1 **SpRY greatly expands the genome editing scope in rice with highly flexible PAM**

2 **recognition**

3

4 Ziyan Xu[1,2#], Yongjie Kuang[1,2#], Bin Ren[1,2,3#], Daqi Yan[1,4], Fang Yan[1], Carl Spetz[5],

5 Wenxian Sun[4,6], Guirong Wang[1,3], Xueping Zhou[1,7]* & Huanbin Zhou[1,2]*

6

7 [1]State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant

8 Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, China.

9 [2]Scientific Observing and Experimental Station of Crop Pests in Guilin, Ministry of

10 Agriculture and Rural Affairs, Guilin 541399, China.

11 [3]Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Genome

12 Analysis Laboratory of the Ministry of Agriculture, Agricultural Genomics Institute at

13 Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518120, China.

14 [4]College of Plant Protection, China Agricultural University, Beijing 100193, China.

15 [5]Norwegian Institute of Bioeconomy Research, Aas 1432, Norway.

16 [6]College of Plant Protection, Jilin Agricultural University, Changchun, 130118, China.

17 [7]State Key Laboratory of Rice Biology, Institute of Biotechnology, Zhejiang University,

18 Hangzhou 310058, China.

19

20 #These authors contributed equally to this work.

21 *Corresponding authors: Huanbin Zhou; E-mail: hbzhou@ippcaas.cn; Phone: 86-10-

22 62815914; Fax: 86-10-62894642; Xueping Zhou; E-mail: zzhou@zju.edu.cn; Phone:

23 86-10-62815906; Fax: 86-10-62894642

24

25 **Running title**: CRISPR/SpRY-mediated genome editing in rice

26   **Abstract**

27   Plant genome engineering mediated by various CRISPR-based tools requires specific

28   protospacer adjacent motifs (PAMs), such as the well-performed NGG, NG, NNG, *etc.*,

29   to initiate the target recognition, which notably restricts the editable range of the plant

30   genome. In this study, we thoroughly investigated the nuclease activity and the PAM

31   preference of two structurally-engineered SpCas9 variants (SpG and SpRY) in

32   transgenic rice. Our study shows that SpG nuclease favors NGD PAMs, albeit less

33   efficiently than the previously described SpCas9-NG and that SpRY nuclease achieves

34   efficient editing across a wide range of genomic loci, exhibiting a preference of NGD

35   as well as NAN PAMs. Furthermore, SpRY-fused cytidine deaminase hAID*Δ and

36   adenosine deaminase TadA8e were generated, respectively. These constructs efficiently

37   induced C-to-T and A-to-G conversions in the target genes toward various non-

38   canonical PAMs, including non-G PAMs. Remarkably, high-frequency self-editing

39   events (indels and DNA fragments deletion) in the integrated T-DNA fragments as a

40   result of the nuclease activity of SpRY were observed, whereas the self-editing of SpRY

41   nickase-mediated base editor was quite low in transgenic rice lines. In conclusion, the

42   broad PAM compatibility of SpRY greatly expands the targeting scope of CRISPR-

43   based tools in plant genome engineering.

44

45   **Keywords**:

46   CRISPR, SpG, SpRY, genome editing, base editing, self-targeting, *Oryza sativa* L.

47

48

Background

CRISPR technologies, which enable a variety of precise genetic modifications through targeted genome editing across nearly all living organisms, have been intensively developed in recent years and revolutionized agricultural studies. CRISPR is derived from the defense system of bacteria and archaea, protecting them from the invasion of bacteriophages and mobile genetic elements [1]. A large collection of CRISPR/Cas systems have been identified by genome sequencing and metagenome studies. Among these, the type II CRISPR/SpCas9 from *Streptococcus pyogenes* has been extensively studied and adapted for genome editing. For example, SpCas9 nuclease can be used to genetically modify a specific target locus within the plant genome. Briefly, SpCas9 catalyzes a double-stranded break (DSB) which subsequently stimulates diverse DNA repair mechanisms [non-homologous end joining (NHEJ), microhomology-mediated end joining (MMEJ), and homology-directed repairs (HDR)] resulting in gene knockout, DNA fragment insertion, deletion, and replacement as specifically required [2-6]. Alternatively, the SpCas9(D10A) nickase can be utilized to drive various engineered nucleoside deaminases to the target region and catalyze the hydrolytic deamination of cytosines to uracils and/or adenosines to inosine within the editing window. As a result, the mismatched DNA base pairs are processed through the base excision repair pathway, resulting in various base transitions and transversions: C-to-D (where D is T, G, or A), G-to-H (where H is A, C, or T), A-to-G, and T-to-C [7-13]. Recently, the SpCas9(H840A) nickase-guided Moloney murine leukemia virus reverse transcriptase (M-MLV RT) has been developed in both rice and wheat, facilitating the introduction of all 12 base-to-base conversions, small insertion as well as deletions in a precise and targeted manner through reverse transcription of the pegRNA [14-17]. Thus, diverse customized CRISPR/SpCas9 tools for precise and versatile genome editing greatly accelerate our understanding of the genetic basis of economic traits and the generation of novel germplasms for crop breeding.

It's well known that SpCas9 recognizes the canonical NGG PAM which is located immediately downstream of the target sequence [18]. Initial NGG PAM binding by SpCas9 triggers DNA strand separation locally and facilitates base pairing between

79   spacer RNA and the target DNA strand [19]. In other words, the presence of a canonical

80   NGG PAM near the target sites is critical for CRISPR/SpCas9-mediated genome

81   editing. SpCas9 and its related tools have been shown to perform efficiently in genetic

82   manipulation of a wide range of plant species [2, 10, 20-23]. However, the specific

83   PAM requirement for SpCas9 recognition strongly restricts the targetable loci of the

84   CRISPR/SpCas9-based editing tools, especially base editors because the targeted point

85   mutation needs the availability of a PAM appropriately positioned [7-9, 24].

86   To expand the genome-targeting scope of CRISPR tools, many efforts have been

87   focused on searching for new Cas variants and orthologs with novel PAM preferences.

88   For example, directed evolution and structure-guided design of SpCas9 resulted in

89   SpCas9-VRQR, xCas9, Cas9-NG variants that recognize non-canonical NGA and NG

90   PAM sites in plant [25-28]. Multiple naturally occurring SpCas9 orthologues have been

91   identified from *Streptococcus canis* (ScCas9), *Staphylococcus aureus* (SaCas9),

92   *Streptococcus thermophiles* (St1Cas9), *Brevibacillus laterosporus* (BlatCas9) and have

93   been demonstrated to edit plant genomic loci bearing NNG, NNGRRT, NNAGAAW,

94   NNNCND PAM sequence, respectively [29-32]. Also, the type V Cas12a and Cas12b

95   from diverse bacterial sources, which are distinct from Cas9, have been characterized

96   with AT-rich PAM specificity and utilized successfully in targeted plant genome editing

97   [33, 34]. Regardless of the fact that these Cas proteins with different PAM sequences

98   significantly improve the targeting range, there are still many agronomic trait-related

99   loci inaccessible for genome editing.

100   Very recently, two variants (SpG and SpRY) engineered from SpCas9-VRQR

101   through structure-guided design have been reported to recognize a wider range of PAM

102   sequences. SpG has higher genome-editing activity toward NG PAM than SpCas9-NG,

103   whereas SpRY is capable of targeting almost all PAMs (NRN>NYN) [35]. Both SpG

104   and SpRY exhibit robust activities with minimal side effects on a wide range of sites in

105   human cells [35]. Furthermore, an evolved adenosine deaminase TadA8e, which

106   catalyzes DNA deamination up to ~1100-fold faster than the previous version TadA7.10,

107   has been reported to substantially improve adenine base editing in human cells [36, 37].

108   However, whether SpG, SpRY, and TadA8e can be utilized to improve genome editing,

4

109   especially base editing, in plants remains unknown. In this study, the efficiency of both

110   SpG and SpRY nucleases toward various PAMs and its application in both cytosine and

111   adenine base editing was investigated in detail using transgenic rice callus. Our study

112   shows that SpG recognizes NG PAM sequences, but it's outperformed by SpCas9-NG

113   in rice. However, albeit with self-targeting activity on transfer T-DNA sequence, SpRY

114   surpasses SpCas9-NG since it is capable of achieving efficient cleavage by its nuclease

115   activity, cytosine base editing with hAID*Δ as well as adenine base editing with

116   TadA8e at more relaxed PAM sites (NRN, where R is G or A) in rice. Therefore, SpRY

117   and TadA8e facilitate the future design of genome editing tools in plants and broaden

118   the possible applications in agriculture and plant biology.

