTGTTTGCCAGGGTTTTAGA WT
no.3,142–5–1
Copy 1: TGTGTTCTGCAGGTTG---GCCAGGGTTTTAGA –3 bp
Copy 2: TGTGTTCTGCAGGTT------CAGGGTTTTAGA –6 bp
Copy 3: TGTGTTCTGCAGGTT----------GTTTTAGA –10 bp
no.3,142–6–1
Copy 1: TGTGTTCTGCAGGTTGTTTG-CAGGGTTTTAGA –1 bp
Copy 2: TGTGTTCTGCAGGTT---------GGTTTTAGA –9 bp
no.3,142–6–2
Copy 1: TGTGTTCTGCAGGT--------A-GGTTTTAGA –9 bp
Copy 2: TGTGTTCTGCAGGTTGTTTGCCAGGGTTTTAGA WT
no.3,142–7–1
Copy 1: TGTGTTCTGCAGGTTGTT----AGGGTTTTAGA –4 bp
Copy 2: TGTGTTCTGCAGGTTGTTTGCCAGGGTTTTAGA WT
no.3,142–7–2
Copy 1: TGTGTTCTGCAGGTTGTT----AGGGTTTTAGA –4 bp
Copy 2: TGTGTTCTGCAGGTTGTTTGCCAGGGTTTTAGA WT
no.3,142–8–1
Copy 1: TGTGTTCTGCAGGTTGTTT--CAGGGTTTTAGA –2 bp
Copy 2: TGTGTTCTGCAGGTT--------GGGTTTTAGA –8 bp
Copy 3: TGTGTTCTGCAGGTTGTTTGCCAGGGTTTTAGA WT
no.3,142–8–2
Copy 1: TGTGTTCTGCAGGTTGTTT--CAGGGTTTTAGA –2 bp
Copy 2: TGTGTTCTGCAGGTT--------GGGTTTTAGA –8 bp
Copy 3: TGTGTTCTGCAGGTTGTTTGCCAGGGTTTTAGA WT
no.3,142–10–2
Copy 1: TGTGTTCTGCAGGTTGTT----AGGGTTTTAGA –4 bp
Copy 2: TGTGTTCTGCAGGTTGTTTGCCAGGGTTTTAGA WT
no.3,142–11–1
Copy 1: TGTGTTCTGCAGGTTG------AGGGTTTTAGA –6 bp
Copy 2: TGTGTTCTGCAGGTTGTTTGCCAGGGTTTTAGA WT
no.3,142–11–2
Copy 1: TGTGTTCTGCAGGTTGTTTG-CAGGGTTTTAGA –1 bp
Copy 2: TGTGTTCTGCAGGTTGTTTGCCAGGGTTTTAGA WT

Fig. 2 | Comparison of genome editing and self-cleavage by SpRY in stable rice lines. a, Summary of genome editing and vector self-cleavage editing for 14 T-DNA constructs in rice T_0 lines. The data presented here include editing of two PAM sites (*OsPDS*-AGG-02 and *OsPDS*-CAG-01) by both Cas9 and SpRY as well as editing at ten relaxed PAM sites by SpRY. **b**, Examples of albino phenotype due to biallelic editing of *OsPDS*. Scale bar, 2 cm. **c**, Schematic representation of genotyping results of both genome editing and self-cleavage events for four SpRY constructs in T_0 rice lines. **d**,**e**, Genotypes of rice T_0 lines with genome editing and/or self-cleavage editing at the *OsPDS*-AGG-02 (**d**) and *OsPDS*-ACT-01 (**e**) sites by SpRY. The protospacer is highlighted in blue. The PAM is highlighted in red. Note the T_0 lines highlighted in red contain simultaneous genome editing and self-editing.



b

Target gene	Reagents	sgRNA + PAM	Tested T_0 lines	Mutated T_0 lines (number; ratio)	C-to-T T_0 lines (number; ratio)	Other SNP T_0 lines (number; ratio)	Self-cleavage T_0 lines (number; ratio)
OsDEP1	SpRY-PmCDA1_*OsDEP1*-GAT-02	GACCTGTGCGGCCGGCGGCGGAT	20	2; 10.0%	2; 10.0%	0; 0.0%	0; 0.0%
OsDEP1	SpRY-PmCDA1_*OsDEP1*-GGT-01	GCCGCCGGCCGCACAGGTCCGGGT	19	8; 42.1%	5; 26.3%	3; 15.8%	2; 10.5%
OsALS	SpRY-PmCDA1_*OsALS*-sgRNA22	CCCCACTTGGGATCATAGGCAGC	18	13; 72.2%	6; 33.3%	7; 38.9%	2; 11.1%

