TARGETED GENOME EDITING IN CROPS

EDITED BY: Huanbin Zhou, Lanqin Xia, Yong Zhang and Seiichi Toki

PUBLISHED IN: Frontiers in Genome Editing

frontiers Research Topics

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-88971-554-1
DOI 10.3389/978-2-88971-554-1

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

TARGETED GENOME EDITING IN CROPS

Topic Editors:
Huanbin Zhou, Chinese Academy of Agricultural Sciences, China
Lanqin Xia, Chinese Academy of Agricultural Sciences, China
Yong Zhang, University of Electronic Science and Technology of China, China
Seiichi Toki, National Institute of Agrobiological Sciences, Japan

Citation: Zhou, H., Xia, L., Zhang, Y., Toki, S., eds. (2021). Targeted Genome Editing in Crops. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88971-554-1

Table of Contents

05 ***Editorial: Targeted Genome Editing in Crops***
Huanbin Zhou, Yong Zhang, Lanqin Xia and Seiichi Toki

07 ***Precision Genome Engineering Through Cytidine Base Editing in Rapeseed (*Brassica napus*. L)***
Limin Hu, Olalekan Amoo, Qianqian Liu, Shengli Cai, Miaoshan Zhu, Xiaoxiao Shen, Kaidi Yu, Yungu Zhai, Yang Yang, Lei Xu, Chuchuan Fan and Yongming Zhou

19 ***A Universal System of CRISPR/Cas9-Mediated Gene Targeting Using All-in-One Vector in Plants***
Ayako Nishizawa-Yokoi, Masafumi Mikami and Seiichi Toki

31 ***Repurposing of Anthocyanin Biosynthesis for Plant Transformation and Genome Editing***
Yubing He, Min Zhu, Junhua Wu, Lejun Ouyang, Rongchen Wang, Hui Sun, Lang Yan, Lihao Wang, Meilian Xu, Huadong Zhan and Yunde Zhao

41 ***Precision Genome Engineering for the Breeding of Tomatoes: Recent Progress and Future Perspectives***
Tien Van Vu, Swati Das, Mil Thi Tran, Jong Chan Hong and Jae-Yean Kim

57 ***Genome Engineering in Plant Using an Efficient CRISPR-xCas9 Toolset With an Expanded PAM Compatibility***
Chengwei Zhang, Guiting Kang, Xinxiang Liu, Si Zhao, Shuang Yuan, Lu Li, Yongxing Yang, Feipeng Wang, Xiang Zhang and Jinxiao Yang

67 ***Establishment of CRISPR/Cas9 Genome Editing in Witloof (*Cichorium intybus *var.* foliosum*)***
Charlotte De Bruyn, Tom Ruttink, Tom Eeckhaut, Thomas Jacobs, Ellen De Keyser, Alain Goossens and Katrijn Van Laere

79 ***CRISPR-Cas9-Mediated Mutagenesis of the Rubisco Small Subunit Family in* Nicotiana tabacum**
Sophie Donovan, Yuwei Mao, Douglas J. Orr, Elizabete Carmo-Silva and Alistair J. McCormick

94 ***Precise Genome Editing in miRNA Target Site via Gene Targeting and Subsequent Single-Strand-Annealing-Mediated Excision of the Marker Gene in Plants***
Namie Ohtsuki, Keiko Kizawa, Akiko Mori, Ayako Nishizawa-Yokoi, Takao Komatsuda, Hitoshi Yoshida, Katsuyuki Hayakawa, Seiichi Toki and Hiroaki Saika

106 ***Genome Editing and Protoplast Regeneration to Study Plant–Pathogen Interactions in the Model Plant* Nicotiana benthamiana**
Chen-Tran Hsu, Wen-Chi Lee, Yu-Jung Cheng, Yu-Hsuan Yuan, Fu-Hui Wu and Choun-Sea Lin

115 ***Protein Phosphatase 2A Catalytic Subunit PP2A-1 Enhances Rice Resistance to Sheath Blight Disease***
Qiu Jun Lin, Jin Chu, Vikranth Kumar, De Peng Yuan, Zhi Min Li, Qiong Mei and Yuan Hu Xuan

124 ***CRISPR/Cas9-Mediated Multi-Allelic Gene Targeting in Sugarcane Confers Herbicide Tolerance***
Mehmet Tufan Oz, Angelika Altpeter, Ratna Karan, Aldo Merotto and Fredy Altpeter

frontiers in Genome Editing

EDITORIAL
published: 09 September 2021
doi: 10.3389/fgeed.2021.757916

Editorial: Targeted Genome Editing in Crops

Huanbin Zhou[1]*, *Yong Zhang*[2], *Lanqin Xia*[3] and *Seiichi Toki*[4,5,6]

[1]Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China, [2]Department of Biotechnology, School of Life Sciences and Technology, University of Electronic Science and Technology of China, Chengdu, China, [3]Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China, [4]Department of Plant Life Science, Ryukoku University, Otsu, Japan, [5]Institute of Agrobiological Sciences, National Agriculture and Food Research Organization (NARO), Tsukuba, Japan, [6]Kihara Institute for Biological Research, Yokohama City University, Yokohama, Japan

Keywords: genome editing, gene targeting, CRISPR, cas protein, gene function study, crop improvement

Editorial on the Research Topic

Targeted Genome Editing in Crops

In the past 4 decades, tremendous progress has been made in our understanding of plant molecular biology, especially the molecular basis of a wide range of agronomic traits, such as seed development, nutrient acquisition, stress tolerance, disease resistance, etc. However, compared to the active researches in model plants, our current knowledge about crops lags far behind due to the limited availability of mutants for gene function study. In addition, translational research, which applies the fundamental knowledge into improved agronomic traits in crops, is also poor due to technical problems. Fortunately, the newly-developed, cutting-edged genome editing technologies, which enable precise sequence modifications in the crop genomes, provide effective tools to address these challenges. Nowadays, various loss-of-function mutants and novel gain-of-function alleles of genes of interest can be easily generated *via* genome editing. These new gene editing technologies bridge the gaps between gene function study and crop improvement.

So far, engineered meganucleases, zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly interspersed short palindromic repeats (CRISPR)-CRISPR-associated protein (Cas) have been successfully adapted for targeted genome editing in crops. Of these, the CRISPR/Cas systems, characterized by the RNA-guided DNA nucleases, have been widely employed in a large number of crop species for multiple applications since it is simple-to-design, easy-to-use, and highly-efficient. Generally speaking, the CRISPR/Cas nucleases generate a double-strand DNA break at the target site in the crop genome, resulting in indel mutations *via* the error-prone non-homologous end joining pathway or DNA fragment replacement in the presence of donor template through the homology-directed repair pathway (Sukegawa et al., 2021). In addition, different kind of effectors can be engineered with CRISPR/Cas nickase system for precise nucleotide editing. For example, cytidine deaminases or adenine deaminases could be guided by CRISPR/Cas to the target sequences, inducing C to T and A to G conversion, respectively (Ren et al., 2018; Yan et al., 2018), and CRISPR/Cas9-fused M-MLV reverse transcriptase could introduce all possible nucleotide substitutions and combinations through reverse transcription (Li et al., 2020; Tang et al., 2020).

In this special topic of frontiers in genome editing, we have compiled the recent advances in the development and applications of targeted genome editing technology in crops. Ten original research papers on the special topic of genome editing, including DNA fragment replacement, gene knockout and base editing, in a number of crops are included in this issue. Seiichi Toki's lab present two powerful gene targeting (GT) methods utilizing a SSA (Single-Strand-Annealing)-mediated marker excision system and a CRISPR/Cas9-mediated strategy with all-in-one vector. In their first paper, Ohtsuki et al. used SSA-mediated marker excision system to introduce three and seven multiple discontinuous bases into the microRNA miR172

OPEN ACCESS

Edited and reviewed by:
Bing Yang,
University of Missouri, United States

****Correspondence:***
Huanbin Zhou
hbzhou@ippcaas.cn

Specialty section:
This article was submitted to Genome Editing in Plants, a section of the journal Frontiers in Genome Editing

Received: *13 August 2021*
Accepted: *26 August 2021*
Published: *09 September 2021*

Citation:
Zhou H, Zhang Y, Xia L and Toki S (2021) Editorial: Targeted Genome Editing in Crops. Front.Genome Ed. 3:757916. doi: 10.3389/fgeed.2021.757916

target site of *OsCly1* gene in rice to achieve the cleistogamous flowering phenotype. In their second paper, Nishizawa-Yokoi et al. developed a CRISPR/Cas9-mediated GT strategy utilizing a single, all-in-one vector containing a CRISPR/Cas9 expression construct, a selectable marker and a donor template. In their study, Nishizawa-Yokoi and colleagues utilized this novel system for homology-directed repair (HDR) in rice and tobacco, successfully modifying specific target genes such as *OsALS*, *NtALS*, *NtEPSPS*, etc.

Oz and colleagues from Fredy Altpeter's lab present novel advances in sugarcane gene editing. In their study, Oz and colleagues present a detailed description on the generation of herbicide-tolerant sugarcane plants by co-editing multiple alleles of *ALS via* HDR-mediated repair of double-strand DNA break induced by CRISPR/Cas9.

PAM compatibility is a limiting factor in the CRISPR/Cas9 system. To address this issue, Zhang and colleagues from Jinxiao Yang's lab provide an evidence that tRNA-esgRNA broadens the PAM compatibility of xCas9 nuclease, enabling it to function at NG, GAA, GAT, and GAG PAM sites in rice. Moreover, Zhang and colleagues further enrich the CRISPR toolbox by developing an xCas9-based cytosine base editor (CBE) capable of editing NG and GA PAM sites.

Bruyn and colleagues' paper brings into light what is required for efficient gene modification of witloof. In their study, Bruyn and colleagues report a highly-efficient CRISPR/Cas9-mediated genome editing workflow in traditional Belgian crop witloof based on PEG-mediated protoplast transfection, whole plant regeneration and HiPlex amplicon sequencing. By using this platform, Bruyn and colleagues successfully edit *CiGAS*, *CiGAO*, and *CiCOS*, which are of significant importance to control the biosynthesis of sesquiterpene lactones and hence the bitterness of witloof.

Nicotiana benthamiana is one of the most utilized model plant species in plant molecular biology. Thus, the paper presented by Hsu and colleagues from Choun-Sea Lin's lab is a noticeable contribution in this research area. In their paper, Hsu and colleagues report a simple, highly robust genome-editing in *N. benthamiana* protoplasts by using SaCas9, SpCas9, FnCas12a, and nCas9-based CBE, and successfully regenerated stable lines from protoplast. Moreover, Hu and colleagues report that CBEs can enable precise C-to-T substitutions at endogenous loci in rapeseed.

To facilitate the generation of transgene-free edited lines, He and colleagues from Yunde Zhao's lab report a simple method to detect transgenic events using a visual marker. In their study, He and colleagues develop an anthocyanin-marker assisted CRISPR system that enabled the identification of transgene-free and target gene-edited plants in T1 generation based on anthocyanin accumulation.

Advances in the use of CRISPR/Cas in *N. tabacum*, rice and tomato are also presented in this special issue. Donovan et al. and Lin et al. describe the successful application of the CRISPR/SpCas9 system in editing *rbcS* homologs in *N. tabacum* and *PP2A-1* in rice. Finally, Vu et al. presents a comprehensive review on the latest developments in precision genome editing in tomatoes and prospects its applications in breeding.

CRISPR is a technology which is constantly evolving. As presented in the articles on this special issue, CRISPR/Cas is developing into a "game-changing" technology in plant science with diverse applications ranging from basic studies to the more applied translational research. We hope that the topics covered in this issue will boost the benefit of genome editing technology, not only for gene discovery but also for crop improvement.

AUTHOR CONTRIBUTIONS

HZ wrote the original draft. All authors participated in discussion and revision of the manuscript, and approved it for publication.

FUNDING

HZ received funding from the National Natural Science Foundation of China (31871948) and the Central Public-interest Scientific Institution Basal Research Fund (Y2020PT26).

ACKNOWLEDGMENTS

We thank Prof. Bing Yang, Chief Editor of Frontiers in Genome Editing, for inviting us to compile this Special Issue, and all contributors and reviewers for their input.

REFERENCES

Li, H., Li, J., Chen, J., Yan, L., and Xia, L. (2020). Precise Modifications of Both Exogenous and Endogenous Genes in rice by Prime Editing. *Mol. Plant.* 13 (5), 671–674. doi:10.1016/j.molp.2020.03.011

Ren, B., Yan, F., Kuang, Y., Li, N., Zhang, D., Zhou, X., et al. (2018). Improved Base Editor for Efficiently Inducing Genetic Variations in Rice With CRISPR/Cas9-guided Hyperactive hAID Mutant. *Mol. Plant.* 11 (4), 623–626. doi:10.1016/j.molp.2018.01.005

Sukegawa, S., Saika, H., and Toki, S. (2021). Plant Genome Editing: Ever More Precise and Wide Reaching. *Plant J.* 106, 1208–1218. doi:10.1111/tpj.15233

Tang, X., Sretenovic, S., Ren, Q., Jia, X., Li, M., Fan, T., et al. (2020). Plant Prime Editors Enable Precise Gene Editing in Rice Cells. *Mol. Plant.* 13 (5), 667–670. doi:10.1016/j.molp.2020.03.010

Yan, F., Kuang, Y., Ren, B., Wang, J., Zhang, D., Lin, H., et al. (2018). Highly Efficient A·T to G·C Base Editing by Cas9n-Guided tRNA Adenosine Deaminase in Rice. *Mol. Plant.* 11 (4), 631–634. doi:10.1016/j.molp.2018.02.008

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Zhou, Zhang, Xia and Toki. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genome Editing

ORIGINAL RESEARCH
published: 20 November 2020
doi: 10.3389/fgeed.2020.605768

Precision Genome Engineering Through Cytidine Base Editing in Rapeseed (*Brassica napus*. L)

Limin Hu, Olalekan Amoo, Qianqian Liu, Shengli Cai, Miaoshan Zhu, Xiaoxiao Shen, Kaidi Yu, Yungu Zhai, Yang Yang, Lei Xu, Chuchuan Fan and Yongming Zhou*

National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, China

Rapeseed is one of the world's most important sources of oilseed crops. Single nucleotide substitution is the basis of most genetic variation underpinning important agronomic traits. Therefore, genome-wide and target-specific base editing will greatly facilitate precision plant molecular breeding. In this study, four CBE systems (BnPBE, BnA3A-PBE, BnA3A1-PBE, and BnPBGE14) were modified to achieve cytidine base editing at five target genes in rapeseed. The results indicated that genome editing is achievable in three CBEs systems, among which BnA3A1-PBE had the highest base-editing efficiency (average 29.8% and up to 50.5%) compared to all previous CBEs reported in rapeseed. The editing efficiency of BnA3A1-PBE is ∼8.0% and fourfold higher, than those of BnA3A-PBE (averaging 27.6%) and BnPBE (averaging 6.5%), respectively. Moreover, BnA3A1-PBE and BnA3A-PBE could significantly increase the proportion of both the homozygous and biallelic genotypes, and also broaden the editing window compared to BnPBE. The cytidine substitution which occurred at the target sites of both *BnaA06.RGA* and *BnaALS* were stably inherited and conferred expected gain-of-function phenotype in the T1 generation (i.e., dwarf phenotype or herbicide resistance for weed control, respectively). Moreover, new alleles or epialleles with expected phenotype were also produced, which served as an important resource for crop improvement. Thus, the improved CBE system in the present study, BnA3A1-PBE, represents a powerful base editor for both gene function studies and molecular breeding in rapeseed.

Keywords: ***Brassica napus*, base-editing, cytidine deaminase, gain-of-function mutation, crop improvement**

OPEN ACCESS

Edited by:
Lanqin Xia,
Chinese Academy of Agricultural Sciences, China

Reviewed by:
Hongliang Zhu,
China Agricultural University, China
Hao Lin,
Chinese Academy of Agricultural Sciences, China

****Correspondence:***
Chuchuan Fan
fanchuchuan@mail.hzau.edu.cn

Specialty section:
This article was submitted to Genome Editing in Plants, a section of the journal Frontiers in Genome Editing

Received: *13 September 2020*
Accepted: *27 October 2020*
Published: *20 November 2020*

Citation:
Hu L, Amoo O, Liu Q, Cai S, Zhu M, Shen X, Yu K, Zhai Y, Yang Y, Xu L, Fan C and Zhou Y (2020) Precision Genome Engineering Through Cytidine Base Editing in Rapeseed (Brassica napus. L). Front. Genome Ed. 2:605768. doi: 10.3389/fgeed.2020.605768

INTRODUCTION

Rapeseed (*Brassica napus* L., AACC, 2n = 38) is one of the most important sources of oilseed crops in the world, accounting for ∼16% of the entire global vegetable oil production (Woodfield et al., 2017). Achieving high yields through genetic improvements has always been the major goal in rapeseed production. The constant creation and use of novel genetic variation are important to both genetic research and plant trait improvement. In order to optimize the agronomic traits of crops, breeders applied various methods such as chemical compounds and irradiation to produce heritable mutations. However, these traditional techniques are not target-specific and require genome-scale screening, which is time-and-labor-consuming (Russell et al., 1958; Sega, 1984). As an allotetraploid species, rapeseed has a complicated genome in which most genes have several homologous copies (Chalhoub et al., 2014). Thus, obtaining mutations at all homologous

copies is challenging by traditional mutagenesis. With the rapid progress in molecular biology, genome-editing technologies have proven to be a powerful tool to address this issue.

In recent years, CRISPR/Cas9 systems have been proven to be very efficient in improving agronomic traits, especially yield-related traits of rapeseed through genome editing (Braatz et al., 2017; Yang et al., 2018; Zhai et al., 2020). The traditional CRISPR/Cas9 system prefers to generate small insertions and deletions (indels) and is best suited to create knockout mutations. This makes the traditional CRISPR/Cas9 system ineffective when precise base substitutions are needed. However, many desired agronomic traits involve only single nucleotide variants within genes, such as the reported cytidine (C) to thymidine (T) replacement at particular sites of *BnaALS*, *BnaRGA*, and *BnaA3.IAA7* that conferred gain-of-function mutations with valuable benefits for agricultural applications in rapeseed (Liu et al., 2010; Li et al., 2015, 2019). Therefore, precise base editing has great potential for the production of desired alleles and trait improvement. Using this method, desirable traits can be introgressed into elite lines without compromise, and the resulting lines with targeted improvement will be utilized for practical production.

Recently, base editors, including cytidine base editors (CBEs) and adenine base editors (ABEs), enable precise base alterations in the genome without inducing DNA double-stranded breaks (DSBs) (Komor et al., 2016; Nishida et al., 2016; Gaudelli et al., 2017). CBEs using a Cas9 variant fused with cytidine deaminase have enabled C-to-T conversion without requiring DSBs formation and homology-directed repair in mammalian cells (Komor et al., 2016). Currently, the most commonly used CBE, named BE3, consists of the rat cytidine deaminase APOBEC1 (rAPOBEC1) and uracil DNA glycosylase inhibitor (UGI) fused to Cas9 nickase (nCas9) (Li et al., 2017; Lu and Zhu, 2017; Ren et al., 2017; Zong et al., 2017). The BE3 system typically allows C-T substitution within a small editing window from C4 to C8 of the protospacer (Komor et al., 2016; Li et al., 2017). Several studies have reported the successful applications of CBE in several crop species including rice, maize, wheat, tomato, cotton, and rapeseed (Li et al., 2017; Zong et al., 2017, 2018; Qin et al., 2019; Wu et al., 2020). In addition, many other base editing systems have been developed in plants to improve gene editing accuracy and efficiency. A *Petromyzon marinus* cytidine deaminase (PmCDA1)-based CBE has resulted in efficient editing in rice and tomato (Shimatani et al., 2017). Moreover, a separate study has shown that PmCDA1 has higher base editing activity than rAPOBEC1 in rice (Tang et al., 2018). Zong et al. (2018) further improved CBEs by using the more effective human APOBEC3A (named A3A-PBE) which worked efficiently in wheat, rice and potato with a 17-nucleotide editing window, independent of sequence context.

During the preparation of our manuscript, Wu et al. (2020) reported the successful application of cytosine base-editing in rapeseed using rat cytidine deaminase APOBEC1. The editing efficiency was 1.8%, which is relatively lower when compared to other crops, and only one copy of *BnaALS* gene was edited (Wu et al., 2020). Cheng et al. (2020) successfully used A3A-PBE system to target *ALS*, *RGA*, and *IAA7* genes with an averaging editing efficiency of 23.6%, which also needs further improvement in the editing efficiency. In addition, they provided very limited information on the editing feature of the A3A-PBE system (Cheng et al., 2020). Therefore, further studies are required to establish more effective CBE systems in rapeseed based on the commonly used cytidine deaminases like rAPOBEC1, PmCDA1, and APOBEC3A.

In this study, we modified four CBE systems to achieve cytidine base editing at different genome sites in rapeseed. Five important genes with well-known functions, including *BnaCLV3*, *BnaRGA*, *BnaA3.IAA7*, *BnaDA1*, and *BnaALS*, were selected for precise base editing to improve agronomic traits in rapeseed (Liu et al., 2010; Li et al., 2015, 2019; Wang et al., 2017; Yang et al., 2018). Our results indicated that BnA3A1-PBE represents the best CBE editor in rapeseed at present, with the highest base-editing efficiency (up to 50.5%) and higher proportion of both homozygous and biallelic genotypes. The cytidine substitution that occurred at the target sites of *BnaRGA* and *BnaALS* were stably inherited and conferred expected phenotype in the T1 generation, indicating its powerful application prospect in rapeseed improvement.

METHODS

Vector Construction

To construct BnPBE and BnA3A-PBE vectors, cytidine deaminase (rAPOBEC1 or APOBEC3A), nCas9 and UGI units were amplified from pnCas9-PBE or A3A-PBE template plasmid (Zong et al., 2017), while the 35S promoter and ccdB units were amplified from PYLCRISPRCas9P35s-H (Ma et al., 2015). The resulting polymerase chain reaction (PCR) products were inserted into the *Pme*I/*BamH*I sites of binary vector PYLCRISPRCas9P35s-H through a Pro Ligation-Free Cloning Kit (Applied Biological Materials Inc, Canda, Cat.No.E086/E087). Furthermore, the cereal plant APOBEC3A sequences were codon-optimized for dicotyledon plant and synthesized commercially (Nanjing, China, GenScript) to create BnA3A1-PBE. The multiple sgRNA constructs were generated following a previous protocol used in combining sgRNAs to PYLCRISPRCas9P35s-H (Yang et al., 2018). Then, the multiple sgRNAs were amplified from the generating vector and the resulting PCR product was inserted into the *Asc*I sites of BnPBE, BnA3A-PBE and BnA3A1-PBE, using the Pro Ligation-Free Cloning Kit. To construct BnPBGE14 vector, nCas9 and PmCDA1 (Shimatani et al., 2017) were codon-optimized for dicotyledon plant and then synthesized commercially to replace Cas9 in PYLCRISPRCas9P35s-H. The multiple target sequences were synthesized and ligated to the *Bsa*I sites of PYLCRISPRCas9P35s-H. Primers used for vector construction are listed in **Supplementary Table 1**.

Agrobacterium-Mediated Rapeseed Transformation

Following verification of the fused constructs via sequencing, the CBE expressing binary vectors were transformed into

an elite cultivar (J9707) via the *Agrobacterium tumefaciens*-mediated hypocotyl method (Zhou and Fowke, 2002). Hygromycinselection (25 mg/L) was used to screen the transgenic plants.

On-Target Mutation Analysis by Targeted Deep Sequencing

Genomic DNA was extracted from the T_0 transgenic and wild type rapeseed plants using the CTAB method. The positive transgenic plants were screened by PCR using the specific primer pairs PB-L/PB-R (**Supplementary Table 1**). Then, the targeted mutations were determined in transgenic plants using the high-throughput tracking of mutations (Hi-TOM) platform (Liu et al., 2019). The sequencing analysis was conducted following the approach previously described by Zhai et al. (2020). The targets specific primer sets are listed in **Supplementary Table 1**.

Off-Target Analysis

Putative off-target sites, which contained 2–4-nucleotide mismatches relative to the *BnaCLV3* and *BnaRGA* target sites, were identified using Cas-OFFinder and CRISPR-P software (Bae et al., 2014; Liu et al., 2017). These potential off-target sites were detected in all edited T_0 transgenic plants using targeted deep sequencing. For each target gene, mixed genomic DNA from all T_0 editing plants was used as the template, and DNA of wild type plant was included as a control. All PCR products were purified and mixed in equal amounts (50 ng for each) as one sample. The DNA library construction, sequencing using the Illumina HiSeq 3000 system and data analysis were conducted according to the methods previously described by Yang et al. (2018). The independent sequence reads of each off-target site were aligned to the genomic wild type sequence, which covered each off-target site as a reference.

Herbicide Resistance Test

The T_1 mutants and wild type plants grown in the greenhouse (23°C, 16 h light/20°C, 8 h dark) were treated with commercial sulfonylureas at 1, 2, and 4 times field-recommended concentration (200, 400, and 800 mg/L). Representative pictures were taken 3 weeks after treatment.

RESULTS AND DISCUSSION

Design of Four CBE Systems and sgRNA Expression Cassettes

In this study, we adopted the base editor units (cytidine deaminase, nCas9 and UGI) from the PBE and A3A-PBE plasmid (Zong et al., 2018) to replace Cas9 in the pYLCRISPR/Cas9P35S-H binary vector (Yang et al., 2018), leading to the BnPBE and BnA3A-PBE systems, respectively. A codon-optimized APOBEC3A for *Brassica* plants was synthesized to optimize A3A-PBE, resulting in the creation of BnA3A1-PBE (**Figure 1**). PmCDA1 and nCas9 sequences (Shimatani et al., 2017) were codon-optimized for *Brassica* plants and synthesized to replace Cas9 in the pYLCRISPR/Cas9P35S-H binary vector, leading to the BnPBGE14 system (**Figure 1**). Thus, four CBE systems were modified to test cytidine base editing in *B. napus*. Three of them, including BnPBE, BnA3A-PBE, and BnA3A1-PBE, used the 35S promoter and the AtU3/AtU6 promoters to express the base editor unit and sgRNAs, respectively; while, BnPBGE14 use a 35S promoter to express both the base editor unit and sgRNAs in one ORF, in which multiple sgRNAs were further released using the tRNA-processing system-based strategy (**Figure 1**).

To investigate the feasibility and efficacy of these CBE systems in rapeseed, we designed 10 sgRNAs for five endogenous genes: sgRNA1 (S1) and S2 for *BnaCLV3*, S3 for *BnaRGA*, S4 for *BnaA3.IAA*, S5 to S7 for *BnaDA1*, and S8 to S10 for *BnaALS*. Then, four constructs for each CBE system were generated and introduced separately into the rapeseed variety J9707 through *Agrobacterium*-mediated transformation. An average of 78 independent T_0 transgenic lines were generated for each of the 15 CRISPR constructs (**Table 1**, **Supplementary Table 2**).

Detection of the Base-Editing of Different CBE Systems

Base-editing of the generated plants was assessed by Hi-TOM through sequencing of the sgRNA target sites (Liu et al., 2019). The observed mutations at the S1 and S2 sites in this study were considered as one due to the overlapping of the sgRNA sequences. It showed that three of CBEs (BnPBE, BnA3A-PBE, and BnA3A1-PBE) were active at all sgRNAs, except BnPBE showed no editing at the S4-to-S7 sites (**Figure 2**). And the BnPBGE14 was inactive at all sgRNAs (**Supplementary Table 2**). BnA3A1-PBE had the highest base-editing efficiency, with an average editing efficiency of 29.8%, which is ∼8.0% and fourfold higher than those of BnA3A-PBE (averaging 27.6%) and BnPBE (averaging 6.5%), respectively (**Figure 2**; **Table 1**). The C-to-T substitution efficiencies reached up to 50.5% in BnA3A1-PBE (**Figure 2**; **Table 1**), which is comparable with the efficiency in other crops (Qin et al., 2019). Recently, there were two reports about the successful application of different CBE systems in rapeseed [i.e., a PBE system (1.8% editing efficiency) reported by Wu et al. (2020) and an A3A-PBE system (averaging 23.6% editing efficiency) reported by Cheng et al. (2020)]. The performance of BnA3A1-PBE was much better than these two reported CBE systems, and thus represents the best CBE editor in rapeseed at present.

The overall base-editing efficiencies of the three CBE systems showed a similar trend at all sgRNAs [i.e., a higher editing efficiency at S1–S3, S9, and S10 and a lower editing efficiency at S4–S7 (**Figure 2**)]. In accordance with previous results, the difference in editing efficiency at these sgRNAs might be due to their nucleotide composition, GC content or promoter activities. Based on the fact that base editing with rAPOBEC1 is limited to a narrow deamination window and is inefficient in the GC contexts (Komor et al., 2016; Zong et al., 2018), this might be the reason why BnPBE has no editing activity at the S4–S7 sites.

Comparison of the Mutation Features in Different CBE Systems

Analyses of the base-editing efficiencies at every protospacer position across different sgRNAs revealed that the deamination window for BnA3A1-PBE spanned 17 nucleotides from

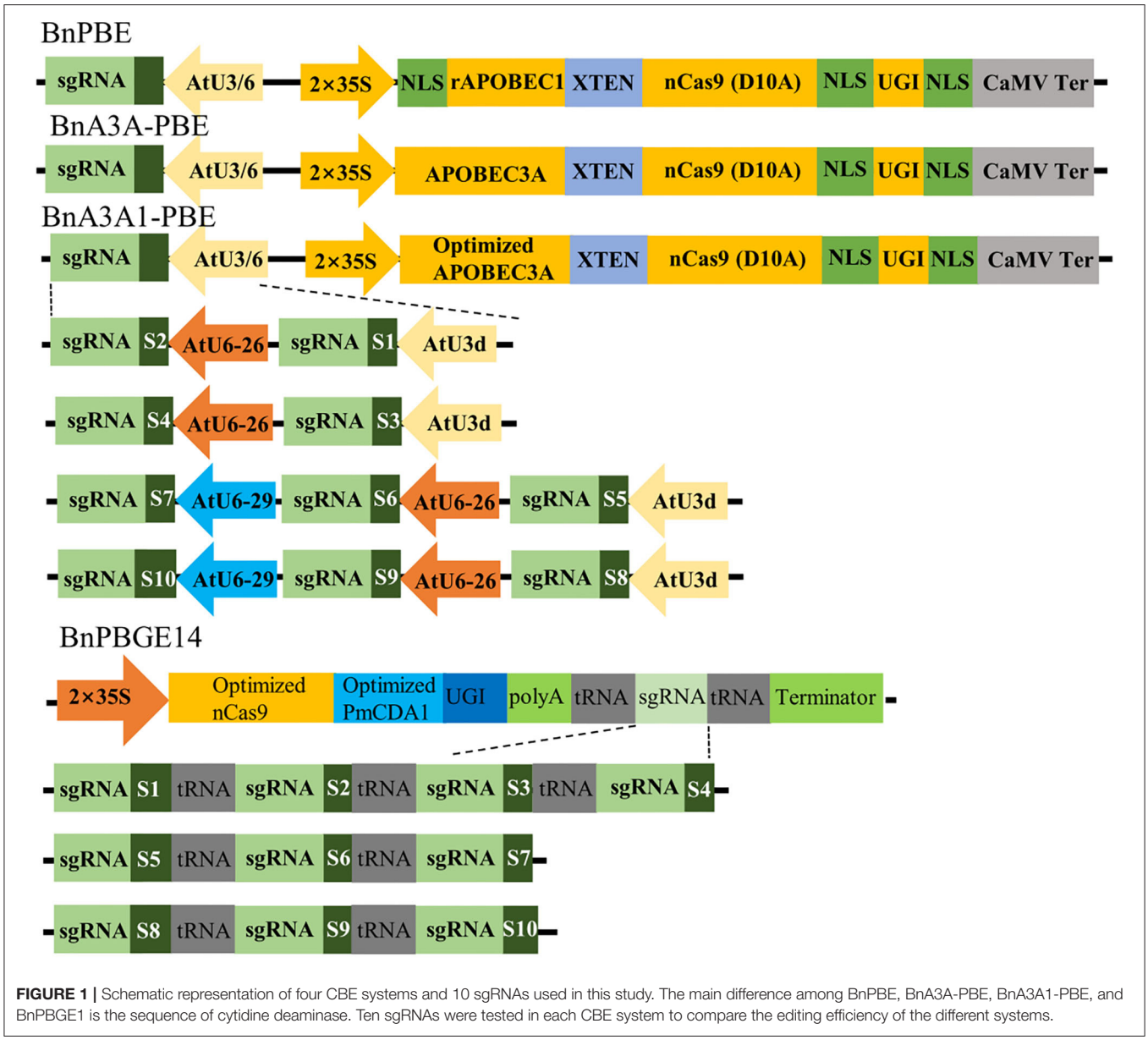


FIGURE 1 | Schematic representation of four CBE systems and 10 sgRNAs used in this study. The main difference among BnPBE, BnA3A-PBE, BnA3A1-PBE, and BnPBGE1 is the sequence of cytidine deaminase. Ten sgRNAs were tested in each CBE system to compare the editing efficiency of the different systems.

protospacer positions 2–18, compared with 2–16 for BnA3A-PBE, and 3–8 for BnPBE (**Figure 3**), which is consistent with previous reports (Zong et al., 2018). Furthermore, we found that the on-target editing products were different among the active three CBE systems: BnA3A-PBE and BnA3A1-PBE preferred to substitute more C into T simultaneously compared with BnPBE. For example, at the S10 site, BnPBE created four types of mutations with one to three substitutions (C6, C6C7 or C6C7C8), while BnA3A-PBE and BnA3A1-PBE created six types of mutations, where simultaneous editing of three or four Cs (C6C7C8 and C6C7C8C10) occurred more frequently (**Supplementary Table 3**). These results suggested that there was an obvious difference in the main mutation genotypes between rat-APOBEC1-based BnPBE and human-APOBEC3A-based BnA3A-PBE or BnA3A1-PBE. Thus, BnA3A-PBE and BnA3A1-PBE could increase the production of novel alleles with diverse genetic variations because of their broad editing window. Whereas, BnPBE could reduce the possibility of introducing undesired mutations at specific sites because of the narrow editing window. However, BnPBE has a lower editing efficiency than BnA3A-PBE and BnA3A1-PBE (**Figures 2**, **3**). Therefore, it is critical to fully understand the characteristics of these editing systems for better utilization.

By analyzing the ratio of different mutation genotypes in base-edited plants, the three active CBE systems produced mutants with a similar trend [i.e, heterozygous (Hetero) > homozygous (Homo) > biallelic (Bi) > chimerism (Ch) (**Figure 4**)]. Compared with BnPBE, BnA3A-PBE, and BnA3A1-PBE could

TABLE 1 | Detail information of the numbers of T_0 plants with different mutation types.

Name of vector	Target gene	sgRNA	No. of plants examined	No. of plants with INDEL	No. of plants with C-A/G base editing	No. of T_0 plants with C-T base editing				The ratio of edited T_0 plants
						HE	HO	Bi-allelic	Chimeric	
BnPBE-1	*BnaA04.CLV3*	sgRNA1 sgRNA2	107	0 (0.0%)	0 (0.0%)	9	0	0	0	7/107 (6.5%)
	BnaC04.CLV3	sgRNA1 sgRNA2	112	0 (0.0%)	0 (0.0%)	2	1	0	0	3/112 (2.7%)
BnA3A-PBE1	*BnaA04.CLV3*	sgRNA1 sgRNA2	81	0 (0.0%)	7 (9.1%)	21	9	12	5	47/81 (58.0%)
	BnaC04.CLV3	sgRNA1 sgRNA2	77	3 (3.9%)	5 (6.5%)	11	11	20	2	44/77 (57.1%)
BnA3A1-PBE1	*BnaA04.CLV3*	sgRNA1 sgRNA2	86	1 (1.2%)	7 (8.1%)	15	9	12	6	42/86 (48.8%)
	BnaC04.CLV3	sgRNA1 sgRNA2	84	2 (2.4%)	2 (2.4%)	9	14	8	3	34/84 (40.5%)
BnPBE-2	*BnaA06.RGA*	sgRNA3	87	0 (0.0%)	0 (0.0%)	21	1	6	0	28/87 (32.2%)
BnA3A-PBE2	*BnaA06.RGA*	sgRNA3	80	4 (5.0%)	2 (2.5%)	7	12	11	5	35/80 (43.8%)
BnA3A1-PBE2	*BnaA06.RGA*	sgRNA3	95	11 (11.6%)	4 (4.2%)	8	10	24	6	48/95 (50.5%)
BnPBE-2	*BnaA03.IAA7*	sgRNA4	86	0 (0.0%)	0 (0.0%)	1	0	0	0	1/86 (1.2%)
BnA3A-PBE2	*BnaA03.IAA7*	sgRNA4	83	1 (1.2%)	0 (0.0%)	10	0	3	1	14/83 (16.8%)
BnA3A1-PBE2	*BnaA03.IAA7*	sgRNA4	95	7 (7.4%)	2 (2.1%)	13	2	5	1	21/95 (22.1%)
BnPBE-3	*BnaA06DA1*	sgRNA5	93	0 (0.0%)	0 (0.0%)	0	0	0	0	0 (0.0%)
	BnaC05DA1	sgRNA6	93	0 (0.0%)	0 (0.0%)	0	0	0	0	0 (0.0%)
	BnaA08.DA1	sgRNA7	93	0 (0.0%)	0 (0.0%)	0	0	0	0	0 (0.0%)
	BnaC08.DA1	sgRNA5	93	0 (0.0%)	0 (0.0%)	0	0	0	0	0 (0.0%)
BnA3A-PBE3	*BnaA06.DA1*	sgRNA5	79	3 (3.8%)	0 (0.0%)	4	1	4	0	9/79 (11.4%)
	BnaC05.DA1	sgRNA6	79	4 (5.1%)	1 (1.3%)	4	1	0	1	6/79 (7.6%)
	BnaA08.DA1	sgRNA7	83	1 (1.2%)	0 (0.0%)	2	0	0	2	4/83 (4.8%)
	BnC08.DA1	sgRNA5	83	1 (1.2%)	0 (0.0%)	9	2	0	0	11/83 (13.3%)
BnA3A1-PBE3	*BnaA06.DA1*	sgRNA5	49	1 (2.0%)	0 (0.0%)	3	1	2	0	6/49 (12.2%)
	BnaC05.DA1	sgRNA6	49	2 (4.1%)	0 (0.0%)	3	1	1	0	5/49 (10.2%)
	BnaA08.DA1	sgRNA7	49	1 (2.0%)	0 (0.0%)	2	3	1	0	6/49 (12.2%)
	BnaC08.DA1	sgRNA5	49	1 (2.0%)	0 (0.0%)	7	0	1	0	8/49 (16.3%)
BnPBE-4	*BnaALS3*	sgRNA8	159	0 (0.0%)	0 (0.0%)	2	0	0	0	2/159 (1.3%)
	BnaALS3	sgRNA10	159	0 (0.0%)	0 (0.0%)	8	0	2	0	10/159 (6.3%)
	BnaALS1	sgRNA9	159	0 (0.0%)	0 (0.0%)	8	1	0	0	9/159 (5.7%)
	BnaALS1	sgRNA10	159	2 (1.3%)	0 (0.0%)	9	0	0	2	11/159 (7.0%)
BnA3A-PBE4	*BnaALS3*	sgRNA8	107	1 (0.9%)	0 (0.0%)	9	0	0	1	10/107 (9.3%)
	BnaALS3	sgRNA10	106	7 (6.6%)	0 (0.0%)	22	11	9	2	44/106 (41.5%)
	BnaALS1	sgRNA9	108	3 (2.8%)	1 (0.9%)	25	15	0	0	40/108 (37.0%)
	BnaALS1	sgRNA10	108	10 (9.3%)	1 (0.9%)	21	13	11	2	47/108 (43.5%)
BnA3A1-PBE4	*BnaALS3*	sgRNA8	80	0 (0.0%)	0 (0.0%)	14	4	0	0	18/80 (22.5%)
	BnaALS3	sgRNA10	80	0 (0.0%)	0 (0.0%)	18	6	6	0	30/80 (37.5%)
	BnaALS1	sgRNA9	81	0 (0.0%)	0 (0.0%)	18	9	3	0	30/81 (37.0%)
	BnaALS1	sgRNA10	81	1 (1.2%)	0 (0.0%)	16	6	10	1	33/81 (40.7%)

significantly reduce the proportion of heterozygous genotypes and increase the proportion of both homozygous and biallelic genotypes (**Figure 4**).

We compared the base-editing efficiencies of these sgRNA sites that could target both the A and C subgenomes of rapeseed in this study. The data showed that all the CBEs had uniform editing rate between the two subgenomes for six sgRNAs, whereas an obvious bias of base editing at S8 and S9 was observed in the C-subgenome (26.7% on average) compared to that in the A-subgenome (11.0% on average). More than 80% of base editing in BnPBE occurred at only C subgenomes, whereas more than 56.7 and 61.7% of base editing occurred simultaneously at both

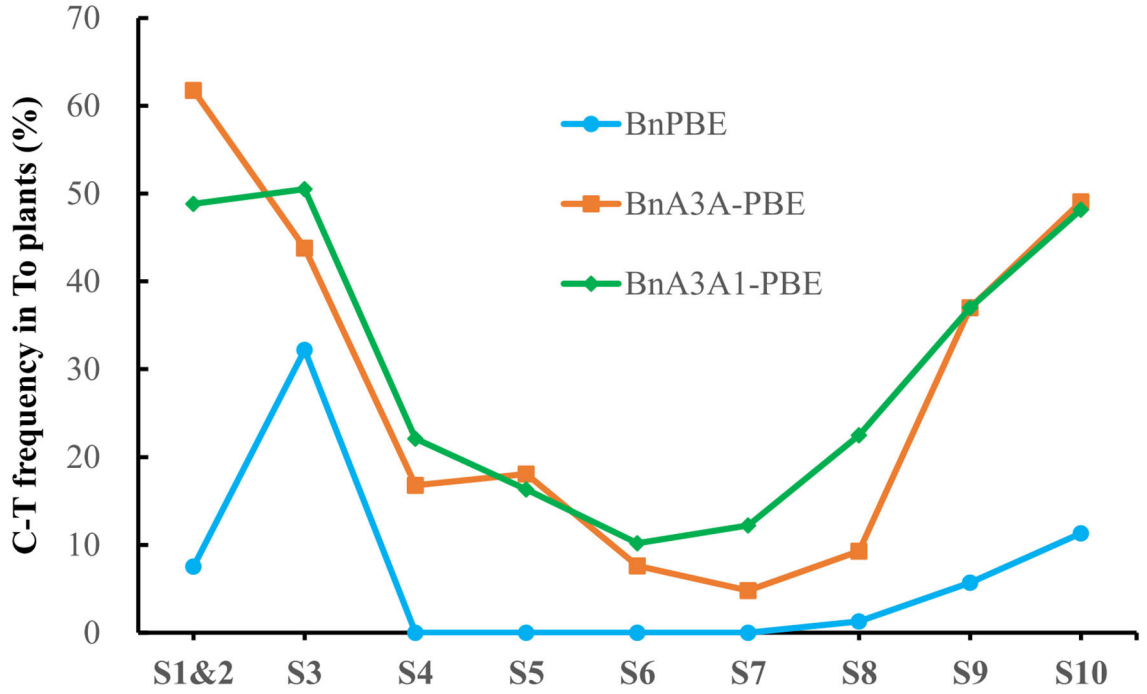


FIGURE 2 | Frequencies of C-to-T conversion using three active CBE systems at 10 sgRNAs in T_0 plants.

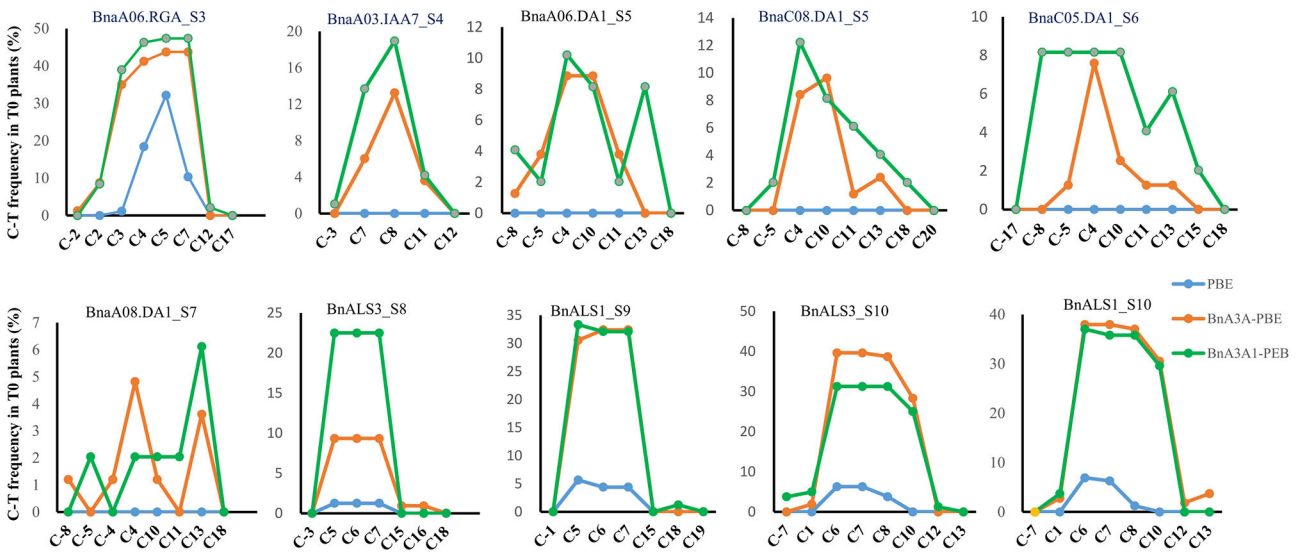


FIGURE 3 | Frequencies of single C-to-T conversions by BnPBE, BnA3A-PBE, and BnA3A1-PBE at 10 target sites in T_0 plants. Substitution efficiency was calculated by the ratio of the edited plants in the total transgenic positive plants.

subgenomes for BnA3A-PBE and BnA3A1-PBE, respectively (**Table 1**).

The overall unexpected nucleotide changes and indels in the putative editing window occurred with much lower frequencies than C-to-T base editing, and BnA3A-PBE and BnA3A1-PBE yielded relatively higher frequencies of these undesired edits than that those observed in BnPBE (**Table 1**). This showed that the frequency of undesired edits was positively correlated with the editing efficiency of CBE systems.

Off-Target Activity of the CBE Systems in T_0 Transgenic Rapeseed Plants

To detect any potential off-target effects of the CBE systems reported here, we selected the target sites with the highest editing efficiency corresponding to the three active CBE systems

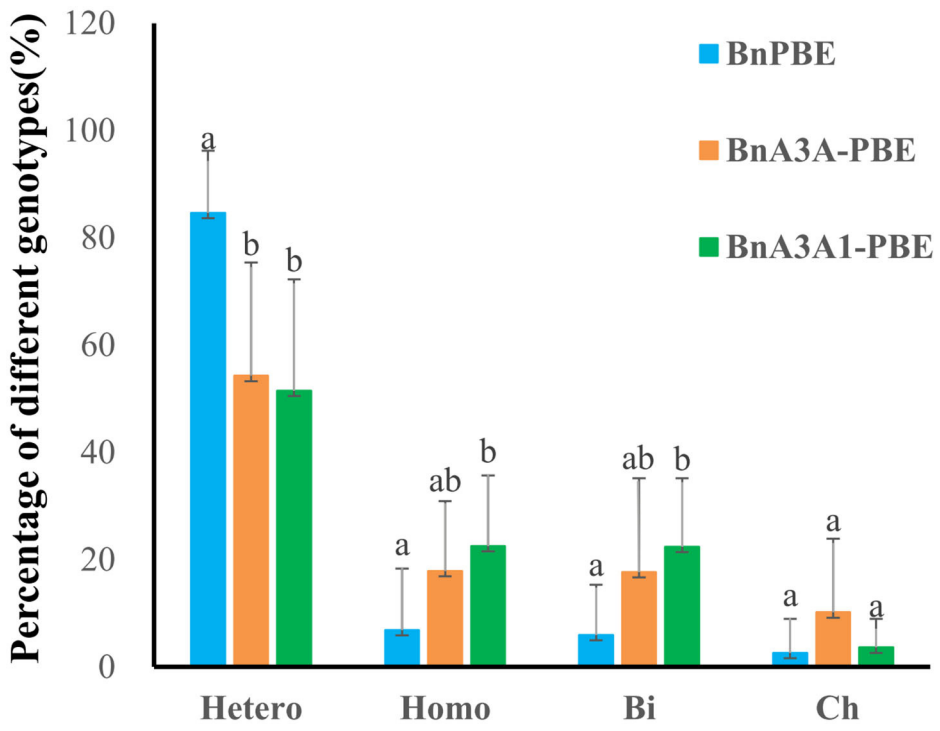


FIGURE 4 | Percentage of four genotypes of base mutations in 10 targets using three CBE systems. Uppercase letters indicate a significant difference at the 0.05 probability level among the three CBE systems based on a multiple comparison test. Hetero, heterozygous; Homo, homozygous; Bi, biallelic; Ch, chimerism.

to detect the off-target efficiency. There were 23 and four potential off-target sites identified for *BnaCLV3* and *BnaRGA* sites, respectively, with up to 4-nucleotide mismatches (Bae et al., 2014; Liu et al., 2017). High-throughput sequencing of the PCR products of these potential off-target sites revealed that no significant difference was observed in the off-target ratio between the base-edited and wild-type plants (**Supplementary Table 4**). These results revealed that the three active CBE systems have a high specificity for targeted mutagenesis in rapeseed, which is consistent with previous reports in animals and plants (Kim et al., 2017; Qin et al., 2019; Cheng et al., 2020).

The Base Editing of the *BnaRGA* Gene Produced an Expected Dwarf Phenotype

The S3 target site was fully matched with a functional copy of *BnaA06.RGA* (*BnaA06g34810D*) which encodes a DELLA protein, serving as a Gibberellins (GA) signaling repressor. A C-to-T substitution conferred a mutation (P91L) in its TVHYNP motif and resulted in a dwarf phenotype (Liu et al., 2010). Sequencing analysis revealed that 108 (41.2%) of the T_0 plants contained a C-to-T substitution which occurred at C2, C3, C4, C5, C7, and C12 from the protospacer position at the S3 site (**Table 1**, **Supplementary Table 3**). A total of 14 different mutation genotypes were detected from the 108 edited lines in the T_0 generation, and the homozygous substitutions at the conserved P91 in three different lines showed obvious dwarf phenotypes (**Figure 5A**).

To obtain stable homozygous mutants and test whether the base-editing mutants are inherited, two independent heterozygous lines (BnPBE3-2-8 and BnA3A-PBE-2-86) were self-pollinated. Then, their respective T_1 progeny were genotyped via Hi-TOM sequencing at the S3 site. In the T_1 progeny from the BnPBE3-2-8 and BnA3A-PBE-2-86 lines, the segregation ratio observed for the heterozygous, homozygous and wild type genotype was ∼1:2:1 (χ^2 = 0.15 and 0.11, P > 0.05; **Figure 5C**). These results indicated that the produced base substitution was successfully transmitted from T_0 to T_1 generation, with an expected monogenic segregation pattern. Furthermore, the PCR assay was performed to detect exogenous T-DNA using the PB-L/R primer pairs (**Figure 5B**). Twenty edited mutant plants, including seven homozygous mutants without exogenous T-DNA were obtained in the T_1 generation (**Figure 5C**).

Indeed, several T_1 plants with the expected P91L or novel P91F&A92V substitutions showed a decreased plant height compared with wild-type plants (**Figures 5D–F**). The significant reduction in height was due to a lower first branch position and shorter internodes compared with wild-type plants. Besides, we found that the heterozygous mutants also show a significant reduction in plant height. Previous report showed that the target substitution in *BnaC09.RGA* conserved domain generated dwarf phenotype (Cheng et al., 2020). Altogether, we can conclude that both the functional copies of *BnaA06.RGA* and *BnaC09.RGA* can achieve gain-of-function mutations at the conserved P91 through CBE system. The utilization of these semi-dwarf mutants

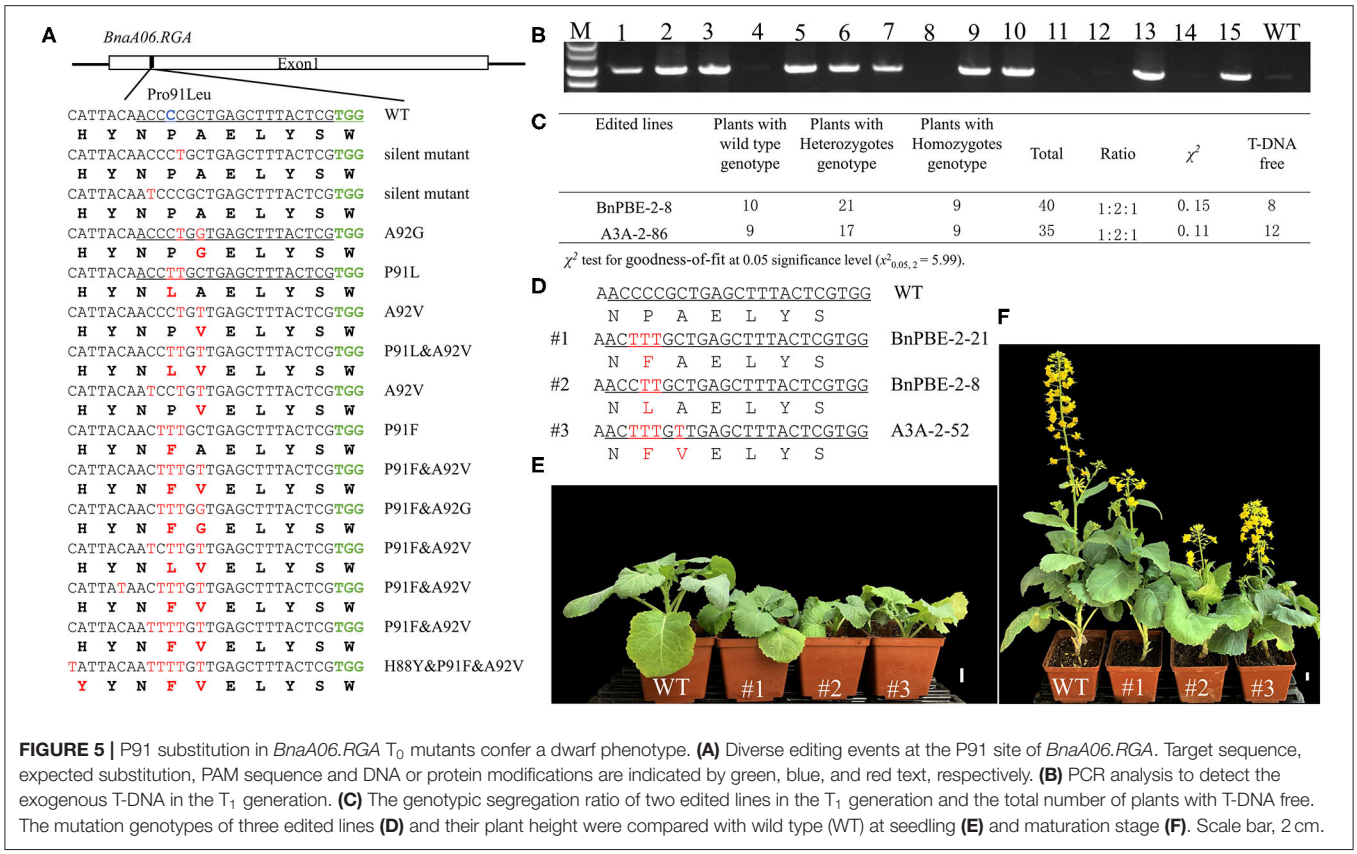


FIGURE 5 | P91 substitution in *BnaA06.RGA* T_0 mutants confer a dwarf phenotype. **(A)** Diverse editing events at the P91 site of *BnaA06.RGA*. Target sequence, expected substitution, PAM sequence and DNA or protein modifications are indicated by green, blue, and red text, respectively. **(B)** PCR analysis to detect the exogenous T-DNA in the T_1 generation. **(C)** The genotypic segregation ratio of two edited lines in the T_1 generation and the total number of plants with T-DNA free. The mutation genotypes of three edited lines **(D)** and their plant height were compared with wild type (WT) at seedling **(E)** and maturation stage **(F)**. Scale bar, 2 cm.

produced in the study could improve the lodging resistance in rapeseed breeding.

The Base Editing of the *BnaALS* Gene Produced Herbicide Resistance Rapeseed

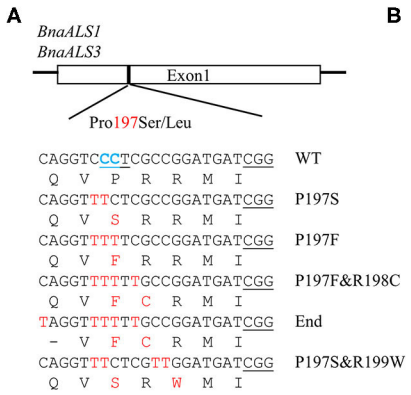
In the edited plants, we were excited to obtain expected substitutions at the conserved P197 site of the acetolactate synthase gene (*BnaALS*) targeted by S10, which probably confer resistance to sulfonylurea herbicides (Li et al., 2015). Sequencing results revealed that 101 (29.0%) of the T_0 plants contained a C-to-T substitution at the S10 site, which occurred at C1, C6, C7, C8, and C10 from the protospacer position (**Table 1**). Among these mutant plants, 57 harbored missense mutations in both functional copies of *BnaALS*, among which 21 and 23 had missense mutations in *BnaC01g25380D* (*BnaALS1*) and *BnaA01g20380D* (*BnaALS3*), respectively. Diverse editing events were detected at the target sites of *BnaALS1* and *BnaALS3* (**Figure 6A**). The expected amino acid substitution (P197S or P197L) was rarely detected in these mutants, whereas most of the editing lines carried P197F and P197F&R198C substitutions (**Supplementary Table 3**).

To test whether the observed mutations are stably inherited and obtain stable homozygous mutants, four independent T_0 editing lines of *BnaALS* were self-pollinated to produce T_1 progeny. The target mutations of progenies from these T_0 lines were verified by Hi-TOM sequencing analysis of the target sites. As expected, the observed base substitutions were transmitted to the T_1 generation, and different single and double mutants were obtained (**Figure 6B**).

To determine whether these P197 missense mutations in the *BnaALS* conferred sulfonylureas herbicide resistance, the T_1 mutants with homozygous P197-substitutions at a single (AAcc, aaCC) or double (aacc) copies of *BnaALS* gene were treated with various field application levels of tribenuron-methyl at the four-leaf stage. The mutants carrying the P197F edited alleles grew better at 1, 2, and 4 times field-recommended rates (200, 400, and 800 mg/L) over three weeks compared with wild type (**Figures 6C,D**), the resistance of aacc mutant was the best, followed by AAcc and aaCC mutants at the field application levels of 200 to 800 mg/L sulfonylurea herbicides. Thus, both copies of *BnaALS* gene likely confer herbicide resistance with a similar effect and work in an additive manner (**Figure 6D**), which is different from the report that *BnaALS3* confer better herbicide resistance than *BnaALS1* (Cheng et al., 2020). Thus, the P197F substitution represents a novel allele which confers herbicide resistance in rapeseed.

Utilization of Base Editor as a Toolkit for the Insertion of Stop Codon

There are five homologous copies of *BnaIAA7* in the *B. napus* genome, and the S4 was designed to fully target four out of the five gene copies, with *BnaC05.IAA7* having one base mismatched (**Supplementary Figure 1A**). A G-to-A mutation changed the glycine at the 84th position to glutamic acid (G84E)



B

Plant ID	Generation	Genotype at targets of *BnaA01.ALS*	Genotype at targets of *BnaC01.ALS*
A3A1-4-88	T_0	Bi-allelic (C6C7C8>T6T7T8; C6C7C8C10>T6T7T8T10)	Bi-allelic (C6C7C8>T6T7T8; C6C7C8C10>T6T7T8T10)
A3A1-4-88-2	T_1	Homo (C6C7C8C10>T6T7T8T10)	Homo (C6C7C8C10>T6T7T8T10)
A3A1-4-88-3	T_1	Bi-allelic (C6C7C8>T6T7T8; C1C6C7C8C10>T1T6T7T8T10)	Homo (C6C7C8C10>T6T7T8T10)
A3A1-4-66	T_0	WT	Bi-allelic (C6C7C8>T6T7T8; C6C7C8C10>T6T7T8T10)
A3A1-4-66-1	T_1	WT	Homo (C6C7C8>T6T7T8)
A3A1-4-66-3	T_1	WT	Homo (C6C7C8C10>T6T7T8T10)
A3A-4-64	T_0	Hetero (C6C7C8>T6T7T8)	Hetero (C6C7C8C10>T6T7T8T10)
A3A-4-64-1	T_1	WT	Homo (C6C7C8C10>T6T7T8T10)
A3A-4-64-5	T_1	Homo (C6C7C8>T6T7T8)	WT
A3A-4-64-18	T_1	Homo (C6C7C8>T6T7T8)	Homo (C6C7C8C10>T6T7T8T10)
A3A-4-63	T_0	Hetero (C6C7C8C10>T6T7T8T10)	WT
A3A-4-63-2	T_1	Homo (C6C7C8C10>T6T7T8T10)	WT

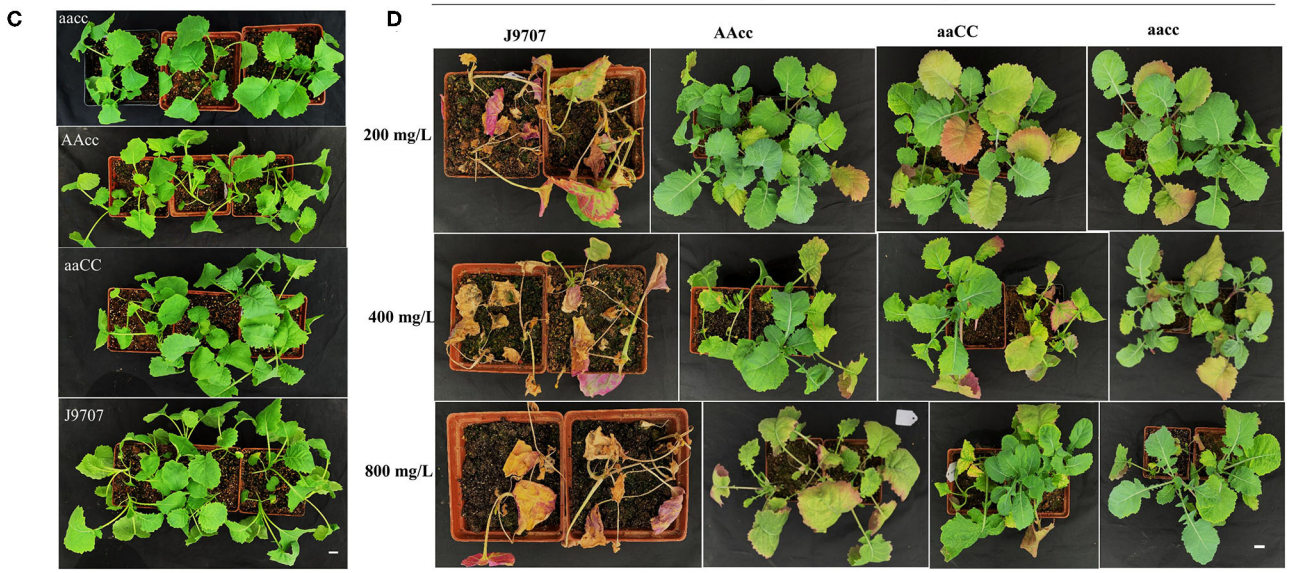


FIGURE 6 | P197F substitution in *BnaALS* T_0 mutants confer resistance to sulfonylureas herbicides. **(A)** Diverse co-editing events at the P197 site of *BnaALS*. The BnaALS-P197 site is conserved in the **(A,C)** subgenomes and 197 was numbered according to the corresponding sequence of Arabidopsis. The PAM were underlined in black line **(B)** Representative P197F substitution in *BnaALS* T_0 mutants and their transmission to T_1 generation. Hetero, heterozygous; Homo, homozygous; Bi, biallelic; WT, wild type. **(C,D)** Phenotypes of base-edited plants with different alleles and wild type before and after being tested with 200, 400, and 800 mg/L tribenuron-methyl herbicide in the T_1 generation. "aaCC," "AAcc," and "aacc" represent homozygous mutations of the target gene in *BnaALS3*, *BnaALS1* and both copies, respectively. Scale bar, 2 cm.

in *BnaA3.IAA7*, which contributes to reducing the length of internodes and branch angles in rapeseed (Li et al., 2019). As expected, this substitution caused the conversion of the conserved Gly84 to glutamic acid (**Supplementary Figure 1B**). However, all the edited plants carry a G-to-A mutation at C8, which results in the insertion of a stop codon at the 85th position (**Supplementary Figure 1B**). Our results suggested that base editing can also be utilized to create knockout mutations by insertion of a stop codon which results in premature termination.

The NG Protospacer Adjacent Motif Greatly Broaden the Targeting Scope of Base Editing in Rapeseed

In the *B. napus* genome, there are four copies of *BnaDA1* gene. S5 was designed to target both *BnaA06.DA1* and *BnaC08.DA1*, while S6 and S7 were designed to target *BnaC05.DA1* and *BnaA08.DA1*, respectively. We designed these targets intending to obtain base substitutions at the conserved Arg358 of the *BnaDA1* (DA means big in Chinese) targeted by S5–S7. The A358K conversion of *BnaDA1* probably contributes to the improvement of seed weight in rapeseed (Wang et al., 2017). S8 and S9 were designed to target *BnaALS3* and *BnaALS1*, respectively. The alanine at position 122 of *ALS* is converted to valine, which endow mutant resistance to imidazolinone herbicide (Li et al., 2008; Sala et al., 2008; Han et al., 2012). In the T_0 edited plants, only 11 edited plants contained the intended R358K conversion at S5–S7 and only 1 edited plant contained the intended A122V conversion at S8–S9 since the target bases (C13 or C15) are located outside of the hot spot of the deamination window (**Figure 3**; **Supplementary Table 5**). Thus, it is imperative to develop an engineered SpCas9 variant

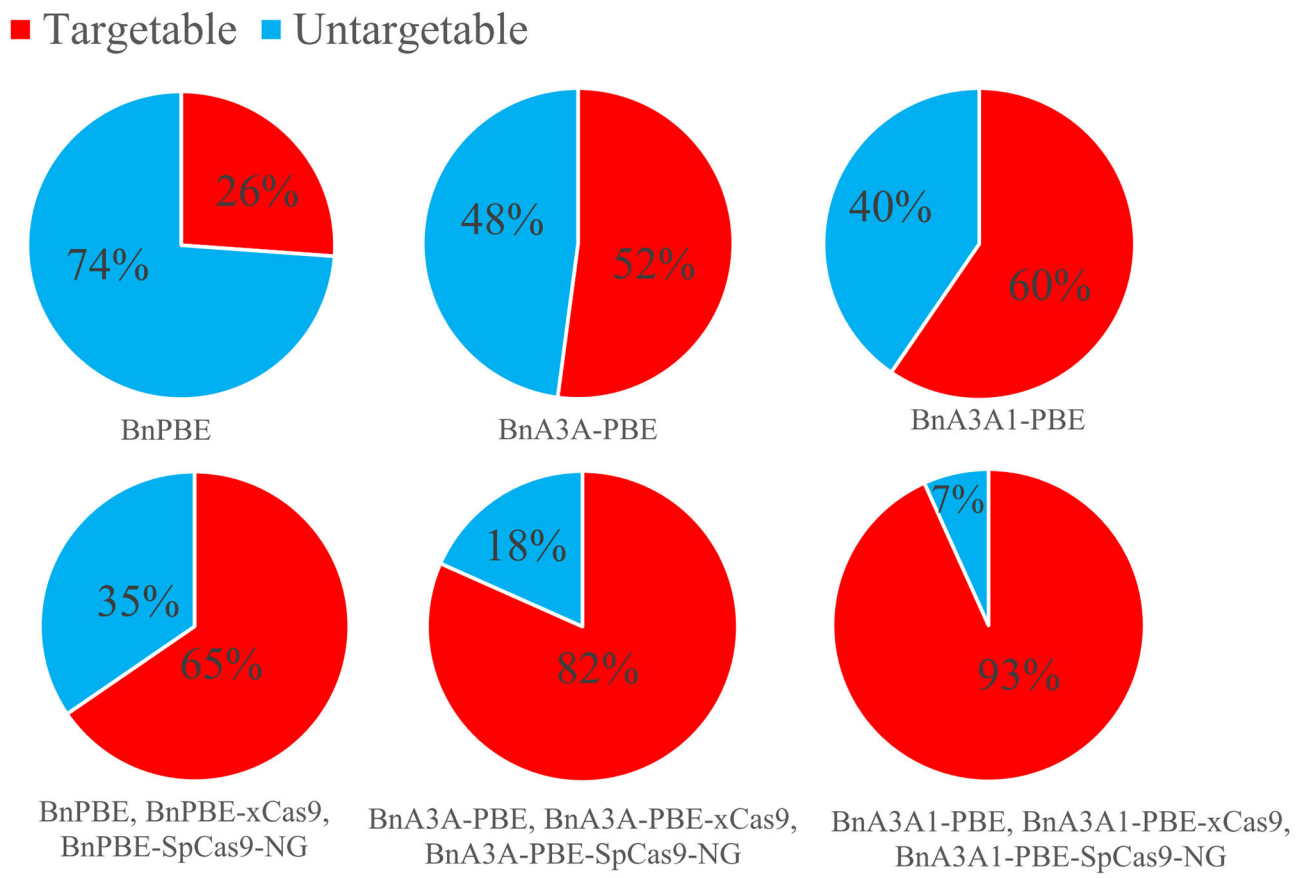


FIGURE 7 | Computational analysis of the rapeseed genome revealed that Cytidines (NGG PAM) and guanidines (CCN PAM) can potentially be edited by BnPBE, BnA3A-PBE, and BnA3A1-PBE. The above three CBEs together with the various Cas9 variants (SpCas9 and xCas9) promise to increase the scope of base editing of targeted cytidines and guanidines in rapeseed genome.

that recognizes not only the NGG protospacer adjacent motif (PAM), to ensure that the desired target bases are located within the hot spot of the editing window. The seed weight of the 11 edited plants containing the intended R358K conversion at S5–S7 will be tested in the next generation since there were limited seeds to conduct a field experiment in the T_1 plants.

Streptococcus pyogenes Cas9 (SpCas9) recognizes a very simple NGG PAM, making it the most commonly used CRISPR-Cas9 system. The canonical NGG PAM limits its targeting scope in a genome, especially for applications that require precise Cas9 positioning such as base editing (Wang et al., 2019). Recently, the engineered SpCas9 variant, SpCas9-NG, which recognize NG PAMs are more efficient than the xCas9 variant (Hu et al., 2018; Nishimasu et al., 2018), Moreover, SpCas9-NG coupled with the activation-induced cytidine deaminase (AID) can mediate the C-to-T conversion at target sites with NG PAMs in human cells (Nishimasu et al., 2018). Endo et al. (2019) reported that the SpCas9-NG can efficiently mutagenize endogenous target sites with NG PAMs in rice and Arabidopsis genomes. For *B. napus* genome, the 17-nt editing window of BnA3A1-PBE theoretically increases up to 15.3% and 1.3-fold the number of genomic cytidines and guanidine available for base editing when compared to BnA3A-PBE and BnPBE, respectively (**Figure 7**). When combined with SpCas9, xCas9, and other variants with NG PAM, BnA3A1-PBE theoretically targets 93% of the cytidines and guanidine in the rapeseed genome (**Figure 7**), which makes it as an ideal editing system for further improvement in future research.

In conclusion, the three CBEs described here can efficiently and specifically perform precise C-to-T substitutions across a broad range of endogenous genomic loci in rapeseed. The improved BnA3A1-PBE performed efficiently as a base editor with higher editing efficiency, a more broadened editing window coupled with a higher proportion of homozygous and biallelic genotypes compared with BnA3A-PBE and BnPBE systems. When compared with the two latest studies which reported the successful application of cytosine base-editing in rapeseed (Wu et al., 2020, Cheng et al., 2020), BnA3A1-PBE has a high editing efficiency, which is ∼16-fold than PBE (Wu et al., 2020), and 26% higher than those of A3A-PBE (Cheng

et al., 2020). Thus, BnA3A1-PBE represents a powerful base editor for both gene function study and molecular breeding in rapeseed.

DATA AVAILABILITY STATEMENT

Sequencing data has been deposited in NCBI Sequence Read Archive (SRA) database under the accession no. PRJNA663415.

AUTHOR CONTRIBUTIONS

LH, CF, and YZho: conceived and designed the experiments. LH, QL, OA, SC, MZ, XS, KY, YZha, YY, and LX: performed the experiments. LH and CF: wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the National Natural Science Foundation of China (Nos. 31371240 and 31671279). This study was also supported by Grants 2015CB150200 from the National Key Basic Research Program of China.

ACKNOWLEDGMENTS

We thank Dr. Caixia Gao of the Institute of Genetics and Developmental Biology, CAS, for her kind gift of the plasmid A3A-PBE for this research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgeed.2020.605768/full#supplementary-material

Supplementary Figure 1 | The desired substitution accompany with a stop codon occurred at the S4 target site. **(A)** The target sequences are shown with the PAM underlined in red line. **(B)** Diverse co-editing events at the BnaIAA7-Gly84 site, all of these new alleles introduce an early stop codon.

Supplementary Table 1 | The primers used in this study.

Supplementary Table 2 | Detail information on the numbers of T_0 plants in BnPBG14 system.

Supplementary Table 3 | Detail information of the mutation of all edited T_0 plants in this study.

Supplementary Table 4 | Summary of deep sequencing data for most likely off-target sites of S1, S2, and S3.

Supplementary Table 5 | The edited lines contain the intended C13 to T13 conversion at S5-S7.

REFERENCES

Bae, S., Park, J., and Kim, J. S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* 30, 1473–1475. doi: 10.1093/bioinformatics/btu048

Braatz, J., Harloff, H. J., Mascher, M., Stein, N., Himmelbach, A., and Jung, C. (2017). CRISPR-Cas9 targeted mutagenesis leads to simultaneous modification of different homoeologous gene copies in polyploid oilseed rape (*Brassica napus*). *Plant Physiol*. 174, 935–942. doi: 10.1104/pp.17.00426

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., et al. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* 345, 950–953. doi: 10.1126/science.1253435

Cheng, H., Hao, M., Ding, B., Mei, D., Wang, W., Wang, H., et al. (2020). Base editing with high efficiency in allotetraploid oilseed rape by A3A-PBE base-editing system. *Plant Biotechnol. J.* doi: 10.1111/pbi.13444. [Epub ahead of print].

Endo, M., Mikami, M., Endo, A., Kaya, H., Itoh, T., Nishimasu, H., et al. (2019). Genome editing in plants by engineered CRISPR-Cas9 recognizing NG PAM. *Nat. Plants*. 5, 14–17. doi: 10.1038/s41477-018-0321-8

Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I., et al. (2017). Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage. *Nature* 551, 464–471. doi: 10.1038/nature24644

Han, H., Yu, Q., Purba, E., Li, M., Walsh, M., Friesen, S., et al. (2012). A novel amino acid substitution Ala-122-Tyr in ALS confers high-level and broad resistance across ALS-inhibiting herbicides. *Pest Manag. Sci.* 68, 1164–1170. doi: 10.1002/ps.3278

Hu, J. H., Miller, S. M., Geurts, M. H., Tang, W., Chen, L., Sun, N., et al. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556, 57–63. doi: 10.1038/nature26155

Kim, D., Lim, K., Kim, S. T., Yoon, S. H., Kim, K., Ryu, S. M., et al. (2017). Genome-wide target specificities of CRISPR RNA-guided programmable deaminases. *Nat. Biotechnol*. 35, 475–480. doi: 10.1038/nbt.3852

Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. doi: 10.1038/nature17946

Li, D., Barclay, I., Jose, K., Stefanova, K., and Appels, R. (2008). A mutation at the Ala122 position of acetohydroxyacid synthase (AHAS) located on chromosome 6D of wheat: improved resistance to imidazolinone and a faster assay for marker-assisted selection. *Mol. Breeding* 22, 217–225. doi: 10.1007/s11032-008-9168-4

Li, H., Li, J., Song, J., Zhao, B., Guo, C., Wang, B., et al. (2019). An auxin signaling gene *BnaA3.IAA7* contributes to improved plant architecture and yield heterosis in rapeseed. *New Phytol*. 222, 837–851. doi: 10.1111/nph.15632

Li, H., Li, J., Zhao, B., Wang, J., Yi, L., Liu, C., et al. (2015). Generation and characterization of tribenuron-methyl herbicide-resistant rapeseed (*Brasscia napus*) for hybrid seed production using chemically induced male sterility. *Theor. Appl. Genet*. 128, 107–118. doi: 10.1007/s00122-014-2415-7

Li, J., Sun, Y., Du, J., Zhao, Y., and Xia, L. (2017). Generation of targeted point mutations in rice by a modified CRISPR/Cas9 system. *Mol. Plant*. 10, 526–529. doi: 10.1016/j.molp.2016.12.001

Liu, C., Wang, J., Huang, T., Wang, F., Yuan, F., Cheng, X., et al. (2010). A missense mutation in the VHYNP motif of a DELLA protein causes a semi-dwarf mutant phenotype in *Brassica napus*. *Theor. Appl. Genet*. 121, 249–258. doi: 10.1007/s00122-010-1306-9

Liu, H., Ding, Y., Zhou, Y., Jin, W., Xie, K., and Chen, L. L. (2017). CRISPR-P 2.0: an improved CRISPR-Cas9 tool for genome editing in plants. *Mol. Plant*. 10, 530–532. doi: 10.1016/j.molp.2017.01.003

Liu, Q., Wang, C., Jiao, X., Zhang, H., Song, L., Li, Y., et al. (2019). Hi-TOM: a platform for high-throughput tracking of mutations induced by CRISPR/Cas systems. *Sci. China Life Sci*. 62, 1–7. doi: 10.1007/s11427-018-9402-9

Lu, Y., and Zhu, J. (2017). Precise editing of a target base in the rice genome using a modified CRISPR/Cas9 system. *Mol. Plant*. 10, 523–525. doi: 10.1016/j.molp.2016.11.013

Ma, X., Zhang, Q., Zhu, Q., Liu, W., Chen, Y., Qiu, R., 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

Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., et al. (2016). Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* 353:aaf8729. doi: 10.1126/science.aaf8729

Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., et al. (2018). Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* 361, 1259–1262. doi: 10.1126/science.aas9129

Qin, L., Li, J., Wang, Q., Xu, Z., Sun, L., Alariqi, M., et al. (2019). High efficient and precise base editing of C· G to T· A in the allotetraploid cotton (*Gossypium hirsutum*) genome using a modified CRISPR/Cas9 system. *Plant Biotechnol. J.* 18, 45–56. doi: 10.1111/pbi.13168

Ren, B., Yan, F., Kuang, Y., Li, N., Zhang, D., Lin, H., et al. (2017). A CRISPR/Cas9 toolkit for efficient targeted base editing to induce genetic variations in rice. *Sci. China Life Sci.* 60, 516–519. doi: 10.1007/s11427-016-0406-x

Russell, W. L., Russell, L. B., and Kelly, E. (1958). Radiation dose rate and mutation frequency. *Science* 128, 1546–1550. doi: 10.1126/science.128.3338.1546

Sala, C., Bulos, M., Echarte, M., Whitt, S. R., and Ascenzi, R. (2008). Molecular and biochemical characterization of an induced mutation conferring imidazolinone resistance in sunflower. *Theor. Appl. Genet.* 118, 105–112. doi: 10.1007/s00122-008-0880-6

Sega, G. A. (1984). A review of the genetic effects of ethyl methanesulfonate. *Mutat. Res.* 134, 113–142. doi: 10.1016/0165-1110(84)90007-1

Shimatani, Z., Kashojiya, S., Takayama, M., Terada, R., Arazoe, T., Ishii, H., et al. (2017). Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. *Nat. Biotechnol.* 35, 441–443. doi: 10.1038/nbt.3833

Tang, X., Ren, Q., Yang, L., Bao, Y., Zhong, Z., He, Y., et al. (2018). Single transcript unit CRISPR 2.0 systems for robust Cas9 and Cas12a mediated plant genome editing. *Plant Biotechnol. J.* 17, 1431–1445. doi: 10.1111/pbi.13068

Wang, J., Tang, M., Chen, S., Zhen, X., Mo, H., Li, S., et al. (2017). Down-regulation of *BnDA1*, whose gene locus is associated with the seeds weight, improves the seeds weight and organ size in *Brassica napus*. *Plant Biotechnol. J.* 15, 1024–1033. doi: 10.1111/pbi.12696

Wang, M., Wang, Z., Mao, Y., Lu, Y., Yang, R., Tao, X., et al. (2019). Optimizing base editors for improved efficiency and expanded editing scope in rice. *Plant Biotechnol. J.* 17, 1697–1699. doi: 10.1111/pbi.13124

Woodfield, H. K., Sturtevant, D., Borisjuk, L., Munz, E., Guschina, I. A., Chapman, K., et al. (2017). Spatial and temporal mapping of key lipid species in *Brassica napus* seeds. *Plant Physiol.* 173, 1998–2009. doi: 10.1104/pp.16.01705

Wu, J., Chen, C., Xian, G., Liu, D., Lin, L., Yin, S., et al. (2020). Engineering herbicide-resistant oilseed rape by CRISPR/Cas9-mediated cytosine base-editing. *Plant Biotechnol. J.* doi: 10.1111/pbi.13368

Yang, Y., Zhu, K., Li, H., Han, S., Meng, Q., Khan, S. U., et al. (2018). Precise editing of *CLAVATA* genes in *Brassica napus,* L. regulates multilocular silique development. *Plant Biotechnol. J.* 16, 1322–1335. doi: 10.1111/pbi.12872

Zhai, Y., Yu, K., Cai, S., Hu, L., Amoo, O., Xu, L., et al. (2020). Targeted mutagenesis of *BnTT8* homologs controls yellow seed coat development for effective oil production in *Brassica napus L. Plant Biotechnol. J.* 18, 1153–1168. doi: 10.1111/pbi.13281

Zhou, Y., and Fowke, L. C. (2002). Control of petal and pollen development by the plant cyclin-dependent kinase inhibitor *ICK1* in transgenic Brassica plants. *Planta* 215, 248–257. doi: 10.1007/s00425-002-0752-2

Zong, Y., Song, Q., Li, C., Jin, S., Zhang, D., Wang, Y., et al. (2018). Efficient C-to-T base editing in plants using a fusion of nCas9 and human APOBEC3A. *Nat. Biotechnol.* 36, 950–953. doi: 10.1038/nbt.4261

Zong, Y., Wang, Y., Li, C., Zhang, R., Chen, K., Ran, Y., et al. (2017). Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. *Nat. Biotechnol.* 35, 438–440. doi: 10.1038/nbt.3811

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Hu, Amoo, Liu, Cai, Zhu, Shen, Yu, Zhai, Yang, Xu, Fan and Zhou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genome Editing

ORIGINAL RESEARCH
published: 25 November 2020
doi: 10.3389/fgeed.2020.604289

A Universal System of CRISPR/Cas9-Mediated Gene Targeting Using All-in-One Vector in Plants

Ayako Nishizawa-Yokoi[1,2], *Masafumi Mikami*[3] *and Seiichi Toki*[1,3,4]*

[1] Plant Genome Engineering Research Unit, Institute of Agrobiological Sciences, National Agriculture and Food Research Organization (NARO), Tsukuba, Japan, [2] Precursory Research for Embryonic Science and Technology (PRESTO), Japan Science and Technology Agency (JST), Kawaguchi, Japan, [3] Graduate School of Nanobioscience, Yokohama City University, Yokohama, Japan, [4] Kihara Institute for Biological Research, Yokohama City University, Yokohama, Japan

Homologous recombination-mediated genome editing, also called gene targeting (GT), is an essential technique that allows precise modification of a target sequence, including introduction of point mutations, knock-in of a reporter gene, and/or swapping of a functional domain. However, due to its low frequency, it has been difficult to establish GT approaches that can be applied widely to a large number of plant species. We have developed a simple and universal clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9)-mediated DNA double-strand break (DSB)-induced GT system using an all-in-one vector comprising a CRISPR/Cas9 expression construct, selectable marker, and GT donor template. This system enabled introduction of targeted point mutations with non-selectable traits into several target genes in both rice and tobacco. Since it was possible to evaluate the GT frequency on endogenous target genes precisely using this system, we investigated the effect of treatment with Rad51-stimulatory compound 1 (RS-1) on the frequency of DSB-induced GT. GT frequency was slightly, but consistently, improved by RS-1 treatment in both target plants.

Keywords: gene targeting, homologous recombination, CRISPR/Cas9, genome editing, rice

OPEN ACCESS

Edited by:
Jimmy R. Botella, The University of Queensland, Australia

Reviewed by:
Yanfei Mao, Shanghai Institutes for Biological Sciences (CAS), China
Zhaohui Zhong, University of Electronic Science and Technology of China, China

***Correspondence:**
Seiichi Toki
stoki@affrc.go.jp

Specialty section:
This article was submitted to Genome Editing in Plants, a section of the journal Frontiers in Genome Editing

Received: *09 September 2020*
Accepted: *19 October 2020*
Published: *25 November 2020*

Citation:
Nishizawa-Yokoi A, Mikami M and Toki S (2020) A Universal System of CRISPR/Cas9-Mediated Gene Targeting Using All-in-One Vector in Plants. Front. Genome Ed. 2:604289. doi: 10.3389/fgeed.2020.604289

INTRODUCTION

Genome editing techniques have come to be required for both the development of basic research and plant molecular breeding in recent years. Targeted mutagenesis using sequence-specific nucleases (SSNs) such as transcription activator-like effector nucleases (TALENs) and the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) is one of the genome editing techniques that has become available for several plant species and crops (Voytas, 2013; Puchta, 2017; Vats et al., 2019). However, it is difficult to deliberately introduce a desired mutation into target locus using this method because the mutations occur randomly in the process of repair *via* the non-homologous end joining (NHEJ) pathway of DNA double-strand breaks (DSBs) induced by SSNs. On the other hand, homologous recombination (HR)-mediated gene targeting (GT) allows precise genome engineering (the introduction of nucleotide substitutions, swapping of functional domains, and knock-in of reporter genes, etc.) of endogenous target genes *via* "copy and paste" of sequences from a repair template. This approach is essential

to create gain-of-function-type mutants. However, establishing a universal GT method that can be applied to a large number of plant species remains a challenge.

It is well-known that the frequency of GT in higher plants is very low (Hrouda and Paszkowski, 1994; Gallego et al., 1999; Puchta, 2002). The positive–negative selection method has been used widely to enrich rare GT cells at least in rice (Terada et al., 2002; Yamauchi et al., 2009; Nishizawa-Yokoi et al., 2015). Using this method, a positive selection marker expression cassette is introduced into the target locus together with the desired mutations. Accordingly, we have used the *piggyBac* transposon, which transposes without leaving a footprint at the excised site, to allow precise marker elimination from a GT locus in rice (Nishizawa-Yokoi et al., 2015). We succeeded in introducing targeted point mutations or reporter genes into multiple target genes *via* GT with positive–negative selection and subsequent excision of the positive selection marker from the target locus using *piggyBac* transposition. However, such positive–negative selection has not been applied successfully to the selection of GT cells in higher plants other than rice.

The induction of SSN-mediated DSBs at specific target loci has been explored as a major strategy for the improvement of GT frequency in several plant species (Fauser et al., 2012, 2014; Cermák et al., 2015; Gil-Humanes et al., 2017; Miki et al., 2018). In almost all of these studies, GT cells have been enriched from a large number of non-GT cells using antibiotic, herbicide, and/or visible selection, resulting from the introduction of targeted point mutations conferring resistance to herbicide, knock-in of antibiotic resistance, or the introduction of a reporter gene into a target gene *via* GT. Thus, GT frequency is difficult to evaluate accurately using this system because the total number of cells transfected with the GT template is unclear.

On the other hand, various kinds of small-molecule compounds that have been developed to modulate DSB repair pathways (inhibit NHEJ or activate HR) have been utilized to improve the frequency of CRISPR/Cas9-mediated GT in mammalian cells (Danner et al., 2017; Hengel et al., 2017). It has been reported that Rad51-stimulatory compound 1 (RS-1) led to an increase in the CRISPR/Cas9- and TALEN-mediated knock-in frequency in rabbit, whereas SCR7 (a DNA ligase 4 inhibitor) had minimal effect (Jayathilaka et al., 2008; Pinder et al., 2015; Song et al., 2016).

We reveal here that a simple and easy-to-use CRISPR/Cas9-mediated GT method can introduce point mutations with a non-selectable trait into several target genes in rice and tobacco using an all-in-one GT vector carrying CRISPR/Cas9, a selectable marker gene, and a GT template. Furthermore, we evaluated the GT frequency on several endogenous target genes accurately using this method and investigated the effect of RS-1 on the frequency of GT with an all-in-one vector in rice and tobacco.

MATERIALS AND METHODS

Plant Materials

Rice (*Oryza sativa* L. cv. Nipponbare) calli were induced from mature seeds as described previously (Saika and Toki, 2010) and were used for *Agrobacterium*-mediated transformation. Tobacco (*Nicotiana tabacum* L. cv. Petit Havana, SR-1) plants were grown in soil in a greenhouse (16 h light/8 h dark) at 21°C, and mature leaves of 3- or 4-week-old plants were used for *Agrobacterium*-mediated transformation.

Vector Construction

All-in-one GT vectors are based on the CRISPR/Cas9 expression vectors described in a previous study as pZH_OsU6sgRNA_SpCas9-wt (Endo et al., 2019) or pZH_OsU6gRNA_MMCas9 (Mikami et al., 2015) for rice and pDe_Cas9_KAN (Fauser et al., 2014) for tobacco. The 20-nt annealed oligonucleotide pairs for the target sequences shown in **Supplementary Table 1** were cloned into the *Bbs*I site of the single guide RNA (sgRNA) expression vector described as pU6gRNA-oligo (Mikami et al., 2015) for rice and pEn-Chimera (Fauser et al., 2014), pMR217, and pMR218 (Ritter et al., 2017) for tobacco, respectively.

An all-in-one GT vector for the introduction of point mutations into the *OsALS* gene was constructed as follows. The GT donor template (683 bp) containing a partial *OsALS* coding sequence with W548L/S627I mutations was amplified by PCR with the vector used in our previous study (Endo et al., 2007) using the primers shown in **Supplementary Table 1**. Two sgRNA expression cassettes (OsU6::sgOsALS_W548 and OsU6::sgOsALS_S627) were amplified by PCR using the primers shown in **Supplementary Table 1**. Three DNA fragments of the GT donor, OsU6::sgOsALS_W548, and OsU6::sgOsALS_S627 were introduced simultaneously into the *Asc*I site of pZH_OsU6sgRNA_SpCas9-wt by an in-fusion reaction (Takara). An all-in-one GT vector for the modification of *OsCly1* gene was constructed as follows. The sgRNA expression cassette (OsU6::sgCly1) was digested and ligated into the *Asc*I/*Pac*I site in pZH_OsU6gRNA_MMCas9. To insert the GT donor template into the vector, an artificially synthesized DNA fragment (1,001 bp) containing a partial *OsCly1* coding sequence with mutations in the miR172 targeting site was cloned into the *Asc*I site in pZH_OsU6sgOsCly1_MMCas9. For construction of an all-in-one GT vector in tobacco, the sgRNA expression cassettes (AtU6::sgNtALS in pEn-Chimera, AtU6::sgNtEPSPS_T176 in pMR217, and AtU6::sgNtEPSPS_P180 in pMR218) were generated *via* a cut-ligation reaction with *Bbs*I as described above. Single (for *NtALS*) or double sgRNA modules (for *NtEPSPS*) were combined with pDe_Cas9_KAN using a Gateway LR reaction (Thermo Fisher Scientific). GT donor templates for the modification of *NtALS* (1,040 bp DNA fragments containing a partial *NtALS-B* coding sequence with W568L/S647I mutations) or *NtEPSPS* (1,018 bp DNA fragments carrying a partial *NtEPSPS-B* with T176I/P180S mutations) were synthesized artificially and introduced into the *Pac*I site of pDe_Cas9_AtU6sgNtALS or pDe_Cas9_AtU6sgNtEPSPSx2, respectively.

Agrobacterium-Mediated Transformation

All-in-one GT vectors were transformed into rice and tobacco by *Agrobacterium*-mediated transformation as described previously (Horsch et al., 1985; Toki et al., 2006). Rice calli were immersed and shaken gently in AAM medium with 25 μM RS-1 or 0.05%

dimethyl sulfoxide (DMSO) as a control for 30 min before *Agrobacterium* infection. Transgenic rice calli were selected on medium N6D containing 50 mg/L hygromycin, 25 mg/L meropenem (Wako Pure Chemical Industries), and 25 μM RS-1 or 0.05% DMSO for 2 weeks at 33°C. Hygromycin-resistant calli were transferred to N6D containing 50 mg/L hygromycin and 25 mg/L meropenem without RS-1 or DMSO and cultured for 2 more weeks. To evaluate GT frequency targeting the *OsALS* locus without CRISPR/Cas9-mediated DSB induction, a vector harboring an 8-kb fragment of genomic DNA encoding the 5′ truncated *OsALS* gene carrying W548L and S627I mutations was transformed into rice calli by *Agrobacterium*-mediated transformation (Horsch et al., 1985; Toki et al., 2006). Transgenic rice calli were cultured on N6D medium containing 25 mg/L meropenem, 0.75 μM herbicide bispyribac (BS), and 25 μM RS-1 or 0.05% DMSO for 3 weeks. BS-tolerant calli were transferred to N6D medium containing 25 mg/L meropenem and 0.75 μM BS without RS-1 or DMSO and cultured for a further 2 weeks.

Tobacco leaf discs were punched from surface-sterilized leaves with a cork borer (6 mm in diameter) and treated with 25 μM RS-1 or 0.05% DMSO as a control for 30 min in Murashige and Skoog (MS) medium (Murashige and Skoog, 1962). Leaf discs were then inoculated with *Agrobacterium* strain EHA105 (Hood et al., 1993) harboring an all-in-one GT vector. After 3 days of co-cultivation with *Agrobacterium*, leaf discs were transferred to regeneration medium with 100 mg/L kanamycin, 25 mg/L meropenem, and 25 μM RS-1 or 0.05% DMSO for 2 weeks and cultured under a 16-h light/8-h dark cycle at 27°C. Kanamycin-resistant calli were transferred to fresh medium with 100 mg/L kanamycin and 25 mg/L meropenem without RS-1 or 0.05% DMSO and cultured for 2 more weeks.

Chemicals

RS-1 and the Lig4 inhibitor, SCR7, were purchased commercially (Cosmo Bio). A portion of these compounds as used in this study was synthesized as described previously (Jayathilaka et al., 2008; Greco et al., 2016).

