nature biotechnology

Efficient C-to-T base editing in plants using a fusion of nCas9 and human APOBEC3A

Yuan Zong[1,2,4], Qianna Song[1,2,4], Chao Li[1,2], Shuai Jin[1,2], Dingbo Zhang[1,2], Yanpeng Wang[1], Jin-Long Qiu[3] & Caixia Gao[1,2]

Base editors (BEs) have been used to create C-to-T substitutions in various organisms. However, editing with rat APOBEC1-based BE3 is limited to a 5-nt sequence editing window and is inefficient in GC contexts. Here, we show that a base editor fusion protein composed of Cas9 nickase and human APOBEC3A (A3A-PBE) converts cytidine to thymidine efficiently in wheat, rice and potato with a 17-nucleotide editing window at all examined sites, independent of sequence context.

Many single nucleotide variants have been associated with important agronomic traits and used in crop improvement[1,2]. Genetic engineering of single nucleotide polymorphisms in plants represents a great advance in molecular breeding[3,4]. Base editors generate programmed base changes without requiring DNA double-strand breaks (DSBs) or exogenous DNA[5,6]. The most commonly used base editor, BE3, consists of the rat cytidine deaminase APOBEC1 fused with a Cas9 nickase (nCas9 (D10A)) and the uracil glycosylase inhibitor (UGI). It converts targeted cytidine to thymidine in DNA[7], and its protospacer-adjacent motif requirement, efficiency and specificity have been increased[8–11]. Recently, a human APOBEC3A cytidine deaminase –based base editor has been reported to improve specificity[12] and efficiency in mammalian cells[13]. However, the scope of base editing remains restricted by a narrow deamination window[7] (5 nt) and much higher activity in a TC than a GC context[7]. Here we show that a base editor, A3A-PBE, consisting of the human A3A cytidine deaminase fused with a Cas9 nickase (**Fig. 1a**), is very effective in generating targeted C-to-T mutations in wheat, rice and potato.

To improve our earlier base editor, nCas9-PBE[14] (referred to hereafter as PBE), we replaced rat APOBEC1 with human APOBEC3A optimized for cereal plant codons to create A3A-PBE (**Fig. 1b** and **Supplementary Sequences**). We also fused another UGI and the Mu protein coding region to A3A-PBE, generating A3A-Gam (**Fig. 1b** and **Supplementary Sequences**) to increase editing efficiency and product purity, as in mammalian cells[9]. We first used a reporter system in plant protoplasts to compare these BEs, by examining conversion of BFP to GFP when C_4 of the BFP single guide RNA (sgRNA) targeting sequences was converted to T_4 (ref. 14). PBE, A3A-PBE and A3A-Gam were separately cotransfected with pUbi-BFPm and pOsU3-BFP-sgRNA into rice protoplasts by polyethylene glycol–mediated transformation[14]. Flow cytometry showed that A3A-PBE yielded the highest percentage of GFP-expressing cells (24.5%), about 12-fold higher than PBE, with A3A-Gam of intermediate efficiency (**Fig. 1c**,**d**).

To examine base editing of endogenous genes, we designed four sgRNAs for three wheat genes and six sgRNAs for six rice genes (**Fig. 1e**,**f** and **Supplementary Table 1**). A control employed Cas9 to produce indels. C-to-T base editing of the target genes in protoplasts was assessed by amplicon deep sequencing with 19,000–140,000 reads per locus. As in the reporter system, A3A-PBE had the highest base editing efficiency, averaging 13.1%, about 13-fold and fivefold higher, respectively, than those of PBE (average 1.0%) and A3A-Gam (average 2.8%) (**Fig. 1e**,**f**). By analyzing editing efficiencies at every protospacer position across all ten target sites, we found that the deamination window for A3A-PBE spanned 17 nt, from protospacer positions 1 to 17, compared with 3 to 9 for PBE (**Fig. 1e**,**f**). Also, none of the constructs showed any sign of undesired editing upon analysis of the amplicon deep sequencing data at any of the on-target loci (<0.1%) (**Supplementary Figs. 1** and **2**), and all yielded much lower frequencies of undesired on-target indels (<0.1%) than Cas9 (3.7–21.6%) (**Supplementary Fig. 3**).

