Robust genome editing of CRISPR-Cas9 at NAG PAMs in rice

Xiangbing Meng[1†], Xixun Hu[2†], Qing Liu[2], Xiaoguang Song[1], Caixia Gao[3], Jiayang Li[1,4*] & Kejian Wang[2*]

[1]State Key Laboratory of Plant Genomics and National Center for Plant Gene Research (Beijing), Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China;
[2]State Key Laboratory of Rice Biology, China National Rice Research Institute, Chinese Academy of Agricultural Sciences, Hangzhou 310006, China;
[3]State Key Laboratory of Plant Cell and Chromosome Engineering, and Center for Genome Editing, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China;
[4]University of Chinese Academy of Sciences, Beijing 100049, China

Dear Editor,

The CRISPR-Cas9 (clustered regularly interspaced short palindromic repeats/Cas9) system has been widely used for a variety of applications, including targeted gene knockout, gene insertion, gene replacement and base editing. Despite its wide use, the genome editing using CRISPR-Cas9 is performed almost exclusively at sites containing canonical NGG protospacer adjacent motifs (PAMs). To overcome the PAM constraint of the CRISPR-Cas9 system, many attempts have been made to develop various Cas9 orthologs and variants with altered PAM specificities (Kleinstiver et al., 2015; Hu et al., 2016). Here we reveal that the most widely used wild-type SpCas9 is robust in recognising both NAG and NGG PAMs in rice. The NAG PAM could be chosen alone or together with NGG PAM for efficient genome editing, and have relatively low off-target effect.

For genome editing in rice, we used the CRISPR-Cas9 system with highly expressed SpCas9 protein and the structure-modified sgRNA (Hu et al., 2017)(Figure S1 in Supporting Information) to disrupt the *Bph14* gene. The genomic site with target sequence followed by NGG PAM was edited successfully. During screening potential off-target events, we found that a genomic locus with the same target sequence followed by a NAG PAM was also edited effectively (17.86%, 10 of 56 transgenic lines)(Table S1 in Supporting Information). The findings inspired us to evaluate and compare the editing efficiency of NAG and NGG PAMs in rice.

Large segments on the 11th and 12th chromosomes (Chr11 and Chr12) are the result of the most recent large-scale duplication in the rice genome and show high sequence similarity. We designed the sgRNA targeting sites in Chr11 and Chr12 with the same recognition sequence but different PAMs, NGG and NAG. Two genes, *RAD51A* and *DMC1*, were selected, each of which has two copies on Chr11 and Chr12, respectively, and two sgRNAs targeting them were designed, in which D1 targeting *RAD51A* and D2 targeting *DMC1* that allow us to perform a direct comparison of the efficiency at NGG and NAG PAMs (Figure S2 and Table S2 in Supporting Information). Through *Agrobacterium*-mediated transformation, 56 plants targeting *RAD51A1* and *RAD51A2* and 35 seedlings targeting *DMC1A* and *DMC1B* were generated. Via DNA sequencing, we found that the

---

†Contributed equally to this work. X.H., X.M., J.L., and K.W. designed the studies. X.H., X.S., and X.M. performed the experiments; Q.L. conducted the bioinformatic analyses. K.W., J.L., C.G., X.H., and X.M. wrote the manuscript.
*Corresponding authors (Kejian Wang, email: wangkejian@caas.cn; Jiayang Li, email: jyli@genetics.ac.cn)

levels of mutation efficiency at the NGG sites were 87.50% (D1) and 74.29% (D2), while those at the NAG sites were 67.86% (D1) and 34.29% (D2) (Figure S3A in Supporting Information). To further compare the efficiency of NAG and NGG PAMs, we screened the rice genome *in silico* and constructed three more sgRNAs (D3 to D5). Each sgRNA was designed to target two genomic sites with NGG and NAG PAMs located within 5 kb on the same chromosome (Table S2 in Supporting Information). By detecting 48 transgenic calluses in each individual transformation, we found that the efficiency at sites with NAG PAM was fairly high (68.75% at D3, 70.83% at D4 and 85.42% at D5) and only slightly lower than the efficiency at sites with NGG PAM (Figure S3B in Supporting Information). By calculating and comparing the efficiency in the five pairs of target sites used above, we found that the editing efficiency at sites with NAG PAM obtained from rice genetic transformation was about 76.44% of that at sites with NGG PAM, indicating that the sites containing NAG PAM, like the canonical NGG PAM, might be chosen for normal genome editing in rice.

