Unusual Removal of T-DNA in T1 Progenies of Rice after Agrobacterium-mediated CRISPR/Cas9 Editing

Qin Yang

Jeonju University

Park Tae-Sung

NAAS: National Academy of Agricultural Sciences

Lee Bumkyu

Jeonju University

Lim Myung-Ho (✉ mlim312@korea.kr)

NAAS: National Academy of Agricultural Sciences

Research Article

Keywords: marker-free, CRISPR/Cas9 editing, T-DNA removal, rice

Posted Date: January 10th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1066224/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

In the present study, we attempted to knock out the *bar* gene selection marker in the fixed Bt- and herbicide-resistant transgenic line BT-T07 (T_8 generation) to generate a marker-free Bt-resistant rice line. A binary vector containing a CRISPR/Cas9 system targeting the 108–130 bp region of *bar* was transformed into rice embryonic calli, and plantlets were regenerated under non-selective conditions. Three T_0 plants were observed to have non-target point mutations and deletions in the targeted gene and were putatively heterozygous and chimeras. One T_0 plant, 130-4, was confirmed to have a 76-nt deletion, from 140 bp to 225 bp, and it showed the segregation of *bar* in its T_1 progenies, with 16 *bar*-knockout lines and seven normal *bar*-expressing lines. However, the CRISPR/Cas9 editing vector sequences were not detected in any of the T_1 plants. In addition, unusual removal of pre-existing T-DNA was observed in all *bar*-knockout T_1 plants. Illumina sequencing of a *bar*-knockout line, 130-4-36, revealed a small fraction of read residues of pre-existing T-DNA from the *bar* sequence to the right border at the original junction site. We speculate that this rare phenomenon might be directed by the homology between pre-existing T-DNA and CRISPR/Cas9 editing vector sequences during meiotic recombination. We report imprecise modifications and unpredictable outcomes of gene-editing techniques, providing valuable perspectives on gene-editing applications.

Key Message

We report the “escape” of a stable homozygous T-DNA from T_1 progenies of a CRISPR/Cas9-edited T_0 rice plant. Such unpredictable outcomes should be considered in development of gene-edited crops.

Introduction

Genome editing, a powerful tool for fast and accurate modification of plant genomes, is mainly represented by two distinct technologies: site-directed nucleases (SDNs) and oligonucleotide-directed mutagenesis (ODM). The former has been developed using four types of engineered nucleases, namely, meganuclease (MN), zinc finger nuclease (ZFN), transcription activator-like effector nuclease (TALEN), and type II clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein (Cas). These genome-editing technologies introduce DNA double-strand breaks (DSBs) and repair by either non-homologous end-joining (NHEJ) or homology-directed repair to generate different targeted DNA mutation types such as site-directed deletions, allele replacement, and site-directed insertions (or SDN-1/2/3, according to the terminology of Podevin et al. 2013). Among the SDN-based techniques, the CRISPR/Cas9 system has been most widely applied in genome editing to enhance important agronomic traits in cereal crops such as maize, sorghum, barley, rice, and wheat (Jiang et al. 2013; Liang et al. 2014, 2017; Zhang et al. 2014; Kapusi et al. 2017; El-Mounadi et al. 2020). Cas protein can be targeted for specific genome sequences by an easily designed 20-bp RNA guide sequence with a protospacer adjacent motif (PAM) that binds to DNA targets through base pairing. To compare ZFN and TALEN, the

CRISPR/Cas genome-editing system can be easily constructed, and it modifies multiple genes simultaneously at a relatively low cost (Osakabe et al. 2015).

With the increased applications of genome-editing technologies in crop improvement for national food and feed security, there has been considerable debate over environmental protection and the need to strictly regulate the development and launch of genome-edited crops such as genetically modified (GM) crops. Countries like Argentina, Brazil, Chile, Japan, and the USA have established biosafety regulations for genome-edited crops or have declared their deregulation. In most countries, including Korea, government policymakers, regulators, and risk assessors continue to discuss the adoption of product-based, process-based, or dual product and process approaches to regulate genome-editing crops and foods.

Mutation locations and scales, potential off-target modifications, complexity of the introduced changes, and novelty of the developed traits call for rigorous research on genome-editing applications and reliable techniques for risk assessments of genome-edited plants to address associated biosafety and societal concerns. Imprecise editing efficiency by CRISPR/Cas9 editing systems in plants result in undesirable non-targeted mutations or large deletions of target genes. The intrinsic feature of the target site, e.g., G or C of the -4 nucleotide from PAM sequence, and two or more repeat nucleotides at the cutting site, may cause off-target DNA cleavage, repetitive unit deletion, and indels of various sizes (Zhang et al. 2014; Chakarbarti et al. 2019; Manghwar et al. 2020; Mollar and Yang 2020). Irregular deletion may depend on the type of repair mechanism, in which DSB repair by microhomology-mediated end joining (MMEJ) frequently produces long deletions of up to 87 bp at the DSB site (Kapusi et al. 2017). Geminivirus-mediated CRISPR/Cas9 expression systems, such as 2X35S and ubiquitin promoters, can cause high-dosage CRISPR/Cas9 complexes in cells transiently, resulting in a significant increase in off-target mutation frequency (Ordon et al. 2017; Zhang et al. 2018). Furthermore, the CRISPR/Cas9 editing system has been confirmed to have a high frequency of integration into the target site, resulting in large deletions (42–926 nt) in target sites (Lee et al. 2019).

In the present study, the CRISPR/Cas9 editing system was applied to modify one locus of *bar* in a transgenic rice line BT-T07, which harbors transgenes of the modified Bt-resistance gene *mcry1Ac1* and the herbicide-resistance gene *bar*. Glufosinate (or phosphinothricin) is a non-selective herbicide that competitively inhibits glutamine synthetase, thereby leading to the accumulation of ammonia, which inhibits photosystem I and II reactions (Kumar et al. 2020). The bacterial gene *bar* from *Streptomyces hygroscopicus* encodes phosphinothricin acetyltransferase (PAT), which detoxifies this herbicide by acetylation, and it has been widely used for developing glufosinate-resistant transgenic crops. Moreover, *bar* is the most widely used selection marker in the transformation process, increasing selection efficiency during the development of commercial GM crops, such as wheat, rice, maize, soybean, barley, and sorghum (Brasileiro 2001). Therefore, this target is chosen not only to develop marker-free Bt-resistant rice plants but also to discover potential risks in both CRISPR/Cas9 application and derived modified plants.