We also tested A3A-PBE in tetraploid potato (*Solanum tuberosum*), whose tetrasomic inheritance makes traditional crossbreeding difficult[15]. We used the 35S promoter to express A3A-PBE and PBE, and the AtU6 promoter to express sgRNAs (**Supplementary Fig. 4a**): four to target *StALS* (StALS-T1 to StALS-T4) and six to target *StGBSS* (StGBSS-T1 to StGBSS-T6) (**Fig. 1g** and **Supplementary Fig. 4b**). The appropriate sgRNAs were transformed along with the A3A-PBE or PBE construct into potato protoplasts. C-to-T conversion by A3A-PBE at the ten sites averaged about 11-fold higher than by PBE; the deamination window again spanned positions 1 to 17 in the protospacer (**Fig. 1g**), and the frequency of indels was very low (**Supplementary Fig. 4c**).

We next tested A3A-PBE at target sites in GC contexts, using seven sgRNAs targeting three wheat genes and three rice genes (**Fig. 2a**), and compared its editing activities with PBE, which is very inefficient at targeting cytidines immediately downstream of guanidines[7]. We found that A3A-PBE displayed editing efficiencies of up to 41.2% at all seven target sites (**Fig. 2a**), whereas PBE produced virtually no C-to-T editing events (<0.2%) (**Fig. 2a**). Thus A3A-PBE, unlike PBE, edits cytidines almost equally well regardless of context (**Fig. 2b**).

[1]State Key Laboratory of Plant Cell and Chromosome Engineering, Center for Genome Editing, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China. [2]University of Chinese Academy of Sciences, Beijing, China. [3]State Key Laboratory of Plant Genomics, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China. [4]These authors contributed equally to this work. Correspondence should be addressed to C.G. (cxgao@genetics.ac.cn).

Received 14 June; accepted 25 August; published online 1 October 2018; doi:10.1038/nbt.4261