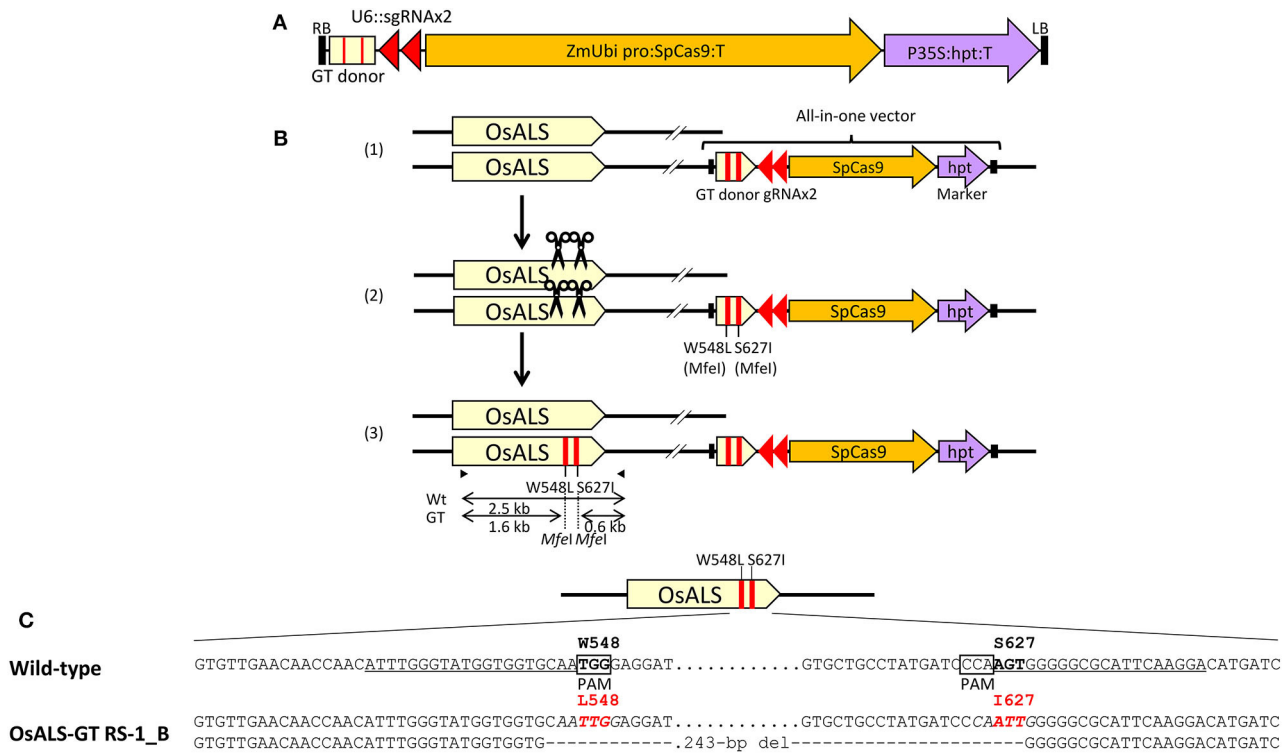


FIGURE 1 | Introduction of point mutations into the rice *ALS* locus *via* clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9)-mediated gene targeting (GT) with an all-in-one vector. **(A)** Schematic diagram of the all-in-one GT vector for modification of the rice acetolactate synthase (*OsALS*) locus. The vector carries two sgRNA expression units (OsU6::sgOsALS_W548 and OsU6::sgOsALS_S627), SpCas9 expression cassette (maize ubiquitin promoter::OsADH 5′UTR-SpCas9-SV40 NLS::terminator), selection marker (CaMV35S promoter::hpt::terminator), and GT donor template (ca. 700 bp) containing a partial *OsALS* coding sequence with W548L and S627I mutations. **(B)** CRISPR/Cas9-mediated GT strategy in rice. (1) The all-in-one GT vector was integrated into the host genome *via Agrobacterium*-mediated transformation. (2) Following the expression of CRISPR/Cas9 from the GT vector on the genome, a double-strand break (DSB) was induced at the target site during the selection of transgenic cells. (3) Repair of DSB *via* homologous recombination (HR) between the target gene and GT donor then introduces the desired mutations into the target gene. The primer sets used for cleaved amplified polymorphic sequence (CAPS) analysis that identify transgenic calli in which a GT event occurred at the *OsALS* locus are shown as black arrows. **(C)** Nucleotide sequences of the *OsALS* gene in wild-type (top) and regenerated plants from OsALS-GT RS-1_B calli (bottom). Target sites of sgRNA are underlined, W548L and S627I mutations *via* GT are in red, and the recognition sites of *Mfe*I (CAATTG) used for CAPS analysis are in italics.

Screening of Gene Targeting Candidate Calli by Cleaved Amplified Polymorphic Sequence Analysis

Genomic DNA extracted from small pieces of clonally propagated hygromycin-resistant rice calli or kanamycin-resistant tobacco calli using Agencourt chloropure (Beckman Coulter) according to the manufacturer's protocol was subjected to cleaved amplified polymorphic sequence (CAPS) analysis. PCR amplifications were performed with KOD FX neo or KOD ONE (TOYOBO) using the primer sets shown in **Supplementary Table 1**. PCR products were digested with restriction enzyme *Mfe*I for *OsALS* and *NtALS-B*, *Xba*I for *OsCly1*, and *Hin*dIII for *NtEPSPS-B* and analyzed with MultiNA microchip electrophoresis system (Shimadzu).

PCR fragments derived from CAPS-positive calli or plants were cloned into pCR-BluntII-TOPO (Invitrogen) and subjected to sequence analysis using an ABI3130 sequencer (Applied Biosystems).

Southern Blot Analysis

Genomic DNA was extracted from leaves of seedlings using the Nucleon Phytopure extraction kit (Cytiva) according to the manufacturer's protocol. Two micrograms of rice genomic DNA or 15 μg of tobacco genomic DNA was digested with the same restriction enzyme as that used in CAPS analysis and fractionated in a 0.8% agarose gel. Southern blot analysis was performed according to the digoxigenin (DIG) Application Manual (Sigma-Aldrich). Specific DNA probes for the *hpt*, *nptII*, and *OsCly1* locus were synthesized with a PCR DIG probe synthesis kit (Sigma-Aldrich) according to the manufacturer's protocol using the primers shown in **Supplementary Table 1**.

Herbicide-Susceptibility Testing in Rice and Tobacco

Seeds of wild-type rice (Nipponbare) and T_1 progeny of OsALS-GT RS-1_B were sown on 1/2 MS medium with or without 1.5 μM BS and were grown in a growth chamber at 27°C under a 16 h photoperiod. Seeds of wild-type tobacco (SR-1) and T_1 progeny of NtALS_GT RS-1_B#1 were sown on 1/2 MS medium with or without 500 nM chlorsulfuron (CS) and grown in a growth chamber (16 h light/8 h dark) at 27°C.

RESULTS

Establishing a CRISPR/Cas9-Mediated Gene Targeting Method With an All-in-One Vector in Rice

Previously, we established a system to select infrequent GT cells by introducing two amino acid substitutions (W548L and S627I) conferring tolerance to the herbicide bispyribac sodium (BS) into the rice *acetolactate synthase* (*OsALS*) gene (Endo et al., 2007). The *ALS* gene has been used as a model target for development of a GT system in various plant species (Zhang et al., 2013; Nishizawa-Yokoi et al., 2015; Wolter et al., 2018). We also succeeded in introducing two amino acid substitutions (W548L/S627I) into the *OsALS* gene by GT with positive–negative selection (Nishizawa-Yokoi et al., 2015) rather than using herbicide selection. Thus, an all-in-one GT vector harboring two sgRNAs targeting the *OsALS* gene, Cas9, *hygromycin phosphotransferase* (*hpt*) gene expression cassettes, and a GT donor template (ca. 700 bp) carrying W548L and S627I, in which these point mutations generate recognition sites for the restriction enzyme *Mfe*I, CAATTG, was constructed and transformed into rice calli (**Figures 1A,B**). The desired point mutations were located on the protospacer adjacent motif (PAM) sequences following the DNA region targeted by SpCas9, indicating that the GT donor template on the vectors was not cleaved by the CRISPR/Cas9 system. To examine the cytotoxicity of RS-1 in rice calli, wild-type calli were cultured on medium containing 25 or 50 μM RS-1 for 2 weeks. No significant differences in the growth of calli were observed between the treatment groups with DMSO as a control, 25 or 50 μM RS-1 (**Supplementary Figure 1**). Therefore, 25 μM RS-1 was used for subsequent experiments. Rice calli transformed with an all-in-one GT vector were selected on medium containing hygromycin B, not herbicide BS, with or without RS-1, to investigate the ratio of GT cells to *Agrobacterium*-mediated transformed cells in rice. After a 2-week selection period, transgenic callus lines were cloned and further propagated for another 2 weeks. Genomic DNA was extracted from 960 callus lines each of the control (DMSO treatment) and 25 μM RS-1 treatment groups. CAPS analysis, i.e., PCR analysis coupled with *Mfe*I digestion, revealed that there were no GT positive callus lines from among 960 independent transgenic calli harboring all-in-one GT vector treated with DMSO (**Table 1**), while *Mfe*I-digested PCR fragments were detected in two independent callus lines, designated as OsALS-GT RS-1_A and B, among the 960 all-in-one GT vector transgenic calli treated with 25 μM RS-1, suggesting that, in these callus lines, W548L/S627I mutations were introduced into the *OsALS* locus by HR between the GT vector that could be integrated into the genome and the target locus. Thus, the frequency of GT calli to transgenic calli analyzed by CAPS was 0.21% (2/960) under treatment with RS-1 (**Table 1**). Sequencing analysis of PCR fragments derived from the *OsALS* locus in OsALS-GT RS-1_A and B revealed that the proportion

TABLE 1 | Summary of gene targeting (GT) experiments targeting OsALS locus using an all-in-one GT vector.

Experiments	Treatment	No. of hygromycin-resistant calli analyzed	No. of calli with W548L/S627I mutations on OsALS	GT frequency (%)
A	DMSO	576	0	0
	25 μM RS-1	576	1*	0.17
B	DMSO	384	0	0
	25 μM RS-1	384	1**	0.26
Total	DMSO	960	0	0
	RS-1	960	2	0.21

*GT positive callus lines were designated as follows: *, OsALS-GT RS-1_A; **, OsALS-GT RS-1_B.*

TABLE 2 | Inheritance of target mutations and segregation of an all-in-one gene targeting (GT) vector in T_1 plants.

Line no.	No. of T_1 plants analyzed	No. of progeny plants carrying W548L/S627I mutations via GT in *OsALS* locus						The existence of all-in-one vector	
		wild-type	%	Hetero	%	Homo	%	+	–
OsALS-GT RS-1_B#1	36	0	0	0	0	36	100	34	2
OsALS-GT RS-1_B#2	31	0	0	1	3.2	30	96.8	30	1

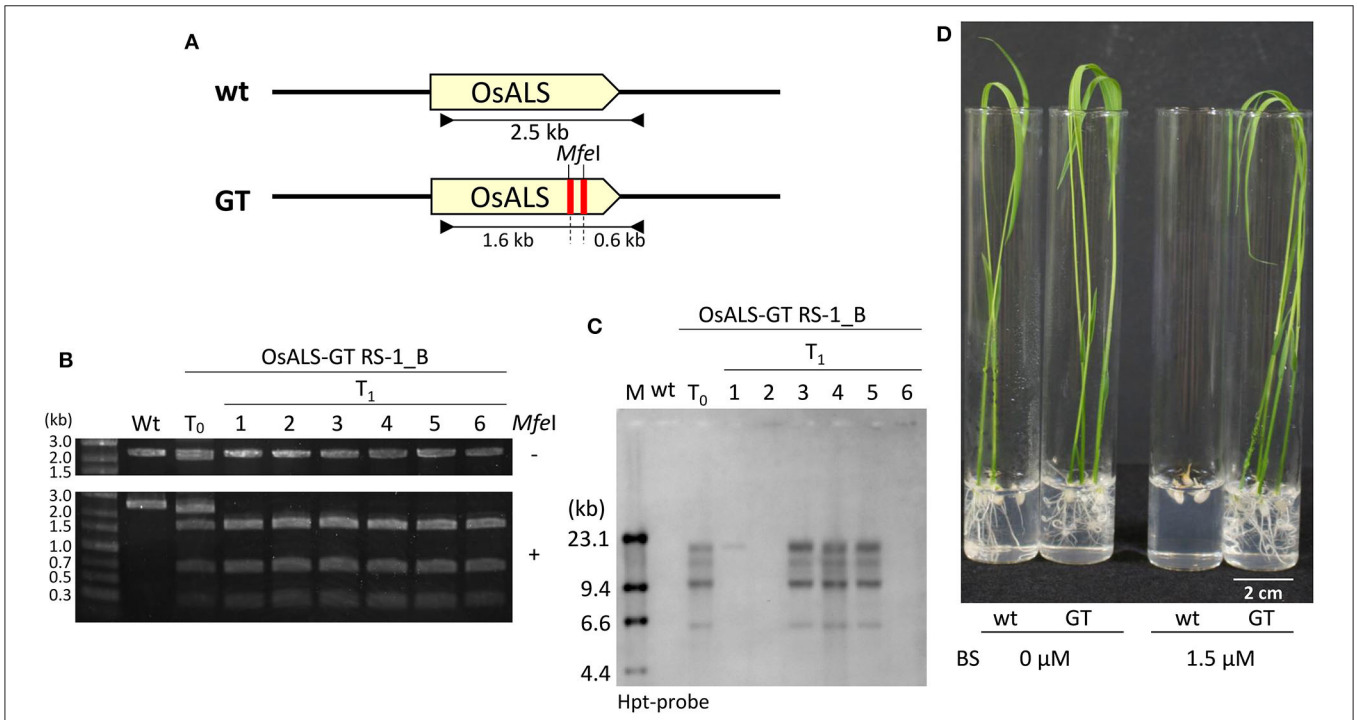


FIGURE 2 | Functional analysis of modified *OsALS* gene in T_1 progeny. **(A)** Genomic structure of wild-type *OsALS* locus (top) and the modified *OsALS* locus in T_1 plants from OsALS-GT RS-1_B calli (bottom). The primer sets used for cleaved amplified polymorphic sequence (CAPS) analysis are shown as black arrows. **(B)** CAPS analysis with genomic DNA of wild type, T_0 regenerated plants, and T_1 progeny of OsALS-GT RS-1_B using *OsALS*-specific primers shown in **(A)**. PCR fragments were digested with (lower) or without (upper) *Mfe*I (*Mfe*I + and –, respectively). **(C)** Southern blot analysis with the hpt-specific probe using *Mfe*I-digested genomic DNA of wild type, T_0 regenerated plants, and T_1 progeny of OsALS-GT RS-1_B. **(D)** Herbicide bispyribac (BS)-tolerant phenotype of T_1 plants of OsALS-GT RS-1_B.

of GT cells carrying a GT-modified *OsALS* locus within the transgenic calli analyzed was 8.3% (2/24) and 37.5% (9/24), respectively (**Supplementary Table 2**). These results suggest that transgenic cells with W548L/S627I mutations *via* GT and without these mutations at the *OsALS* locus existed in a chimeric state in a single callus clone. OsALS-GT RS-1_B calli were transferred to regeneration medium to obtain regenerated plants, and the *OsALS* locus of T_0 regenerated plants was genotyped using CAPS and sequencing analysis. All regenerated plants (OsALS-GT RS-1_B#1-#24) from OsALS-GT RS-1_B analyzed (24/24, 100%) had an identical mutation pattern, which was the desired point mutation introduced *via* GT and the long deletion (243 bp) between the two target sequences of CRISPR/Cas9 in each *OsALS* allele, respectively (**Figure 1C**). Furthermore, we obtained T_1 plants from self-pollinating regenerated plants (OsALS-GT RS-1_B#1 and #2) of the OsALS-GT RS-1_B line and conducted genotyping of the *OsALS* gene and the BS-sensitivity test. We found that all progeny plants from OsALS-GT RS-1_B#1 T_0 plants carried homozygous point mutations *via* GT in the *OsALS* gene (**Table 2**; **Figures 2A,B**). In another line, OsALS-GT RS-1_B#2, 96.8% (30/31) of progeny plants were homozygous mutants harboring the desired point mutations *via* GT in the *OsALS* gene, while only one plant (3.2%) was a heterozygous mutant carrying GT-mediated point mutations and the CRISPR/Cas9-mediated deletion in each allele of *OsALS* (**Table 2**). These results indicate that defects in the *OsALS* gene might cause a sterility phenotype in rice even when its mutations are introduced into one allele; in other words, it was likely that the loss of the *OsALS* gene induced the defect of gametogenesis. What is more important is that progeny plants carrying W548L/S627I mutations in the *OsALS* gene and lacking all-in-one vectors could be obtained by genetic segregation *via* self-crossing of regenerated plants (**Figures 2B,C**). These T_1 plants showed a herbicide BS-tolerant phenotype compared with

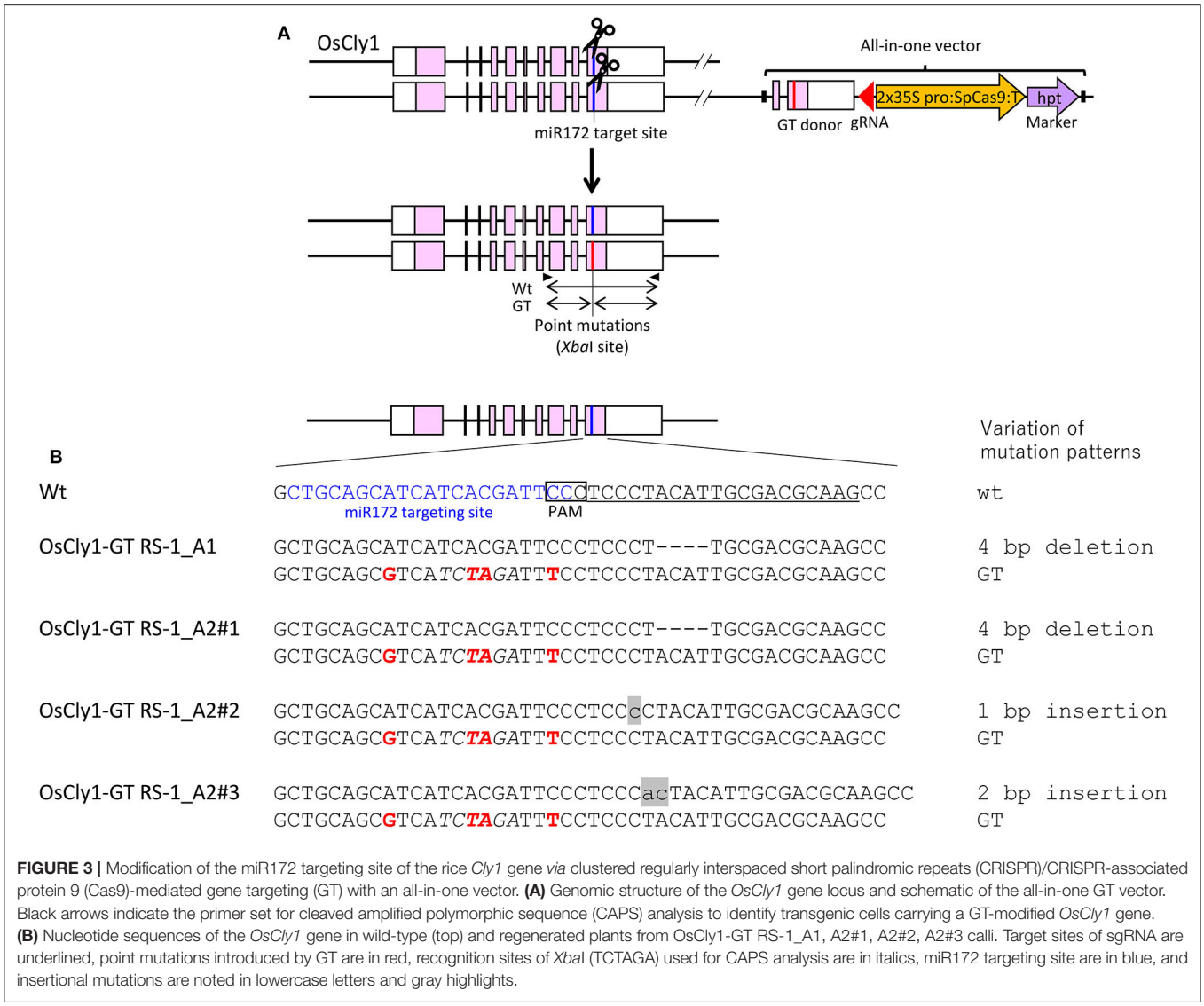


FIGURE 3 | Modification of the miR172 targeting site of the rice *Cly1* gene *via* clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9)-mediated gene targeting (GT) with an all-in-one vector. **(A)** Genomic structure of the *OsCly1* gene locus and schematic of the all-in-one GT vector. Black arrows indicate the primer set for cleaved amplified polymorphic sequence (CAPS) analysis to identify transgenic cells carrying a GT-modified *OsCly1* gene. **(B)** Nucleotide sequences of the *OsCly1* gene in wild-type (top) and regenerated plants from OsCly1-GT RS-1_A1, A2#1, A2#2, A2#3 calli. Target sites of sgRNA are underlined, point mutations introduced by GT are in red, recognition sites of *Xba*I (TCTAGA) used for CAPS analysis are in italics, miR172 targeting site are in blue, and insertional mutations are noted in lowercase letters and gray highlights.

wild-type plants (**Figure 2D**). To prove the effect of treatment with RS-1 on GT efficiency in rice, we evaluated GT frequency by introducing W548L/S627I substitutions into the *OsALS* gene without CRISPR/Cas9-mediated DSB induction. Rice calli were transformed with the vector harboring an 8-kb fragment of genomic DNA encoding the 5′ truncated *OsALS* gene carrying W548L/S627I mutations (Endo et al., 2007) and cultured on medium containing BS with or without RS-1 for 3 weeks. We obtained 38 and 71 callus lines of 6,561 and 6,289 transgenic calli treated with DMSO and RS-1, respectively (**Table 3**). CAPS analysis revealed that 35 and 64 callus lines treated with DMSO and RS-1, respectively, were true GT lines carrying two point mutations in the *OsALS* gene *via* GT. The proportion of true GT callus lines to transgenic calli was 0.5 and 1.0% under treatment with DMSO and RS-1, respectively (**Table 3**). These data provide support for the positive effects of RS-1 on GT.

To confirm that CRISPR/Cas9-mediated GT with an all-in-one vector is a universal approach in rice, we applied this approach to modify the microRNA miR172 targeting site of the rice *cleistogamy 1* (*OsCly1*) gene (Chen, 2004; Nair et al., 2010)—a model gene for GT experiments in rice in our hands (Nishizawa-Yokoi et al., 2015). The miR172 targeting site of the *OsCly1* gene was edited by GT, producing a recognition site for the restriction enzyme *Xba*I, allowing isolation of GT cells by CAPS analysis (wild-type: CTGCAGCATCATCACGATTCC, GT: CTGCAGCgTCA*TCtaGA*TTtC, *Xba*I site in italics), with no change to amino acid sequences (**Figure 3A**). CAPS analysis revealed that each of two independent lines of 768 hygromycin-resistant calli were GT lines, with treatment with DMSO (OsCly1-GT DSMO_B1 and B2) as a control or RS-1 (OsCly1-GT RS-1_A1 and A2), respectively, indicating that the GT frequency was 0.26% in each experimental group (**Table 4**). Regenerated plants were obtained from these GT lines, and the genotype of the *OsCly1* gene was analyzed by CAPS and sequencing analysis. Out of 24 regenerated plants from OsCly1-GT RS-1_A1 and A2 callus lines, nine (40%) and 12 (50%), respectively, were true

TABLE 3 | Summary of gene targeting (GT) experiments targeting the *OsALS* locus without clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9)-mediated double-strand break (DSB) induction.

Experiments	Treatment	No. of *Agrobacterium* infected calli	No. of BS-resistant calli	True GT calli	
				No. of calli	(%)
A	DMSO	843	2	2	0.24
	25 μM RS-1	688	10	9	1.31
B	DMSO	1,209	2	2	0.17
	25 μM RS-1	1,352	11	9	0.67
C	DMSO	1,966	11	11	0.56
	25 μM RS-1	1,846	23	22	1.19
D	DMSO	2,543	23	20	0.79
	25 μM RS-1	2,403	27	24	1.00
Total	DMSO	6,561	38	35	0.53
	25 μM RS-1	6,289	71	64	1.02

TABLE 4 | Summary of gene targeting (GT) experiments targeting miR172 site of *OsCly1* gene using an all-in-one GT vector.

Experiments	Treatment	No. of hygromycin-resistant calli analyzed	No. of calli with mutations on the microRNA targeting site of *OsCly1* gene	GT frequency (%)
A	DMSO	384	0	0
	25 μM RS-1	384	2*	0.52
B	DMSO	384	2**	0.52
	25 μM RS-1	384	0	0
Total	DMSO	768	2	0.26
	25 μM RS-1	768	2	0.26

*GT positive callus lines were designated as follows: *, OsCly1-GT RS-1_A1 and A2; **, OsCly1-GT DMSO_B1 and B2.*

GT plants containing the desired GT-mediated mutations and CRISPR/Cas9-mediated indels in each allele of the *OsCly1* gene, while all regenerated plants from OsCly1-GT DSMO_B1 and B2 callus lines lacked GT-mediated point mutations in the *OsCly1* locus (0/24) (**Supplementary Table 3**). All GT plants from the OsCly1-GT RS-1_A1 line (9/9) contained an identical mutation pattern (mutations *via* GT/4 bp deletion) in the *OsCly1* gene (**Figure 3B**), while different mutation patterns were observed between individual GT plants from the OsCly1-GT RS-1_A2 line. These different mutation patterns were classified into three groups (mutations *via* GT/4 bp deletion, mutations *via* GT/1 bp insertion, and mutations *via* GT/2 bp insertion) (**Figure 3B**). Furthermore, we verified that progeny plants carrying biallelic GT-mediated point mutations in the *OsCly1* gene and lacking the all-in-one vector could be obtained *via* genetic segregation (**Supplementary Figure 2**).

Precise Gene Editing *via* Gene Targeting With All-in-One Vector in Tobacco

To demonstrate that this CRISPR/Cas9-mediated GT approach with an all-in-one vector can be applied to other plant species, we designed an experiment aimed at substituting two amino acids (W568L/S647I) conferring tolerance to the herbicide CS in tobacco ALS-B (*NtALS-B*) using an all-in-one vector. A vector carrying an sgRNA targeting region proximal to S647 in the *NtALS-B* gene, Cas9, nptII expression cassettes, and GT donor template 1.0 kb *NtALS-B* coding region with mutations (W568L/*Mfe*I site and S647I) was introduced into tobacco leaf discs with or without RS-1 by *Agrobacterium*-mediated transformation (**Figure 4A**). After coculture with *Agrobacterium*, tobacco leaf discs were transferred to the selection medium containing kanamycin and DMSO or RS-1 and were cultured for 2 weeks. Kanamycin-resistant calli were propagated for another 2 weeks on selection medium without DMSO or RS-1. CAPS analysis revealed that 0.07% (1/1,357; Line no. NtALS_GT DMSO_C) and 0.22% (3/1,382; Line no. NtALS_GT RS-1_A1, A2, and B) of transgenic calli treated with DMSO and RS-1, respectively, harbored a GT-modified *NtALS-B* gene (**Table 5**). Regenerated plants obtained from CAPS-positive calli were subjected to CAPS and sequencing analysis to confirm the GT-mediated introduction of W568L and S647I mutations in the *NtALS-B* gene. In CAPS analysis, *Mfe*I-digested fragments were detected in 6.9% (4/58, NtALS_GT RS-1_B#1 to B#4) and 7.8% (4/51, NtALS_GT_DMSO_C#1 to C#4) of T0 plants regenerated from NtALS_GT RS-1_B and DMSO_C callus lines, respectively, but not from NtALS_GT RS-1_A1 and A2. Sequencing analysis of PCR fragments from CAPS-positive plants (NtALS_GT RS-1_B#1-B#4 and NtALS_GT_DMSO_C#1-C#4) showed that W568L (*Mfe*I recognition site) and S647I mutations had been introduced into the targeted *NtALS-B* gene in the biallelic state in NtALS_GT_RS-1_B#1-B#4 or in the chimeric state in DMSO_C#1–C#4, respectively. However, in plants NtALS_GT_RS-1_B#2–B#4 and DMSO_C#1–C#4, unexpected mutations (insertion or deletion) also occurred at the CRISPR/Cas9 target region, suggesting that CRISPR/Cas9 recognized and digested again at the target sequence even after the introduction of S647I mutations *via* GT (**Supplementary Figure 3**), whereas only one plant (NtALS_GT RS-1_B#1) carried biallelic W568L/S647I mutations in the *NtALS-B* gene without CRISPR/Cas9-mediated dispensable mutations (**Figure 4B**, **Supplementary Figure 3**).

To confirm that W568L/S647I amino acid substitutions introduced *via* GT were stable and heritable, we obtained T_1 progenies from NtALS_GT RS-1_B#1 regenerated plants with homozygous desired mutations in the *NtALS-B* gene. Sequencing analysis revealed that all progeny plants harbored a GT-modified *NtALS-B* gene with homozygous W568L/S647I mutations (34/34); however, additional mutations were also generated by CRISPR/Cas9 in 37.5% of T_1 progenies (**Table 6**). We found diverse patterns (insertions or deletions) of additional mutation by CRISPR/Cas9 at the *NtALS-B* gene target site in siblings of NtALS_GT RS-1_B#1 T_1 plants (**Supplementary Figure 4**). Furthermore, the all-in-one vector segregated out in 44.1% of progeny plants with homozygous W568L/S647I mutations and without additional CRISPR/Cas9-mediated mutations

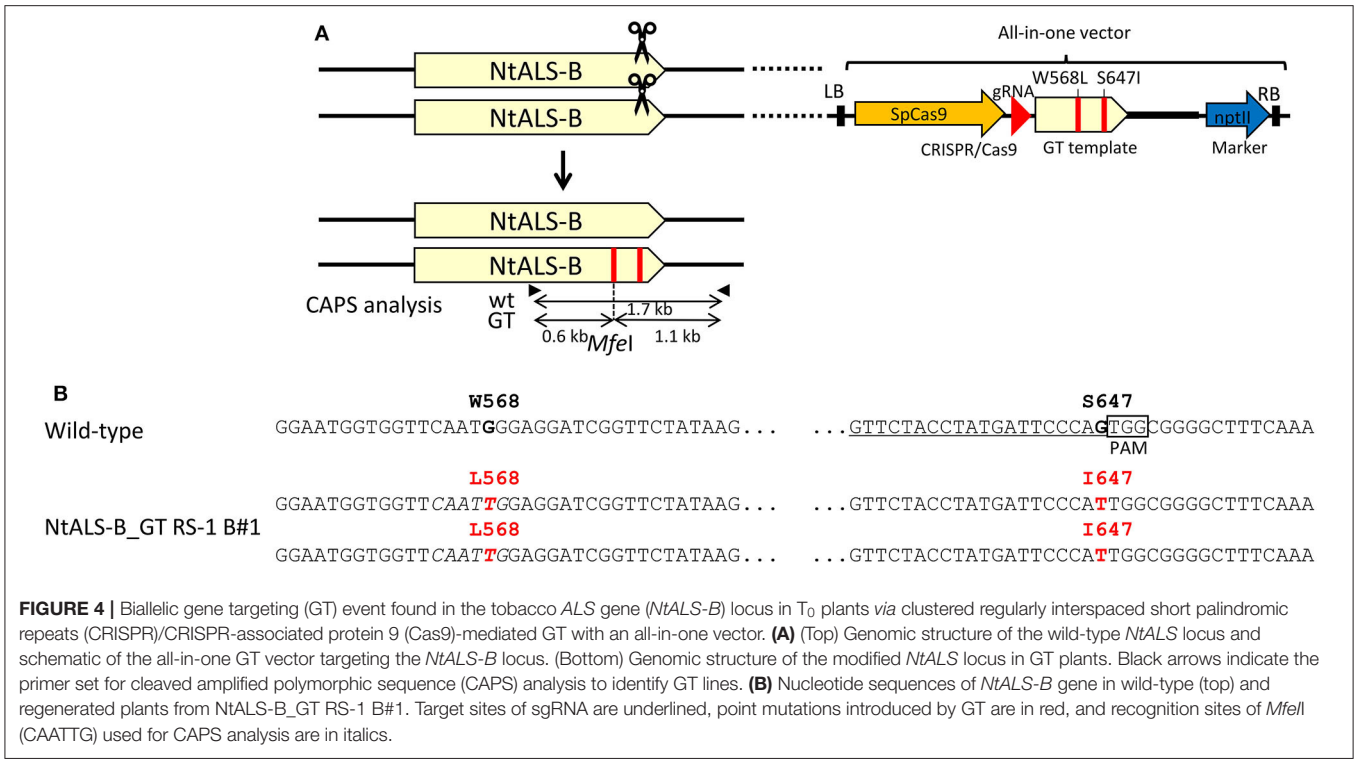


FIGURE 4 | Biallelic gene targeting (GT) event found in the tobacco *ALS* gene (*NtALS-B*) locus in T_0 plants *via* clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9)-mediated GT with an all-in-one vector. **(A)** (Top) Genomic structure of the wild-type *NtALS* locus and schematic of the all-in-one GT vector targeting the *NtALS-B* locus. (Bottom) Genomic structure of the modified *NtALS* locus in GT plants. Black arrows indicate the primer set for cleaved amplified polymorphic sequence (CAPS) analysis to identify GT lines. **(B)** Nucleotide sequences of *NtALS-B* gene in wild-type (top) and regenerated plants from NtALS-B_GT RS-1 B#1. Target sites of sgRNA are underlined, point mutations introduced by GT are in red, and recognition sites of *Mfe*I (CAATTG) used for CAPS analysis are in italics.

TABLE 5 | Summary of gene targeting (GT) experiments targeting *NtALS-B* gene using an all-in-one GT vector.

Experiments	Treatment	No. of Kanamycin-resistant calli analyzed	No. of CAPS-positive callus lines	%
A	DMSO	661	0	0
	25 μM RS-1	631	2*	0.32
B	DMSO	408	0	0
	25 μM RS-1	463	1**	0.22
C	DMSO	288	1***	0.35
	25 μM RS-1	288	0	0
Total	DMSO	1,357	1	0.07
	25 μM RS-1	1,382	3	0.22

*GT positive callus lines were designated as follows: *, NtALS_GT RS-1_A1 and A2; **, NtALS_GT RS-1_B; ***, NtALS_GT DMSO_C.*

(**Figures 5A,B**). We also confirmed that the introduction of W568L/S647I mutations in *NtALS-B via* GT conferred a CS-tolerant phenotype on tobacco plants (**Figure 5C**).

The next target for precise modification *via* GT in tobacco was the gene encoding endogenous *3-phosphoshikimate 1-carboxyvinyltransferase* (*EPSPS*), a well-known target of the herbicide glyphosate that catalyzes an essential step in the shikimate pathway common to aromatic amino acid biosynthesis (Steinrücken and Amrhein, 1980). Yu et al. (2015) showed that two amino acid substitutions [T102I and P106S (TIPS)] in a conserved region of the *EPSPS* gene led to high-level glyphosate resistance in goose grass (*Eleusine indica*). Thus, we attempted to introduce TIPS (T176I and P180S) amino acid substitutions into tobacco *EPSPS-B* (*NtEPSPS-B*) using a CRISPR/Cas9-mediated GT approach (**Supplementary Figures 5A,B**). Tobacco leaf discs were infected with *Agrobacterium* harboring an all-in-one vector carrying an sgRNA targeting the nearby S176 and P180 in the *NtEPSPS-B* gene, Cas9, nptII expression cassettes, and a GT donor template 1.0 kb *NtEPSPS-B* coding region with mutations (T176I and P180S) with or without RS-1 treatment. We isolated GT callus lines by CAPS analysis (P180S mutations carry a new *Hin*dIII restriction enzyme recognition site). CAPS-positive bands were detected in one (NtEPSPS_GT DMSO_1) and two (NtEPSPS_GT RS-1_1 and 2) independent lines from 384 transgenic calli treated with DMSO and RS-1, indicating that GT frequency was 0.26% and 0.52% in the treatment of DMSO and RS-1, respectively (**Supplementary Table 4**). PCR products amplified from CAPS-positive calli were cloned and sequenced to confirm the introduction of TIPS mutations in the *NtEPSPS-B* gene even at low efficiency (**Supplementary Table 5**). Unfortunately, mutant plants that carried the TIPS mutations in *NtEPSPS-B* gene *via* GT could not be regenerated from these CAPS-positive callus lines.

DISCUSSION

Positive–negative selection to enrich GT cells has been developed and applied successfully to the modification of several endogenous target genes in rice *via* GT (Terada et al., 2002;

TABLE 6 | Inheritance of target mutations *via* gene targeting (GT) into NtALS_GT RS-1_B#1 progenies.

No. of plants analyzed	No. of plants with W548L via GT in *NtALS-B*	%	No. of plants with mutations via CRISPR/Cas9 at S647 in *NtALS-B*	%	No. of null segregants of all-in-one vector	%
34	34	100	9	37.5	15	44.1

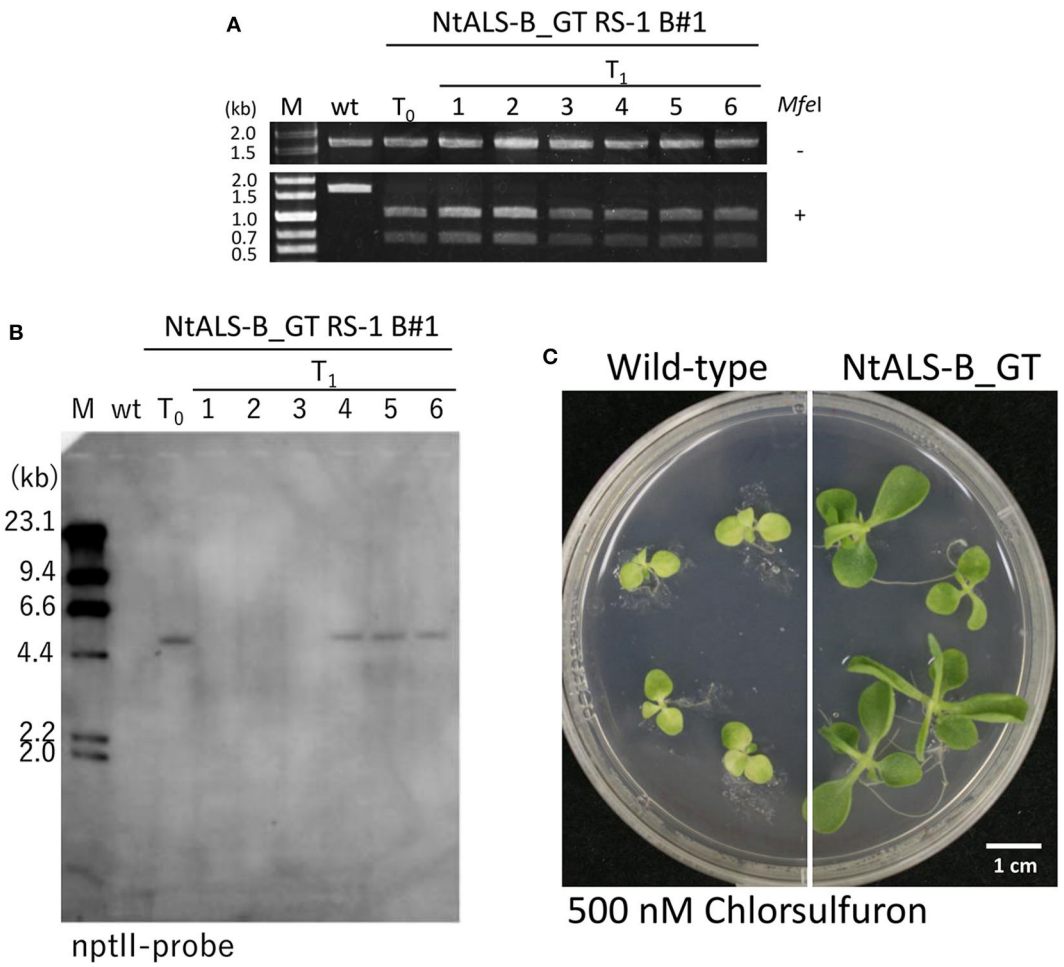


FIGURE 5 | Modified *NtALS-B* gene with W568L/S647I mutations confers a chlorsulfuron (CS)-tolerant phenotype in T_1 progeny of NtALS-B_GT RS-1 B#1. **(A)** Cleaved amplified polymorphic sequence (CAPS) analysis with genomic DNA of wild type, T_0 regenerated plants, and T_1 progeny of NtALS-B_GT RS-1_B#1 using *NtALS* specific primers shown in **Figure 4A**. PCR fragments were digested with (lower) or without (upper) *Mfe*I (*Mfe*I + and –, respectively). **(B)** Southern blot analysis with an nptII-specific probe using wild type, T_0, and T_1 plants of NtALS-B_GT RS-1 B#1. Line nos. 1–3 of T_1 plants were found to be transgene-null segregants. **(C)** Herbicide CS-tolerant phenotype of T_1 plants of NtALS-B_GT RS-1 B#1.

Yamauchi et al., 2009, 2014; Ono et al., 2012; Nishizawa-Yokoi et al., 2015). Using this method, because transgenic cells carrying T-DNA that had integrated randomly could be excluded by expression of the negative selection marker, no transgenes were left in the host genome other than at the target locus. However, to date, GT with positive–negative selection has been applied exclusively in rice. In addition, it is difficult to construct the large GT vector required, which has two negative selection marker cassettes at both ends of the vector and a positive selection marker cassette within a 6-kb DNA donor template, in which a long homology arm is needed for HR between the GT vector and target locus by use of spontaneous DSBs. In contrast, our novel all-in-one vector comprises a conventional CRISPR/Cas9 vector with an extra 1-kb DNA donor template; thus, unlike GT with positive–negative selection, long PCRs (>3–4 kb) are not needed to identify GT candidates *via* the CRISPR/Cas9-mediated GT method. The CRISPR/Cas9-mediated GT method developed in the present

study provides a very simple and user-friendly approach. In addition, we attempted to establish an efficient CRISPR/Cas9-mediated GT system using a combination of an all-in-one vector and treatment with the small-molecule RS-1, which is known to enhance HR in mammalian cells. Evaluation of the ratio of GT callus to transgenic callus carrying the randomly integrated all-in-one vector revealed that the GT frequency under treatment with RS-1 was slightly, but consistently, higher than that of controls in both rice and tobacco. Although the impact of treatment with RS-1 on GT efficiency was limited even in rice and tobacco, our findings suggest that RS-1 has the potential to improve the frequency of CRISPR/Cas9-mediated GT in plants. Whereas, we also tested the effect of SCR7—a DNA ligase 4 inhibitor—on GT efficiency in rice, there were no significant differences with and without SCR7 treatment (data not shown). Structural optimization of RS-1 to plant Rad51, or structural modification of RS-1 to enhance delivery efficiency into plant cells, is expected to lead the development of more effective approaches to establish a universal GT system for various types of crops.

In tobacco, we found that CRISPR/Cas9-mediated additional mutations were also generated at the target site in T_1 progenies with the all-in-one vector, but not in siblings in which the all-in-one vector was segregated out (**Supplementary Figure 4**). These results suggest that Cas9-mediated DNA cleavage occurred in tobacco somatic cells even after the introduction of a single point mutation *via* GT within the seed region of the target sequence. To completely abolish induction of these *de novo* mutations in T_1 tobacco plants, it is essential that multiple mismatches, or a single mismatch, are introduced into the sgRNA target site or PAM sequence *via* GT, at least in dicot plants.

CRISPR/Cas9-mediated GT strategies have been applied for precise gene modification in plants by several research groups. Geminivirus-based replicons have been used for transient expression of SSNs and delivery of DNA donor template, resulting in successful knock-in of reporter or antibiotic-resistance genes, etc., into an endogenous target gene in tobacco (Baltes et al., 2014), tomato (Cermák et al., 2015), potato (Butler et al., 2016), cassava (Hummel et al., 2018), rice (Wang et al., 2017), and wheat (Gil-Humanes et al., 2017). Puchta and colleagues (Fauser et al., 2012; Schiml et al., 2014) developed an *in planta* GT strategy in which DSBs were induced by SSNs at both ends of a GT donor template on a GT vector integrated into the host genome and at an endogenous target gene, resulting in repair of the target gene *via* HR using a released linear GT donor from a GT vector in *Arabidopsis*. Using these strategies, although GT frequency was still low in higher plants, GT cells were enriched by counter selection derived from the introduction of targeted point mutations conferring resistance to herbicide, knock-in of antibiotic resistance or a reporter gene into the target gene *via* GT. Miki et al. (2018) revealed that not only in-frame reporter gene knock-in but also the introduction of amino acid substitutions, i.e., non-selectable traits, into the endogenous target genes was achieved by the expression of Cas9 under the egg cell-specific promoter in *Arabidopsis*. This tool provides a powerful, but thus far *Arabidopsis*-specific, approach.

Here, we have established a CRISPR/Cas9-mediated GT strategy using an all-in-one vector in combination with RS-1 treatment in rice and tobacco. Although removal of an all-in-one vector remains challenging in vegetative propagation of plants, we are hopeful that our approach will become widely applicable for precise genome modification in a variety of crops. Therefore, we are currently screening several chemicals with the aim of improving GT efficiency in plants.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

AN-Y designed the research and wrote the manuscript. AN-Y and MM conducted the experiments. ST commented on the research and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was also supported by JST, PRESTO (Grant No. JPMJPR16QA), Research Grant from NARO Gender Equality Program, Cross-Ministerial Strategic Innovation Promotion Program (SIP) Technologies for Smart Bio-industry and Agriculture (funding agency: Bio-oriented Technology Research Advancement Institution, NARO), and JSPS KAKENHI (the Grant-in-Aid for Research Activity Start-up, Grant No. JP19K21185).

ACKNOWLEDGMENTS

We thank Ms. K. Iida, Ms. A. Mori, Ms. K. Amagai, Ms. C. Furusawa, Ms. A. Nagashii, Ms. A. Sugai, and Ms. R. Takahashi for experimental technical support. We are grateful to Dr. Y. Takaoka (Tohoku University) for his help with the synthesis of chemicals used in this study, Drs. F. Fauser, S. Schiml, and H. Puchta (Karlsruhe Institute of Technology) for providing the vectors (pDe_Cas9_KAN and pEn_Chimera), Drs. M. Ron and A. Britt (University of California, Davis) for providing pMR217 and pMR218, Drs. H. Saika, M. Endo, A. Endo, H. Kaya, and S. Hirose (NARO) for insightful discussions and comments, and Dr. H. Rothnie for English editing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgeed.2020.604289/full#supplementary-material

REFERENCES

Baltes, N. J., Gil-Humanes, J., Cermak, T., Atkins, P. A., and Voytas, D. F. (2014). DNA replicons for plant genome engineering. *Plant Cell* 26, 151–163. doi: 10.1105/tpc.113.119792

Butler, N. M., Baltes, N. J., Voytas, D. F., and Douches, D. S. (2016). Geminivirus-mediated genome editing in potato (*Solanum tuberosum* L.) using sequence-specific nucleases. *Front. Plant Sci*. 7:1045. doi: 10.3389/fpls.2016.01045

Cermák, T., Baltes, N. J., Cegan, R., Zhang, Y., and Voytas, D. F. (2015). High-frequency, precise modification of the tomato genome. *Genome Biol*. 16:232. doi: 10.1186/s13059-015-0796-9

Chen, X. (2004). A microRNA as a translational repressor of APETALA2 in *Arabidopsis* flower development. *Science* 303, 2022–2025. doi: 10.1126/science.1088060

Danner, E., Bashir, S., Yumlu, S., Wurst, W., Wefers, B., and Kühn, R. (2017). Control of gene editing by manipulation of DNA repair mechanisms. *Mamm. Genome*. 28, 262–274. doi: 10.1007/s00335-017-9688-5

Endo, M., Mikami, M., Endo, A., Kaya, H., Itoh, T., Nishimasu, H., et al. (2019). Genome editing in plants by engineered CRISPR-Cas9 recognizing NG PAM. *Nat. Plants* 5, 14–17. doi: 10.1038/s41477-018-0321-8

Endo, M., Osakabe, K., Ono, K., Handa, H., Shimizu, T., and Toki, S. (2007). Molecular breeding of a novel herbicide-tolerant rice by gene targeting. *Plant J*. 52, 157–166. doi: 10.1111/j.1365-313X.2007.03230.x

Fauser, F., Roth, N., Pacher, M., Ilg, G., Sanchez-Fernandez, R., Biesgen, C., et al. (2012). In planta gene targeting. *Proc. Natl. Acad. Sci. U.S.A*. 109:7535–7540. doi: 10.1073/pnas.1202191109

Fauser, F., Schiml, S., and Puchta, H. (2014). Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. *Plant J*. 79, 348–359. doi: 10.1111/tpj.12554

Gallego, M. E., Sirand-Pugnet, P., and White, C. I. (1999). Positive-negative selection and T-DNA stability in *Arabidopsis* transformation. *Plant Mol. Biol*. 39, 83–93. doi: 10.1023/A:1006192225464

Gil-Humanes, J., Wang, Y., Liang, Z., Shan, Q., Ozuna, C. V., Sánchez-León, S., et al. (2017). High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. *Plant J*. 89, 1251–1262. doi: 10.1111/tpj.13446

Greco, G. E., Conrad, Z. A., Johnston, A. M., Li, Q. Y., and Tomkinson, A. E. (2016). Synthesis and structure determination of SCR7, a DNA ligase inhibitor. *Tetrahedron Lett*. 57, 3204–3207. doi: 10.1016/j.tetlet.2016.06.037

Hengel, S. R., Spies, M. A., and Spies, M. (2017). Small-molecule inhibitors targeting DNA repair and DNA repair deficiency in research and cancer therapy. *Cell Chem. Biol*. 2, 1101–1119. doi: 10.1016/j.chembiol.2017.08.027

Hood, E. E., Gelvin, S. B., Melchers, L. S., and Hoekema, A. (1993). New *Agrobacterium* helper plasmids for gene-transfer to plants. *Transgenic Res*. 2, 208–218. doi: 10.1007/BF01977351

Horsch, R. B., Fry, J. E., Hoffmann, N. L., Eichholtz, D., Rogers, S. G., and Fraley, R. T. (1985). A simple and general method for transferring genes into plants. *Science* 227, 1229–1231. doi: 10.1126/science.227.4691.1229

Hrouda, M., and Paszkowski, J. (1994). High fidelity extrachromosomal recombination and gene targeting in plants. *Mol. Gen. Genet*. 243, 106–111. doi: 10.1007/BF00283882

Hummel, A. W., Chauhan, R. D., Cermak, T., Mutka, A. M., Vijayaraghavan, A., Boyher, A., et al. (2018). Allele exchange at the EPSPS locus confers glyphosate tolerance in cassava. *Plant Biotechnol. J*. 16, 1275–1282. doi: 10.1111/pbi.12868

Jayathilaka, K., Sheridan, S. D., Bold, T. D., Bochenska, K., Logan, H. L., Weichselbaum, R. R., et al. (2008). A chemical compound that stimulates the human homologous recombination protein RAD51. *Proc. Natl. Acad. Sci. U.S.A*. 105, 15848–15853. doi: 10.1073/pnas.0808046105

Mikami, M., Toki, S., and Endo, M. (2015). Comparison of CRISPR/Cas9 expression constructs for efficient targeted mutagenesis in rice. *Plant Mol. Biol*. 88, 561–572. doi: 10.1007/s11103-015-0342-x

Miki, D., Zhang, W., Zeng, W., Feng, Z., and Zhu, J. K. (2018). CRISPR/Cas9-mediated gene targeting in *Arabidopsis* using sequential transformation. *Nat. Commun*. 9:1967. doi: 10.1038/s41467-018-04416-0

Murashige, T., and Skoog, F. (1962). A revised medium for the rapid growth and bioassay with tobacco tissue cultures. *Physiol. Plant* 15, 473–497. doi: 10.1111/j.1399-3054.1962.tb08052.x

Nair, S. K., Wang, N., Turuspekov, Y., Pourkheirandish, M., Sinsuwongwat, S., Chen, G., et al. (2010). Cleistogamous flowering in barley arises from the suppression of microRNA-guided HvAP2 mRNA cleavage. *Proc. Natl. Acad. Sci. U.S.A*. 107, 490–495. doi: 10.1073/pnas.0909097107

Nishizawa-Yokoi, A., Endo, M., Ohtsuki, N., Saika, H., and Toki, S. (2015). Precision genome editing in plants via gene targeting and piggyBac-mediated marker excision. *Plant J*. 8, 160–168. doi: 10.1111/tpj.12693

Ono, A., Yamaguchi, K., Fukada-Tanaka, S., Terada, R., Mitsui, T., and Iida, S. (2012). A null mutation of ROS1a for DNA demethylation in rice is not transmittable to progeny. *Plant J*. 71, 564–574. doi: 10.1111/j.1365-313X.2012.05009.x

Pinder, J., Salsman, J., and Dellaire, G. (2015). Nuclear domain 'knock-in' screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing. *Nucleic Acids Res*. 43, 9379–9392. doi: 10.1093/nar/gkv993

Puchta, H. (2002). Gene replacement by homologous recombination in plants. *Plant Mol. Biol*. 48, 173–182 doi: 10.1023/A:1013761821763

Puchta, H. (2017). Applying CRISPR/Cas for genome engineering in plants: the best is yet to come. *Curr. Opin. Plant Biol*. 36, 1–8. doi: 10.1016/j.pbi.2016.11.011

Ritter, A., Iñigo, S., Fernández-Calvo, P., Heyndrickx, K. S., Dhondt, S., Shi, H., et al. (2017). The transcriptional repressor complex FRS7-FRS12 regulates flowering time and growth in *Arabidopsis*. *Nat. Commun*. 8:15235. doi: 10.1038/ncomms15235

Saika, H., and Toki, S. (2010). Mature seed-derived callus of the model indica rice variety kasalath is highly competent in *Agrobacterium*-mediated transformation. *Plant Cell Rep*. 29, 1351–1364. doi: 10.1007/s00299-010-0921-x

Schiml, S., Fauser, F., and Puchta, H. (2014). The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in *Arabidopsis* resulting in heritable progeny. *Plant J*. 80, 1139–1150. doi: 10.1111/tpj.12704

Song, J., Yang, D., Xu, J., Zhu, T., Chen, Y. E., and Zhang, J. (2016). RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nat. Commun*. 7:10548. doi: 10.1038/ncomms10548

Steinrücken, H. C., and Amrhein, N. (1980). The herbicide glyphosate is a potent inhibitor of 5-enolpyruvyl-shikimic acid-3-phosphate synthase. *Biochem. Biophys. Res. Commun*. 94, 1207–1212. doi: 10.1016/0006-291X(80)90547-1

Terada, R., Urawa, H., Inagaki, Y., Tsugane, K., and Iida, S. (2002). Efficient gene targeting by homologous recombination in rice. *Nat. Biotechnol*. 20, 1030–1034. doi: 10.1038/nbt737

Toki, S., Hara, N., Ono, K., Onodera, H., Tagiri, A., Oka, S., et al. (2006). Early infection of scutellum tissue with *Agrobacterium* allows high-speed transformation of rice. *Plant J*. 47, 969–976. doi: 10.1111/j.1365-313X.2006.02836.x

Vats, S., Kumawat, S., Kumar, V., Patil, G. B., Joshi, T., Sonah, H., et al. (2019). Genome editing in plants: exploration of technological advancements and challenges. *Cells* 8:1386. doi: 10.3390/cells8111386

Voytas, D. F. (2013). Plant genome engineering with sequence-specific nucleases. *Annu. Rev. Plant Biol*. 64, 327–350. doi: 10.1146/annurev-arplant-042811-105552

Wang, M., Lu, Y., Botell,a, J. R., Mao, Y., Hua, K., and Zhu, J.K. (2017). Gene targeting by homology-directed repair in rice using a geminivirus-based CRISPR/Cas9 system. *Mol. Plant* 10, 1007–1010. doi: 10.1016/j.molp.2017.03.002

Wolter, F., Klemm, J., and Puchta, H. (2018). Efficient in planta gene targeting in *Arabidopsis* using egg cell-specific expression of the Cas9 nuclease of *Staphylococcus aureus*. *Plant J*. 94, 735–746. doi: 10.1111/tpj.13893

Yamauchi, T., Johzuka-Hisatomi, Y., Fukada-Tanaka, S., Terada, R., Nakamura, I., and Iida, S. (2009). Homologous recombination-mediated knock-in targeting of the MET1a gene for a maintenance DNA methyltransferase reproducibly reveals dosage-dependent spatiotemporal gene expression in rice. *Plant J*. 60, 386–396. doi: 10.1111/j.1365-313X.2009.03947.x

Yamauchi, T., Johzuka-Hisatomi, Y., Terada,. R., Nakamura, I., and Iida, S. (2014). The MET1b gene encoding a maintenance DNA methyltransferase is indispensable for normal development in rice. *Plant Mol. Biol*. 85, 219–232. doi: 10.1007/s11103-014-0178-9

Yu, Q., Jalaludin, A., Han, H., Chen, M., Sammons, R. D., and Powles, S. B. (2015). Evolution of a double amino acid substitution in the 5-enolpyruvylshikimate-3-phosphate synthase in *Eleusine indica* conferring high-level glyphosate resistance. *Plant Physiol*. 167, 1440–1447. doi: 10.1104/pp.15.00146

Zhang, Y., Zhang, F., Li, X., Baller, J. A., Qi, Y., Starker, C. G., et al. (2013). Transcription activator-like effector nucleases enable efficient plant genome engineering. *Plant Physiol*. 161:20–27. doi: 10.1104/pp.112.205179

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Nishizawa-Yokoi, Mikami and Toki. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genome Editing

ORIGINAL RESEARCH
published: 03 December 2020
doi: 10.3389/fgeed.2020.607982

Repurposing of Anthocyanin Biosynthesis for Plant Transformation and Genome Editing

Yubing He[1,2†], Min Zhu[2†], Junhua Wu[2†], Lejun Ouyang[3], Rongchen Wang[4], Hui Sun[2], Lang Yan[2], Lihao Wang[2], Meilian Xu[2], Huadong Zhan[1] and Yunde Zhao[5]*

[1] State Key Laboratory of Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing, China, [2] National Key Laboratory of Crop Genetic Improvement and National Center of Plant Gene Research (Wuhan), Huazhong Agricultural University, Wuhan, China, [3] Guangdong Laboratory for Lingnan Modern Agricultural Science and Technology, Guangdong University of Petrochemical Technology, Maoming, China, [4] Key Laboratory of Plant Resource Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China, [5] Section of Cell and Developmental Biology, University of California, San Diego, La Jolla, CA, United States

OPEN ACCESS

Edited by:
Lanqin Xia,
Chinese Academy of Agricultural Sciences, China

Reviewed by:
Yiping Qi,
University of Maryland, United States
Kan Wang,
Iowa State University, United States

***Correspondence:**
Yubing He
yubinghe@njau.edu.cn

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Genome Editing in Plants, a section of the journal Frontiers in Genome Editing

Received: *18 September 2020*
Accepted: *09 November 2020*
Published: *03 December 2020*

Citation:
He Y, Zhu M, Wu J, Ouyang L, Wang R, Sun H, Yan L, Wang L, Xu M, Zhan H and Zhao Y (2020) Repurposing of Anthocyanin Biosynthesis for Plant Transformation and Genome Editing. Front. Genome Ed. 2:607982. doi: 10.3389/fgeed.2020.607982

CRISPR/Cas9 gene editing technology has been very effective in editing genes in many plant species including rice. Here we further improve the current CRISPR/Cas9 gene editing technology in both efficiency and time needed for isolation of transgene-free and target gene-edited plants. We coupled the CRISPR/Cas9 cassette with a unit that activates anthocyanin biosynthesis, providing a visible marker for detecting the presence of transgenes. The anthocyanin-marker assisted CRISPR (AAC) technology enables us to identify transgenic events even at calli stage, to select transformants with elevated *Cas9* expression, and to identify transgene-free plants in the field. We used the AAC technology to edit *LAZY1* and *G1* and successfully generated many transgene-free and target gene-edited plants at T1 generation. The AAC technology greatly reduced the labor, time, and costs needed for editing target genes in rice.

Keywords: CRISPR, transgene-free, anthocyanin, rice, AAC

INTRODUCTION

CRISPR/Cas9 genome editing technology has been widely used to generate targeted modifications of genes in many plant species (Feng et al., 2013; Li et al., 2013a; Mao et al., 2013; Miao et al., 2013; Nekrasov et al., 2013; Shan et al., 2013; Xie and Yang, 2013; Gao et al., 2015, 2016). Gene editing efficiency correlates with the expression level of *Cas9* and higher expression of *Cas9* usually leads to an increase in editing efficiency. Selecting transgenic plants with elevated *Cas9* expression by analyzing Cas9 protein concentrations is laborious and time-consuming, Therefore, *Cas9* expression levels are often monitored indirectly. Coupling a fluorescence marker with the *Cas9* expression unit provides an effective approach for identify plants with elevated Cas9 concentrations (Gao et al., 2016; Wang and Chen, 2020). Another creative approach was to couple *Cas9* expression with a guide RNA that can lead to a visible phenotype when the target gene is edited. Genes involved in trichome development have been targeted to visibly monitor gene editing efficiency (Wang et al., 2015; Miki et al., 2018). Whereas, the aforementioned methods have been effective, fluorescence markers need special equipment and are not suitable for field conditions. Targeting trichomes is very useful for Arabidopsis, but may not be effective for monocots such as rice. Additional markers that can be indicative of the levels of transgene expression will be very useful in conducting gene editing experiments.

In addition to improving gene editing efficiency, a main emphasis in editing plants/crops is to generate edited plants without any transgene residuals. Crops with any CRISPR/Cas9 component residual will unlikely receive approval for commercial applications from government regulatory agencies. The continuous presence of gene editing elements in plants may cause genetic instability and off-target events. Several strategies have been reported for isolating transgene-free and target gene-edited plants. Transient expression of *Cas9* and gRNA genome editing complex through Agrobacterium-mediated infiltration has led to the identification of target gene-edited plants without any transgene residues in tobacco (Chen et al., 2018). Cas9 and gRNA can also be assembled into ribonucleoprotein (RNP) complexes and then delivered into plant cells using nano particles (Doyle et al., 2019; Landry and Mitter, 2019) or bombardment (Svitashev et al., 2016; Zhang et al., 2016). Because RNP contains no DNA, any edited plants are considered transgene-free, However, RNP method is extremely inefficient because the majority of the regenerated plants are not transformed due to a lack of selection pressure. Identification of transgene-free, target gene-edited plants can also be assisted by fluorescence markers (Gao et al., 2016; He et al., 2017a; Yu and Zhao, 2019; Ouyang et al., 2020) and positive or negative selection against chemicals (Lu et al., 2017; Wu et al., 2019; Li et al., 2020), Moreover, we developed a CRISPR/Cas9 gene editing system in which the transgenes undergo automatic self-elimination after the target gene has been edited (transgene killer CRISPR (TKC), greatly improving the efficiency for isolating the desired plants (He et al., 2018; 2019). Whereas, TKC system reduces labor and time needed for conducting gene editing in rice, it does not have a proxy for indicating *Cas9* expression levels.

Anthocyanins are a large class of secondary metabolites (Tanaka et al., 2008), which are widely distributed in various tissues including flowers, stems, leaves, and fruits. Anthocyanins are water-soluble polyphenol pigments with vivid red, blue, purple, and other colors, Therefore, anthocyanins potentially can be used as a visible marker for visualizing transgenic events or the presence of transgenes, However, anthocyanin biosynthetic pathway is very complex and consists of at least eight genes (Zhu et al., 2017; Zheng et al., 2019). It is not realistic to couple the entire anthocyanin biosynthetic pathway with Cas9/gRNA units in a plasmid because the resulting plasmid would be too big for efficient transformation or cloning. Fortunately, anthocyanin biosynthesis pathway is under the control of key transcription factors including MYB, bHLH, and WD40 family genes (Zhang et al., 2014; Liu et al., 2015; Xu et al., 2015). Almost all of the anthocyanin biosynthesis structural genes are activated (or enhanced) by the MYB–bHLH–WD40 (MBW) complex (Xu et al., 2015; Zhu et al., 2017; Zheng et al., 2019), and the MYB protein is believed to be the key component in the allocation of specific gene expression patterns (Jaakola, 2013; Xu et al., 2015). Based on the number of imperfect repeats (R) domain(s), MYB genes are divided into four major groups: 1R (R1/2, R3-MYB), 2R (R2R3-MYB), 3R (R1R2R3-MYB), and 4R (four R1/R2-like repeats) (Liu et al., 2015). The R2R3-MYB is an activator for the synthesis of anthocyanins (Borevitz et al., 2000; Liu et al., 2015).

The *ZmC1* gene, encoding an R2R3 MYB, regulates the synthesis of anthocyanins in the corn aleurone layer (Cone et al., 1986; Paz-Ares et al., 1986). Its ortholog, *OsC1* gene controls the color change in leaf sheath and stigma in rice. Ectopic expression of *OsC1* in different rice varieties results in accumulation of anthocyanins in various tissues and organs (Gao et al., 2011; Chin et al., 2016; Zhao et al., 2016), Therefore, ectopic expression of *OsC1* may be used to activate anthocyanin biosynthesis, thus providing a visible marker in rice. Here, we show that the expression of *OsC1* under the control of the rice *ACTIN* promoter leads to the accumulation

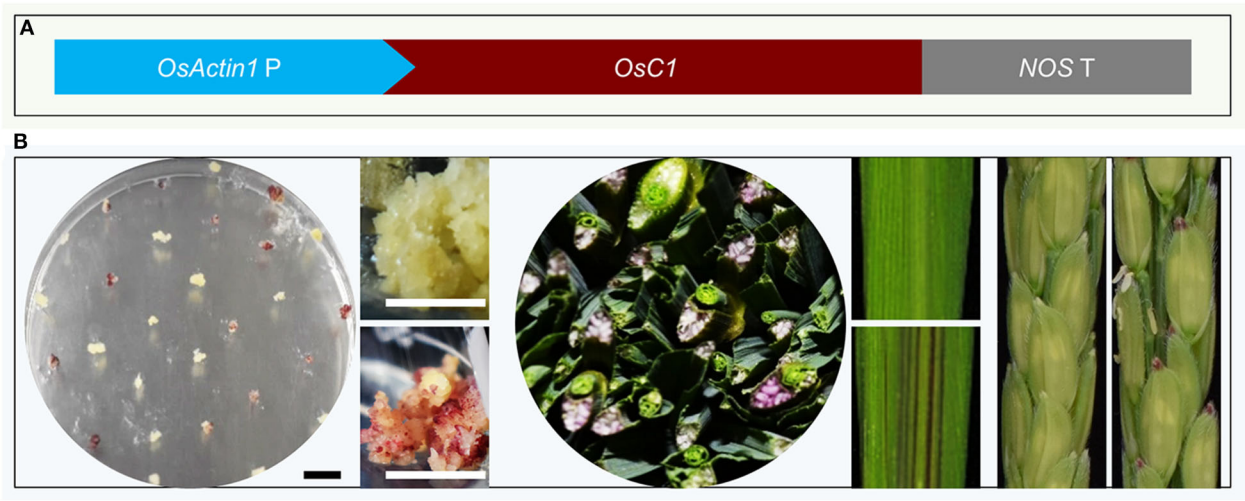


FIGURE 1 | The *pACTIN-OsC1* activates anthocyanin biosynthesis in rice. **(A)** A schematic description of the *pACTIN-OsC1* unit in the vector used for rice transformation. The anthocyanin regulation gene *OsC1* is under the control of the rice *OsACTIN1* promoter. *NOS* T refers to the terminator of the nopaline synthase gene from *Agrobacterium tumefaciens*. **(B)** The *AAC* plasmid induced obvious accumulation of anthocyanin (purple color) in rice calli, vascular tissue in stems (transverse section), leaves, and the tip of grains. Scale bars 1 cm in tissue culture images.

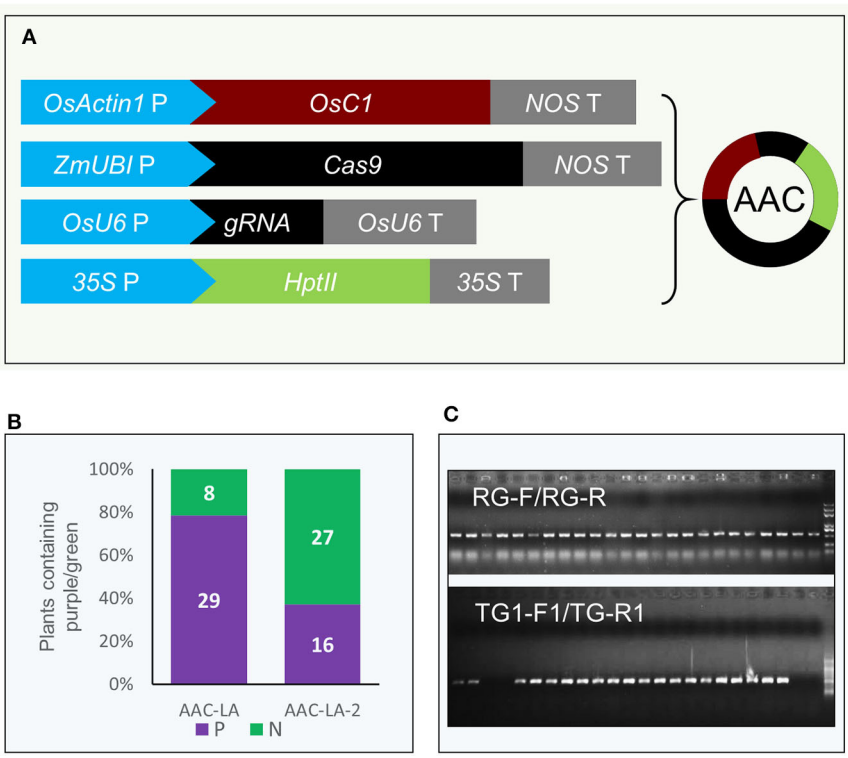


FIGURE 2 | Anthocyanin accumulation as a marker for the presence of transgenes in gene editing experiments. **(A)** Coupling the *pACTIN-OsC1* unit with the CRISPR/Cas9 genome editing units in the *AAC* (Anthocyanin-marker Assisted CRISPR) plasmid. The anthocyanin regulatory gene *OsC1* is under the control of the rice *OsACTIN1* promoter. *NOS* T refers to the terminator of the nopaline synthase gene from *Agrobacterium tumefaciens*. The rice codon-optimized *Cas9* was placed under the control of the maize *UBIQUITIN* promoter. *OsU6* P and *OsU6* T represent rice *U6* promoter and terminator, respectively. gRNA refers to guide RNA. The transcription of *hygromycin phosphotransferase II* gene *HptII* is under the control of the *CaMV 35S* promoter and the transcription was terminated with a *CaMV 35S* terminator. **(B)** Not all antibiotic resistant plants generated purple color in T0. AAC-LA was transgenic event screening from the plate with two rounds hygromycin selection (10 days in a round). AAC-LA-2 was regenerated from the calli on normal plate after 7 days hygromycin selection (only one round). "P" and "N" represent purple and normal color plants, respectively. **(C)** Verification of the transgenes in the rice plants by PCR. "RG-F/RG-R" and "TG-F1/TG-R1" refer to the primer pairs used for detection the presence of rice genomic DNA and the T-DNA of *AAC* plasmid, respectively. The primer information was described in **Supplementary Table 3**.

of anthocyanin in calli, young seedlings, leaf vascular tissue, and grains, providing a visible marker for selecting transgenic plants. When the *OsC1* unit is coupled with CRISPR/Cas9 units, target genes in almost all of the anthocyanin positive plants have been edited, suggesting that the threshold of visible anthocyanin accumulation is a good indication of high level *Cas9* expression, Moreover, anthocyanin accumulation allows a visual differentiation between transgenic and non-transgenic plants at T1 generation without the need of conducting PCR and other molecular analysis, greatly accelerating the isolation of transgene free and target-gene edited plants.

RESULTS

Activation of Anthocyanin Biosynthesis Serves as a Marker for the Presence of Transgenes

Ectopic expression of the *R2R3-MYB* gene in rice and other plants activated anthocyanin biosynthesis (Borevitz et al., 2000; Shin et al., 2006; Han et al., 2009; Li et al., 2017; Zhang et al., 2019). We chose the *OsC1* (GenBank ID: MK636605), an *R2R3-MYB* gene, from the rice cultivar Heishuai (Zheng et al., 2019), which displays obvious purple color in leaves, stems, and grains. We placed *OsC1* cDNA under the control of the rice *ACTIN1* promoter, a constitutively activated promoter (**Figure 1A**). When plasmids containing the *pACTIN-OsC1* unit were transformed into rice, we observed that anthocyanin accumulated in calli, leaves, and the tip of grains (**Figure 1B**), demonstrating that the *pACTIN-OsC1* unit can be used to produce a visible color, which can facilitate the identification of transgenic events in calli and transgenic plants.

Coupling Anthocyanin Biosynthesis With Genome Editing Units

We placed the *pACTIN-OsC1* unit in adjacent to the cassettes for *Cas9* expression and gRNA production, resulting in the *AAC* plasmid (*A*nthocyanin-marker *A*ssisted *C*RISPR) (**Figure 2A**). We first tested whether our *AAC* plasmids can achieve efficient editing of the *LAZY1* gene, which leads to agravitropic growth when compromised (Li et al., 2007), providing a visible

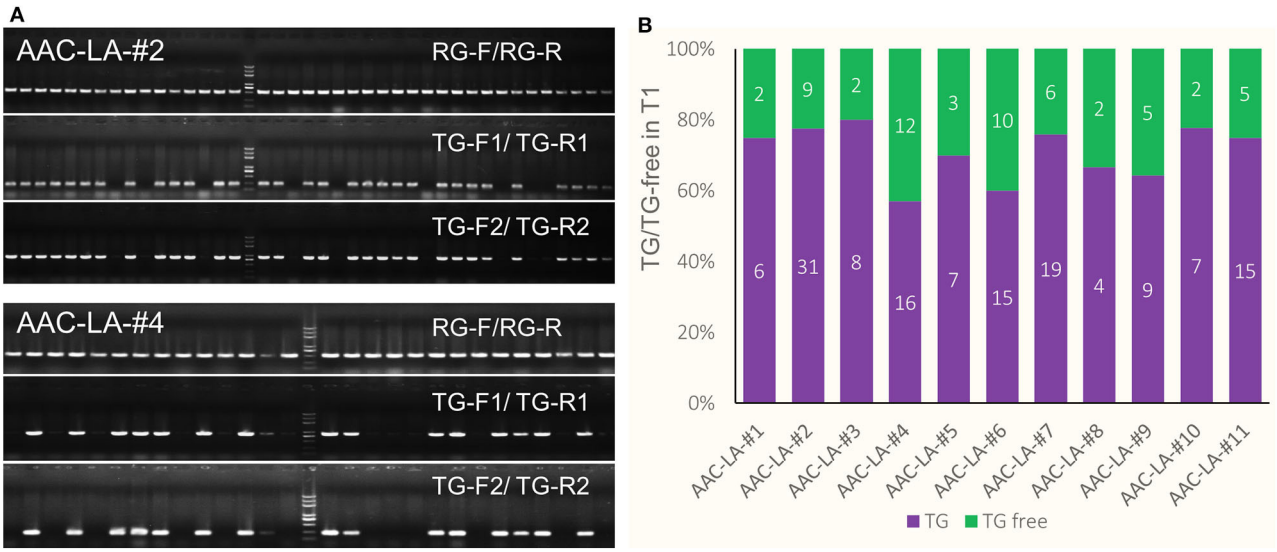


FIGURE 3 | Isolation of transgene-free T1 plants by anthocyanin-marker assisted strategy. **(A)** The transgene-free plants from the progeny of the T0 plants that had purple color were identified by PCR and visual screen. AAC-LA–#2 and AAC-LA-#4 represent two AAC-LA T0 plants with purple color, respectively. **(B)** The segregation of the transgene (TG) and transgene-free (TG-free) plants from 11 AAC-LA T0 lines. The numbers on purple and green columns represent the numbers of purple and normal colored plants at T1 generation. The results were consistent with the results identified by PCR. "RG-F/RG-R" refers to primer pair used for checking the quality of rice genome DNA. "TG-F1/TG-R1," and "TG-F2/TG-R2" represent two primer pairs used for detection the presence of T-DNA, respectively. The primers information was described in **Supplementary Table 3**.

phenotype for gene editing events. We cloned the *LAZY1* gRNA unit into the *AAC* plasmid to generate the *AAC-LA* plasmid, which was subsequently transformed into rice through *Agrobacterium*-mediated transformation. We found that almost 80% of the T0 plants regenerated from the plate with two rounds hygromycin selection (10 days in a round) had purple color tissues (**Figure 2B**). Furthermore, we noticed that transferring the calli to antibiotic-free media after 7 days hygromycin selection (only one round), resulted in about 40% regenerated plants with purple color tissues (**Figure 2B**), suggesting that activation of anthocyanin can be used for selecting transgenic events with a decreased usage of antibiotic. Whereas, antibiotic or herbicide resistant markers are potentially detrimental to the environment, the native anthocyanin biosynthesis pathway offers a good alternative for selecting transformants.

We observed that all of the T0 plants that displayed purple color contained the *OsC1* expression cassette revealed by our PCR analyses (**Figure 2C**). We also found that some of the hygromycin-resistant plants without the purple color also contained the transgene, suggesting that a threshold level expression of the *OsC1* is required in order to produce a visible color. The results suggest that we might be able to use the accumulation of anthocyanin as a proxy for monitoring Cas9 expression levels.

Facilitating the Isolation of Transgene-Free and Target Gene-Edited T1 Plants Using the Anthocyanin-Marker

We hypothesized that we might be able to visually identify transgene free T1 plants generated from the T0 plants that had obvious anthocyanin accumulation. We also used PCR-based assays to further confirm the absence of transgenes. We used two pairs of primers, TG-F1/TG-R1 and TG-F2/TG-R2 to amply specific DNA fragments from the *AAC* plasmid (**Figure 3A**). We found that the PCR results matched perfectly with the color based visual screen. For example, among 40 and 28 T1 plants from the AAC-LA-#2 and AAC-LA-#4 T0 plants, 9 and 12 plants did not display obvious anthocyanin accumulation, respectively (**Figure 3B**). Our PCR results identified the same plants as transgene-free (**Figure 3A**). Further analyses of the T1 plants from the AAC-LA–#1 to # 11 T0 plants demonstrated that the transgene-free T1 plants identified on basis of anthocyanin accumulation matched with those transgene-free plants identified by PCR assays (**Figure 3B**). Our results demonstrated that the anthocyanin-based visual screen for transgene-free T1 plants was effective and accurate. Such a visual assay greatly reduced labor and time for identifying transgene-free plants.

Molecular Characterization of the Transgene-Free and Gene-Edited Plants

To analyze the molecular lesions in the *LAZY1* gene in the T1 plants generated from the AAC-LA T0 plants that had displayed anthothyanin accumulation, we directly sequenced the *LAZY1* fragment amplified using the primers of LAZY1-GTF/ LAZY1-GTR (**Figure 4A**). We analyzed T1 plants from 11 independent T0 plants, which had obvious anthocyanin accumulation (**Table 1**, **Supplementary Table 1**). Every single plant contained mutations at the target site except that a few plants failed to generate quality sequencing data. The T1

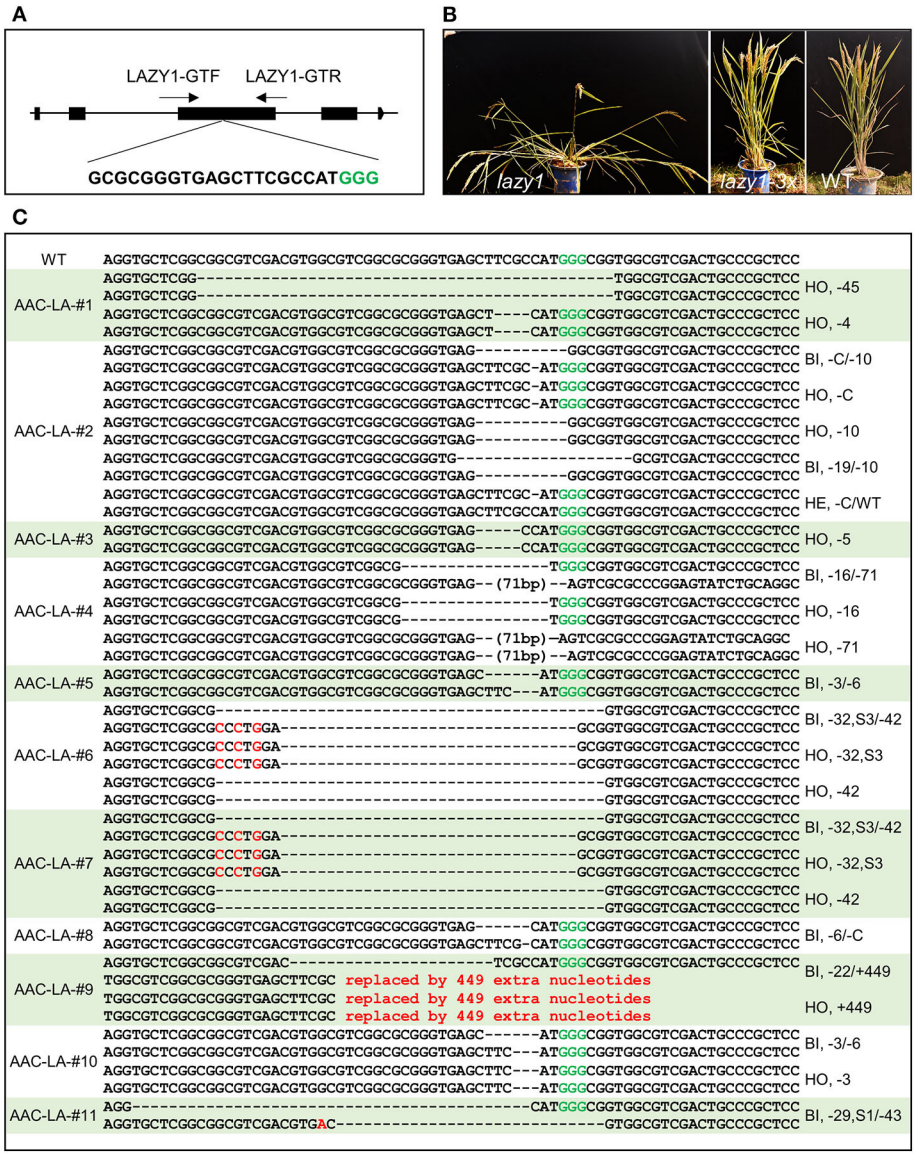


FIGURE 4 | Identification of transgene-free and target gene-edited *lazy1* mutants. **(A)** Target site of *LAZY1* used in *AAC-LA*. A target sequence in the *LAZY1* gene includes the PAM site GGG (marked green) is shown. LAZY1-GTF and LAZY1-GTR were the genotyping primer pair of *LAZY1* CRISPR plants. Primers were listed in **Supplementary Table 3**. **(B)** Loss-of-function *lazy1* mutants displayed an obvious tiller angle phenotype. WT, *lazy1* and *lazy1-3x* refer to wild type, loss-of-function *lazy1* mutant, and in-frame (multiple of three base pairs deletion) *lazy1* mutant plants, respectively. **(C)** The mutation forms of the transgene-free and CRISPR-edited *lazy1* mutants. The PAM site "GGG" in green required for Cas9 cleavage is marked in green. "WT" refers to the wild type plants. "HO," "HE," and "BI" represent homozygous, heterozygous, and bi-allelic genotypes, respectively. "-" refers to a deletion of one base pair. "S3" and "S1" means three and one base pair(s) substitution, respectively.

plants analyzed were either homozygous or bi-allelic except one T1 plant from AAC-LA-#2, which was heterozygous, demonstrating the power of AAC technology in gene editing (**Table 1**, **Supplementary Table 1**). For comparison, our previously experiments using our regular CRISPR vectors usually had the editing efficiency <80% (He et al., 2017b).

Further analysis of the mutations revealed some interesting patterns. We observed four *lazy* alleles in the progeny of AAC-LA−2 T0 plant: a "C" deletion, a 10-base pair deletion, a 40-base pair deletion, and a 19-base pair deletion. The *lazy* alleles resulted in six combinations of genotypes: eighteen "–C/−10" biallelic mutants, twelve "–C" homozygous mutants, seven "-10" homozygous mutants, one "−40/−19" biallelic, one "−19/−10" biallelic mutant, and a wild type plant in the T1 plants (**Table 1**, **Supplementary Table 1**). The first three genotypes were predominant, accounting for 92.5% (37/40) of the T1 plants. Because of the complex

TABLE 1 | Genotypes and segregation patterns in the T1 plants generated from the AAC-LA.

Plants No.	Genotypes	Segregation ratio	Transgene-free ratio
AAC-LA–#1	HO,−45	6/8	1/6
	HO,−4	2/8	1/2
AAC-LA–#2	BI,–C/−10	18/40	3/18
	HO,–C	12/40	2/12
	HO,−10	7/40	2/7
	BI,−40/−10	1/40	0/1
	BI,−19/−10	1/40	1/1
	HE,–C/WT	1/40	1/1
AAC-LA–#3	HO,−411	4/9	0/4
	HO,−5	5/9	1/5
AAC-LA–#4	BI,−16/−71	15/26	5/15
	HO,−16	3/26	1/3
	HO,−71	8/26	5/8
AAC-LA–#5	BI,−3/−6	3/7	1/3
	HO,−3	2/7	0/2
	HO,−6	2/7	0/2
AAC-LA–#6	BI,−32,S3/−42	11/23	4/11
	HO,−32,S3	7/23	3/7
	HO,−42	5/23	2/5
AAC-LA–#7	BI,−32,S3/−42	13/25	5/13
	HO,−32,S3	5/25	1/5
	HO,−42	7/25	1/7
AAC-LA–#8	BI,−6/–C	3/5	1/3
	HO,−6	1/5	0/1
	HO,–C	1/5	0/1
AAC-LA–#9	BI,−22/+449	5/13	1/5
	HO,−22	2/13	0/2
	HO, +449	6/13	2/6
AAC-LA–#10	BI,−3/−6	4/8	1/4
	HO,−3	3/8	1/3
	HO,−6	1/8	0/1
AAC-LA–#11	BI,−29,S1/−43	11/18	4/11
	HO,−29,S1	6/18	0/6
	HO,−43	1/18	0/1

"HO," "HE," and "BI" represent homozygous, heterozygous, and bi-allelic genotype, respectively. The numbers in the column of "Genotypes" refer to the numbers of base pair changes in each line. The details of the mutations of each genotype were shown in ***Supplementary Table 1****. "Segregation ratio" refers the ratio of the number of the plants shown the genotype in the column of "Genotypes" to the total number of the plants in the T1 line.*

genotype (likely mosaic) in T0 plants, it is extremely important to analyzed T1 plants that no longer have the *Cas9* transgenes.

Using anthocyanin as a visible marker for transgenes (**Figure 3**) of the T1 plants, we easily identified transgene-free *lazy1* mutants of at the T1 generation from multiple independent T0 plants (**Figures 4B,C**). Some of T0 plants did not show the *lazy* mutant phenotype, but they harbored mutations with 3x of base pair deletion, such as the lines from AAC-LA–#1, 5, 6, 7, 8, 10 (**Figure 4C**, **Table 1**, **Supplementary Table 1**). Such 3x mutations did not cause frameshift and likely produced functional protein.

Testing the AAC Gene Editing System With Another Target

We assembled another construct *AAC-G1* to target the *G1* gene, which takes part in suppressing the development of the sterile lemma. When *G1* is disrupted, the length of the sterile lemma increased dramatically, providing an easily scorable phenotype (Yoshida et al., 2009). We analyzed 88 T1 plants from 9 AAC-G1 T0 plants that displayed anthocyanin accumulation. The results of our visual screen for transgene-free plants (**Figure 5A**) were consistent with the results identified by PCR using primer pairs TG-F1/TG-R1 and TG-F2/TG-R2 (**Figure 5B**). For example, there were 5 and 2 transgene-free plants in the progeny of AAC-G1–#5 and AAC-G1–#6, respectively (**Figures 5A,B**). We also identified multiple alleles of *g1* (**Figures 5C,D**, **Supplementary Table 2**). It was clear that all of the T1 plants sequenced were either homozygous or bi-allelic (**Supplementary Table 2**). Our results demonstrated that the anthocyanin-based screening for transgene-free and edited plants were very effective (**Figures 5D,E**).

DISCUSSION

Plant transformation through tissue culture usually takes weeks before a positive result is obtained. The ability to identify transgenic events at the earliest stages of plant transformation offers advantages. We showed that anthocyanin biosynthesis can be activated at calli stage, providing a visual and functional assay for positive transformants very early on (**Figure 6**). Anthocyanin biosynthesis does not require exogenous substrates. It allows continuous monitoring transgene activities under sterile conditions, which is very useful in tissue culture. Because anthocyanin is a plant pigment existent in many plants including some rice varieties, it is not toxic to plants and causes fewer environmental impacts compared to antibiotic and herbicide-resistance screening. In addition, anthocyanins can absorb excessive ultraviolet lights and visible lights and remove free radicals to protect plants from ultraviolet rays, thereby providing protection for plants (Guo et al., 2008). Anthocyanins can also enhance plant resistance to drought and low temperature, and strengthen the ability to resist pathogenic bacteria, thereby protecting plants from biotic and abiotic stresses (Ahmed et al., 2014), Moreover, anthocyanins can also respond to external trauma by preventing oxidation (Gould et al., 2002), Therefore, the T0 plants generated using our AAC system might have better chance to survive after transferring them from the sterile environment to natural field.

We used the rice *ACTIN* promoter, which is a strong and ubiquitously active promoter, to drive *OsC1* expression. Interestingly, anthocyanin accumulation was not ubiquitously distributed (**Figure 1**). We observed obvious purple color in calli, the vascular tissue of the stems and leaves, and the grains. But some other tissues did not display the purple color (**Figure 1B**). Previous studies showed that the anthocyanin biosynthesis pathway is regulated by the MYB-bHLH-WD40 (MBW) complex (Xu et al., 2015; Zhu et al., 2017; Zheng et al., 2019), and MYB

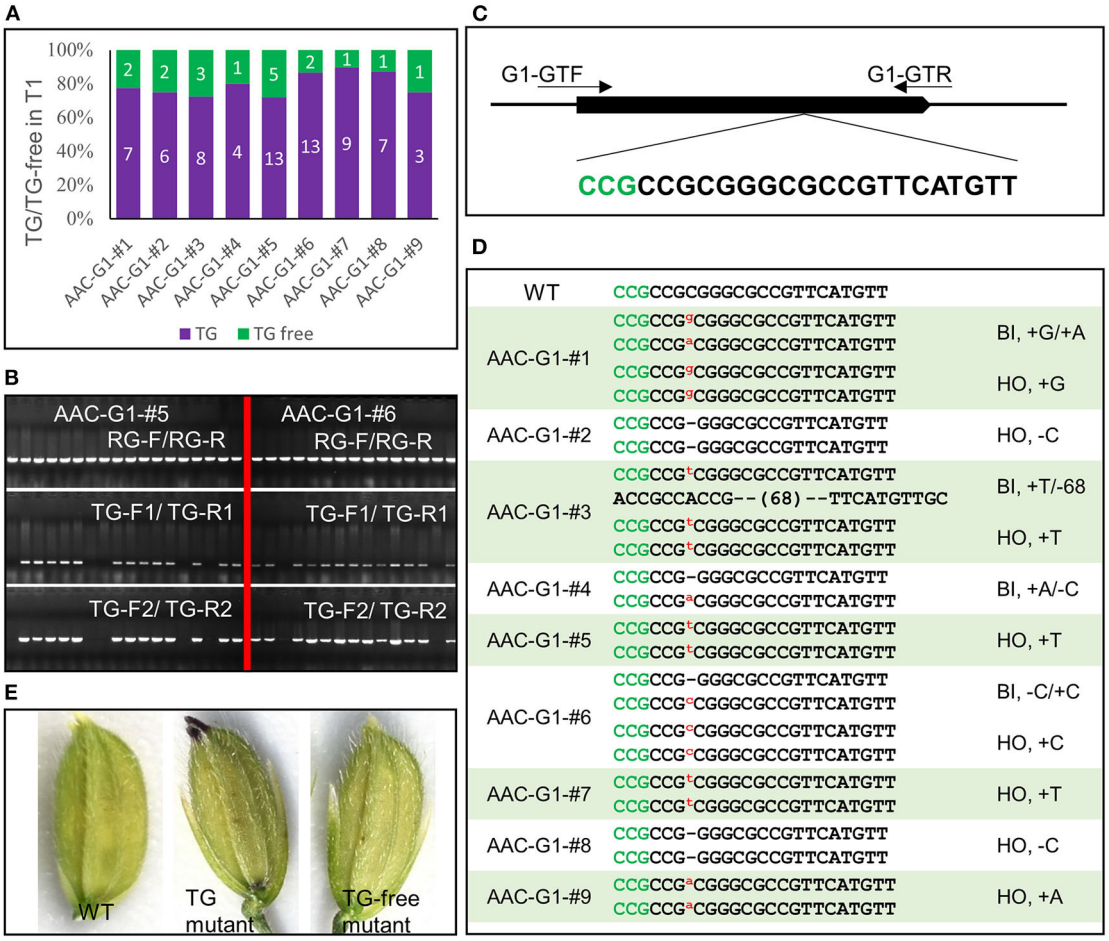


FIGURE 5 | Identification transgene-free and gene-edited *g1* mutants. **(A)** The segregation of the transgene (TG) and transgene-free (TG-free) plants from 9 AAC-G1 T1 lines. The numbers on purple and green columns represent the numbers of purple and normal color plants in T1 lines which were consistent with the results identified by PCR, respectively. **(B)** The transgene-free plants from the progeny of the purple-colored T0 plants were identified by PCR. AAC-G1–#5 and AAC-G1-# 6 refer to two AAC-G1 T0 plants containing the purple color, respectively. "RG-F/RG-R" refers to primer pair used for checking the quality of rice genome DNA. "TG-F1/TG-R1" and "TG-F2/TG-R2" represent two primer pairs used for detection the presence of T-DNA, respectively. The primers information was described in **Supplementary Table 3**. **(C)** Target site of *G1* used in *AAC-G1* plasmid. A target sequence including the PAM site CCG (marked green) was chosen. G1-GTF and G1-GTR are the genotyping primer pair of *G1* CRISPR plants. Primers were listed in **Supplementary Table 3**. **(D)** The mutation forms of the transgene-free and CRISPR-edited *g1* mutants. The PAM site "CCG" required for Cas9 cleavage is marked in green. "WT" refers to the wild-type plants. "HO" and "BI" represent homozygous and bi-allelic genotypes, respectively. "-" refers to a deletion of one base pair. "a," "g," "c" and "t" in red and superscript refers to an insertion of an "A," "G," "C," and "T," respectively. **(E)** The *g1* mutant florets with (TG) or without transgene (TG-free) generated by AAC-G1.

protein is the key component (Jaakola, 2013; Xu et al., 2015). Our activation of anthocyanin biosynthesis by expressing the transcription factor OsC1 only works in plants that have all functional anthocyanin biosynthetic genes and sufficient amount of bHLH and WD40. Many popular white rice varieties have defects in some of the anthocyanin genes (Zheng et al., 2019), rendering them not suitable for our AAC technology. Recently, we developed a novel color reporter RUBY (He et al., 2020), which can generate red color in all eukaryotic cells include rice. For those rice cultivars that AAC is not suitable, RUBY can be used for color-assisted CRISPR technology. We introduced the *pACTIN-OsC1* cassette to the rice variety Chao2-10 (Li et al., 2013b), which was previously shown to have all functional anthocyanin biosynthetic genes (Li et al., 2013b, 2017; Zheng et al., 2019), but lacked the *MYB* gene in the MBW complex. From a practical point view, the non-ubiquitous accumulation of anthocyanin actually is better because of its minimal impact on plant growth and development.

In our previous studies, we used antibiotics to screen the candidate gene-edited plants and found that some plants regenerated from antibiotic-resistance calli did not have target site mutations (He et al., 2017b). In this study, we found that all of the purple color plants were transgenic plants (**Figures 2**, **3**, **5A,B**) and all of the purple color plants had mutation at the target sites (**Table 1**, **Supplementary Tables 1**, **2**). Our interpretation is that anthocyanin accumulation may be indicative of elevated Cas9 expression, which directly determines gene editing efficiency.

