nature biotechnology

ARTICLES

https://doi.org/10.1038/s41587-022-01254-w

Check for updates

An engineered prime editor with enhanced editing efficiency in plants

Yuan Zong[1,6], Yijing Liu[1,2,6], Chenxiao Xue[1,2,6], Boshu Li[1,2], Xiangyang Li[3], Yanpeng Wang[1], Ji Li[2,4], Guanwen Liu[1], Xingxu Huang[3], Xiaofeng Cao[2,4,5] ✉ and Caixia Gao[1,2] ✉

Prime editing is a versatile genome-editing technology, but it suffers from low editing efficiency. In the present study, we introduce optimized prime editors with substantially improved editing efficiency. We engineered the Moloney–murine leukemia virus reverse transcriptase by removing its ribonuclease H domain and incorporated a viral nucleocapsid protein with nucleic acid chaperone activity. Each modification independently improved prime editing efficiency by ~1.8–3.4-fold in plant cells. When combined in our engineered plant prime editor (ePPE), the two modifications synergistically enhanced the efficiency of base substitutions, deletions and insertions at various endogenous sites by on average 5.8-fold compared with the original PPE in cell culture. No significant increase in byproducts or off-target editing was observed. We used the ePPE to generate rice plants tolerant to sulfonylurea and imidazolinone herbicides, observing an editing frequency of 11.3% compared with 2.1% using PPE. We also combined ePPE with the previously reported dual-prime editing guide (peg) RNAs and engineered pegRNAs to further increase efficiency.

Development of powerful genome-editing tools that precisely and efficiently manipulate the genome of living systems is critical for biomedical research, agricultural breeding, pharmaceuticals and therapeutic applications[1,2]. Nevertheless, precise targeted mutagenesis conferring one or more nucleotide conversions, insertions or deletions is challenging due to low homology-directed repair-mediated genome editing, and only a subset of single base conversions can be achieved using base editors[1–5]. Prime editing (PE) is a newly developed genome-editing tool that can precisely enable the installation of all 12 nucleotide substitutions, short insertions and short deletions[6]. Prime editors are protein complexes comprising a nickase Cas9 (His840Ala) and a Moloney–murine leukemia virus reverse transcriptase (M-MLV RT), and editing events are encoded by a pegRNA that serves as a template for reverse transcription directly into the genome of living cells[6]. PE has been demonstrated as successful in editing the genomes of mammalian cells, organoids, rice, wheat, maize, *Drosophila* spp., mouse, zebrafish and rabbits[7–13]. Despite this versatility, the editing efficiency of current prime editors is low and often variable between different target sites and cell types[2,6,9]. Recent efforts to improve prime editing efficiencies are mainly focused on pegRNA engineering such as the enhancement of pegRNA expression through the use of polycistronic transfer RNAs and ribozymes[9,13], designing the pegRNA sequence based on melting temperature preferences[14], using dual-pegRNAs[14], pegRNA processing by the RNA endoribonuclease Csy4 (ref. [15]) and enhancing pegRNA stability using engineered (e)pegRNAs[16]. In the present study, we developed a series of new prime editors through engineering the nCas9–RT fusion protein. We found that the combination of deleting the M-MLV RT ribonuclease H (RNase H) domain and the addition of a viral nucleocapsid (NC) protein synergistically and broadly improves prime editing efficiency at a variety of target sites in rice and wheat, increasing the flexibility and applicability of PE.

Results

Optimized PPEs by engineering reverse transcriptase and fusion of viral proteins. We reasoned that the very-low-to-modest editing efficiency of existing prime editors may be a result of low M-MLV RT activity, so we anticipate that engineering the M-MLV RT may significantly improve enzyme activity and DNA synthesis efficiency during PE (Fig. 1a–c). M-MLV RT is composed of fingers, palm, thumb and connection domains, each having a unique role in nucleotide incorporation during DNA synthesis (Supplementary Fig. 1). There is also a RNase H domain that functions as a processive endonuclease cleaving the RNA strand in RNA–DNA heteroduplexes[17–19]. We first introduced four single amino-acid substitutions previously demonstrated to enhance DNA polymerization activity and increase overall thermostability[20–22] into the M-MLV RT finger or palm domain of the PPE[9], resulting in PPE-F155Y, PPE-F155V, PPE-F156Y and PPE-N200C (the original Asp200 had been mutated to Asn200 in PE2 (ref. [6])) (Fig. 1a,b, Supplementary Sequences and Supplementary Notes 1 and 2). We hypothesized that inactive or abolished M-MLV RT RNase H activity may inhibit RNase H-directed degradation of the RNA strand in a single guide (sg)RNA–DNA heteroduplex and improve the overall stability of the prime editor complex. Toward this end, we engineered three RNase H-inactivated complexes through the introduction of an inactivating Asp524Asn substitution into the RNase H domain[23], deletion of the whole RNase H domain or simultaneous deletion of the RNase H domain and the connection domain (linked to the RNase H domain)[17,19], resulting in three new PPE variants PPE-D524N, PPE-ΔRNase H and PPE-ΔRNase H-Δconnection, respectively

[1]State Key Laboratory of Plant Cell and Chromosome Engineering, Center for Genome Editing, Institute of Genetics and Developmental Biology, Innovation Academy for Seed Design, Chinese Academy of Sciences, Beijing, China. [2]College of Advanced Agricultural Sciences, University of Chinese Academy of Sciences, Beijing, China. [3]School of Life Science and Technology, ShanghaiTech University, Shanghai, China. [4]State Key Laboratory of Plant Genomics, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China. [5]Center for Excellence in Molecular Plant Sciences, Chinese Academy of Sciences, Beijing, China. [6]These authors contributed equally: Yuan Zong, Yijing Liu, Chenxiao Xue. ✉e-mail: xfcao@genetics.ac.cn; cxgao@genetics.ac.cn

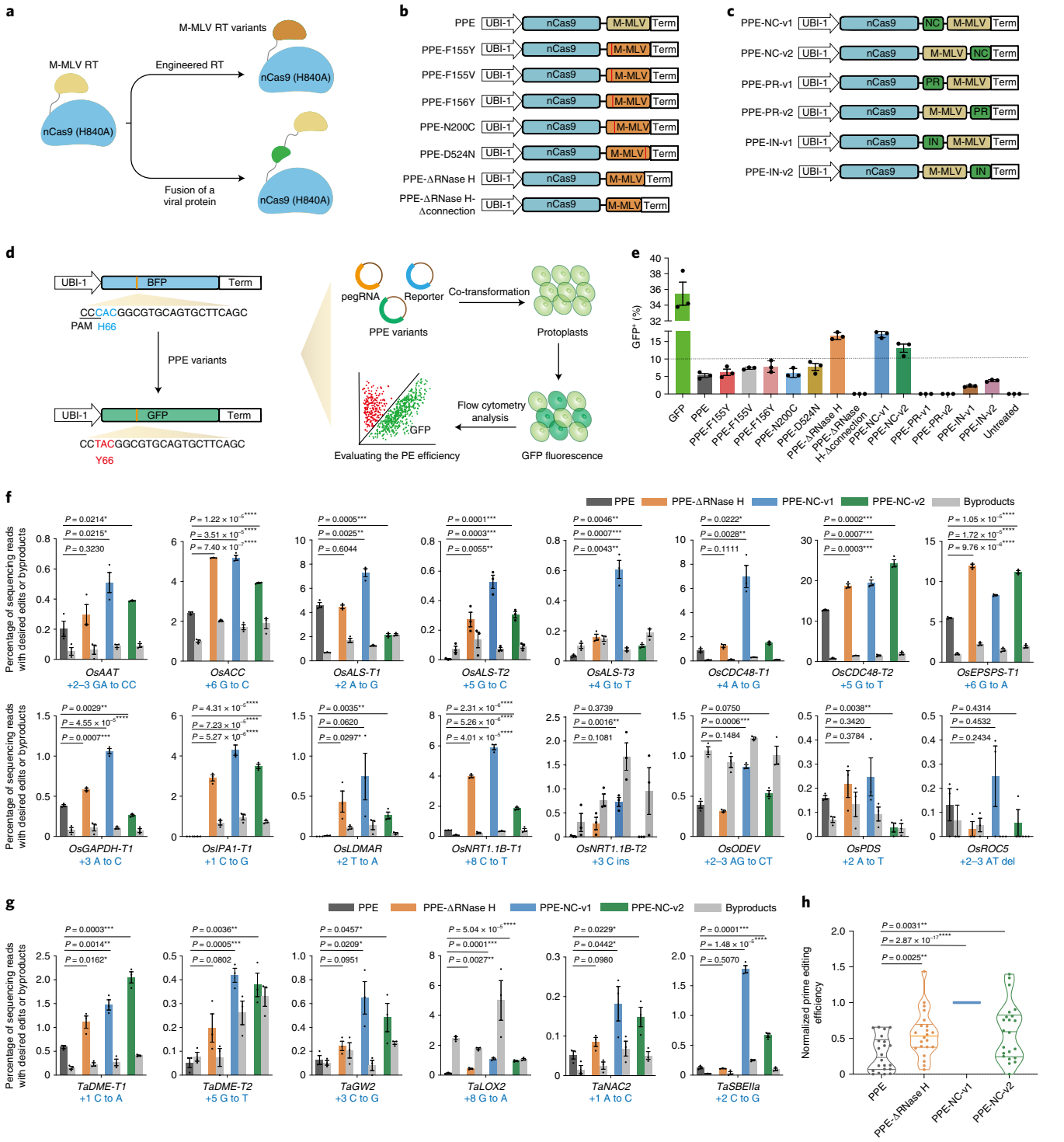


Fig. 1 | Improvement of prime editing efficiency by removing the RT RNase H domain and an addition of a viral NC protein in plant cells. a, Schematic diagram of engineering a prime editor by two methods: amino-acid substitutions or truncation of the M-MLV RT, and fusion of the viral proteins. **b**, Schematic representation of the PPE, PPE-F155Y, PPE-F155V, PPE-F156Y, PPE-N200C, PPE-D524N, PPE-ΔRNase H and PPE-ΔRNase H-Δconnection constructs. **c**, Representation of the PPE-NC-v1, PPE-PR-v1 and PPE-IN-v1 (fusion of NC, PR and IN proteins between nCas9 and RT), and PPE-NC-v2, PPE-PR-v2 and PPE-IN-v2 (fusion of NC, PR and IN proteins at the C terminus of RT) constructs. **d**, Schematic diagram of the BFP-to-GFP reporter system for prime editing through flow cytometry (FCM) analysis after transformation into plant protoplasts. The prime editors could change the BFP to GFP by changing CAC (histidine, H) to TAC (tyrosine, Y). The PAM motif of pegRNA is underlined. **e**, Frequencies (%) of BFP-to-GFP conversion in rice protoplasts measured by FCM. **f**,**g**, Frequencies of prime editing and byproducts induced by PPE, PPE-ΔRNase H, PPE-NC-v1 and PPE-NC-v2 at 16 rice target sites (**f**) and 6 wheat target sites (**g**). Frequencies (mean ± s.e.m.) were calculated from three independent experiments ($n = 3$) in **e**–**g**. ins, insertion; del, deletion. **h**, Overall editing frequencies induced by PPE, PPE-ΔRNase H, PPE-NC-v1 and PPE-NC-v2. The average editing frequencies using PPE-NC-v1 for each target were normalized to 1, and the frequencies using PPE, PPE-ΔRNase H and PPE-NC-v2 for each target were adjusted accordingly. *P* values were obtained using the two-tailed Student's *t*-test: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

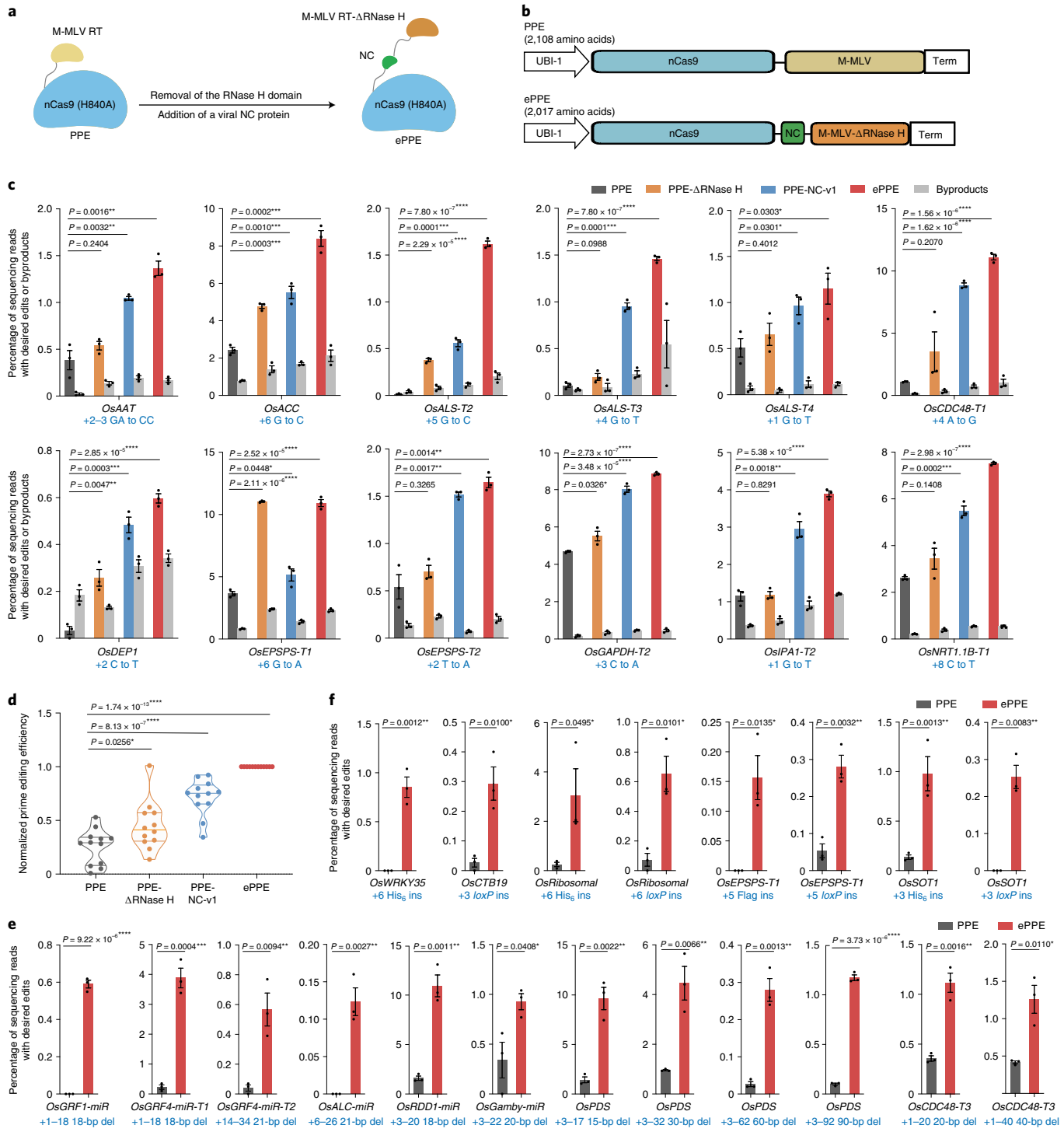


Fig. 2 | Engineered prime editors for precise genome editing in plant cells. a, Schematic diagram of the ePPE by deleting the RNase H domain of M-MLV RT and simultaneously fusing with a viral NC protein. **b**, Representation of the ePPE and PPE constructs, fusing the NC protein between nCas9 (with XTEN linker) and deleted RNase H domain RT (with 32-amino-acid linker). **c**, Comparison of the prime editing efficiencies and byproduct efficiencies of four different prime editors (PPE, PPE-ΔRNase H, PPE-NC-v1 and ePPE) at 12 target sites in rice protoplasts. **d**, Overall editing frequencies induced by PPE, PPE-ΔRNase H, PPE-NC-v1 and ePPE. The average editing frequencies using ePPE for each target were normalized to 1, and the frequencies using PPE, PPE-ΔRNase H and PPE-NC-v1 for each target were adjusted accordingly. **e**,**f**, Comparison of targeted precise deletions of ~15- to 90-bp (**e**) and precise insertions of ~18- to 34-bp (**f**) induced by ePPE and PPE. Frequencies (mean ± s.e.m.) were calculated from three independent experiments ($n = 3$) in **c**, **e** and **f**. *P* values were obtained using the two-tailed Student's *t*-test: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

(Fig. 1a,b). We first compared the editing efficiency of these seven PPE variants with the original PPE system in plant cells using a blue fluorescent protein (BFP)-to-green fluorescent protein (GFP) reporter system (Fig. 1d). A pegRNA-targeting BFP, with an RT template encoding the conversion of codon 66 from CAC (histidine) to TAC (tyrosine), was designed to convert BFP to GFP. We introduced

Fig. 3 | Effect of prime editors on off-target prime editing. a, Effect of mismatched pegRNAs on prime editing in rice protoplasts. This shows the editing efficiencies of prime editors paired with either an on-target pegRNA or a mismatched pegRNA at six target sites in rice protoplasts. PBS, primer binding site. **b**, Editing efficiencies of prime editors at 11 endogenous on-target and 29 off-target (OT) sites with 1–3 mismatches identified by Cas-OFFinder. The mismatched nucleotides are shown in red. Frequencies (mean ± s.e.m.) were calculated from three independent experiments ($n = 3$).

a

Target site	Prime editor	Edited type	No. of resistant calli	No. of mutant calli	PE frequency (%)
OsIPA1-T1	PPE	+1 C to G	191	0	0
	ePPE		174	8	4.6
OsALS-T6	PPE	+4–5 TG to AT	384	4	1.0
	ePPE		346	11	3.2
OsCDC48-T1	PPE	+4 A to G	132	4	3.0
	ePPE		112	12	10.7
OsRDD1-miR	PPE	+3–20 18-bp del	108	3	2.8
	ePPE		146	46	31.5

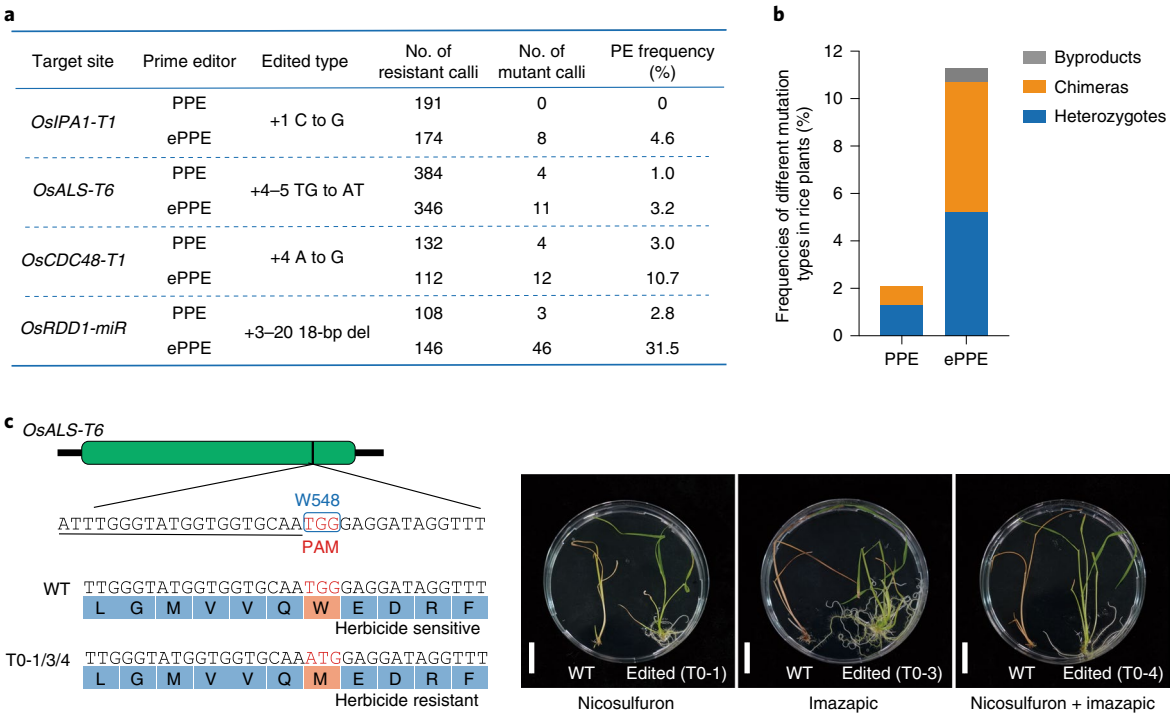


Fig. 4 | Engineered prime editors generated mutations in plants. a, Comparison of editing efficiencies induced by PPE and ePPE in resistant rice calli at four target sites. **b**, Frequencies of the different mutation types among the T0 mutants at *OsALS-T6* induced by PPE and ePPE. **c**, *OsALS-T6* with Trp548Met substitution conferring resistance to herbicides. The sequence alignment compares WT *OsALS-T6* with that in the T0-1, T0-3 and T0-4 mutants. PAM is highlighted in red and the protospacer is underlined. Phenotypes of WT (left) and heterozygous *OsALS* W548M mutants (right) grown in the regeneration medium supplemented with 1.10 μl l^{-1} of imazapic, 0.09 mg l^{-1} of nicosulfuron or both imazapic and nicosulfuron. Scale bars, 1 cm. One biological experiment was performed.

this pegRNA with each of the PPE constructs into rice protoplasts to compare relative editing efficiency (Fig. 1d). Flow cytometry analyses showed that the PPE-ΔRNase H construct yielded the highest percentage of GFP-expressing cells (16.6%), which reflects a 3.1-fold improvement compared with the original PPE (5.3%), followed by PPE-F156Y, PPE-D524N, PPE-F155V, PPE-F155Y and PPE-N200C, the efficiencies of which were slightly higher than or comparable to the editing efficiency of PPE (Fig. 1e, Supplementary Fig. 2 and Supplementary Data). When tested at endogenous genes, these five point-mutation prime editors displayed comparable or decreased editing efficiency compared with PPE (Extended Data Fig. 1). Therefore, these amino-acid substitutions were not considered in any further study. Surprisingly, we found that the PPE-ΔRNase H-Δconnection gave no editing efficiency (Fig. 1e), suggesting that the connection sequences are essential for M-MLV RT activity, and solely removing the RNase H domain can largely improve prime editing efficiency in rice protoplasts.

Reverse transcription of viral genomic RNA requires an RT enzyme and is further facilitated either directly or indirectly by other viral proteins[24], including an NC protein that has nucleic acid chaperone activity affecting a variety of RT-related functions[25,26], a protease (PR) that is essential for replication and cleaves polyproteins leading to virus maturation[27] and a viral integrase (IN) that integrates the newly synthesized DNA into the host cell genome[25]. Guided by these natural functions, we first codon optimized these proteins for expression in cereals and subsequently fused them either between the nCas9 and the M-MLV RT (named variant 1, v1) or at the C terminus of the nCas9–RT protein (named variant 2, v2), resulting in a total of six constructs including PPE-NC-v1, PPE-NC-v2, PPE-PR-v1, PPE-PR-v2, PPE-IN-v1 and PPE-IN-v2 (Fig. 1a,c and Supplementary Sequences). We evaluated these new PPE constructs using the above-described BFP-to-GFP reporter system and observed that PPE-NC-v1 and PPE-NC-v2 displayed a 3.2-fold (17.0%) and a 2.5-fold (13.1%) improvement in editing efficiency compared with the original PPE, respectively (Fig. 1e, Supplementary Fig. 2 and Supplementary Data). However, fusion of the IN or PR viral proteins reduced prime editing activity (Fig. 1e), suggesting that only the NC protein and not the IN or PR proteins can chaperone reverse transcription during prime editing. Collectively, these results suggest that the PPE-ΔRNase H, PPE-NC-v1 and PPE-NC-v2 can efficiently enhance plant prime editing efficiency.

Next, we evaluated prime editing efficiency of PPE-ΔRNase H, PPE-NC-v1 and PPE-NC-v2 at endogenous sites by delivering these editors into protoplasts, and analyzed their editing efficiency using targeted amplicon sequencing (Fig. 1f,g). We evaluated a total of 22 genomic sites using pegRNAs designed with previously published general design guidelines[6] or PlantPegDesigner[14], including 16 genomic sites in rice and 6 genomic sites in wheat (Supplementary Table 1). Prime editing efficiencies were all significantly improved using these three newly optimized PPEs. PPE-NC-v1 showed the highest editing efficiencies, ranging from 0.2% to 19.5% in both rice and wheat protoplasts, followed by PPE-NC-v2 and PPE-ΔRNase H (Fig. 1f,g). These three new PPEs demonstrated an average of 3.4-fold, 2.1-fold and 2.0-fold improved editing efficiency, respectively, compared with the original PPE (Fig. 1h). Strikingly, PPE-NC-v1 was greatly improved at sites *OsALS-T2* (+5 G-to-C), *OsIPA1-T1* (+1 C-to-G), *OsLDMAR* (+2 T-to-A) and *OsNRT1.1B-T2* (+3 C insertions (ins)) (up to 4.3%), in all of which the original PPE resulted in undetectable levels of editing (Fig. 1f). Furthermore, although the exact values of byproducts marginally increased at some target sites, the ratio edit:byproduct at most

tested sites was increased or comparable using these three new PPE constructs compared with the original PPE (Fig. 1f,g and Extended Data Fig. 2). Together, these results demonstrate that the removal of the RNase H domain or incorporation of an NC peptide between the nCas9 and M-MLV RT domains can result in large improvements to prime editing in plants.