Here, we report an unusual case wherein a stable homozygous transgene (T-DNA) escaped from some T₁ progenies of a *bar* gene-edited T₀ rice plant by the application of the CRISPR/Cas9 system. We show that CRISPR/Cas9 applications can create imprecise modifications and undesirable outcomes, and we propose and discuss possible scenarios for such deletion events.

Materials And Methods

Plant materials

A fixed Bt-resistant and herbicide-resistant transgenic rice line, BT-T07 (T₈ generation), was used as a CRISPR/Cas9 editing recipient line to knock out the selection marker *bar*. BT-T07 harbored an *mcry1Ac1* expression cassette and a *bar* expression cassette under the control of the RuBisCO promoter *Prbc* and cauliflower mosaic virus (CaMV) 35S promoter, respectively (Fig. 1). BT-T07 was developed by Dr. Myung-Ho Lim's group and was confirmed to have one copy T-DNA integration into 32,692,617 bp on chromosome 4 of the donor rice variety Dongjin (*Oryza sativa* L. ssp. *japonica*) in the reverse direction at the intergenic site with 9-nt deletions of genomic DNA.

Transformation vector construction

Based on the *bar* sequence, we designed sgRNA targeting 108~130 bp of the 5'-end of *bar* using the CRISPR RGEN tool (<http://www.rgenome.net/cas-designer/>; Park et al. 2015). An OsU3::gRNA scaffold::poly T cassette (468 bp) was amplified from the vector pRGE31 (Addgene plasmid #51295, a gift from Dr. Yinong Yang's laboratory) and inserted into pENTR to construct the psgRNA vector (Xie et al. 2013). The sgRNA was cloned into the psgRNA vector to synthesize psgRNA-*bar*130. We PCR-amplified 4,140 bp of *Atcas9* from the vector pAtCas9_1 (Addgene plasmid #91156, a gift from Dr. Daniel Voytas's laboratory) and inserted it into the pPZP vector to synthesize the binary vector pPZP-P35AtCas9 (Cermak et al. 2017). The sgRNA expression cassette was digested by *EcoRI* from the plasmid psgRNA-*bar*130 and cloned into the binary vector pPZP-P35AtCas9 to construct binary vectors of pPZP-*Atcas9Bar*130 (primer information is shown in Supplementary Table S1) for rice transformation.

Rice transformation

Agrobacterium-mediated transformation of embryogenic calli from BT-T07 was performed as described by Main et al. (2015) with some modifications. To knock out *bar*, transformed calli were grown on non-selective media until plantlet formation. Briefly, *A. tumefaciens* LBA4404 carrying the binary vector of pPZP-*Atcas9Bar*130 was grown on YEP medium with kanamycin (50 µg/mL) and spectinomycin (100 µg/mL) at 28 °C for 1–2 days. *Agrobacterium* cells were suspended in infection medium containing 100 µM acetosyringone, and the cell density at OD₆₀₀ was adjusted to 0.05. One-month-old rice calli were collected and immersed in *Agrobacterium* cultures for 20 min following transfer to the co-culture medium for 3 days at 25 °C in the dark. The infected calli were transferred to fresh N6 medium (marker-free) containing cefotaxime (250 mg/mL) for 2 weeks in the dark to remove the residual *Agrobacterium*. All

calli were then transferred to regeneration MS medium at 28 °C with a 16 h light/8 h dark cycle to produce shoots and roots. The regenerated rice plantlets were grown in soil in a greenhouse.

CRISPR/Cas9-mediated transgenic line selection and molecular analysis

Genomic DNA was extracted from the young leaves of 2-week-old regenerated T₀ plants using a DNase-secure Plant Kit (TianGen Biotech. Co., Ltd, Beijing, China). A set of specific primers (CAS-1749F/CAS-2341R) from *Atcas9* of the transformation vector was designed, and PCR was performed to screen potential CRISPR/Cas9-mediated transformed lines. Furthermore, DNA mutation analysis of the targeted *bar* for all selected T₀ plants was carried out using a proofreading polymerase (Phusion High Fidelity DNA Polymerase, Thermo Fisher Scientific, Inc., Waltham, MA, USA). PCR products were eluted and cloned into the pGEM-T vector (Promega Co., Madison, WI, USA), followed by Sanger sequencing of 24 positive clones per line. For T₁ plants, we evaluated *bar* expression using ImmunoStrips (Agdia Inc., Elkhart, IN, USA), where two red bands indicate PAT protein expression. Long-range PCR with LA Taq DNA polymerase (Takara Bio. Inc., Shiga, Japan) was used to amplify pre-existing T-DNA using flanking primers Q1CR-F02/Q1CR4-R01, which are respectively located at 0.5 kb from 5' and 3' of the T-DNA insertion site (T-DNA vector PZRT3 in Fig. 1). All primer information is provided in Supplementary Tables S2 and S3.

NGS sequencing and reference read mapping

Genomic DNA libraries were prepared using the TruSeq Library Kit (Illumina, Inc., San Diego, CA, USA), following the manufacturer's recommended protocol. Four separate indexed libraries, including the two T₁ lines 130-4-36 and 130-4-42, BT-T07 (recipient), and Dongjin (donor), were produced before sequencing. Sequencing was performed using the Illumina HiSeq platform (Illumina, Inc., San Diego, CA, USA), following the manufacturer's instructions (Kovalic et al. 2012). Sufficient numbers of sequence reads from each pooled sample library were obtained with a range of 117.7~149.8X genome coverage (Supplementary Table S4).

Raw fastaq data were created and imported into the CLC Genomics Workbench, version 9.0. Read mapping to a reference of the japonica variety Nipponbare (Os-Nipponbare-Reference-IRGSP-1.0) for the four materials was performed with a tool of "Map reads to reference," followed by track creation of four genomes. Sequence reads from the donor Dongjin, recipient BT-T07, and two CRISPR/Cas9-edited T₁ rice lines 130-4-36 and 130-4-42 were mapped to the T-DNA transformation vector of PZRT3 (pre-existing) and CRISPR/Cas9 editing vector of pPZP-*Atcas9Bar130*, respectively. Reference mapping parameters were set as follows: match score = 1, mismatch cost = 2, insertion cost = 3, deletion cost = 3, length fraction = 0.6, and similarity fraction = 0.9.