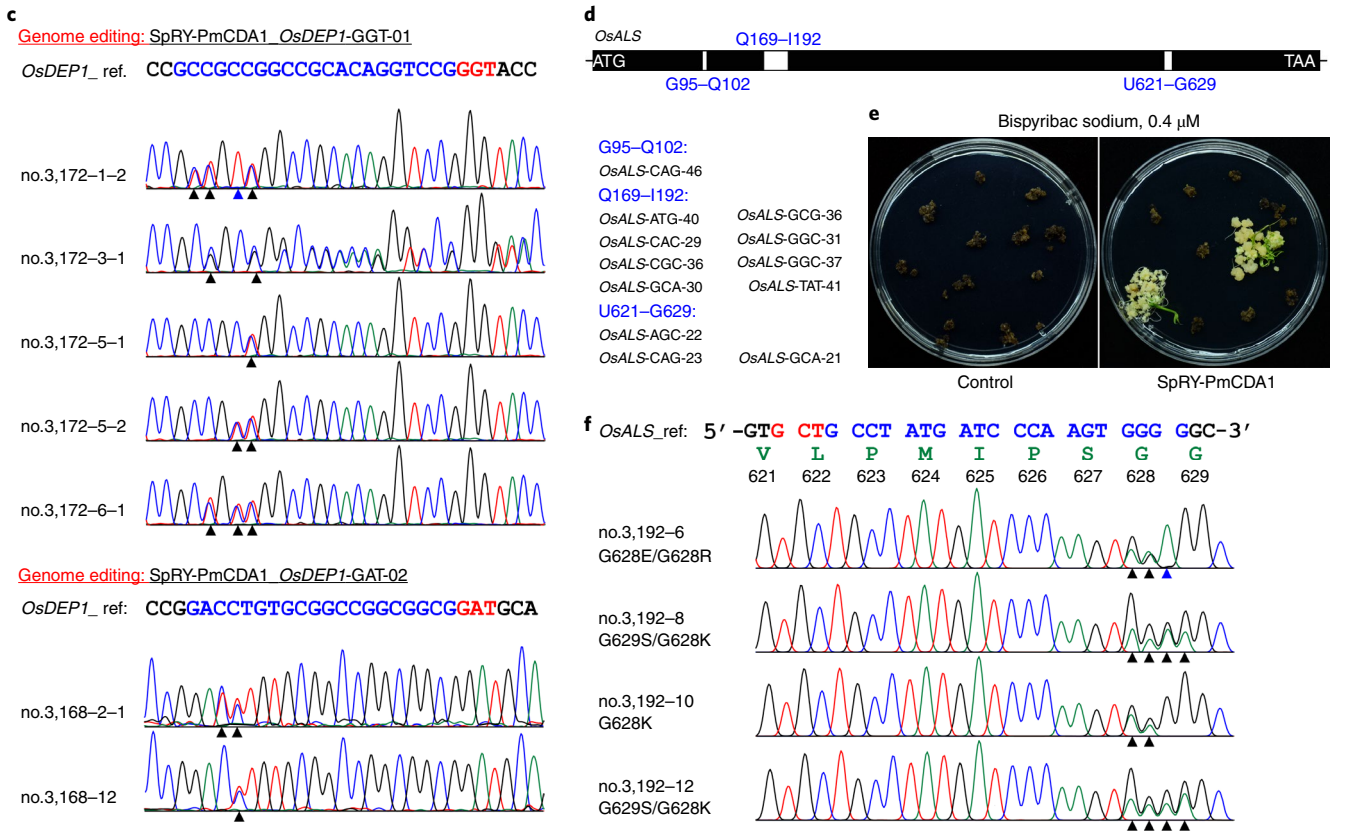
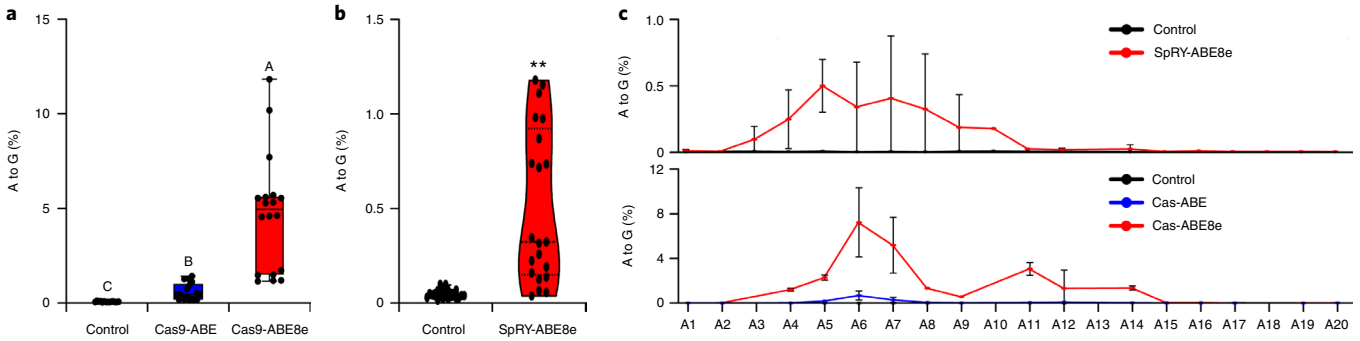