Engineered PPE with enhanced prime editing in protoplasts. As PPE-ΔRNase H and PPE-NC-v1 independently increased editing activity at the tested target sites, we speculated that combining these two strategies might further improve plant prime editing activity. To test this hypothesis, we constructed a new ePPE by fusing the NC protein between the nCas9 and M-MLV–ΔRNase H domain with an XTEN linker and a 32-amino-acid linker, respectively, resulting in a new complex with a total size that is 91 amino acids smaller than the PPE (Fig. 2a,b and Supplementary Sequences). We designed 12 target sites across 9 rice genes to compare the activity of ePPE, PPE-ΔRNase H, PPE-NC-v1 and PPE (Supplementary Table 1). Targeted amplicon sequencing demonstrated that ePPE displayed a substantial improvement and gave the highest editing efficiency compared with the other complexes, resulting in a 1.9- to 121.5-fold (average 3.9-fold) improvement to editing compared with PPE, a 1.0- to 7.3-fold (average 2.2-fold) improvement compared with PPE-ΔRNase H and up to a 4.2-fold (average 1.4-fold) improvement compared with PPE-NC-v1 across all tested genomic sites (Fig. 2c,d). The ePPE efficiently generated G-to-C, A-to-G, G-to-T, C-to-T, G-to-A, T-to-A, C-to-A and GA-to-CC base substitutions at an average efficiency of 4.9% (Fig. 2c). In addition, most sites appeared to show no apparent change in the proportion of byproducts realized, and perhaps even resulted in higher edit:byproduct ratios using ePPE compared with the original PPE at all sites except the *OsAAT* and *OsGAPDH-T2* targets (Fig. 2c and Extended Data Fig. 3). To summarize, these results indicate that combining the removal of the RNase H domain and the addition of an NC protein resulted in a synergistic effect to further enhance the installation of precise edits using prime editing in plants.

To ensure that this improvement in PE efficiency was not limited to small edits, such as base substitutions, we tested 20 additional pegRNAs that encoded 12 larger deletions ranging from 15 bp to 90 bp and 8 larger insertions ranging from 18 bp to 34 bp across multiple genomic sites (Supplementary Table 1). We observed that ePPE enabled editing efficiency averaging 2.9% (up to 10.9%), which is on average a 6.5-fold improvement compared with PPE for 15-, 18-, 20-, 21-, 30-, 40-, 60- and 90-bp deletions (Fig. 2e and Extended Data Fig. 4a). We also used ePPE to perform precise insertions, including a His_6 tag (18 bp, up to 3.1% efficiency), a Flag epitope tag (24 bp, 0.2%) and an extended Cre recombinase *loxP* site (34 bp, ~0.3%) into multiple genomic sites at which PPE produced almost undetectable editing levels (Fig. 2f). The ePPE is successful at generating precise large deletions and insertions, which makes the manipulation of regulatory elements possible. Collectively, the prime editing experiments described above demonstrate that the use of ePPE resulted in, on average, a 5.8-fold improvement in prime editing efficiency comprising various base substitutions, small insertions and deletions, and large precise insertions and deletions compared with PPE across 32 genomic sites (Extended Data Fig. 4b). These results establish that this engineered plant prime editor is a remarkably versatile genome-editing technology.

Expanding the scope of ePPE using SpG Cas9. To expand the targeting scope of prime editors, we replaced the nCas9 (His840Ala) domain in ePPE with a codon-optimized SpG (His840Ala) variant to produce ePPE–SpG[28] (Extended Data Fig. 5a and Supplementary Sequences). We found that ePPE–SpG maintains a broad targeting range to prime edit at NG protospacer adjacent motif (PAM) sequences including NGC, NGA and NGG, with efficiencies ranging from 0.4% to 7.5% across four target sites (Extended Data Fig. 5b), which substantially expands the targeting scope and capabilities of prime editing.

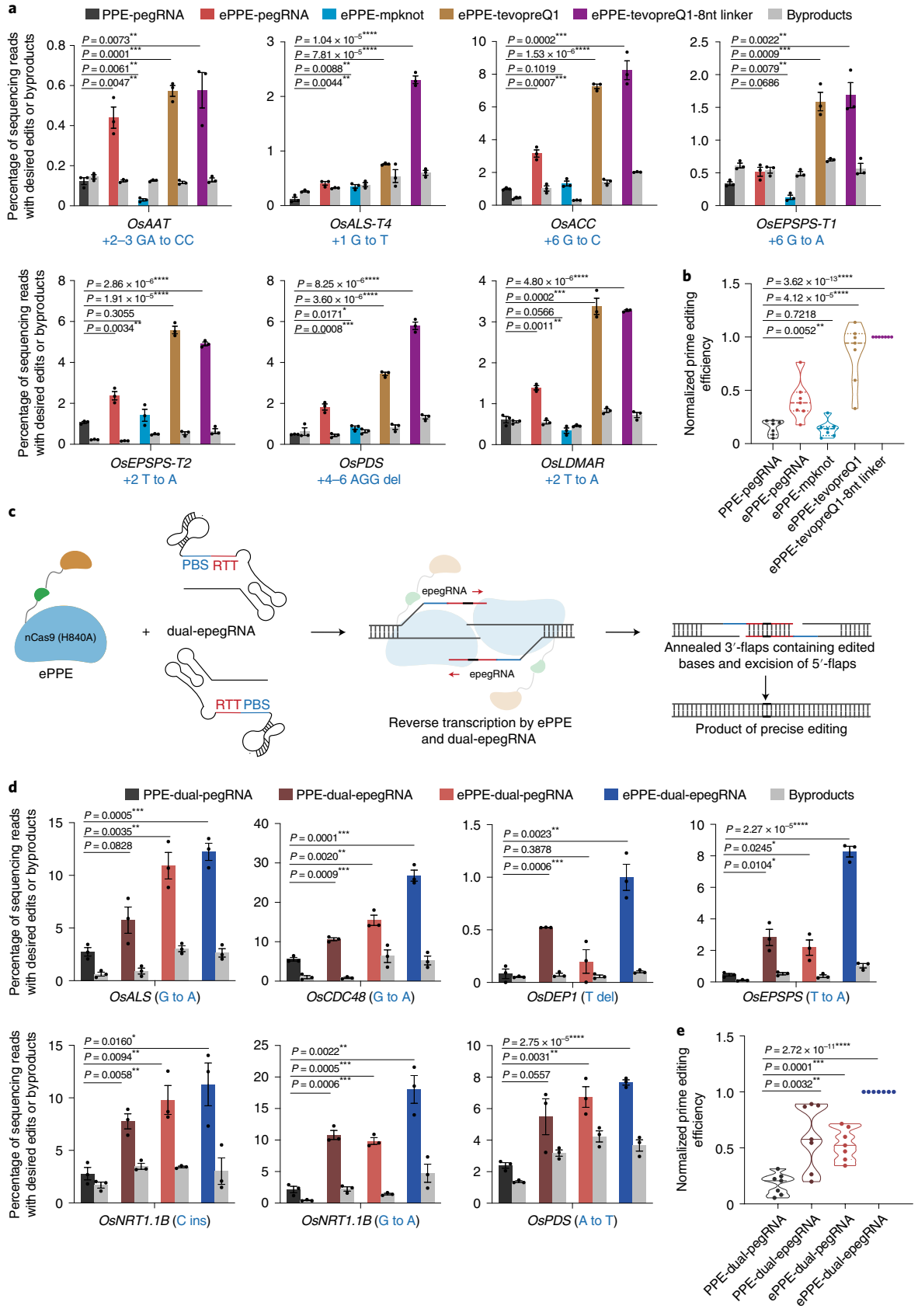
Effect of ePPE on off-target prime editing. Our group previously observed that PPE resulted in low levels of pegRNA-dependent off-target edits in plants[29]. To determine whether ePPE significantly changed the extent of off-target editing, we first tested ePPE's tolerance to mismatches in the pegRNA, including in the spacer or both in the primer binding site and spacer in rice protoplasts. Overall, ePPE resulted in only slightly higher off-target editing efficiency at certain sites (Fig. 3a and Supplementary Table 2). We also tested the off-target effect of ePPE at endogenous sites comprising 1–3 mismatches in the spacers for 11 pegRNAs, resulting in a total of 29 off-target sites. Deep sequencing revealed that ePPE and PPE both exhibited very low off-target prime editing efficiency at all examined sites except one, namely *OsCDC48-T1* (OT-22), which showed higher editing efficiency induced by ePPE compared with PPE (Fig. 3b and Supplementary Table 3). These results suggest that ePPE does not result in a significantly increased level of off-target editing at most target sites compared with the canonical PPE.

Prime editing compared with base editing. In addition, we compared PEs with cytosine base editors (CBEs)[30] or adenine base editors (ABEs)[31] at three or four genomic loci, respectively. For overall editing efficiency, A3A-PBE or PABE8 displayed higher editing efficiency compared with ePPE when the target C or A is positioned at the center of the editing window, averaging a 2.4-fold increase to editing, but demonstrated lower editing efficiency when the edited base was positioned outside the optimal base-editing window (Extended Data Fig. 6a and Supplementary Table 4). Due to limitations in PAM targeting, it is not always possible to position target bases in the most optimal base-editing window. The efficiency of prime editing greatly exceeds that of base editing when installing precise edits without any bystander editing events (Extended Data Fig. 6b). Collectively, these results indicate that prime editing and base editing offer complementary strengths and weaknesses for making targeted point mutations.

The ePPE-induced desired mutations in resistant rice calli. Next, we targeted four rice loci and constructed a binary expression vector, pH-ePPE (Extended Data Fig. 7a). We first designed the pH-ePPE

Fig. 5 | Prime editing efficiency is enhanced by combining the engineered prime editors with optimized pegRNAs in plant cells. a, Frequencies of prime editing induced by different PPEs with different epegRNAs across seven rice target sites. **b**, Overall editing frequencies induced by different PPEs with different epegRNA forms. The average editing frequencies using ePPE-tevopreQ1-8nt linker for each target were normalized to 1 and the frequencies using other complexes for each target were adjusted accordingly. **c**, Diagram of ePPE using the dual-pegRNA strategy based on the epegRNA scaffold. The dual-epegRNA approach results in the same edit on both DNA strands. RTT, reverse transcription template. **d**, Frequencies of prime editing induced by different PPEs with canonical pegRNAs or epegRNAs based on the dual-pegRNA strategy at seven rice target sites. The edits are named based on the DNA forward strand. Frequencies (mean ± s.e.m.) were calculated from three independent experiments ($n = 3$) in **a** and **d**. **e**, Overall editing frequencies induced by different PPEs with canonical pegRNAs or epegRNAs based on the dual-pegRNA strategy. The average editing frequencies using ePPE-dual-epegRNA for each target were normalized to 1 and the frequencies for the other complexes were adjusted accordingly. *P* values were obtained using the two-tailed Student's *t*-test: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

vectors to harbor each of four unique pegRNAs, and then introduced them into rice calli using *Agrobacterium*-mediated transformation. Resistant calli were selected and identified by polymerase chain reaction-restriction enzyme (PCR-RE) and Sanger sequencing. We found that, in contrast to PPE, ePPE-induced prime editing frequencies of the desired C-to-G edit at the *OsIPA1-T1* site improved from 0% to 4.6%, TG-to-AT at the *OsALS-T6* site improved from 1.0% to 3.2%, A-to-G at the *OsCDC48-T1* site improved from 3.0%



to 10.7% and an 18-bp deletion at the *OsRDD1-miR* site improved from 2.8% to 31.5% (Fig. 4a). In addition, we detected no off-target editing events in rice calli (Supplementary Table 5). Collectively, ePPE resulted in an 8.6-fold improvement in editing efficiency compared with PPE across four genomic sites, suggesting that ePPE is effective at inducing specific mutations in a highly precise manner in rice-resistant calli.

