Characterization of Substitution Mutations of eIF4G Gene Generated through Adenine Base Editors in Rice

Yaiphabi Kumam1, Rajadurai G2, Kumar K K1, Varanavasiappan S3, Ravendran M1, Manonmani S2, Gopalakrishnan C3, Arul L1, Kokiladevi E1 and Sudhakar D1*

1Department of Plant Biotechnology, Centre for Plant Molecular Biology & Biotechnology, Tamil Nadu Agricultural University, Coimbatore-641 003
2Department of Plant Genetic Resources, Centre for Plant Breeding & Genetics, Tamil Nadu Agricultural University, Coimbatore-641 003
3Department of Rice, Centre for Plant Breeding & Genetics, Tamil Nadu Agricultural University, Coimbatore-641 003

ABSTRACT

Adenine base editor (ABE) creates A to G transitions within its editing window. In the present study, an ABE was used to target a stretch of six amino acid residues, VLFPNL in translation initiation factor four gamma (eIF4G) gene of rice. Agrobacterium-mediated transformation of rice cultivar ASD16 resulted in T0 events with high mutation efficiency of 89.29%. Substitution mutations of A > G occurred within the editing window of four to eight bases at A7 > G10 (74.67 %) and A7 > G10 (2.46 %). Non-canonical substitutions of G > C/A were also observed at G9 > C15 (9.29 %) and G9 > A11 (1.15 %). A total of 15 missense base substitution events affecting the target residue was identified. Taken together, the present study showed that ABEs create unexpected base substitutions besides efficient canonical editing of A > G in the rice genome.

Keywords: Rice; Adenine base editors; eIF4G

INTRODUCTION

Genome editing by CRISPR/Cas technology has never ceased to evolve, making it the most attractive tool of the 21st century. It has proved to be remarkably successful in creating targeted insertion and deletion mutants across species (Feng et al., 2013; Yang et al., 2014; Kang et al., 2016; Sun et al., 2017; Roberts and Barrangou, 2020). These indel mutations result from double-stranded breaks at the target site that are predominantly repaired by non-homologous end joining (Maruyama et al., 2015; Kosicki et al., 2013). The repair, albeit precise, is error-prone, with no control over the number of nucleotide bases being deleted or inserted, leading to unanticipated mutants (Manghwar et al., 2019; Maruyama et al., 2015; Piergentili et al., 2021). More recently, base-editing technology, comprising of cytosine base editors (CBEs) and adenine base editors (ABEs) have refined CRISPR/Cas9 by performing irreversible substitutions of C-G to T-A (Komor et al., 2013) and A-T to G-C (Gaudelli et al., 2017) base pairs respectively. These substitutions take place exclusively within the editing window and thus are highly precise in targeting human cells (Komor et al., 2013), mouse cells (Gaudelli et al., 2017), bacterial cells (Zheng et al., 2018) and crop species (Qin et al., 2020; Veillet et al., 2020; Wang et al., 2020; Wu et al., 2020).

Both CBEs and ABEs have made their significant contributions in editing essential genes that govern the agronomic performance in rice, including nutritional improvement (Li et al., 2018), plant architecture and grain yield (Zong et al., 2017; Hua et al., 2018, 2019) and high nitrogen use efficiency (Lu and Zhu, 2017). In addition to these, they have been used to target genes responsible for imparting herbicide resistance (Shimatani et al., 2017, Li et al., 2018) and blast resistance (Ren et al., 2018) in rice. Of these two systems of base editors, CBEs have been reported to have a higher off-target mutation that arises from cytosine deaminases coupled with low editing efficiency. Thus, ABE is the preferred choice for gene editing in rice (Hao et al., 2019; Jin et al., 2019).