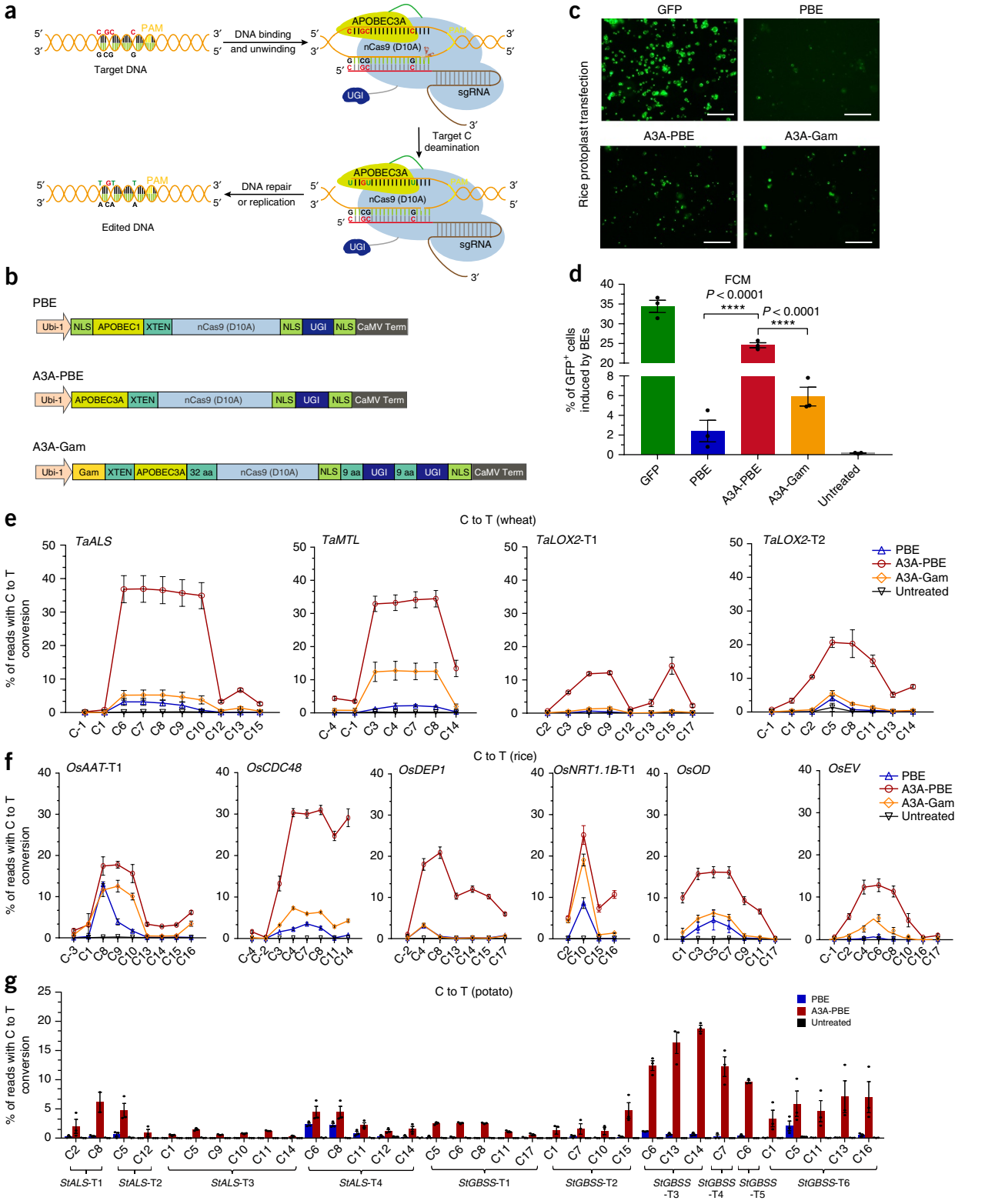

Figure 1 Comparison of C-to-T base editing by A3A-PBE and PBE. (**a**) Scope of cytidine base editing by A3A-PBE. (**b**) Representation of the three cytidine base editors. (**c**) Microscope views of rice BFP to GFP conversion using the three base editors, showing fields of protoplasts transformed with the relevant cytidine base editors together with pUbi-BFPm[14] and pOsU3-BFP-sgRNA[14]. GFP and an untreated protoplast sample were controls. Scale bars, 150 μm. (**d**) Frequencies (in percent) of C-to-T conversion in the target region of the BFP coding sequence measured by flow cytometry ($n = 3$ independent experiments). Values represent means ± s.e.m. $^{****}P < 0.0001$. Statistical differences between control and the treatments were tested using two-tailed Student's *t*-test. (**e**,**f**) Frequencies of single C-to-T conversions by PBE, A3A-PBE and A3A-Gam at four target sites in wheat protoplasts (**e**) and six target sites in rice protoplasts (**f**). (**g**) Frequencies of targeted single C-to-T conversions introduced by PBE and A3A-PBE at ten target sites in potato protoplasts. An untreated protoplast sample served as control. Frequencies (mean ± s.e.m.) were calculated from three independent experiments ($n = 3$).

The wide deamination window and efficiency of A3A-PBE suggested it could be used to mutate multiple sites in gene regulatory regions. We therefore investigated whether A3A-PBE could generate a diversity of mutations when combined with multiple sgRNAs. The *TaVRN1-A1* promoter contains several regulatory sites—namely, a VRN box, CArG box and putative AG hybrid box (**Fig. 2c**)—associated with wheat

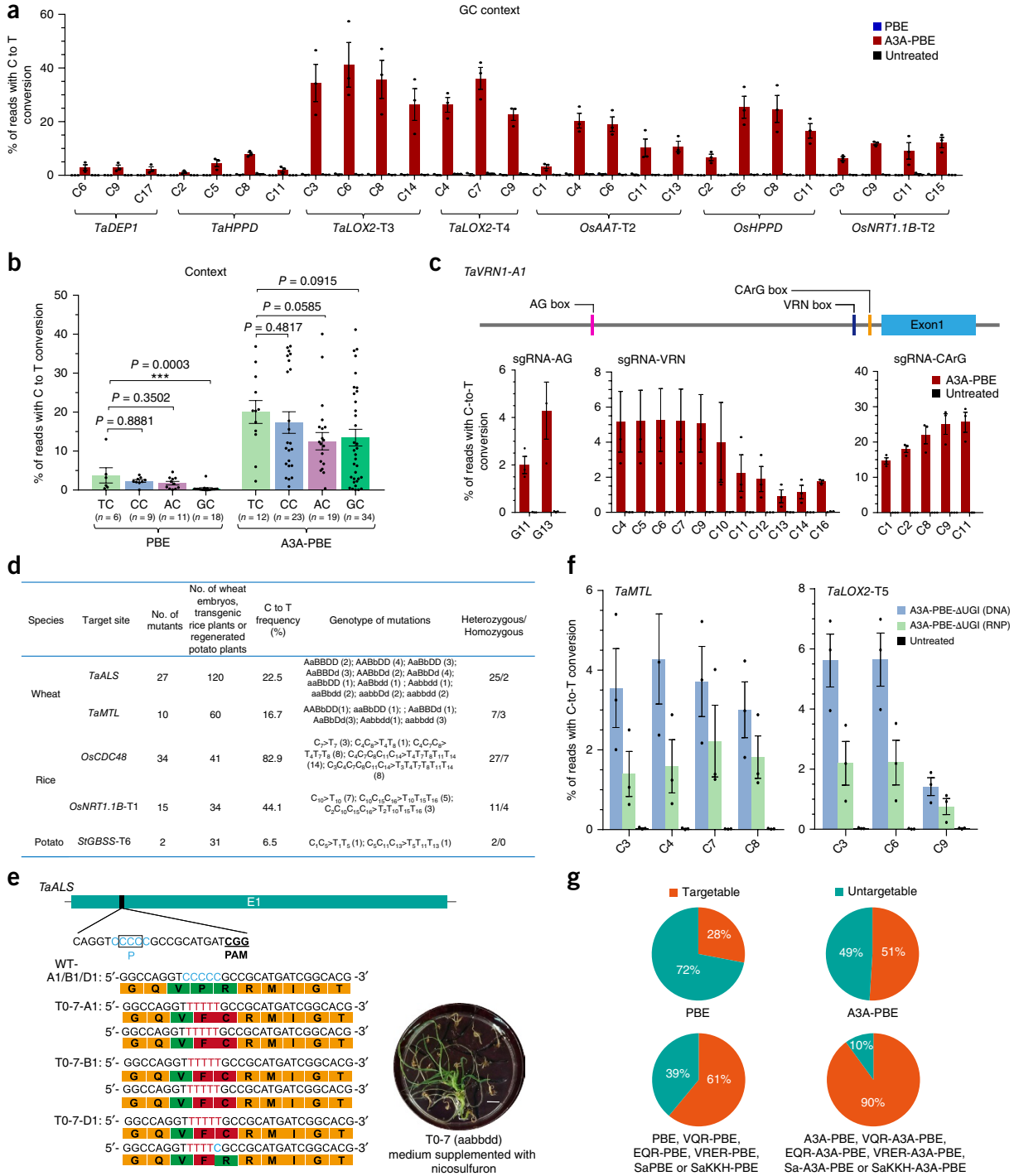

Figure 2 A3A-PBE is widely useful for C-to-T base editing. (**a**) Comparison of C-to-T base editing by A3A-PBE and PBE in GC contexts. (**b**) Effect of sequence context on base editing by PBE (window from 3 to 9) and A3A-PBE (window from 1 to 17). Frequencies (mean ± s.e.m.) were calculated using the data in **Figures 1e,f** and **2a**. (**c**) Frequencies of single C-to-T conversions introduced by A3A-PBE in *cis* elements of the promoter of *TaVRN1-A1*. (**d**) Frequencies of mutations induced by A3A-PBE in T0 wheat, rice and potato plants. (**e**) Amino acid substitutions in *TaALS* confer resistance to herbicide. Amino acid sequence alignment of wild-type (WT) *TaALS* and the T0-7 mutant. Growth of T0-7 in regeneration medium with 0.254 p.p.m. nicosulfuron after 3 weeks. A1/B1/D1 represents three homoeologs of *TaALS*; P represents proline. Scale bar, 1 cm. (**f**) The C-to-T base editing efficiencies of A3A-PBE-ΔUGI (DNA) and A3A-PBE-ΔUGI (RNP). Frequencies (mean ± s.e.m.) were calculated from three independent experiments in **a**, **c** and **f**. (**g**) Cytidines (NGG PAM) and guanidines (CCN PAM) in the rice genome that can potentially be edited by PBE and A3A-PBE. PBE and A3A-PBE together with the various Cas9 variants (VQR, EQR, VRER, SaCas9 and SaKKH) promise to increase the scope of base editing of targeted cytidines and guanidines in the rice genome.

vernalization[16,17]. Three sgRNAs were designed to target these binding sites (**Fig. 2c**). Deep sequencing of amplicons from protoplasts treated with A3A-PBE identified reads containing a variety of mutations in all three *cis* elements (with efficiencies ranging from 1.2% to 27.7%). For example, at the VRN-box site, A3A-PBE edited all the cytidines in positions 4 to 16 of the protospacer, which should disrupt binding of bZIP transcription factors (**Fig. 2c**). These findings indicate that A3A-PBE is useful for manipulating promoters and other regulatory elements.