An amino-acid substitution for herbicide resistance in rice. Herbicide resistance in crop plants is critical for integrated weed management in agriculture. It has previously been shown that a TG-to-AT (Trp548Met) replacement in acetolactate synthase (ALS), the first enzyme in the biosynthesis of branched-chain amino acids[32], in rice could endow plants with broad-spectrum resistance against ALS-inhibiting herbicides[33]. Resistant calli from PPE or ePPE treatment that generated this Trp548Met edit (*OsALS-T6*) were transferred to regeneration medium to obtain transgenic plants. Examination of 346 pH-ePPE-transformed lines revealed 39 mutants harboring the TG-to-AT substitution at the target genomic site. The observed mutation efficiency of 11.3% is 5.4-fold higher than that of PPE-mediated editing (2.1%, 5 heterozygous mutants and 3 chimeras in 384 tested plants) (Fig. 4b and Extended Data Fig. 7b,c). Among the 39 mutants treated with ePPE, 18 (5.2%, 18/346) contained heterozygous TG-to-AT substitutions, 19 (5.5%, 19/346) contained chimeric substitutions and the remaining 2 (0.6%, 2/346) contained byproducts (Fig. 4b). Furthermore, we detected no mutations at the predicted off-target sites (Supplementary Table 5). We then assessed the herbicide resistance of mutants carrying the heterozygous Trp548Met substitutions across different herbicides. After 10 d of growth on regeneration medium supplemented with 1.10 μl l^{-1} of imazapic, 0.09 mg l^{-1} of nicosulfuron or both imazapic and nicosulfuron, we found that mutant plants produced new rootlets and had normal phenotypes when grown with both herbicides, whereas wild-type (WT) plants displayed withered leaves and an absence of any rootlet growth (Fig. 4c). These results indicate that ALS-W548M mutants generated by prime editing in rice endows high levels of resistance to a broad spectrum of sulfonylurea- and imidazolinone-type herbicides, which hold great promise for addressing the worsening weed problems in rice cultivation.

The ePPE with optimized pegRNAs further improves prime editing. To further improve plant prime editing efficiency, we combined ePPE with our previously reported dual-pegRNA strategy[14]. We evaluated six targets in rice protoplasts and found that the ePPE–dual-pegRNA strategy substantially improved editing outcomes, averaging a 8.6-, 3.2- and 2.8-fold improvement to editing levels compared with PPE with individual pegRNAs, ePPE with individual pegRNAs or PPE with dual-pegRNAs, respectively (Extended Data Fig. 8 and Supplementary Table 6). In addition, we evaluated combination of the recently reported epegRNAs[16], including tevopreQ1, tevopreQ1-8nt linker and mpknot, with our ePPE complex. We observed that the ePPE–tevopreQ1-8nt linker resulted in the largest improvement, averaging a 2.5-fold improvement to editing efficiency compared with ePPE–pegRNA, and a 6.5-fold improvement compared with PPE–pegRNA, followed by ePPE–tevopreQ1, which reflects 2.1-fold higher editing levels than ePPE–pegRNA and 5.5-fold higher than PPE–pegRNA (Fig. 5a,b). However, the ePPE–mpknot displayed the lowest editing efficiency (Fig. 5a,b). In addition, most epegRNAs, except for the ePPE–mpknot treatment, showed no decrease in edit:byproduct ratios (Extended Data Fig. 9). Prime editing efficiency could be further improved by combining ePPE with a dual-epegRNA approach containing the 3′-tevopreQ1-8nt linker motif (ePPE–dual-epegRNA) (Fig. 5c,d and Extended Data Fig. 10), resulting in an average of 7.9-, 2.3- and 2.0-fold improvement to editing levels compared with PPE–dual-pegRNA, ePPE–dual-pegRNA and PPE–dual-epegRNA, respectively (Fig. 5e). These results demonstrate that prime editing efficiency could be further improved when combining our ePPE with optimized pegRNAs, especially ePPE with dual-epegRNAs.

Discussion

The low editing efficiency of original prime editors severely limits the utility of prime editing[2,6,9]. In contrast to previous studies optimizing the pegRNA, in the present study we engineered the protein component of prime editors. We demonstrated that two unique approaches, deleting the RT RNase H domain and the addition of a viral NC protein, stimulate much higher prime editing efficiency in plants. Combining both methods (resulting in ePPE) cooperatively improved prime editing efficiency of various base substitutions, up to 90-bp deletions and 34-bp insertions in rice and wheat when compared with the original prime editor. In addition, ePPE can induce significantly higher editing efficiency compared with PPE in resistant calli and stable transgenic plants, and can also facilitate the generation of rice plants with herbicide resistance against sulfonylurea- and imidazolinone-type herbicides. We speculate that the synergistic improvement using these two methods is a result of two independent mechanisms. The removal of the RNase H domain further stabilizes the heteroduplex between the sgRNA–DNA and the nCas9–RT–pegRNA complex, whereas the NC viral protein serves as a chaperone during the reverse transcription process via its nucleic acid-annealing activities and its interactions with the RT enzyme[17,26]. Of note, plant prime editing efficiency can be further improved when combining ePPE with a dual-epegRNA strategy. Unfortunately, the editing efficiency of the engineered prime editors was comparable to that of the original prime editor when tested in a variety of human cells (Supplementary Fig. 3), which is consistent with previously reported results that deleting the RNase H domain showed comparable editing efficiencies in HEK293 cells compared with the original PE2 (ref. [34]). We speculate that this may reflect a difference between the reverse transcription processes in mammalian cells and plant cells. Thus, additional engineering efforts, such as the recently reported addition of a DNA-binding domain[35], manipulating cellular determinants of editing outcomes (PE4/PE5)[36] or TwinPE strategies[37], are needed to further enhance the efficiency of prime editors across all cell types. We anticipate that the engineered prime editors described in the present study will propel the field of plant genome editing and provide a new and improved tool for use across a wide range of research and agricultural applications.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41587-022-01254-w.

Received: 19 August 2021; Accepted: 8 February 2022;
Published online: 24 March 2022

References

1. Gao, C. Genome engineering for crop improvement and future agriculture. *Cell* **184**, 1621–1635 (2021).
2. Anzalone, A. V., Koblan, L. W. & Liu, D. R. Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* **38**, 824–844 (2020).
3. Zhu, H., Li, C. & Gao, C. Applications of CRISPR–Cas in agriculture and plant biotechnology. *Nat. Rev. Mol. Cell Biol.* **21**, 661–677 (2020).
4. Chen, K., Wang, Y., Zhang, R., Zhang, H. & Gao, C. CRISPR/Cas genome editing and precision plant breeding in agriculture. *Annu. Rev. Plant. Biol.* **70**, 667–697 (2019).
5. Ran, Y., Liang, Z. & Gao, C. Current and future editing reagent delivery systems for plant genome editing. *Sci. China Life Sci.* **60**, 490–505 (2017).

6. Anzalone, A. V. et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019).
7. Newby, G. A. & Liu, D. R. In vivo somatic cell base editing and prime editing. *Mol. Ther.* **29**, 3107–3124 (2021).
8. Bosch, J. A., Birchak, G. & Perrimon, N. Precise genome engineering in *Drosophila* using prime editing. *Proc. Natl Acad. Sci. USA* **118**, e2021996118 (2021).
9. Lin, Q. et al. Prime genome editing in rice and wheat. *Nat. Biotechnol.* **38**, 582–585 (2020).
10. Petri, K. et al. CRISPR prime editing with ribonucleoprotein complexes in zebrafish and primary human cells. *Nat. Biotechnol.* **40**, 189–193 (2022).
11. Qian, Y. et al. Efficient and precise generation of Tay–Sachs disease model in rabbit by prime editing system. *Cell Discov.* **7**, 50 (2021).
12. Liu, Y. et al. Efficient generation of mouse models with the prime editing system. *Cell Discov.* **6**, 27 (2020).
13. Jiang, Y. Y. et al. Prime editing efficiently generates W542L and S621I double mutations in two ALS genes in maize. *Genome Biol.* **21**, 257 (2020).
14. Lin, Q. et al. High-efficiency prime editing with optimized, paired pegRNAs in plants. *Nat. Biotechnol.* **39**, 923–927 (2021).
15. Liu, Y. et al. Enhancing prime editing by Csy4-mediated processing of pegRNA. *Cell Res.* **31**, 1134–1136 (2021).
16. Nelson, J. W. et al. Engineered pegRNAs improve prime editing efficiency. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-021-01039-7 (2021).
17. Das, D. & Georgiadis, M. M. The crystal structure of the monomeric reverse transcriptase from Moloney murine leukemia virus. *Structure* **12**, 819–829 (2004).
18. Rein, A. Murine leukemia viruses: objects and organisms. *Adv. Virol.* **2011**, 403419 (2011).
19. Lim, D. et al. Crystal structure of the moloney murine leukemia virus RNase H domain. *J. Virol.* **80**, 8379–8389 (2006).
20. Gao, G., Orlova, M., Georgiadis, M. M., Hendrickson, W. A. & Goff, S. P. Conferring RNA polymerase activity to a DNA polymerase: a single residue in reverse transcriptase controls substrate selection. *Proc. Natl Acad. Sci. USA* **94**, 407–411 (1997).
21. Boyer, P. L., Sarafianos, S. G., Arnold, E. & Hughes, S. H. Analysis of mutations at positions 115 and 116 in the dNTP binding site of HIV-1 reverse transcriptase. *Proc. Natl Acad. Sci. USA* **97**, 3056–3061 (2000).
22. Katano, Y. et al. Generation of thermostable Moloney murine leukemia virus reverse transcriptase variants using site saturation mutagenesis library and cell-free protein expression system. *Biosci. Biotechnol. Biochem.* **81**, 2339–2345 (2017).
23. Blain, S. W. & Goff, S. P. Effects on DNA synthesis and translocation caused by mutations in the RNase H domain of Moloney murine leukemia virus reverse transcriptase. *J. Virol.* **69**, 4440–4452 (1995).
24. Herschhorn, A. & Hizi, A. Retroviral reverse transcriptases. *Cell. Mol. Life Sci.* **67**, 2717–2747 (2010).
25. Katz, R. A. & Skalka, A. M. The retroviral enzymes. *Annu. Rev. Biochem.* **163**, 133–173 (1994).
26. Mougel, M., Houzet, L. & Darlix, J. L. When is it time for reverse transcription to start and go? *Retrovirology* **6**, 24 (2009).
27. Cannon, K., Qin, L., Schumann, G. & Boeke, J. D. Moloney murine leukemia virus protease expressed in bacteria is enzymatically active. *Arch. Virol* **143**, 381–388 (1998).
28. Walton, R. T. et al. Unconstrained genome targeting with near-PAMless engineered CRISPR–Cas9 variants. *Science* **368**, 290–296 (2020).
29. Jin, S. et al. Genome-wide specificity of prime editors in plants. *Nat. Biotechnol.* **39**, 1292–1299 (2021).
30. Zong, Y. et al. Efficient C-to-T base editing in plants using a fusion of nCas9 and human APOBEC3A. *Nat. Biotechnol.* **36**, 950–953 (2018).
31. 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).
32. Powles, S. B. & Yu, Q. Evolution in action: plants resistant to herbicides. *Annu. Rev. Plant Biol.* **61**, 317–347 (2010).
33. Chen, L. et al. Trp548Met mutation of acetolactate synthase in rice confers resistance to a broad spectrum of ALS-inhibiting herbicides. *Crop J.* **9**, 750–758 (2021).
34. Zheng, C. et al. A flexible split prime editor using truncated reverse transcriptase improves dual-AAV delivery in mouse liver. *Mol. Ther.* https://doi.org/10.1016/j.ymthe.2022.01.005 (2022).
35. Song, M. et al. Generation of a more efficient prime editor 2 by addition of the Rad51 DNA-binding domain. *Nat. Commun.* **12**, 5617 (2021).
36. Chen, P. J. et al. Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell* **184**, 5635–5652.e5629 (2021).
37. Anzalone, A. V. et al. Programmable deletion, replacement, integration and inversion of large DNA sequences with twin prime editing. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-021-01133-w (2021).

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

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2022, corrected publication 2022

Methods

Plasmid construction. The plasmids of PPE-F155Y, PPE-F155V, PPE-F156Y, PPE-N200C and PPE-D524N were mutated by mismatch PCR; PPE-ΔRNase H and PPE-ΔRNase H-Δconnection were amplified to the desired domain of M-MLV RT, and then the resultant fragment was cloned into the PPE vector backbone. To construct vectors of PPE-NC-v1, PPE-NC-v2, PPE-PR-v1, PPE-PR-v2, PPE-IN-v1 and PPE-IN-v2, NC, PR and IN proteins were codon optimized for cereal plants and synthesized commercially (GENEWIZ). and the fusion protein sequences were cloned into the PPE vector backbone. To construct vectors of ePPE and ePPE–SpG, the fused XTEN-NC-32aa–M-MLV RT-ΔRNase H sequences were cloned into the PPE[9] and PPE–SpG[14] backbone, respectively. To construct the PABE8 vector, ABE8 was codon optimized for cereal plants and synthesized commercially (GENEWIZ), and the protein sequences were cloned into the vector PABE7 (ref. [38]) backbone, yielding the PABE8 plasmid (Supplementary Sequences). To construct the binary vector pH-ePPE for *Agrobacterium*-mediated rice transformation, ePPE expression cassettes were cloned into the pH-PPE-v2 (ref. [14]) backbones using a ClonExpressII One Step Cloning Kit (Vazyme). The pegRNA and epegRNA expression vectors were constructed as reported previously[9,16]. PCR was performed using TransStart FastPfu DNA Polymerase (TransGen Biotech).

Protoplast transfection. A *Japonica* rice variety Zhonghua11 and winter wheat variety Kenong199 were used to isolate protoplasts. The isolation of protoplasts was as described previously[39,40]. Plasmids used for protoplast transformation were extracted using the Wizard Plus Midipreps DNA Purification System (Promega). Plasmids, 5 μg, were introduced by PEG-mediated transfection. The mean transformation efficiency was 30–50% by flow cytometry. Transfected protoplasts were incubated at 26 °C for 48 h. After incubation, protoplasts were collected for DNA extraction[9,39].

Flow cytometry analysis. We used FACSAria III (BD Biosciences) for flow cytometry analysis. Rice protoplasts transfected with pegRNA expression plasmids, fluorophore reporter expression plasmids and prime editor expression plasmids were prepared for analysis. All samples were sorted for GFP-positive cells. FACSDiva v.6.1.3 software was used for flow cytometry result analysis. Gating of all samples can be found in the Supplementary Data.

DNA extraction. The gDNA of protoplasts and leaves of each plant was extracted with DNA Quick Plant System (Tiangen Biotech). The extracted gDNA was quantified with NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific).

Amplicon deep sequencing and data analysis. Specific primers with a barcode at the 5′-end were designed to amplify the targeted sequence. Amplicons were purified with the EasyPure PCR Purification Kit (TransGen Biotech) and quantified with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific). Equal amounts of PCR product were pooled and sequenced commercially (Novogene) using the NovaSeq platform. For all prime editing yield quantification, prime editing efficiency was calculated as: percentage (no. of reads with the desired edit without byproducts)/(no. of total reads). The percentages of byproducts during installation of point mutations were calculated as: percentage (no. of reads with imprecise or undesired edits)/(no. of total reads). The percentages of byproducts during installation of deletions or insertions were calculated as: percentage (no. of indel-containing reads except for desired indel mutation)/(no. of total reads). The editing efficiencies induced by base editors were calculated as described previously[30]. Amplicon sequencing was repeated three times for each target site using gDNA extracted from three independent protoplast samples. The primers are listed in Supplementary Table 7.

***Agrobacterium*-mediated transformation of rice callus cells.** Binary plasmid pH-ePPE or pH-PPE-containing pegRNA and the prime editor expression cassette were introduced into *Agrobacterium tumefaciens* strain AGL1 by electroporation (400 ng per transformation). *Agrobacterium*-mediated transformation of callus cells of Zhonghua11 was conducted as reported[38,40]. Hygromycin (50 μg ml^{-1}) was used to select transgenic plants.

Prediction of pegRNA spacer-like off-target edits. The pegRNA spacer-like off-target sites were predicted with an offline version of Cas-OFFinder[41]. The high-quality Zhonghua11 genome was used as a reference genome[42]. The maximum mismatch was set at three.

Mutant identification by PCR-RE assays and Sanger sequencing. PCR-RE digestion assays and Sanger sequencing were used to identify rice mutants with desired conversions in target regions, as described previously[39]. The plants regenerated from rice callus were examined individually. At least two leaves of each plant were used to extract gDNA. The National Center for Biotechnology Information's (NCBI's) primer blast was used to design specific primers (https://www.ncbi.nlm.nih.gov/tools/primer-blast). Target sequences were amplified with 2× Rapid Taq Master Mix (Vazyme Biotech).

Herbicide resistance test. The *OsALS-T6* with Trp548Met heterozygous mutants and the WT were transferred to a plate containing rooting medium with 1.10 μl l^{-1} of imazapic, 0.09 mg l^{-1} of nicosulfuron or both imazapic and nicosulfuron, and cultured in a growth chamber (23 °C, 16-h light:8-h dark). The pictures were taken 10 d after treatment.

Statistical analysis. GraphPad Prism 8 software was used to analyze the data. All numerical values are presented as mean ± s.e.m. Differences between control and treatments were tested using two-tailed Student's *t*-tests.

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

Data availability

All data supporting the findings of the present study are available in the article, extended data and supplementary figures and tables, or are available from the corresponding author on request. The deep sequencing data have been deposited in an NCBI BioProject database (accession no. PRJNA802997). The Zhonghua11 genome is available at NCBI BioProject database (accession no. PRJNA602608). Plasmids encoding ePPE, ePPE–SpG and pH-ePPE are available from Addgene (plasmids 183095, 183096, 183097). Source data are provided with this paper.

References

38. Li, C. et al. Expanded base editing in rice and wheat using a Cas9–adenosine deaminase fusion. *Genome Biol.* **19**, 59 (2018).
39. Shan, Q. et al. Targeted genome modification of crop plants using a CRISPR–Cas system. *Nat. Biotechnol.* **31**, 686–688 (2013).
40. Shan, Q. et al. Rapid and efficient gene modification in rice and *Brachypodium* using TALENs. *Mol. Plant* **6**, 1365–1368 (2013).
41. 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).
42. Jin, S. et al. Rationally designed APOBEC3B cytosine base editors with improved specificity. *Mol. Cell* **79**, 728–740 (2020).

Acknowledgements

This work was supported by grants from the National Natural Science Foundation of China (no. 31788103) and the Strategic Priority Research Program of the Chinese Academy of Sciences (nos. XDA24020102 to C.G, XDB27030201 to X.C. and XDA24020310 to Y.W.).

Author contributions

Y.Z., Y.L., C.X., X.H. and C.G. designed the project. Y.Z., Y.L., C.X., B.L., X.L., J.L. and G.L. performed the experiments. Y.W. prepared the figures. Y.Z., Y.L., C.X., X.C. and C.G. wrote the manuscript. C.G. supervised the project.

Competing interests

The authors have submitted a patent application based on the results reported in this paper.

Additional information

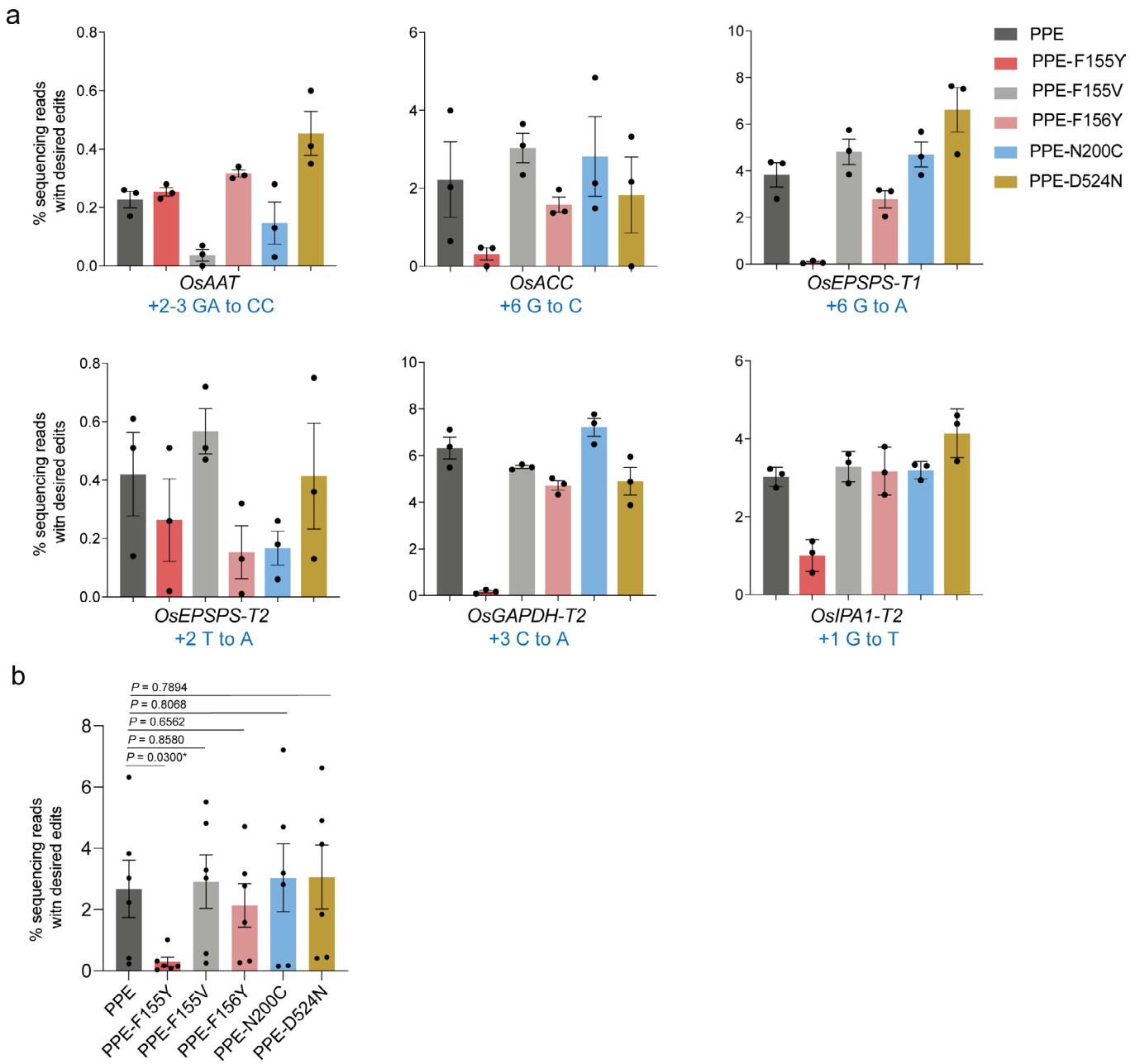
Extended data Extended data are available for this paper at https://doi.org/10.1038/s41587-022-01254-w.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41587-022-01254-w.

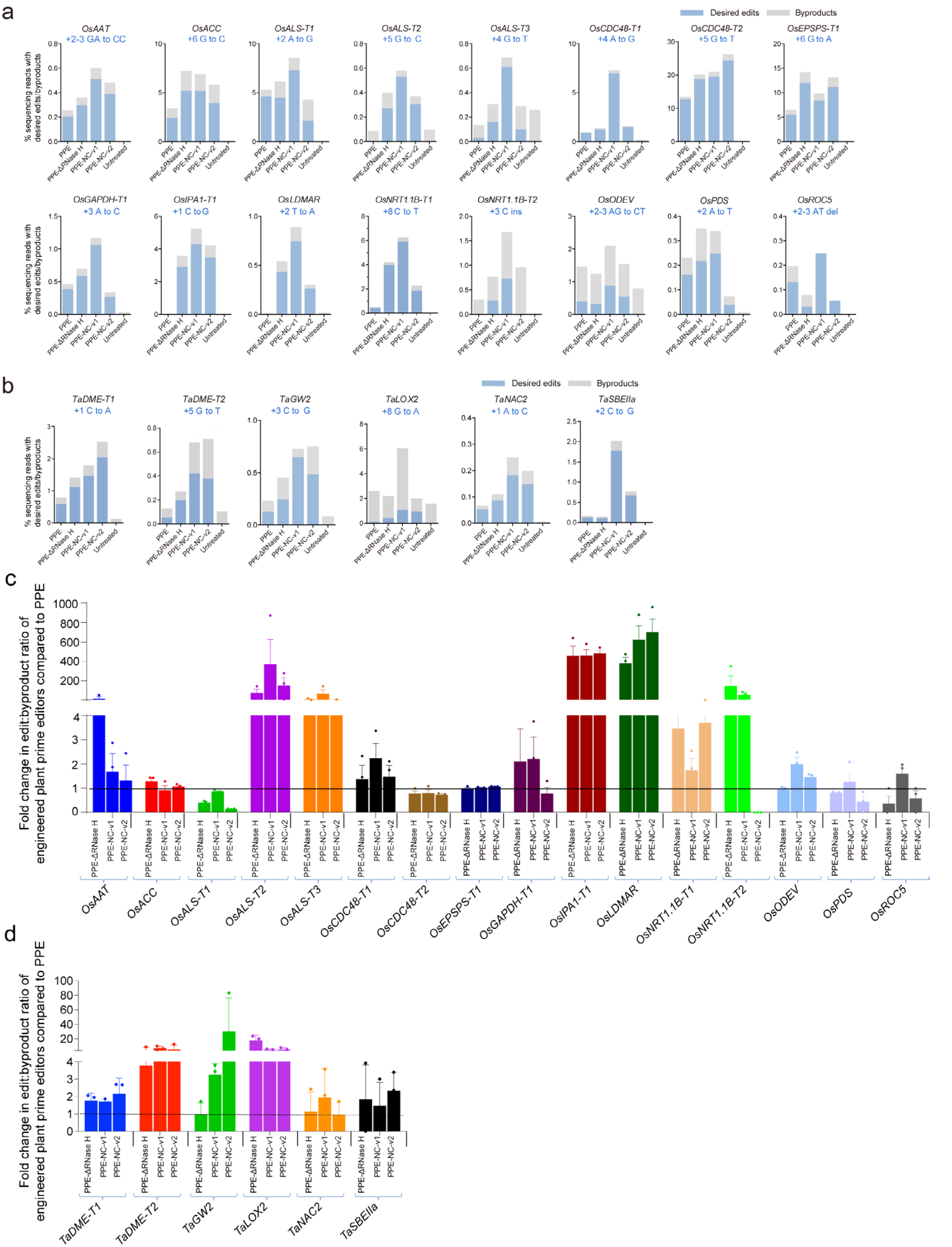
Correspondence and requests for materials should be addressed to Xiaofeng Cao or Caixia Gao.

Peer review information *Nature Biotechnology* thanks Hyongbum Henry Kim and 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.

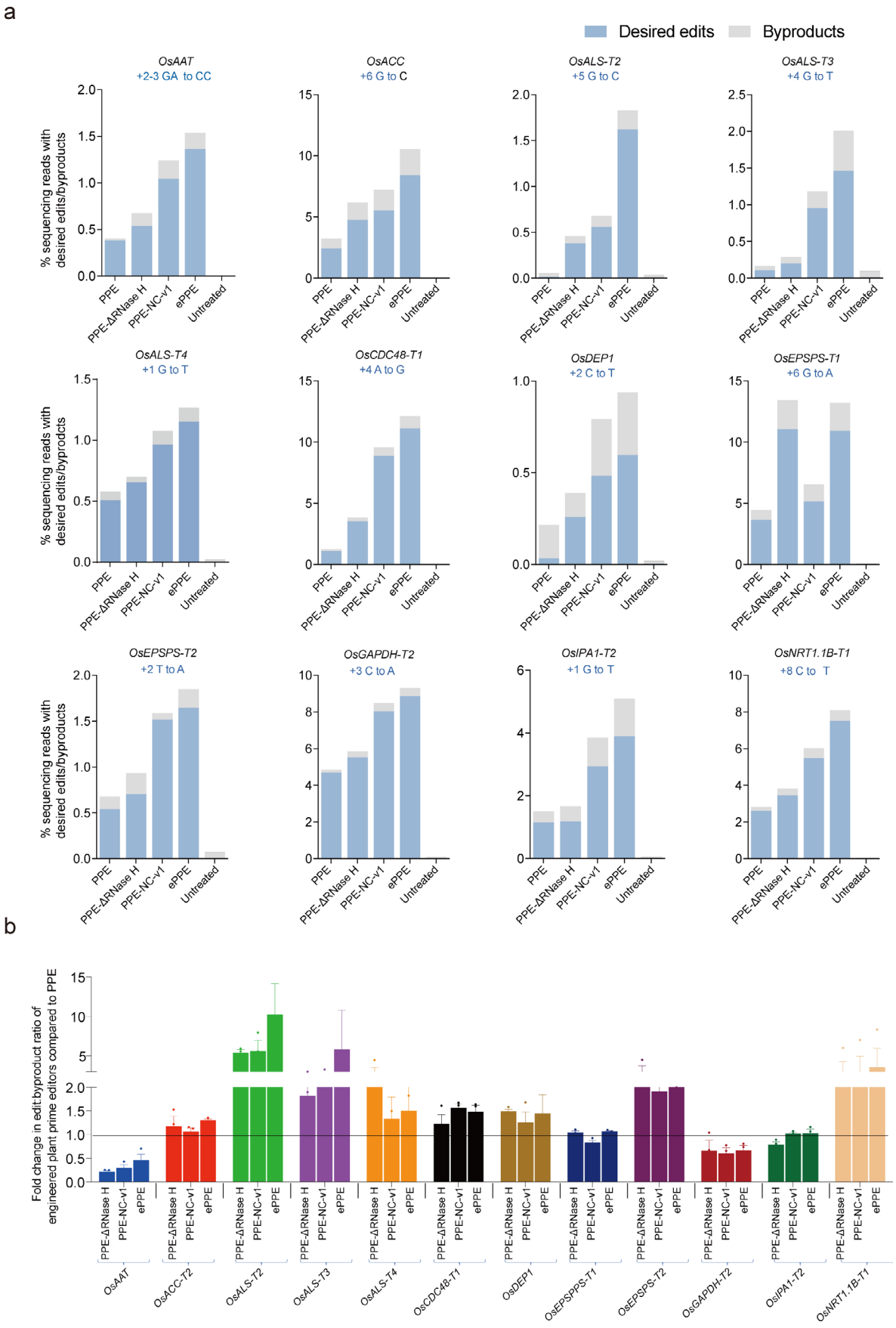


Extended Data Fig. 1 | Prime editing induced by PPE, PPE-F155Y, PPE-F155V, PPE-F156Y, PPE-N200C and PPE-D524N. (**a**) Frequencies of prime editing induced by PPE, PPE-F155Y, PPE-F155V, PPE-F156Y, PPE-N200C, PPE-D524N at six rice target sites. (**b**) The average editing frequencies induced by PPE, PPE-F155Y, PPE-F155V, PPE-F156Y, PPE-N200C and PPE-D524N across six targets. Frequencies (mean ± s.e.m.) were calculated using the data in **a**. *P* values were obtained using two-tailed Student's *t*-tests. $^{*}P < 0.05$.



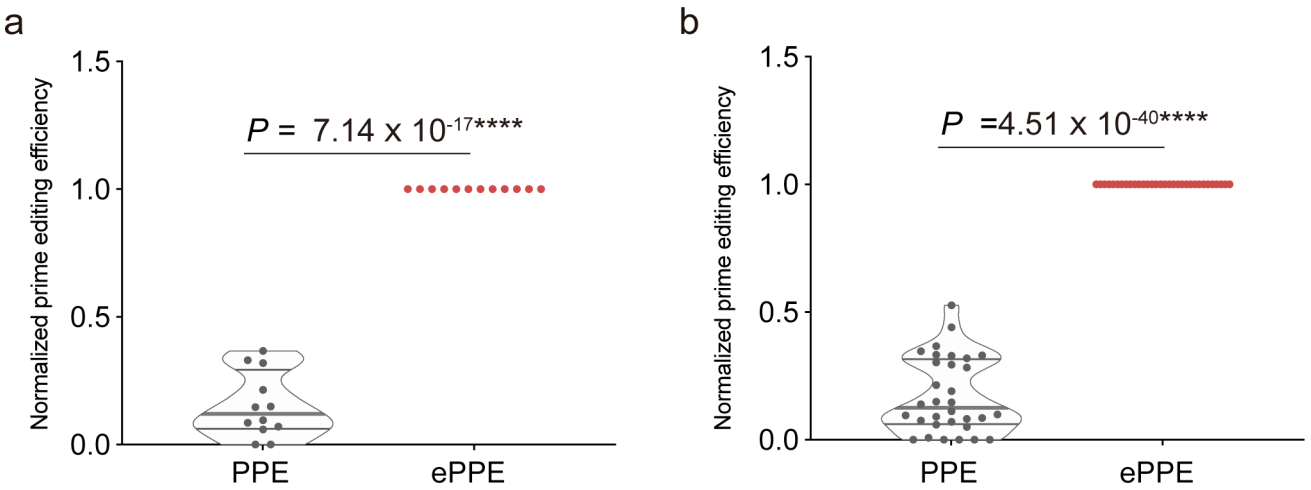
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Product purity for PPE, PPE-ΔRNase H, PPE-NC-v1 and PPE-NC-v2. Frequencies of prime editing and undesired byproducts induced by PPE, PPE-ΔRNase H, PPE-NC-v1 and PPE-NC-v2 at 16 endogenous sites in rice protoplasts (**a**) and six target sites in wheat protoplasts (**b**). Fold-change in the observed prime editing edit:byproduct ratio for rice target sites (**c**), and for wheat targets (**d**). Values were calculated from the data presented in Fig. 1f and 1g respectively. Data and error bars reflect the mean and standard deviation of three independent biological replicates. Frequencies (means ± s.e.m.) were calculated from three independent experiments ($n = 3$).

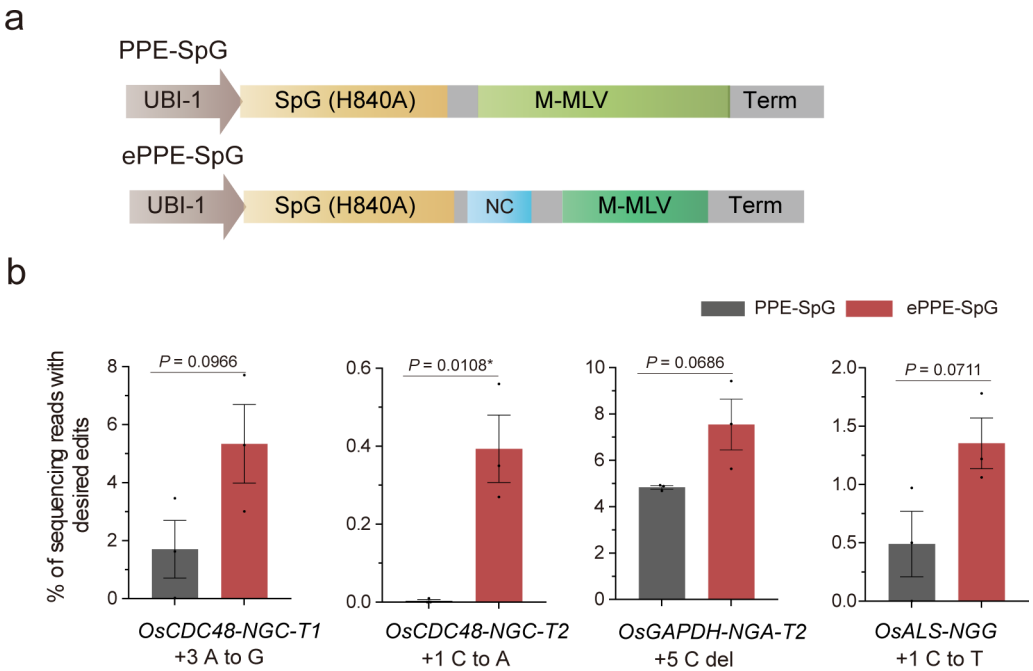


Extended Data Fig. 3 | See next page for caption.