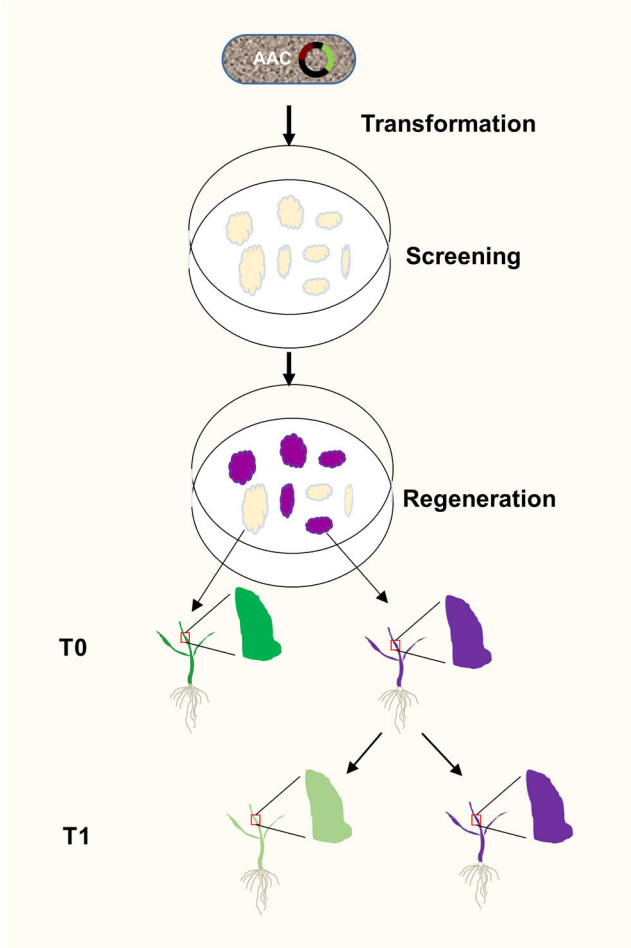


FIGURE 6 | A flow chart of AAC-mediated isolation of transgene-free and target gene-edited rice plants. The *AAC* plasmid was transformed into rice calli through *Agrobacterium*-mediated transformation. At the calli stage, the *OsC1* gene was expressed to generate purple callus, and the target gene is presumably being edited by Cas9/sgRNA complex. Consequently, at T1 stage, the transgene free plants without purple were selected from the progeny of the purple T0 plants.

In summary, we have repurposed the anthocyanin biosynthesis pathway for serving as a visible marker for selecting plant transformation events and as a proxy for the presence of transgenes (**Figure 6**). Our system offers an alternative to fluorescence markers and antibiotic selections. It is especially powerful when used in tissue culture and in combination with gene editing machinery. The extend of anthocyanin accumulation might be indicative of the expression levels of other linked transgenes such as *Cas9*, thus providing a robust and speedy method for identifying plants with elevated gene editing efficiency.

MATERIALS AND METHODS

The AAC Plasmid Construction

Primers used in this study were listed in the **Supplementary Table 3**. Our *AAC* plasmid contained two main expression cassettes: *OsACTIN1 promoter-OsC1* and *UBIQUITIN promoter-Cas9*. *OsACTIN1* promoter was amplified using primers of pCXUN-Act1PF and pCXUN-Act1PR from the plasmid *pCB2006*, which was kindly provided by Professor Lizhong Xiong (Xiao et al., 2009). The PCR product was cloned into *pCXUN-Cas9* (He et al., 2017b) at the *Kpn* I site. The *OsC1* gene was amplified from the cDNA plasmid provided by Dr. Hao Chen and Dr. Jie Zheng (Huazhong Agricultural University, Wuhan, China) by using primers of C1-Act1P-PCA9F and C1-Act1P-PCA9R, and the PCR fragment was inserted into the *Kpn* I site behind the *OsACTIN1* promoter to complete the *AAC* (Anthocyanin-marker Assisted CRISPR) plasmid construction. Guide RNA production cassette was inserted into the *Pme* I site of the *AAC* plasmid. We generated the plasmid *AAC-LA* and *AAC-G1*, which produce guide RNAs from the rice *U6* and *U3* promoter to target the rice *LAZY1* and *G1* gene by overlapping PCR, respectively (He et al., 2019). The correct clones were confirmed by sequencing with primer AAC-PmeI-seqF.

Plant Transformation

The *AAC-LA* and *AAC-G1* plasmids were transformed into Chao2-10 through *Agrobacterium*-mediated plant transformation following a protocol that was previously described (Hiei et al., 1994). T0 plants were visually scored for color phenotype at calli culture stage and at different growing stages in natural field. Seeds from each individual T0 plants containing purple color in vascular tissue of stems and leaves were harvested separately.

Characterization of T1 Plants

We randomly selected the T1 progenies of 11 T0 plants of AAC-LA and 9 independent AAC-G1 T0 plants to determine the efficiency of AAC system in editing the target genes. We used two primer pairs to amplify the specific regions of the AAC plasmids. TG-F1/TG-R1 was used to detect the presence of the *OsC1* expression cassette, and TG-F2/TG-R2 was used to amply part of the *CAS9* expression cassette. To check the quality of our genomic DNA samples, we used RG-F/RG-R for PCR reactions.

We used primer pairs LAZY1-GTF/LAZY1-GTF and G1-GTF/G1-GTR to amplify part of the *LAZY1* and *G1* genes from the AAC-LA and AAC-G1 T1 plants, respectively. We also directly sequenced the PCR products by using primers LAZY1-seq and G1-seq, respectively. For heterozygous or bi-allelic plants, the overlapping peaks were resolved using the publicly available Dsdecode site (http://skl.scau.edu.cn/dsdecode/) (Xie et al., 2017).

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

YH and YZ conceived the idea and wrote the first draft of the manuscript. YH, MZ, and JW conducted most of the experiments. MZ, JW, LO, RW, HS, LY, LW, MX, and HZ

contributed to manuscript revision. All authors contributed to the article and approved the submitted version.

FUNDING

This work was partially supported by grants from the National Transgenic Science and Technology Program (2019ZX08010-001 and 2019ZX08010-003) to YH and HZ, the Fundamental Research Funds for the Central Universities (JCQY201903) and a startup fund from the Nanjing Agricultural University.

ACKNOWLEDGMENTS

We thank Professor Lizhong Xiong for providing *pCB2006* plasmid. We are grateful to Dr. Hao Chen and Dr. Jie Zheng for kindly providing the cDNA of *OsC1* and the seeds of Chao2-10.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgeed.2020.607982/full#supplementary-material

REFERENCES

Ahmed, N. U., Park, J. I., Jung, H. J., Yang, T. J., Hur, Y., and Nou, I. S. (2014). Characterization of dihydroflavonol 4-reductase (DFR) genes and their association with cold and freezing stress in *Brassica rapa*. *Gene* 550, 46–55. doi: 10.1016/j.gene.2014.08.013

Borevitz, J. O., Xia, Y., Blount, J., Dixon, R. A., and Lamb, C. (2000). Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell* 12, 2383–2394. doi: 10.1105/tpc.12.12.2383

Chen, L., Li, W., Katin-Grazzini, L., Ding, J., Gu, X., Li, Y., et al. (2018). A method for the production and expedient screening of CRISPR/Cas9-mediated non-transgenic mutant plants. *Hortic Res.* 5:13. doi: 10.1038/s41438-018-0023-4

Chin, H. S., Wu, Y. P., Hour, A. L., Hong, C. Y., and Lin, Y. R. (2016). Genetic and evolutionary analysis of purple leaf sheath in Rice. *Rice* 9:8. doi: 10.1186/s12284-016-0080-y

Cone, K. C., Burr, F. A., and Burr, B. (1986). Molecular analysis of the maize anthocyanin regulatory locus C1. *Proc. Natl. Acad. Sci. U.S.A.* 83, 9631–9635. doi: 10.1073/pnas.83.24.9631

Doyle, C., Higginbottom, K., Swift, T. A., Winfield, M., Bellas, C., Benito-Alifonso, D., et al. (2019). A simple method for spray-on gene editing in planta. *bioRxiv [Prprint]*. doi: 10.1101/805036

Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D. L., Wei, P., et al. (2013). Efficient genome editing in plants using a CRISPR/Cas system. *Cell Res.* 23, 1229–1232. doi: 10.1038/cr.2013.114

Gao, D., He, B., Zhou, Y., and Sun, L. (2011). Genetic and molecular analysis of a purple sheath somaclonal mutant in japonica rice. *Plant Cell Rep.* 30, 901–911. doi: 10.1007/s00299-011-1004-3

Gao, X., Chen, J., Dai, X., Zhang, D., and Zhao, Y. (2016). An effective strategy for reliably isolating heritable and Cas9-free arabidopsis mutants generated by CRISPR/Cas9-mediated genome editing. *Plant Physiol.* 171, 1794–1800. doi: 10.1104/pp.16.00663

Gao, Y., Zhang, Y., Zhang, D., Dai, X., Estelle, M., and Zhao, Y. (2015). Auxin binding protein 1 (ABP1) is not required for either auxin signaling or arabidopsis development. *Proc. Natl. Acad. Sci. U.S.A.* 112, 2275–2280. doi: 10.1073/pnas.1500365112

Gould, K. S., McKelvie, J., and Markham, K. R. (2002). Do anthocyanins function as antioxidants in leaves? Imaging of H2O2 in red and green leaves after mechanical injury. *Plant Cell Environ.* 25, 1261–1269. doi: 10.1046/j.1365-3040.2002.00905.x

Guo, J., Han, W., and Wang, M. (2008). Ultraviolet and environmental stresses involved in the induction and regulation of anthocyanin biosynthesis: a review. *Afr. J. Biotechnol.* 7, 4966-4972. Available online at: https://www.ajol.info/index.php/ajb/article/view/59709

Han, Y. J., Kim, Y. M., Lee, J. Y., Kim, S. J., Cho, K. C., Chandrasekhar, T., et al. (2009). Production of purple-colored creeping bentgrass using maize transcription factor genes Pl and Lc through Agrobacterium-mediated transformation. *Plant Cell Rep.* 28, 397–406. doi: 10.1007/s00299-008-0648-0

He, Y., Wang, R., Dai, X., and Zhao, Y. (2017a). On Improving CRISPR for editing plant genes: ribozyme-Mediated Guide RNA production and fluorescence-based technology for isolating transgene-free mutants generated by CRISPR. *Prog. Mol. Biol. Transl. Sci.* 149, 151–166. doi: 10.1016/bs.pmbts.2017.03.012

He, Y., Zhang, T., Sun, H., Zhan, H., and Zhao, Y. (2020). A reporter for non-invasively monitoring gene expression and plant transformation. *Hortic Res.* 7:152. doi: 10.1038/s41438-020-00390-1

He, Y., Zhang, T., Yang, N., Xu, M., Yan, L., Wang, L., et al. (2017b). Self-cleaving ribozymes enable the production of guide RNAs from unlimited choices of promoters for CRISPR/Cas9 mediated genome editing. *J. Genet. Genomics* 44, 469–472. doi: 10.1016/j.jgg.2017.08.003

He, Y., Zhu, M., Wang, L., Wu, J., Wang, Q., Wang, R., et al. (2018). Programmed self-elimination of the CRISPR/Cas9 construct greatly accelerates the isolation of edited and transgene-free rice plants. *Mol. Plant* 11, 1210–1213. doi: 10.1016/j.molp.2018.05.005

He, Y., Zhu, M., Wang, L., Wu, J., Wang, Q., Wang, R., et al. (2019). Improvements of TKC technology accelerate isolation of transgene-free CRISPR/Cas9-edited rice plants. *Rice Sci.* 26, 109–117. doi: 10.1016/j.rsci.2018.11.001

Hiei, Y., Ohta, S., Komari, T., and Kumashiro, T. (1994). Efficient transformation of rice (*Oryza sativa* L.) mediated by Agrobacterium and sequence analysis of the boundaries of the T-DNA. *Plant J.* 6, 271–282. doi: 10.1046/j.1365-313X.1994.6020271.x

Jaakola, L. (2013). New insights into the regulation of anthocyanin biosynthesis in fruits. *Trends Plant Sci.* 18, 477–483. doi: 10.1016/j.tplants.2013.06.003

Landry, M. P., and Mitter, N. (2019). How nanocarriers delivering cargos in plants can change the GMO landscape. *Nat. Nanotechnol.* 14, 512–514. doi: 10.1038/s41565-019-0463-5

Li, J., Qin, R., Zhang, Y., Xu, S., Liu, X., Yang, J., et al. (2020). Optimizing plant adenine base editor systems by modifying the transgene selection system. *Plant Biotechnol. J.* 18, 1495–1497. doi: 10.1111/pbi.13304

Li, J. F., Norville, J. E., Aach, J., McCormack, M., Zhang, D., Bush, J., et al. (2013a). Multiplex and homologous recombination-mediated genome editing in Arabidopsis and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat Biotechnol.* 31, 688–691. doi: 10.1038/nbt.2654

Li, P., Wang, Y., Qian, Q., Fu, Z., Wang, M., Zeng, D., et al. (2007). LAZY1 controls rice shoot gravitropism through regulating polar auxin transport. *Cell Res.* 17, 402–410. doi: 10.1038/cr.2007.38

Li, Y., Lin, Y., and Chen, H. (2017). Function of transcription factor OsC1 and OsPAC1 regulating biosynthesis of rice anthocyanin. *J. Huazhong Agric. Univ.* 1, 1–9. doi: 10.13300/j.cnki.hnlkxb.2017.01.001

Li, Y., Zhang, T., Shen, Z. W., Xu, Y., and Li, J. Y. (2013b). Overexpression of maize anthocyanin regulatory gene Lc affects rice fertility. *Biotechnol Lett.* 35, 115–119. doi: 10.1007/s10529-012-1046-9

Liu, J., Osbourn, A., and Ma, P. (2015). MYB transcription factors as regulators of phenylpropanoid metabolism in plants. *Mol. Plant* 8, 689–708. doi: 10.1016/j.molp.2015.03.012

Lu, H. P., Liu, S. M., Xu, S. L., Chen, W. Y., Zhou, X., Tan, Y. Y., et al. (2017). CRISPR-S: an active interference element for a rapid and inexpensive selection of genome-edited, transgene-free rice plants. *Plant Biotechnol. J.* 15, 1371–1373. doi: 10.1111/pbi.12788

Mao, Y., Zhang, H., Xu, N., Zhang, B., Gou, F., and Zhu, J. K. (2013). Application of the CRISPR-Cas system for efficient genome engineering in plants. *Mol. Plant* 6, 2008–2011. doi: 10.1093/mp/sst121

Miao, J., Guo, D., Zhang, J., Huang, Q., Qin, G., Zhang, X., et al. (2013). Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Res.* 23, 1233–1236. doi: 10.1038/cr.2013.123

Miki, D., Zhang, W., Zeng, W., Feng, Z., and Zhu, J. K. (2018). CRISPR/Cas9-mediated gene targeting in Arabidopsis using sequential transformation. *Nat. Commun.* 9:1967. doi: 10.1038/s41467-018-04416-0

Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J. D., and Kamoun, S. (2013). Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* 31, 691–693. doi: 10.1038/nbt.2655

Ouyang, L., Ma, M., and Li, L. (2020). An efficient transgene-free DNA-editing system for Arabidopsis using a fluorescent marker. *Biotechnol. Lett.* 42, 313–318. doi: 10.1007/s10529-019-02778-z

Paz-Ares, J., Wienand, U., Peterson, P. A., and Saedler, H. (1986). Molecular cloning of the c locus of *Zea mays*: a locus regulating the anthocyanin pathway. *Embo J.* 5, 829–833. doi: 10.1002/j.1460-2075.1986.tb04291.x

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., et al. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 686–688. doi: 10.1038/nbt.2650

Shin, Y. M., Park, H. J., Yim, S. D., Baek, N. I., Lee, C. H., An, G., et al. (2006). Transgenic rice lines expressing maize C1 and R-S regulatory genes produce various flavonoids in the endosperm. *Plant Biotechnol. J.* 4, 303–315. doi: 10.1111/j.1467-7652.2006.00182.x

Svitashev, S., Schwartz, C., Lenderts, B., Young, J. K., and Mark Cigan, A. (2016). Genome editing in maize directed by CRISPR-Cas9 ribonucleoprotein complexes. *Nat. Commun.* 7:13274. doi: 10.1038/ncomms13274

Tanaka, Y., Sasaki, N., and Ohmiya, A. (2008). Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant J.* 54, 733–749. doi: 10.1111/j.1365-313X.2008.03447.x

Wang, J., and Chen, H. (2020). A novel CRISPR/Cas9 system for efficiently generating Cas9-free multiplex mutants in Arabidopsis. *aBIOTECH.* 1, 6–14. doi: 10.1007/s42994-019-00011-z

Wang, Z. P., Xing, H. L., Dong, L., Zhang, H. Y., Han, C. Y., Wang, X. C., et al. (2015). Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. *Genome Biol.* 16:144. doi: 10.1186/s13059-015-0715-0

Wu, T. M., Huang, J. Z., Oung, H. M., Hsu, Y. T., Tsai, Y. C., and Hong, C. Y. (2019). H(2)O(2)-based method for rapid detection of transgene-free rice plants from segregating CRISPR/Cas9 genome-edited progenies. *Int. J. Mol. Sci.* 20:3885. doi: 10.3390/ijms20163885

Xiao, B. Z., Chen, X., Xiang, C. B., Tang, N., Zhang, Q. F., and Xiong, L. Z. (2009). Evaluation of seven function-known candidate genes for their effects on improving drought resistance of transgenic rice under field conditions. *Mol. Plant.* 2, 73–83. doi: 10.1093/mp/ssn068

Xie, K., and Yang, Y. (2013). RNA-guided genome editing in plants using a CRISPR-Cas system. *Mol. Plant* 6, 1975–1983. doi: 10.1093/mp/sst119

Xie, X., Ma, X., Zhu, Q., Zeng, D., Li, G., and Liu, Y. G. (2017). CRISPR-GE: a convenient software toolkit for CRISPR-based genome editing. *Mol. Plant* 10, 1246–1249. doi: 10.1016/j.molp.2017.06.004

Xu, W., Dubos, C., and Lepiniec, L. (2015). Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. *Trends Plant Sci.* 20, 176–185. doi: 10.1016/j.tplants.2014.12.001

Yoshida, A., Suzaki, T., Tanaka, W., and Hirano, H. Y. (2009). The homeotic gene long sterile lemma (G1) specifies sterile lemma identity in the rice spikelet. *Proc. Natl. Acad. Sci. U.S.A.* 106, 20103–20108. doi: 10.1073/pnas.0907896106

Yu, H., and Zhao, Y. (2019). Fluorescence marker-assisted isolation of Cas9-free and CRISPR-edited Arabidopsis plants. *Methods Mol. Biol.* 1917, 147–154. doi: 10.1007/978-1-4939-8991-1_11

Zhang, R., Liu, J., Chai, Z., Chen, S., Bai, Y., Zong, Y., et al. (2019). Generation of herbicide tolerance traits and a new selectable marker in wheat using base editing. *Nat. Plants* 5, 480–485. doi: 10.1038/s41477-019-0405-0

Zhang, Y., Butelli, E., and Martin, C. (2014). Engineering anthocyanin biosynthesis in plants. *Curr. Opin. Plant Biol.* 19, 81–90. doi: 10.1016/j.pbi.2014.05.011

Zhang, Y., Liang, Z., Zong, Y., Wang, Y., Liu, J., Chen, K., et al. (2016). Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA. *Nat. Commun.* 7:12617. doi: 10.1038/ncomms12617

Zhao, S., Wang, C., Ma, J., Wang, S., Tian, P., Wang, J., et al. (2016). Map-based cloning and functional analysis of the chromogen gene C in rice (*Oryza sativa* L.). *J. Plant Biol.* 59, 496–505. doi: 10.1007/s12374-016-0227-9

Zheng, J., Wu, H., Zhu, H., Huang, C., Liu, C., Chang, Y., et al. (2019). Determining factors, regulation system, and domestication of anthocyanin biosynthesis in rice leaves. *New Phytol.* 223, 705–721. doi: 10.1111/nph.15807

Zhu, Q., Yu, S., Zeng, D., Liu, H., Wang, H., Yang, Z., et al. (2017). Development of "purple endosperm rice" by engineering anthocyanin biosynthesis in the endosperm with a high-efficiency transgene stacking system. *Mol. Plant* 10, 918–929. doi: 10.1016/j.molp.2017.05.008

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a past co-authorship with several of the authors YH, MX, and YZ.

Copyright © 2020 He, Zhu, Wu, Ouyang, Wang, Sun, Yan, Wang, Xu, Zhan and Zhao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genome Editing

REVIEW
published: 15 December 2020
doi: 10.3389/fgeed.2020.612137

Precision Genome Engineering for the Breeding of Tomatoes: Recent Progress and Future Perspectives

Tien Van Vu [1,2], Swati Das [1], Mil Thi Tran [1,3], Jong Chan Hong [1] and Jae-Yean Kim [1,4]**

[1] Division of Applied Life Science (BK21 Four Program), Plant Molecular Biology and Biotechnology Research Center, Gyeongsang National University, Jinju, South Korea, [2] National Key Laboratory for Plant Cell Biotechnology, Agricultural Genetics Institute, Hanoi, Vietnam, [3] Crop Science and Rural Development Division, College of Agriculture, Bac Lieu University, Bac Lieu, Vietnam, [4] Division of Life Science, Gyeongsang National University, Jinju, South Korea

Currently, poor biodiversity has raised challenges in the breeding and cultivation of tomatoes, which originated from the Andean region of Central America, under global climate change. Meanwhile, the wild relatives of cultivated tomatoes possess a rich source of genetic diversity but have not been extensively used for the genetic improvement of cultivated tomatoes due to the possible linkage drag of unwanted traits from their genetic backgrounds. With the advent of new plant breeding techniques (NPBTs), especially CRISPR/Cas-based genome engineering tools, the high-precision molecular breeding of tomato has become possible. Further, accelerated introgression or *de novo* domestication of novel and elite traits from/to the wild tomato relatives to/from the cultivated tomatoes, respectively, has emerged and has been enhanced with high-precision tools. In this review, we summarize recent progress in tomato precision genome editing and its applications for breeding, with a special focus on CRISPR/Cas-based approaches. Future insights and precision tomato breeding scenarios in the CRISPR/Cas era are also discussed.

Keywords: CRISPR/Cas, gene editing, precision genome engineering, tomato breeding, precision breeding, new plant breeding techniques

OPEN ACCESS

Edited by:
Seiichi Toki,
National Institute of Agrobiological Sciences, Japan

Reviewed by:
Keishi Osakabe,
Tokushima University, Japan
Satoko Nonaka,
University of Tsukuba, Japan

****Correspondence:***
Jae-Yean Kim
kimjy@gnu.ac.kr
orcid.org/0000-0002-1180-6232
Tien Van Vu
tienvu.agi@gmail.com
orcid.org/0000-0002-6369-7664

Specialty section:
This article was submitted to Genome Editing in Plants, a section of the journal Frontiers in Genome Editing

Received: *30 September 2020*
Accepted: *24 November 2020*
Published: *15 December 2020*

Citation:
Vu TV, Das S, Tran MT, Hong JC and Kim J-Y (2020) Precision Genome Engineering for the Breeding of Tomatoes: Recent Progress and Future Perspectives. Front. Genome Ed. 2:612137. doi: 10.3389/fgeed.2020.612137

INTRODUCTION

The domestication of wild plants, in which plant parts or seeds with desirable characteristic(s) are kept for the next cropping seasons, is the first step of plant breeding (Lin et al., 2014). Post-domestication, the selection of more desirable traits from the domesticated plants can generate novel varieties with added value. Traditionally, plant breeding approaches have been based on the selection of visibly desirable traits from the cultivated crops. This practice was subsequently extended to the selection of offspring from two distinct parental plants, in the so-called cross-breeding or hybrid crossing technique, after Mendel discovered phenotype-associated alleles and genetic inheritance rules during cross-pollination of pea plants. The process is time-consuming and laborious (Bai and Lindhout, 2007). In the modern era, the selection of desirable traits that usually links to one or several quantitative trait loci (QTL) has been assisted by molecular markers, thereby shortening cross-breeding time and labor (Collard and Mackill, 2008; Foolad and Panthee, 2012). One of the major limitations to the traditional crossbreeding technique is linkage drag, which can introduce undesirable traits from a parental donor in addition to the desirable ones. Genetic engineering approaches, such as transgenesis, have efficiently helped overcome this limitation by

introducing only the genes/alleles of interest into an elite plant. However, due to the need to introduce a selection marker, which is usually isolated from non-plant sources, or the *de novo* integration of single or multiple copies of foreign DNA into a targeted plant, the products of the process have been tightly regulated and require many lengthy and costly trials and biosafety assessments before the strain can be released to the environment (Bai and Lindhout, 2007).

Another major technology for crop breeding is the generation of random mutations in a plant by chemicals or physical agents, such as gamma rays. The chemicals or radiation randomly induces large amounts of DNA damage in the genome of a plant, such as nucleotide chemical modifications or double-stranded breaks (DSBs), thereby generating many mutant strains. Extensive screening and selection of the mutants are required to obtain a plant with traits of interest. However, although time-consuming and repetitive, back crossing is often needed to remove non-desirable mutations, and many unexpected modifications may also be fixed in the genome of the mutant plants (Shelake et al., 2019). Nevertheless, plants generated by random mutation approaches have been as accepted as those from conventional breeding approaches. Recently, new plant breeding technologies (NPBTs), especially clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein (Cas)-based approaches, have been emerging as the superior precision plant breeding technologies for crop genetic improvement and bringing hope to our future agriculture (Belhaj et al., 2013; Chen et al., 2019). The power of CRISPR/Cas systems is the ability to specifically introduce theoretically any genetic modification of interest to any genomic site of a plant without any linkage drag phenomenon. We can now edit the genomes of crops by native CRISPR/Cas complexes, including the induction of targeted single base transition/transversion by a range of base editors, customization of DNA changes by prime editors, or precise replacement of a single base to several kilobases by homologous recombination (HR)-based knock-in (HKI) with the assistance of the CRISPR/Cas system (Cermak et al., 2015; Rees and Liu, 2018; Chen et al., 2019; Lin et al., 2020; Vu et al., 2020). All of these approaches are site-specific and can be controlled to be free of off-target or transgene effects. At the targeted sites, the levels of precise sequence modifications are determined by the DNA repair pathway that is directed to repair DNA damage induced by the CRISPR/Cas complex.

The above mentioned NPBTs have been successfully adapted for tomato genome editing and the subsequent precision breeding of new varieties without linkage drag. Appropriate applications of the NPBTs also helped accelerate the introgression of novel traits from wild relatives of tomato into their elite cultivars and made it feasible with *de novo* domestication of wild tomato (Zsogon et al., 2017, 2018; Li et al., 2018c). In this review, we summarize recent progress in precision breeding of tomato using CRISPR/Cas-based approaches and further discuss the future perspectives within this field.

Abbreviations: BE, Base editing; cNHEJ, Canonical nonhomologous end joining; Cas, CRISPR-associated proteins; CRISPR, Clustered regularly interspaced short palindromic repeats; DSBs, Double-stranded breaks; HR, Homologous recombination; HKI, HR-based knock-in; NPBTs, New plant breeding techniques; NHEJ, Non-homologous end joining.

CURRENT STATUS OF TOMATO BREEDING

Conventional Approaches

The domestication of the cultivated tomato (*Solanum lycopersicum* var. *lycopersicum*) is believed to have started in the Andean region of Central America and has undergone two intermediate stages, represented first by *S. pimpinellifolium* and then by *S. lycopersicum* var. *cerasiforme* as the direct ancestor. During domestication, the evolution/selection force was fruit size (Lin et al., 2014). The other characteristics that can be used to distinguish domesticated tomato and its wild relatives are growth parameters and other fruit traits (Bai and Lindhout, 2007). The modern cultivated tomato was thought to have initially spread to the Old World from Mexico to Europe and later from Europe to the rest of the world (Jenkins, 1948). Following the spreading and selection of tomato varieties for adaptation to each specific geographical area, the genetic diversity of the domesticated tomato has been substantially reduced thanks to genetic drift (Bai and Lindhout, 2007; Lin et al., 2014).

The major conventional approaches for tomato breeding include pedigree, hybrid, and backcross breedings that focus on combinations of various traits for different consumption and market requirements (Bai and Lindhout, 2007). Tomato breeders can be individuals such as farmers or institutions (public and private sectors) and may have different goals for breeding programs. Due to the reduced genetic diversity among inbred populations of tomato resulting from the long period of selective domestication (Ranc et al., 2008), cross-hybridization among the populations is the simplest and fastest way to obtain genetic variations and subsequent selection of new varieties exhibiting novel traits. The pedigree method keeps performance records of all the progenies in many generations of a hybrid from genetically distant parents, thereby supporting the selection of varieties with new traits of interest. The new traits can arise only from the gene pools of the parental populations. To obtain novel traits, such as biotic or abiotic stress tolerance, from wild relatives, tomato breeders use a backcross breeding method to introgress new alleles into cultivated lines and recurrently backcross progenies with the parental cultivated lines to recover their genetic backgrounds. These conventional breeding approaches require extensive observation and selection of the best progenies in many generations and are therefore time-consuming and laborious.

In the genomics era, the selection of specific genotypes can be assisted by molecular markers through so-called marker-assisted selection (MAS) (Collard and Mackill, 2008). Usually, DNA markers that are tightly linked to the QTLs of interest are used to track the presence of QTLs in hybrid offspring by PCR or sequencing. With the ability to sequence the whole genome of a plant at minimal cost, plant breeding by the conventional method has become much more efficient.

New Plant Breeding Approaches

NPBTs represent the newly emerging molecular techniques applied to plant breeding in the genomics and genome editing era, including the CRISPR/Cas nucleases. The NPBTs emphasize engineering plant genomes with a high degree of precision. In particular, with the revolutionary advent and applications of CRISPR/Cas systems for plant genome engineering, plant breeding at the molecular level has become more efficient and precise (Chen et al., 2019). With CRISPR/Cas complexes, many options are available for specifically modifying gene sequences of interest from a single base by base editors and prime editors to several kilobases with HKI (Van Vu et al., 2019).

CRISPR/Cas-Based Genome Editing

CRISPR/Cas-based DNA interference is a phenomenon of prokaryote defense against infectious phages (Barrangou et al., 2007). In general, a single-unit Cas nuclease, such as SpCas9, is activated by complexing with a single CRISPR guide RNA (sgRNA), and the Cas-sgRNA complex "scans" for a dsDNA target that contains a complementary protospacer sequence. An NGG protospacer adjacent motif (PAM, N=A, T, G or C), binds to it and then cleaves both the strands (**Figure 1A**) (Jinek et al., 2012). Due to the DSB-forming nature of the CRISPR/Cas nucleases, they have been used for specifically inducing targeted mutations within a genome of interest (Belhaj et al., 2013; Li et al., 2013; Nekrasov et al., 2013; Shan et al., 2013). The DNA damage triggers the repair system to maintain the integrity of the genomes, thereby avoiding the fatal effects that a single unrepaired DSB can induce. The major pathways involved in DSB repair are the non-homologous end-joining (NHEJ) and cell cycle-dependent HR pathways (**Figure 1A**) (Puchta, 2005; Lieber, 2010; Chapman et al., 2012). The NHEJ mechanism facilitates the repair of the two DSB terminal ends by direct ligation with the activity of DNA ligase IV. Under unfavorable conditions, NHEJ may be erroneous and hence result in small DNA mutations (i.e., deletions or insertions). In plant somatic cells, the HR pathway repairs the DSBs by recombining the sequences flanking the broken ends with homologous sequences from DNA donors. HR can be divided into at least two major subpathways: single-strand annealing (SSA) and synthesis-dependent strand annealing (SDSA) (Puchta, 2005; Van Vu et al., 2019).

Recent studies have revealed multiple CRISPR/Cas systems that can be used to edit RNA (Cox et al., 2017) or stimulate nucleobase damage by deaminases to induce single base transitions (Komor et al., 2016; Nishida et al., 2016; Gaudelli et al., 2017) or transversions (Kurt et al., 2020; Zhao et al., 2020) (**Figure 1B**). Another interesting approach for genome editing without inducing DSBs is a prime editor that uses a reverse transcriptase to copy genetic information from an extended sgRNA (**Figure 1C**). The priming extended gRNA (pegRNA) primes the Cas9-fused RT by binding to the sequence upstream of the nicked site on the untargeted sequence, thereby triggering reverse transcription from the 3′ OH of the nicked end using the pegRNA as a template (Anzalone et al., 2019). The prime editor system was successfully applied in rice and wheat (Lin et al., 2020), but it seems that the performance of the present prime editor in plants was limited. Thus, there is a research opportunity for improvement for further applications in plants (Butt et al., 2020; Hua et al., 2020).

Tomato Precision Genome Engineering

CRISPR/Cas-Based Targeted Mutagenesis

The CRISPR/Cas revolution has paved the way for powerful precision plant breeding using molecular tools. Starting in 2013, the early publications regarding CRISPR/Cas-based targeted mutagenesis in tomato focused on the feasibility of efficient use of the tool and its applicability for studying gene function with knock-out approaches (**Figure 2**).

Engineering for Growth Habits

The first CRISPR/Cas9-based tomato genome editing data were published in 2014 by Brooks and coworkers and showed highly efficient CRISPR/Cas9-based targeted mutations of four different loci (**Figure 2**; **Table 1**). Wiry leaf phenotypes (recessive) were revealed in the first generation of transformants of which both the alleles of the tomato ARGONAUTE7 (SlAGO7) gene were mutated (Brooks et al., 2014). The genotype and phenotype were also shown to be inherited in the next generations, confirming the huge potential of CRISPR/Cas technology in tomato genetic improvement and breeding.

The regulation of plant growth and inflorescence development for higher yield and better fruit production has been a traditional priority in tomato domestication and breeding. In 2015, CRISPR/Cas9 was used in a functional study of novel genes involved in the Clavata (CLV)-Wushel (WUS) circuit in the regulation of shoot meristem size. Targeted mutagenesis of tomato CLV homologs and an arabinosyltransferase generated mutant plants with phenotypes resembling that of natural mutants (**Table 1**). The work advanced our knowledge about the regulation of the CLV pathway, which may be very helpful in the customization of mutant alleles for tomato breeding (Xu et al., 2015). Efficiently targeted knock-out of SlPDS produced biallelic KO allele-carrying strains with an albino phenotype as a visible marker but also led to the suppression of their growth (Pan et al., 2016). Genes involved in inflorescence growth and maturation were also extensively studied using CRISPR/Cas9-based targeted mutagenesis (Xu et al., 2016; Roldan et al., 2017; Soyk et al., 2017). Inflorescence maturation was positively linked to the tomato BLADE-ON-PETIOLE transcriptional cofactors (SlBOPs). SlBOP knock-out mutants showed flowering defects as they produced inflorescences with only a single flower (Xu et al., 2016). Further investigations into the inflorescence structure, growth, and maturation led by Soyk et al. shed light on the roles of the other MADS-box transcription factors in these processes. The J2, EJ2, and LIN genes play roles in controlling inflorescence branching and hence flower numbers in a quantitative manner (Roldan et al., 2017; Soyk et al., 2017). Dosing the mutations of the genes by combining their mutated alleles in various groupings may help to design desirable inflorescences for production goals (Soyk et al., 2017). During fruit setting, the fruit shape is determined by the activity of several proteins, including OVATE and SUPPRESSOR OF OVATE1 (SOV1), members of the OVATE FAMILY PROTEIN (OFP) family. Mutated alleles of OVATE and SOV1 led to the production of elongated fruits (oval or pear

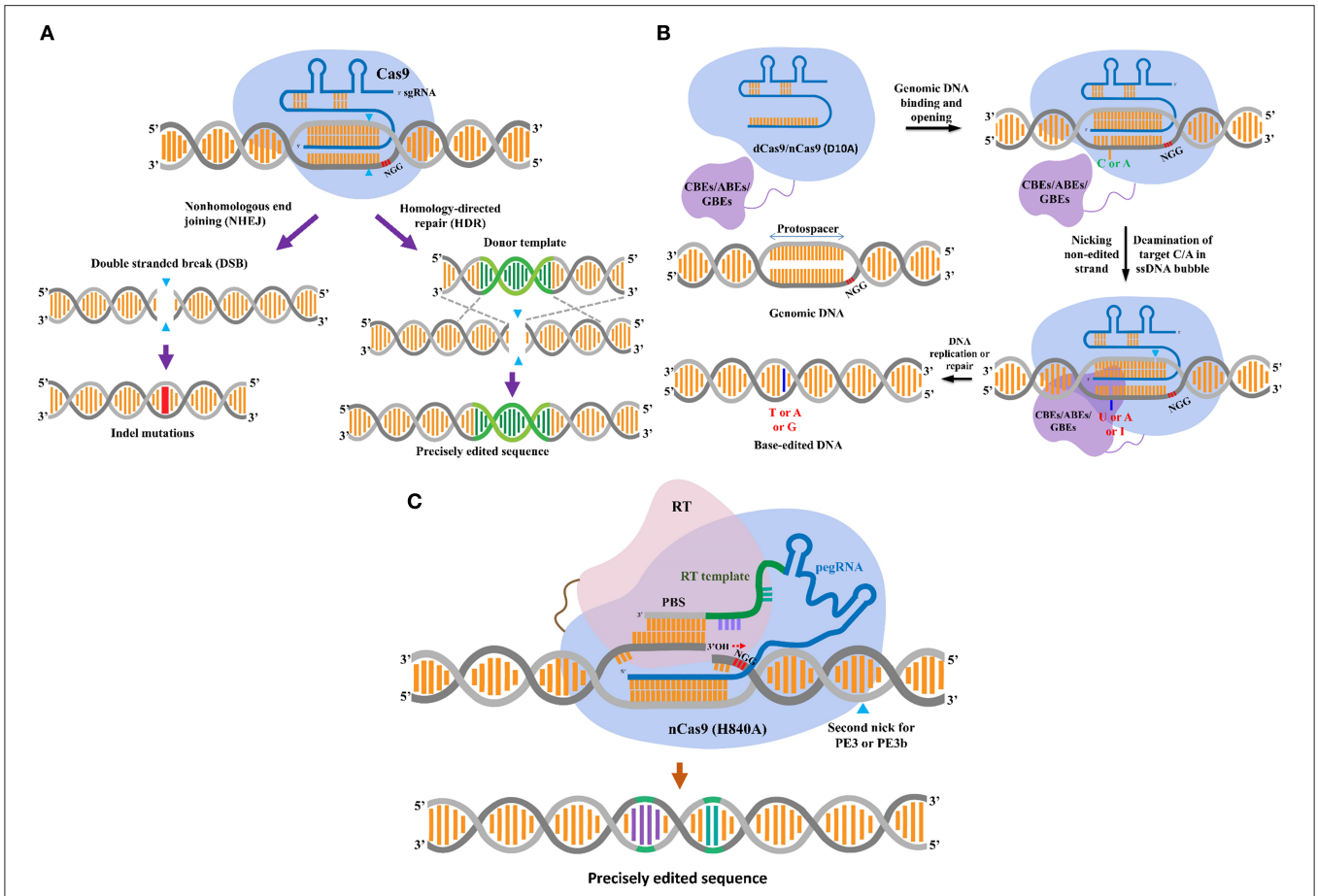


FIGURE 1 | CRISPR/Cas-based plant gene-editing tools. **(A)** CRISPR/Cas9-mediated double-stranded break (DSB) repair mechanism. The CRISPR-associated enzyme Cas9-sgRNA complex breaks down the target DNA to create a DSB, that leads to gene editing via two major DSB repair pathways, i.e., non-homologous end joining (NHEJ) and homology-directed repair (HDR) pathway. The blue triangles indicate locations of the cut site. The NHEJ pathway usually results in small insertions and deletions, as indicated by the red line. The HDR pathway uses a DNA donor template with two homologous ends to the DSB terminals to precisely introduce the edited sequence (green helix) into the genomic site. **(B)** DNA base editors. From left to right and top to down: base editors are created by fusing a dead/nickase Cas9 (d/nCas9) (light blue) to a nucleobase deaminase (in case of cytosine base editors (CBEs) and adenine base editors (ABEs), painted in purple), and a uracil-DNA glycosylase (Ung) [for glycosylase base editors (GBEs)]. The d/nCas9 activated by complexing with a guide RNA (gRNA, in blue color) "scans" for and binds to an NGG PAM and its upstream adjacent sequence that is complementary to the spacer of gRNA. The d/nCas9-gRNA-DNA bound site forms an R-loop that exposes the non-complementary strand as a disordered, unprotected ssDNA. Subsequently, the fused ssDNA-specific nucleobase deaminase works on the exposed strand and deaminates its favorable substrates (C for CBEs, A for ABEs, and C/A for GBEs). In the case of nCas9, a nick would be introduced to the non-edited strand (the blue triangle) for enhancing the editing efficiency. Ultimately, CBEs convert C to T via a U intermediate; and ABEs deaminate A to inosine (I) that is recognized as a G and then fixed as a G after DNA repair or replication (highlighted as the dark blue stick). GBEs consist of a Cas9 nickase, a cytidine/adenine deaminase, and a uracil-DNA glycosylase (Ung). Ung excises the deaminated product (U/I) formed by the deaminase, creating an apurinic/apyrimidinic (AP) site that initiates the process of repairing DNA. In E. coli, the activation-induced cytidine deaminase (AID) was used to construct AID-nCas9-Ung for C to A conversion. In mammalian cells, rat APOBEC1 substituted AID (APOBEC-nCas9-Ung) for C-to-G conversions. **(C)** The prime editor. The prime editing system consists of a CRISPR/ Cas complex by fusion of reverse transcriptase (RT, pink) domain to a C-terminal of Cas 9 (H840A) nickase, and a prime editing gRNA [pegRNA, composed of a gRNA with a scaffold (blue), an RT template (green) and a prime binding sequence (PBS, gray)] with a 3′ extended PBS that binds to the 3′ nicked site of the target DNA having the PAM site. The nicked terminal's free 3′-OH provides a prime for the RT to copy the genetic information from the RT template (the orientation is indicated by the discontinuous red arrow). The copied information is then fixed into the genomic site via a complex process that includes flapping of the competitively original strand and subsequent integration of the edited strand via DNA repair that may not be fully understood in plants.

shape), and the shape could be rescued by knocking out TRM5, a member of the TONNEAU1 Recruiting Motif (TRM) family, which was shown to work downstream and in close contact with *ovate* and *sov1* (Wu et al., 2018). The floral organ and locule number were also enhanced in the KO mutant of EXCESSIVE NUMBER OF FLORAL ORGANS (SlENO). CRISPR/Cas9-based targeted mutagenesis of the SlENO locus was extremely efficient (Yuste-Lisbona et al., 2020). Plant vegetative growth could be regulated by gibberrellins (GAs) via their interaction with DELLA protein, a negative regulator of plant growth, thereby subjecting it to degradation. A mutation in the interacting site of PROCERA, a DELLA protein, blocked GA binding and thus suppressed plant growth, resulting in tomato dwarfism (Tomlinson et al., 2019).

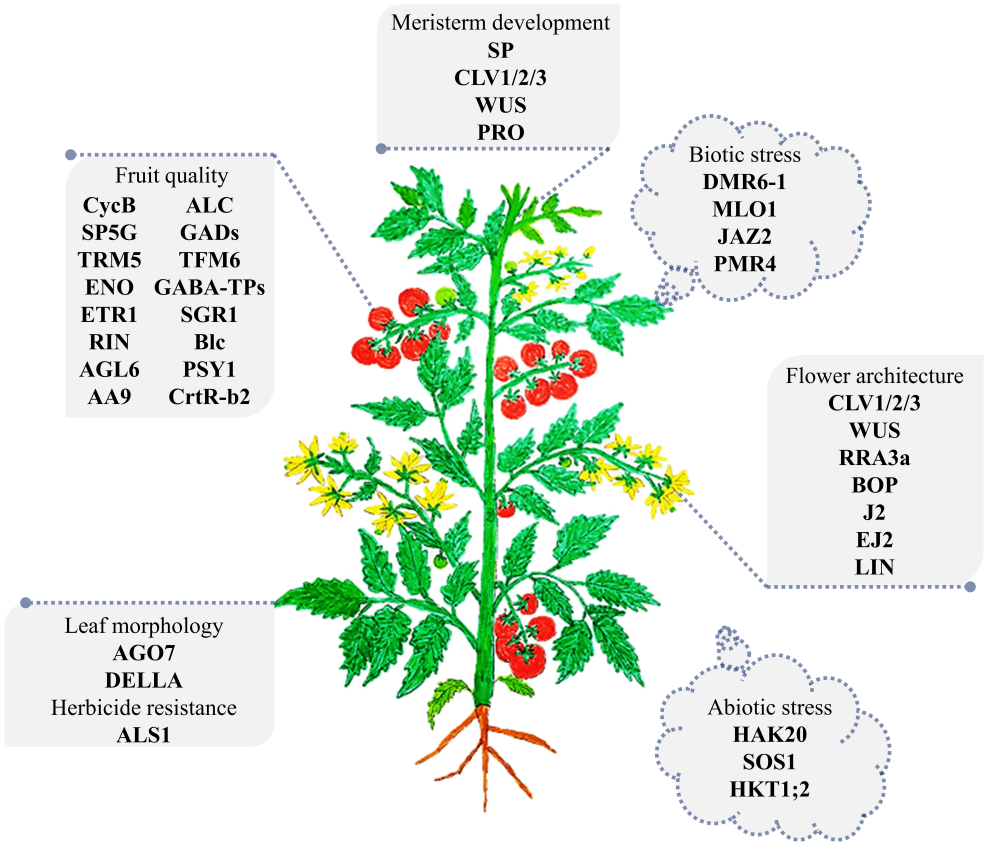


FIGURE 2 | The major genes targeted by CRISPR/Cas-based targeted mutagenesis in tomato. A representative tomato plant at the vegetative growth stage was drawn. The CRISPR/Cas edited genes are illustrated in association with the organs they affect. The abiotic stress gene set is shown surrounding the plant, indicating their global impacts on the plant. The detail of related studies is summarized in **Table 1**. The plant is not drawn to scale.

Fruit Quality Improvement

Tomato fruit quality is increasing of interest in modern markets, especially for fresh uses. However, cultivated tomatoes have lost many fruit qualitative characteristics, such as flavors (Tieman et al., 2017), due to the yield-based domestication and industrialization of tomato production. Studying the gene functions and their mutants in the fruit setting and maturation is important for customizing tomato fruit for higher quality by CRISPR/Cas9 technology (**Figure 2**). Fruit shelf life is an important parameter of fruit quality and storage capability. A MADS-box transcription factor gene (SlRIN, **Table 1**) involved in fruit ripening was destroyed by CRISPR/Cas9, resulting in an incomplete ripening stage and extended shelf life (Ito et al., 2015). However, the *rin* mutants, even in the heterozygous form, may lead to a reduction of lycopene content in tomato fruits (Herner and Sink, 1973). Another well-characterized mutant allele in tomato breeding for long shelf life is *alc* (*alcobaca*), which may introduce fewer side effects (Yu et al., 2017). The seedless tomato is an interesting example with fresh uses as well as uses in processing. The parthenocarpic tomato lines were created by EMS mutagenesis of the tomato AGAMOUS-LIKE 6 (SlAGL6) gene that promoted fruit development without fertilization, especially under heat stress. The *slagl6* allele could also be specifically generated with CRISPR/Cas9 complexes. Interestingly, if fertilization is successful under normal conditions, then seeds are still normally developed, suggesting that the traits could be used in practical breeding (Klap et al., 2017). An alternative approach for breeding of parthenocarpic tomato was through targeted mutation of AUXIN-INDUCED 9 (SlIAA9), a repressor of fruit development without fertilization. Knock-out mutation of SlIAA9 led to the production of parthenocarpic fruits but also abnormal leaf morphology (Ueta et al., 2017) that could ultimately affect yield. Malic acid (malate salt) is an intermediate metabolite of C4 plants, and it plays roles in plant growth, fruit quality, and Al detoxification in roots (Ye et al., 2017). Tomato fruit malate on chromosome6 (TFM6) was shown to be associated with tomato fruit malate content. CRISPR/Cas9-based deletion of the TFM6 gene led to a reduction in fruit malate content. However, a 3-bp deletion in the binding site of a WRKY transcription repressor (SlWRKY42) in the promoter of TFM6 enhanced the accumulation of malate in fruit (Ye et al., 2017).

ɣ-Aminobutyric acid (GABA), a non-protein amino acid, acts as an inhibitory neurotransmitter and hypotensive relief

TABLE 1 | Precision genome editing in tomato using CRISPR/Cas complexes.

Trait category	Target gene	Accession ID	Editing tool	Major mutant phenotype	Repair pathway	Editing efficiency (%)	References
CRISPR/Cas-BASED TARGETED MUTAGENESIS							
Growth habit	SlAGO7 (Argonaute 7)	Solyc01g010970	CRISPR/SpCas9	Typical compound flat leaves become needle like or wiry	cNHEJ	48.0	Brooks et al., 2014
	SlHPAT homolog	Solyc08g041770		Multiple aspects of tomato reproductive development		75.0	
	SlHPAT homolog	Solyc07g021170				100.0	
	SlHPAT homolog	Solyc12g044760				100.0	
	SlCLV3 (Clavata 3)	Solyc11g071380	CRISPR/SpCas9	Branched inflorescences with fasciated flowers	cNHEJ	57.1	Xu et al., 2015
	SlCLV1	Solyc04g081590		Weak branching and fasciated flowers		100.0	
	SlCLV2	Solyc04g056640		Weak branching and fasciated flowers		83.3	
	SlRRA3a (Reduced residual arabinose 3a)	Solyc04g080080		Branched inflorescences with fasciated flowers		66.7	
	SlPDS (phytoene desaturase)	Solyc03g123760	CRISPR/SpCas9	Albino	cNHEJ	71.4–100.0	Pan et al., 2016
	SlPIF4 (Phytochrome interacting factor 4)	Solyc07g043580		No obvious abnormal phenotype		84.0–89.5	
	SlBOPs (Blade-on-petiole)	SlBOP1 (Solyc04g040220), SlBOP2 (Solyc10g079460), SlBOP3 (Solyc10g079750)	CRISPR/SpCas9	Flowering defect	cNHEJ	-	Xu et al., 2016
	SlJ2 (Jointless-2)	Solyc12g038510	CRISPR/SpCas9	Jointless unbranched inflorescences	cNHEJ	-	Roldan et al., 2017; Soyk et al., 2017
	SlEJ2 (Enhancer-of-jointless2)	Solyc03g114840		Exceptionally large sepals and pear-shaped fruits		-	
	SlLIN (Long inflorescence)	Solyc04g005320		Moderately branched inflorescences and increased flower production		-	
	SlTRM5 (TONNEAU1 Recruiting Motif5)	Solyc07g008670	CRISPR/SpCas9	Slightly flatter fruit	cNHEJ	-	Wu et al., 2018
	SlENO (Excessive number of floral organs)	Solyc03g117230	CRISPR/SpCas9	Increased number of floral organs and multilocular fruits	cNHEJ	100.0	Yuste-Lisbona et al., 2020
	SlPRO (Procera)	Solyc11g011260	CRISPR/SpCas9	Dwarf/gibberellin-responsive dominant dwarf DELLA allele	cNHEJ	17.4	Tomlinson et al., 2019
Fruit quality	SlRIN (Ripening inhibitor)	Solyc05g012020	CRISPR/SpCas9	Incomplete-ripening fruits, extended shelf life	cNHEJ	11.8–50.0	Ito et al., 2015
	SlAGL6 (Agamous-like 6)	Solyc01g093960	CRISPR/SpCas9	Seedless	cNHEJ	-	Klap et al., 2017
	SlIAA9 (Auxin-induced 9)	Solyc04g076850	CRISPR/SpCas9	Seedless	cNHEJ	42.9–100.0	Ueta et al., 2017
	SlALC (alcobaca)	FJ404469	CRISPR/SpCas9	Long-shelf life	cNHEJ	72.7	Yu et al., 2017
	SlGAD2 (glutamate decarboxylase 2)	B1Q3F1	CRISPR/SpCas9	Increased GABA accumulation: In T0 fruits: 1.5–10.0 folds	cNHEJ	68.8–78.6	Nonaka et al., 2017

(Continued)

TABLE 1 | Continued

Trait category	Target gene	Accession ID	Editing tool	Major mutant phenotype	Repair pathway	Editing efficiency (%)	References
	SlGAD3 (glutamate decarboxylase 3)	B1Q3F2		Increased GABA accumulation: In T0 fruits: 1.0–5.0 folds In T1 fruits: 7.0–15.0 folds		25.7	
	SlTFM6 (Tomato fruit malate on chromosome6)	Solyc06g072910	CRISPR/SpCas9	Reduced fruit malate content.	cNHEJ	-	Ye et al., 2017
	SlGABA-TP1 (pyruvate-dependent γ-aminobutyric acid transaminase 1)	AY240229	CRISPR/SpCas9	2.85-fold increased GABA accumulation	cNHEJ	50.0–56.8	Li et al., 2018a
	SlGABA-TP2	AY240230		-		0.0	
	SlGABA-TP3	AY240231		3.5-fold increased GABA accumulation (together with SlGABA-TP1 mutation)		46.6	
	SlCAT9 (Cationic amino acid transporter 9)	XM_004248503		No fruit		6.8	
	SlSSADH (Succinate semialdehyde dehydrogenase)	NM_001246912		No fruit		9.1	
	SlSGR1 (Stay green 1)	DQ100158	CRISPR/SpCas9	5.1-fold increase of lycopene content in fruit	cNHEJ	41.7–95.8	Li et al., 2018c
	SlLCY-E (Lycopene ε-cyclase)	EU533951		-		8.3	
	SlBlc (Beta-lycopene cyclase)	XM_010313794		1.83-fold increase of lycopene content in fruit		91.7	
	SlLCY-B1 (Lycopene β-cyclase 1)	EF650013		-		0.0	
	SlLCY-B2 (Lycopene β-cyclase 2)	AF254793		-		4.2	
	SlPSY1 (Phytoene synthase 1)	P08196	CRISPR/SpCas9	Yellow-flesh fruit	cNHEJ	75.0–84.0	D'Ambrosio et al., 2018
	SlCrtR-b2 (Beta-carotene hydroxylase 2)	Q9S6Y0		White-flower		75.0–100.0	
Biotic stress tolerance	SlDMR6-1 (Downy mildew resistance 6-1)	Solyc03g080190	CRISPR/SpCas9	Disease resistance against different pathogens, including *P. syringae*, *P. capsici*, and *Xanthomonas* spp.	cNHEJ	-	Paula de Toledo Thomazella et al., 2016
	SlMLO1 (Mildew Locus O 1)	Solyc04g049090	CRISPR/SpCas9	Powdery mildew disease resistance	cNHEJ	80.0	Nekrasov et al., 2017
	SlJAZ2 (Jasmonate Zim Domain 2)	Solyc12g009220	CRISPR/SpCas9	Resistance against *P. syringae* pv. tomato DC3000	cNHEJ	65.2	Ortigosa et al., 2019
	SlPMR4 (Powdery Mildew Resistance 4)	Solyc07g053980	CRISPR/SpCas9	Reduced susceptibility to Powdery mildew disease	cNHEJ	45.9	Santillan Martinez et al., 2020
Abiotic stress tolerance	SlMAPK3 (Mitogen activated protein kinase 3)	AY261514	CRISPR/SpCas9	Reduced drought tolerance	cNHEJ	41.8	Wang et al., 2017
	SlNPR1 (Nonexpressor of pathogenesis-related gene 1)	KX198701	CRISPR/SpCas9	Reduced drought tolerance	cNHEJ	33.3–46.7	Li et al., 2019
	SlHAK20 (High-affinity K+ 20)	Solyc04g008450	CRISPR/SpCas9	Hypersensitivity to salt stress	cNHEJ	-	Wang et al., 2020a
	SlSOS1 (Salt Overly Sensitive 1)	Solyc11g044540	CRISPR/SpCas9	Increased salt sensitivity	cNHEJ	-	Wang et al., 2020b
	SlCBF1 (C-repeat/dehydration-responsive element binding factor 1)	AAS77820	CRISPR/SpCas9	Reduced chilling tolerance	cNHEJ	25.0–57.5	Li et al., 2018b
	SlBZR1 (brassinazole-resistant 1)	Solyc04g079980	CRISPR/SpCas9	Reduced heat tolerance	cNHEJ	-	Yin et al., 2018
	SlHyPRP1 (Hybrid proline-rich protein 1)	Solyc12g009650	CRISPR/SpCas9	Salinity tolerance	cNHEJ	4.5–20.0	Tran et al., 2020

(Continued)

TABLE 1 | Continued

Trait category	Target gene	Accession ID	Editing tool	Major mutant phenotype	Repair pathway	Editing efficiency (%)	Reference
CRISPR/Cas-BASED PRECISE DNA CHANGES/REPLACEMENTS							
Growth habit	SlDELLA	Solyc11g011260	nSpCas9-PmCDA1	Reduced serrated leaflets	BER/NER	54.5	Shimatani et al., 2017
	SlETR1 (Ethylene receptor 1)	Solyc12g011330	nSpCas9-PmCDA1	Insensitivity to ethylene	BER/NER	70.0	
Growth habit	SlANT1 (Anthocyanin 1)	Solyc10g086260	CRISPR/SpCas9	Global anthocyanin over-accumulation	HDR	11.7	Cermak et al., 2015
Growth habit	SlANT1	Solyc10g086260	CRISPR/LbCas12a	Global anthocyanin over-accumulation	HDR	12.8	Vu et al., 2020
Abiotic stress tolerance	SlHKT1;2 (High-affinity potassium transporter 1;2)	Solyc07g014680		Salinity tolerance		0.7	
Herbicide resistance	SlALS1 (Acetolactate synthase 1)	Solyc03g044330	nSpCas9-PmCDA1	Chlorsulfuron resistance	BER/NER	71.0	Veillet et al., 2019
Herbicide resistance	SlALS1	Solyc03g044330	CRISPR/SpCas9	Chlorsulfuron resistance	HDR	12.7	Danilo et al., 2019

factor that benefits human health (Takayama et al., 2015; Li et al., 2018a). However, even produced at relatively high levels compared to other crops, the GABA content in tomato fruits is still very low. To enhance GABA production in tomato fruits, Nonaka and coworkers used CRISPR/Cas9 to target the autoinhibitory C-terminal coding sequences of two genes, glutamate decarboxylase 2 and 3 (SlGAD2 and SlGAD3, respectively), that are involved in GABA synthesis in fruit stages. The SlGAD3 mutated plants carrying a premature stop codon before the autoinhibitory domain produced GABA in T1 fruits at 7–15-fold higher levels than non-edited plants (**Table 1**) (Nonaka et al., 2017). Lee and coworkers generated hybrid lines of the best event (TG3C37) obtained from the work of Nonaka et al. The heterozygous state of the alleles also accumulated a high level of GABA in fruits while having minimal effects on plant growth and fruit development (Lee et al., 2018). A similar approach was also taken by Li and coworkers, but they targeted five genes (**Table 1**) involved in the GABA conversions of the GABA shunt. The highest accumulation of GABA in red fruits was 3.5-fold higher than that of WT fruits (Li et al., 2018a). However, higher GABA accumulation in leaves (∼6–20 folds) (Li et al., 2018a), and fruits (11–12 folds, transgenic plants over-expressing C-terminal truncated SlGAD3 in fruits) (Takayama et al., 2017) led to reduced growth and prolonged flowering time, and changing fruit contents, respectively. Two possible explanations, which are not mutually exclusive, are currently available. The first hypothesis is that GABA overaccumulation causes a deficit of glutamate, the precursor of GABA. The second explanation is that GABA itself functions in the plant signaling pathway; thus, its overaccumulation is detrimental to plant growth. Therefore, together with GABA overaccumulation, strategies to enhance the glutamate synthesis pathway or sequester GABA into vacuoles might be future options.

Carotenoids produced during tomato fruit growth and maturation contribute to a large portion of fruit pigments and volatiles as the byproducts of their metabolism (metabolites). Lycopene, produced in tomato fruit, is a carotenoid that contributes to the color of the fruit and is hypothesized to possess potential health effects (Story et al., 2010). Therefore, tomato breeding for enhancement of lycopene accumulation in fruit is of interest. In total, five genes involved in lycopene metabolism at the early processes (SlSGR1) and lycopene cyclization stages (LCY-E, Blc, LCY-B1, and LCY-B2) were targeted for the enhancement of its production (**Table 1**). Up to a 5.1-fold increase in lycopene content was recorded after the single mutation of SlSGR1, while additional mutated genes caused a reduction of the lycopene amount compared to that in the SlSGR1 mutant, though the content was still much higher compared to that of non-edited fruits (Li et al., 2018d). Regulation of carotenoid accumulation in tomato fruit has also been used to customize fruit color. Targeting the genes encoding enzymes involved in the carotenoid pathway is the major approach using CRISPR/Cas9 complexes. Targeted knock-out of the phytoene synthase 1 (SlPSY1) led to the abolishment of lycopene production and thus resulted in tomato fruits with yellow flesh (D'Ambrosio et al., 2018).

Biotic Stress Tolerance Enhancement

Environmental conditions are continuously changing; under the pressure of arable land shortages, sustaining food production to feed increasing populations will be a challenge by 2050 (Hickey et al., 2019). The breeding of resilient crops for stress tolerance is a major solution to help meet this challenge. CRISPR/Cas9 has been efficiently used to target genes encoding negative regulators of biotic as well as abiotic stress response pathways (Paula de Toledo Thomazella et al., 2016; Nekrasov et al., 2017; Ortigosa et al., 2019; Santillan Martinez et al., 2020) (**Figure 2** and **Table 1**).

Bacterial speck disease caused by *Pseudomonas syringae* pv. tomato is one of the major threats to tomato production since it can lead to losses in yield and fruit quality (Cai et al., 2011). In an early application of CRISPR/Cas9 targeting a knock-out

of a positive regulator of the disease, mutant alleles of a tomato ortholog of *Arabidopsis* downy mildew resistance 6 (DMR6) were generated. Initial tests of the mutants showed resistance against *P. syringae* pv. tomato DC3000 (*Pto* DC3000), *Phytophthora capsici,* and *Xanthomonas* spp. (Paula de Toledo Thomazella et al., 2016) that may be highly useful sources for tomato breeding. Another interesting approach to prevent *P. syringae* colonization was via regulation of stomata opening/closing. *P. syringae* produces coronatine (COR), a mimic of jasmonic acid-isoleucine (JA-Ile), during infection, and COR subsequently stimulates stomata opening and triggers degradation of a major COR coreceptor, JASMONATE ZIM DOMAIN 2 (JAZ2). CRISPR/Cas9-based truncation of the C-terminal Jas domain of SlJAZ2 generated enhanced *Pto* DC3000-resistant plants without altering the resistance against the pathogenic fungus *Botrytis cinerea* (Ortigosa et al., 2019). The fungal pathogen *Oidium neolycopersici* is the causal agent of powdery mildew disease, which leads to serious yield losses in tomato production and fruit quality reduction (Jones et al., 2001). Some members of the transmembrane protein Mildew Locus O (MLO) family are responsible for susceptibility to *O. neolycopersici* infection. Among the 16 MLOs in tomato, the SlMLO1 was shown to be the major susceptibility gene, and its natural loss-of-function mutants exhibited powdery mildew disease resistance (Zheng et al., 2016). CRISPR/Cas9-based mutant strains carrying homozygous *Slmlo1* alleles that are 48-bp truncated versions of the WT SlMLO1 showed complete resistance to *O. neolycopersici* infection. Interestingly, the *Slmlo1* plants were free of any foreign T-DNA sequence and therefore were indistinguishable from natural *Slmlo1*-mutated plants (Nekrasov et al., 2017). An alternative approach for combatting powdery mildew disease is through the regulation of unicellular hyperresponsiveness (HR) in the penetration site of the fungi, and knocking out tomato Powdery Mildew Resistance 4 (SlPMR4) could make this possible (Santillan Martinez et al., 2020). SlPMR4 encodes an enzyme that catalyzes callose synthesis in response to environmental stress. Overexpression of PMR4 in *Arabidopsis* showed complete resistance to powdery mildew by blocking fungal penetration at the papilla sites (Huckelhoven, 2014). Surprisingly, *pmr4* mutants that do not produce callose at the papillae also resisted *O. neolycopersici* infection, which might have resulted from HR-like cell death at the infection site (Santillan Martinez et al., 2020).

Abiotic Stress Tolerance Engineering

CRISPR/Cas-based genome editing for abiotic stress tolerance in tomato breeding is promising for the creation of resilient cultivars for the sustainable production of tomato fruits. The most important abiotic stresses studied using CRISPR/Cas9 or Cas12a (Cpf1) tools in tomato have been drought (Wang et al., 2017; Li et al., 2019), salinity (Tran et al., 2020; Vu et al., 2020; Wang et al., 2020a,b) and temperature (Li et al., 2018b; Yin et al., 2018) (**Figure 2** and **Table 1**). Two drought stress-responsive genes, mitogen-activated protein kinase 3 (SlMAPK3) (Wang et al., 2017) and non-expressor of pathogenesis-related gene 1 (SlNPR1) (Li et al., 2019), were knocked out by CRISPR/Cas9, but neither of them showed improvement in drought tolerance. The data indicate that SlMAPK3 and SlNPR1 may positively contribute to drought stress responses in tomato. Salinity-tolerant alleles were revealed from functional studies of genes relating to the perception of salt during plant growth. Another protein involved in K^+/Na^+ homeostasis in tomato is SlHAK20, a member of the high-affinity K^+/K^+ uptake/K^+ (HAK/KUP/KT) transporter that was functionalized for salinity responses. The mutated *slhak20* allele contributed to the hypersensitivity to salinity (Wang et al., 2020b). Tomato Salt Overly Sensitive 1 (SlSOS1) is a Na^+/H^+ antiporter that helps to control Na^+ levels in root epidermal cells. Blocking the activity of SlSOS1, therefore, reduced salt tolerance performance (Wang et al., 2020a). A very recent work conducted by our team revealed a strong salt-tolerant allele obtained by CRISPR/Cas9-based precise removal of a proline-rich domain of tomato hybrid proline-rich protein 1 (SlHyPRP1) (Tran et al., 2020). Another important environmental factor for the growth of tomato is temperature. Climate changes accompanying wider temperature changes may affect tomato cropping. Understanding the roles of genes involved in temperature responses is critical for engineering and breeding temperature-tolerant tomatoes. To this end, tomato C-repeat binding factor 1 (SlCBF1), a chilling-related gene, and Brassinazole Resistant 1 (SlBZR1), a heat-responsive factor, were knocked out by CRISPR/Cas9. The data showed that both genes were positively involved in temperature tolerance since the mutant alleles of *slcbf1* and *slbzr1* led to reduced chilling (Li et al., 2018b) and heat (Yin et al., 2018) stress tolerance, respectively. Further works are needed in order to reveal abiotic stress tolerance alleles for tomato breeding, especially those negatively affecting recessive alleles, or genome editing technologies to introduce dominant alleles.

CRISPR/Cas-Based Precise DNA Changes/Replacements

Base Substitutions

The uses of CRISPR/Cas complexes in tomato genome editing have not been limited to targeted mutagenesis but have been extended to precise changes of every base up to long DNA sequences. The direct evolution of nucleotide deaminases for fitting with CRISPR/Cas9-guides for base editing (BE) has been extensively conducted by Liu's group at Harvard University. For an extensive review, please refer to Ree and Liu's review (Rees and Liu, 2018). There are a limited number of published data applying base editors in tomato breeding. The *Petromyzon marinus* cytidine deaminase (PmCDA1) was fused to CRISPR/Cas9 [death Cas9 or nickase Cas9 (D10A)] for BE in tomato (**Figure 1B**). Efficient base substitutions in the Della gene led to a loss-of-function mutation that produced a *procera* phenotype with reduced serrated leaflets. Similar tools were also used for base substitutions in SlETR1, an ethylene receptor, to produce ethylene insensitive strains (Shimatani et al., 2017) (**Table 1**). In another paper, BE-based GE was used to produce herbicide tolerant tomato. Proline-197 in *Arabidopsis* ALS was shown to be the key a for conferring chlorsulfuron resistance when it was changed to another a.a. Thus, the corresponding proline-186 in tomato ALS1 was subjected to changes to other amino acids by nCas9-PmCDA1. Strains with substituted bases (C to T or C to

G) exhibited strong resistance to the treatment of chlorsulfuron (Veillet et al., 2019).

Homologous Recombination (HR)-Based Knock-in

HKI in plants is an all-in-one precision technique for the replacement of SNPs or large DNA sequences (**Figure 1A**). However, due to the low efficiency of natural HR, the number of HKI applications in plants is limited and not currently available for practical use. With the advent of CRISPR/Cas complexes as molecular scissors for generating DSBs at specific genomic sites, HKI frequency has been improved dramatically (Cermak et al., 2015; Dahan-Meir et al., 2018; Merker et al., 2020; Vu et al., 2020). In tomato, HKI was engineered with CRISPR/Cas9 or Cas12a proteins and geminiviral replicons for generating anthocyanin overaccumulating events by inserting the CaMV 35S promoter upstream of SlANT1, an R2R3-MYB transcription factor. The HKI frequency was improved 10–30 times compared to that of the T-DNA cargos and was several orders of magnitude higher than that of spontaneous HKI (Cermak et al., 2015; Vu et al., 2020). A single amino acid substitution (N207D) in the polypeptide sequence of the *Arabidopsis* HIGH-AFFINITY K^+ TRANSPORTER1 (AtHKT1) led to salt tolerance. Using the CRISPR/Cas12a-mediated HKI approach, we have successfully generated a salt-tolerant strain carrying a tomato ortholog of the AtHKT1 N207D, namely, SlHKT1;2 N217D, without using an allele-associated selection marker (Vu et al., 2020). CRISPR/Cas9-mediated HKI was also efficiently applied to generate herbicide-tolerant tomatoes by targeting SlALS1 for P186A modification (Danilo et al., 2019).

FUTURE INSIGHTS INTO TOMATO BREEDING

As discussed in the opening of this writing, the tomato domestication and selective breeding processes led to the reduction of genetic diversity in nowadays-cultivated tomatoes (Ranc et al., 2008; Lin et al., 2014). The cultivated tomato appears with high yield and compact architectures but tends to be more vulnerable to environmental attacks by physical as well as biological agents from the environment. Tomato production has become more difficult, especially in the face of global climate changes. Therefore, in the new scenario of tomato breeding, new margins of traits and techniques have to be fully considered. Alleles that determine abiotic and biotic stress-tolerant traits or fruit flavor have been widely lost in modern tomatoes but widely available in their wild relatives (Bai et al., 2018). Those alleles have been re-introduced into the cultivated tomatoes by conventional breeding without or with MAS. However, the traditional breeding approaches are time-consuming and laborious, especially for pyramiding polygenic traits or multiple monogenic traits of interest. The CRISPR/Cas-based *de novo* domestication and/or accelerated allele introgression could not only help to reduce the time and labor but also allow us to precisely control the genetic modification types (Fernie and Yan, 2019). Further, CRISPR/Cas-based pyramiding of polygenic traits or multiple monogenic traits could be feasible in a timely breeding program.

Another conventional approach in tomato breeding is random mutagenesis using the chemical as well as physical agents to disrupt the tomato genome and select for new alleles and acceptable mutant lines for further breeding purposes. Due to the random and extensive nature of the induced mutations, mutant lines show severely defective traits, and hence, cannot be suitable for crop production. Even if we can obtain a useful allele, its genetic background might have been dramatically changed from its parental origins. Thus, the approach is also time-consuming and laborious for the selection of usable alleles from a huge number of mutations. Again, CRISPR/Cas-based technologies appear to be revolutionized solutions for mining useful alleles for tomato breeding. The accumulated data showed high potential and efficient advances of multiplexed editing that can be used for discovering novel alleles based on the extensively released omics databases (Chen et al., 2019; Pramanik et al., 2020). With the multiplexed editing, engineering an intact metabolic pathway is also possible at high loci-specific precision (Li et al., 2018a,d). It is also wonderful to be able to obtain the homozygous edited alleles at the first generation of genome-edited events by haploid inducer-based genome editing (Kelliher et al., 2019). There are still hurdles in the selection and regeneration of edited events, especially those resulting from the allele-associated marker-free conditions in case of the low-efficiency HKI (Van Vu et al., 2019). Tackling the issue, some *in planta* transformation approaches mediated by *Agrobacterium* (Maher et al., 2020) or nanoparticles (Cunningham et al., 2018) could be applied for the genome editing process. Those approaches may help to significantly reduce tomato breeding time and labor that a small-scale enterprise can afford for contribution to the field.

Accelerated Allele Introgression and *de novo* Domestication

Food production with current technologies is predicted to not meet the demands of a dramatic increase in population by 2050 (Ray et al., 2013). Strategic breeding programs in the field have been initiated to reverse these catastrophic prospects, and integral solutions for sustainable agriculture will be key to overcoming food production barriers. The development of ideal/super crops should be a major goal (Zsogon et al., 2017). Engineering new crops by redomestication or *de novo* domestication from wild relatives or semidomesticated plants would also offer more options to cope with challenges in feeding people by 2050 (Fernie and Yan, 2019; Hickey et al., 2019). The most powerful applications of CRISPR/Cas technology in plant breeding may be the ability to accelerate the introgression of novel alleles into elite cultivars and the *de novo* domestication of wild plants for cultivation (**Figure 3**). The success of these processes is strongly dependent on the precision and efficacy of the CRISPR/Cas-based technologies. *De novo* domestication of wild tomato or orphan crops has been illustrated (Lemmon et al., 2018; Li et al., 2018c; Zsogon et al., 2018), thus paving revolutionary paths toward a new era of tomato breeding.

The lengthy domestication and conventional breeding processes have reduced some important qualitative and stress-tolerance traits (Liu et al., 2020; Wang et al., 2020a,b).

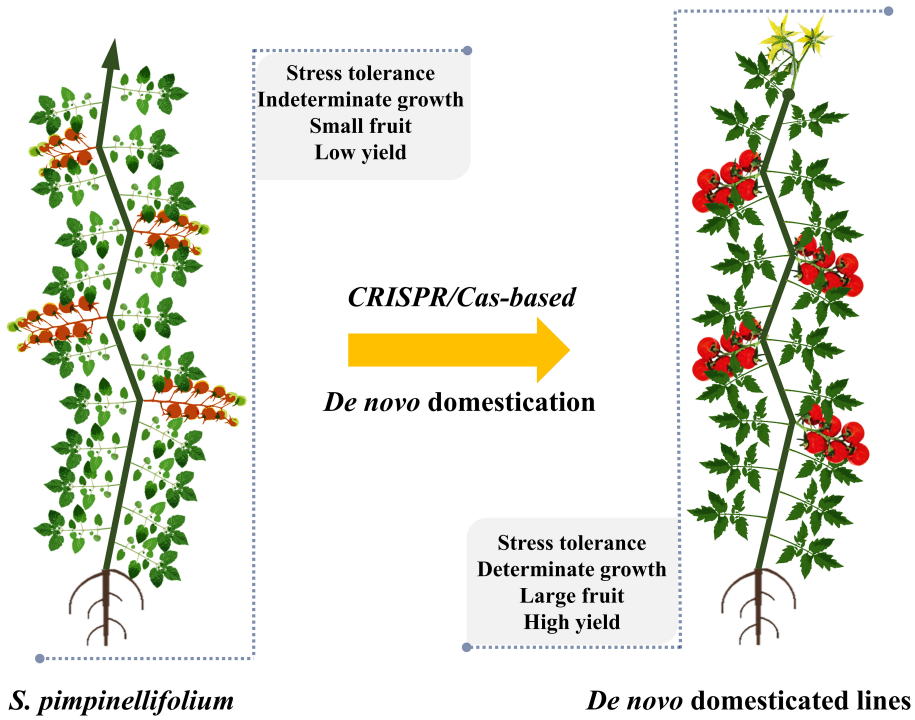


FIGURE 3 | *De novo* domestication of *S. pimpinellifolium* by CRISPR/Cas technology. A representative *S. pimpinellifolium* plant and its *de novo* domesticated form are shown, and the CRISPR/Cas-mediated domesticated traits are indicated. The plants are not drawn to scale.

Therefore, the *de novo* domestication of wild relatives or semicultivated plants by precisely introducing selected traits from their domesticated cultivars may yield potentially novel crops as alternative options for food supply. The ideas of redomestication and *de novo* domestication are strongly supported by CRISPR/Cas technology, which eases the manipulation of theoretically any genomic site of a genome of interest, including those of wild relatives and semicultivated plants. There are merely 15 plant species in use for food production, but thousands of semicultivated species are considered orphan crops and used to produce foods locally (Lemmon et al., 2018; Fernie and Yan, 2019). For tomato, the wild relative *S. pimpinellifolium* was used as a parental plant for *de novo* domestication of elite traits (**Figure 3**) (Li et al., 2018c; Zsogon et al., 2018). The most important traits selected for *de novo* domestication in the studies were growth habit (SELF PRUNING, SP and SP5G), fruit setting (OV, CLV3, FW2.2; MULT and WUS), and fruit quality (CycB and SlGGP1). The most desirable traits were obtained when knock-out mutations were precisely introduced in the selected genes in the genome of *S. pimpinellifolium* (**Table 2**) (Li et al., 2018c; Zsogon et al., 2018).

Proposing a different strategy for *de novo* domestication using genome editing technology, Lemmon and coworkers introduced targeted mutations (orthologs of the tomato domesticated alleles) into the genome of an orphan crop (*Physalis pruinosa*), a sweet ground cherry of the *Solanaceae* originating from Central and South America. The determinate growth, higher fruit number, and higher locule number traits were precisely added into the plant (**Table 2**), and generated phenotypes similar to those of their tomato orthologs (**Table 1**) (Lemmon et al., 2018).

Taken together, the above studies have paved a novel path toward obtaining *de novo* domesticated/redomesticated plants at the fastest rates for the breeding of resilient tomato to cope with environmental changes by the wild genetic background while enhancing/sustaining productivity and fruit quality.

More Precise Tomato Breeding at the "Speed of Light"

CRISPR/Cas-based targeted mutagenesis is highly flexible and efficient for targeting theoretically any desirable site. However, its precision is at the locus/gene level and is not controllable at every single base. Therefore, the approach can be readily applied for targeted knock-outs within coding sequences or random changes of non-coding sequences, such as *cis*-elements, for random promoter engineering (Rodriguez-Leal et al., 2017; Li et al., 2020).

Precision editing at every single base by base editors (**Figure 1B**) has been extensively conducted in animals and plants but is still limited to nucleotide transitions or C to G transversion. Most of the base substitutions shown in plants were [C/G to T/A] or [A/T to G/C] (Mishra et al., 2020), and transversion base editing has not been demonstrated in plants, thus limiting the applications of base editors for crop improvement. Nevertheless, BE has been used at limited scales in tomato (**Table 1**) (Shimatani

TABLE 2 | *De novo* domestication of Solanaceae using multiplexed CRISPR/Cas tools.

Species	Trait category	Target gene	Accession/Contig ID	Major mutant phenotype	Editing efficiency (%)	References
Solanum pimpinellifolium	Growth habit	SlSP (Self-prunning)	Solyc06g074350	Determinate growth	30.0	Zsogon et al., 2018
	Fruit shape	SlOV (Ovate)	Solyc02g085500	Oval fruit	30.0	
	Fruit size	SlFAS (Fasciated/Yabby)	Solyc11g071810	-	0.0–66.7	
		SlFW2.2 (Fruit weight 2.2)	Solyc02g090730	No obvious phenotype	30.0–66.7	
		SlCLV3	Solyc11g071380	Higher fruit locule number and fruit weight	66.7	
	Fruit number	SlMULT (Multiflora)	Solyc02g077390	Higher number of fruits per truss	0.0–66.7	
	Fruit quality	SlCycB (Lycopene beta cyclase)	Solyc04g040190	Higher accumulation of lycopene	30.0	
	Growth habit	SlSP5G (Self-prunning 5G)	Solyc05g053850	Determinate growth	32.1–57.1	Li et al., 2018c
	Growth habit	SlSP	Solyc06g074350	Determinate growth		
	Fruit size	SlCLV3	Solyc11g071380	Very slighly higher locule number and fruit size		
	Fruit size	SlWUS (Wushel)	Solyc02g083950	Higher locule number and fruit size		
	Fruit quality	SlGGP1 (GDP-L-Galactose phosphorylase)	Solyc02g091510	Increased foliar ascorbic acid content	-	
Physalis pruinosa	Growth habit	PpAGO7	Ppr-t_75930 through Ppr-t_75944	Narrower leaves and petals	-	Lemmon et al., 2018
	Growth habit	PpSP	Ppr-t_24561	Severe determinate plant	-	
	Growth habit	PpSP5G	Ppr-g_k141_668713	Determinate growth, higher fruit number	-	
	Growth habit	PpJ2 (Jointless-2)	Ppr-t_50452	Jointless unbranched inflorescences	-	
	Fruit size	PpCLV1	Ppr-t_75296	Higher locule number and fruit size	-	

et al., 2017; Veillet et al., 2019). Recently, novel approaches have been explored for substitutions of any base of interest, such as prime editing (PE) (**Figure 1C**), or precise editing at a medium DNA length using microhomology-mediated end joining (Tan et al., 2020; Van Vu et al., 2020). Although prime editor appeared to be efficient in animals (Anzalone et al., 2019) and well-adapted to monocot plants (Lin et al., 2020), its application in dicots remains limited and needs further improvement (Lu et al., 2020). While an allele that can be created by substituting just a few SNPs within a particular editing window is achievable by base editors or prime editors, the more base changes and the wider the DNA window that are required, the more complicated and challenging it is for base editors or prime editors to edit precisely (Rees and Liu, 2018; Hua et al., 2020; Lin et al., 2020; Mishra et al., 2020).

The HKI approach may be the last option for precise gene editing for crop plants due to its low frequency and complexity in design, but it can be used to precisely edit most, if not all, of the types of base/DNA changes of interest. HKI-mediated precision editing in tomato can range in size from a single base to thousands of base pairs (Cermak et al., 2015; Yu et al., 2017; Dahan-Meir et al., 2018; Danilo et al., 2019; Vu et al., 2020). HKI frequency has been continuously improved from the trace level in nature to a level that can realistically and affordably be used for crop precision breeding. The milestones in tomato HKI improvements came from the use of CRISPR/Cas for DSB formation (Yu et al., 2017; Danilo et al., 2019), the combination of CRISPR/SpCas9 with geminiviral replicons (Cermak et al., 2015; Dahan-Meir et al., 2018); and CRISPR/LbCas12a (LbCpf1) with multireplicons (Vu et al., 2020). HKI could be further improved by fine-engineering components of the CRISPR/Cas complexes, such as temperature-tolerant LbCas12a (Merker et al., 2020), to reach a true "speed-of-light" (Wolter et al., 2019) precision genome editing technology for tomato breeding.

CONCLUDING REMARKS

The cultivated tomato was domesticated and selected to retain favorable traits for consumption and/or processing. However, the domestication process and subsequent breeding dramatically

reduced the genetic diversity among the modern commonly used tomato cultivars (Lin et al., 2014; Liu et al., 2020; Wang et al., 2020a,b). Global and local climate change has put pressure on tomato growers to sustain production and, at the same time, to diversify their products, such as those with more favorable colors, flavors, or higher nutritional/health quality. Conventionally, to introgress an elite allele into a cultivated variety, breeders have to perform hybrid crossing with a donor source, usually a wild relative. The crossing helps to generate a hybrid genome with the allele of interest but also leads to the introduction of undesired genetic background or linkage drag from the donor parent. The most undesired traits can be removed by backcrossing several times to the parental elite line and selection for the interested allele in each generation of offspring. However, linkage drag makes this more challenging. Therefore, conventional breeding usually requires years to obtain a new tomato variety for cultivation, even with MAS approaches.

The emergence of CRISPR/Cas technology (**Figure 1**), one of the ultimate NPBTs, has spurred a revolution in crop breeding, including tomato breeding. CRISPR/Cas studies conducted on tomato as a model system have been extensively reported (**Figure 2** and **Table 1**). Since then, allele introgression via targeted mutagenesis, as well as the more precise BE and HKI in tomato, has become easier, and the time required to produce a new variety has been vastly reduced, to months. CRISPR/Cas tools have been widely used for generating tomato strains with better growth habits, improved fruit quantity, and quality for higher productivity with better nutrition and health properties (**Table 1**). Moreover, research that focuses on the ability to withstand environmental stresses has also been extensively released (**Table 1**).

Recently, a new trend in using CRISPR/Cas in tomato breeding has been to accelerate the so-called *de novo* domestication of new tomato varieties by introducing elite traits evolved during domestication and selective breeding into its wild relative/ancestor *S. pimpinellifolium* (**Figure 3** and **Table 2**). The resulting plants carried improved traits, such as better growth, larger fruits, and higher productivity and quality, but they still retain important wild traits, including stress tolerance, especially those traits determined by multiple genes/alleles (Li et al., 2018c; Zsogon et al., 2018). This approach would help to save years in breeding super tomato cultivars that are resilient to climate change. Another idea is to *de novo* domesticate tomato-like orphan crops through CRISPR/Cas-based introgression of orthologous genes/alleles of cultivated tomato for improved growth and higher yield (Lemmon et al., 2018). It would be interesting to apply this approach to many potential tomato-like orphan crops to increase the production capability of local growers.

CRISPR/Cas-based targeted mutagenesis itself is much more precise than random mutagenesis technologies using chemicals or radiation. However, a large portion of important traits in tomato is encoded by complex alleles that require precise base/sequence replacements. The recent advancement of CRISPR/Cas applications has created more precise editing tools, such as BE, PE, and HKI. These tools are continuously improving to become more efficient and precise for an era of faster tomato breeding.

AUTHOR CONTRIBUTIONS

TV and J-YK: conceptualization and supervision. TV: methodology and writing—original draft. TV, SD, MT, JH, and J-YK: writing—review and editing. JH and J-YK: funding acquisition. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Research Foundation of Korea (Grant NRF 2020M3A9I4038352 and 2020R1A6A1A03044344), the Next-Generation BioGreen 21 Program (SSAC, Grant PJ01322601), and the Program for New Plant Breeding Techniques (NBT, Grant PJ01478401), Rural Development Administration (RDA), Republic of Korea.

ACKNOWLEDGMENTS

We apologize to colleagues whose work could not be included owing to space constraints. We thank Ngan Thi Nguyen for helping to prepare the figures.

REFERENCES

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., et al. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576, 149–157. doi: 10.1038/s41586-019-1711-4

Bai, Y., Kissoudis, C., Yan, Z., Visser, R. G. F., and van der Linden, G. (2018). Plant behaviour under combined stress: tomato responses to combined salinity and pathogen stress. *Plant J.* 93, 781–793. doi: 10.1111/tpj.13800

Bai, Y., and Lindhout, P. (2007). Domestication and breeding of tomatoes: what have we gained and what can we gain in the future? *Ann. Bot.* 100, 1085–1094. doi: 10.1093/aob/mcm150

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., et al. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709–1712. doi: 10.1126/science.1138140

Belhaj, K., Chaparro-Garcia, A., Kamoun, S., and Nekrasov, V. (2013). Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods* 9:39. doi: 10.1186/1746-4811-9-39

Brooks, C., Nekrasov, V., Lippman, Z. B., and Van Eck, J. (2014). Efficient gene editing in tomato in the first generation using the clustered regularly interspaced short palindromic repeats/CRISPR-associated9 system. *Plant Physiol.* 166, 1292–1297. doi: 10.1104/pp.114.247577

Butt, H., Rao, G. S., Sedeek, K., Aman, R., Kamel, R., and Mahfouz, M. (2020). Engineering herbicide resistance via prime editing in rice. *Plant Biotechnol J.* 18:2370–2. doi: 10.1111/pbi.13399

Cai, R., Lewis, J., Yan, S., Liu, H., Clarke, C. R., Campanile, F., et al. (2011). The plant pathogen *Pseudomonas syringae* pv. tomato is genetically monomorphic and under strong selection to evade tomato immunity. *PLoS Pathog.* 7:e1002130. doi: 10.1371/journal.ppat.1002130

Cermak, T., Baltes, N. J., Cegan, R., Zhang, Y., and Voytas, D. F. (2015). High-frequency, precise modification of the tomato genome. *Genome Biol.* 16:232. doi: 10.1186/s13059-015-0796-9

Chapman, J. R., Taylor, M. R., and Boulton, S. J. (2012). Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell* 47, 497–510. doi: 10.1016/j.molcel.2012.07.029

Chen, K., Wang, Y., Zhang, R., Zhang, H., and Gao, C. (2019). CRISPR/Cas genome editing and precision plant breeding in agriculture. *Annu. Rev. Plant Biol.* 70, 667–697. doi: 10.1146/annurev-arplant-050718-100049

Collard, B. C., and Mackill, D. J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 363, 557–572. doi: 10.1098/rstb.2007.2170

Cox, D. B. T., Gootenberg, J. S., Abudayyeh, O. O., Franklin, B., Kellner, M. J., Joung, J., et al. (2017). RNA editing with CRISPR-Cas13. *Science* 358, 1019–1027. doi: 10.1126/science.aaq0180

Cunningham, F. J., Goh, N. S., Demirer, G. S., Matos, J. L., and Landry, M. P. (2018). Nanoparticle-mediated delivery towards advancing plant genetic engineering. *Trends Biotechnol.* 36, 882–897. doi: 10.1016/j.tibtech.2018.03.009

Dahan-Meir, T., Filler-Hayut, S., Melamed-Bessudo, C., Bocobza, S., Czosnek, H., Aharoni, A., et al. (2018). Efficient in planta gene targeting in tomato using geminiviral replicons and the CRISPR/Cas9 system. *Plant J.* 95, 5–16. doi: 10.1111/tpj.13932

D'Ambrosio, C., Stigliani, A. L., and Giorio, G. (2018). CRISPR/Cas9 editing of carotenoid genes in tomato. *Transgenic Res.* 27, 367–378. doi: 10.1007/s11248-018-0079-9

Danilo, B., Perrot, L., Mara, K., Botton, E., Nogue, F., and Mazier, M. (2019). Efficient and transgene-free gene targeting using Agrobacterium-mediated delivery of the CRISPR/Cas9 system in tomato. *Plant Cell Rep.* 38, 459–462. doi: 10.1007/s00299-019-02373-6

Fernie, A. R., and Yan, J. (2019). *De novo* domestication: an alternative route toward new crops for the future. *Mol. Plant* 12, 615–631. doi: 10.1016/j.molp.2019.03.016

Foolad, M. R., and Panthee, D. R. (2012). Marker-assisted selection in tomato breeding. *Crit. Rev. Plant Sci.* 31, 93–123. doi: 10.1080/07352689.2011.616057

Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I., et al. (2017). Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* 551, 464–471. doi: 10.1038/nature24644

Herner, R. C., and Sink, K. C. (1973). Ethylene production and respiratory behavior of the rin tomato mutant. *Plant Physiol.* 52, 38–42. doi: 10.1104/pp.52.1.38

Hickey, L. T., Robinson, H., Jackson, S. A., Leal-Bertioli, S. C. M., Tester, M. (2019). Breeding crops to feed 10 billion. *Nat. Biotechnol.* 37, 744–754. doi: 10.1038/s41587-019-0152-9

Hua, K., Jiang, Y., Tao, X., and Zhu, J. K. (2020). Precision genome engineering in rice using prime editing system. *Plant Biotechnol J.* 18, 2167–2169. doi: 10.1111/pbi.13395

Huckelhoven, R. (2014). The effective papilla hypothesis. *N. Phytol.* 204, 438–440. doi: 10.1111/nph.13026

Ito, Y., Nishizawa-Yokoi, A., Endo, M., Mikami, M., and Toki, S. (2015). CRISPR/Cas9-mediated mutagenesis of the RIN locus that regulates tomato fruit ripening. *Biochem. Biophys. Res. Commun.* 467, 76–82. doi: 10.1016/j.bbrc.2015.09.117

Jenkins, J. A. (1948). The origin of the cultivated tomato. *Economic Bot.* 2, 379–392. doi: 10.1007/BF02859492

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi: 10.1126/science.1225829

Jones, H., Whipps, J. M., and Gurr, S. J. (2001). The tomato powdery mildew fungus *Oidium neolycopersici*. *Mol. Plant Pathol.* 2, 303–309. doi: 10.1046/j.1464-6722.2001.00084.x

Kelliher, T., Starr, D., Su, X., Tang, G., Chen, Z., Carter, J., et al. (2019). One-step genome editing of elite crop germplasm during haploid induction. *Nat. Biotechnol.* 37, 287–292. doi: 10.1038/s41587-019-0038-x

Klap, C., Yeshayahou, E., Bolger, A. M., Arazi, T., Gupta, S. K., Shabtai, S., et al. (2017). Tomato facultative parthenocarpy results from SlAGAMOUS-LIKE 6 loss of function. *Plant Biotechnol. J.* 15, 634–647. doi: 10.1111/pbi.12662

Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. doi: 10.1038/nature17946

Kurt, I. C., Zhou, R., Iyer, S., Garcia, S. P., Miller, B. R., Langner, L. M., et al. (2020). CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nat. Biotechnol.* doi: 10.1038/s41587-020-0609-x. [Epub ahead of print].

Lee, J., Nonaka, S., Takayama, M., and Ezura, H. (2018). Utilization of a genome-edited tomato (*Solanum lycopersicum*) with high gamma aminobutyric acid content in hybrid breeding. *J. Agric. Food Chem.* 66, 963–971. doi: 10.1021/acs.jafc.7b05171

Lemmon, Z. H., Reem, N. T., Dalrymple, J., Soyk, S., Swartwood, K. E., Rodriguez-Leal, D., et al. (2018). Rapid improvement of domestication traits in an orphan crop by genome editing. *Nat. Plants* 4, 766–770. doi: 10.1038/s41477-018-0259-x

Li, J. F., Norville, J. E., Aach, J., McCormack, M., Zhang, D., Bush, J., et al. (2013). Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. *Nat. Biotechnol.* 31, 688–691. doi: 10.1038/nbt.2654

Li, Q., Sapkota, M., and van der Knaap, E. (2020). Perspectives of CRISPR/Cas-mediated cis-engineering in horticulture: unlocking the neglected potential for crop improvement. *Hortic Res* 7, 36.

Li, R., Li, R., Li, X., Fu, D., Zhu, B., Tian, H., et al. (2018a). Multiplexed CRISPR/Cas9-mediated metabolic engineering of gamma-aminobutyric acid levels in Solanum lycopersicum. *Plant Biotechnol. J.* 16, 415–427. doi: 10.1111/pbi.12781

Li, R., Liu, C., Zhao, R., Wang, L., Chen, L., Yu, W., et al. (2019). CRISPR/Cas9-mediated SlNPR1 mutagenesis reduces tomato plant drought tolerance. *BMC Plant Biol.* 19:38. doi: 10.1186/s12870-018-1627-4

Li, R., Zhang, L., Wang, L., Chen, L., Zhao, R., Sheng, J., et al. (2018b). Reduction of tomato-plant chilling tolerance by CRISPR-Cas9-mediated SlCBF1 mutagenesis. *J. Agric. Food Chem.* 66, 9042–9051. doi: 10.1021/acs.jafc.8b02177

Li, T., Yang, X., Yu, Y., Si, X., Zhai, X., Zhang, H., et al. (2018c). Domestication of wild tomato is accelerated by genome editing. *Nat. Biotechnol.* (2018) 36:1160–3. doi: 10.1038/nbt.4273

Li, X., Wang, Y., Chen, S., Tian, H., Fu, D., Zhu, B., et al. (2018d). Lycopene is enriched in tomato fruit by CRISPR/Cas9-mediated multiplex genome editing. *Front Plant Sci.* 9:559. doi: 10.3389/fpls.2018.00559

Lieber, M. R. (2010). The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev. Biochem.* 79, 181–211. doi: 10.1146/annurev.biochem.052308.093131

Lin, Q., Zong, Y., Xue, C., Wang, S., Jin, S., Zhu, Z., et al. (2020). Prime genome editing in rice and wheat. *Nat. Biotechnol.* 38, 582–585. doi: 10.1038/s41587-020-0455-x

Lin, T., Zhu, G., Zhang, J., Xu, X., Yu, Q., Zheng, Z., et al. (2014). Genomic analyses provide insights into the history of tomato breeding. *Nat. Genet.* 46, 1220–1226. doi: 10.1038/ng.3117

Liu, D., Yang, L., Zhang, J.-z., Zhu, G.-t. Lü, H.-j., et al. (2020). Domestication and breeding changed tomato fruit transcriptome. *J. Integrat. Agricult.* 19, 120–132. doi: 10.1016/S2095-3119(19)62824-8

Lu, Y., Tian, Y., Shen, R., Yao, Q., Zhong, D., Zhang, X., et al. (2020). Precise genome modification in tomato using an improved prime editing system. *Plant Biotechnol. J.* doi: 10.1111/pbi.13497. [Epub ahead of print].