119

120   **Results**

121   **SpCas9-NG nuclease outperforms SpG on the minimal NG PAM in transgenic rice**

122   **lines**

123   We first compared the DNA-cleavage capability and the PAM specificity of SpG

124   nuclease to that of SpCas9-NG (which is well known for NG PAM recognition) in

125   targeted genome editing in transgenic rice. *SpCas9* carrying mutations

126   D1135L/S1136W/G1218K/E1219Q/R1335Q/T1337R was rice codon-optimized

127   (Table S1) and used to replace the SpCas9-NG gene in the binary vector pUbi:SpCas9-

128   NG [28], resulting in pUbi:SpG in which *SpG* is under the control of the maize ubiquitin

129   1 promoter (Fig. 1a). The rice U6 promoter-driven sgRNA expression cassette can be

130   shuttled into both binary vectors through Gateway recombination reaction as previously

131   described [28]. The nuclease activity of SpG and SpCas9-NG was tested side-by-side

132   with the same *sgRNA*s targeting the endogenous *OsPAL5*, *OsGSK4*, *OsCERK1*,

133   *OsETR2*, and *OsRLCK185* genes toward four types of NGN PAMs through

134   *Agrobacterium*-mediated rice transformation, respectively (Fig. S1; Fig. S2). Target

135   regions in T0 independent transgenic callus lines were PCR amplified and directly

136   subjected to Sanger sequencing.

137   Analysis of 40 independent lines for each NGG PAM site tested with SpG nuclease

138   revealed twelve lines that carried indel mutations in *OsPAL5* and four mutant lines of

139 *OsGSK4*, representing 30% and 10% editing efficiency, respectively (Fig. 1b, c). By

140 contrast, SpCas9-NG achieved 62.50% (25 out of 40 lines) editing efficiency in *OsPAL5*,

141 a 1.1-fold increase over that of SpG; and 20.51% (8 out of 39 lines) editing efficiency

142 in *OsGSK4*, a 1.1-fold increase over SpG (Fig. 1b, c). At the NGA PAM site tested,

143 SpG generated much fewer indel mutations (< 4 fold) than SpCas9-NG, 11.11% of lines

144 edited by SpG carried indels in *OsCERK1* as compared to 46.34% of the lines edited

145 by SpCas9-NG (Fig. 1b, c). At the NGT PAM sites tested, the editing efficiency of SpG

146 was 36.17% (17 out of 47 lines) in *OsGSK4* and 2.22% (1 out of 45 lines) in *OsETR2*;

147 whereas the editing efficiency of SpCas9-NG was higher, being 52.63% and 5.26% in

148 *OsGSK4* and in *OsETR2*, respectively (Fig. 1b, c). In the case of the NGC PAM, which

149 has been reported to be less efficiently recognized by SpCas9-NG [28, 38], SpG edited

150 *OsCERK1* and *OsRLCK185* at the corresponding sites with comparable efficiency to

151 SpCas9-NG (Fig. 1b, c). It should be noted that both di-allelic and mono-allelic

152 mutations were identified in T0 mutant lines generated by SpG toward all PAM

153 sequences, being the ratios of mono-allelic edits relatively higher (Fig. 1b, c; Fig. S1;

154 Fig. S2). Overall, these results indicate that the newly-engineered SpG nuclease is

155 capable of recognizing minimal NG PAM in rice, albeit less efficiently than SpCas9-

156 NG. Consequently, the activity of SpG in plants is different from the previously

157 reported SpG-mediated genome editing efficiency in human cells [35].

158

159 **SpRY nuclease preferentially recognizes NAN and NGN PAM sequences in**

160 **transgenic rice lines**

161 Five other point mutations consisting of A61R, L1111R, N1317R, A1322R, and

162 R1333P were further introduced into SpG to generate SpRY (Table S1; Fig S3a).

163 Following the same experimental procedure mentioned above, the nuclease activity and

164 PAM preference of SpRY were thoroughly investigated in transgenic rice calli. Thirty-

165 two endogenous genomic sites bearing all 16 possible alternative PAM sequences,

166 which vary at the second and the third positions, were used (Fig. S3-S16). Each

167 transformation construct used harbored two *sgRNAs* targeting the same gene or

168 different genes to the extent that multiplex genome editing and/or DNA fragment

6

169   deletion could be achieved and investigated in rice cells.

170   Genotyping of all the T0 transgenic lines for the NGN PAM sites showed that

171   SpRY induces indel mutations predominantly at the NGG PAM site (63.83% editing),

172   followed by the NGA (> 74% editing) and NGT (> 47% editing) PAM sites. The NGC

173   PAM site is the least recognized, achieving only a 2.13% editing efficiency (Fig. 2a, b).

174   Thus, SpRY nuclease favors the recognition of NGD > NGC PAM in rice. On the NAN

175   PAM sites tested, the editing efficiency of SpRY was highly variable, but it reached as

176   high as 20.00% for a NAG PAM site in *OsMPK8*; 33.33% for a NAA site in *OsMPK9*;

177   51.28% for a NAT site in *OsMPK10*; and 29.79% for a NAC site in *OsCPK2* (Fig. 2a,

178   b), suggesting that SpRY moderately recognizes NAN PAM sequences in rice, albeit in

179   a locus-dependent manner. Across the rest 16 genomic sites with NYN PAMs, SpRY

180   exhibited editing at 4 sites (25%), with high activity on a NTC PAM site in *OsCPK28*

181   (73.91% efficiency) and weak activity on a NCG site in *OsCPK20* (2.94% editing), a

182   NCT site in *OsMPK3* (4.26% editing), and a NCC site in *OsMPK4* (7.69% editing) (Fig.

183   2a, b). It should be mentioned that screening of all target genes revealed that the

184   majority of edited lines contained mono-allelic mutations at the target sites, and that

185   only 2 out of 47 independent lines (4.26% ratio) had genomic DNA fragment deletions

186   in *OsCPK2* (Fig. 2a; Fig S3-S16). Combining all the data (Fig. 2b), we conclude that

187   the structurally-engineered SpRY is capable of recognizing NRN and some NYN PAMs

188   in rice, but with impaired nuclease activity as compared to that of SpCas9 reported

189   previously [2].

190   The broad PAM compatibility of SpRY nuclease implies that SpRY might be

191   similar to SpCas9-NG which was reported to self-edit the single-guide RNA region of

192   the transferred T-DNA in transgenic rice cells [39]. Therefore, the identities of the

193   *sgRNA* transgenes in some *SpRY*-transgenic lines were examined by direct PCR

194   sequencing. For the high-efficiency *OsCPK5-sgRNA1*, *OsCPK1-sgRNA1*, and

195   *OsCPK1-sgRNA2*, no more than 4.26% lines were detected with modified *sgRNAs* (Fig.

196   2a, b). On the other hand, mutation frequencies at the other high-efficiency *sgRNA* sites

197   (such as *OsCPK3-sgRNA1*, *OsCPK3-sgRNA2*, *OsMPK10-sgRNA1*, *OsCPK28-sgRNA1*)

198   were much higher, ranging from 23.08% to 100% (Fig. 2a, b). Also, varying self-editing

199 frequencies (0-80.65%) were observed in the low-efficiency *sgRNA*s (i.e. *OsCPK5-*

200 *sgRNA2*, *OsCPK14-sgRNA1*, *OsCPK14-sgRNA2*, *OsCPK27-sgRNA1*) (Fig. 2a, b). It

201 should be mentioned that on-target editing and self-editing of SpRY were detected

202 simultaneously or alone in the individual transgenic line, and DNA fragment deletion

203 events between two *sgRNAs* in the T-DNA regions were detected in targeting *OsCPK4*

204 and *OsCPK28*, respectively (Fig. S6; Fig. S13). Combining all the data, we conclude

205 that SpRY elicits complex self-editing events that might occur before and after gene

206 editing of on-targets in transgenic rice.

207 A previous study has shown that utilizing a GCCC-type sgRNA scaffold alleviates

208 the self-targeting property of SpCas9-NG without substantial loss of on-target activity

209 in rice [39]. Therefore, we replaced the original GTTT-type sgRNA in our

210 CRISPR/SpRY system with the GCCC-type sgRNA scaffold (Table S1). Subsequently,

211 we tested the nuclease activity of our modified CRISPR/SpRY system in transgenic rice

212 callus using the *sgRNAs* with high self-editing frequencies. To our surprise, replacing

213 the GTTT-type scaffold with the GCCC-type scaffold did not show any improvement

214 in on-target editing (Fig. 2a). On the contrary, the editing efficiency of *OsCPK3*,

215 *OsMPK10*, and *OsCPK28* was significantly reduced (Fig. 2a). We further investigated

216 the identities of each sgRNA transgene in all lines and found that, out of 12 sgRNAs

217 examined, eleven lines carried mutations at high frequencies (Fig. 2a). Based on our

218 data, we conclude that the GCCC-type sgRNA scaffold elicits stronger self-editing

219 events in sgRNA transgenes, which results in a decrease in the efficiency of on-target

220 editing events in rice, as compared to the naturally occurring (original) GTTT-type

221 sgRNA scaffold.