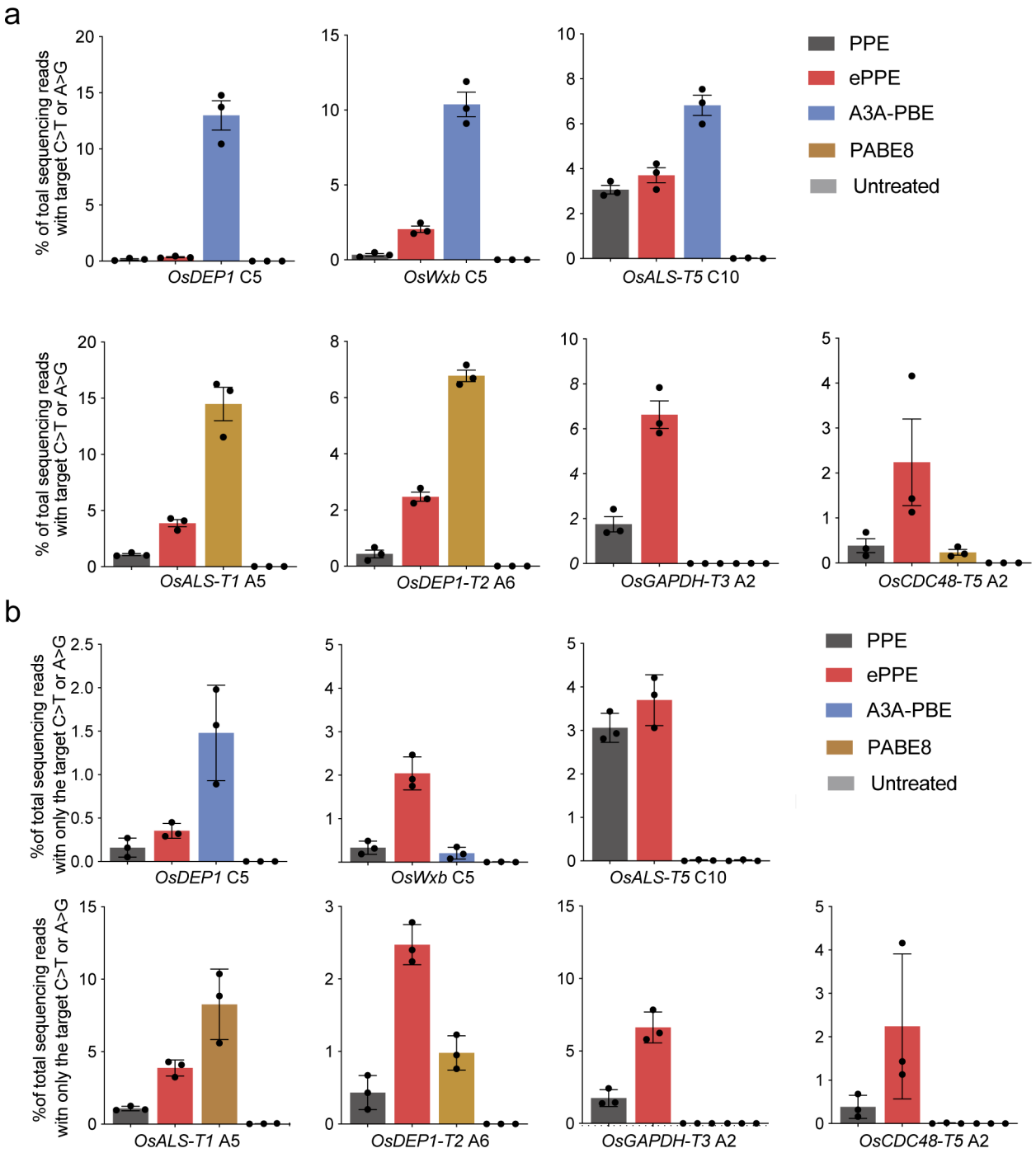
Extended Data Fig. 3 | Product purity for PPE, PPE-ΔRNase H, PPE-NC-v1, and ePPE. (a) Product purity in prime editing by PPE, PPE-ΔRNase H, PPE-NC-v1, and ePPE at 12 endogenous sites in rice protoplasts. Frequencies (means ± s.e.m.) were calculated from three independent experiments ($n = 3$). (**b**) Fold-change in the observed prime editing edit:byproduct ratio for 12 rice target sites. Values were calculated from the data presented in Fig. 2c. Data and error bars reflect the mean and standard deviation of three independent biological replicates. Frequencies (means ± s.e.m.) were calculated from three independent experiments ($n = 3$).



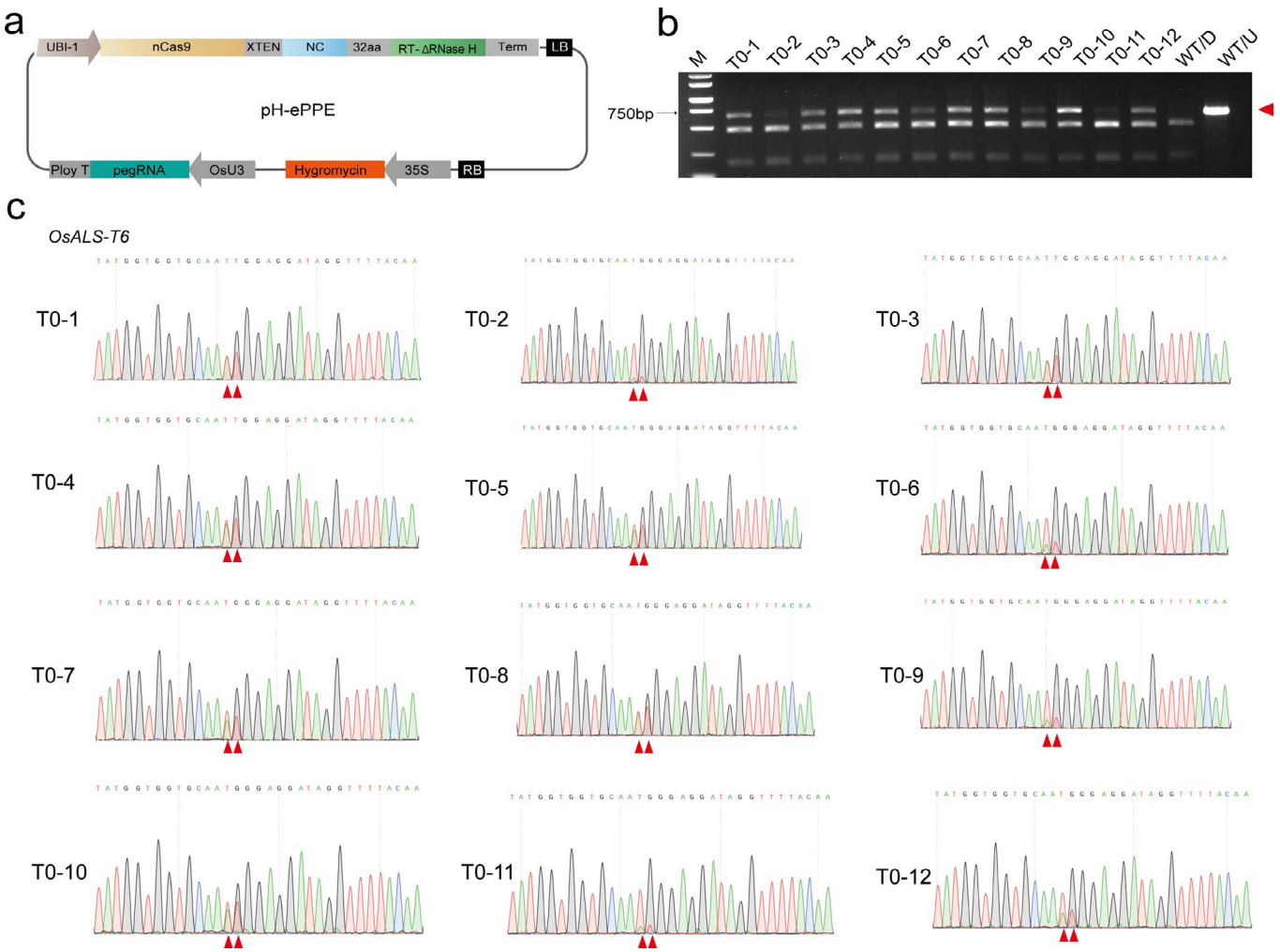
Extended Data Fig. 4 | Overall editing frequencies induced by PPE and ePPE. The overall editing frequencies induced by PPE and ePPE at 12 target sites in Fig. 2e (**a**) and at 32 target sites in Fig. 2c,e,f (**b**). The average of editing frequencies using ePPE for each target were normalized to 1, and the frequencies using PPE for each target were adjusted accordingly ($n = 3$ independent experiments). *P* values were obtained using the two-tailed Student's *t*-test. ****$P < 0.0001$.



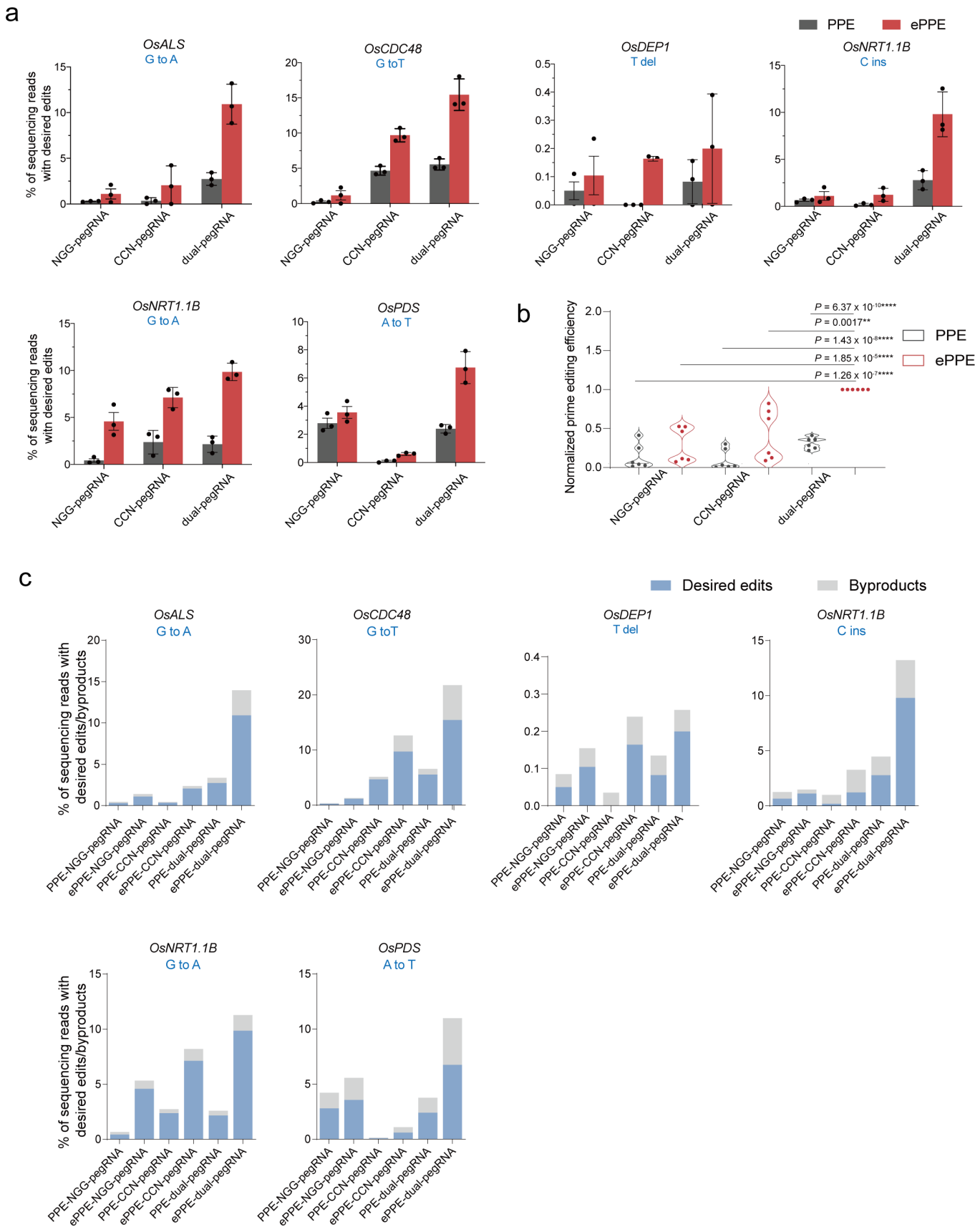
Extended Data Fig. 5 | Prime editing induced by PPE-SpG and ePPE-SpG in rice protoplasts. (**a**) Schematic representation of PPE-SpG and ePPE-SpG. (**b**) Frequencies of prime editing induced by PPE-SpG and ePPE-SpG at four target sites. Frequencies (mean ± s.e.m.) were calculated from three independent experiments ($n = 3$). *P* values were obtained using two-tailed Student's *t*-tests. $^{*}P < 0.05$.



Extended Data Fig. 6 | Comparison of prime editing and base editing. The total editing efficiency (**a**), and the precise C>T or A>G editing efficiency without bystander edits (**b**) at the seven targets induced by prime editors and base editors. Frequencies (mean ± s.e.m.) were calculated from three independent experiments ($n = 3$).

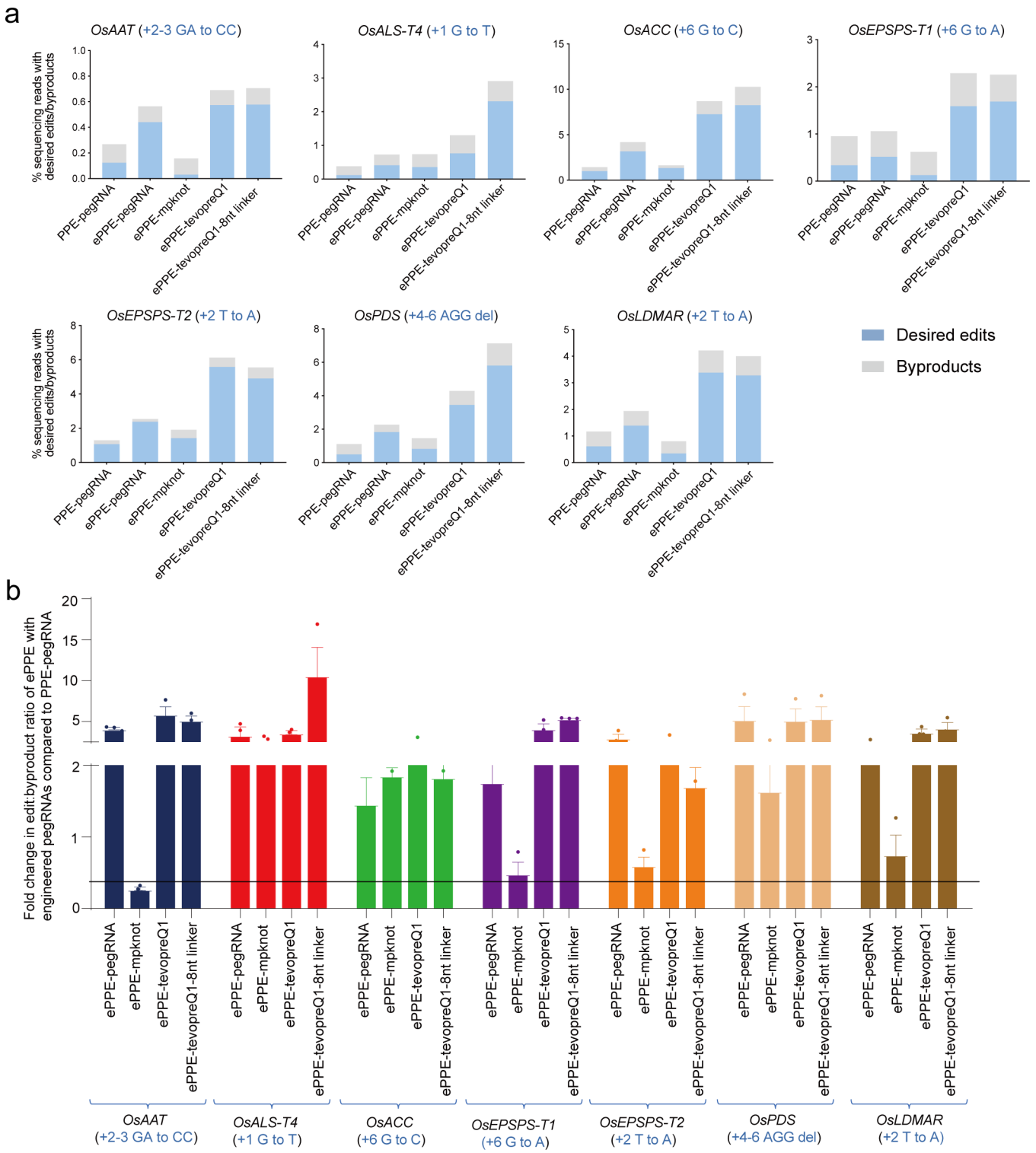


Extended Data Fig. 7 | Genotypes of prime-edited OsALS-T6 rice mutants. (a) Schematic representation of pH-ePPE. **(b)** The results of PCR-RE assays analyzing 12 representative *OsALS-T6* plantlets (T0-1 to T0-12). restriction enzyme. 'M' represents marker. 'WT/D' represents digested PCR products of wild-type. 'WT/U' represents undigested PCR products of wild-type (Untreated). Arrowheads indicate the bands anticipated from BsrDI restriction enzyme. **(c)** Sanger sequencing chromatograms of representative seven prime-edited heterozygous and five chimera mutants. Red arrows represent the desired edits. One biological experiment was performed.

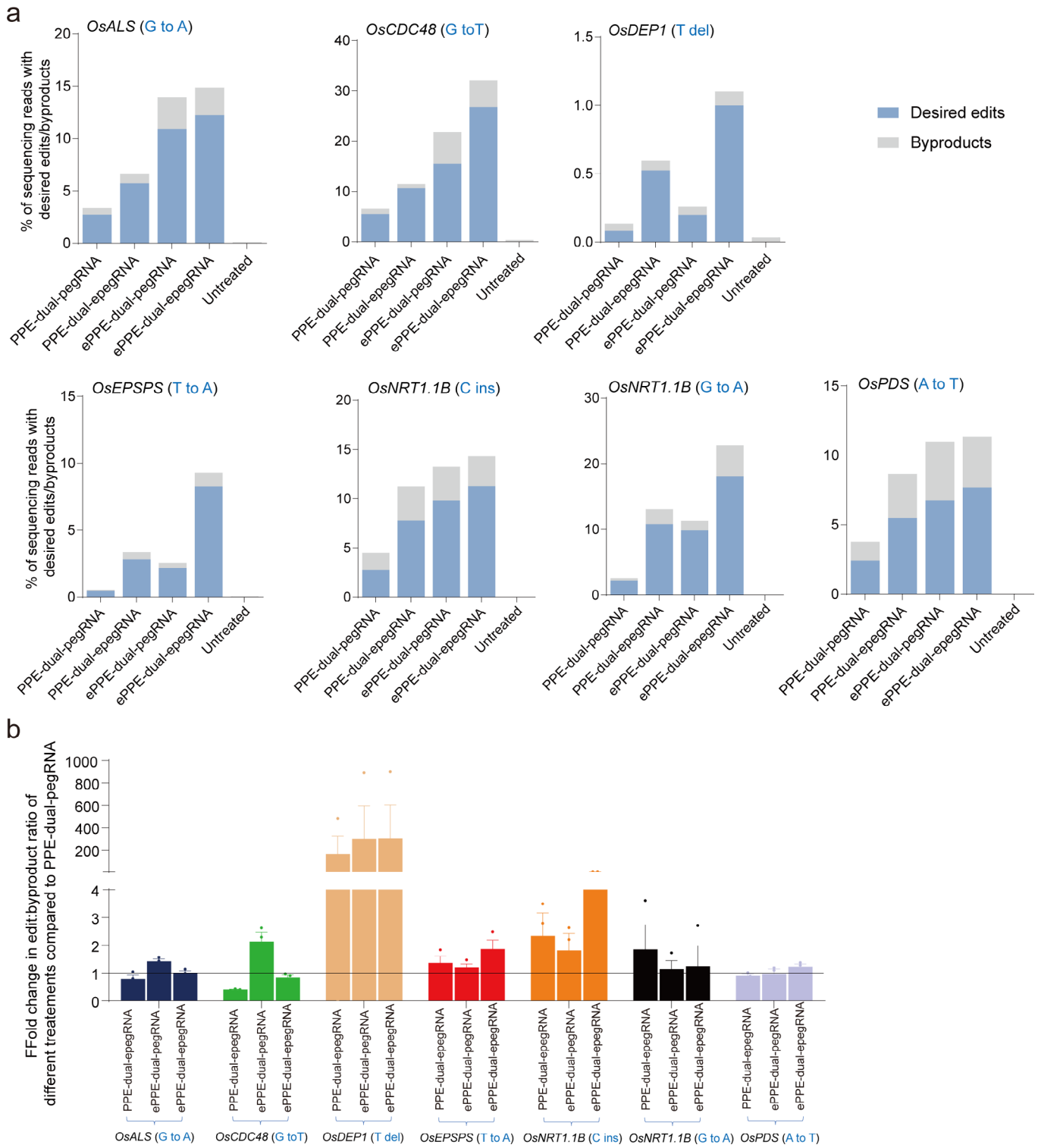


Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Comparison of the prime editing efficiency induced by PPE or ePPE with NGG-pegRNA, CCN-pegRNA and dual-pegRNA strategies. (a) Frequencies of prime editing induced by PPE and ePPE at six rice target sites using NGG-pegRNA, CCN-pegRNA and dual-pegRNA strategies. The edits were referred to the base on the DNA forward strand. (**b**) Overall editing frequencies induced by PPE and ePPE containing NGG-pegRNA, CCN-pegRNA and dual-pegRNA. The average editing frequencies using ePPE-dual-pegRNA for each target were normalized to 1, and the frequencies using others for each target were adjusted accordingly. **(c)** Product purity in prime editing by PPE and ePPE using NGG-pegRNA, CCN-pegRNA and dual-pegRNA strategies. Data and error bars reflect the mean and standard deviation of three independent biological replicates. Frequencies (means ± s.e.m.) were calculated from three independent experiments ($n = 3$). *P* values were obtained using two-tailed Student's *t*-tests. $^{**}P < 0.01$, $^{****}P < 0.0001$.



Extended Data Fig. 9 | Product purity induced by different PPEs and different engineered pegRNA forms. (a) Product purity in prime editing by different PPEs and different engineered pegRNA forms at seven endogenous sites in rice protoplasts. Frequencies (means ± s.e.m.) were calculated from three independent experiments ($n = 3$). (**b**) Fold-change in the observed prime editing edit:byproduct ratio for seven rice target sites. Values were calculated from the data presented in Fig. 5a. Data and error bars reflect the mean and standard deviation of three independent biological replicates. Frequencies (means ± s.e.m.) were calculated from three independent experiments ($n = 3$).



Extended Data Fig. 10 | Product purity induced by different PPEs and different pegRNA forms. (a) Product purity in prime editing by different PPEs and different pegRNA forms at seven endogenous sites in rice protoplasts. Frequencies (means ± s.e.m.) were calculated from three independent experiments ($n = 3$). (**b**) Fold-change in the observed prime editing edit:byproduct ratio for seven rice target sites. Values were calculated from the data presented in Fig. 5d. Data and error bars reflect the mean and standard deviation of three independent biological replicates. Frequencies (means ± s.e.m.) were calculated from three independent experiments ($n = 3$).

nature research

Corresponding author(s): Caixia Gao

Last updated by author(s): Feb 3, 2022

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	Illumina NovaSeq platform was used to collect the amplicon deep sequencing data. BD FACSAriaIII was used to do flow cytometry.
Data analysis	Amplicon sequencing data of prime-editing processivity was analyzed using the published code as previously described in reference 9. The custom Perl script to analyze types of mutational reads and amino acid substitutions will be made available upon request. Graphpad prism 8 was used to analyze the data. FACSDiva Version 6.1.3 software was used for flow cytometry result analysis. Cas-OFFinder was used to predict pegRNA spacer-like 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

The authors declare that all data supporting the findings of this study are available in the article, extended data and its Supplementary Information files or are available from the corresponding author on request. The deep sequencing data have been deposited in a National Center for Biotechnology Information BioProject database (accession code PRJNA802997). The Zhonghua11 genome can be available at NCBI BioProject database (accession code PRJNA602608).

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	The experiments of protoplasts were performed with three biological repeats. About 500,000 protoplasts were used for each transfection. The number of protoplasts in each transfection was measured by thrombocytometry. The experiment in rice resistant calli and regenerated plants were performed once. All the regenerated seedlings were sampled, the number of mutants were confirmed by Sanger sequencing.
Data exclusions	No data exclusion.
Replication	All attempts for replication were successful. For the experiments in rice and wheat protoplasts, a minimum of three independent experiments were included.
Randomization	Rice and wheat protoplasts were isolated and randomly separated to each transformation.
Blinding	Not applicable. As samples were processed identically through standard and in some cases automated procedures (DNA sequencing, transfection, DNA isolation) that should not bias outcomes.

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

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK293T, HCT116, HELA, U2OS.
Authentication	Cells were authenticated by the supplier using STR analysis.
Mycoplasma contamination	All cell lines tested negative for mycoplasma.
Commonly misidentified lines (See ICLAC register)	None Used.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Rice protoplasts were isolated from the stem of rice seedlings, transfected as described in the Mehtods and incubated in 2 ml WI solution for 2 days.
Instrument	BD FACSAriaIII
Software	FACSDiva Version 6.1.3 software was used for analysis.
Cell population abundance	The abundance of cells for flow cytometry analysis was 5,000 for each sample.
Gating strategy	Negative control (untreated) and fluorophore-positive cells were used to establish gates for each cell type. Gates were drawn to collect cells expressing either fluorophore. See the provided examples for gates used in Supplementary Data.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.