Fig. 3 | PAM-less C-to-T base editing in rice. a, Summary of C-to-T base editing by SpRY-PmCDA1 at 26 relaxed PAM sites in rice protoplasts. Each dot represents a biological replicate and each sgRNA contains three biological replicates. Independent-sample two-tailed *t*-tests, $^{*}P < 0.05$ and $^{**}P < 0.01$. Samples with the same uppercase letters have no significant difference. The median, IQR and 1.5× IQR are shown. The maxima, centre and minima each box refer to the upper quartile (Q3), median (Q2) and lower quartile (Q1). The maxima and minima of whiskers refer to Q3 + 1.5× IQR, Q1 − 1.5× IQR. **b**, Summary of SpRY-PmCDA1 based C-to-T editing at three relaxed PAM sites and at self-editing sites in rice T_0 lines. **c**, Examples of C-to-T base edited T_0 lines by SpRY-PmCDA1 at the *OsDEP1*-GGT-01 and *OsDEP1*-GAT-02 sites. Monoallelic editing bases are indicated by black arrows and biallelic editing bases are indicated by blue arrows. The protospacer is highlighted in blue and the PAM is highlighted in red. **d**, Schematic representation of three target regions in *OsALS* by SpRY-PmCDA1 with 12 sgRNAs for targeted evolution of herbicide resistance. **e**, Examples of herbicide resistant calli and seedlings on SpRY-PmCDA1 base editing of *OsALS*. The non-SpRY-PmCDA1-treated control calli were killed by the herbicide (0.4 μM bispyribac sodium). **f**, Genotypes of herbicide resistant calli and seedlings. Monoallelic editing bases are indicated by black arrows and biallelic editing bases are indicated by blue arrows.

of SpRY-PmCDA1 in rice protoplasts at 26 target sites demonstrated PAM-less base editing as it resulted in C-to-T base editing at many sites with NAN, NGN, NCN and NTN PAMs, although the editing efficiencies varied (Fig. 3a and Supplementary Fig. 18). Consistent with NHEJ mutagenesis data, SpRY-PmCDA1 showed higher editing efficiency at NRN PAMs than at NYN PAMs (Fig. 3a). Analysis of protoplast editing data at the 26 sites revealed an editing window spanning from the first nucleotide to the sixth nucleotide in the protospacer from the 5′ end (Supplementary Fig. 19), consistent with previous PmCDA1 studies[25,26]. Further analysis of three SpRY-PmCDA1 constructs in stable transgenic rice plants demonstrated efficient base editing at all three relaxed PAM sites (10.0, 42.1



d

Target gene	Reagents	sgRNA + PAM	Tested T_0 lines	Mutated T_0 lines (number; ratio)	C-to-T T_0 lines (number; ratio)	Other SNP T_0 lines (number; ratio)	Self-cleavage T_0 lines (number; ratio)
OsPDS	SpRY-ABE8e_*OsPDS*-TAA-04	AGAAACAGTGAACAACCCACTAA	19	15; 79.0%	15; 79.0%	0; 0.0%	8; 42.1%
OsPDS	SpRY-ABE8e_*OsPDS*-TTG-01	GCAATAATAGAGTACAAAGGTTG	20	9; 45.0%	9; 45.0%	0; 0.0%	14; 70.0%
OsPDS	SpRY-ABE8e_*OsPDS*-CCA-02	GCAAACAACCTGCAGAAAAGCCA	18	0; 0.0%	0; 0.0%	0; 0.0%	5; 27.8%