Maher, M. F., Nasti, R. A., Vollbrecht, M., Starker, C. G., Clark, M. D., and Voytas, D. F. (2020). Plant gene editing through de novo induction of meristems. *Nat. Biotechnol.* 38, 84–89. doi: 10.1038/s41587-019-0337-2

Merker, L., Schindele, P., Huang, T. K., Wolter, F., and Puchta, H. (2020). Enhancing in planta gene targeting efficiencies in Arabidopsis using temperature-tolerant CRISPR/LbCas12a. *Plant Biotechnol J.* 18:2382–4. doi: 10.1111/pbi.13426

Mishra, R., Joshi, R. K., and Zhao, K. (2020). Base editing in crops: current advances, limitations and future implications. *Biotech J.* 18, 20–31. doi: 10.1111/pbi.13225

Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J. D., and Kamoun, S. (2013). Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* 31, 691–693. doi: 10.1038/nbt.2655

Nekrasov, V., Wang, C., Win, J., Lanz, C., Weigel, D., and Kamoun, S. (2017). Rapid generation of a transgene-free powdery mildew resistant tomato by genome deletion. *Sci. Rep.* 7:482. doi: 10.1038/s41598-017-0 0578-x

Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., et al. (2016). Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* 353:aaf8729. doi: 10.1126/science.aaf8729

Nonaka, S., Arai, C., Takayama, M., Matsukura, C., and Ezura, H. (2017). Efficient increase of -aminobutyric acid (GABA) content in tomato fruits by targeted mutagenesis. *Sci. Rep.* 7:7057. doi: 10.1038/s41598-017-06400-y

Ortigosa, A., Gimenez-Ibanez, S., Leonhardt, N., and Solano, R. (2019). Design of a bacterial speck resistant tomato by CRISPR/Cas9-mediated editing of SlJAZ2. *Plant Biotechnol. J.* 17, 665–673. doi: 10.1111/pbi.13006

Pan, C., Ye, L., Qin, L., Liu, X., He, Y., Wang, J., et al. (2016). CRISPR/Cas9-mediated efficient and heritable targeted mutagenesis in tomato plants in the first and later generations. *Sci. Rep.* 6:24765. doi: 10.1038/srep 24765

Paula de Toledo Thomazella, D., Brail, Q., Dahlbeck, D., and Staskawicz, B. (2016). CRISPR-Cas9 mediated mutagenesis of a DMR6 ortholog in tomato confers broad-spectrum disease resistance. *bioRxiv* 2016:064824. doi: 10.1101/064824

Pramanik, D., Shelake, R. M., Kim, M. J., and Kim, J. Y. (2020). CRISPR-mediated engineering across the central dogma in plant biology for basic research and crop improvement. *Mol. Plant*. doi: 10.1016/j.molp.2020.11.002. [Epub ahead of print].

Puchta, H. (2005). The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *J. Exp. Bot.* 56, 1–14. doi: 10.1093/jxb/eri025

Ranc, N., Munos, S., Santoni, S., and Causse, M. (2008). A clarified position for *Solanum lycopersicum* var. cerasiforme in the evolutionary history of tomatoes (solanaceae). *BMC Plant Biol.* 8:130. doi: 10.1186/1471-2229-8-130

Ray, D. K., Mueller, N. D., West, P. C., and Foley, J. A. (2013). Yield trends are insufficient to double global crop production by 2050. *PLoS ONE* 8:e66428. doi: 10.1371/journal.pone.0066428

Rees, H. A., and Liu, D. R. (2018). Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* 19, 770–788. doi: 10.1038/s41576-018-0059-1

Rodriguez-Leal, D., Lemmon, Z. H., Man, J., Bartlett, M. E., and Lippman, Z. B. (2017). Engineering quantitative trait variation for crop improvement by genome editing. *Cell* 171, 470–480.e8.

Roldan, M. V. G., Perilleux, C., Morin, H., Huerga-Fernandez, S., Latrasse, D., Benhamed, M., et al. (2017). Natural and induced loss of function mutations in SlMBP21 MADS-box gene led to jointless-2 phenotype in tomato. *Sci. Rep.* 7:4402. doi: 10.1038/s41598-017-04556-1

Santillan Martinez, M. I., Bracuto, V., Koseoglou, E., Appiano, M., Jacobsen, E., Visser, R. G. F., et al. (2020). CRISPR/Cas9-targeted mutagenesis of the tomato susceptibility gene PMR4 for resistance against powdery mildew. *BMC Plant Biol.* 20:284. doi: 10.1186/s12870-020-02497-y

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., et al. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 686–688. doi: 10.1038/nbt.2650

Shelake, R. M., Pramanik, D., and Kim, J. Y. (2019). Exploration of plant-microbe interactions for sustainable agriculture in CRISPR Era. *Microorganisms* 7:269. doi: 10.3390/microorganisms7080269

Shimatani, Z., Kashojiya, S., Takayama, M., Terada, R., Arazoe, T., Ishii, H., et al. (2017). Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. *Nat. Biotechnol.* 35, 441–443. doi: 10.1038/nbt.3833

Soyk, S., Lemmon, Z. H., Oved, M., Fisher, J., Liberatore, K. L., Park, S. J., et al. (2017). Bypassing negative epistasis on yield in tomato imposed by a domestication gene. *Cell* 169:e1112. doi: 10.1016/j.cell.2017.04.032

Story, E. N., Kopec, R. E., Schwartz, S. J., and Harris, G. K. (2010). An update on the health effects of tomato lycopene. *Annu. Rev. Food Sci. Technol.* 1, 189–210. doi: 10.1146/annurev.food.102308.124120

Takayama, M., Koike, S., Kusano, M., Matsukura, C., Saito, K., Ariizumi, T., et al. (2015). Tomato glutamate decarboxylase genes SlGAD2 and SlGAD3 play key roles in regulating gamma-aminobutyric acid levels in tomato (*Solanum lycopersicum*). *Plant Cell Physiol.* 56, 1533–1545. doi: 10.1093/pcp/pcv075

Takayama, M., Matsukura, C., Ariizumi, T., and Ezura, H. (2017). Activating glutamate decarboxylase activity by removing the autoinhibitory domain leads to hyper gamma-aminobutyric acid (GABA) accumulation in tomato fruit. *Plant Cell Rep.* 36, 103–116. doi: 10.1007/s00299-016-2061-4

Tan, J., Zhao, Y., Wang, B., Hao, Y., Wang, Y., Li, Y., et al. (2020). Efficient CRISPR/Cas9-based plant genomic fragment deletions by microhomology-mediated end joining. *Plant Biotechnol. J.* 18, 2161–2163. doi: 10.1111/pbi.13390

Tieman, D., Zhu, G., Resende, M. F. Jr., Lin, T., Nguyen, C., Bies, D., et al. (2017). A chemical genetic roadmap to improved tomato flavor. *Science* 355, 391–394. doi: 10.1126/science.aal1556

Tomlinson, L., Yang, Y., Emenecker, R., Smoker, M., Taylor, J., Perkins, S., et al. (2019). Using CRISPR/Cas9 genome editing in tomato to create a gibberellin-responsive dominant dwarf DELLA allele. *Plant Biotechnol. J.* 17, 132–140. doi: 10.1111/pbi.12952

Tran, M. T., Doan, D. T. H., Kim, J., Song, Y. J., Sung, Y. W., Das, S., et al. (2020). CRISPR/Cas9-based precise excision of SlHyPRP1 domain(s) to obtain salt stress-tolerant tomato. *Plant Cell Rep*. doi: 10.1007/s00299-020-02622-z. [Epub ahead of print].

Ueta, R., Abe, C., Watanabe, T., Sugano, S. S., Ishihara, R., Ezura, H., et al. (2017). Rapid breeding of parthenocarpic tomato plants using CRISPR/Cas9. *Sci. Rep.* 7:507. doi: 10.1038/s41598-017-00501-4

Van Vu, T., Sung, Y. W., Kim, J., Doan, D. T. H., Tran, M. T., and Kim, J. Y. (2019). Challenges and perspectives in homology-directed gene targeting in monocot plants. *Rice* 12:95. doi: 10.1186/s12284-019-0355-1

Van Vu, T., Thi Hai Doan, D., Kim, J., Sung, Y. W., Thi Tran, M., Song, Y. J., et al. (2020). CRISPR/Cas-based precision genome editing via microhomology-mediated end joining. *Plant Biotechnol J.* 24, 93–99. doi: 10.1111/pbi.13490

Veillet, F., Perrot, L., Chauvin, L., Kermarrec, M. P., Guyon-Debast, A., Chauvin, J. E., et al. (2019). Transgene-free genome editing in tomato and potato plants using agrobacterium-mediated delivery of a CRISPR/Cas9 cytidine base editor. *Int. J. Mol. Sci.* 20:402. doi: 10.3390/ijms20020402

Vu, T. V., Sivankalyani, V., Kim, E. J., Doan, D. T. H., Tran, M. T., Kim, J., et al. (2020). Highly efficient homology-directed repair using CRISPR/Cpf1-geminiviral replicon in tomato. *Plant Biotechnol. J.* 18, 2133–43. doi: 10.1111/pbi.13373

Wang, L., Chen, L., Li, R., Zhao, R., Yang, M., Sheng, J., et al. (2017). Reduced drought tolerance by CRISPR/Cas9-mediated SlMAPK3 mutagenesis in tomato plants. *J. Agric. Food Chem.* 65, 8674–8682. doi: 10.1021/acs.jafc.7bv02745

Wang, Z., Hong, Y., Li, Y., Shi, H., Yao, J., Liu, X., et al. (2020a). Natural variations in SlSOS1 contribute to the loss of salt tolerance during tomato domestication. *Plant Biotechnol. J.* doi: 10.1111/pbi.13443. [Epub ahead of print].

Wang, Z., Hong, Y., Zhu, G., Li, Y., Niu, Q., Yao, J., et al. (2020b). Loss of salt tolerance during tomato domestication conferred by variation in a Na(+) /K(+) transporter. *EMBO J.* 39:e103256. doi: 10.15252/embj.2019103256

Wolter, F., Schindele, P., and Puchta, H. (2019). Plant breeding at the speed of light: the power of CRISPR/Cas to generate directed genetic diversity at multiple sites. *BMC Plant Biol.* 19:176. doi: 10.1186/s12870-019-1775-1

Wu, S., Zhang, B., Keyhaninejad, N., Rodriguez, G. R., Kim, H. J., Chakrabarti, M., et al. (2018). A common genetic mechanism underlies morphological diversity in fruits and other plant organs. *Nat. Commun.* 9:4734. doi: 10.1038/s41467-018-07216-8

Xu, C., Liberatore, K. L., MacAlister, C. A., Huang, Z., Chu, Y. H., Jiang, K., et al. (2015). A cascade of arabinosyltransferases controls shoot meristem size in tomato. *Nat. Genet.* 47, 784–792. doi: 10.1038/ng.3309

Xu, C., Park, S. J., Van Eck, J., and Lippman, Z. B. (2016). Control of inflorescence architecture in tomato by BTB/POZ transcriptional regulators. *Genes Dev.* 30, 2048–2061. doi: 10.1101/gad.288415.116

Ye, J., Wang, X., Hu, T., Zhang, F., Wang, B., Li, C., et al. (2017). An InDel in the promoter of Al-ACTIVATED MALATE TRANSPORTER9 selected during tomato domestication determines fruit malate contents and aluminum tolerance. *Plant Cell* 29, 2249–2268. doi: 10.1105/tpc.17.00211

Yin, Y., Qin, K., Song, X., Zhang, Q., Zhou, Y., Xia, X., et al. (2018). BZR1 transcription factor regulates heat stress tolerance through FERONIA receptor-like kinase-mediated reactive oxygen species signaling in tomato. *Plant Cell Physiol.* 59, 2239–2254. doi: 10.1093/pcp/pcy146

Yu, Q. H., Wang, B., Li, N., Tang, Y., Yang, S., Yang, T., et al. (2017). CRISPR/Cas9-induced targeted mutagenesis and gene replacement to generate long-shelf life tomato lines. *Sci. Rep.* 7:11874. doi: 10.1038/s41598-017-12262-1

Yuste-Lisbona, F. J., Fernández-Lozano, A., Pineda, B., Bretones, S., Ortíz-Atienza, A., García-Sogo, B., et al. (2020). ENO regulates tomato fruit size through the floral meristem development network. *Proc. Natl. Acad. Sci. U.S.A.* 117, 8187–8195. doi: 10.1073/pnas.1913688117

Zhao, D., Li, J., Li, S., Xin, X., Hu, M., Price, M. A., et al. (2020). Glycosylase base editors enable C-to-A and C-to-G base changes. *Nat. Biotechnol.* doi: 10.1038/s41587-020-0592-2. [Epub ahead of print].

Zheng, Z., Appiano, M., Pavan, S., Bracuto, V., Ricciardi, L., Visser, R. G., et al. (2016). Genome-wide study of the tomato SlMLO gene family and its functional characterization in response to the powdery mildew fungus *Oidium neolycopersici*. *Front. Plant Sci.* 7:380. doi: 10.3389/fpls.2016.00380

Zsogon, A., Cermak, T., Naves, E. R., Notini, M. M., Edel, K. H., Weinl, S., et al. (2018). *De novo* domestication of wild tomato using genome editing. *Nat Biotechnol*. 36:1211–6. doi: 10.1038/nbt.4272

Zsogon, A., Cermak, T., Voytas, D., and Peres, L. E. (2017). Genome editing as a tool to achieve the crop ideotype and *de novo* domestication of wild relatives: case study in tomato. *Plant Sci.* 256, 120–130. doi: 10.1016/j.plantsci.2016.12.012

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Vu, Das, Tran, Hong and Kim. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genome Editing

ORIGINAL RESEARCH
published: 16 December 2020
doi: 10.3389/fgeed.2020.618385

Genome Engineering in Plant Using an Efficient CRISPR-xCas9 Toolset With an Expanded PAM Compatibility

*Chengwei Zhang, Guiting Kang, Xinxiang Liu, Si Zhao, Shuang Yuan, Lu Li, Yongxing Yang, Feipeng Wang, Xiang Zhang and Jinxiao Yang**

Beijing Key Laboratory of Maize DNA Fingerprinting and Molecular Breeding, Beijing Academy of Agriculture & Forestry Sciences, Beijing, China

The CRISPR-Cas9 system enables simple, rapid, and effective genome editing in many species. Nevertheless, the requirement of an NGG protospacer adjacent motif (PAM) for the widely used canonical *Streptococcus pyogenes* Cas9 (SpCas9) limits the potential target sites. The xCas9, an engineered SpCas9 variant, was developed to broaden the PAM compatibility to NG, GAA, and GAT PAMs in human cells. However, no knockout rice plants were generated for GAA PAM sites, and only one edited target with a GAT PAM was reported. In this study, we used tRNA and enhanced sgRNA (esgRNA) to develop an efficient CRISPR-xCas9 genome editing system able to mutate genes at NG, GAA, GAT, and even GAG PAM sites in rice. We also developed the corresponding xCas9-based cytosine base editor (CBE) that can edit the NG and GA PAM sites. These new editing tools will be useful for future rice research or breeding, and may also be applicable for other related plant species.

Keywords: genome editing, xCas9, cytosine base editor, tRNA-sgRNA, rice

OPEN ACCESS

Edited by:
Huanbin Zhou, Chinese Academy of Agricultural Sciences, China

Reviewed by:
Kabin Xie, Huazhong Agricultural University, China
Kejian Wang, Chinese Academy of Agricultural Sciences, China

****Correspondence:***
Jinxiao Yang
yangjinxiao@maizedna.org

Specialty section:
This article was submitted to Genome Editing in Plants, a section of the journal Frontiers in Genome Editing

Received: *16 October 2020*
Accepted: *24 November 2020*
Published: *16 December 2020*

Citation:
Zhang C, Kang G, Liu X, Zhao S, Yuan S, Li L, Yang Y, Wang F, Zhang X and Yang J (2020) Genome Engineering in Plant Using an Efficient CRISPR-xCas9 Toolset With an Expanded PAM Compatibility. Front. Genome Ed. 2:618385. doi: 10.3389/fgeed.2020.618385

INTRODUCTION

The clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated nuclease 9 (Cas9) system derived from microbial adaptive immune systems has facilitated diverse genomic manipulations, including targeted gene disruption (Bortesi and Fischer, 2015; Ma et al., 2015; Xie et al., 2015), transcriptional activation or repression (Lowder et al., 2015; Piatek et al., 2015), and base substitutions (Li et al., 2017; Lu and Zhu, 2017; Zong et al., 2017), in various organisms and cell types (Komor et al., 2017; Ge et al., 2019). The application of these genomic modifications has led to substantial advances in research regarding plant biology as well as crop breeding (Yin et al., 2017; Hille et al., 2018). However, to be recognized by Cas9, a target site requires a short protospacer adjacent motif (PAM) sequence at its 3′ end (Mojica et al., 2009). The widely used Cas9 from *Streptococcus pyogenes* (SpCas9) mainly recognizes an NGG PAM sequence (Sternberg et al., 2014), thereby restricting the targetable sites in the genome.

To address this limitation, researchers have used natural CRISPR nucleases with different PAM requirements, including those of *Neisseria meningitides* (NmeCas9) (Esvelt et al., 2013), *Streptococcus thermophilus* (StCas9) (Xu et al., 2015), *Staphylococcus aureus* (SaCas9) (*Ran et al., 2015*), *Campylobacter jejuni* (CjeCas9) (Kim et al., 2017), and *Geobacillus thermodenitrificans* (ThermoCas9 and GeoCas9) (Harrington et al., 2017; Mougiakos et al., 2017). However,

the PAMs recognized by these Cas9s are relatively complex, restricting the widespread use of these nucleases for genome editing. To date, only SaCas9 has been commonly applied in plants (Kaya et al., 2016; Qin et al., 2019). In addition to the Cas9 proteins, class 2 type V CRISPR-Cas systems involving Cas12a (or Cpf1) and Cas12b (or C2c1) have been adopted for modifying genomes at AT-rich PAM sequences (Zetsche et al., 2015; Teng et al., 2018). For example, LbCpf1 and FnCpf1 can modify the genomes of many plant species (Endo et al., 2016; Tang et al., 2017).

Another way to address the limitations of the CRISPR-Cas9 system related to the PAM sequence involves altering the PAM-interacting domain of Cas9. Several engineered SpCas9 and SaCas9 variants that can recognize NGA (VQR), NGCG (VRER), NGAG (EQR), and NNNRRT (SaKKH-Cas9) PAM sequences have been obtained (Kleinstiver et al., 2015a, Kleinstiver et al., 2015b). Nevertheless, the PAMs recognized by these Cas9 proteins are still relatively complex. Some engineered Cas9 variants with an increased PAM compatibility have recently been reported, including the SpCas9 variants, xCas9 and Cas9-NG, which enable researchers to target simple non-canonical NG PAM sites; these variants have robust editing activities without sacrificing DNA specificity in human cells (Hu et al., 2018; Nishimasu et al., 2018). Moreover, xCas9 is functional over a relatively broad range of PAM sequences, including GAA and GAT, which are not recognized by Cas9-NG. At least six groups have employed xCas9 to manipulate the rice genome (Hua et al., 2019; Li et al., 2019; Ren et al., 2019; Wang et al., 2019a; Zhong et al., 2019; Zeng et al., 2020). However, the xCas9-based gene disruption system is unexpectedly inefficient for targeting sites with GAA and GAT PAM sequences in transgenic T_0 plants. In the current study, we developed an efficient CRISPR-xCas9 system in rice that can recognize GAA, GAT, and even GAG PAM sites without sacrificing NG PAM recognition. We also developed the corresponding xCas9-based cytosine base editor (CBE), which enables the efficient C-to-T conversion at GA and NG PAM sites in rice.

MATERIALS AND METHODS

Plasmid Construction

In this study, the codons encoding the SpCas9 protein described by Cong et al. (2013) were first optimized for rice by GenScript Corp. (Nanjing, China). The following mutations were introduced by PCR to obtain xCas9: A262T/R324L/S409I/E480K/E543D/M694I/E1219V (**Supplementary Data**). The SpCas9n & PmCDA1 & UGI & T35s sequence (Wu et al., 2019) was replaced by the xCas9 & T35s fusion sequence between the SnabI and AvrII restriction enzyme sites in the SpCas9n-pBE-basic vector to generate pxCas9-basic-M. Next, pxCas9-basic-M was digested with BsaI and HindIII, after which the larger fragment lacking the sgRNA & *OsU3* terminator was purified and ligated to the esgRNA & poly-T fragment, which was digested with BbsI and HindIII to generate pxCas9-basic. Using the method described by Ma et al. (2015), targets were added to pxCas9-basic to generate pxCas9 constructs. Each pxCas9 construct comprised three target sequences, respectively, under the control of the *OsU3*, *OsU6c*, and *OsU6a* promoters. The pxCas9 constructs were digested with SnabI and AvrII, after which the xCas9 & T35s fragment was replaced by xCas9n & PmCDA1 & UGI & T35s and ecTadA & ecTadA* & xCas9n & T35s sequences, ultimately generating the xCas9n-CBE and xCas9n-ABE (adenine base editor) constructs, respectively. Regarding the vector for pxCas9 without a tRNA, pxCas9-basic-M was digested with BamHI and HindIII, after which the larger fragment lacking the tRNA & sgRNA & OsU3 terminator was purified and ligated to the esgRNA & poly-T fragment, which was digested with BamHI and HindIII to generate pxCas9-no tRNA-basic. Three target sequences were then added using primers lacking a tRNA sequence as described by Ma et al. (2015). All of the primers used in this study are listed in **Supplementary Table 1**.

Rice Transformation

All of the constructed binary vectors were inserted into *Agrobacterium tumefaciens* strain EHA105 cells using a freeze/thaw method. The transformed *A. tumefaciens* cells were then used to infect rice embryogenic calli induced from mature Nipponbare rice seeds as previously described (Hiei and Komari, 2008). After 10 min, the calli were recovered for 3 days and then cultured on selection medium containing 50 μg/mL hygromycin for 4 weeks to obtain hygromycin-resistant calli. The transgenic calli were then cultured on regeneration medium for ∼1 month to induce shoot development. Shoots 4–5 cm long were then cultured on rooting medium for ∼2 weeks to induce root development and obtain T_0 plants.

Identification of Transgenic T_0 Plants

Genomic DNA samples were extracted from T_0 plants using the DNA-quick Plant System kit (Tiangen Biotech, Beijing, China). Target loci were amplified by PCR using specific xCas9 primers (**Supplementary Table 2**). Transgenic T_0 plants were identified based on the amplification of an 854 bp fragment, which was detected by agarose gel electrophoresis.

Mutant Identification

Several transgenic T_0 plants were analyzed to identify gene mutations and C-to-T or A-to-G conversions. Target loci were amplified by PCR using specific primers (**Supplementary Table 2**). The PCR products were subjected to Sanger sequencing (General Biol, Anhui, China). The mutation types at the target sites (i.e., double peaks in the sequencing chromatograms) were determined using an online tool (http://skl.scau.edu.cn/dsdecode/) (Liu et al., 2015; Xie et al., 2017). The insertions/deletions (Indels) at or near the target sites were defined as gene mutations. Additionally, the frequency (%) of the C-to-T or A-to-G conversion was calculated based on the number of mutants with any target C-to-T or A-to-G substitution among all tested transgenic T_0 plants. The frequency (%) of different mutation types was calculated based on the number of mutants with the same mutation type among all of the mutants. A mutant in which all of the C-to-T conversions at the target site were homozygous was considered to be a homozygous mutant.

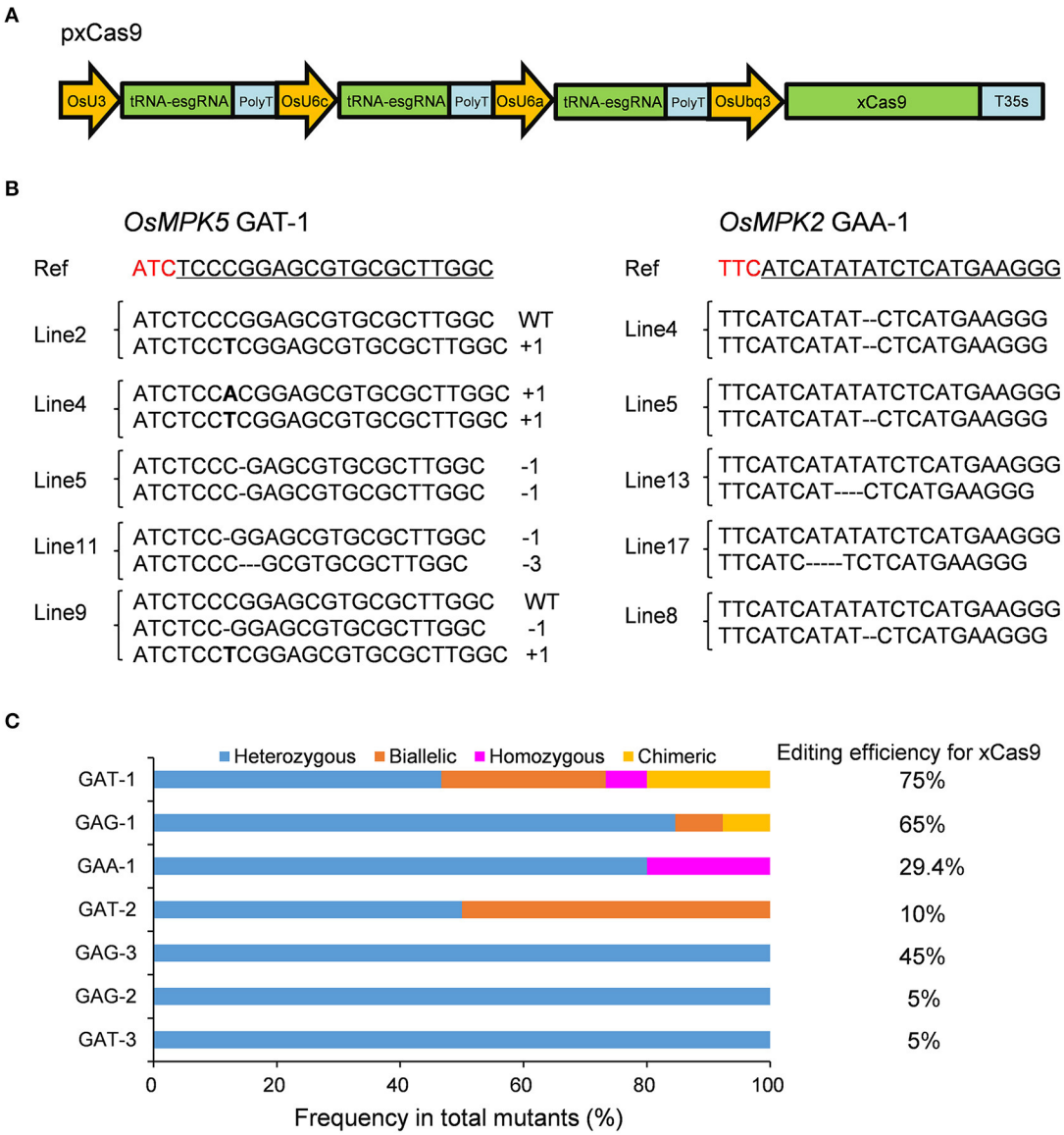


FIGURE 1 | Gene mutagenesis at GAD PAM sites induced by xCas9 in rice T_0 plants. **(A)** Schematic illustration of the pxCas9 construct for the CRISPR-xCas9 system. **(B)** Mutation types at GAA-1 and GAT-1 target sites of some representative stable mutant lines. Ref sequence is the wild-type sequence. Target sequences are underlined and PAM sequences are indicated in red. **(C)** Frequencies of different mutation types at GAD PAM sites mutated by xCas9.

RESULTS

xCas9 Induces Gene Mutations at GAD (Where D is A, T, or G) PAM Sites in Rice Plants

In our previous study, we revealed that the tRNA-esgRNA system might help the xCas9-based CBE to efficiently edit target sites with a GA PAM in rice (Zhang et al., 2020). Therefore, we used tRNA and esgRNA to develop a CRISPR-xCas9 system to assess the cleavage activity of xCas9 in rice (**Figure 1A**). The rice codon-optimized xCas9 sequence with A262T/R324L/S409I/E480K/E543D/M694I/E1219V mutations (derived from SpCas9) (Zhang et al., 2020) under the control of the *Oryza sativa* ubiquitin (*OsUbq*) promoter was used in this study (**Figure 1A**). Each tRNA together with esgRNA was placed under the control of the rice *U3* or *U6* promoter (**Figure 1A**).

Because xCas9 can edit GAA and GAT PAM target sites in human cells (Hu et al., 2018), we first tested the editing activity of xCas9 in rice at three GAA PAM target sites in the *OsMPK2*, *OsMPK5*, and *OsNRT1.1B* genes (GAA-1, GAA-2, and GAA-3, respectively) (**Supplementary Table 3**) as well as at three GAT PAM sites in the *OsMPK5* or *OsWaxy* gene (GAT-1, GAT-2, and GAT-3) (**Table 1** and **Supplementary Table 3**). In the T_0 plants, all three GAT PAM sites were edited, with

TABLE 1 | Summary of the mutation frequencies and mutation types at GA PAM sites targeted by xCas9.

PAM sequence	Target site	Target gene	Target sequence	Tested T_0 plants	Edited T_0 plants	Mutation frequency (%)	Mutation types
GAA	GAA-1	*OsMPK2*	CCCTTCATGAGATATATGATGAA	17	5	29.4	4He+1Ho
	GAA-2	*OsMPK5*	CACCTTCAACCCGCTGCAGAGAA	19	0	0	
	GAA-3	*OsNRT1.1B*	ACCAGGAGGTACAGGTTCAAGAA	18	0	0	
GAT	GAT-1	*OsMPK5*	GCCAAGCGCACGCTCCGGGAGAT	20	15	75	7He+4Bi+1Ho+3Chi
	GAT-2	*OsWaxy*	CTCATGCAGGAGGACGTCCAGAT	20	2	10	1He+1Bi
	GAT-3	*OsWaxy*	CTGCTCCTTGAAGAGCCTGAGAT	20	1	5	1He
GAG	GAG-1	*OsWaxy*	CGTCATTCCTGGAGAAGGTGGAG	20	13	65	11He+1Bi+1Chi
	GAG-2	*OsALS*	CCAACCACCTCTTCCGCCACGAG	20	1	5	1He
	GAG-3	*OsALS*	GCTTCCTCATGAACATTCAGGAG	20	9	45	9He
GAC	GAC-1	*OsWaxy*	TACCGAGGAGAGATCACCATGAC	20	0	0	
	GAC-2	*OsGRF4*	CACTTTCGTTCTTTGGGAACGAC	20	0	0	
	GAC-3	*OsNRT1.1B*	GCGACCACCATCATGTTCTGGAC	20	0	0	
	GAC-4	*OsALS*	ACCTTGTCCTTGATGTGGAGGAC	20	0	0	
	GAC-5	*OsWaxy*	CCACCGGCTTCGGCATCGCCGAC	20	0	0	
	GAC-6	*OsALS*	ACCTCGTGTCCGCGCTCGCCGAC	20	0	0	
	GAC-7	*OsNRT1.1B*	GCTTCGGCTCCGACCAGTTCGAC	20	0	0	
	GAC-8	*OsPDS*	TCCTGCAGAGGAATGGGTTGGAC	20	0	0	

The PAM sequence is underlined. He, heterozygous mutation; Bi, biallelic mutation; Ho, homozygous mutation; Chi, chimeric mutation.

frequencies ranging from 5 to 75% (**Figure 1B** and **Table 1**). Of the three GAA PAM sites, only GAA-1 was edited, with a frequency of 29.4% (**Figure 1B** and **Table 1**). Accordingly, xCas9 can efficiently induce gene mutations at sites with GAA and GAT PAMs in rice.

Considering the ability of xCas9 to edit sites with GAA and GAT PAMs, we speculated that it might also be able to modify GAG and GAC PAM sites in rice. Thus, three target sites with GAG PAM sites in the *OsWaxy* or *OsALS* gene (GAG-1, GAG-2, and GAG-3) (**Supplementary Table 3**) and three sites with GAC PAM sites in the *OsWaxy*, *OsGRF4*, and *OsNRT1.1B* genes (GAC-1, GAC-2, and GAC-3, respectively), were analyzed (**Table 1** and **Supplementary Table 3**). Although xCas9 edited all three targets with GAG PAM sites, with frequencies ranging from 5 to 65%, no mutations were detected at the three GAC PAM sites (**Table 1**). To further evaluate the editing capability of xCas9 at GAC PAM sites, we tested another five target sites (GAC-4 to -8), all of which were not mutated by xCas9 (**Table 1**). These results suggest that xCas9 can efficiently mutate target sites with GAG PAM sites.

These findings indicate that the CRISPR-xCas9 system is useful for expanding the potential editing sites in the rice genome to sequences including a GAD PAM. We subsequently characterized the mutation types generated by xCas9 by analyzing the sequencing chromatograms for the edited target sites. Four mutation types were identified, namely heterozygous, biallelic, homozygous, and chimeric (**Figure 1C**). Only heterozygous mutations were produced at GAT-3, GAG-2, and GAG-3 target sites, whereas two distinct mutation types were produced at GAT-2 and GAA-1 sites (**Figure 1C**). At two target sites that were efficiently edited, GAG-1 (65%) and GAT-1 (75%), three and four mutation types were identified, respectively (**Figures 1B,C**). We detected a high proportion of heterozygous mutations at all target sites. Additionally, the number of mutation types tended to increase as the editing efficiency increased.

Finally, to clarify the effect of the tRNA on the efficiency of the CRISPR-xCas9 system, we constructed a new vector without a tRNA sequence to target all three GAG PAM sites edited by xCas9 relatively efficiently. The resulting system mutated all three target sites less efficiently than the system with tRNA, with decreases in the editing frequency from 65 to 35% at the GAG-1 site, 5 to 0% at the GAG-2 site, and 45 to 38.9% at the GAG-3 site (**Table 1** and **Supplementary Table 4**). These results imply that tRNA may not be required for the CRISPR-xCas9 system, but it increases the editing efficiency of the system.

xCas9 Induces Gene Mutations at NG PAM Sites in Rice Plants

To assess whether the CRISPR-xCas9 system can edit target sites harboring NG PAMs, we tested two NGG, NGA, and NGT PAM sites and three NGC PAM sites (NGC-1, NGC-2, and NGC-3) in the *OsMPK2*, *OsMPK5*, *OsGRF4*, *OsNRT1.1B*, or *OsWaxy* genes (**Table 2** and **Supplementary Table 3**). Analyses of rice T_0 plants confirmed the robust editing activities of xCas9 at target sites with NGG, NGA, and NGT PAMs (**Table 2**). The editing efficiencies exceeded 90% at NGG-1 (95%) and NGT-1 (94.7%) (**Figure 2A** and **Table 2**). The NGA-1 and NGT-2 sites were also edited highly efficiently (∼70%) (**Figure 2A** and **Table 2**). The NGG-2 site was also mutated, albeit less efficiently (45%) (**Table 2**). In contrast, the editing efficiency at NGA-2 was low (26.3%) (**Table 2**) and xCas9 did not modify any of the three NGC PAM sites. Another three NGC PAM sites in *OsMPK5*, *OsMPK2*, and *OsWaxy* were tested (NGC-4, NGC-5, and NGC-6, respectively) (**Table 2** and **Supplementary Table 3**). Only the NGC-4 target site was mutated by xCas9, with an editing efficiency of 30%

TABLE 2 | Summary of the mutation frequencies and mutation types at NG PAM sites targeted by xCas9.

PAM sequence	Target site	Target gene	Target sequence	Tested T_0 plants	Edited T_0 plants	Mutation frequency (%)	Mutation types
NGG	NGG-1	*OsMPK2*	ACACTGCAGCTATTGATATCTGG	20	19	95	1He+10Bi+4Ho+4Chi
	NGG-2	*OsMPK5*	CGACATGATGACGGAGTACGTGG	20	9	45	3He+1Bi+5Chi
NGA	NGA-1	*OsGRF4*	GCATTCTCATCAGCGAGGTCTGA	20	14	70	4He+1Bi+9Chi
	NGA-2	*OsNRT1.1B*	GCTCTACCTGGGGCTCTACCTGA	19	5	26.3	5He
NGT	NGT-1	*OsMPK2*	CAACGCCCGCAGATATGTGAGGT	19	18	94.7	3He+10Bi+5Ho
	NGT-2	*OsMPK5*	CATCCGCTCCAACCAAGAACTGT	19	13	68.4	7He+2Bi+1Ho +3Chi
NGC	NGC-1	*OsMPK5*	TCAGGCCGACGATGACGCACGGC	20	0	0	
	NGC-2	*OsMPK2*	AGACCTCAGGCCAAGTAATTTGC	20	0	0	
	NGC-3	*OsWaxy*	GGCACACTGGCCCACTGGCGAGC	20	0	0	
	NGC-4	*OsMPK5*	AGCCGCCCATCATGCCCATTGGC	20	6	30	5He+1Chi
	NGC-5	*OsMPK2*	CCACCTTCTTCGATCAAACCAGC	20	0	0	
	NGC-6	*OsWaxy*	TCGGCCACCGGCTTCGGCATCGC	20	0	0	

The PAM sequence is underlined. He, heterozygous mutation; Bi, biallelic mutation; Ho, homozygous mutation; Chi, chimeric mutation.

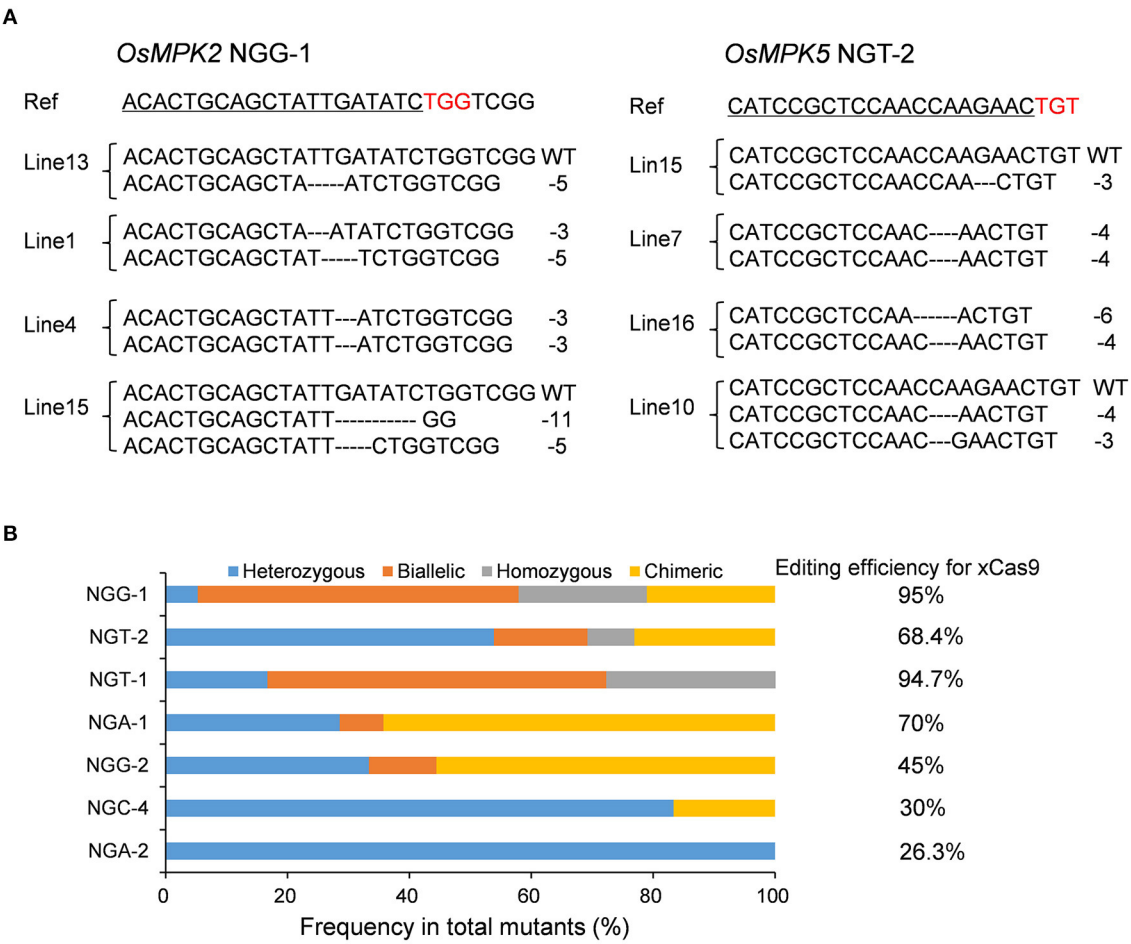


FIGURE 2 | Gene mutagenesis at NG PAM sites induced by xCas9 in rice T_0 plants. **(A)** Mutation types at NGG-1 and NGT-2 target sites of some representative stable mutant lines. Ref sequence is the wild-type sequence. Target sequences are underlined and PAM sequences are indicated in red. **(B)** Frequencies of different mutation types at NG PAM sites mutated by xCas9.

(**Table 2**). Collectively, these results suggest the CRISPR-xCas9 system can induce gene mutations at NGD PAM sites in the rice genome. It can also modify NGC PAM sites, but less efficiently. Moreover, on average, xCas9 can edit NG PAM sites more efficiently than GAD PAM sites.

A subsequent examination identified four mutation types generated by xCas9 at the above edited NG PAM target sites (**Figure 2B**). Only heterozygous mutations were detected at the NGA-2 site (**Figure 2B**), in contrast to the heterozygous and chimeric mutations at the NGC-4 site (**Figure 2B**). Three or four mutation types were detected at the other five sites that were edited relatively efficiently (**Figure 2B**). We observed that heterozygous mutations could be generated at all analyzed target sites. Furthermore, the number of mutation types generally increased as the editing efficiency increased, which was consistent with the results for the GAD PAM sites. However, unlike

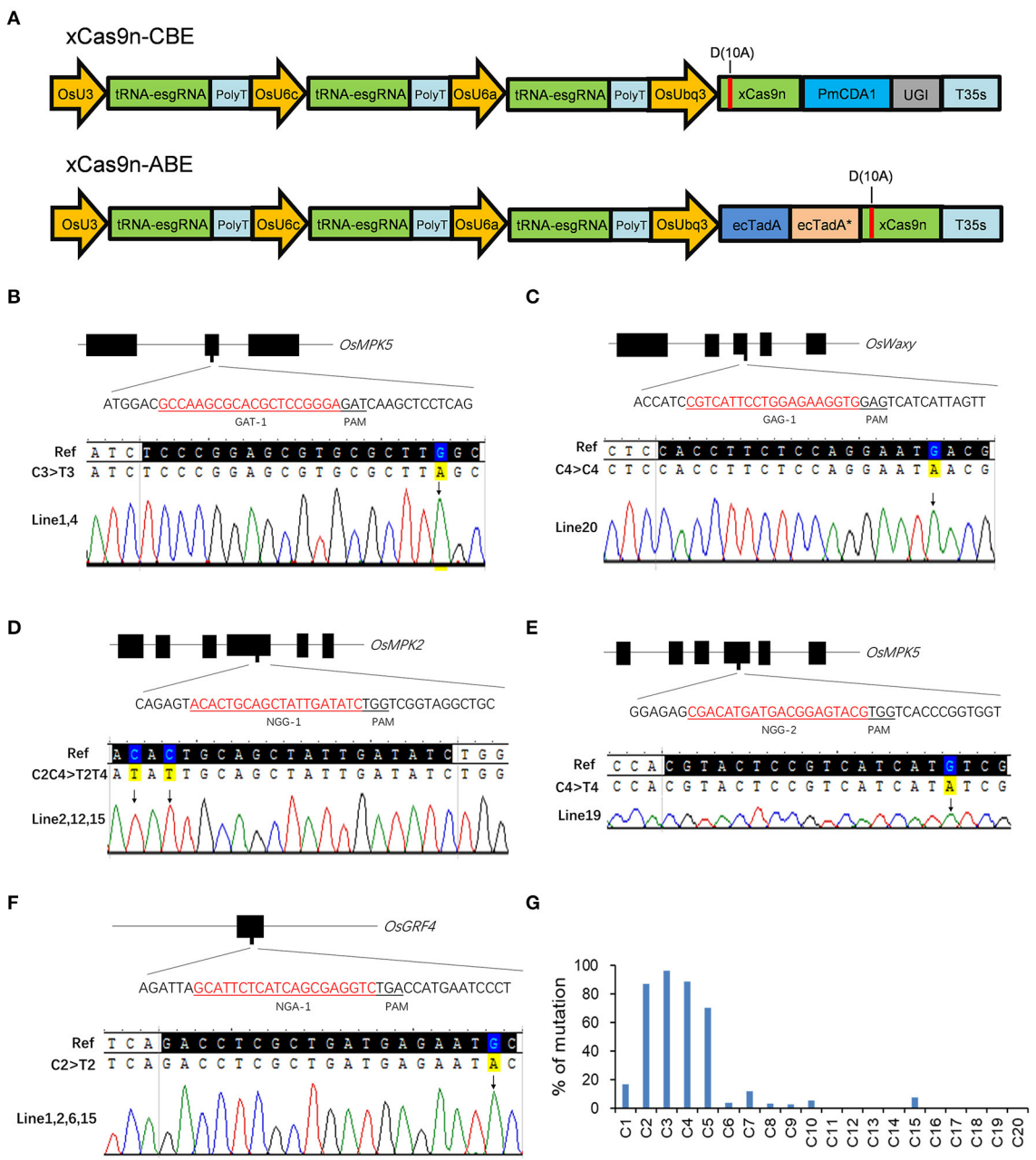


FIGURE 3 | Base editing by xCas9 in rice T_0 plants. **(A)** Schematic view of xCas9n-CBE and xCas9n-ABE. **(B–F)** Sequencing chromatograms of all homozygous mutant lines in which xCas9n-CBE modified the GAT-1 **(B)**, GAG-1 **(C)**, NGG-1 **(D)**, NGG-2 **(E)**, and NGA-1 **(F)** target sites. The mutated bases are marked by arrows. **(G)** Ratio of T_0 plants mutated at a certain position to all genome-edited T_0 plants containing C at that position. The x-axis presents the position of C from the 5′ end of the target.

TABLE 3 | C-to-T mutations induced by xCas9n-CBE in rice T_0 plants.

PAM sequence	Target site	Tested T_0 plants	Edited T_0 plants	C-to-T frequency (%)	Genotypes	Homozygous
GAA	GAA-1	12	4	33.3	C2>T2(1);C2C3>T2T3(1); C1C2C3>T1T2T3(2)	0
	GAA-2	12	0	0		
	GAA-3	12	0	0		
GAT	GAT-1	19	9	47.4	C3>T3(3);C2C3>T2T3(6)	2
	GAT-2	20	0	0		
	GAT-3	20	0	0		
GAG	GAG-1	20	14	70	C4>T4(10);C1C4>T1T4(3); C4C9>T4T9(1)	2
	GAG-2	20	2	10	C15>T15(2)	0
	GAG-3	20	4	20	C2>T2(2);C2C5>T2T5(2)	0
NGG	NGG-1	19	13	68.4	C2>T2(1);C2C4>T2T4(10); C2C4C7>T2T4T7(2)	3
	NGG-2	20	5	25	C4>T4(5)	1
NGA	NGA-1	20	13	65	C2>T2(12);C2C6C8>T2T6T8(1)	4
	NGA-2	19	1	5.3	C4>T4(1)	0
NGT	NGT-1	18	12	66.7	C1>T1(1);C4>T4(5);C7>T7(2); C1C4>T1T4(1);C4C7>T4T7(1); C6C7>T6T7(1); C1C7C10>T1T7T10(1)	0
	NGT-2	20	10	50	C5>T5(1);C4C5>T4T5(8); C4C5C10>T4T5T10(1)	0
NGC	NGC-4	27	14	51.9	C3C4>T3T4(13);C3C4C8>T3T4T8(1)	0
	NGC-5	29	0	0		
	NGC-6	30	0	0		
GAC	GAC-1	19	2	10.5	C3>T3(1);C3C4>T3T4(1)	0
	GAC-2	20	0	0		
	GAC-3	20	1	5	C2>T2(1)	0

the GAD PAM sites, heterozygous mutations were not the predominant mutation at all target sites. The ratio of chimeric or biallelic mutations was relatively high, and these mutations accounted for a large proportion of the modifications at the NGG-2, NGA-1, NGT-1, and NGG-1 sites (**Figure 2B**).

Base Editing by xCas9 in Rice

On the basis of the efficient editing of the GAD and NG PAMs by the CRISPR-xCas9 system, we decided to develop a new xCas9-based CBE and ABE using vectors that were structurally similar to those used to produce the CRISPR-xCas9 system for editing the bases at the same PAM sites in rice. First, xCas9 was mutated to xCas9(D10A) nickase (xCas9n) via PCR and then fused to *Petromyzon marinus* cytidine deaminase1 (PmCDA1) and uracil DNA glycosylase inhibitor (UGI) to generate xCas9n-CBE (**Figure 3A**). The 18 target sites with GAA, GAT, GAG, NGG, NGA, NGT, and NGC sequences used for testing the CRISPR-xCas9 system were selected to evaluate the C-to-T base editing activities of xCas9n-CBE (**Table 3**).

As expected, xCas9n-CBE efficiently mediated the C-to-T base conversion at the tested target sites (**Table 3**). An examination of the T_0 plants indicated that of the three GAA PAM sites, GAA-1 was edited, with a frequency of 33.3%, whereas GAA-2 and GAA-3 were not modified (**Table 3**), which was consistent with the editing results for the CRISPR-xCas9 system. Among the three target sites with the GAT PAM, only GAT-1 was edited by xCas9n-CBE, with a mutation rate of 47.4% (**Figure 3B** and **Table 3**). Base mutations were undetectable at GAT-2 and GAT-3, although Indel mutations were induced by xCas9 (**Tables 1**, **3**). Additionally, xCas9n-CBE edited all three targets with GAG PAMs, with frequencies ranging from 10 to 70% (**Figure 3C** and **Table 3**). For the nine NG PAM sites, xCas9n-CBE induced C-to-T substitutions at all target sites that were edited by the CRISPR-xCas9 system (**Figures 3D–F**, **Tables 1**, **3**).

These results imply that xCas9n-CBE and the CRISPR-xCas9 system are two distinct editing systems, with some differences in the editable target sites. Therefore, although xCas9 inefficiently edits GAC PAM sites, because we previously confirmed that the tRNA-esgRNA system (xCas9n-epBE) can efficiently edit these PAM sites (Zhang et al., 2020), we chose the same target sites (GAC-1, GAC-2, and GAC-3) to determine whether xCas9n-CBE can modify GAC PAM sites. Surprisingly, base mutations were detected at GAC-1 and GAC-3, with frequencies of 10.5 and 5%, respectively, whereas no base mutations were detected at GAC-2 (**Table 3**). These findings suggest that xCas9n-CBE can efficiently mediate the C-to-T base conversion even at target sites with GAC PAMs, albeit with a relatively low editing efficiency.

Considered together, our results indicate that xCas9n-CBE induces the C-to-T base conversion at GA and NG PAM sites in rice. Thus, it can target a wider range of sequences than the CRISPR-xCas9 system, which inefficiently edits GAC PAM sites. To further characterize xCas9n-CBE regarding its editing window, editing preference, and mutation types, all 104 genome-edited T_0 plants for the above-mentioned 14 edited target sites were analyzed together. The editing window typically spanned positions 1 to 10 within the protospacer and occasionally extended to position 15 (**Figure 3G**). In the editing window, C2, C3, C4, and C5 were edited more frequently than C1 and C7 (C3 > C2 ≈ C4 > C5 >> C1 ≈ C7), whereas C6, C8, C9, C10, and C15 were mutated relatively infrequently (< 10%) (**Figure 3G**). In the T_0 plants, single- and double-base substitutions were the predominant mutations at the edited targets. Single C-to-T mutations were detected at almost all edited sites, and no other types of base substitutions were detected at three target sites with a mutation frequency of 10% or less (**Table 3**). Almost all of the triple-base substitutions were detected at targets with an editing efficiency of 50% or more (**Figure 3G**). Furthermore, five of the 14 edited target sites had homozygous mutations (**Figures 3B–F**, **Table 3**).

We also fused the wild-type adenine deaminase ecTadA and its variant ecTadA* to the N-terminus of xCas9n to generate xCas9n-ABE in rice (**Figure 3A**). All 21 target sites used to test xCas9n-CBE were also used for assessing the editing capability of xCas9n-ABE. Unfortunately, the sequencing results revealed a lack of A-to-G mutations in the T_0 rice plants (data not shown).

DISCUSSION

In this study, the utility of xCas9 for targeted gene mutation was thoroughly investigated by analyzing its editing activities at several endogenous target sites with GA and NG PAMs in rice. The analyses revealed that xCas9 can recognize a broad range of PAM sequences, including NG, GAA, and GAT, which is consistent with the results of a recent study involving human cells (Hu et al., 2018). Additionally, xCas9 can also induce Indel mutations at GAG PAM sites, which have not been observed in human cells. Although there have been many studies on the editing activities of xCas9 in rice, no edited target sites with GAA and GAG PAM sequences have been detected in T_0 plants (Hua et al., 2019; Li et al., 2019; Ren et al., 2019; Wang et al., 2019a; Zhong et al., 2019; Zeng et al., 2020). Therefore, the CRISPR-xCas9 system developed in the current study will be useful for expanding the range of potential gene mutations by targeting these PAM sites in rice.

We used tRNA and esgRNA in our CRISPR-xCas9 system. We determined that tRNA increased the editing efficiency of xCas9 (**Table 1** and **Supplementary Table 4**), which was consistent with the results of earlier studies in which the tRNA-sgRNA system enhanced the activity of high-fidelity Cas9 variants in human cells and rice (Zhang et al., 2017; He et al., 2019). Our previous research proved that the tRNA-esgRNA system enables xCas9n-epBE to efficiently edit GA and NG PAM sites (Zhang et al., 2020). We speculated that tRNA-esgRNA is also important for facilitating efficient gene mutations via the CRISPR-xCas9 system.

The xCas9n-CBE system described herein differs from our previously reported xCas9n-epBE system in terms of the associated vector architecture (Zhang et al., 2020). They belong to different multiplex editing systems. Three independent tRNA-esgRNAs under the control of three different promoters were employed in xCas9n-CBE, whereas one or more tRNA-esgRNAs in xCas9n-epBE were controlled by a common promoter. Both xCas9n-CBE and xCas9n-epBE can efficiently modify bases at GA and NGD (where D is G, T, or A) PAM sites. However, xCas9n-CBE can efficiently edit NGC PAM sites, whereas xCas9n-epBE cannot. Additionally, on average, xCas9n-CBE can edit NG PAM sites more efficiently than xCas9n-epBE (**Table 2**). Moreover, xCas9n-CBE has a wider editing window than xCas9n-epBE because it can target C8, C9, and C15 (**Figure 3G**). Thus, xCas9n-CBE might be a better alternative to xCas9n-epBE.

Regrettably, the adenine base editor (xCas9n-ABE) developed in our study failed to modify the tested target sites, which was in contrast to the results of a recent study involving human cells (Hu et al., 2018). Although Hua et al. (2019) identified one NGT PAM site edited by ABE-P6, the editing efficiency was quite low (4.8%) and none of the other five tested PAM sites were edited. Additionally, no target sites were edited by the xCas9-based ABE in rice in another previous study (Zeng et al., 2020). Considering that the editing efficiency of xCas9 at the same target sites contrasted with the relative lack of editing by the xCas9-based ABE, the editing activity of xCas9n-ABE in rice will need to be improved in future studies.

Because some of the target sites were edited relatively inefficiently by the CRISPR-xCas9 system or xCas9n-CBE, further research applying highly efficient nuclear localization signals and surrogate systems is required to increase the editing efficiency (Wang et al., 2019b, 2020; Xu et al., 2020). Only one of six NGC PAM sites was successfully mutated by xCas9 (**Table 2**). This is similar to the results reported by Hua et al. (2019). Moreover, SpCas9-NG also had limited activity at NGC PAM sites in human cells and rice (Nishimasu et al., 2018; Ren et al., 2019). Accordingly, the editing efficiency at NGC PAM sites will need to be increased. Three new SpCas9 variants, SpCas9-NRRH, SpCas9-NRTH, and SpCas9-NRCH, were recently reported to recognize non-G PAMs in human cells (Miller et al., 2020). Furthermore, a near PAM-less SpCas9 variant (SpRY) was also developed (Walton et al., 2020). The utility of these new Cas9 variants should be tested in rice to further expand the genome editing scope.

CONCLUSION

In this study, we developed an efficient CRISPR-xCas9 system that can expand the potential target genome sequences to include GAD PAM sites in rice. It can also mutate genes at NG PAM sites. We also developed xCas9n-CBE with a similar vector architecture. The efficient base-editing activities of xCas9n-CBE at GA and NG PAM sites were confirmed, with the main deamination window comprising protospacer positions 1 to 10.

These new genome engineering tools will be useful for future basic rice research and crop improvement.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

JY and CZ designed the experiments. XL, SZ, SY, LL, YY, and FW performed all the experiments. CZ, GK, and XZ analyzed the results. CZ and JY wrote the manuscript. JY supervised the project. All authors read and approved the final manuscript.

FUNDING

This work was supported by Beijing Academy of Agriculture & Forestry Sciences. Funding was provided by the Beijing Municipal Natural Science Foundation (6204041), the Youth Research Fund of Beijing Academy of Agriculture and Forestry Sciences (QNJJ202028), and the Beijing Scholars Program (BSP041).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgeed.2020.618385/full#supplementary-material

REFERENCES

Bortesi, L., and Fischer, R. (2015). The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnol. Adv.* 33, 41–52. doi: 10.1016/j.biotechadv.2014.12.006

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819–823. doi: 10.1126/science.1231143

Endo, A., Masafumi, M., Kaya, H., and Toki, S. (2016). Efficient targeted mutagenesis of rice and tobacco genomes using Cpf1 from *Francisella novicida*. *Sci. Rep.* 6:38169. doi: 10.1038/srep38169

Esvelt, K. M., Mali, P., Braff, J. L., Moosburner, M., Yaung, S. J., and Church, G. M. (2013). Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods* 10, 1116–1121. doi: 10.1038/nmeth.2681

Ge, Z., Zheng, L., Zhao, Y., Jiang, J., Zhang, E. J., Liu, T., et al. (2019). Engineered xCas9 and SpCas9-NG variants broaden PAM recognition sites to generate mutations in Arabidopsis plants. *Plant Biotechnol. J.* 17, 1865–1867. doi: 10.1111/pbi.13148

Harrington, L. B., Paez-Espino, D., Staahl, B. T., Chen, J. S., Ma, E., Kyrpides, N. C., et al. (2017). A thermostable Cas9 with increased lifetime in human plasma. *Nat. Commun.* 8:1424. doi: 10.1038/s41467-017-01408-4

He, X., Wang, Y., Yang, F., Wang, B., Xie, H., Gu, L., et al. (2019). Boosting activity of high-fidelity CRISPR/Cas9 variants using a $tRNA^{Gln}$-processing system in human cells. *J. Biol. Chem.* 294, 9308–9315. doi: 10.1074/jbc.RA119.007791

Hiei, Y., and Komari, T. (2008). Agrobacterium-mediated transformation of rice using immature embryos or calli induced from mature seed. *Nat. Protoc.* 3, 824–834. doi: 10.1038/nprot.2008.46

Hille, F., Richter, H., Wong, S. P., Bratovic, M., Ressel, S., and Charpentier, E. (2018). The biology of CRISPR-Cas: backward and forward. *Cell* 172, 1239–1259. doi: 10.1016/j.cell.2017.11.032

Hu, J. H., Miller, S. M., Geurts, M. H., Tang, W., Chen, L., Sun, N., et al. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556, 57–63. doi: 10.1038/nature26155

Hua, K., Tao, X., Han, P., Wang, R., and Zhu, J. K. (2019). Genome engineering in rice using Cas9 variants that recognize NG PAM sequences. *Mol. Plant* 12, 1003–1014. doi: 10.1016/j.molp.2019.03.009

Kaya, H., Mikami, M., Endo, A., Endo, M., and Toki, S. (2016). Highly specific targeted mutagenesis in plants using *Staphylococcus aureus* Cas9. *Sci. Rep.* 6:26871. doi: 10.1038/srep26871

Kim, E., Koo, T., Park, S. W., Kim, D., Kim, K., Cho, H. Y., et al. (2017). *In vivo* genome editing with a small Cas9 orthologue derived from *Campylobacter jejuni*. *Nat. Commun.* 8:14500. doi: 10.1038/ncomms14500

Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Topkar, V. V., Zheng, Z., et al. (2015a). Broadening the targeting range of staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition. *Nat. Biotechnol.* 33, 1293–1298. doi: 10.1038/nbt.3404

Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Topkar, V. V., Nguyen, N. T., Zheng, Z., et al. (2015b). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* 523, 481–485. doi: 10.1038/nature14592

Komor, A. C., Badran, A. H., and Liu, D. R. (2017). CRISPR-based technologies for the manipulation of eukaryotic genomes. *Cell* 168, 20–36. doi: 10.1016/j.cell.2016.10.044

Li, J., Luo, J., Xu, M., Li, S., Zhang, J., Li, H., et al. (2019). Plant genome editing using xCas9 with expanded PAM compatibility. *J. Genet. Genomics* 46, 277–280. doi: 10.1016/j.jgg.2019.03.004

Li, J., Sun, Y., Du, J., Zhao, Y., and Xia, L. (2017). Generation of targeted point mutations in rice by a modified CRISPR/Cas9 system. *Mol. Plant* 10, 526–529. doi: 10.1016/j.molp.2016.12.001

Liu, W., Xie, X., Ma, X., Li, J., Chen, J., and Liu, Y. G. (2015). DSDecode: a web-based tool for decoding of sequencing chromatograms for genotyping of targeted mutations. *Mol. Plant* 8, 1431–1433. doi: 10.1016/j.molp.2015.05.009

Lowder, L. G., Zhang, D., Baltes, N. J., Paul, J. W. III, Tang, X., Zheng, X., et al. (2015). A CRISPR/Cas9 toolbox for multiplexed plant genome editing and transcriptional regulation. *Plant Physiol.* 169, 971–985. doi: 10.1104/pp.15.00636

Lu, Y., and Zhu, J. K. (2017). Precise editing of a target base in the rice genome using a modified CRISPR/Cas9 system. *Mol. Plant* 10, 523–525. doi: 10.1016/j.molp.2016.11.013

Ma, X., Zhang, Q., Zhu, Q., Liu, W., Chen, Y., Qiu, R., 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

Miller, S. M., Wang, T., Randolph, P. B., Arbab, M., Shen, M. W., Huang, T. P., et al. (2020). Continuous evolution of SpCas9 variants compatible with non-G PAMs. *Nat. Biotechnol.* 38, 471–481. doi: 10.1038/s41587-020-0412-8

Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155, 733–740. doi: 10.1099/mic.0.023960-0

Mougiakos, I., Mohanraju, P., Bosma, E. F., Vrouwe, V., Finger Bou, M., Naduthodi, M. I. S., et al. (2017). Characterizing a thermostable Cas9 for bacterial genome editing and silencing. *Nat. Commun.* 8:1647. doi: 10.1038/s41467-017-01591-4

Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., et al. (2018). Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* 361, 1259–1262. doi: 10.1126/science.aas9129

Piatek, A., Ali, Z., Baazim, H., Li, L., Abulfaraj, A., Al-Shareef, S., et al. (2015). RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. *Plant Biotechnol. J.* 13, 578–589. doi: 10.1111/pbi.12284

Qin, R., Li, J., Li, H., Zhang, Y., Liu, X., Miao, Y., et al. (2019). Developing a highly efficient and wildly adaptive CRISPR-SaCas9 toolset for plant genome editing. *Plant Biotechnol. J.* 17, 706–708. doi: 10.1111/pbi.13047

Ran, F. A., Cong, L., Yan, W. X., Scott, D. A., Gootenberg, J. S., Kriz, A. J., et al. (2015). *In vivo* genome editing using *Staphylococcus aureus* Cas9. *Nature* 520, 186–191. doi: 10.1038/nature14299

Ren, B., Liu, L., Li, S., Kuang, Y., Wang, J., Zhang, D., et al. (2019). Cas9-NG greatly expands the targeting scope of genome-editing toolkit by recognizing NG and other atypical PAMs in rice. *Mol. Plant* 12, 1015–1026. doi: 10.1016/j.molp.2019.03.010

Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C., and Doudna, J. A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* 507, 62–67. doi: 10.1038/nature13011

Tang, X., Lowder, L. G., Zhang, T., Malzahn, A. A., Zheng, X., Voytas, D. F., et al. (2017). A CRISPR-Cpf1 system for efficient genome editing and transcriptional repression in plants. *Nat. Plants* 3:17018. doi: 10.1038/nplants.2017.18

Teng, F., Cui, T., Feng, G., Guo, L., Xu, K., Gao, Q., et al. (2018). Repurposing CRISPR-Cas12b for mammalian genome engineering. *Cell Discov.* 4:63. doi: 10.1038/s41421-018-0069-3

Walton, R. T., Christie, K. A., Whittaker, M. N., and Kleinstiver, B. P. (2020). Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* 368, 290–296. doi: 10.1126/science.aba8853

Wang, F., Zhang, C., Xu, W., Yuan, S., Song, J., Li, L., et al. (2020). Developing high-efficiency base editors by combining optimized synergistic core components with new types of nuclear localization signal peptide. *Crop J.* 8, 408–417. doi: 10.1016/j.cj.2020.01.003

Wang, J., Meng, X., Hu, X., Sun, T., Li, J., Wang, K., et al. (2019a). xCas9 expands the scope of genome editing with reduced efficiency in rice. *Plant Biotechnol. J.* 17, 709–711. doi: 10.1111/pbi.13053

Wang, M., Wang, Z., Mao, Y., Lu, Y., Yang, R., Tao, X., et al. (2019b). Optimizing base editors for improved efficiency and expanded editing scope in rice. *Plant Biotechnol. J.* 17, 1697–1699. doi: 10.1111/pbi.13124

Wu, Y., Xu, W., Wang, F., Zhao, S., Feng, F., Song, J., et al. (2019). Increasing cytosine base editing scope and efficiency with engineered Cas9-PmCDA1 fusions and the modified sgRNA in rice. *Front. Genet.* 10:379. doi: 10.3389/fgene.2019.00379

Xie, K., Minkenberg, B., and Yang, Y. (2015). Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3570–3575. doi: 10.1073/pnas.1420294112

Xie, X., Ma, X., Zhu, Q., Zeng, D., Li, G., and Liu, Y. G. (2017). CRISPR-GE: a convenient software toolkit for CRISPR-based genome editing. *Mol. Plant* 10, 1246–1249. doi: 10.1016/j.molp.2017.06.004

Xu, K., Ren, C., Liu, Z., Zhang, T., Zhang, T., Li, D., et al. (2015). Efficient genome engineering in eukaryotes using Cas9 from *Streptococcus thermophilus*. *Cell Mol. Life Sci.* 72, 383–399. doi: 10.1007/s00018-014-1679-z

Xu, W., Yang, Y., Liu, Y., Kang, G., Wang, F., Li, L., et al. (2020). Discriminated sgRNAs-based SurroGate system greatly enhances the screening efficiency of plant base-edited cells. *Mol. Plant* 13, 169–180. doi: 10.1016/j.molp.2019.10.007

Yin, K., Gao, C., and Qiu, J. L. (2017). Progress and prospects in plant genome editing. *Nat. Plants* 3:17107. doi: 10.1038/nplants.2017.107

Zeng, D., Li, X., Huang, J., Li, Y., Cai, S., Yu, W., et al. (2020). Engineered Cas9 variant tools expand targeting scope of genome and base editing in rice. *Plant Biotechnol. J.* 18, 1348–1350. doi: 10.1111/pbi.13293

Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-cas system. *Cell* 163, 759–771. doi: 10.1016/j.cell.2015.09.038

Zhang, C., Xu, W., Wang, F., Kang, G., Yuan, S., Lv, X., et al. (2020). Expanding the base editing scope to GA and relaxed NG PAM sites by improved xCas9 system. *Plant Biotechnol. J.* 18, 884–886. doi: 10.1111/pbi.13259

Zhang, D., Zhang, H., Li, T., Chen, K., Qiu, J. L., and Gao, C. (2017). Perfectly matched 20-nucleotide guide RNA sequences enable robust genome editing using high-fidelity SpCas9 nucleases. *Genome Biol.* 18:191. doi: 10.1186/s13059-017-1325-9

Zhong, Z., Sretenovic, S., Ren, Q., Yang, L., Bao, Y., Qi, C., et al. (2019). Improving plant genome editing with high-fidelity xCas9 and noncanonical PAM-targeting Cas9-NG. *Mol. Plant* 12, 1027–1036. doi: 10.1016/j.molp.2019.03.011

Zong, Y., Wang, Y., Li, C., Zhang, R., Chen, K., Ran, Y., et al. (2017). Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. *Nat. Biotechnol.* 35, 438–440. doi: 10.1038/nbt.3811

Conflict of Interest: The authors submitted patent applications based on the results reported in this paper.

Copyright © 2020 Zhang, Kang, Liu, Zhao, Yuan, Li, Yang, Wang, Zhang and Yang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genome Editing

ORIGINAL RESEARCH
published: 21 December 2020
doi: 10.3389/fgeed.2020.604876

Establishment of CRISPR/Cas9 Genome Editing in Witloof (*Cichorium intybus* var. *foliosum*)

Charlotte De Bruyn[1,2,3]*, Tom Ruttink[1], Tom Eeckhaut[1], Thomas Jacobs[2,3], Ellen De Keyser[1], Alain Goossens[2,3] and Katrijn Van Laere[1]

[1] Plant Sciences Unit, Flanders Research Institute for Agriculture, Fisheries and Food (ILVO), Melle, Belgium, [2] Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium, [3] Center for Plant Systems Biology, Flanders Institute for Biotechnology (VIB), Ghent, Belgium

Cichorium intybus var. *foliosum* (witloof) is an economically important crop with a high nutritional value thanks to many specialized metabolites, such as polyphenols and terpenoids. However, witloof plants are rich in sesquiterpene lactones (SL) which are important for plant defense but also impart a bitter taste, thus limiting industrial applications. Inactivating specific genes in the SL biosynthesis pathway could lead to changes in the SL metabolite content and result in altered bitterness. In this study, a CRISPR/Cas9 genome editing workflow was implemented for witloof, starting with polyethylene glycol (PEG) mediated protoplast transfection for CRISPR/Cas9 vector delivery, followed by whole plant regeneration and mutation analysis. Protoplast transfection efficiencies ranged from 20 to 26 %. A CRISPR/Cas9 vector targeting the first exon of the *phytoene desaturase* (*CiPDS*) gene was transfected into witloof protoplasts and resulted in the knockout of *CiPDS*, giving rise to an albino phenotype in 23% of the regenerated plants. Further implementing our protocol, the SL biosynthesis pathway genes *germacrene A synthase (GAS), germacrene A oxidase (GAO)*, and *costunolide synthase (COS)* were targeted in independent experiments. Highly multiplex (HiPlex) amplicon sequencing of the genomic target loci revealed plant mutation frequencies of 27.3, 42.7, and 98.3% in regenerated plants transfected with a CRISPR/Cas9 vector targeting *CiGAS*, *CiGAO*, and *CiCOS*, respectively. We observed different mutation spectra across the loci, ranging from consistently the same +1 nucleotide insertion in *CiCOS* across independent mutated lines, to a complex set of 20 mutation types in *CiGAO* across independent mutated lines. These results demonstrate a straightforward workflow for genome editing based on transfection and regeneration of witloof protoplasts and subsequent HiPlex amplicon sequencing. Our CRISPR/Cas9 workflow can enable gene functional research and faster incorporation of novel traits in elite witloof lines in the future, thus facilitating the development of novel industrial applications for witloof.

Keywords: gene editing, protoplast transfection, *Cichorium intybus*, sesquiterpene lactones, phytoene desaturase, HiPlex amplicon sequencing

OPEN ACCESS

Edited by:
Lanqin Xia, Chinese Academy of Agricultural Sciences, China

Reviewed by:
Bastian Minkenberg, University of California, Berkeley, United States
Chidananda Nagamangala Kanchiswamy, Fondazione Edmund Mach, Italy

***Correspondence:**
Charlotte De Bruyn
charlotte.debruyn@ilvo.vlaanderen.be

Specialty section:
This article was submitted to Genome Editing in Plants, a section of the journal Frontiers in Genome Editing

Received: *10 September 2020*
Accepted: *24 November 2020*
Published: *21 December 2020*

Citation:
De Bruyn C, Ruttink T, Eeckhaut T, Jacobs T, De Keyser E, Goossens A and Van Laere K (2020) Establishment of CRISPR/Cas9 Genome Editing in Witloof (Cichorium intybus var. foliosum). Front. Genome Ed. 2:604876. doi: 10.3389/fgeed.2020.604876

INTRODUCTION

Witloof belongs to the species *Cichorium intybus*, a member of the Asteraceae family. Within the *Cichorium* genus, three groups are distinguished: root chicory, leaf chicory and Witloof (Raulier et al., 2015). Root chicory, also known as industrial chicory (*C. intybus* var. *sativum*), is characterized by a large tap root and is mainly grown for inulin production. The leaf chicory group can be divided into three subgroups: Sugarloaf (*C. intybus* var. *porphyreum*), Radicchio (*C. intybus* var. *latifolium*), and Catalogne (*C. intybus* var. *intybus* or *C. intybus* var. *sylvestre*); they consist of leafy vegetables that can be consumed fresh or cooked. Witloof, also referred to as Witlof, Belgian endive, Chicon, or *C. intybus* var. *foliosum*, is a biennial plant that, without forcing conditions, produces a taproot and a rosette of leaves in the first year of growth and, following a period of cold exposure, develops a floral meristem in the second year of growth. By cultivating the tap root under artificial conditions, a vegetable composed of tightly packed white leaves is produced, known as witloof. It is a traditional Belgian crop rich in nutritionally relevant compounds, such as polyphenols and terpenoids, which have positive effects on human health because of the biological and pharmacological activities of their specialized metabolites (Street et al., 2013). Consequently, it is a crop with an important economic value, which has led to large scale cultivation (2.100 ha annually in Belgium). However, the bitter taste limits the use of witloof as it has a negative influence on consumer acceptability (Drewnowski and Gomez-Carneros, 2000). Modifying the bitterness in witloof can lead to product differentiation by creating a more diverse range of flavors enabling consumers to choose between less or more bitter cultivars, maximizing the acceptance and economic impact and creating new market opportunities.

Sesquiterpene lactones (SLs) are the specialized metabolites responsible for the bitter taste and play an important role in plant defense against herbivores and pathogens (Peters et al., 1997). Within the SLs, different classes are recognized, with the guaianolides being the most important with regard to bitterness (de Kraker et al., 1998). This class comprises lactucin, deoxylactucin, lactucopicrin, and their derivatives (Chadwick et al., 2013). They all originate from the cytosolic mevalonate pathway that leads to the production of the building blocks isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP). When DMAPP reacts with two units of IPP, farnesyl diphosphate (FDP) is formed, which is further converted to costunolide via the enzymes germacrene A synthase (GAS), germacrene A oxidase (GAO), and costunolide synthase (COS) (de Kraker et al., 1998; Liu et al., 2011). The genes encoding GAS, GAO and COS have previously been cloned from chicory (Bouwmeester, 2002; Liu et al., 2011) and are of significant importance to control the biosynthesis of SLs and hence the bitterness of witloof.

In traditional witloof breeding, the stable integration of specific traits in elite lines takes a minimum of 10 years, making breeding programs to improve yield and nutritional properties time-consuming and labor intensive. CRISPR/Cas9 [clustered regularly interspaced short palindromic repeat (CRISPR) associated nuclease (CAS)] genome editing, which has recently been applied on a number of plant species (Manghwar et al., 2019; Zhang et al., 2019), can be a valuable tool to study the function of genes involved in specific specialized metabolite biosynthetic pathways, such as the SL pathway, and thus identify the genomic target loci underlying crop improvement. Moreover, CRISPR/Cas9 genome editing may offer more straightforward breeding opportunities to generate new varieties within a shorter time period by targeted introduction of functional sequence diversity in elite lines (Zhang et al., 2019). The CRISPR/Cas9 system enables the alteration of specific DNA sequences to achieve gene modifications. The Cas9 endonuclease uses a guide RNA (gRNA) with a 20 nucleotide spacer sequence to recognize a complementary target DNA site (the protospacer) upstream of a protospacer adjacent motif (PAM) in the genomic DNA. Upon recognition, Cas9 generates a double-stranded DNA break (DSB) between the 3rd and 4th nucleotides on the 5′ side of the PAM (Jinek et al., 2012). These DSBs are typically repaired through non-homologous end joining (NHEJ), whereby imperfect repair results in small insertions/deletions (indels) and/or substitutions (SNPs) at the target region (Manghwar et al., 2019). Such mutations in protein coding sequences may result in premature stop codons downstream of the indel, inducing the elimination of the protein product. Changes in the regulation of gene expression or the activity of the encoded protein are also possible.

The CRISPR/Cas9 system is often delivered through a vector into plant cells. Hereby, two main delivery methods are commonly used: *Agrobacterium*-mediated transformation and protoplast transfection. Using *Agrobacterium* transformation, the CRISPR/Cas9 system is typically stably integrated into the plant genome, whereas protoplast transfection allows for transient expression of the CRISPR/Cas9 system (Zhang et al., 2019). As protoplasts are cells without a cell wall, DNA uptake can readily occur through the plasma membrane using polyethylene glycol (PEG) (Zhang et al., 2019). As a result, a high quantity (>50,000) of cells can be simultaneously transfected and plants can be regenerated from single cells (Jaganathan et al., 2018). Furthermore, the chance of stable integration of vector DNA into the genome is reduced and off-target effects are decreased (Zhang et al., 2016). Working with protoplasts also allows the implementation of a DNA-free transfection method using pre-assembled ribonucleoprotein complexes (RNPs) instead of vector DNA (Wook Woo et al., 2015). Because the delivery vectors are not stably integrated into the plant genome, using a transient transfection method does not allow the use of typical plant selection markers (e.g., kanamycin, hygromycin, *bar* gene). Hence, high protoplast transfection and mutation efficiencies are required. Isolating and, especially, regenerating protoplasts can be very challenging, thus hampering the strategy in many genera and species. However, a successful protoplast isolation and regeneration method has previously been developed for *Cichorium* (Deryckere et al., 2012). Recently, protoplasts of *C. intybus* have been successfully transfected and mutated in the *phytoene desaturase* (*CiPDS*) gene, using a CRISPR/Cas9 PEG-mediated transfection protocol, resulting

in 4.5% of the regenerated plants with an albino phenotype (Bernard et al., 2019).

In this study, we report an efficient CRISPR/Cas9 genome editing workflow for witloof based on PEG-mediated protoplast transfection, transient expression of CRISPR/Cas9 vectors and whole plant regeneration. We first developed a protoplast transfection and regeneration protocol using a CRISPR/Cas9 vector to induce mutations in the *CiPDS* gene, leading to the regeneration of several independent albino plantlets. To further implement the genome editing technique, we used our CRISPR/Cas9 protocol to induce mutations in *CiGAS, CiGAO*, and *CiCOS*, three genes known to be involved in the SL biosynthesis pathway. Highly multiplex (HiPlex) amplicon sequencing was used to analyze the genomic target loci in the regenerated plants and revealed a variety of mutated alleles and targeted knockouts, indicating the potential of our CRISPR/Cas9 workflow.

MATERIALS AND METHODS

Plant Material

Plant materials of witloof *C. intybus* var. *foliosum* "Van Hamme" and "Topmodel" were provided by COSUCRA (Belgium). Roots of *in vivo* plants of the selected *Cichorium* varieties were rinsed with water, grated on the outside and cut into slices of 1 cm. The slices were rinsed for 1 min in 70% ethanol, sterilized in 2.5% NaOCl, and rinsed in sterile water. Next, the slices were cut into pieces of 1–2 cm^3 and transferred to solid plant medium [4.4 $g.L^{-1}$ Murashige and Skoog medium (Murashige and Skoog, 1962), 45 $g.L^{-1}$ sucrose, 8 $g.L^{-1}$ plant tissue culture agar No. 4 (Neogen, Lansing, Michigan, United States), pH 6] at 23 ± 2°C under a 16/8 h (light/dark) photoperiod at 40 $\mu mol.m^{-2}.s^{-1}$ photosynthetic active radiation. After shoot induction, plants were transferred to new solid plant medium (4.4 $g.L^{-1}$ Murashige and Skoog medium + vitamins, 20 $g.L^{-1}$ sucrose, 7 $g.L^{-1}$ plant tissue culture agar No. 4, pH 6) and subcultured every 6 weeks.

CRISPR/Cas9 Vector Construction

Guide RNAs for *CiPDS* (MK455771), *CiGAS* (AF498000.1), *CiGAO* (ADF43080), and *CiCOS* (G3GBK0) were designed in the first half of the CDS using Geneious 10.2.6 (http://www.geneious.com) and were selected based on high on-target activity scores (Doench et al., 2014).

An overview of de vector constructions can be found in **Supplementary Figure 1**. A first step in the construction of a *Cas9* destination vector, was to "domesticate" a common high-copy entry vector by removing a *Bbs*I restriction site in the backbone. Hence, digestion of pEN-L1-AG-L2 (Houbaert et al., 2018) with *Apa*I and *Bbs*I was followed by ligation with a double-stranded linker (**Supplementary Table 1**) with a mutated *Bbs*I site, generating the vector pEN-L1-AG-L2,\ (-Bbs1). In a second step, six entry clones pGGA004, pGGB003, pGGD002, pGGE001 (Lampropoulos et al., 2013), pGG-C-Cas9PTA*-D, pGG-F-AtU6-26-BbsI-BbsI-G (Decaestecker et al., 2019), and annealed *Bsa*I oligos 9 and 10 (**Supplementary Table 1**) were assembled into pEN-L1-AG-L2,\(-Bbs1), generating the Cas9 destination vector pCDB-Cas9. Prior to assembly, pGG-F-AtU6-26-BbsI-BbsI-G was digested with *Bbs*I. The Golden Gate reaction was performed as previously described (Decaestecker et al., 2019). The pCDB-Cas9 destination vector contains two *Bsa*I restriction sites between the AtU6-26 promotor and the gRNA scaffold to enable one-step Golden Gate assembly of new gRNAs. Next, the *ccdB* gene and chloramphenicol resistance marker (CmR) (Decaestecker et al., 2019) was added between the AtU6-26 and scaffold elements to further streamline the cloning of new gRNAs. pCDB-Cas9 was digested with *Bsa*I, after which the ccdB/CmR insert was ligated to obtain the unarmed gRNA destination vector pCDB-Cas9-ccdB.

A similar destination vector, pCDB-Cas9-GFP-ccdB was generated containing a Green Fluorescent Protein–Nuclear Localization Signal (*GFP-NLS*)-tag translationally fused to the *Cas9* C-terminus. The entry vectors pGGA004, pGGB003, pGGD001, pGGE001 (Lampropoulos et al., 2013), pGG-C-Cas9PTA-D, pGG-F-AtU6-26-AarI-AarI-G (Decaestecker et al., 2019), were combined into the vector pEN-L1-AG-L2,\(-Bbs1) to construct the vector pCDB-Cas9-GFP. The vector was digested with *Aar*I and the ccdB/CmR insert was added to obtain the unarmed gRNA destination vector pCDB-Cas9-GFP-ccdB.

Oligos 1−8 (**Supplementary Table 1**) containing the gRNA sequences and 5′ overlap sequences (5′-ATTG-N_{20}-3′ and 5′-AAAC-N_{20}(reverse complement)-3′) were annealed and cloned into the unarmed gRNA destination vectors pCDB-Cas9-cddB and pCDB-Cas9-GFP-ccdB, as previously described (Decaestecker et al., 2019). This yielded the vectors pCDB-Cas9-PDS, pCDB-Cas9-GFP-PDS, pCDB-Cas9-GAS, pCDB-Cas9-GAO, and pCDB-Cas9-COS (**Supplementary Table 2**). All vectors were verified using colony PCR with primer20 and primer24 (**Supplementary Table 1**), followed by a *Hinc*II restriction digest. Vector DNA was extracted using QIAGEN Plasmid Maxi Kit and the vector gRNA sites were analyzed using Sanger sequencing with primer20. All vectors and their size are listed in **Supplementary Table 2** and are available via the Gateway vector website https://gatewayvectors.vib.be.

Protoplast Isolation, Transfection, and Regeneration

Protoplast Isolation and Transfection

Witloof protoplasts were isolated from young and healthy leaves from *in vitro* maintained plants as previously described (Deryckere et al., 2012). Protoplast suspensions were diluted to 500,000 protoplasts.mL^{-1} and 100 μL was added to a minimum of 10 μg of vector. Next, 120 μL PEG3350 solution [400 $g.L^{-1}$ PEG3350, 72.8 $g.L^{-1}$ mannitol, 23.6 $g.L^{-1}$ $Ca(NO_3)_2.4H_2O$, pH 6] was added to the solution, gently mixed and samples were incubated in the dark for 10 min at room temperature. The transfection reaction was stopped by adding 600 μL of W5 medium (8.77 $g.L^{-1}$ NaCl, 18.38 $g.L^{-1}$ $CaCl_2.2H_2O$, 0.37 $g.L^{-1}$ KCl, and 0.9 $g.L^{-1}$ glucose, pH 5.8) and mixed by inverting the tubes five times. The samples were centrifuged for 5 min at 80 g in a swing out centrifuge (Eppendorf™ 5810R Centrifuge) and the supernatant was removed.

Determination of Transfection Efficiencies Using Fluorescence Microscopy

Protoplasts were transfected with 10 or 20 μg pCDB-Cas9-GFP-PDS or pKAR6 in at least three independent experiments. The pKAR6 vector (Robert Blanvillain, unpublished data) (Thomson et al., 2011; **Supplementary Table 2**) expresses GFP under a 35S promotor and was used as a positive control (PC) for protoplast transfection. Protoplasts transfected without vector (Negative Control 1; NC1) and protoplasts without the addition of both PEG and vector (Negative Control 2; NC2), were used as negative controls. After transfection, 1 mL of 0.5 M mannitol was added to the protoplast pellet and the protoplast suspension was transferred into a 6-well plate and cultured in the dark at 23 ± 2°C on an orbital shaker (30 rpm, 20 h). Next, the protoplast suspension was transferred to an Eppendorf tube, centrifuged for 5 min at 80 g in a swing out centrifuge and supernatant was removed. Twenty μL of the protoplast suspension was transferred to a Bürker chamber and analyzed with a Zeiss AxioImager M2 fluorescence microscope equipped with an Axiocam MRm camera and ZEN software and magnification 200× (Carl Zeiss MicroImaging, Belgium). Transfection efficiencies were calculated as the ratio of the number of GFP expressing protoplasts (GFP Zeiss filter set 10 (excitation 489 nm, emission 509 nm) to the total number of living protoplasts (based on round shape under bright field microscopy).

Regeneration of Transfected Protoplasts

Protoplasts were transfected using 20 μg pCDB-Cas9-PDS in three independent experiments. Subsequent protoplast transfection experiments used a minimum of 10 μg pCDB-Cas9-GAS, pCDB-Cas9-GAO, or pCDB-Cas9-COS. Protoplasts transfected without vector (NC1) and protoplasts without the addition of both PEG and vector (NC2) were used as negative controls. After transfection, 600 μL of regeneration medium (½ Murashige & Skoog macro elements (without NH_4NO_3 and KNO_3) (Murashige and Skoog, 1962) with Heller micro elements (Heller, 1953) and Morel & Wetmore vitamins (Morel and Wetmore, 1951), 18.3 $mg.L^{-1}$ FeNA-EDTA, 750 $mg.L^{-1}$ KCl, 100 $mg.L^{-1}$ inositol, 750 $mg.L^{-1}$ glutamine, 10 $g.L^{-1}$ sucrose, 60 $g.L^{-1}$ mannitol, 0.5 $mg.L^{-1}$ NAA, 0.5 $mg.L^{-1}$ BAP, pH 5.5) was added to the protoplast pellet and the protoplasts were regenerated into plants following the protocol described by Deryckere et al. (2012). After 4 to 5 months, the pCDB-Cas9-GAS, pCDB-Cas9-GAO, pCDB-Cas9-COS transfected shoots and respective control shoots were acclimatized for 4 weeks under a fog tunnel construction with plastic covering (temperature ± 25 °C, 70–80% relative humidity). Afterwards, plantlets were transferred to pots (Ø; 9 cm) and grown in a peat based substrate (1.5 $kg.m^{-3}$ fertilizer: 12N:14P:24K + trace elements, pH 5.0–6.5, EC 450 $\mu S.cm^{-1}$, Van Israel, Geraardsbergen, Belgium) under greenhouse conditions (temperature ± 20°C, 60–65% relative humidity). The frequency of the albino phenotype of the pCDB-Cas9-PDS transfected plants was calculated by dividing the number of albino plantlets by the total number of regenerated plants, maintained *in vitro*.

Ploidy Level Analyses

Ploidy levels were analyzed on 442 *in vitro* regenerated plants (pCDB-Cas9-PDS transfected plants and NC1 and NC2 control plants) and 182 acclimatized greenhouse plants (pCDB-Cas9-GAS, pCDB-Cas9-GAO, pCDB-Cas9-COS transfected plants, and NC1 control plants). Approximately 1 cm^2 of leaf tissue of both the sampled witloof plant and *Pisum sativum*, as an internal standard, were prepared together. The leaf samples were ground with one 3 mm zirconium bead (VWR, Leuven, Belgium) in a 2 mL Eppendorf tube in 250 μL of buffer 1 (0.1 M citric acid monohydrate and 0.5% Tween 20) (Otto, 1997) using a TissueLyser II (Retsch Qiagen, Aartselaar, Belgium) at 20 Hz during 2 min. The ground material was filtered over a 50 μm filter (CellTrics, Sysmex) and stained with 750 μL of buffer 2 [0.4 M $Na_2HPO_4.12H_2O$ and 2 $mg.L^{-1}$ 4′, 6-diamidino-2-phenylindole (DAPI)] (Otto, 1997). Ploidy analysis was performed using a CyFlow Space flow cytometer equipped with a UV Light Emitting Diode (365 nm) (Sysmex, Münster, Germany) and Flomax 2.11 software (Quantum Analysis, Münster, Germany). Ploidy levels were derived from the ratio between the peak position of the sample and the internal standard on the histograms and compared to ratios obtained from the analysis of control witloof plants with known (diploid) ploidy level.

Molecular Analyses

Genomic DNA of 557 regenerated plants (10 albino pCDB-Cas9-PDS transfected plants and 540 pCDB-Cas9-GAS, pCDB-Cas9-GAO, pCDB-Cas9-COS transfected plants, and NC1 plants) was extracted from ± 50 mg fresh leaf material using a CTAB method (Doyle and Doyle, 1990). Per sample, DNA concentration was measured using the Nanodrop ND1000 (Isogen Lifescience B.V.) and samples were diluted to obtain a final DNA concentration of maximum 40 $ng.\mu l^{-1}$. Primers were designed for *CiPDS*, *CiGAS, CiGAO,* and *CiCOS* (**Supplementary Table 3**) flanking the gRNA target site and the 100–150 bp amplicons were amplified using a highly multiplex (HiPlex) PCR reaction, while attaching sample-specific barcodes. Amplicons from all samples were pooled and ligated to Illumina TruSeq sequencing adapters using the KAPA Hyper prep PCR-free ligation kit according to manufacturer directions (Kapa Biosystems, United States). HiPlex amplification reactions and library preparations were performed by Floodlight Genomics LLC (Knoxville, TN, United States). The libraries were sequenced with 150 PE on a HiSeq3000 instrument (Admera, United States). Forward and reverse reads were merged with PEAR (v0.9.8) (Zhang et al., 2014), sample-specific barcodes were used for sample demultiplexing with custom python scripts and sample-specific barcodes and linker sequences introduced during library preparation were removed. The following steps were performed per sample, and processed in parallel. Reads were sorted per gene by mapping [BWA-MEM with default parameters Li and Durbin, 2009] to the reference gene sequences for *CiPDS*, *CiGAS*, *CiGAO*, and *CiCOS,* and the original fastq read files with all HiPlex reads per sample were split into subsets of reads per gene per sample using the readID. The gene-specific amplification primers were removed by trimming the reads with Cutadapt (Martin, 2011) and the remaining sequence window (defined

as the entire sequence between the HiPlex primers per gene) was considered as an allele per gene. All unique read sequences, including any potential novel (non-reference) alleles originating from genome editing, were counted per gene per sample. After processing all samples, an integrated table was created listing all read counts per allele per gene per sample across the sample set. Next, the relative allele frequency was calculated as the number of reads per allele per gene per sample divided by the total number of reads per gene per sample. Low frequency alleles were removed using a minimal allele frequency threshold of 2% for analyzing *CiPDS* and 6% for analyzing *CiGAS*, *CiGAO*, and *CiCOS*. This frequency threshold was calibrated based on empirical observations of the distribution of allele frequencies of alternative (non-reference) sequences in wild type (non-mutated) loci, which were assumed to be derived from PCR artifacts, including sequence jumps; low-frequency sequencing errors, such as base calling errors inherent to Illumina short-read sequencing, read mapping errors, etc. A detailed illustration of the workflow is presented in **Supplementary Figure 2**. To search for possible insertions of fragments of the transfection vector at the genome-edited loci, any alleles containing a long insertion compared to the reference sequence (>15 nucleotides), and independent of relative allele frequency, were mapped to the pCDB-Cas9 vector. Seven out of 540 samples (1.3%; 6 pCDB-Cas9-GAS and 1 pCDB-Cas9-GAO transfected plants) showed irregular sequencing results and were excluded from the analysis. The overall plant mutation frequency per target gene was calculated by dividing the number of regenerated plants containing at least one observed mutated (non-reference) allele by the total number of regenerated plants (i.e., wild type plants + mutated plants). The gene knockout frequency per target gene was calculated by dividing the number of observed frameshift mutations leading to the premature truncation of the protein, by the total number of regenerated plants (i.e., wild type plants + mutated plants).

CiPDS Copy Number Analysis With ddPCR

Droplet digital PCR (ddPCR) was used to quantify the copy number of *CiPDS* in a diploid "Van Hamme" witloof plant according to the protocol described in Desmet et al. (2020) with some minor assay-specific modifications. Two reference genes were selected: the single copy gene *PP2AA3* and *UBQ10*, present in a double amount of copies (Delporte et al., 2015); primers are described in **Supplementary Table 3**. For the amplification of the *CiPDS*, primer11 and primer12 (**Supplementary Table 3**) were designed in the same region as the primers used for HiPlex amplicon sequencing of the *CiPDS* target region. Four DNA samples of Van Hamme (1 μg) were digested with EcoRI (10U) and MspI (50U). Ten ng of the digest was used as input material for ddPCR and each sample, including no template controls (NTC), was analyzed in duplicate on the QX200™ (BioRad). Annealing temperature of the PCR was 56°C for these assays and Quantasoft version 1.7.4.0917 (Bio-Rad) was used for calculating concentrations (copies.μL^{-1}). Haploid *CiPDS* copy number was calculated as [*CiPDS*] / [reference gene] X 1 or 2 (for *PP2AA3* or *UBQ10*, respectively). Finally mean (+/– stdev) copy number was calculated over both reference genes and all 4 samples.

RESULTS

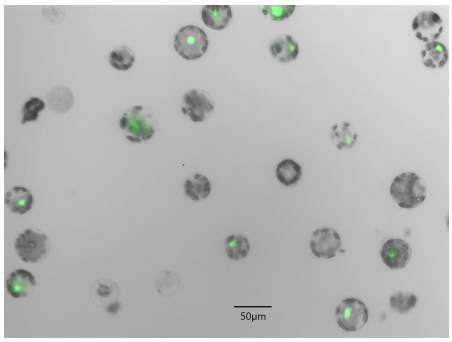
A CRISPR/Cas9-Vector Based Protoplast Transfection and Regeneration Protocol

Protoplasts of two witloof varieties ("Van Hamme" and "Topmodel") were transfected with 10 or 20 μg of the pCDB-Cas9-GFP-PDS vector to test the effect of vector concentration on protoplast transfection efficiency (**Table 1**). Protoplast transfection efficiencies with pCDB-Cas9-GFP-PDS were between 20 and 26% in both varieties. For "Van Hamme" protoplasts, this was similar to the transfection efficiency when using the positive control vector pKAR6, while in "Topmodel" protoplasts the transfection efficiency with pKAR6 was markedly higher (44.2%). No GFP fluorescence was observed in any of the negative control treatments. Overall, these results show that 10 μg of vector DNA is sufficient to obtain efficient protoplast transfection.