ABE7.10, an adenine base editor, is widely used for editing A·T to G·C base pair with high fidelity within an activity window of 4 - 8 bp of the sgRNA sequence (with reference to the protospacer adjacent motif (PAM) at positions 21 to 23) (Kim et al., 2019). Li et al. (2018) modified this vector to yield maximum A to G editing by using the construct pH-PABE-7-essgRNA in rice and thus achieved high editing efficiency, up to 59.1 % in a japonica variety, Zhonghua 11. Earlier reports on mutations in translation initiation factor four gamma (eIF4G) gene, leading to substitution or in-frame deletions at
amino acid residues viz., Y^{1059} V^{1060} have been shown to confer resistance against rice tungro spherical virus (RTSV) in naturally occurring RTSV resistant genotypes (Lee et al., 2010). Subsequent in-frame mutations upstream of the YV residue spanning across SVLFPNLAGKS (amino acid positions 1046 to 1058) and especially N^{1051} L^{1052} imparted resistance against RTSV (Macovei et al., 2018). Thus, in the present study, we attempted to create targeted novel modifications in VLFNPNL residues in an elite tungro susceptible indica cultivar, ASD16 using the ABE construct, pH-PABE-7-esgRNA.

MATERIAL AND METHODS

**Design and development of plant adenine base editing vector**

A single guide RNA (sgRNA) in translation initiation factor 4 gamma (elf4G) gene of indica variety (BG109GA25931; Ensemble Plants), targeting the VLFNPNL residue was designed using the Chop-Chop tool (Labun et al., 2016). To the sgRNA sequences (Top strand: 5’-ACAAAATGGGAACAAAGACGC-3’ and Bottom strand: 5’-GGCTTCGTGCTCCCTAATTGT-3’), BsaI adaptors (5’ GGCG in the top strand and 5’ AAAC in the bottom strand) were added for DNA oligomer synthesis (Eurofins, Bengaluru). The synthesized DNA oligomers were duplexed and cloned into the BsaI restriction site of the binary vector, pH-PABE7-esgRNA, a gift from Dr. Caixia Gao (Addgene plasmid # 115620; http://n2t.net/addgene:115620; RRID: Addgene_115620) (Li et al., 2018). This clone was mobilized into Agrobacterium strain, LBA4404. T-DNA of the binary plant expression vector, harboring the sgRNA is represented in Fig. 1.

**Agrobacterium-mediated transformation of ASD16**

ASD16, an elite medium duration indica rice variety that is widely grown in Tamil Nadu was chosen as a target genotype. Immature embryos were used for Agrobacterium-mediated transformation of ASD16 (Hiei and Komari, 2008) with the construct based on the plant ABE vector. Friable yellow calli that survived two rounds of stringent selection in 50 mgL⁻¹ hygromycin antibiotic successfully regenerated into complete plants. The regenerated plants were hardened and maintained in transgenic greenhouse. The transformation efficiency (in percentage) of the construct used was calculated using the formula,

\[
\text{Transformation efficiency} = \frac{\text{Number of co-cultivated embryos that produced plants}}{\text{Total number of embryos co-cultivated}} \times 100
\]

**Molecular characterization of putative T₀ mutants**

Plant genomic DNA from young leaves of putative mutants and ASD16 wild type were isolated using the CTAB method (Porebski et al., 1997). Molecular analyses by PCR for T-DNA presence using sequence-specific primers for hpt (hygromycin phosphotransferase) and cas9 genes (Table 1) were performed. The target region encompassing the sgRNA sequence in the PCR positive mutants was amplified using elf4G gene-specific primers (Table 1). The PCR amplicons were purified (Nucleospin Gel and PCR Purification Kit, Machery Nagel) and sequenced using Sanger sequencing method (Eurofins, Bengaluru).

**Figure 1. T-DNA of binary vector pH-PABE-7-esgRNA with sgRNA**

Results obtained from sequencing were analyzed using web-based softwares, DSDecodeM (http://skl.scau.edu.cn/dsdecode/) (Xie et al., 2017; Liu et al., 2015) to decode substitutions in the target region in both alleles of the gene and CRISPR-ID (http://crispid.gbiomed.kuleuven.be/) (Dehairs et al., 2016) to identify the localization of the mutation along the entire length of the amplified sequence. Besides these softwares, the percentage contribution of a base in substituting its target base was calculated by using an online web-based tool, EditR (http://baseeditr.com) (Kluesner et al., 2018). The corresponding protein sequences were translated using an online translation tool, Expasy (https://www.expasy.org). Mutants were identified from analysis of the sequencing results and mutation efficiency (in percentage) was calculated as given below,

\[
\text{Mutation efficiency} = \frac{\text{Number of events with mutations}}{\text{Total number of events generated}} \times 100
\]

**RESULTS AND DISCUSSION**

**Agrobacterium-mediated transformation of ASD16**

Agrobacterium-mediated transformation of ASD16 was performed using the Agrobacterium strain, LBA4404 harboring pH-PABE7-esgRNA-elf4G construct. Thirteen batches of co-cultivation, consisting of 1391 immature embryos were performed. A total of 112 independent events was generated, giving transformation efficiency of 8.20 % (Table 2).

**Characterization of mutants generated**

Molecular analysis by PCR for the presence of cas9 and hpt genes in the 112 independent events confirmed that all the mutants were positive for these genes (Fig. 2a and 2b). Sanger sequencing analysis of these PCR positive events identified 100
mutants out of 112, thus giving a high mutation efficiency of 89.29% (Table 2). However, the majority of the mutants, comprising of 85 events had silent mutations, and only 15 had missense mutations. These 15 missense mutants had base substitutions affecting the SVLFPNLAGKS residues. Earlier reports by Macovei et al. (2018) suggested that mutations affecting the above stretch of 11 amino acid residues can impart resistance against tungro disease. Hence, the 15 missense mutants were the promising outcome of the experiment (Table 2). A > G substitutions occurred at two positions of the sgRNA sequence, A7 and A4, where 74.68% and 2.4% of adenosines were converted to guanosines respectively (Fig. 3 & 4). In addition to this, A7 > G was observed in homozygous conditions in 42 T0 independent events, while only mono-allelic substitution was observed at A4. The mechanism attributing to this exceptionally biased preference of base substitution of A7 yield large number of homozygous mutants in T0 generation remains unclear. However, this may partly be attributed to the sequences that are present immediately adjacent to the adenine residue, as sgRNAs from different genomic loci respond distinctively to adenine base editing (Li et al., 2018). The sgRNA used in the present study has A7 succeeded by three Gs viz., G8G9G10. This suggests the possible influence of flanking bases on the performance of the base editors.

Table 1. Primers and PCR conditions used in the study

<table>
<thead>
<tr>
<th>Name of gene</th>
<th>Forward (F) and reverse (R) primers (5’ to 3’)</th>
<th>Amplicon size (in bp)</th>
<th>PCR conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>hpt</td>
<td>hpt F: GCTGTTATGCGGCCATTGGTC hpt R: GCCTCCAGAAGAAGATGTG</td>
<td>686</td>
<td>94°C for 5 min 94°C for 1 min 58°C for 30 s 72°C for 30 s 72°C for 2 min 35 cycles</td>
</tr>
<tr>
<td>cas9</td>
<td>cas9 F: ACTAACTCTGTGGGCTGGGC cas9 R: GCGCAATGAGATCCTCGAAC</td>
<td>694</td>
<td>95°C for 5 min 95°C for 45 s 58°C for 45 s 72°C for 45 s 72°C for 10 min 30 cycles</td>
</tr>
<tr>
<td>eIF4G</td>
<td>eIF4G F: AAGACTTTCCGGCAATTA eIF4G R: TAATTGGCCCGCAAGTCTT</td>
<td>577</td>
<td>95°C for 5 min 95°C for 45 s 53°C for 1 min 72°C for 45 s 72°C for 2 min 30 cycles</td>
</tr>
</tbody>
</table>

More recently, ABEs have been found to induce conversion of cytosine residue to guanine and thymine within its activity window in human and mouse cells (Lee et al., 2018; Kim et al., 2019, Jeong et al., 2020). This has questioned the exemption of ABEs from yielding unanticipated mutants, although in negligible percentage (Li et al., 2018). In support of this, an intriguing observation based on this experiment is the non-canonical substitution of C • G to G • C other than the expected A > G substitution. Unanticipated mutations in the sgRNA region were as well observed at positions G15 > C15 (9.29%) and G8 > A8 (1.15%) (Fig. 3 and Fig. 5). Examination of previous reports on ABE 7.10 series of adenine base editors in editing non-target bases showed that the deaminase enzyme used in the construction of

![Figure 2a. PCR analysis of putative mutants for the presence of hpt gene](image)

![Figure 2b. PCR analysis of putative mutants for the presence of cas9 gene](image)
ABE 7.10 from *E. coli* (ecTadA, *E. coli* tRNA Specific Adenosine deaminase) harbors common catalytic site for deamination of both cytosine and adenine residues (Jeong et al., 2020). As a result, cytosine conversions to G/T/A residues were observed using an ABE 7.10 in human and mouse cells. This, however was restricted to a narrow editing window of 5-7 bp, provided that the cytosine base was positioned in a specific TC*N* residue (Lee et al., 2018, Kim et al., 2019). Lee et al. (2018) have also discussed the concept of opposite strand editing, as they observed high G·A mutations, which may be due to C·T conversions on the opposite strand when they used CBEs. This opposite strand editing,

![Figure 3. Percent nucleotide changes at each base position of the target sequence](image)

(The arrows indicate the expected base substitution of A > G)

![Figure 4. Mutations observed at different nucleotide positions and corresponding protein sequences of the missense events generated. (Substitutions are denoted in red. A1: Allele 1; A2: Allele 2; WT: Wild Type)](image)
however, occurred only outside the editing window, also termed as bystander editing. Our observations of $G_{15} > C_{15}$ and $G_8 > C_8$ in the target strand corresponds to $C_{15} > G_{15}$ and $C_8 > G_8$ editing on the opposite strand. This can be considered as a case of opposite strand bystander editing. The editing at $G_8$ however does not follow the TC*N trend. Instead, it was observed in a N*CT fashion. Insights on unexpected edits with base editors, presumably regarded to be highly specific have opened up newer possibilities in genome editing. Jeong’s group has acknowledged that bystander editing of cytosines using ABEs are not out of the ordinary. They proposed that the probability of such an occurrence could be reduced by engineering the ecTadA enzyme. Their work on a series of ecTadA enzyme mutants with key modifications at specific amino acid positions have shown that cystine deaminase activity of the enzyme could be fine-tuned to increase or decrease cystine catalysis activity (Jeong et al., 2020). Being in the preliminary stage of application, dedicated research to understand the molecular mechanisms of actions of ecTadA enzyme has to be carried out.

CONCLUSION

The present study, based on the application of an adenine base editing vector to target a host translation initiation factor, eIF4G gene in indica rice ASD16, successfully identified 15 mutants with missense mutations. These 15 missense mutants had nucleotide substitutions affecting the VLFPNL residue and are promising candidates that will impart resistance against tungro disease, based on earlier reports of Lee et al. (2010) and Macovei et al. (2018). Progeny analysis needs to be carried out to identify homozygous mutants by raising subsequent T1 and T2 generations. Bioassay on these homozygous mutants would prove the level of resistance imparted by the substitution of the target residues.

FUNDING AND ACKNOWLEDGMENT

The authors would like to thank ICAR-NASF (ICAR/CRIISRPR-Cas-7003/2017-18) for the funding and Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore for providing facilities. YK also thank ICAR-NASF for the fellowship.

Table 2. Agrobacterium-mediated transformation of rice cultivar ASD16 and mutations events generated

<table>
<thead>
<tr>
<th>No. of batches co-cultivated</th>
<th>No. of immature embryos co-cultivated</th>
<th>No. of events generated</th>
<th>Transformation efficiency (%)</th>
<th>Events with missense mutations</th>
<th>Events with silent mutations</th>
<th>Mutation Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>1391</td>
<td>112</td>
<td>8.20</td>
<td>15</td>
<td>85</td>
<td>89.29</td>
</tr>
</tbody>
</table>

Figure 5. C substitutions on the opposite strand (3’ to 5’)

<table>
<thead>
<tr>
<th>Target strand 5’ to 3’</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>T5</th>
<th>T6</th>
<th>A7</th>
<th>C8(C&gt;A)</th>
<th>G9</th>
<th>G10</th>
<th>A11</th>
<th>A12</th>
<th>C12</th>
<th>A14</th>
<th>G15(G&gt;C)</th>
<th>A16</th>
<th>A17</th>
<th>C18</th>
<th>A19</th>
<th>G19</th>
<th>C20</th>
<th>T21/G22/G23</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complementary 3’ to 5’</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>AOC</td>
</tr>
</tbody>
</table>

(\(G_8 > C_8\) and \(G_8 > C_8\) when analysed on the complimentary strand corresponds to base substitutions of \(C > T\) and \(C > G\).)

REFERENCES