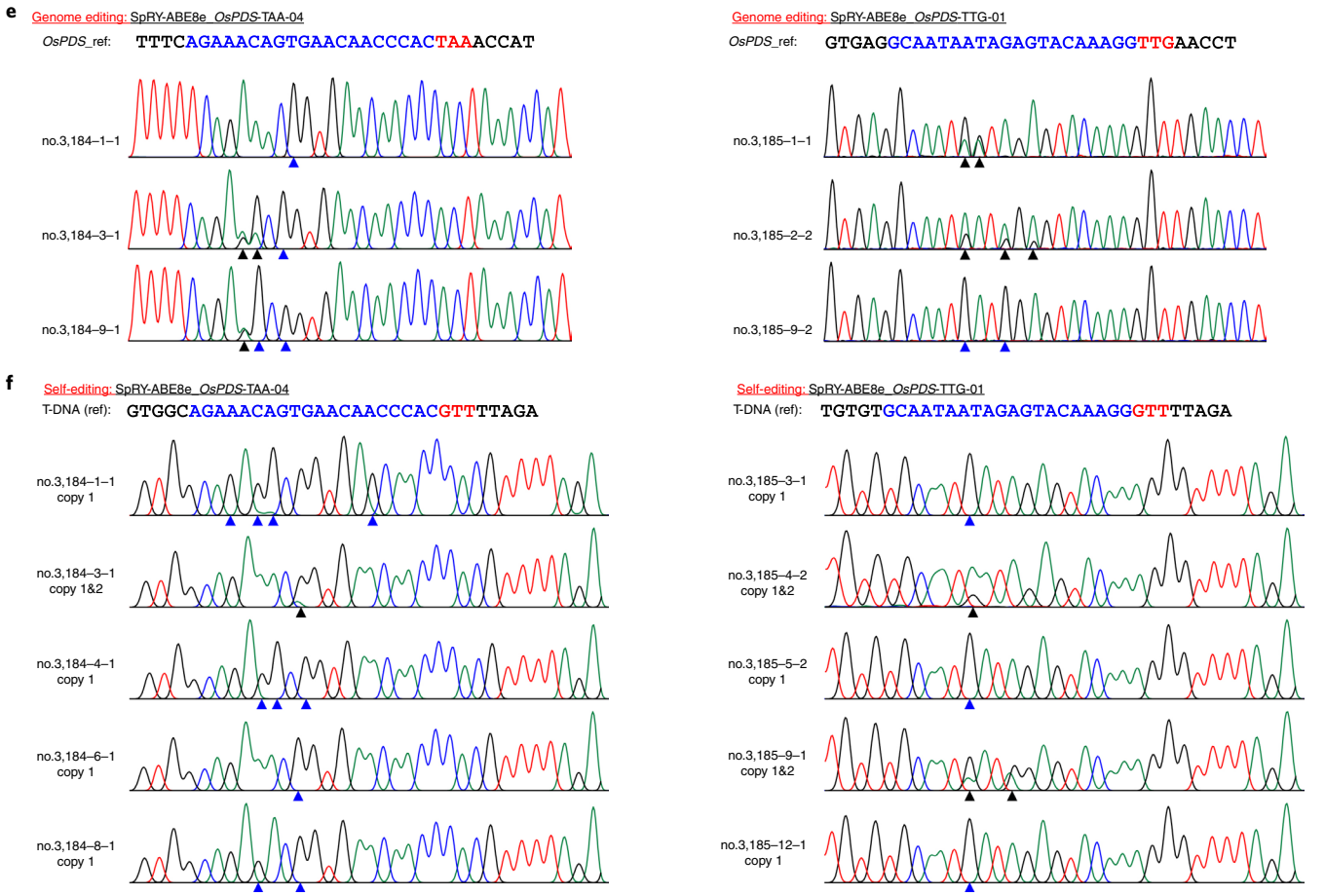


Fig. 4 | PAM-less A-to-G base editing in rice. a, Comparison of Cas9-ABE and Cas9-ABE8e for A-to-G base editing at NGG PAM sites in rice protoplasts. The NGS data include six independent target sites with each containing three biological replicates. Independent-sample two-tailed *t*-tests, $^{*}P < 0.05$ and $^{**}P < 0.01$. Samples with the same uppercase letters have no significant difference. The median, IQR and 1.5× IQR are shown. The maxima, centre and minima of each box refer to the upper quartile (Q3), median (Q2) and the lower quartile (Q1). The maxima and minima of whiskers refer to Q3 + 1.5× IQR, Q1 − 1.5× IQR. **b**, SpRY-ABE8e-mediated A-to-G base editing in rice protoplasts. The NGS data include seven independent relaxed PAM sites, with three biological replicates for each PAM site. Independent-sample two-tailed *t*-tests, $^{*}P < 0.05$ and $^{**}P < 0.01$. **c**, Analysis of A-to-G base editing windows for SpRY-ABE8e, Cas9-ABE8e and Cas9-ABE using data derived from **a** and **b**. Data are presented as mean values ± s.d. $n = 21$ for SpRY-ABE8e system. $n = 18$ for Cas-ABE/ABE8e system. **d**, Summary of SpRY-ABE8e based A-to-G editing at three relaxed PAM sites and at self-editing sites in rice T_0 lines. **e**,**f**, Example of A-to-G base edited T_0 lines at the *OsPDS*-TAA-04 and *OsPDS*-TTG-01 sites as well as at the vector self-editing sites: genome editing (**e**) and self-editing (**f**). Monoallelic editing bases are indicated by black arrows and biallelic editing bases are indicated by blue arrows.

and 72.2%, respectively) (Fig. 3b and Supplementary Fig. 20). At the OsDEP1-GGT-01 site, three T_0 lines (15.8%) contained byproduct deletions (Fig. 3b) and two T_0 lines (10.5%) displayed vector self-editing (Supplementary Fig. 20). At the OsALS-sgRNA22 site, seven T_0 lines (38.9%) had byproduct editing and two lines showed T-DNA self-editing (Fig. 3b). At all three target sites, on-target base editing was much favoured over T-DNA self-editing (Fig. 3b). On the basis of the data in rice protoplasts and stable plants, we concluded that SpRY confers precise C-to-T base editing without PAM restrictions.