222

223 **SpRY mediates cytosine base editing at non-canonical PAM sites in the rice**

224 **genome**

225 PAM preferences of Cas proteins largely limit the targeting scope of base editors [7-9,

226 28, 29]. Considering the greatly relaxed PAM compatibility of SpRY, we tested whether

227 it could improve cytosine base editing in transgenic rice when fusing to the hyperactive

228 cytidine deaminase hAID*Δ. The SpRY nuclease was mutated into the SpRY(D10A)

229   nickase and employed to replace the *SpCas9n* gene in the previously-reported cytosine

230   base editor rBE9 which enables C-to-D editing [7], resulting in rBE66 (Fig. 3a).

231   The base-editing activity of rBE66 was then investigated in transgenic rice callus

232   with three sgRNAs targeting the endogenous *OsCOI2* gene toward non-canonical NGC,

233   NGT, and NAG PAMs, respectively. Intriguingly, of the 50 independent transgenic lines

234   confirmed by sequencing, thirteen lines were identified with nucleotide mutations at

235   the NGC PAM site which is presumably less efficient for SpRY targeting (Fig 3b, e, h).

236   Among these thirteen lines, twelve lines (24% efficiency) carried nucleotide

237   substitutions (predominantly the C-to-T conversion) in the typical editing window

238   whereas six lines (12% frequency) carried indels (Fig 3e). These indels likely result

239   from both the deaminase activity of hAID*Δ and the nickase activity of SpRY based on

240   the position of the indels. For the other two PAMs tested in *OsCOI2*, while the NGT

241   PAM was inefficient, the 1 bp-shifted NAG PAM resulted in a base editing efficiency

242   of 34.15% (14 out of 41 lines) and indel frequency of 4.88% (2 out of 41 lines) (Fig 3c,

243   f, h). rBE66 was also tested with two sgRNAs targeting the endogenous *BSR-K1* gene

244   at NGA and NAG PAM sites, respectively. After genotyping the transgenic lines

245   obtained for both transformation constructs, we only identified a single heterozygous

246   line (2.13%) with base editing events and a single heterozygous line (2.13%) carrying

247   indel at the NAG PAM site (Fig 3d, g, h). Finally, the self-editing effect of rBE66 was

248   also investigated. We identified 4 lines (9.76% ratio) and 2 lines (4.26% ratio) carrying

249   nucleotide substitutions in *OsCOI2-sgRNA* (NAG PAM) and *BSR-K1-sgRNA* (NAG

250   PAM) transgenes, respectively (Fig. 3h, Fig. S17a). These data, combined, indicate that

251   SpRY is compatible with hAID*Δ, greatly expanding the targeting scope of hAID*Δ-

252   mediated cytosine base editor in rice. Thus, SpRY has very promising prospects in

253   targeted base editing in plants due to the broad PAM compatibility and negligible self-

254   editing effect.

255

256   **SpRY-fused TadA8e monomer enables efficient adenine base editing at non-G**

257   **PAM sites in the rice genome**

258   The CRISPR/Cas-guided adenine deaminase heterodimer TadA:TadA7.10 is capable of

9

259   introducing A-to-G conversion in rice [9, 28, 29]. Very recently, an evolved version

260   TadA8e with a more robust enzymatic activity has been characterized in human cells

261   [36, 37]. Therefore, we investigated whether SpRY could improve adenine base editing

262   in   rice   by   fusing   it   to   TadA8e.   *TadA7.10*   carrying   mutations

263   A109S/T111R/D119N/H122N/Y147D/F149Y/T166I/D167N   was   rice   codon-

264   optimized (Table S1) and fused to the 5'-terminus of *SpRY(D10A)*. The chimeric gene

265   was utilized to replace the previously-reported adenine base editor gene *rBE14* in the

266   binary vector pUbi:rBE14 [9], resulting in pUbi:rBE62 (Fig. 4a).

267   The endogenous *OsMPK13* and *OsGS1* gene were utilized to test rBE62's activity

268   toward the NAA and NAT PAM in transgenic rice calli, respectively. Strikingly, 14 out

269   of 47 lines for the NAA PAM and 45 out of 48 lines for the NAT PAM were identified

270   with a pure A-to-G conversion at the target sites, representing 29.79% and 93.75%

271   adenine editing efficiency, respectively (Fig 4b-f). All mutant lines carried mono-allelic

272   mutations and no indel mutations were detected (Fig 4b, c). Furthermore, we genotyped

273   the *sgRNA* transgenes in all lines to evaluate the self-targeting effect of rBE62. Only

274   one self-edited line (2.13% ratio) in which the *sgRNA* transgene carried A-to-G

275   conversion was detected (Fig S17b). Nevertheless, no modification was observed in the

276   target region since the entire *OsMPK13* was intact. These data, combined, indicate that

277   the adenine base editor rBE62 (engineered with SpRY and TadA8e monomer as

278   mentioned above) is capable of efficiently inducing A-to-G conversions in a wide

279   editable range of the rice genome, including non-G PAM sites. To our knowledge, this

280   is the first time a base editor which recognized non-G PAM sites in plants is described.

281

282   **Discussion**

283   PAM preference is the key limitation to each CRISPR-based tool for targeted genome

284   editing, especially for base editing since it requires the precise positioning of Cas

285   proteins at a given site. Therefore, reducing or eliminating the PAM requirement of Cas

286   proteins will substantially advance various CRISPR technologies. To date, several other

287   Cas proteins with altered or relaxed PAM specificities besides SpCas9, including

288   SpCas9-VQR (NGA), SpCas9-VRER (NGA), SpCas9-NG (NG), ScCas9 (NNG), and

289   SaKKH-Cas9 (NNNRRT), have been successfully adopted in both cytosine and/or

290   adenine base editing in plants [8, 28, 29, 38]. The latest reported SpCas9 variants, SpG

291   and SpRY with relaxed PAM specificity [35], might further optimize plant CRISPR

292   tools for targeted base editing. However, different from the original report that SpG

293   nuclease shows high-activity editing across all NGN PAM sites in human cells [35], we

294   observed a different PAM preference (NGN/NAN/NTN>NCN) in transgenic rice lines.

295   Moreover, our data also shows that the editing efficiency of SpG is lower than that of

296   SpCas9-NG in transgenic rice lines. Therefore, we believe that SpCas9-NG is still the

297   best genome editing player for NGN PAM sites in plants.

298   SpRY exhibits a similar preference for NGN PAM sequences as SpG. However,

299   SpRY nuclease is also capable of processing many NAN, NTC, NCT, NCC, NCG PAM

300   sites with varying efficiency, which have previously been inaccessible with CRISPR

301   tools. SpRY thereby greatly increases the number of targeting sites previously not

302   available for genome editing in rice. However, the broadened PAM compatibility of

303   SpRY nuclease can also result in complex and unpredictable self-targeting of sgRNA

304   transgenes as shown in our study, which presumably affects the on-target editing at both

305   early- and late-transformation-stage. Nevertheless, we show that the naturally occurring

306   GTTT-type of guide RNA scaffold outperforms the modified GCCC-type [39] when

307   using SpRY to edit some endogenous genomic sites, and we do not recommend SpRY

308   for multiple gene knocking out since it causes loss of sgRNA transgenes in some cases.

309   More GTN- and GCN-type sgRNA scaffolds in self-editing and the heritability of

310   edited genes in generations should be investigated in the future.

311   The SpRY nickase was compatible with both hAID*Δ and TadA8e deaminases

312   and achieved efficient cytosine and adenine base editing at various non-canonical PAM

313   sites, especially non-G PAM sites. Intriguingly, compared to the high-frequency self-

314   editing caused by the SpRY nuclease in transgenic rice, significantly low ratios of self-

315   editing of *sgRNA* transgenes was detected in our base editing assay, in which mutations

316   are the result of the activity of the TadA8e deaminase and not the SpRY nickase. In our

317   study, we also show that the newly evolved TadA8e, as reported in human cells [36],

318   enables efficient A-to-G conversion in rice as well. Therefore, both SpRY and Tad8e

11

319　can be presumably utilized to multiplex base editing in rice and further optimize other

320　plant base editors in terms of targeting scope and editing efficiency.

321　　　The broadened PAM compatibility of SpRY raises the possibility of increased off-

322　target editing across all plant genomes. To address this question, the fidelity-enhancing

323　substitutions of SpCas9-HF1 could be introduced into the SpRY nuclease to alleviate

324　its off-target effect in rice as reported in human cells [35]. Alternatively, careful design

325　of the sgRNA is another way to avoid off-target cleavage. For example, off-target

326　mutations were detected only for *sgRNAs* targeting *OsCPK3* and *OsCPK28* in our study

327　(Fig. S18). Regarding the SpRY nickase-mediated base editing with specific sgRNAs,

328　hAID*Δ and TadA8e-dependent off-target base editing are the major issues, which

329　might be minimized through protein engineering in the future. Nevertheless, timely

330　isolation of the transgene-free, gene-edited lines in T1 progenies to prevent continual

331　editing is highly recommended, and backcrossing it with the recipient material to purify

332　the genetic background can be carried out if needed.