To regenerate base-edited plants, we targeted the wheat acetolactate synthase gene (*ALS*), the first enzyme in biosynthesis of the branched-chain amino acids. Substitutions at the conserved Pro197 of *Lolium rigidum* ALS confer resistance to the herbicide nicosulfuron[18], and its conserved Pro197 residue corresponds to Pro174 in wheat *TaALS*. We delivered A3A-PBE and pTaU6-ALS-sgRNA constructs into immature wheat embryos by particle bombardment, and identified 27 mutants harboring at least one C-to-T conversion in plants regenerated, without herbicide or antibiotic selection[14], from approximately 120 bombarded embryos (**Fig. 2d** and **Supplementary Fig. 5**). Base editing occurred at positions –7, 6, 7, 8, 9, 10, 12 and 13 of the protospacer (**Fig. 2d** and **Supplementary Fig. 5**). Among the 27 mutants, we identified multiple combinations of amino acid substitutions, including 12 mutants with substitutions in all three subgenomes (**Supplementary Table 2**). Notably, six alleles were simultaneously edited in two of the mutants (T0-7 and T0-9), and the deduced proteins all contained amino acid substitutions (**Fig. 2d**,**e** and **Supplementary Table 2**). We assessed the nicosulfuron resistance of the T0-7 mutant and found it to be resistant (**Fig. 2e**).

MTL encodes a sperm-specific phospholipase[19]. Loss of function of *MTL* triggers haploid induction in maize[19]. We attempted to generate wheat *TaMTL* loss-of-function mutants by substituting Gln95 (CAG) with a stop codon (TAG) by base editing with A3A-PBE. Ten base editing mutants were identified (frequency 16.7%), three being homozygous for all six alleles (**Fig. 2d** and **Supplementary Table 3**). No indels were detected at these two target sites (**Supplementary Tables 2** and **3**). Furthermore, nine transgene-free plants for *TaALS* and six for *TaMTL* were identified by PCR screening with five primer sets specific for A3A-PBE and pTaU6-sgRNA, respectively (**Supplementary Fig. 6** and **Supplementary Tables 2** and **3**).

We also used A3A-PBE to generate base-edited rice and potato plants, targeting rice *OsCDC48* and *OsNRT1.1B-T1* sites via *Agrobacterium*-mediated transformation and targeting potato StGBSS-T6 via polyethylene glycol–mediated protoplast transfection. We obtained various combinations of substitutions with efficiencies of 82.9% (34 of 41) for *OsCDC48* and 44.1% (15 of 34) for *OsNRT1.1B-T1*, including seven homozygous mutants for *OsCDC48* and four for *OsNRT1.1B-T1* (**Fig. 2d**). Two independent heterozygous mutant potato plants at the StGBSS-T6 target site were recovered from 31 regenerated plants (**Fig. 2d**). Once more, none of the transgenic rice and potato plants contained indels or undesired edits at the target sites (**Fig. 2d**), and we detected no mutations in triply mismatched off-target rice regions (**Supplementary Tables 4** and **5**). To the best of our knowledge, this is the first report of editing by cytidine deamination in potato; it paves the way for widespread use of A3A-PBE in dicotyledons.

Finally, we expressed A3A-PBE without UGI (A3A-PBE-ΔUGI) in *Escherichia coli*, purified it (**Supplementary Fig. 7**) and combined it with *in vitro*–transcribed sgRNA in ribonucleoprotein complexes (RNPs) used to target two wheat genes (**Fig. 2f**). Amplicon deep sequencing revealed average frequencies of C-to-T conversion of 1.8% at the two sites, somewhat lower than with the A3A-PBE-ΔUGI vector (average 3.9%) (**Fig. 2f**). It should be possible to improve A3A-PBE-ΔUGI RNPs to create mutant plants free of foreign DNA, which would greatly facilitate the application of base editing to plant breeding and the commercialization of edited plants.

Computational analysis of the rice reference genome (Os-Nipponbare reference IRGSP-1.0) revealed that the 17-nt editing window of A3A-PBE theoretically increases 1.8-fold the number of genomic cytidines and guanidines available for base editing compared with PBE (**Fig. 2g**). When combined with SpCas9, SaCas9 and their variants with NGG, NGA, NGCG, NNGRRT and NNNRRT protospacer-adjacent motifs, A3A deaminase fusions could potentially target 90% of the cytidines and guanidines in the rice genome (**Fig. 2g**).