Witloof "Van Hamme" was selected for subsequent transfection and regeneration experiments because of its higher regeneration capacity (data not shown) and more consistent transfection efficiencies (**Table 1**). Protoplasts were transfected with pCDB-Cas9-PDS to induce mutations in *CiPDS* which can result in albino plantlets (Zhang et al., 2017). Protoplast transfection with pCDB-Cas9-PDS and subsequent regeneration yielded a total of 186 regenerated plants over three independent experiments, among which we observed a total of 55 albino plantlets during *in vitro* culture. The frequency of observed albino phenotypes to the total number of regenerated plants varied from 21 to 27% (mean 23 ± 3%; **Table 2**) and no albino plantlets were observed in the 355 negative control regenerated plants.

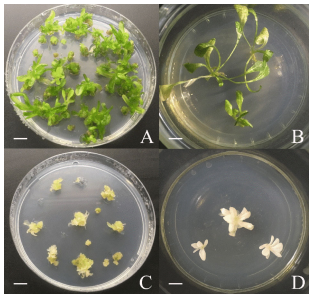
HiPlex amplicon sequencing was used to determine the genomic DNA sequence at the *CiPDS* target loci of ten albino plantlets. All other 533 regenerated plants (see description of *CiGAS*, *CiGAO*, and *CiCOS* targeted plants below in 3.2) were used as controls for *CiPDS* sequencing. In all plants (control and albino), the *CiPDS* primers consistently amplified two different sequence variants, differentiated by a single SNP [localized 39 nucleotides upstream from the gRNA target site (**Supplementary Table 4**)]. The ten albino plants showed different combinations of two mutated alleles derived from *CiPDS* reference locus 1 (**Tables 3**, **4**, listed as *CiPDS* locus 1) and all seven unique mutated alleles carried frameshift mutations that lead to premature termination and functional knockout of the encoded protein. These observations can be explained by the hypothesis that these alleles are derived from a locus that is homozygous in diploid wild type plants and can give rise to two alternative alleles upon genome-editing. The second reference locus (**Tables 3**, **4**, listed as *CiPDS* locus 2) was also edited, giving rise to a total of 14 unique mutated alleles. The number of unique mutations varied between the 10 albino plants and, strikingly, gave rise to up to six different mutated alleles in a single diploid plant, suggesting three additional paralogous copies of the gRNA target site and flanking regions (Plant A4, **Table 3**). All mutations were positioned at the predicted Cas9 cut site (**Table 4**), consistent with DSB followed by imperfect NHEJ repair. The alleles with large insertions contained relatively large fragments

TABLE 1 | Fluorescence microscopy image of witloof "Van Hamme" protoplasts transfected with pCDB-Cas9-GFP-PDS (left panel) and transfection efficiencies with the pCDB-Cas9-GFP and pKAR6 vector (right panel).

Fluorescence microscopy image	Variety	Vector	DNA	Efficiency (%)
	Witloof "Van Hamme"	pKAR6	10 µg	22.9 ± 2.6
		pCDB-Cas9-GFP-PDS	10 µg	20.5 ± 4.2
		pCDB-Cas9-GFP-PDS	20 µg	24.0 ± 5.9
		NC1	–	0.0
		NC2	–	0.0
	Witloof "Topmodel"	pKAR6	10 µg	44.2 ± 4.0
		pCDB-Cas9-GFP-PDS	10 µg	20.7 ± 7.2
		pCDB-Cas9-GFP-PDS	20 µg	26.0 ± 1.9
		NC1	–	0.0
		NC2	–	0.0

NC1, protoplasts transfected with PEG but without vector; NC2, protoplasts without addition of both PEG and vector (mean +/– stdev; n = 3–5).

TABLE 2 | Regeneration of witloof protoplasts and frequency of albino phenotypes.



Exp.	# Green	# Albino	Frequency (%)
1	30	8	21.0
2	45	17	27.4
3	111	30	21.3
Total	186	55	**23.3 ± 3.0**
NC1	127	0	0.0
NC2	228	0	0.0

*Left Panel: **(A)** Wild-type (WT) shoots emerging from callus, **(B)** WT plants, **(C)** CiPDS mutated albino shoots, and **(D)** CiPDS mutated albino plants. Scale bar, 1 cm. Right Panel: Witloof "Van Hamme" plants were transfected with 20 µg pCDB-Cas9-PDS. Frequency of albino phenotypes are given per experiment (1, 2, 3); Frequency is given as mean +/– stdev across the three experiments (n = 3, in bold). NC1, protoplasts transfected with PEG but without vector; NC2, protoplasts without addition of both PEG and vector.*

TABLE 3 | Genotype of the ten albino plantlets and their corresponding ploidy level.

Plant	Ploidy	*CiPDS* Locus 1	*CiPDS* Locus 2
A1	Diploid	I1A/I1T	WT/I1T/D10I88
A2	Diploid	I1A/I1T	WT/I1T/D10I88
A3	Diploid	I1A/I1T	WT/I1T/D10I88
A4	Diploid	I1C/I1T	D9/D9I2/D3/D2/WT/I1A
A5	Diploid	D8I6/I1T	D10/D2/D9I30/D1I25
A6	Diploid	D5/I1T	D13/D3/WT/I70
A7	Tetraploid	D2/I1A	WT
A8	Tetraploid	D2/I1A	WT
A9	Tetraploid	I1T	WT/I1G
A10	Tetraploid	D14/I1T	D3/D2/D1/WT/I1T

I, Insertion; D, Deletion; WT, Wild Type. Mutation types and plant genotypes with mutations leading to early stop codons, truncating the protein translation are underlined.

(21, 26, 70, and 87 nucleotides) of the transfection vector. The copy number of the *CiPDS* gRNA target region was estimated using droplet digital PCR (ddPCR) and resulted in the presence of at least four copies (4.54 +/− 0.18; **Supplementary Table 5**). Taken together, the observation of up to eight different mutated *CiPDS* alleles in a single diploid plant can be explained by the presence of at least four paralogous copies of the *CiPDS* gRNA target site in plant variety "Van Hamme," in line with the ddPCR results. Our data further show that multiple loci may be edited simultaneously by a common gRNA targeting a conserved sequence and their parallel detection was possible with the common HiPlex *CiPDS* primers.

As plant regeneration from protoplasts may lead to changes in ploidy level (Larkin and Scowcroft, 1981), we scored the frequency of ploidy changes under our conditions of protoplast transfection and regeneration. The ploidy level of 442 *in vitro* regenerated plants (both pCDB-Cas9-PDS transfected plants and NC1 and NC2 control plants), were analyzed by flow cytometry, showing 77.2% diploid plants, 21.5% tetraploid plants and the remaining 1.3% consisted of haploids, hexaploids and mixoploids.

Taken together, these results show that our protocol for genome editing based on transfection and regeneration of witloof protoplasts without the use of typical plant selection markers (e.g., kanamycin, hygromycin, *bar* gene), yielded about 23% albino plantlets amongst all regenerated plants (**Table 2**).

CRISPR/Cas9 Induced Mutations in the SL Biosynthesis Pathway Genes GAS, GAO, or COS

Our genome editing protocol was then used to target the *CiGAS*, *CiGAO*, and *CiCOS* genes, known to be involved in the SL biosynthesis pathway (Bouwmeester, 2002; Liu et al., 2011). Witloof "Van Hamme" protoplasts were transfected with either the pCDB-Cas9-GAS, pCDB-Cas9-GAO, or pCDB-Cas9-COS vectors. A total of 533 plants were regenerated and genetically characterized at all target loci, including 11

TABLE 4 | Overview of the mutation types and relative contribution of each mutation type in the ten mutated albino plants.

***CiPDS*—Locus 1**		**Relative contribution %**
WT	AAAAAGATGTACTCTCATTTGGATGCTGTGATGCCATGGGTCACAG	0.0
I1T	AAAAAGATGTACTCTCATTTGGATGCTGTGATGCTCATGGGTCACAG	42.1
I1A	AAAAAGATGTACTCTCATTTGGATGCTGTGATGCACATGGGTCACAG	26.3
D2	AAAAAGATGTACTCTCATTTGGATGCTGTGAT--CATGGGTCACAG	10.5
I1C	AAAAAGATGTACTCTCATTTGGATGCTGTGATGCCCATGGGTCACAG	5.3
D8I6	AAAAAGATGTACTCTCATTTGGATGC--ATCACACATGGGTCACAG	5.3
D5	AAAAAGATGTACTCTCATTTGGATGCTGT-----CATGGGTCACAG	5.3
D14	AAAAAGATGTACTCTCATTT----------CATGGGTCACAG	5.3
***CiPDS*—Locus 2**		**Relative contribution %**
WT	AAAAAGATGTACTCTCATTTGGATGCTGTGATGCCATGGGTCACAG	28.1
I1T	AAAAAGATGTACTCTCATTTGGATGCTGTGATGCTCATGGGTCACAG	12.5
D2	AAAAAGATGTACTCTCATTTGGATGCTGTGAT--CATGGGTCACAG	9.4
D3	AAAAAGATGTACTCTCATTTGGATGCTGTGA---CATGGGTCACAG	9.4
D10I88	AAAAAGATGTACTCTCATTTGGATCN$_{87}$CATGGGTCACAGA	9.4
I1A	AAAAAGATGTACTCTCATTTGGATGCTGTGATGCACATGGGTCACAG	3.1
D13	AAAAAGATGTACTCTCATTTGGATGC----------ATCACAG	3.1
D10	AAAAAGATGTACTCTCATTTGGATGC----------TGGGTCACAG	3.1
D9	AAAAAGATGTACTCTCATTTGGATGC---------ATGGGTCACAG	3.1
D9I2	AAAAAGATGTACTCTCATTTGGATGC-------ATATGGGTCACAG	3.1
D1	AAAAAGATGTACTCTCATTTGGATGCTGTGATG-CATGGGTCACAG	3.1
I1G	AAAAAGATGTACTCTCATTTGGATGCTGTGATGCGCATGGGTCACAG	3.1
D9I30	AAAAAGATGTACTCTCATTTGGATGCATCAN$_{26}$CATGGGTCACAG	3.1
D1I25	AAAAAGATGTACTCTCATTTGGATGCTGTGATGATGGN$_{21}$CATGGGTCACAG	3.1
I70	AAAAAGATGTACTCTCATTTGGATGCTGTGATGCN$_{70}$CATGGGTCACAG	3.1

Relative contribution frequencies contain small rounding errors. Dashed line, cut site; purple, PAM site; orange, inserted nucleotides; underlined mutation type, indel leading to early stop codon, truncating the protein translation (gene knockout); underlined sequence, exact vector insert; I, Insertion; D, Deletion; N#, amount of inserted nucleotides; WT, Wild Type.

control plants (NC1), 374 plants transfected with pCDB-Cas9-GAS, 89 plants transfected with pCDB-Cas9-GAO and 59 plants transfected with pCDB-Cas9-COS. The ploidy level was analyzed for 182 regenerated plants showing 80.8% diploid plants, 18.1% tetraploid plants and 1.1% consisting of haploids, hexaploids, and mixoploids. Out of the 522 regenerated plants after transfection, mutation analysis revealed 324 wild type plants, 137 plants with a monoallelic mutation (one reference allele and one mutated allele), 5 plants with a single type of mutated allele (but no reference allele detected) and 56 plants with biallelic mutations (two different mutated alleles) (**Table 5**). Furthermore, 19 plants contained indels resulting in a premature truncation of the protein (presumably gene knockout) in all observed alleles per plant (**Table 5**). This resulted in an overall plant mutation frequency of 37.9% and gene knockout frequency of 3.6%. More specifically, in pCDB-Cas9-GAS transfected plants, the plant mutation frequency was 27.3% (**Table 5**) with a total of five mutated plant genotypes (**Table 6**). The pCDB-Cas9-GAO transfected plants showed a plant mutation frequency of 42.7% (**Table 5**) and a total of 18 mutated plant genotypes (**Table 6**). In pCDB-Cas9-COS transfected plants, the plant mutation frequency was 98.3% (**Table 5**) and only one mutated plant genotype was observed (M24) (**Table 6**). Additionally, nine *CiGAO* mutated plant genotypes (M7, M9, M12, M13, M17, M18, M20, M21, and M23) showed premature truncation of the CiGAO protein due to a frameshift mutation in all observed alleles, resulting in a *CiGAO* gene knockout frequency of 21.4%, while no homozygous knockouts were created in *CiGAS* or *CiCOS* mutated plants (**Table 6**).

Table 7 gives a detailed overview of the detected mutation types and their relative contribution within the *CiGAS*, *CiGAO,* and *CiCOS* mutated plants. The *CiGAS* amplicon reads revealed six different mutation types across all *CiGAS* mutated plants (**Table 7**). The *CiGAO* amplicon reads revealed a total of 20 different mutation types across all *CiGAO* mutated plants (**Table 7**). Notably, five of the *CiGAO* mutated plants showed, in at least one of the alleles, an insertion which was part of the Cas9 vector with fragment lengths of 30, 43, 47, 90, and 104 nucleotides. The *CiCOS* amplicon reads revealed a mutation type containing a single "A" nucleotide insertion across all *CiCOS* mutated plants (**Table 7**). No mutations were detected among the 11 control plants and in the observed loci that were not targeted by the gRNAs. These results show that our protocol for genome editing based on transfection and regeneration of witloof protoplasts, can create a variety of mutation types and mutated plant genotypes. Furthermore, gene knockouts can be created that could be able to change the SL metabolite content and result in altered bitterness.

TABLE 5 | Overview of the mutation events in the pCDB-Cas9-GAS, pCDB-Cas9-GAO, and pCDB-Cas9-COS protoplast transfected and regenerated plants.

Genes	GAS	GAO	COS	Total	Control
Total number of regenerants	374	89	59	522	11
WT	272	51	1	324	11
WT/Indel	71	8	58	137	0
Indel (KO)	0 (0)	5 (3)	0 (0)	5 (3)	0 (0)
Indel/Indel (KO/KO)	31 (0)	25 (16)	0 (0)	56 (16)	0 (0)
Plant mutation frequency[a]	27.3%	42.7%	98.3%	37.9%	0.0%
Gene knockout frequency[b]	0.0%	21.4%	0.0%	3.6%	0.0%

WT, Wild Type plant; WT/Indel, Monoallelic mutation; Indel, Single observed mutation; Indel/Indel, Biallelic mutation; KO, Gene knockout (indel leading to early stop codon, truncating the protein translation). [a]*Calculated by dividing the sum of WT/Indel, Indel and Indel/Indel by the total number of regenerants: (WT/Indel + Indel + Indel/Indel)/TotReg.* [b]*Calculated by dividing the sum of KO and KO/KO by the total number of regenerants: (KO + KO/KO)/TotReg.*

TABLE 6 | Overview and number of *CiGAS*, *CiGAO* and *CiCOS* mutated plant genotypes.

Target gene	Plant	Plant genotype	# plants
CiGAS	M1	WT/D11	25
	M2	WT/D10	12
	M3	WT/D7	19
	M4	WT/D6	15
	M5	D6*/D8I3	31
CiGAO	M6	D27/D10	1
	M7	D26	1
	M8	D12/D12*	2
	M9	D4	1
	M10	D4/D2I92	1
	M11	D3/D14	5
	M12	D2/I47	2
	M13	D1/D4	2
	M14	WT/D2	2
	M15	WT/I1T	6
	M16	D12	1
	M17	I1T/D17	6
	M18	I1A/D13	1
	M19	I1T/D3	1
	M20	I1T/D2	2
	M21	I1A	1
	M22	I30	1
	M23	D12I55/I104	2
CiCOS	M24	WT/I1A	58

*WT/Indel, Heterozygous mutation; Indel, Single observed mutation; Indel/Indel, Compound heterozygous mutation, I, Insertion, D, Deletion, *, different mutation type; WT, Wild Type. Mutation types and plant genotypes with mutations leading to early stop codons, truncating the protein translation are underlined.*

DISCUSSION

CRISPR/Cas9 genome editing is a powerful tool for both gene function research and plant breeding and was already successfully applied on several crops such as maize, rice, tomato, among others (Manghwar et al., 2019; Zhang et al., 2019). We evaluated CRISPR/Cas9 in protoplast transfection and regeneration experiments for witloof and analyzed protoplast transfection efficiencies, CRISPR/Cas9 mutation efficiencies, mutation spectra and *in vitro* tissue culture observations (ploidy changes and clonal propagation).

Firstly, we confirmed CRISPR/Cas9 vector expression in the protoplasts using PEG transfection and a CRISPR/Cas9 vector containing a GFP marker in two witloof varieties "Van Hamme" and "Topmodel." Protoplast transfections resulted in transfection efficiencies around 20 to 26%. Previous reports have shown variable PEG mediated protoplast transfection efficiencies, for instance, in cherry (33.4%) (Yao et al., 2016), orchid (41.7%) (Li et al., 2018), rice (53–75%) (Zhang et al., 2011) and poinsettia (>70%). Those studies showed the impact of multiple parameters such as species, plant tissue, PEG concentration, vector size, vector concentration and transfection duration. In addition, those studies were aimed at achieving maximal transfection efficiencies, without taking into consideration long-term protoplast viability and protoplast regeneration capacity. In our study, we obtained 20 to 26% transfection efficiency while still obtaining successful regeneration. Hence, we established an efficient transfection protocol in witloof protoplasts, which makes further applications for witloof genome editing possible.

The mutation efficiency in witloof using CRISPR/Cas9 was initially estimated by targeting the *CiPDS* gene using a single gRNA and without the use of selective media. The *PDS* gene is often used as a target during the development of genome editing techniques as it allows direct visual screening for albino phenotypes resulting from CRISPR/Cas9 induced mutations. The *PDS* gene has been targeted in many species such as wheat and tobacco (Upadhyay et al., 2013), rice (Shan et al., 2013), poplar (Fan et al., 2015), apple (Nishitani et al., 2016), watermelon (Tian et al., 2017), cassava (Odipio et al., 2017), banana (Kaur et al., 2018; Naim et al., 2018), and recently also chicory (Bernard et al., 2019). In our study, a CRISPR/Cas9 vector targeting the first exon of *CiPDS* was delivered into witloof protoplasts and resulted in albino plants in 23% of the regenerated plants. DdPCR analysis revealed at least four paralogous copies of the *CiPDS* gRNA target sequence, while HiPlex amplicon sequencing analysis showed two types of *CiPDS* gRNA flanking sequences in the "Van Hamme" genome, which can be distinguished based on a SNP 39 nucleotides away from the gRNA target site. All tested plants with an albino phenotype carried a knockout mutation in all observed alleles in locus 1 of the *CiPDS* gene. The other loci displayed variable numbers of mutated alleles, reaching a total of up to eight different mutation types in the *CiPDS* alleles of an albino individual. Taken together, these data show that a single gRNA that binds a conserved sequence can effectively and simultaneously induce mutations at multiple genomic loci. PEG-mediated protoplast transfection targeting *PDS* in other species showed highly variable mutation frequencies, probably because of numerous experimental parameters, such as gene target, target design, species transfection and regeneration ability and ploidy level. For instance, the reported mutation frequency in the *PDS*

TABLE 7 | Overview of the mutation types and relative contribution of each mutation type in the plants transfected with pCDB-Cas9 vector targeting *CiGAS*, *CiGAO*, and *CiCOS*.

CiGAS		**Relative contribution %**
WT	ACGGGAACGAGGCTTACCACATGGCAGTTTGTGACCAAGGTGTCTG	34.8
D8I3	ACGGGAACGAGGCTTACCACATGGCAA------ACCCAAGGTGTCTG	15.2
D6	ACGGGAACGAGGCTTACCACATGGCAG------ACCAAGGTGTCTG	15.2
D11	ACGGGAACGAGGCTTACCACATGG-----------CAAGGTGTCTG	12.3
D7	ACGGGAACGAGGCTTACCACATGGCAG-------CCAAGGTGTCTG	9.3
D6*	ACGGGAACGAGGCTTACCACATGGCAGTT------CAAGGTGTCTG	7.4
D10	ACGGGAACGAGGCTTACCACATG----------ACCAAGGTGTCTG	5.9
CiGAO		**Relative contribution %**
WT	TCAAAGAGAGTTTACTGAGATTGTTAAAGAGATCTTGAGGCAAACT	11.3
I1T	TCAAAGAGAGTTTACTGAGATTGTTAAAGAGATCTTTGAGGCAAACT	21.1
D3	TCAAAGAGAGTTTACTGAGATTGTTAAAGAG---TTGAGGCAAACT	8.5
D17	TCAAAGAGAGTTTACTG----------TTGAGGCAAACT	8.5
D14	TCAAAGAGAGTTTACTGAGA----------TTGAGGCAAACT	7.0
D4	TCAAAGAGAGTTTACTGAGATTGTTAAAGA----TTGAGGCAAACT	5.6
D2	TCAAAGAGAGTTTACTGAGATTGTTAAAGAGATC--GAGGCAAACT	5.6
D12	TCAAAGAGAGTTTACTGAGATT----------TTGAGGCAAACT	4.2
D1	TCAAAGAGAGTTTACTGAGATTGTTAAAGAGAT-TTGAGGCAAACT	2.8
*D2**	TCAAAGAGAGTTTACTGAGATTGTTAAAGAGA--TTGAGGCAAACT	2.8
D12*	TCAAAGAGAGTTTACTGAGAT----------CTTGAGGCAAACT	2.8
I1A	TCAAAGAGAGTTTACTGAGATTGTTAAAGAGATCATTGAGGCAAACT	2.8
D12I55	TCAAAGAGAGTTTACTGAGAT----------N_{55}CTTGAGGCAAACT	2.8
I47	TCAAAGAGAGTTTACTGAGATTGTTAAAGAGATN_{47}CTTGAGGCAAACT	2.8
I104	TCAAAGAGAGTTTACTGAGATTGTTAAAGAGATN_{104}CTTGAGGCAAACT	2.8
D10	TCAAAGAGAGTTTACTGAGATTGTTAAAG----------GCAAACT	1.4
D13	TCAAAGAGAGTTTACTGAGA----------CTTGAGGCAAACT	1.4
D26	TCAAAGAGAGTTT----------GCAAACT	1.4
D27	TCAAAGA----------TTGAGGCAAACT	1.4
I30	TCAAAGAGAGTTTACTGAGATTGTTAAAGAGATCN_{30}TTGAGGCAAACT	1.4
D2I92	TCAAAGAGAGTTTACTGAGATTGTTAAAGAGATC--N_{92}GAGGCAAACT	1.4
CiCOS		**Relative contribution %**
WT	GCCAAAAAGGTTCGATCTTTCAGCTCCATTCGCGAAGAGGAGCTTA	50.4
I1	GCCAAAAAGGTTCGATCTTTCAGCTCCATTCGCGAAAGAGGAGCTTA	49.6

*Relative contribution frequencies contain small rounding errors. Dashed vertical line, cut site; purple, PAM site; orange, inserted nucleotides, underlined mutation type, indel leading to early stop codon, truncating the protein translation (gene knockout); underlined sequence, exact vector insert; I, Insertion; D, Deletion; *, different mutation type; N#, number of inserted nucleotides; WT, Wild Type.*

gene ranged from 1.1 to 5.6% in PEG-transfected *Arabidopsis* protoplasts and was around 37% in tobacco protoplasts (Li et al., 2013). Additionally, parameters related to mutation analysis techniques, such as mutation detection techniques, mutation threshold values and mutation frequency calculations, can also play a role in the reported variability of mutation frequencies. For instance, a mutation frequency of 6.6% was reported in bamboo protoplasts using band intensity calculations of gel electrophoresis images, while a mutation frequency of 12.5% was reported using NGS sequencing and read depth analysis (Lin et al., 2018). Furthermore, protoplasts of tetraploid potato yielded mutations in all four alleles of the *GBSS* gene in up to 2% of regenerated lines, whereas 2–12% of regenerated lines showed mutations in at least one allele of the gene (Andersson et al., 2017). Therefore, comparison of mutation efficiencies should be based on the same quantification analysis. We have further reported our mutation efficiencies in terms of (1) plant mutation frequency to analyze the efficiency of obtaining plants with single, monoallelic, and biallelic mutations, and (2) gene knockout frequency to analyze the efficiency of obtaining plants with a knockout of all observed alleles. Furthermore, these mutation efficiencies can be of interest for gene function analysis and plant breeding purposes. Acceptable mutation efficiencies thus depend on plant species and the objective of the CRISPR/Cas9 genome editing technique. In our research, we demonstrated that genome editing in witloof protoplasts is promising, with transfection rates of at least 20% and frequencies of the albino phenotype of at least 23% which is more than sufficient for subsequent screening

and analysis of mutated plant lines without the use of stable transformation and typical plant selection markers.

To further demonstrate the potential of our genome editing approach, three previously described SL biosynthesis pathway genes were targeted using our CRISPR/Cas9 vectors. Characterization of the witloof mutated plants revealed single, monoallelic, and biallelic mutations. Interestingly, in *CiGAO* mutated plants, many different mutation types were observed and 21.4% of the *CiGAO* transfected plants had a mutation in all observed alleles leading to a premature stop codon (knockout). These plants containing knockout mutations in all observed alleles are also the ideal material for studying gene function. Furthermore, self-fertilization of the plants containing a heterozygous knockout mutation (WT/KO) could also generate a homozygous knockout mutation (KO/KO) after Mendelian segregation in the progeny. Further research and creation of next generation plants through self-pollination will provide information about the segregation in the progeny and to verify the success rate of creating these homozygous knockout plants. Overall mutation screening of the 522 transfected greenhouse plants revealed a plant mutation frequency of 38% containing a wide spectrum of mutation types. Mostly a single deletion, insertion, or nucleotide substitution was induced, although sometimes a combination of deletion and insertion was observed. In case the Cas9 enzyme produces a DSB at the target site, a deletion can be followed by an insertion (e.g., D8I3, **Table 7**). This is supported by previous research on the repair of DSBs in plants, which shows different combinations can occur during the repair process (Puchta, 2005). However, most mutation analysis reports, typically only list the difference in total sequence length (expressed in number of nucleotides) at the mutated target site compared to the corresponding reference sequence length. We also observed instances of CRISPR/Cas9 DNA vector fragment insertions at the site of the DSBs in the genome of mutated plants. Once the vectors are introduced in the protoplasts, they are digested by endogenous nucleases and yield fragments that can be integrated in the plant genome. This phenomenon has previously been reported, amongst others, in potato (Andersson et al., 2017), tobacco (Lin et al., 2018), and chicory (Bernard et al., 2019). Generally, HiPlex amplicon sequencing allows to screen for vector fragment inserts at the site of the DSBs, but large indels may not be detected if it affects the primer binding site. Alternatively, size-exclusion during library preparation and sequencing also creates a bias against amplicons with a substantial change in length after mutation, resulting from relatively large insertions (>150 bp) or deletions (>90 bp) between the primer binding sites. Further mutation analysis revealed that the number of produced mutation types seems to be linked to the target sequence site. Previously, it has been suggested that the variability of mutation types could also be linked to the intrinsic DNA repair mechanism of the species, transformation method and/or culture conditions (Allen et al., 2019). However, as the target sequence site was the only variable parameter in our research, we can determine the target sequence site as an influencing factor in creating this variable mutation spectrum. This has already been described by Shen et al. (2018) and Liu et al. (2020), implementing a computational method that predicts DNA repair outcomes at DSBs induced by CRISPR/Cas9 resulting from NHEJ. However, using the online computational method inDelphi (https://indelphi.giffordlab.mit.edu/) (Shen et al., 2018) on our target genes, suggested different mutation spectra and contributions. As the model was trained on mammalian cell types, it is not expected to generalize well to bacteria, plants, and non-mammalian eukaryotes.

Working with *in vitro* tissue culture comprising protoplast transfection and regeneration, questions regarding ploidy level changes (Larkin and Scowcroft, 1981) and plant clones arise. Therefore, we analyzed ploidy level changes during protoplast regeneration and observed the formation of around 20 % tetraploids during pCDB-Cas9-PDS and pCDB-Cas9-GAS, pCDB-Cas9-GAO, and pCDB-Cas9-COS protoplast transfection and regeneration experiments. Hereby, we only observed up to two different mutated alleles in *CiGAS*, *CiGAO*, and *CiCOS* mutated plants, hypothesizing that ploidy level changes occur after the CRISPR/Cas9 mutation event. Ultimately, using CRISPR-Cas9 mutated plants in witloof breeding requires to screen and select for diploid regenerated plants, and it is important to monitor ploidy changes in an early stage after plant regeneration. We thereafter analyzed the frequency with which clonal lines are obtained from callus. Using our mutation data, we observed the presence of a high number (18) of unique mutated plant genotypes in all (38) *CiGAO* mutated plants (**Table 6**), suggesting a low occurrence of clonal plant lines. Nevertheless, two plant genotypes (M12 and M23; **Table 6**) carrying the same vector fragment insert (e.g., the exact same sequence of 47 nucleotides), illustrate that in our experiments occasionally clonal lines were obtained from individual calli. As it is unlikely that the same vector fragment insert was introduced into the genome of more than one independent protoplast, we speculate that some separately analyzed plants actually originate from the same transfected and mutated protoplast, resulting in multiple plant clonal genotypes. While during regeneration, each callus can produce multiple shoots, selection of only one shoot per callus will reduce the number of clonal plant genotypes. However, the presence of a high number of plants containing the same mutated plant genotype (e.g., M24, **Table 6**) could also be the result of a preferred DNA repair outcome at DSBs induced by CRISPR/Cas9 resulting from NHEJ. The high mutation frequency and low mutation variation observed for pCAS9-COS transfected plants (M24) could thus result from the preferred "A" nucleotide insertion during NHEJ, possibly combined with some level of clonal propagation.

We have shown that the CRISPR/Cas9 technology is very valuable to induce targeted mutations in four genes of witloof. To further implement a DNA-free genome editing technique, it would be interesting to use pre-assembled ribonucleoprotein complexes (RNPs) instead of vector DNA to deliver Cas9/gRNA into *Cichorium* protoplasts. RNPs have been used in *Arabidopsis*, tobacco, lettuce, and rice protoplasts, through PEG-mediated transfection, using the same conditions as with vector DNA transfection, leading to higher mutation frequencies and eliminating the chance of vector fragment insertions (Wook Woo et al., 2015). In conclusion, CRISPR/Cas9 genome editing is of significant importance for future witloof breeding as it comprises a powerful tool for investigating gene functions and altering agronomical traits in commercially interesting witloof varieties.

AUTHOR CONTRIBUTIONS

CD, KV, TE, TR, ED, TJ, and AG: study conception. TE and CD: design of experiments and protoplast assays. TJ and CD: vector design. CD: production of gene edited plants and wrote the manuscript. TR and CD: NGS data analysis and interpretation. KV and CD: overall overview of experiments. KV and TR: revised the manuscript drafts. All authors contributed to the manuscript revision, read, and approved the submitted version.

FUNDING

The authors gratefully acknowledge the financial support of the Research Foundation Flanders (FWO SB Grant No. 1S01520N).

ACKNOWLEDGMENTS

The authors gratefully acknowledge Laurens Pauwels (VIB-UGent) for providing the pKAR6 vector, Ronald Van den Oord, and Michiel Carron (ILVO) for technical support in plant tissue culture, Leen Leus and Hilde Carlier (ILVO) for technical support in flow cytometry, Veerle Buysens (ILVO) for technical support in DNA-extraction and Dries Schaumont (ILVO) for technical support in bio-informatics.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgeed.2020.604876/full#supplementary-material

REFERENCES

Allen, F., Crepaldi, L., Alsinet, C., Strong, A. J., Kleshchevnikov, V., De Angeli, P., et al. (2019). Predicting the mutations generated by repair of Cas9-induced double-strand breaks. *Nat. Biotechnol.* 37, 64–82. doi: 10.1038/nbt.4317

Andersson, M., Turesson, H., Nicolia, A., Fält, A. S., Samuelsson, M., and Hofvander, P. (2017). Efficient targeted multiallelic mutagenesis in tetraploid potato (*Solanum tuberosum*) by transient CRISPR-Cas9 expression in protoplasts. *Plant Cell Rep.* 36, 117–128. doi: 10.1007/s00299-016-2062-3

Bernard, G., Gagneul, D., Alves, H., Santos, D., Etienne, A., Hilbert, J., et al. (2019). Efficient genome editing using CRISPR/Cas9 technology in Chicory. *Int. J. Mol. Sci.* 20:1155. doi: 10.3390/ijms20051155

Bouwmeester, H. J. (2002). Isolation and characterization of two germacrene a synthase cDNA clones from Chicory. *Plant Physiol.* 129, 134–144. doi: 10.1104/pp.001024

Chadwick, M., Trewin, H., Gawthrop, F., and Wagstaff, C. (2013). Sesquiterpenoids lactones: benefits to plants and people. *Int. J. Mol. Sci.* 14, 12780–12805. doi: 10.3390/ijms140612780

de Kraker, J. W., Franssen, M. C. R., de Groot, A., Konig, W. A., and Bouwmeester, H. J. (1998). (+)-Germacrene a biosynthesis - the committed step in the biosynthesis of bitter sesquiterpene lactones in chicory. *Plant Physiol.* 117, 1381–1392. doi: 10.1104/pp.117.4.1381

Decaestecker, W., Andrade Buono, R., Pfeiffer, M., Vangheluwe, N., Jourquin, J., Karimi, M., et al. (2019). CRISPR-TSKO: a technique for efficient mutagenesis in specific cell types, tissues, or organs in arabidopsis. *Plant Cell* 31:2868–87. doi: 10.1105/tpc.19.00454

Delporte, M., Legrand, G., Hilbert, J.-L., and Gagneul, D. (2015). Selection and validation of reference genes for quantitative real-time PCR analysis of gene expression in *Cichorium intybus*. *Front. Plant Sci.* 6:651. doi: 10.3389/fpls.2015.00651

Deryckere, D., Eeckhaut, T., Van Huylenbroeck, J., and Van Bockstaele, E. (2012). Low melting point agarose beads as a standard method for plantlet regeneration from protoplasts within the *Cichorium genus*. *Plant Cell Rep.* 31, 2261–2269. doi: 10.1007/s00299-012-1335-8

Desmet, S., Dhooghe, E., De Keyser, E., Quataert, P., Eeckhaut, T., Van Huylenbroeck, J., et al. (2020). Segregation of rol genes in two generations of sinningia speciosa engineered through wild type *Rhizobium rhizogenes*. *Front. Plant Sci.* 11:859. doi: 10.3389/fpls.2020.00859

Doench, J. G., Hartenian, E., Graham, D. B., Tothova, Z., Hegde, M., Smith, I., et al. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9- mediated gene inactivation. *Nat. Biotechnol.* 32, 1262–1267. doi: 10.1038/nbt.3026.Rational

Doyle, J., and Doyle, J. (1990). DNA isolation from small amount of plant tissue. *Phytochem. Bull.* 57, 13–15. doi: 10.2307/4119796

Drewnowski, A., and Gomez-Carneros, C. (2000). Bitter taste, phytonutrients, and the consumer: a review. *Am. J. Clin. Nutr.* 72, 1424–1435. doi: 10.1093/ajcn/72.6.1424

Fan, D., Liu, T., Li, C., Jiao, B., Li, S., Hou, Y., et al. (2015). Efficient CRISPR/Cas9-mediated targeted mutagenesis in populus in the first generation. *Sci. Rep.* 5:12217. doi: 10.1038/srep12217

Heller, R. (1953). Recherche sur la nutrition minérale des tissus végétaux cultivés *in vitro*. *Ann. Sci. Nat. Bot. Biol. Veg.* 14, 1–223

Houbaert, A., Zhang, C., Tiwari, M., Wang, K., de Marcos Serrano, A., Savatin, D. V., et al. (2018). POLAR-guided signalling complex assembly and localization drive asymmetric cell division. *Nature* 563, 574–578. doi: 10.1038/s41586-018-0714-x

Jaganathan, D., Ramasamy, K., Sellamuthu, G., Jayabalan, S., and Venkataraman, G. (2018). CRISPR for crop improvement: an update review. *Front. Plant Sci.* 9:985. doi: 10.3389/fpls.2018.00985

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi: 10.1126/science.1225829

Kaur, N., Alok, A., Shivani, Kaur, N., Pandey, P., Awasthi, P., et al. (2018). CRISPR/Cas9-mediated efficient editing in phytoene desaturase (PDS) demonstrates precise manipulation in banana cv. rasthali *genome. Funct. Integr. Genomics* 18, 89–99. doi: 10.1007/s10142-017-0577-5

Lampropoulos, A., Sutikovic, Z., Wenzl, C., Maegele, I., Lohmann, J. U., and Forner, J. (2013). GreenGate - a novel, versatile, and efficient cloning system for plant transgenesis. *PLoS ONE* 8:e83043. doi: 10.1371/journal.pone.0083043

Larkin, P. J., and Scowcroft, W. R. (1981). Somaclonal variation - a novel source of variability from cell cultures for plant improvement. *Theor. Appl. Genet.* 60, 197–214. doi: 10.1007/BF02342540

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, J., Liao, X., Zhou, S., Liu, S., Jiang, L., and Wang, G. (2018). Efficient protoplast isolation and transient gene expression system for Phalaenopsis hybrid cultivar 'Ruili Beauty'. *Vitr. Cell. Dev. Biol. - Plant* 54, 87–93. doi: 10.1007/s11627-017-9872-z

Li, J.-F., Norville, J. E., Aach, J., McCormack, M., Zhang, D., Bush, J., et al. (2013). Multiplex and homologous recombination–mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.* 31, 688–691. doi: 10.1038/nbt.2650

Lin, C.-S., Hsu, C.-T., Yang, L.-H., Lee, L.-Y., Fu, J.-Y., Cheng, Q.-W., et al. (2018). Application of protoplast technology to CRISPR/Cas9 mutagenesis: from single cell mutation detection to mutant plant regeneration. *Plant Biotechnol. J.* 16, 1295–310. doi: 10.1111/pbi.12870

Liu, H. J., Jian, L., Xu, J., Zhang, Q., Zhang, M., Jin, M., et al. (2020). High-throughput CRISPR/Cas9 mutagenesis streamlines trait gene identification in maize. *Plant Cell* 32, 1397–1413. doi: 10.1105/tpc.19.00934

Liu, Q., Majdi, M., Cankar, K., Goedbloed, M., Charnikhova, T., Verstappen, F. W., et al. (2011). Reconstitution of the costunolide

biosynthetic pathway in yeast and *Nicotiana benthamiana*. *PLoS ONE* 6:e23255. doi: 10.1371/journal.pone.0023255

Manghwar, H., Lindsey, K., Zhang, X., and Jin, S. (2019). CRISPR/Cas system: recent advances and future prospects for genome editing. *Trends Plant Sci*. 24, 1102–1125. doi: 10.1016/j.tplants.2019.09.006

Martin, M. (2011). Cutadapt removes adapter sequences from high-throuhput sequencing reads. *EMBnet. J*. 1–3. doi: 10.14806/ej.17.1.200

Morel, G., and Wetmore, R. (1951). Fern callus tissue culture. *Am. J. Bot*. 38, 141–143.

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bio addays with tobacco tissue cultures. *Physiol. Plant*. 15, 473–497.

Naim, F., Dugdale, B., Kleidon, J., Brinin, A., Shand, K., Waterhouse, P., et al. (2018). Gene editing the phytoene desaturase alleles of Cavendish banana using CRISPR/Cas9. *Transgenic Res*. 27, 451–460. doi: 10.1007/s11248-018-0083-0

Nishitani, C., Hirai, N., Komori, S., Wada, M., Okada, K., Osakabe, K., et al. (2016). Efficient genome editing in apple using a CRISPR/Cas9 system. *Sci. Rep*. 6:31481. doi: 10.1038/srep31481

Odipio, J., Alicai, T., Ingelbrecht, I., Nusinow, D. A., Bart, R., and Taylor, N. J. (2017). Efficient CRISPR/Cas9 genome editing of phytoene desaturase in cassava. *Front. Plant Sci*. 8:1780. doi: 10.3389/fpls.2017.01780

Otto, F. (1997). DAPI staining of fixed cells for high-resolution flow cytometly of nuclear DNA. *Methods Cell Biol*. 33, 105–110.

Peters, A. M., Haagsma, N., and Van Amerongen, A. (1997). A pilot study on the effects of cultivation conditions of chicory (C*ichorium intybus L.)* roots on the levels of sesquiterpene lactones in chicons. *Eur. Food Res. Technol*. 205, 143–147. doi: 10.1007/s002170050142

Puchta, H. (2005). The repair of double-strand breaks in plants: Mechanisms and consequences for genome evolution. *J. Exp. Bot*. 56, 1–14. doi: 10.1093/jxb/eri025

Raulier, P., Maudoux, O., Notté, C., Draye, X., and Bertin, P. (2015). Exploration of genetic diversity within *Cichorium endivia* and *Cichorium intybus* with focus on the gene pool of industrial chicory. *Genet. Resour. Crop Evol*. 63, 243–259. doi: 10.1007/s10722-015-0244-4

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., et al. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotchnol*.31, 686–688. doi: 10.1038/nbt.2652

Shen, M. W., Arbab, M., Hsu, J. Y., Worstell, D., Culbertson, S. J., Krabbe, O., et al. (2018). Predictable and precise template-free editing of pathogenic variants. *Nature* 563, 645–651. doi: 10.1038/s41586-018-0686-x

Street, R. A., Sidana, J., and Prinsloo, G. (2013). *Cichorium intybus*: traditional uses, phytochemistry, pharmacology, and toxicology. *Evid. Based. Complement. Alternat. Med*. 2013:579319. doi: 10.1155/2013/579319

Thomson, J. G., Cook, M., Guttman, M., Smith, J., and Thilmony, R. (2011). Novel sulI binary vectors enable an inexpensive foliar selection method in *Arabidopsiss*. *BMC Res. Notes* 4:44. doi: 10.1186/1756-0500-4-44

Tian, S., Jiang, L., Gao, Q., Zhang, J., Zong, M., Zhang, H., et al. (2017). Efficient CRISPR/Cas9-based gene knockout in watermelon. *Plant Cell Rep*. 36, 399–406. doi: 10.1007/s00299-016-2089-5

Upadhyay, S. K., Kumar, J., Alok, A., and Tuli, R. (2013). RNA-Guided genome editing for target gene mutations in wheat. *G3* 3, 2233–2238. doi: 10.1534/g3.113.008847

Wook Woo, J., Kim, J., Kwon, S., Corvalan, C., Woo Cho, S., Kim, H., et al. (2015). DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat. Biotechnol*. 33, 1162–1164. doi: 10.1038/nbt.3389

Yao, L., Liao, X., Gan, Z., Peng, X., Wang, P., Li, S., et al. (2016). Protoplast isolation and development of a transient expression system for sweet cherry (Prunus avium L.). *Sci. Hortic*. 209, 14–21. doi: 10.1016/j.scienta.2016.06.003

Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014). PEAR: a fast and accurate Illumina paired-End reAd mergeR. *Bioinformatics* 30, 614–620. doi: 10.1093/bioinformatics/btt593

Zhang, K., Raboanatahiry, N., Zhu, B., and Li, M. (2017). Progress in genome editing technology and its application in plants. *Front. Plant Sci*. 8:177. doi: 10.3389/fpls.2017.00177

Zhang, Y., Liang, Z., Zong, Y., Wang, Y., Liu, J., Chen, K., et al. (2016). Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA. *Nat. Commun*. 7:12617. doi: 10.1038/ncomms12617

Zhang, Y., Malzahn, A. A., Sretenovic, S., and Qi, Y. (2019). The emerging and uncultivated potential of CRISPR technology in plant science. *Nat. Plants* 5, 778–794. doi: 10.1038/s41477-019-0461-5

Zhang, Y., Su, J., Duan, S., Ao, Y., Dai, J., Liu, J., et al. (2011). A highly efficient rice green tissue protoplast system for transient gene expression and studying light/chloroplast-related processes. *Plant Methods* 7, 1–14. doi: 10.1186/1746-4811-7-30

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 De Bruyn, Ruttink, Eeckhaut, Jacobs, De Keyser, Goossens and Van Laere. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genome Editing

ORIGINAL RESEARCH
published: 23 December 2020
doi: 10.3389/fgeed.2020.605614

CRISPR-Cas9-Mediated Mutagenesis of the Rubisco Small Subunit Family in *Nicotiana tabacum*

*Sophie Donovan[1], Yuwei Mao[1], Douglas J. Orr[2], Elizabete Carmo-Silva[2] and Alistair J. McCormick[1]**

[1] SynthSys and Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh, Edinburgh, United Kingdom, [2] Lancaster Environment Centre, Lancaster University, Lancaster, United Kingdom

Engineering the small subunit of the key CO_2-fixing enzyme Rubisco (SSU, encoded by *rbcS*) in plants currently poses a significant challenge, as many plants have polyploid genomes and SSUs are encoded by large multigene families. Here, we used CRISPR-Cas9-mediated genome editing approach to simultaneously knock-out multiple *rbcS* homologs in the model tetraploid crop tobacco (*Nicotiana tabacum cv.* Petit Havana). The three *rbcS* homologs *rbcS_S1a, rbcS_S1b* and *rbcS_T1* account for at least 80% of total *rbcS* expression in tobacco. In this study, two multiplexing guide RNAs (gRNAs) were designed to target homologous regions in these three genes. We generated tobacco mutant lines with indel mutations in all three genes, including one line with a 670 bp deletion in *rbcS-T1*. The Rubisco content of three selected mutant lines in the T_1 generation was reduced by *ca.* 93% and mutant plants accumulated only 10% of the total biomass of wild-type plants. As a second goal, we developed a proof-of-principle approach to simultaneously introduce a non-native *rbcS* gene while generating the triple SSU knockout by co-transformation into a wild-type tobacco background. Our results show that CRISPR-Cas9 is a viable tool for the targeted mutagenesis of *rbcS* families in polyploid species and will contribute to efforts aimed at improving photosynthetic efficiency through expression of superior non-native Rubisco enzymes in plants.

Keywords: chloroplast, *Chlamydomas reinhardtii*, photosynthesis, agroinfiltration, SpCas9, tobacco

OPEN ACCESS

Edited by:
Seiichi Toki,
National Institute of Agrobiological Sciences, Japan

Reviewed by:
Anshu Alok,
Panjab University, India
Huawei Zhang,
Peking University, China

***Correspondence:**
Alistair J. McCormick
alistair.mccormick@ed.ac.uk

Specialty section:
This article was submitted to Genome Editing in Plants, a section of the journal Frontiers in Genome Editing

Received: *12 September 2020*
Accepted: *27 November 2020*
Published: *23 December 2020*

Citation:
Donovan S, Mao Y, Orr DJ, Carmo-Silva E and McCormick AJ (2020) CRISPR-Cas9-Mediated Mutagenesis of the Rubisco Small Subunit Family in Nicotiana tabacum. Front. Genome Ed. 2:605614. doi: 10.3389/fgeed.2020.605614

INTRODUCTION

The assimilation of CO_2 in photosynthetic organisms is primarily catalyzed by the bi-functional enzyme ribulose-1,5-biphosphate carboxylase/oxygenase (Rubisco). In plants, Rubisco has a relatively slow carboxylation rate (k^c_{cat}) and a competitive oxygenase activity that results in yield limitations, particularly in C_3 plants, which include important crops such as *Oryza sativa* (rice) and *Triticum aestivum* (wheat). Variations in the catalytic properties of Rubisco between different species [e.g., carboxylation turnover rate (k^c_{cat}) and the specificity of Rubisco for CO_2 vs. O_2 ($S_{C/O}$)] suggest that Rubisco could be engineered to improve the efficiency of CO_2 assimilation in plants (Zhu et al., 2004; Galmés et al., 2014; Orr et al., 2016; Young et al., 2016; Martin-Avila et al., 2020).

Rubisco in plants (i.e., Form I Rubisco, L_8S_8) is composed of eight chloroplast-encoded (*rbcL* gene) large subunits (LSUs) that form the active sites, and eight small subunits (SSUs) that are nuclear-encoded by a family of *rbcS* genes (Spreitzer, 2003). Although structurally distant

from the active site, SSUs are known to affect the catalytic properties of Rubisco (Genkov and Spreitzer, 2009; Ishikawa et al., 2011; Esquivel et al., 2013; Fukayama et al., 2019; Orr et al., 2020). Arabidopsis mutants that lack up to three out of four *rbcS* homologs have proven useful models for the expression of non-native SSUs to examine the effect of divergent sequences on Rubisco catalysis (Izumi et al., 2012; Atkinson et al., 2017; Khumsupan et al., 2020). However, replacing the native SSU family remains difficult in polyploid species (i.e., most crops), which can have up to 22 *rbcS* homologs (e.g., wheat) and tend to produce near-identical mature SSU peptides. Recently, a family of phylogenetically distinct *rbcS* homologs were identified in *Nicotiana tabacum* (tobacco), rice, and several other species that produce Rubisco with altered catalytic properties, including an increased k^{c}_{cat} and decreased $S_{C/O}$ (Morita et al., 2014, 2016; Laterre et al., 2017; Pottier et al., 2018). Although these *rbcS* homologs are typically expressed in non-photosynthetic tissues, overexpression could lead to changes in the catalytic properties of the Rubisco pool in leaves, provided that the remaining *rbcS* family members are sufficiently suppressed (Morita et al., 2016).

Tobacco and rice plants with reduced Rubisco content through antisense suppression of *rbcS* have offered an insight into the extent of Rubisco limitation on photosynthesis and growth and in response to different light intensities, nitrogen availability, and temperatures (Stitt and Schulze, 1994; Makino et al., 1997, 2000). For example, *rbcS* antisense tobacco mutants have shown that Rubisco content could be decreased to 40% of wild-type levels before impairment of growth and photosynthesis under controlled growth conditions (*ca.* 300 μmol photons m^{-2} s^{-1}) (Quick et al., 1991a; Stitt et al., 1991). Although antisense studies have greatly advanced our understanding of Rubisco limitation in plants, the effectiveness of supression varies between plants, tissues, and developmental stages, and a loss of supression can occur in later generations (Quick et al., 1991b; Mitchell et al., 2004). As a result, the Rubisco content of each plant needs to be determined in every experiment, which affects extending this approach to engineer additional crop species and limits the potential to test large suites of candidate SSUs. Therefore, there is a need to generate lines with a stable suppression of endogenous SSUs to provide a platform to test SSU engineering approaches in crops.

Rubisco also requires several chaperone proteins for assembly, with significant progress made in recent years in establishing the assembly process in plants. For example, identifying and characterizing the roles of chaperone proteins has improved the efficiencies of producing chimeric Rubisco enzymes in tobacco and allowed the assembly of Rubisco from *Arabidopsis thaliana* (Arabidopsis) in *Escherichia coli* (Whitney et al., 2015; Aigner et al., 2017). Furthermore, enhancing plant productivity and robustness by increasing Rubisco abundance is now achievable in *Zea mays* (maize) through nuclear overexpression of native LSU, SSU and the RAF1 chaperone (Salesse-Smith et al., 2018, 2020) and in *O. sativa* (rice) through co-expressing an additional native SSU (Yoon et al., 2020). Recent work by Martin-Avila et al. (2020) described a next-generation tobacco mutant line (tobRrΔs) in which native Rubisco production was substituted with Rubisco from *Rhodospirillum rubrum* and native *rbcS* gene expression was blocked. Although the tobRrΔs mutant is an exciting screening platform for non-native Rubisco variants, routine expression of non-native Rubisco variants in wild-type crop plant backgrounds remains a significant challenge.

This goal is now feasible owing to the development of RNA-guided endonucleases (RGENs), such as CRISPR-Cas9, which facilitate the editing of multiple genes simultaneously in polyploid species (Morineau et al., 2017; Wolabu et al., 2020). Several toolkits have been developed for assembling plasmid vectors carrying multiple gRNA expression cassettes to target different genes (Xing et al., 2014; Lowder et al., 2015; Ma et al., 2015). Alternatively, gene families that share high nucleotide identity can be edited using one or more "promiscuous" gRNAs that target homologous regions (Endo et al., 2015). This approach was recently used to successfully induce frameshift mutations in *rbcS* genes in diploid rice (Matsumura et al., 2020). Here, we designed a CRISPR-Cas9 approach targeting the three predominant *rbcS* homologs in tobacco to explore the potential application of RGEN-mediated multigene editing of *rbcS* genes in a large, allotetraploid crop genome. The tobacco *rbcS* family comprises at least 13 homologs, and the three genes *rbcS-T1*, *rbcS-S1a,* and *rbcS-S1b* are reported to account for over 80% of total *rbcS* transcripts (Lin et al., 2020). We targeted these three genes and generated a tobacco triple SSU knockout mutant with reduced Rubisco content as a platform for heterologous SSU expression studies. We then tested a co-transformation strategy to simultaneously introduce a non-native *rbcS* gene while generating the triple SSU knockout, which reduces the need for multiple rounds of transformation and screening, and could benefit similar approaches in crop species that take longer to transform.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

Seeds of wild-type tobacco (*N. tabacum* cv. Petite Havana) and transgenic lines generated in this study were sown on a compost and sand mix (F2+ S; Levington, UK). Seeds were germinated in a controlled environment growth chamber (AR-36L3; Percival Scientific, USA) at 25°C and 60–70% relative humidity in a 16-h photoperiod with cool white fluorescent bulbs (170 μmol photons m^{-2} s^{-1}). Fourteen-day-old seedlings were transplanted to pots (3L capacity) and maintained in a greenhouse (20–21°C day; 18°C night) in a 15-h photoperiod under natural light supplemented with 300 μmol photons m^{-2} s^{-1} of light. Plant positions were rotated every 2 days to allow consistent access to light and supplemented weekly with Hoagland solution (Hoagland and Snyder, 1933).

gRNA Design

Two gRNAs with target sites common to exons 1 (gRNA1) and 4 (gRNA4) in *rbcS-S1a* (KM025316.1), *rbcS-S1b* (KM025317.1) and *rbcS-T1* were identified by the Cas-Designer tool (www.rgenome.net/cas-designer) as potential gRNA sites for editing by Cas9 from *Streptococcus pyogenes* (*spCas9*) using a dual gRNA approach (Bae et al., 2014; Park et al., 2015). We checked for potential off-target sites in the tobacco genome using

TABLE 1 | Genomic locations of thirteen Rubisco small subunit genes in tobacco.

			N. tabacum TN90[b]		*N. tabacum* v1.0[c]	
No.	**Gene**	**Accession[a]**	**Scaffold**	**Location (bp)**	**Chr.**	**Location (bp)**
1	*rbcS-S1a*	KM025316.1	SS1336	810318–811169	Nt21	11725909–11726259
2	*rbcS-S1b*	KM025317.1	Maps to same region as *rbcS-S1a*	-	-	-
3	*rbcS-S2*	KM025319.1	SS4468	754873–755617	Nt03	46963643–46964387
4	*rbcS-S3*	KM025321.1	SS4468	554204–554937	Nt03	46873734–46874467
5	*rbcS-S4*	KM025323.1	SS4468	399989–400743	Nt03	46911953–46912707
6	*rbcS-S5*	KM025325.1	SS4468	456463–457081	Nt03	46771193–46772295
7	*rbcS-T1*	KM025327.1	SS2179	301404–302119	Nt14	90863242–90863574
8	*rbcS-T2*	KM025329.1	SS17012	102405–102957	Nt17	208193244–208193991
9	*rbcS-T3a*	KM025331.1	Maps to same region as *rbcS-T2*	-	Nt17	208121436–208122184
10	*rbcS-T3b*	KM025332.1	Maps to same region as *rbcS-T2*	-	Nt17	Maps to same region as *rbcS-T3a*
11	*rbcS-T4a*	KM025334.1	SS17012	88923–89705	Nt17	208180177–208180959
12	*rbcS-T4b*	KM025335.1	Maps to same region as *rbcS-T4a*			
13	*rbcS-T5*	KM025337.1	SS17012	138156–139280	Nt17	208193500–208193991

The tobacco (N. tabacum) genome is allotetraploid with component diploid maternal and paternal genomes S and T, respectively, likely arising from hybridization between diploid N. sylvestris (S) and N. tomentosiformis (T) ancestors. Six Rubisco small subunit (rbcS) genes are on the S genome, and seven are on the T genome. Partial-coding sequences from (Gong et al., 2014[a]) were used to BLAST search two genome assemblies on the Sol Genomics database. The regions on the TN90 (Sierro et al., 2014[b]) and v.10 (Edwards et al., 2017[c]) genomes with the highest nucleotide identity to the target sequences were used to design gene-specific primers for this study.

the Cas-OFFinder tool (www.rgenome.net) and confirmed that the two gRNA sequences had no complementarity to any of the other ten *rbcS* homologs (**Table 1**) (Bae et al., 2014). A total of three potential off-target sites with two mismatches to the gRNA sequences were identified (**Supplementary Table 1**). The potential off-target sites were not evaluated further for off-target mutations as the mismatches were located in the 8–12 nt region proximal to the PAM site (Hahn and Nekrasov, 2019).

Plasmid Design and Construction

All cloning reactions were performed in a 20 μL volume following the Golden Gate assembly protocol previously described (Engler et al., 2014). Plasmids pICSL90010 (Addgene #117520), pICSL90002 (Addgene #68261), pEPOR0SP0013 (Addgene #117521), and The MoClo Plant Parts Kit (Addgene kit # 1000000044) were gifts from Nicola Patron (Earlham Institute, UK) (Engler et al., 2014; Lawrenson et al., 2015; Raitskin et al., 2019). Full-length gRNA sequences were amplified from a template plasmid that contained the gRNA scaffold sequence (pICSL90010) and primers (IDT, Germany) that included a 19-nt protospacer region preceding the PAM site using Q5 High-Fidelity DNA Polymerase (M0491S; New England BioLabs, USA) as per the manufacturer's instructions (**Supplementary Table 2**). Each full-length gRNA sequence was assembled with the Arabidopsis *U6* gene promoter (pICSL90002) into a Level 1 entry vector. A coding sequence for *spCas9* optimized for expression in plants (pEPOR0SP0013) was assembled with the Arabidopsis *ubiquitin 10* promoter and 5′ untranslated region (UTR) (pICSL12015) and *heat shock protein 18.2* (HSP) terminator and 3′ UTR into a Level 1 entry vector (Nagaya et al., 2010). The four Level 1 assemblies carrying expression cassettes for spCas9, each of the two gRNAs, and kanamycin resistance were combined into a single Level 2 binary vector (pAGM4723) to produce the plasmid for plant transformation (pGRNA14) (**Supplementary Figure 1**). The vector used to express *rbcS2* from Chlamydomonas (*CrrbcS2*) was assembled by cloning the *CrrbcS2* coding sequence fused with the *rbcS1A* chloroplastic transit peptide sequence from Arabidopsis (Atkinson et al., 2017) into a Level 1 entry vector with the *S. lycopersicum rbcS2* (*SlrbcS2*) gene promoter and 5′ UTR (pICH71301) and HSP terminator into a Level 1 entry vector. The Level 1 *CrrbcS2* expression cassette was assembled with a hygromycin resistance cassette into a Level 2 binary vector for plant transformation (pRBCS-Cr) (**Supplementary Figure 2**).

Tobacco Transformation

Vectors were transformed into *Agrobacterium tumefaciens* strain AGL1 by electroporation and colonies were verified by PCR and sequencing using insert-specific primers (**Supplementary Table 2**). For transient expression in tobacco leaves, a 15 mL culture was prepared, resuspended in 10 mM $MgCl_2$ to an OD_{600} of 0.8. Diluted cultures were syringe-infiltrated into the youngest fully expanded leaves of four-week-old plants. Stable CRISPR-Cas9 lines were produced by germinating sterile wild-type seeds in Magenta GA-7 boxes (V8505; Sigma Aldrich, UK) on 0.8% (w/v) agar (pH 5.8) containing Murashige and Skoog (MS) medium (M5524; Sigma Aldrich) and 3% (w/v) sucrose. A 150 mL suspension of AGL1 containing vector pGRNA14 was prepared and resuspended in the same volume of 1x liquid MS. Leaves from 6-week-old plants were cut into 2 cm^2 pieces, incubated for 30 min in the AGL1 suspension, and placed abaxial side up on MS medium containing 0.1 mg/L indole-3-butyric acid (IBA) (57310; Sigma Aldrich) and 1 mg/L 6- benzylaminopurine (B3408; Sigma Aldrich). After 2 days of co-cultivation with *Agrobacterium*, leaf discs were washed three times in liquid

MS and cultured on selective media with 500 mg/L augmentin and 100 mg/L kanamycin to select for pGRNA14. Shoots were excised after 4–5 weeks and placed on MS medium with 100 mg/L kanamycin in Magenta GA-7 boxes for root induction. Following the appearance of roots, kanamycin-resistant plantlets were transferred to pots (9 cm diameter) and leaf tissue was harvested to screen for mutations in T_0 lines by PCR (see section Mutation Screening). Seeds from T_0 plants were germinated on soil and screened for mutations by PCR to obtain T_1 plants for the growth analysis.

Stable CRISPR-Cas9 lines overexpressing *CrrbcS2* were generated as previously described except that explants were cultured on media that contained 100 mg/L kanamycin and 30 mg/L hygromycin to select for pGRNA14 and pRBCS-Cr, respectively. The T_1 generation of plants was obtained by germinating seeds on MS that contained 30 mg/L hygromycin to select for pRBCS-Cr and screened for mutations by PCR.

Mutation Screening

Genomic DNA was extracted as previously described (Khumsupan et al., 2020). Kanamycin-resistant plantlets were first screened by PCR using primers for *spCas9* to confirm the presence of the transgene. Gene-specific primers for the *rbcS* genes were designed based on the tobacco draft genomes and used to amplify *rbcS-T1* (rbcS-T1_F1 and rbcS-T1_R1) and *rbcS-S1a/b* (rbcS-S1_F1 and rbcS-S1_R1) (Sierro et al., 2014; Edwards et al., 2017). All primer sequences are given in **Supplementary Table 2** (Gong et al., 2014). Sanger sequencing of PCR amplicons was performed by Edinburgh Genomics (Edinburgh, UK). Mutations were identified by pairwise sequence alignment with the respective wild-type genomic DNA sequences using EMBOSS Needle (EMBL-EBI, UK) (Madeira et al., 2019). Mutation frequencies were determined from the sequencing chromatograms using TIDE (http://tide.deskgen.com) (Brinkman et al., 2014).

RNA Extraction and qRT-PCR

Total RNA was isolated from leaf tissue using an RNeasy Plant Mini Kit (#74904, QIAGEN) and treated with RNase free DNase I (#79254, QIAGEN) according to the manufacturer's protocol. For cDNA synthesis, 1 μg of RNA was reverse-transcribed in a 20 μL reaction according to the protocol for the GoScript Reverse Transcription System (A5003, Promega, USA). Quantitative reverse transcription PCR (qRT-PCR) reactions were prepared in a 10 μL volume that contained 4 μL cDNA (8 ng/μL), 1 μL of each primer (10 μM) and 5 μL of SYBR Mastermix (B0701, Eurogentec, Belgium) and performed on a LightCycler 480 (05015278001, Roche, Switzerland) with the following thermal cycling parameters: 95°C for 3 min, 40 cycles of 95°C for 10 s, 60°C for 20 s, 72°C for 30 s followed by a dissociation curve (66–95°C) at the end of each run. Relative expression of the target genes was calculated according to the $2^{-\Delta\Delta Ct}$ method using the tobacco *ribosomal protein L25* gene (GenBank: L18908) for normalization (Schmidt and Delaney, 2010). All primer sequences are given in **Supplementary Table 3**.

Protein Extraction and Western Blotting

Chlamydomonas reinhardtii (Chlamydomonas) cultures were provided as a gift from Attila Molnar (University of Edinburgh, UK). A cell lysate was prepared from Chlamydomonas cells according to (Atkinson et al., 2019). Leaf samples (7.9 cm^2) were harvested and immediately frozen and total protein was extracted in 25 mM Tris-HCl (pH 7.5), 1 mM EDTA, 10% (v/v) glycerol, 0.1% (v/v) TritonX-100, 150 mM NaCl, 1 mM DTT and cOmplete^{TM} EDTA-free protease inhibitor cocktail (COEDTAF, Roche). The sample was centrifuged at 5,000 g (4°C) for 5 min and a sub-sample of the supernatant (10 μL) was combined with Pierce 660 nm Protein Assay Reagent (22660, ThermoFisher, UK) to measure total soluble protein against BSA pre-diluted standards (23208, ThermoFisher). The remaining sample was mixed with 1% (w/v) LDS and 1 μL (per 100 μL) β-mercaptoethanol, and heated to 100°C for 1 min. Total soluble protein was separated by SDS-PAGE on 12% Bis-Tris gels (NP0342, Invitrogen, USA) and transferred to a PVDF membrane using iBlot2 gel transfer (IB21001, Invitrogen). Membranes were probed with rabbit serum raised against wheat Rubisco (Howe et al., 1982) at a 1:10000 dilution, RbcS2 from Chlamydomonas (CrRbcS2) (raised to the C-terminal region of the SSU (KSARDWQPANKRSV) by Eurogentec, Southampton, UK) at 1:1000 dilution, histone H3 (ab18521, Abcam, UK) at a 1:10000 dilution, or actin (60008-1-1G, Proteintech, USA) at a 1:1000 dilution. A 1:10000 dilution of IRDye 800CW goat anti-rabbit IgG (LI-COR, USA) was used to visualize bands on an Odyssey Clx Imager (LI-COR) that were quantified with Image Studio Lite software (v. 5.2.5, LI-COR).

Rubisco Content

Leaf samples (5.9 cm^2) were collected from the youngest fully expanded leaves of 4-week-old plants, frozen on liquid nitrogen and stored at −80 °C before extraction. Samples were ground rapidly in an ice-cold mortar and pestle in 250 μL of extraction buffer (50 mM Bicine-NaOH pH 8.2, 20 mM MgCl2, 1 mM EDTA, 2 mM benzamidine, 5 mM ε-aminocaproic acid, 50 mM β-mercaptoethanol, 10 mM dithiothreitol, 1% (v/v) protease inhibitor cocktail (Sigma-Aldrich, Mo, USA), and 1 mM phenylmethylsulphonyl fluoride) for *ca.* 1 min followed by 1 min centrifugation (14,700 g at 4°C). The supernatant (100 μL) was then mixed with 100 μL of carboxyarabinitol-1,5-bisphosphate (CABP) binding buffer which contained 100 mM Bicine-NaOH (pH 8.2), 20 mM $MgCl_2$, 20 mM $NaHCO_3$, 1.2 mM (37 kBq μmol^{-1}) [^{14}C]CABP, incubated at RT for 25 min, and Rubisco content determined via [^{14}C]CABP binding (Sharwood et al., 2016). Bradford assay was used to determine total soluble protein in the same supernatant as prepared for Rubisco content analysis (Bradford, 1976).

Photosynthesis Measurements

The response of photosynthesis (A) to different levels of photosynthetic active radiation (PAR) (1,800, 1,500, 1,000, 500, 200, 100, 50, and 20 μmol photons m^{-2} s^{-1}) was measured at 400 μmol CO_2 mol^{-1} using a LI-COR 6400-XT portable gas exchange system (LI-COR) on 41–44 day-old plants. The response of A to the intercellular CO_2 concentration (C_i) was

measured under a saturating light intensity (1,500 μmol photons m^{-2} s^{-1}). The first measurement of each A/C_i curve was taken at an external CO_2 concentration (C_a) of 400 μmol mol^{-1} and then decreased to 50 μmol mol^{-1} in increments of 50 μmol mol^{-1}. The upper part of the A/C_i response was measured from 500 to 2,000 μmol mol^{-1} in increments of 200 μmol mol^{-1}. All measurements were taken with leaf temperature held at 25 °C and under a relative humidity of 60–70%. The maximum rate of photosynthesis at ambient levels of CO_2 (A_{sat}) was estimated from the A/PAR response as described in Monteith (1991). The maximum carboxylation rate of Rubisco (V_{cmax}) was estimated by fitting A/C_i data to a C_3 photosynthesis model (Ethier and Livingston, 2004). Dark-adapted leaves were used to determine the maximum quantum yield of photosystem II (F_v/F_m) using a Handy PEA chlorophyll fluorimeter (Hansatech Instruments, UK). F_v/F_m measurements were taken on the final day of growth experiments prior to harvesting.

Chlorophyll Content

Chlorophyll was extracted on a leaf area basis (58.9 mm^2) as described in Khumsupan et al. (2020) and quantified according to Porra et al. (1989).

Growth Measurements

Plants in the growth experiments were harvested at 45-days old for growth measurements and the leaves and stems were separated immediately and weighed to determine the fresh weight. Images of separated stems and leaves were analyzed to determine total leaf area and stem height using iDIEL Plant software and ImageJ, respectively (Schneider et al., 2012; Dobrescu et al., 2017). Samples were then dried in an oven (80°C for 7 days) and weighed to determine dry weight.

Statistical Analysis

All statistical analyses were performed using GraphPad Prism 8 software (GraphPad Software, USA). Significant differences between two groups were identified using Student's t-test ($P < 0.05$) and more than two groups were evaluated using a one-way ANOVA followed by Tukey's honestly significant difference (HSD) test ($P < 0.05$).

RESULTS

CRISPR-Cas9 Was Highly Efficient in Tobacco Transient Expression Assays

Two gRNAs (gRNA1 and gRNA4) were designed to target DNA sequence regions found only in the three most highly expressed *rbcS* homologs in tobacco (*rbcS-T1*, *rbcS-S1a*, and *rbcS-S1b*) (**Figures 1A,B**) (Lin et al., 2020). Gene-specific primers for screening edits in *rbcS-S1a* or *rbcS-S1b* were not available when the gRNA sites were initially designed, as the only available reference genes were partially sequenced, 98.5% identical, and did not map to unique regions on the tobacco KN90 genome (Sierro et al., 2014) (**Table 1**, **Supplementary Figure 3**). To try to overcome this limitation, we used a paired gRNA approach to generate a large deletion in each *rbcS* gene that could be screened using primers common to both *rbcS-S1a* and *rbcS-S1b*.

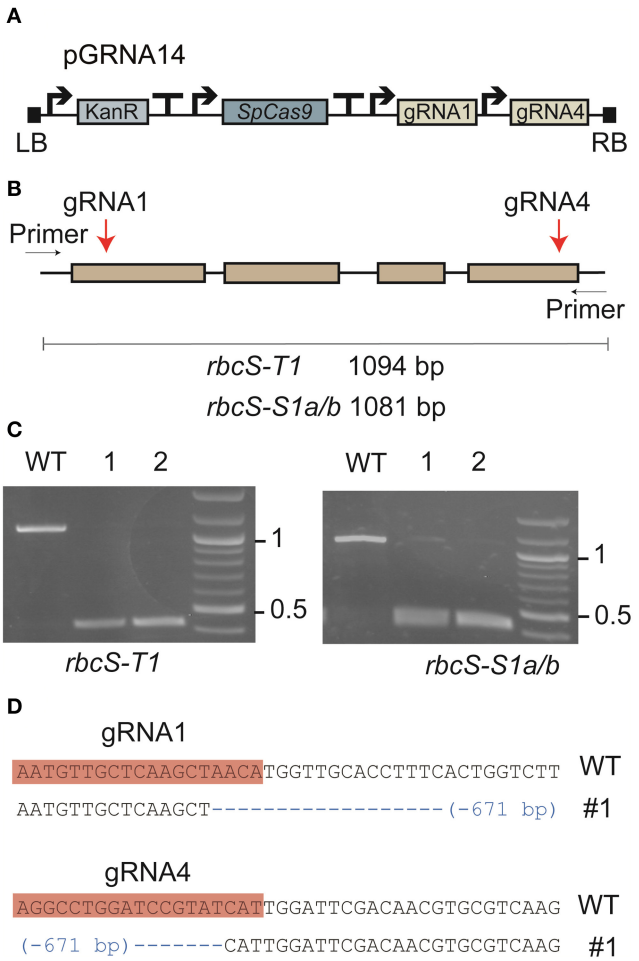


FIGURE 1 | Expression of a Cas9-gRNA vector targeting three Rubisco small subunit genes in tobacco. **(A)** Cas9-gRNA expression construct with *Cas9* from *Streptococcus pyogenes* (SpCas9) and two gRNA sequences (gRNA1 and gRNA4) (see **Supplementary Figure 1** for details of vector pGRNA14). **(B)** The target sites of gRNA1 and gRNA4 were located on the first and fourth exons, respectively, in *rbcS-T1*, *rbcS-S1a*, and *rbcS-S1b*. **(C)** PCR screen of leaf tissue with pGRNA14 transiently expressed showed deletion events (i.e., bands at <500 bp) in *rbcS-T1* (left) and *rbcS-S1a/b* (right). **(D)** Example of a pairwise sequence comparison of the deletion band from sample 1 (#1) for *rbcS-T1* showed a 671 bp deletion between the gRNA1 and gRNA4 target sites compared to wild-type (WT) *rbcS-T1*.

Vector pGRNA14 (carrying expression cassettes for gRNA1, gRNA4, and *SpCas9*) was transiently expressed in tobacco leaves by agroinfiltration to test the efficiency of the gRNAs. Subsequent amplification of DNA from agro-infiltrated leaves showed that the expected amplicons for wild-type *rbcS-T1* and *rbcS-S1a/b* were absent or barely detectable compared to the wild-type control (**Figure 1C**). Instead, smaller amplicons were observed that were consistent with a large deletion event between the gRNA 1 and gRNA four sites. The substantial reduction in intensity of the wild-type amplicons for *rbcS-S1a/b* suggested that both *rbcS-S1a* and *rbcS-S1b* had been edited. Sequencing of *rbcS-S1a/b* amplicons was not performed because

sequence-specific primers to distinguish the two orthologs were unavailable. However, sequencing of the smaller sized band for *rbcS-T1* confirmed that a 671 bp deletion had occurred between the gRNA1 and gRNA4 sites (**Figure 1D**). Therefore, transient expression assays clearly showed that gRNA1 and gRNA4 were functional and appeared highly efficient in tobacco.

Stable and Chimeric Mutations Were Identified in the T_0 Generation

Tobacco leaf disks were transformed with vector pGRNA14 and cultured on selective media to obtain kanamycin-resistant plantlets. Eight T_0 plants with varying leaf phenotypes were transferred to soil and confirmed to contain the *SpCas9* transgene by PCR using primers specific for the *SpCas9* expression cassette. Four of these had visibly smaller and paler leaves than the non-transformed tissue culture control (i.e., wild-type), which is typical of Rubisco-deficient mutants (Khumsupan et al., 2020; Martin-Avila et al., 2020) (**Figure 2A**). The remaining four plants had a mixed pale and wild-type leaf phenotype indicative of chimeric mutations. All eight plants were screened for mutations in *rbcS-T1* and *rbcS-S1a/b* by PCR using the same primers as used for the transient expression assays (**Figure 2B**). In contrast to the results for the latter, only a single plant (line 4) showed a deletion band for *rbcS-T1*, while all other amplifications were similar in size to wild-type amplicons. Sequencing of the *rbcS-T1* deletion band from line 4 confirmed a 670 bp deletion between the two gRNA sites (**Supplementary Figure 4**). Therefore, line 4 was considered homozygous for the 670 bp deletion because only a single allele was identified by PCR and sequencing.

The wild-type sized amplicons for *rbcS-T1* from the remaining seven plants were sequenced to assess if small mutations were present. Pairwise-sequence alignments between the *rbcS-T1* amplicons from the transgenic plants and wild-type suggested that more than one *rbcS-T1* allele was present in each plant. The TIDE tool was used to identify and determine the frequency of different mutations in samples that likely had more than one allele by analyzing the sequencing chromatograms from mutant and wild-type plants (**Figure 2C**) (Brinkman et al., 2014). Plants with one mutated allele in addition to the wild-type allele were considered heterozygous, plants with two mutated alleles were considered bi-allelic, and those with more than two alleles were considered chimeric (**Figure 2D**). The TIDE analysis identified mutations at the gRNA4 target site that ranged in size from +1 bp to −17 bp between the seven lines. No mutations were observed at the gRNA1 target site. Line 14 appeared to have a bi-allelic mutation in *rbcS-T1* (i.e., a 1 bp deletion and 1 bp insertion) and had pale leaves similar to line 4. The remaining six plants had three or more alleles. The wild-type allele was identified in lines 2, 3, 12, and 21, which was consistent with the chimeric leaf phenotype seen for these plants. In contrast, only mutant alleles were identified in lines 1 and 9, both of which had homogenous pale leaves.

Similar to the transient expression assays, it was not possible to inspect for mutations in *rbcS-S1a* and *rbcS-S1b* by sequencing. T_0 plant lines 1 and 14 did not establish following transfer to soil, but all other lines were progressed to the T_1 generation based on the observed phenotypes and evidence of mutations in *rbcS-T1*.

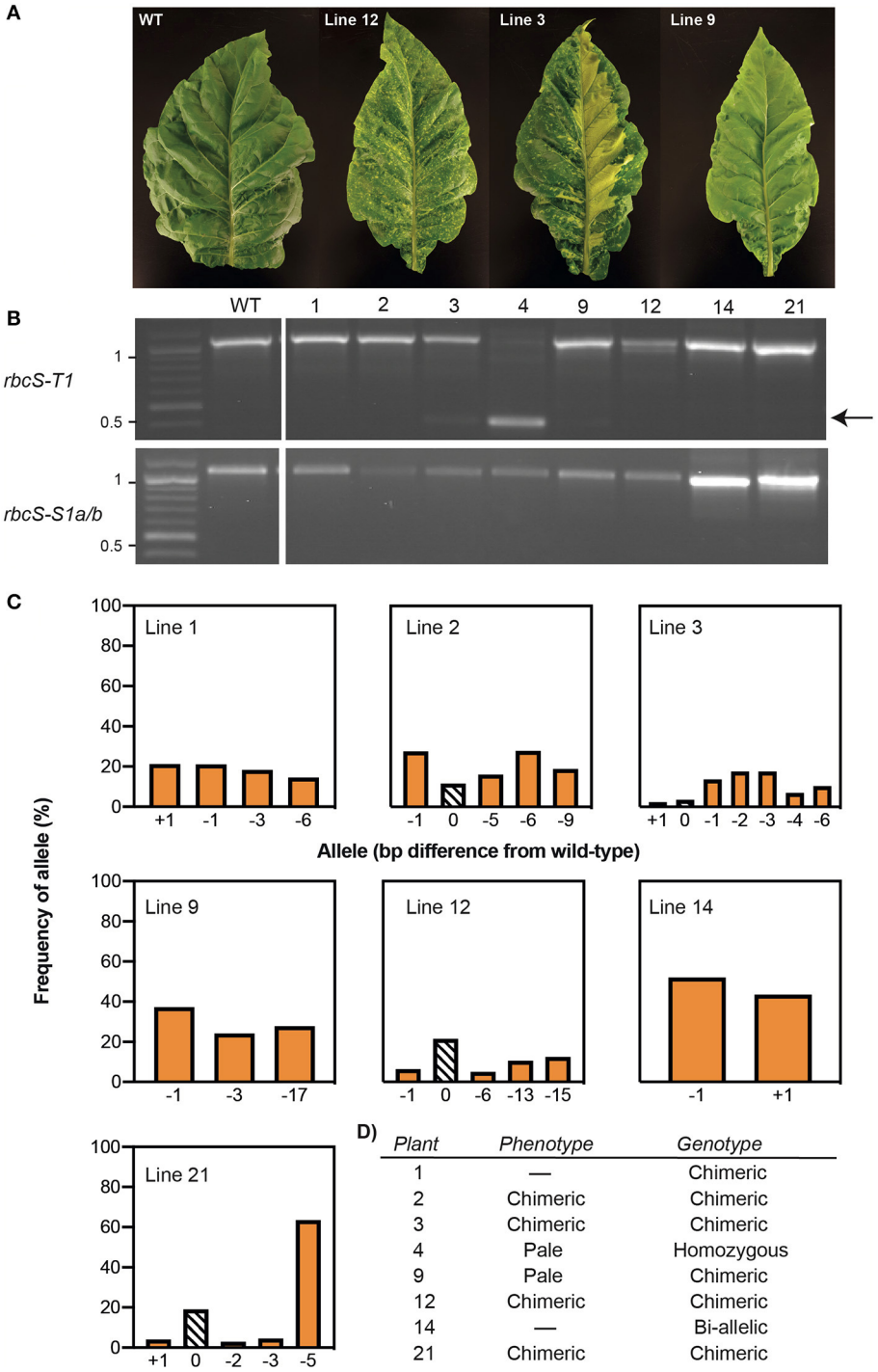
Transgene-Free Mutants Were Identified in the T_1 Generation

To identify heritable mutations in *rbcS-T1* in the T_1 generation, we first screened for the absence of the *SpCas9* transgene in the progeny of lines 2, 3, 4, 9, 12, and 21 (**Table 2**). Transgene-free T_1 plants accounted for 15% of line 2 (2/13), 79% of line 3 (15/19), 32% of line 4 (5/19), 56% of line 9 (10/18), 11% of line 12 (2/19). No transgene-free plants were identified in line 21. A range of phenotypes were observed among transgenic and transgene-free T_1 plants in the different lines. All plants for line 21 had a chimeric phenotype. Plants from lines 2 and 3 that retained *SpCas9* had pale leaves but the transgene-free progeny of these lines appeared similar to wild-type. However, all plants from lines 4, 9, and 12 had a pale leaf phenotype regardless of the absence or presence of *SpCas9*.

Owing to the consistent pale leaf phenotype, we screened for mutations in *rbcS-T1* in T_1 plants of lines 4, 9, and 12 (**Table 2**). All line four plants appeared homozygous for the 670 bp deletion allele of *rbcS*-T1. Consistent with the T_0 generation, only wild-type-sized amplicons were observed for lines 9 and 12. Sequencing the amplicons of four line nine plants showed a variety of inherited mutations around the gRNA4 target site: two plants had a 1 bp homozygous deletion, two plants had different bi-allelic mutation (i.e., a 17 bp deletion and 3 bp deletion, and a 1 bp deletion and 6 bp deletion, respectively). Amplicons of the two line 12 plants also revealed bi-allelic mutations: one plant contained a 1 bp deletion and 6 bp deletion, and the other had a 9 bp deletion and 4 bp deletion.

Given the pale leaf phenotype of T_1 plants for lines 4, 9, and 12, and the confirmed mutations in *rbcS-T1*, more in-depth characterisations were carried out on the T_1 plants for these lines. Initially, this was by determining the relative abundance of Rubisco-encoding transcripts via qRT-PCR (**Figures 3A,B**). For the three *rbcS* genes targeted for editing by *SpCas9*, transcript levels for *RbcS-T1* and *RbcS-S1a/b* in line 4 were decreased by 98% and 35%, respectively (**Figure 3A**). In contrast, expression levels were not reduced in lines 9 and 12. Although the relative expression levels of *rbcS-S2*, *rbcS-S3*, *rbcS-S4* and *rbcS-T5* were increased in some or all the three mutant lines, this had no significant impact on total relative *rbcS* abundances (**Figure 3B**). Overall, the total abundance of *rbcS* transcripts for line 4 was also reduced by 25%, while lines 9 and 12 showed no changes in *rbcS* transcript abundance compared to wild-type. All three mutant lines had a 20% reduction in transcripts encoding the large subunit of Rubisco (*rbcL*) relative to wild-type (**Figure 3B**). Therefore, the reduction in Rubisco content suggested that the three *rbcS* genes had loss-of-function mutations in all three lines.

Remarkably, the leaf Rubisco content in all three mutant lines was decreased by *ca*. 93% relative to wild-type, which corresponded to an 85 and 60% reduction in SSU and LSU, respectively (**Figures 3C,E**). Total leaf soluble protein content was reduced in lines 4, 9 and 12 by 70–80% (**Figure 3D**).



Plant	Phenotype	Genotype
1	—	Chimeric
2	Chimeric	Chimeric
3	Chimeric	Chimeric
4	Pale	Homozygous
9	Pale	Chimeric
12	Chimeric	Chimeric
14	—	Bi-allelic
21	Chimeric	Chimeric

FIGURE 2 | Phenotypes and genotypes of tobacco plants transformed with Cas9-gRNA vector pGRNA14. **(A)** T_0 plants had paler leaves than wild-type (WT) (e.g., line 9) or chimeric leaf phenotypes (e.g., lines 3 and 12). **(B)** A PCR screen of *rbcS-T1* and *rbcS-S1a/b* in eight T_0 lines identified a deletion between gRNA1 and gRNA4 in *rbcS-T1* in line 4 (arrow) and wild-type-sized amplicons for the remaining plants. **(C)** TIDE analysis of the sequencing chromatograms for wild-type-sized amplicons of *rbcS-T1* showed the frequency of wild-type (0 bp) or mutant [insertion (+N bp) or deletion –N bp)] alleles. **(D)** Summary of phenotypes and *rbcS-T1* genotypes of T_0 plants.

Our results showed that wild-type plants invested one third of total soluble protein into the Rubisco pool. Thus, the observed decreases in soluble protein for lines 4, 9, and 12 could not be accounted for by the reduction in Rubisco content alone, which indicated that the synthesis of proteins other than Rubisco was also reduced. Furthermore, the mutant lines also had a significant

TABLE 2 | Inheritance of CRISPR-Cas9 mutations in the Rubisco small subunit *rbcS-T1*.

Line	T_0 generation	T_1 generation	Alleles (T_1)
2	Chimeric	3bi;5chi	d1/d5 (2)
			d9/d6 (1)
3	Chimeric	1bi;4het;10chi	d2/i1 (1)
			d2/WT (3)
			i1/WT (1)
4	Homozygous	16hom	d670 (16)
9	Chimeric	2hom;2bi;4chi	d1 (2)
			d17/d3 (1)
			d1/**d6** (1)
12	Chimeric	6bi; 6chi	d2/**d7** (2)
			d1/d6 (1)
			d9/**d4** (3)
21	Chimeric	n/a	n/a

The number of T_1 progeny with homozygous (hom), heterozygous (het), bi-allelic (bi) or chimeric (chi) mutations is shown. Non-chimeric alleles in the T_1 generation are described as deletions (d) or insertions (i) followed by the number of base-pairs compared to wild-type (WT). Alleles that were not identified in the T_0 progenitor are shown in bold lettering. The number of progeny with a single genotype is subsequently shown in brackets.

reduction in chlorophyll per leaf area compared to wild-type (**Supplementary Figure 5**).

Decreased Rubisco Resulted in Reduced Biomass Accumulation and Lower CO_2 Assimilation Rates

The growth of lines 4, 9, and 12 was compared with that of a non-transformed tissue culture control line (i.e., wild-type plants) (**Figures 4A–D**). All three mutant lines grew slowly and accumulated <8% of the total biomass (dry weight) of wild-type plants after 45 days of growth. This was associated with a 92% reduction in height, and an 81–93% reduction in total leaf area.

At ambient CO_2 concentrations (i.e., 400 μmol mol $^{-1}$) lines 4, 9, and 12 showed similar reductions in the light-saturated rate of CO_2 assimilation at ambient CO_2 (A_{sat}) (*ca.* 42% of wild-type) (**Figure 4E**, **Table 3**). CO_2 assimilation rates also plateaued at a lower light intensity (400 μmol photons m^{-2} s^{-1}) compared to wild-type plants (1,000 μmol photons m^{-2} s^{-1}). Under saturating light (1,500 μmol photons m^{-2} s^{-1}), the response of A to changes in C_i was also affected in the three mutant lines (**Figure 4F**). The initial slope of the A/C_i response curve is associated with the carboxylation efficiency of Rubisco, and was significantly lower in the mutant lines compared to wild-type (Long and Bernacchi, 2003). Furthermore, the maximum rate of Rubisco carboxylation (V_{cmax}) was more than 50% lower than wild-type in the mutant lines.

Co-transformation Facilitated Simultaneous Knockout and Introduction of Rubisco Small Subunits

To test if we could simultaneously reduce native Rubisco content and introduce a novel Rubisco SSU, we co-transformed wild-type tobacco with the plasmid vector pGRNA14 and a second vector carrying an expression cassette for the *CrrbcS2* gene from the green alga Chlamydomonas (pRBCS-Cr) (**Supplementary Figure 2**). We first designed a suitable expression cassette for the heterologous SSU by testing three common high strength promoters in tobacco protoplasts using a dual-luciferase assay (**Supplementary Figure 6A**). The *SlrbcS2* gene promoter showed significantly higher activity than the Arabidopsis *rbcS1A* gene promoter, and the Arabidopsis *rbcS3B* gene promoter produced the lowest expression. Therefore, the *SlRbcS2* promoter was chosen to drive *CrrbcS2* expression. The *CrrbcS2* gene was previously modified for expression in higher plants, where the mature peptide was fused to a *rbcS1A* chloroplast transit peptide (Atkinson et al., 2017). Agroinfiltration of tobacco leaves with the modified *CrrbcS2* fused to a GFP-tag at the C-terminus confirmed that the heterologous SSU localized to the chloroplast (**Supplementary Figure 6B**).

Vectors pGRNA14 and pRBCS-Cr were co-transformed into wild-type tobacco and the explants were cultured on selective media containing two antibiotics (i.e., selective for each T-DNA insertion) (**Figures 5A,B**). We obtained a small number of plants that had both T-DNA cassettes integrated ($n = 5$) (**Figure 5C**). Two T_0 plants (co-transformed (CT) lines CT-3 and CT-4) had a pale leaf phenotype compared to wild-type. Neither of these two lines had a large deletion in *rbcS-T1* (**Figure 5D**) but sequencing of the *rbcS-T1* amplicon revealed multiple mutated alleles in CT-3 and two mutated alleles in CT-4 (1 bp deletion and 1 bp insertion). Therefore, the CT-3 line was chimeric and the CT-4 line was either chimeric or bi-allelic.

We germinated T_1 seeds from CT-4 on soil to screen for progeny that contained the *CrrbcS2* transgene but lacked the *SpCas9* transgene. Inheritance of the *CrrbcS2* transgene followed a 3:1 Mendelian segregation (13/20 plants). However, all CT-4 plants screened contained *SpCas9*, indicating multiple copy insertions of the *SpCas9* transgene (Tizaoui and Kchouk, 2012). As a proof-of-principle, we evaluated the mutations in *rbcS-T1* in eleven plants. The wild-type allele was not identified in any of the CT-4 plants that were screened. Four out of eleven plants contained a homozygous 1 bp insertion in *rbcS-T1* (CT-4-i4), while the remaining seven plants had a bi-allelic mutation (9 bp deletion and 1 bp insertion; CT-4-d9) (**Supplementary Figure 7A**).

Based on a new reference genome for tobacco (Edwards et al., 2017), it was then possible to design primers to differentiate between mutations in *rbcS-S1a* and *rbcS-S1b* (**Supplementary Table 2**). We re-germinated T_1 CT-4 seeds and examined four CT-4 plants, in which we identified mutations in all three target *rbcS* genes near the gRNA4 target site, including deletions in *rbcS-S1a* (1-4 bp) and *rbcS-S1b* (2-12 bp), and a bi-allelic 1 bp insertion in *rbcS-T1* (**Supplementary Figure 7B**). We also re-evaluated tissue samples from each T_0 line, four T_1 CRISPR-Cas9 line 4 plants and three T_1 line 9 plants (see section Transgene-Free Mutants Were Identified in the T1 Generation). In the T_0 generation, we again identified mutations near the gRNA4 target site, including deletions in *rbcS-S1a* and *rbcS-S1b* (1–4 bp) (**Supplementary Figure 7C**). Similarly, we found mutations in the T_1 generation for line 4 and line 9,

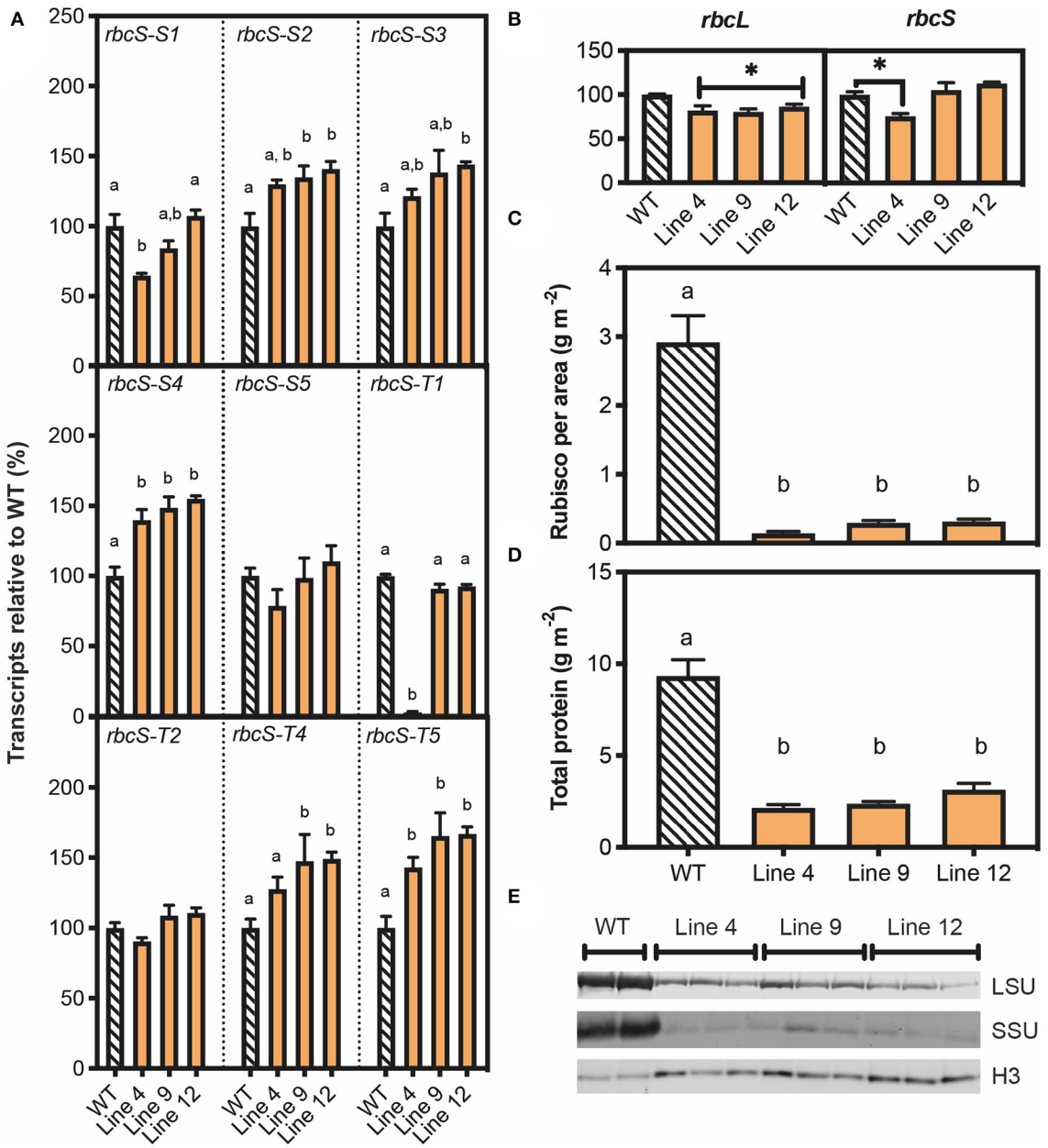


FIGURE 3 | Molecular analyses of T_1 plants with *SpCas9*-induced mutations in *rbcS-T1*. **(A)** RNA was extracted from 45-day-old plants and the abundance of transcripts for nine *rbcS* homologs was measured by quantitative PCR (qPCR) with gene-specific primers and the transcript level is shown relative to wild-type (Supplementary Information 1 in **Supplementary Material**). **(B)** Abundance of Rubisco large subunit transcripts (*rbcL)* and total *rbcS* transcripts. The latter was calculated by combining the results for the nine individual *rbcS* homologs. **(C)** Rubisco content per leaf area and **(D)** total protein per leaf area measured by Bradford assay. **(E)** Immunoblot of total soluble protein to detect Rubisco large subunit (LSU) and small subunit (SSU) contents, with a histone H3 (H3) loading control. Protein extracted from an equivalent amount of leaf area was loaded. Each lane represents a biological replicate (two for WT and three for each mutant line). All data are the mean ± SEM of three biological replicates. Significant differences between groups ($P < 0.05$) as identified by one-way ANOVA and Tukey's HSD test are shown by different letters.

including deletions in *rbcS-S1a* (1–5 bp) and *rbcS-S1b* (3-12 bp) (**Supplementary Figure 7D**). Together, these results confirmed that CT-4, line 4 and line 9 had loss-of-function mutations in the T_1 generation in all three target genes.