333

334　Conclusions

335　Overall, this article is the first report describing efficient base editing at minimal NA

336　PAM sites in the plant genome. The structurally-engineered SpRY efficiently

337　recognizes NAN in addition to NGN PAM sequences, greatly expanding the targetable

338　range of the rice genome. With the new SpRY-based tools added into the toolbox, plant

339　biologists can easily carry out high-resolution genome editing now, and we recommend

340　using SpCas9 for target sites carrying NGG PAMs, SpCas9-NG for NGH PAMs (where

341　H is A, C, or T), ScCas9 for NHG PAMs, and SpRY for NAH PAMs.

342

343　Methods

344　Rice materials

345　Rice cultivars Kitaake (*Oryza sativa L*. ssp. *Geng*) was used in this study and kept in

346　our lab. Rice plants were cultivated in the paddy field under natural conditions in

347　normal rice growing seasons and immature seeds were harvested from expanded

348　panicles for use in rice transformation.

349

Vector design and plasmid construction

The coding region of *SpG* gene [35] flanked by two nuclear localization signal (*NLS*) sequences was rice codon-optimized, the 5' region (*SpG-fg1*), and the 3' region (*SpG-fg2*) were separately synthesized by Tsingke (Beijing) (Table S1). The full-length *SpG* was assembled using the overlapping-extension PCR-based method with the high-fidelity DNA polymerase Phusion (NEB) and the primers listed in Table S3, and directly cloned downstream of the maize ubiquitin 1 promoter and upstream of the *NOS* terminator by replacing the *Cas9-NG* gene in pUbi:Cas9-NG [28] through BamHI/SpeI digestion and DNA ligation, resulting in the binary vector pUbi:SpG.

For the construction of the binary vector of SpRY nuclease, a point mutation of GCG to AGG (A61R) was first introduced into *SpG-fg1* by PCR amplification of the whole plasmid pUC57:SpG-fg1 with the primer pairs SpRY-F1/SpRY-R1. The amplicon was self-ligated and confirmed by sequencing, resulting in pUC57:SpRY-fg1. Next, the 3' region of *SpRY* [35] fused with an *NLS* sequence was rice codon-optimized (*SpRY-fg2*) and subjected to synthesizing by Tsingke (Beijing) (Table S1). Finally, both *SpRY-fg1* and *SpRY-fg2* were assembled and pUbi:SpRY was constructed in the same manner as mentioned for *SpG*.

For the construction of cytosine adenine base editors, the gene fusion of *hAID*Δ*, *SpRY*, and *UGI* fragment was carried out with a simple PCR-based cloning strategy. Briefly, the full-length *SpRY* gene was PCR amplified using the primer pairs OsCas9-Fg1-F1/OsCas9-Fg1-R1 and pUbi:SpRY as the template, and ligated with the backbone of pUC19:rBE9 [7] carrying both *hAID*Δ* and *UGI* fragment, which was amplified with the primer pairs UGI-F1/rAPO-R1, resulting in pUC19:rBE66. For the construction of adenine base editors, the *TadA8e* fragment was codon-optimized and subjected to synthesizing by Tsingke (Beijing) (Table S1), and fused to the 5' end of *SpRY* using the overlapping-extension PCR-based method with the primers listed in Table S3, resulting in *rBE62* fragment. Finally, rBE66 and rBE62 were cloned into the binary vector by BamHI/SpeI digestion as described above, resulting in pUbi:rBE66 for cytosine base editing and pUbi:rBE62 for adenine base editing in rice, respectively.

13

379   For the construction of guide RNA gene, the GCCC-type sgRNA [39] was

380   subjected to synthesizing by Tsingke (Beijing) (Table S1), amplified with the primer

381   pairs gRNA4-NG-F2/gRNA4-NG-R2, and used to replace the GTTT-type sgRNA in

382   pENTR4:sgRNA4 using In-Fusion HD Enzyme Premix (Clontech), resulting in the

383   entry vector pENTR4:sgRNA4-NG.

384   For the construction of the final T-DNA transformation plasmids, the 20-bp

385   complementary oligos (Table S2), corresponding to each target site (Table S3) and

386   carrying appropriate 4-bp adaptor, were phosphorylated, annealed, and inserted into the

387   BsaI-digested pENTR4:sgRNA4 or pENTR4:sgRNA4-NG. Each sgRNA expression

388   cassette was then shuffled into the respective binary vector by LR clonase (Invitrogen).

389   All primers used here are listed in Table S3.

390   The identities of *SpG*, *SpRY*, *rBE66*, and *rBE62* in vectors and oligo insertions in

391   each construct were confirmed by Sanger sequencing.

392

Agrobacterium-mediated rice transformation

394   T-DNA transformation plasmids harboring gene-targeting *sgRNA* were transferred into

395   the *A. tumefaciens* strain EHA105 by electroporation. Rice callus derived from

396   immature seeds of Kitaake were used for stable transformation following a protocol as

397   described previously [24].

398

Genotyping of transgenic rice lines

400   Genomic DNA was isolated from each T0 transgenic callus line using the

401   hexadecyltrimethylammonium bromide (CTAB) method [40]. PCR amplification of the

402   targeted genomic region was carried out with specific primers listed in Table S3, and

403   the PCR products were subjected to Sanger sequencing. Besides, a selection of PCR

404   products with low quality of sequencing data was cloned into the pEASY-Blunt cloning

405   vector (TransGen Biotech), and colonies with insertion were randomly chosen for

406   Sanger sequencing.

407

Acknowledgements

411

Authors' contributions

413   HZ, XZ, GW, CS, SW, and BR designed the research; ZX, YK, BR, DY, and FY

414   conducted the experiments and analyzed the data; HZ, XZ, and GW supervised the

415   research; HZ, XZ, and CS wrote the original draft; all authors participated in discussion

416   and revision of the manuscript.

417

425

Ethics approval and consent to participate

427   Not applicable.

428

Competing interests

430   The authors have filed a patent application based on the results reported in this study.

431

Author details

433   [1]State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant

434   Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, China.

435   [2]Scientific Observing and Experimental Station of Crop Pests in Guilin, Ministry of

436   Agriculture and Rural Affairs, Guilin 541399, China. [3]Shenzhen Branch, Guangdong

437   Laboratory for Lingnan Modern Agriculture, Genome Analysis Laboratory of the

438   Ministry of Agriculture, Agricultural Genomics Institute at Shenzhen, Chinese

439   Academy of Agricultural Sciences, Shenzhen 518120, China. [4]College of Plant

440   Protection, China Agricultural University, Beijing 100193, China. [5]Norwegian Institute

441   of Bioeconomy Research, Aas 1432, Norway. [6]College of Plant Protection, Jilin

442   Agricultural University, Changchun, 130118, China. [7]State Key Laboratory of Rice

443   Biology, Institute of Biotechnology, Zhejiang University, Hangzhou 310058, China.

444

445   **References**

446   1.   Hille F, Richter H, Wong SP, Bratovic M, Ressel S, Charpentier E: **The biology**

447   **of CRISPR-Cas: backward and forward.** *Cell* 2018, **172:**1239-1259.

448   2.   Zhou H, Liu B, Weeks DP, Spalding MH, Yang B: **Large chromosomal**

449   **deletions and heritable small genetic changes induced by CRISPR/Cas9 in**

450   **rice.** *Nucleic Acids Res* 2014, **42:**10903-10914.

451   3.   Li J, Meng X, Zong Y, Chen K, Zhang H, Liu J, Li J, Gao C: **Gene replacements**

452   **and insertions in rice by intron targeting using CRISPR-Cas9.** *Nat Plants*

453   2016, **2:**16139.

454   4.   Lu Y, Tian Y, Shen R, Yao Q, Wang M, Chen M, Dong J, Zhang T, Li F, Lei M,

455   Zhu J-K: **Targeted, efficient sequence insertion and replacement in rice.** *Nat*

456   *Biotechnol* 2020, DOI: 10.1038/s41587-020-0581-5.

457   5.   Tan J, Zhao Y, Wang B, Hao Y, Wang Y, Li Y, Luo W, Zong W, Li G, Chen S, et

458   al: **Efficient CRISPR/Cas9-based plant genomic fragment deletions by**

459   **microhomology-mediated end joining.** *Plant Biotechnol J* 2020, DOI:

460   10.1111/pbi.13390.

461   6.   Sun Y, Zhang X, Wu C, He Y, Ma Y, Hou H, Guo X, Du W, Zhao Y, Xia L:

462   **Engineering herbicide-resistant rice plants through CRISPR/Cas9-**

463   **mediated homologous recombination of acetolactate synthase.** *Mol Plant*

464   2016, **9:**628-631.

465   7.   Ren B, Yan F, Kuang Y, Li N, Zhang D, Zhou X, Lin H, Zhou H: **Improved**

466   **base editor for efficiently inducing genetic variations in rice with**

467   **CRISPR/Cas9-guided hyperactive hAID mutant.** *Mol Plant* 2018, **11:**623-

468   626.

8.   Ren B, Yan F, Kuang Y, Li N, Zhang D, Lin H, Zhou H: **A CRISPR/Cas9 toolkit for efficient targeted base editing to induce genetic variations in rice.** *Sci China Life Sci* 2017, **60:**516-519.

9.   Yan F, Kuang Y, Ren B, Wang J, Zhang D, Lin H, Yang B, Zhou X, Zhou H: **Highly efficient A.T to G.C base editing by Cas9n-guided tRNA adenosine deaminase in rice.** *Mol Plant* 2018, **11:**631-634.

10.   Li C, Zong Y, Wang Y, Jin S, Zhang D, Song Q, Zhang R, Gao C: **Expanded base editing in rice and wheat using a Cas9-adenosine deaminase fusion.** *Genome Biol* 2018, **19:**59.

11.   Li C, Zong Y, Jin S, Zhu H, Lin D, Li S, Qiu J-L, Wang Y, Gao C: **SWISS: multiplexed orthogonal genome editing in plants with a Cas9 nickase and engineered CRISPR RNA scaffolds.** *Genome Biol* 2020, **21:**141.

12.   Zhao D, Li J, Li S, Xin X, Hu M, Price MA, Rosser SJ, Bi C, Zhang X: **Glycosylase base editors enable C-to-A and C-to-G base changes.** *Nat Biotechnol* 2020, DOI: 10.1038/s41587-020-0592-2.

13.   Kurt IC, Zhou R, Iyer S, Garcia SP, Miller BR, Langner LM, Grunewald J, Joung JK: **CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells.** *Nat Biotechnol* 2020, DOI: 10.1038/s41587-020-0609-x.

14.   Lin Q, Zong Y, Xue C, Wang S, Jin S, Zhu Z, Wang Y, Anzalone AV, Raguram A, Doman JL, et al: **Prime genome editing in rice and wheat.** *Nat Biotechnol* 2020, **38:**582-585.

15.   Li H, Li J, Chen J, Yan L, Xia L: **Precise modifications of both exogenous and endogenous genes in rice by prime editing.** *Mol Plant* 2020, **13:**671-674.

16.   Tang X, Sretenovic S, Ren Q, Jia X, Li M, Fan T, Yin D, Xiang S, Guo Y, Liu L, et al: **Plant prime editors enable precise gene editing in rice cells.** *Mol Plant* 2020, **13:**667-670.

17.   Xu R, Li J, Liu X, Shan T, Qin R, Wei P: **Development of plant prime-editing systems for precise genome editing.** *Plant Commun* 2020, **1:**100043.

18.   Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y,

499 Fine EJ, Wu X, Shalem O, et al: **DNA targeting specificity of RNA-guided**
500 **Cas9 nucleases.** *Nat Biotechnol* 2013, **31:**827-832.

501 19. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA: **DNA**
502 **interrogation by the CRISPR RNA-guided endonuclease Cas9.** *Nature* 2014,
503 **507:**62-67.

504 20. Li Z, Liu ZB, Xing A, Moon BP, Koellhoffer JP, Huang L, Ward RT, Clifton E,
505 Falco SC, Cigan AM: **Cas9-guide RNA directed genome editing in soybean.**
506 *Plant Physiol* 2015, **169:**960-970.

507 21. Lawrenson T, Shorinola O, Stacey N, Li CD, Ostergaard L, Patron N, Uauy C,
508 Harwood W: **Induction of targeted, heritable mutations in barley and**
509 ***Brassica oleracea* using RNA-guided Cas9 nuclease.** *Genome Biology* 2015,
510 **16**:258.

511 22. Barone P, Wu E, Lenderts B, Anand A, Gordon-Kamm W, Svitashev S, Kumar
512 S: **Efficient gene targeting in maize using inducible CRISPR-Cas9 and**
513 **marker-free donor template.** *Mol Plant* 2020,13:1219-1227.

514 23. Gosavi G, Yan F, Ren B, Kuang Y, Yan D, Zhou X, Zhou H: **Applications of**
515 **CRISPR technology in studying plant-pathogen interactions: overview and**
516 **perspective.** *Phytopathol Res* 2020, **2:**21.

517 24. Kuang Y, Li S, Ren B, Yan F, Spetz C, Li X, Zhou X, Zhou H: **Base-editing-**
518 **mediated artificial evolution of *OsALS1 in planta* to develop novel**
519 **herbicide-tolerant rice germplasms.** *Mol Plant* 2020, **13:**565-572.

520 25. Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, Gonzales
521 AP, Li Z, Peterson RT, Yeh JR, et al: **Engineered CRISPR-Cas9 nucleases**
522 **with altered PAM specificities.** *Nature* 2015, **523:**481-485.

523 26. Hu JH, Miller SM, Geurts MH, Tang W, Chen L, Sun N, Zeina CM, Gao X,
524 Rees HA, Lin Z, Liu DR: **Evolved Cas9 variants with broad PAM**
525 **compatibility and high DNA specificity.** *Nature* 2018, **556:**57-63.

526 27. Nishimasu H, Shi X, Ishiguro S, Gao L, Hirano S, Okazaki S, Noda T,
527 Abudayyeh OO, Gootenberg JS, Mori H, et al: **Engineered CRISPR-Cas9**
528 **nuclease with expanded targeting space.** *Science* 2018, **361:**1259-1262.

529   28.   Ren B, Liu L, Li S, Kuang Y, Wang J, Zhang D, Zhou X, Lin H, Zhou H: **Cas9-**
530   **NG greatly expands the targeting scope of the genome-editing toolkit by**
531   **recognizing NG and other atypical PAMs in rice.** *Mol Plant* 2019, **12:**1015-
532   1026.

533   29.   Wang M, Xu Z, Gosavi G, Ren B, Cao Y, Kuang Y, Zhou C, Spetz C, Yan F,
534   Zhou X, Zhou H: **Targeted base editing in rice with CRISPR/ScCas9 system.**
535   *Plant Biotechnol J* 2020, **18:**1645-1647.

536   30.   Hua K, Tao X, Zhu JK: **Expanding the base editing scope in rice by using**
537   **Cas9 variants.** *Plant Biotechnol J* 2019, **17:**499-504.

538   31.   Steinert J, Schiml S, Fauser F, Puchta H: **Highly efficient heritable plant**
539   **genome engineering using Cas9 orthologues from *Streptococcus***
540   ***thermophilus* and *Staphylococcus aureus*.** *Plant J* 2015, **84:**1295-1305.

541   32.   Karvelis T, Gasiunas G, Young J, Bigelyte G, Silanskas A, Cigan M, Siksnys V:
542   **Rapid characterization of CRISPR-Cas9 protospacer adjacent motif**
543   **sequence elements.** *Genome Biol* 2015, **16:**253.

544   33.   Tang X, Lowder LG, Zhang T, Malzahn AA, Zheng X, Voytas DF, Zhong Z,
545   Chen Y, Ren Q, Li Q, et al: **A CRISPR-Cpf1 system for efficient genome**
546   **editing and transcriptional repression in plants.** *Nat Plants* 2017, **3:**17103.

547   34.   Ming M, Ren Q, Pan C, He Y, Zhang Y, Liu S, Zhong Z, Wang J, Malzahn AA,
548   Wu J, et al: **CRISPR–Cas12b enables efficient plant genome engineering.**
549   *Nat Plants* 2020, **6:**202-208.

550   35.   Walton RT, Christie KA, Whittaker MN, Kleinstiver BP: **Unconstrained**
551   **genome targeting with near-PAMless engineered CRISPR-Cas9 variants.**
552   *Science* 2020, **368:**290-296.

553   36.   Richter MF, Zhao KT, Eton E, Lapinaite A, Newby GA, Thuronyi BW, Wilson
554   C, Koblan LW, Zeng J, Bauer DE, et al: **Phage-assisted evolution of an**
555   **adenine base editor with improved Cas domain compatibility and activity.**
556   *Nat Biotechnol* 2020, **38:**883-891.

557   37.   Lapinaite A, Knott GJ, Palumbo CM, Lin-Shiao E, Richter MF, Zhao KT, Beal
558   PA, Liu DR, Doudna JA: **DNA capture by a CRISPR-Cas9–guided adenine**

19

559   **base editor.** *Science* 2020, **369:**566-571.

560   38.   Hua K, Tao XP, Zhu JK: **Expanding the base editing scope in rice by using**

561   **Cas9 variants.** *Plant Biotechnol J* 2019, **17:**499-504.