We reasoned PAM-less base editing would greatly facilitate direct protein evolution in vivo. To demonstrate this capability, we pooled a small library of 12 sgRNAs for targeting three select regions in *OsALS* with SpRY-PmCDA1 for evolving herbicide resistance in rice (Fig. 3d). *Agrobacteria* carrying SpRY-PmCDA1 and 12 sgRNAs were used to transform rice calli followed by herbicide selection on 0.4 μM bispyribac sodium MS medium. Multiple resistant calli or seedlings emerged on the sgRNA-positive plates, while no surviving calli were found on the sgRNA-negative plates (Fig. 3e). Sequencing of independent surviving calli revealed the molecular basis of new herbicide resistant *OsALS* alleles, where C-to-T base changes on the non-coding strand resulted in missense mutations G628E/R/K and G629S (Fig. 3f). All these events were induced by an sgRNA with a relaxed AGC PAM (Fig. 3f), and the C-to-T conversion happened in the PmCDA1 editing window (first to the sixth nucleotide of the protospacer)[18,25–27]. This experiment demonstrates the powerful application of SpRY-PmCDA1 PAM-less base editor for directed protein evolution in plants[28–31].

Last but not least, we wanted to develop an efficient SpRY adenine base editing (ABE) system for PAM-less A-to-G base editing. Recently, an improved ABE8e was reported to have very high A-to-G editing efficiency in human cells[9], which catalyses deamination more than 1,000 times faster than early ABEs[32]. We first generated ABE8e and ABEmax (hereafter ABE)[33] based on Cas9 and compared both ABEs at the canonical NGG PAM sites in rice protoplasts (Supplementary Fig. 21). Indeed, we observed much higher A-to-G base editing with ABE8e than with ABE (Fig. 4a,b and Supplementary Fig. 21). Next, we made the SpRY version of ABE8e (SpRY-ABE8e) (Supplementary Fig. 1), which showed detectable A to G editing across many relaxed PAM sites in rice protoplasts (Fig. 4a,b and Supplementary Fig. 22a). The high activity editing window for SpRY-ABE8e appeared to span from the fourth to the eighth nucleotide in the protospacer from the 5′ end (Fig. 4c and Supplementary Fig. 22b), consistent with ABE8e editing data in human cells[9]. It is of note that the efficiency of A-to-G base editing is generally much lower than C-to-T base editing in protoplasts[18,30], which could be due to low activity of A-to-G base editing pathway in non-replicating cells. To further assess SpRY-ABE8e, we generated stable transgenic rice T_0 lines for two constructs and found 79.0% A-to-G editing at the OsPDS-TAA04 site and 45.0% A-to-G editing at the OsPDS-TTG-01 site (Fig. 4d and Supplementary Fig. 23). Notably, no byproducts were found among edited lines (Fig. 4d) and homozygous edited lines were easily identified at different relaxed PAM sites (Fig. 4e and Supplementary Fig. 23). High-frequency T-DNA self-editing was also observed for these two constructs (42.1 and 70.0%) (Fig. 4f). In addition, testing at a third target site (OsPDS-CCA-02) by SpRY-ABE8e revealed only T-DNA self-editing (27.8%) (Fig. 4d). Such prevalent self-editing could potentially contribute to secondary off-targeting due to de novo generated sgRNAs with 20-nt protospacers. To investigate this secondary off-targeting effect, we genotyped 39 top off-targeting sites in 13 T_0 lines derived from three independent SpRY-ABE8e constructs. Two secondary off-target events were detected in two independent T_0 lines (nos. 3185-1-1 and 3185-14-1) and these events were caused by the same de novo generated sgRNA that contained only one nucleotide mismatch at the core 2–20 nt protospacer sequence to the off-target site (Supplementary Table 2). Both events carried a single A-to-G mutation at the off-target sites (Supplementary Fig. 24). No secondary off-target mutations were detected at the remaining 37 putative off-target sites. Thus, SpRY-ABE8e offers robust A-to-G base editing at relaxed PAM sites in stable transgenic rice plants, albeit with frequent self-editing.

This study demonstrates a comprehensive SpRY toolbox for targeted mutagenesis and base editing in a nearly PAM-less manner in plants. Despite vector self-editing, the SpRY editing tools displayed high editing efficiency and specificity. SpRY vector self-editing seems to alter NHEJ editing profiles towards slightly larger deletions, which may aid certain genome editing applications. We applied SpRY for efficient genome editing in gymnosperms. Furthermore, we successfully applied the SpRY-PmCDA1 C-to-T base editor for directed evolution of herbicide resistance in rice. Finally, the SpRY-ABE8e base editor was very efficient in generating A-to-G base editing with high product purity in stable rice plants. The potential secondary off-target effects of SpRY were assessed in targeted mutagenesis and A-to-G base editing. Given the tendency of SpRY for generating larger deletions, many de novo generated sgRNAs may fail to function due to truncated protospacers. By contrast, high-purity SpRY base editing systems will generate fully functional new sgRNAs with roughly 20-nt protospacers that can contribute to secondary off-target mutations in the genome. Since both targeted mutagenesis and base editing outcomes can be readily predicted[34,35], the secondary off-targeting effects should be predicted and considered when designing and implementing the SpRY-based genome editing experiments.

On the basis of the data in this study, we make a practical recommendation that Cas9 should be used for editing the canonical NGG PAM sites and SpRY should be used for editing all other non-canonical PAM sites due to its more robust performance over Cas9-NG and xCas9 at such sites (Supplementary Fig. 10). Hence, augmented by its PAM-less feature, the SpRY genome editing toolbox developed here will have many promising applications in plant biology.