Immunoblotting revealed that CrRbcS2 was expressed in T_1 CT-4 plants (**Figure 5E**). The size of the band was consistent with that for CrRbcS2 expressed in Arabidopsis [i.e., previously generated line $S2_{Cr}$ (Atkinson et al., 2017)] and Chlamydomonas. Although the T_1 CT-4 plants expressing CrRbcS2 still contained the *SpCas9* transgene, we performed a preliminary growth analysis to compare this line to three CRISPR-Cas9 SSU lines (i.e., lines 4, 9, and 12) and wild-type

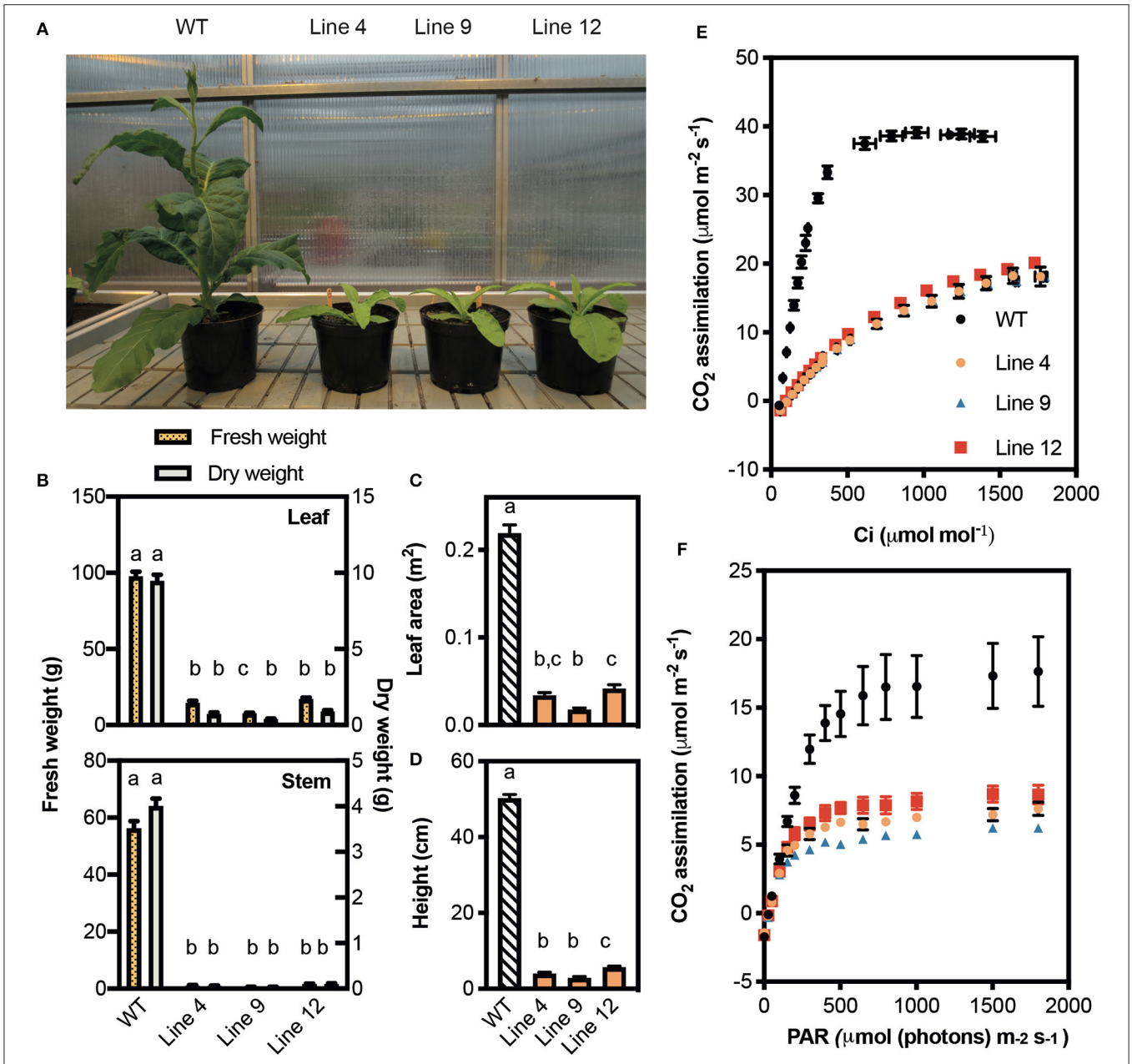


FIGURE 4 | Growth and photosynthetic rates of CRISPR-Cas9 Rubisco mutants in the T_1 generation. **(A)** Phenotypes of wild-type (WT) and three Rubisco mutant lines after 45 days of growth in a greenhouse with a 15-h photoperiod. **(B)** Fresh (left y-axis) and dry (right y-axis) weights of leaves (top panel) and stems (bottom panel). **(C)** Total leaf area and **(D)** plant height. Growth data show the mean ± SEM ($n = 6$–8) with significant differences ($P < 0.05$) indicated by different letters determined by ANOVA and Tukey's HSD tests. **(E)** The response of A to changing irradiance (PAR) measured at 25°C under 400 μmol mol^{-1} CO_2. **(F)** The response of CO_2 assimilation (A) to intracellular CO_2 concentration (C_i) measured at 25°C under saturating irradiance (1,500 μmol photons m^{-2} s^{-1}). Each data point represents the mean ± SEM of four leaves each from separate plants ($n = 4$).

plants (**Supplementary Figure 8**). This analysis showed that the CT-4 line remained smaller that wild-type but accumulated significantly more biomass relative to the CRISPR-Cas9 lines (**Supplementary Figures 8A–D**). Furthermore, the response of A to PAR was significantly higher (*ca.* 45% of wild-type) compared to the CRISPR-Cas9 lines (**Supplementary Figures 8E,F**). Together, these results indicated that expression of a non-native SSU from Chlamydomonas likely rescued reduction in Rubisco content caused by mutations in three native *rbcS* genes in tobacco. However, further analyses of additional independent CT lines lacking the *SpCas9* transgene are needed to strengthen these findings.

TABLE 3 | Photosynthetic parameters of CRISPR-Cas9 plants with reduced rubisco contents.

	WT	Line 4	Line 9	Line 12
A_{sat} (μmol CO_2 m^{-2} s^{-1})	21.6 ± 0.8[a]	9.3 ± 0.3[b]	9.4 ± 0.3[b]	9.8 ± 0.4[b]
V_{cmax} (μmol CO_2 m^{-2} s^{-1})	98.9 ± 1.2[a]	41.2 ± 5.4[b]	48.7 ± 2.2[b]	43.8 ± 7.5[b]
Γ (μmol CO_2 mol^{-1})	53.0 ± 0.5[a]	96.7 ± 1.8[b]	92.1 ± 1.8[b]	95.7 ± 2.9[b]
Initial slope (A/C_i)	0.137 ±0.004[a]	0.027±0.005[b]	0.026±0.004[b]	0.030±0.001[b]
F_v/F_m	0.85 ± 0.03[a]	0.76 ± 0.02[b]	0.67 ± 0.02[c]	0.77 ± 0.02[b]

For measurements of net photosynthetic CO_2 assimilation (A) response to photosynthetically active radiation (PAR) ambient CO_2 levels were maintained at 400 μmol mol^{-1}. Measurements of A in response to sub-stomatal CO_2 concentration (C_i) were done under a constant illumination of 1,500 μmol photons m^{-2} s^{-1}. All measurements were taken with leaf temperature held at 25°C and under a relative humidity of 60–70%. Values from leaf gas exchange measurements represent the mean ± SEM (n = 4). Dark-adapted leaves were used for F_v/F_m measurements (n = 10). Different letters (i.e. a, b or c) after each value indicate significant differences determined by ANOVA followed by Tukey's HSD tests (P < 0.05). Γ, CO_2 compensation point; A_{sat}, light-saturated CO_2 assimilation rate at ambient CO_2; F_v/F_m, maximum quantum yield of photosystem II (PSII); V_{cmax}, maximum rate of Rubisco carboxylation; WT, wild-type.

DISCUSSION

In this study we generated tobacco mutants with decreased amounts of Rubisco by targeting three *rbcS* homologs with CRISPR-Cas9. These lines have a similar decrease in Rubisco content as previous antisense tobacco lines and demonstrate the potential for RGEN-mediated editing of *rbcS* families in polyploid crop species (Khumsupan et al., 2020; Martin-Avila et al., 2020). We also co-transformed tobacco with CRISPR-Cas9 to reduce expression of the native rbcS alongside expression of a non-native rbcS, demonstrating the usefulness of this approach for efforts aimed at improving the efficiency of carbon assimilation through better Rubiscos.

Our strategy was designed to create a 670 bp deletion in three target genes to facilitate simpler screening for multiple mutations by PCR. Transient expression assays in tobacco protoplasts indicated that the dual gRNA approach created large deletions in the target genes with high efficiency. However, *in planta* the large deletion only occurred at a frequency of 12.5% (1/8 plants) in a single target gene (*rbcS-T1*). Our results showed that indel mutations at a single gRNA site were more common than a deletion between both sites in tobacco plants, which is consistent with previous studies in Arabidopsis, *Nicotiana benthamiana, Z. mays,* and *O. sativa* (Zhou et al., 2014; Ordon et al., 2017; Durr et al., 2018; Doll et al., 2019; Khumsupan et al., 2020). Transient assays offer a useful system to pre-select gRNA candidates before stable expression, as differences in the mutation efficiency of each gRNA can reduce the efficiency of deletions between two target sites (Zhou et al., 2014; Doll et al., 2019). However, in agreement with our findings, other studies have reported a lower frequency of large deletions in transgenic plants with gRNAs that appeared highly efficient in transient assays (Zhou et al., 2014; Khumsupan et al., 2020). Therefore, large deletions are feasible but a higher abundance of indels limits the use of paired gRNAs to streamline screening approaches. Reducing the size of the deletion to 50–100 bp could improve the frequency of deletions between two gRNA sites and allow detection by PCR (Ordon et al., 2017).

Few studies have reported RGEN-mediated editing in tobacco and the germline transmission rate of mutations has not yet been described (Gao et al., 2015; Xie et al., 2017). In line with previous studies, we found that most of the *rbcS-T1* alleles in T_0 plants had mutations at a single gRNA target site. Although homozygous, heterozygous, and bi-allelic mutants have been reported in several species, somatic mutations are more frequently detected (Brooks et al., 2014; Zhang et al., 2014). Non-somatic mutations are likely detected at variable frequencies because of genomic differences in the target site and the timing of DSB repair (Zhang et al., 2014). However, chimeric plants can transmit heritable mutations in germline cells to the next generation (Feng et al., 2014; Zhang et al., 2014). We found complex and variable segregation patterns of mutations in independent lines and identified heritable mutations in the transgene-free progeny of two lines (line 9 and line 12) that were chimeric in the T_0 generation. Although only a single line (line 4) had the expected 3:1 segregation ratio for *SpCas9*, our results suggest that plants with somatic mutations can be bypassed by screening for mutations in *SpCas9* segregants. As a result, it is possible to increase the likelihood of obtaining transgene-free homozygous mutants in the T_1 generation in complex polyploid species.

In addition to describing the germline transmission of mutations in tobacco, our work has generated results that offer a useful comparison between RGEN-mediated approaches and antisense technology. Consistent with the reports for antisense *rbcS* tobacco, we found that a severe reduction in Rubisco content reduced photosynthetic rates and biomass accumulation in the CRISPR-Cas9 lines (Quick et al., 1991b; Stitt et al., 1991; Masle et al., 1993; Martin-Avila et al., 2020). Evidently, the antisense tobacco lines had significantly less *rbcS* mRNA than wild-type and Rubisco content was correspondingly decreased. Although our approach produced lines with a more severe decrease in Rubisco content, total *rbcS* transcripts in the CRISPR-Cas9 lines were equivalent to wild-type except for line 4. We hypothesized that the observed reductions in *rbcS* mRNA in line 4 were due to the large deletion, which removed the forward primer binding site and could have disrupted transcription. In contrast, lines 9 and 12 had small indels in *rbcS-T1* that were unlikely to affect transcription and primer binding. The effect of the large deletion on gene transcription could be further investigated in line 4 by designing primers that anneal outside of the deletion region.

The CRISPR-Cas9 lines also had a slight suppression in *rbcL* transcripts that did not seem to be linked to the amount of *rbcS* mRNA. Antisense tobacco lines with *ca.* 12% of wild-type *rbcS* had no observable changes in the amount of *rbcL* mRNA

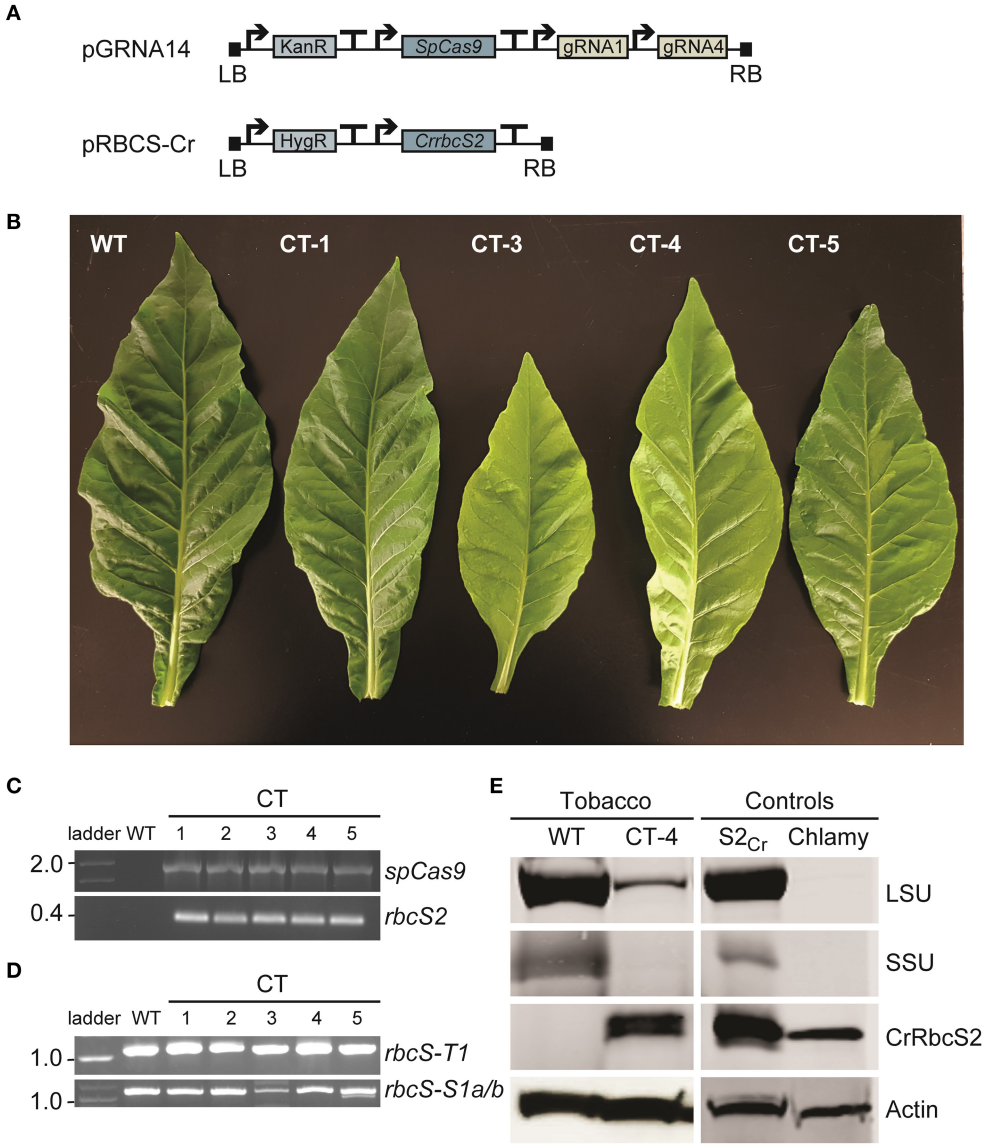


FIGURE 5 | Phenotypes and molecular analyses of co-transformed (CT) plants. **(A)** Wild-type plants were transformed with a Cas-gRNA vector (vector pGRNA14) and gene expression cassette for *CrrbcS2* from Chlamydomonas (vector pRBCS-Cr) (see **Supplementary Figures 1, 2** for vector details). **(B)** Leaves of a wild-type (WT) plant and four CT lines with phenotypes ranging from wild-type-like (CT-1) to pale (CT-3 and CT-4) and chimeric (CT-5) are shown. **(C)** PCR amplification of cDNA showing stable integration and expression of *SpCas9* and *rbcS2* in T_0 CT lines. **(D)** PCR amplification of *rbcS-T1* and *rbcS-S1a/b* (as in **Figure 2B**). **(E)** Immunoblots of total soluble protein to detect the Rubisco large subunit (LSU) and small subunit (SSU) from tobacco and SSU from Chlamydomonas (CrRbcS2) in CT-4 (T_1 generation). The two controls were from Arabidopsis expressing CrRbcS2 ($S2_{Cr}$; Atkinson et al., 2017) and Chlamydomonas. An actin loading control is shown. No LSU and SSU bands were visible for the Chlamydomonas extract using the primary antibody against wheat Rubisco.

but less LSU protein was produced (Rodermel et al., 1996). In contrast, the CRISPR-Cas9 lines generally had wild-type *rbcS* levels, with the exception of line 4. Therefore, the transcription of *rbcL* in the CRISPR-Cas9 plants is likely affected by different regulatory mechanisms than in the antisense plants because of the lack of, or a relatively small, suppression of *rbcS* mRNA levels (Wostrikoff and Stern, 2007). Alternatively, inhibition of *rbcL* transcription could occur in the CRISPR-Cas9 lines owing to increased degradation of truncated or non-functional SSU peptides, as has been reported for the *polygalacturonase* (PG) gene in tomato (Smith et al., 1990). Similarly, a significant reduction in *rbcL* mRNA was observed in CRISPR-Cas9 and T-DNA insertion Arabidopsis mutants with 3–4% of wild-type *rbcS* (Khumsupan et al., 2020).

Recently, an RNAi-rbcS tobacco master line (tobRrΔS) was described that enables the expression of homogenous non-native Rubisco enzymes by introducing an rbcL-rbcS operon into the plastome of tobacco (Martin-Avila et al., 2020). Our study generated a complementary tobacco line with a stable decrease in Rubisco content to use as a platform for heterologous

SSU expression by nuclear transformation. Furthermore, our approach could be extended to crop species that are not amenable to chloroplast transformation. We also co-transformed wild-type tobacco with CRISPR-Cas9 and an heterologous SSU expression vector to examine if it was feasible to simultaneously remove and replace native SSUs, as transformation is time-consuming in non-model and/or crop species (Martin-Avila et al., 2020; Matsumura et al., 2020). Further analyses are required in the next generation of CT-4 plants to determine the amount of Rubisco in these lines and confirm that expression of the Chlamydomonas SSU complemented the reduced growth phenotype produced from simultaneously mutating three *rbcS* genes. However, our finding demonstrate the practicality of a co-transformation approach to circumvent lethal deletions in crops (e.g., when attempting to knockout an entire native *rbcS* family). In conclusion, this proof-of-principle study provides an approach to partially or fully replace entire gene families in a single step in complex polyploid species.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

AM and SD planned and designed the research and wrote the manuscript. YM and DO performed additional experimental work. All authors assisted with editing the manuscript.

FUNDING

This work was supported by the UK Biotechnology and Biological Sciences Research Council (BB/M006468/1 and BB/S015531/1 to AM, and BB/I024488/1 to EC-S) and the Leverhulme Trust (grant RPG-2017-402 to AM). SD was funded by the BBSRC East of Scotland Bioscience (EASTBIO) Doctoral Training Partnership program.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgeed.2020.605614/full#supplementary-material

REFERENCES

Aigner, H., Wilson, R. H., Bracher, A., Calisse, L., Bhat, J. Y., Hartl, F. U., et al. (2017). Plant RuBisCo assembly in *E. coli* with five chloroplast chaperones including BSD2. *Science* 358, 1272–1278. doi: 10.1126/science.aap9221

Atkinson, N., Leitão, N., Orr, D. J., Meyer, M. T., Carmo-Silva, E., Griffiths, H., et al. (2017). Rubisco small subunits from the unicellular green alga Chlamydomonas complement Rubisco-deficient mutants of Arabidopsis. *New Phytol.* 214, 655–667. doi: 10.1111/nph.14414

Atkinson, N., Velanis, C. N., Wunder, T., Clarke, D. J., Mueller-Cajar, O., and McCormick, A. J. (2019). The pyrenoidal linker protein EPYC1 phase separates with hybrid Arabidopsis–chlamydomonas rubisco through interactions with the algal rubisco small subunit. *J. Exp. Bot.* 70, 5271–5285. doi: 10.1093/jxb/erz275

Bae, S., Park, J., and Kim, J.-S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* 30, 1473–1475. doi: 10.1093/bioinformatics/btu048

Bradford, M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. doi: 10.1006/abio.1976.9999

Brinkman, E. K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* 42:e168. doi: 10.1093/nar/gku936

Brooks, C., Nekrasov, V., Lippman, Z. B., and Eck, J. V. (2014). Efficient gene editing in tomato in the first generation using the clustered regularly interspaced short palindromic repeats/CRISPR-associated9 system. *Plant Physiol.* 166, 1292–1297. doi: 10.1104/pp.114.247577

Dobrescu, A., Scorza, L. C. T., Tsaftaris, S. A., and McCormick, A. J. (2017). A "do-it-yourself" phenotyping system: measuring growth and morphology throughout the diel cycle in rosette shaped plants. *Plant Methods* 13:95. doi: 10.1186/s13007-017-0247-6

Doll, N. M., Gilles, L. M., Gérentes, M.-F., Richard, C., Just, J., Fierlej, Y., et al. (2019). Single and multiple gene knockouts by CRISPR–Cas9 in maize. *Plant Cell Rep.* 38, 487–501. doi: 10.1007/s00299-019-02378-1

Durr, J., Papareddy, R., Nakajima, K., and Gutierrez-Marcos, J. (2018). Highly efficient heritable targeted deletions of gene clusters and non-coding regulatory regions in *Arabidopsis* using CRISPR-Cas9. *Sci. Rep.* 8:4443. doi: 10.1038/s41598-018-22667-1

Edwards, K. D., Fernandez-Pozo, N., Drake-Stowe, K., Humphry, M., Evans, A. D., Bombarely, A., et al. (2017). A reference genome for *Nicotiana tabacum* enables map-based cloning of homeologous loci implicated in nitrogen utilization efficiency. *BMC Genomics* 18:448. doi: 10.1186/s12864-017-3791-6

Endo, M., Mikami, M., and Toki, S. (2015). Multigene knockout utilizing off-target mutations of the CRISPR/Cas9 system in rice. *Plant Cell Physiol.* 56, 41–47. doi: 10.1093/pcp/pcu154

Engler, C., Youles, M., Gruetzner, R., Ehnert, T.-M., Werner, S., Jones, J. D. G., et al. (2014). A golden gate modular cloning toolbox for plants. *ACS Synth. Biol.* 3, 839–843. doi: 10.1021/sb4001504

Esquivel, M. G., Genkov, T., Nogueira, A. S., Salvucci, M. E., and Spreitzer, R. J. (2013). Substitutions at the opening of the rubisco central solvent channel affect holoenzyme stability and CO2/O2 specificity but not activation by rubisco activase. *Photosynth Res.* 118, 209–218. doi: 10.1007/s11120-013-9916-0

Ethier, G. J., and Livingston, N. J. (2004). On the need to incorporate sensitivity to CO_2 transfer conductance into the farquhar–von caemmerer–berry leaf photosynthesis model. *Plant Cell Environ.* 27, 137–153. doi: 10.1111/j.1365-3040.2004.01140.x

Feng, Z., Mao, Y., Xu, N., Zhang, B., Wei, P., Yang, D.-L., et al. (2014). Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in *Arabidopsis*. *PNAS* 111, 4632–4637. doi: 10.1073/pnas.1400822111

Fukayama, H., Kobara, T., Shiomi, K., Morita, R., Sasayama, D., Hatanaka, T., et al. (2019). Rubisco small subunits of C4 plants, napier grass and guinea grass confer C4-like catalytic properties on rubisco in rice. *Plant Prod. Sci.* 22, 296–300. doi: 10.1080/1343943X.2018.1540279

Galmés, J., Andralojc, P. J., Kapralov, M. V., Flexas, J., Keys, A. J., Molins, A., et al. (2014). Environmentally driven evolution of rubisco and improved photosynthesis and growth within the C3 genus *Limonium* (Plumbaginaceae). *New Phytol.* 203, 989–999. doi: 10.1111/nph.12858

Gao, J., Wang, G., Ma, S., Xie, X., Wu, X., Zhang, X., et al. (2015). CRISPR/Cas9-mediated targeted mutagenesis in *Nicotiana tabacum*. *Plant Mol. Biol.* 87, 99–110. doi: 10.1007/s11103-014-0263-0

Genkov, T., and Spreitzer, R. J. (2009). Highly conserved small subunit residues influence rubisco large subunit catalysis. *J. Biol. Chem.* 284, 30105–30112. doi: 10.1074/jbc.M109.044081

Gong, L., Olson, M., and Wendel, J. F. (2014). Cytonuclear evolution of rubisco in four allopolyploid lineages. *Mol. Biol. Evol.* 31, 2624–2636. doi: 10.1093/molbev/msu207

Hahn, H., and Nekrasov, V. (2019). CRISPR/Cas precision: do we need to worry about off-targeting in plants? *Plant Cell Rep.* 38, 437–441. doi: 10.1007/s00299-018-2355-9

Hoagland, D. R., and Snyder, W. C. (1933). Nutrition of strawberry plant under controlled conditions. *Proc. Am. Soc. Hortic. Sci.* 30, 288–294.

Howe, C. J., Auffret, A. D., Doherty, A., Bowman, C. M., Dyer, T. A., and Gray, J. C. (1982). Location and nucleotide sequence of the gene for the proton-translocating subunit of wheat chloroplast ATP synthase. *Proc Natl Acad Sci U.S.A.* 79, 6903–6907. doi: 10.1073/pnas.79.22.6903

Ishikawa, C., Hatanaka, T., Misoo, S., Miyake, C., and Fukayama, H. (2011). Functional Incorporation of sorghum small subunit increases the catalytic turnover rate of rubisco in transgenic rice. *Plant Physiol.* 156, 1603–1611. doi: 10.1104/pp.111.177030

Izumi, M., Tsunoda, H., Suzuki, Y., Makino, A., and Ishida, H. (2012). RBCS1A and RBCS3B, two major members within the arabidopsis RBCS multigene family, function to yield sufficient Rubisco content for leaf photosynthetic capacity. *J. Exp. Bot.* 63, 2159–2170. doi: 10.1093/jxb/err434

Khumsupan, P., Kozlowska, M. A., Orr, D. J., Andreou, A. I., Nakayama, N., Patron, N., et al. (2020). Generating and characterizing single- and multigene mutants of the Rubisco small subunit family in arabidopsis. *J. Exp. Bot.* 71, 5963–5975. doi: 10.1093/jxb/eraa316

Laterre, R., Pottier, M., Remacle, C., and Boutry, M. (2017). Photosynthetic trichomes contain a specific rubisco with a modified pH-dependent activity. *Plant Physiol.* 173, 2110–2120. doi: 10.1104/pp.17.00062

Lawrenson, T., Shorinola, O., Stacey, N., Li, C., Østergaard, L., Patron, N., et al. (2015). Induction of targeted, heritable mutations in barley and brassica oleracea using RNA-guided Cas9 nuclease. *Genome Biol.* 16:258. doi: 10.1186/s13059-015-0826-7

Lin, M. T., Stone, W. D., Chaudhari, V., and Hanson, M. R. (2020). Enzyme kinetics of tobacco rubisco expressed in *Escherichia coli* varies depending on the small subunit composition. *Nat. Plants* 6, 1289–1299. doi: 10.1038/s41477-020-00761-5

Long, S. P., and Bernacchi, C. J. (2003). Gas exchange measurements, what can they tell us about the underlying limitations to photosynthesis? Procedures and sources of error. *J. Exp. Botany* 392, 2393–2401. doi: 10.1093/jxb/erg262

Lowder, L. G., Zhang, D., Baltes, N. J., Paul, J. W., Tang, X., Zheng, X., et al. (2015). A CRISPR/Cas9 toolbox for multiplexed plant genome editing and transcriptional regulation. *Plant Physiol.* 169, 971–985. doi: 10.1104/pp.15.00636

Ma, X., Zhang, Q., Zhu, Q., Liu, W., Chen, Y., Qiu, R., Wang, B., 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

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 47, W636–W641. doi: 10.1093/nar/gkz268

Makino, A., Nakano, H., Mae, T., Shimada, T., and Yamamoto, N. (2000). Photosynthesis, plant growth and N allocation in transgenic rice plants with decreased Rubisco under CO_2 enrichment. *J. Exp. Botany* 51, 383–389. doi: 10.1093/jexbot/51.suppl_1.383

Makino, A., Shimada, T., Takumi, S., Kaneko, K., Matsuoka, M., Shimamoto, K., et al. (1997). Does decrease in ribulose-1,5-bisphosphate carboxylase by antisense RbcS lead to a higher N-use efficiency of photosynthesis under conditions of saturating CO_2 and light in rice plants? *Plant Physiol.* 114, 483–491. doi: 10.1104/pp.114.2.483

Martin-Avila, E., Lim, Y.-L., Birch, R., Dirk, L. M. A., Buck, S., Rhodes, T., et al. (2020). Modifying plant photosynthesis and growth via simultaneous chloroplast transformation of rubisco large and small subunits. *Plant Cell.* 32, 2898–2916. doi: 10.1105/tpc.20.00288

Masle, J., Hudson, G. S., and Badger, M. R. (1993). Effects of ambient CO_2 concentration on growth and nitrogen use in tobacco (*Nicotiana tabacum*) plants transformed with an antisense gene to the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase. *Plant Physiol.* 103, 1075–1088. doi: 10.1104/pp.103.4.1075

Matsumura, H., Shiomi, K., Yamamoto, A., Taketani, Y., Kobayashi, N., Yoshizawa, T., et al. (2020). Hybrid rubisco with complete replacement of rice rubisco small subunits by sorghum counterparts confers C4-plant-like high catalytic activity. *Mol. Plant* 13, 1570–1581. doi: 10.1016/j.molp.2020.08.012

Mitchell, R. A. C., Joyce, P. A., Rong, H., Evans, V. J., Madgwick, P. J., and Parry, M. A. J. (2004). Loss of decreased-rubisco phenotype between generations of wheat transformed with antisense and sense rbcS. *Ann. Appl. Biol.* 145, 209–216. doi: 10.1111/j.1744-7348.2004.tb00377.x

Monteith, J. (1991). "Plant and crop modelling–a mathematical approach to plant and crop physiology," in *Agricultural Systems*, eds J. H. M. Thornley, and I. R. Johnson (Oxford: Clarendon Press), 660.

Morineau, C., Bellec, Y., Tellier, F., Gissot, L., Kelemen, Z., Nogué, F., et al. (2017). Selective gene dosage by CRISPR-Cas9 genome editing in hexaploid *Camelina sativa*. *Plant Biotechnol. J.* 15, 729–739. doi: 10.1111/pbi.12671

Morita, K., Hatanaka, T., Misoo, S., and Fukayama, H. (2014). Unusual small subunit that is not expressed in photosynthetic cells alters the catalytic properties of rubisco in rice. *Plant Physiol.* 164, 69–79. doi: 10.1104/pp.113.228015

Morita, K., Hatanaka, T., Misoo, S., and Fukayama, H. (2016). Identification and expression analysis of non-photosynthetic Rubisco small subunit, OsRbcS1-like genes in plants. *Plant Gene* 8, 26–31. doi: 10.1016/j.plgene.2016.09.004

Nagaya, S., Kawamura, K., Shinmyo, A., and Kato, K. (2010). The HSP terminator of arabidopsis thaliana increases gene expression in plant cells. *Plant Cell Physiol.* 51, 328–332. doi: 10.1093/pcp/pcp188

Ordon, J., Gantner, J., Kemna, J., Schwalgun, L., Reschke, M., Streubel, J., et al. (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

Orr, D. J., Alcântara, A., Kapralov, M. V., Andralojc, P. J., Carmo-Silva, E., and Parry, M. A. J. (2016). Surveying rubisco diversity and temperature response to improve crop photosynthetic efficiency. *Plant Physiol.* 172, 707–717. doi: 10.1104/pp.16.00750

Orr, D. J., Worrall, D., Lin, M. T., Carmo-Silva, E., Hanson, M. R., and Parry, M. A. J. (2020). Hybrid cyanobacterial-tobacco rubisco supports autotrophic growth and procarboxysomal aggregation. *Plant Physiol.* 182, 807–818. doi: 10.1104/pp.19.01193

Park, J., Bae, S., and Kim, J.-S. (2015). Cas-designer: a web-based tool for choice of CRISPR-Cas9 target sites. *Bioinformatics* 31, 4014–4016. doi: 10.1093/bioinformatics/btv537

Porra, R. J., Thompson, W. A., and Kriedemann, P. E. (1989). Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochim. Biophys. Acta* 975, 384–394. doi: 10.1016/S0005-2728(89)80347-0

Pottier, M., Gilis, D., and Boutry, M. (2018). The hidden face of rubisco. *Trends Plant Sci.* 23, 382–392. doi: 10.1016/j.tplants.2018.02.006

Quick, W. P., Schurr, U., Fichtner, K., Schulze, E.-D., Rodermel, S. R., Bogorad, L., and Stitt, M. (1991a). The impact of decreased rubisco on photosynthesis, growth, allocation and storage in tobacco plants which have been transformed with antisense rbcS. *Plant J.* 1, 51–58. doi: 10.1111/j.1365-313X.1991.00051.x

Quick, W. P., Schurr, U., Scheibe, R., Schulze, E. D., Rodermel, S. R., Bogorad, L., and Stitt, M. (1991b). Decreased ribulose-1,5-bisphosphate carboxylase-oxygenase in transgenic tobacco transformed with "antisense" rbcS : I. Impact on photosynthesis in ambient growth conditions. *Planta* 183, 542–554. doi: 10.1007/BF00194276

Raitskin, O., Schudoma, C., West, A., and Patron, N. J. (2019). Comparison of efficiency and specificity of CRISPR-associated (Cas) nucleases in plants: an expanded toolkit for precision genome engineering. *PLoS ONE* 14:e0211598. doi: 10.1371/journal.pone.0211598

Rodermel, S., Haley, J., Jiang, C. Z., Tsai, C. H., and Bogorad, L. (1996). A mechanism for intergenomic integration: abundance of ribulose bisphosphate carboxylase small-subunit protein influences the translation of the large-subunit mRNA. *PNAS* 93, 3881–3885. doi: 10.1073/pnas.93.9.3881

Salesse-Smith, C. E., Sharwood, R. E., Busch, F. A., Kromdijk, J., Bardal, V., and Stern, D. B. (2018). Overexpression of rubisco subunits with RAF1 increases rubisco content in maize. *Nat. Plants* 4, 802–810. doi: 10.1038/s41477-018-0252-4

Salesse-Smith, C. E., Sharwood, R. E., Busch, F. A., and Stern, D. B. (2020). Increased rubisco content in maize mitigates chilling stress and speeds recovery. *Plant Biotechnol. J.* 18, 1409–1420. doi: 10.1111/pbi.13306

Schmidt, G. W., and Delaney, S. K. (2010). Stable internal reference genes for normalization of real-time RT-PCR in tobacco (*Nicotiana tabacum*) during development and abiotic stress. *Mol. Genet. Genomics* 283, 233–241. doi: 10.1007/s00438-010-0511-1

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. doi: 10.1038/nmeth.2089

Sharwood, R. E., Sonawane, B. V., Ghannoum, O., and Whitney, S.M. (2016). Improved analysis of C4 and C3 photosynthesis via refined *in vitro* assays of their carbon fixation biochemistry. *J. Exp. Bot.* 67, 3137–3148, doi: 10.1093/jxb/erw154

Sierro, N., Battey, J. N. D., Ouadi, S., Bakaher, N., Bovet, L., Willig, A., et al. (2014). The tobacco genome sequence and its comparison with those of tomato and potato. *Nat. Commun.* 5:3833. doi: 10.1038/ncomms4833

Smith, C. J. S., Watson, C. F., Bird, C. R., Ray, J., Schuch, W., and Grierson, D. (1990). Expression of a truncated tomato polygalacturonase gene inhibits expression of the endogenous gene in transgenic plants. *Mol. Gen. Genet*. 224, 477–481. doi: 10.1007/BF00262443

Spreitzer, R. J. (2003). Role of the small subunit in ribulose-1,5-bisphosphate carboxylase/oxygenase. *Arch. Biochem. Biophys.* 414, 141–149. doi: 10.1016/S0003-9861(03)00171-1

Stitt, M., Quick, W. P., Schurr, U., Schulze, E.-D., Rodermel, S. R., and Bogorad, L. (1991). Decreased ribulose-1,5-bisphosphate carboxylase-oxygenase in transgenic tobacco transformed with "antisense" rbcS. *Planta* 183, 555–566. doi: 10.1007/BF00194277

Stitt, M., and Schulze, D. (1994). Does rubisco control the rate of photosynthesis and plant growth? An exercise in molecular ecophysiology. *Plant Cell Environ.* 17, 465–487. doi: 10.1111/j.1365-3040.1994.tb00144.x

Tizaoui, K., and Kchouk, M. E. (2012). Genetic approaches for studying transgene inheritance and genetic recombination in three successive generations of transformed tobacco. *Genet. Mol. Biol.* 35, 640–649. doi: 10.1590/S1415-47572012000400015

Whitney, S. M., Birch, R., Kelso, C., Beck, J. L., and Kapralov, M. V. (2015). Improving recombinant rubisco biogenesis, plant photosynthesis and growth by coexpressing its ancillary RAF1 chaperone. *PNAS* 112, 3564–3569. doi: 10.1073/pnas.1420536112

Wolabu, T. W., Cong, L., Park, J.-J., Bao, Q., Chen, M., Sun, J., et al. (2020). Development of a highly efficient multiplex genome editing system in outcrossing tetraploid alfalfa (*Medicago sativa*). *Front. Plant Sci.* 11:1063. doi: 10.3389/fpls.2020.01063

Wostrikoff, K., and Stern, D. (2007). Rubisco large-subunit translation is autoregulated in response to its assembly state in tobacco chloroplasts. *PNAS* 104, 6466–6471. doi: 10.1073/pnas.0610586104

Xie, X., Qin, G., Si, P., Luo, Z., Gao, J., Chen, X., et al. (2017). Analysis of *Nicotiana tabacum* PIN genes identifies NtPIN4 as a key regulator of axillary bud growth. *Physiol. Plant.* 160, 222–239. doi: 10.1111/ppl.12547

Xing, H.-L., Dong, L., Wang, Z.-P., Zhang, H.-Y., Han, C.-Y., Liu, B., et al. (2014). A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol.* 14:327. doi: 10.1186/s12870-014-0327-y

Yoon, D.-K., Ishiyama, K., Suganami, M., Tazoe, Y., Watanabe, M., Imaruoka, S., et al. (2020). Transgenic rice overproducing rubisco exhibits increased yields with improved nitrogen-use efficiency in an experimental paddy field. *Nat. Food* 1, 134–139. doi: 10.1038/s43016-020-0033-x

Young, J. N., Heureux, A. M. C., Sharwood, R. E., Rickaby, R. E. M., Morel, F. M. M., and Whitney, S. M. (2016). Large variation in the rubisco kinetics of diatoms reveals diversity among their carbon-concentrating mechanisms. *J. Exp. Bot.* 67, 3445–3456. doi: 10.1093/jxb/erw163

Zhang, H.ui, Zhang, J., Wei, P., Zhang, B., Gou, F., Feng, Z., Mao, Y., et al. (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

Zhou, H., Liu, B., Weeks, D. P., Spalding, M. H., and Yang, B. (2014). Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Res.* 42, 10903–10914. doi: 10.1093/nar/gku806

Zhu, X.-G., Portis, A. R., and Long, S. P. (2004). Would transformation of C3 crop plants with foreign Rubisco increase productivity? A computational analysis extrapolating from kinetic properties to canopy photosynthesis. *Plant Cell Environ.* 27, 155–165. doi: 10.1046/j.1365-3040.2004.01142.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Donovan, Mao, Orr, Carmo-Silva and McCormick. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genome Editing

ORIGINAL RESEARCH
published: 12 January 2021
doi: 10.3389/fgeed.2020.617713

Precise Genome Editing in miRNA Target Site via Gene Targeting and Subsequent Single-Strand-Annealing-Mediated Excision of the Marker Gene in Plants

Namie Ohtsuki[1†], Keiko Kizawa[2], Akiko Mori[1], Ayako Nishizawa-Yokoi[1,3], Takao Komatsuda[4], Hitoshi Yoshida[1], Katsuyuki Hayakawa[2], Seiichi Toki[1,5,6] and Hiroaki Saika[1*]*

[1] Institute of Agrobiological Sciences, National Agriculture and Food Research Organization (NARO), Tsukuba, Japan, [2] Nisshin Flour Milling Inc., Tsukuba, Japan, [3] Japan Science and Technology Agency (JST), Precursory Research for Embryonic Science and Technology (PRESTO), Kawaguchi, Japan, [4] Institute of Crop Sciences, NARO, Tsukuba, Japan, [5] Graduate School of Nanobioscience, Yokohama City University, Yokohama, Japan, [6] Kihara Institute for Biological Research, Yokohama City University, Yokohama, Japan

OPEN ACCESS

Edited by:
Yiping Qi,
University of Maryland, United States

Reviewed by:
Keunsub Lee,
Iowa State University, United States
Kunling Chen,
Chinese Academy of Sciences, China

***Correspondence:**
Seiichi Toki
stoki@affrc.go.jp
Hiroaki Saika
saika@affrc.go.jp

†Present address:
Namie Ohtsuki,
Biotechnology Research Center, The University of Tokyo, Tokyo, Japan

Specialty section:
This article was submitted to Genome Editing in Plants, a section of the journal Frontiers in Genome Editing

Received: 15 October 2020
Accepted: 10 December 2020
Published: 12 January 2021

Citation:
Ohtsuki N, Kizawa K, Mori A, Nishizawa-Yokoi A, Komatsuda T, Yoshida H, Hayakawa K, Toki S and Saika H (2021) Precise Genome Editing in miRNA Target Site via Gene Targeting and Subsequent Single-Strand-Annealing-Mediated Excision of the Marker Gene in Plants. Front. Genome Ed. 2:617713. doi: 10.3389/fgeed.2020.617713

Gene targeting (GT) enables precise genome modification—e.g., the introduction of base substitutions—using donor DNA as a template. Combined with clean excision of the selection marker used to select GT cells, GT is expected to become a standard, generally applicable, base editing system. Previously, we demonstrated marker excision via a *piggyBac* transposon from GT-modified loci in rice. However, *piggyBac*-mediated marker excision has the limitation that it recognizes only the sequence TTAA. Recently, we proposed a novel and universal precise genome editing system consisting of GT with subsequent single-strand annealing (SSA)-mediated marker excision, which has, in principle, no limitation of target sequences. In this study, we introduced base substitutions into the microRNA miR172 target site of the *OsCly1* gene—an ortholog of the barley *Cleistogamy1* gene involved in cleistogamous flowering. To ensure efficient SSA, the GT vector harbors 1.2-kb overlapped sequences at both ends of a selection marker. The frequency of positive–negative selection-mediated GT using the vector with overlapped sequences was comparable with that achieved using vectors for *piggyBac*-mediated marker excision without overlapped sequences, with the frequency of SSA-mediated marker excision calculated as ∼40% in the T_0 generation. This frequency is thought to be adequate to produce marker-free cells, although it is lower than that achieved with *piggyBac*-mediated marker excision, which approaches 100%. To date, introduction of precise substitutions in discontinuous multiple bases of a targeted gene using base editors and the prime editing system based on CRISPR/Cas9 has been quite difficult. Here, using GT and our SSA-mediated marker excision system, we succeeded in the precise base substitution not only of single bases but also of artificial discontinuous multiple bases in the miR172 target site of the *OsCly1* gene.

Precise base substitution of miRNA target sites in target genes using this precise genome editing system will be a powerful tool in the production of valuable crops with improved traits.

Keywords: gene targeting, precise genome modification, single-strand annealing, cleistogamy 1, *oryza sativa*, miRNA target site

INTRODUCTION

Biological species have developed repair systems for DNA double-strand breaks (DSBs) as such repairs are critical to life. DSB repair systems have been classified traditionally into two pathways: non-homologous end joining (NHEJ) and homologous recombination (HR) (Chapman et al., 2012; Hustedt and Durocher, 2016). The former is a rapid but error-prone response that results in some inserted and/or deleted bases due to the simple ligation of both ends of a DSB site. The latter is an accurate repair system that uses a homologous region of the sister chromatid as a template at the DSB site.

Gene targeting (GT) is a powerful genome engineering technology that can be used to introduce various types of mutation into a target gene locus by HR using a donor DNA as a template. The first demonstration of GT in higher plants was reported as far back as 1988 (Paszkowski et al., 1988). Much later, a GT procedure applied to an endogenous gene was first reported in the *WAXY* gene in rice (Terada et al., 2002). Since then, knock-out as well as knock-in mutants of several genes have been produced using GT techniques (Shimatani et al., 2015). Although the CRISPR/Cas9 system is now used commonly for gene knock-out in various plant species, including rice (Mikami et al., 2015a,b), it can introduce insertion and/or deletion of a small number of bases in the target gene, thus precise genome modifications—such as base substitutions—are still difficult using CRISPR/Cas9. Base editor systems using Cas9 nickase fused to cytidine and adenosine deaminase have been developed recently; these can introduce C to T (G to A) and A to G (T to C) substitutions, respectively (Komor et al., 2016; Nishida et al., 2016; Gaudelli et al., 2017). Very recently, it was reported that C to A and C to G substitutions can be introduced by use of a new base editor consisting of Cas9 nickase fused to cytidine deaminase and glycosylase in *Escherichia coli* and human cells, respectively (Kurt et al., 2020; Zhao et al., 2020). However, the window, i.e., the possible region of base substitution, is narrow, and bystander substitution of bases adjacent to the target base occurs often. In addition, it was shown that prime editing, consisting of Cas9 nickase fused to reverse transcriptase and a prime editing guide RNA consisting of a guide RNA and RNA homologous to the target DNA, enables the introduction of small mutations, including base substitutions, in human cells (Anzalone et al., 2019). This system has been applied to rice and wheat, where it was found that not only single bases but also discontinuous, up to 4-base, substitutions could be introduced into a target gene (Lin et al., 2020; Xu et al., 2020). However, as yet, there are no reports of successful substitution of several discontinuous bases. Thus, it is not always possible to introduce the desired substitutions into a target gene using the above-mentioned systems, and the development of novel and improved GT systems is an important step toward solving this problem.

In the positive–negative-selection-mediated GT system, a positive selection marker located between both homology arms confers drug resistance to GT cells, while negative selection markers located outside the two homology arms act to kill cells in which the GT vector has integrated randomly in the genome. Desired mutations are introduced into a target site concomitant with the insertion of a positive selection marker by HR between the donor and genomic DNA, and subsequent excision of the positive selection marker from the GT locus leaves only the desired mutations (Shimatani et al., 2015). For marker excision, site-specific recombinases such as Cre-*loxP* (Sauer and Henderson, 1990), FLP-*FRT* (Golic and Lindquist, 1989), and R-RS (Onouchi et al., 1991) have been used. A marker excision system using Cre-*loxP* has been applied to removing the positive selection marker gene from the GT locus in rice (Terada et al., 2010; Dang et al., 2013). In this latter system, the "footprint," which can be several tens of bases for recombinase recognition, remained at the target site after marker excision. In contrast, *piggyBac* transposon, derived from the cabbage looper moth, removed the selectable marker without leaving any footprint in human cells (Yusa et al., 2011; Morioka et al., 2014; Sun and Zhao, 2014). We have previously demonstrated that *piggyBac* could be applied successfully with high efficiency to remove a positive selection marker gene without leaving any footprint in rice (Nishizawa-Yokoi et al., 2015a). However, due to the *piggyBac* transposon's requirement for a TTAA sequence for transposase-dependent integration and excision, the site of the positive marker gene integration site on the GT vector must contain that motif.

Single-strand annealing (SSA) is a DSB repair system in many organisms. DNA repair by SSA occurs between homologous sequences located on both sides of the DSB site. The intervening sequences between homologous regions are eliminated by annealing single-stranded DNA of the two homologous sequences at the DSB site. Several reports have demonstrated elimination of the fragment between homologous sequences on genomic DNA via SSA attributed to DSBs in plants, including in rice (Puchta and Hohn, 1991; Kwon et al., 2012). This system had been applied to marker excision at the GT locus in mice and yeast nearly 30 years ago (Hasty et al., 1991; Valancius and Smithies, 1991). Recently, we reported precise genome editing using GT and subsequent marker excision via SSA in rice (Endo et al., 2020).

Barley *cleistogamy 1* (*cly1*) has been isolated as an essential factor for cleistogamy—an unconventional pollinating style with

discontinuously closed flower on some commercial cultivars—in barley (**Figure 1A**). *cly1* transcript levels are regulated by a microRNA (miRNA), miR172, binding at a complementary 21-bp site encoded on the 10th exon (Nair et al., 2010; Anwar et al., 2018). The *OsCly1* gene (Os04g0649100) is a homolog of barley *cly1* in rice (Zhu and Helliwell, 2011). Rice plants overexpressing the *oscly1* mutant and OsmiR172b frequently exhibit enlarged lodicules and unclosing lemma (Zhu et al., 2009; Zhou et al., 2012). These results suggest the possibility that, as in barley, miR172-mediated downregulation of *OsCly1* is involved in flower closing in rice, and that it might be possible to change opened flowering to closed flowering in rice by substitution of conserved miR172 target sequences in *OsCly1*. Moreover, we have already reported the successful introduction via GT of base substitutions at the miR172 target site in *OsCly1* and subsequent *piggyBac* transposon-mediated marker excision (Nishizawa-Yokoi et al., 2015a). Thus, the *OsCly1* gene is a suitable target gene for this study.

Here, we introduced mutations via GT then compared mutation frequencies between subsequent marker excision via either *piggyBac* transposon or SSA. In addition, we attempted to introduce not only a precise single base substitution, but also a 2-base substitution (both these 1- and 2-bp changes are found naturally among barley varieties) as well as a 7-base artificial discontinuous substitution into the miRNA target site in the *OsCly* gene.

MATERIALS AND METHODS

Vector Construction

The GT vectors for the *OsCly* gene shown in **Figure 1A** were constructed as follows. To construct a vector harboring 4.9-kb of 5′ homology sequence for the *OsCly1* locus with overlapped sequences, fragments amplified by PCR using gene-specific primers (listed in **Supplementary Table 1**) and rice genome DNA as a template were inserted into *Asc*I/*Pme*I-digested pE(L1-L4) vector, yielding pE(L1-L4)5′OsCly1. The substitution in GT-*OsCly1* variation 1 (**Figure 1C**) into pE(L1-L4)5′OsCly1 was performed using a QuickChange II XL site-directed mutagenesis kit (Stratagene, USA) according to the manufacturer's protocol with the primer sets listed in **Supplementary Table 1**, yielding pE(L1-L4)5′OsCly1-var1. The 3.7-kb of 5′ homology sequence for *OsCly1* locus without overlapped sequences was constructed using a similar method. To construct a vector harboring 5.5-kb of 3′ homology sequence for *OsCly1* locus, fragments amplified by PCR using gene-specific primers (listed in **Supplementary Table 1**) and rice genome DNA were inserted into *Bam*HI/*Xho*I-digested pE(L3-L2) vector, yielding pE(L3-L2)3′OsCly1. The LR reaction for the introduction of entry vectors described above and the pE(R4-R3)I-SceITactHyg vector containing the rice actin terminator and *htp* expression cassette into the destination vector, pKOD4 (Nishizawa-Yokoi et al., 2015a) was performed using LR clonase II (Life Technologies, USA), yielding a GT vector, GT-OsCly1 variation 1. To construct GT vectors, GT-OsCly1 variation 2 and 3, 1.8-kb fragments amplified by PCR using gene-specific primers (listed in **Supplementary Table 1**) were replaced with GT-OsCly1 variation 1 using *Not*I/*Asc*I.

The GT vector for the *OsALS* gene was constructed as follows. To construct a vector harboring 3.4-kb of 5′ homology sequence for *OsALS* locus, pZAmGFP containing mutated OsALS harboring W548L/S627I mutations (Osakabe et al., 2005), and GFP expression cassette in a derivative of pPZP202 (Hajdukiewicz et al., 1994) was digested with *Asc*I and *Hin*dIII (blunt-ended with T4 DNA polymerase [TOYOBO, Japan]), and inserted into *Asc*I and *Pac*I (blunt-ended with T4 DNA polymerase)-digested pE(L1-L4) vector. To construct a vector harboring 6.2-kb of 3′ homology sequence for the *OsALS* locus, NBALSGT(AscI/PacI)+pPZP2028 vector (Saika et al., 2015) was digested with *Asc*I and *Pac*I, and inserted into *Asc*I/*Pac*I-digested pE(L3-L2) vector. To construct a positive selection marker, pCAMBIA1390-sGFP (Toki et al., 2006) was digested with *Xba*I and *Pst*I (blunt-ended with T4 DNA polymerase). The resulting 0.7-kb fragment of GFP was inserted into *Xba*I and *Bam*HI (blunt-ended with T4 DNA polymerase)-digested pE(R4-R3)I-SceITactHygcodA vector containing the rice actin terminator and *htp* expression cassette. The LR reaction for the introduction of the three vectors described above into the destination vector, pKOD4 (Nishizawa-Yokoi et al., 2015a) was performed using LR clonase II (Life Technologies).

The I-*Sce*I expression vector shown in **Figure 3A** was constructed by LR reaction for introduction of the following vectors into the destination vector pZD202 (Kwon et al., 2012), pE(L1-L4)Pg3pDsRed2Tg3p containing a 2.4-kb fragment of the rice glyceraldehyde-3-phosphate (G3P) promoter, a 0.8-kb fragment of intron-DsRed and a 0.8-kb fragment of the rice G3P terminator, pE(R4-R3)TactP35SnptIIThsp (Nishizawa-Yokoi et al., 2015b) and pE(L3-L2)P2X35S::I-SceI::Thsp (Kwon et al., 2012). For construction of the control vector, pE(L3-L2)T17.3 containing a 1.2-kb fragment of rice heatshock protein 17.3 terminator was used for the LR reaction instead of pE(L3-L2)P2X35S::I-SceI::Thsp.

Agrobacterium-Mediated Transformation

GT vectors were transformed into *Agrobacterium tumefaciens* strain EHA105 (Hood et al., 1993) by the electroporation method as shown schematically in **Supplementary Figure 1**. Rice (*Oryza sativa*. L cv. Nipponbare) was used for *Agrobacterium*-mediated transformation as described previously (Toki, 1997; Toki et al., 2006). Briefly, 3-week-old secondary calli transformed with *Agrobacterium* harboring pKOD4/OsCly1 were selected on N6D medium solidified with 0.4% gelrite containing 50 mg/L hygromycin and 25 mg/L meropenem. GT candidate calli confirmed as below were transferred to regeneration medium with 25 mg/L meropenem, and shoots arising from callus were transferred to MS medium without phytohormones. For marker excision, GT calli in the T_0 generation or induced from mature seeds in the T_1 generation were transformed with *Agrobacterium* harboring the I-*Sce*I expression vector. Transformed calli were selected on N6D medium containing 35 mg/L G418 (Geneticin) and 25 mg/L meropenem. Marker-free calli confirmed as below were transferred to regeneration medium.

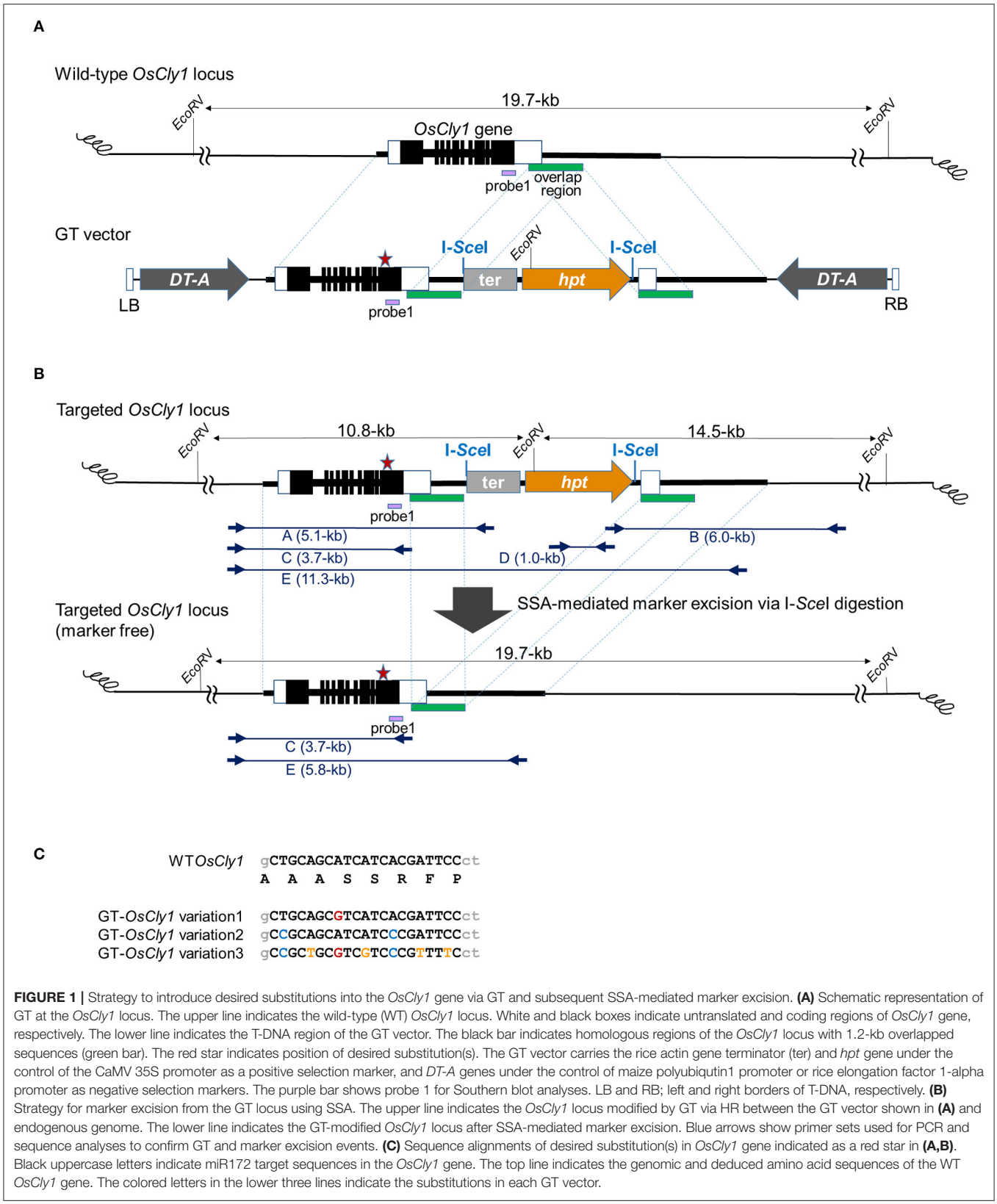


FIGURE 1 | Strategy to introduce desired substitutions into the *OsCly1* gene via GT and subsequent SSA-mediated marker excision. **(A)** Schematic representation of GT at the *OsCly1* locus. The upper line indicates the wild-type (WT) *OsCly1* locus. White and black boxes indicate untranslated and coding regions of *OsCly1* gene, respectively. The lower line indicates the T-DNA region of the GT vector. The black bar indicates homologous regions of the *OsCly1* locus with 1.2-kb overlapped sequences (green bar). The red star indicates position of desired substitution(s). The GT vector carries the rice actin gene terminator (ter) and *hpt* gene under the control of the CaMV 35S promoter as a positive selection marker, and *DT-A* genes under the control of maize polyubiqutin1 promoter or rice elongation factor 1-alpha promoter as negative selection markers. The purple bar shows probe 1 for Southern blot analyses. LB and RB; left and right borders of T-DNA, respectively. **(B)** Strategy for marker excision from the GT locus using SSA. The upper line indicates the *OsCly1* locus modified by GT via HR between the GT vector shown in **(A)** and endogenous genome. The lower line indicates the GT-modified *OsCly1* locus after SSA-mediated marker excision. Blue arrows show primer sets used for PCR and sequence analyses to confirm GT and marker excision events. **(C)** Sequence alignments of desired substitution(s) in *OsCly1* gene indicated as a red star in **(A,B)**. Black uppercase letters indicate miR172 target sequences in the *OsCly1* gene. The top line indicates the genomic and deduced amino acid sequences of the WT *OsCly1* gene. The colored letters in the lower three lines indicate the substitutions in each GT vector.

Screening of GT and Marker Excision Events

Genomic DNA was extracted from hygromycin resistant calli after 4–5 weeks selection and from leaves of regenerated plants by Agencourt Chloropure (Bechman Coulter, USA) according to the manufacturer's protocol. PCR analysis was performed with PrimeSTAR GXL DNA Polymerase (TAKARA BIO) or KOD FX neo (TOYOBO, Japan) using the primer sets listed in **Supplementary Table 1**. For direct sequence analysis, amplified fragments were purified with a QIAquick Gel Extraction Kit (Qiagen, Germany). Sequences of purified PCR fragments were read with an ABI3130 sequencer (ABI, USA) and analyzed with Sequence Scanner.

Southern Blot Analysis

Genomic DNA was extracted from leaves of GT candidate plants using a Nucleon Phytopure Extraction Kit (GE Healthcare, USA) according to the manufacturer's protocol. Genomic DNA (2 μg) was digested with *Eco*RV or *Msc*I and gel electrophoresis performed in a 0.8% gel with around 30 V. Specific DNA probes were prepared using a PCR digoxigenin (DIG) probe synthesis kit (Roche Diagnostics, Switzerland) according to the manufacturer's protocol using the primer sets listed in **Supplementary Table 1**. Southern blot analyses were performed by following a conventional protocol.

Observation of Floral Tissues by Optical Microscopy

GT homozygous plants in the T_1 and T_2 generations, GT#34-6-53 and #441-113-115-38, respectively, were grown in a greenhouse. Floral tissues of 2.5-month-old plants were observed with a microscope (LEICA DFC310 FX, Germany) as previously described (Yoshida et al., 2007; Lombardo et al., 2017).

Observation of GFP and DsRed Fluorescence

GFP and DsRed fluorescence from rice calli was observed using a fluorescence microscope with GFP2 and DsRed filters, respectively (MZ FLIII).

RESULTS AND DISCUSSION

Precise Modification of the miR172 Target Site in the *OsCly1* Gene via Positive–Negative Selection-Mediated GT

The T-DNA structures in GT vectors used in this study are illustrated in **Figure 1A**. In the GT vectors, endogenous rice genomic sequence from the *OsCly1* locus with desirable substitutions at the miR172 target site (GT-*OsCly1* variation 1, 2, and 3 in **Figure 1C**) was interrupted by the positive selection marker consisting of the cauliflower mosaic virus (CaMV) 35S promoter, *hygromycin phosphotransferase* (*hpt*) gene, and rice actin gene terminator. The purpose of the rice actin terminator was to help prevent transcriptional drive-through from the *OsCly1* gene to the downstream *hpt* gene. I-*Sce*I meganuclease recognition sequences were placed at both ends of the *hpt* selection marker cassette. Partially overlapped sequence of the *OsCly1* gene of 1.2-kb in length was located at the 3′ end of the *hpt* cassette to induce break-induced SSA for excision of the *hpt* cassette. *Diphtheria toxin A subunit* (*DT-A*) gene expression cassettes as a negative selection marker were located just inside the left and right borders to suppress growth of hygromycin-resistant cells in which the GT vector is integrated randomly into the rice genome.

First, we performed GT experiments using the vector GT-OsCly1 variation 1 to introduce the single base substitution found in the *cly1* gene of cleistogamous barley varieties, which is the same substitution as our previous report (Nishizawa-Yokoi et al., 2015a). The A to G substitution in GT-*OsCly1* variation 1 is located at the 8th position of the miRNA172 target sequence in the *OsCly1* gene (Nair et al., 2010). Rice calli transformed with GT-OsCly1 variation 1 were cultured on medium containing hygromycin B for 4 weeks. A total of 1,476 hygromycin-resistant calli were obtained from 8,239 (∼56 g) pieces of *Agrobacterium*-infected calli (**Table 1**). To screen GT calli, PCR analyses with primer sets A and B to amplify 5′ and 3′ regions of the targeted locus shown in **Figure 1B** were performed. Both 5′ and 3′ junction fragments were detected in a total of 30 independent lines (**Table 1**), indicating that the *hpt* gene was introduced into the *OsCly1* locus by HR between the GT vector and endogenous target sequences. Regenerated plants from these 30 lines of GT-positive calli were analyzed. Direct sequence analyses of PCR fragments amplified with primer set C showed that a heterozygous base substitution A/G, at the 8th position of the miRNA172 target site in the *OsCly1* gene, was found in eight lines of regenerated plants (**Figure 2A**), suggesting that true GT events had occurred in these plants. Southern blot analyses of *Eco*RV-digested DNA from these eight lines, using probe 1 recognizing the endogenous *OsCly1* gene (**Figure 1**), showed that wild-type (WT) bands (19.7-kb) and bands corresponding to the GT allele (10.8-kb) were detected in five lines of regenerated plants, although only 19.7-kb bands were detected in non-transformed plants (**Figure 2B**). Taken together, these molecular analyses showed that precise introduction of the *hpt* gene and desired substitution into the *OsCly1* gene via GT had occurred successfully in a total of five independent plants.

Effect on GT Efficiency of Overlapping Sequence in the GT Vector

We expected that the *hpt* marker cassette would be excised from the GT locus by I-*Sce*I-dependent break-induced SSA using the 1.2-kb overlapped sequence (**Figure 1B**). We previously reported that SSA occurs spontaneously in rice, although at low frequency (Kwon et al., 2012). Thus, the marker cassette could be removed from the GT vector or GT locus in the absence of I-*Sce*I through spontaneously induced DSBs and subsequent DSB repair by SSA. If the *hpt* marker cassette is removed before GT, GT candidate cells cannot be selected with hygromycin. Similarly, if the *hpt* marker cassette is removed from the GT locus, GT cells cannot survive on medium containing hygromycin. To assess whether spontaneously occurring SSA could decrease the efficiency of GT cell selection, we compared the frequencies with which GT lines

TABLE 1 | Summary of GT experiments for *OsCly1* locus.

GT vector	Substitution (same as Figure 1C)	Overlapped region	No. of *Agrobacterium*-infected calli		No. of hygromycin-resistant calli	No. of targeted calli	No. of targeted calli with desired mutations
GT-*OsCly1* variation 1	CTGCAGCGTCATCA CGATTCC	No overlap		2,069 (46.69 g)	179	5	2
		1.2-kb of overlapped sequence	Exp.1	2,336 (22.64 g)	495	10	2
			Exp.2	2,658 (15.11 g)	645	10	2
			Exp.3	3,245 (18.31 g)	336	10	1
			Total	8,239 (56.06 g)	1476	30	5
GT-*OsCly1* variation 2	CCGCAGCATCATCC CGATTCC	1.2-kb of overlapped sequence		1404 (9.56 g)	107	6	3
GT-*OsCly1* variation 3	CCGCTGCGTCGTCC CGTTTTC	1.2-kb of overlapped sequence		1560 (11.33 g)	120	2	1

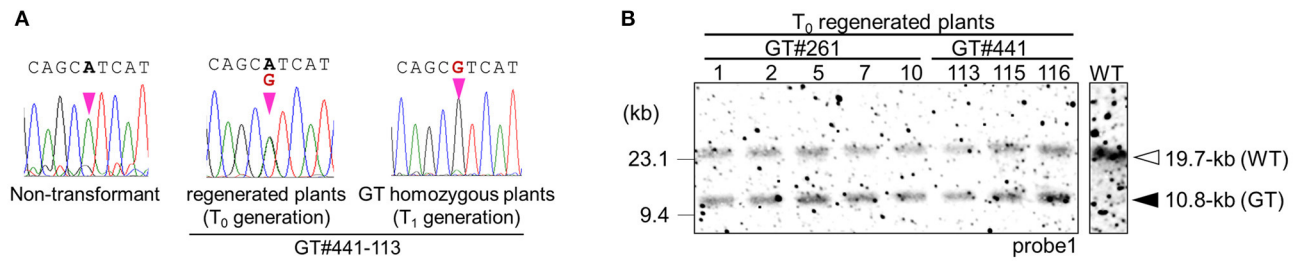


FIGURE 2 | Molecular analyses confirming GT events in GT#261 and GT#441. **(A)** Sequencing chromatograms at the mutation site on the *OsCly1* gene in a non-transformed plant, and in GT plants from the T_0 and T_1 generations. **(B)** Southern blot analyses with probe1 (**Figure 1**) using *Eco*RV-digested genomic DNA of regenerated plants from GT calli of GT#261 and GT#441. White and black triangles indicated bands derived from WT and GT allele, respectively.

were obtained between GT-OsCly1 variation 1 vectors without overlapped sequences, as shown in **Supplementary Figure 2A**. GT experiments and molecular analyses of calli and regenerated plants showed that two lines of true GT regenerated plants (#34 and 62) were obtained from 2,069 pieces of *Agrobacterium*-infected calli (**Table 1**; **Supplementary Figures 2B,C**). Moreover, in our previous study, it was shown that two GT calli carrying an A/G mutation in the miR172 targeting site of the *OsCly1* gene were obtained from 5,139 pieces of *Agrobacterium*-infected calli (Nishizawa-Yokoi et al., 2015a). These results showed that the frequency, i.e., the ratio of the number of GT lines to that of *Agrobacterium*-infected calli using a GT vector without overlapping sequences, was estimated as 0.1%, which is not greatly different from that using a GT vector with overlapping sequences (0.06%) (**Table 1**). Moreover, in this experiment, the GT frequency, i.e., the ratio of GT cells to transformed cells (e.g., hygromycin-resistant cells), is estimated as 2.8 and 2.3% using a GT vector with and without overlapping sequences, respectively, which is similar to that reported previously (generally 0.1–10%; Shimatani et al., 2015). These results suggest that spontaneous SSA-mediated marker excision occurred only rarely in our experiments. We have recently reported a successful example of GT- and SSA-mediated marker excision using 30-bp overlapped sequences in a GT vector (Endo et al., 2020). Here, the frequency (as defined above) was similar between GT vectors with/without 1.2-kb overlapped sequences, suggesting that the use of short overlapped sequences may not be necessary in this experiment.

One of the difficulties of performing GT in higher plants is its very low frequency, due mainly to the low HR frequency. In rice, GT cells via naturally occurring HR can be screened if not using a sequence-specific nuclease such as CRISPR/Cas9. Just recently, we reported a CRISPR/Cas9-mediated DNA DSB-induced GT system using a vector harboring a CRISPR/Cas9 expression construct, selectable marker, and GT donor template (Nishizawa-Yokoi et al., 2020). Moreover, our previous report showed that DSB induction via CRISPR/Cas9, in combination with a deficiency of Ligase 4—a key enzyme in NHEJ competing

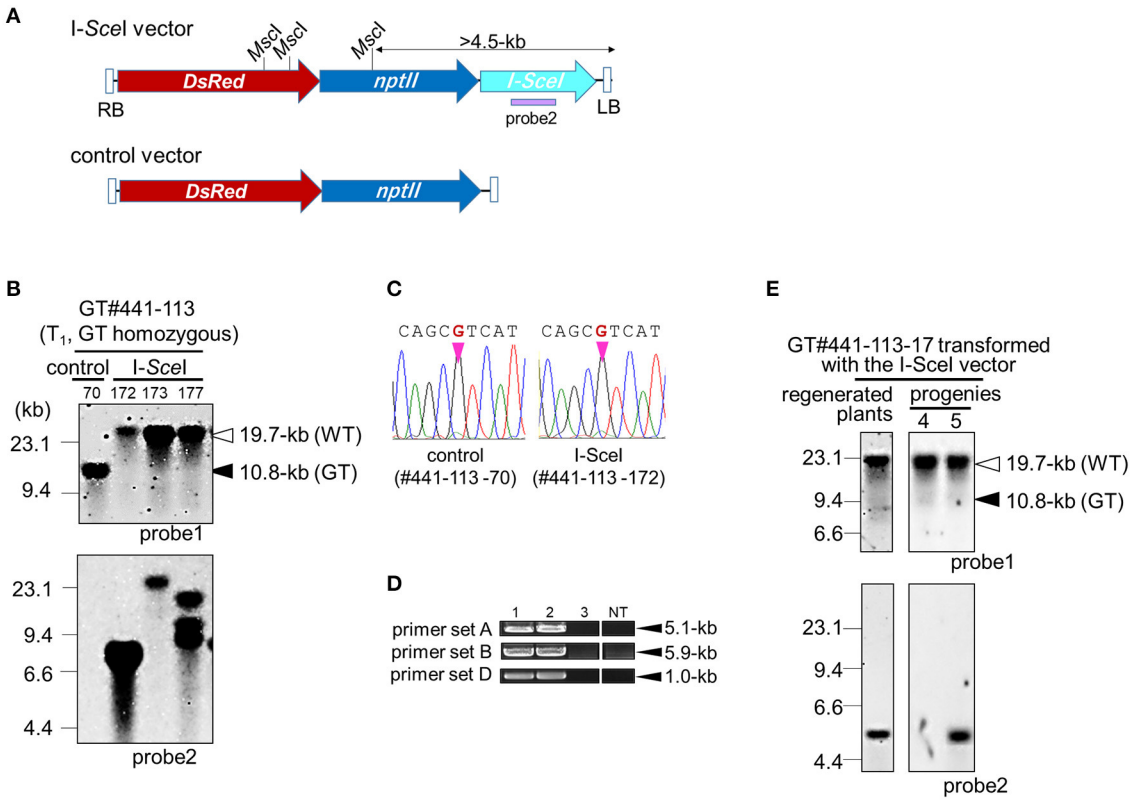


FIGURE 3 | Marker excision from GT locus in GT#441. **(A)** Schematic diagrams of T-DNA for the I-*Sce*I expression vector. The I-SceI vector (top) carries the I-*Sce*I gene under the control of CaMV 35S promoter, *npt* gene, and *DsRed* gene. The control vector (bottom) carries no I-*Sce*I expression cassette. A purple bar indicates probe 2 for Southern blot analysis. **(B)** Southern blot analyses with probe 1 or 2 using *Eco*RV or *Msc*I-digested genomic DNA, respectively. Samples are GT homozygous plants in the T_1 generation transformed with the I-SceI vector or control vector. White and black triangles indicated bands derived from the WT and marker-excised GT allele, and the GT allele, respectively. **(C)** Sequencing chromatograms at the mutation site on the *OsCly1* gene in GT#441-113-70 and GT#441-113-172. **(D)** PCR analysis to confirm marker excision. Sample 1: regenerated plants of GT#441-113; samples 2 and 3: GT homozygous plants of GT#441-113-70 and GT#441-113-172 transformed with control and I-SceI vector, respectively. NT; non-transformant. **(E)** Southern blot analyses with probe 1 or 2 using *Eco*RV- or *Msc*I-digested genomic DNA, respectively. Samples are regenerated plants and the progenies of GT homozygous plants of GT#441-113-17 (T_1 generation) transformed with the I-SceI vector. Details as in **(B)**.

with HR—could enhance GT frequency in rice (Endo et al., 2016). DSB induction via CRISPR/Cas9 will be used to improve positive–negative selection-mediated GT frequency in this experiment also.

Precise Elimination of a Positive-Marker Cassette From the GT Locus via I-*Sce*I-Mediated Break-Induced SSA

As the *hpt* gene is no longer needed after selection of true GT cells, the *hpt* gene cassette was excised from the GT locus by I-*Sce*I-mediated break-induced SSA. Here, two lines, GT#261 and GT#441 (**Figure 2**), were used for marker excision experiments, as shown schematically in **Figure 1B** and **Supplementary Figure 1**. Homozygous or heterozygous GT callus lines derived from T_1 seeds of GT#261 and GT#441 were infected individually with *Agrobacterium* harboring an I-*Sce*I expression vector driven by a double CaMV 35S promoter (Kwon et al., 2012) as shown in **Figure 3A**. *Agrobacterium*-infected calli were selected on medium containing G418. To screen cells in which the positive selection marker had been excised successfully from the GT locus in G418-resistant calli, PCR analysis with primer sets B and E (**Figure 1B**) was performed. Primer set B amplifies a 6.0-kb band in GT lines if the positive selection marker remains in the *OsCly1* locus, but not in marker-excised lines; primer set E amplifies a 11.3-kb band in GT lines still containing the positive selection marker in the *OsCly1* locus, while a 5.8-kb band is amplified in marker-excised lines and WT. As summarized in **Table 2**, the positive selection marker was excised from the *OsCly1* locus in over 25 and 90% of calli heterozygous and homozygous for the GT allele, respectively. Interestingly, bi-allelic marker excision was detected in 38 and 20% of calli in GT#261 and GT#441, respectively. In contrast, marker excision was not found in calli transformed with a control vector lacking the I-*Sce*I expression construct. In general, G418-resistant callus is a mosaic of marker excised and non-excised

TABLE 2 | Frequency of SSA-mediated marker excision from *OsCly1* locus.

GT Vector	GT calli lines		Vector	Marker excision from *OsCly1* locus			
	Line No. (generation)	GT allele		Without marker		With marker	Total
				Mono-allelic	Bi-allelic		
GT-*OsCly1* variation 1	GT#441-113 (T_1)	homozygous	Control	0	0	16	16
			I-*Sce*I	18	12	2	32
		heterozygous	Control	0	-	15	16
			I-*Sce*I	18	-	14	32
	GT# 261-7 (T_1)	homozygous	Control	0	0	16	16
			I-*Sce*I	22	6	2	30
		heterozygous	Control	0	-	16	16
			I-*Sce*I	8	-	24	32
		heterozygous	Control	0	-	16	16
			I-*Sce*I	5	-	11	16
	GT# 137 (T_0)	homozygous	Control	0	-	11	11
			I-*Sce*I	18	-	25	43

cells; thus, PCR fragments could be amplified using primer set B from cells neighboring those in which a positive marker was excised. Consequently, PCR analysis might underestimate marker excision frequency.

Regenerated plants were obtained from homozygous GT calli in which the positive selection marker had been excised successfully from the GT locus. In GT#441, Southern blot analysis of *Eco*RV-digested DNA from GT homozygous plants with probe 1 revealed a 19.7-kb band in both WT and a marker-excised GT line transformed with the I-*Sce*I vector, although bands for the GT allele (10.8-kb) were detected in regenerated plants transformed with a control vector (**Figure 3B**). This result suggests that the positive selection marker was completely excised from mono-allelic or bi-allelic *OsCly1* loci as expected. Direct sequence analyses of PCR fragments amplified with primer set E revealed that the desired mutations found in calli were maintained in these plants (**Figure 3C**). Moreover, PCR analyses of plants of GT#441-172 and GT#441-85 confirmed marker excision. Primer sets A, B, and D amplify fragments in case of successful targeted integration of the positive selection marker in *OsCly1* gene but not in WT or marker-excised lines (**Figure 1B**). As expected, fragments were not amplified in these lines using these primer sets (**Figure 3D**). In addition, Southern blot analysis using probe 2 revealed that the copy number of the I-*Sce*I vector was low in these plants (**Figures 3B,E**). In the next generation, line #441-113-17, in which a single copy of T-DNA was integrated, T-DNA of the I-*Sce*I vector inserted into the rice genome was segregated, and marker-excised plants without the I-*Sce*I vector were obtained successfully in both lines (**Figures 3B,E**). We confirmed successful marker excision in GT#261 also (**Supplementary Figure 3**). Thus, following marker excision, plants harboring precise genome editing with the desired point mutation in the miR172 target site in the *OsCly1* gene were obtained successfully by a combination of a positive–negative selection-mediated GT approach and subsequent SSA-mediated precise excision of the positive selectable marker.

Marker Excision Before Regeneration in the GT T_0 Generation

Next, to shorten the total experimental time, we attempted to excise the positive marker gene cassette immediately after GT. The T_0 callus lines GT#137 and GT#441, confirmed as GT events by PCR and Southern blot analysis, were used in this experiment (Exp. 3 in **Table 1**; **Supplementary Figure 1**). Three months after the first transformation of the GT vector, calli were infected with *Agrobacterium* harboring the I-*Sce*I vector (**Figure 3A**). One and half months after onset of G418 selection, excision of the positive selection marker was confirmed by PCR analyses with primer pair A or B. A total of 18 calli from 43 I-*Sce*I-transformed calli were seen to have lost the positive marker gene, whereas there were no marker-free calli in 11 lines transformed with a control vector (**Table 2**). Several plants regenerated from those GT#137 T_0 calli were analyzed further. Direct sequencing of PCR fragments amplified using primer set C or E showed the simultaneous detection of superposing signals of A and G at the 8th positions at the miRNA172 target site in the *OsCly1* gene in plants transformed with the I-*Sce*I vector (**Figures 4A,B**; **Supplementary Figure 4**). The results of Southern blot analyses of *Eco*RV- or *Msc*I- digested genomic DNA with probe 1 or 2, respectively, also supported the loss of the positive selection marker from the GT locus accompanying I-*Sce*I expression in this generation (**Figure 4B**; **Supplementary Figure 4B**). Thus, we had again successfully introduced the desired substitutions into the rice acetolactate synthase (*OsALS*) gene by GT and SSA-mediated marker excision (**Supplementary Figure 5**; **Supplementary Table 2**).

Notably, the SSA-mediated repair system does not have any limitations regarding the donor sequence on the GT vector, whereas the *piggyBac* system needs the "TTAA" recognition sequence to allow transposase PBase to remove the positive marker (Nishizawa-Yokoi et al., 2015a). Thus, for some specific genomic regions, application of the *piggyBac* system would be troublesome. On the other hand, we found that the efficiency

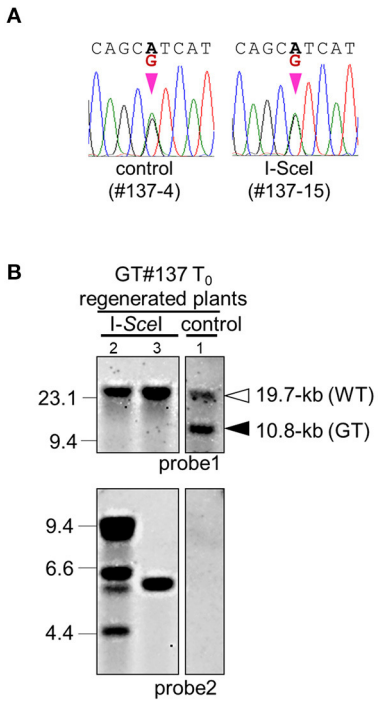


FIGURE 4 | Molecular analysis to confirm GT and marker excision events in GT#137. **(A)** Sequencing chromatograms of regenerated plants GT#137-4 and GT#137-15 transformed with control and I-*Sce*I vector, respectively. Sequences were determined using fragments amplified with primer set **(C)** in **Figure 1B**. **(B)** Southern blot analyses with probe 1 or 2 using *Eco*RV or *Msc*I-digested genomic DNA, respectively. Samples are regenerated plants of GT#137 transformed with the I-*Sce*I vector or control vector. Details as in **Figure 3B**.

of SSA-mediated marker excision was lower than that of the *piggyBac*-mediated system: in T_0 calli, the efficiency of the *piggyBac*-mediated system was nearly 100%, while that of the SSA-mediated system was around 40% (**Table 2**). The latter is thought to be adequate to produce marker-free rice plants. However, marker excision frequency would need to be further improved for plant species in which it is difficult to separate marker-excised cells from a mosaic of marker-excised and non-excised cells. It has been reported that SSA is not the sole DNA repair pathway in rice and Arabidopsis, even if overlapped sequences surround the DSB site (Kwon et al., 2012; Vu et al., 2014). Marker excision frequency might be improved by enhancement of the SSA pathway or suppression of DSB repair pathways other than SSA. We previously demonstrated that SSA can be enhanced by overexpression of rice exonuclease, OsExo1 and/or OsRecQl4 helicase (Kwon et al., 2012). Thus, overexpression of OsExo1 and OsRecQl4 would be expected to improve SSA-mediated marker excision.

Introduction of Multiple Substitutions in the *OsCly1* Gene

It is expected that the lower homology between the miR172 sequence and its target sequence in the *OsCly1* gene would result in greater tolerance to miR172-mediated downregulation of *OsCly1*. Therefore, we attempted to introduce multiple substitutions at the miR172 target site in the *OsCly1* gene. Another two GT vectors designed to introduce multiple substitutions at the miR172 target site were constructed (**Figure 1C**). GT-OsCly1 variation 2 also mimicked natural variations in the *cly1* gene of cleistogamous barley varieties. The substitutions in GT-*OsCly1* variation 2, T to C and A to C, are located at the 2nd and 14th positions of the miRNA172 target site in the *OsCly1* gene, respectively (Nair et al., 2010). On the other hand, GT-*OsCly1* variation 3 harbors not only three substitutions, located at the 2nd, 8th, and 14th positions at miRNA172 target site found as natural variations in barley, but also four artificial substitutions at all the triplet codon 3rd positions of the *OsCly1* gene. In designing the four artificial substitutions, care was taken not to create "rare codons." It is expected these substitutions will only affect transcript levels regulated by miRNA172 target because there are no base substitutions altering amino acid residues of the OsCly1 protein.

GT experiments using vectors GT-*OsCly1* variation 2 and 3 were performed as described above. Finally, 3 and 1 true independent GT lines with the desired substitutions were obtained from 1,404 and 1,560 calli transformed with GT-OsCly1 variation 2 and 3 vectors, respectively (**Table 1**). GT frequencies using GT-*OsCly1* variation 2 and 3 were thought to be comparable to those using GT-*OsCly1* variation 1 (**Table 1**). To confirm precise genome editing in regenerated plants obtained from GT calli, molecular analyses were performed in GT#75 and GT#220 in GT-*OsCly1* variations 2 and 3, respectively. Direct sequence analysis revealed that desired substitutions of 2 and 7 bases at the miR172 target site in the *OsCly1* gene were introduced successfully in T_0 plants (**Figure 5A**). PCR analysis also showed successful marker excision from the GT allele in regenerated plants (**Figure 5B**). In sum, we demonstrate that our system could be used for precise rice genome modifications, from single base substitutions to multiple discontinuous base changes.

Phenotype of *OsCly1*-Edited Rice Plants

In barley, variations in miR172 target sequences in the *cly1* gene are involved in the cleistogamous phenotype (Nair et al., 2010). In addition, the rice *oscly1* mutant frequently showed enlarged lodicules (Zhou et al., 2012). Here, we observed the floral organs in GT homozygous plants with A to G substitution at the 8th position at miRNA172 target sequences in *OsCly1* gene (GT-*OsCly1* variation 1, **Figure 1C**). We grew GT homozygous plants harboring the positive selection marker, GT#34-6-53 and #441-113-115-38, in a greenhouse under natural long-day conditions. In GT#34-6-53 (see **Supplementary Figure 2C**, T_1 generation), the size of lodicules in GT plants was significantly smaller than that of WT lodicules (**Figure 6A**), similar to the phenotype observed in the recessive *cly1* homozygous barley plants (Nair et al., 2010). Furthermore, GT plants showed much less stamen exertion compared with WT plants (**Figure 6B**). Interestingly, closed flowers were observed in #441-113-115-38 (a progeny of #441-113 in **Figure 2B**, T_2 generation), although

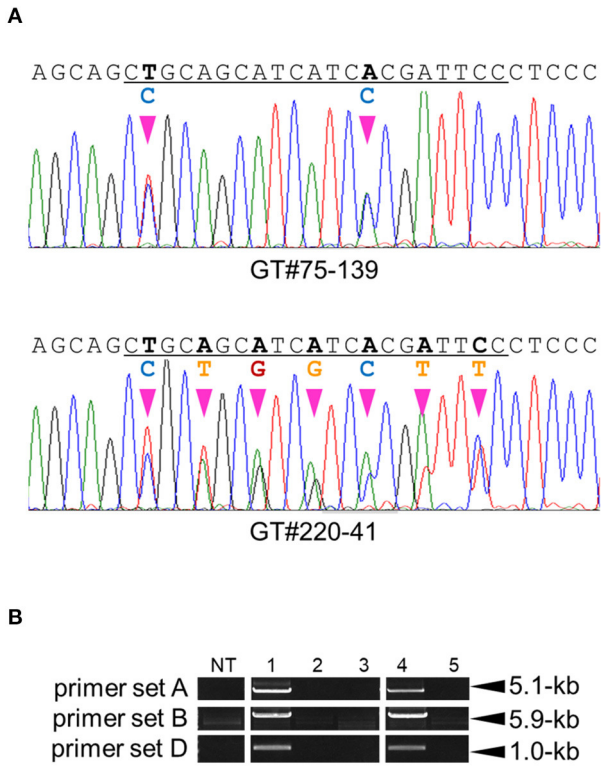


FIGURE 5 | Molecular analysis to confirm GT and marker excision events in GT#75 and 220. **(A)** Sequencing chromatograms of regenerated plants, GT#75-139 and GT#220-41 derived from GT calli transformed with GT vectors harboring mutations induced by GT-OsCly1 variation 2 and 3, respectively. The miR172 target site region is underlined. **(B)** PCR analysis to confirm marker excision in GT#75 and GT#220. Samples 1–3: GT regenerated plants (GT#75) transformed with control vector (1), and I-*Sce*I vector (2, 3). Samples 4, 5: GT regenerated plants (GT#220) transformed with control vector (4), and I-*Sce*I vector (5). NT; non-transformant. Details as in **Figure 3D**.

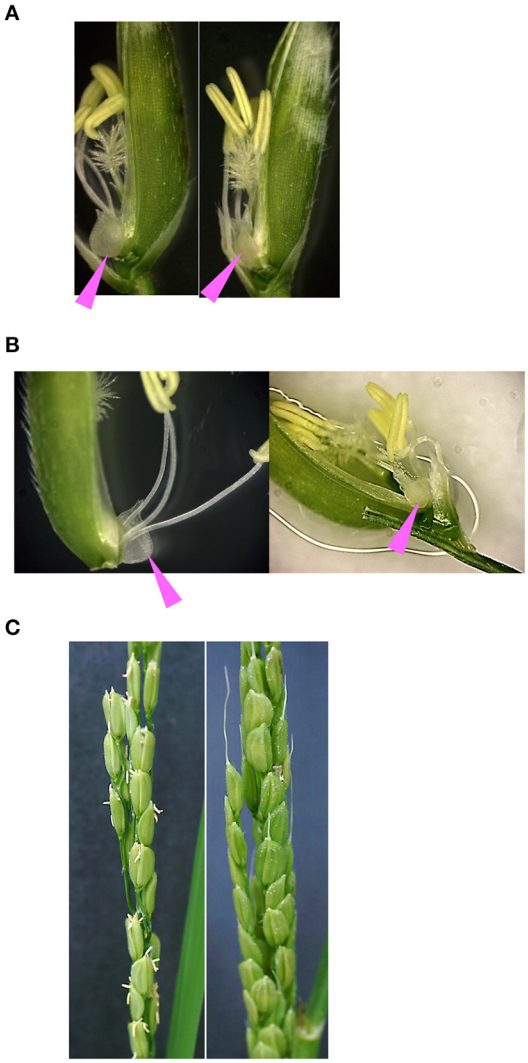


FIGURE 6 | Phenotype of floral organs in GT homozygous plants. Photos of segregated WT (left) and GT homozygous (right) plants in the T_1 generation harboring an A to G substitution at the 8th position (variation 1) at the miR172 target site in the *OsCly1* gene possessing a positive selection marker. Pink arrowheads indicate lodicules. **(A,B)** Spikelets (WT and GT plants) of #34-6-53 observed under a light microscope. **(C)** Panicles (WT and GT plants) of #441-113-115-38 at flowering time.

flowers opened in segregated WT (**Figure 6C**), suggesting that sequence variation (variation 1) at the miR172 target site in the *OsCly1* gene would be involved in cleistogamous flowering, as in the case of barley *cly1*. Further observation of GT plants without a positive selection marker and with other substitutions (GT-*OsCly1* variation 2 and 3) is planned in the near future.

Future Prospects

Here, we demonstrated that GT and SSA-mediated marker excision allows desired mutations such as substitution of 2 and 7 discontinuous bases to be introduced into a target gene in rice. An earlier report showed that 18 single-base substitutions and 3 single-base deletions were introduced simultaneously at sites within 12.2-kb target sequences in rice via positive–negative-mediated GT (Johzuka-Hisatomi et al., 2008). As discussed in the *Introduction*, this GT and marker excision system could be a powerful tool to precisely modify target sequences that are difficult to access using conventional mutagenesis, base editors and prime editing. To induce SSA, we transformed GT calli with an I-*Sce*I expression vector. This might cause more somaclonal mutations, with resultant reduced regeneration ability, due to the nature of long de-differentiated callus culture. Inducible I-SceI expression is a possible approach to prevent this problem, although strict ON/OFF regulation of I-SceI expression would be necessary. Moreover, we succeeded in producing rice plants with the desired phenotype via precise mutagenesis of the miRNA target site in the *OsCly1* gene. miRNAs regulate important agronomical traits such as grain number, filling rate, fertility, and leaf inclination in rice

(Peng et al., 2019). For example, a single substitution in the miR156 target site of *OsSPL14* gene involves *OsSPL14* mRNA level regulated by miR156, resulting in an increase in grain yield (Jiao et al., 2010; Miura et al., 2010). Substitutions to inhibit miRNA binding to its target gene via the system presented in this study could produce valuable rice plants. Moreover, in general, there are homologs that show highly conserved sequences in some miRNAs (Reinhart et al., 2002). Precise modification by our system enables the expression levels of miRNA-targeted genes to be regulated more strictly and specifically.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Materials**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

TK, KH, ST, and HS designed the experiments. NO, KK, AM, AN-Y, HY, and HS performed the experiments. NO and HS wrote the article with contributions of all the authors. ST supervised and complemented the writing.

FUNDING

This work was supported by NARO grant-in-aid, 20902, the Program for Promotion of Basic and Applied Researches for Innovations in Bio-oriented Industry, JSPS KAKENHI (Grant Number 24658011) to HS, and the Cabinet Office, Government of Japan, Cross-ministerial Strategic Innovation Promotion Program (SIP), Technologies for creating next-generation agriculture, forestry, and fisheries and Technologies for Smart Bio-industry and Agriculture (funding agency: Bio-oriented Technology Research Advancement Institution, NARO) to ST.

ACKNOWLEDGMENTS

We thank K. Osakabe, H. Kaya, A. Endo, M. Endo, S. Hirose, K. Abe, M. Mikami in NARO for advice and discussion, and K. Amagai, R. Aoto, C. Furusawa, A. Nagashii, F. Suzuki and R. Takahashi in NARO for general experimental technical support. We also thank Dr. H. Rothnie for English editing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgeed.2020.617713/full#supplementary-material

REFERENCES

Anwar, N., Ohta, M., Yazawa, T., Sato, Y., Li, C., Tagiri, A., et al. (2018). miR172 downregulates the translation of cleistogamy 1 in barley. *Ann. Bot.* 122, 251–265. doi: 10.1093/aob/mcy058

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., et al. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576, 149–157. doi: 10.1038/s41586-019-1711-4

Chapman, J. R., Taylor, M. R., and Boulton, S. J. (2012). Playing the end game: DNA double-strand break repair pathway choice. *Mol. Cell* 47, 497–510. doi: 10.1016/j.molcel.2012.07.029

Dang, T. T., Shimatani, Z., Kawano, Y., Terada, R., and Shimamoto, K. (2013). Gene editing a constitutively active OsRac1 by homologous recombination-based gene targeting induces immune responses in rice. *Plant Cell Physiol.* 54, 2058–2070. doi: 10.1093/pcp/pct147

Endo, M., Iwakami, S., and Toki, S. (2020). Precision genome editing in plants via gene targeting and subsequent break-induced single-strand annealing. *Plant Biotechnol. J.* doi: 10.1111/pbi.13485. [Epub ahead of print].

Endo, M., Mikami, M., and Toki, S. (2016). Biallelic gene targeting in rice. *Plant Physiol.* 170, 667–677. doi: 10.1104/pp.15.01663

Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I., et al. (2017). Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* 551, 464–471. doi: 10.1038/nature24644

Golic, K. G., and Lindquist, S. (1989). The FLP recombinase of yeast catalyzes site-specific recombination in the Drosophila genome. *Cell* 59, 499–509. doi: 10.1016/0092-8674(89)90033-0

Hajdukiewicz, P., Svab, Z., and Maliga, P. (1994). The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. *Plant Mol. Biol.* 25, 989–994. doi: 10.1007/BF00014672

Hasty, P., Ramirez-Solis, R., Krumlauf, R., and Bradley, A. (1991). Introduction of a subtle mutation into the *Hox-2.6* locus in embryonic stem cells. *Nature* 350, 243–246. doi: 10.1038/350243a0

Hood, E. E., Gelvin, S. B., Melchers, L. S., and Hoekema, A. (1993). New *Agrobacterium* helper plasmids for gene transfer to plants. *Transgenic Res.* 2, 208–218. doi: 10.1007/BF01977351

Hustedt, N., and Durocher, D. (2016). The control of DNA repair by the cell cycle. *Nat. Cell Biol.* 19, 1–9. doi: 10.1038/ncb3452

Jiao, Y., Wang, Y., Xue, D., Wang, J., Yan, M., Liu, G., et al. (2010). Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. *Nat. Genet.* 42, 541–544. doi: 10.1038/ng.591

Johzuka-Hisatomi, Y., Terada, R., and Iida, S. (2008). Efficient transfer of base changes from a vector to the rice genome by homologous recombination: involvement of heteroduplex formation and mismatch correction. *Nucleic Acids Res.* 36, 4727–4735. doi: 10.1093/nar/gkn451

Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. doi: 10.1038/nature17946.

Kurt, I. C., Zhou, R., Iyer, S., Garcia, S. P., Miller, B. R., Langner, L. M., et al. (2020). CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nat. Biotechnol.* doi: 10.1038/s41587-020-0609-x. [Epub ahead of print].

Kwon, Y. I., Abe, K., Osakabe, K., Endo, M., Nishizawa-Yokoi, A., Saika, H., et al. (2012). Overexpression of OsRecQl4 and/or OsExo1 enhances dsb-induced homologous recombination in rice. *Plant Cell Physiol.* 53, 2142–2152. doi: 10.1093/pcp/pcs155

Lin, Q., Zong, Y., Xue, C., Wang, S., Jin, S., Zhu, Z., et al. (2020). Prime genome editing in rice and wheat. *Nat. Biotechnol.* 38, 582–585. doi: 10.1038/s41587-020-0455-x

Lombardo, F., Kuroki, M., Yao, S. G., Shimizu, H., Ikegaya, T., Kimizu, M., et al. (2017). The superwoman1-cleistogamy2 mutant is a novel resource for gene containment in rice. *Plant Biotechnol. J.* 15, 97–106. doi: 10.1111/pbi.12594

Mikami, M., Toki, S., and Endo, M. (2015a). Comparison of CRISPR/Cas9 expression constructs for efficient targeted mutagenesis in rice. *Plant Mol. Biol.* 88, 561–572. doi: 10.1007/s11103-015-0342-x

Mikami, M., Toki, S., and Endo, M. (2015b). Parameters affecting frequency of CRISPR/Cas9 mediated targeted mutagenesis in rice *Plant Cell Rep.* 34, 1807–1815. doi: 10.1007/s00299-015-1826-5

Miura, K., Ikeda, M., Matsubara, A., Song, X. J., Ito, M., Asano, K., et al. (2010). OsSPL14 promotes panicle branching and higher grain productivity in rice. *Nat. Genet.* 42, 545–549. doi: 10.1038/ng.592

Morioka, Y., Fujihara, Y., and Okabe, M. (2014). Generation of precise point mutation mice by footprintless genome modification. *Genesis* 52, 68–77. doi: 10.1002/dvg.22727

Nair, S. K., Wang, N., Turuspekov, Y., Pourkheirandish, M., Sinsuwongwat, S., Chen, G., et al. (2010). Cleistogamous flowering in barley arises from the suppression of microRNA-guided HvAP2 mRNA cleavage. *Proc. Natl. Acad. Sci. U.S.A.* 107, 490–495. doi: 10.1073/pnas.0909097107

Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., et al. (2016). Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* 353:aaf8729. doi: 10.1126/science.aaf8729

Nishizawa-Yokoi, A., Endo, M., Ohtsuki, N., Saika, H., and Toki, S. (2015a). Precision genome editing in plants via gene targeting and *piggyBac*-mediated marker excision. *Plant J.* 81, 160–168. doi: 10.1111/tpj.12693

Nishizawa-Yokoi, A., Mikami, M., and Toki, S. (2020). A universal system of CRISPR/Cas9-mediated gene targeting using all-in-one vector in plants. *Front. Genome Ed.* 2:604289. doi: 10.3389/fgeed.2020.604289

Nishizawa-Yokoi, A., Nonaka, S., Osakabe, K., Saika, H., and Toki, S. (2015b). A universal positive-negative selection system for gene targeting in plants combining an antibiotic resistance gene and its antisense RNA. *Plant Physiol.* 169, 362–370. doi: 10.1104/pp.15.00638

Onouchi, H., Yokoi, K., Machida, C., Matsuzaki, H., Oshima, Y., Matsuoka, K., et al. (1991). Operation of an efficient site-specific recombination system of Zygosaccharomyces rouxii in tobacco cells. *Nucleic Acids Res.* 19, 6373–6378. doi: 10.1093/nar/19.23.6373

Osakabe, K., Endo, M., Kawai, K., Nishizawa, Y., Ono, K., Abe, K., et al. (2005). The mutant form of acetolactate synthase genomic DNA from rice is an efficient selectable marker for genetic transformation. *Mol. Breed.* 16, 313–320. doi: 10.1007/s11032-005-0999-y

Paszkowski, J., Baur, M., Bogucki, A., and Potrykus, I. (1988). Gene targeting in plants. *EMBO J.* 7, 4021–4026. doi: 10.1002/j.1460-2075.1988.tb03295.x

Peng, T., Teotia, S., Tang, G., and Zhao, Q. (2019). MicroRNAs meet with quantitative trait loci: small powerful players in regulating quantitative yield traits in rice. *Wiley Interdiscip. Rev. RNA* 10:e1556. doi: 10.1002/wrna.1556

Puchta, H., and Hohn, B. (1991). A transient assay in plant cells reveals a positive correlation between extrachromosomal recombination rates and length of homologous overlap. *Nucleic Acids Res.* 19, 2693–2700. doi: 10.1093/nar/19.10.2693

Reinhart, B. J., Weinstein, E. G., Rhoades, M. W., Bartel, B., and Bartel, D. P. (2002). MicroRNAs in plants. *Genes Dev.* 16, 1616–1626. doi: 10.1101/gad.1004402

Saika, H., Mori, A., Endo, M., Osakabe, K., and Toki, S. (2015). Rapid evaluation of the frequency of gene targeting in rice via a convenient positive-negative selection method. *Plant Biotechnol.* 32, 169–173. doi: 10.5511/plantbiotechnology.15.0427a

Sauer, B., and Henderson, N. (1990). Targeted insertion of exogenous DNA into the eukaryotic genome by the Cre recombinase. *New Biol.* 2, 441–449.

Shimatani, Z., Nishizawa-Yokoi, A., Endo, M., Toki, S., and Terada, R. (2015). Positive-negative-selection-mediated gene targeting in rice. *Front. Plant Sci.* 5:748. doi: 10.3389/fpls.2014.00748

Sun, N., and Zhao, H. (2014). Seamless correction of the sickle cell disease mutation of the HBB gene in human induced pluripotent stem cells using TALENs. *Biotechnol. Bioeng.* 111, 1048–1053. doi: 10.1002/bit.25018

Terada, R., Nagahara, M., Furukawa, K., Shimamoto, M., Yamaguchi, K., and Iida, S. (2010). Cre-loxP mediated marker elimination and gene reactivation at the waxy locus created in rice genome based on strong positive-negative selection. *Plant Biotechnol.* 27, 29–37. doi: 10.5511/plantbiotechnology.27.29

Terada, R., Urawa, H., Inagaki, Y., Tsugane, K., and Iida, S. (2002). Efficient gene targeting by homologous recombination in rice. *Nat. Biotechnol.* 20, 1030–1034. doi: 10.1038/nbt737

Toki, S. (1997). Rapid and efficient *Agrobacterium*-mediated transformation in rice. *Plant Mol. Biol. Rep.* 15, 16–21. doi: 10.1007/BF02772109

Toki, S., Hara, N., Ono, K., Onodera, H., Tagiri, A., Oka, S., et al. (2006). Early infection of scutellum tissue with *Agrobacterium* allows high-speed transformation of rice. *Plant J.* 47, 969–976. doi: 10.1111/j.1365-313X.2006.02836.x

Valancius, V., and Smithies, O. (1991). Testing an "in-out" targeting procedure for making subtle genomic modifications in mouse embryonic stem cells. *Mol. Cell. Biol.* 11, 1402–1408. doi: 10.1128/MCB.11.3.1402

Vu, G. T. H., Cao, H. X., Watanabe, K., Hensel, G., Blattner, F. R., Kumlehn, J., et al. (2014). Repair of site-specific DNA double-strand breaks in barley occurs via diverse pathways primarily involving the sister chromatid. *Plant Cell* 26, 2156–2167. doi: 10.1105/tpc.114.126607

Xu, W., Zhang, C., Yang, Y., Zhao, S., Kang, G., He, X., et al. (2020). Versatile nucleotides substitution in plant using an improved prime editing system. *Mol. Plant* 13, 675–678. doi: 10.1016/j.molp.2020.03.012

Yoshida, H., Itoh, J., Ohmori, S., Miyoshi, K., Horigome, A., Uchida, E., et al. (2007). superwoman1-cleistogamy, a hopeful allele for gene containment in GM rice. *Plant Biotechnol. J.* 5, 835–846. doi: 10.1111/j.1467-7652.2007.00291.x

Yusa, K., Rashid, S. T., Strick-Marchand, H., Varela, I., Liu, P. Q., Paschon, D. E., et al. (2011). Targeted gene correction of alpha(1)-antitrypsin deficiency in induced pluripotent stem cells. *Nature* 478, 391–394. doi: 10.1038/nature10424

Zhao, D., Li, J., Li, S., Xin, X., Hu, M., Price, M. A., et al. (2020). New base editors change C to A in bacteria and C to G in mammalian cells. *Nat. Biotechnol.* doi: 10.1038/s41587-020-0592-2. [Epub ahead of print].

Zhou, Y., Lu, D., Li, C., Luo, J., Zhu, B. F., Zhu, J., et al. (2012). Genetic control of seed shattering in rice by the APETALA2 transcription factor shattering abortion1. *Plant Cell* 24, 1034–1048. doi: 10.1105/tpc.111.094383

Zhu, Q. H., and Helliwell, C. A. (2011). Regulation of flowering time and floral patterning by miR172. *J. Exp. Bot.* 62, 487–495. doi: 10.1093/jxb/erq295

Zhu, Q. H., Upadhyaya, N. M., Gubler, F., and Helliwell, C. A. (2009). Over-expression of miR172 causes loss of spikelet determinacy and floral organ abnormalities in rice (*Oryza sativa*). *BMC Plant Biol.* 9:149. doi: 10.1186/1471-2229-9-149

Conflict of Interest: ST has filed a patent application related to this work. Authors KK and KH were employed by the company, Nisshin Flour Milling Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Ohtsuki, Kizawa, Mori, Nishizawa-Yokoi, Komatsuda, Yoshida, Hayakawa, Toki and Saika. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genome Editing

ORIGINAL RESEARCH
published: 21 January 2021
doi: 10.3389/fgeed.2020.627803

Genome Editing and Protoplast Regeneration to Study Plant–Pathogen Interactions in the Model Plant *Nicotiana benthamiana*

Chen-Tran Hsu†, Wen-Chi Lee†, Yu-Jung Cheng, Yu-Hsuan Yuan, Fu-Hui Wu and Choun-Sea Lin*

Agricultural Biotechnology Research Center, Academia Sinica, Taipei, Taiwan

Biotic diseases cause substantial agricultural losses annually, spurring research into plant pathogens and strategies to mitigate them. *Nicotiana benthamiana* is a commonly used model plant for studying plant–pathogen interactions because it is host to numerous plant pathogens and because many research tools are available for this species. The clustered regularly interspaced short palindromic repeats (CRISPR) system is one of several powerful tools available for targeted gene editing, a crucial strategy for analyzing gene function. Here, we demonstrate the use of various CRISPR-associated (Cas) proteins for gene editing of *N. benthamiana* protoplasts, including *Staphylococcus aureus* Cas9 (SaCas9), *Streptococcus pyogenes* Cas9 (SpCas9), *Francisella novicida* Cas12a (FnCas12a), and nCas9-activation-induced cytidine deaminase (nCas9-Target-AID). We successfully mutated *Phytoene Desaturase* (*PDS*) and *Ethylene Receptor 1* (*ETR1*) and the disease-associated genes *RNA-Dependent RNA Polymerase 6* (*RDR6*), and *Suppressor of Gene Silencing 3* (*SGS3*), and confirmed that the mutated alleles were transmitted to progeny. *sgs3* mutants showed the expected phenotype, including absence of *trans-acting siRNA3* (*TAS3*) siRNA and abundant expression of the GFP reporter. Progeny of both *sgs3* and *rdr6* null mutants were sterile. Our analysis of the phenotypes of the regenerated progeny indicated that except for the predicted phenotypes, they grew normally, with no unexpected traits. These results confirmed the utility of gene editing followed by protoplast regeneration in *N. benthamiana*. We also developed a method for *in vitro* flowering and seed production in *N. benthamiana*, allowing the regenerants to produce progeny *in vitro* without environmental constraints.

Keywords: FnCas12a, nCas9-Target-AID, RDR6, SaCas9, SGS3, SpCas9

OPEN ACCESS

Edited by:
Lanqin Xia,
Chinese Academy of Agricultural Sciences, China

Reviewed by:
Anshu Alok,
Panjab University, India
Yanfei Mao,
Chinese Academy of Sciences (CAS), China

****Correspondence:***
Choun-Sea Lin
cslin99@gate.sinica.edu.tw
orcid.org/0000-0001-9566-2952

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Genome Editing in Plants, a section of the journal Frontiers in Genome Editing

Received: *10 November 2020*
Accepted: *28 December 2020*
Published: *21 January 2021*

Citation:
Hsu C-T, Lee W-C, Cheng Y-J, Yuan Y-H, Wu F-H and Lin C-S (2021) Genome Editing and Protoplast Regeneration to Study Plant–Pathogen Interactions in the Model Plant Nicotiana benthamiana. Front. Genome Ed. 2:627803. doi: 10.3389/fgeed.2020.627803

INTRODUCTION

Nicotiana benthamiana is a host to many plant pathogens, especially viruses, and is widely used to study plant–pathogen interactions (Goodin et al., 2008). Many tools for functional genomics are available for this species, including viral vectors, RNA interference (RNAi), ethyl methanesulfonate mutagenesis, agroinfiltration, protoplast transfection, and *Agrobacterium*-mediated stable transformation. These tools are useful for research in genomics, biochemistry, metabolomics, cell biology, and pathology, as well as other topics in agriculture (Derevnina et al., 2019).

Notwithstanding its many advantages, the fact that *N. benthamiana* is allotetraploid, with a very large genome (3.1 Gb) (Bombarely et al., 2012), makes it difficult to edit the genome of this plant and to obtain mutants for plant biological and gene functional studies. We chose to address this problem by using the powerful genome-editing tool CRISPR-Cas (clustered regularly interspaced short palindromic repeats-CRISPR-associated protein). CRISPR-Cas core technology involves programmable DNA cleavage by the Cas protein at DNA sites specified by the targeting sequence in a guide RNA (gRNA; review by Yue et al., 2020). The use of CRISPR-Cas has greatly accelerated plant research and crop breeding in recent years (Li et al., 2013; Nekrasov et al., 2013; Shan et al., 2013; Li and Xia, 2020; Yue et al., 2020).

Most genome editing studies in plants, including *N. benthamiana*, have involved Agrobacterium-mediated stable transformation to deliver DNA into target cells in order to express Cas protein and gRNA. However, mutant plants derived from Agrobacterium-mediated transformation could be considered genetically modified organisms (GMOs), especially for vegetatively propagated crops in which the transgenes cannot be removed from the genome by crossing. In dicots, however, most transformants are chimeric, and the edited allele cannot be transmitted to the progeny when the edited cells exist only in vegetative organs. Thus, as a less controversial alternative, plasmids encoding the Cas and gRNA sequences or pre-assembled Cas:gRNA ribonucleoprotein complexes (RNPs) can be delivered directly into protoplasts using transient transfection. Because the protoplast is a single cell, once gene editing has been performed, the entire regenerant derived from this edited protoplast will contain the same edited gene (Woo et al., 2015; Lin et al., 2018; Hsu et al., 2019). Although a similar type of delivery can also be achieved by particle bombardment, polyethylene glycol (PEG)-mediated protoplast transfection offers high transfection efficiency and high viability for robust gene editing while generating recombinant-DNA-free plants to circumvent GMO issues (Woo et al., 2015; Andersson et al., 2018; Lin et al., 2018).

The main bottleneck of this strategy, however, is protoplast regeneration. We previously established a protoplast regeneration system and a CRISPR-Cas gene editing system for polyploid tobacco (*N. tabacum*) (Lin et al., 2018; Hsu et al., 2019). Here we report a simple, highly robust protocol for streamlined CRISPR-mediated genome editing in *N. benthamiana*. This protocol, together with CRISPR genome editing and improved genomics resources, ushers in a new era of forward and reverse genetic analyses of this valuable model plant system.

MATERIALS AND METHODS

Plant Materials

Sterile *N. benthamiana* plantlets were propagated by cutting and grown in half-strength Murashige and Skoog (1/2 MS) medium supplemented with 30 mg/L sucrose and 1% agar, pH 5.7. These plantlets were incubated in a 26°C culture room (12 h light /12 h dark cycle) with a light density of 75 μmol m^{-2} s^{-1}. The plantlets were subcutlured into fresh medium every month. For comparison with the seedlings derived from protoplasts and seed propagation, seeds were sown in 3-inch pots with peat moss, vermiculite and perlite in a ratio of 5:1:1. Each treatment had five repeats.

Protoplast Isolation and Transfection

The protoplast isolation and transfection followed our previously published method with minor modification (Hsu et al., 2019). The protoplasts were isolated from the mature leaves of *in vitro* plantlets. Five to seven leaves (about 0.2–0.25 g) were used for 10^6 protoplast isolation. These leaves were put into a 6-cm glass petri dish with 10 ml digestion solution (1/4 MS liquid medium containing 1% cellulose, 0.5% Macerozyme, 3% sucrose and 0.4 M mannitol, pH 5.7) and cut into 0.5 cm-wide strips. The solution was incubated at room temperature in the dark overnight. The digested solution was diluted with 10 ml W5 (154 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl, 2 mM MES, and 5 mM glucose) solution and filtered by 40 μm nylon mesh. The solution was centrifuged at low-speed (200 × g) for 3 min to collect the protoplasts. The protoplasts were purified with 20% sucrose solution and washed in W5 solution three times. The protoplasts were transferred to a transfection buffer (1/2 MS solution supplemented with 3% sucrose, 0.4 M mannitol, 1 mg/L naphthaleneacetic acid (NAA), and 0.3 mg/L kinetin, 5 mM MES, pH 5.7) and the cell concentration was adjusted to 3×10^5/mL.

The protoplasts were transfected with plasmids by PEG-mediated transfection (Woo et al., 2015; Lin et al., 2018). CRISPR reagent DNA (40 μg in 40 μl) was added to 400 μl (1.2×10^5 protoplasts) and mixed carefully. Then the same volume of PEG solution was added and mixed, then left to stand for 30 min. To end the reaction, 3 ml of W5 was added and mixed well. The transfected protoplasts were collected by centrifugation at 200 × g for 3 min. The protoplasts were washed in 3 ml of W5 by centrifugation at 200 × g for 3 min.

Plasmids

Several target sites in *N. benthamiana* whose editing efficiencies have been confirmed in *N. tabacum* (Hsu et al., 2019) as well as new constructions were used in this study. The following Cas proteins and target genes were tested in *N. benthamiana*:

1. SaCas9: The binary plasmid (gPDA_Sa) was published by Kaya et al. (2016). The target gene is *Phytoene Desaturase 1 (NbPDS-1)*, the target site is TTGCGATGCCTAACAAGC CAG.
2. FnCas12a: The binary plasmids (*crNtPDS-1* and *crNtPDS-2*) were published by Endo et al. (2016). Target genes are *NbPDS-1* and *NbPDS-2*, and the target sites are TCATCCAGTCCTT AACACTTAAAC*(crNtPDS-1)*, and ACATGGCAATGAACA CCTCATCTG *(crNtPDS-1)*.
3. nCas9-Target-AID: The plasmid (pDicAID_nCas9-PmCDA-2A-NptII_ETR) is published in Shimatani et al. (2017) (Addgene ID: 91695). The target genes are *NbETR1-1* and *NbETR1-2*. The target site is TGCACAAGAACCCATCTATA.
4. SpCas9: the vector commonly used for dicot transformation (pYLCRISPR/Cas9P35S-N) is used (Ma et al., 2015). The target genes were *RNA-dependent RNA Polymerase 6*

(NbRDR6-1 and *NbRDR6-2*), and *Suppressor of Gene Silencing 3 (NbSGS3-1* and *NbSGS3-2*). For convenience, to validate the presence and efficiency of the mutations, double sgRNAs were present in a single construct for each gene *(NbSGS3*: AAGC AGTGCTGGGAAGCAAT, CTCATGCCACGATGGCCTTG; *NbRDR6*: GCCATGGCCTTCTCAAAGCT, GCAGTTCTA TAGAAAACCAA). The sgRNAs were cloned into vectors.

Protoplast Regeneration

Pooled protoplast DNA was used as a template to amplify the target genes for validation by sequencing. The putatively edited protoplasts were transferred to 5 cm diameter Petri dishes containing 3 ml 1/2 MS liquid medium supplemented with 3% sucrose, 0.4 M mannitol, 1 mg/L NAA, and 0.3 mg/L kinetin (1N0.3K) for plant regeneration. Callus formation occurred using protoplasts after 1 month of incubation in the dark. The calluses were subcultured in 9 cm diameter Petri dishes containing fresh medium with 1 mg/L 6-benzylaminopurine (1B) for 3–4-weeks in the light. Calluses that had turned green were then transferred to solid medium containing the same plant growth regulators. The explants were subcultured every 4 weeks until shoots formed after several subcultures. The shoots were subcultured in solid root medium (HB1: 3 g/L Hyponex No. 1, 2 g/L tryptone, 20 g/L sucrose, 1 g/L activated charcoal, 10 g/L Agar, pH 5.2). Adventitious roots formed at the bottoms of the containers (**Figure 1**).

Genotype Analysis of Regenerated Plants

Two pairs of primers were designed to amplify the sgRNA-targeted DNA region for each target gene. PCR conditions were 94°C for 5 min, 35 cycles of denaturing (94°C for 30 s), annealing (55°C for 30 s), and polymerization (72°C for 30 s), followed by an extension reaction at 72°C for 5 min. The PCR product was sequenced by the Sanger method to determine the mutagenesis. The multiple sequences derived from mutated regenerated plants were separated using Poly Peak Parser (http://yosttools.genetics.utah.edu/PolyPeakParser/; Hill et al., 2014) or further confirmed by sequential T/A cloning and sequencing.

RESULTS

N. benthamiana Protoplast Regeneration

For protoplast regeneration, we placed protoplasts isolated from the leaves of *in vitro*-grown shoots (**Figure 1a**) in 1N0.3K liquid medium, incubated them in the dark for 1 month (**Figure 1b**), and then transferred them to fresh 1N0.3K medium and incubated them in the dark for another month (**Figure 1c**). Unlike in our previous method described for tobacco (Lin et al., 2018), we incubated *N. benthamiana* calluses directly in liquid 1B medium in light without embedding (**Figure 1d**). This step avoids the difficulty associated with embedding; however, it has the disadvantage that the calluses stick together and sometimes cannot be distinguished. After 1 month, we transferred the calluses larger than 3 mm to 1B solid medium and incubated them in light (**Figure 1e**). After several subcultures, shoots formed on the surface of the calluses (**Figure 1f**); this took more time for *N. benthamiana* than it does for *N. tabacum* (Lin et al., 2018). We then subcultured the shoots in solid HB1 medium and observed that adventitious roots formed without the need for plant growth regulators (**Figure 1g**). These plants could be further incubated successfully in test tubes, where they flowered and produced seeds (**Figure 1h**), or were transferred to a growth chamber for further growth (**Figure 1i**). The time required from protoplast isolation to regeneration was ∼4–6 months.

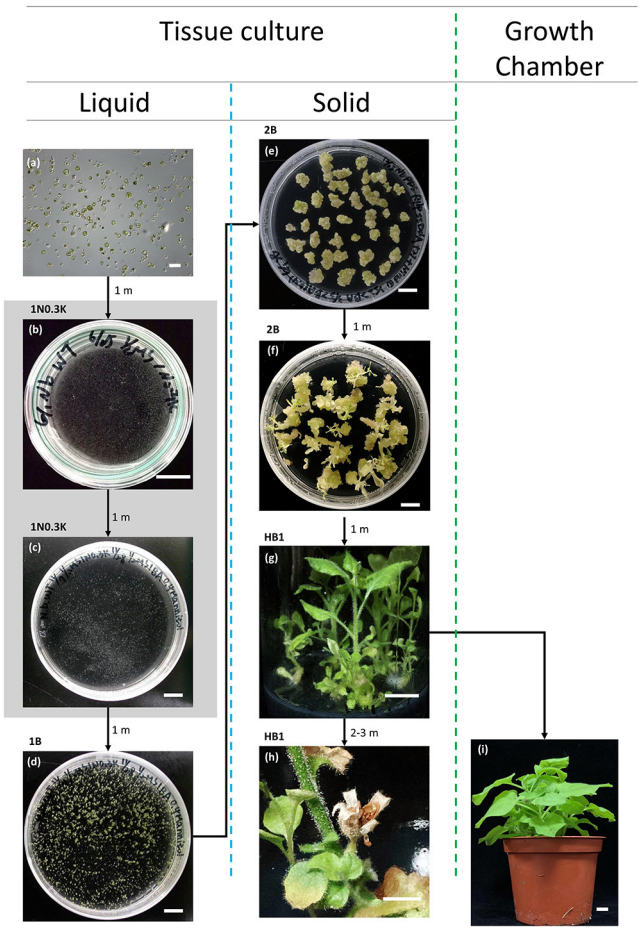


FIGURE 1 | Protoplast regeneration in *N. benthamiana*. The gray background indicates that the explants were incubated in the dark. All liquid media were supplemented with 0.4 M mannitol. The abbreviated names of the media are shown in the top left corner of the images. 1N0.3K: 1/2 MS basal medium supplemented with 1 mg/L NAA and 0.3 mg/L kinetin. 1B: ½ MS basal medium supplemented with 1 mg/L BA and 20 g/L sucrose. All solid media were solidified with 7 g/L Phytagel. 2B: 1/2 MS basal medium supplemented with 2 mg/L BA. HB1: 3 g/L Hyponex No. 1 (N:P:K = 7:6:19), 2 g/L tryptone, 20 g/L sucrose, and 1 g/L activated charcoal. **(a)** Transfected mesophyll protoplasts. Bar = 200 μm. **(b)** Microcallus formation after 1 month of incubation in 1N0.3K. Bar = 1 cm. **(c)** Microcallus amplification. **(d)** Microcalli transferred to 1B medium and incubated in the light. Bar = 1 cm. **(e)** Callus amplification in 2B. Bar = 1 cm. **(f)** Shoot formation in 2B. Bar = 1 cm. **(g)** Shoots incubated in HB1 for root formation. **(h)** *In vitro* seed formation. Bar = 5 mm. **(i)** Plantlet incubated in a growth chamber. Bar = 1 cm.

To look for unexpected phenotypes in the regenerants, we randomly selected three regenerated plants (protoplasts #1, #2, and #3) and harvested their seeds. We grew the progeny in a

growth chamber for 40 days and compared them to seedlings derived from seed propagation (seed #1, #2, and #3). We observed no significant differences in plant height between the regenerants and seed-derived plants (**Figure 2**). All plants flowered and produced seeds normally.

CRISPR Efficiency

To demonstrate that this protocol can be used for CRISPR-mediated gene editing, we performed protoplast transfections using plasmids previously shown to be effective in *N. tabacum* (Hsu et al., 2019). We successfully used SaCas9, FnCas12a, and nCas9-Target-AID to obtain regenerated plants for target gene editing using this protocol. The efficiency of *N. benthamiana* transformation was similar to that of *N. tabacum* (**Figure 3**, Hsu et al., 2019). As in *N. tabacum*, three different Cas proteins were successfully used to edit different target genes simultaneously in a single *N. benthamiana* protoplast. The target site of SaCas9 has a mismatch in *NbPDS-2*, and there was 10.0% off-target editing of *NbPDS-2* (**Figure 3**). In nCas9-Target-AID, we only observed mutation and no C to T editing regenerant was obtained. We designed *NbSGS3* and *NbRDR6* sgRNAs that can be used in both *N. tabacum* and *N. benthamiana* and introduced them into the SpCas9 plasmid. The regeneration results indicated that, except for sgRNA 2 in *RDR6*, these sgRNAs had target mutagenesis efficiencies in *N. benthamiana* (**Figure 3**) and *N. tabacum* (data not shown).

Phenotypes of Regenerants Following Targeted Editing

For the *PDS* gene study, we used SaCas9 and FnCas12a for targeted mutagenesis. When we used SaCas9 alone, since there was a mismatch in *NbPDS-2* (Niben101Scf14708Ctg003), we obtained no *NbPDS-1* and *NbPDS-2* double knock out mutants, and found *nbpds-1* null mutants with wild-type or heterozygous *NbPDS-2*. These mutants did not appear albino, the usual phenotype for this mutant, because *NbPDS-2* was still functional, unlike our previous findings with *N. tabacum* in which double knock out mutants were obtained using the same plasmid (Lin et al., 2018). Although the mutagenesis efficiency was low, because the target sequences of *NbPDS-1* and *NbPDS-2* are identical, we still obtained albino *nbpds-1 nbpds-1/nbpds-2 nbpds-2* double null mutants in regenerants derived from FnCas12a-mediated transfection. We also obtained *NbPDS-1 nbpds-1/nbpds-2 nbpds-2* heterozygous T_0 plants when the protoplasts were transfected with three plasmids (SaCas9, FnCas12a, and Target-AID). Albino mutants were detected in the T_1 offspring, and their proportions and genotypes were as expected.

The homozygous *nbrdr6-1 nbrdr6-2* double mutant derived from protoplast regeneration (*nbrdr6*#C13) was sterile, as are genome-edited mutants obtained via Agrobacterium-mediated transformation (Ludman and Fátyol, 2019; Matsuo and Atsumi, 2019), because they fail to produce seeds. Interestingly, two *N. benthamiana sgs3-1 sgs3-2* biallelic mutants (*nbsgs3*-14 and *nbsgs3*-16) both produced seeds. We identified four editing "scars" in *nbsgs3*-14 (**Figure 4A**): a 1-bp substitution (E) and a 1-bp insertion (a) in *NbSGS3-1*, and a 5-bp deletion (d) and a 1-bp insertion in *NbSGS3-2* (**Figure 4B**). Progeny with the EE/aa genotype could produce seeds, but dd/aa plants bore abnormal flowers that failed to produce fertile seeds (**Figure 4C**), as did the *nbsgs3-14* aa/at progeny. All *nbsgs3-14* progeny, regardless of their genotype, exhibited lower expression of the trans-acting secondary siRNA TAS3 than the wild type (**Figure 4D**). In sterile progeny (dd/aa in *nbsgs3-14*, aa/at in *nbsgs3-16*), no TAS3 siRNA was detected. These results indicate that RNA silencing was aberrant in the *nbsgs3* mutants. *RDR6* and *SGS3* function in RNA silencing by reducing foreign gene expression. Similar to the enhanced transgene expression observed in Agrobacterium-infiltrated *N. benthamiana rdr6*

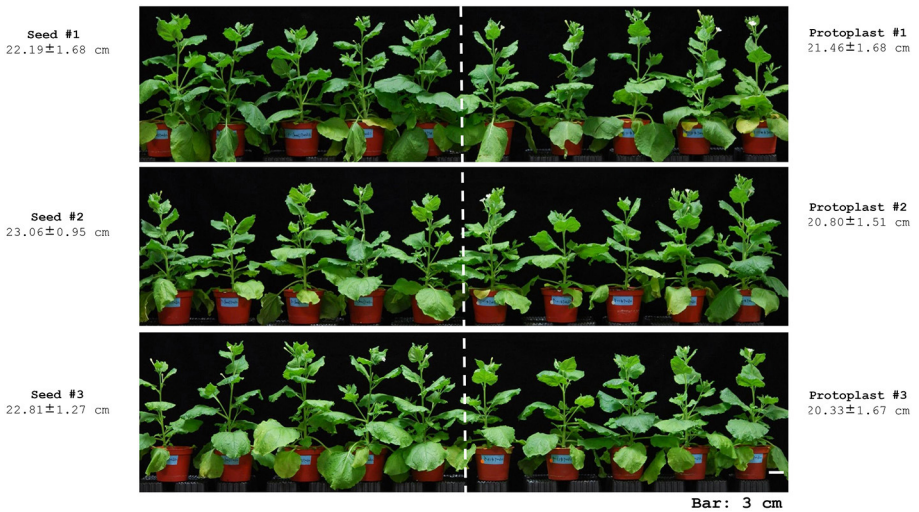
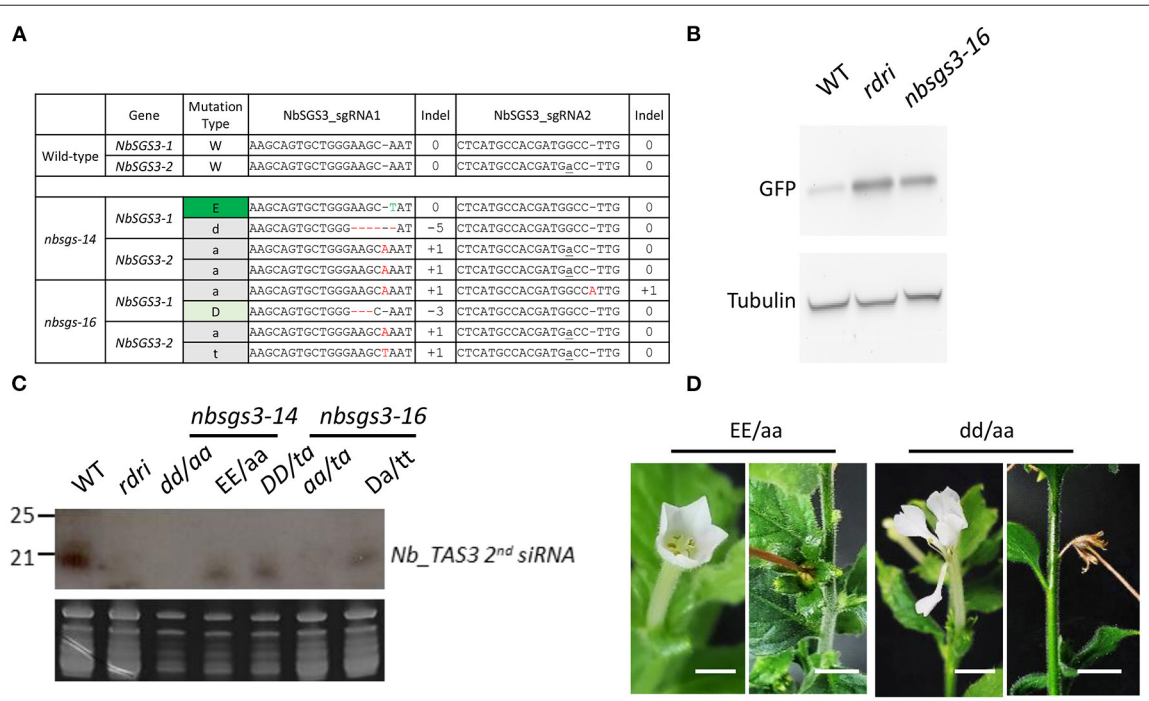


FIGURE 2 | Phenotypes of seed-derived seedlings and seedlings regenerated from protoplasts.

Cas	Gene		Accession no.	sgRNA		Efficiency	
						Mu./Total	%
FnCas12a	*PDS*	*NbPDS-1*	Niben101Scf01283Ctg022	1	TCATCCAGTCCTTAACACTTAAA	3/10	30.0
				2	ACATGGCAATGAACACCTCATCTG	0/10	0.0
		NbPDS-2	Niben101Scf14708Ctg003	1	TCATCCAGTCCTTAACACTTAAA	2/10	20.0
				2	ACATGGCAATGAACACCTCATCTG	0/10	0.0
SaCas9	*PDS*	*NbPDS-1*	Niben101Scf01283Ctg022	1	TTGCGATGCCTAACAAGCCAG	7/10	70.0
		NbPDS-2	Niben101Scf14708Ctg003		TTGCGATGCCTAACAAGCC G	1/10	10.0
nCas9-Target-AID	*ETR1*	*NbETR1-1*	Niben101Scf11767Ctg031	1	TGCACAAGAACCCATCTATA	4/10*	40.0
		NbETR1-2	Niben101Scf02971Ctg007		TGCACAAGAACCCATCTATA	0/10	0.0
SpCas9	*RDR6*	*NbRDR6-1*	Niben101Scf12609Ctg016	1	GCCATGGCCTTCTCAAAGCT	0/10	0.0
				2	GCAGTTCTATAGAAAACCAA	3/10	30.0
		NbRDR6-2	Niben101Scf03832Ctg041	1	GCCATGGCCTTCTCAAAGCT	0/10	0.0
				2	GCAGTTCTATAGAAAACCAA	2/10	20.0
	SGS3	*NbSGS3-1*	Niben101Scf03392Ctg069	1	AAGCAGTGCTGGGAAGCAAT	7/10	70.0
				2	CTCATGCCACGATGGCCTTG	3/10	30.0
		NbSGS3-2	Niben101Scf05468Ctg070	1	AAGCAGTGCTGGGAAGCAAT	6/10	60.0
				2	CTCATGCCACGATGGCCTTG	0/10	0.0

FIGURE 3 | Target mutagenesis efficiencies. Mu, mutants. Gray: sgRNA mismatch. *: mutation. There were 10 regenerated plants (Total) analysis in each transfection. The mutants were confirmed by Sanger sequencing. % = No. of mutants/total no. of regenerated plants analyzed) X 100.



	Gene	Mutation Type	NbSGS3_sgRNA1	Indel	NbSGS3_sgRNA2	Indel
Wild-type	*NbSGS3-1*	W	AAGCAGTGCTGGGAAGC-AAT	0	CTCATGCCACGATGGCC-TTG	0
	NbSGS3-2	W	AAGCAGTGCTGGGAAGC-AAT	0	CTCATGCCACGATGaCC-TTG	0
nbsgs-14	*NbSGS3-1*	E	AAGCAGTGCTGGGAAGC-TAT	0	CTCATGCCACGATGGCC-TTG	0
		d	AAGCAGTGCTGGG------AT	-5	CTCATGCCACGATGGCC-TTG	0
	NbSGS3-2	a	AAGCAGTGCTGGGAAGCAAAT	+1	CTCATGCCACGATGaCC-TTG	0
		a	AAGCAGTGCTGGGAAGCAAAT	+1	CTCATGCCACGATGaCC-TTG	0
nbsgs-16	*NbSGS3-1*	a	AAGCAGTGCTGGGAAGCAAAT	+1	CTCATGCCACGATGGCCATTG	+1
		D	AAGCAGTGCTGGG---C-AAT	-3	CTCATGCCACGATGGCC-TTG	0
	NbSGS3-2	a	AAGCAGTGCTGGGAAGCAAAT	+1	CTCATGCCACGATGaCC-TTG	0
		t	AAGCAGTGCTGGGAAGCTAAT	+1	CTCATGCCACGATGaCC-TTG	0

FIGURE 4 | Mutants regenerated from CRISPR-Cas-edited *Nicotiana benthamiana* protoplasts. **(A)** Genotypes of F_1 *N. benthamiana sgs3* mutants used for small RNA analysis. *NbSGS3-1*: Niben101Scf03392Ctg069; *NbSGS3-2*: Niben101Scf05468Ctg070. Red -: deleted nucleotide. Letter in gray: inserted nucleotide. Letter in green: edited nucleotide. **(B)** Five-week-old *N. benthamiana* plants subjected to Agrobacterium-mediated transient infiltration with 1 OD Agrobacterium cultures (OD_{600} = 1) harboring binary vector with *Green fluorescence protein* (*GFP*) driven by the cauliflower mosaic virus *35S* promoter. Leaves were harvested 3 d after infiltration. GFP and tubulin levels were analyzed by immunoblot analysis. **(C)** RNA gel blot analysis of the progeny of the *Nbsgs* mutants. WT, wild type; *rdri*, RNAi line of *NbRDR6*; *D*, 3-bp deletion. T insertion. **(D)** The progeny of *Nbsgs3-14*. Uppercase letters: in frame; lowercase letters: out of frame. *E*: base editing. *a*: A insertion. *d*: 5-bp deletion. Bar = 5 mm.

mutants (Ludman and Fátyol, 2019; Matsuo and Atsumi, 2019), GFP accumulated to higher levels in *nbsgs3-16* than in wild type (**Figure 4B**).

In vitro Flowering

We incubated *N. benthamiana* plants regenerated from protoplasts in the same medium used for *in vitro* flowering of the orchid *Erycina pusilla* (Chiu et al., 2011). The *N. benthamiana* regenerants were able to flower (**Figure 5a**) and produce seeds (**Figure 5b**) *in vitro*. The seeds matured normally (**Figure 5c**) and germinated (**Figure 5d**). To investigate whether this medium can be widely used, we also incubated plants regenerated from protoplasts of other species in the same medium, including protoplasts from tobacco, broccoli, cauliflower, Arabidopsis, and rapid cycling *Brassica oleracea*. The tobacco, broccoli, and cauliflower plants did not flower *in vitro*, whereas the Arabidopsis and rapid cycling *Brassica oleracea* plants flowered but failed to produce seeds.

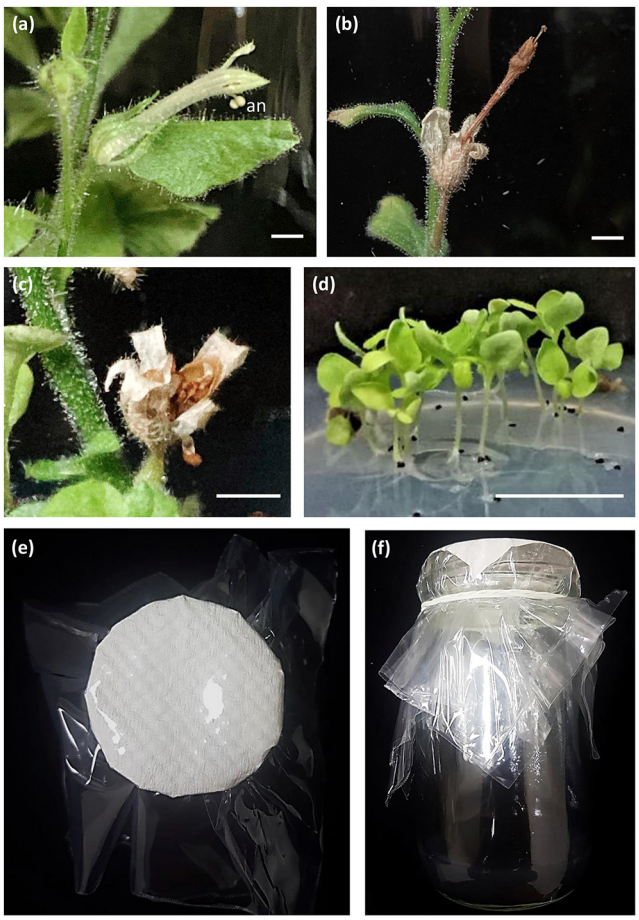


FIGURE 5 | *In vitro* flowering and seed production in *N. benthamiana*. **(a)** *In vitro* flower. an: anther. Bar = 5 mm. **(b)** Mature fruit. Bar = 0.5 mm. **(c)** Open capsule. Bar = 0.5 mm. **(d)** Seedlings derived from *in vitro* seeds. Bar = 1 cm. **(e)** Bottle lip sealed with our newly developed sealing material. Bar = 1 cm. **(f)** Subculture in HB1 medium. Bar = 1 cm.

To increase seed production, we attempted to reduce the humidity inside the flask. For this purpose, we designed a new sealing material. A small piece of paper larger than the diameter of the bottle lip was placed inside two pieces of plastic film (**Figure 5e**), sealed over the lip, tied with a rubber band (**Figure 5f**), and sterilized by autoclaving. *N. benthamiana* produced seeds normally under these conditions, whereas the other species were still unable to produce seeds.

DISCUSSION

During this study, we determined that protoplasts derived from *N. tabacum* are easier to regenerate than protoplasts from *N. benthamiana*. The bottleneck in regenerating *N. benthamiana* plants occurs during the step from callus to shoot formation: whereas most calluses derived from *N. tabacum* can be regenerated into shoots after a single subculture, *N. benthamiana* requires two or three successive subcultures before shoots form, and the proportion of shoots that form in each subculture step is unpredictable. Nonetheless, *N. benthamiana* still has many advantages. Compared to *N. tabacum*, *N. benthamiana* requires less space, and it can flower all-year round, whereas in subtropical regions, *N. tabacum* cannot grow and produce flowers in the greenhouse in summer.

Protoplast regeneration has been used since 2016 for transcription activator-like effector nucleases (TALEN)-mediated gene editing (Li et al., 2016). Although this strategy has many advantages with regard to gene editing procedures, it is often avoided. Instead, protoplasts have mainly been used for protoplast fusion and mutation. Furthermore, protoplast regeneration is thought to result in many unanticipated mutations. Indeed, a study involving whole-genome sequencing of potato plants regenerated from protoplasts suggested that protoplast regeneration can cause numerous mutations and even chromosome rearrangements (Fossi et al., 2019). Because the tobacco genome is so large, we have not yet sequenced the entire genomes of gene-edited *N. benthamiana* regenerants, and thus we have not directly investigated their levels of mutations. However, our comparison of regenerated plants with seed-propagated offspring indicated that this protocol does not produce plants with phenotypic differences from the wild type.

In fact, unexpected mutations can occur in any tissue culture process and even under natural conditions (Lin and Chang, 1998; Yue et al., 2020). In crop breeding, even if mutations occur, desired offspring can be identified through selection from a wide range of gene-edited regenerated plants, without the ethical problems associated with the human application of genome editing (Tang et al., 2019). In plant research, the problem of unexpected mutations could be resolved by generating multiple mutations of the same gene, such as in Arabidopsis and rice knockout lines, or by transferring the edited gene to a wild-type plant by crossing.

Given concerns about the use of genetically modified crops, it is important to be able to produce genetically-edited crops without introducing foreign genes. In particular, although

transgenes introduced via stable transformation can be removed from many plants through crossing, this is not the case for the many important crops that are propagated asexually, such as potato. The delivery of CRISPR reagents into cells by transient transfection, however, is widely regarded as transgene-free gene editing. Using protoplast regeneration, RNPs (Woo et al., 2015) or plasmids (Lin et al., 2018; Hsu et al., 2019) can be used as CRISPR reagents for transgene-free gene editing; this is the main reason that we use protoplast regeneration for gene editing of crops. In our experience, both RNPs and plasmids are effective for gene editing. When choosing gene-editing reagents, if no documented gene target sites are available to confirm editing efficiency, we use plasmids, which allow us to use multiple targets at once and are relatively cost-effective. For target insertion, we strongly recommend using RNPs, as plasmids may act as donor DNAs.

In addition, RNPs can be used to validate novel Cas proteins when the expected results are not obtained using plasmids. Since our Cas protein was translated and RNP-confirmed *in vitro*, we directly introduced these RNPs into protoplasts to validate that this protein functions in the species of interest. We used RNPs to monitor the efficiency of Cas12a proteins in *N. benthamiana*. Cas12a has a high target mutagenesis efficiency in Poaceae (Li et al., 2019). In dicots, LbCas12a has a higher editing efficiency than AsCas12a in soybean and tobacco protoplasts when delivered as RNP molecules (Kim et al., 2017). Similarly, when tested in rice, Arabidopsis, and maize, LbCas12a but not AsCas12a successfully edited target genes when these nuclease plasmids were delivered into protoplasts (Kim et al., 2017; Malzahn et al., 2019).

More importantly, the efficiency of Cas12a proteins is temperature dependent (Malzahn et al., 2019). These enzymes have high activity at 37°C, the temperature used for human cell culture, whereas plant transformation is performed at ∼28°C, a temperature at which LbCas12a activity is reduced (Moreno-Mateos et al., 2017; Malzahn et al., 2019). Hence, it is likely that the absence of edited plants was due to the lower temperature along with the lower overall activity of AsCas12a. In both *N. tabacum* and *N. benthamiana*, the efficiencies of FnCas12a when using plasmids are <10%. When FnCas12a and AsCas12a RNPs (which were confirmed to have cleavage activity *in vitro*) were used, target mutagenesis did not occur, even when we increased the amount of RNP, raised the temperature, or changed the medium composition. Therefore, we suggest that the low target mutagenesis efficiency of Cas12a is due not to the low expression of this protein but to intracellular conditions unsuitable for its activity.

Because transient transfection can deliver multiple plasmids into the same protoplast at one time, Cas protein and sgRNAs do not necessarily need to be encoded by the same vector. The Agrobacterium transformation vector we used in the current study is a low copy number vector, which makes plasmid DNA extraction more difficult. Because the DNA does not need to be inserted into the chromosome, there is no need to clone these genes into the T-DNA vector. In addition, protoplast transfection is highly efficient and does not require a selectable marker for screening. Thus, to refine our method, we could simplify the vector structure and use a high copy number vector for CRISPR-mediated protoplast transfection. Alternatively, we could co-transfect *in vitro* transcribed sgRNA with the overexpression Cas protein vector to reduce the labor involved in construction (Zhang et al., 2016). When designing sgRNAs, we will not only use software to predict the efficiencies of the sgRNAs, but also select the relative positions of sgRNAs that have been successful for other species. For example, sgRNAs that were designed by this strategy based on the *RDR6* and *SGS3* sgRNAs that are effective in *N. benthamiana* were also effective in *Solanum peruvianum*. If there are many effective sgRNA candidates, we will choose the one that can be used in the largest number of species to increase usage. *ETR1* sgRNA, which we used successfully in *N. benthamiana* and *N. tabacum*, is derived from tomato (Shimatani et al., 2017). Applying such a principle for sgRNA design to simultaneously induce mutations in multiple genes or gene families in heterozygous or polyploid plants is difficult because of mismatches. Because this is not a problem for off-target crops, it can instead be exploited for this type of multiplex gene mutation (Endo et al., 2015).

Most explants grow to the vegetative stage *in vitro*. By manipulating the medium and culture conditions, however, many plants can also be induced to flower *in vitro*. For example, bamboo has a juvenile period of several decades in the natural environment but can flower within 1 year in medium containing cytokinin (Lin and Chang, 1998). Plant species that can flower and be successfully pollinated, form fruit, and complete all growth stages *in vitro* are potentially good model plants for further study. For example, *E. pusilla*, which has these capabilities, serves as a model plant of the orchid family (Chiu et al., 2011). Here we demonstrated that *N. benthamiana* can bear fruit and produce seeds *in vitro* and that the use of HB1 medium and sealing film that we developed can increase fruiting and seed production. Although a speed breeding method has been developed to accelerate a plant's growth cycle and achieve year-round production (Ghosh et al., 2018), this method is quite expensive. It is important to develop an economical, space-saving method that can be used by all laboratories. The *in vitro* method developed in the study represents an alternative strategy for achieving this goal. However, this method cannot be used for all crops, an issue that will need to be addressed.

CONCLUSIONS

Although various protocols have been published for Agrobacterium-mediated stable transformation or DNA-free plant genome editing of *N. benthamiana* using virally delivered CRISPR-Cas (Ma et al., 2020), these techniques pose several problems, including issues related to the regulation of transgenic crops and the production of genetic chimeras. Protoplast regeneration represents an alternative approach for high-efficiency gene editing that avoids these complications. With this method, no foreign DNA is integrated into the chromosomes, the regenerated plants are derived from single cells, and all of the edited alleles can be passed on to the offspring. We also used this procedure to transfer large amounts of donor

DNA to increase the efficiency of target DNA insertion. We believe that this system and the resulting mutants represent excellent tools for researchers using *N. benthamiana* for crop pathogen-related research.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

C-SL conceived and designed the experiments. C-TH, and Y-HY performed the CRISPR-Cas9 experiments. C-TH, Y-HY, and C-SL conducted the protoplast regeneration. C-TH, W-CL, Y-HY, and F-HW performed the molecular biology experiments and targeted mutagenesis analysis. W-CL performed small RNA Northern analysis and Western analysis. C-SL wrote the manuscript with input from all co-authors. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We thank Qiao-Wei Cheng for help with tissue culture and protoplast transfection, Prof. Ho-Ming Chen for small RNA analysis, and Profs. Jen Sheen and Ming-Che Shih for discussion. This research was supported by Academia Sinica Innovative Translational Agricultural Research Administrative Office (AS-KPQ-107-ITAR-10; AS-KPQ-108-ITAR-10; AS-KPQ-109-ITAR-10), and the Ministry of Science and Technology (105-2313-B-001-007-MY3; 108-2313-B-001—011-; 109-2313-B-001—011-), Taiwan.

REFERENCES

Andersson, M., Turesson, H., Olsson, N., Fält, A. S., Ohlsson, P., Gonzalez, M. N., et al. (2018). Genome editing in potato via CRISPR-Cas9 ribonucleoprotein delivery. *Physiol. Plant*. 164, 378–384. doi: 10.1111/ppl.12731

Bombarely, A., Rosli, H. G., Vrebalov, J., Moffett, P., Mueller, L. A., and Martin, G. B. (2012). A draft genome sequence of Nicotiana benthamiana to enhance molecular plant microbe biology research. *Mol. Plant Microbe Interact.* 25, 1523–1530. doi: 10.1094/MPMI-06-12-0148-TA

Chiu, Y. T., Lin, C. S., and Chang, C. (2011). *In vitro* fruiting and seed production in *Erycina pusilla* (L.) NH Williams and MW Chase. *Propag. Ornam. Plants* 11, 131–136.

Derevnina, L., Kamoun, S., and Wu, C. h. (2019). Dude, where is my mutant? *Nicotiana benthamiana* meets forward genetics. *New Phytol.* 221, 607–610. doi: 10.1111/nph.15521

Endo, A., Masafumi, M., Kaya, H., and Toki, S. (2016). Efficient targeted mutagenesis of rice and tobacco genomes using Cpf1 from *Francisella novicida*. *Sci. Rep*. 6:38169. doi: 10.1038/srep38169

Endo, M., Mikami, M., and Toki, S. (2015). Multi-gene knockout utilizing off-target mutations of the CRISPR/Cas9 system in rice. *Plant Cell Physiol*. 56,41–47. doi: 10.1093/pcp/pcu154

Fossi, M., Amundson, K., Kuppu, S., Britt, A., and Comai, L. (2019). Regeneration of Solanum tuberosum plants from protoplasts induces widespread genome instability. *Plant Physiol*. 180, 78–86. doi: 10.1104/pp.18.00906

Ghosh, S., Watson, A., Gonzalez-Navarro, O. E., Ramirez-Gonzalez, R. H., Yanes, L., Mendoza-Suárez, M., et al. (2018). Speed breeding in growth chambers and glasshouses for crop breeding and model plant research. *Nat. Protoc*. 13, 2944–2963. doi: 10.1038/s41596-018-0072-z

Goodin, M. M., Zaitlin, D., Naidu, R. A., and Lommel, S. A. (2008). *Nicotiana benthamiana*: Its history and future as a model for plant-pathogen interactions. *Mol. Plant Microbe Interact*. 21: 1015–1026. doi: 10.1094/MPMI-21-8-1015

Hill, J. T., Demarest, B. L., Bisgrove, B. W., Su, Y. C., Smith, M., and Yost, H. J. (2014). Poly peak parser: method and software for identification of unknown indels using Sanger sequencing of PCR products. *Dev. Dyn*. 243, 1632–1636. doi: 10.1002/dvdy.24183

Hsu, C. T., Cheng, Y. J., Yuan, Y. H., Hung, W. F., Cheng, Q. W., Wu, F. H., et al. (2019). Application of Cas12a and nCas9-activation-induced cytidine deaminase for genome editing and as a non-sexual strategy to generate homozygous/multiplex edited plants in the allotetraploid genome of tobacco. *Plant Mol. Biol*. 101, 355–371. doi: 10.1007/s11103-019-00907-w

Kaya, H., Mikami, M., Endo, A., Endo, M., and Toki, S. (2016). Highly specific targeted mutagenesis in plants using *Staphylococcus aureus* Cas9. *Sci. Rep*. 6:26871. doi: 10.1038/srep26871

Kim, H., Kim, S., Ryu, J., Kang, B. C., Kim, J. S., and Kim, S. G. (2017). CRISPR/Cpf1-mediated DNA-free plant genome editing. *Nat. Commun*. 8:14406. doi: 10.1038/ncomms14406

Li, J., Stoddard, T. J., Demorest, Z. L., Lavoie, P. O., Luo, S., Clasen, B. M., et al. (2016). Multiplexed, targeted gene editing in *N. benthamiana* for glyco-engineering and monoclonal antibody production. *Plant Biotechnol. J*. 14, 533–542. doi: 10.1111/pbi.12403

Li, J. F., Norville, J. E., Aach, J., McCormack, M., Zhang, D. D., Bush, J., et al. (2013). Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol*. 31, 688–691. doi: 10.1038/nbt.2654

Li, S., and Xia, L. (2020). Precise gene replacement in plants through CRISPR/Cas genome editing technology: current status and future perspectives. *aBIOTECH* 1, 58–73. doi: 10.1007/s42994-019-00009-7

Li, S., Zhang, Y., Xia, L., and Qi, Y. (2019). CRISPR-Cas12a enables efficient biallelic gene targeting in rice. *Plant Biotechnol. J*. 18, 1351–1353. doi: 10.1111/pbi.13295

Lin, C.-S., Hsu, C.-T., Yang, L.-H., Lee, L.-Y., Fu, J.-Y., Cheng, Q.-W., et al. (2018). Application of protoplast technology to CRISPR/Cas9 mutagenesis: from single-cell mutation detection to mutant plant regeneration. *Plant Biotechnol. J*. 16,1295–1310. doi: 10.1111/pbi.12870

Lin, C. S., and Chang, W. C. (1998). Micropropagation of *Bambusa edulis* through nodal explants of field-grown culms and flowering of regenerated plantlets. *Plant Cell Rep*. 17, 617–620. doi: 10.1007/s002990050453

Ludman, M., and F?tyol, K. (2019). The virological model plant, Nicotiana benthamiana expresses a single functional RDR6 homeolog. *Virology*. 537, 143–148. doi: 10.1016/j.virol.2019.08.017

Ma, X., Zhang, Q., Zhu, Q., Liu, W., Chen, Y., Qiu, R., 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

Ma, X., Zhang, X., Liu, H., and Li, Z. (2020). Highly efficient DNA-free plant genome editing using virally delivered CRISPR–Cas9. *Nat. Plants*, 6, 773–779. doi: 10.1038/s41477-020-0704-5

Malzahn, A. A., Tang, X., Lee, K., Ren, Q., Sretenovic, S., Zhang, Y., et al. (2019). Application of CRISPR-Cas12a temperature sensitivity for improved genome editing in rice, maize, and *Arabidopsis*. *BMC Biol*. 17:9. doi: 10.1186/s12915-019-0629-5

Matsuo, K., and Atsumi, G. (2019). CRISPR/Cas9-mediated knockout of the RDR6 gene in Nicotiana benthamiana for efficient transient expression of recombinant proteins. *Planta*. 250, 463–473. doi: 10.1007/s00425-019-03180-9

Moreno-Mateos, M. A., Fernandez, J. P., Rouet, R., Vejnar, C. E., Lane, M. A., Mis, E., et al. (2017). CRISPR-Cpf1 mediates efficient homology-directed repair and temperature-controlled genome editing. *Nat. Commun*. 8:2024. doi: 10.1038/s41467-017-01836-2

Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J. D. G., and Kamoun, S. (2013). Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nat. Biotechnol*. 31, 691–693. doi: 10.1038/nbt.2655

Shan, Q. W., Wang, Y. P., Li, J., Zhang, Y., Chen, K. L., Liang, Z., et al. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol*. 31, 686–688. doi: 10.1038/nbt.2650

Shimatani, Z., Kashojiya, S., Takayama, M., Terada, R., Arazoe, T., Ishii, H., et al. (2017). Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. *Nat. Biotechnol*. 35,441–443. doi: 10.1038/nbt.3833

Tang, J., Chen, L., and Liu, Y. G. (2019). Off-target effects and the solution. *Nat. Plants* 5:341–342. doi: 10.1038/s41477-019-0406-z

Woo, J. W., Kim, J., Kwon, S. I., Corvalan, C., Cho, S. W., Kim, H., et al. (2015). DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat. Biotechnol*. 33, 1162–1164. doi: 10.1038/nbt.3389

Yue, J. J., Hong, C. Y., Wei, P., Tsai, Y. C., and Lin, C. S. (2020). How to start your monocot CRISPR/Cas project: plasmid design, efficiency detection, and offspring analysis. *Rice*. 13:9. doi: 10.1186/s12284-019-0354-2

Zhang, Y., Liang, Z., Zong, Y., Wang, Y., Liu, J., Chen, K., et al. (2016). Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA. *Nat. Commun*. 7:12617. doi: 10.1038/ncomms12617

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Hsu, Lee, Cheng, Yuan, Wu and Lin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genome Editing

ORIGINAL RESEARCH
published: 25 February 2021
doi: 10.3389/fgeed.2021.632136

Protein Phosphatase 2A Catalytic Subunit PP2A-1 Enhances Rice Resistance to Sheath Blight Disease

Qiu Jun Lin[1†], Jin Chu[2†], Vikranth Kumar[3†], De Peng Yuan[1], Zhi Min Li[1], Qiong Mei[1] and Yuan Hu Xuan[1*]*

[1] College of Plant Protection, Shenyang Agricultural University, Shenyang, China, [2] Institute of Plant Protection, Liaoning Academy of Agricultural Sciences, Shenyang, China, [3] Division of Applied Life Science (BK21 Program), Plant Molecular Biology and Biotechnology Research Center (PMBBRC), Gyeongsang National University, Jinju, South Korea

OPEN ACCESS

Edited by:
Huanbin Zhou,
Chinese Academy of Agricultural Sciences, China

Reviewed by:
Chuanxiao Xie,
Chinese Academy of Agricultural Sciences, China
Xuli Wang,
Chinese Academy of Agricultural Sciences, China

***Correspondence:**
Qiong Mei
meiqiong@syau.edu.cn
Yuan Hu Xuan
xuanyuanhu115@syau.edu.cn

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Genome Editing in Plants, a section of the journal Frontiers in Genome Editing

Received: 22 November 2020
Accepted: 08 February 2021
Published: 25 February 2021

Citation:
Lin QJ, Chu J, Kumar V, Yuan DP, Li ZM, Mei Q and Xuan YH (2021) Protein Phosphatase 2A Catalytic Subunit PP2A-1 Enhances Rice Resistance to Sheath Blight Disease. Front. Genome Ed. 3:632136. doi: 10.3389/fgeed.2021.632136

Rice (*Oryza sativa*) production is damaged to a great extent by sheath blight disease (ShB). However, the defense mechanism in rice against this disease is largely unknown. Previous transcriptome analysis identified a significantly induced eukaryotic *protein phosphatase 2A catalytic subunit 1* (*PP2A-1*) after the inoculation of *Rhizoctonia solani*. Five genes encoding PP2A exist in rice genome, and these five genes are ubiquitously expressed in different tissues and stages. Inoculation of *R. solani* showed that the genome edited *pp2a-1* mutants using the CRISPR/Cas9 were more susceptible to ShB than the wild-type control, but other *PP2A* gene mutants exhibited similar response to ShB compared to wild-type plants. In parallel, *PP2A-1* expression level was higher in the activation tagging line, and *PP2A-1* overexpression inhibited plant height and promoted the resistance to ShB. PP2A-1-GFP was localized in the cytoplasm and nucleus. In addition, *R. solani*-dependent induction kinetics of pathogen-related genes *PBZ1* and *PR1b* was lower in *pp2a-1* mutants but higher in *PP2A-1* activation line compared to those in the wild-type. In conclusion, our analysis shows that PP2A-1 is a member of protein phosphatase, which regulates rice resistance to ShB. This result broadens the understanding of the defense mechanism against ShB and provides a potential target for rice breeding for disease resistance.

Keywords: PP2A-1, sheath blight, resistance, enhance, rice

INTRODUCTION

R. solani is the causative agent of ShB in rice (Savary et al., 1995; Suryadi et al., 2013), and which damages rice during the entire growth period, and predominantly targets the leaves, sheaths, and panicles, eventually resulting in the withering and lodging of the entire plant. A severe form of ShB can lower the rice produce by ∼50% (Savary et al., 2000). The rapid variation, wide host range, and high survival ability of the pathogen can make the disease control more challenging (Taheri and Tarighi, 2011; Yellareddygari et al., 2014; Singh et al., 2019). Currently, there is a dearth of ShB-resistant cultivars, therefore, the strategies to quell ShB involve the use of fungicides (Savary et al., 2000). However, fungicides directly affect the living environment of other microorganisms and increase the cost of cultivation. Thus, to develop ShB resistance in rice, it is necessary to isolate resistant cultivars and understand their underlying defense mechanisms against ShB.

Extensive studies have been performed to investigate the mechanism of rice defense against ShB. Overexpression of chitinase, β-1,3-glucanase, or OsPGIP1 (polygalacturonase-inhibiting protein) (Shah et al., 2009; Mao et al., 2014; Zhu et al., 2019), OsACS2 (key enzyme in ethylene synthesis) (Helliwell et al., 2013), OsGSTU5 (tau class glutathione-*S*-transferase 5) (Tiwari et al., 2020), and Os2H16 (Li et al., 2013, 2018) were found to promote rice resistance to ShB. In addition, *BSR2* (*broad-spectrum resistance 2*) (Maeda et al., 2019) or a transcription factor complex including LPA1 (indeterminate domain 14, IDD14) and IDD13 (Sun et al., 2019, 2020) were reported to positively regulate rice resistance to ShB while *SWEET11* (*sugar will eventually be exported transporter 11*) (Gao et al., 2018) exhibited a negative regulation. The transcription factor OsWRKYs also plays an important role in resistance to sheath blight (Peng et al., 2012, 2016; Wang et al., 2015; Jimmy and Babu, 2019; Yuan et al., 2020). In addition, salicylic acid-dependent immunity showed a positive regulation in ShB resistance in rice and *Brachypodium distachyon* (Kouzai et al., 2018).

Protein phosphatase also plays an important role in plant defense response. The protein phosphatases (PPs) with a vast array of structures and functions are mainly categorized as serine/threonine (Ser/Thr) PPs and protein tyrosine phosphatases (PTPs). PP1, PP2A, PP2B, and PP2C account for the sub-divisions of the protein tyrosine phosphatase group. The PP2A complex comprises three subunits: A, B, and C with scaffolding, regulatory, and catalytic roles, respectively (Yu et al., 2005; Durian et al., 2016). The role of PP2A protein in plant abiotic stress signal transduction has been confirmed. For instance, drought and elevated salinity induce high levels of *OsPP2A-1* and *OsPP2A*-3, the closely associated genes coding for the C-subunit of PP2A (Yu et al., 2003). In *Arabidopsis*, the growth of roots and shoots is augmented by *PP2A-C5* overexpression in the presence of several salts indicating the vital function of protein in growth to combat salinity (Hu et al., 2017). *AtPP2A* is involved in acclimation to light as well as when responding to pathogens, both based on the regulation of ROS (Rahikainen et al., 2016; Máthé et al., 2019). Exposure of wheat to *R. cerealis* or hydrogen peroxide showed elevated *TaPP2Ac-4B* and *TaPP2Ac-4D* RNA levels revealing the involvement of PP2A in the biotic stress response. Silencing of *TaPP2A* in wheat boosted the expression of ROS-scavenging and pathogenesis-related (PR) RNA molecules (Zhu et al., 2018). Resistance to *Botrytis cinerea* and leaf senescence in *Arabidopsis* involves the role of PP2A-B$'\gamma$. The swift induction of the gene coding for the heterotrimeric PP2A catalytic subunit, *LePP2A-1* was observed when resistant tomato plants were challenged with *Pseudomonas syringae* pv. *tomato* (a virulent strain) (He et al., 2004). A mutation which was isolated from rice blast fungus was inserted into the promoter region of MoPPG1, a ser/thr-PP2A catalytic subunit (PP2Ac) gene, which made the mutant defective in the growth of vegetative mycelium and could not cause disease (Du et al., 2013). *Fusarium graminearum* contains three kinds of PP2A (FgPp2A, FgSit4, and FgPg1), which play a key role in the growth, development, and pathogenicity of fungi (Liu et al., 2018). Our recent transcriptomic study showed the sensitivity of *PP2A-1* expression to *R. solani* infection (Yuan et al., 2020). However, PP2A function in rice defense to ShB is unknown.

In this study, *PP2A-1* was significantly induced following *R. solani* inoculation. Further bioinformatics, genetic, and molecular analyses were performed to identify the function of PP2A family members in rice defense to ShB. Our results broaden the knowledge of the underlying ShB defense mechanisms and provide a potential target for resistant breeding in rice.

MATERIALS AND METHODS

Plant and Fungal Materials

Four rice lines/cultivars, including Japonica rice cultivar Dongjin (DJ), Zhonghua11 (ZH11), *pp2a-1* CRISPR/Cas-9 genome editing mutants in ZH11 background, and *PP2A-1* activation tagging line (*PP2A-1 OX*) in DJ background were used in this study. All the rice lines used in this study were grown in a greenhouse in natural light. The type strain used in this study was *R. solani* AG1-IA.

Construction of the CRISPR/Cas9 Plasmids

The human codon-optimized hSpCas9 (Cong et al., 2013) was linked to the maize ubiquitin promoter (UBI) in an intermediate plasmid followed by its insertion into a binary pCAMBIA1300 vector (Cambia, Australia) harboring the *HPT* (*hygromycin B phosphotransferase*) gene. A point mutation kit (Transgen, China) was used to eliminate the original *Bsa*I site in the backbone of pCAMBIA1300. A OsU6 promoter fragment (Feng et al., 2013), *ccdB*, a gene for negative selection flanked by two *Bsa*I sites, and a pX260- derived sgRNA (Cong et al., 2013) were inserted employing an In-Fusion cloning kit (Takara, Japan) into this vector to produce the CRISPR/Cas9 binary vector pBGK032 (**Figure 1**). The vector was maintained in *Escherichia coli* strain DB3.1.

The targeting specificity and the selection of the 23-bp targeting sequences (including PAM) was done employing a BLAST search (http://blast.ncbi.nlm.nih.gov/Blast.cgi) against the rice genome (Hsu et al., 2013). The designed targeting sequences were annealed to synthesize the oligo adaptors. The vector pBGK032 was restriction digested using *Bsa*I and purified employing a DNA purification kit (Tiangen, China). This was ligated with 0.05 mM of oligo adaptor (10 μL) resulting in CRISPR/Cas9 plasmids, which were directly transformed in competent *E. coli* cells.

Transformation of Plants and Mutation Detection

Agrobacterium tumefaciens strain EHA105 was transformed with the CRISPR/Cas9 plasmids followed by rice transformation following an earlier published protocol (Nishimura et al., 2006). The genomic DNA from these transformants was extracted, and PCR was conducted employing primer pairs flanking the designed target site. The Degenerate Sequence Decoding approach was applied to directly sequence and identified the PCR products (300–500 bp) method (Ma et al., 2015).

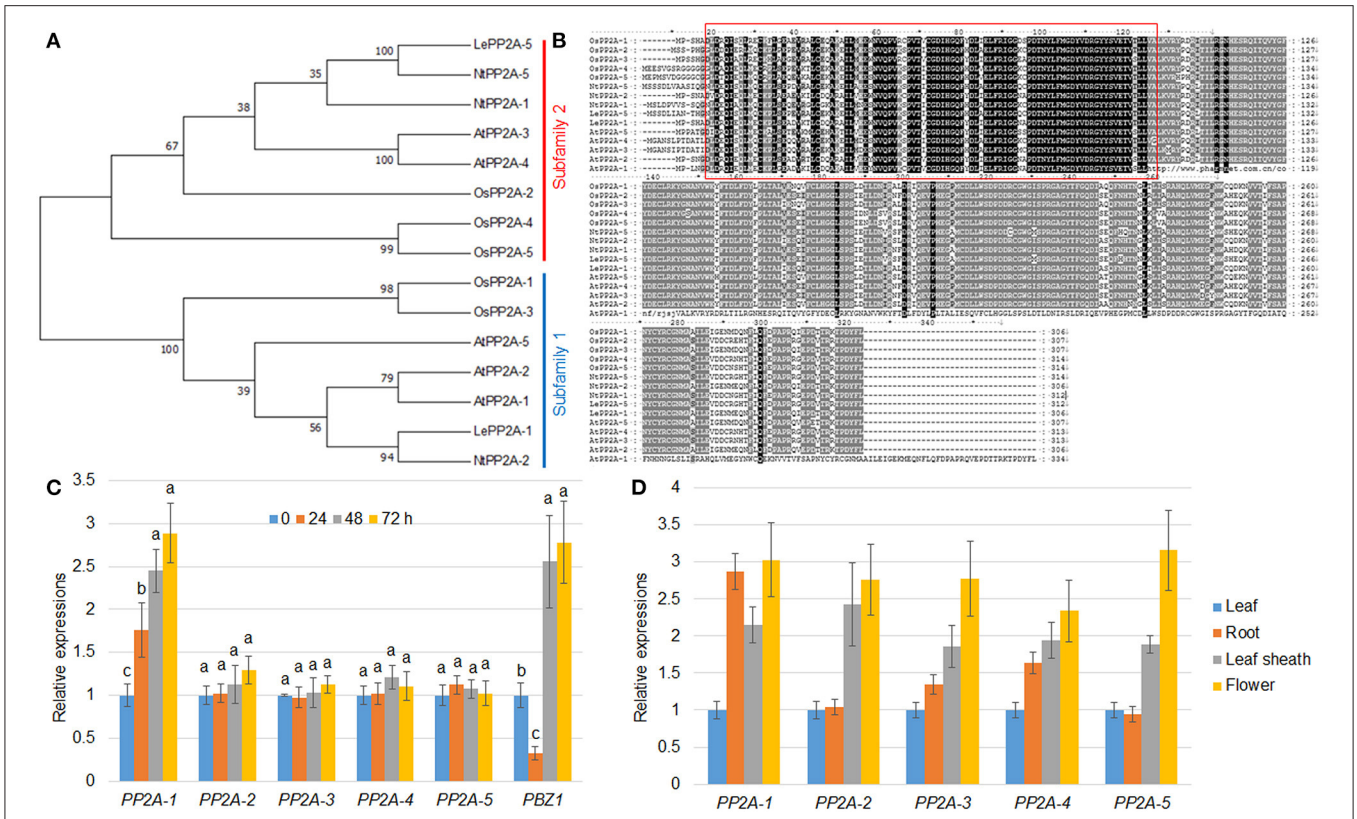


FIGURE 1 | Phylogenetic and expression pattern analysis. **(A)** The maximum likelihood was utilized for the phylogenetic tree construction. The phylogenetic tree was generated by MEGA 7.0 on the basis of PP2A amino acid sequences from *Oryza sativa* (OsPP2As), *Arabidopsis thaliana* (AtPP2As), *Solanum lycopersicum* (LePP2As), and *Nicotiana tabacum* (NtPP2As). The PP2As were classified into subfamily 1 and 2 that were marked with blue and red, respectively. **(B)** The protein sequence identity of the PP2As listed in **(A)** was analyzed by the Clustal W program. The red box indicates a highly conserved region between PP2As. **(C)** *R. solani* inoculation-mediated *PP2A* expression. The five *PP2A* and *PBZ1* gene expression patterns were evaluated after 0, 24, 48, and 72 h following *R. solani* inoculation. Normalization of expression was done with reference to the internal control *ubiquitin*. The experiments were done in triplicate. Different letters indicate significant differences at $P < 0.05$. **(D)** Tissue-specific expression patterns of five *PP2A* genes were analyzed using the RNA extracted from the leaf, root, leaf sheath, and flower. *Ubiquitin* was used as the internal control to normalize expression levels. The experiments were done in triplicate.

RNA Extractions

The total cellular RNA was extracted from the 1-month-old plant sheath, roots, leaves, or 3-month-old flower tissues. For analysis of *R. solani*-mediated gene expression, Trizol reagent (Invitrogen, China) was used to isolate the total RNA from 0.1 g of rice leaves, roots, leaf sheath, and flowers. Then the product was purified using the RNeasy mini kit (Promega, China) and RNase-Free DNase I (Promega, China) following manufacturer's instructions. The cDNAs were synthesized with M-MLV Reverse Transcriptase (Promega) kit following manufacturer's instructions.

Sequence Analysis of PP2A

The PP2A amino acid sequences were isolated from *Arabidopsis*, rice, tobacco, and tomato to perform phylogenetic analysis. MEGA7 software was used for multiple sequence alignment of the original sequences. The comparison results were edited by GeneDoc to export the multi-sequence alignment results graph. MEGA7 software was used for phylogenetic tree construction using the nearest neighbor-joining method (Kumar et al., 2016).

cDNA Synthesis and qRT–PCR

Reverse transcription using 2 μg of each purified RNA sample was done using a Prime Script TMRT Reagent Kit with gDNA Eraser (TaKaRa, China) in accordance with the provided instructions. qRT-PCR was performed on the ABI 7500 RT-PCR system (Applied Biosystems, United States). The composition of the mix was: 10 μL 2× SYBR Premix Ex Taq, 0.4 μL 50× ROX Reference Dye II, 0.4 μM of each primer, and 5 μL of the cDNA template (50-fold dilution) in a net volume of 20 μL. The conditions were: 95°C for 30 s; 95°C for 5 s, 58°C for 15 s, and 72°C for 34 s for 40 cycles. The $2^{[-\mathrm{DeltaDeltaC(T)}]}$ approach was employed to estimate the expression levels of target gene(s) relative (Livak and Schmittgen, 2001). *Ubiquitin* was used as an internal reference. **Table 1** presents the primers in this study.

Inoculation With *R. solani* and Scoring Response of Rice Plants

Rice plants were grown in the glasshouse for 1 month prior to inoculation with the pathogen *R. solani* AG1-IA. The second leaf of the main tiller was cut into 10-cm slices, placed on wet

TABLE 1 | Sequence of the primers used in this study.

Primer	Sequence
Ubiquitin F	CACGGTTCAACAACATCCAG
Ubiquitin R	TGAAGACCCTGACTGGGAAG
PP2A-1 F	CACGGTGTTCAGCGCCCCAAAC
PP2A-1 R	CGCGTTGTGTCCGGCTCTATTTG
PP2A-2 F	GCTAGAGCTCACCAGTTGGTCATG
PP2A-2 R	TACATCTGGCTCTCCCCTTCTTG
PP2A-3 F	CTCTCATCTCAAGGGCACATCAAC
PP2A-3 R	TGTGTCTGGTTCAATTTGCCGAGGAG
PP2A-4 F	CGAACAAAAGGTCGTGACCATATTC
PP2A-4 R	ATCAGGTGTTCTCCGTGTCACATC
PP2A-5 F	TAGCTCGGGCTCATCAACTAGTTATG
PP2A-5 R	AAATAATCGGGCGTCCTCCGTGTCAC
PBZ1 F	CCCTGCCGAATACGCCTAA
PBZ1 R	CTCAAACGCCACGAGAATTTG
PR1b F	GCGTCTTCATCACATGCAACTA
PR1b R	ACCTGAAACAGAAAGAAACAGAGG
PP2A-1 GFP F	CCATGGATGCCGTCGCACGCGGATCTGGAC
PP2A-1 GFP R	AGATCTCAAAAAGTAGTCGGGGGTCTTGCGC

TABLE 2 | *R. solani* inoculation-mediated expression patterns of *PP2A* genes.

Gene	Locus number	Description	Log_2FC	*P*-value
OsPP2A-1	Os06g0574500	Protein Phosphatase 2A catalytic subunit 1	1.02878	0.00063
OsPP2A-2	Os03g0805300	Protein Phosphatase 2A catalytic subunit 2	N/A	N/A
OsPP2A-3	Os02g0217600	Protein Phosphatase 2A catalytic subunit 3	N/A	N/A
OsPP2A-4	Os10g0410600	Protein Phosphatase 2A catalytic subunit 4	N/A	N/A
OsPP2A-5	Os03g0167700	Protein Phosphatase 2A catalytic subunit 5	N/A	N/A

filter paper and stored in a culture dish (36 × 36 × 2.5 cm). In a completely randomized design, five leaves were placed in each plate, with a total of three replicate plates for each treatment. The fungal plug (7 mm in diameter) was cut from the Potato Dextrose Agar (PDA) plate with *R. solani* and placed on the back of the leaf. The leaves were cultured for 72 h at 25°C under continuous light, and the moisture of the filter paper was maintained with sterile water (Gao et al., 2018). Measurement from 0 (no lesion) to 9 (lesions occupying 90–100% of the leaf surface) was done after visual observation. Scores from one to eight represented 10–80% diseased leaf area (Prasad and Eizenga, 2008).

Rice plants that were cultured in a greenhouse for 1 month prior to the tillering stage were used for inoculation. The sheath of the first leaf of the main stem was inoculated with *R. solani* AG1-IA. The PDA fungal plug was inoculated into rice leaf, sprayed with sterile water, and the severity of the disease was determined after 24, 48, and 72 h.

Construction of PP2A-1-GFP Plasmid and Its Subcellular Localization

PP2A-1 ORF region was amplified by PCR and moved into pCAMBIA1302 vector to create *PP2A-1-GFP* plasmid. The *Agrobacterium*-mediated transient expression approach was followed to introduce the fusion proteins into *Nicotiana benthamiana* (Kim et al., 2009). The location of the protein was monitored via GFP fluorescence with a confocal microscope (SP5; Leica, Solms, Germany).

Statistical Analysis

The significant differences between different groups were analyzed using Microsoft Excel to compute the mean, standard deviation, and the Student's *t*-test. Dunnett's test was done employing the SPSS 19.0 statistical software.

RESULTS

Inoculation of *R. solani* Significantly Induced *PP2A-1* Expression

Our previous transcriptome analysis identified that *PP2A-1* expression was induced by the inoculation of *R. solani* AG1-IA (Yuan et al., 2020). Rice genome harbors five PP2A isoforms, and a phylogenetic analysis of PP2A proteins from *Arabidopsis*, rice, tobacco, and tomato revealed that OsPP2A-1 clustered with OsPP2A-3, NtPP2A-2, AtPP2A-1, AtPP2A-2, AtPP2A-5, and LePP2A-1, all of which belong to Subfamily II, while OsPP2A-5, OsPP2A-4, OsPP2A-2, AtPP2A-3, AtPP2A-4, NtPP2A-1, NtPP2A-5, and LePP2A-5 belong to Subfamily I (**Figure 1A**). The homologous sequence alignment of PP2As showed that OsPP2A-1 and OsPP2A-3 shared 98% similarity. The red box indicates a highly conserved region between the five PP2As (**Figure 1B**). qRT-PCR of the 5 *PP2A* genes was done for the verification of the transcriptome data post-inoculation with *R. solani* after 0, 24, 48, and 72 h. The results indicated that only *PP2A-1* expression was induced by *R. solani* infection, and *PP2A-1* expression was the highest 72 h after inoculation, while the other four *PP2A* genes did not respond to *R. solani*. *PBZ1*, a marker gene was used for evaluating pathogen infection, its expression was down-regulated at 24 h after inoculation, while it was up-regulated after 48 h and 72 h of inoculation (**Figure 1C**; **Table 2**). In addition, tissue-specific expression of *PP2As* was examined by qRT-PCR. All *PP2As* were expressed in root, leaf sheath, leaf, and flower tissues, while *PP2A-1* was expressed highly in root and flower, indicating that *PP2As* were ubiquitously expressed in different tissues and developmental stages (**Figure 1D**).

pp2a-1 Mutants Are Susceptible to ShB

To analyze the function of *PP2A* genes in rice defense to ShB, Crispr/Cas9 induced genome editing mutants for *PP2As* were generated. The *PP2A* genes consist of multiple exons and introns in the genome (**Figure 2A**). The sequencing of *PP2A* genome editing mutants revealed that *pp2a-1* mutants have a genomic lesion in the first exon with 1 or 2-bp insertions (*pp2a-1-1*, *pp2a-1-2*) and the *pp2a-2* mutant has a 1-bp deletion in the 11th

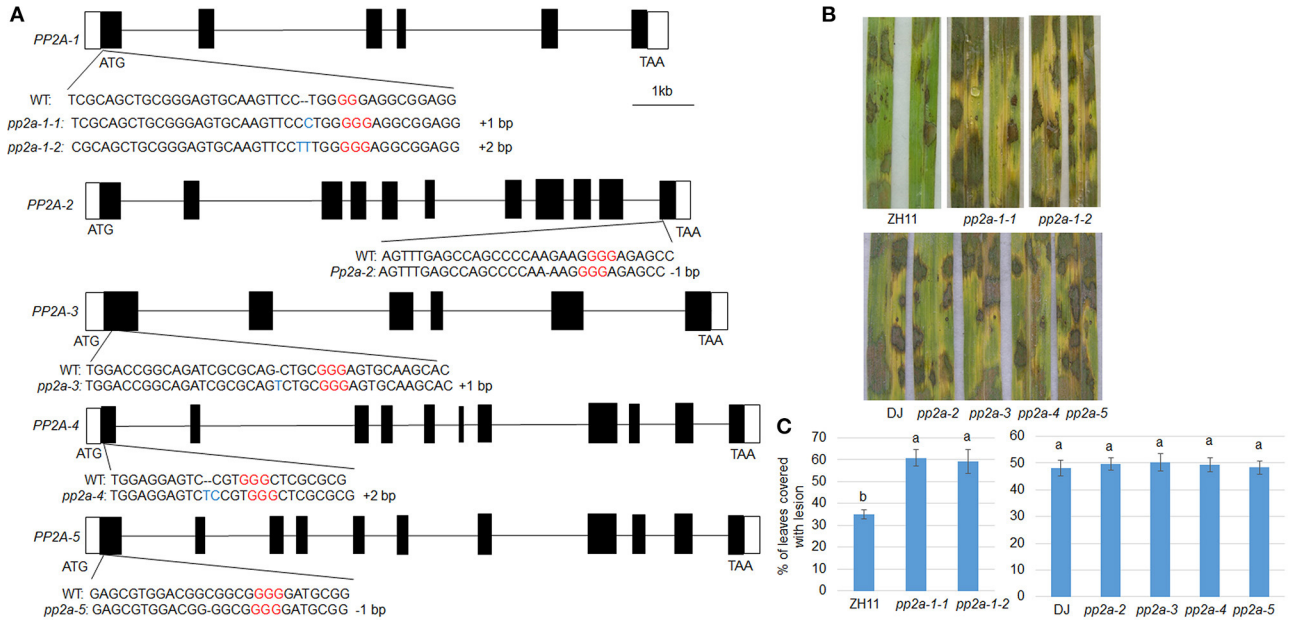


FIGURE 2 | The genomic structure and defense response of *PP2As*. **(A)** Black and white boxes indicate the exon and UTR regions, respectively, while the lines indicate introns. The wild-type (WT) and CRISPR/Cas9-induced edited sequences are shown below the exon. **(B)** The leaves from the ZH11, *pp2a-1* (#*1*, #*2*), DJ, *pp2a-2*, *pp2a-3*, *pp2a-4*, and *pp2a-5* were inoculated with *R. solani* AG1-IA. **(C)** Shows the leaf surface lesion areas. Data show the average ± SE ($n > 10$). Different letters above the bars denote statistically significant differences ($P < 0.05$).

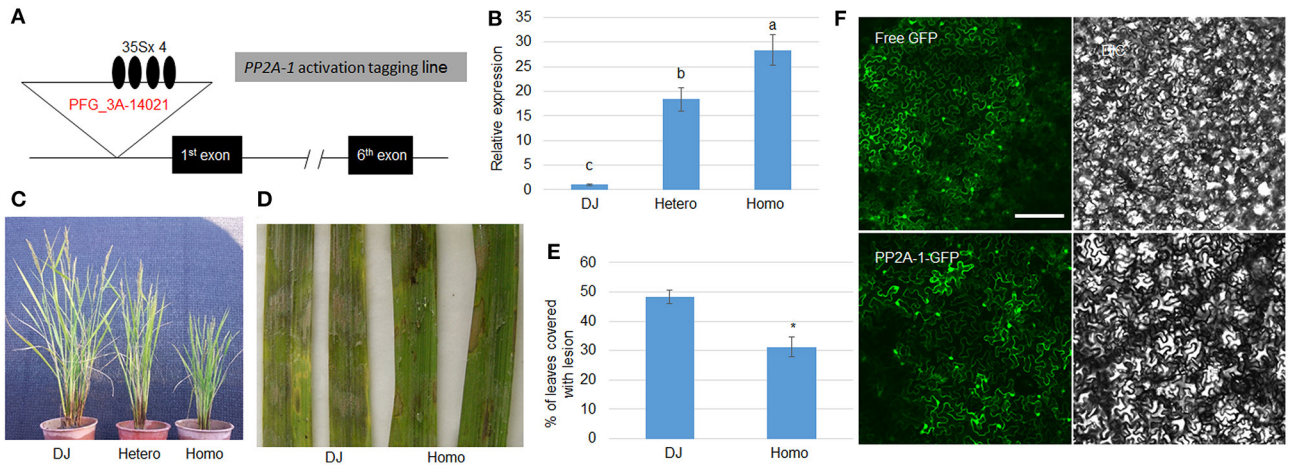


FIGURE 3 | Genomic structure and defense response of *PP2A-1* overexpressor and PP2A-1 subcellular localization. **(A)** Genomic structure of *PP2A-1* activation tagging (*PP2A-1 OX*) line. Black boxes are demonstrative of exons and lines for introns. The T-DNA insertion site is shown as the white triangle and the black oval represents the 35S promoter. PFG_3A-14021 is the ID number of the mutant in the SALK database. **(B)** *PP2A-1* expression levels in DJ (wild-type), *PP2A-1* heterozygote, and homozygote plants. Normalization of expression was done with reference to the internal control *ubiquitin*. The experiments were done in triplicate. Significant changes are illustrated by various letters ($P < 0.05$). **(C)** Three-month-old DJ, *PP2A-1* heterozygous, and homozygous plants. **(D)** Leaves from the DJ and *PP2A-1* homozygous plants were inoculated with *R. solani* AG1-IA. **(E)** The lesion areas on the leaf surfaces are shown in **(D)**. Data show the average ± SE ($n > 10$). Different letters above the bars denote statistically significant differences ($P < 0.05$). **(F)** Free GFP and PP2A-1-GFP were expressed in tobacco leaves. GFP and bright-field channels Scale bar = 20 μm.

exon. The *pp2a-3*, *pp2a-4*, and *pp2a-5* mutants contained edited sequenced in the first exon with a 1-bp insertion, 2-bp insertion, and 1-bp deletion, respectively (**Figure 2A**). After inoculation with *R. solani* AG1-IA, *pp2a-1* genome editing mutants in ZH11 background were more susceptible than ZH11 plants, showing obvious chlorosis (**Figure 2B**), while other *pp2a* genome editing mutants in DJ background had no obvious disease grade differences compared with that in wild-type plants (DJ)

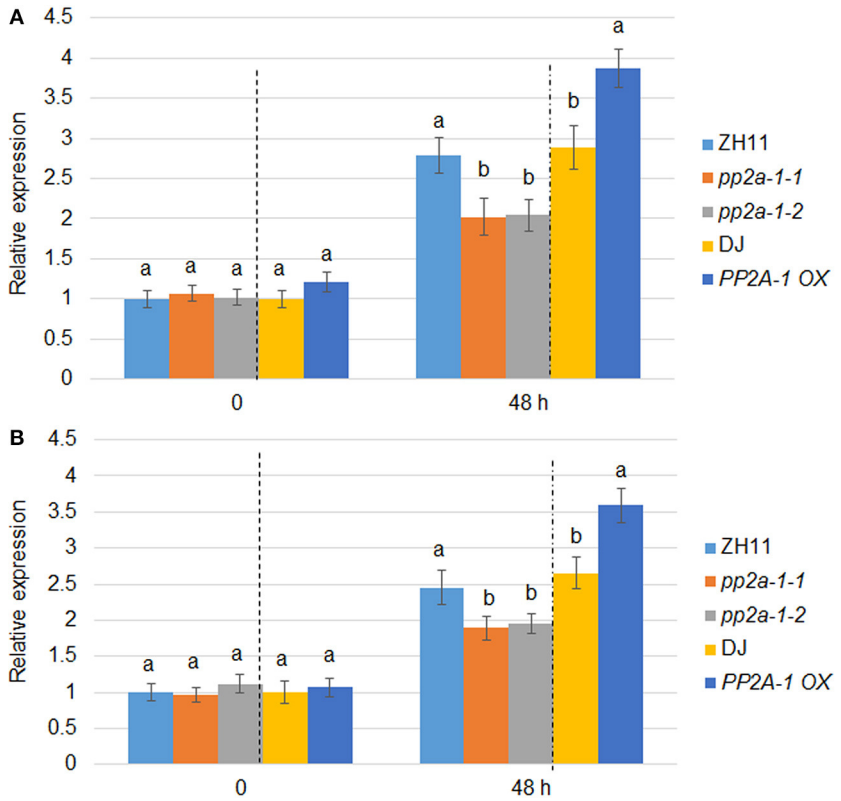


FIGURE 4 | *PBZ1* and *PR1b* expression patterns in wild-type, *pp2a-1*, and *PP2A-1* overexpression plants. **(A)** *PBZ1* and **(B)** *PR1b* expression after *R. solani* inoculation. The *PBZ1* and *PR1b* gene expression patterns were evaluated after 0 and 48 h following *R. solani* inoculation. ZH11 is the control of *pp2a-1-1* and *pp2a-1-2*, while DJ is the control of *PP2A-1 OX* plants. Normalization of expression was done with reference to the internal control *ubiquitin*. The experiments were done in triplicate. Significant changes are illustrated by various letters ($P < 0.05$).

(**Figure 2B**). The lesion coverage on leaves of ZH11 (wild-type), *pp2a-1-1*, and *pp2a-1-2* were 34.1, 60.8, and 59.2%, respectively. However, the lesion area on DJ, *pp2a-2, pp2a-3, pp2a-4*, and *pp2a-5* were 48.2, 49.3, 50.1, 49.5, and 48.4%, respectively (**Figure 2C**).

PP2A-1 Overexpression Enhanced Rice Resistance to ShB

Since *pp2a-1* mutants were susceptible to ShB, the *PP2A-1* overexpression plants were further examined in response to ShB. We isolated a *PP2A-1* activation tagging line (*PP2A-1 OX*) in a T-DNA insertional library (Jeong et al., 2002). In the activation tagging line, T-DNA was inserted in the promoter region in which four copies of the 35S promoter activated the *PP2A-1* expression (**Figure 3A**). The qRT-PCR results indicated that *PP2A-1* expression was significantly higher in hetero- and homozygous *PP2A-1* activation tagging plants than wild type and the *PP2A-1* expression level was significantly higher in homozygous compared to heterozygous plants (**Figure 3B**). *PP2A-1 OX* lines displayed a semi-dwarf phenotype, with homozygous plants significantly shorter than heterozygous and wild-type plants (**Figure 3C**). Next, we selected the homozygous *PP2A-1 OX* to inoculate *R. solani* AG1-IA. The lesion area of *PP2A-1* homozygous overexpression plants was smaller than that of DJ (wild-type) after 48 h of inoculation (**Figure 3D**). The lesion coverage on leaves was 48.1 and 30.8%, respectively (**Figure 3E**), indicating that overexpression of *PP2A-1* enhanced rice resistance to ShB. In addition, PP2A-1-GFP and free GFP were expressed in tobacco leaves, and PP2A-1-GFP signal was detected in the cytoplasm and nucleus (**Figure 3F**).

PP2A-1 Positively Regulates Defense Gene Expression

PP2A-1 OX plants were less susceptible while *pp2a-1* mutants were more susceptible to ShB compared to the wild-type control. The expression patterns of defense genes *PBZ1* and *PR1b* in wild-type, *pp2a-1*, and *PP2A-1 OX* plants were examined following *R. solani* inoculation. qPCR results showed that there was no significant difference in the expression levels of *PBZ1* and *PR1b* among wild-type, genome editing mutants, and overexpression lines with no *R. solani* inoculation. However, the expression level of *PBZ1* in *pp2a-1* mutants was significantly lower than that in control ZH11, while it was higher in *PP2A-1 OX* plants than that in wild-type DJ after 48 h of inoculation (**Figure 4A**). The expression pattern of *PR1b* was similar to that of *PBZ1* at 48 h of inoculation, which showed lower and higher induction kinetics

in *pp2a-1* mutants and *PP2A-1 OX* compared to that in wild-type plants, respectively (**Figure 4B**).

DISCUSSION

Rice ShB caused by *R. solani*, bacterial leaf blight caused by *Xanthomonas oryzae,* and rice blast caused by *Magnaporthe oryzae* are three major diseases that significantly affect rice yield in China (Suryadi et al., 2013). The use of resistant varieties has been the primary means of disease control. However, due to the lack of resistant varieties and that sclerotium has a strong reproductive capacity, ShB control is challenging. Therefore, it is necessary to screen disease resistance genes and understand their resistance mechanism. Previous studies showed that PP2A regulates the development of lateral and primary roots, seed germination, and abiotic stress response against high concentration of sugar, salt, and drought (Yu et al., 2003; País et al., 2009; Liu et al., 2014; Hu et al., 2017). PP2A also plays important roles in biotic stress resistance. For example, AtPP2A is involved in regulation of PTI (pathogen-associated molecular pattern-triggered immunity) upon response to *P. syringae* pv. *tomato* (Pto) DC3000 infection; *LePP2A* gene was rapidly induced by inoculating with the model strain *P. syringae* pv. *tomato; TaPP2A-4B* and *TaPP2A-4D* may negatively regulate wheat defense response to *R. cerealis* infection by triggering the generation of ROS and PTI-mediated induction of PR genes (He et al., 2004; Segonzac et al., 2014; Durian et al., 2016; Zhu et al., 2018), suggesting that PP2A may be a key regulator of PAMP induced immunity. In rice, the induction of okadaic acid-dependent AMY3 and RCht2 (rice chitinase) transcription are regulated via the PP2A signal transduction pathway (Luan et al., 1993; Kim et al., 1998). However, the role of PP2A in rice disease resistance still remains unclear.

Our results indicate that the protein phosphatase 2A catalytic subunit OsPP2A-1 enhances resistance to sheath blight disease in rice. In our previous transcriptome analysis, *OsPP2A-1* was significantly induced by *R. solani* inoculation (Yuan et al., 2020). The CRISPR/Cas9-mediated genome editing lines revealed higher susceptibility of *pp2a-1* mutants to ShB, compared to wild-type control and other *PP2A* mutants (**Figure 2**). While *PP2A-1* OX lines displayed a semi-dwarf phenotype, homozygous plants were significantly shorter than heterozygous and wild-type plants (**Figure 3C**). Inoculation of *R. solani* AG1-IA demonstrated that the *PP2A-1 OX* lines were less susceptible to ShB. The *PP2As* are ubiquitously expressed in different stages and tissues, and PP2A-1-GFP was localized at the cytosol and nucleus in tobacco leaves. The *PP2A-1* expression level was significantly higher in the activation tagging line, with higher *PP2A-1* expression inhibiting plant height while promoting ShB resistance.

The vital involvement of PP2A in responding to pathogens by plants has been demonstrated in recent studies (He et al., 2004; Zhu et al., 2018). For example, in *A. thaliana*, RLKs FLS2 (flagellin sensing receptor2) recognizes and EFR (EF-Tu receptor) is capable of recognizing the EF- Tu (elongation factor), both are PAMPs (pathogen-associated molecular pattern) of bacterial pathogens. The autophosphorylation and functioning of BAK1 (BRI1-associated kinase 1) is limited by PP2A-holoenzyme (Segonzac et al., 2014). *PP2A-c4* and *PP2A-a1* gene knockout mutants display a stronger resistance to virulent *P. syringae* pv. *tomato* DC3000 (Segonzac et al., 2014). The BSMV-VIGS (barley stripe mosaic virus–induced gene silencing) approach was applied to augment *R. cerealis* resistance in wheat attributed to *TaPP2Ac-4B* and *TaPP2Ac-4D* knock-down, suggesting the negative regulation of TaPP2A to wheat sharp eyespot. In many species, PP2A appears to function as a negative regulator, while OsPP2A-1 was found to positively regulate resistance to *R. solani* in this study. It seemed that the same gene may play different functions in response to infection of different types of pathogens. For example, WRKY transcription factors were more resistant to the hemibiotrophic bacterial pathogen *P. syringae*, but more susceptible to necrotrophic fungal pathogen *B. cinerea* in *Arabidopsis* (Xu et al., 2006), implying that PP2A-1 might play diverse functions when experiencing different stimuli. In our study, we confirmed that the CRISPR/Cas9-induced *pp2a-1* genome editing mutants more susceptible to *R. solani*, while the other *pp2as* were similar to wild-type plants in response to *R. solani* infection. It may be valuable to dissect the associated molecular mechanism in the future research.

As mentioned earlier, PP2As comprises three subunits (A, B, and C). In the *A. thaliana* genome, these subunits are encoded by five genes of subunit C, three genes of subunit A, and 17 genes of subunit B (Farkas et al., 2007) to establish a minimum of 255 novel forms of the molecule. Immunity in plants is influenced by subunits A and B (with scaffolding and regulatory functioning, respectively). Resistance to *P. syringae* pv. *tomato* was augmented due to a subunit B-B′θ deficient mutation (Kataya et al., 2015). AtPP2A-B′γ enhances the negatively regulated defense against *Myzus persicae* (green peach aphid) and *B. cinerea* (a necrotrophic fungus) (Trotta et al., 2011; Rasool et al., 2014). PR protein phosphorylation (PR1, PR2–PR5) is augmented by mutations in subunit *AtPP2A-B′γ* (Trotta et al., 2011). The constitutive expression of *PR1a, PR1b*, and *PR5* was induced by *NbNPP4-1* and *NbNPP4-2* silencing in *N. benthamiana* (He et al., 2004). *PR2* levels were up-regulated by *TaPP2A* silencing (Zhu et al., 2018). *PBZ1*, a *PR10* family protein accumulates in rice tissues which are in the process of cell mortality (Huang et al., 2016; Moselhy et al., 2016). In this study, *PBZ1* and *PR1b* genes were up-regulated by *R. solani* infection, suggesting that that *PBZ1* and *PR1* play a role in ShB resistance in rice. The results indicated that *R. solani*-induced *PBZ1* and *PR1b* expressions are under control of PP2A-1, suggesting that the expression of PP2A-1 might be through the activation of *PR* genes to promote rice defense.

Taken together, our findings suggest that the protein phosphatase 2A catalytic subunit, PP2A-1, regulates the defense response in rice to *R. solani* infection. This study revealed a new function of the rice *PP2A* in immune response, which provided a potential target for breeding ShB-resistant lines.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

QL, QM, and YX conceived and designed the studies and wrote the manuscript. QL, JC, VK, ZL, and DY collected and analyzed the data. All authors have read and approved the final version of the manuscript.

FUNDING

This work was supported by the Support Program for Science and Technology Innovation Talents of Shenyang (RC190489), and the Natural Science Foundation of Liaoning Province (2020-YQ-05).

REFERENCES

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823. doi: 10.1126/science.1231143

Du, Y., Shi, Y., Yang, J., Chen, X., Xue, M., Zhou, W., et al. (2013). A serine/threonine-protein phosphatase PP2A catalytic subunit is essential for asexual development and plant infection in *Magnaporthe oryzae*. *Curr. Genet.* 59, 33–41. doi: 10.1007/s00294-012-0385-3

Durian, G., Rahikainen, M., Alegre, S., Brosche, M., and Kangasjarvi, S. (2016). Protein phosphatase 2A in the regulatory network underlying biotic stress resistance in plants. *Front. Plant Sci.* 7:812. doi: 10.3389/fpls.2016.00812

Farkas, I., Dombradi, V., Miskei, M., Szabados, L., and Koncz, C. (2007). *Arabidopsis* PPP family of serine/threonine phosphatases. *Trends Plant Sci.* 12, 169–176. doi: 10.1016/j.tplants.2007.03.003

Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D. L., Wei, P., et al. (2013). Efficient genome editing in plants using a CRISPR/Cas system. *Cell Res.* 23, 1229–1232. doi: 10.1038/cr.2013.114

Gao, Y., Zhang, C., Han, X., Wang, Z. Y., Ma, L., Yuan, D. P., et al. (2018). Inhibition of OsSWEET11 function in mesophyll cells improves resistance of rice to sheath blight disease. *Mol. Plant Pathol.* 19, 2149–2161. doi: 10.1111/mpp.12689

He, X., Anderson, J. C., Pozo, O.d., Gu, Y.-Q., Tang, X., and Martin, G. B. (2004). Silencing of subfamily I of protein phosphatase 2a catalytic subunits results in activation of plant defenseresponses and localized cell death. *Plant J.* 38, 563–577. doi: 10.1111/j.1365-313X.2004.02073.x

Helliwell, E. E., Wang, Q., and Yang, Y. (2013). Transgenic rice with inducible ethylene production exhibits broad-spectrum disease resistance to the fungal pathogens *Magnaporthe oryzae* and *Rhizoctonia solani*. *Plant Biotechnol. J.* 11, 33–42. doi: 10.1111/pbi.12004

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31, 827–832. doi: 10.1038/nbt.2647

Hu, R., Zhu, Y., Wei, J., Chen, J., Shi, H., Shen, G., et al. (2017). Overexpression of PP2A-C5 that encodes the catalytic subunit 5 of protein phosphatase 2A in *Arabidopsis* confers better root and shoot development under salt conditions. *Plant Cell Environ.* 40, 150–164. doi: 10.1111/pce.12837

Huang, L. F., Lin, K. H., He, S. L., Chen, J. L., Jiang, J. Z., Chen, B. H., et al. (2016). Multiple patterns of regulation and overexpression of a ribonuclease-like pathogenesis-related protein gene, *OsPR10a*, conferring disease resistance in rice and *Arabidopsis*. *PLoS ONE* 11:e0156414. doi: 10.1371/journal.pone.0156414

Jeong, D. H., An, S., Kang, H. G., Moon, S., Han, J. J., Park, S., et al. (2002). T-DNA insertional mutagenesis for activation tagging in rice. *Plant physiology* 130, 1636–1644. doi: 10.1104/pp.014357

Jimmy, J. L., and Babu, S. (2019). Gene network mediated by WRKY13 to regulate resistance against sheath infecting fungi in rice (*Oryza Sativa* L.). *Plant Sci.* 280, 269–282. doi: 10.1016/j.plantsci.2018.12.017

Kataya, A. R. A, Heidari, B., and Lillo, C. (2015). Protein phosphatase 2A regulatory subunits affecting plant innate immunity, energy metabolism, and flowering time—joint functions among B'η subfamily members. *Plant Signal. Behav.* 10:e1026024. doi: 10.1080/15592324.2015.1026024

Kim, C. Y., Gal, S. W., Choe, M. S., Jeong, S. Y., Lee, S. I., Cheong, Y. H., et al. (1998). A new class II rice chitinase, *Rcht2,* whose induction by fungal elicitor is abolished by protein phosphatase 1 and 2A inhibitor. *Plant Mol. Biol.* 37, 523–534. doi: 10.1023/A:1005960313459

Kim, J.-G., Li, X., Roden, J. A., Taylor, K. W., Aakre, C. D., Su, B., et al. (2009). Xanthomonas T3S effector XopN suppresses PAMP-triggered immunity and interacts with a tomato atypical receptor-like kinase and TFT1. *Plant Cell* 21, 1305–1323. doi: 10.1105/tpc.108.063123

Kouzai, Y., Kimura, M., Watanabe, M., Kusunoki, K., Osaka, D., Suzuki, T., et al. (2018). Salicylic acid-dependent immunity contributes to resistance against *Rhizoctonia solani*, a necrotrophic fungal agent of sheath blight, in rice and *Brachypodium distachyon*. *New Phytol.* 217, 771–783. doi: 10.1111/nph.14849

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for Bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Li, N., Kong, L., Zhou, W., Zhang, X., Wei, S., Ding, X., et al. (2013). Overexpression of Os2H16 enhances resistance to phytopathogens and tolerance to drought stress in rice. *Plant Cell Tissue Organ Cult.* 115, 429–441. doi: 10.1007/s11240-013-0374-3

Li, N., Wei, S., Chen, J., Yang, F., Kong, L., Chen, C., et al. (2018). OsASR2 regulates the expression of a defence-related gene, *Os2H16*, by targeting the GT-1 *cis*-element. *Plant Biotechnol. J.* 16, 771–783. doi: 10.1111/pbi.12827

Liu, D., Li, A., Mao, X., and Jing, R. (2014). Cloning and characterization of TaPP2Ab B"-Alpha, a member of the PP2A regulatory subunit in wheat. *PLoS ONE* 9:e94430. doi: 10.1371/journal.pone.0094430

Liu, Z., Liu, N., Jiang, H., Yan, L., Ma, Z., and Yin, Y. (2018). The activators of type 2A phosphatases (PP2A) regulate multiple cellular processes via PP2A-dependent and -independent mechanisms in *Fusarium graminearum*. *Mol. Plant Microbe Interact.* 31, 1121–1133. doi: 10.1094/MPMI-03-18-0056-R

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ Method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Luan, S., Li, W., Rusnak, F., Assmann, S. M., and Schreiber, S. L. (1993). Immunosuppressants implicate protein phosphatase regulation of K+ channels in guard cells. *Proc. Natl. Acad. Sci. U.S.A.* 90, 2202–2206. doi: 10.1073/pnas.90.6.2202

Ma, X., Chen, L., Zhu, Q., Chen, Y., and Liu, Y. G. (2015). Rapid decoding of sequence-specific nuclease-induced heterozygous and biallelic mutations by direct sequencing of PCR products. *Mol. Plant* 8, 1285–1287. doi: 10.1016/j.molp.2015.02.012

Maeda, S., Dubouzet, J. G., Kondou, Y., Jikumaru, Y., Seo, S., Oda, K., et al. (2019). The rice CYP78A gene BSR2 confers resistance to *Rhizoctonia solani* and affects seed size and growth in *Arabidopsis* and Rice. *Sci. Rep.* 9:587. doi: 10.1038/s41598-018-37365-1

Mao, B., Liu, X., Hu, D., and Li, D. (2014). Co-expression of RCH10 and AGLU1 confers rice resistance to fungal sheath blight *Rhizoctonia solani* and blast *Magnorpathe oryzae* and reveals impact on seed germination. *World J. Microbiol. Biotechnol.* 30, 1229–1238. doi: 10.1007/s11274-013-1546-3

Máthé, C., Garda, T., Freytag, C., and, Hamvas, M. M. (2019). The role of serine-threonine protein phosphatase PP2A in plant oxidative stress signaling—facts and hypotheses. *Int. J. Mol. Sci.* 20:3028. doi: 10.3390/ijms20123028

Moselhy, S. S., Asami, T., Abualnaja, K. O., Al-Malki, A. L., Yamano, H., Akiyama, T., et al. (2016). Spermidine, a polyamine, confers resistance to rice blast. *J. Pestic. Sci.* 41, 79–82. doi: 10.1584/jpestics.D16-008

Nishimura, A., Aichi, I., and Matsuoka, M. (2006). A protocol for agrobacterium-mediated transformation in rice. *Nat. Protoc.* 1, 2796–2802. doi: 10.1038/nprot.2006.469

País, S. M., González, M. A., Téllez-Iñón, M. T., and Capiati, D. A. (2009). Characterization of potato (*Solanum tuberosum*) and tosdmato (*Solanum lycopersicum*) protein phosphatases type 2A catalytic subunits and their involvement in stress responses. *Planta* 230, 13–25. doi: 10.1007/s00425-009-0923-5

Peng, X., Hu, Y., Tang, X., Zhou, P., Deng, X., Wang, H., et al. (2012). Constitutive expression of rice WRKY30 gene increases the endogenous jasmonic acid accumulation, PR gene expression and resistance to fungal pathogens in rice. *Planta* 236, 1485–1498. doi: 10.1007/s00425-012-1698-7

Peng, X., Wang, H., J.-,Jang, C., Xiao, T., He, H., Jiang, D., et al. (2016). OsWRKY80-OsWRKY4 module as a positive regulatory circuit in rice resistance against *Rhizoctonia solani*. *Rice* 9:63. doi: 10.1186/s12284-016-0137-y

Prasad, B., and Eizenga, G. C. (2008). Rice sheath blight disease resistance identified in *Oryza* spp. Accessions. *Plant Dis.* 92, 1503–1509. doi: 10.1094/PDIS-92-11-1503

Rahikainen, M., Pascual, J., Alegre, S., Durian, G., and Kangasjärvi, S. (2016). PP2A phosphatase as a regulator of ROS signaling in plants. *Antioxidants* 5:8. doi: 10.3390/antiox5010008

Rasool, B., Karpinska, B., Konert, G., Durian, G., Denessiouk, K., Kangasjarvi, S., et al. (2014). Effects of light and the regulatory B-subunit composition of protein phosphatase 2A on the susceptibility of *Arabidopsis thaliana* to aphid (*Myzus persicae*) infestation. *Front. Plant Sci.* 5:405. doi: 10.3389/fpls.2014.00405

Savary, S., Castilla, N. P., Elazegui, F. A., McLaren, C. G., Ynalvez, M. A., and Teng, P. S. (1995). Direct and indirect effects of nitrogen supply and disease source structure on rice sheath blight spread. *Phytopathology* 85, 959–965. doi: 10.1094/Phyto-85-959

Savary, S., Willocquet, L., Elazegui, F. A., Castilla, N. P., and Teng, P. S. (2000). Rice pest constraints in tropical asia: quantification of yield losses due to rice pests in a range of production situations. *Plant Dis.* 84, 357–369. doi: 10.1094/PDIS.2000.84.3.357

Segonzac, C., Macho, A. P., Sanmartin, M., Ntoukakis, V., Sanchez-Serrano, J. J., and Zipfel, C. (2014). Negative control of BAK1 by protein phosphatase 2a during plant innate immunity. *EMBO J.* 33, 2069–2079. doi: 10.15252/embj.201488698

Shah, J. M., Raghupathy, V., and Veluthambi, K. (2009). Enhanced sheath blight resistance in transgenic rice expressing an endochitinase gene from *Trichoderma virens*. *Biotechnol. Lett.* 31, 239–244. doi: 10.1007/s10529-008-9856-5

Singh, P., Mazumdar, P., Harikrishna, J. A., and Babu, S. (2019). Sheath blight of rice: a review and identification of priorities for future research. *Planta* 250, 1387–1407. doi: 10.1007/s00425-019-03246-8

Sun, Q., Li, D. D., Chu, J., Yuan, D. P., Li, S., Zhong, L. J., et al. (2020). Indeterminate domain proteins regulate rice defense to sheath blight disease. *Rice* 13:15. doi: 10.1186/s12284-020-0371-1

Sun, Q., Li, T. Y., Li, D. D., Wang, Z. Y., Li, S., Li, D. P., et al. (2019). Overexpression of loose plant architecture 1 increases planting density and resistance to sheath blight disease via activation of PIN-FORMED 1a in rice. *Plant Biotechnol. J.* 17, 855–857. doi: 10.1111/pbi.13072

Suryadi, Y., Susilowati, D. N., Kadir, T. S., Zaffan, Z. R., Hikmawati, N. and Mubarik, N. R. (2013). Bioformulation of antagonistic bacterial consortium for controlling blast, sheath blight and bacterial blight diseases on rice. *Asian J. Plant Pathol.* 7, 92–108. doi: 10.3923/ajppaj.2013.92.108

Taheri, P., and Tarighi, S. (2011). Cytomolecular aspects of rice sheath blight caused by *Rhizoctonia solani*. *Eur. J. Plant Pathol.* 129, 511–528. doi: 10.1007/s10658-010-9725-7

Tiwari, M., Srivastava, S., Singh, P. C., Mishra, A. K., and Chakrabarty, D. (2020). Functional characterization of *tau* class glutathione-S-transferase in rice to provide tolerance against sheath blight disease. *3 Biotech* 10:84. doi: 10.1007/s13205-020-2071-3

Trotta, A., Wrzaczek, M., Scharte, J., Tikkanen, M., Konert, G., Rahikainen, M., et al. (2011). Regulatory subunit B'gamma of protein phosphatase 2A prevents unnecessary defense reactions under low light in *Arabidopsis*. *Plant physiol.* 156, 1464–1480. doi: 10.1104/pp.111.178442

Wang, H., Meng, J., Peng, X., Tang, X., Zhou, P., Xiang, J., et al. (2015). Rice WRKY4 acts as a transcriptional activator mediating defense responses toward *Rhizoctonia solani*, the causing agent of rice sheath blight. *Plant Mol. Biol.* 89, 157–171. doi: 10.1007/s11103-015-0360-8

Xu, X., Chen, C., Fan, B., and Chen, Z. (2006). Physical and functional interactions between pathogen-induced *Arabidopsis* WRKY18, WRKY40, and WRKY60 transcription factors. *Plant cell* 18, 1310–1326. doi: 10.1105/tpc.105.037523

Yellareddygari, S, Reddy, M. S., Kloepper, J. W., Lawrence, K. S., and Fadamiro, H. (2014). Rice sheath blight: a review of disease and pathogen management approaches. *J. Plant Pathol. Microbiol.* 05:1000241. doi: 10.4172/2157-7471.1000241

Yu, R. M., Wong, M. M., Jack, R. W., and Kong, R. Y. (2005). Structure, evolution, and expression of a second subfamily of protein phosphatase 2A catalytic subunit genes in the rice plant (*Oryza sativa* L.). *Planta* 222, 757–768. doi: 10.1007/s00425-005-0018-x

Yu, R. M. K., Zhou, Y., Xu, Z.-F., Chye, M.-L., and Kong, R. Y. C. (2003). Two genes encoding protein phosphatase 2A catalytic subunits are differentially expressed in rice. *Plant Mol. Biol.* 51, 295–311. doi: 10.1023/a:1022006023273

Yuan, D. P., Xu, X. F., Hong, W.-J., Wang, S. T., Jia, X. T., Liu, Y., et al. (2020). Transcriptome analysis of rice leaves in response to *Rhizoctonia solani* infection and reveals a novel regulatory mechanism. *Plant Biotechnol. Rep.* 14, 559–573. doi: 10.1007/s11816-020-00630-9

Zhu, G., Liang, E., Lan, X., Li, Q., Qian, J., Tao, H., et al. (2019). ZmPGIP3 gene encodes a polygalacturonase-inhibiting protein that enhances resistance to sheath blight in rice. *Phytopathology* 109, 1732–1740. doi: 10.1094/PHYTO-01-19-0008-R

Zhu, X., Wang, Y., Su, Z., Lv, L., and Zhang, Z. (2018). Silencing of the wheat protein phosphatase 2A catalytic subunit TaPP2Ac enhances host resistance to the necrotrophic pathogen *Rhizoctonia cerealis*. *Front. Plant Sci.* 9:1437. doi: 10.3389/fpls.2018.01437

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Lin, Chu, Kumar, Yuan, Li, Mei and Xuan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genome Editing

ORIGINAL RESEARCH
published: 08 July 2021
doi: 10.3389/fgeed.2021.673566

CRISPR/Cas9-Mediated Multi-Allelic Gene Targeting in Sugarcane Confers Herbicide Tolerance

Mehmet Tufan Oz[1,2†], *Angelika Altpeter*[1], *Ratna Karan*[1†], *Aldo Merotto*[1†] *and Fredy Altpeter*[1,2*]

[1] Agronomy Department, Plant Molecular and Cellular Biology Program, Genetics Institute, University of Florida, IFAS, Gainesville, FL, United States, [2] DOE Center for Advanced Bioenergy and Bioproducts Innovation, Gainesville, FL, United States

OPEN ACCESS

Edited by:
Lanqin Xia,
Chinese Academy of Agricultural Sciences, China

Reviewed by:
Tomas Cermak,
Inari Agriculture, United States
Sergei Svitashev,
Corteva Agriscience™, United States

***Correspondence:**
Fredy Altpeter
altpeter@ufl.edu

†Present address:
Mehmet Tufan Oz,
Earlham Institute, Norwich Research Park, Norwich, United Kingdom
Ratna Karan,
NCF Diagnostics and DNA Technologies, Alachua, FL, United States
Aldo Merotto,
Crop Science Department, School of Agriculture, Federal University of Rio Grande do Sul, Porto Alegre, Brazil

Specialty section:
This article was submitted to Genome Editing in Plants, a section of the journal Frontiers in Genome Editing

Received: 27 February 2021
Accepted: 28 May 2021
Published: 08 July 2021

Citation:
Oz MT, Altpeter A, Karan R, Merotto A and Altpeter F (2021) CRISPR/Cas9-Mediated Multi-Allelic Gene Targeting in Sugarcane Confers Herbicide Tolerance. Front. Genome Ed. 3:673566. doi: 10.3389/fgeed.2021.673566

Sugarcane is the source of 80% of the sugar and 26% of the bioethanol produced globally. However, its complex, highly polyploid genome ($2n = 100 - 120$) impedes crop improvement. Here, we report efficient and reproducible gene targeting (GT) in sugarcane, enabling precise co-editing of multiple alleles *via* template-mediated and homology-directed repair (HDR) of DNA double strand breaks induced by the programmable nuclease CRISPR/Cas9. The evaluation of 146 independently transformed plants from five independent experiments revealed a targeted nucleotide replacement that resulted in both targeted amino acid substitutions W574L and S653I in the acetolactate synthase (ALS) in 11 lines in addition to single, targeted amino acid substitutions W574L or S653I in 25 or 18 lines, respectively. Co-editing of up to three ALS copies/alleles that confer herbicide tolerance was confirmed by Sanger sequencing of cloned long polymerase chain reaction (PCR) amplicons. This work will enable crop improvement by conversion of inferior alleles to superior alleles through targeted nucleotide substitutions.

Keywords: homologous recombination, CRISPR/Cas9, sugarcane (*Saccharum* hybrid complex), genome editing, gene targeting, homology directed repair

INTRODUCTION

Sugarcane (*Saccharum* spp. hybrid) is the source of 80% of the world's sugar and 26% of its bioethanol, and it is an exceptionally productive crop due to its superior light conversion and water- and nitrogen-use efficiencies (Byrt et al., 2011). The genome of sugarcane is the most complex of any domesticated agricultural species ($2n = 100$–120) (Piperidis and D'Hont, 2020). Modern sugarcane cultivars are derived from hybridization between *Saccharum officinarum* ($2n = 80$, $x = 10$) and *Saccharum spontaneum* ($2n = 40$–128, $x = 8$); these are responsible for high sugar content and stress tolerance or vigor, respectively (Piperidis and D'Hont, 2020). Elevated sugar production was achieved by backcrossing the hybrid to *S. officinarum*. The resulting cultivars are aneuploid, highly heterozygous, and highly polyploid, with 100–120 chromosomes. Most chromosomes are derived from *S. officinarum*, depending on the cultivar, with 10–20% originating from *S. spontaneum* and ∼10% from interspecific recombinants (D'Hont et al., 1996). Modern cultivars typically have 12 copies of each of the first four basic chromosomes, while parent species tend to differ in those basic chromosomes. One to four of these copies correspond to entire *S. spontaneum* chromosomes or interspecific recombinant chromosomes. In addition, inter-chromosomal translocations are also present (Garsmeur et al., 2018; Piperidis and D'Hont, 2020). Unsurprisingly, elite cultivars

require vegetative propagation to maintain their quality and agronomic performance. Cultivar development by conventional breeding must overcome the challenges of photoperiod sensitivity in floral induction, lack of pollen fertility, and synchrony of flowering in most parental sugarcane clones (Horsley and Zhou, 2013).

Genome editing with sequence-specific nucleases is revolutionizing crop breeding (Voytas and Gao, 2014; Zhang et al., 2018) and has promising applications for sugarcane and other vegetatively propagated polyploid crops with complex genomes (Jung and Altpeter, 2016; Weeks, 2017). DNA cleavage through sequence-specific nucleases, including CRISPR/Cas9, is followed by the use of cellular repair mechanisms, including non-homologous end joining (NHEJ) or homology-directed repair (HDR), to rectify double-strand breaks (DSBs). Non-homologous end joining enables the construction of knockout alleles through frameshift mutations. By contrast, HDR-mediated gene targeting (GT) allows the introduction of precise genetic modifications, including single-nucleotide substitutions, gene replacements, and large insertions (Huang and Puchta, 2019). Gene targeting requires recombination with a repair template with homology to the break site. The repair template may be the sister chromatid or a co-introduced targeting vector, containing a desired sequence modification for incorporation into the break site (Puchta, 1998; Bortesi and Fischer, 2015; Que et al., 2019).

Homology-directed repair-mediated GT in plant genomes remains a challenge, resulting in a small number of successful studies (Chen et al., 2019; Sedeek et al., 2019). In contrast to knock-outs with up to a 100% mutation frequency (Brooks et al., 2014), precise gene replacement frequencies are typically in the range of 0.1 to a few percent, with large, targeted insertions being the most challenging (Huang and Puchta, 2019; Mao et al., 2019). Because GT is typically an inefficient process, mutations that confer a selectable phenotype, such as herbicide resistance, have been favored as initial targets for recovering the events that caused them (Shukla et al., 2009; Svitashev et al., 2015, 2016; Butler et al., 2016; Sun et al., 2016). The acetolactate synthase (ALS) enzyme catalyzes the biosynthesis of essential branched-chain amino acids (Smith et al., 1989) and is strongly inhibited by several herbicides, such as sulfonylureas, imidazolinones, triazolopyrimidines, pyrimidinyloxybenzoates, and sulfanilamide-carbonyl-thiazolidinones (Smith et al., 1989; Shaner and O'Connor, 1991; Devine et al., 1992; Duggleby et al., 2008; Powles and Yu, 2010). Resistance to ALS-inhibiting herbicides is controlled by specific mutations of the *ALS* gene at amino acid positions Ala122, Pro197, Ala205, Asp376, Arg377, Trp574, Ser653, and Ser654 (Tan et al., 2006; Li et al., 2008; Merotto et al., 2009; Rodríguez-Suárez et al., 2009). One of these mutations, Trp574, confers high-level resistance and cross-resistance to all ALS-inhibiting herbicides (Tan et al., 2006). The Pro197 mutation results in the greatest tolerance to sulfonylureas. Mutations at Ala122, Ala205, and Ser653 provide resistance to imidazolinones (Tan et al., 2006). Mutated herbicide-resistant *ALS* alleles are semi-dominant, and increased resistance has been ascribed to two or more resistance-conferring codons and/or resistance genes or alleles in a given genotype (Pozniak and Hucl, 2004; Shimizu et al., 2005).

Here, we report efficient and reproducible multi-allelic GT in sugarcane. We also provide an evaluation of two alternative sgRNAs used alone or in combination for altering one or two codons (W574L and/or S653I) that are known to confer herbicide tolerance. We also compared HDR frequency following the biolistic delivery of different quantities of the repair template.

MATERIALS AND METHODS

Construction of Plasmid Vectors Carrying DNA-Editing Tools

Five plasmids carrying various DNA elements were constructed (**Supplementary Figure 1**, **Supplementary Table 1**). *Cas9* from *Streptococcus pyogenes* was codon-optimized for sugarcane and custom-synthesized, including nuclear localization signals from SV40 and nucleoplasmin, by GENEWIZ, Inc. (South Plainfield, NJ, USA). The selectable marker *neomycin phosphotransferase* II (*npt*II) and codon-optimized *Cas9* (*coCas9*) were under the transcriptional control of cauliflower mosaic virus 35S RNA (CaMV35S) promoter with a 70 kDa heat-shock protein (HSP70) intron and CaMV35S terminator or *Arabidopsis thaliana* heat-shock protein (AtHSP) terminator, respectively. Benchling (https://www.benchling.com) was used for the gRNA design process. The sgRNAs that direct cleavage near amino acid residue 574 (sgRNA1) and 653 (sgRNA2) were placed under the transcriptional control of U6 promoter from *Oryza sativa*. The 20 bp target sequences of sgRNA1 and sgRNA2 are 5′tcactgggaggttctcaatt and 5′gtcaaagaaaggcagggagg, respectively. The sgRNA constructs with U6 promoters were custom-synthesized by GENEWIZ. A repair template (**Supplementary Table 2**) was designed based on the sugarcane *ALS* sequence, with nucleotide modifications to introduce W574L and S653I, along with two modified PAM sites (PAM1 and PAM2) to prevent cleavage of the template by sgRNA1 or sgRNA2. Homology arms of 1,007 bp at the 5′-end and 447 bp at the 3′-end relative to the targeted DSBs were included in the design of the 1,833 bp double-stranded DNA template custom-synthesized by GENEWIZ.

Cas9 *in vitro* Cleavage Assay

The targeting efficiency of the sgRNAs was investigated *in vitro* using a commercial Cas9 protein (PNA Bio, CA). A 1,174 bp target region of the *ALS* gene was amplified using Q5 high-fidelity DNA polymerase (NEB, MA) with the primer pair UP5 and DO1 (**Supplementary Table 3**) in a conventional polymerase chain reaction (PCR). The amplicon was isolated using gel electrophoresis and purified using the Monarch® DNA gel extraction kit (NEB). Two sgRNAs were transcribed *in vitro* using the HiScribe™ T7 quick high-yield RNA synthesis kit (NEB) from a DNA template generated with PCR. For this approach, a T7 promoter was fused upstream of the sgRNAs, using Q5 DNA polymerase with primer pair C1F and CR for sgRNA1 and primer pair C2F and CR for sgRNA2. After DNase I (RNase-free) treatment, synthesized RNA was purified through phenol:chloroform extraction followed by ethanol precipitation. An *in vitro* cleavage assay was performed at 37°C overnight in the presence of 450 ng target *ALS* fragment, 150 ng Cas9 protein,

300 ng sgRNA, and 100 μg/mL BSA buffered with Buffer 3 (NEB). Cleavage products were isolated using agarose gel electrophoresis to determine the sgRNA targeting efficiency.

Generation of Genome-Edited Sugarcane

Plasmids carrying expression cassettes and donor template were introduced into sugarcane callus through biolistic gene transfer for indirect embryogenesis, as described earlier (Taparia et al., 2012). The amounts of DNA used per shot and the molar ratios of editing components are shown in **Table 1**. For treatments 1, 2, 3, and 5, 1.5 ng per kilobase DNA per shot (ng/kb and shot) was used (Wu et al., 2015). For treatment 4, 0.5 ng was used. The editing components were precipitated onto 1.8 mg gold particles using a protocol described previously (Sandhu and Altpeter, 2008). Approximately 35 calli per shot were placed without gaps to cover a circular target area 30 mm in diameter (**Supplementary Figure 2B**). Following biolistic gene transfer and callus selection with 20 mg/L Geneticin, calli were transferred to regeneration media containing 90.5 μg/L bispyribac sodium (BS) for shoot elongation for 60 days, followed by rooting on media without BS and growth regulators (**Supplementary Figure 2**). The targeted mutations W574L and S653I in the ALS gene confer resistance to several herbicides, including BS, as previously demonstrated in sugarcane with a transgenic approach (Dermawan et al., 2016).

Nucleic Acid Isolation and PCR

Following Geneticin selection and again following regeneration on culture media containing BS, approximately 0.1–0.2 g tissue was sampled from each line and frozen in liquid nitrogen. The frozen tissue was ground using TissueLyser II (Qiagen, Germany). Genomic DNA was extracted using cetyltrimethylammonium bromide as described earlier (Murray and Thompson, 1980). DNA was dissolved in 50–100 μl nuclease-free water and quantified using NanoDrop™ One (Thermo Fisher Scientific, MA, USA). The presence of transgenes in genomic DNA extracts was investigated using PCR with gene-specific primers (**Supplementary Table 3**).

Verification of Target Mutations

The initial screening of target mutations was completed *via* restriction enzyme (RE) digestion. Long PCR amplicons (1,913 bp) were generated with the primers DO1 and UP6 (**Supplementary Table 3**), the latter located outside (upstream) of the template sequence to prevent the amplification of randomly inserted templates. Primers F1 and R1 were used to generate shorter, 455 bp amplicons in a nested PCR for subsequent screening *via* RE digestion. The targeted mutation W574L is expected to introduce an *Mme*I RE recognition site. Mutation S653I is expected to eliminate the *Bfa*I RE recognition site (**Figure 1**). Enzyme digestion was performed following the manufacturer's instructions.

TaqMan® Probe-Based Genotyping Assays

Intended mutations were further verified using fluorescent-labeled TaqMan® MGB probes, designed to detect wild-type and mutant alleles at mutation sites W574L and S653I in two HDR-genotyping assays (**Supplementary Figure 3**). Fluorophore VIC® was employed for wild-type alleles and fluorophore FAM was used for mutant alleles (Thermo Fisher Scientific, MA, USA). Thermal-cycling conditions constituted initial denaturation at 95°C for 10 min and 45 cycles of denaturation and extension at 95°C for 15 s and at 62°C for 1 min, respectively. The relative fluorescence units (RFUs) from probes were detected in a CFX Connect Real-Time PCR Detection System (Bio-Rad, CA, USA) and used to generate allelic discrimination plots. Sugarcane lines were assigned as events with intended mutations provided if $RFU^{mut} > 200$ and $RFU^{mut}/RFU^{wt} > 0.1$ or if $200 > RFU^{mut} > 20$ and $RFU^{mut}/RFU^{wt} > 0.4$. RFU^{mut} and RFU^{wt} are mutant and wild-type allele RFU-values, respectively. Example discrimination plots in the HDR detection assay are displayed in **Supplementary Figure 3**.

TABLE 1 | Frequency of allele replacement with homology-directed repair (HDR) following five different treatments for biolistic delivery of genome editing components.

Treatment	Plasmid DNA amount (ng/kb and shot)					Shots	Lines PCR + for *npt*II	Lines after BS herbicide selection	Number of lines with intended mutations			
	Cas9	*npt*II	gRNA1	gRNA2	Repair template				W574L only (%)†	S653I only (%)†	W574L and S653I (%)†	W574L and/or S653I (%)†
1	1.5	1.5	1.5	0.0	6.0	10	84	28	8 (9.5%)	4 (4.8%)	1 (1.2%)	13 (15.5%)
2	1.5	1.5	0.0	1.5	6.0	10	91	27	7 (7.7%)	3 (3.3%)	6 (6.6%)	16 (17.6%)
3	1.5	1.5	1.5	1.5	6.0	10	97	25	5 (5.2%)	3 (3.1%)	1 (1.0%)	9 (9.3%)
4	0.5	0.5	0.5	0.5	2.0	10	102	32	4 (3.9%)	6 (5.9%)	3 (2.9%)	13 (12.7%)
5	1.5	1.5	1.5	1.5	12.0	10	93	34	1 (1.1%)	2 (2.2%)	0 (0%)	3 (3.2%)
Total							**467**	**146**	**25 (5.4%)**	**18 (3.9%)**	**11 (2.4%)**	**54 (11.6%)**

Number of lines with intended mutations W574L and/or S653I at target locations in sugarcane ALS are shown. Introduction of mutations with HDR were determined using a TaqMan® probe-based genotyping assay. Cas9, codon optimized CRISPR-associated gene 9; nptII, neomycin phosphotransferase II; gRNA, single guide RNA; ALS, acetolactate synthase; BS, bysparibac sodium.

† Frequencies are calculated as percentage of lines with intended mutations W574L and/or S653I in lines positive for nptII and prior selection with the herbicide bysparibac sodium (BS). All lines that regenerated from media containing BS were analyzed.

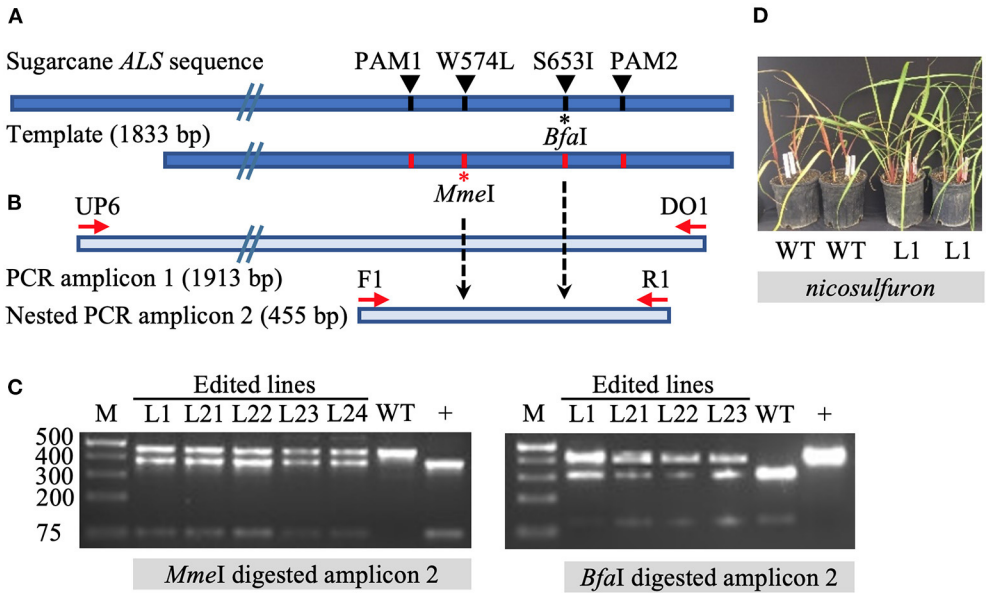


FIGURE 1 | Identification of gene targeting in sugarcane by restriction endonuclease assays and herbicide application. **(A)** Repair template compared to sugarcane acetolactate synthase (*ALS*) gene. The 1,833 bp template carries four nucleotide substitutions, one at the codon positions W574L (introducing the *Mme*I recognition site) and S653I (eliminating *Bfa*I recognition site), as well as two modified PAM sites (PAM1 and PAM2) (preventing template cleavage by Cas9 complexed with sgRNA1 or sgRNA2). Amino acid residue numbering in the sugarcane *ALS* gene follows the *Arabidopsis* nomenclature. **(B)** Schematic representation of primer annealing positions. Primers DO1 and UP6 were used to amplify a 1,913 bp fragment. Using primer UP6 prevented amplification of randomly inserted template by annealing outside the repair template. Primers F1 and R1 were used to generate a 455 bp nested PCR amplicon for analyses by restriction enzyme digestion. **(C)** Restriction-digestion pattern of nested PCR amplicons from wild-type (WT) and edited lines after digestion with *Mme*I or *Bfa*I. + indicates positive control that includes the *Mme*I site and lacks the *Bfa*I site. **(D)** Vegetatively propagated edited line L1 with multi allele conversion of W574L and S653I was tolerant to application of the herbicide nicosulfuron (Accent® DuPont) at 95 g ha^{-1}, in contrast to non-edited WT plants.

Amplicon Sequencing Using the Sanger Chain Termination Method

Cloned *ALS* alleles from transgenic plants with CRISPR/Cas9-mediated modifications were sequenced using the Sanger method. Briefly, long PCR amplicons (1,913 bp) generated with primers DO1 and UP6 (**Supplementary Table 3**) were cloned in pGEM®-T Easy Vector (Promega, WI, USA) and sequenced with M13 forward and reverse sequencing primers at Eurofins Genomics (Louisville, KY, USA). The sequences of 15–88 cloned PCR amplicons from *ALS* alleles of randomly selected events that were positive in TaqMan® assays for mutation sites W574L and/or S653I (**Table 2**) were also compared with the repair template to identify the presence of naturally occurring single nucleotide polymorphisms (SNPs; **Table 3**). This was done to evaluate the low probability of template switching between sugarcane *ALS* alleles and randomly inserted repair templates during PCR amplification and to distinguish between alleles.

RESULTS

In vitro Cleavage Assay of Target DNA With Cas9 Nuclease Confirms sgRNA Activity

The amino acid substitutions in the *ALS* gene known to confer herbicide tolerance include W574L and S653I. The expression cassettes for two sgRNAs were designed and custom-synthesized to induce DSBs by Cas9 near these two codons. The two sgRNAs were tested *in vitro* with the commercially available Cas9 protein. The electrophoretic separation of *in vitro* cleavage products confirmed that both sgRNAs effectively targeted Cas9 to induce DSBs at the targeted sites in the *ALS* gene (**Supplementary Figure 4**).

PCR Analyses of Sugarcane Events Following Co-transformation With Genome Editing Tools and Selectable Marker

The biolistic transfer of five combinations and ratios of unlinked selectable *npt*II marker and genome editing components was carried out as shown in **Table 1** and in **Supplementary Figures 1, 2**. Following selection with Geneticin 467 independent callus lines were confirmed as PCR-positive for *npt*II (**Table 1**). Following selection with BS, 146 independent events (2–3 independent events per shot) were regenerated to plants (**Table 1**).

Detection of Mutations by Altered Restriction Enzyme Digestion Patterns

The introduction of precision nucleotide substitutions at two codons associated with herbicide resistance and the corresponding PAM sites was targeted using a 1,833 bp double-stranded DNA repair template (**Figure 1A**,

TABLE 2 | Edited sugarcane lines with intended mutations W574L and S653I.

	TaqMan® probe-based HDR-detection assay				Amplicon sequencing with Sanger chain termination method‡						
	Mutation frequency at W574 (%)*	W574L†	Mutation frequency at S653 (%)*	S653I†	Total number of amplicons	Number of amplicons with intended mutations				Mutation frequency at W574 (%)**	Mutation frequency at S653 (%)**
						W574L	S653I	W574L and/or S653I	W574L and S653I		
L1	21.6	+	46.9	+	88	27	47	69	5	30.7	53.4
L2	55.0	+	0	–	24	13	0	13	0	54.2	0
L3	1.7	–	25.7	+	23	0	6	6	0	0	26.1
L4	49.5	+	56.0	+	17	6	6	11	1	35.3	35.3
L5	41.8	+	47.2	+	17	5	7	13	0	29.4	41.2
L6	71.2	+	54.0	+	16	8	5	11	2	50.0	31.3
L7	0	–	67.3	+	15	0	5	5	0	0	33.3
L9	68.9	+	64.5	+	67	32	38	62	8	47.8	56.7
L10	45.5	+	45.1	+	24	9	14	21	2	37.5	58.3
L11	24.3	+	30.7	+	24	7	10	17	0	29.2	41.7
L13	68.9	+	0	–	22	15	1	16	0	68.2	4.5
L15	52.7	+	51.7	+	21	9	11	19	1	42.9	52.4
L16	48.5	+	38.0	+	22	8	8	16	0	36.4	36.4
L17	39.1	+	9.1	+	22	8	3	10	1	36.4	13.6
L18	22.0	+	30.8	+	23	6	8	14	0	26.1	34.8
L19	85.4	+	0	–	23	16	2	16	2	69.6	8.7
Total					**448**	**169**	**171**	**319**	**22**		
Average	**43.51**		**35.44**							**37.11**	**32.98**
WT	0	–	0	–	4	0	0	0	0	0	0

Intended mutations were determined using a TaqMan® probe-based HDR-detection assay and amplicon sequencing, with the Sanger chain termination method.
†Presence (+) or absence (–) of intended mutations were determined according to allelic discrimination plots and relative fluorescent unit (RFU) values obtained using a TaqMan® probe-based HDR detection assay.
‡PCR amplicons 1,913 bp in length were generated from genomic DNA from edited sugarcane lines and sequenced bi-directionally after cloning.
**Frequency was calculated as percentage of fluorescence signal from probe targeting the mutations W574L or S653I to total fluorescence signal.*
***Frequency was calculated as percentage of amplicons with intended mutations W574L or S653I to total number of amplicons sequenced.*

Supplementary Figure 1). This template differed from endogenous sugarcane *ALS* in four nucleotides, including two PAM sites and two targeted codons W574L and S653I. The PCR amplification of the target *ALS* amplicon from putatively edited sugarcane lines excludes the randomly inserted template using primer UP6, annealing to the *ALS* coding region upstream of the donor template (**Figure 1A**). Targeted mutation W574L introduces an *Mme*I RE recognition site, and mutation S653I eliminates a *Bfa*I site (**Figure 1B**). In contrast to the wild-type control, edited lines displayed both digested and undigested *ALS* amplicons (**Figure 1C**).

Edited Lines Identified *via* TaqMan® Genotyping Assays

To allow for analyses of more than 100 edited events, a high-throughput TaqMan® genotyping assay was developed for the discrimination of DNA edits in the sugarcane ALS alleles as a result of HDR (**Table 2**, **Supplementary Figure 3**). For the HDR detection assays, the mutant probes were designed to recognize and bind sequences containing point mutations in the W574L or S653I codons and produce a signal. In the absence of DNA editing with HDR, the fluorescence signal was produced by the wild-type probe. The high level of polyploidy and variation in multi-allelic editing in sugarcane resulted in a range of fluorescent signals from different events. Representative allelic discrimination plots constructed with RFU from probes that detect wild-type or mutant alleles are presented in **Supplementary Figure 3**. These plots were used to identify edited plants with intended mutations. The number of sugarcane lines with intended mutations W574L and/or S653I are presented in **Table 1**. Frequencies of recovered lines with W574L and/or S653I mutations in treatments 1–5 ranged from 3.2 to 17.6% of the *npt*II positive callus lines prior selection with BS. The BS selection was not very stringent and more than 40% of the lines that regenerated on BS containing medium did not display edits. The frequency at which both mutations (W574L and S653I) were introduced into the same edited line ranged between 0 and 6.6% of the *npt*II positive callus lines prior selection with BS. The highest percentage of lines with both mutations was observed after treatment 2, where only sgRNA2 was delivered (**Table 1**). The lowest percentage of sugarcane lines with intended mutations was observed after treatment 5. No edited line carrying both mutations was recovered from this treatment. If a single sgRNA was used, editing was also observed at the distant target site 317 and 300

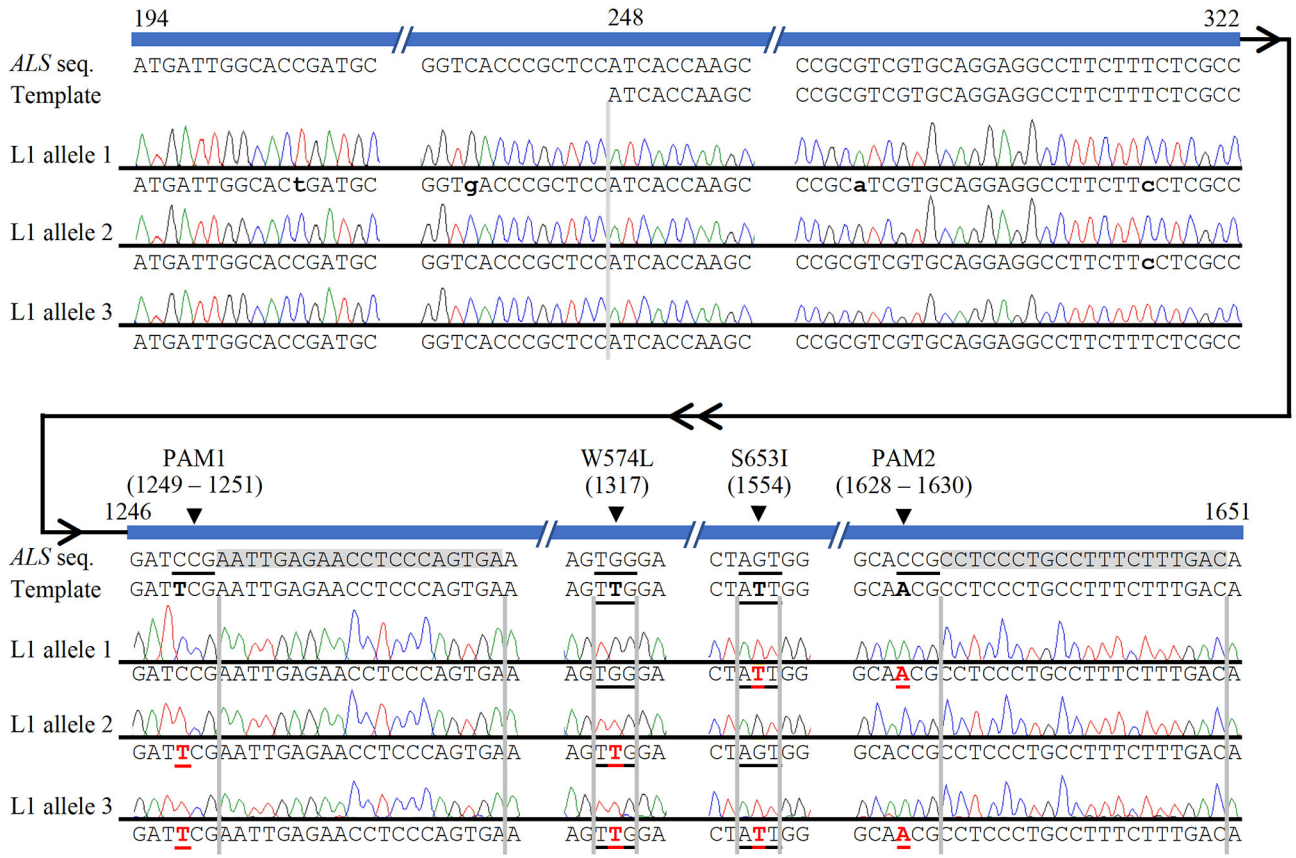


FIGURE 2 | Confirmation of multi-allelic gene targeting in sugarcane by Sanger sequencing of cloned PCR amplicons from *ALS* alleles. The first 247 bp (194–247 bp shown here) of each cloned *ALS* amplicons were used to eliminate possible amplification and sequencing of donor template. Naturally occurring non-target single nucleotide polymorphisms, which were used to identify alleles, are indicated with lower case black letters. HDR-mediated CRISPR/Cas9 gene targeting edits are shown in red bold font and underlined. The four targeted nucleotide substitutions are highlighted in the template in black bold font for contrast with the wild-type *ALS* sequence. The two sgRNA sequences are highlighted in gray. HDR-mediated nucleotide substitutions of S653I and PAM2 (allele 1), W574L and PAM1 (allele 2), or both herbicide resistance sites and both PAM sites (allele 3) are shown in edited line L1. Numbering follows the 1,913 bp PCR amplicon of *ALS* alleles.

nt from the targeted DSB for W574L and S653I, respectively (**Table 1**).

Sanger Sequencing of Cloned Amplicons Confirms Edited Lines Identified *via* the TaqMan® Genotyping Assay

The high level of polyploidy in sugarcane ($2n = 100 - 120$) results in multiple copies/alleles of the *ALS* gene. Sanger sequencing and analysis of naturally occurring SNPs of cloned PCR amplicons of the *ALS* gene from non-modified sugarcane cultivar CP 88 – 1,762 allowed us to distinguish 23 sequence variants that represent different copies or alleles (**Supplementary Table 4**). Sanger sequencing of the cloned PCR amplicons was used for validation of the TaqMan® genotyping assay and to analyze *ALS* alleles with targeted precision nucleotide substitutions and naturally occurring SNPs. Sanger sequencing of 16 independent lines confirmed GT at the W574L and/or S653I codons in all 16 lines analyzed, which were selected according to positive TaqMan® probe-based HDR-detection assays (**Table 2**). For most lines, the frequencies of targeted nucleotide substitutions in amplicon sequencing corresponded to frequencies indicated by the TaqMan® HDR-detection assay. Co-editing with targeted nucleotide substitutions in three unique alleles was confirmed in one line analyzed by Sanger sequencing (**Figure 2**). Several lines including L5, L11, L16, and L18 displayed each of the targeted nucleotide substitutions in different alleles (W574L in one allele and S653I in another allele, **Table 2**). Sequence comparison of the PCR amplicons from edited *ALS* alleles with the repair template revealed the presence of naturally occurring SNPs in these amplicons (**Table 3**). When a single sgRNA was used (Treatment 1, **Table 1**), Sanger sequencing also confirmed editing at the distant S653I target and its corresponding PAM site, 300 or 374 nt from the targeted DSB in two independent lines (**Supplementary Figure 5**). In these two events, the PAM and W574L target site proximal to the DSB were not edited.

Evaluation of Vegetative Progenies of Edited Sugarcane for Herbicide Resistance

Line L1, with multi-allelic precision nucleotide substitutions, was vegetatively propagated under greenhouse conditions to

TABLE 3 | Naturally occurring SNPs distinguishing the different *ALS* alleles from the repair template in edited sugarcane *ALS* alleles.

Nucleotide location	205	238	296	316	451	548	616	619	790	925	1,075	1,089	1,151	1,187	1,203	1,249 (PAM1)	1,317 (W574L)	1,483	1,513	1,554 (S653I)	1,628 (PAM2)	1,709	1,759	1,791	1,817	1,860
Line and edited allele number																										
L1 allele 1	T	G	A	C		C			G									C		T	A				G	
L1 allele 2				C		C	T	C	G	A						T	T						T	A	G	C
L1 allele 3																T	T			T	A	T			G	
L4 allele 1				C		C	T	C	G			T				T	T			T	A	T			G	
L6 allele 1											C					T	T			T	A	T			G	
L6 allele 2				C		C	T	C					A			T	T						T	A	G	C
L10 allele 1																T	T			T	A	T			G	
L10 allele 2				C		C	T	C	G	A						T	T						T	A	G	C
L10 allele 3				C	T	C	T	C	G							T	T						T	A	G	C
L10 allele 4	T	G	A	C	T	C						T		C	G			C	T	T	A				G	
L15 allele 1																T	T			T	A	T			G	
L17 allele 1	T	G	A	C		C			G							T	T			T	A	T			G	
L19 allele 1				C		C	T	C	G			T				T	T			T	A	T			G	

Intended edits introduced with HDR in ALS alleles are shown in red. Naturally occurring SNPs distinguishing the different ALS alleles from the repair template are shown in black. Analysis of edited ALS allele amplicons from seven selected lines (L1, L4, L6, L10, L15, L17, L19) and presence of naturally occurring SNPs in these amplicons suggests the absence of template switching by the DNA polymerase between sugarcane ALS alleles and the randomly inserted repair template during PCR amplification.

investigate herbicide resistance. Both wild-type plants and four edited sugarcane plants with multi-allele edits of the *ALS* gene were sprayed with nicosulfuron (Accent® DuPont) at the four-leaf stage with twice the labeled rate, 4 weeks after the stem cuttings were planted. One month after the application of herbicide, the wild-type sugarcane plants showed severe necrosis, in contrast to the edited plants, which had green leaves (**Figure 1D**). Then 2 months after herbicide application, all four edited plants were growing vigorously, whereas the wild-type sugarcane plants had died (data not shown).

DISCUSSION

Highly efficient and reproducible GT in sugarcane (*Saccharum* spp. hybrid) is reported here for the first time. Sugarcane has the highest genetic complexity of any crop, due to the interspecific origin of its highly polyploid ($2n = 100$–120) genome (D'Hont et al., 1996; Piperidis and D'Hont, 2020). Precision genome editing by HDR is an urgently needed tool for crop improvement (Huang and Puchta, 2019) and will support the targeted conversion of inferior to superior alleles. Template-mediated repair of DNA DSBs induced by the programmable nuclease CRISPR/Cas9 introduced multi-allelic precision nucleotide substitutions in the *ALS* gene of sugarcane, as confirmed by Sanger sequencing, conferring resistance to the herbicide nicosulfuron.

Precision genome editing of crops *via* HDR-mediated nucleotide substitutions has rarely been reported (Huang and Puchta, 2019). By contrast, loss-of-function mutations *via* the error-prone NHEJ repair pathway have been described in many crops (Yin et al., 2017; Huang and Puchta, 2019; Mao et al., 2019) including sugarcane (Jung and Altpeter, 2016; Kannan et al., 2018; Eid et al., 2021). Homology-directed repair frequency, as calculated as a percentage of GT events to all generated events, is typically in the range of 0.1 to a few percent (D'Halluin et al., 2013; Schiml et al., 2014; Svitashev et al., 2015; Endo et al., 2016; Hahn et al., 2018), which is more than an order of magnitude below the frequencies reported for NHEJ-mediated knock-outs of gene function (Yin et al., 2017; Hua et al., 2019; Huang and Puchta, 2019; Mao et al., 2019). This has been attributed to the superior efficiency of NHEJ for DNA repair in somatic cells (Puchta, 2005).

At least one targeted nucleotide substitution (W574L and/or S653I) in the *ALS* gene was detected in 54 of the 146 independent lines which regenerated from 467 transgenic callus lines in five different experiments. The HDR frequency for at least one targeted nucleotide substitution (W574L and/or S653I) in regenerated events ranged between 3.2 and 17.6% of the *npt*II positive lines prior BS selection. Plants with two resistance-conferring codons display a higher level of herbicide tolerance (Pozniak and Hucl, 2004; Shimizu et al., 2005). Two targeted nucleotide substitutions (W574L and S653I) were identified in 11 lines, accounting for 2.4% of the 467 lines prior BS selection.

Polymerase chain reaction artifacts from random head-to-tail inserts of the repair template and template switching of the DNA polymerase during PCR amplification have been described as potentially compromising the accurate evaluation of HDR frequency (Won and Dawid, 2017; Skryabin et al., 2020). The primers were designed to anneal to conserved regions for the amplification of all *ALS* alleles determined in this study. To prevent the PCR amplification of randomly integrated template DNA and selectively amplify the sugarcane *ALS* alleles, one of

the chosen primers annealed to the 5′ *ALS* coding sequence, which was not part of the template. In addition, the sequence comparison of the PCR amplicons from the edited *ALS* allele with the repair template indicates the presence of naturally occurring SNPs in the PCR amplicons. This suggests the absence of template switching by the DNA polymerase between sugarcane *ALS* alleles and the randomly inserted repair template during PCR amplification. Evaluation of more than 400 bidirectional Sanger sequencing reads did not reveal any chimeric PCR amplicons representing random head-to-tail insertions of the repair template.

Double-strand breaks enhance recombination frequencies between homologous plant chromosomes or between sequence repeats (Puchta, 1998; Hayut et al., 2017; Taagen et al., 2020; Zhao et al., 2021). In diploid tomato, allele-dependent HDR reached 14% of all detectable DSB repair events (Hayut et al., 2017). The high polyploidy level in sugarcane may have contributed to the frequent occurrence of the HDR reported here. Twenty-three allelic variants of the ALS gene were identified by SNP analysis following Sanger sequencing of cloned PCR amplicons. The large number of target copies per cell may increase the chance that homologous recombination will occur. This provides a unique opportunity for crop improvement and genetic studies using GT.

Homology-directed repair efficiency in plants is highly correlated with the amount, size, and type of the repair template delivered to the target cells (Huang and Puchta, 2019). Biolistic gene transfer offers the ability to control the quantity of the delivered DNA during the particle coating reaction, influencing the frequency of GT and the number of transgene copies that are randomly inserted into the genome (Sandhu and Altpeter, 2008; Lowe et al., 2009; Sun et al., 2016). Elevating the molar ratio of DNA repair template four times over the other genome-editing components resulted in the highest HDR frequencies in this study. Recently, the use of geminiviral replicons as HDR vectors has increased the copy number of the template DNA delivered by *Agrobacterium tumefaciens* to target crops (Baltes et al., 2014; Butler et al., 2016; Wang et al., 2017). Elevating the template quantity with geminiviral replicons increased the HDR frequency to 34.5% in tetraploid potato (Butler et al., 2016) and to 25% in diploid tomato (Dahan-Meir et al., 2018). For HDR in sugarcane, the optimum template amount was 6 ng per kb of delivered repair template DNA. Doubling the delivered repair template quantity resulted in much lower editing efficiencies. Large numbers of randomly inserted repair template copies can lead to false-positive HDR results (Lawrenson et al., 2021). Our results from comparing the delivery of different repair template quantities suggest that excessive amounts of exogenously supplied DNA can have a negative impact on the repair process and that this did not inflate the HDR results in the developed assays.

The design of the repair template with long homology arms (Zhang et al., 2013) may also have contributed to the high HDR frequency (11.6% of the *npt*II positive callus lines and 36.9% of the lines that regenerated from BS medium on average of five experiments) in sugarcane reported here. Different repair outcomes with both W574L and S653I substitutions or one to four targeted nucleotide substitutions in one or multiple alleles were observed. Several lines displayed each of the targeted nucleotide substitutions in different alleles. The vast majority of targeted nucleotide substitutions at the W574L or S653I site occurred simultaneously with the corresponding proximal silent mutation of the specific PAM site. This supports the conclusion that the portion of the donor that is used in strand invasion, Holliday junction formation, and branch migration determines the outcome of HDR (Puchta and Fauser, 2015). Similarly, in rice, three haplotypes of HDR were observed at the S627 locus after a double-stranded DNA donor fragment was introduced (Sun et al., 2016). Sanger sequencing revealed multiplexed co-editing of up to three of the unique, cloned sugarcane *ALS* amplicons displaying the targeted nucleotide substitutions.

Editing was observed at both codon 574 and 653 at distances of 63 or 80 nt from the targeted DSBs. If a single sgRNA was used, edits were also observed 300 or 317 nt from the targeted DSB. Homology-directed repair repair frequencies are typically negatively correlated with distance to the DSB site (Baur et al., 1990; Paquet et al., 2016). The limiting step for HDR-mediated edits at more distant sites is likely 5′- to 3′-end resection. The factors involved in eukaryotic end resection have been thoroughly investigated in *Saccharomyces cerevisiae*, revealing that the number of DSBs per cell plays an important role in the activation of end resection. Four DSBs in yeast cells induced end resection of at least 300 nt in the G1 phase in 22% of the induced DSBs. In contrast, a single DSB resulted in only 8% of its induced DSBs, displaying an at least 300 nt end resection (Zierhut and Diffley, 2008). Single nucleotide polymorphisms analyses of *ALS* amplicons from non-edited sugarcane target cultivar CP 88–1762 suggest the presence of 23 *ALS* copies/alleles. The high number of *ALS* copies will result in an elevated frequency of DSBs per cell, potentially leading to more efficient end resection. This should enable both high HDR frequencies and the generation of HDR events distant from the cut site. However, co-delivering both sgRNAs instead of a single one did not elevate the frequency of the targeted nucleotide substitutions. Simultaneous cleavage at two sgRNA target sites can result in large deletions, conferring loss of function of the allele. This may have a deleterious effect on plant regeneration due to the essential role of *ALS* in the synthesis of branched chain amino acids in plants (Miflin and Cave, 1972). In addition, the use of two sgRNAs may elevate the frequency of NHEJ repair which may lead to a lower number of target sites that can be repaired using the donor template.

Surprisingly, the editing frequency was higher at the W574L codon than at the S653I codon, regardless of whether a sgRNA that cleaves 63 or 317 nt away from it was used. The 5′ homology arm of the repair template was more than twice as long as the 3′ homology arm and may have contributed to this outcome. In addition, herbicide selection in tissue culture and the major contribution of the W574L mutation to the level of herbicide resistance (Tan et al., 2006) may have also created a bias for the recovery of events with the W574L mutation.

Repair outcomes with the nucleotide edit only at a site that is further away from the DSB, leaving the closer one as WT, were also observed. This could be explained by the use of a fragmented template during the HDR-mediated repair

process. DNA shearing is prominent with particle bombardment, generating a range of DNA fragments (Banakar et al., 2019). A fragmented repair template that is suitable for HDR-mediated repair would have a single homology arm. Single homology arms direct HDR in mammalian cells and are more prone to local repair with template switching (Basiri et al., 2017; Paix et al., 2017; Suzuki et al., 2019). This may favor the use of the NHEJ pathway to repair the gap on the side with no homology arm. The coupling of homologous and non-homologous repair mechanisms to preserve genomic integrity has been documented in mammalian cells (Richardson and Jasin, 2000).

Further improvements to multiplex precision nucleotide substitutions in sugarcane may be enabled by manipulating the competing DNA repair pathways (Even-Faitelson et al., 2011; Qi et al., 2013; Endo et al., 2016) or increasing HDR frequency by tethering the DNA repair template to the genome editing tool (Aird et al., 2018). Base editors and prime editors have emerged as alternative strategies to template-mediated HDR for precision nucleotide substitutions (Huang and Puchta, 2021). Relative to the template-mediated HDR employed in this study, base editors are limited by their narrow target range editing window of approximately 10 bp at the target site (Rees and Liu, 2018). Similarly, prime editors are limited by their template size (10 and 20 bp). In addition, the efficiency of prime editors varies greatly between target sites and the frequent generation of unintended indels is a concern (Huang and Puchta, 2021).

Multi-allelic precision nucleotide substitutions in sugarcane conferred herbicide tolerance. All progeny plants displayed herbicide resistance in their entire foliage suggesting the absence of chimerism. Introduction of herbicide resistance by GT will be a very effective tool for selecting events following the co-editing of multiple target genes. By comparison, conventional sugarcane breeding is complicated by poor male fertility, difficulty synchronizing the flowering time of the parental lines, poor seed germination, and the disruptive effects of meiotic recombination on the predictable performance of the progenies (Scortecci et al., 2012). This established protocol for targeted nucleotide substitutions allows fast and efficient introduction of the selected gene variants into elite sugarcane cultivars without crossing and associated linkage drag in support of crop improvement and genetic studies.

CONCLUSIONS

We report efficient and reproducible GT in sugarcane, enabling precise co-editing of multiple alleles *via* template- and homology-mediated repair of DNA DSBs induced by the CRISPR/Cas9 programmable nuclease. This work will enable crop improvement by modifying inferior alleles through multiplexed GT.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation. Sanger sequencing reads of cloned PCR amplicons of the ALS gene were submitted to NCBI with GenBank accession IDs MZ268741-MZ268818 (**Supplementary Table 5**).

AUTHOR CONTRIBUTIONS

FA conceived, designed, and managed the research project. MO carried out the construction of vectors carrying sgRNAs and confirmed their activity *via in vitro* cleavage assay, carried out sugarcane tissue culture and transformation, confirmed regenerated transgenic sugarcane lines by PCR, cloned PCR amplicons from edited sugarcane lines, and analyzed them using the restriction digestion assay, Taqman® genotyping, and Sanger sequencing. AA coordinated and carried out sugarcane tissue culture and transformation and contributed to the molecular characterization of genome-edited plants. RK contributed the template design and molecular characterization of genome-edited plants. AM developed the herbicide application protocol. MO and FA wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research) under Award Number DE-SC0018420. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy. This work was also supported by the USDA National Institute of Food and Agriculture, Hatch project 1020425. AM was supported by a fellowship from the National Council for Scientific and Technological Development (CNPq), Brazil.

ACKNOWLEDGMENTS

The authors thank Dr. Hardev Sandhu, Everglades Research and Education Center, UF-IFAS, for providing sugarcane tops and Sun Gro Horticulture, Apopka, FL, for the donation of potting mix.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgeed.2021.673566/full#supplementary-material

Supplementary Figure 1 | Plasmids carrying various genome-editing elements. Expression of codon-optimized *Cas9* and *npt*II were under transcriptional control of CaMV35S promoter. An intron from HSP70 was included between the promoter and coding sequence. Termination signals from AtHSP and CaMV35S were located downstream of co*Cas9* and *npt*II, respectively. sgRNA were placed under the control of U6 promoter from *Oryza sativa*. Donor template with nucleotide modifications to introduce W574L and S653I, and two modified PAM sites (PAM1 and PAM2) to prevent cleavage of the template by sgRNA1 or sgRNA2, were designed according to sugarcane *ALS* sequence. Homology arms are indicated with double-headed arrows. Vector components are not given to scale. *npt*II, neomycin phosphotransferase II; CRISPR, clustered regularly interspaced short palindromic repeats; *coCas9*, codon optimized CRISPR-associated gene 9; RNA;

sgRNA, single guide RNA; *ALS*, acetolactate synthase; $P_{CaMV35S}$, promoter of cauliflower mosaic virus 35S RNA; I_{HSP70}, intron of 70 kDa heat-shock protein; $T_{CaMV35S}$, terminator of cauliflower mosaic virus 35S RNA; T_{AtHSP}, terminator of *Arabidopsis thaliana* heat-shock protein; NLS1, nuclear localization signal from SV40; NLS2, nuclear localization signal from nucleoplasmin; P_{U6}, U6 promoter from *Oryza sativa*; PAM, protospacer adjacent motif; kb, kilobases.

Supplementary Figure 2 | Sugarcane tissue culture, plant regeneration, and genetic transformation. **(A)** General outline of sugarcane tissue culture and genetic transformation. **(B)** Calli placed at the center of a Petri dish for bombardment. **(C)** Selection of transgenic calli expressing *npt*II on geneticin containing culture medium. **(D)** Regeneration of plantlets on culture medium supplemented with the herbicide byspyridbac sodium.

Supplementary Figure 3 | High-throughput TaqMan® probe-based genotyping assay. **(A)** Schematic drawing of the TaqMan® probe-based genotyping strategy to detect HDR events. Large PCR amplicons (1,913 bp) were generated with primers DO1 and UP6, the latter located outside of the template sequence to prevent amplification of randomly inserted templates. Fluorescent labeled TaqMan® probes were designed to detect wild-type and mutant alleles (fluorophore VIC® for wild-type and fluorophore FAM for mutant allele) at both amino acid positions 574 and 653 in two HDR-detection assays. **(B)** Expected allelic discrimination plot in a probe-based HDR-detection assay. **(C,D)** Allelic discrimination plots constructed with relative fluorescent units from probes detecting wild-type or mutant alleles were used to identify edited plants with intended mutations. Sugarcane lines with targeted mutations W574L or S653I are indicated with green triangles whereas lines showing only wild-type allele signals are indicated with blue squares. Positive control plasmids (+) are shown with orange circles. No template control did not produce fluorescent signals.

Supplementary Figure 4 | Electrophoretic separation of cleavage products after *in vitro* Cas9 nuclease assay. Cleavage with commercially available Cas9 protein confirmed that both sgRNAs ($sgRNA_{574}$ and $sgRNA_{653}$) effectively targeted Cas9 to induce double strand breaks at the target sites in the *ALS* gene. M, marker.

Supplementary Figure 5 | Targeted nucleotide substitutions in the sugarcane *ALS* gene in lines L3 and L7, 300 and 374 nt from the DNA double strand break, respectively, targeted with a single gRNA as determined by Sanger sequencing of cloned PCR amplicons. Homology-directed repair-mediated CRISPR/Cas9 gene targeting edits are shown in red bold font and underlined. The four targeted nucleotide substitutions were highlighted in the template in black bold font compared to the wild-type ALS sequence. Naturally occurring non-target single nucleotide polymorphisms (SNPs), used to identify alleles, are indicated with lowercase black letters. The single sgRNA sequence used in this treatment (Treatment 1, **Table 2**) is highlighted in gray. Homology-directed repair-mediated nucleotide substitutions of S653I and PAM2 are shown in edited lines L3 and L7. Numbering follows the 1,913 bp PCR amplicon of *ALS* alleles.

Supplementary Table 1 | Sequence of expression cassettes carrying gRNAs.

Supplementary Table 2 | Sequence of sugarcane *ALS* gene and repair template.

Supplementary Table 3 | Primer and probe sequences.

Supplementary Table 4 | Sugarcane wild-type ALS alleles from CP 88–1762 determined with amplicon sequencing using the Sanger chain termination method.

Supplementary Table 5 | GenBank accession IDs for sequence reads of cloned PCR amplicons of the ALS gene with intended mutations W574L and/or S653I from gene-edited lines.

REFERENCES

Aird, E. J., Lovendahl, K. N., St Martin, A., Harris, R. S., and Gordon, W. R. (2018). Increasing Cas9-mediated homology-directed repair efficiency through covalent tethering of DNA repair template. *Commun. Biol.* 1:54. doi: 10.1038/s42003-018-0054-2

Baltes, N. J., Gil-Humanes, J., Cermak, T., Atkins, P. A., and Voytas, D. F. (2014). DNA replicons for plant genome engineering. *Plant Cell* 26, 151–163. doi: 10.1105/tpc.113.119792

Banakar, R., Eggenberger, A. L., Lee, K., Wright, D. A., Murugan, K., Zarecor, S, et al. (2019). High-frequency random DNA insertions upon co-delivery of CRISPR-Cas9 ribonucleoprotein and selectable marker plasmid in rice. *Sci. Rep.* 9:19902. doi: 10.1038/s41598-019-55681-y

Basiri, M., Behmanesh, M., Tahamtani, Y., Khalooghi, K., Moradmand, A., and Baharvand, H. (2017). The convenience of single homology arm donor DNA and CRISPR/Cas9-nickase for targeted insertion of long DNA fragment. *Cell J.* 8, 532–539. doi: 10.22074/cellj.2016.4719

Baur, M., Potrykus, I., and Paszkowski, J. (1990). Intermolecular homologous recombination in plants. *Mol. Cell Biol.* 10, 492–500. doi: 10.1128/mcb.10.2.492

Bortesi, L., and Fischer, R. (2015). The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnol. Adv.* 33, 41–52. doi: 10.1016/j.biotechadv.2014.12.006

Brooks, C., Nekrasov, V., Lippman, Z. B., and Van Eck, J. (2014). Efficient gene editing in tomato in the first generation using the clustered regularly interspaced short palindromic repeats/CRISPR-associated9 system. *Plant Physiol.* 166, 1292–1297. doi: 10.1104/pp.114.247577

Butler, N. M., Baltes, N. J., Voytas, D. F., and Douches, D. S. (2016). Geminivirus-mediated genome editing in potato (*Solanum tuberosum* L.) using sequence-specific nucleases. *Front. Plant Sci.* 7:1045. doi: 10.3389/fpls.2016.01045

Byrt, C. S., Grof, C. P. L., and Furbank, R. T. (2011). C-4 Plants as biofuel feedstocks: optimising biomass production and feedstock quality from a lignocellulosic perspective. *J. Integr. Plant Biol.* 53, 120–135. doi: 10.1111/j.1744-7909.2010.01023.x

Chen, K., Wang, Y., Zhang, R., Zhang, H., and Gao, C. (2019). CRISPR/Cas genome editing and precision plant breeding in agriculture. *Annu. Rev. Plant Biol.* 70, 667–697. doi: 10.1146/annurev-arplant-050718-100049

Dahan-Meir, T., Filler-Hayut, S., Melamed-Bessudo, C., Bocobza, S., Czosnek, H., Aharoni, A., et al. (2018). Efficient in planta gene targeting in tomato using geminiviral replicons and the CRISPR/Cas9 system. *Plant J.* 95, 5–16. doi: 10.1111/tpj.13932

Dermawan, H., Karan, R., Jung, J. H., Zhao, Y., Parajuli, S., Sanahuja, G., et al. (2016). Development of an intragenic gene transfer and selection protocol for sugarcane resulting in resistance to acetolactate synthase-inhibiting herbicide. *Plant Cell Tiss. Org. Cult.* 126, 459–468. doi: 10.1007/s11240-016-1014-5

Devine, M., Duke, S. O., and Fedtke, C. (1992). *Physiology of Herbicide Action*. Englewood Cliffs, NJ: PTR Prentice Hall.

D'Halluin, K., Vanderstraeten, C., Van Hulle, J., Rosolowska, J., Van Den Brande, I., Pennewaert, A., et al. (2013). Targeted molecular trait stacking in cotton through targeted double-strand break induction. *Plant Biotechnol. J.* 11, 933–941. doi: 10.1111/pbi.12085

D'Hont, A., Grivet, L., Feldmann, P., Rao, S., Berding, N., and Glaszmann, J. C. (1996). Characterisation of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. *Mol. Gen. Genet.* 250, 405-413. doi: 10.1007/BF02174028

Duggleby, R. G., McCourt, J. A., and Guddat, L. W. (2008). Structure and mechanism of inhibition of plant acetohydroxyacid synthase. *Plant Physiol. Biochem.* 46, 309–324. doi: 10.1016/j.plaphy.2007.12.004

Eid, A., Mohan, C., Sanchez, S., Wang, D., and Altpeter, F. (2021). Multiallelic, targeted mutagenesis of magnesium chelatase with CRISPR-Cas9 provides a rapidly scorable phenotype in highly polyploid sugarcane. *Front. Genome Ed.* 3:654996. doi: 10.3389/fgeed.2021.654996

Endo, M., Mikami, M., and Toki, S. (2016). Biallelic gene targeting in rice. *Plant Physiol.* 170, 667–677. doi: 10.1104/pp.15.01663

Even-Faitelson, L., Samach, A., Melamed-Bessudo, C., Avivi-Ragolsky, N., and Levy, A. A. (2011). Localized egg-cell expression of effector proteins for targeted modification of the Arabidopsis genome. *Plant J.* 68, 929–937. doi: 10.1111/j.1365-313X.2011.04741.x

Garsmeur, O., Droc, G., Antonise, R., Grimwood, J., Potier, B., Aitken, K., et al. (2018). A mosaic monoploid reference sequence for the highly complex genome of sugarcane. *Nat. Commun.* 9:2638. doi: 10.1038/s41467-018-05051-5

Hahn, F., Eisenhut, M., Mantegazza, O., and Weber, A. P. M. (2018). Homology-directed repair of a defective glabrous gene in Arabidopsis with

Cas9-based gene targeting. *Front. Plant Sci.* 9:424. doi: 10.3389/fpls.2018.00424

Hayut, F. S., Melamed Bessudo, C., and Levy, A. (2017). Targeted recombination between homologous chromosomes for precise breeding in tomato. *Nat. Commun.* 8:15605. doi: 10.1038/ncomms15605

Horsley, T., and Zhou, M. (2013). Effect of photoperiod treatments on pollen viability and flowering at the South African Sugarcane Research Institute. *Proc. South Afr. Sug. Techn. Assoc.* 86, 286–290. doi: 10.37580/JSR.2019.2.9.138-149

Hua, K., Zhang, J. S., Botella, J. R., Ma, C., Kong, F., Liu, B., et al. (2019). Perspectives on the application of genome-editing technologies in crop breeding. *Mol. Plant* 12, 1047–1059. doi: 10.1016/j.molp.2019.06.009

Huang, T. K., and Puchta, H. (2019). CRISPR/Cas-mediated gene targeting in plants: finally a turn for the better for homologous recombination. *Plant Cell Rep.* 38, 443–453. doi: 10.1007/s00299-019-02379-0

Huang, T. K., and Puchta, H. (2021). Novel CRISPR/Cas applications in plants: from prime editing to chromosome engineering. *Transgenic Res.* doi: 10.1007/s11248-021-00238-x. [Epub ahead of print].

Jung, J. H., and Altpeter, F. (2016). TALEN mediated targeted mutagenesis of the caffeic acid O-methyltransferase in highly polyploid sugarcane improves cell wall composition for production of bioethanol. *Plant Mol. Biol.* 92, 131–142. doi: 10.1007/s11103-016-0499-y

Kannan, B., Jung, J. H., Moxley, G. W., Lee, S. M., and Altpeter, F. (2018). TALEN-mediated targeted mutagenesis of more than 100 COMT copies/alleles in highly polyploid sugarcane improves saccharification efficiency without compromising biomass yield. *Plant Biotechnol. J.* 16, 856–866. doi: 10.1111/pbi.12833

Lawrenson, T., Hinchliffe, A., Clarke, M., Morgan, Y., and Harwood, W. (2021). In-planta gene targeting in barley using Cas9, with and without geminiviral replicons. *Front. Genome Ed.* doi: 10.3389/fgeed.2021.663380. [Epub ahead of print].

Li, D., Barclay, I., Jose, K., Stefanova, K., and Appels, R. (2008). A mutation at the Ala122 position of acetohydroxyacid synthase (AHAS) located on chromosome 6D of wheat: improved resistance to imidazolinone and a faster assay for marker assisted selection. *Mol. Breed.* 22, 217–225. doi: 10.1007/s11032-008-9168-4

Lowe, B. A., Prakash, N. S., Way, M., Mann, M. T., Spencer, T. M., and Boddupalli, R. S. (2009). Enhanced single copy integration events in corn *via* particle bombardment using low quantities of DNA. *Transgenic Res.* 18, 831–840. doi: 10.1007/s11248-009-9265-0

Mao, Y., Botella, J. R., Liu, Y., and Zhu, J.-K. (2019). Gene editing in plants: progress and challenges. *Natl. Sci. Rev.* 6, 421–437. doi: 10.1093/nsr/nwz005

Merotto, A., Jasieniuk, M., Osuna, M. D., Vidotto, F., Ferrero, A., and Fischer, A. J. (2009). Cross-resistance to herbicides of five ALS-inhibiting groups and sequencing of the ALS gene in *Cyperus difformis* L. *J. Agric. Food Chem.* 57, 1389–1398. doi: 10.1021/jf802758c

Miflin, B. J., and Cave, P. R. (1972). The control of leucine, isoleucine, and valine biosynthesis in a range of higher plants. *J. Exp. Bot.* 23, 511–516. doi: 10.1093/jxb/23.2.511

Murray, M. G., and Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8, 4321–4325. doi: 10.1093/nar/8.19.4321

Paix, A., Folkmann, A., Goldman, D. H., Kulaga, H., Grzelak, M. J., Rasoloson, D., et al. (2017). Precision genome editing using synthesis-dependent repair of Cas9-induced DNA breaks. *Proc. Natl. Acad. Sci. U.S.A.* 114, E10745–E10754. doi: 10.1073/pnas.1711979114

Paquet, D., Kwart, D., Chen, A., Sproul, A., Jacob, S., Teo, S., et al. (2016). Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* 533, 125–129. doi: 10.1038/nature17664

Piperidis, N., and D'Hont, A. (2020). Sugarcane genome architecture decrypted with chromosome-specific oligo probes. *Plant J.* 103, 2039–2051. doi: 10.1111/tpj.14881

Powles, S. B., and Yu, Q. (2010). Evolution in action: plants resistant to herbicides. *Annu. Rev. Plant Biol.* 61, 317–347. doi: 10.1146/annurev-arplant-042809-112119

Pozniak, C. J., and Hucl, P. J. (2004). Genetic analysis of imidazolinone resistance in mutation-derived lines of common wheat. *Crop Sci.* 44, 23–30. doi: 10.2135/cropsci2004.2300

Puchta, H. (1998). Repair of genomic double-strand breaks in somatic plant cells by one-sided invasion of homologous sequences. *Plant J.* 13, 331–339. doi: 10.1046/j.1365-313X.1998.00035.x

Puchta, H. (2005). The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *J. Exp. Bot.* 56, 1–14. doi: 10.1093/jxb/eri025

Puchta, H., and Fauser, F. (2015). "Double-strand break repair and its application to genome engineering in plants," in *Advances in New Technology for Targeted Modification of Plant Genomes,* eds F. Zhang, H. Puchta, and J. G. Thomson (New York, NY: Springer New York), 1–20.

Qi, Y. P., Zhang, Y., Zhang, F., Baller, J. A., Cleland, S. C., Ryu, Y., et al. (2013). Increasing frequencies of site-specific mutagenesis and gene targeting in Arabidopsis by manipulating DNA repair pathways. *Genome Res.* 23, 547–554. doi: 10.1101/gr.145557.112

Que, Q., Chen, Z., Kelliher, T., Skibbe, D., Dong, S., and Chilton, M. D. (2019). Plant DNA repair pathways and their applications in genome engineering. *Methods Mol. Biol.* 1917, 3–24. doi: 10.1007/978-1-4939-8991-1_1

Rees, H. A., and Liu, D. R. (2018). Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* 19, 770–788. doi: 10.1038/s41576-018-0059-1

Richardson, C., and Jasin, M. (2000). Coupled homologous and nonhomologous repair of a double-strand break preserves genomic integrity in mammalian cells. *Mol. Cell Biol.* 20, 9068–9075. doi: 10.1128/mcb.20.23.9068-9075.2000

Rodríguez-Suárez, C., Ramírez, M. C., Martínez, C., Nadal, S., Martín, A., and Atienza, S. G. (2009). Selection and molecular characterization of imidazolinone resistant mutation-derived lines of Tritordeum HT621. *Mol. Breed.* 23, 565–572. doi: 10.1007/s11032-009-9256-0

Sandhu, S., and Altpeter, F. (2008). Co-integration, co-expression and inheritance of unlinked minimal transgene expression cassettes in an apomictic turf and forage grass (*Paspalum notatum* Flugge). *Plant Cell Rep.* 27, 1755–1765. doi: 10.1007/s00299-008-0599-5

Schiml, S., Fauser, F., and Puchta, H. (2014). The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. *Plant J.* 80, 1139–1150. doi: 10.1111/tpj.12704

Scortecci, K. C., Creste, S., and Calsa, T. Jr., Xavier, M. A., Landell, M. G. A., Figueira, A. et al. (2012). "Challenges, opportunities and recent advances in sugarcane breeding," in *Plant Breeding*, ed I. Abdurakhmonov (London: InTech), 267–296.

Sedeek, K. E. M., Mahas, A., and Mahfouz, M. (2019). Plant genome engineering for targeted improvement of crop traits. *Front. Plant Sci.* 10:114. doi: 10.3389/fpls.2019.00114

Shaner, D. L., and O'Connor, S. L. (1991). *The Imidazolinone Herbicides*. Boca Raton, FL: CRC Press.

Shimizu, T., Kaku, K., Kawai, K., Miyazawa, T., and Tanaka, Y. (2005). "Molecular characterization of acetolactate synthase in resistant weeds and crops," in *Environmental Fate and Safety Management of Agrochemicals,* Vol. 899 (Washington, DC: American Chemical Society), 255–271.

Shukla, V. K., Doyon, Y., Miller, J. C., DeKelver, R. C., Moehle, E. A., Worden, S. E., et al. (2009). Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature* 459, 437–441. doi: 10.1038/nature07992

Skryabin, B. V., Kummerfeld, D. M., Gubar, L., Seeger, B., Kaiser, H., Stegemann, A., et al. (2020). Pervasive head-to-tail insertions of DNA templates mask desired CRISPR-Cas9-mediated genome editing events. *Sci. Adv.* 6:eaax2941. doi: 10.1126/sciadv.aax2941

Smith, J. K., Schloss, J. V., and Mazur, B. J. (1989). Functional expression of plant acetolactate synthase genes in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 86, 4179–4183. doi: 10.1073/pnas.86.11.4179

Sun, Y., Zhang, X., Wu, C., He, Y., Ma, Y., Hou, H., et al. (2016). Engineering herbicide-resistant rice plants through CRISPR/Cas9-mediated homologous recombination of acetolactate synthase. *Mol. Plant* 9, 628–631. doi: 10.1016/j.molp.2016.01.001

Suzuki, K., Yamamoto, M., Hernandez-Benitez, R., Li, Z., Wei, C., Soligalla, R. D., et al. (2019). Precise *in vivo* genome editing *via* single homology arm donor mediated intron-targeting gene integration for genetic disease correction. *Cell Res.* 29, 804–819. doi: 10.1038/s41422-019-0213-0

Svitashev, S., Schwartz, C., Lenderts, B., Young, J. K., and Cigan, A. M. (2016). Genome editing in maize directed by CRISPR-Cas9 ribonucleoprotein complexes. *Nat. Commun.* 7:13274. doi: 10.1038/ncomms13274

Svitashev, S., Young, J. K., Schwartz, C., Gao, H. R., Falco, S. C., and Cigan, A. M. (2015). Targeted mutagenesis, precise gene editing, and site-specific gene insertion in maize using Cas9 and guide RNA. *Plant Physiol.* 169, 931–945. doi: 10.1104/pp.15.00793

Taagen, E., Bogdanove, A. J., and Sorrells, M. E. (2020). Counting on crossovers: controlled recombination for plant breeding. *Trends Plant Sci.* 25, 455–465. doi: 10.1016/j.tplants.2019.12.017

Tan, S., Evans, R., and Singh, B. (2006). Herbicidal inhibitors of amino acid biosynthesis and herbicide-tolerant crops. *Amino Acids* 30, 195–204. doi: 10.1007/s00726-005-0254-1

Taparia, Y., Gallo, M., and Altpeter, F. (2012). Comparison of direct and indirect embryogenesis protocols, biolistic gene transfer and selection parameters for efficient genetic transformation of sugarcane. *Plant Cell Tiss. Org. Cult.* 111, 131–141. doi: 10.1007/s11240-012-0177-y

Voytas, D. F., and Gao, C. (2014). Precision genome engineering and agriculture: opportunities and regulatory challenges. *PLoS Biol.* 12:e1001877. doi: 10.1371/journal.pbio.1001877

Wang, M. G., Lu, Y. M., Botella, J. R., Mao, Y. F., Hua, K., and Zhu, J. K. (2017). Gene targeting by homology-directed repair in rice using a geminivirus-based CRISPR/Cas9 system. *Mol. Plant* 10, 1007–1010. doi: 10.1016/j.molp.2017.03.002

Weeks, D. P. (2017). Gene editing in polyploid crops: wheat, camelina, canola, potato, cotton, peanut, sugar cane, and citrus. *Prog. Mol. Biol. Transl. Sci.* 149, 65–80. doi: 10.1016/bs.pmbts.2017.05.002

Won, M., and Dawid, I. B. (2017). PCR artifact in testing for homologous recombination in genomic editing in zebrafish. *PLoS ONE* 12:e0172802. doi: 10.1371/journal.pone.0172802

Wu, H., Awan, F. S., Vilarinho, A., Zeng, Q., Kannan, B., Phipps, T., et al. (2015). Transgene integration complexity and expression stability following biolistic or Agrobacterium-mediated transformation of sugarcane. *In Vitro Cell. Dev. Biol. Plant* 51, 603–611. doi: 10.1007/s11627-015-9710-0

Yin, K., Gao, C., and Qiu, J. L. (2017). Progress and prospects in plant genome editing. *Nat. Plants* 3:17107. doi: 10.1038/nplants.2017.107

Zhang, Y., Massel, K., Godwin, I. D., and Gao, C. X. (2018). Applications and potential of genome editing in crop improvement. *Genome Biol.* 19:210. doi: 10.1186/s13059-018-1586-y

Zhang, Y., Zhang, F., Li, X., Baller, J. A., Qi, Y., Starker, C. G., et al. (2013). Transcription activator-like effector nucleases enable efficient plant genome engineering. *Plant Physiol.* 161, 20–27. doi: 10.1104/pp.112.205179

Zhao, Y., Karan, R., and Altpeter, F. (2021). Error-free recombination in sugarcane mediated by only 30 nucleotides of homology and CRISPR/Cas9 induced DNA breaks or Cre-recombinase. *Biotechnol. J.* 12:2000650. doi: 10.1002/biot.202000650

Zierhut, C., and Diffley, J. F. (2008). Break dosage, cell cycle stage and DNA replication influence DNA double strand break response. *EMBO J.* 27, 1875–1885. doi: 10.1038/emboj.2008.111

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Oz, Altpeter, Karan, Merotto and Altpeter. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership