nature biotechnology

Tuning plant phenotypes by precise, graded downregulation of gene expression

Chenxiao Xue[1,2], Fengti Qiu[1], Yuxiang Wang[1,2], Boshu Li[1,2], Kevin Tianmeng Zhao [3], Kunling Chen[1] & Caixia Gao [1,2]✉

The ability to control gene expression and generate quantitative phenotypic changes is essential for breeding new and desired traits into crops. Here we report an efficient, facile method for downregulating gene expression to predictable, desired levels by engineering upstream open reading frames (uORFs). We used base editing or prime editing to generate de novo uORFs or to extend existing uORFs by mutating their stop codons. By combining these approaches, we generated a suite of uORFs that incrementally downregulate the translation of primary open reading frames (pORFs) to 2.5–84.9% of the wild-type level. By editing the 5′ untranslated region of *OsDLT*, which encodes a member of the GRAS family and is involved in the brassinosteroid transduction pathway, we obtained, as predicted, a series of rice plants with varied plant heights and tiller numbers. These methods offer an efficient way to obtain genome-edited plants with graded expression of traits.

Variable gene expression can generate diverse plant phenotypes. The ability to fine-tune gene expression levels is, therefore, critical for improving crop traits while balancing complex tradeoffs caused by gene pleiotrophy[1,2]. Substantial efforts have been made to regulate gene expression. Widely used genetic tools, such as CRISPR–Cas[3], CRISPR interference (CRISPRi)[4] and RNA interference (RNAi)[5], generally result in a unique change of gene expression level that ranges from a given level of downregulation to complete absence[6]. Genome editing of promotors can produce a wide range of gene expression levels, thus generating quantitative phenotypic changes for breeding purposes[2,7,8] and offering a ground-breaking method for regulating gene expression at the transcriptional level. The use of base editors to mutate splice sites provides an effective approach for manipulating pre-mRNA splicing[9,10]. Upstream open reading frames (uORFs) are short protein-coding elements located in the 5′ untranslated regions (UTRs) of primary open reading frames (pORFs)[11] and are common in the mRNAs of eukaryotes[12–14]. In plants, 24–30% of coding mRNAs contain uORFs in their 5′ UTRs[12,14]. The presence of uORFs is associated with reduced mRNA translation[13]. Bioinformatics resources, such as uORFSCAN[15], uORFlight (http://uorflight.whu.edu.cn)[16] and PsORF (http://psorf.whu.edu.cn/)[17], have been used to search for uORFs in plants. Many factors have been reported to impact the inhibitory effects of uORFs[18–22].

In 2018, it was shown that knocking out endogenous uORFs is an efficient and tunable method for upregulating protein expression[23–25]. However, very few methods are available for fine-downregulating endogenous gene expression on the translational level in plants. Here we describe a simple, predictable and universal method for incrementally downregulating protein expression using precision genome editing to generate uORFs with different inhibitory activities, thus enlarging the toolbox for manipulating gene translation in plants.

Results

Repressing protein expression by generating de novo uORFs

Because we had previously reported that disrupting uORFs could upregulate protein expression[23–25], we hypothesized that introducing de novo uORFs might downregulate protein expression (Fig. 1a). To test this, we selected two genes, *AtABI1* and *OsBRI1*, and introduced ATG start codons into their 5′ UTRs to generate uORFs encoding at least two amino acids (Extended Data Fig. 1a). To assess the effects of these uORFs, 5′ UTRs with or without the corresponding uORFs were cloned upstream of the firefly luciferase (LUC) coding region in the dual-luciferase reporter system. The resulting constructs also harbored a cassette expressing *Renilla reniformis* luciferase (REN) as an internal vector control (Extended Data Fig. 1b). Constructs for each gene

[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]College of Advanced Agricultural Sciences, University of Chinese Academy of Sciences, Beijing, China. [3]Qi Biodesign, Beijing, China. ✉e-mail: cxgao@genetics.ac.cn
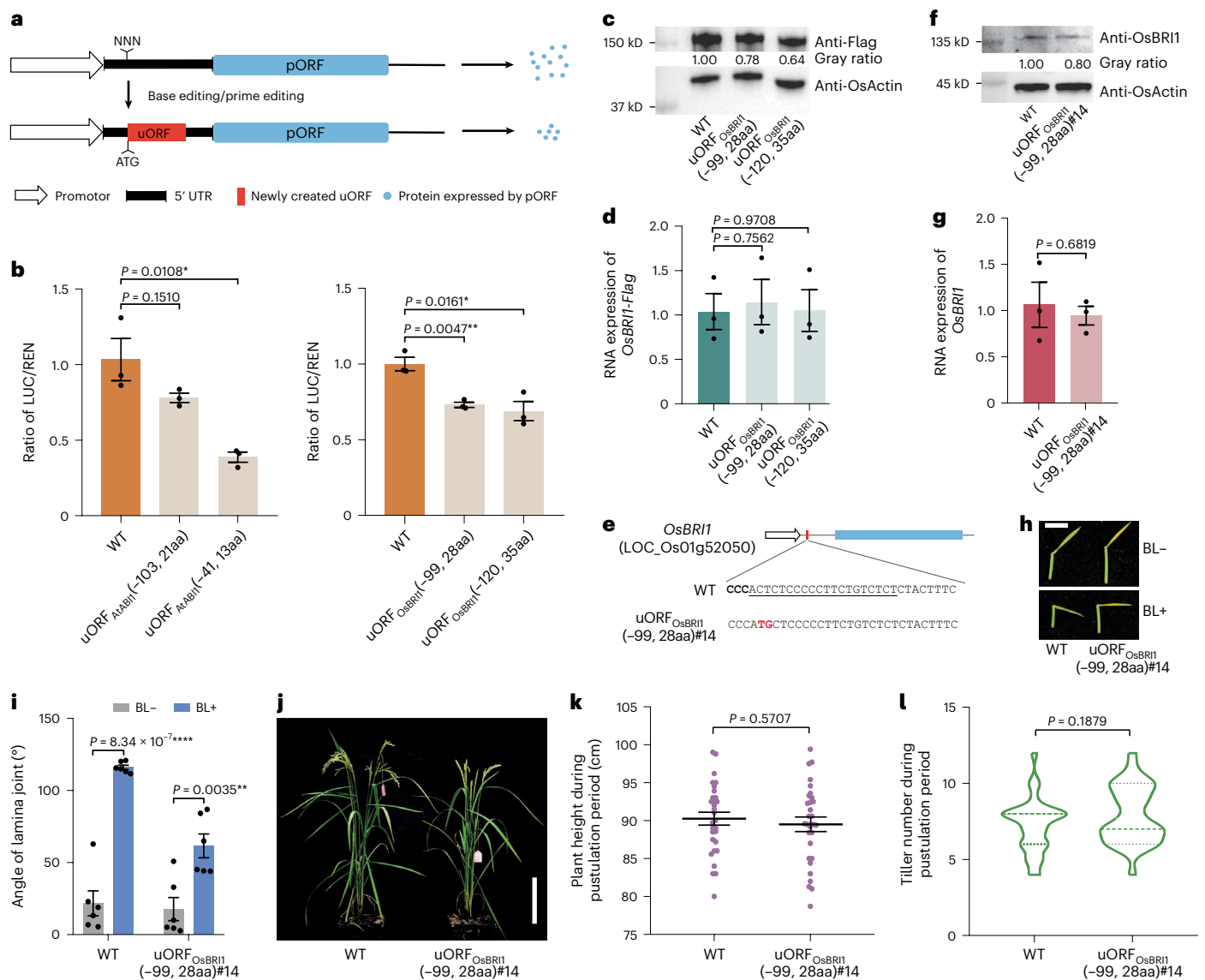
Fig. 1 | Introduced uORFs repress protein expression in protoplasts and plants. a, Schematic diagram of uORFs introduced to repress gene expression. **b**, Effects of artificial insertion of uORFs on LUC/REN activity ratios ($n = 3$). **c**, Effects of de novo uORFs on OsBRI1-Flag levels in protoplasts. **d**, Effects of expression of *OsBRI1-Flag* on RNA levels in protoplasts ($n = 3$). **e**, A T_0 homozygous mutant of *OsBRI1* carrying uORF$_{OsBRI1}$(−99, 28 aa). The target sequence is underlined. The PAM is in red and bold. The mutant base is in red and bold. **f,g**, Expression of *OsBRI1* at the protein (**f**) and RNA (**g**) ($n = 3$) level in flag leaves of WT and the T_1 progenies from the mutant uORF. **h**, BR sensitivity of WT and the T_1 progenies from the uORF$_{OsBRI1}$(−99, 28 aa)#14 assayed by the lamina joint inclination method. Images were taken 72 hours after immersion in water or 10^{-6} M epi-brassinolide (BL) solution. Scale bar, 1 cm. **i**, Statistical data

for lamina joint bending angles as described in **h** ($n = 6$ biologically independent samples). **j**, Grass morphology of WT and T_1 progenies of the mutant uORF during the pustulation period. Scale bar, 25 cm. **k,l**, Plant heights (**k**) and tiller numbers (**l**) of WT and the T_1 mutant progenies during the pustulation period ($n = 30$ biologically independent samples). In the dual-luciferase reporter system, a construct with the WT 5′ UTR was used as control, and the data were normalized to the average LUC/REN activity ratio of the control ($n = 3$). *OsActin* was used as loading control and as internal control in the immunoblot and quantitative RT−PCR assays. The angles of lamina joint were measured using ImageJ software. All data are presented as mean ± s.e.m. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ and ****$P < 0.0001$ by two-tailed Student's *t*-test.

were introduced into rice protoplasts, and then LUC activity relative to REN (LUC/REN) and LUC/REN mRNA levels were determined. The 5′ UTRs with newly introduced uORFs were found to reduce the LUC/REN activity ratios to 38.7–78.0% of the original value (Fig. 1b). Quantitative RT−PCR assays revealed that the levels of mRNA transcribed from the various constructs did not differ considerably (Extended Data Fig. 1c).

We next used *OsBRI1* as target[26] to see whether newly introduced uORFs retained their inhibitory activity on translation of the *OsBRI1* primary coding sequence (CDS). We created uORF$_{OsBRI1}$(−99, 28 aa) (the uORF starts at position −99 relative to the *OsBRI1* pORF start codon and is 28 amino acids (aa) long) and uORF$_{OsBRI1}$(−120, 35 aa) by

introducing ATG triplets 99 bp and 120 bp upstream of the pORF, generating CT-to-TG and C-to-G mutations, respectively (Extended Data Fig. 1a). We constructed vectors comprising an CDS of *OsBRI1-Flag* fusion with the wild-type (WT) 5′ UTR or one of the two 5′ UTR mutants. Western blotting (Fig. 1c and Source Data) and quantitative RT−PCR assays (Fig. 1d) revealed that the newly introduced uORFs inhibited OsBRI1-Flag production to the same extents as was measured by the dual-luciferase reporter system, without any effect on mRNA levels.

To test whether an introduced uORF could reduce endogenous protein expression in rice, we used prime editing to obtain mutants carrying uORF$_{OsBRI1}$(−99, 28 aa) and uORF$_{OsBRI1}$(−120, 35 aa) in the 5′ UTR

of *OsBRI1*. Using prime editing, we generated three prime editing guide RNAs (pegRNAs) (Supplementary Table 7) targeting each site. By next-generation sequencing, we found that *OsBRI1-T4* pegRNA, used together with the plant prime editor (PPE2)[27], was the most effective in introducing uORF$_{OsBRII}$(−99, 28 aa) (Extended Data Fig. 2a). Using a pH-ePPE binary vector harboring the corresponding engineered pegRNA (epegRNA)[28,29] (Extended Data Fig. 2b), we obtained 17 rice mutants with CT-to-TG mutations in uORF$_{OsBRII}$(−99, 28 aa) (Extended Data Fig. 2c and Supplementary Table 1).

OsBRI1 encodes a receptor for the phytohormone brassinosteroid (BR) in rice[26]. The T$_1$ progenies from a T$_0$ homozygous mutant uOR-F$_{OsBRII}$(−99, 28 aa)#14 were used for further phenotypic experiments (Fig. 1e and Supplementary Table 5). As expected, the level of OsBRI1 protein in the progenies of uORF$_{OsBRII}$(−99, 28 aa)#14 was reduced to the same extends as was seen in the transient expression system (Fig. 1f and Source Data), but the mRNA level of *OsBRI1* did not differ between WT and mutants (Fig. 1g). Lamina joint inclination assays showed that the lamina joint bending angles of the T$_1$ progenies from the uORF$_{OsBRII}$(−99, 28 aa)#14 were smaller than those of WT plants, indicating that the mutants were less sensitive to BR (Fig. 1h,i). The T$_1$ progenies from the uORF$_{OsBRII}$(−99, 28 aa)#14 also had compact stat-ures, but their heights and tiller numbers were similar to WT (Fig. 1j–l). This phenotype is consistent with that of the previously reported *OsBRI1* knockdown mutants generated by RNAi in which expression of *OsBRI1* was inhibited to nearly 70% (ref. [30]) (which is similar to the extent of downregulation of *OsBRI1* in the T$_1$ progenies from the uOR-F$_{OsBRII}$(−99, 28 aa)#14 as demonstrated in the transient assay (Fig. 1b,c)). These results confirm that introducing new uORFs can inhibit mRNA translation quantitatively in plants.

Reducing protein expression by extending an endogenous uORF

uORF size and intercistronic length have both been reported to impact the inhibitory ability of uORFs[20,21]. We used a strategy to extend existing uORFs involving mutating the stop codons of the original uORFs to lengthen the uORF coding sequences while simultaneously shorten-ing the intercistronic distance (Fig. 2a). The 5′ UTRs of *AtABI1*, *AtPYR1*, *AtBRI1*, *OsGW7*, *OsDLT* and *OsCKX2* were chosen to evaluate the effect of this approach (Extended Data Fig. 3a). Dual-luciferase assays (Extended Data Fig. 3b) showed that extending a uORF enhanced its inhibitory activity and reduced the LUC/REN activity ratio to 9.5–86.9% of its original level (Fig. 2b) but had no effect on the LUC/REN mRNA ratio (Extended Data Fig. 4). These results indicate that the inhibitory effects of uORFs can be exploited to reduce gene expression.

We selected *OsDLT* as a target to further evaluate whether uOR-F$_{OsDLT}$(−589, 56 aa) retained its inhibitory activity when combined with the *OsDLT* primary CDS. We created uORF$_{OsDLT}$(−589, 56 aa) by artifi-cially mutating two putative stop codons through A-to-G mutations at positions −580 and −571 relative to the pORF start codon in the 5′ UTR of *OsDLT*, thus extending the original 3-aa uORF to 56 amino acids (Extended Data Fig. 3a). When we used an *OsDLT-Flag* expression cassette together with the WT or mutant *OsDLT* 5′ UTR, western blot-ting (Fig. 2c and Source Data) and quantitative RT–PCR (Fig. 2d) assays revealed that uORF$_{OsDLT}$(−589, 56 aa) inhibited OsDLT-Flag production to a level similar to that seen in the dual-luciferase assay while not affecting mRNA levels. We designed three single guide RNAs (sgRNAs) (Supplementary Table 6) targeting each of the two stop codons and used an evolved adenine base editor (ABE8e)[31] to generate the desired A-to-G mutations and obtain rice mutants carrying the uORF$_{OsDLT}$(−589, 56 aa) in the endogenous gene. Next-generation sequencing showed that *OsDLT-m1-1-T1* and *OsDLT-m1-2-T1* sgRNAs had the highest editing efficiencies (Extended Data Fig. 5a), so we used a pH-ABE8e-spG vec-tor[31,32] (Extended Data Fig. 5b and Supplementary Sequences) and these two sgRNAs and obtained 78 rice mutants carrying uORF$_{OsDLT}$(−589, 56 aa) (Extended Data Fig. 5c and Supplementary Table 2).

OsDLT encodes a plant-specific GRAS family member and is involved in the BR transduction pathway in rice[33,34]. We tested the BR sensitivity of WT and the T$_1$ progenies of the uORF$_{OsDLT}$(−589, 56 aa)#8 (Fig. 2e and Supplementary Table 5) using the lamina joint inclina-tion assay. The lamina joint bending angles of the T$_1$ progenies of the mutant uORF were smaller than those of WT plants (Fig. 2f,g), whereas the levels of *OsDLT* transcripts were similar (Fig. 2h). The T$_1$ progenies from the mutant uORF had compact statures (Fig. 2i), and their heights and tiller numbers were much lower than those of WT plants (Fig. 2j,k). Because the phenotype of the T$_1$ progenies from the mutant uORF was consistent with that of *OsDLT* loss-of-function mutants[33], we conclude that its phenotype is due to reduced translation level of the *OsDLT* mRNA. These results indicate that manipulating the inhibitory activity of uORFs can control gene expression in plants.

Graded downregulation of protein expression

Based on the strategies described above, we thought that generat-ing uORFs with different inhibitory activities in the 5′ UTR of a gene might allow us to control the extent of downregulation of a gene prod-uct. We selected *OsTB1*, *OsTCP19* and *OsDLT* to test this expectation. Because there was no existing uORF in the 5′ UTR of *OsTB1*, we chose to insert separately four de novo uORFs (uORF$_{OsTB1}$(−293, 10 aa), uOR-F$_{OsTB1}$(−58, 4 aa), uORF$_{OsTB1}$(−176, 30 aa) and uORF$_{OsTB1}$(−75, 23 aa)) (Fig. 3a and Extended Data Fig. 6). The dual-luciferase assay showed that uORF$_{OsTB1}$(−58, 4 aa), uORF$_{OsTB1}$(−176, 30 aa) and uORF$_{OsTB1}$(−75, 23 aa) reduced LUC/REN activity incrementally to 47.9%, 37.9% and 31.9%, respectively, whereas uORF$_{OsTB1}$(−293, 10 aa) had almost no effect on LUC protein levels (Fig. 3a). Quantitative RT–PCR assays showed that most of these newly created uORFs had no effect on LUC/REN mRNA ratios (Fig. 3a). The 5′ UTR of *OsTCP19* containing a uORF encoding two amino acids was located 17 bp upstream of the pORF. We generated four additional uORFs by inserting three new uORFs (uORF$_{OsTCP19}$(−44, 11 aa), uORF$_{OsTCP19}$(−52, 213 aa) and uORF$_{OsTCP19}$(−173, 18 aa)) or extending the original uORF to encode 91 amino acids (uORF$_{OsTCP19}$(−17, 91 aa)) (Fig. 3b and Extended Data Fig. 6). These uORFs reduced LUC/REN activity progressively by 97.7%, 83.5%, 52.8% and 22.6% but had no effect on LUC/REN mRNA levels (Fig. 3b). Thus, we were able to induce incre-mental inhibition of gene expression by generating a series of uORFs, demonstrating that such a series of uORF variants can produce specific and subtle changes of gene expression level and of the corresponding quantitative traits.