Structural variant detection

Structural variants, including single-nucleotide variations (SNVs), small-scale insertions and deletions (1~15 bp), and replacement and multi-nucleotide variations (MNVs), were identified from Dongjin (donor),

BT-T07 (recipient), 130-4-36 (T_1), and 130-4-42 (T_1) on the CLC Genomics Workbench according to the tutorial description (<https://digitalinsights.qiagen.com/support/tutorials>). An option of “basic variant detection” in the resequencing analysis tool was used to detect variants, with parameters set as follows: maximum expected variations (ploidy) = 2, ignore position with coverage above 60X, ignore broken pairs, coverage and count filters by minimum coverage = 10, minimum count = 2, minimum frequency (%) = 35.0, base quality filter in neighborhood radius = 5, minimum neighborhood quality = 15, minimum central quality = 20, relative read direction filter significance (%) = 1.0, remove pyro-error variants in homopolymer regions with minimum length = 3, and frequency below 0.8. After variant detection, we performed “compare variants” for variant tracks of BT-T07, 130-4-36, and 130-4-42, comparing the variant track of Dongjin to retain different variants. To further narrow down specific variants of three lines, an option of “filter against control reads” was carried out to filter variants against the mapped sequencing reads of Dongjin. Annotation of the extracted specific variants of the three lines was performed by using the option “amino acid change” to identify variants that cause a change in the amino acid composition of a protein. To compare BT-T07 and exclude the same variants between the two T_1 lines, line-specific variants of 130-4-36 and 130-4-42 were extracted for further statistical analysis.

Results

Experimental design, transformed T_0 plant selection, and mutation analysis

Considering the high GC content of *bar* and high identity to the rice genome, a 20-bp-long sgRNA targeting 108–130 bp of *bar* with 70% GC content and low off-target potential in the rice genome was selected to construct a binary vector of pPZP-Atcas9Bar130 (Fig. 1). *In vitro* cleavage of PZRT3 (*mcry1Ac1::bar*) with purified Cas9 protein and transcribed sgRNA was performed, and the on-target cleavage of sgRNA was confirmed (Supplementary Fig. 1a). We did not *design a selection marker expression cassette* in the CRISPR/Cas9 editing transformation vector, which indicated that transformed calli were grown on non-selective medium until plantlet regeneration. We performed PCR-based screening and selected three potential CRISPR/Cas9-edited T_0 plantlets, including 130-4, 130-37, and 130-115, from 150 regenerated plantlets (Supplementary Fig. 1b).

For the three lines, nucleotide modifications of *bar* were analyzed by high-fidelity polymerase PCR, followed by cloning and Sanger sequencing. We observed one SNP (C/T) at 226 bp and 23 of the wild-type *bar* for line 130-37. For line 130-115, one SNP (T/A) at 171 bp, one MNP (GT/AA) at 190–191 bp, and 22 wild types were identified. The T_0 line 130-4 was confirmed as having *bar* mutations, with one 76-nt nucleotide deletion from 140 bp to 215 bp, one SNP (G/A) at 380 bp, and 22 wild-type sequences (Fig. 2a). These results indicated that the three T_0 lines mostly had non-target mutations of *bar* and were putatively heterozygous or chimeras. The C-to-T base substitution at 226 bp changed proline to serine in the 130-37 line, but T-to-A and GT-to-AA base substitutions generated an early stop codon, mutating valine to lysine in the 130-115 line. The 76-bp deletion caused a frame shift and deletion of the protein sequence, whereas the G-to-A base substitution at 380 bp changed glycine to glutamic acid in the 130-4

line. PAT protein expression indicated active *bar* expression in all three T₀ lines, implying a failure to produce homozygous *bar*-mutated lines in T₀ plants by CRISPR/Cas9 editing.

Molecular characterization of T₁ plants

A total of 23, 24, and 9 T₁ plants from 130-4, 130-37, and 130-115 lines, respectively, were examined for PAT protein expression. As shown in Fig. 2b, a T₁ population consisting of 23 plants from self-pollinated seeds of 130-4 was observed in the PAT protein expression segregation, consisting of 16 *bar*-knockout T₁ plants and seven *bar*-expressing T₁ plants. However, the two T₁ populations from self-pollinated seeds of 130-37 and 130-115 showed completely active *bar* expression. Perhaps nucleotide modifications were produced in somatic cells of the T₀ plant but did not reach the germline and were not transmitted to the T₁ generations. When Cas9-specific PCR was performed for T₁ plants from 130-4 line to identify the CRISPR/Cas9 gene construct (pPZP-Atcas9Bar130), none of the T₁ plants harbored the Cas9 editing vector (data not shown). After PCR amplification of *bar* in T₁ plants from 130-4, we discovered a rare phenomenon in which PCR failed to amplify the *bar* gene in all *bar*-knockout T₁ plants (Fig. 2c). Cry protein expression was examined by ImmunoStrip tests, and the results were consistent with the plants expressing PAT protein (Fig. 2d). We considered the possibility that *bar* and *mcry1Ac1* expression cassettes were removed from these *bar*-knockout T₁ lines. This supposition was confirmed by full-length T-DNA PCR amplification of the T₁ plants. Based on the T-DNA-flanking sequence information of the BT-T07 line, PCR products amplified by flanking sequences located 0.5 kb from the 5'- and 3'-ends of the T-DNA insertion site were observed upon separation of ~1 kb and ~7 kb products from 23 T₁ plants of 130-4 (Fig. 2e). Among them, 16 T₁ *bar*-knockout plants consistently presented the same amplified product size of ~1 kb. Analysis of all ~1 kb products confirmed the flanking sequence in the genome without any trace of T-DNA integration. Further, two homozygous genotypes of 23 T₁ progenies were found based on the T-DNA site as described above, without heterozygous genotypes.