Methods

Construction of Gateway-compatible SpRY vectors. Details about construction of SpRY Gateway entry vectors are described in the Supplementary Methods. All target sites were listed in Supplementary Table 3. The oligos and gBlocks in this study were summarized in Supplementary Table 4.

Assembly of T-DNA expression vectors. The T-DNA expression vectors were assembled from a single Multi-site Pro LR reaction (1-5-2) with the attR1-attR2 destination vector pYPQ203 (Addgene no. 86207), an attL1-attR5 Cas9 entry clone and an attL5-attL2 CRISPR RNA expression entry clone using Gateway LR clonase II (Invitrogen). Additional T-DNA vectors were also made by conventional cloning. The detailed procedure is described in Supplementary Methods. The resulting 146 T-DNA vectors used in this study were listed in Supplementary Table 5.

Protoplast transformation and stable transformation. The Japonica rice cultivar Nipponbare was used in this study. Polyethylene glycol (PEG) transfection of rice protoplasts was performed at 32 °C based on our previously published protocol[36–38]. For Larix protoplast transformation, Larix seeds were in dark for 14 d at 26 °C to induce callus. Larix callus tissues were transferred into the enzyme solution and hydrolysed for 6 h. The enzyme/protoplasts solution was filtered with 70-μm nylon mesh. The protoplasts were centrifuged at 60*g* for 2 min. Supernatant was removed and the protoplasts were resuspended in W5 solution. The protoplast cells were rested for 30 min. Then, the W5 solution was removed and the protoplasts were resuspended at 2×10^5 ml^{-1} in mannitol magnesium solution. Vector DNA and mannitol magnesium were added to make up 30 μl, and then added to 200 μl of protoplasts (2×10^5 protoplasts). The mixture was incubated in PEG solution for 30 min. The transfection mixture was diluted with 1 ml of W5 solution and centrifuged at 60*g* for 2 min to remove the supernatant. The protoplasts were then gently suspended with the W5 solution in each well of a 12-well tissue culture plate. Protoplast cells were collected for detection after 2 d of dark incubation at 28 °C. Rice stable transformation was carried out by following the same procedure that we published previously[39,40].

Mutagenesis analysis. For assessing mutagenesis in protoplasts, protoplasts of rice, tobacco or larch were collected 48 h after transfection. DNA was extracted using the cetyl trimethylammonium bromide method as previous reported[37,38]. Cas9-induced mutations were generally first detected and quantified by restriction fragment length polymorphism (RFLP) analysis and then followed by NGS. For NGS, the genomic regions flanking the target sites were PCR-amplified using barcoded primers. The PCR amplicons were sequenced by Novogene with an Illumina HiseqX platform. CRISPRMatch[41] was used to analyse the sequencing data. For assessing mutagenesis in stable transgenic rice T_0 lines, single strand conformation polymorphism, RFLP and Sanger sequencing were used as in our previous studies[21,39,40,42]. Sanger sequencing was also used to detect possible mutations for vector self-editing as well as off-target mutations at putative off-target sites, which were predicted by Cas-OFFinder[24].

Screen for herbicide resistant rice lines. After *Agrobacterium*-mediated transformation, the rice calli were selected on 50 mg l^{-1} of hygromycin medium for 2 weeks at 32 °C in light. Actively grown calli were selected on medium containing 50 mg l^{-1} of hygromycin and 0.4 μmol l^{-1} of bispyribac sodium at 28 °C with a 16 h light/8 h dark cycle. After 3–4 weeks, transgenic and herbicide resistant seedlings were identified.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Regarding accession codes, the five Gateway-compatible Cas9 entry vectors are available from Addgene: pYPQ166-SpRY (no. 161520, zSpRY), pYPQ266E (no. 161521, SpRY-D01A-PmCDA1-UGI), pYPQ262m (no. 161522, wtTadA-TadA*-zSpCas9-D10A), pYPQ262-ABE8e (no. 161523, TadA8e-zSpCas9-D10A) and pYPQ262B-ABE8e (no. 161524, TadA8e-zSpRY-D10A). The high-throughput sequencing data sets have been submitted to the National Center for Biotechnology information database under Sequence Read Archive Bio Project ID PRJNA665932.