Because *OsDLT* is a pleiotropic gene influencing multiple agro-nomic traits, including plant height, leaf angle, tiller number and grain shape[35], the ability to fine-tune OsDLT levels by genome editing should facilitate crop improvements. Besides WT-uORF$_{OsDLT}$(−589, 3 aa), there is another uORF (WT-uORF$_{OsDLT}$(−540, 32 aa)) in the 5′ UTR of *OsDLT*. We generated six uORFs by introducing five further uORFs (uORF$_{OsDLT}$(−514, 31 aa), uORF$_{OsDLT}$(−402, 27 aa), uORF$_{OsDLT}$(−220, 22 aa), uORF$_{OsDLT}$(−141, 42 aa) and uORF$_{OsDLT}$(−105, 30 aa)) or extending the original uORF (uORF$_{OsDLT}$(−540, 73 aa)) (Fig. 3c and Extended Data Fig. 6). These uORFs reduced LUC/REN activity progressively to 84.9%, 73.5%, 47.2%, 3.8%, 2.5% and 21.9% of its starting value but had no effect on LUC/REN mRNA levels (Fig. 3c). To generate these uORFs in the endogenous 5′ UTR of *OsDLT*, we designed a suite of sgRNAs (Supplementary Table 6) and pegRNAs (Supplementary Table 7) to be used in conjunction with ABE8e and PPE2, respectively. We introduced each editing construct into rice protoplasts and measured its editing efficiency. The *OsDLT-T15* pegRNA targeting uORF$_{OsDLT}$(−402, 27 aa), *OsDLT-m2-T4* sgRNA target-ing uORF$_{OsDLT}$(−540, 73 aa), *OsDLT-T7* pegRNA targeting uORF$_{OsDLT}$(−141, 42 aa) and *OsDLT-T11* pegRNA targeting uORF$_{OsDLT}$(−105, 30 aa) had relatively high editing efficiencies (Extended Data Fig. 7a,b), so these modified alleles were chosen for further study. Expression vectors with the relevant 5′ UTRs of *OsDLT* cloned upstream of a *OsDLT-Flag* CDS were used to confirm the inhibitory effects of the four mutated uORFs. Western blotting showed that all four uORFs reduced OsDLT-Flag levels
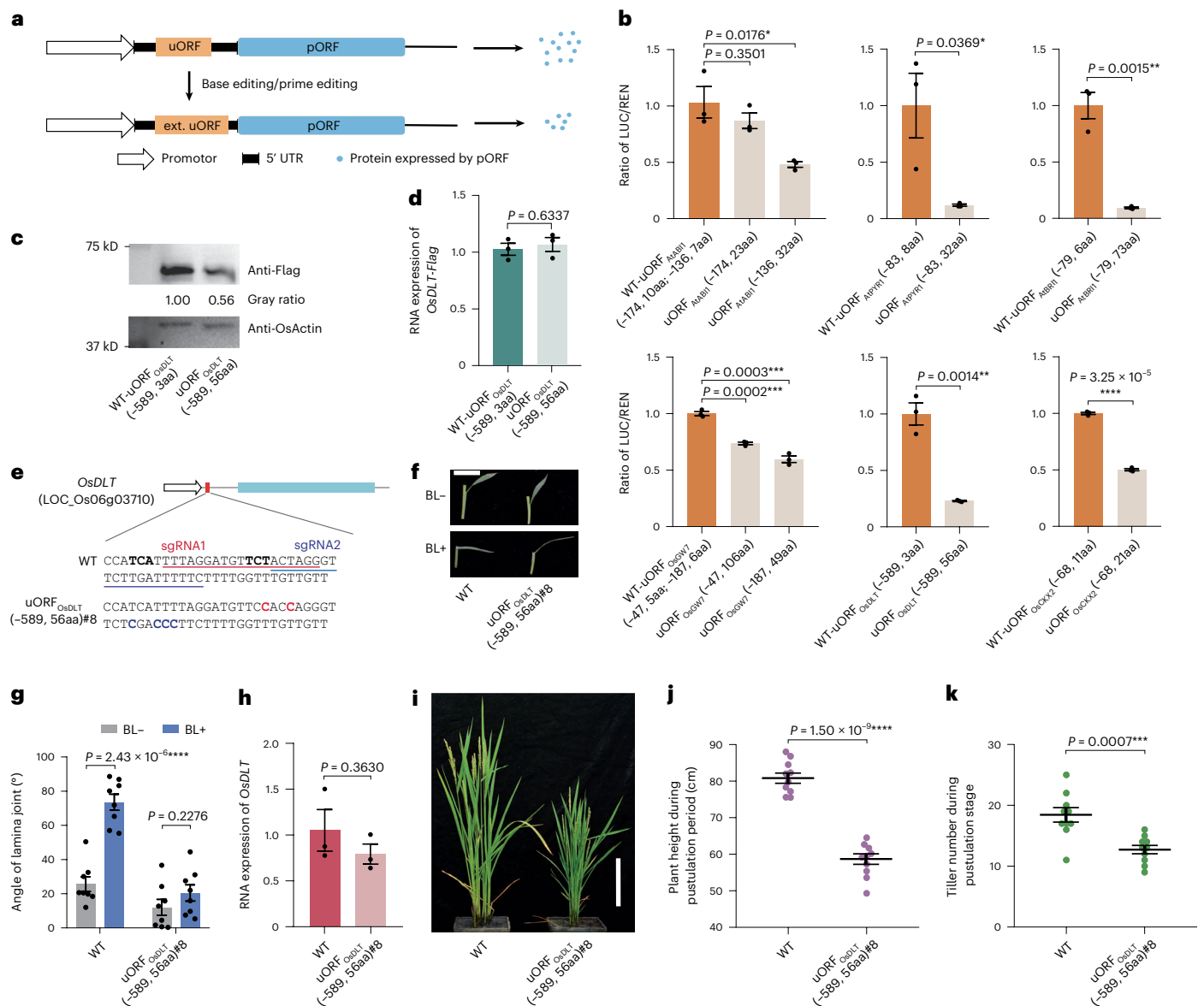
Fig. 2 | Extending original uORFs by base editing reduces protein expression in protoplasts and plants. a, Schematic representation of procedure for extending a uORF by mutating its stop codon. **b**, The effect of the extended uORF on LUC/REN activity ($n = 3$). **c**, The effect of extending the uORF on production of OsDLT-Flag in protoplasts. **d**, Expression of *OsDLT-Flag* at the RNA level in protoplasts ($n = 3$). **e**, The homozygous T_0 mutant of *OsDLT* carrying uORF$_{OsDLT}$(−589, 56 aa). The target sequence is underlined. The PAM is black and bold. The red and bold represents the mutant bases targeted by sgRNA 1, and the blue and bold represents the mutant bases targeted by sgRNA 2. **f**, BR sensitivity of WT and T_1 progenies from the mutant uORF assayed by the lamina joint inclination method. Images were taken 48 hours after immersion in water or 10^{-6} M BL solution. Scale bar, 1 cm. **g**, Statistical data for the lamina joint bending angles in **f** ($n = 8$ biologically independent samples). **h**, Expression of *OsDLT* in WT and T_1 progenies from the mutant uORF at the RNA level in 4-day-old seedlings ($n = 3$). **i**, Grass morphology of WT and T_1 progenies from the mutant uORF during the flowering period. Scale bar, 25 cm. **j,k**, Plant heights (**j**) and tiller numbers (**k**) of WT and the T_1 progenies of the mutant uORF during the pustulation period ($n = 10$ biologically independent samples). In the dual-luciferase reporter system, a construct with the WT 5′ UTR was used as control, and the data were normalized to the average LUC/REN activity of the control ($n = 3$). The angles of lamina joints were measured using ImageJ software. *OsActin* was used as loading control and internal control in the immunoblot and quantitative RT–PCR assays. All data are presented as mean ± s.e.m. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ and ****$P < 0.0001$ by two-tailed Student's *t*-test.

(Fig. 3d and Source Data) and that the quantitative effects were consistent with those obtained using the dual-luciferase reporter system (Fig. 3c), whereas mRNA levels were not affected (Fig. 3e). We then constructed PE (pH-ePPE-epegRNA) (Extended Data Fig. 2b) and ABE (pH-ABE8e) binary vectors (Extended Data Fig. 7c and Supplementary Sequences) harboring the corresponding epegRNAs or sgRNAs and introduced these vectors into rice calli by *Agrobacterium*-mediated transformations. Plants regenerated from resistant calli were examined by Sanger sequencing to obtain mutants carrying each of the

four uORFs in the *OsDLT* 5′ UTR region (Extended Data Fig. 8 and Supplementary Tables 3 and 4), and the T_1 progenies of homozygotes harboring these mutant uORFs were used in further experiments (Fig. 3f and Supplementary Table 5).

We first evaluated the BR sensitivity of these T_1 progenies. Their lamina joint angles were found to decline to varying extents as expected from the corresponding reductions of OsDLT levels (Fig. 4a,b), whereas the levels of *OsDLT* transcripts in these edited plants were similar to that of WT plants (Fig. 4c). The statures of the mutants were
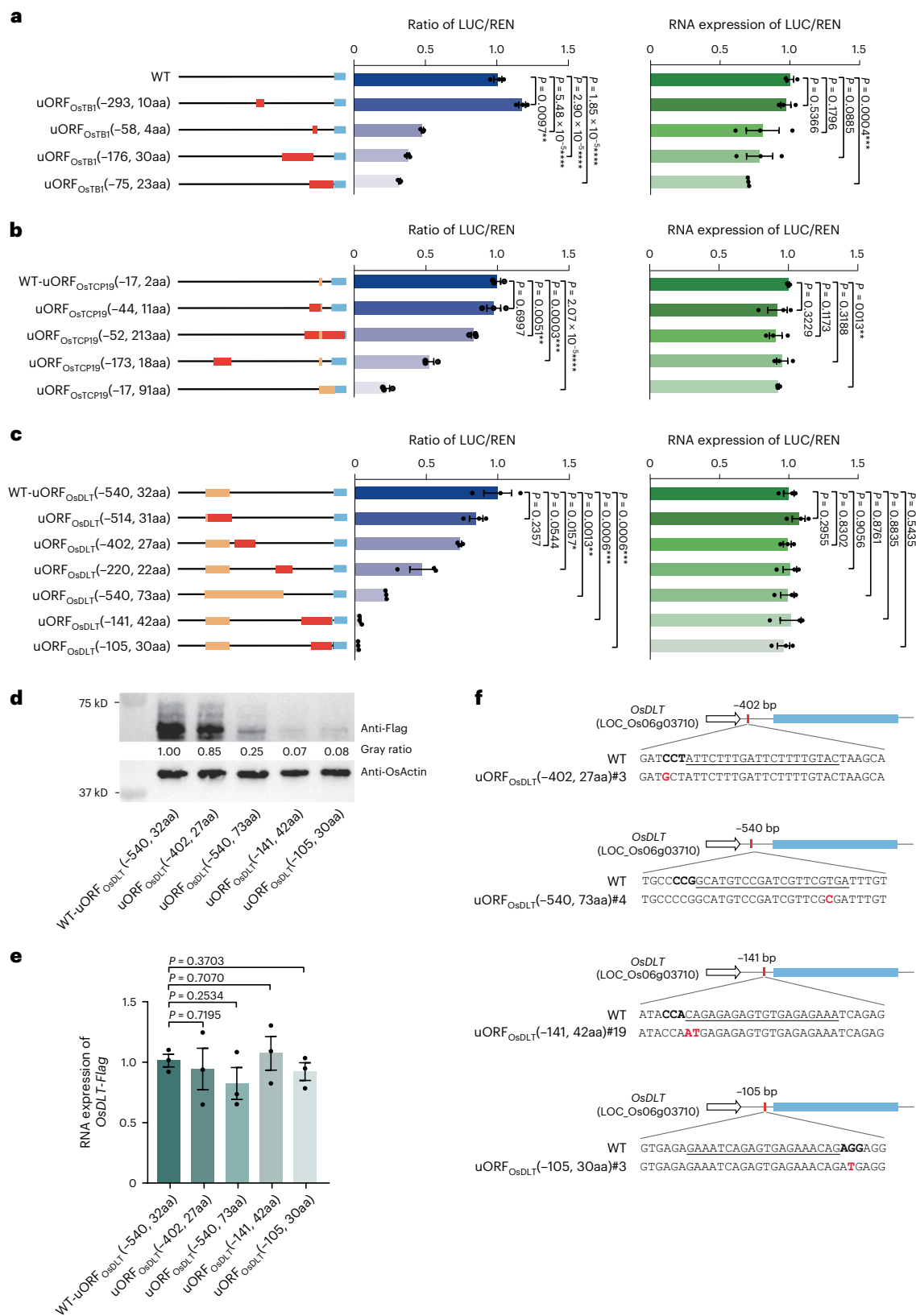
Fig. 3 | Producing uORFs with diverse inhibitory activities to downregulate protein expression in a graded fashion. a–c, Schematic representation of the uORFs generated in the 5′ UTRs of *OsTB1* (**a**), *OsTCP19* (**b**) and *OsDLT* (**c**) (left) and the effects of these uORFs on LUC/REN activity (middle) and RNA levels (right) in dual-luciferase assays (*n* = 3). The orange square represents the endogenous uORF; the red square represents the newly created uORF; and the blue square represents the pORF. **d,e**, Effects of different forms of uORF on OsDLT-Flag (**d**) and *OsDLT-Flag* mRNA (**e**) (*n* = 3) levels. *OsActin* was used as loading control and internal control in the immunoblot and quantitative RT–PCR assays. **f**, The T₀ homozygous mutants of *OsDLT* carrying uORF_OsDLT(−402, 27 aa), uORF_OsDLT(−540, 73 aa), uORF_OsDLT(−141, 42 aa) and uORF_OsDLT(−105, 30 aa). The target sequence is underlined. The PAM is black and bold. The mutant base is red and bold. In the dual-luciferase reporter system, a construct with the WT 5′ UTR was used as control, and the data were normalized to the average LUC/REN activity and mRNA levels of the control (*n* = 3). All data are presented as mean ± s.e.m. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ and ****$P < 0.0001$ by two-tailed Student's *t*-test.
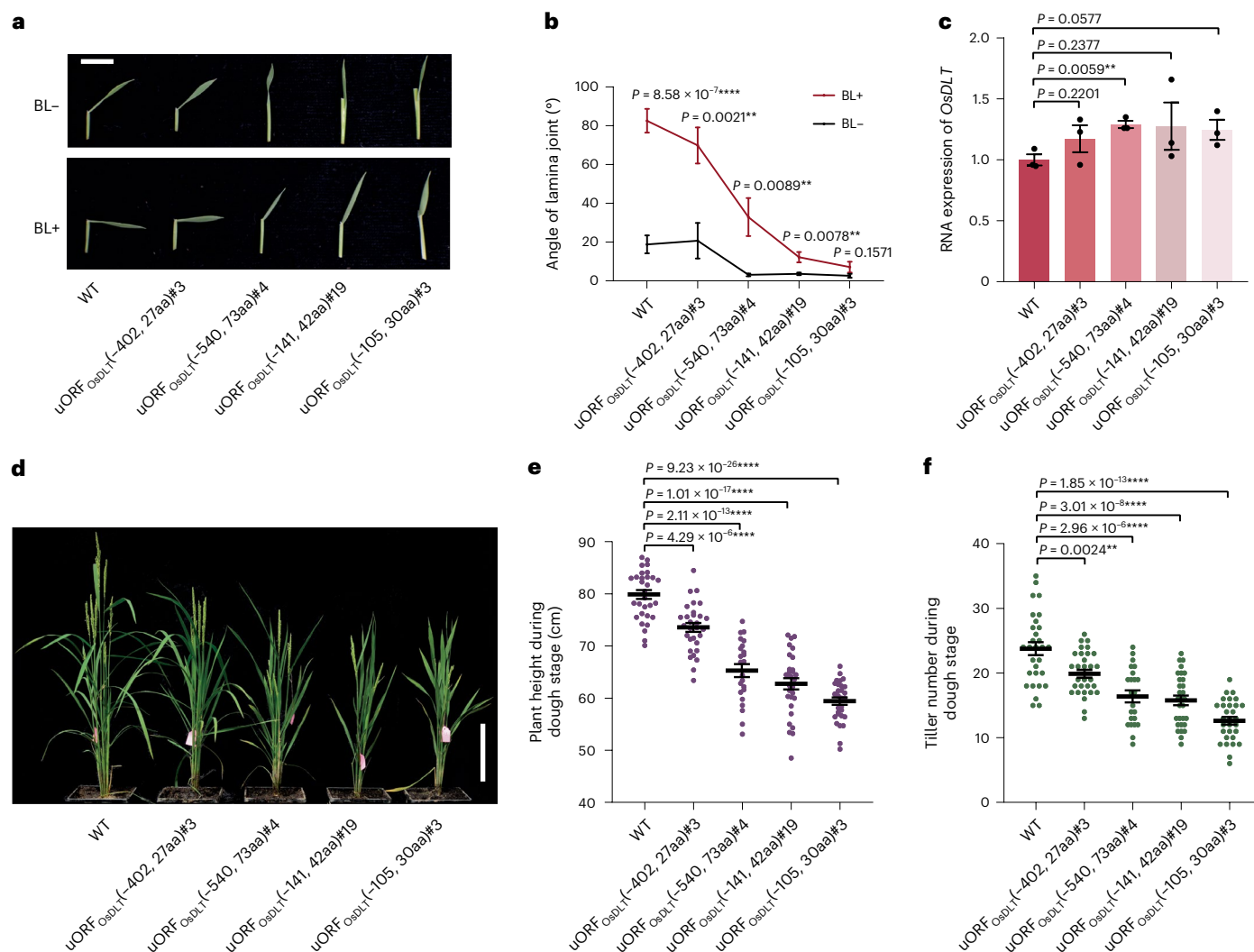
Fig. 4 | Obtaining mutants with the expected quantitative traits.
a, BR sensitivity of WT and T_1 progenies from the uORF$_{OsDLT}$(−402, 27 aa)#3, uORF$_{OsDLT}$(−540, 73 aa)#4, uORF$_{OsDLT}$(−141, 42 aa)#19 and uORF$_{OsDLT}$(−105, 30 aa)#3 in lamina inclination experiments. Images were taken 48 hours after immersion in water or 10^{-6} M BL solution. Scale bar, 1 cm. **b**, Statistical data for the lamina joint bending angles in **a** (n = 8 biologically independent samples). **c**, Expression of *OsDLT* at the RNA level in WT and the T_1 progenies from the mutant uORFs in 4-day-old seedlings (n = 3). *OsActin* was used as internal control. **d**, Grass morphology of WT and T_1 progenies from the mutant uORFs during the flowering stage. Scale bar, 25 cm. **e**,**f**, Plant heights (**e**) and tiller numbers (**f**) of WT and the T_1 progenies from the mutant uORFs during the dough period (n = 30, 30, 23, 30 and 30 biologically independent T_1 progenies from the WT, uORF$_{OsDLT}$(−402, 27 aa)#3, uORF$_{OsDLT}$(−540, 73 aa)#4, uORF$_{OsDLT}$(−141, 42 aa)#19 and uORF$_{OsDLT}$(−105, 30 aa)#3)). The angles of lamina joints were measured using ImageJ software. The data are mean ± s.e.m. *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001 by two-tailed Student's t-test.

also progressively more compact and shorter at the flowering stage (Fig. 4d). Furthermore, statistical analyses confirmed that the plant heights and tiller numbers of the mutants at the dough stage decreased progressively (Fig. 4e,f). Because *OsDLT* transcript levels were all similar (Fig. 4c), we conclude that different efficiences of translation of *OsDLT* transcripts were responsible for the phenotypic differences between the mutants, which also were consistent with the corresponding reductions in OsDLT levels seen in the transient reporter system (Fig. 3c,d). These findings highlight our ability to quantitatively and incrementally manipulate mRNA translation by creating uORFs with different inhibitory activities without affecting mRNA transcription.

Discussion

The ability to generate quantitative changes of gene expression is very important for obtaining new phenotypes. Over the years, our laboratory has reported that eliminating uORFs by genome editing can upregulate mRNA translation and so provide an efficient method for stimulating gene expression[23–25]. In this study, we demonstrate the use of base editing and prime editing to generate sets of short uORFs in the 5′ UTRs of coding genes to achieve the converse effect—that is, to quantitatively downregulate endogenous gene expression. By analyzing the 5′ UTR of genes with website tools[15–17], it is possible to identify those genes without uORFs in their 5′ UTR. One can then generate several de novo uORFs upstream of these genes by mutating or inserting 1–3 bases to create upstream ATGs. For genes with their own uORFs, one can extend the original uORF or generate additional uORFs. Using transient reporter systems, such as the dual-luciferase assay, one then can identify uORFs with the desired inhibitory effects and obtain mutants carrying these uORFs by base editing or prime editing.

Many factors have been reported to influence the inhibitory abilities of uORFs[18–22], so it is difficult to predict the effect of the uORFs on pORFs translation and the phenotypes. In this study, we highlight the fact that the variable extents of phenotypic changes observed in the edited plants reflect the changes in pORF expression measured using

a simple transient reporter system. Therefore, we conclude that the transient luciferase reporter system can rapidly and reliably predict the phenotypes of uORF mutants and their effects on gene expression in genome-edited plants. Moreover, transgene-free lines with the desired quantitative trait levels are readily obtained (Extended Data Fig. 9, Supplementary Table 5 and Source Data) and should expedite crop improvement. Notably, downregulating gene expression through manipulating or introducing new uORFs had no effect on mRNA transcription levels, in agreement with previous findings[23,24]. We think that the ability to use precision genome editing to edit uORFs in the 5' UTRs of coding genes in plants represents an exciting and widely applicable approach to generating quantitative changes in gene expression in crop breeding.

Online content

Any methods, additional references, Nature Portfolio 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/s41587-023-01707-w.