In summary, A3A-PBE generated cytidine to thymidine conversions efficiently at a broad range of endogenous genomic loci with diverse sequences within a 17-nt deamination window. A3A-PBE thus represents a promising tool for improving crops by genome engineering.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank M. Andersson for technical support on potato protoplast regeneration and Y. Huo for technical support on flow cytometry. This work was supported by grants from the National Key Research and Development Program of China (2016YFD0101804), the National Natural Science Foundation of China (31788103 and 31420103912), the National Transgenic Science and Technology Program (2016ZX08010002, 2018ZX0800102B-001 and 2018ZX0801002B-002) and the Chinese Academy of Sciences (QYZDY-SSW-SMC030 and GJHZ1602).

AUTHOR CONTRIBUTIONS

Y.Z. and C.G. designed the experiments; Y.Z., Q.S. and C.L. performed most of the experiments; D.Z. purified the protein. S.J. and Y.W. analyzed the results; C.G. supervised the project; Y.Z., J.-L.Q. and C.G. wrote the manuscript.

COMPETING INTERESTS

The authors have submitted a patent application (application no. 201810816603.7) based on the results reported in this paper.

1. Zhao, K. *et al. Nat. Commun.* **2**, 467 (2011).
2. Henikoff, S. & Comai, L. *Annu. Rev. Plant Biol.* **54**, 375–401 (2003).
3. Voytas, D.F. & Gao, C. *PLoS Biol.* **12**, e1001877 (2014).
4. Gao, C. *Nat. Rev. Mol. Cell Biol.* **19**, 275–276 (2018).
5. Hess, G.T., Tycko, J., Yao, D. & Bassik, M.C. *Mol. Cell* **68**, 26–43 (2017).
6. Yang, B., Li, X., Lei, L. & Chen, J. *J. Genet. Genomics* **44**, 423–437 (2017).
7. Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A. & Liu, D.R. *Nature* **533**, 420–424 (2016).
8. Kim, Y.B. *et al. Nat. Biotechnol.* **35**, 371–376 (2017).
9. Komor, A.C. *et al. Sci. Adv.* **3**, eaao4774 (2017).
10. Kim, K. *et al. Nat. Biotechnol.* **35**, 435–437 (2017).
11. Rees, H.A. *et al. Nat. Commun.* **8**, 15790 (2017).
12. Gerhke, J.M. *et al. Nat. Biotechnol.* https://doi.org/10.1038/nbt.4199 (2018).
13. St. Martin, A. *et al. Nucleic Acids Res.* https://doi.org/10.1093/nar/gky332 (2018).
14. Zong, Y. *et al. Nat. Biotechnol.* **35**, 438–440 (2017).
15. Obidiegwu, J.E., Flath, K. & Gebhardt, C. *Theor. Appl. Genet.* **127**, 763–780 (2014).
16. Li, C. & Dubcovsky, J. *Plant J.* **55**, 543–554 (2008).
17. Kippes, N. *et al. Proc. Natl. Acad. Sci. USA* **112**, E5401–E5410 (2015).
18. Powles, S.B. & Yu, Q. *Annu. Rev. Plant Biol.* **61**, 317–347 (2010).
19. Kelliher, T. *et al. Nature* **542**, 105–109 (2017).

ONLINE METHODS

Plasmid construction. To construct vectors A3A-PBE, A3A-Gam and A3A-PBE-ΔUGI, the Gam and deaminase human *APOBEC3A* sequences were codon-optimized for cereal plants and synthesized commercially (Genewiz, Suzhou, China), and the fusion protein sequences were cloned into the backbone of vector pJIT163. The constructs pOsU3-sgRNA, pTaU6-sgRNA and pAtU6-sgRNA were made as previously described[20–24]. The multiple sgRNA constructs were made as previously reported[25]. To construct the binary vector for *Agrobacterium*-mediated rice transformation, A3A-PBE and the sgRNA expression cassettes were cloned into the pHUE411 backbone[25] by using a ClonExpressII One Step Cloning Kit (Vazyme, Nanjing, China). All primer sets used in this work are listed in **Supplementary Table 5** and were synthesized by Beijing Genomics Institute (BGI).

Protoplast transfection. We used the winter wheat variety Kenong199, the *Japonica* rice variety Zhonghua11 and the Désirée potato variety to prepare the protoplasts used in this study. Wheat and rice protoplast isolation and transformation were performed as described[21,24]. Potato protoplasts were isolated and transformed as before[25] and incubated in 2 mL of culture medium E. Plasmid DNA (10 μg per construct) was introduced by PEG-mediated transfection, the mean transformation efficiency being 30–50% by flow cytometry. The transfected protoplasts were incubated at 23 °C. At 48 h after transfection, the protoplasts were collected to extract genomic DNA for deep amplicon sequencing and PCR restriction enzyme digestion assays (PCR-RE assays) (see below).

Flow cytometry analysis. Flow cytometry analysis was carried out using a FACSAria III (BD Biosciences). Rice protoplast cells were transfected with guide RNA expression plasmids, fluorophore expression plasmids and editor expression plasmids. Both samples were sorted for GFP-positive cells. Gating for all samples can be found in the **Supplementary Data**.

DNA extraction. Genomic DNA was extracted with a DNA Quick Plant System (Tiangen Biotech, Beijing, China). The targeted sequences were amplified with specific primers, and the amplicons were purified with an EasyPure PCR Purification Kit (TransGen Biotech, Beijing, China) and quantified with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA).

Amplicon deep sequencing and data analysis. Genomic DNA was extracted from the protoplast samples 48 h after transfection and used as template. In the first-round PCR, the target region was amplified with the Fast Mutagenesis System (TransGen Biotech, Beijing, China) using site specific primers (**Supplementary Table 5**). In the second, both forward and reverse barcodes were added to the ends of the PCR products for library construction (**Supplementary Table 5**). Equal amounts of the PCR products were pooled and samples were sequenced commercially (MyGenostics, Beijing, China) using the Illumina NextSeq 500 platform. The sgRNA target sites in the sequenced reads were examined for C-to-T substitutions and indels. The amplicon sequencing was repeated three times for each target site, using genomic DNA extracted from three independent protoplast samples. Analyses of base-editing processivity were performed as previously described[7].

***Agrobacterium*-mediated transformation of rice callus cells.** *Agrobacterium tumefaciens* strain AGL1 was transformed with the binary vectors by electroporation. *Agrobacterium*-mediated transformation of callus cells of Zhonghua11 was conducted as reported[20]. Hygromycin (50 μg/ml) was used to select transgenic plants.

Biolistic delivery of DNA constructs into wheat immature embryo cells. The DNAs of plasmids A3A-PBE and pTaU6-sgRNA were simultaneously delivered into immature embryos of Kenong199 via particle bombardment, as previously described[26,27]. After bombardment, the embryos were cultured for plantlet regeneration on medium without selective agent.

Potato protoplast regeneration. After transfection, protoplasts were embedded in alginate until calli were formed as previously described[28] Approximately 4 weeks after transfection, the calluses were released and incubated in liquid medium for an additional 2 to 4 weeks to allow further development. The enlarged calluses were then transferred to solid medium for shoot development.

Mutant identification by PCR-RE assays and Sanger sequencing. PCR-RE assays and Sanger sequencing were used to identify rice, wheat and potato mutants with C-to-T conversions in target regions, as described previously[21,24,26]. For rice and potato, T0 plants were examined individually (at least two leaves from each plant samples) by Sanger sequencing. For wheat, because the plants were regenerated from callus without herbicide selection, we combined the plantlets regenerated from the same immature embryos as a pool (each pool usually containing three to four plantlets) to detect the mutations using PCR-RE assay or Sanger sequencing. All the plantlets in the pools that gave positive PCR-RE or Sanger sequencing signals were sampled again and tested one by one with homoeolog-specific primers and confirmed by Sanger sequencing.

Detection of off-target mutations. Likely off-targets were predicted using the online tool CRISPR-P[29]. Off-target sites in rice *OsCDC48* and *OsNRT1.1B* T1 were identified and examined in this study.

Expression and purification of A3A-PBE-ΔUGI protein. For exogenous expression of A3A-PBE-ΔUGI, the A3A-PBE-ΔUGI coding sequence was cloned into pET42b to generate pET42b-A3A-PBE-ΔUGI, and then the plasmids were transformed into BL21 Star *E. coli* cells. Protein expression was induced with 0.5 mM IPTG at 18 °C for 14–16 h. After induction, cells were harvested and resuspended in lysis buffer (25 mM Tris-HCl (pH 8.0), 500 mM NaCl, 1 mM DTT, 10 mM imidazole, 0.1% Triton X-100, 1 mM PMSF). Cells were broken with a sonicator and then centrifuged at 25,000*g* for 30 min. The supernatant was incubated with Ni-NTA beads with rotation at 4 °C for 1 h, and the protein were purified according to the manufacturer's manual (GE Healthcare). Purified protein was concentrated using Amicon Ultra-15 with a 30-kDa cut off (Millipore), and was flash-frozen in liquid nitrogen and stored at –80 °C.

***In vitro* transcription of sgRNA.** sgRNAs DNA fragments were amplified from their appropriate expression plasmids using the relevant primers, meanwhile introducing a T7 promoter (**Supplementary Table 5**). Transcription was accomplished with an HiScribe T7 High Yield RNA Synthesis Kit (New England Biolabs) according to the manufacturer's instructions.

Statistical analysis. All numerical values are presented as means ± s.e.m. Statistical differences between control and the 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 this study are available in the article and its supplementary figures and tables or are available from the corresponding author on request. For sequence data, rice LOC_Os identifiers (http://rice.plantbiology.msu.edu/): LOC_Os01g55540 (*OsAAT*), LOC_Os03g05730 (*OsCDC48*), LOC_Os09g26999 (*OsDEP1*), LOC_Os10g40600 (*OsNRT1.1B*), LOC_Os02g11010 (*OsOD*, *OsEV*), LOC_Os02g07160 (*OsHPPD*); NCBI GenBank: AY210405 (*TaALS*), LS992089(*TaMTL*), GU167921 (*TaLOX2*), FJ039902 (*TaDEP1*), AM084898 (*TaHPPD*), MH264470 (*TaVRN1*), HM114275 (*StALS*), A23741 (*StGBSS*); NCBI Sequence Read Archive: SRR7586558, SRR7586559 and SRR7586560. Plasmids encoding A3A-PBE, A3A-Gam, A3A-PBE-ΔUGI, p35S-PBE, p35S-A3A-PBE, pAtU6-sgRNA and pH-A3A-PBE will be available from Addgene.

20. Shan, Q. *et al. Mol. Plant* **6**, 1365–1368 (2013).
21. Shan, Q. *et al. Nat. Biotechnol.* **31**, 686–688 (2013).
22. Zhang, D. *et al. Genome Biol.* **18**, 191 (2017).
23. Li, C. *et al. Genome Biol.* **19**, 59 (2018).
24. Wang, Y. *et al. Nat. Biotechnol.* **32**, 947–951 (2014).
25. Xing, H.L. *et al. BMC Plant Biol.* **14**, 327 (2014).
26. Liang, Z. *et al. Nat. Commun.* **8**, 14261 (2017).
27. Zhang, Y. *et al. Nat. Commun.* **7**, 12617 (2016).
28. Andersson, M. *et al. Plant Cell Rep.* **36**, 117–128 (2017).
29. Lei, Y. *et al. Mol. Plant* **7**, 1494–1496 (2014).

doi:10.1038/nbt.4261

nature research

Corresponding author(s): Caixia Gao

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.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
☐	☒	An indication of 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 statistics 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
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Illumina NextSeq 500 platform was used to collect the amplicon deep sequencing data. BD FACSAriaIII was used to do flow cytometry.
Data analysis	Graphpad prism 7 was used to analyze the data. FACSDiva Version 6.1.3 software was used for flow cytometry result analysis.

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 upon request. 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 and its Supplementary Information files or are available from the corresponding author on request.

Nature Biotechnology: doi:10.1038/nbt.4261

Field-specific reporting

Please select 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/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

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

Sample size	All the experiments were performed with three biological repeats. For wheat, rice and potato protoplast assays, about 500,000 protoplasts were used for each transfection. The number of mutants were confirmed by Sanger sequencing.
Data exclusions	No data exclusion.
Replication	All attempts for replication were successful. A minimum of 3 biological replicates were included.
Randomization	Wheat, rice and potato 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

Materials & experimental systems

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

Methods

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

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 10,000 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

Nature Biotechnology: doi:10.1038/nbt.4261

Gating strategy collect cells expressing either fluorophore. See the provided examples for gates used.

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

Nature Biotechnology: doi:10.1038/nbt.4261