To assess the off-target effect of SpCas9 when targeting the sites harbouring NAG PAMs, we introduced a series of mismatches at the protospacer (D1)(Table S3 in Supporting Information). As D1 has two target sites bearing NAG and NGG PAMs, we could directly compare the off-target patterns between the two PAMs. After genetic transformation, we sequenced the target sites of each transgenic callus and analysed the data. Regarding the NGG PAM, when a single-nucleotide mismatch was located at the PAM-distal region (20 to 11 nt), the editing frequency (67%–71%) was comparable to that of the original protospacer. When mismatches were generated at the PAM-proximal region (10 to 2 nt), the editing frequency was reduced (0%–52%) (Figure S4 in Supporting Information). However, when the mismatch occurred at the 3′ end of the protospacer, the editing efficiency was as high as that of the original protospacer, suggesting that SpCas9 is highly tolerant of mismatch occurring at the first base of the protospacer. For the pattern of NAG PAMs, when a single mismatch was introduced at the PAM-distal region (20 to 11 nt), the editing efficiency was significantly impaired. Moreover, when a point mutation was generated at the PAM-proximal region (10 to 2 nt) of the protospacer, the editing was completely disrupted (Figure S4 in Supporting Information). Different from NGG PAM, a single mismatch just before the PAM largely impaired the editing efficiency of NAG PAM. The results of double or triple mismatches introduced at the protospacer are basically consistent with the above conclusions, although the editing efficiency was more severely affected by multiple mismatches (Figure S4 in Supporting Information). Overall, the off-target patterns of SpCas9 at the sites containing NGG and NGA PAMs basically conform to the rule that the closer the mismatch to PAM, the lower the possibility that off-target events occur.

Furthermore, by comparing the two patterns, the genome editing at NAG PAM exhibited greater sensitivity to mismatches than that at NGG PAMs, which may be related to fit mechanism of non-canonical PAM (Jiang and Doudna, 2017).

We next investigated whether SpCas9 can be employed for genome editing at sites containing NAG PAMs; for this, we selected four genes (*NAL1*, *LPA1*, *LG1* and *GL1-1*) and designed eight sgRNAs (NAG1–2 for *NAL1*, NAG3–4 for *LPA1*, NAG5–7 for *LG1* and NAG8 for *GL1-1*) targeting the sites covering restriction enzyme sites in front of NAG PAMs (Figure S5 and Table S4 in Supporting Information). Each sgRNA-containing construct was individually co-transferred with plasmids expressing SpCas9 into rice protoplasts. By PCR/restriction enzyme (PCR/RE) assays, we detected mutations at all eight targets with editing efficiencies ranging from 0.95% (NAG5) to 8.45% (NAG2) in rice protoplasts (Figure 1A).

To test whether SpCas9 can be used for multiplex genome editing at sites harbouring NAG PAMs during stable genetic transformation, we assembled four sgRNAs with an expression cassette of SpCas9 into one binary vector (NAG1, NAG3, NAG6 and NAG8 were assembled into V1 vector; NAG2, NAG4, NAG5 and NAG7 were assembled into V2 vector) and performed genetic transformation. After genetic transformation, a total of 58 and 56 plants transformed with V1 and V2 vectors were obtained, respectively. DNA sequencing of the target sites of all plants revealed that the mutation efficiency of the sites ranged from 15.52% (NAG6) to 82.14% (NAG4) (Figure 1B and D, Figure S6 in Supporting Information). In addition, biallelic mutations were identified at all target sites, especially in NAG4 where 19 of 56 plants (33.93%) exhibited phenotypes of greater tiller angle and leaf angle due to the disruption of both copy of *LPA1* gene (Figure 1C). Moreover, 37.93% (22 of 58) of V1 and 71.43% (40 of 56 plants) of V2 harboured mutations at more than two genes (Figure 1E). The potential off-target sites of four relatively efficient sgRNAs (NAG2, NAG4, NAG7 and NAG8) were surveyed, but no off-target events were found (Table S5 in Supporting Information). The results suggest SpCas9 can be applied for efficient multiplex genome editing at sites with NAG PAMs during genetic transformation. By analysing the surrounding nucleotides around the PAMs of the eight targets (NAG1 to NAG8) (Figure 1B), we found that all of them were effectively edited and no noteworthy surrounding nucleotide preference was discovered, indicating no extra sequence requirement in the recognition of NAG PAMs in rice.

To determine whether SpCas9 can be designed for simultaneous editing at sites containing NAG and NGG PAMs, we assembled NAG2 and NAG4 with NGG1, which targets the sites of the *DL* gene with an NGG PAM, into a binary vector (V3). By genetic transformation, 54 plants
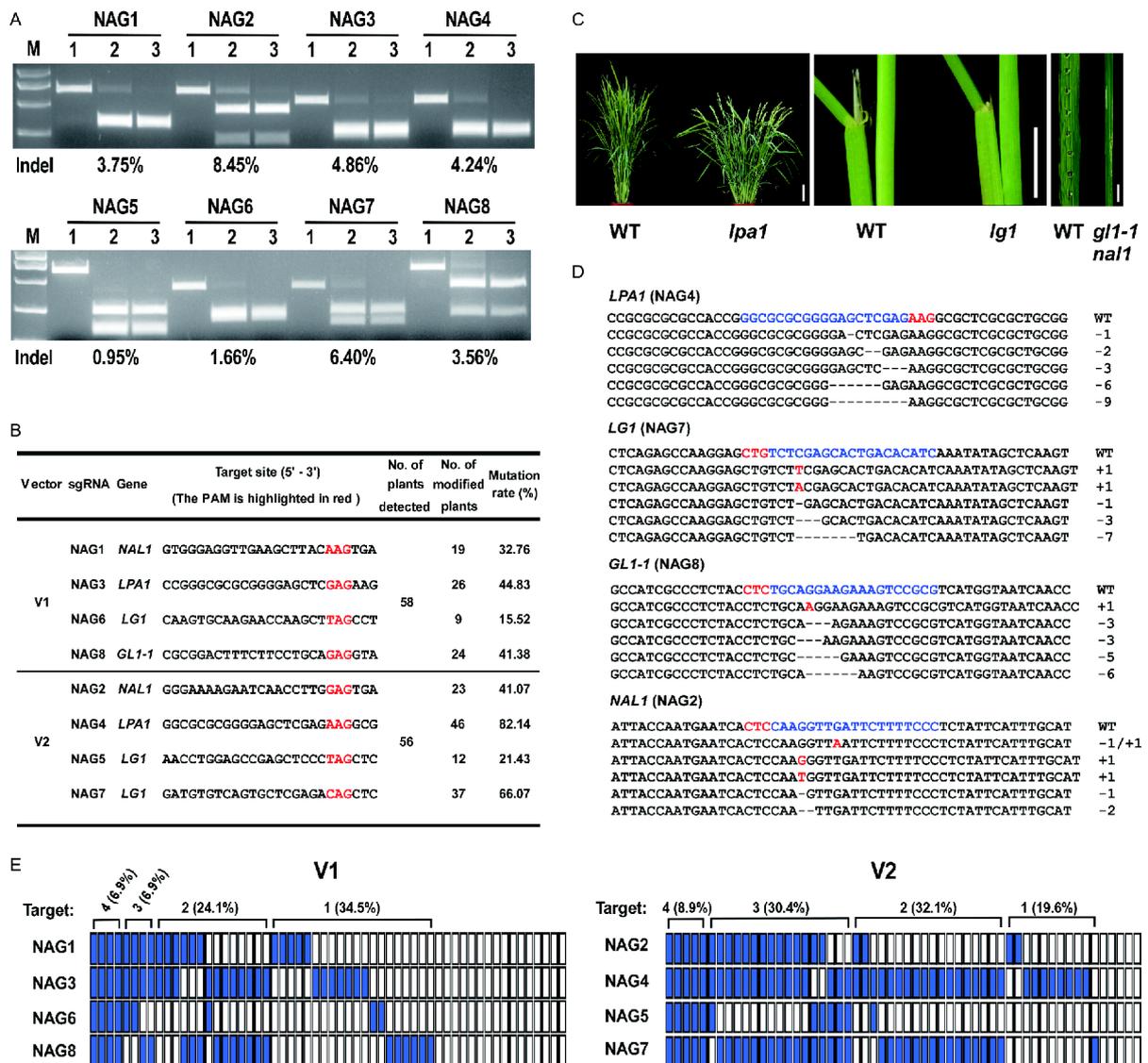
Figure 1 Genome editing at sites containing NAG PAMs in rice. A, PCR/RE assay of NAG1–NAG8 in protoplasts. Lane 1, undigested wild-type controls. Lane 2, sample treated with Cas9 and sgRNAs. The mutation frequencies measured by band intensities are listed at the bottom. Lane 3, digested wild-type controls. B, Genome editing of NAG1–NAG8 target sites in seedlings. PAMs are highlighted in red and the three bases behind PAMs are exhibited as well. C, The mutational phenotypes of related genes. The wild type (WT) is exhibited on the left. *lpa1* shows loose plant architecture. *lg1* shows a lack of auricle and ligule. *gl1-1nal1* exhibits narrow leaf and loss of cuticular wax. The bar in the image of whole plants represents 10 cm and the bars in images of local features represent 1 cm. D, The sequence of part of mutations. PAMs and targets are highlighted in red and blue, respectively. E, Multiplex editing of V1 and V2. Each column represents one seedling. Blue rectangles indicate target sites with mutations. The numbers on the top indicate the number of co-edited sites and the frequency.

were obtained and all targets were found to be mutated with high efficiency (79.63% in NGG1, 20.37% in NAG2 and 62.96% in NAG4) (Figure S7A and B in Supporting Information). A total of 23 of 54 plants (42.59%) had mutations at two genes and 11 of 54 (20.37%) seedlings bore mutations in all three genes (Figure S7C in Supporting Information). Besides, 14 double mutants (25.93%) that were modified at sites containing both NGG and NAG PAMs were identified (Figure S7D in Supporting Information). Screening for off-target mutations was also performed, but no off-target events were found (Table S5 in Supporting Information). These results confirm that SpCas9 could be applied for multiplex editing in a strategy involving mixed PAMs.

Understanding the scope of genome editing and off-target effects is critical for basic research and applications. Previous studies in bacterial and human cells revealed that SpCas9 can recognise NAG PAM, but its efficiency is only about a fifth of NGG PAM sites, which hinders its application for genome editing (Hsu et al., 2013; Jiang et al., 2013). Here we reveal that SpCas9 *per se* is robust in recognising NAG PAM in addition to the widely used NGG PAM in rice. We speculate that the high efficiency of SpCas9 at NGA PAMs might be caused by highly expressed SpCas9 protein, modified sgRNA, or merely a relatively long time of tissue

culture during rice transformation, further experiments are required to distinguish between these possibilities. In summary, the findings not only roughly double the editing scope of SpCas9 in the rice genome (Figure S8 in Supporting Information), but also demonstrate that more attention should be paid to off-target effects at NAG PAM sites when using CRISPR-Cas9 for genome editing in rice.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., Cradick, T.J., Marraffini, L.A., Bao, G., and Zhang, F. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol 31, 827–832.

Hu, X., Wang, C., Fu, Y., Liu, Q., Jiao, X., and Wang, K. (2016). Expanding the range of CRISPR/Cas9 genome editing in rice. Mol Plant 9, 943–945.

Hu, X., Meng, X., Liu, Q., Li, J., and Wang, K. (2017). Increasing the efficiency of CRISPR-Cas9-VQR precise genome editing in rice. Plant Biotechnol J, doi: 10.1111/pbi.12771.

Jiang, F., and Doudna, J.A. (2017). CRISPR–Cas9 structures and mechanisms. Annu Rev Biophys 46, 505–529.

Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L.A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol 31, 233–239.

Kleinstiver, B.P., Prew, M.S., Tsai, S.Q., Topkar, V.V., Nguyen, N.T., Zheng, Z., Gonzales, A.P.W., Li, Z., Peterson, R.T., Yeh, J.R.J., Aryee, M. J., and Joung, J.K. (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature 523, 481–485.

SUPPORTING INFORMATION

Supplementary Methods

Figure S1 The CRISPR-Cas9 system and sequence of codon-optimized Cas9 used in this study.

Figure S2 The target locations of D1 and D2.

Figure S3 The mutation efficiency of NGG and NAG PAMs.

Figure S4 The assay of mismatch tolerance within the protospacer.

Figure S5 The target locations of NAG1 to NAG8.

Figure S6 Mutational types of rice seedlings with NAG PAMs.

Figure S7 The strategy of mixed PAMs for multiple editing.

Figure S8 The target sites of NGG and NAG in rice genome.

Table S1 The off-target effects with NAG PAM during *Bph14* editing

Table S2 The location of double targets of sgRNAs

Table S3 The editing events using sgRNA with mismatches in the protospacer

Table S4 Target sites of AG1–AG8 for genome editing

Table S5 The results of off-target detection

Table S6 Primers used in the study

The supporting information is available online at http://life.scichina.com and https://link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.