Efficient C-to-G editing in rice using an optimized base editor

Yifu Tian1,†, Rundong Shen1,†, Zuren Li2,+, Qi Yao1, Xuening Zhang1, Dating Zhong1, Xinhang Tan1, Minglei Song1, Han Han3, Jian-Kang Zhu1,4,5,6,*, and Yuming Lu1,3,*

1 Shanghai Center for Plant Stress Biology, Center for Excellence in Molecular Plant Sciences, Chinese Academy of Sciences, Shanghai, China
2 Hunan Academy of Agricultural Sciences, Changsha, China
3School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai, China
4 Institute of Advanced Bioindustry Technologies, and Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China
5Hainan Yazhou Bay Seed Lab, Sanya, Hainan, China

*These authors contribute equally.
†These authors contribute equally.

Keywords: Cas9, base editing, rice, CGBE, UNG.

Cytosine and adenine base editors (CBE and ABE) have been vigorously developed in plants, but the base conversion types are limited (Ren et al., 2018, 2021; Xu et al., 2021). A new base editor CGBE with a uracil DNA N-glycosylase (UNG) has recently been reported that enables efficient C-to-G editing in mammalian cells (Kurt et al., 2021). In plants, generating more base substitution types can expand its application, and help create new germplasm resources (Bharat et al., 2020; Butt et al., 2020; Molla and Yang, 2019). Currently, CGBEs applicable to plants have yet to be developed. Here, we established a CGBE system in rice that enables efficient C-to-G editing.

We firstly developed a reporter system to evaluate C-to-G/A editing. As shown in Figure 1a, the nano luciferase (nLuc) reporter was inactivated by a stop-codon mutation. CBE induced C-to-T (TAG to TAA) cannot restore its function, but C-to-G/A edit would restore it to tyrosine (TAT/TAC), making it possible to evaluate the C-to-G/A activity. Accordingly, three sgRNAs were designed and the CBE editors PCBE4 was initially used for evaluation. Protoplasts-based assays showed that this conventional CBE could not mediate C-to-G/A editing, but at low efficiencies (0.11–0.14%, corresponding to 100% for the positive reporter, Figure 1b,c). This result illustrated the demand for efficiency improvement of C-to-G, and also validated the feasibility of this reporter system.

Next, we synthesized three codon-optimized UNGs from human (hUNG), E. coli (eUNG) and Mycobacterium smegmatis (UDGX), and constructed them into PCBE4 to replace the UGI and generated three CGBE vectors (Table S1). Protoplasts-based assays revealed a substantial increase in the C-to-G/A editing efficiency compared with CBE (0.49% vs 0.12%), indicating C-to-G/A editing feasibility (Figure 1c). Since UNG is a conserved class of proteins, by sequence comparison, we identified LOC_Os04g57730 as a candidate (OsUNG) to generate another CGBE, CGBE-OsUNG (Figure S1). Luciferase-based assay showed that its C-to-G/A editing efficiency was further improved. We named CGBE-OsUNG as OsCGBE01 for subsequent optimization.

Since CGBE may also depend on the efficiency of cytosine deamination, three highly active deaminases, hAID, hA3A, and Anc689, were selected to replace rAPOBEC1 in OsCGBE01. Luciferase assays showed that hAID or hA3A did not increase the efficiency, but Anc689 (named OsCGBE02, Addgene#183807) significantly improved the efficiency, with a 9.6-fold enhancement at C6. We then tested another optimization on Anc689 with R33A, which resulted in a further improvement (OsCGBE03, Addgene#183808), reaching up to 1.69% at C6, 12.1 times higher than the original PCBE4 (Figure 1d).

Next, we tested OsCGBE03 on four rice genes (OsIPA1, OsZIPS5, OsSLR1, and OsALST) to obtain stable edited plants (Figure 1f). A total of 128 T0 plants were firstly generated for OsIPA1 and OsZIPS5. DSDecode analysis on Sanger sequencing results (Liu et al., 2015) showed that 21 and 16 plants contained C-to-G and C-to-T editing, respectively. We later used Hi-TOM (Liu et al., 2019) to examine the 94 T0 plants for SLR1 and ALS1. The results showed that more than half (65%) contained either C-to-G or C-to-T conversion (reads proportion >1%) was counted, referred to as chimerism rate. To ensure heritability, we included only those plants with a chimerism rate >10% as valid edited plants. Accordingly, 24 plants (25.5%) were identified to harbour C-to-G conversions and 21 plants (22.3%) harbour C-to-T conversions. Biallelic plants with both conversions (C-to-T&G) were also frequently detected. Collectively, we achieved both C-to-G and C-to-T editing in a total of 222 plants at a high frequency (30.2%). When all base conversion types at C6 were analysed, we found that the proportion of C-to-G was dramatically increased from 3.6% to 45.9% (Figure 1e). These results demonstrated the feasibility of CGBE.

To explore its value in creating genetic diversity, we further edited ALS1 and NRT1.1B. ALS1 is responsible for herbicide resistant and sgRNA8 was designed targeting G628 (Figure 1g). Hi-Tom sequencing for the 59 OsCGBE03-edited plants showed that 31 lines (52.5%) contained C-to-G/GT, resulting in a total of 8 types of amino acid substitutions (Figure 1h). Imazethapyr treatment (0.03%) showed that the G628E and G628D lines exhibited...
C-to-G editing in rice

(a) pDual-LucM

(b) PCBE4

(c) Relative C-to-G base conversion efficiency (%)

(d) Relative C-to-G/A base conversion efficiency (%)

(e) Locus | CGBE tool | Target | Total T0 | T0 only contains Indel | Total C-to-G/T | C-to-G | C-to-T | C-to-G/A
|---------|-------------|--------|--------|----------------|--------------|--------|--------|--------|
| iPAS | OsCGBE03  | sgRNA1: ACACACAGTAGAGAATTGAGCG | 70 | 12 (17.1%) | 20 (28.6%) | 10 | 2 | 10 | 0 | 2 | 0 |
| B2IPS | osCGBE03  | sgRNA1: CCACACAGCGCTGGAAGGTG | 58 | 9 (15.5%) | 14 (24.1%) | 7 | 2 | 6 | 0 | 1 | 0 |
| SLR1  | osCGBE03  | sgRNA1: TCCAGCCTGCTCGCTGCTGACGGG | 52 | 13 (25%) | 17 (32.7%) | 5 | 7 | 8 | 4 | 3 | 4 |
| ALS1  | osCGBE03  | sgRNA1: CAGGTCCCACCGCCGATGATGG | 42 | 5 (11.9%) | 16 (38.1%) | 7 | 5 | 6 | 3 | 4 | 1 |
| ALS1  | osCGBE03  | sgRNA1: CGGGCCACCTCGGGACATCGA | 59 | 9 (15.3%) | 31 (52.5%) | 7 | 11 | 3 | 16 | 2 | 4 |

(g) ALS1

(h) Amino acid | DNA (From GGGGCG) | Propportion |
|--------------|------------------|------------|
| G52S9        | GGAGAAG, GGGGAG, GGGGAC, GGAGAC | 51.30%    |
| G52S9R       | GGAGGCG, GGAGAC | 20.70%    |
| G52S9N       | GGGGAG | 3.10%    |
| G52S1        | GGAGATC | 3.70%    |
| G52S3E       | GAGGCG | 2.90%    |
| G52S3D       | GAGGCG | 10.10%    |
| G52S3B       | GAGGCG | 3.50%    |
| G52S3A       | GACGGC | 4.70%    |

(i) WT G628E G628D

(k) ACG ATG GCG AGG AGG T327R
herbicide-resistance phenotypes (Figure 1). OsCGBE02, OsCGBE03, and an ABE with sgRNA-9 were applied to edit NRT1.1B (Figure 1). As expected, OsCGBE02 gave a higher C-to-G/T conversion frequency than OsCGBE03 at C7 (40.6% vs 9.7%), while both produced a more diverse editing pattern than ABE. Compared with the wild type, the novel variations (T327R or T327A) gave an obvious dark-green phenotype similar to that of the natural-occurring T327M allele, possibly reflecting the altered nitrogen use efficiency (Figure 1k). These case studies further demonstrated the value of CGBE in creating novel base-substitutions for plant breeding.

Taken together, our CGBE editing on five rice genes produced a total of 376 T0 plants and C-to-G conversions were successful at all loci tested, with an average frequency of 21.3%. Genetic analysis of 11 lines confirmed a heritability ranging from 4.2% to 62.5% (Table S2). Our CGBE showed greater stability and higher editing efficiency in rice than the recently reported systems (Sretenovic et al., 2021). Compared to that in mammalian cells, it gave higher indel frequencies (111 lines, 29.5%, Figure S2), indicating that our CGBE needs to be further optimized in the future by inhibiting the base excision repair process.

Author contributions
Y.L. and J.-K.Z. designed the study. Y.T., R.S., and Z.L. performed most experiments. Z.L., Q.Y., and D.Z performed rice transformation. X.T., X.Z., and M.S genotyped the transgenic lines. Y.T., R.S., and Z.L. performed rice transformation. H.H., Y.L., and J.-K.Z. wrote the manuscript.

Acknowledgement
Supported by the National Key R&D Program of China (No. 2021YFD1201300) and the National Natural Science Foundation of China (No. 32070396) to Y.L., and the CAS Strategic Priority Research Program (No. XDB27040101) to J.-K.Z.

Conflict of interests
The authors declare no competing interests.

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

Butt, H., Zaidi, S.S., Hassan, N. and Mahfouz, M. (2020) CRISPR-based directed evolution for crop improvement. Trends Biotechnol. 38, 236–240.

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

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

Liu, W., Xie, X., Ma, X., Li, J., Chen, J. and Liu, Y.G. (2015) DSDecode: a web-based tool for decoding of sequencing chromatograms for genotyping of targeted mutations. Mol. Plant. 8, 1431–1433.

Molla, K.A. and Yang, Y. (2019) CRISPR/Cas-mediated base editing: technical considerations and practical applications. Trends Biotechnol. 37, 1121–1142.

Ren, B., Yan, F., Kuan, Y., Li, N., Zhang, D., Zhou, X., Lin, H. et al. (2018) Improved base editor for efficiently inducing genetic variations in rice with CRISPR/Cas9-guided hyperactive hAID mutant. Mol. Plant. 11, 623–626.

Ren, Q., Sretenovic, S., Liu, G., Zhong, Z., Wang, J., Huang, L., Tang, X. et al. (2021) Improved plant cytosine base editors with high editing activity, purity, and specificity. Plant Biotechnol. J. 19, 2052–2068.

Sretenovic, S., Liu, S., Li, G., Cheng, Y., Fan, T., Xu, Y., Zhou, J. et al. (2021) Exploring C-To-G base editing in rice, tomato, and poplar. Front. Genome Ed. 3, 756766.

Xu, R., Kong, F., Qin, R., Li, J., Liu, X. and Wei, P. (2021) Development of an efficient plant dual cytosine and adenine editor. J. Integr. Plant Biol. 63, 1600–1605.

Supporting information
Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1 Sequence alignment for eUNG, 28 hUNG, UDGX and OsUNG.
Figure S2 OsCGBE03 induced small fragment deletions in transgenic plants.
Table S1 DNA sequences of related 39 vectors and genes.
Table S2 Heritability analysis 42 on T1 progenies.
Table S3 Oligos 44 used in this study.