Received: 30 September 2020; Accepted: 3 December 2020;
Published online: 4 January 2021

References

1. Zhang, Y., Malzahn, A. A., Sretenovic, S. & Qi, Y. The emerging and uncultivated potential of CRISPR technology in plant science. *Nat. Plants* **5**, 778–794 (2019).
2. Tang, X. et al. A CRISPR–Cpf1 system for efficient genome editing and transcriptional repression in plants. *Nat. Plants* **3**, 17018 (2017).
3. Ming, M. et al. CRISPR–Cas12b enables efficient plant genome engineering. *Nat. Plants* **6**, 202–208 (2020).
4. Zhu, H., Li, C. & Gao, C. Applications of CRISPR–Cas in agriculture and plant biotechnology. *Nat. Rev. Mol. Cell Biol.* **21**, 661–677 (2020).
5. Schindele, A., Dorn, A. & Puchta, H. CRISPR/Cas brings plant biology and breeding into the fast lane. *Curr. Opin. Biotechnol.* **61**, 7–14 (2020).
6. Tang, X. et al. A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. *Genome Biol.* **19**, 84 (2018).
7. Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
8. Walton, R. T., Christie, K. A., Whittaker, M. N. & Kleinstiver, B. P. Unconstrained genome targeting with near-PAMless engineered CRISPR–Cas9 variants. *Science* **368**, 290–296 (2020).
9. Richter, M. F. et al. Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat. Biotechnol.* **38**, 883–891 (2020).
10. Steinert, J., Schiml, S., Fauser, F. & Puchta, H. Highly efficient heritable plant genome engineering using Cas9 orthologues from *Streptococcus thermophilus* and *Staphylococcus aureus*. *Plant J.* **84**, 1295–1305 (2015).
11. Kaya, H., Mikami, M., Endo, A., Endo, M. & Toki, S. Highly specific targeted mutagenesis in plants using *Staphylococcus aureus* Cas9. *Sci. Rep.* **6**, 26871 (2016).
12. Kleinstiver, B. P. et al. Engineered CRISPR–Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481–485 (2015).
13. Hu, X. et al. Expanding the range of CRISPR/Cas9 genome editing in rice. *Mol. Plant* **9**, 943–945 (2016).
14. Endo, M. et al. Genome editing in plants by engineered CRISPR–Cas9 recognizing NG PAM. *Nat. Plants* **5**, 14–17 (2019).
15. Hua, K., Tao, X., Han, P., Wang, R. & Zhu, J. K. Genome engineering in rice using Cas9 variants that recognize NG PAM sequences. *Mol. Plant* **12**, 1003–1014 (2019).
16. Ren, B. et al. 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 (2019).
17. Zhong, Z. Improving plant genome editing with high-fidelity xCas9 and non-canonical PAM-targeting Cas9-NG. *Mol. Plant* **12**, 1027–1036 (2019).
18. Sretenovic, S. et al. Expanding plant genome editing scope by an engineered iSpyMacCas9 system targeting the A-rich PAM sequences. *Plant Commun.* https://doi.org/10.1016/j.xplc.2020.100101 (2020).
19. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR–Cas systems. *Nat. Biotechnol.* **31**, 233–239 (2013).
20. Zhong, Z. et al. Plant genome editing using FnCpf1 and LbCpf1 nucleases at redefined and altered PAM sites. *Mol. Plant* **11**, 999–1002 (2018).
21. Zhou, J. et al. CRISPR–Cas9 based genome editing reveals new insights into microRNA function and regulation in rice. *Front. Plant Sci.* **8**, 1598 (2017).
22. Rodriguez-Leal, D., Lemmon, Z. H., Man, J., Bartlett, M. E. & Lippman, Z. B. Engineering quantitative trait variation for crop improvement by genome editing. *Cell* **171**, 470–480 (2017).
23. Hu, J. H. et al. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**, 57–63 (2018).
24. Bae, S., Park, J. & Kim, J. S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **30**, 1473–1475 (2014).
25. Shimatani, Z. et al. Targeted base editing in rice and tomato using a CRISPR–Cas9 cytidine deaminase fusion. *Nat. Biotechnol.* **35**, 441–443 (2017).
26. Tang, X. et al. Single transcript unit CRISPR 2.0 systems for robust Cas9 and Cas12a mediated plant genome editing. *Plant Biotechnol. J.* **17**, 1431–1445 (2019).
27. Nishida, K. et al. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* **353**, aaf8729 (2016).
28. Kuang, Y. et al. Base-editing-mediated artificial evolution of OsALS1 in planta to develop novel herbicide-tolerant rice germplasms. *Mol. Plant* **13**, 565–572 (2020).
29. Zhang, Y. & Qi, Y. CRISPR enables directed evolution in plants. *Genome Biol.* **20**, 83 (2019).
30. Li, C. et al. Targeted, random mutagenesis of plant genes with dual cytosine and adenine base editors. *Nat. Biotechnol.* **38**, 875–882 (2020).
31. Butt, H. et al. CRISPR-directed evolution of the spliceosome for resistance to splicing inhibitors. *Genome Biol.* **20**, 73 (2019).
32. Lapinaite, A. et al. DNA capture by a CRISPR–Cas9-guided adenine base editor. *Science* **369**, 566–571 (2020).
33. Koblan, L. W. et al. Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat. Biotechnol.* **36**, 843–846 (2018).
34. Allen, F. et al. Predicting the mutations generated by repair of Cas9-induced double-strand breaks. *Nat. Biotechnol.* **37**, 64–72 (2018).
35. Arbab, M. et al. Determinants of base editing outcomes from target library analysis and machine learning. *Cell* **182**, 463–480 (2020).
36. Tang, X. et al. A single transcript CRISPR-Cas9 system for efficient genome editing in plants. *Mol. Plant* **9**, 1088–1091 (2016).
37. Zhong, Z. et al. Intron-based single transcript unit CRISPR systems for plant genome editing. *Rice* **13**, 8 (2020).
38. Ren, Q. et al. Bidirectional promoter-based CRISPR-Cas9 systems for plant genome editing. *Front Plant Sci.* **10**, 1173 (2019).
39. Zhou, J. et al. Multiplex QTL editing of grain-related genes improves yield in elite rice varieties. *Plant Cell Rep.* **38**, 475–485 (2019).
40. Wang, B. et al. Targeted mutagenesis of NAC transcription factor gene, *OsNAC041*, leading to salt sensitivity in rice. *Rice Sci.* **26**, 98–108 (2019).
41. You, Q. et al. CRISPRMatch: an automatic calculation and visualization tool for high-throughput CRISPR genome-editing data analysis. *Int. J. Biol. Sci.* **14**, 858–862 (2018).
42. Zheng, X. et al. Effective screen of CRISPR/Cas9-induced mutants in rice by single-strand conformation polymorphism. *Plant Cell Rep.* **35**, 1545–1554 (2016).