562   39.   Qin R, Li J, Liu X, Xu R, Yang J, Wei P: **SpCas9-NG self-targets the sgRNA**

563   **sequence in plant genome editing.** *Nat Plants* 2020, **6:**197-201.

564   40.   Porebski S, Bailey LG, Baum BR: **Modification of a CTAB DNA extraction**

565   **protocol for plants containing high polysaccharide and polyphenol**

566   **components.** *Plant Mol Biol Report* 1997, **15:**8-15.
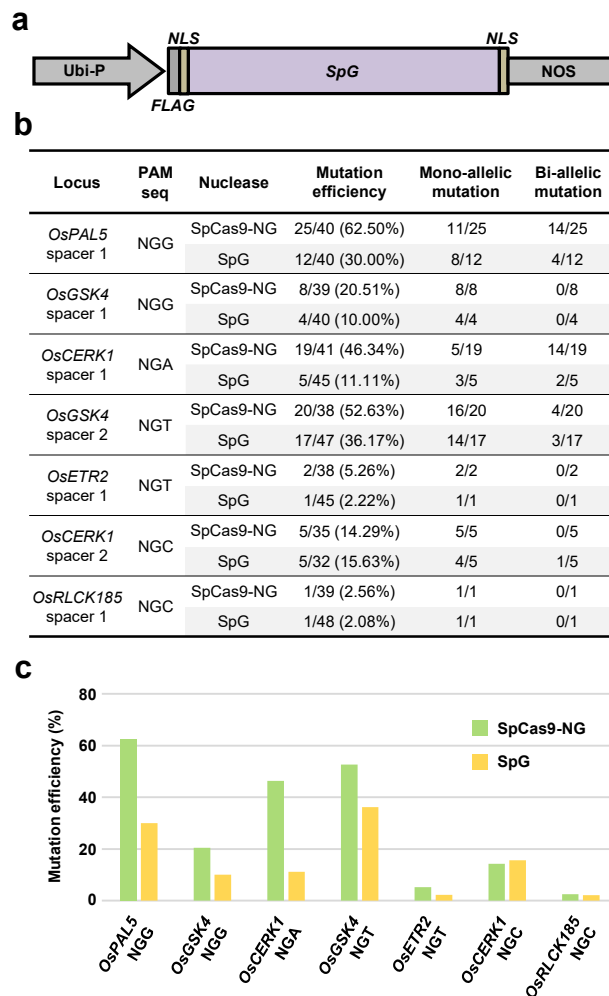
567

568

Figure legends

Fig. 1 Analysis of SpG nuclease activity on different NGN PAMs in transgenic rice. **a** The gene construct of SpG nuclease used for genome editing in transgenic rice. Ubi-P, maize ubiquitin 1 promoter; *NLS*, nuclear localization sequence. **b** Summary of the genotyping results on T0 transgenic rice lines generated with SpG and SpCas9-NG nuclease. **c** Comparison of the mutation frequency by SpG and SpCas9-NG nucleases toward 4 NGN PAMs in T0 transgenic rice lines.

Fig. 2 Analysis of SpRY nuclease activity on 16 possible NNN PAMs in transgenic rice. **a** Summary of the genotyping results on T0 transgenic rice lines generated with SpRY nuclease using different PAM sequences and sgRNA scaffolds. NA, not available. **b** Comparison of the on-target mutation efficiency and the self-editing frequency of *sgRNA* transgenes by SpRY nuclease toward different PAMs in T0 transgenic rice lines.

Fig. 3 Efficient cytosine base editing mediated by SpRY-fused hAID*Δ toward non-canonical PAMs in rice. **a** The gene construct of *rBE66* used for cytosine base editing in transgenic rice. Ubi-P, maize ubiquitin 1 promoter; *hAID*Δ*, truncated version of hyperactive human *AID* gene; *UGI*, uracil glycosylase inhibitor; *NLS*, nuclear localization sequence. **b**, **c**, and **d** Representative edited alleles of *OsCOI2* toward a NGC PAM (**b**) and a NAG PAM (**c**), and *BSR-K1* toward a NAG PAM (**d**) generated by rBE66 in T0 transgenic rice lines. The PAM sequences for each target site are shown in the pale purple arrows; the candidate bases in the putative editing window for editing and detected nucleotide changes are highlighted in red and green, respectively; the nucleotide substitutions are underlined in the sequencing chromatograms. **e**, **f** and **g** Summary of nucleotide changes in the editing window of *OsCOI2* toward a NGC PAM (**e**) and a NAG PAM (**f**), and *BSR-K1* toward a NAG PAM (**g**) caused by rBE66 in T0 transgenic lines. The PAM sequences and the detected nucleotide changes are highlighted in green and red, respectively. **h** Summary of the genotyping results on T0 transgenic rice lines generated with rBE66. NA, not available.

598   **Fig. 4** Efficient adenine base editing mediated by SpRY-fused TadA8e monomer toward

599   non-G PAMs in rice. **a** The gene construct of *rBE62* used for adenine base editing in

600   transgenic rice. Ubi-P, maize ubiquitin 1 promoter; *TadA8e*, the evolved *E.coli*

601   *TadA7.10* gene; *XTEN,* 16 amino acid flexible linker; *NLS*, nuclear localization

602   sequence. **b** and **c** Representative edited alleles of *OsMPK13* (**b**) and *OsGS1* (**c**)

603   generated by rBE62 in T0 transgenic rice lines. The PAM sequences for each target site

604   are shown in the pale purple arrows; the candidate bases in the putative editing window

605   for editing and detected nucleotide changes are highlighted in red and green,

606   respectively; the nucleotide substitutions are underlined in the sequencing

607   chromatograms. **d** and **e** Summary of nucleotide changes in the editing window of

608   *OsMPK13* (**d**) and *OsGS1* (**e**) caused by rBE62 in T0 transgenic rice lines. The PAM

609   sequences and the detected nucleotide changes are highlighted in green and red,

610   respectively. **f** Summary of the genotyping results on T0 transgenic rice lines generated

611   with rBE62. The PAM sequences and the target regions are highlighted in green and

612   bold, respectively.

613

a

b

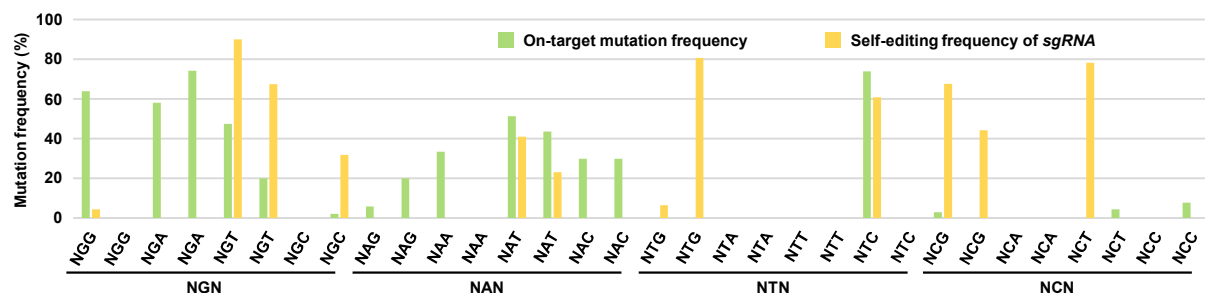| Locus	PAM seq	Nuclease	Mutation efficiency	Mono-allelic mutation	Bi-allelic mutation
OsPAL5 spacer 1	NGG	SpCas9-NG	25/40 (62.50%)	11/25	14/25
		SpG	12/40 (30.00%)	8/12	4/12
OsGSK4 spacer 1	NGG	SpCas9-NG	8/39 (20.51%)	8/8	0/8
		SpG	4/40 (10.00%)	4/4	0/4
OsCERK1 spacer 1	NGA	SpCas9-NG	19/41 (46.34%)	5/19	14/19
		SpG	5/45 (11.11%)	3/5	2/5
OsGSK4 spacer 2	NGT	SpCas9-NG	20/38 (52.63%)	16/20	4/20
		SpG	17/47 (36.17%)	14/17	3/17
OsETR2 spacer 1	NGT	SpCas9-NG	2/38 (5.26%)	2/2	0/2
		SpG	1/45 (2.22%)	1/1	0/1
OsCERK1 spacer 2	NGC	SpCas9-NG	5/35 (14.29%)	5/5	0/5
		SpG	5/32 (15.63%)	4/5	1/5
OsRLCK185 spacer 1	NGC	SpCas9-NG	1/39 (2.56%)	1/1	0/1
		SpG	1/48 (2.08%)	1/1	0/1

c

Fig. 1 Analysis of SpG nuclease activity on different NGN PAMs in transgenic rice. **a** The gene construct of SpG nuclease used for genome editing in transgenic rice. Ubi-P, maize ubiquitin 1 promoter; *NLS*, nuclear localization sequence. **b** Summary of the genotyping results on T0 transgenic rice lines generated with SpG and SpCas9-NG nuclease. **c** Comparison of the mutation frequency by SpG and SpCas9-NG nucleases toward 4 NGN PAMs in T0 transgenic rice lines.
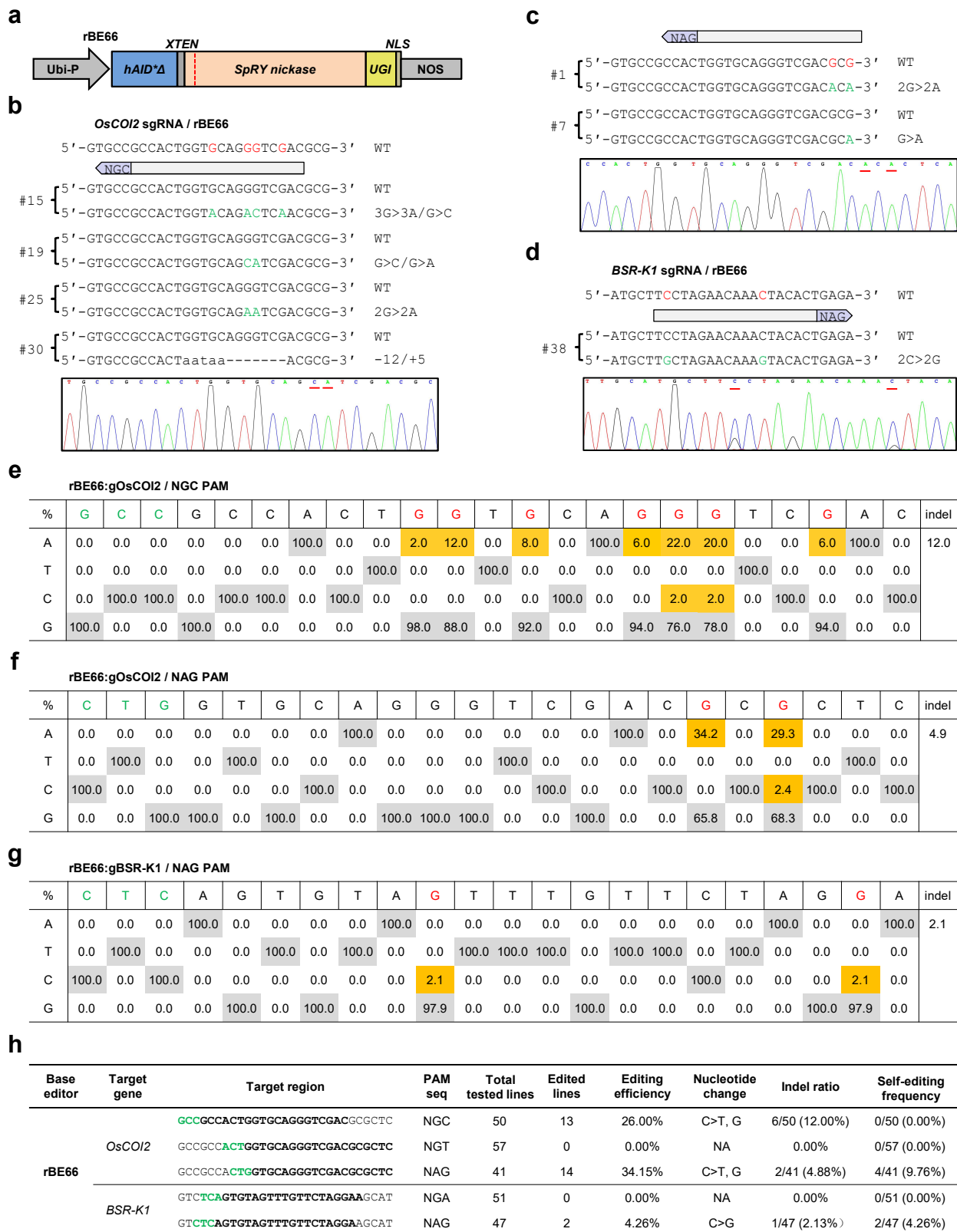
a

Construct	sgRNA type	Spacer	PAM seq	Mutation efficiency	Mono-allelic mutation	Di-allelic mutation	Self-editing frequency	Spacer distance	Deletion ratio
OsCPK5	GTTT	1	NGG	30/47 (63.83%)	25/30	5/30	2/47 (4.26%)	216 bp	0
		2	NGG	0/47 (0.00%)	0	0	0/47 (0.00%)		
OsCPK1	GTTT	1	NGA	18/31 (58.06%)	18/18	0	0/31 (0.00%)	207 bp	0
		2	NGA	23/31 (74.19%)	23/23	0	0/31 (0.00%)		
OsCPK3	GTTT	1	NGT	19/40 (47.50%)	19/19	0	36/40 (90.00%)	600 bp	0
		2	NGT	8/40 (20.00%)	8/8	0	27/40 (67.50%)		
	GCCC	1	NGT	5/30 (16.67%}	5/5	0	30/30 (100.00%)		0
		2	NGT	1/30 (3.33%)	1/1	0	30/30 (100.00%)		
OsCPK4	GTTT	1	NGC	0/41 (0.00%)	0	0	0/41 (0.00%)	243 bp	0
		2	NGC	1/41 (2.13%)	1/1	0	13/41 (31.71%)		
	GCCC	1	NGC	0/42 (0.00%)	0	0	30/42 (71.43%)		0
		2	NGC	0/42 (0.00%)	0	0	42/42 (100.00%)		
OsMPK8	GTTT	1	NAG	2/35 (5.71%)	2/2	0	NA	203 bp	0
		2	NAG	7/35 (20.00%)	7/7	0	NA		
OsMPK9	GTTT	1	NAA	17/51 (33.33%)	11/17	6/17	NA	5,251 bp	0
		2	NAA	0/51 (0.00%)	0	0	NA		
OsMPK10	GTTT	1	NAT	20/39 (51.28%)	20/20	0	16/39 (41.03%)	3,760 bp	0
		2	NAT	17/39 (43.58%)	17/17	0	9/39 (23.08%)		
	GCCC	1	NAT	1/40 (2.50%)	1/1	0	33/40 (82.50%)		0
		2	NAT	6/40 (15.00%)	6/6	0	32/40 (80.00%)		
OsCPK2	GTTT	1	NAC	14/47 (29.79%)	11/14	3/14	NA	378 bp	2/47(4.26%)
		2	NAC	14/47 (29.79%)	12/14	2/14	NA		
OsCPK14	GTTT	1	NTG	0/31 (0.00%)	0	0	2/31 (6.45%)	318 bp	0
		2	NTG	0/31 (0.00%)	0	0	25/31 (80.65%)		
	GCCC	1	NTG	0/38 (0.00%)	0	0	27/38 (71.05%)		0
		2	NTG	0/38 (0.00%)	0	0	0/38 (0.00%)		
OsCPK8	GTTT	1	NTA	0/34 (0.00%)	0	0	NA	181 bp	0
		2	NTA	0/34 (0.00%)	0	0	NA		
OsCPK27	GTTT	1	NTT	0/31 (0.00%)	0	0	0/31 (0.00%)	327 bp	0
		2	NTT	0/31 (0.00%)	0	0	0/31 (0.00%)		
OsCPK20	GTTT	1	NCG	1/34 (2.94%)	1/1	0	23/34 (67.65%)	849 bp	0
		2	NCG	0/34 (0.00%)	0	0	15/34 (44.12%)		
	GCCC	1	NCG	0/38 (0.00%)	0	0	34/38 (89.47%)		0
		2	NCG	0/38 (0.00%)	0	0	32/38 (84.21%)		
OsCPK21	GTTT	1	NCA	0/40 (0.00%)	0	0	NA	251 bp	0
		2	NCA	0/40 (0.00%)	0	0	NA		
OsCPK28	GTTT	1	NTC	34/46 (73.91%)	30/34	4/34	28/46 (60.87%)	172 bp	0
		2	NCT	0/46 (0.00%)	0	0	36/46 (78.26%)		
	GCCC	1	NTC	3/37 (7.50%)	3/3	0	37/37 (100.00%)		0
		2	NCT	0/37 (0.00%)	0	0	36/37 (97.30%)		
OsMPK3 /OsMPK4	GTTT	1	NCT	2/47 (4.26%)	2/2	0	NA	NA	0
		2	NCC	0/47 (0.00%)	0	0	NA		
OsMPK3 /OsMPK4	GTTT	1	NTC	0/52 (0.00%)	0	0	NA	NA	0
		2	NCC	4/52 (7.69%)	4/4	0	NA		

b

Fig. 2 Analysis of SpRY nuclease activity on 16 possible NNN PAMs in transgenic rice. **a** Summary of the genotyping results on T0 transgenic rice lines generated with SpRY nuclease using different PAM sequences and sgRNA scaffolds. NA, not available. **b** Comparison of the on-target mutation efficiency and the self-editing frequency of *sgRNA* transgenes by SpRY nuclease toward different PAMs in T0 transgenic rice lines.

a

rBE66

Ubi-P — hAID*Δ — XTEN — SpRY nickase — UGI (NLS) — NOS

b

OsCOI2 sgRNA / rBE66

```
5'-GTGCCGCCACTGGTGCAGGGTCGACGCG-3'   WT
         NGC

#15  5'-GTGCCGCCACTGGTGCAGGGTCGACGCG-3'   WT
     5'-GTGCCGCCACTGGTACAGACTCAACGCG-3'   3G>3A/G>C

#19  5'-GTGCCGCCACTGGTGCAGGGTCGACGCG-3'   WT
     5'-GTGCCGCCACTGGTGCAGCATCGACGCG-3'   G>C/G>A

#25  5'-GTGCCGCCACTGGTGCAGGGTCGACGCG-3'   WT
     5'-GTGCCGCCACTGGTGCAGAATCGACGCG-3'   2G>2A

#30  5'-GTGCCGCCACTGGTGCAGGGTCGACGCG-3'   WT
     5'-GTGCCGCCACTaataa-------ACGCG-3'   -12/+5
```

c

```
                                    NAG

#1  5'-GTGCCGCCACTGGTGCAGGGTCGACGCG-3'   WT
    5'-GTGCCGCCACTGGTGCAGGGTCGACACA-3'   2G>2A

#7  5'-GTGCCGCCACTGGTGCAGGGTCGACGCG-3'   WT
    5'-GTGCCGCCACTGGTGCAGGGTCGACGCA-3'   G>A
```

d

BSR-K1 sgRNA / rBE66

```
5'-ATGCTTCCTAGAACAAACTACACTGAGA-3'   WT
                              NAG

#38  5'-ATGCTTCCTAGAACAAACTACACTGAGA-3'   WT
     5'-ATGCTTGCTAGAACAAAGTACACTGAGA-3'   2C>2G
```

e

rBE66:gOsCOI2 / NGC PAM

%	G	C	C	G	C	C	A	C	T	G	G	T	G	C	A	G	G	G	T	C	G	A	C	indel
A	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	2.0	12.0	0.0	8.0	0.0	100.0	6.0	22.0	20.0	0.0	0.0	6.0	100.0	0.0	12.0
T	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	
C	0.0	100.0	100.0	0.0	100.0	100.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	2.0	2.0	0.0	100.0	0.0	0.0	100.0	
G	100.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	98.0	88.0	0.0	92.0	0.0	0.0	94.0	76.0	78.0	0.0	0.0	94.0	0.0	0.0	

f

rBE66:gOsCOI2 / NAG PAM

%	C	T	G	G	T	G	C	A	G	G	G	T	C	G	A	C	G	C	G	C	T	C	indel
A	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	34.2	0.0	29.3	0.0	0.0	0.0	4.9
T	0.0	100.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	
C	100.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	100.0	0.0	100.0	2.4	100.0	0.0	100.0	
G	0.0	0.0	100.0	100.0	0.0	100.0	0.0	0.0	100.0	100.0	100.0	0.0	0.0	100.0	0.0	0.0	65.8	0.0	68.3	0.0	0.0	0.0	

g

rBE66:gBSR-K1 / NAG PAM

%	C	T	C	A	G	T	G	T	A	G	T	T	T	G	T	T	C	T	A	G	G	A	indel
A	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	100.0	2.1
T	0.0	100.0	0.0	0.0	0.0	100.0	0.0	100.0	0.0	0.0	100.0	100.0	100.0	0.0	100.0	100.0	0.0	100.0	0.0	0.0	0.0	0.0	
C	100.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	2.1	0.0	
G	0.0	0.0	0.0	0.0	100.0	0.0	100.0	0.0	0.0	97.9	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	100.0	97.9	0.0	

h

Base editor	Target gene	Target region	PAM seq	Total tested lines	Edited lines	Editing efficiency	Nucleotide change	Indel ratio	Self-editing frequency
rBE66	*OsCOI2*	GCCGCCACTGGTGCAGGGTCGACGCGCTC	NGC	50	13	26.00%	C>T, G	6/50 (12.00%)	0/50 (0.00%)
		GCCGCCACTGGTGCAGGGTCGACGCGCTC	NGT	57	0	0.00%	NA	0.00%	0/57 (0.00%)
		GCCGCCACTGGTGCAGGGTCGACGCGCTC	NAG	41	14	34.15%	C>T, G	2/41 (4.88%)	4/41 (9.76%)
	BSR-K1	GTCTCAGTGTAGTTTGTTCTAGGAAGCAT	NGA	51	0	0.00%	NA	0.00%	0/51 (0.00%)
		GTCTCAGTGTAGTTTGTTCTAGGAAGCAT	NAG	47	2	4.26%	C>G	1/47 (2.13%)	2/47 (4.26%)

Fig. 3 Efficient cytosine base editing mediated by SpRY-fused hAID*Δ toward non-canonical PAMs in rice. **a** The gene construct of *rBE66* used for cytosine base editing in transgenic rice. Ubi-P, maize ubiquitin 1 promoter; *hAID*Δ*, truncated version of hyperactive human *AID* gene; *UGI*, uracil glycosylase inhibitor; *NLS*, nuclear localization sequence. **b**, **c** and **d** Representative edited alleles of *OsCOI2* toward a NGC PAM (**b**) and a NAG PAM (**c**), and *BSR-K1* toward a NAG PAM (**d**) generated by rBE66 in T0 transgenic rice lines. The PAM sequences for each target site are shown in the pale purple arrows; the candidate bases in the putative editing window for editing and detected nucleotide changes are highlighted in red and green, respectively; the nucleotide substitutions are underlined in the sequencing chromatograms. **e**, **f** and **g** Summary of nucleotide changes in the editing window of *OsCOI2* toward a NGC PAM (**e**) and a NAG PAM (**f**), and *BSR-K1* toward a NAG PAM (**g**) caused by rBE66 in T0 transgenic lines. The PAM sequences and the detected nucleotide changes are highlighted in green and red, respectively. **h** Summary of the genotyping results on T0 transgenic rice lines generated with rBE66. NA, not available.

a rBE62

b OsMPK13 sgRNA / rBE62

c OsGS1 sgRNA / rBE62

d rBE62:gOsMPK13 / NAA PAM

%	G	G	A	C	A	T	G	G	A	G	T	T	C	T	T	T	A	C	G	G	A	A	indel
A	0.0	0.0	100.0	0.0	70.2	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	100.0	100.0	0.0
T	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	100.0	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	
C	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	
G	100.0	100.0	0.0	0.0	29.8	0.0	100.0	100.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0.0	

e rBE62:gOsGS1 / NAT PAM

%	A	C	A	C	C	A	A	C	T	A	C	A	G	G	T	G	A	G	G	G	A	T	indel
A	100.0	0.0	8.3	0.0	0.0	54.2	100.0	0.0	0.0	100.0	0.0	100.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	100.0	0.0	0.0
T	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	
C	0.0	100.0	0.0	100.0	100.0	0.0	0.0	100.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
G	0.0	0.0	91.7	0.0	0.0	45.8	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	100.0	0.0	100.0	100.0	100.0	0.0	0.0	

f

Base editor	Target gene	PAM seq	Total tested lines	Edited lines	Editing efficiency	Nucleotide change	Indel ratio	Self-editing frequency
rBE62	OsMPK13	NAA	47	14	29.79%	A>G	0.00%	1/47 (2.13%)
	OsGS1	NAT	48	45	93.75%	A>G	0.00%	0/48 (0.00%)

Fig. 4 Efficient adenine base editing mediated by SpRY-fused TadA8e monomer toward non-G PAMs in rice. **a** The gene construct of *rBE62* used for adenine base editing in transgenic rice. Ubi-P, maize ubiquitin 1 promoter; *TadA8e*, the evolved *E.coli TadA7.10* gene; *XTEN*, 16 amino acid flexible linker; *NLS*, nuclear localization sequence. **b** and **c** Representative edited alleles of *OsMPK13* (**b**) and *OsGS1* (**c**) generated by rBE62 in T0 transgenic rice lines. The PAM sequences for each target site are shown in the pale purple arrows; the candidate bases in the putative editing window for editing and detected nucleotide changes are highlighted in red and green, respectively; the nucleotide substitutions are underlined in the sequencing chromatograms. **d** and **e** Summary of nucleotide changes in the editing window of *OsMPK13* (**d**) and *OsGS1* (**e**) caused by rBE62 in T0 transgenic rice lines. The PAM sequences and the detected nucleotide changes are highlighted in green and red, respectively. **f** Summary of the genotyping results on T0 transgenic rice lines generated with rBE62. The PAM sequences and the target regions are highlighted in green and bold, respectively.