References

1. Gao, C. Genome engineering for crop improvement and future agriculture. *Cell* **184**, 1621–1635 (2021).
2. Song, X. et al. Targeting a gene regulatory element enhances rice grain yield by decoupling panicle number and size. *Nat. Biotechnol.* **40**, 1403–1411 (2022).
3. Wang, Y. et al. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol.* **32**, 947–951 (2014).
4. Qi, L. S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).
5. Hannon, G. J. RNA interference. *Nature* **418**, 244–251 (2002).
6. Bowman, E. K. et al. Bidirectional titration of yeast gene expression using a pooled CRISPR guide RNA approach. *Proc. Natl Acad. Sci. USA* **117**, 18424–18430 (2020).
7. Rodríguez-Leal, D., Lemmon, Z. H., Man, J., Bartlett, M. E. & Lippman, Z. B. Engineering quantitative trait variation for crop improvement by genome editing. *Cell* **171**, 470–480 (2017).
8. Hendelman, A. et al. Conserved pleiotropy of an ancient plant homeobox gene uncovered by *cis*-regulatory dissection. *Cell* **184**, 1724–1739 (2021).
9. Xue, C., Zhang, H., Lin, Q., Fan, R. & Gao, C. Manipulating mRNA splicing by base editing in plants. *Sci. China Life Sci.* **61**, 1293–1300 (2018).
10. Yuan, J. et al. Genetic modulation of RNA splicing with a CRISPR-guided cytidine deaminase. *Mol. Cell* **72**, 380–394 (2018).
11. Barbosa, C., Peixeiro, I. & Romao, L. Gene expression regulation by upstream open reading frames and human disease. *PLoS Genet.* **9**, e1003529 (2013).
12. Srivastava, A. K., Lu, Y., Zinta, G., Lang, Z. & Zhu, J. K. UTR-dependent control of gene expression in plants. *Trends Plant Sci.* **23**, 248–259 (2018).
13. Zhang, H. et al. Genome-wide maps of ribosomal occupancy provide insights into adaptive evolution and regulatory roles of uORFs during *Drosophila* development. *PLoS Biol.* **16**, e2003903 (2018).
14. Zhang, T., Wu, A., Yue, Y. & Zhao, Y. uORFs: important *cis*-regulatory elements in plants. *Int. J. Mol. Sci.* **21**, 6238 (2020).
15. Tran, M. K., Schultz, C. J. & Baumann, U. Conserved upstream open reading frames in higher plants. *BMC Genomics* **9**, 361 (2008).
16. Niu, R. et al. uORFlight: a vehicle toward uORF-mediated translational regulation mechanisms in eukaryotes. *Database (Oxford)* **2020**, baaa007 (2020).
17. Chen, Y. et al. PsORF: a database of small ORFs in plants. *Plant Biotechnol. J.* **11**, 2158–2160 (2020).
18. Ferreira, J. P., Overton, K. W. & Wang, C. L. Tuning gene expression with synthetic upstream open reading frames. *Proc. Natl Acad. Sci. USA* **110**, 11284–11289 (2013).
19. Lin, Y. et al. Impacts of uORF codon identity and position on translation regulation. *Nucleic Acids Res.* **47**, 9358–9367 (2019).
20. Kozak, M. Effects of intercistronic length on the efficiency of reinitiation by eucaryotic ribosomes. *Mol. Cell. Biol.* **7**, 3438–3445 (1987).
21. Kozak, M. Constraints on reinitiation of translation in mammals. *Nucleic Acids Res.* **29**, 5226–5232 (2001).
22. Wang, J., Zhang, X., Greene, G. H., Xu, G. & Dong, X. PABP/ purine-rich motif as an initiation module for cap-independent translation in pattern-triggered immunity. *Cell* **185**, 3186–3200 (2022).
23. Zhang, H. et al. Genome editing of upstream open reading frames enables translational control in plants. *Nat. Biotechnol.* **36**, 894–898 (2018).
24. Xing, S. et al. Fine-tuning sugar content in strawberry. *Genome Biol.* **21**, 230 (2020).
25. Si, X., Zhang, H., Wang, Y., Chen, K. & Gao, C. Manipulating gene translation in plants by CRISPR–Cas9-mediated genome editing of upstream open reading frames. *Nat. Protoc.* **15**, 338–363 (2020).
26. Yamamuro, C. et al. Loss of function of a rice *brassinosteroid insensitive1* homolog prevents internode elongation and bending of the lamina joint. *Plant Cell* **12**, 1591–1605 (2000).
27. Lin, Q. et al. Prime genome editing in rice and wheat. *Nat. Biotechnol.* **38**, 582–585 (2020).
28. Zong, Y. et al. An engineered prime editor with enhanced editing efficiency in plants. *Nat. Biotechnol.* **40**, 1394–1402 (2022).
29. Nelson, J. W. et al. Engineered pegRNAs improve prime editing efficiency. *Nat. Biotechnol.* **40**, 402–410 (2022).
30. Morinaka, Y. et al. Morphological alteration caused by brassinosteroid insensitivity increases the biomass and grain production of rice. *Plant Physiol.* **141**, 924–931 (2006).
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. Walton, R. T., Christie, K. A., Whittaker, M. N. & Kleinstiver, B. P. Unconstrained genome targeting with near-PAMless engineered CRISPR–Cas9 variants. *Science* **368**, 290–296 (2020).
33. Tong, H. et al. DWARF AND LOW-TILLERING, a new member of the GRAS family, plays positive roles in brassinosteroid signaling in rice. *Plant J.* **58**, 803–816 (2009).
34. Tong, H. et al. DWARF AND LOW-TILLERING acts as a direct downstream target of a GSK3/SHAGGY-like kinase to mediate brassinosteroid responses in rice. *Plant Cell* **24**, 2562–2577 (2012).
35. Tong, H. & Chu, C. Functional specificities of brassinosteroid and potential utilization for crop improvement. *Trends Plant Sci.* **23**, 1016–1028 (2018).

Methods

Plasmid construction

To construct plasmids used in the dual-luciferase assay, constructs containing the *35S* promotor fused with artificially designed sequences with *Bsa*I restriction enzyme sites were synthesized commercially (GenScript) (Supplementary Sequences). They were cloned into pGreenII0800-LUC vector[36] digested with *Hin*dIII and *Nco*I to construct pGreen-1562 vector. The WT and mutated 5′ UTRs of each gene were amplified by PCR and then cloned into the pGreen-1562 vector backbone between *Bsa*I restriction enzyme sites. To construct a vector for pH-ABE8e, TadA8e was amplified from PABE8 (ref. [28]) and cloned into pH-PABE7-sgRNA[37]. For pH-ABE8e-spG, spG was amplified from ePPE-SpG[28] and used to replace part of Cas9 in pH-ABE8e. To construct a vector for pH-ePPE-epegRNA, the artificially designed sequence containing two *Bsa*I restriction enzyme sites and the tevopreQ1 were synthesized commercially and replaced the sequence between *Bsa*I and *Hin*dIII restriction enzyme sites in pH-ePPE[28]. To construct the sgRNA expression vectors, primers containing the target spacer were annealed and then cloned into the OsU3-sgRNA vector digested with *Bsa*I. The dual sgRNA expression cassette in pH-ABE8e-spG was constructed as previously reported[37]. To construct the pegRNA expression vectors, using OsU3-sgRNA plasmid as template, pegRNAs are amplified using primers containing the target spacer in the forward primer and the PBS + RT sequences in the reverse primer and cloned into the OsU3-sgRNA vector digested with *Bsa*I and *Hin*dIII[38,39]. epegRNAs were amplified using primers carrying the target spacer in the forward primer and the PBS + RT sequences in the reverse primer and cloned into pH-ePPE-epegRNA digested with *Bsa*I. PCR was performed using TransStart FastPfu DNA Polymerase (TransGen Biotech). A One-Step Cloning Kit (Nanjing Vazyme Biotech) was used for vector cloning. Primers used in this study are listed in Supplementary Table 8.

Protoplast transfection

The *Japonica* rice variety Zhonghua 11 and *Arabidopsis* Columbia were used to generate protoplasts. About 14-day-old rice or *Arabidopsis* seedlings cultured at 27 °C or 24 °C on MS medium with a 16-hour light/8-hour dark cycle were used for protoplast isolation. For rice protoplasts, health and fresh rice sheath were cut into fine strips and digested in the enzyme solution (1.5% Cellulase R-10, 0.75% Macerozyme R-10, 0.8 M mannitol, 10 mM MES at pH 5.7, 10 mM CaCl$_2$ and 0.1% BSA), followed by vacuum infiltration for 30 minutes in the dark using a vacuum pump at approximtely −15 to −20 (in Hg). After a 5–6-hour digestion with gentle shaking (60–80 r.p.m.), protoplasts were released by filtering through 40-μm nylon meshes into round-bottom tubes. The pellets were collected by centrifugation at 250*g* for 3 minutes. After washing with W5 solution (154 mM NaCl, 125 mM CaCl$_2$, 5 mM KCl and 2 mM MES at pH 5.7), the pellets were then resuspended in MMG solution (0.4 M mannitol, 15 mM MgCl$_2$ and 4 mM MES at pH 5.7). The protoplast transformation was carried out in PEG solution (40% (w/v) PEG 4000, 0.2 M mannitol and 0.1 M CaCl$_2$). The transformation system (plasmid DNA mixed with 200 μl of protoplasts in 220 μl of PEG solution) was gently mixed. After 20-minute incubation at room temperature in the dark, protoplast cells were harvested and washed by 880 μl of W5 solution. The cells were centrifuged and then resuspended in 1 ml of WI solution (0.5 M mannitol, 20 mM KCl and 4 mM MES at pH 5.7) and cultured under dark at room temperature for 48 hours (refs. [40,41]). The isolation and transfection of *Arabidopsis* protoplasts is following the protocol reported previously[42]. Plasmids used for protoplasts transformation were extracted with a Wizard Plus Midipreps DNA Purification System (Promega). Then, 5 μg of plasmid DNA was used for PEG-mediated transfections, and 3 μg was used for western blotting. After incubation, protoplasts were collected for DNA, RNA or protein extraction, and the LUC/REN activity was measured using the Dual-Luciferase Reporter Assay System (Promega).

Protein extraction and protein gel blot analysis

Protein was extracted from protoplasts with extraction buffer containing 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% NP40, 4 M urea and 1 mM PMSF. Gel blot analysis was performed with anti-Flag (1:3,000 dilution), anti-BRI1 (1:2,000 dilution) or anti-OsActin antibody (1:5,000 dilution). The secondary antibody was goat anti-mouse antibody conjugated to horseradish peroxidase (1:10,000 dilution), and reaction signals were visualized by enhanced chemiluminescence (Millipore). The gray levels of each band were calculated using the ImageJ tool. Gray ratios were normalized to the construct with WT 5′ UTR.

DNA extraction

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

Amplicon deep sequencing and data analysis

Specific primers with 5′ barcodes were designed to amplify the targeted sequences. Amplicons were purified with EasyPure PCR Purification Kits (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 follows: percentage (number of reads with the desired edit) / (number of total reads)[28]. For all base editing yield quantification, base editing efficiency was calculated as follows: percentage (number of reads with A-to-G substitutions at the expected sites) / (number of total reads)[28]. Amplicon sequencing was repeated at least two times for each target site using genomic DNA extracted from at least two independent protoplast samples. Primers are listed in Supplementary Table 8.

Agrobacterium-mediated transformation of rice callus cells

Binary plasmids pH-ePPE-epegRNA, pH-ABE8e-SpG and pH-ABE8e containing sgRNAs or epegRNAs were introduced into *Agrobacterium tumefaciens* strain AGL1 by electroporation (400 ng per transformation). *Agrobacterium*-mediated transformation of callus cells of the *Japonica* rice variety Kitaake was conducted according to Hiei et al.[43]. Hygromycin (50 μg ml^{-1}) was used to select transgenic plants.

Mutant identification by Sanger sequencing

Plants regenerated from rice calluses were examined individually. At least two leaves of each plant were used for genomic DNA extraction. Target sequences were amplified with 2× Rapid Taq Master Mix (Nanjing Vazyme Biotech). Sanger sequencing was used to detect mutants.

RNA preparation and quantitative real-time PCR

Total RNA was extracted from protoplasts or plant samples with a Takara MiniBEST Plant RNA Extraction Kit. Reverse transcription was performed using M-MLV Reverse Transcriptase (Promega). Subsequently, quantitative real-time PCR was performed using a ChamQ Universal SYBR qPCR Master Mix (Nanjing Vazyme Biotech). The primers used are listed in Supplementary Table 8.

Lamina joint inclination assay

Next, 2-cm segments containing the second leaf lamina joint, leaf blade and leaf sheath were excised from 8-day-old rice seedlings. The excised samples were floated on sterile water for 10 minutes and then transferred to BL solution or sterile water. After incubation for 48 hours or 72 hours at 28 °C, lamina joint angles were measured with ImageJ software (https://imagej.nih.gov/ij/).

Statistical analysis

GraphPad Prism 8 and Microsoft Excel 2016 software were used to analyze the data. All numerical values are presented as mean ± s.e.m.

Differences between control and treatments were tested by two-tailed Student's *t*-test.

Reporting Summary

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

Data availability

All data supporting the findings of this study are available in the article, extended data figures and supplementary information or are available from the corresponding author upon reasonable request. Sequence data are present in The Arabidopsis Information Resource (https://seqviewer.arabidopsis.org/) or Phytozome databases (https://phytozome-next.jgi.doe.gov/) under the following accession numbers: *AtABI1* (AT4G26080), *AtPYR1* (AT4G17870), *AtBRI1* (AT4G39400), *OsBRI1* (LOC_Os01g52050), *OsGW7* (LOC_Os07g41200), *OsDLT* (LOC_Os06g03710), *OsCKX2* (LOC_Os01g10110), *OsTCP19* (LOC_Os06g12230) and *OsTB1* (LOC_Os03g49880). The deep sequencing data have been deposited in a National Center for Biotechnology Information BioProject database (accession code PRJNA931443)[44]. Plasmids for pH-ABE8e and pH-ABE8e-spG will be made available through Addgene. Source data are provided with this paper.

References

36. Hellens, R. P. et al. Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. *Plant Methods* **1**, 13 (2005).
37. Li, C. et al. Expanded base editing in rice and wheat using a Cas9–adenosine deaminase fusion. *Genome Biol.* **19**, 59 (2018).
38. Lin, Q. et al. High-efficiency prime editing with optimized, paired pegRNAs in plants. *Nat. Biotechnol.* **39**, 923–927 (2021).
39. Jin, S., Lin, Q., Gao, Q. & Gao, C. Optimized prime editing in monocot plants using PlantPegDesigner and engineered plant prime editors (ePPEs). *Nat. Protoc.* https://doi.org/10.1038/s41596-022-00773-9 (2022).
40. 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).
41. Shan, Q. et al. Rapid and efficient gene modification in rice and *Brachypodium* using TALENs. *Mol. Plant* **6**, 1365–1368 (2013).
42. Zhai, Z., Jung, H. I. & Vatamaniuk, O. K. Isolation of protoplasts from tissues of 14-day-old seedlings of *Arabidopsis thaliana*. *J. Vis. Exp.* **30**, e1149 (2009).
43. Hiei, Y., Ohta, S., Komari, T. & Kumashiro, T. Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. *Plant J.* **6**, 271–282 (1994).
44. Xue, C. et al. Tuning plant phenotypes by precise, graded downregulation of gene expression. *National Center for Biotechnology Information (NCBI)* https://dataview.ncbi.nlm.nih.gov/object/PRJNA931443 (2023).

Author contributions

C.X., K.C. and C.G. designed the project. C.X., F.Q. and Y.W. performed the experiments. B.L. performed rice transformation. C.X., K.T.Z. and C.G. wrote the manuscript. C.G. supervised the project. All authors reviewed the manuscript.

Competing interests

The authors have submitted a patent application based on the results reported in this paper. K.T.Z. is a founder and employee at Qi Biodesign.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41587-023-01707-w.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41587-023-01707-w.

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

Peer review information *Nature Biotechnology* thanks the anonymous reviewers 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 | Creating uORFs in 5′ UTRs. (a) 5′ UTR and part of CDS of *AtABI1* and *OsBRI1*. Lowercase is the non-uORF sequence in 5′ UTR; black uppercase is the sequence of uORF; gold uppercase is the sequence of pORF; red bold base is the upstream ATG (uATG) sites to be created. **(b)** Schematic diagram of the dual-luciferase reporter system with or without de novo ATG in 5′ UTR upstream the CDS of *LUC*. **(c)** RNA expression of *LUC* relative to *REN* in protoplasts. The data were normalized to control ($n = 3$). All data are presented as mean ± s.e.m. *$P < 0.05$ by two-tailed Student's *t*-test.

a

b

c

uORF$_{OsBRI1}$(-99, 28aa)

Extended Data Fig. 2 | Genotypes of prime-edited rice mutants carrying uORF$_{OsBRI1}$(−99, 28aa). (a) Editing efficiencies of pegRNAs with PPE2 used to generate uORF$_{OsBRI1}$(−120, 35aa) or uORF$_{OsBRI1}$(−99, 28aa) at the endogenous 5' UTR of *OsBRI1* in protoplasts ($n = 2$). **(b)** Schematic representation of the pH-ePPE-epegRNA vector. The black arrows indicate three pairs of primers used to detect transgene-free mutants. **(c)** Sanger sequencing chromatograms of representative prime-edited mutants carrying uORF$_{OsBRI1}$(−99, 28aa). Red arrows represent the desired edits.

a

AtABI1

gaagcaattgttgcattagcctacccatttcctccttctttctctcttctatctgtgaacaaggcacattagaactcttcttttcaacttttttaggtgtatatagatgaatctagaaa
tagttttatagttggaaattaattgaagagagagagatattactacaccaatcttttcaagaggtcctaacgaattacccacaatccaggaaacccttattgaaattcaattca
tttctttctttctgtgtttgtgattttcccgggaaatatttttgggtatATGTCTCTCTGTTTTTTGCTTTCCTTTTTTCA**TAG**GAGTCATGTG
TTTCTTCTTGTCTTCC**TAG**CTTCTTCTAATAAAGTCCTTCTCTTGTGAAAATCTCTCGAATTTTCATTTTT
GTTCCATTGGAGCTATCTTATAGatcacaaccagagaaaaagatcaaatctttaccgtta

uORF$_{AtABI1}$(-174, 23aa) uORF$_{AtABI1}$(-136, 32aa)

AtPYR1

uORF$_{AtPYR1}$(-83, 32aa)

gagagagtctaaaagctcgtcgtcgtcttcaATGGTGAATCTCAAACCAATTGGA**TAA**ATAAAACAAAACAAAAAAAAC
TTCACAAAAAAAAAAAAAAAGATCCAAATTACAACCATGCCTTCGGAGTTAACACCAGAAGAACGATC

AtBRI1

uORF$_{AtBRI1}$(-79, 73aa)

gtcttttgactctcttctctctctcctctctttctgctttcctcaatctctctctttctatctctagagcttccacttcctctctaATGGTGGAACCAAAACCC**TAG**
ATTCCCCCTTTCATCTTCTCTACTTCCCACACTTTTCTCTCTCACAAACTCTTGAGAAATGAAGACTTTT
TCAAGCTTCTTTCTCTCTGTAACAACTCTCTTCTTCTTCTCCTTCTTTTCTCTTTCATTTCAAGCTTCACC
ATCTCAGTCTTTATACAGAGAAATCCATCAGCTTATAAGCTTCAAAGACGTTCTTCCTGACAAGAATCT

OsDLT

caccacaccggagaggtgagtgagagtgagagagtgagagcagagaccaccaccaccggagaggttagtgagagaggagtggtaatggtgaggcaacaagag
taggttccatttcatatcatcactaggatagcgtagtttgtaggctgcatctccatctccatcgccattgattcgcattgcatccatcattttaggATGTTCTAC**TAG**
GGTTCT**TGA**TTTTTCTTTTGGTTTGTTGTTTTGACGAATGGAGGTATTGTTGGGATTCGCCGCCTGCTG
CTCGTCGTCGTCGTCGCCGATGAGGAGGCCGTGCGGGCTCTGCCCCGGCATGTCCGATCGTTCGTGAT
TTGTTTTTTCTACATGTTTTAGggcccatttgttcttgatcctattctttgattcttttgtactaagcattctaaggcgaagccacccattctttcctgcatat
atacttacaaacacatagcccccatctgatctcacaaacattatttctctctctttttttctcagttttttctttgttgatttactgaccaaattctttggaagaacaacaagatcatc
tggtttttatctgctcattcttttgtacatcgaatcatatacatttccattccaccaaagccttagccagataccacagagagagtgtgagagaaatcagagtgagaaacag
aggaggaagaagaagaagaagacgaggaggaggaggaggaggaggagcaggaggaggaggaggtctcttcttggcacgtcgcgttccggcgagtgacgtgtctcc
ggg

OsCKX2

uORF$_{OsCKX2}$(-68, 21aa)

tctagctatctatcagctgccttccatcgtcagcacacaaactacacaagaatctgcttatttataggccaccttgtcccttctacaATGGTGCAAGAACACA
CAAATTCACACACACAC**TGA**CACACACAAACCGATCGATTGATTGATTGAta

OsGW7

uORF$_{OsGW7}$(-187, 49aa)

accctgcATGGCATTCGCACGCCAT**TGA**CACATATCTCATCATACCATCACCATACCACATCTCATCTCACG
CGTCCCAAAGAAAGTGATAGTGTACGTACCTCTGCTTGCTCCTTGGCTTTCTGAGCTGAGCTGCTCTG

uORF$_{OsGW7}$(-47, 106aa)

CAGTTGGATGTGCGTGAGCTGG**TGA**TCTGGGAGGAGTCGGAGTCGGAGGAGGAGATGCCTCCGGCG
AGGGTGCTCGGCGGCGGTGGTGGAGGAGGGGGGGTTGGGGGACGAGGCGCCGGAGCTGGAGAGGCA
GATGGGGTGCATGGCCGGCATCTTCCAGATCTTCGACCGCCGCCAGCGCCTGCTCACCGCGCGCCGCC
GCCGCCCGCCGCCCAAGATGCTGCCCCCGGGCCCAGGCCATACTCTTCCAAGAAGCAGCAGCAATGT
TGCAGCGCAGAGCTCAAGTACCTCCAAGATCGTTCTGGAGAAAACATTCAGCAAGAGCATGACCGAG

b

Extended Data Fig. 3 | Extending uORFs in 5′ UTRs. (a) 5′ UTR and part of the CDS of *AtABI1*, *AtPYR1*, *AtBRI1*, *OsDLT*, *OsCKX2* and *OsGW7*. Lowercase is the non-uORF sequence in 5′ UTR; underlined uppercase is the CDS of uORF; gold uppercase is the CDS of pORF; red bold base is the stop codons to be mutated. **(b)** Schematic diagram of the dual-luciferase reporter system with original or extended uORF in 5′ UTR upstream the CDS of *LUC*.

Extended Data Fig. 4 | Effects of extended uORFs on LUC/REN mRNA levels in dual-luciferase assay. RNA expression of *LUC* relative to *REN* in protoplasts. The data were normalized to control (*n* = 3). All data are presented as mean ± s.e.m. *P < 0.05 by two-tailed Student's *t*-test.

Extended Data Fig. 5 | Genotypes of base-edited rice mutants containing uORF$_{OsDLT}$(−589, 56aa). (a) Editing efficiencies of sgRNAs with ABE8e to generating uORF$_{OsDLT}$(−589, 56aa) at the endogenous 5′ UTR of *OsDLT* in protoplasts ($n = 3$). All data are presented as mean ± s.e.m. **(b)** Schematic representation of the pH-ABE8e-spG vector. **(c)** Sanger sequencing chromatograms of representative base-edited mutants containing uORF$_{OsDLT}$(−589, 56aa). Red arrows indicate the desired edits.

OsDLT

caccacaccggagaggtgagtgagagtgagagagtgagagcagagaccaccaccaccggagaggttagtgagagaggagtggtaatggtgaggcaacaa
gagtaggttccatttcatatcatcactaggatagcgtagtttgtaggctgcatctccatctccatcgccattgattcgcattgcatccatcattttaggatgttctactagg

uORF_{OsDLT}(-540, 73aa)

gttcttgattttctttggtttgttgttttgacga**ATG**GAGGTATTGTTGGGATTCGCCGC**CTG**CTGCTCGTCGTCGTCGTC
GCCGATGAGGAGGCCGTGCGGGCTCTGCCCCGGCATGTCCGATCGTTCG**TGA**TTTGTTTTTTCTACA

uORF_{OsDLT}(-514, 31aa)

TGTTTTAGGGCCCATTTGTTCTTG**ATC**CTATTCTTTGATTCTTTTGTACTAAGCATTCTAAGGCGAAG
CCACCCATTCTTTCCTGCATATATACTTACAAACACATAGcccccatctgatctcacaaacattatttctctctcttttttctcagttt

uORF_{OsDLT}(-402, 27aa)

tttctttgttgatttactgaccaaattctttggaagaacaacaagatc**ATC**TGGTTTTTATCTGCTCATTCTTTTGTACATCGAATCA
TATACATTTCCATTCCACCAAAGCCTTAGccagatacca**CAG**AGAGAGTGTGAGAGAAATCAGAGTGAG

uORF_{OsDLT}(-220, 22aa)

uORF_{OsDLT}(-141, 42aa)

AAACAG**AGG**AGGAAGAAGAAGAAGAAGACGAGGAGGAGGAGGAGGAGGAGCAGGAGGAGGAG
GAGGTCTCTTCTTGGCACGTCGCGTTCCGGCGAGTGAcgtgtctccggg

uORF_{OsDLT}(-105, 30aa)

OsTCP19

aatcccgacaaagtacacacaccatcgtctcccctcctgcctcgatgctagctctcccattccatctctccactctccctcctatcatatgctcctcccttctccactct
tcttctcaaatcctctctcttcctcttcgtcctcctcctcttgcaaccaaaaaacaccataaaatagtaatcccaatccacaagctcacctctctcttgcacacacacacac
ctgacacaacacttacacatcataccagaagaaaaaccaattcgattttgccaaattacctagctatagcatataggtaggtagctaggtagcttggttttccctaatag

uORF_{OsTCP19}(-173, 18aa)

ctctagcttgatctctttctagctctacc**GGC**GGCGTGGCGGCGGCGGCGGCGGCCGCACCACTCTACATCATCACGTA
CACCTGAccagcttagcgggaagcccacaagaagagtttgtaggtcaccaatcagatcatcagttcatct**TTG**TGGTT**GTG**TGGGTGTGT

uORF_{OsTCP19}(-52, 213aa)   uORF_{OsTCP19}(-44, 11aa)

GTGTATATATATACCATGGCG**TGA**GAG**TAG**CCATGGATGTCACCGGAGACGGCGGAGGAGGAGGGCA
ACGGCCCAATTTCCCCCTGCAGCTCCTCGGGAAGAAGGAGGAGCAGACGTGCTCGACGTCGCAGAC
TGCCGGGGCGGGCGGCGGCGGCGTCGTGGGCGCGAATGGGTCGGCGGCGGCGGCGCCGCCGAAG
CGGACGTCGACGAAGGACCGGCACACGAAGGTGGACGGGCGGGGGCGGCGCATCCGGATGCCGGC
GATCTGCGCCGCGCGGGTGTTCCAGCTGACGCGGGAGCTCGGGCACAAGACCGACGGCGAGACCA
TCGAGTGGCTGCTGCAGCAGGCGGAGCCGGCGGTGATCGCGGCGACCGGGACGGGCACCATCCCG
GCCAACTTCACCTCCCTCAACATCTCCCTCCGCTCCTCCGGCTCGTCGCTCTCCATCCCTTCTCACCT
CCGCCTTGCCGGCTTGGCTGGCCCTCGCTTCGGCGGCGGCGCGCGGGCGGCGGACGCGTGGGACCG
CGTCGTCGGCCTCGGGTTCGGCGGTGCGGCCGACGCCCCGTCCTCCGCCACCTCCTCCTCCTCGTCG
CCGCTTCTGCTGAGCTTCCACTCCGGTAGCGTCGGCCTTGACGTGTCGCCGCCGTCGGCGTCGACCT

uORF_{OsTCP19}(-17, 91aa)

OsTB1

aagatggcaacaccctgatctctagcttagctgcagaggggagaggaacctcacatccaaactcctagctacaacttgtactagcatcctaagcaaccaagcaca
accaaagcaagcaagcacgaacaattctttcttcctctctacctctagctgctgcctgcctcctaatcctcctacccaccactccacatgagcccatgctgtgtgcc

uORF_{OsTB1}(-293, 10aa)

TGTGTCTGTGTGTGTGTGTTCTACTCCTACCATGAgagaagagaccaagcatcaaccaagctagctagctcgtcctctcctcgatctcta
cttctctctcccacacaagctgagcgcc**CAG**GTAGGCTGCCTGCTAGGTCTCGTGCATGGCCGGACACATCTGATCAT

uORF_{OsTB1}(-176, 30aa)

AGCCCACTACGGCACTATTCCCCCCTTCCGCCTCGCACGCTGAgaggtggcc**GGA**GAGGGAGGGAGGCC

uORF_{OsTB1}(-75, 23aa)

AGCGAGCAGCAGTAGCAGCAGCAACGCGGCTAGGAGTAAGGAGTCCCATCAGTAAagcATGCTTCCT

uORF_{OsTB1}(-58, 4aa)

Extended Data Fig. 6 | 5′ UTR and part of the CDS of *OsDLT*, *OsTCP19* and *OsTB1*. Lowercase is non-uORF sequence in 5′ UTR; underlined uppercase is the CDS of uORF; gold uppercase is the CDS of pORF; red bold base is the uATG site to be created or stop codon to be mutated.

Extended Data Fig. 7 | Editing efficiencies of pegRNAs and sgRNAs used to generate uORF$_{OsDLT}$(−402, 27aa), uORF$_{OsDLT}$(−540, 73aa), uORF$_{OsDLT}$(−141, 42aa) and uORF$_{OsDLT}$(−105, 30aa) in the endogenous 5′ UTR of *OsDLT*. (a) Editing efficiencies of pegRNAs with plant prime editor (PPE2) used to generate uORF$_{OsDLT}$(−402, 27aa), uORF$_{OsDLT}$(−141, 42aa) and uORF$_{OsDLT}$(−105, 30aa) in the endogenous 5′ UTR of *OsDLT* in protoplasts ($n = 2$). (b) Editing efficiencies of sgRNAs with adenine base editor (ABE8e) used to generate uORF$_{OsDLT}$(−540, 73aa) in the endogenous 5′ UTR of *OsDLT* in protoplasts ($n = 3$). The data are presented as mean ± s.e.m. (c) Schematic representation of the pH-ABE8e vector.
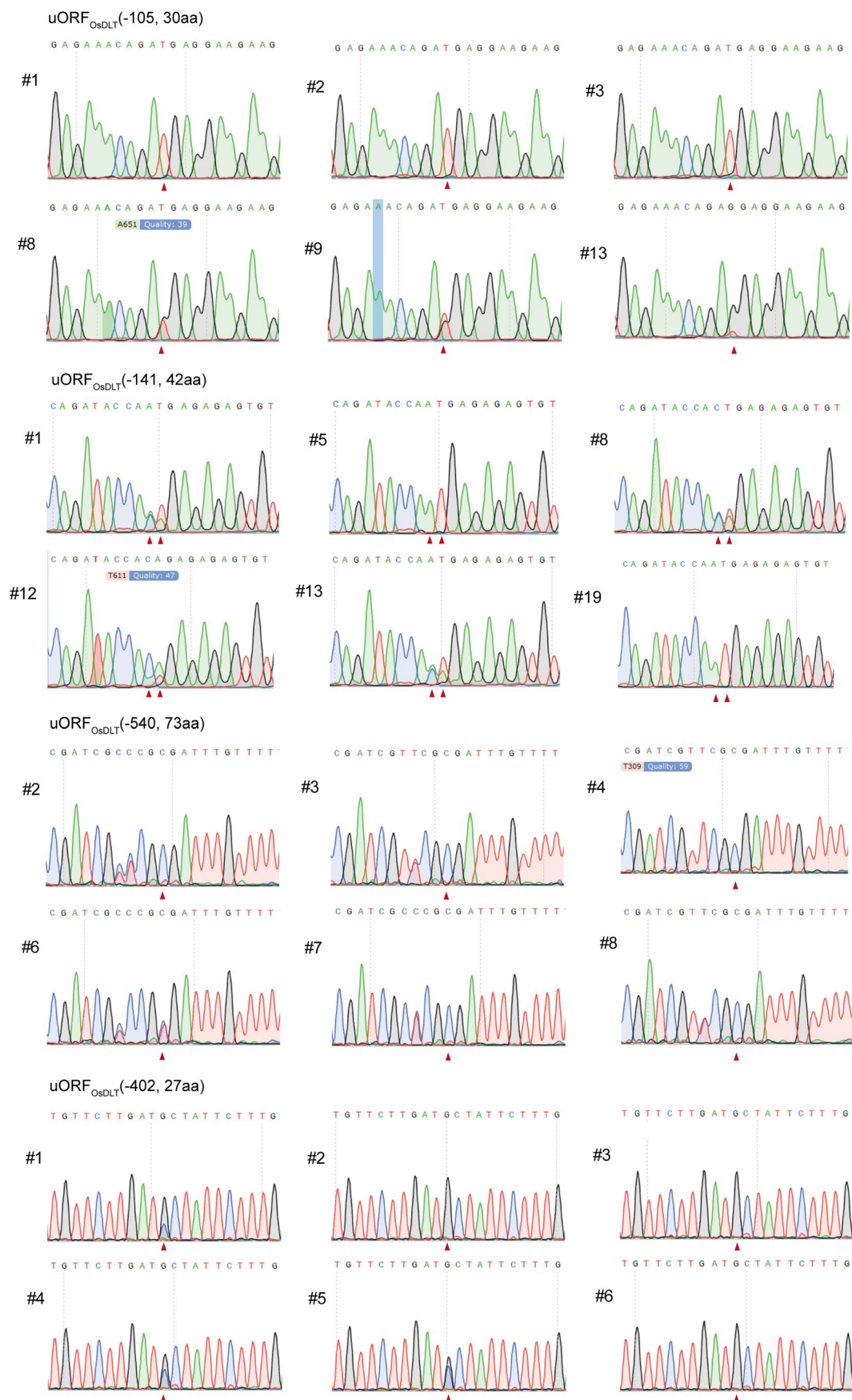
Extended Data Fig. 8 | Sanger sequencing chromatograms of representative mutants containing uORF$_{OsDLT}$(−402, 27aa), uORF$_{OsDLT}$(−540, 73aa), uORF$_{OsDLT}$(−141, 42aa) and uORF$_{OsDLT}$(−105, 30aa), respectively. Red arrows indicate the desired edits.
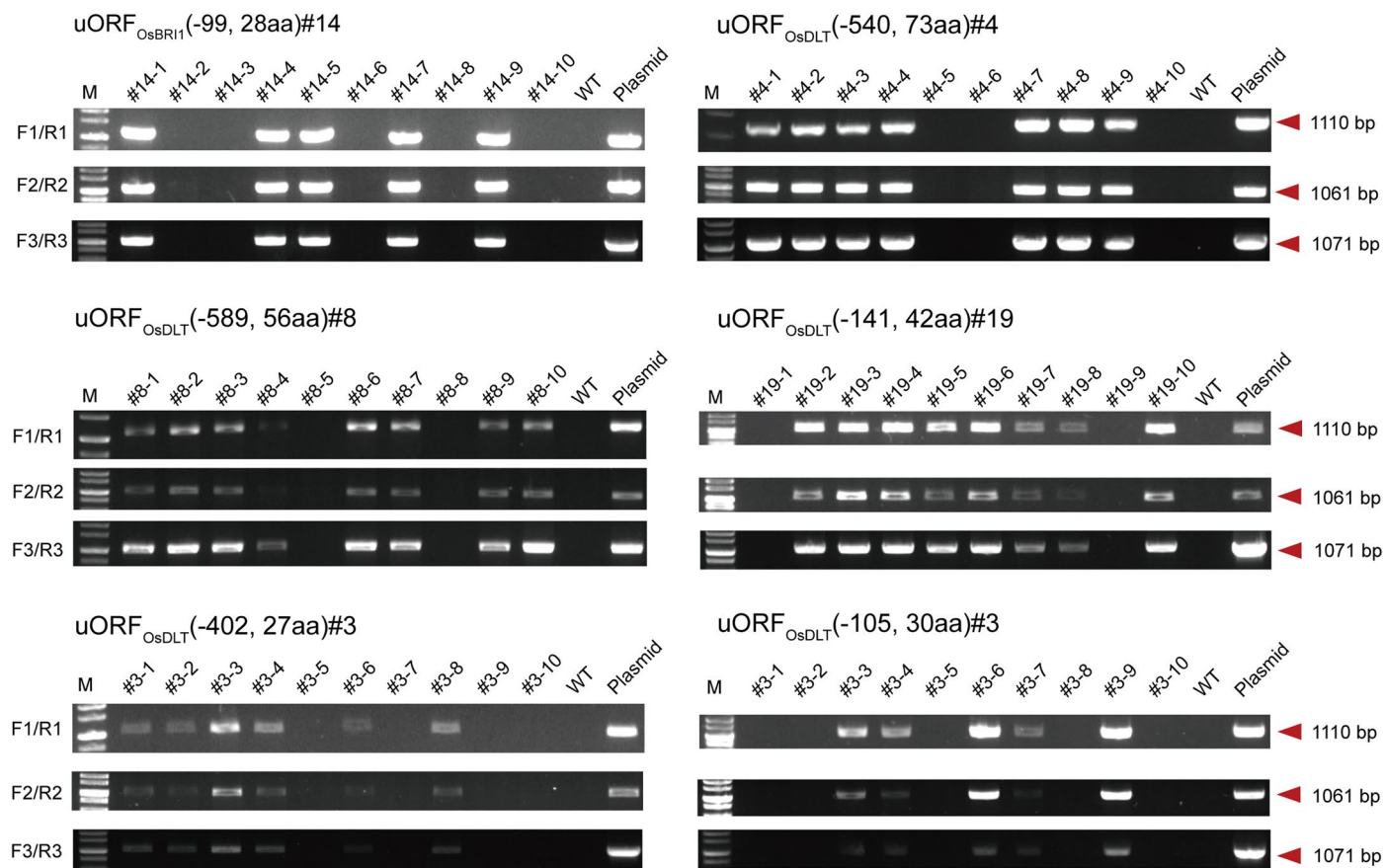
Extended Data Fig. 9 | Detection of transgene-free mutants with three pairs of primers based on the pH-ePPE-epegRNA, pH-ABE8e-spG and pH-ABE8e binary vector. Lanes with no bands generated by the three pairs of primers indicate transgene-free T_1 mutants. M represents a DNA molecular weight ladder.

Corresponding author(s): Caixia Gao

Last updated by author(s): Feb 5, 2023

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. The original Image J was used to measure the angles of lamina joint inclination and the gray ratio of each band in western blot.
Data analysis	Amplicon sequencing data of prime-editing and base-editing processivity was analyzed using the published code as previously described in reference 28. Graphpad prism 8 and Microsoft Excel 2016 were used to analyze the data.

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 figures and supplementary tables, or are available from the corresponding author on request and all data supporting the findings of this study is available in a publicly accessible repository. The deep sequencing data have been deposited in a National Center for Biotechnology Information (NCBI) BioProject database (accession code PRJNA931443). Plasmids for pH-ABE8e and pH-ABE8e-spG will be available through Addgene. Source data are provided with this paper.

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	About 500,000 protoplasts were used for each transfection. The number of protoplasts in each transfection was measured by thrombocytometry. At least six lamina joint from six individual plants were used for BR sensitive experiment. At least 23 plants were used for plant height and tiller number measurement in Fig.1 and Fig.4. 10 plants were used for plant height and tiller number measurement in Fig.2. The number of protoplasts used for transfection was determined following reference 40. The sample size in BR sensitive experiment and phenotype evaluation was determined following the study in reference 33 and 34. The sample sizes used in the study could detect significant changes and produce reproducible results supporting meaningful conclusions.
Data exclusions	No data exclusion.
Replication	Quantitative RT-PCR and dual-luciferase assay were performed with three biological repeats independently. The experiments to test editing efficiencies of sgRNAs and pegRNAs with base editor or prime editor in protoplasts were performed with at least two biological repeats independently. The western blot was performed once. The experiment in rice regenerated plants was performed once. Experiment findings were reliably reproduced.
Randomization	Rice and Arabidopsis protoplasts were isolated and randomly separated to each transformation. Rice plants are randomly separated to each experiment groups.
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

Antibodies

Antibodies used	The anti-plant actin (Abclonal, Cat#AC009, 1:5000 dilution), anti-flag (Sigma, Cat#F1804, 1:3000 dilution), anti-plant BRI1 (Abmart, ZW028039S, 1:2000 dilution) and peroxidase-conjugated goat anti-mouse IgG secondary antibody (Sigma, Cat# A4416, 1:10 000 dilution) are used in this study.
Validation	The detail about anti-plant actin antibody is in https://abclonal.com/catalog-antibodies/PlantactinMonoclonalAntibody/AC009; The detail about anti-flag antibody is in https://www.sigmaaldrich.cn/CN/en/product/sigma/f1804; The detail about anti-plant BRI antibody is in http://www.ab-mart.com.cn/page.aspx?node=%2089%20&id=%2017951 and anti-plant BRI1 antibody has been used in the previous reported study (Li et al. Plant Cell, https://doi.org/10.1093/plcell/koac364, 2023)