However, a weak PCR band of *bar* could be observed in several *bar*-knockout T₁ plants, (e.g., #13, #18, #36, #41, #43, and #45 lines), implying a small amount of nucleotide residues of *bar* (Fig. 2c). To confirm this speculation, we carried out long-range PCR from the mid-region of *bar* to the 5'- or 3'-flanking region of T-DNA (primer location can be confirmed in Fig. 1 and Supplementary Table S3) for one *bar*-knockout line 130-4-36. Compared to Dongjin and BT-T07, 130-4-36 showed a failure to amplify the region from *bar* to Q1CR4-R01, which consists of the *P35S* promoter, *mcry1Ac1* expression cassette, and left border (LB) sequence. However, a weak band amplified the region from *bar* to Q1CR4-F02 (Fig. 2f). Sequencing results also revealed an integration of the right border (RB) sequence and genomic DNA, accompanied by multiple sequencing peaks. The above results support the hypothesis that a minority of partial T-DNA from *bar* to the RB was retained in the genome of the T₁ line 130-4-36.

Whole-genome analysis of T₁ plants vis-à-vis their recipient and donor lines

To further understand the T-DNA integration of *bar*-knockout T₁ plants, we performed whole-genome sequencing of the T₁ line 130-4-36. For comparisons, a 130-4-42 plant with active *bar/mcry1Ac1* expression, the recipient line BT-T07, and the non-transgenic donor variety Dongjin were also analyzed. Based on the T-DNA transformation vector (PZRT3), reference mapping was implemented for these four lines (Fig. 3). Due to the endogenous multigene family of RuBisCO in the rice genome, a large number of reads were captured by the *Prbc* promoter region for all four lines. Uniform distributions of sequence reads were aligned to the T-DNA regions of *mcry1Ac1*, PinII terminator, and P35S::*bar*::Tnos cassette of BT-T07 and 130-4-42. No read sequences were mapped to vector sequences in the Dongjin genome, but a few reads were distributed in the *bar* and *Tnos* regions in 130-4-36. We compared the junction region of T-DNA and genomic DNA at 32,692,617 bp of chromosome 4 for the four lines (Supplementary Fig. 2). Only one read was observed in the right border (RB) of the T-DNA sequence joined to the genome sequence in line 130-4-36. However, many LB and RB sequences of T-DNA joined to genome sequences could be confirmed in BT-T07 and 130-4-42. These results indicated that most of the T-DNA of 130-4-36 had been removed from the integration site, with a few read residues from *bar* to the RB at the same junction region. This is consistent with the results of the PCR analysis. Read mapping to the CRISPR/Cas9 editing vector (pPZP-Atcas9Bar130) was carried out for 130-4-36 and 130-4-42 by whole-genome sequence analysis and compared to BT-T07 and Dongjin (Supplementary Fig. 3a). The lines 130-4-36 and Dongjin were observed on similar mapping graphs in which sequence reads were only captured by the *OsU3* promoter region (an endogenous gene in the rice genome). In comparison, BT-T07 and 130-4-42 showed additional read alignment for the *P35S* promoter and *Tnos* terminator that shared the same sequence as pre-existing T-DNA (PZRT3 vector). Therefore, we concluded that there was no *Cas9* gene construct integration into the rice genome of the two T₁ plants, which is consistent with the results of the PCR analysis.

We also found a chromosome break on chromosome 4 of 130-4-36, unlike the other three lines (Supplementary Fig. 3b). This break occurs between 10,688,843 bp and 10,689,431 bp of chromosome 4, with a nucleotide deletion of 588 bp, including a 5'-UTR region of the TE-related gene *Os04g19250*. To exclude mapping errors, PCR was performed to confirm this deletion (Supplementary Fig. 4a). We also analyzed 23 T₁ plants from 130-4 and confirmed that this deletion was only detected in five of 16 *bar*-knockout T₁ plants, indicating progeny segregation at this site (Supplementary Fig. 4b). Genotypes based on this 588-bp deletion were estimated, and the results revealed three genotypes, namely, homozygous without deletion, heterozygous genotype, and homozygous with deletion, with a segregation ratio of 2:9:5 for 16 *bar*-knockout T₁ plants (Supplementary Fig. 4c). For *bar*-knockout T₁ plants of 130-4-36, this deletion was confirmed to be homozygous and inherited in the T₂ progenies (Supplementary Fig. 4d). However, we did not detect the deletion from seven T₁ plants with active *bar/mcry1Ac1* expression. The above results support the possibility that T-DNA removal might be one of the reasons for this chromosome break in chromosome 4.

Detection of genome structural variations

We identified 21,599 and 12,423 sequence variants from 130-4-36 and 130-4-42, respectively, compared to their recipient line BT-T07, in which a majority of somatic variations arose during callus formation, plantlet regeneration, or *Agrobacterium* transformation. To exclude the same variants, we calculated 11,852 and 2,667 line-specific variations from 130-4-36 and 130-4-42, respectively (Supplementary Table S6). There were 11,363 synonymous variants and 489 non-synonymous variations in 130-4-36 and 2,489 and 171 for 130-4-42, respectively. We observed significant chromosome-bias distributions in which the highest numbers of 5,451 synonymous and 261 non-synonymous variants were distributed on chromosome 4 for 130-4-36. However, the highest number (512 synonymous and 87 non-synonymous variants) occurred on chromosome 11 for 130-4-42 (Fig. 4a). For 130-4-36, homozygous non-synonymous variations were mainly concentrated in chromosome 4 and on average accounted for 53.5% of the total variants. In comparison, approximately 28.1% of homozygous non-synonymous variants were mainly distributed on chromosome 11 for 130-4-42 (Fig. 4b). SNVs, deletions, insertions, replacements, and MNVs, respectively, accounted for 80.6%, 4.7%, 3.7%, 1.2%, and 9.8% of all non-synonymous variants in 130-4-36, but 59.1%, 13.5%, 2.9%, 0.6%, and 24.0% in 130-4-42. The lines 130-4-36 and 130-4-42 showed similar numbers of MNVs and deletions, but 130-4-36 had more SNVs, insertions, and substitutions (Fig. 4c). These results further confirmed that chromosomal rearrangement caused by pre-existing T-DNA removal might be limited to chromosome 4 but not the whole genome for T₁ plants of 130-4-36.

Discussion

We report the generation of targeted mutations in the herbicide-resistance gene *bar* in the fixed homozygous Bt- and herbicide-tolerant transgenic rice line BT-T07 using a CRISPR/Cas9 nuclease platform. This approach is considered promising for producing marker-free Bt-resistant transgenic rice lines in the early generations. Therefore, we regenerated CRISPR/Cas9-edited, vector-transformed rice calli using a non-selective medium, followed by PCR-based screening for Cas9-edited lines. We regenerated 150 T₀ plantlets and screened three lines with Cas9-editing vector integration. However, non-target point mutations and deletions of *bar* were identified, and very low ratios of 1:23 or 2:22 of mutated PCR amplicons to wild-type amplicons were obtained, showing heterozygous or chimeric genotypes in T₀ plants. Furthermore, we found a T₀ line, 130-4, harboring a 76-nt deletion from 140 bp to 215 bp near the target site of the 108–130 bp region of *bar*. Previous studies have shown that the mutation efficiency of CRISPR/Cas9 varies widely and can be affected by plant species, target sequence, *Cas9* promoter, sgRNA sequence selection, and transformation methods (Ma et al. 2015, 2016). Moreover, 50–70% of the GC content of the target sequence usually results in a high editing efficiency; however, it may also be associated with a greater risk of off-target effects. In this study, 68% of GC content occupied *bar*, and 70% of the GC content of sgRNA was selected for target editing. Likewise, non-target and point mutations in *bar* were major variations in our previous study, in which we edited *bar* in a herbicide-tolerant transgenic rice line using a TALEN platform (Qin et al. 2021). We tested on-target cleavage efficiency by *in vitro* transcription of selected sgRNA and Cas9 protein to avoid putative off-target effects on pre-existing T-DNA sequences. However, *in vitro* precise on-target cleavage did not work efficiently in the rice

genome during CRISPR/Cas9 transformation, suggesting that DNA repair mechanisms in plants could affect the targeting outcomes.

Pathak et al. (2019) edited three genes, including the selection marker *GUS*, by a dual-targeting CRISPR/Cas9 editing system in the rice genome. They observed variable point mutagenesis instead of targeted deletions on target loci. They explained that the Cas9:sgRNA complex could bind to the broken termini of DNA and prevent the removal of the free fragment from the site, eventually leading to simultaneous DSBs with point mutations introduced by NHEJ. A study using human cells revealed that editing precision is an intrinsic feature of the target site and depends on 10–12 nt of the seed sequence of NGG PAM-proximal regions (Chakrabarti et al. 2019). They suggested inducing predictable mutations at regions of interest by targeting nucleotide –4 from the PAM sequence, in which G or C shows mostly imprecise targets, while choosing target sites with A or T at this position. Off-target DNA cleavage often occurs in DNA sequences with up to 3–5-bp mismatches in the PAM-distal region of sgRNA-targeting sequences (reviewed by Manghwar et al. 2020). The imprecise targeting results in a high proportion of deletions and yields a range of indel sizes, with deletions as long as 2,315 bp (Chakrabarti et al. 2019). Zhang et al. (2014) tested 11 target genes for their amenability to CRISPR/Cas9-induced editing and found that the mutation types and frequency might be related to the nature of the target genes. Several studies have reported that the presence of two or more repeating nucleotides at the cut site frequently results in the deletion of repeating units (reviewed by Molla and Yang 2020). From another viewpoint, Cas9-induced DSBs follow error-prone NHEJ or MMEJ DNA repair, the latter frequently producing longer deletions of up to 87 bp at the DSB site (Kapusi et al. 2017). Irregular deletion may depend on the type of repair mechanism. The editing efficiency of SpCas9 in geminivirus-mediated CRISPR/Cas9 expression systems—such as those driven by the 2X35S promoter and ubiquitin promoter—can be greatly strengthened because DNA replicons harboring Cas9 and sgRNA cassettes transiently amplify hundreds of copies in plant cells, causing high-dosage CRISPR/Cas9 complexes in the cells and markedly enhancing the off-target mutation frequency (Ordon et al. 2017; Zhang et al. 2018). Lee et al. (2019) adapted a CRISPR/Cas9 system to demonstrate targeted T-DNA integration in the rice genome and identified five plants with T-DNAs integrated into the target site. They found at least 43-nt truncated LBs and large deletions (42–926 nt) at the targeting sites. These large deletions were likely due to delayed joining of LB–gDNAs due to the persistent binding of Cas9 to DSBs. These findings provide some reasonable explanations for our results. Two gene editing systems were used to consistently achieve non-target and point mutations, and different outcomes from rice plants arose from *in vitro* tests; therefore, we considered the inherent properties of the target gene *bar* and the sgRNA targeting sequence. The *bar* gene has a high GC content; short sequence repeats, including two repeats of CCGCAGGA at the sgRNA-targeting site (Fig. 2a), four repeats of GGACGG, and six repeats of CCGAG; and high identities to the rice genome, which could be mainly responsible for the observed off-target effects in the rice genetic background. We hypothesized that homologous sequences were used to repair open DNA strands to generate an out-of-frame 76-nt deletion in *bar* based on the MMEJ mechanism of DNA repair. Furthermore, the large-scale deletion was likely caused by the constitutive promoter 2X35S, which

induced a high dosage of the Cas9/sgRNA complex and caused excessive cleavage in error reorganization of the repeat sequences.

Primary regenerated plants mostly harbor chimeric mutational effects after CRISPR/Cas9 editing in rice (Zhang et al. 2014; Lee et al. 2019; Pathak et al. 2019), soybean (Michno et al. 2020), and barley (Kapuski et al. 2017). Chimeric mutations generally form when mutations occur after the division of the transformed embryogenic cell. Zhang et al. (2014) investigated putative genotypes of T_0 plants on target genes and found more than 40% of chimera genotypes in all mutated plants. T_1 genotypes from T_0 plants harboring chimeric mutations were irregular and mostly did not follow Mendel's law of segregation. Our data are consistent with these previous reports. As shown in Fig. 5, we tested all 23 T_1 plants and revealed only two segregation groups with 16 *bar*-knockout lines and seven wild type plants, similar to the recipient BT-T07, but without heterozygotes harboring both genotypes. Thereafter, we speculated that only a small sector of transformed calli was target-mutated and produced a chimeric T_0 plant 130-4. The segregation ratio and genotype patterns further supported this hypothesis. However, in an undesirable situation, not only the targeted *bar* but also the full pre-existing T-DNA in 16 *bar*-knockout T_1 lines were confirmed to be removed from their genomes. Whole-genome resequencing of one of the *bar*-knockout lines 130-4-36 revealed only a small quantity of T-DNA residue from *bar* to the RB joined with the genome sequence at the same T-DNA integration site. We found only one read at the junction between the RB and genome sequence at the T-DNA integration site in the chromosomal mapping view (Supplementary Fig. 1), and the recipient line BT-T07 was confirmed to have a 9-nt deletion resulting from T-DNA integration. However, there was no deletion at this site in the genome of 130-4-36, which showed the same nucleotides as those of the non-transgenic donor rice variety Dongjin. We also found an additional 588-nt deletion at 10,688,843–10,689,431 bp in the same chromosome where pre-existing T-DNA was integrated only in the 130-4-36 genome. In the F_2 population from 130-4-36, the residual pre-existing T-DNA sequences disappeared, but an additional 588-bp deletion proved the homozygous and heritable properties of this site (data not shown). In light of these results, we try to address some of the following questions: (i) Is it possible that T-DNA removal and the additional 588-bp deletion were caused by off-target editing of CRISPR/Cas9? (ii) Is the T-DNA removal caused by DNA recombination between the CRISPR/Cas9 editing gene construct and the pre-existing T-DNA? (iii) Is this a kind of “fighting back” response inherent to the rice genome to remove T-DNA?

First, 16 potential off-target sites were predicted through *in silico* off-target analysis of sgRNA in the rice genome, with the parameter setting of 2-nt DNA or RNA bulge and 2-nt mismatch, but 542 potential off-target sites were predicted with a 3-nt mismatch. We found no relationship between these potential off-target sites and loci of T-DNA elimination or 588-bp deletion. We investigated eight potential off-target sites involved in the functional elements of the eight genes and identified some mutations, including SNPs and 1–2-bp insertion/deletions, in five genes for both T_1 lines 130-4-36 and 130-4-42 and their recipient line BT-T07 (Supplementary Table S7). However, these mutations are likely not off-target effects of CRISPR/Cas9 editing but spontaneous mutations and somaclonal variations caused by tissue culture and *Agrobacterium*-mediated transformation.

Second, CRISPR/Cas9 editing vector sequences were not detected in T₁ plants in this study, indicating that T-DNA might be removed before or during seed production. Nevertheless, we confirmed the editing vector integration in T₀ plants by PCR in an earlier analysis. Accordingly, we hypothesized that T-DNA deletions might occur during meiosis. At this stage, programmed DSB formation and repair by homologous recombination leads to the establishment of connections between homologous chromosomes for the proper segregation of homologs during the first meiotic division for fertility (Borde and Massy 2015). Su et al. (2013) studied the MAP kinase kinase kinase gene *MEKK1* and its tandemly duplicated gene family using *Arabidopsis* T-DNA insertion lines. They identified a deletion mutant from 2,000 seedlings that appeared to be the result of recombination between a T-DNA insertion in *MEKK1* and a second T-DNA located 21 kb away in *MEKK3*, due to homology within the same T-DNA vector sequence. This phenomenon was termed “deletions caused by T-DNA recombination.” Extended research has confirmed that this type of deletion may be a relatively rare event; no events were found despite screening 28,314 meioses (Seagrist et al. 2018). Hu et al. (2017) demonstrated chromosome rearrangements, including a large deletion (155 kb) in an *Arabidopsis* T-DNA insertion mutant, which was caused by heterogeneous T-DNA integration but not recombination between two T-DNA insertions. The two large deletions described above, which occurred in T-DNA insertion lines, showed two different mechanisms. In one, meiotic crossing over between T-DNAs leads to intervening genomic DNA deletion. In the other, two ends of a T-DNA molecule integrate into two independent DSBs, leading to chromosomal rearrangements by incorrect DSB repair. In the present study, the CRISPR/Cas9 editing transformation vector had several homologous sites to those of pre-existing T-DNA sequences, including the same LB and RB sequences, *Tnos* terminator, and partial *35S* promoter. We considered that the T-DNA removal most likely occurred during meiotic recombination, which was directed by the homology between pre-existing T-DNA and the CRISPR/Cas9 editing vector sequence.

However, we cannot fully explain our findings such as partial sequence residues in the T₁ plant line 130-4-36. We considered the possibility that the CRISPR/Cas9 editing vector was inserted into the targeting site of CRISPR/Cas9 and the sgRNA complex, or one of the homologous sites of pre-existing T-DNA. Generally, T-DNA integrates at induced DNA DSBs, together with the typical NHEJ footprints, which support a model of T-DNA integration via the DSB repair pathway (Gelvin 2017). In addition, the T-DNA junction with the genome was not randomly under non-selective conditions, with preferential formation in regions that share microhomology to the T-DNA borders and with some enrichment in AT-rich regions (Shilo et al. 2017). Several studies have reported high-frequency T-DNA insertion into CRISPR/Cas9 cleavage sites in *Arabidopsis* (Zhang et al. 2018), rice (Lee et al. 2019), and soybean (Michno et al. 2020). Therefore, DSBs induced by CRISPR/Cas9 at the target site might be repaired by T-DNA integration during the transformation process, which results in a single T-DNA strand harboring a chimeric CRISPR/Cas9 editing vector sequence. During meiosis, several microhomology sites between two vector sequences might create deletions in the intervening sequence or homology sequence through DSB repair by homologous recombination. Considering this possibility, we analyzed T₀ plants (after ratooning) by NGS sequencing to identify the T-DNA integration site. Unexpectedly, we failed to identify the CRISPR/Cas9 editing vector insertion for 130-4. The most likely explanation for this finding is that the tillers harboring

the CRISPR/Cas9-mediated mutant sector died after ratooning. We also found an additional deletion of 588 bp on chromosome 4, the same chromosome in which the pre-existing T-DNA was inserted in the *bar*-knockout T₁ line 130-4-36. This deletion site showed a segregation ratio of 2:9:5 based on 16 *bar*-knockout T₁ plants (Fig. 5). Although the sample size was small, an approximately 1:2:1 segregation pattern was observed, following Mendel's segregation law. Thus, it can be seen that this deletion might occur with the emergence of CRISPR/Cas9 editing vector integration. Based on these combined results, a high frequency of structural variants was discovered, implying complex rearrangements on chromosome 4 of the T₁ line of 130-4-36. This deletion could be a rare event, but it remains unclear whether or not it is recovered by cell cycle checkpoints and is inherited by T₁ plants.

Third, it is also suspected that the removal of pre-existing T-DNA could be a kind of retaliation of the inherent rice genome. Transgenic plants are frequently observed to be unstable following *in vitro* propagation, leading to the loss of newly acquired traits. The newly integrated DNA might be recognized as foreign because its insertion disrupts normal chromatin structure; alternatively, defined GC content and sequence properties of the integrated DNA distinguish it from surrounding genome environments, resulting in the inactivation and methylation of this region (Finnegan and McElroy 1994). However, transgene instabilities often result from T-DNA methylation by multicopy integration, promoter inactivation, or co-suppression of ectopic and endogenous genes, but not from deletion or mutation of the introduced DNA. In this study, a 9-nt deletion of genomic DNA occurred in the recipient BT-T07 line; however, the fragment was observed in the T₁ plant 130-4-36 at the T-DNA integration site, with the sequence approximately equivalent to that of the non-transgenic donor rice variety Dongjin. Therefore, we speculated that cells may have some "memory" or maintenance systems to preserve genes in the genome. They can identify T-DNA as foreign because of the same T-DNA borders or similar invasion methods; they fight it out and then recover the genome to the original state by using DNA polymerase. Alternatively, the deleted nucleotides might never be lost from the "memory" of the rice genome even after eight generations since the first T-DNA integration, just waiting for a chance to recover. This unusual removal of T-DNA in T₁ progenies is rare and has not been reported in such detail in experiments using Cas9 or other nucleases. We believe that this study provides the preliminary experience to set up an effective CRISPR/Cas9 editing system focusing on target genes and sgRNA selections, vector construction, transformation methods, selective conditions, etc. More intelligent experimental designs are also needed when using the CRISPR/Cas9 system to modify readymade transgenic plants.

Abbreviations

SDN, site-directed nuclease; ODM, oligonucleotide-directed mutagenesis; MN, meganuclease; ZFN, zinc finger nuclease; TALEN, transcription activator-like effector nuclease; CRISPR, clustered regularly interspaced short palindromic repeat; Cas, CRISPR-associated protein; DSB, double-strand break; NHEJ, non-homologous end-joining; PAM, protospacer adjacent motif; MMEJ, microhomology-mediated end joining; PAT, phosphinothricin acetyltransferase; GM, genetically modified; SNV, single-nucleotide variation; MNV, multi-nucleotide variation; LB, left border; RB, right border

Declarations

Statements and Declarations

Author contributions: YQ and MHL conceived the experiments; YQ performed the experiments, data analysis, and writing; MHL, TSP and BL considered and discussed the results; MHL: handled the project administration; all authors reviewed the manuscript.

Compliance with ethical standards

Disclosure of potential conflicts of interest: The authors declare that there are no conflicts of interest.

Data Availability Statement

NGS sequencing data of the non-transgenic donor variety Dongjin, the recipient line BT-T07, 130-4 (T0), 130-4-36 (T1) and 130-4-42 (T1) described in this article had been deposited in an openly available database of National Agricultural Biotechnology Information Center, Korea, with accession numbers of NN-7768, NN-7765, NN-7769, NN-7766 and NN-7767, respectively.

Acknowledgment

This work was supported by a grant (PJ01432206) from the National Institute of Agricultural Sciences (Rural Development Administration, Republic of Korea).

References

- Borde V, de Massy B (2015) Meiosis: early DNA double-strand breaks pave the way for inter-homolog repair. *Dev Cell* 32:663-664. doi: 10.1016/j.devcel.2015.03.011.
- Brasileiro ACM (2001) Marker genes for in vitro selection of transgenic plants. *J Plant Biotechnol* 3(3):113-121.
- Chakrabarti AM, Henser-Brownhill T, Monserrat J, Poetsch AR, Luscombe NM, Scaffidi P (2019) Target-specific precision of CRISPR-mediated genome editing. *Mol Cell* 73:699-713. <https://doi.org/10.1016/j.molcel.2018.11.031>
- Finnegan J, McElroy D (1994) Transgene inactivation: plants fight back! *Nature Biotechnol* 12:883-888. <https://doi.org/10.1038/nbt0994-883>
- Gelvin SB (2017) Integration of *Agrobacterium* T-DNA into the plant genome. *Annu Rev Genet* 51:195-217. doi: 10.1146/annurev-genet-120215-035320
- Hu Y, Chen Z, Zhuang C, Huang J (2017) Cascade of chromosomal rearrangements caused by a heterogeneous T-DNA integration supports the double-stranded break repair model for T-DNA integration.

Plant J 90:954-965. doi: 10.1111/tpj.13523

Jiang WZ, Xhou HB, BI HH, Fromm M, Yang B, Weeks DP (2013) Demonstration of 515 CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum 516 and rice. Nucleic Acids Res 41:e188. doi: 10.1093/nar/gkt780

Kapusi E, Corcuera-Gómez M, Melnik S, Stoger E (2017) Heritable genomic fragment deletions and small indels in the putative ENGase gene induced by CRISPR/Cas9 in barley. Front Plant Sci 8:540. doi: 10.3389/fpls.2017.00540

Kumar K, Gambhir G, Dass A, Tripathi AK, Singh A, Jha AK, Yadava P, Choudhary M, Rakshit S (2020) Genetically modified crops: current status and future prospects. Planta 251(4):91. <https://doi.org/10.1007/s00425-020-03372-8>

Lee K, Eggenberger AL, Banakar R, McCaw ME, Zhu HL, Main M, Kang M, Gelvin SB, Wang K (2019) CRISPR/Cas9-mediated targeted T-DNA integration in rice. Plant Mol Biol 99:317-328. <https://doi.org/10.1007/s11103-018-00819-1>

Liang Z, Zhang K, Chen K, Gao C (2014) Targeted mutagenesis in *Zea mays* using TALENs and the CRISPR/Cas system. J Genet Genomics 41:63-68. doi: 10.1016/j.jgg.2013.12.001

Liang Z, Chen K, Li T, Zhang Y, Wang Y, Zhao Q, Liu JX, Zhang HW, Liu C, Ran Y, Gao C (2017) Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. Nat Commun 8:14261. <https://doi.org/10.1038/ncomms14261>

Ma X, Zhang Q, Zhu Q, et al. (2015) A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. Mol Plant 8:1274-1284. doi: 10.1016/j.molp.2015.04.007

Main M, Frame B, Wang K (2015) Rice, Japonica (*Oryza sativa* L.). In: Wang K (ed) *Agrobacterium* protocols, 3rd edn. Springer, New York, pp 169-180.

Manghwar H, Li B, Ding X, Hussain A, Lindsey K, Zhang X, Jin S (2020) CRISPR/Cas system in genome editing: methodologies and tools for sgRNA design, off-target evaluation, and strategies to mitigate off-target effects. Adv Sci 7:1902312. <https://doi.org/10.1002/adv.201902312>

Michno JM, Virdi K, Stec AO, Liu J, Wang X, Xiong Y, Stupar RM (2020) Integration, abundance, and transmission of mutations and transgenes in a series of CRISPR/Cas9 soybean lines. BMC Biotechnol 20:10. <https://doi.org/10.1186/s12896-020-00604-3>

Molla KA, Yang Y (2020) Predicting CRISPR/Cas9-induced mutations for precise genome editing. Trends Biotechnol 38 (2):136-141. doi: 10.1016/j.tibtech.2019.08.002

- Ordon J, Gantner J, Kemna J, Schwalgun L, Reschke M, Streubel J, Boch J, Stuttmann J (2017) Generation of chromosomal deletions in dicotyledonous plants employing a user-friendly genome editing toolkit. *Plant J* 89:155-168. doi: 10.1111/tpj.13319
- Osakabe Y, Osakabe K (2015) Genome editing with engineered nucleases in plants. *Plant Cell Physiol* 56(3):389-400. <https://doi.org/10.1093/pcp/pcu170>
- Pathak B, Zhao S, Manoharan M, Srivastava V (2019) Dual-targeting by CRISPR/Cas9 leads to efficient point mutagenesis but only rare targeted deletions in the rice genome. *3 Biotech* 9:158. doi: 10.1007/s13205-019-1690-z
- Podevin N, Davies HV, Hartung F, Nogue F, Casacuberta JM (2013) Site-directed 563 nucleases: a paradigm shift in predictable, knowledge-based plant breeding. *Trends Biotechnol* 31:375-383. doi: 10.1016/j.tibtech.2013.03.004
- Qin Y, Park TS, Cho YS, Lim MH (2021) TALEN-mediated bar-knockout rice production and transcriptome profiling. *Plant Breed Biotech* 9(1):32-44.
- Seagrist JF, Su SH, Krysan PJ (2018) Recombination between T-DNA insertions to cause chromosomal deletions in Arabidopsis is a rare phenomenon. *Peer J* 6:e5076. doi: 10.7717/peerj.5076
- Shilo S, Tripathi P, Melamed-Bessudo C, Tzfadia O, Muth TR, Levy AA (2017) T-DNA-genome junctions form early after infection and are influenced by the chromatin state of the host genome. *PLoS Genet* 13(7):e1006875. doi: 10.1371/journal.pgen.1006875
- Srivastava V, Underwood JL, Zhao S (2017) Dual-targeting by CRIPSR/Cas9 for precise excision of transgenes from rice genome. *Plant Cell Tiss Organ Cult* 129:153-160. <https://doi.org/10.1007/s11240-016-1166-3>
- Su SH, Bush SM, Zaman N, Stecker K, Sussman MR, Krysan P (2013) Deletion of a tandem gene family in Arabidopsis: increased MEKK2 abundance triggers autoimmunity when the MEKK1/2-MPK4 signaling cascade is disrupted. *Plant Cell* 25:1895-1910. doi: 10.1105/tpc.113.112102
- Zhang H, Zhang JS, Wei PL, Zhang BT, Gou F, Feng ZY, Mao YF, Yang L, Zhang H, Xu NF, Zhu JK (2014) The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnol J* 12:797-807. doi: 10.1111/pbi.12200
- Zhang Q, Xing HL, Wang ZP, Zhang HY, Yang F, Wang XC, Chen QJ (2018) Potential high-frequency off-target mutagenesis induced by CRISPR/Cas9 in Arabidopsis and its prevention. *Plant Mol Biol* 96:445-456. doi: 10.1007/s11103-018-0709-x

Figures

Figure 1

Schematic diagram of the experimental design. The CRISPR/Cas9 and sgRNA complex targets on the selection marker *bar* (yellow box with a red outline) for the recipient rice line BT-T07, a T_8 fixed line harboring a T-DNA (PZRT3 vector) integrated into 32,692,617 bp on chromosome 4

Figure 2

Molecular characterization of the CRISPR/Cas9-edited T₀ plant 130-4 and its T₁ progenies. (a) Sequence deletions (76 bp) of *bar* of a 130-4 T₀ plant; (b) PAT protein expression test by ImmunoStrips for all 23 T₁ plants from 130-4. (c) PCR amplification targeting *bar* in 23 T₁ plants. (d) Cry (Bt-resistant *mcry1Ac1*) protein expression test by ImmunoStrips for 23 T₁ plants. (e) Long-range PCR analysis for pre-existing T-DNA in BT-T07 line confirmation of 23 T₁ plants by using the primer set Q1CR4-F02/Q1CR4-R01; (i): PCR amplification containing full-length T-DNA; (ii): PCR amplification without T-DNA. (f) PCR amplified from *bar* to 5'- or 3'-regions of genomic DNAs, and Sanger sequencing for products amplified using the primer set Q1CR4-F02/Bar-f3

Figure 3

Reference mapping of two T₁ Cas9-edited lines 130-4-36 and 130-4-42, the recipient line Bt-T07, and the non-transgenic donor variety Dongjin based on the pre-existing T-DNA vector sequence

Figure 4

Structural variation analysis for two T₁ plants (130-4-36 and 130-4-42) using a CLC genomic workbench platform. (a) Distribution of synonymous and non-synonymous variations on chromosomes of both 130-4-36 and 130-4-42. (b) Comparisons between non-synonymous homozygous and heterozygous variations on each chromosome between 130-4-36 and 130-4-42. (c) Chromosome distribution of single-nucleotide variations (SNVs), deletions, insertions, replacements, and multiple nucleotide variations (MNVs) of 130-4-36 and 130-4-42, respectively

Figure 5

Overall schematic presentations of the experimental procedure and pedigree of the CRISPR/Cas9-edited T_0 rice plant 130-4

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterials.docx](#)