Acknowledgements

This work was supported by the National Transgenic Major Project (award no. 2018ZX08020-003), the National Natural Science Foundation of China (award nos. 31771486, 32072045 and 31960423), the State Key Laboratory of Rice Biology (20200205) and the Science Strength Promotion Program of the University of Electronic Science and Technology of China (UESTC) to Yong Zhang and X.Z. This work was also supported by the National Science Foundation Plant Genome Research Program grants (award nos. IOS-1758745 and IOS-2029889), the US Department of Agriculture Biotechnology Risk Assessment Grant Program competitive grants (award nos. 2018-33522-28789 and 2020-33522-32274) and the Emergency Citrus Disease Research and Extension Program (award no. 2020-70029-33161) to Y.Q. S.S. is a Foundation for Food and Agriculture Research Fellow. Y.C. was supported by a scholarship from China Scholarship Council. The content of this publication is solely the responsibility of the authors and does not necessarily represent the official views of these funding agencies.

Author contributions

Y.Q. and Yong Zhang designed the experiments. S.S., Q.R., S.L., Y.C., D.Y., C.P. and Yingxiao Zhang generated all the constructs. Q.R. and S.L. carried out rice protoplast transformation and data analysis. X.T. performed the data analysis. L.H. performed the Dahurian larch protoplast transformation and analysis. Q.R., S.L., Y.H., L.L. and Y.G. conducted rice stable transformation. Q.R. and S.L. analysed rice transgenic lines. L.L., Z.Z., G.L. and X.Z. helped with rice phenotype and genotype data analysis. W.L., L.Q. and C.L. collected the Dahurian larch material and developed the larch callus culture protocol. Y.Q., Yong Zhang, Q.R. and S.S. wrote the paper with input from other authors. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41477-020-00827-4.

Correspondence and requests for materials should be addressed to Y.Q. or Y.Z.

Peer review information *Nature Plants* thanks Arjun Khakhar the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Yiping Qi, Yong Zhang

Last updated by author(s): 2020/12/01

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	None used
Data analysis	CRISPRmatch was used for analysis of PCR amplicon data; Cas-OFFinder was used to identify top off-target sites.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Accession codes. The five Gateway® compatible Cas9 entry vectors are available from Addgene: pYPQ166-SpRY (# 161520, zSpRY), pYPQ266E (# 161521, SpRY-D01A-PmCDA1-UGI), pYPQ262m (# 161522, wtTadA-TadA*-zSpCas9-D10A), pYPQ262-AVE8e (# 161523, TadA8e-zSpCas9-D10A), and pYPQ262B-ABE8e (# 161524, TadA8e-zSpRY-D10A). The high-throughput sequencing data sets have been submitted to the National Center for Biotechnology information (NCBI) database under Sequence Read Archive (SRA) Bio Project ID: PRJNA665932.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For rice protoplasts assays, three biological replicates were used for each sample. Specifically, 360 µL rice protoplasts (2x1000,000 cells/mL) were used for PEG-mediated transfection per sample. For larix protoplasts assays, 230 µ larix protoplasts (2×100,000 cells/mL) were used for PEG-mediated transfection per sample. For stable rice transformation, 15 to 24 individual transgenic lines were evaluated in T0 rice plants for each construct.
Data exclusions	No data was excluded.
Replication	For rice and larix protoplasts assays, three biological replicates were used for each construct. The biological replicates were done at intervals ranging from days to weeks between experiments and performed by up to three different researchers. All findings have been replicated.
Randomization	Rice and larix protoplasts were isolated and randomly separated to each transformation. The constructs used for transfection didn't need to follow any specific order.
Blinding	Not applicable, as no bias will be introduced by knowing the identity of each sample. All samples were processed identically through standard procedures (e.g. DNA transfection, DNA isolation, and DNA sequencing).

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging