CRISPR-CAS IN AGRICULTURE: OPPORTUNITIES AND CHALLENGES

EDITED BY: Sandeep Kumar, Hiroshi Ezura, Vladimir Nekrasov and Linda Ann Rymarquis

PUBLISHED IN: Frontiers in Plant Science

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88966-815-1
DOI 10.3389/978-2-88966-815-1

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

CRISPR-CAS IN AGRICULTURE: OPPORTUNITIES AND CHALLENGES

Topic Editors:
Sandeep Kumar, Corteva Agriscience™, United States
Hiroshi Ezura, University of Tsukuba, Japan
Vladimir Nekrasov, Rothamsted Research, United Kingdom
Linda Ann Rymarquis, Bayer Crop Science (United States), United States

Citation: Kumar, S., Ezura, H., Nekrasov, V., Rymarquis, L. A., eds. (2021). CRISPR-Cas in Agriculture: Opportunities and Challenges. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88966-815-1

Table of Contents

frontiers in Plant Science

EDITORIAL
published: 26 March 2021
doi: 10.3389/fpls.2021.672329

Editorial: CRISPR-Cas in Agriculture: Opportunities and Challenges

Sandeep Kumar[1†], *Linda Ann Rymarquis*[2†], *Hiroshi Ezura*[3,4†] *and Vladimir Nekrasov*[5*†]

[1] Corteva Agriscience, Johnston, IA, United States, [2] Bayer Crop Science, Chesterfield, MO, United States, [3] Faculty of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Japan, [4] Tsukuba Plant Innovation Research Center (T-PIRC), University of Tsukuba, Tsukuba, Japan, [5] Plant Sciences Department, Rothamsted Research, Harpenden, United Kingdom

Keywords: agriculture, genome editing, crop, CRISPR, Cas9, Cas12a, policy, regulation

Editorial on the Research Topic

CRISPR-Cas in Agriculture: Opportunities and Challenges

OPEN ACCESS

Edited and reviewed by:
Raúl Alvarez-Venegas, Centro de Investigaciones y Estudios Avanzados, Instituto Politécnico Nacional de México (CINVESTAV), Mexico

***Correspondence:**
Vladimir Nekrasov
vladimir.nekrasov@rothamsted.ac.uk

†ORCID:
Sandeep Kumar
orcid.org/0000-0001-8614-9700
Linda Ann Rymarquis
orcid.org/0000-0003-0297-8371
Hiroshi Ezura
orcid.org/0000-0003-1443-2210
Vladimir Nekrasov
orcid.org/0000-0001-9386-1683

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *25 February 2021*
Accepted: *02 March 2021*
Published: *26 March 2021*

Citation:
Kumar S, Rymarquis LA, Ezura H and Nekrasov V (2021) Editorial: CRISPR-Cas in Agriculture: Opportunities and Challenges. Front. Plant Sci. 12:672329. doi: 10.3389/fpls.2021.672329

CRISPR-CAS TECHNOLOGY: STATE OF THE ART, POLICY, AND REGULATION

CRISPR-Cas genome editing technology is developing at a rapid pace and new molecular tools, such as CRISPR nucleases, are becoming regularly available. As part of this Research Topic, Bandyopadhyay et al. provide a comprehensive overview of Cas12a, a CRISPR nuclease formerly known as Cpf1. In their review article, the authors cover structural and mechanistic aspects of Cas12a in comparison with Cas9, the most commonly used CRISPR nuclease. They also highlight uses of Cas12a for the purpose of improving agriculturally important traits in various crops. An overview of Cas9 genome editing applications in plants is provided by El-Mounadi et al. who introduce the reader to the mechanism of Cas9 activity, methods of its delivery to plant cells (i.e., transformation techniques), give examples of improving crop traits using CRISPR-Cas9, and touch on biosafety and regulatory aspects associated with genome editing. A number of countries (e.g., the USA, Brazil, Argentina, and Japan) have already exempted genome edited crops, which do not carry transgenic DNA or novel combination of genetic material (i.e., not similarly achievable through conventional breeding), from being regulated similarly to Genetically Modified Organisms (GMOs) as genetically engineered (GE) organisms (Schmidt et al., 2020). Although the above-mentioned countries have passed legislation allowing cultivation of genome edited crops without GE regulation, the public dialogue and policy developments on the issue are evolving. In the case of Japan, Tabei et al. analyze Twitter conversation on genome-edited foods and their labeling over the period from May to October 2019. The analysis reveals that 54.5% of relevant tweets were statements opposed to food produced using genome edited crops, while only 7% were statements in favor of it. The remaining 38.5% of tweets were statements deemed neutral. Although the analysis was not necessarily representative of the wider Japanese society due to bias among Twitter users, the study underlines the importance of a continuous public dialogue on the issue of genome edited crops in Japan and the rest of the world.

CRISPR-CAS AS A TOOL FOR GENE FUNCTION STUDIES AND CROP TRAIT IMPROVEMENT

One of the factors impacting the efficiency of CRISPR-Cas is the expression level of the gene encoding the respective nuclease during different developmental stages of the plant. For example,

CRISPR-Cas mutagenesis in Arabidopsis often results in chimerism in the T1 generation due to low expression of *Cas9* (when driven by a promoter, such as 35S) during the zygote and early embryo developmental stages (Feng et al., 2014). To address the chimerism problem, egg cell-specific promoters could be a good choice to drive CRISPR nuclease gene expression for increasing the rates of CRISPR-Cas-induced germline mutations, which are inherited by the next generation (Wang et al., 2015; Yan et al., 2015; Mao et al., 2016). Zheng et al. test four different egg cell-specific promoters (two from Arabidopsis and two from soybean) to drive expression of *Cas9* in Arabidopsis and soybean transgenic lines. Out of the four promoters, AtEC1.2e1.1p, which is an adaptation of the previously published *AtEC1.1* and *AtEC1.2* promoter fusion (Wang et al., 2015), seems to perform best in both plant species. The study by Zheng et al. therefore contributes new molecular tools for efficient targeted mutagenesis in a model plant, such as Arabidopsis, and an economically important crop, such as soybean.

Virdi et al. and Zhang et al. highlight the use the CRISPR-Cas for gene function analysis. Virdi et al. studied the soybean *KASI* gene, which is crucial for conversion of sucrose to oil. They demonstrate that CRISPR-Cas induced knockout and in-frame deletion of *GmkasI* alleles, have an increase in seed sucrose content and a decrease in total seed oil content relative to wild type. These phenotypes are consistent with what was observed in the mutant line where the *GmKASI* gene is disrupted by a reciprocal chromosomal translocation. Thus, the authors prove that the phenotype of the line carrying the chromosomal translocation is indeed due to disruption of the *GmKASI* gene.

Zhang et al. report on knocking out 63 genes involved in immune response in tomato. The authors performed a detailed analysis of the types of mutations generated at an average frequency of 68%, which is similar to previously reported rates for CRISPR-Cas mutagenesis in tomato (Brooks et al., 2014; Nekrasov et al., 2017). They demonstrate that the mutations were transmitted through the germline to the next generation. The off-target analysis they performed for 12 guide RNAs showed no mutations at off-target sites with up to four mismatches and, indicating the high precision of CRISPR-Cas in tomato, this was consistent with what was previously reported (Nekrasov et al., 2017; Hahn and Nekrasov, 2019). The knockout lines are cataloged in the online Plant Genome Editing Database (PGED; http://plantcrispr.org; Zheng et al., 2019).

The CRISPR-Cas technology is a versatile genome editing tool that has been used to improve agriculturally important crop traits, such as quality, disease resistance, and herbicide tolerance. In potato, enzymatic browning is a serious problem for both growers and the industry as it decreases the quality of both the fresh and processed product. González et al. report on a successful application of CRISPR ribonucleoproteins for the purpose of reducing enzymatic browning in potato tubers by targeting the Polyphenol Oxidase 2 gene (*StPPO2*), one of the five potato PPO genes. By disrupting all four copies of *StPPO2* the authors achieved a dramatic reduction in tuber PPO activity (up to 69%) and enzymatic browning (73%). The findings presented by González et al. are consistent with the reported reduction in potato browning achieved by silencing the *StPPO2* gene using RNAi (Richael, 2021).

In addition to dicot crops, CRISPR-Cas has been extensively used for trait improvement in cereals, such as rice and maize. As an example, Zafar et al. are reporting on enhancing disease resistance to *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), a pathogen causing bacterial blight in rice, by editing the promoter of a susceptibility (S) gene. *Xoo* secretes transcription activator-like effectors (TALEs) that activate host S genes, such as Os*SWEET* family members. Using CRISPR-Cas, the authors introduced deletions overlapping with effector binding elements (EBEs) recognized by AvrXa7/PthXo3 or TalF TALEs within the promoter of the *OsSWEET14* gene in the Super Basmati elite cultivar. Mutant rice lines carrying deletions in the AvrXa7/PthXo3 EBEs showed enhanced resistance to the *Xoo* strain carrying *AvrXa7* in agreement with previously published reports (Li et al., 2012; Blanvillain-Baufumé et al., 2017; Oliva et al., 2019; Xu et al., 2019).

In another report, Komatsu et al. address the problem of "volunteer rice" that emerges from seeds falling into fields during the harvest season and then spontaneously germinates the next spring. If volunteer rice originates from a feed variety, it can compromise the quality of rice meant for human consumption, which is grown during the next season. As many japonica rice cultivars are resistant to beta-triketone herbicides (bTH), such as benzobicyclon (BBC), the authors tested the feasibility of engineering BBC susceptibility in japonica rice (cv. Nipponbare) by targeting the *HIS1* gene using the cytosine base editor (CBE). They successfully generated a number of *his1* knockout lines by eliminating the start codon or introducing premature stop codons within the *HIS1* coding sequence. The *his1* loss-of-function lines appear to be susceptible to BBC and other beta-triketone pesticides, paving a way to controlling volunteer rice in the field by applying the same strategy to BBC-resistant feed rice cultivars.

In maize, Gao et al. report on a CRISPR-Cas9- and recombinase-mediated strategy for stacking biotech traits within complex trait loci (CTLs). Each CTL spans 4–5 cM and includes 12–30 pre-selected sites used for insertion of a landing pad via homology-directed repair (HDR) using CRISPR-Cas9. As a result, the authors generated a set of individual transgenic lines, each carrying a landing pad at one of the preselected sites within one of the four CTLs. At the following step, the landing pad-carrying lines were used for integration of trait genes using the FLP recombinase. Finally, integrated trait genes were stacked on the same chromosome by crossing respective individual transgenic lines and selecting recombinants. The study by Gao et al. therefore presents a modular and flexible way of stacking biotech traits, as compared to previously reported strategies involving recombinases or zinc finger/homing endonucleases (Ow, 2011; D'Halluin et al., 2013; Kumar et al., 2015), due to the possibility for trait genes to be easily combined or separated (e.g., in case one of them loses efficiency) by conventional breeding.

This volume highlights the many opportunities that the CRISPR-Cas systems hold for Agriculture. Both Cas9 and Cas12a have been proven to drive edits in plants and new improvements, such as using germline-specific

promoters that increase heritability, will only enhance their potential (Zheng et al.). The CRISPR-Cas systems have been successfully implemented to forward basic research, such as gene discovery of oil (Virdi et al.) and disease genes (Zhang et al.), as well as improve agricultural outcomes e.g., via decreasing potato browning (González et al.), improving disease resistance (Zafar et al.), mitigating volunteer rice (Komatsu et al.), and stacking biotech traits (Gao et al.). In order for this potential to be fully realized, Tabei et al. showed that work must be done to gain public acceptance and ensure implementation of favorable public policy.

AUTHOR CONTRIBUTIONS

VN wrote the paper. SK, LR, and HE made direct contributions to the work, reviewed, and approved it for publication. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

VN receives grant-aided support from the Biotechnology and Biological Sciences Research Council (BBSRC) Designing Future Wheat (DFW) programme (grant number BBS/E/C/000I0220) and Newton Fund (grant number BB/R019827/1).

REFERENCES

Blanvillain-Baufumé, S., Reschke, M., Solé, M., Auguy, F., Doucoure, H., Szurek, B., et al. (2017). Targeted promoter editing for rice resistance to *Xanthomonas oryzae* pv. *oryzae* reveals differential activities for SWEET14-inducing TAL effectors. *Plant Biotechnol. J.* 15, 306–317. doi: 10.1111/pbi.12613

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

D'Halluin, K., Vanderstraeten, C., Hulle, J. V., Rosolowska, J., Brande, I. V. D., 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

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. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4632–4637. doi: 10.1073/pnas.1400822111

Hahn, F., 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

Kumar, S., AlAbed, D., Worden, A., Novak, S., Wu, H., Ausmus, C., et al. (2015). A modular gene targeting system for sequential transgene stacking in plants. *J. Biotechnol.* 207, 12–20. doi: 10.1016/j.jbiotec.2015.04.006

Li, T., Liu, B., Spalding, M. H., Weeks, D. P., and Yang, B. (2012). High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat. Biotechnol.* 30, 390–392. doi: 10.1038/nbt.2199

Mao, Y., Zhang, Z., Feng, Z., Wei, P., Zhang, H., Botella, J. R., et al. (2016). Development of germ-line-specific CRISPR-Cas9 systems to improve the production of heritable gene modifications in Arabidopsis. *Plant Biotechnol. J.* 14, 519–532. doi: 10.1111/pbi.12468

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/s41598-017-00578-x

Oliva, R., Ji, C., Atienza-Grande, G., Huguet-Tapia, J. C., Perez-Quintero, A., Li, T., et al. (2019). Broad-spectrum resistance to bacterial blight in rice using genome editing. *Nat. Biotechnol.* 37, 1344–1350. doi: 10.1038/s41587-019-0267-z

Ow, D. W. (2011). Recombinase-mediated gene stacking as a transformation operating system. *J. Integr. Plant Biol.* 53, 512–519. doi: 10.1111/j.1744-7909.2011.01061.x

Richael, C. M. (2021). "Development of the Genetically Modified Innate® Potato," in *Plant Breeding Reviews*, ed I. Goldman (Hoboken, NJ: John Wiley & Sons, Ltd.), 57–78. doi: 10.1002/9781119717003.ch3

Schmidt, S. M., Belisle, M., and Frommer, W. B. (2020). The evolving landscape around genome editing in agriculture. *EMBO Rep.* 21:e50680. doi: 10.15252/embr.202050680

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

Xu, Z., Xu, X., Gong, Q., Li, Z., Li, Y., Wang, S., et al. (2019). Engineering broad-spectrum bacterial blight resistance by simultaneously disrupting variable TALE-binding elements of multiple susceptibility genes in rice. *Mol. Plant* 12, 1434–1446. doi: 10.1016/j.molp.2019.08.006

Yan, L., Wei, S., Wu, Y., Hu, R., Li, H., Yang, W., et al. (2015). High-efficiency genome editing in Arabidopsis using YAO promoter-driven CRISPR/Cas9 system. *Mol. Plant* 8, 1820–1823. doi: 10.1016/j.molp.2015.10.004

Zheng, Y., Zhang, N., Martin, G. B., and Fei, Z. (2019). Plant genome editing database (PGED): a call for submission of information about genome-edited plant mutants. *Mol. Plant* 12, 127–129. doi: 10.1016/j.molp.2019.01.001

Conflict of Interest: SK is an employee of Corteva Agriscience™. LR is an employee of Bayer.

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 09 January 2020
doi: 10.3389/fpls.2019.01649

Reduced Enzymatic Browning in Potato Tubers by Specific Editing of a Polyphenol Oxidase Gene *via* Ribonucleoprotein Complexes Delivery of the CRISPR/Cas9 System

Matías Nicolás González[1,2*], *Gabriela Alejandra Massa*[1,2,3], *Mariette Andersson*[4], *Helle Turesson*[4], *Niklas Olsson*[4], *Ann-Sofie Fält*[4], *Leonardo Storani*[1,2], *Cecilia Andrea Décima Oneto*[2], *Per Hofvander*[4†] *and Sergio Enrique Feingold*[2†]

[1] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina, [2] Laboratorio de Agrobiotecnología, INTA - EEA Balcarce, Balcarce, Argentina, [3] Facultad de Ciencias Agrarias, Universidad Nacional de Mar del Plata, Balcarce, Argentina, [4] Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, Sweden

OPEN ACCESS

Edited by:
Vladimir Nekrasov, Rothamsted Research (BBSRC), United Kingdom

Reviewed by:
Thomas Benjamin Jacobs, VIB-UGent Center for Plant Systems Biology, Belgium
Fabien Nogué, INRA Centre Versailles-Grignon, France

***Correspondence:**
Matías Nicolás González
gonzalez.matiasn@inta.gob.ar

[†]*These authors share senior authorship*

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *11 September 2019*
Accepted: *22 November 2019*
Published: *09 January 2020*

Citation:
González MN, Massa GA, Andersson M, Turesson H, Olsson N, Fält A-S, Storani L, Décima Oneto CA, Hofvander P and Feingold SE (2020) Reduced Enzymatic Browning in Potato Tubers by Specific Editing of a Polyphenol Oxidase Gene via Ribonucleoprotein Complexes Delivery of the CRISPR/Cas9 System. Front. Plant Sci. 10:1649. doi: 10.3389/fpls.2019.01649

Polyphenol Oxidases (PPOs) catalyze the conversion of phenolic substrates to quinones, leading to the formation of dark-colored precipitates in fruits and vegetables. This process, known as enzymatic browning, is the cause of undesirable changes in organoleptic properties and the loss of nutritional quality in plant-derived products. In potato (*Solanum tubersoum* L.), PPOs are encoded by a multi-gene family with different expression patterns. Here, we have studied the application of the CRISPR/Cas9 system to induce mutations in the *StPPO2* gene in the tetraploid cultivar Desiree. We hypothesized that the specific editing of this target gene would result in a lower PPO activity in the tuber with the consequent reduction of the enzymatic browning. Ribonucleoprotein complexes (RNPs), formed by two sgRNAs and Cas9 nuclease, were transfected to potato protoplasts. Up to 68% of regenerated plants contained mutations in at least one allele of the target gene, while 24% of edited lines carried mutations in all four alleles. No off-target mutations were identified in other analyzed *StPPO* genes. Mutations induced in the four alleles of *StPPO2* gene, led to lines with a reduction of up to 69% in tuber PPO activity and a reduction of 73% in enzymatic browning, compared to the control. Our results demonstrate that the CRISPR/Cas9 system can be applied to develop potato varieties with reduced enzymatic browning in tubers, by the specific editing of a single member of the *StPPO* gene family.

Keywords: CRISPR/Cas9, potato, genome editing, enzymatic browning, polyphenol oxidase, ribonucleoprotein complexes

INTRODUCTION

Polyphenol Oxidases (PPOs; E.C.1.10.3.1, E.C.1.10.3.2, or E.C.1.14.18.1) are copper-containing enzymes, widely distributed among higher plants (Yoruk and Marshall, 2003), that catalyze the oxidation of an extensive range of phenolic compounds to their respective quinones. The quinones

generated by action of PPOs can undergo self-polymerization or react with amino acids or free radicals in proteins leading to the formation of dark-colored precipitates (Mayer, 2006). This process, known as Enzymatic Browning, is the cause of reduction in quality that alters the color, taste, texture and nutritional value of several fresh and processed fruits and vegetables (Jukanti, 2017). In addition, the oxidation of polyphenolic compounds by PPOs in plant derived products for human consumption is highly undesirable, since polyphenols are natural antioxidants with possible protective effects against cancer and cardiovascular diseases (Shahidi and Ambigaipalan, 2015).

In potato (*Solanum tuberosum* L.), enzymatic browning is a serious problem for both, producers and the industry, because the tubers can be affected during harvest and post-harvest procedures such as shipping, storage, distribution and blanching (Bachem et al., 1994). This undesired process is controlled in industry by using chemical and/or physical agents (Zhang et al., 2018b). However, these methods have important disadvantages including alterations of organoleptic and nutritional quality of the final products and some of them can even represent a potential risks for human health (Tinello and Lante, 2018). Therefore, the development of new technologies to control PPOs activity in planta is the most promising and safest approach to avoid undesirable browning compounds in fresh and processed potato derived products.

In most of plant species, PPOs are encoded by multi-gene families, which suggests their implication in a variety of cell processes (Tran et al., 2012). PPOs have been associated with several metabolic and biosynthetic processes (Jukanti, 2017) as well as with plant defense responses (Li and Steffens, 2002; Thipyapong et al., 2004; Wang and Constabel, 2004; Kampatsikas et al., 2019). Five *PPO* genes have been originally described in potato (*StPPO*), each one having a special pattern of tissue induction and expression (Thygesen et al., 1995). Once the potato genome sequence data was available (Potato Genome Sequencing Consortium, 2011), a genome-wide survey revealed nine *StPPO*-like genes (named *StPPO1* to *9*), with differential prevalence of ESTs found from different potato tissues (Chi et al., 2014). Several reports have described the use of different RNA silencing technologies to down regulate *StPPO* genes, in order to reduce the enzymatic browning in the tubers (Bachem et al., 1994; Coetzer et al., 2001; Rommens et al., 2006; Llorente et al., 2011). Most of these reports are based on down-regulation of multiple *StPPO* genes, which could have a negative impact on other functions of the enzyme in the plant. Moreover, with this strategy, the gene constructs of the silencing machinery need to be stably inserted into the genome, which represents a drawback considering the time-consuming and costly process of deregulation of a Genetically Modified Organism (GMO) in several countries (Eckerstorfer et al., 2019).

Chi et al. (2014) studied the contribution of each member of the *StPPO* gene family to the total PPO protein activity in the potato tuber. By using artificial micro-RNAs (amiRNAs) authors down-regulated *StPPO* genes individually or in combinations, concluding that four genes are the main responsible for PPO activity in the tuber. *StPPO2* (PGSC0003DMG400018916) gene is the principal contributor to PPO total protein content, with 55% of the total enzyme, followed by *StPPO1* (PGSC0003DMG400029575) with 25–30% and *StPPO4* (PGSC0003DMG400018917) and *StPPO3* (PGSC0003DMG400018914), together with less than 15%.

Genome editing using the CRISPR/Cas9 system is a powerful tool for crop improvement and has been applied to add or modify several traits in many economically important plant species (Arora and Narula, 2017; Baltes et al., 2017; Scheben et al., 2017; Gao, 2018). In its simplest form, the Cas9 nuclease is guided by one or more RNA molecule/s (sgRNA/s) to a specific target site in the host genome to introduce a double stranded break (DSB) in the DNA (Jinek et al., 2012). Following the induction of this DSB, mutations are introduced by the error-prone DNA repair mechanism of Non Homologous End Joining (NHEJ), (Puchta, 2005). When performed in an exon, this can produce a loss of gene function due to frame shifts or deletions of specific fragments of the coding sequence. Cas9 and sgRNAs can be directly delivered to the cell as a Ribonucleoprotein complex (RNPs), (Woo et al., 2015) an approach that avoids foreign DNA insertions in the plant genome. This strategy has been successfully applied to modify genes in several important crops like maize (Svitashev et al., 2016), bread wheat (Liang et al., 2017) and, more recently, potato (Andersson et al., 2018). Considering the current criteria for the determination of the regulatory status of genome edited crops in Argentina and other countries (Whelan and Lema, 2015; Lema, 2019), this approach could result in the development of crop varieties not subjected to the cumbersome GMO regulation process, and treated under the same regulatory framework as varieties obtained by conventional breeding, which includes chemical or radiation mutagenesis (Eckerstorfer et al., 2019).

In this work, we have studied the editing of the *StPPO2* gene in the tetraploid cultivar Desiree, by using the CRISPR/Cas9 system. The reagents for genome editing were delivered in the form of RNPs into potato protoplasts, aiming to avoid the insertion of foreign DNA. Regenerated lines were screened for induced mutations in the target gene and potential off target activity on other members of *StPPO* gene family. Selected lines with mutations in the four alleles of the target gene were grown and assayed for enzymatic browning and PPO activity levels in tubers.

RESULTS

SgRNA Design on *StPPO2* Gene and Off Target Prediction

In order to find targets to direct Cas9 nuclease to the *StPPO2* gene, a fragment covering the 5′end of the coding sequence was amplified from *S. tuberosum* cv. Desiree and sequenced. The amplified fragment was predicted to encode the N-terminal of the enzyme, including the first copper-binding site (CuA; **Supplementary Figure S1**), which forms part of the active site (Marusek et al., 2006). Two sgRNAs were selected on the

resulting sequence with strict absence of allelic variation and named sgRNA157 and sgRNA564 (**Figure 1A**). The expected cutting sites for Cas9 on each target were estimated to be separated by 111 bp on the *StPPO2* sequence (**Supplementary Figure S1**).

In order to avoid inducing mutations in other *StPPO* genes, the two selected sgRNAs were analyzed for possible off target activity. Considering up to four mismatches (Hahn and Nekrasov, 2019) *StPPO1* and *StPPO4* genes were identified as possible off targets of sgRNA564 (**Figure 1B** and **Supplementary Figures S2** and **S3**). Four mismatches at positions +1, +2, +8, and +13 from the Protospacer Adjacent Motif (PAM) were identified in the potential off target site on *StPPO1* and four mismatches at positions +1, +8, +13, and +20, in the potential off target site on *StPPO4* (**Figure 1B**).

No putative off targets on *StPPO* genes were found for sgRNA157 considering four or less mismatches. **Figure 1C** shows the alignment of sgRNA157 with the corresponding sequences of *StPPO1* and *StPPO4*. Although not considered as possible off targets according to the mentioned parameters, both regions were included for further analysis. As highlighted, five mismatches were identified between *StPPO1* and sgRNA157. In addition, a non-canonical PAM sequence (NAG) was found at the 3′ end of the *StPPO1* gene sequence (**Figure 1C**). Eight mismatches were identified between sgRNA157 and the corresponding sequence of the *StPPO4* gene (**Figure 1C**).

Protoplast Transfection With RNPs and Mutation Screening of Regenerated Lines

CRISPR/Cas9 was delivered in the form of Ribonucleoprotein complexes (RNPs, Andersson et al., 2018) into protoplasts by transfections with 25 or 40% Polyethylenglycol 4000 (PEG) and incubations times of 3 or 30 min, respectively. After regeneration, the identification of edited lines was carried out using the High Resolution Fragment Analysis (HRFA, **Figure 2A**). Based on the analysis of 64 lines regenerated from the 25% PEG transfection, the genome editing efficiency was 27%, defined as the percentage of analyzed lines carrying mutations in at least one allele of the target gene. On the other hand, from the 40% PEG transfection, 28 regenerated lines were analyzed and 68% were found to carry mutations. Taking both transfections together, nine edited lines displayed mutations in all the four alleles of the target gene, with eight of these lines originated from the 40% PEG transfection (**Table 1**). The majority of mutations were small deletions, but in several lines, larger deletions from 102 to 118 nucleotides were observed (**Table 1**), suggesting that Cas9 nuclease introduced cuts at both targets sites, leading to the elimination of the fragment in between. In addition, insertions ranging from 22 to 302 bp were identified in nine lines (**Table 1**). Finally, more than four allelic variants suggesting chimerism was not observed in any of the 92 analyzed lines (**Table 1**).

Sequence Analysis of *StPPO2* in Selected Lines

Sequence analysis was performed on selected lines to confirm HRFA results (**Figure 2B** and **Supplementary Figure S4**). In lines M07056, M08001 and M08002, small deletions were identified, which in most alleles were the product of mutations induced at both target sites, without the elimination of the fragment in between (**Figure 2B** and **Supplementary Figure S4**). In the case of M07056, all mutations are predicted to change the reading frame of the *StPPO2* coding sequence (**Supplementary Figure S4**). The loss of the fragment spanned by the two sgRNAs target sites, was confirmed in alleles of lines

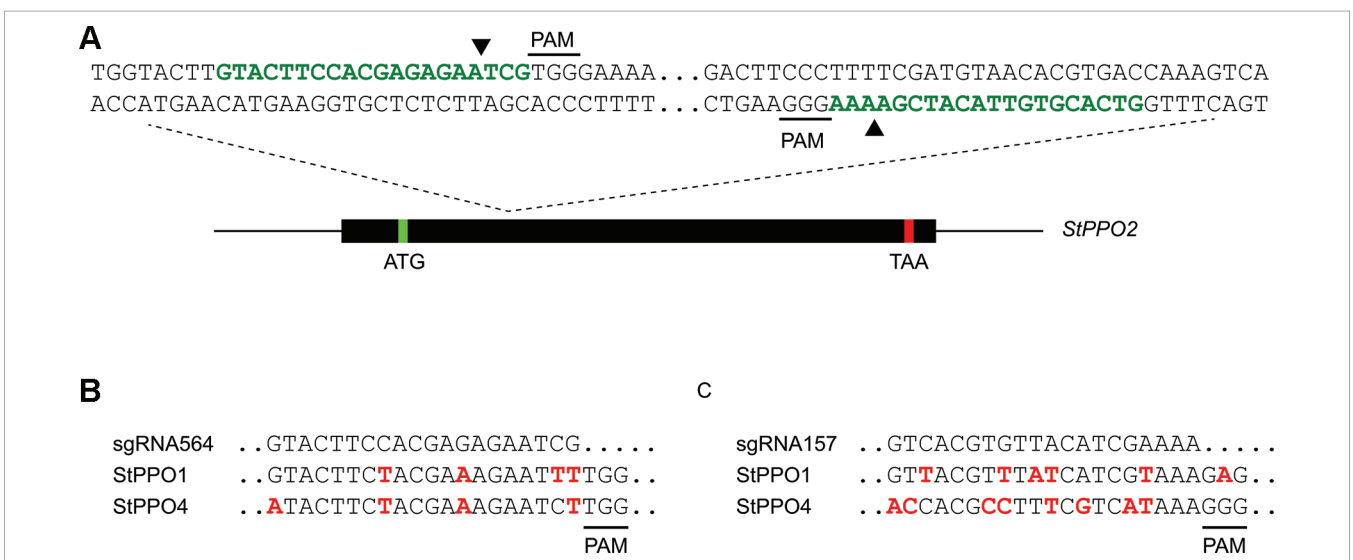


FIGURE 1 | sgRNA design on the *StPPO2* gene and off target prediction **(A)** The structure of *StPPO2* gene is shown with the theoretical positions of the start (ATG, green box) and stop codons (TAA, red box). The partial sequence used for sgRNA design is shown above the gene structure. Targets sites for sgRNAs are marked in green letters and the PAM (5′-NGG-3′) of each target site is indicated. Black arrowheads indicate the predicted cut site for the Cas9 nuclease **(B)** Predicted off target sites for sgRNA564 on *StPPO1* and *StPPO4* genes, with mismatches marked in red letters **(C)** Alignment of sgRNA157 with *StPPO1* and *StPPO4* genes, with mismatches marked in red letters.

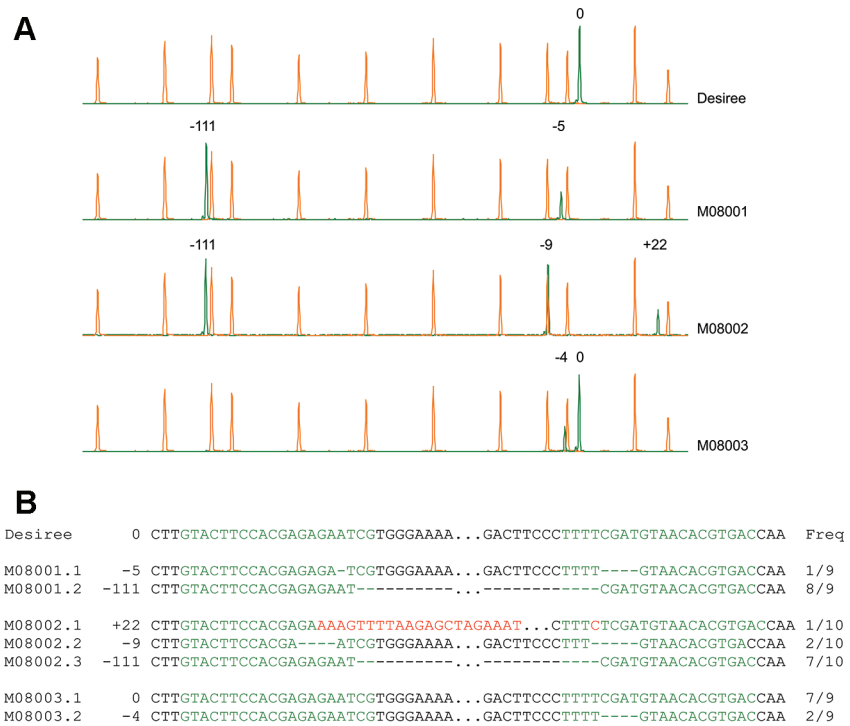


FIGURE 2 | Identification of edited lines using High Resolution Fragment Analysis (HRFA) and characterization of mutations by sequencing **(A)** Electropherograms of HRFA obtained for wild type Desiree and lines M08001, M08002, and M08003. The orange peaks correspond to the elution points of the size standard and green peaks correspond to elution of the *StPPO2* gene fragments. The elution of the wild type fragment is set to 0 and the number of bases inserted (+) or deleted (–) in each fragment is indicated above the respective peak **(B)** Sequencing of a partial fragment of the *StPPO2* alleles in selected lines. Target sites for the sgRNAs are marked in green letters. Deleted nucleotides are indicated as hyphens and inserted bases are marked in red letters. The frequencies obtained during Sanger analysis are indicated, as the number of clones carrying each allelic variant related to the total number of sequenced clones.

M08001, M08002 (**Figure 2B**), and M08008 (**Supplementary Figure S4**), as was indicated by the HRFA results. The presence of the wild type allele was confirmed in line M08003, along with at least one allele carrying a deletion of 4 bp on the target site of sgRNA157 (**Figure 2B**). Moreover, the lower prevalence of the mutated allele in comparison with the wild type in the sequence analysis, suggest that M08003 possess multiple copies of the wild type allele (**Figure 2B**).

Finally, insertions observed in the HRFA were analyzed in lines M08008 and M08002. The larger insertions were found to correspond to fragments of genomic DNA of potato as well as elements of DNA used for the *in vitro* transcription of the sgRNAs (**Figure 2B** and **Supplementary Figure S4**).

Analysis of Off Target Mutations in *StPPO* Genes

Unexpected mutations have been reported in plants using CRISPR/Cas9 as a genome editing tool (Zhang et al., 2018a). With the aim of analyzing the presence of off target mutations on other *StPPO* genes, HRFA was performed on *StPPO1* and *StPPO4* genes in selected lines carrying mutations in all the four alleles of *StPPO2* gene. The electropherograms analysis revealed no differences in fragments length between the edited lines and the control (**Figures 3A**, **B** and **Supplementary Figure S5**), indicating no insertions or deletions introduced on the possible recognition sites for the two sgRNAs.

Enzymatic Browning and PPO Activity Analysis in Tubers

Selected lines carrying mutations in all four alleles of the *StPPO2* gene were subjected to phenotypic analysis of enzymatic browning and PPO activity in tubers. A wild type line obtained from the regeneration of non-transfected protoplasts were used as a control (Desiree RC). Line M08003 was also included, since it presents a mutation in at least one allele of the target gene, together with at least one copy of the wild type allele (**Figure 2B**). All lines were grown in a growth chamber and displayed no evident phenotypic abnormalities during plant development.

For phenotype analysis, the tubers were cut, exposed to air and discoloration development was registered at times 0, 24, and 48 h after cutting (**Figure 4**). After 24 h of air exposure, the typical brown discoloration related to oxidation was visible in lines Desiree RC and M08003, but not in the rest of the analyzed lines (**Figure 4**). The same pattern, but with stronger differences between lines was observed after 48 h of air exposure. Lines Desiree RC and M08003 developed the brown discoloration in a shorter time and over a larger area of the tuber surface (**Figure 4**).

TABLE 1 | Edited lines per experiment detected by HRFA.

Line	Transfection	Allelic Variants
Desiree RC	–	0
M07006	25% PEG	–4/0
M07009	25% PEG	–4/–1/0
M07014	25% PEG	–4/0
M07020	25% PEG	–102/–4/–1/0
M07028	25% PEG	0/+46
M07029	25% PEG	–6/0
M07030	25% PEG	–5/0
M07031	25% PEG	–111/0/+1
M07032	25% PEG	–4/0
M07036	25% PEG	–5/0
M07046	25% PEG	–1/0
M07051	25% PEG	0/+1
M07053	25% PEG	–6/–4/0/+1
M07056	25% PEG	–7/–4
M07057	25% PEG	–1/0
M07062	25% PEG	–1/0/+1
M07063	25% PEG	–3/0/+1
M07066	25% PEG	0/+1
M08001	40% PEG	–111/–5
M08002	40% PEG	–111/–9/+22
M08003	40% PEG	–4/0
M08007	40% PEG	–9/–5/–3/0
M08008	40% PEG	–111/–15/+121/+302
M08009	40% PEG	–112/–5/+44
M08012	40% PEG	0/+1
M08013	40% PEG	–115/–8/–5/+55
M08014	40% PEG	–5/–4/0
M08015	40% PEG	–111/0/+98
M08016	40% PEG	–4/0/+45
M08017	40% PEG	–14/–8/0
M08018	40% PEG	–114/–11/–5/0
M08020	40% PEG	–1/0/+48
M08024	40% PEG	–116/–111/–55/+58
M08025	40% PEG	–18/–1/0
M08026	40% PEG	–4/–2/0
M08027	40% PEG	–118/–111
M08028	40% PEG	–113/–5/+1

The transfection conditions and the allelic variants found in the analysis of all four alleles of StPPO2 *are indicated per line. The number 0 indicates the presence of the wild type allele. Alleles with insertions or deletions are indicated as the number of base pairs with a minus (–) or a plus (+) sign, respectively.*

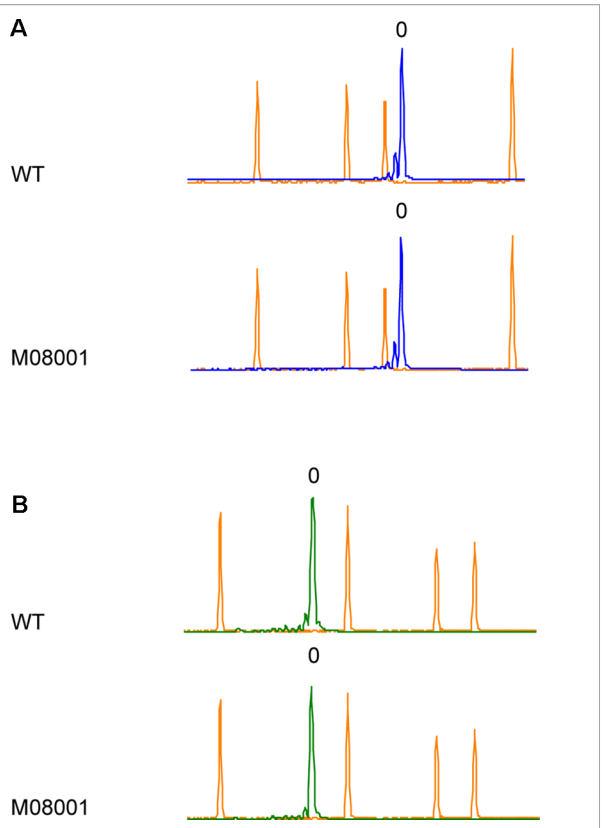


FIGURE 3 | Screening for off target mutations on *StPPO1* and *StPPO4* genes by HRFA. Electropherograms of wild type Desiree and line M08001 are shown. The orange peaks correspond to the elution points of the size standard and the elution of the respective wild type fragment is set to 0 (A) Blue peaks correspond to elution of the *StPPO1* gene fragments (B) Green peaks correspond to elution of the *StPPO4* gene fragments.

The enzymatic browning was measured for each line and related to that of the control Desiree RC (**Figure 5A**). The relative enzymatic browning was significantly lower in all the edited lines in comparison to the control line (**Table 2**), with the exception of line M08003. The relative enzymatic browning in lines M08001 and M08002 ranged between 0.26 and 0.27, demonstrating a reduction of around 73% related to the control line (**Figure 5A**). Lines M07056, M08008 and M08027, displayed a middle reduction of 68, 67, and 66%, respectively, compared to the control (**Figure 5A**).

The PPO activity was measured for each line and made relative to the control Desiree RC (**Figure 5B**). Relative PPO activity was significantly lower in lines M08001 and M08002 in comparison to the control line Desiree RC (**Table 2**) with reductions of 64 and 69%, respectively (**Figure 5B**). Even though not significantly statistical differences, a middle reduction of 39, 28, and 41% in relative PPO activity was observed for lines M07056, M08008, and M08027, respectively, when compared to the control (**Figure 5B**).

In order to determine whether the relative enzymatic browning was correlated to the relative PPO activity in our study, the Spearman correlation coefficient (ρ) was determined between the two variables. As expected, a significantly positive correlation was found ($\rho = 0.63$, $p < 0.005$).

DISCUSSION

Enzymatic browning caused by the activity of PPOs leads to alterations in color and organoleptic properties of fresh and processed fruits and vegetables, which is perceived as a serious quality deficiency for industry and consumers (Yoruk and Marshall, 2003). In our study, the CRISPR/Cas9 system was applied in potato to induce mutations in the *StPPO2* gene, responsible for most of the PPO activity and enzyme content in tubers (Chi et al., 2014). We hypothesized that the specific editing of this target gene would result in a lower PPO activity in the tuber and the consequent reduction of the enzymatic browning.

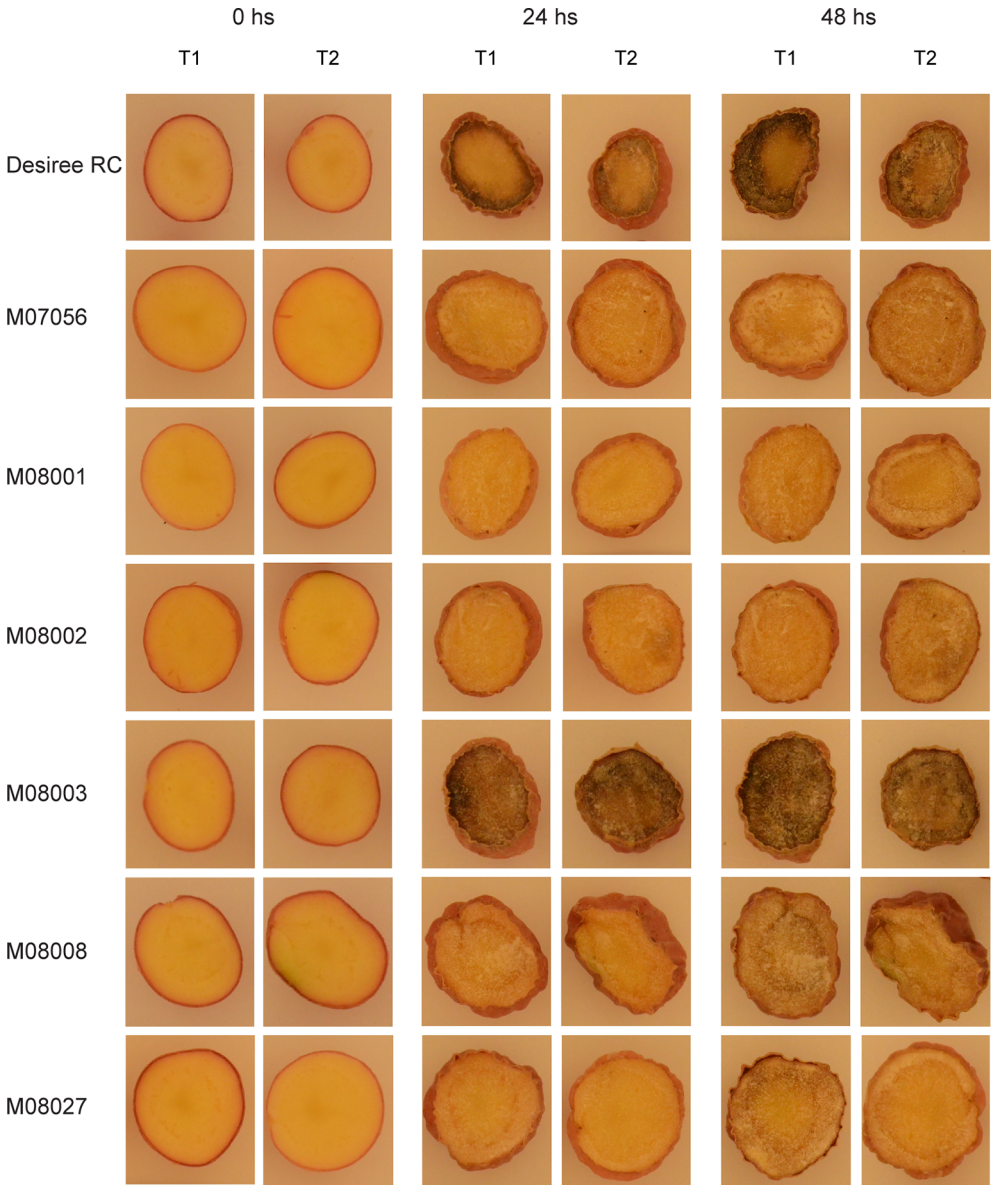


FIGURE 4 | Discoloration development of selected edited lines at times 0, 24, and 48 h after cutting. Two tubers were randomly selected for each edited line and the control, cut and exposed to the air for 48 h at room temperature (24°C). Photos were taken immediately after cutting (0hs), 24 and 48 h later. T1 and T2 indicate Tuber 1 and Tuber 2 of each line, respectively.

For the CRISPR/Cas9 system delivery, we utilized Ribonucleoprotein complexes (RNPs) to transfect potato protoplasts and further whole plant regeneration (Nicolia et al., 2015). The genome editing efficiency of 27 and 68% obtained in this study was higher than previously reported using RNPs in potato (9–25%, Andersson et al., 2018). The efficiency in genome editing is largely affected by the target gene as well as the sgRNAs sequence used to direct the Cas9 nuclease (Kumlehn et al., 2018). On the other hand, the activity of CRISPR/Cas9 would be influenced by the transfection efficiency of the reagents into the protoplasts, which could vary between potato varieties. In addition, the combination of two sgRNAs on one target gene used in our study, could explain the increase in the efficiency obtained. Such strategy not only increased the possibilities of inducing mutations in the target gene, but also led to the elimination of larger, specific fragments from the coding sequence as was previously reported in tomato (Brooks et al., 2014), rice (Zhou et al., 2014), barley (Kapusi et al., 2017) and potato (Tuncel et al., 2019; Veillet et al., 2019).

The HRFA performed in our study has shown lines with multiple alleles of *StPPO2* carrying the same type of mutation. Although a less frequent pathway than NHEJ, the DSB repair *via* homologous recombination (HR) is a mechanism observed in plant somatic cells (Puchta, 2005; Shi et al., 2017; Yu et al., 2017). The availability of a mutated homologue allele as a donor template

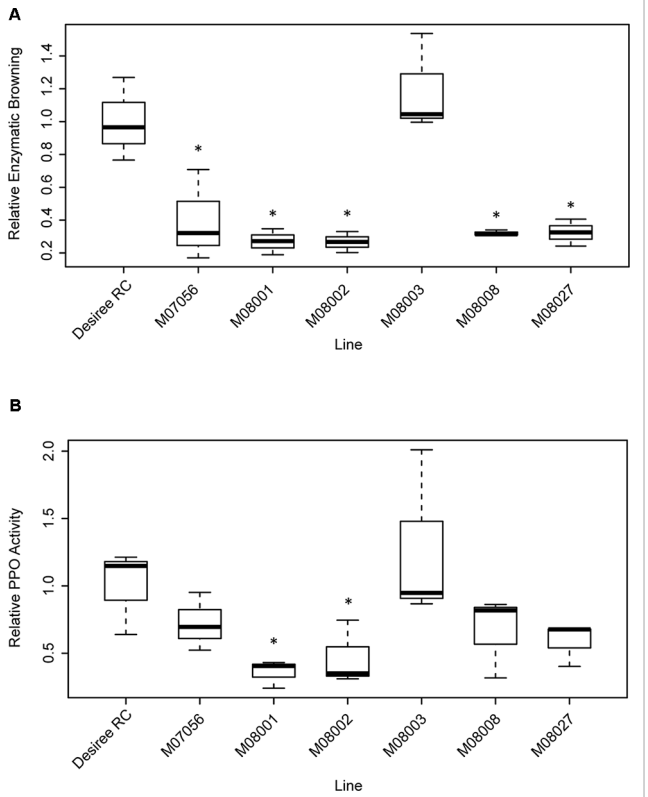


FIGURE 5 | Analysis of Relative Enzymatic Browning **(A)** and Relative PPO Activity **(B)** in tubers of selected edited lines. Each box represents data of three biological replicates of the edited lines and the control Desiree RC, the line across the box represents the median. The box represents the 25th and the 75th percentiles and whiskers represent the maximum and minimum value. Data are relative to the control line Desiree RC. Statistical differences with the control line Desiree RC are denoted *($p < 0.05$).

TABLE 2 | Effect of each line on Relative Enzymatic Browning and Relative PPO Activity variables.

Response variable	Fixed effect	Estimate ± SE	*p*
Relative Enzymatic Browning	Intercept	1.00 ± 0.15	<0.001
	M07056	−0.60 ± 0.22	0.01
	M08001	−0.73 ± 0.15	<0.001
	M08002	−0.73 ± 0.15	<0.001
	M08003	0.19 ± 0.22	0.40
	M08008	−0.68 ± 0.14	<0.001
	M08027	−0.67 ± 0.15	<0.001
Relative PPO Activity	Intercept	0.48 ± 0.08	<0.001
	M07056	−0.13 ± 0.10	0.2295
	M08001	−0.31 ± 0.09	0.0047
	M08002	−0.25 ± 0.11	0.0355
	M08003	0.13 ± 0.19	0.5137
	M08008	−0.16 ± 0.12	0.2057
	M08027	−0.19 ± 0.09	0.0615

Parameters estimates and p *values were taken from Linear Mixed Models with the Line as a fixed effect and the biological replicate as a random effect.* $p < 0.05$ *indicates significant differences with the Desiree RC control line.*

during DSB repair could result in a bias towards homozygous mutations, as observed in the mentioned lines. Nevertheless, our results are not sufficient to confirm such mechanism and further experiments would be necessary to confirm this hypothesis.

Foreign DNA integration into the plant genome is a major concern in genome editing techniques, and is preferably avoided when applied for commercial breeding purposes (Eckerstorfer et al., 2019). This is of special importance in a tetraploid and highly heterozygous crop like potato, since backcrossing techniques to eliminate inserted foreign DNA would lead to the loss of allelic combination in an elite variety (Nadakuduti et al., 2018). In some of the mutated lines identified in our study, insertions were observed in the target region, which corresponded to fragments of the DNA template used in the *in vitro* transcription of the sgRNAs, or potato genomic fragments. Although, the latter cannot be consider as a foreign DNA integration, the first type of insertions could be avoided by using synthetic sgRNA instead, as previously reported by Andersson et al. (2018). Nevertheless, the percentage of insertions detected was very low (9 out of 37 lines) and, in addition, we obtained a majority of multi-allelic edited lines with no evident DNA insertions into the target sites. The confirmation of the absence of foreign DNA in such lines could result in plants considered not different from conventionally bred varieties, taking into account the actual criteria for determining the regulatory status of genome edited products in Argentina and other countries (Whelan and Lema, 2015; Eckerstorfer et al., 2019; Lema, 2019).

Off target activity, i.e. introduction of unintended mutations, have been reported using the CRISPR/Cas9 system in plants (Zhang et al., 2018a). Assaying all possible off target mutations induced by the selected sgRNAs would only be possible throughout whole genome sequencing of the edited lines (Li et al., 2019), a goal that is beyond the objectives of our study. Nevertheless, we aimed to confirm that our selected edited lines displayed mutations only in the *StPPO2* gene, with no alteration in the coding sequences of other members of the *StPPO* gene family, as paralogs may share a considerably degree of sequence similarity (Chi et al., 2014). Only two possible off target sites were found on other *StPPO* genes for sgRNA564, considering up to four mismatches. The HRFA of the selected lines indicated no insertions or deletions in *StPPO1* and *StPPO4* genes. The presence of multiple mismatches into the seed region (defined as the 8–12 nt proximal to the PAM) between the selected sgRNAs and the rest of *StPPO* genes could explain their specificity for *StPPO2* (Hahn and Nekrasov, 2019). On the other hand, the use of RNPs as delivery method for the CRISPR/Cas9 system has been proposed to reduce the incidence of off targets effects, due to the rapid degradation of the Cas9 nuclease and the sgRNAs in the cell (Nadakuduti et al., 2018; Zhang et al., 2018a; Hahn and Nekrasov, 2019).

Earlier studies have reported the use of different RNA silencing technologies to down-regulate the expression of *StPPO* genes in potato tubers (Bachem et al., 1994; Rommens et al., 2006; Llorente et al., 2011). The approach taken in those reports was to reduce the expression of several members of the *StPPO* gene family, which led to a reduction in the enzyme content and enzymatic browning reactions. The contribution of the different members of *StPPO* genes to the total PPO activity

was latter established in potato tubers using amiRNA technology (Chi et al., 2014). Despite amiRNAs proved to be efficient in regulating the expression of *StPPO* genes individually or in combination, several off targets effects were observed with lines displaying a moderate to high reduction of non-targeted *StPPO* genes expression (Chi et al., 2014). The reduction in PPO activity was 15–95%, while the reduction in enzymatic browning was 10–65%, depending on the combination of *StPPO* genes down regulated. The greatest reduction, however, occurred when *StPPO1* to *4* were all suppressed. For unknown reasons, the authors could not obtain lines expressing the amiRNA directed to *StPPO2* gene alone. Nevertheless, correlations studies indicated that the expression of *StPPO2* gene was strongly correlated with the levels of PPO activity and enzyme content in tuber. In the present study, we have demonstrated that lines carrying mutations in all the four alleles of *StPPO2* gene displayed a reduction up to 69% and 73% in the PPO activity and enzymatic browning, respectively. Our result not only corroborate the previous report pointing out *StPPO2* as the major contributor to PPO activity in tubers, but also demonstrate that non-browning potatoes can be obtained by the sole induction of mutations in that gene, without affecting other members of the gene family. Our approach could be advantageous in order to avoid the downside effects of reducing the expression of other members of the *StPPO* gene family, affecting their potential involvement in important cell functions (Yoruk and Marshall, 2003; Jukanti, 2017). Furthermore, we have demonstrated that the CRISPR/Cas9 system is a highly efficient tool for inducing mutations in a specific member of a gene family that shares a high identity of nucleotide sequence (Thygesen et al., 1995; Chi et al., 2014).

The phenotypes observed in the selected lines, were correlated with the mutations found in the *StPPO2* gene. Thus, the frame shift mutations (deletions of 4 or 7 bp) in all alleles of the *StPPO2* gene, is the most likely cause for the reduced PPO activity and concurrent reduced enzymatic browning in line M07056. Similar phenotypic effects were observed in lines M08001, M08002, M08008, and M08027. In addition to alleles carrying mutations that produced frame shifts in the coding sequence, alleles with deletions of 111 bp were introduced in *StPPO2* of those lines. Even though this mutation is not expected to produce a frame shift, a large deletion introduced in the coding sequence near the first copper-binding domain, might affect the functionality of the enzyme, if translated. PPOs from a large number of plant species share a conserved structure in the N-terminal domain, which is critical for the function of the enzyme (García-Borrón and Solano, 2002; Marusek et al., 2006; Tran et al., 2012). Similarly, line M08002 presented one allele carrying a deletion of 9 bp, because of a deletion of 4 bp in the target site of sgRNA564 and a deletion of 5 bp in the target site of sgRNA157. Although no frame shifts were detected for the rest of the coding sequence, the frame shift in the region spanning between both target sites may be related to reduction in the enzyme activity, similar to the effect produced by the elimination of such fragment.

It is not established if all the alleles of the *StPPO2* gene contribute equally to the protein activity in the tuber. Based on our sequencing results, line M08003 contained at least one mutated allele of *StPPO2* most likely in combination with multiple copies of wild type allele. This line displayed a PPO activity and enzymatic browning levels almost identical to the control Desiree RC, which indicates that the remaining wild type alleles were sufficient for a normal enzyme function. Previous studies in other plant species have shown allelic variations in *PPO* genes to be associated with differences in the levels of PPO activity (Taketa et al., 2010; Beecher et al., 2012). A more detailed analysis is needed in the case of potato *PPO* genes.

For phenotypic characterization, we performed two analyses on selected edited lines, i.e. Relative Enzymatic Browning and Relative PPO Activity. The two methods produced similar results, with the selected edited lines displaying a reduction in both parameters. In addition, both variables presented a significantly positive correlation between them. Even though a clear reduction in relative PPO activity for lines M07056, M08008, and M08027, no statistical differences were observed relative to the control. This may be due to a higher variance of the values for these lines. Despite the small disparity in the statistical analysis between the variables for the mentioned lines, our results were consistent with all lines displaying a reduced enzyme activity, which turned to a reduced enzymatic browning in the tuber. Lines M08001 and M08002 values were statistically significant with both determinations.

Argentina is one of a few countries to develop legislation to assess regulatory matters regarding genome edited organisms (Whelan and Lema, 2015; Eckerstorfer et al., 2019; Lema, 2019). In this report, we have studied the application of the CRISPR/Cas9 system to produce edited potato plants with a reduced PPO activity and enzymatic browning in tubers. Our system proved to be specific for the target gene, without affecting the coding sequence of other *StPPO* family members and, consequently, their roles in other cell functions. Considering the current criteria for the determination of the regulatory status of genome edited crops in Argentina and other countries, application of this technology could result in plants that do not fall under strict GMOs regulation, which might represent a major advantage in comparison with previous strategies taken for the improvement of the same trait in potato. We consider that our study represents an important step towards the development of potato varieties that maintain the organoleptic, antioxidant and nutritional properties during harvest and post-harvest procedures, without the utilization of potentially harmful browning controlling agents. This advantage results in benefits for the farmer, the potato processing industry, and finally the consumer.

MATERIALS AND METHODS

SgRNA Design on *StPPO2* Gene of *S. tuberosum* cv. Desiree

The available sequences of PGSC0003DMG400018916 (Potato Genome Sequencing Consortium, 2011) and POT32 (GenBank: U22921.1, Thygesen et al., 1995) were aligned and used for primer design, in order to amplify the *StPPO2* gene in *S. tubersoum* cv Desiree (Chi et al., 2014). Primers F_StPPO2 and R_StPPO2

(**Supplementary Table S1**) were used to amplify a fragment from the 5′ end of the target gene, using 10 ng of genomic DNA as a template in a reaction with Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, MA, USA). Reaction conditions were 98°C for 1 min, 30 cycles of 98°C 30 s, 60°C 20 s, 72°C 30 s and a final extension of 72°C for 7 min.

PCR products were cloned into the pJET1.2 vector using the CloneJET PCR Cloning Kit (Thermo Fisher Scientific) and transformed to One Shot TOP10 Chemically Competent *E. coli* (Thermo Fisher Scientific), according to manufacturer instructions. Twelve randomly picked colonies were selected for plasmid purification and Sanger sequencing using the primers provided by the CloneJET PCR Cloning Kit. The resulting sequences were aligned to avoid allelic variation during sgRNA design and further High Resolution Fragment Analysis (HRFA) primer design (**Supplementary Figure S1**).

The Cas-Designer Tool[1] was used for sgRNA design, using one of the sequences obtained for *StPPO2* as a query and *S. tuberosum* (PGSC v4.03) as a target genome (Park et al., 2015). sgRNA157 and sgRNA564 (**Figure 1A**) were selected according to the Out of Frame Score (Bae et al., 2014) and the strict absence of allelic variation along the target sequence (**Supplementary Figure S1**).

Off Target Sites Prediction on *StPPO* Genes

Cas-OFFinder Tool[2] was used for possible off targets site identification on other members of the *StPPO* gene family within the genome database of *S. tuberosum* (PGSC v4.03). Searching for sequences with up to 4 mismatches (Hahn and Nekrasov, 2019) with the selected sgRNAs and a 5′-NRG-3′(R = A or G) as PAM sequence, putative off targets were found in the genome at positions 45631511 and 45870133 of the chromosome 8 of potato for sgRNA564. Using the genome browser available on Sol Genomics Network[3] genes PGSC0003DMG400029575 (*StPPO1*) and PGSC0003DMG400018917 (*StPPO4*) were identified as the only two putative off targets on *StPPO* genes with expression data (Potato Genome Sequencing Consortium, 2011; Chi et al., 2014).

The available sequences of both genes were used for primer design (**Supplementary Table S1**) in order to sequence the putative off targets sites in *S. tuberosum* cv Desiree and confirm the *in silico* analysis. Amplification, cloning and sequencing was performed as previously explained for the target gene. The resulting sequences were aligned (**Supplementary Figures S2** and **S3**) and used for HRFA primer design (**Supplementary Table S1**) for off target analysis. The amplified region for HRFA on each gene included both, the predicted off target sites for sgRNA564 (**Figure 1B**) and the region that aligns with sgRNA157 (**Figure 1C**).

Ribonucleoprotein Complexes Assembly

The sgRNAs were *in vitro* transcribed (Andersson et al., 2018) using the GeneArt Precision gRNA Synthesis Kit (Thermo Fisher Scientific), according to the manufacturer instructions, with minor modifications. The DNA templates for *in vitro* transcription of sgRNA157 and sgRNA564, were obtained using Fw_IVT157/Rv_IVT157 primers and Fw_IVT564/Rv_IVT564 primers, respectively (**Supplementary Table S1**). After assembly, both DNA templates were purified using the GeneJET PCR Purification Kit (Thermo Fisher Scientific) and quantified using a Trinean DropSense 16 (Techtum, Nacka, Sweden). Thirty ng of DNA were used in each case for *in vitro* transcription for 3 h. After transcription, sgRNAs were treated with 1 unit of DNAse I for 15 min following the instructions of GeneArt Precision gRNA Synthesis Kit, afterwards purified, and quantified using the Trinean DropSense 16.

Right before transfections into potato protoplasts, 5 μg of each sgRNA was mixed with 0.03 nmol of GeneArt Platinum Cas9 Nuclease (Thermo Fisher Scientific) in a final volume of 5 μl and incubated for 15 min at room temperature.

Protoplasts Transfection and Plant Regeneration

Protoplasts were isolated from 5-week old plantlets according to Nicolia et al. (2015). For transfections, 100,000 protoplasts were incubated with RNPs and 25% Polyethylenglycol (PEG) 4000 (Duchefa Biochemie, Haarlem, The Netherlands) for 3 min, or with RNPs and 40% PEG4000 for 30 min. A regeneration control was included, which consisted of the same number of protoplasts incubated with 40% PEG but no RNPs, for 30 min. After transfections, all protoplasts were embedded in sodium alginate and cultured for calli regeneration, according to Nicolia et al. (2015).

Green calli were released from alginate blobs after 21 days of culture, and subcultured for shoot growth induction. To ensure the analysis of independent lines, one shoot was picked per callus and transferred for root development. Samples from leaves of the full regenerated plantlets were picked for genomic DNA extraction and further analysis.

Identification of Edited Lines and Sequencing Analysis

Genomic DNA of regenerated plants was extracted from leaves in a 96-Deep well plate. The sampled tissue was homogenized with 500 μl of 100 mM Tris HCl, 50 mM EDTA and 1% SDS, pH 9.0 and 5mm steel beads, using a Retsch Mixer Mill MM400 for 30 s at 30 Hz (Retsch, Haan, Germany). After centrifugation of the tissue debris, DNA was extracted from 200 μl of the cleared lysate, in a QIAcube HT extraction robot using a QIAamp 96 DNA QIAcube HT Kit (QIAGEN, Hilden, Germany) according to the manufacturer instructions.

The presence of mutations in the target gene was determined by High Resolution Fragment Analysis (HRFA), according to Andersson et al. (2017). Primers PPO2_2Bf-HEX and PPO2_2Br (**Supplementary Table S1**) were designed for amplification of the region spanning both sgRNAs target sites on the *StPPO2* gene, taking into account the absence of allelic variation in primers annealing sites in the target gene (**Supplementary Figure S1**). Primers were used to amplify a fragment of 228 bp

[1] www.rgenome.net/cas-designer
[2] www.rgenome.net/cas-offinder
[3] www.solgenomics.net

of the target gene, using Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific). Reaction conditions were 98°C for 1 min, 30 cycles of 98°C 30 s, 60°C 20 s, 72°C 15 s, and a final extension of 72°C for 7 min.

Labelled PCR products were analyzed in an Applied Biosystems 3500 Genetic Analyzer (Thermo Fisher Scientific), according to the instructions of manufacturer, using GeneScan 600 LIZ Dye Size Standard (Thermo Fisher Scientific) as internal lane size standard. Fragments length were determined with GeneMarker Software (SoftGenetics, State College, PA, USA) and insertions or deletions were identified comparing each line electropherogram versus the control.

StPPO2 gene was sequenced by Sanger in selected edited lines to confirm the HRFA results. Primers PPO2_2Bf and PPO2_2Br (**Supplementary Table S1**) were used for PCR amplification of the fragment with the same conditions mentioned above, and the products cloned using the CloneJET PCR Cloning Kit (Thermo Fisher Scientific), as previously. Twelve randomly picked clones were sequenced per line for mutations characterization.

Off Target Analysis

The presence of putative off target mutations in *StPPO1* and *StPPO4* genes was determined by HRFA as described above. Primers PPO1_OT564_F-6-FAM and PPO1_OT564_R (**Supplementary Table S1**) were used for the analysis of the *StPPO1* gene (**Supplementary Figure S2**). PPO4_OT564_F-HEX and PPO4_OT564_R primers (**Supplementary Table S1**) were used for the analysis of the *StPPO4* gene (**Supplementary Figure S3**).

Plant Growth Conditions and Tubers Harvesting

Selected *in vitro*-regenerated plantlets were transferred to 1 L pots with substrate and placed in a growth chamber, at a constant temperature of 24°C in a photoperiod of 16 h (120 μmol m^{-2} s^{-1}) light–8 h dark. Three biological replicates were grown for each edited line and the control line Desiree RC. Tubers were harvested after 120 days of culture, right before plants senescence.

Enzymatic Browning and PPO Activity Determinations

Enzymatic Browning and PPO activity were measured according to Chi et al. (2014), with minor modifications. Tubers were randomly selected per each edited line and the control Desiree RC and triple biological replicates were used for the determinations.

For enzymatic browning assay, slices were manually cut from the center of the tubers and immediately frozen in liquid nitrogen. The frozen samples were processed with 5 ml of cold PPO extraction buffer (100 mM sodium phosphate buffer pH 6.0, 2% TX-100, 2% PVPP) using an Ultra-Turrax T-25 (IKA, Königswinter, Germany) at 11,000 rpm for 30 s. Homogenates were allowed to oxidize for 1 h at room temperature, and afterwards aliquots were transferred to 1.5 ml centrifuge tubes and centrifuged for 10 min at 11,000 rpm. The absorbance at 475 nm (A475nm) was measured in 300 μl of the supernatant in a 96 wells plate, using an Epoch Microplate Spectrophotometer (Bioteck, Winooski, VT, USA), with three technical replicates. The total protein concentration of the homogenates were determined using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific) in the same spectrophotometer and the Enzymatic Browning calculated as the A475nm/mg of total protein. Finally, the Relative Enzymatic Browning was calculated as the value of each line related to the control Desiree RC.

For PPO activity assay, the frozen samples were processed with 5 ml of cold PPO extraction buffer as above, and the homogenates were transferred to 2 ml centrifuge tubes and centrifuged at 11,000 rpm, 4°C during 30 min. The supernatants were transferred to new tubes and kept in ice until PPO activity measurements. PPO activity was measured adding 100 μl of sample into a quartz cuvette and 900 μl of PPO assay buffer (50 mM sodium phosphate buffer pH 6.0, 0.1% SDS and 15 mM 4-Methylcatechol). A SmartSpec3000 Spectrophotometer (Bio-Rad, Hercules, CA, USA) was used to measure the absorbance increase at 400 nm (A400nm) every 5 s for 1 min at 25°C. Three technical replicates were performed for the determinations and one unit (1U) of PPO enzymatic activity was defined as the amount of enzyme necessary to change A400nm in 0.001/min at 25°C. The total protein concentration of each sample was determined using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific) as before, and enzymatic activity was calculated as U/mg of total protein. Finally, Relative PPO Activity was calculated as the value of each line related to the control Desiree RC.

Statistical Analysis

Linear Mixed Models were used to test the effect of the different lines in the Relative PPO Activity and Relative Enzymatic Browning variables (Linear Mixed-Effects Models: Basic Concepts and Examples, 2000). We considered each line as a fixed-effect. In order to take in consideration possible variation in the individual plants, we considered the identity of each biological replicate as a random effect. All possible models were evaluated and we compared competitive ones using Akaike Information Criterion (AIC). All the analysis were performed in R[4] using nlme package.

Spearman's correlation analysis was performed in R, using the measured data of the variables "Relative Enzymatic Browning" and "Relative PPO Activity."

DATA AVAILABILITY STATEMENT

All datasets for this study are included in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

MG, GM, MA, SF, and PH designed the study. MG, MA, HT, NO, and A-SF planned and conducted the protoplasts

[4] www.R-project.org

transfection, plant regeneration and mutations analysis. MG, GM, LS, and CD planned and conducted the tuber production and phenotype analysis. MG wrote the manuscript, which was revised by GM, MA, SF, and PH. All authors read and approved the final version of the manuscript.

FUNDING

This work was funded by INTA PNBIO1131024 "Desarrollo de sistemas alternativos de generación y utilización de variabilidad genética y su aplicación al mejoramiento de los cultivos" and INTA-Fondo de Valorización Tecnológica "Variedades de papa editadas con mayor calidad industrial y nutricional". In addition, this study was partially financed by Trees and Crops for the Future (TC4F), a Strategic Research Area at SLU, supported by the Swedish Government. MG visit to SLU was possible thanks to a training grant provided by Becar program from the Argentinian Ministry of Education and Sports.

ACKNOWLEDGMENTS

We want to thank Professor Anders Carlsson and all the staff of the Department of Plant Breeding of the Swedish University of Agricultural Sciences (SLU-Alnarp) for providing the facilities and a warm welcome to MG during this work. We want to thank Martín Castellote and Antonella Gorosábel for advising in genomic and statistical data analysis, respectively, and to Silvina Divito and Pia Ohlsson for technical assistance at the labs. Finally, we want to thank to the research group of "Patología Vegetal" of INTA-EEA Balcarce for providing the growth chambers for plant and tuber production.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at:https://www.frontiersin.org/articles/10.3389/fpls.2019.01649/full#supplementary-material

REFERENCES

Andersson, M., Turesson, H., Nicolia, A., Fält, A.-S., Hofvander, P., and Samuelsson, M. (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

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

Arora, L., and Narula, A. (2017). Gene editing and crop improvement using CRISPR-Cas9 system. *Front. Plant Sci.* 8, 1932. doi: 10.3389/fpls.2017.01932

Bachem, C. W. B., Speckmann, G. J., Van der Linde Piet, C. G., Verheggen, F. T. M., Hunt, M. D., Steffens, J. C., et al. (1994). Antisense expression of polyphenol oxidase genes inhibits enzymatic browning in potato tubers. *Bio/Technology* 12, 1101–1105. doi: 10.1038/nbt1194-1101

Bae, S., Kweon, J., Kim, H. S., and Kim, J.-S. (2014). Microhomology-based choice of Cas9 nuclease target sites. *Nat. Methods* 11, 705–706. doi: 10.1038/nmeth.3015

Baltes, N. J., Gil-Humanes, J., and Voytas, D. F. (2017). "Chapter one - genome engineering and agriculture: opportunities and challenges," in *Gene editing in plants progress in molecular biology and translational science*. Eds. D. P. Weeks and B. Yang (Academic Press), 1–26. doi: 10.1016/bs.pmbts.2017.03.011

Beecher, B. S., Carter, A. H., and See, D. R. (2012). Genetic mapping of new seed-expressed polyphenol oxidase genes in wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 1463–1473. doi: 10.1007/s00122-012-1801-2

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

Chi, M., Bhagwat, B., Lane, W. D., Tang, G., Su, Y., Sun, R., et al. (2014). Reduced polyphenol oxidase gene expression and enzymatic browning in potato (*Solanum tuberosum* L.) with artificial microRNAs. *BMC Plant Biol.* 14, 1–18. doi: 10.1186/1471-2229-14-62

Coetzer, C., Corsini, D., Love, S., Pavek, J., and Tumer, N. (2001). Control of enzymatic browning in potato (Solanum tuberosum L.) by sense and antisense RNA from tomato polyphenol oxidase. *J. Agric. Food Chem.* 49, 652–657. doi: 10.1021/jf001217f

Eckerstorfer, M. F., Engelhard, M., Heissenberger, A., and Simon, S. (2019). Plants developed by new genetic modification techniques — comparison of existing regulatory frameworks in the EU and non-EU countries. *Front. Bioeng. Biotechnol.* 7, 26. doi: 10.3389/fbioe.2019.00026

Gao, C. (2018). The future of CRISPR technologies in agriculture. *Nat. Rev. Mol. Cell Biol.* 19, 275. doi: 10.1038/nrm.2018.2

García-Borrón, J. C., and Solano, F. (2002). Molecular anatomy of tyrosinase and its related proteins: Beyond the histidine-bound metal catalytic center. *Pigment Cell Res.* 15, 162–173. doi: 10.1034/j.1600-0749.2002.02012.x

Hahn, F., 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

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA – guided. *Science* (80-.), 816–822. doi: 10.1126/science.1225829

Jukanti, A. (2017). "Distribution, localization, and structure of plant polyphenol oxidases (PPOs)," in *Polyphenol oxidases (PPOs) in plants*. (Singapore: Springer Singapore), 11–32. doi: 10.1007/978-981-10-5747-2_2

Kampatsikas, I., Bijelic, A., and Rompel, A. (2019). Biochemical and structural characterization of tomato polyphenol oxidases provide novel insights into their substrate specificity. *Sci. Rep.* 9, 1–13. doi: 10.1038/s41598-019-39687-0

Kapusi, E., Corcuera-Gómez, M., Melnik, S., and Stoger, E. (2017). Heritable genomic fragment deletions and small indels in the putative ENGase gene induced by CRISPR/Cas9 in Barley. *Front. Plant Sci.* 8, 1–11. doi: 10.3389/fpls.2017.00540

Kumlehn, J., Pietralla, J., Hensel, G., Pacher, M., and Puchta, H. (2018). The CRISPR/Cas revolution continues: from efficient gene editing for crop breeding to plant synthetic biology. *J. Integr. Plant Biol.* 60, 1127–1153. doi: 10.1111/jipb.12734

Lema, M. A. (2019). Regulatory aspects of gene editing in Argentina. *Transgenic Res.* 28, 147–150. doi: 10.1007/s11248-019-00145-2

Li, L., and Steffens, J. C. (2002). Overexpression of polyphenol oxidase in transgenic tomato plants results in enhanced bacterial disease resistance. *Planta* 215, 239–247. doi: 10.1007/s00425-002-0750-4

Li, J., Manghwar, H., Sun, L., Wang, P., Wang, G., Sheng, H., et al. (2019). Whole genome sequencing reveals rare off-target mutations and considerable inherent genetic or/and somaclonal variations in CRISPR/Cas9-edited cotton plants. *Plant Biotechnol. J.* 17, 858–868. doi: 10.1111/pbi.13020

Liang, Z., Chen, K., Li, T., Zhang, Y., Wang, Y., Zhao, Q., et al. (2017). Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. *Nat. Commun.* 8, 1–5. doi: 10.1038/ncomms14261

Llorente, B., Alonso, G. D., Bravo-Almonacid, F., Rodríguez, V., López, M. G., Carrari, F., et al. (2011). Safety assessment of nonbrowning potatoes: opening the discussion about the relevance of substantial equivalence on next generation biotech crops. *Plant Biotechnol. J.* 9, 136–150. doi: 10.1111/j.1467-7652.2010.00534.x

Marusek, C. M., Trobaugh, N. M., Flurkey, W. H., and Inlow, J. K. (2006). Comparative analysis of polyphenol oxidase from plant and fungal species. *J. Inorg. Biochem.* 100, 108–123. doi: 10.1016/j.jinorgbio.2005.10.008

Mayer, A. M. (2006). Polyphenol oxidases in plants and fungi: going places? *Rev. Phytochem.* 67, 2318–2331. doi: 10.1016/j.phytochem.2006.08.006

Nadakuduti, S. S., Buell, C. R., Voytas, D. F., Starker, C. G., and Douches, D. S. (2018). Genome editing for crop improvement – applications in clonally propagated polyploids with a focus on potato Solanum tuberosum L. *Front. Plant Sci.* 9, 1–11. doi: 10.3389/fpls.2018.01607

Nicolia, A., Proux-Wéra, E., Åhman, I., Onkokesung, N., Andersson, M., Andreasson, E., et al. (2015). Targeted gene mutation in tetraploid potato through transient TALEN expression in protoplasts. *J. Biotechnol.* 204, 17–24. doi: 10.1016/j.jbiotec.2015.03.021

Park, J., Bae, S., and Kim, J. (2015). Sequence analysis Cas-designer : a web-based tool for choice of CRISPR-Cas9 tar- get sites. *Bioinformatics* 31, 1–3. doi: 10.1101/005074.Bae

Pinheiro, J., and Bates, D. (2000). *Linear Mixed-Effects Models: Basic Concepts and Examples* (New York, NY: Springer New York). doi: 10.1007/0-387-22747-4_1

Potato Genome Sequencing Consortium (2011). Genome sequence and analysis of the tuber crop potato. *Nature* 475, 189–195. doi: 10.1038/nature10158

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

Rommens, C. M., Ye, J., Richael, C., and Swords, K. (2006). Improving potato storage and processing characteristics through all-native DNA transformation. *J. Agric. Food Chem.* 54, 9882–9887. doi: 10.1021/jf062477l

Scheben, A., Wolter, F., Batley, J., Puchta, H., and Edwards, D. (2017). Towards CRISPR/Cas crops - bringing together genomics and genome editing. *New Phytol.*, 682–698. doi: 10.1111/nph.14702

Shahidi, F., and Ambigaipalan, P. (2015). Phenolics and polyphenolics in foods, beverages and spices : antioxidant activity and health effects. *J. Funct. Foods* 18, 820–897. doi: 10.1016/j.jff.2015.06.018

Shi, J., Gao, H., Wang, H., Lafitte, H. R., Archibald, R. L., Yang, M., et al. (2017). ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. *Plant Biotechnol. J.* 15, 207–216. doi: 10.1111/pbi.12603

Svitashev, S., Schwartz, C., Lenderts, B., Young, J. K., and Cigan, A. M. (2016). ribonucleoprotein complexes. *Nat. Commun.* 7, 1–7. doi: 10.1038/ncomms13274

Taketa, S., Matsuki, K., Amano, S., Saisho, D., Himi, E., and Shitsukawa, N. (2010). Duplicate polyphenol oxidase genes on barley chromosome 2H and their functional differentiation in the phenol reaction of spikes and grains. *J. Exp. Bot.* 61, 3983–3993. doi: 10.1093/jxb/erq211

Thipyapong, P., Hunt, M. D., and Steffens, J. C. (2004). Antisense downregulation of polyphenol oxidase results in enhanced disease susceptibility. *Planta* 220, 105–117. doi: 10.1007/s00425-004-1330-6

Thygesen, P. W., Dry, I. B., and Robinson, S. P. (1995). Polyphenol oxidase in potato (a multigene family that exhibits differential expression patterns). *Plant Physiol.* 109, 525–531. doi: 10.1104/pp.109.2.525

Tinello, F., and Lante, A. (2018). Recent advances in controlling polyphenol oxidase activity of fruit and vegetable products. *Innov. Food Sci. Emerg. Technol.* 50, 73–83. doi: 10.1016/j.ifset.2018.10.008

Tran, L. T., Taylor, J. S., and Constabel, C. P. (2012). The polyphenol oxidase gene family in land plants: Lineage-specific duplication and expansion. *BMC Genomics*. 13, 395. doi: 10.1186/1471-2164-13-395

Tuncel, A., Corbin, K. R., Ahn-Jarvis, J., Harris, S., Hawkins, E., Smedley, M. A., et al. (2019). Cas9-mediated mutagenesis of potato starch-branching enzymes generates a range of tuber starch phenotypes. *Plant Biotechnol. J.* 17, 2259–2271. doi: 10.1111/pbi.13137

Veillet, F., Chauvin, L., Kermarrec, M. P., Sevestre, F., Merrer, M., Terret, Z., et al. (2019). The Solanum tuberosum GBSSI gene: a target for assessing gene and base editing in tetraploid potato. *Plant Cell Rep.* 38, 1065–1080. doi: 10.1007/s00299-019-02426-w

Wang, J., and Constabel, C. P. (2004). Polyphenol oxidase overexpression in transgenic Populus enhances resistance to herbivory by forest tent caterpillar (Malacosoma disstria). *Planta* 220, 87–96. doi: 10.1007/s00425-004-1327-1

Whelan, A. I., and Lema, M. A. (2015). Regulatory framework for gene editing and other new breeding techniques (NBTs) in Argentina. *GM Crops Food* 6, 253–265. doi: 10.1080/21645698.2015.1114698

Woo, J. W., Kim, J., Kwon, S. I., Corvalán, C., Cho, S. W., Kim, H., et al. (2015). DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat. Biotechnol.* 33, 1162–1165. doi: 10.1038/nbt.3389

Yoruk, R. R., and Marshall, M. (2003). Physicochemical properties and function of plant polyphenol oxidase: a review1. *J. Food Biochem.* 27, 361–422. doi: 10.1111/j.1745-4514.2003.tb00289.x

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, 1–9. doi: 10.1038/s41598-017-12262-1

Zhang, Q., Xing, H. L., Wang, Z. P., Zhang, H. Y., Yang, F., Wang, X. C., et al. (2018a). Potential high-frequency off-target mutagenesis induced by CRISPR/Cas9 in Arabidopsis and its prevention. *Plant Mol. Biol.* 96, 445–456. doi: 10.1007/s11103-018-0709-x

Zhang, Z., Wang, J., Zhang, X., Shi, Q., Xin, L., Fu, H., et al. (2018b). Effects of radio frequency assisted blanching on polyphenol oxidase, weight loss, texture, color and microstructure of potato. *Food Chem.* 248, 173–182. doi: 10.1016/j.foodchem.2017.12.065

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

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 04 February 2020
doi: 10.3389/fpls.2020.00010

Generation and Molecular Characterization of CRISPR/Cas9-Induced Mutations in 63 Immunity-Associated Genes in Tomato Reveals Specificity and a Range of Gene Modifications

Ning Zhang[1†], *Holly M. Roberts*[1†], *Joyce Van Eck*[1,2†] *and Gregory B. Martin*[1,3*†]

[1] Boyce Thompson Institute for Plant Research, Ithaca, NY, United States, [2] Plant Breeding and Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, NY, United States, [3] Plant Pathology and Plant-Microbe Biology Section, School of Integrative Plant Science, Cornell University, Ithaca, NY, United States

The CRISPR/Cas9 system is a powerful tool for targeted gene editing in many organisms including plants. However, most of the reported uses of CRISPR/Cas9 in plants have focused on modifying one or a few genes, and thus the overall specificity, types of mutations, and heritability of gene alterations remain unclear. Here, we describe the molecular characterization of 361 T0 transgenic tomato plants that were generated using CRISPR/Cas9 to induce mutations in 63 immunity-associated genes. Among the T0 transformed plants, 245 carried mutations (68%), with 20% of those plants being homozygous for the mutation, 30% being heterozygous, 32% having two different mutations (biallelic), and 18% having multiple mutations (chimeric). The mutations were predominantly short insertions or deletions, with 87% of the affected sequences being smaller than 10 bp. The majority of 1 bp insertions were A (50%) or T (29%). The mutations from the T0 generation were stably transmitted to later generations, although new mutations were detected in some T1 plants. No mutations were detected in 18 potential off-target sites among 144 plants. Our study provides a broad and detailed view into the effectiveness of CRISPR/Cas9 for genome editing in an economically important plant species.

Keywords: CRISPR/Cas9, genome editing, immunity-associated genes, tomato, Off-target mutation

Edited by:
Hiroshi Ezura,
University of Tsukuba, Japan

Reviewed by:
Yiping Qi,
University of Maryland, College Park, United States
Bing Yang,
University of Missouri, United States

****Correspondence:***
Gregory B. Martin
gbm7@cornell.edu

†ORCID:
Ning Zhang
orcid.org/0000-0003-2775-1755
Holly M. Roberts
orcid.org/0000-0002-8640-9469
Joyce Van Eck
orcid.org/0000-0002-8005-365X
Gregory B. Martin
orcid.org/0000-0003-0044-6830

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *08 November 2019*
Accepted: *07 January 2020*
Published: *04 February 2020*

Citation:
Zhang N, Roberts HM, Van Eck J and Martin GB (2020) Generation and Molecular Characterization of CRISPR/Cas9-Induced Mutations in 63 Immunity-Associated Genes in Tomato Reveals Specificity and a Range of Gene Modifications. Front. Plant Sci. 11:10. doi: 10.3389/fpls.2020.00010

INTRODUCTION

Derived from a native adaptive immune system in eubacteria and archaea, the CRISPR/Cas system enables the alteration of DNA sequences in many organisms to achieve precise gene modifications (Jaganathan et al., 2018). The most widely used *Streptococcus pyogenes* Cas9 (SpCas9) requires the 20-bp spacer sequence of a guide RNA (gRNA) to recognize a complementary target DNA site upstream of a protospacer adjacent motif (PAM) and generates a double-stranded break (DSB) near the target

region (Xie and Yang, 2013). DSBs are repaired through either non-homologous end joining (NHEJ) or homology-directed recombination (HDR) resulting in small insertions/deletions (indels) or substitutions at the target region, respectively (Jinek et al., 2012). Compared to other genome editing tools such as zinc finger nucleases (ZFNs; Kim et al., 1996) and transcription activator-like effector nucleases (TALENs; Bogdanove and Voytas, 2011), CRISPR/Cas is more robust in that the Cas protein can theoretically bind to any genomic region preceding a PAM site and, importantly, target multiple sites simultaneously. However, the possibility of off-target mutations caused by CRISPR/Cas is a potential concern in both basic and applied research in plants, although it has been reported that off-target effects of CRISPR/Cas occur at a much lower frequency in plants than in mammals (Fu et al., 2013; Kuscu et al., 2014). The most effective way to minimize off-target mutations is to select a gRNA target with little or no homology to other genomic regions (Baltes and Voytas, 2015). Other methods to reduce off-target mutations include using paired Cas9 nickases (Ran et al., 2013) or paired fusions of a catalytically dead Cas9 nuclease to the *Fok*I cleavage domain (Guilinger et al., 2014; Tsai et al., 2014).

As one of the most important vegetable crops in the world (Kimura and Sinha, 2008), tomato (*Solanum lycopersicum*) is an important source of health-promoting nutrients including vitamin C and E, minerals, and carotenes such as ß-carotene and lycopene (Canene-Adams et al., 2005). However, tomato production is threatened by many infectious diseases, including bacterial speck disease caused by *Pseudomonas syringae* pv. tomato (*Pst*), which can result in severe economic losses due to reduced yield and quality (Xin and He, 2013). A large number of candidate immunity-associated genes have been identified in tomato, but validation of the functional importance of these genes had been technically challenging before the emergence of CRISPR/Cas technology (Pedley and Martin, 2003; Oh and Martin, 2011; Rosli et al., 2013; Pombo et al., 2014). Although CRISPR/Cas has been used to modify genes with key roles in growth, development, and biotic and abiotic stresses in plants (Ito et al., 2015; Jacobs et al., 2017; Nekrasov et al., 2017; Rodriguez-Leal et al., 2017; Shimatani et al., 2017; Yu et al., 2017; D'Ambrosio et al., 2018; Hashimoto et al., 2018; Li et al., 2018a; Li et al., 2018b; Tashkandi et al., 2018; Zsogon et al., 2018; Ortigosa et al., 2019; Wang et al., 2019; Xu et al., 2019), all of the reported studies have focused on one or a few genes and could not provide broad insights into the specificity, types of mutations and heritability of genome editing by CRISPR/Cas9 in tomato.

Recently, we developed the Plant Genome Editing Database (PGED; http://plantcrispr.org/cgi-bin/crispr/index.cgi) which provides information about a collection of tomato lines with CRISPR/Cas9-induced mutations in immunity-associated genes (Zheng et al., 2019). In the present study, we molecularly characterized 361 T0 transgenic tomato plants that were generated using CRISPR/Cas9 to induce mutations in 63 candidate immunity-associated genes. To enhance the mutation efficiency and reduce the number of transformations needed, we evaluated gRNA efficiency by transient expression in tomato leaves and conducted tomato transformation with *Agrobacterium* pools containing 2–4 Cas9/gRNA constructs. This initial evaluation of gRNAs allowed us to select the most efficient ones for tomato stable transformation while using "*agrobacterium* pools" with various Cas9/gRNAs constructs shortened the time for generating multiple tomato mutant lines. We established an efficient CRISPR/Cas9 system to generate a large number of primary transgenic lines and we report for the first time a systematic investigation of the specificity of targeting, the types of mutations generated and the heritability of the mutations through multiple generations of tomato. Our CRISPR/Cas9-induced tomato mutant plants provide a powerful resource for better understanding the molecular mechanisms of plant-microbe interactions in the future.

MATERIALS AND METHODS

Guide RNA (gRNA) Design and Construct Development

All 20-nt gRNAs specific for the target genes were designed using the software Geneious R11 as described previously (Jacobs et al., 2017). The tomato (*Solanum lycopersicum*) reference genome sequence (SL2.5 or SL3.0) was used as an off-target database to score each gRNA (GN_{19} or gN_{19}; "g" represents a manually added "G" to accommodate the transcription initiation requirement of the U6 promoter if the first nucleotide is not a G at target sites) preceding a PAM (NGG) sequence. For each gene, 2–3 gRNA targets with minimum off-target scores were designed. Single or multiple gRNA cassettes were cloned into a binary vector p201N: Cas9 by Gibson assembly as described previously (Jacobs and Martin, 2016). Colonies containing correct gRNA sequences were confirmed by PCR and Sanger sequencing.

Evaluation of gRNA Efficiency by Agroinfiltration in Tomato and *Nicotiana benthamiana* Leaves

Each Cas9/gRNA vector was transformed into the *Agrobacterium tumefaciens* strain 1D1249 (Wroblewski et al., 2005) by electroporation. For agroinfiltration into tomato leaves, the bacterial cells containing different gRNA plasmids were grown in liquid YEP medium with 150 mg/L kanamycin overnight at 30°C. The bacterial pellet was collected and resuspended in an induction buffer containing 10 mM $MgCl_2$, 10 mM MES (pH 5.7), and 200 μM acetosyringone (Sigma-Aldrich). Bacterial suspensions were adjusted to OD_{600} = 0.3 and incubated at room temperature for 2–5 h. The third and fourth leaves of 4-week-old tomato plants were infiltrated with needleless syringes and the whole plant was then placed in a growth chamber with a temperature of 22°C–24°C, 16 h light/8 h dark photoperiod and 65% relative humidity. Three days later, a pool of six leaf discs were collected from three individual plants (two leaf discs from each of three plants) that had been infiltrated with the tested Cas9/gRNA vector, and used for genomic DNA extraction, PCR and sequencing. The web-based tool TIDE (https://tide.deskgen.com) was used to determine the mutation frequency induced by corresponding Cas9/gRNA vectors.

Tomato Transformation

Tomato transformation was performed either at the plant transformation facility at the Boyce Thompson Institute (BTI) or North Carolina State University (NCSU) (Gupta and Van Eck, 2016; Van Eck et al., 2019). Modifications of the transformation methods were made for Rio Grande (RG), including using 100 mg/L kanamycin for selection, and adding 0.1 mg/L indole-3-acetic acid (IAA) to the plant regeneration media (2Z, 1Z) and 1 mg/L IAA to the rooting medium. Each Cas9/gRNA vector was first electrotransformed into *Agrobacterium tumefaciens* LBA4404 (BTI), AGL1 (BTI), or GV3101(pMP90) (NCSU). In most cases, 2–4 *Agrobacterium* culture preparations (of the same *Agrobacterium* strain), each carrying a different Cas9/gRNA construct, were pooled together and used for transformation to minimize the number of experiments. Tomato genotypes RG-PtoR or RG-prf3 were used for transformation if not specifically labeled (**Table 1**).

Genotyping and Mutation Analysis

Genomic DNA was extracted from tomato cotyledons or young leaves using a modified CTAB method (Murray and Thompson, 1980). The existence of T-DNA was confirmed by PCR using primers listed in **Table S6**. To determine the mutation specificity, genomic regions flanking the target site of each gene were amplified with specific primers (see PGED; http://plantcrispr.org/) and sequenced by Sanger sequencing. TIDE was used to rapidly evaluate the mutated allelic sequences using the sequencing files (.ab1 format), especially for PCR amplicons of biallelic, heterozygous, or chimeric mutants whose mutation length and frequency cannot be determined due to superimposed sequencing chromatograms.

Off-Target Evaluation

To evaluate potential off-target mutations caused by gRNAs in CRISPR-induced mutant plants, twelve gRNAs were selected and used as queries to search for potential off-target sites across the tomato genome with up to 4-nt mismatches and 1-nt indel by the software Geneious R11 or with up to 3-nt mismatches by a web tool Cas-OFFinder. Each off-target site was given a score based on how similar it was to the spacer sequence of gRNAs. A higher score for an off-target site indicated a higher similarity to the original target site and a higher likelihood to cause off-target mutations. A shortlist of potential off-target sites of each gRNA queried was generated by selecting their relatively high-scoring off-target sites predicted by either Geneious R11 or Cas-OFFinder (**Table 3**). Similar to mutation genotyping described above, genomic regions flanking the putative off-target sites were amplified with specific primers (**Table S7**) and PCR amplicons were sequenced to detect if off-target mutations were induced in those regions.

RESULTS

CRISPR/Cas9 Targeting of Immunity-Associated Genes in Tomato

To study the efficiency and specificity of genome editing in tomato by CRISPR/Cas9 and to better understand plant-pathogen interactions, we generated a collection of tomato lines with targeted CRISPR/Cas9-induced mutations in genes that have been implicated in the immune response. Candidate genes were selected based on previous studies involving RNA-Seq, biochemical approaches, virus-induced gene silencing (VIGS), or yeast two-hybrid (Y2H) screens (Zeng et al., 2012; Rosli et al., 2013; Pombo et al., 2014; Giska and Martin, 2019); orthologs of immunity-associated genes reported in other plant species such as Arabidopsis and rice were also included (Shimizu et al., 2010; Hutin et al., 2015; Xin et al., 2016; Yamada et al., 2016; Stegmann et al., 2017).

For each candidate gene, at least two gRNAs targeting different DNA sites were designed and separately cloned into a Cas9-expressing binary vector p201N:Cas9 (Jacobs and Martin, 2016). The gRNAs were designed to be highly specific at target sites and their predicted off-target sites contained at least one nucleotide mismatch in the seed sequence (the last 12 nucleotides preceding the PAM) or two nucleotide mismatches against the full 20-bp protospacer (although some gRNAs were designed to intentionally modify multiple homologs simultaneously). Most Cas9/gRNA constructs in this study had only one gRNA expression cassette per plasmid, except one construct that contained three gRNA cassettes targeting three *Aquaporin transporter* genes (**Table 1**).

Evaluation of gRNA Effectiveness by Agroinfiltration in Tomato and *Nicotiana benthamiana* Leaves

To initially evaluate the effectiveness of gRNAs and subsequently enhance the mutation rate in stably transformed tomato plants, 195 gRNAs were tested for their ability to cause mutations using *Agrobacterium* infiltration (agroinfiltration) in tomato leaves (**Figure 1A**; **Table S1**). After agroinfiltration, DNA was isolated from the leaf tissue and the genomic region spanning each target site was PCR amplified, sequenced, and analyzed with a web-based tool called Tracking of Indels by Decomposition (TIDE; https://tide.deskgen.com; Brinkman et al., 2014) to calculate the mutation frequency. gRNAs with mutation frequency >0 ($p < 0.0001$) were considered to be effective in inducing mutations in this assay while those with mutation rate = 0 were considered ineffective. A total of 61.5% of the tested gRNAs were effective in inducing transient mutations in tomato leaves (**Figure 1B**). Among these, 96% had a mutation rate greater than 0% but less than 10% in this assay, while only five gRNAs (4%) had a mutation frequency over 10% (**Figure 1C**).

Agroinfiltration of tomato leaves is not very efficient and to test whether this affected our estimate of gRNA mutation efficacy, we evaluated the mutation efficiency of two of the gRNAs (Bti9ab-1 and Drm3-1) which each have identical target sites in both tomato and *Nicotiana benthamiana* (**Figure S1**). The mutation frequency induced by these two gRNAs in tomato was much lower than in *N. benthamiana* (**Figure S1A**). In addition to these two gRNAs, we also tested another four gRNAs that each target two of the four homologs of the *Mai1* gene in *N. benthamiana* (Roberts et al., 2019; **Figure S1B**). In *N. benthamiana*, the majority of the gRNA targets showed a

TABLE 1 | Mutation rates and mutation types in T0 transgenic plants. See also **Table S3**.

Target genes	Solyc #	# of transgenic plants	# of edited plants	Mutation rate (%)[a]	Mutation types[b]
ADE	Solyc05g005700	10	5	50	3 homo; 1 hetero; 1 chimeric
AOX	Solyc08g075550	9	8	89	2 homo; 2 biallelic; 3 hetero; 2 chimeric
APE	Solyc11g018800	1	1	100	1 biallelic
Aquaporin 1/2/3[c]	Solyc11g069430 (Aquaporin 1) Solyc06g074820 (Aquaporin 2) Solyc08g066840 (Aquaporin 3)	12	10	83	5 homo; 2 biallelic; 10 hetero; 2 chimeric
BHLH	Solyc03g114230	3	3	100	3 biallelic
BSK830	Solyc12g099830	8	6	75	2 biallelic; 3 hetero; 1 chimeric
BSK830 (Hawaii 7981)[d]	Solyc12g099830	6	5	83	3 biallelic; 1 hetero; 1 chimeric
Bti9-interactor	Solyc09g008010	1	1	100	1 homo
Bti9ab	Solyc07g049180	4	3	75	2 homo; 1 biallelic
CathepsinB1	Solyc02g076980	29	20	69	8 homo; 8 biallelic; 2 hetero; 2 chimeric
CathepsinB2	Solyc02g077040	11	8	73	2 homo; 3 biallelic; 3 hetero
CORE	Solyc03g096190	7	6	86	1 homo; 4 biallelic; 1 hetero
Drm-3	Solyc01g099840	3	3	100	2 biallelic; 1 hetero
EDS1	Solyc06g071280	2	2	100	1 homo; 1 chimeric
ERF5	Solyc03g093560	7	5	71	2 biallelic; 1 hetero; 2 chimeric
Fen	Solyc05g013290	18	13	72	5 biallelic; 7 hetero; 1 chimeric
Fen (RG-pto11)[e]	Solyc05g013290	15	9	60	2 homo; 4 biallelic; 5 chimeric
Fls2.1	Solyc02g070890	4	4	100	3 homo; 1 hetero
Fls3	Solyc04g009640	11	7	64	5 homo; 2 hetero
LRRXII-1	Solyc06g076910	5	1	20	1 biallelic
LRRXII-2	Solyc04g012100	4	3	75	1 biallelic; 2 hetero
Mai1	Solyc04g082260	4	1	25	1 hetero
Mai5	Solyc10g085990	11	8	73	2 homo; 5 hetero; 1 chimeric
MAP3Ka	Solyc11g006000	6	6	100	4 biallelic; 2 hetero
Min7	Solyc12g017830	5	4	80	1 biallelic; 1 hetero; 2 chimeric
MKK1	Solyc12g009020	2	2	100	1 biallelic; 1 chimeric
MKK2	Solyc03g123800	6	5	83	1 homo; 4 biallelic
MKK4	Solyc03g097920	2	1	50	1 hetero
MKKK15	Solyc02g065110	8	6	75	4 hetero; 2 chimeric
MKKK66	Solyc08g081210	3	1	33	1 hetero
MLO16	Solyc11g069220	4	4	100	1 biallelic; 2 hetero; 1 chimeric
NOD	Solyc11g008200	8	7	88	1 homo; 4 hetero; 2 chimeric
NPR1	Solyc07g040690	5	1	20	1 hetero
NRC1/2/3[f]	Solyc01g090430 (NRC1) Solyc10g047320 (NRC2) Solyc05g009630 (NRC3)	8	5	63	1 homo; 3 biallelic; 1 hetero
PAD4	Solyc02g032850	7	1	14	1 biallelic
PBCP	Solyc03g116690	3	2	67	1 homo; 1 biallelic
PBL-T1	Solyc09g007170	5	3	60	1 biallelic; 2 chimeric
PBL-T2	Solyc01g067400	2	2	100	2 hetero
Peptide Transporter 3	Solyc05g009500	4	3	75	1 biallelic; 1 hetero; 1 chimeric
Permease Transporter	Solyc03g005820	4	2	50	1 homo; 1 biallelic
PGA1[f]	Solyc05g005560 Solyc05g005570	5	2	40	2 hetero
Pic1	Solyc07g066260	5	4	80	1 biallelic; 2 hetero; 1 chimeric
PR1b	Solyc00g174340	6	4	67	3 biallelic; 1 chimeric
Propep1	Solyc04g072310	2	2	100	2 chimeric
RALF1	Solyc01g067900	3	3	100	1 homo; 2 biallelic
RALF2	Solyc01g099520	10	7	70	1 homo; 3 hetero; 3 chimeric
RbohB	Solyc03g117980	8	3	38	1 homo; 2 biallelic
SAG101-1/2[f]	Solyc02g069400 (SAG101-1) Solyc02g067660 (SAG101-2)	3	2	67	1 homo; 1 biallelic
SlMlo1/9[f]	Solyc04g049090 (SlMlo1) Solyc06g010030 (SlMlo9)	10	9	90	1 homo; 2 biallelic; 3 hetero; 6 chimeric
SOBIR/SOBIR-like[f]	Solyc06g071810 (SOBIR) Solyc03g111800 (SOBIR-like)	16	9	56	2 homo; 1 biallelic; 4 hetero; 3 chimeric
Solute Transporter 2	Solyc05g005950	4	2	50	1 biallelic; 1 hetero

(Continued)

TABLE 1 | Continued

Target genes	Solyc #	# of transgenic plants	# of edited plants	Mutation rate (%)[a]	Mutation types[b]
STP13	Solyc09g075820	2	1	50	1 homo
TFT1	Solyc11g010470	3	3	100	1 biallelic; 2 hetero
TFT10	Solyc04g076060	4	2	50	2 chimeric
TFT7	Solyc04g074230	2	2	100	2 hetero
Wak1	Solyc09g014720	5	5	100	2 homo; 2 biallelic; 1 chimeric
WRKY11	Solyc08g006320	2	2	100	2 chimeric
WRKY9b	Solyc08g067360	4	1	25	1 biallelic

[a]Mutation rate (%) = number of plants with mutations/number of total transgenic plants. Mutations were analyzed by Geneious R11 and TIDE. [b]Homo, homozygous mutation; hetero, heterozygous mutation. [c]Three gRNA cassettes were cloned into one p201N:Cas9 plasmid. [d]Tomato genotype Hawaii 7981 was used for transformation. [e]Tomato genotype RG-pto11 was used for transformation. [f]Specific gRNA for each gene was independently cloned into the p201:Cas9 vector and the two or three gRNA/Cas9 constructs were pooled together for tomato transformation.

mutation frequency of 10%–40%, while a small number had a mutation frequency less than 5% (**Figure S1B, C**). These observations suggest that inefficiency of agroinfiltration in tomato leaves probably leads to an underestimate of the true efficacy of gRNAs for generating mutations. This is supported by later observations in which some low-efficient gRNAs were very effective in inducing mutations in stably-transformed tomato plants (**Table S2**). Factors that contribute to this lower activity of gRNAs by agroinfiltration of tomato leaves are unknown but might include possible specific immune responses of tomato against *Agrobacterium* or less efficient infiltration due to leaf morphology differences with *N. benthamiana*.

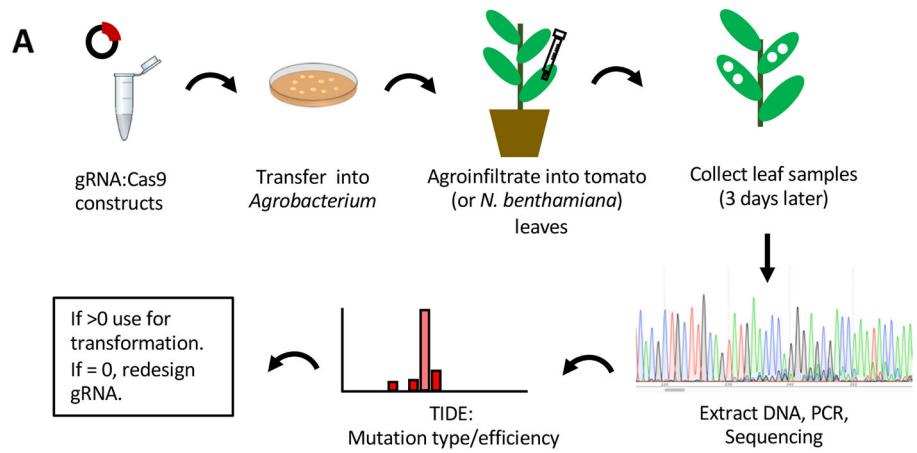
In most cases, we selected the most efficient gRNAs for subsequent stable transformation in tomato, however, some low-efficiency gRNAs were also used if limited gRNAs could be designed for a particular target gene. Most gRNAs that were effective in inducing mutations in the agroinfiltration transient assay also induced mutations in stable transgenic seedlings, with one exception where a gRNA that had a 4.4% mutation frequency in the transient assay did not edit target genes in two stably-transformed plants (**Table S2**). It was not possible, however, to directly compare gRNA efficiency in the transient assay with that in stable transformation, as other factors such as the bias of gRNA transformation into plants using "*Agrobacterium* pools" and the total number of regenerated transgenic seedlings varied from gene to gene in stable transformation.

CRISPR/Cas9-Induced Mutations in T0 Transgenic Plants

A total of 361 putative primary (T0) transgenic tomato plants were regenerated by *Agrobacterium tumefaciens*-mediated stable transformation. To confirm the mutated sequence(s) in each plant, genomic regions spanning the target sites were PCR amplified and sequenced. All five possible genotypes, that is, wild-type, homozygous for the mutation, biallelic (a different mutation in each allele), heterozygous for the mutation, or multiple mutations (chimeric), were detected in our stably transformed tomato plants (**Table 1**). Direct sequencing of PCR amplicons containing biallelic, heterozygous, or chimeric mutations resulted in superimposed sequencing chromatograms, which made it difficult to determine specific mutation types and mutation frequency in those plants. To resolve this problem, TIDE was used to rapidly determine the mutated allelic sequences using the sequencing file (.ab1 format) with superimposed chromatograms (Brinkman et al., 2014), thus avoiding tedious and expensive cloning and multi-clone sequencing for mutation analysis.

Of the 361 T0 plants, 245 were found to have modifications at the target site(s) within 63 genes (**Table 1**; **Table S3**). Most of the lines had only one CRISPR-induced mutation in one gene per plant, while a few had mutations in two or even three genes (the latter cases occurred when using *Agrobacterium* pooling – see below). We identified only one mutant event for some of the targeted genes while for others up to 20 independent mutant events were generated (**Table 1**). Overall, the average editing efficiency (the number of edited plants/the number of transgenic plants) by CRISPR/Cas9 in tomato in our experiments was 68% (**Figure 2A**), although the mutation rate varied over a wide range from 14% to 100% from target gene to target gene in different mutant lines (**Table 1**). All four mutation types (homozygous, biallelic, heterozygous, or chimeric) were observed in several mutant lines that had sufficient independent mutant events, while there was a bias of mutation types in some mutant lines, probably due to the limited number of transgenic events generated (**Table 1**).

We analyzed the distribution of the four mutation types in all the 245 T0 plants at the 267 mutated target sites (some plants had more than one target site) and found that the percentage of homozygous, biallelic, heterozygous, or chimeric mutation was 20%, 32%, 30%, and 18%, respectively (**Figure 2B**). In particular, plants having homozygous or biallelic mutations accounted for 52% of the total. These mutants and their progenies can be used directly for phenotype screening because no wild-type alleles are present, thus speeding the research process by saving time for further genotyping in the next generation. The most common mutation alterations induced by CRISPR/Cas9 were deletions or insertions, with 87% of these modifications at the target sites being less than 10 bp (**Figure 2C**). The proportion of deletion mutations was 77%, and the deletion length spanned a wide range from 1 bp to over 400 bp. Of all the mutations, the most abundant modification was 1-bp deletion or insertion (**Figure 2C**). For these, A- and T-insertions accounted for 79.5%, while G-insertions accounted for only 4.5% (**Figure 2D**). Base substitutions in combination with indels were also detected, but at a much lower frequency. Only three independent mutant events (two were the same mutation type) harbored a nucleotide



B

gRNA efficiency	# of gRNAs	Percentage
> 0	120	61.5%
= 0	75	38.5%

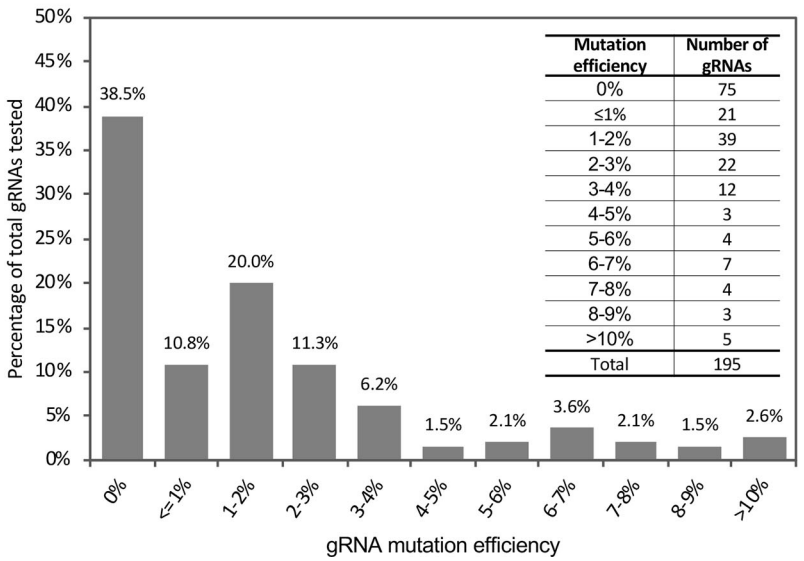


FIGURE 1 | Evaluation of gRNA-mediated mutation efficiency by agroinfiltration in tomato leaves. See also **Figure S1** and **Table S1**. **(A)** Schematic showing the workflow of guide RNA (gRNA) evaluation by agroinfiltration. **(B)** Summary of gRNA efficiency tested by agroinfiltration. **(C)** The distribution of mutation efficiencies of the 195 gRNAs. Inset on the top right shows the number of gRNAs in each mutation efficiency range. TIDE (https://tide.deskgen.com) was used to calculate mutation efficiency by identifying the predominant types of insertions and deletions (indels) in the DNA of a targeted cell pool.

substitution in one copy of the target genes at the positions 5-bp preceding the PAM, along with a short insertion or deletion at the target site (**Figure S2**).

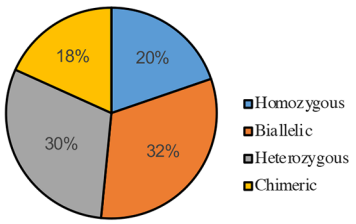
Multiplex editing of three *Aquaporin transporter* (*AquaT*) genes by using one Cas9/gRNA construct was also tested. Three individual promoter-gRNA expression cassettes (in the order *Aqua1-Aqua2-Aqua3*) were assembled into the p201N:Cas9 vector (**Figure S3A**) as previously reported (Jacobs et al., 2017). Ten of the 12 regenerated transgenic plants were edited, including three single mutants, four double mutants, and three triple mutants (having mutations in all three genes). Interestingly, all the three single mutants knocked out *AquaT1*, while three double mutants modified *AquaT1* and *AquaT2* and one edited *AquaT1* and *AquaT3* simultaneously. Three plants had mutations in all the three *AquaT* genes together (**Figure S3B**). These data suggested the position of the gRNA cassette in the vector may affect its mutation rate, considering that all gRNAs were efficient enough to induce mutations once transformed into plants.

A

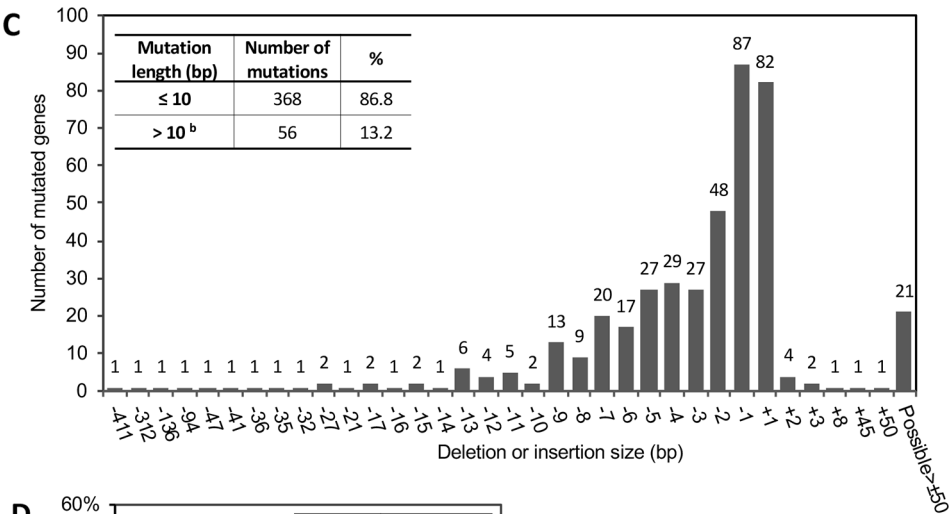
# of total T0 transgenic plants	# of T0 plants with mutations	Overall mutation efficiency
361	245	68%

B

Mutation types	# of mutated genes (T0 plants)[a]
Homozygous	53
Biallelic	84
Heterozygous	81
Chimeric	49
Total #	267



C



D

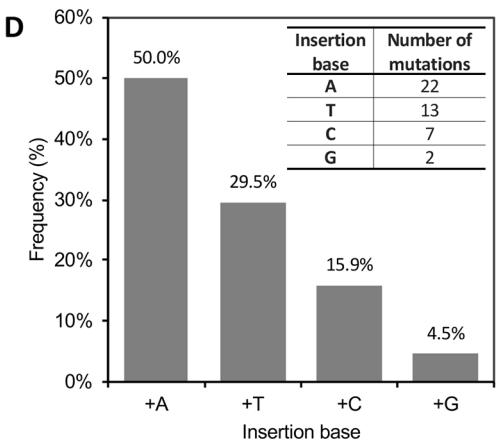


FIGURE 2 | CRISPR/Cas9-induced gene mutations in T0 transgenic plants. See also **Table S3**. **(A)** The average mutation rate induced by CRISPR/Cas9 in T0 plants. **(B)** Summary of CRISPR-induced mutation types and their frequency in T0 plants. Left: Number of genes modified with the corresponding mutation type; Right: Percentage of genes harboring the corresponding mutation type. [a] Some plants have multiple target genes in one plant. **(C)** Frequency of each insertion or deletion mutation. x-axis: number of base pairs (bp) deleted (–) or inserted (+) into target sites. Inset on the left top shows the percentage of mutations ≤10 bp or >10 bp. [b]All "possible > ± 50 bp" in the figure are included in >10 bp. TIDE only calculates mutation length ≤50 bp. **(D)** Percentage of different bases in the 1-bp insertion mutations. Inset at the right top shows the number of mutations with each type of inserted base.

Heritability of the Mutations

To evaluate the heritability of mutations in T0 plants, a large number of T1 and some T2 plants were generated and examined for their genotypes at the target sites. Most of the same genotypes from T0 plants were transmitted to plants in later generations, although we did not record segregation ratios in the progenies. Of note, no new mutations or reversions to wild-type were found in the progeny of any homozygous T0 plants, indicating all the homozygous mutations occurred in the transformed embryogenic cells before the first division. However, we did observe novel genotypes in a small percentage of T1 or T2 plants whose progenitor (T0 plants) harbored biallelic, heterozygous or

chimeric mutations (**Table 2**). In particular, a homozygous mutation (−265 bp) was detected in the progeny of a *NRC2* primary transgenic plant with a "+1 bp/+2 bp" biallelic mutation. It is possible that the new −265 bp modification at the target site was created by further modification of the existing +1 bp or +2 bp mutations in the progenitor as has been reported in rice and carrot (Klimek-Chodacka et al., 2018; Tang et al., 2018). It is also possible that the biallelic T0 plant (+1 bp/+2 bp) was a chimera and the new mutation derives from chimeric tissue of the T0 plant (Zhang et al., 2014).

We also observed new genotypes in the progeny of some T0 plants harboring heterozygous or chimeric mutations (**Table 2**) even though most of the progeny still possessed the same genotypes as the progenitor line. In some of these cases, CRISPR/Cas9 continued to modify the wild-type allele of the target gene in the progeny if the parent plants still contained a wild-type allele and the Cas9/gRNA expression cassette (**Table 2**). In other cases, unexpected genotypes were detected in some mutant lines including Mai1-E10, CathepsinB2-E8, Min7-E6, and ADE-E2 (**Table 2**). For instance, although the ADE-E2 T0 plant was chimeric (−4 bp/−5 bp/−13 bp) without a wild-type allele, we identified one T1 plant that was azygous (two copies of wild-type allele) and another one with a novel biallelic mutation (−4bp/−9 bp). Another example is the CathepsinB2-E8 T0 plant which had a heterozygous (−4 bp/WT) mutation. However, −2 bp homozygous and −1 bp/WT heterozygous mutations were detected in later generations. The unexpected new genotypes discussed above revealed that the one leaf/cotyledon sample may not reveal all the genotypes in the whole plant if the mutant is chimeric. Therefore, for T0 edited plants without any wild-type allele, it will still be useful to perform genotyping in subsequent generations to obtain homozygous mutants without the presence of Cas9/gRNA.

Specificity of CRISPR-Induced Gene Modifications in Tomato

Mutations in unintended sequences (off-target mutations) is a possible concern in both functional genomics studies and plant breeding. To evaluate potential off-target effects by CRISPR/Cas9 in our tomato lines, we first evaluated the specificity of 12 gRNAs of Cas9 by Geneious R11 (https://www.geneious.com; Kearse et al., 2012) and Cas-OFFinder (Bae et al., 2014). The putative off-target sites predicted by Cas-OFFinder were then manually checked using JBrowse (https://solgenomics.net/jbrowse_solgenomics/) to confirm their locations in the tomato genome. The presence of a PAM was required for the site to be considered a candidate site. These gRNAs were selected for off-target

TABLE 2 | New genotypes detected in T1 or T2 plants.

T0 mutation type[a]	Generation	Plants	Gene Solyc#	Mutation[b]	Is T-DNA present?
Bialleleic	**T0**	**NRC2-E4**	**Solyc10g047320**	**+1 bp/+2 bp**	**Yes**
	T1	NRC2-E4-P2	Solyc10g047320	+1 bp/large deletion[c]	No
	T2	NRC2-E4-P2-2	Solyc10g047320	-265 bp/-265 bp	No
	T2	NRC2-E4-P2-4	Solyc10g047320	+1 bp/+1bp	No
Heterozygous	**T0**	**Mai1-E10**	**Solyc04g082260**	**-4 bp/WT**	**Yes**
	T1	Mai1-E10-P20	Solyc04g082260	+1 bp/WT	Yes
Heterozygous	**T0**	**CathepsinB2-E8**	**Solyc02g077040**	**-4 bp/WT**	**Yes**
	T1	CathepsinB2-E8-P2	Solyc02g077040	-2 bp/-2 bp	No
	T1	CathepsinB2-E8-P10	Solyc02g077040	-1 bp/WT	No
Heterozygous	**T0**	**Min7-E6**	**Solyc12g017830**	**+1 bp/WT**	**Yes**
	T1	Min7-E6-P4	Solyc12g017830	+1 bp/-3 bp	Yes
	T1	Min7-E6-P8	Solyc12g017830	-1 bp/WT	No
Heterozygous	**T0**	**NRC2-E1**	**Solyc10g047320**	**-3 bp/WT**	**Yes**
	T1	NRC2-E1-P6	Solyc10g047320	-3 bp/-5 bp	Yes
	T1	NRC2-E1-P15	Solyc10g047320	-3 bp (58%); -6 bp (32%) +1 bp (3.6%)	No
Chimeric	**T0**	**MKKK15-E2**	**Solyc02g065110**	**-1 bp (10.4%); +1 bp (2.1%); WT (85%)**	**Yes**
	T1	MKKK15-E2-P2	Solyc02g065110	-5 bp(60%); -1 bp (31%); +1 bp (4.7%)	Yes
	T1	MKKK15-E2-P3	Solyc02g065110	-5 bp (20%); WT (74%)	No
	T1	MKKK15-E2-P7	Solyc02g065110	+1 bp (77.6%); -5 bp (4.7%); WT (9.5%)	Yes
Chimeric	**T0**	**MKK1-E1**	**Solyc12g009020**	**-7 bp (8.1%); -2 bp (33.3%); WT (48.5%)**	**Yes**
	T1	MKK1-E1-P3	Solyc12g009020	-2 bp/-4 bp	No
Chimeric	**T0**	**Min7-E5**	**Solyc12g017830**	**-1 bp (35.9%); +1 bp (9%); WT (49.4%)**	**Yes**
	T1	Min7-E5-P1	Solyc12g017830	-1 bp/-3 bp	Yes
	T1	Min7-E5-P2	Solyc12g017830	-6 bp (64.5%); +1 bp (19%); -1 bp (8.8%); -7 bp (3.4%)	Yes
	T1	Min7-E5-P3	Solyc12g017830	-5 bp/-6 bp	Yes
	T1	Min7-E5-P7	Solyc12g017830	-9 bp (71.3%); +1 bp (18.2%); -1 bp (3.6%)	Yes
Chimeric	**T0**	**ADE-E2**	**Solyc05g005700**	**-4 bp (20%); -5 bp (40.4%); -13 bp (13.9%)**	**Yes**
	T1	ADE-E2-P1	Solyc05g005700	WT/WT	No
	T1	ADE-E2-P20	Solyc05g005700	-4 bp/-9 bp	No

[a]*For each mutant event, only the plants harboring different genotypes from their T0 progenitors are listed. Unlisted T1 or T2 progeny have the same genotypes as those in T0 plants.*
[b]*Number of base pairs (bp) deleted (-) or inserted (+) into target sites; WT, wild-type. If no percentage is shown, the two genotypes are around 50%:50%.*
[c]*TIDE only detects indels ≤50 bp.*

analysis because morphological defects were observed in one or more mutant lines induced by these gRNAs (**Figure S4**; **Table S4**). A total of 18 possible off-target sites of the 12 gRNAs were identified and off-target mutations were examined in 12 T0 plants, 68 T1 plants and 44 T2 plants by PCR and Sanger sequencing (**Table 3**). No off-target modifications were discovered in the tested plants with or without Cas9, indicating our gRNAs and CRISPR-mediated mutations are highly specific.

Another way to evaluate the specificity of CRISPR/Cas9 is to test the efficiency of gRNAs with a few mismatches against the target sequence in the protospacer. One of our gRNAs targeted two tomato homologs, *Mai5* and *PBL-T1* (bothMai5/PBL-T1; 5'-**g**TAGATCGTAATGGATTGCA-3'; the first nucleotide "C" was converted to "G" to accommodate the transcription initiation requirement of the U6 promoter). The designed 20-bp protospacer sequence exactly matched the target site in *Mai5* but had one mismatch in *PBL-T1* at the third nucleotide from the 5' end (5'-gT**C**GATCGTAATGGATTGCA-3'). We generated five T0 plants that contained the bothMai5/PBL-T1 gRNA construct, all of which had edits in *Mai5* but not in *PBL-T1*, indicating the one mismatch (along with the first nucleotide at the 5' end) in *PBL-T1* appeared to significantly affect Cas9 binding and cleavage activity at the target site. Another gRNA, targeting the tomato *FLS2.2* gene, did not induce targeted modifications in any of the 10 transgenic plants, possibly due to a 1-bp mismatch in the seed region of the gRNA in the Heinz 1706 reference genome (GTCATCAACAT**C**TCGCTTGT) as compared to Rio Grande-PtoR (GTCATCAACAT**T**TCGCTTGT). The reference genome was used for gRNA design and RG-PtoR was used for tomato transformation as it contains the resistance gene *Pto* for investigating *Pto*-mediated immunity in our mutants. This further indicates that CRISPR/Cas9 is highly specific, with even one mismatch in the gRNA rendering the site uncleavable to the Cas9/gRNA complex.

TABLE 3 | Examination of possible off-target mutations caused by 12 selected gRNAs in multiple generations. See also **Tables S4** and **S7**.

gRNAs		Putative off-target locus	Sequence of the putative off-target sites[a]	Predicted by	No. of mis-matches	No. of plants tested	No. of plants with mutations
CathepsinB1-1	#1	SL3.0ch00: 19,581,605 - 19,581,626	gGAATCTG**A**CTTACAA**C**TTATGG	Geneious	2	30 (10 T1 + 20 T2)	**0**
	#2	SL3.0ch08: 52,560,219 - 52,560,240	g**A**AATCT**T**TCTT**C**CAAGTTAGGG	Geneious	3	20 (10 T1 + 10 T2)	**0**
ERF5-1	#1	SL3.0ch03: 56,335,989 - 56,336,010 (CDS)	gGTAT**C**GC**A**A**T**GTTCAGAGATGG	Geneious/ CasOFFinder	3	2 (T0)	**0**
	#2	SL3.0ch03: 56,356,193 - 56,356,214 (CDS)	gGTAT**C**GC**A**A**T**GTTCAGAGATGG	Geneious/ CasOFFinder	3	2 (T0)	**0**
Fen-1	#1	SL3.0ch11: 22,381,618 - 22,381,638	gATGC-T**A**TAACTTGAGTTAGGG	Geneious	2	17 (10 T1 + 7 T2)	**0**
Mai5-2	#1	SL3.0ch05: 66,181,295 - 66,181,316 (CDS)	gTCTACGAATATATGCCA**A**TGGG	Geneious/ CasOFFinder	1	1 (T0)	**0**
Mai5-3	#1[b]	SL3.0ch09: 795,322 - 795,343 (CDS)	gT**C**GATCGTAATGGATTGCAAGG	Geneious/ CasOFFinder	1	3 (1 T0 + 2 T1)	**0**
	#2	SL3.0ch12: 67,901,371 - 67,901,392 (CDS)	gTA**TC**TC**A**TAATGGATTGCAAGG	Geneious	3	3 (1 T0 + 2 T1)	**0**
	#3	SL3.0ch12: 67,814,680 - 67,814,701 (CDS)	gTA**TC**TC**A**TAATGGATTGCAAGG	Geneious	3	3 (1 T0 + 2 T1)	**0**
MKK2-1	#1	SL3.0ch02: 16,695,145 - 16,695,165	gAA**T**ACG-ACAGAATCCT**A**GGGG	Geneious	3	17 (2 T0 + 15 T1)	**0**
MKKK15-1	#1	SL3.0ch02: 48,342,685 - 48,342,706 (intron)	G**C**CCAC-AT**C**GTGTCGACGTGGG	Geneious	3	15 (1 T0 + 7 T1 + 7 T2)	**0**
PBL-T1-2	#1	SL3.0ch11: 16,102,764 - 16,102,785	gATG**T**CGAG**G**C**G**GGTCAAATTGG	Geneious	3	4 (1 T0 + 3 T1)	**0**
RALF2-2	#1	SL3.0ch11: 42,841,387 - 42,841,409	GTTG**A**AG**C**T**T**GGAAGCTCCAAGG	Geneious	3	5 (2 T1 + 3 T2)	**0**
SOBIR-1	#1	SL3.0ch02: 31,876,612 - 31,876,633	gT**C**TATA**C**A**C**CAGAGCTACCAGG	Geneious	3	11 (4 T1 + 7 T2)	**0**
TFT10-1	#1	SL3.0ch11: 42,990,626 - 42,990,647 (Intron)	gATTCACTG**A**TAGTA**T**CAGATGG	Geneious/ CasOFFinder	2	7 (1 T0 + 6 T1)	**0**
	#2	SL3.0ch08: 34,312,506 - 34,312,527 (CDS)	g**C**TTCA**TA**G**T**TAGTAGCAGAAGG	Geneious	4	7 (1 T0 + 6 T1)	**0**
TFT7-2	#1	SL3.0ch11: 17,271,683 - 17,271,703 (Intron)	gAATGT-AT**G**GGAGCAAGAAAGG	Geneious	2	14 (1 T0 + 13 T1)	**0**
	#2	SL2.5chr03: 5,874,879-5,874,902 (CDS)	GAATGT**C**AT**T**GG**T**GCAAGAAGGG	CasOFFinder	3	14 (1 T0 + 13 T1)	**0**

[a]*The PAM motif occurs at the 3' end of each sequence (AGG, GGG, or TGG). Mismatched bases are in bold and underlined; "g" in lower case means the first nucleotide of the putative off-target sequence is not a "G" but was converted to that nucleotide to accommodate the transcription initiation requirement of the U6 promoter.*

[b]*This is not a true off-target as it was intentionally designed to target a Mai5 homolog PBL-T1).*

Tomato Transformation With *Agrobacterium* Pools

Tomato transformation is a lengthy and labor-intensive process. In an approach to minimize the number of transformation experiments needed, three to four *Agrobacterium* culture preparations each carrying a different gRNA construct were pooled and used for a single transformation experiment (**Figure 3A**). T0 plants were genotyped for both the presence of T-DNA carrying the gRNAs and for possible mutations in all the targeted genes in the pool. Of the 79 T0 plants generated, 58 plants (73%) contained precise modifications in one or more of the target genes. In terms of the number of target sites edited by CRISPR/Cas9 with pooled gRNAs, 48 (82.8%) of the T0 plants had mutations in just one gene, while 9 plants (15.5%) had mutations in two and 1 (1.7%) had mutations in three genes (**Figure 3A**). Among these T0 plants, 83.5% contained one gRNA cassette and 15.2% contained two different gRNA cassettes, while no plants recovered contained more than two gRNA cassettes integrated into the genome (**Figure 3B**). Interestingly, one mutant plant did not show detectable integration of the T-DNA sequence (expressing *Cas9* and gRNA) but had a mutated gene, suggesting that transient expression of the Cas9/gRNA occurred in this plant. Additionally, we found another type of transient mutation in 8 T0 plants at 10 different target sites (**Table S5**). In these plants, a Cas9/gRNA expression cassette was integrated into the plant genome, as confirmed by PCR and Sanger sequencing, but the gRNA detected was not the one that induced the mutation in the plant (**Table S5**), suggesting that the mutation was caused by another transiently expressed Cas9/gRNA.

DISCUSSION

Our effort to generate a large number of CRISPR/Cas9-induced tomato mutants targeting immunity-associated genes demonstrates that this mutation approach is efficient and robust for gene editing in tomato. Importantly, gene modifications mostly occurred in germline cells and were stably inherited in subsequent generations, similar to those in rice (Zhang et al., 2014) but not as in Arabidopsis, in which most mutations in T0 plants were somatically modified if a strong constitutive promoter was used to regulate the Cas9 expression (Feng et al., 2014; Feng et al., 2018). However, we detected a greater range of deletion and insertion lengths than observed in rice (Zhang et al., 2014), in which only 1-bp insertions and fewer deletion lengths were found, possibly due to different intrinsic DNA repair mechanisms between these two species. These differences could also be due to other factors including different transformation methods or culture conditions, and different sets of target genes that tolerate different degrees of mutations.

Base substitutions induced by the CRISPR/Cas9 system in tomato were very rare in our study. We frequently observed single nucleotide polymorphisms (SNPs) between Rio Grande (used for transformation) and the tomato reference genome (Heinz 1706), but these SNPs were due to natural variation, not mutagenesis, as confirmed by sequencing of the gene regions from untransformed plants. Most of these SNPs were located outside of the protospacer sequence of the gRNA targets, and to date we have only found one gRNA (targeting *Fls2.2*) which had a mismatch in the seed region that inhibited the Cas9 binding and cleavage at the target site. Therefore, it is reasonable to use the tomato reference genome as

A

Targeted genes	Number of *Agrobacterium* cultures with gRNAs pooled together	# of mutants for each target gene[a]	# of plants genotyped	# of plants with edits	Number of plants with mutations in one, two, or three genes[b]		
					one	two	three
Mai5 and *PBL-T1*	3	*Mai5* (8); *PBL-T1* (3)	14	10	9	1	na
MKKK15; *MKKK20*	4	*MKKK15* (6); *MKKK20* (0)	10	6	6	0	na
TFT1, *TFT7*, *TFT10*	4	*TFT1* (3); *TFT7* (2); *TFT10* (2)	7	5	3	2	0
PROPEP; *P26S*; *ERF5*; *NOD*	4	*Propep1* (1); *NOD* (7); *ERF5* (5); *P26S* (0)	15	11	9	2	0
MKK2 and *MKK4*	4	*MKK2* (5); *MKK4* (1)	8	6	6	0	na
MKKK66; *MAP3Ka*; *MKK1*	4	*MKKK66* (1); *MAPK3Ka* (6); *MKK1* (2)	12	9	9	0	0
AOX; *RALF2*; *WRKY11*	4	*AOX* (8); *RALF2* (7); *WRKY11* (2)	13	11	6	4	1
		Total	79	58	48	9	1

B

Number of gRNAs in the plant	Number of plants
0	1
1	66
2	12

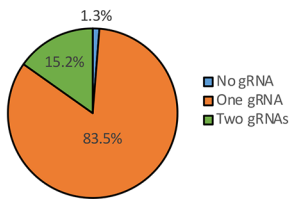


FIGURE 3 | Tomato transformation with "*Agrobacterium* pools." See also **Table S5**. **(A)** An example of a tomato transformation experiment designed to target 2–4 genes by using 3–4 pooled *Agrobacterium* cultures with each culture carrying a different gRNA. [a] Number of plants with mutations in the target gene is shown in parentheses; some plants had mutations in multiple genes. [b] Number of genes modified in the plants. "na" not applicable, since less than 3 genes were targeted in the experiment. **(B)** Left: Number of gRNAs detected in a single mutant plant by PCR and Sanger sequencing. Right: Percentage of T0 plants harboring no, one or two gRNAs.

the template for gRNA design and subsequent mutation genotyping in transgenic Rio Grande and likely other tomato cultivars.

Various morphological phenotypes were detected in some mutants compared to wild-type plants. Some of these abnormal phenotypes were associated with all the mutation events occurring in a specific gene, strongly supporting that the mutation itself is responsible for the altered plant growth or development. However, some mutant lines showed unusual morphology associated with certain mutation events but not all, possibly indicating that another off-target mutation occurred in these plants. We therefore investigated a large number of these plants but did not find any evidence of off-target mutations, suggesting other mutations, if they exist, were either caused by tissue culture or *Agrobacterium* transformation, or spontaneous mutations during seed propagation (Tang et al., 2018). Our observations are consistent with previous reports that CRISPR/Cas9 causes few off-target mutations in plants including Arabidopsis (Feng et al., 2014), rice (Zhang et al., 2014; Tang et al., 2018), tomato (Nekrasov et al., 2017), cotton (Li et al., 2019), and maize (Young et al., 2019). True off-targets reported previously in plants showed high sequence homology to the original spacer sequence of gRNAs (Tang et al., 2018), which can be easily avoided by designing highly specific gRNAs using tools such as Geneious and Cas-OFFinder. Based on our data we devised a rule to avoid off-target effects of CRISPR/Cas9 by designing gRNAs whose highest scored potential off-target sites have at least a 1-nt mismatch in the seed sequence or 2-nt mismatches in the full protospacer sequence.

Surprisingly, we found new mutations in the progeny of some T0 plants that did not contain any wild-type allele. These new mutations did not appear to be derived from existing mutations in the T0 plants, as the Cas9-induced modifications were located within the seed sequence of the gRNA protospacer and as little as 1-bp mismatch in the seed sequence can dramatically impair the Cas9 binding and cleavage activity (Jinek et al., 2012). Therefore, we believe the new mutations were derived from chimeric tissue from the T0 plant that was not detected with the one cotyledon/leaf sample we used for mutation genotyping. We are currently advancing lines that have biallelic or heterozygous mutations, or that were chimeric to develop homozygous plants without the presence of Cas9/gRNA sequence. These plants will be used to investigate whether the mutations affect the plant immune response, especially to *P. syringae* pv. tomato.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the bioRxiv: https://www.biorxiv.org/content/10.1101/835108v1. Primers used to amplify targeted region, the intervening genomic sequence, guide-RNA used, and the mutations found can be found in PGED (http://plantcrispr.org) by searching the relevant gene accession number.

AUTHOR CONTRIBUTIONS

GM and NZ conceived and designed the experiments. NZ designed gRNAs, constructed vectors, performed genotyping in T0 plants and other generations, and analyzed the data. HR performed genotyping of mutants in some T1 and T2 plants. JV optimized the plant transformation protocol and guided the transformation experiments. NZ and GM interpreted the data and wrote the manuscript

FUNDING

Funding was provided by National Science Foundation grants IOS-1732253 (JV) and IOS-1546625 (GM).

ACKNOWLEDGMENTS

We thank Brian Bell, Jay Miller, and Nick Vail for plant care; Patricia Keen, Michelle Tjahjadi, and Sergei Krasnyanski for tomato transformation; and Colleen Conger and Elizabeth Trost for assistance with mutation genotyping. The authors declare no conflicts of interest. This manuscript has been released as a pre-print at bioRxiv (Zhang et al., 2019).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00010/full#supplementary-material

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

Baltes, N. J., and Voytas, D. F. (2015). Enabling plant synthetic biology through genome engineering. *Trends Biotechnol.* 33, 120–131. doi: 10.1016/j.tibtech.2014.11.008

Bogdanove, A. J., and Voytas, D. F. (2011). TAL effectors: customizable proteins for DNA targeting. *Science* 333, 1843–1846. doi: 10.1126/science.1204094

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

Canene-Adams, K., Campbell, J. K., Zaripheh, S., Jeffery, E. H., and Erdman, J. W. Jr. (2005). The tomato as a functional food. *J. Nutr.* 135, 1226–1230. doi: 10.1093/jn/135.5.1226

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

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. *Proc. Natl. Acad. Sci. U. S. A* 111, 4632–4637. doi: 10.1073/pnas.1400822111

Feng, Z., Zhang, Z., Hua, K., Gao, X., Mao, Y., Botella, J. R., et al. (2018). A Highly Efficient Cell Division-Specific CRISPR/Cas9 System Generates Homozygous Mutants for Multiple Genes in Arabidopsis. *Int. J. Mol. Sci.* 19, 3925. doi: 10.3390/ijms19123925

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., et al. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* 31, 822–826. doi: 10.1038/nbt.2623

Giska, F., and Martin, G. B. (2019). PP2C phosphatase Pic1 negatively regulates the phosphorylation status of Pti1b kinase, a regulator of flagellin-triggered immunity in tomato. *Biochem. J.* 476, 1621–1635. doi: 10.1042/BCJ20190299

Guilinger, J. P., Thompson, D. B., and Liu, D. R. (2014). Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol.* 32, 577–582. doi: 10.1038/nbt.2909

Gupta, S., and Van Eck, J. (2016). Modification of plant regeneration medium decreases the time for recovery of Solanum lycopersicum cultivar M82 stable transgenic lines. *Plant Cell Tissue Organ Cult. (PCTOC)* 127, 417–423. doi: 10.1007/s11240-016-1063-9

Hashimoto, R., Ueta, R., Abe, C., Osakabe, Y., and Osakabe, K. (2018). Efficient multiplex genome editing induces precise, and self-ligated type mutations in tomato plants. *Front. Plant Sci.* 9, 916. doi: 10.3389/fpls.2018.00916

Hutin, M., Sabot, F., Ghesquiere, A., Koebnik, R., and Szurek, B. (2015). A knowledge-based molecular screen uncovers a broad-spectrum OsSWEET14 resistance allele to bacterial blight from wild rice. *Plant J.* 84, 694–703. doi: 10.1111/tpj.13042

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

Jacobs, T. B., and Martin, G. B. (2016). High-throughput CRISPR vector construction and generation of tomato hairy roots for validation of DNA modifications and experimental use. *J. Vis. Exp.* 110, e53843. doi: 10.3791/53843

Jacobs, T. B., Zhang, N., Patel, D., and Martin, G. B. (2017). Generation of a collection of mutant tomato lines using pooled CRISPR libraries. *Plant Physiol.* 174, 2023–2037. doi: 10.1104/pp.17.00489

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

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kim, Y. G., Cha, J., and Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. U. S. A* 93, 1156–1160. doi: 10.1073/pnas.93.3.1156

Kimura, S., and Sinha, N. (2008). Tomato (Solanum lycopersicum): A model fruit-bearing crop. *CSH Protoc.* 2008, emo105. doi: 10.1101/pdb.emo105

Klimek-Chodacka, M., Oleszkiewicz, T., Lowder, L. G., Qi, Y., and Baranski, R. (2018). Efficient CRISPR/Cas9-based genome editing in carrot cells. *Plant Cell Rep.* 37, 575–586. doi: 10.1007/s00299-018-2252-2

Kuscu, C., Arslan, S., Singh, R., Thorpe, J., and Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat. Biotechnol.* 32, 677–683. doi: 10.1038/nbt.2916

Li, R., Fu, D., Zhu, B., Luo, Y., and Zhu, H.(2018a). CRISPR/Cas9-mediated mutagenesis of lncRNA1459 alters tomato fruit ripening. *Plant J.* 94,513–524. doi: 10.1111/tpj.13872

Li, T., Yang, X., Yu, Y., Si, X., Zhai, X., Zhang, H., et al. (2018b). Domestication of wild tomato is accelerated by genome editing. *Nat. Biotechnol.* 36, 1160–1163. doi: 10.1038/nbt.4273

Li, J., Manghwar, H., Sun, L., Wang, P., Wang, G., Sheng, H., et al. (2019). Whole genome sequencing reveals rare off-target mutations and considerable inherent genetic or/and somaclonal variations in CRISPR/Cas9-edited cotton plants. *Plant Biotechnol. J.* 17, 858–868. doi: 10.1111/pbi.13020

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

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/s41598-017-00578-x

Oh, C.-S., and Martin, G. B. (2011). Effector-triggered immunity mediated by the Pto kinase. *Trends In Plant Sci.* 16, 132–140. doi: 10.1016/j.tplants.2010.11.001

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

Pedley, K. F., and Martin, G. B. (2003). Molecular basis of Pto-mediated resistance to bacterial speck disease in tomato. *Ann. Rev. Phytopathol.* 41, 215–243. doi: 10.1146/annurev.phyto.41.121602.143032

Pombo, M. A., Zheng, Y., Fernandez-Pozo, N., Dunham, D. M., Fei, Z., and Martin, G. B. (2014). Transcriptomic analysis reveals tomato genes whose expression is induced specifically during effector-triggered immunity and identifies the Epk1 protein kinase which is required for the host response to three bacterial effector proteins. *Genome Biol.* 15, 492. doi: 10.1186/s13059-014-0492-1

Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., et al. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154, 1380–1389. doi: 10.1016/j.cell.2013.08.021

Roberts, R., Hind, S. R., Pedley, K. F., Diner, B. A., Szarzanowicz, M. J., Luciano-Rosario, D., et al. (2019). Mai1 protein acts between Host recognition of pathogen effectors and mitogen-activated protein kinase signaling. *Mol. Plant Microbe Interact.* 32, 1496–1507. doi: 10.1094/MPMI-05-19-0121-R

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 e478. doi: 10.1016/j.cell.2017.08.030

Rosli, H. G., Zheng, Y., Pombo, M. A., Zhong, S., Bombarely, A., Fei, Z., et al. (2013). Transcriptomics-based screen for genes induced by flagellin and repressed by pathogen effectors identifies a cell wall-associated kinase involved in plant immunity. *Genome Biol.* 14, R139. doi: 10.1186/gb-2013-14-12-r139

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

Shimizu, T., Nakano, T., Takamizawa, D., Desaki, Y., Ishii-Minami, N., Nishizawa, Y., et al. (2010). Two LysM receptor molecules, CEBiP and OsCERK1, cooperatively regulate chitin elicitor signaling in rice. *Plant J.* 64, 204–214. doi: 10.1111/j.1365-313X.2010.04324.x

Stegmann, M., Monaghan, J., Smakowska-Luzan, E., Rovenich, H., Lehner, A., Holton, N., et al. (2017). The receptor kinase FER is a RALF-regulated scaffold controlling plant immune signaling. *Science* 355, 287–289. doi: 10.1126/science.aal2541

Tang, X., Liu, G., Zhou, J., Ren, Q., You, Q., Tian, L., et al. (2018). A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. *Genome Biol.* 19, 84. doi: 10.1186/s13059-018-1458-5

Tashkandi, M., Ali, Z., Aljedaani, F., Shami, A., and Mahfouz, M. M. (2018). Engineering resistance against Tomato yellow leaf curl virus *via the* CRISPR/Cas9 system in tomato. *Plant Signal Behav.* 13, e1525996. doi: 10.1080/15592324.2018.1525996

Tsai, S. Q., Wyvekens, N., Khayter, C., Foden, J. A., Thapar, V., Reyon, D., et al. (2014). Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* 32, 569–576. doi: 10.1038/nbt.2908

Van Eck, J., Keen, P., and Tjahjadi, M. (2019). "*Agrobacterium tumefaciens* mediated transformation of tomato," in *Transgenic Plants* (Springer), (New York, NY: Humana Press) 225–234. doi: 10.1007/978-1-4939-8778-8_16

Wang, T., Zhang, H., and Zhu, H. (2019). CRISPR technology is revolutionizing the improvement of tomato and other fruit crops. *Hortic. Res.* 6, 77. doi: 10.1038/s41438-019-0159-x

Wroblewski, T., Tomczak, A., and Michelmore, R. (2005). Optimization of *Agrobacterium*-mediated transient assays of gene expression in lettuce, tomato and Arabidopsis. *Plant Biotechnol. J.* 3, 259–273. doi: 10.1111/j.1467-7652.2005.00123.x

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

Xin, X. F., and He, S. Y. (2013). Pseudomonas syringae pv. tomato DC3000: a model pathogen for probing disease susceptibility and hormone signaling in plants. *Annu. Rev. Phytopathol.* 51, 473–498. doi: 10.1146/annurev-phyto-082712-102321

Xin, X. F., Nomura, K., Aung, K., Velasquez, A. C., Yao, J., Boutrot, F., et al. (2016). Bacteria establish an aqueous living space in plants crucial for virulence. *Nature* 539, 524–529. doi: 10.1038/nature20166

Xu, J., Hua, K., and Lang, Z. (2019). Genome editing for horticultural crop improvement. *Horticult. Res.* 6, 113. doi: 10.1038/s41438-019-0196-5

Yamada, K., Saijo, Y., Nakagami, H., and Takano, Y. (2016). Regulation of sugar transporter activity for antibacterial defense in Arabidopsis. *Science* 354, 1427–1430. doi: 10.1126/science.aah5692

Young, J., Zastrow-Hayes, G., Deschamps, S., Svitashev, S., Zaremba, M., Acharya, A., et al. (2019). CRISPR-Cas9 editing in maize: systematic evaluation of off-target activity and its relevance in crop improvement. *Sci. Rep.* 9, 6729. doi: 10.1038/s41598-019-43141-6

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

Zeng, L., Velasquez, A. C., Munkvold, K. R., Zhang, J., and Martin, G. B. (2012). A tomato LysM receptor-like kinase promotes immunity and its kinase activity is inhibited by AvrPtoB. *Plant J.* 69, 92–103. doi: 10.1111/j.1365-313X.2011.04773.x

Zhang, H., Zhang, J., Wei, P., Zhang, B., Gou, F., Feng, Z., 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

Zhang, N., Roberts, M. H., Van Eck, J., and Martin, G. B. (2019). Generation and molecular characterization of CRISPR/Cas9-induced mutations in 63 immunity-associated genes in tomato reveals specificity and a range of gene modifications. *bioRxiv*. doi: 10.1101/835108

Zheng, Y., Zhang, N., Martin, G. B., and Fei, Z. (2019). Plant Genome Editing Database (PGED): A call for submission of information about genome-edited plant mutants. *Mol. Plant* 12, 127–129. doi: 10.1016/j.molp.2019.01.001

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, 1211–1216. doi: 10.1038/nbt.4272

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.

frontiers in Plant Science

REVIEW
published: 13 February 2020
doi: 10.3389/fpls.2020.00056

Principles, Applications, and Biosafety of Plant Genome Editing Using CRISPR-Cas9

Kaoutar El-Mounadi[1], *María Luisa Morales-Floriano*[2,3] *and Hernan Garcia-Ruiz*[3]*

[1] Department of Biology, Kuztown University of Pennsylvania, Kuztown, PA, United States, [2] Recursos Genéticos y Productividad-Genética, Colegio de Postgraduados, Texcoco, Montecillo, Mexico, [3] Department of Plant Pathology and Nebraska Center for Virology, University of Nebraska-Lincoln, Lincoln, NE, United States

The terms genome engineering, genome editing, and gene editing, refer to modifications (insertions, deletions, substitutions) in the genome of a living organism. The most widely used approach to genome editing nowadays is based on Clustered Regularly Interspaced Short Palindromic Repeats and associated protein 9 (CRISPR-Cas9). In prokaryotes, CRISPR-Cas9 is an adaptive immune system that naturally protects cells from DNA virus infections. CRISPR-Cas9 has been modified to create a versatile genome editing technology that has a wide diversity of applications in medicine, agriculture, and basic studies of gene functions. CRISPR-Cas9 has been used in a growing number of monocot and dicot plant species to enhance yield, quality, and nutritional value, to introduce or enhance tolerance to biotic and abiotic stresses, among other applications. Although biosafety concerns remain, genome editing is a promising technology with potential to contribute to food production for the benefit of the growing human population. Here, we review the principles, current advances and applications of CRISPR-Cas9-based gene editing in crop improvement. We also address biosafety concerns and show that humans have been exposed to Cas9 protein homologues long before the use of CRISPR-Cas9 in genome editing.

Keywords: genome editing, CRISPR-Cas9, Cas9 human exposure, plant breeding, biosafety regulations

OPEN ACCESS

Edited by:
Hiroshi Ezura,
University of Tsukuba, Japan

Reviewed by:
David Gagneul,
Lille University of Science and Technology, France
Vittoria Brambilla,
University of Milan, Italy

***Correspondence:**
Hernan Garcia-Ruiz
hgarciaruiz2@unl.edu

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *16 September 2019*
Accepted: *15 January 2020*
Published: *13 February 2020*

Citation:
El-Mounadi K, Morales-Floriano ML and Garcia-Ruiz H (2020) Principles, Applications, and Biosafety of Plant Genome Editing Using CRISPR-Cas9. Front. Plant Sci. 11:56. doi: 10.3389/fpls.2020.00056

INTRODUCTION

The world population is predicted to reach 10 billion by 2050. While the available farm land and water are being reduced, the global demand for food will increase 25%–70% above current production levels (Hunter et al., 2017). Thus, feeding a rapidly growing population, particularly in the face of climate change, is a big challenge. There is, therefore, an urgent need to improve food production and accelerate sustainable agricultural development.

Long time before the field of genetics was established, humans genetically modified plants through breeding and selection. Without knowledge of genes, mutagenesis, or gene editing, our ancestors influenced the genetic make-up of plants and animals by selecting for traits conducive to food production (Wang et al., 1999; Clark et al., 2005; Li et al., 2013b). A prime example is maize (*Zea mays subsp. mays L.*), which is one of the most produced cereals worldwide. Molecular, cytological, and isozyme profiles have shown that maize is a descendant of an annual species of

teosinte (*Zea mays* ssp. *parviglumis*) native to the Balsas River Valley on the Pacific slopes of the states of Michoacán and Guerrero, Mexico. The process started approximately 9,000 years ago. Teosinte has a popping ability that provided an incentive for its cultivation. Repetitive cycles of selection for traits conducive to kernel production led to the development of the maize plant as we know it (Doebley et al., 1990; Dorweiler et al., 1993; Piperno and Flannery, 2001).

To make plant breeding faster, more predictable, and amendable to a wide range of species, several techniques of plant genetic engineering have been developed. Genome editing through programmable endonucleases is the most recent approach to genetic engineering. Endonucleases are used to specifically induce double strand breaks in target genes of interest. The cellular DNA repair pathway then acts on the double strand break to restore the damage through non-homologous end joining (NHEJ) or homology-directed repair (HDR). In the process, insertions, deletions, substitutions, and DNA recombination may occur (Puchta et al., 1996; Puchta, 2005; Symington and Gautier, 2011).

Three kinds of programmable endonucleases are currently being used for plant genome editing. Zinc finger nucleases, transcription activator-like effector nucleases (TALENs), and CRISPR-Cas9 (Malzahn et al., 2017; Shah et al., 2018; Zhang et al., 2018a; Bao et al., 2019). Zinc finger nucleases are chimeric proteins composed of a synthetic zinc finger DNA binding domain and a DNA cleavage domain. The zinc finger DNA binding domain can be modified to specifically target any long stretch of double stranded DNA of interest (Kim et al., 1996; Cathomen and Joung, 2008). Zinc finger nucleases have been used to edit the genomes of several species, including maize, rice and Arabidopsis (Shukla et al., 2009; Osakabe et al., 2010; Ainley et al., 2013; Gallego-Bartolome et al., 2019).

TALENs are sequence-specific nucleases consisting of transcription activator-like effectors fused to the catalytic domain of the *FokI* endonuclease (Boch et al., 2009; Christian et al., 2010). The DNA-binding domain in TALE monomers in turn is comprised of a central repeat domain (CRD) that directs DNA binding and host specificity. The CRD is formed by tandem repeats of 34 amino acid residues, each binding to one nucleotide in the target nucleotide sequence which allows more flexible target design and increases the number of potential target sites relative to those that can be targeted by zinc finger nucleases (Moscou and Bogdanove, 2009). Genome editing by TALENs has been demonstrated in a wide variety of plants including Arabidopsis (Christian et al., 2013), barley (Budhagatapalli et al., 2015), Brachypodium (Shan et al., 2013), maize (Char et al., 2015), tobacco (Zhang et al., 2013), potato (Clasen et al., 2016; Nicolia et al., 2015), rice (Li et al., 2012; Shan et al., 2013; Shan et al., 2015), soybean (Du et al., 2016), sugarcane (Jung and Altpeter, 2016), tomato (Lor et al., 2014), and wheat (Liang et al., 2014).

The CRISPR-Cas9 system consists of a programmable Cas9 nuclease and a synthetic short guide RNA (sgRNA). DNA target specificity is provided by the guide RNA (**Figure 1**). Thus, the CRISPR-Cas9 system is much easier to be constructed than Zinc finger or TALENs, simple, efficient, has low cost and allows the targeting of multiple genes at once (Cong et al., 2013; Mali et al., 2013).

The CRISPR-Cas9 system has a wide diversity of applications. In medicine, it has been applied in research related to cancer, virus infections, genetic diseases and detection of pathogens. This system has been successfully used in mice to correct mutations in monogenic diseases (Schwank et al., 2013; Ye et al., 2014; Yin et al., 2014; Kang et al., 2015), including the one responsible for Duchenne muscular dystrophy (DMD) (Long et al., 2016; Nelson et al., 2016; Tabebordbar et al., 2016). CRISPR-Cas9 has also been used to disrupt HIV-1 provirus (Ebina et al., 2013), human papillomaviruses (Kennedy et al., 2014) and hepatitis B virus (Kennedy et al., 2015). Furthermore, CRISPR-Cas9 has also been used to target human hereditary liver diseases (Yang et al., 2016; Yin et al., 2016) and has shown great promise for the treatment of cancer (Chen et al., 2019) and Hutchinson–Gilford progeria syndrome (Beyret et al., 2019). In human cells, CRISPR-Cas9 has been used successfully to replace endogenously-encoded antibodies with antibodies protective against Respiratory Syncytial Virus (RSV), human immunodeficiency virus (HIV), influenza virus, and Epstein-Barr virus (EBV) (Moffett et al., 2019). This review is focused on applications of CRISPR-Cas9 in crop improvement.

COMPONENTS OF CRISPR-CAS9

Clustered regularly interspaced short palindromic repeats (CRISPR) are a family of DNA sequences found in the genomes of bacteria and archaea. CRISPRs were first discovered downstream of the alkaline phosphatase isozyme gene (*iap)* in *Escherichia coli* (Ishino et al., 1987). Palindromic repeats are separated by short (32 to 36 bp) sequences derived from the DNA of viruses that have previously infected the cell or its predecessors. These virus-derived sequences integrated into the bacterial genome provide a memory system of previous virus infection. Once integrated into the genome, CRISPRs are transcribed and the virus-derived sequences form short guide RNAs that are bound by CRISPR associated protein 9 (Cas9). Cas9 is a DNA endonuclease. In bacteria and archaea, the natural role of the CRISPR-Cas9 system is to provide adaptive antiviral immunity against DNA viruses. Binary complexes formed by guide RNA-Cas9 recognize and cleave DNA of incoming viruses with sequence similarity to the guide RNA (Garneau et al., 2010; Horvath and Barrangou, 2010; Jinek et al., 2012; Sternberg et al., 2014).

There are several CRISPR-Cas systems in bacteria and archaea. For a comprehensive review, see (Karginov and Hannon, 2010; Sorek et al., 2013). The CRISPR-Cas9 system most frequently used in plant genome editing is an adaptation of the type II CRISPR-Cas system of *Streptococcus pyogenes* (Garneau et al., 2010). *S. pyogenes* is a Gram-positive human-restricted pathogen that colonizes the pharynx and the skin causing an array of diseases ranging from mild sore throat and impetigo to invasive and life-threatening infections (Cunningham, 2000; Rosinski-Chupin et al., 2019). Type II CRISPR-Cas system consists of a Cas9 protein with DNA endonuclease activity and one CRISPR RNA transcript that is

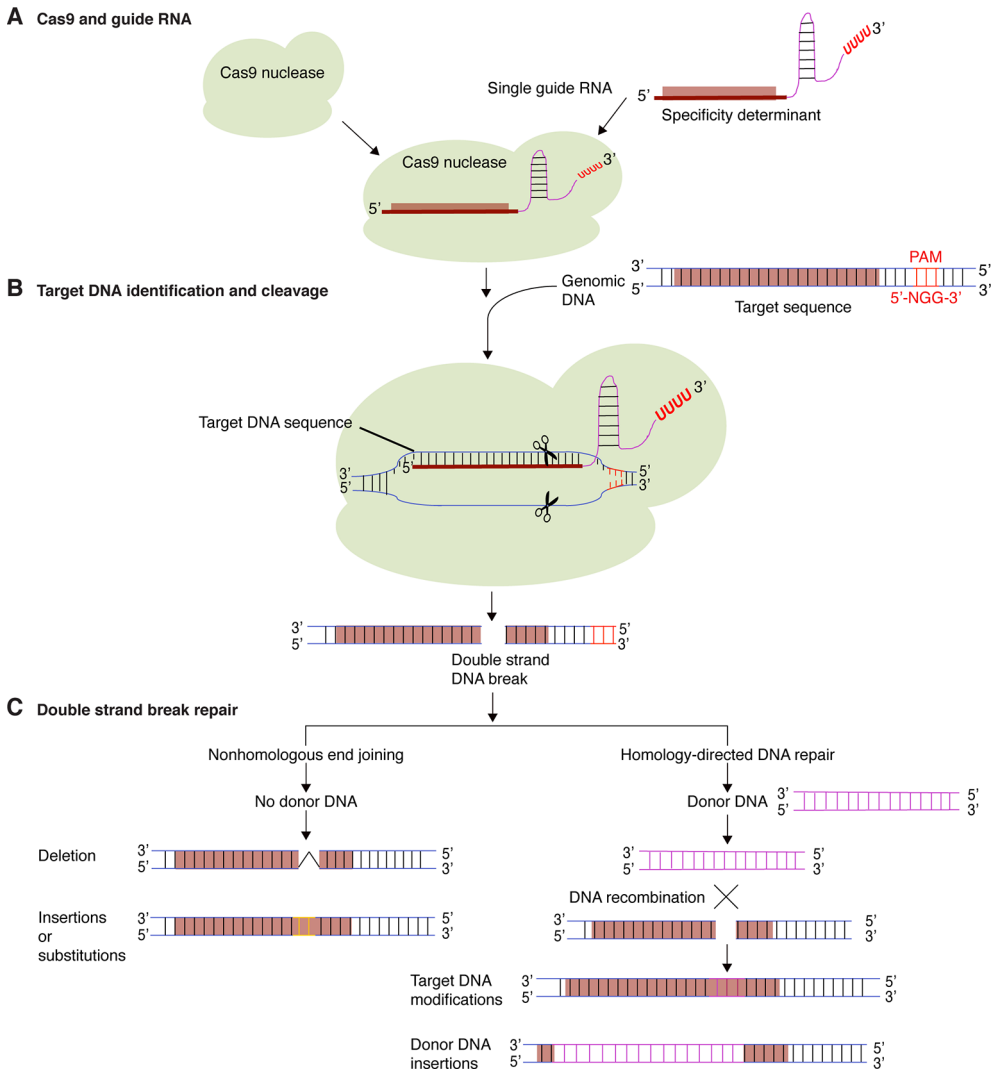


FIGURE 1 | Targeted genome editing using CRISPR-Cas9. **(A)** The CRISPR-Cas9 system consists of a Cas9 protein and one or several guide RNA. Guide RNAs determine target DNA specificity by sequence complementarity**. (B)** Guide RNA and Cas9 protein form a binary complex that specifically cleaves target DNA creating a double-strand DNA break. **(C)** Cellular DNA repair mechanisms, non-homologous end joining (NHEJ) and homology-directed repair (HDR), repairs the double-strand DNA break. In the process, short insertions, deletions, nucleotide substitutions, or gene insertion may occur.

processed to form one or several short guide RNAs that direct Cas9 to the target DNA sequence (**Figure 1**) (Jinek et al., 2012; Lander, 2016; Jiang and Doudna, 2017). In the cell, Cas9 binds to the guide RNA and forms a binary complex that scans the genome for the DNA target for cleavage using Watson-Crick base pairing. The specificity is determined by the guide RNA. Cas9 also requires a specific proto-spacer adjacent motif (PAM) localized on the non-target DNA strand, directly downstream of the target DNA sequence (**Figure 1B**). Cas9 from *S. pyogenes* recognizes NGG as a PAM (Anders et al., 2014; Lander, 2016; Jiang and Doudna, 2017). Cas9 proteins have two signature nuclease domains: HNH and RuvC. The HNH-like nuclease domain cleaves the target DNA strand complementary to the guide RNA sequence. The RuvC-like nuclease domain cleaves the non-target strand (Jinek et al., 2012; Gao et al., 2017; Jiang and Doudna, 2017). This creates a DNA double strand break (DSB) at the target site, which can be subsequently used to introduce modifications by NHEJ or HDR (**Figure 1C**) (Symington and Gautier, 2011). In higher plants, NHEJ occurs most frequently than the more precise HDR, which requires a donor DNA template during homologous recombination to repair the dsDNA breaks. NHEJ does not require a homologous repair template (Puchta et al., 1996; Puchta, 2005). NHEJ has therefore become a popular way to disrupt genes by the creation of small base pair indels (insertions/deletions) at specific points in the target genes, while HDR can be used to precisely introduce specific point mutations and insert or replace desired sequences into the target DNA (**Figure 1C**) (Li et al., 2013a). Currently, it is also possible to precisely edit a single base pair in the genome without the introduction of DSBs

by using engineered Cas9 base editors. The base editors consist of a dead Cas9 domain fused to a cytidine deaminase enzyme that can be programmed with a guide RNA and is able to convert G to A and C to T without inducing dsDNA breaks (Komor et al., 2016). A Cas9 fused with a transfer RNA adenosine deaminase that can mediate conversion of G to A and C to T was also created (Gaudelli et al., 2017). These base editors install point mutations without generating excess undesired editing byproducts. In plants, base editing has been used to efficiently generate point mutations in maize, rice and wheat (Rees and Liu, 2018). These base editors will allow more and better genome modifications and expand the type of cells that can be efficiently edited. In order to make CRISPR-Cas9 a successful genome editing technology in plants, several modifications have been implemented. These include codon modification of the protein Cas9 to ensure its stability in plants, the use of strong constitutive or inducible promotors and the development of versatile DNA cassettes to co-express guide RNAs and Cas9 in the same cells (Li et al., 2013a).

THE GENOME EDITING PROCESS

A fundamental part of the genome editing process is the identification of target genes that determine phenotypes of interest, such as susceptibility to viruses (Garcia-Ruiz, 2018), other pathogens, resistance to herbicides or adverse environmental factors (**Table 1**). Assessment of natural variation or systematic genome-wide screens are also powerful approaches to identify target genes (Kushner et al., 2003; Panavas et al., 2005; Pyott et al., 2016; Giner et al., 2017).

Guide RNAs are artificially designed to specifically direct Cas9 to the target gene to be edited. Bioinformatic programs that generate candidate guide RNAs while accounting for the possibility of off-targets are available (http://crispr.mit.edu/). Dynamic expression vectors have also been designed to clone and co-express guide RNAs and Cas9 (Li et al., 2013a; Fauser et al., 2014; Shimatani et al., 2017). Although variations have been developed recently (Toda et al., 2019), transformation of plant cells to express guide RNAs and Cas9 follows a process similar to these established for the generation of transgenic plants (**Figure 2**). The expression cassettes contain constitutive or inducible promoters, transcription terminators and antibiotic and/or herbicide resistance markers used for selection purposes (Li et al., 2013a; Fauser et al., 2014; Shimatani et al., 2017).

The vector carrying the Cas9 protein and the guide RNA is then introduced into *Agrobacterium tumefaciens* or *Rhizobium rhizogenes* (**Figure 2B**). Colonies containing the CRISPR-Cas9 construct are further used to transform plants by *Agrobacterium*-mediated transformation and first generation transgenic plants are identified by antibiotic or herbicide selection (Li et al., 2013a; Pyott et al., 2016; Veillet et al., 2019). Green fluorescent protein (GFP) has also been used to distinguish cells or calluses containing the CRISPR-Cas9 cassette (Doench et al., 2014). In all cases, sequencing the target gene is required in order to identify the mutations introduced by genome editing. The presence of the CRISPR-Cas9 cassette renders the plants transgenic and thus subject to the corresponding biosafety regulations (Callaway, 2018; Garcia Ruiz et al., 2018; Eckerstorfer et al., 2019). However, in sexually propagated plants, after identification of the genome edited plants, the CRISPR-Cas9 transgene can be eliminated by Mendelian segregation (**Figures 2D–E**) (Zhang et al., 2019a). This key part of the process removes the transgene in the third or subsequent generations resulting in the formation of genome-edited plants without a transgene (Pyott et al., 2016). Because of the absence of the transgene in these plants, they resemble those with mutations generated by natural means or chemical mutagenesis (Lellis et al., 2002; Pyott et al., 2016).

Because the introduction of the CRISPR-Cas9 cassette as a transgene might be controversial under certain regulations in some countries (**Table 2**), protocols have been developed to edit genomes without transgenes using guide RNA-Cas9 ribonucleoprotein complexes (Liang et al., 2017; Toda et al., 2019) or transient expression (Zhang et al., 2016).

Not all plant species are susceptible to *A. tumefaciens*. In species recalcitrant to Agrobacterium-mediated transformation, alternatives include *Rhizobium rhizogenes*-mediated or protoplast transformation. *R. rhizogenes* previously known as *Agrobacterium rhizogenes*, is a soil-borne gram-negative bacteria that causes hairy roots in plants. In most plant species, cells transformed with *R. rhizogenes* and its Ri plasmid differentiate into transformed roots, serving as a visual marker for marker-free screening and selection (Young et al., 2001; Bahramnejad et al., 2019). A comprehensive description of *R. rhizogenes* strains, binary vectors, and plants transformed using them is provided by (Bahramnejad et al., 2019). Examples of plants edited through CRISPR-Cas9 using *R. rhizogenes* include soybean (Du et al., 2016), tomato and rubber producing dandelion *Taraxacum kok-saghyz* (Iaffaldano et al., 2016).

PROTOPLAST OR ZYGOTE TRANSFORMATION

For several plant species, including maize, soybean, wheat, rice, tomato, lettuce, arabidopsis, petunia, grapevine, apple, potato, and tobacco, protocols have been developed to isolate protoplast and transfect them with cassettes carrying CRISPR-Cas9 for genome editing purposes. Protoplast transfection has been used to rapidly optimize CRISPR-Ca9 parameters (Woo et al., 2015; Lin et al., 2018). However, isolation of single-protoplast has been used to regenerate stable transformants after transfection with cassettes carrying CRISPR-Cas9 or with ribonucleoprotein complexes assembled *in vitro* by synthesizing small guide RNAs and Cas9 protein. Gene editing using ribonucleoprotein complexes has the advantage of obtaining mutants without the presence of exogenous DNA. Preassembled Cas9-guide RNA ribonucleoproteins complexes can be delivered into protoplasts using polyethylene glycol-calcium-mediated transfection (Woo et al., 2015; Kim et al., 2017; Liang et al., 2017; Lin et al., 2018). To overcome the low efficiency of this approach, a protocol has been develop to transform plant zygotes by ribonucleoprotein complexes or by biolistic bombardment (Toda et al., 2019).

TABLE 1 | Representative applications of CRISPR-Cas9 in crop breeding.

Group	Crop species	Target gene	Role	Modification	Target trait	Reference
Monocotyledon	Maize	*ZmTMS5*	Causes the TGMS trait	Gene knockout	Thermosensitive genic male sterility	(Li et al., 2017)
Monocotyledon	Sorghum	*k1C*	Encode 22-kD α-kafirin proteins	Genes disruption in N-terminal ER signal peptide region	High Lysine content and increased protein digestibility	(Li et al., 2018a)
Monocotyledon	Wheat	*TaEDR1*	Negative regulator of the defense response against powdery mildew	Knock-down all three homologs of *TaEDR1*	Powdery mildew resistance	(Zhang et al., 2017b)
Monocotyledon	Wheat	*TaGW2-A1, -B1 and -D1.*	Genetic control of grain weight and protein content traits	Homologous genes knockout	Grain weight and protein content increase	(Zhang et al., 2018a)
Monocotyledon	Wheat	*Ms1*	Male fertility gene	Gene knockout	Male sterility	(Okada et al., 2019)
Monocotyledon	Rice	*OsRR22*	Transcription factor	Inactivating mutations	Enhanced salinity tolerance	(Zhang et al., 2019a)
Monocotyledon	Rice	*CAO1 and LAZY1*	Synthesis of Chl b from Chl a and regulating shoot gravitropism, respectively	Genes' disruption	Defective synthesis of Chlorophyll b and tiller-spreading phenotypes	(Miao et al., 2013)
Monocotyledon	Rice	*SBEI and SBEIIb*	Determining the amylose content, fine structure of amylopectin, and physiochemical properties of starch	Genes disruption	Higher proportion of long chains in amylopectin	(Sun et al., 2017)
Monocotyledon	Rice	*Gn1a, DEP1, GS3 and* IPA1	Regulators of grain number, panicle architecture, grain size and plant architecture, respectively	Genes disruption	Enhanced grain number, dense erect panicles, and larger grain size, respectively	(Li et al., 2016)
Monocotyledon	Rice	*OsERF922*	Negative regulator of Rice blast resistance	Gene disruption	Enhanced rice blast resistance	(Wang et al., 2016a)
Monocotyledon	Rice	*OsSWEET13*	Sucrose transporter. Negative regulator of bacterial blight resistance	Gene knockout	Bacterial blight resistance	(Zhou et al., 2015)
Monocotyledon	Rice	*OsMATL*	Encodes a pollen-specific phospholipase	Gene knockout	Haploid seed formation	(Yao et al., 2018)
Monocotyledon	Rice	*ALS*	Acetolactate synthase encoding gene	Gene disruption	Herbicide resistance	(Endo et al., 2016)
Monocotyledon	Rice	*ALS*	Acetolactate synthase encoding gene	Gene replacement	Herbicide resistance	(Sun et al., 2016)
Monocotyledon	Rice	*TMS5*	Thermo-sensitive genic male sterility gene	Gene knockout	Thermo-sensitive genic male sterility	(Zhou et al., 2016)
Monocotyledon	Cavendish banana *Musa acuminata*)	*PDS*	Phytoene desaturase encoding gene	Gene knockout	Albinism phenotype	(Naim et al., 2018)
Monocotyledon	Banana *(Musa spp.)*	Integrated endogenous *banana streak virus* (eBSV) in the B genome of plantain	The eBSV activates into infectious viral particles under stress	Knockout the integrated dsDNA of BSV from the banana genome	Asymptomatic plants to *banana streak virus*	(Tripathi et al., 2019)
Dicotyledon	*Camelina sativa*	*FAD2*	Fatty acids biosynthesis	Genes knockout	Improve seed Oleic acid content	(Jiang et al., 2017)
Dicotyledon	*Arabidopsis thaliana*	The *FWA* and the *SUPERMAN* promoters.	Flowering time gene and a transcriptional regulator of floral homeotic genes	Genes knock in.	Targeted gene activation and DNA methylation in Arabidopsis	(Papikian et al., 2019)
Dicotyledon	*Arabidopsis thaliana*	*CBFs*	C-repeat binding factors encofing genes, key transcription factors in the cold stress response	Genes disruption. Deletions and insertions	Cold tolerance	(Jia et al., 2016b)
Dicotyledon	Tomato	*SlJAZ2*	Important repressor in jasmonate signaling pathway. Key regulator of stomatal aperture during biotic stresses	Gene knock in, lacking the C-	Bacterial speck resistance	(Ortigosa et al., 2018)

(Continued)

TABLE 1 | Continued

Group	Crop species	Target gene	Role	Modification	Target trait	Reference
				terminal Jas domain		
Dicotyledon	Tomato	*SlMlo1*	Confers susceptibility to fungi, causing the powdery mildew disease	Gene disruption. 48 bop deletion	Powdery mildew resistance	(Nekrasov et al., 2017)
Dicotyledon	Tomato	*SP5G*	Florigen paralog and flowering repressor	Gene knockout	Rapid flowering. Early yield	(Soyk et al., 2017)
Dicotyledon	Tomato	*SlAGL6*	Transcription factor. It plays essentials roles, especially in flower meristem and floral organ development	Gene knockout	Parthenocarpic phenotype	(Klap et al., 2017)
Dicotyledon	Tomato	*SlIAA9*	Key gene controlling parthenocarpy	Gene knockout	Parthenocarpic phenotype	(Ueta et al., 2017)
Dicotyledon	Tomato	*SlMAPK3*	Mitogen-activated protein kinases 3 encoding gene, responds to drought stress	Gene knockout.	Drought tolerance	(Wang et al., 2017)
Dicotyledon	Tomato	*CrtR-b2* and *Psy1.*	Key genes of carotenoid biosynthesis	Genes knockout	Changes on carotenoids profile	(D'ambrosio et al., 2018)
Dicotyledon	Wild tomato	*SELF-PRUNING, OVATE, FASCIATED and FRUIT WEIGHT 2.2, MULTIFLORA and LYCOPENE BETA CYCLASE*	Encode general plant growth habit, fruit shape, fruit size, fruit number and nutritional quality, respectively	Genes knockout	Obtain domestication traits (fruit number, size, shape, nutrient content and plant architecture)	(Zsogon et al., 2018)
Dicotyledon	Stress-tolerant wild-tomato	*SP, SP5G, SlCLV3, SlWUS and SlGGP1*	Flowering repressors, small-peptide-encoding gene, homeobox-encoding gene and vitamin C–biosynthetic enzyme encoding gene.	Genes disruption. Insertions, deletions and invertions.	Domesticated phenotypes yet retained parental disease resistance and salt tolerance	(Li et al., 2018b)
Dicotyledon	Potato	*GBSS*	Granule-bound starch synthase encoding gene, is responsible for amylose synthesis	Gene knockouts	Increased amylopectin content	(Andersson et al., 2017)
Dicotyledon	Cucumber	*eIF4E*	Eukaryotic translation initiation factor. Is a central part of the translation machinery	Gene knockout	Cucumber Vein Yellowing Virus, Zucchini yellow mosaic virus and Papaya ring spot mosaic virus-W resistance	(Chandrasekaran et al., 2016)
Dicotyledon	Soybean	*GmFT2a*	Integrator in the photoperiod flowering pathway in soya bean	Gene disruption.1-bp insertion or short deletion	Late flowering	(Cai et al., 2018)
Dicotyledon	Grape	VvWRKY52	Transcription factor gene that plays important roles in plant defense regulatory networks in grape	Gene knockout	*Botrytis cinerea* resistance	(Wang et al., 2018)
Dicotyledon	Oranges	*CsLOB1*	Plays a critical role in promoting pathogen growth and erumpent pustule formation	Disruption of *CsLOB1* promoter. Deletions, insertions and substitutions	Citrus canker resistance	(Peng et al., 2017)
Dicotyledon	Grapefruits	*CsLOB1*	Critical citrus disease susceptibility gene for citrus canker	Disruption the coding region of both alleles of *CsLOB1*	Citrus canker resistance	(Jia et al., 2017)
Dicotyledon	Grapefruits	*CsLOB1*	Plant-specific transcriptional factor in the lateral organ boundaries (LOB) domain family	Disruption of the PthA4 effector binding elements in the Type I CsLOB1 Promoter	Citrus canker alleviated	(Jia et al., 2016a)
Fungus	Mushroom	*PPO*	Enzymes that use molecular oxygen to oxidize *ortho*-diphenols to *ortho*-quinones. These commonly cause browning reactions following tissue damage, and may be important in plant defense. Some PPOs function as hydroxylases	Knockout of one of six *PPO* genes	Non-browning phenotype	(Waltz, 2016b)

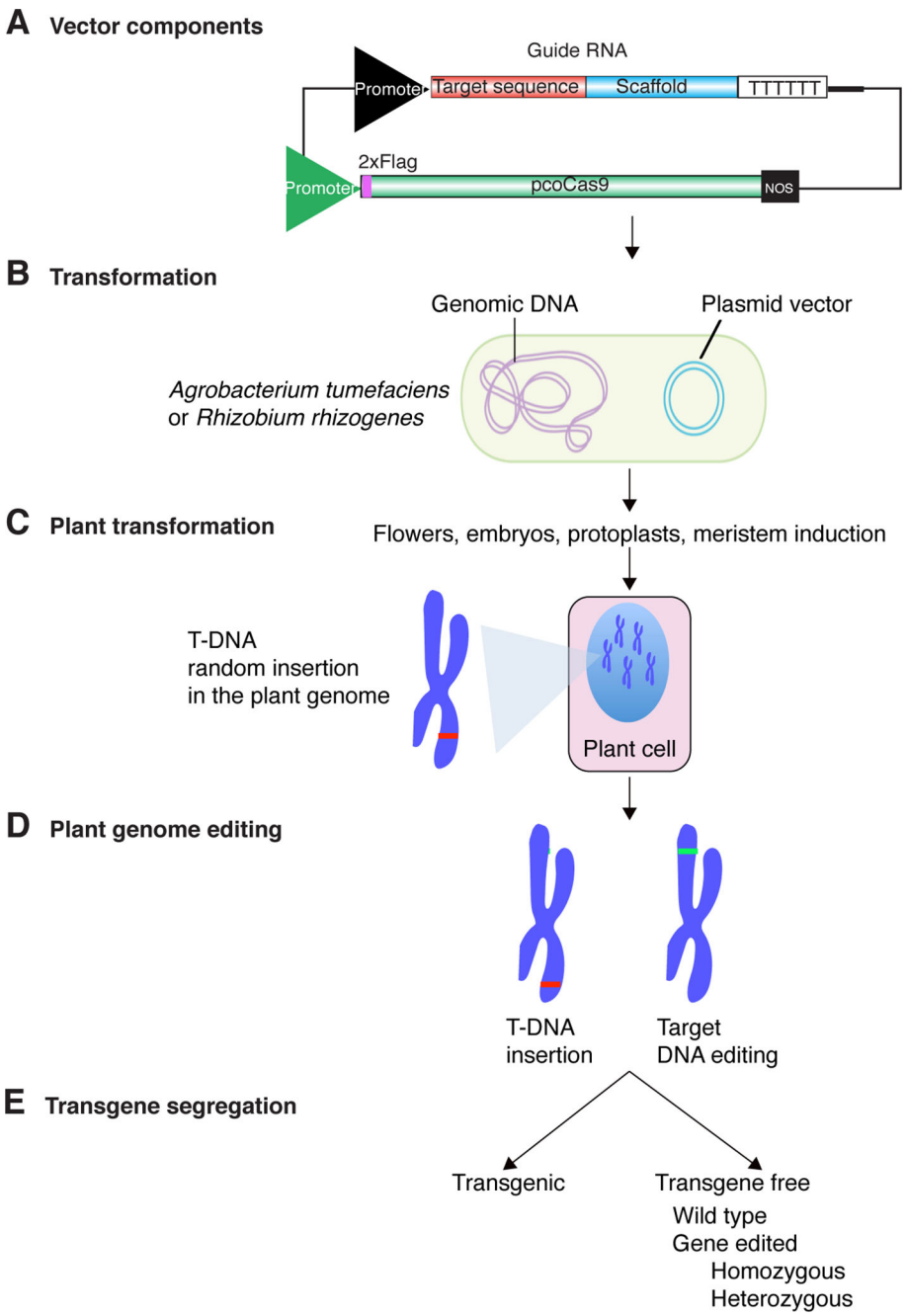


FIGURE 2 | Genome editing process using CRISPR-Cas9 and Agrobacterium tumefaciens. **(A)** Cas9 protein and guide RNAs are cloned into the same plasmid vector containing transfer DNA (T-DNA) signals. Expression is driven by strong constitutive (U6, 35S, or other), inducible or tissue specific promoters. Transcription termination is programmed by addition of terminator such as the U6 or Nopaline synthase (NOS). For plant genome editing purposes, Cas9 has been codon-optimized and might contain an epitope tag to determine expression. **(B)** *A. tumefaciens* or *R. rhizogens* is transformed with the plasmid vector carrying the cassette for Cas9 protein and guide RNAs expression. **(C)** Bacteria is used to transform embryos, ovules in flowers, protoplasts, roots, or cells in leaves. Integration site of the T-DNA is random. **(D)** Expression of Cas9 protein and guide RNAs lead to editing of the target DNA. The T-DNA insertion site and the DNA target are likely not linked. **(E)** The T-DNA insertion and edited part of the genome can be separated by Mendelian segregation.

DE NOVO INDUCTION OF MERISTEMS

Delivering the CRISPR-Cas9 cassette into the germ line or protoplasts is technically challenging and inefficient. However, in dicotyledonous plants, those limitations might be eliminated through *de novo* induction of meristems. Developmental regulators and gene-editing components are delivered into somatic cells of whole plants. From treated tissue, shoots

TABLE 2 | Regulation of genetically modified and genome edited plants across countries.

Country	Genetically modified plants[1]	Genome-edited plants[2]
Argentina	Regulated	Case-by-case, mostly non-regulated
Australia	Regulated	Non-regulated
Brazil	Regulated	Case-by-case, mostly non-regulated
Canada	Regulated	Regulated
Chile	Regulated	Case-by-case, mostly non-regulated
European Union	Regulated/opposed	Regulated/Opposed
India	Regulated	Regulated
Japan	Regulated	Non-regulated
Malaysia	Regulated	Regulated
Mexico	Regulated	Regulated
New Zealand	Regulated	Regulated
South Africa	Regulated	Regulated
Thailand	Regulated	Regulated
United States of America	Regulated	Non-regulated

[1]*Refers to the final product containing transgenes, such as selection markers or other form of foreign DNA used during the process.*
[2]*Refers to the final product lacking transgenes that might have been used during the process.*

emerge that contain the targeted DNA modifications that are transmitted to the next generation (Maher et al., 2019).

APPLICATIONS OF GENOME EDITING IN CROP IMPROVEMENT

Genome editing with CRISPR-Cas9 is amendable to edit any gene in any plant species. Because of its simplicity, efficiency, low cost, and the possibility to target multiple genes, it allows faster genetic modification than other techniques. It also can be used to genetically modify plants that were previously neglected. The potential that this represents for crop breeding and the development of sustainable agriculture is incommensurable (Cong et al., 2013; Mali et al., 2013; Zhang et al., 2017a; Toda et al., 2019; Wurtzel et al., 2019; Zhang et al., 2019b).

Impressive genetic modifications have been achieved with CRISPR-Cas9 to enhance metabolic pathways, tolerance to biotic (fungal, bacterial or viral pathogens), or abiotic stresses (cold, drought, salt), improve nutritional content, increase yield and grain quality, obtain haploid seeds, herbicide resistance, and others (**Table 1**). Notable cases include thermosensitive genic male sterility in maize (Li et al., 2017) and wheat (Okada et al., 2019), improved nutritional properties in sorghum and wheat (Li et al., 2018a; Zhang et al., 2018b), tolerance or resistance to pathogens (Zhang et al., 2017b; Pyott, 2016), and resistance to herbicides (Endo et al., 2016; Sun et al., 2016).

In potato CRISPR-Cas9 was used to knockout the gene encoding granule-bound starch synthase (GBSS) in one round of transfection resulting in the development of potato plants that produce amylopectin starch, a highly desirable commercial trait (Andersson et al., 2017). In cucumber CRISPR-Cas9 system was used to inactivate the eukaryotic translation initiation factor gene *eIF4E*. The resulting non-transgenic homozygotic mutant plants were immune to Cucumber vein yellowing virus (Genus *Ipomovirus*) and resistant to the potyviruses Zucchini yellow mosaic virus and Papaya ring spot mosaic virus (Chandrasekaran et al., 2016). Engineering genetic resistance to viruses and other pathogens has immense potential to manage diseases for which no natural resistance has been detected, such as maize lethal necrosis disease and tomato brown rugose fruit virus (Luria et al., 2017; Garcia-Ruiz, 2018; Wamaitha et al., 2018).

HUMAN EXPOSURE TO CAS9 PROTEINS

In many bacteria and most archaea, CRISPR-Cas provides acquired immunity against viruses and plasmids by targeting nucleic acid in a sequence-specific manner (Horvath and Barrangou, 2010). Comparative genomic analyses revealed that CRISPR and genes coding for their associated proteins were present in diverse bacterial phylogenetic groups (Haft et al., 2005; Lillestol et al., 2006; Makarova et al., 2006). Since this adaptive immune system is useful for bacterial survival, it is likely to be present in all bacteria.

We compared the amino acid sequence of the Cas9 protein from *S. pyogenes* used in plant genome editing to proteins from bacteria to which humans are exposed through food consumption or in the environment. Results showed that Cas9 from *S. pyogenes* has 23% to 58% similarity to Cas9 protein from *Streptococcus thermophilus*, a bacterium widely used as a probiotic and in the production of cheese and yogurt (**Figure 3**). Additionally, Cas9 from *S. pyogenes* shares up to 35% similarity with Cas9 proteins from a wide range of bacteria used in food production such as *Lactobacillus plantarum* used to make cheese, yogurt, kefir and other fermented milk and meat products as well as fermented vegetables and beverages (Coloretti et al., 2007; Zago et al., 2011; Khemariya et al., 2016; Settachaimongkon et al., 2016; Sidira et al., 2017; Behera et al., 2018). *L. plantarum* is frequently encountered as a natural inhabitant of the human gastrointestinal tract, in which it is a transient guest acquirable through the diet (Vesa et al., 2000; De Vries et al., 2006). Additionally, *L. plantarum* is often used as a probiotic and can improve the balance of beneficial intestinal microflora (Nguyen et al., 2007; Nagpal et al., 2012; Kassayova et al., 2014).

Furthermore, Cas9 from *S. pyogenes* has homologues in diverse Gram-positive and Gram-negative bacteria that occupy very diverse niches throughout the human body (Louwen et al., 2014). Some are commensals and others are pathogenic bacteria. More than 80% amino acid sequence similarity was detected between Cas9 from *S. pyogenes* and that from human commensal and pathogenic bacteria such as *Streptococcus dysgalactiae subsp. equisimilis*, *Staphylococcus aureus*, *Klebsiella pneumonia* and *S. canis* (**Figure 3**).

These observations show that humans have been exposed to Cas9 proteins in their food and environment long before the development of genome editing. The biosafety risk of human exposure to the Cas9 used for plant genome editing needs further

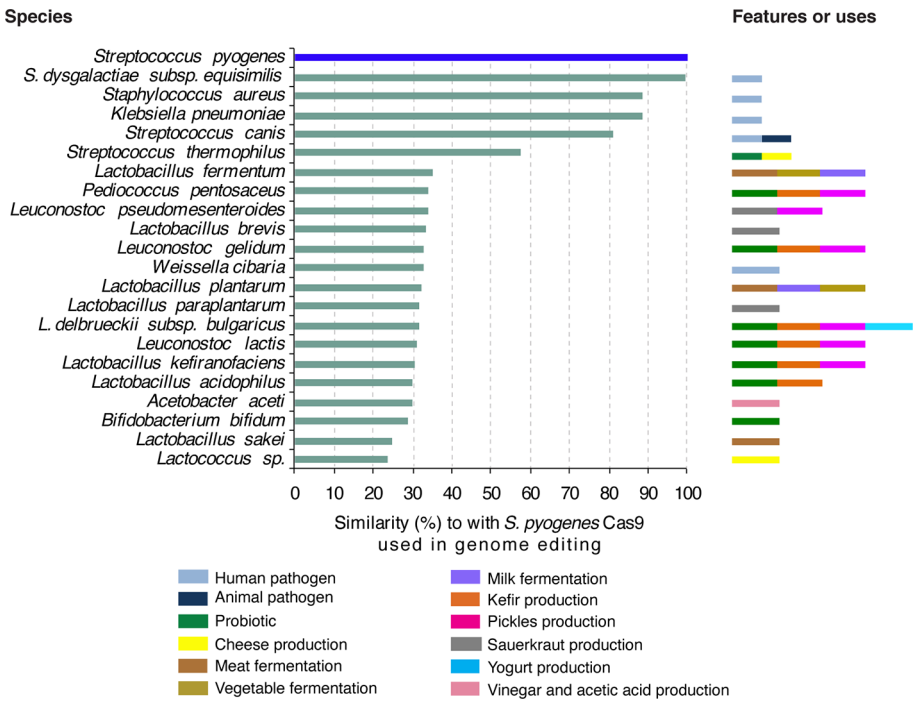


FIGURE 3 | Bacteria frequently in contact with humans and similarity of their proteins to S. pyogenes Cas9 frequently used in plant genome editing. Amino acid sequence of *S. pyogenes* Cas9 was used to search for homologues proteins in GenBank. Proteins with more than 20% similarity are indicated. Features or anthropocentric use of each bacteria species are color-coded.

assessment (Pineda et al., 2019) and our results do not mean that potential human exposure to Cas9 used in genome editing is irrelevant.

BIOSAFETY CONCERNS ABOUT GENOME-EDITED PLANTS

Methodological, biosafety and social concerns remain about the use of genome editing in plants. They mostly are related to target gene site selection, guide RNA design, off-target effects, and the delivery method. The major concern is the risk of generating unwanted genetic changes in plants due to off-target mutations (Liang et al., 2018; Pineda et al., 2019). Fragments of the CRISPR-Cas9 might be degraded into filler DNA and inserted into expected and/or unexpected genomic positions during the DNA repair process (Gorbunova and Levy, 1997; Zhang et al., 2016). However, transgene integration and the risk of off-target mutations can be prevented by delivering *in vitro* pre-assembled CRISPR-Cas9 ribonucleoproteins (Malnoy et al., 2016; Svitashev et al., 2016; Zhang et al., 2016; Liang et al., 2018). This technique has already been used in several crop species but there are still some drawbacks in its application such as low stability, high costs and high levels of technical requirements, which need to be improved (Malnoy et al., 2016; Subburaj et al., 2016; Murovec et al., 2018).

Substantial work has also been done to minimize off-target effects of Cas9 itself, including improving RNA guide–design strategies, ribonucleoprotein delivery, protein engineering, using spatiotemporally controlled Cas9, and/or gRNAs through a plethora of chemical or environmental inducers, or using synthetic genetic circuits that modulate CRISPR function according to predefined logic (Svitashev et al., 2016; Liang et al., 2018). Base editing is also being modified to improve the specificity of base editors by limiting deaminase activity outside of Cas9 binding through the use of different deaminase effectors or rationally engineering the deaminase to decrease its DNA binding ability (Shimatani et al., 2017).

Other concerns about CRISPR-Cas9 technology are related to the Cas9 protein itself as it was shown to induce an immune response when delivered by adeno-associated virus in mice, making immunogenic side effects a concern (Chew et al., 2016). There are also concerns about the specificity of Cas9 and the limited number of sites which can be targeted due to the requirement of the PAM (Spencer and Zhang, 2017). Protein engineering efforts led to the identification of mutations in Cas9 that alter its PAM recognition and enhance its fidelity and recognize other motifs (Kleinstiver et al., 2015; Kleinstiver et al., 2016; Leenay and Beisel, 2017). Further modifications to Cas9 and guide RNA design, such as FokI fusions, paired nicking, and the use of truncated guide RNAs, have provided additional improvements to specificity (Wyvekens et al., 2015). Furthermore, Cas9 variants, Cas9 homologs derived from other

bacteria, or novel Cas proteins such as Cpf1 nucleases can be used (Nakade et al., 2017; Pineda et al., 2019).

The societal concerns about genome editing stem in part from the lack of information about its principles and applications. A fundamental feature here is the distinction between genetically modified plants, transgenic plants, and genome edited plants (Garcia Ruiz et al., 2018; Eckerstorfer et al., 2019). Genome edited plants may or may not be transgenic. As indicated above, the transgene carrying the CRISPR-Cas9 cassette might be removed by gene segregation (**Figure 2**). If this is done, a genome-edited plant might be classified as non-transgenic. Educating the public on the principles of genome editing has the potential to correct and prevent the spread of misconceptions (Garcia Ruiz et al., 2018; Eckerstorfer et al., 2019).

REGULATION OF GENOME-EDITED CROPS

The term genetically modified refers to plants whose genome has been modified in a way that would not have been occurred naturally (Wang et al., 2016b; Duensing et al., 2018; Friedrichs et al., 2019).

In contrast, gene editing refers to DNA modifications similar to those potentially generated naturally (deletions, nt substitutions, insertions) of by conventional plant breeding (Nature Plants Editorial, 2018). The basis to regulate the release and international trade of living genetically modified organisms were established in the Cartagena Protocol on Biosafety. However, production, consumption, and regulation of genetically modified plants have followed contrasting patterns. While some countries reject consumption and ban production, others openly grow and consume them (Garcia Ruiz et al., 2018).

Regulation of genome-edited plants follows two frameworks. Some countries regulate the process, while others regulate characteristics of the final product (Eckerstorfer et al., 2019; Van Vu et al., 2019). While some countries have established biosafety regulations for genome edited plants, or declared their deregulation (**Table 2**), most countries have not yet established their position (Eckerstorfer et al., 2019). Challenges in regulating plant genome editing include market access, and addressing the societal concerns about its biological safety without limiting the development of the technology (Kupferschmidt, 2018; Eckerstorfer et al., 2019). Transgene-free, genome-edited plants are similar to varieties containing genetic variations created naturally (**Figure 2**). Therefore, commercialization of genome edited plants or their products might bypass the strict biosafety regulations required for transgenic plants (Tuteja et al., 2012; Van Vu et al., 2019).

The United States Department of Agriculture (USDA) declared in March 2018 that genome editing is the equivalent of conventional breeding in some instances and therefore does not require regulatory oversight within the American regulatory framework (Waltz, 2016a). A mushroom engineered to resist browning and a waxy corn engineered to contain starch composed exclusively of amylopectin are the first CRISPR edited crops to be approved for commercialization in the USA with no regulations (Waltz, 2016b). The decision not to regulate was based on the fact that no foreign DNA (transgene) was inserted during editing and that the resulting change did not involve resistance to pesticides or herbicides.

Canada, on the other hand, has remained committed to the scientific principles laid down in its domestic regulatory framework for plants with novel traits established 25 years ago. Canada's policy states that any gene editing technology that creates a novel product is subject to additional regulatory oversight on allergenicity, toxicity and impacts on non-target organisms (Smyth, 2017). Two products obtained by gene editing have been approved in Canada, non-browning apples and non-dark spots potatoes (Waltz, 2016b). The approval was granted after a lengthy evaluation process that determined that the changes made to the apples and the potatoes did not pose a greater risk to human health than apples and potatoes currently available on the Canadian market (Waltz, 2016b).

Argentina has developed a functional regulatory system for the approval of genome-edited products (Whelan and Lema, 2015). The regulatory system was developed to be consistent with the Cartagena Protocol on Biosafety and relies on case-by-case assessment. If a transgene technology was used in the development of a product, where the final product is free of the transgene, then this product can be classified as nontransgenic. Chile and Brazil followed Argentina's lead. Chile signed a normative resolution in 2017 while Brazil published a resolution in January 2018 (Duensing et al., 2018). Both regulate gene-edited products on a case-by-case basis and exempt them from regulation when there is no insertion of transgenes.

Meanwhile, European Union (EU) countries remain politically opposed to genetically modified crops (Waltz, 2016b). On July 2018, the Court of Justice of the European Union (ECJ) ruled that gene-edited crops should be subject to the same stringent regulations as conventional genetically modified (GM) organisms. In its ruling, the ECJ determined that only mutagenesis techniques that have conventionally been used in a number of applications and have a long safety record are exempt from this rule.

In Australia, the Gene Technology Act (GT Act), introduced in 2000, stipulates that a GMO is an organism produced by any technique that modifies genes or other genetic material. In 2001, the Gene Technology Regulations were introduced. Schedule 1 of these regulations, specifies that organisms resulting from an exchange of DNA in which the donor species is also the hosts species and the vector DNA does not contain heterogenous DNA as not GMOs. In October 2019, an amendment to schedule 1 came in effect. The amendment excludes organisms modified through CRISPR-Cas9 and other unguided repair of site-directed nuclease activity (SDN), from being regulated as GMOs. The amendment also indicates that organisms generated in the intermediated steps of the SDN method are deemed non GMOs if 1) no nucleic acid template is supplied to guide genome repair through homology-directed recombination, and 2) the organism has no other modifications as a result of the gene technology (Eckerstorfer et al., 2019).

In New Zealand, importation, development, field testing, and release of GMOs genetically modified are regulated by the Hazardous Substances and New Organisms Act 1996 (HSNO Act). The country has the most rigorous and comprehensive process for regulation of GMOs. As a result of that, no GMO commercial crops are grown in the country and no GM meat or fresh produce is sold in the country. Furthermore, processed food that contains imported GM ingredients is tested for safety and should be labeled as so. In 2016, the HSNO Act was amended with an article stating that plant breeding by genome editing is subject to the same regulations as the GMOs (Shimatani et al., 2017).

India's regulatory process for research, development and use of GMOs and their products, including new gene technologies was established in 1989. The Food Safety and Standards Authority of India define genetically engineered or modified food as "any food or food ingredient composed or containing genetically modified or engineered organisms obtained through modern biotechnology, or food and food ingredients produced from but not containing genetically modified or engineered organisms obtained through modern biotechnology". Thus all new technologies including CRISPR-Cas9 gene technologies (including genome editing) are still regulated within the existing regulatory framework (Friedrichs et al., 2019).

Japan's Ministry of Health, Labor and Welfare (MHLW) has recently declared that foods derived from genome editing technologies which do not contain transgenic genes and/or fragments of transgenic genes are not considered GMOs and are not subject to regulations as long as the DNA double-strand break induced by the genetic engineering method is either a base-pair deletion, a naturally occurring gene deletion and/or a concomitant insertion of one to several base pairs. The new MHLW's policy also indicates that off-target mutations in GE foods should not be of concern as they can also be observed in multiple locations in the genome of crops produced by traditional breeding (South et al., 2019; Van Vu et al., 2019).

It seems that the decision to regulate or not regulate GE crops and foods depends mainly on the type of GMO regulatory system already in place in the country. Countries that have adopted a process-based GMO regulatory system and consider that products made using the regulated process are fundamentally different or more risky than similar products made using other methods will likely regulate GE crops and foods under the GMO laws. On the other hand, countries who follow a product-based regulatory system and regulate based on the characteristics of the final product rather than the process by which it was made might not regulate GE crops and foods under GMO laws. Countries, such as Malaysia and Thailand, who adopt a dual product and process approach will also likely to regulate GE crops and foods under GMO laws (Friedrichs et al., 2019).

FUTURE PERSPECTIVES

Due to the many practical applications related to food production, genome editing can and will be used to solve agricultural issues that directly affect food security, such a citrus greening disease (Taylor et al., 2019), and the high yield losses in C3 plants, such as rice and barely, due to inefficient photorespiration in these crops. A recent report described the construction of three synthetic glycolate metabolic pathways in tobacco chloroplasts with the aim of improving the plant's photosynthetic efficiency. Flux through the synthetic pathways was maximized by inhibiting glycolate export from the chloroplast using RNA interference to down-regulate a native chloroplast glycolate transporter. In the transgenic tobacco plants, the photosynthetic yield increased by 20% while biomass productivity improved by more than 40% (South et al., 2019). While this study did not use genome editing technology and was carried out in *Nicotiana tabacum*, a model plant, the concept can be applied easily and successfully in staple crops using CRISPR-Cas9. Successful improvement in photorespiration efficiency in crops such as maize, rice and wheat, has the potential to substantially increase food production for the growing worldwide population while using the same cultivation areas and without having to destroy more forest areas for agricultural purposes. Producing crops with better quality food through genome editing will also help achieve food security (Li et al., 2018b; Narayanan et al., 2019).

CONCLUSIONS

Genome editing in general, and CRISPR-Cas9 in particular, is a revolutionary tool that can impact science, food production, and society. CRISPR-Cas9 has great potential for transforming agriculture by making plants tolerant to biotic and abiotic stresses and improving their nutritional value and yield. These attributes are necessary to meet the demand of an increasing world population. In order to be able to effectively and durably use this technology in crop improvement, the scientific community needs to address the various biosafety and societal concerns about it. There is also a need to re-evaluate the regulations of genome-edited plants and to educate the general public about their properties.

AUTHOR CONTRIBUTIONS

HG-R conceived the study. KE-M, MM-F, and HG-R performed the analysis. KE-M, MM-F, and HG-R wrote the paper.

FUNDING

This research was supported by NIH grant R01GM120108 to HG-R and by the Nebraska Agricultural Experiment Station with funding from the Hatch Act (Accession Number 1007272) through the USDA National Institute of Food and Agriculture. Open access costs were provided by the same grant. MM-F was the recipient of a graduate student fellowship from CONACYT, Mexico.

REFERENCES

Ainley, W. M., Sastry-Dent, L., Welter, M. E., Murray, M. G., Zeitler, B., Amora, R., et al. (2013). Trait stacking via targeted genome editing. *Plant Biotechnol. J.* 11, 1126–1134. doi: 10.1111/pbi.12107

Anders, C., Niewoehner, O., Duerst, A., and Jinek, M. (2014). Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* 513, 569–573. doi: 10.1038/nature13579

Andersson, M., Turesson, H., Nicolia, A., Falt, 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

Bahramnejad, B., Naji, M., Bose, R., and Jha, S. (2019). A critical review on use of Agrobacterium rhizogenes and their associated binary vectors for plant transformation. *Biotechnol. Adv.* 37, 107405. doi: 10.1016/j.biotechadv.2019.06.004

Bao, A., Burritt, D. J., Chen, H., Zhou, X., Cao, D., and Tran, L. P. (2019). The CRISPR/Cas9 system and its applications in crop genome editing. *Crit. Rev. Biotechnol.* 39, 1–16. doi: 10.1080/07388551.2018.1554621

Behera, S. S., Ray, R. C., and Zdolec, N. (2018). Lactobacillus plantarum with functional properties: an approach to increase safety and shelf-life of fermented foods. *BioMed. Res. Int.* 2018, 9361614. doi: 10.1155/2018/9361614

Beyret, E., Liao, H. K., Yamamoto, M., Hernandez-Benitez, R., Fu, Y., Erikson, G., et al. (2019). Single-dose CRISPR-Cas9 therapy extends lifespan of mice with Hutchinson-Gilford progeria syndrome. *Nat. Med.* 25, 419–422. doi: 10.1038/s41591-019-0343-4

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., et al. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326, 1509–1512. doi: 10.1126/science.1178811

Budhagatapalli, N., Rutten, T., Gurushidze, M., Kumlehn, J., and Hensel, G. (2015). Targeted Modification of Gene Function Exploiting Homology-Directed Repair of TALEN-Mediated Double-Strand Breaks in Barley. G3. (Bethesda). 5, 1857–1863.

Cai, Y., Chen, L., Liu, X., Guo, C., Sun, S., Wu, C., et al. (2018). CRISPR/Cas9-mediated targeted mutagenesis of GmFT2a delays flowering time in soya bean. *Plant Biotechnol. J.* 16, 176–185. doi: 10.1111/pbi.12758

Callaway, E. (2018). CRISPR plants now subject to tough GM laws in European Union. *Nature* 560, 16. doi: 10.1038/d41586-018-05814-6

Cathomen, T., and Joung, J. K. (2008). Zinc-finger nucleases: the next generation emerges. *Mol. Ther.* 16, 1200–1207. doi: 10.1038/mt.2008.114

Chandrasekaran, J., Brumin, M., Wolf, D., Leibman, D., Klap, C., Pearlsman, M., et al. (2016). Development of broad virus resistance in non-transgenic cucumber using CRISPR/Cas9 technology. *Mol. Plant Pathol.* 17, 1140–1153. doi: 10.1111/mpp.12375

Char, S. N., Unger-Wallace, E., Frame, B., Brigss, S. A., Main, M., Spalding, M. H., et al. (2015). Heritable site-specific mutagenesis using TALENs in maize. *Plant Biotechnol. J.* 13, 1002–1010.

Chen, M., Mao, A., Xu, M., Weng, Q., Mao, J., and Ji, J. (2019). CRISPR-Cas9 for cancer therapy: opportunities and challenges. *Cancer Lett.* 447, 48–55. doi: 10.1016/j.canlet.2019.01.017

Chew, W. L., Tabebordbar, M., Cheng, J. K., Mali, P., Wu, E. Y., Ng, A. H., et al. (2016). A multifunctional AAV-CRISPR-Cas9 and its host response. *Nat. Methods* 13, 868–874. doi: 10.1038/nmeth.3993

Christian, M., Cermak, T., Doyle, E. L., Schmidt, C., Zhang, F., Hummel, A., et al. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186, 757–761. doi: 10.1534/genetics.110.120717

Christian, M., Qi, Y., Zhang, Y., and Voytas, D. F. (2013). Targeted mutagenesis of Arabidopsis thaliana using engineered TAL effector nucleases. G3. (Bethesda). 3, 1697–1705.

Clark, R. M., Tavare, S., and and Doebley, J. (2005). Estimating a nucleotide substitution rate for maize from polymorphism at a major domestication locus. *Mol. Biol. Evol.* 22, 2304–2312. doi: 10.1093/molbev/msi228

Clasen, B. M., Stoddard, T. J., Luo, S., Demorest, Z. L., Li, J., Cedrone, F., et al. (2016). Improving cold storage and processing traits in potato through targeted gene knockout. *Plant. Biotechnol. J.* 14, 169–176.

Coloretti, F., Carri, S., Armaforte, E., Chiavari, C., Grazia, L., and Zambonelli, C. (2007). Antifungal activity of lactobacilli isolated from salami. *FEMS Microbiol. Lett.* 271, 245–250. doi: 10.1111/j.1574-6968.2007.00723.x

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

Cunningham, M. W. (2000). Pathogenesis of group A streptococcal infections. *Clin. Microbiol. Rev.* 13, 470–511. doi: 10.1128/CMR.13.3.470

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

De Vries, M. C., Vaughan, E. E., Kleerebezem, M., and De Vos, W. M. (2006). Lactobacillus plantarum—survival, functional and potential probiotic properties in the human intestinal tract. *Int. Dairy J.* 16, 1018–1028. doi: 10.1016/j.idairyj.2005.09.003

Doebley, J., Stec, A., Wendel, J., and Edwards, M. (1990). Genetic and morphological analysis of a maize-teosinte F2 population: implications for the origin of maize. *Proc. Natl. Acad. Sci. U. S. A.* 87, 9888–9892. doi: 10.1073/pnas.87.24.9888

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

Dorweiler, J., Stec, A., Kermicle, J., and Doebley, J. (1993). Teosinte glume architecture 1: a genetic locus controlling a key step in maize evolution. *Science* 262, 233–235. doi: 10.1126/science.262.5131.233

Du, H., Zeng, X., Zhao, M., Cui, X., Wang, Q., Yang, H., et al. (2016). Efficient targeted mutagenesis in soybean by TALENs and CRISPR/Cas9. *J. Biotechnol.* 217, 90–97. doi: 10.1016/j.jbiotec.2015.11.005

Duensing, N., Sprink, T., Parrott, W. A., Fedorova, M., Lema, M. A., Wolt, J. D., et al. (2018). Novel features and considerations for ERA and regulation of crops produced by Genome editing. *Front. Bioeng. Biotechnol.* 6, 79. doi: 10.3389/fbioe.2018.00079

Ebina, H., Misawa, N., Kanemura, Y., and Koyanagi, Y. (2013). Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. *Sci. Rep.* 3, 2510. doi: 10.1038/srep02510

Eckerstorfer, M. F., Engelhard, M., Heissenberger, A., Simon, S., and Teichmann, H. (2019). Plants developed by new genetic modification techniques—comparison of existing regulatory frameworks in the EU and non-EU countries. *Front. In Bioeng. Biotechnol.* 7, 26. doi: 10.3389/fbioe.2019.00026

Endo, M., Mikami, M., and Toki, S. (2016). Biallelic gene targeting in rice. *Plant Physiol.* 170, 667–677. doi: 10.1104/pp.15.01663

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

Friedrichs, S., Takasu, Y., Kearns, P., Dagallier, B., Oshima, R., Schofield, J., et al. (2019). Meeting report of the OECD conference on "genome editing: applications in agriculture—implications for health, environment and regulation". *Transgenic Res.* 28, 419–463. doi: 10.1007/s11248-019-00154-1

Gallego-Bartolome, J., Liu, W., Kuo, P. H., Feng, S., Ghoshal, B., Gardiner, J., et al. (2019). Co-targeting RNA polymerases IV and V promotes efficient de novo DNA methylation in *Arabidopsis*. *Cell.* 176, 1068–1082. doi: 10.1016/j.cell.2019.01.029

Gao, L., Cox, D. B. T., Yan, W. X., Manteiga, J. C., Schneider, M. W., Yamano, T., et al. (2017). Engineered Cpf1 variants with altered PAM specificities. *Nat. Biotechnol.* 35, 789–792. doi: 10.1038/nbt.3900

Garcia Ruiz, M. T., Knapp, A. N., and Garcia-Ruiz, H. (2018). Profile of genetically modified plants authorized in Mexico. *GM Crops Food* 9, 152–168. doi: 10.1080/21645698.2018.1507601

Garcia-Ruiz, H. (2018). Susceptibility genes to plant viruses. *Viruses* 10, 484. doi: 10.3390/v10090484

Garneau, J. E., Dupuis, M. E., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., et al. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71. doi: 10.1038/nature09523

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.

Giner, A., Pascual, L., Bourgeois, M., Gyetvai, G., Rios, P., Picó, B., et al. (2017). A mutation in the melon vacuolar protein sorting 41prevents systemic infection of Cucumber mosaic virus. *Sci. Rep.* 7, 10471. doi: 10.1038/s41598-017-10783-3

Gorbunova, V., and Levy, A. A. (1997). Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. *Nucleic Acids Res.* 25, 4650–4657. doi: 10.1093/nar/25.22.4650

Haft, D. H., Selengut, J., Mongodin, E. F., and Nelson, K. E. (2005). A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PloS Comput. Biol.* 1, e60. doi: 10.1371/journal.pcbi.0010060

Horvath, P., and Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327, 167–170. doi: 10.1126/science.1179555

Hunter, M. C., Smith, R. G., Schipanski, M. E., Atwood, L. W., and Mortensen, D. A. (2017). Agriculture in 2050: recalibrating targets for sustainable intensification. *Bioscience* 67, 385–390. doi: 10.1093/biosci/bix010

Iaffaldano, B., Zhang, Y., and Cornish, K. (2016). CRISPR/Cas9 genome editing of rubber producing dandelion Taraxacum kok-saghyz using Agrobacterium rhizogenes without selection. *Ind. Crops Prod.* 89, 356–362. doi: 10.1016/j.indcrop.2016.05.029

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *J. Bacteriol.* 169, 5429–5433. doi: 10.1128/JB.169.12.5429-5433.1987

Jia, H., Orbovic, V., Jones, J. B., and Wang, N. (2016a). Modification of the PthA4 effector binding elements in Type I CsLOB1 promoter using Cas9/sgRNA to produce transgenic Duncan grapefruit alleviating XccDeltapthA4:dCsLOB1.3 infection. *Plant Biotechnol. J.* 14, 1291–1301. doi: 10.1111/pbi.12495

Jia, Y., Ding, Y., Shi, Y., Zhang, X., Gong, Z., and Yang, S. (2016b). The cbfs triple mutants reveal the essential functions of CBFs in cold acclimation and allow the definition of CBF regulons in *Arabidopsis*. *New Phytol.* 212, 345–353. doi: 10.1111/nph.14088

Jia, H., Zhang, Y., Orbovic, V., Xu, J., White, F. F., Jones, J. B., et al. (2017). Genome editing of the disease susceptibility gene CsLOB1 in citrus confers resistance to citrus canker. *Plant Biotechnol. J.* 15, 817–823. doi: 10.1111/pbi.12677

Jiang, F., and Doudna, J. A. (2017). CRISPR-Cas9 structures and mechanisms. *Annu. Rev. Biophys.* 46, 505–529. doi: 10.1146/annurev-biophys-062215-010822

Jiang, W. Z., Henry, I. M., Lynagh, P. G., Comai, L., Cahoon, E. B., and Weeks, D. P. (2017). Significant enhancement of fatty acid composition in seeds of the allohexaploid, Camelina sativa, using CRISPR/Cas9 gene editing. *Plant Biotechnol. J.* 15, 648–657. doi: 10.1111/pbi.12663

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

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

Kang, H., Minder, P., Park, M. A., Mesquitta, W. T., Torbett, B. E., and Slukvin, II (2015). CCR5 disruption in induced Pluripotent stem cells using CRISPR/Cas9 provides selective resistance of immune cells to CCR5-tropic HIV-1 virus. *Mol. Ther. Nucleic Acids* 4, e268. doi: 10.1038/mtna.2015.42

Karginov, F. V., and Hannon, G. J. (2010). The CRISPR system: small RNA-guided defense in bacteria and archaea. *Mol. Cell* 37, 7–19. doi: 10.1016/j.molcel.2009.12.033

Kassayova, M., Bobrov, N., Strojny, L., Kiskova, T., Mikes, J., Demeckova, V., et al. (2014). Preventive effects of probiotic bacteria Lactobacillus plantarum and dietary fiber in chemically-induced mammary carcinogenesis. *Anticancer Res.* 34, 4969–4975.

Kennedy, E. M., Kornepati, A. V., Goldstein, M., Bogerd, H. P., Poling, B. C., Whisnant, A. W., et al. (2014). Inactivation of the human papillomavirus E6 or E7 gene in cervical carcinoma cells by using a bacterial CRISPR/Cas RNA-guided endonuclease. *J. Virol.* 88, 11965–11972. doi: 10.1128/JVI.01879-14

Kennedy, E. M., Bassit, L. C., Mueller, H., Kornepati, A. V. R., Bogerd, H. P., Nie, T., et al. (2015). Suppression of hepatitis B virus DNA accumulation in chronically infected cells using a bacterial CRISPR/Cas RNA-guided DNA endonuclease. *Virology* 476, 196–205. doi: 10.1016/j.virol.2014.12.001

Khemariya, P., Singh, S., Jaiswal, N., and Chaurasia, S. N. S. (2016). Isolation and identification of lactobacillus plantarum from vegetable samples. *Food Biotechnol.* 30, 49–62. doi: 10.1080/08905436.2015.1132428

Kim, Y. G., Cha, J., and Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. U. S. A* 93, 1156–1160. doi: 10.1073/pnas.93.3.1156

Kim, H., Kim, S.-T., 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

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

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

Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z., et al. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529, 490–495. doi: 10.1038/nature16526

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.

Kupferschmidt, K. (2018). EU verdict on CRISPR crops dismays scientists. *Science* 361, 435–436. doi: 10.1126/science.361.6401.435

Kushner, D. B., Lindenbach, B. D., Grdzelishvili, V. Z., Noueiry, A. O., Paul, S. M., and Ahlquist, P. (2003). Systematic, genome-wide identification of host genes affecting replication of a positive-strand RNA virus. *Proc. Natl. Acad. Sci. U. S. A.* 100, 15764–15769. doi: 10.1073/pnas.2536857100

Lander, E. S. (2016). The heroes of CRISPR. *Cell* 164, 18–28. doi: 10.1016/j.cell.2015.12.041

Leenay, R. T., and Beisel, C. L. (2017). Deciphering, communicating, and engineering the CRISPR PAM. *J. Mol. Biol.* 429, 177–191. doi: 10.1016/j.jmb.2016.11.024

Lellis, A. D., Kasschau, K. D., Whitham, S. A., and Carrington, J. C. (2002). Loss-of-susceptibility mutants of *Arabidopsis thaliana* reveal an essential role for eIF (iso)4E during potyvirus infection. *Curr. Biol.* 12, 1046–1051. doi: 10.1016/S0960-9822(02)00898-9

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, Y., Vonholdt, B. M., Reynolds, A., Boyko, A. R., Wayne, R. K., Wu, D. D., et al. (2013b). Artificial selection on brain-expressed genes during the domestication of dog. *Mol. Biol. Evol.* 30, 1867–1876. doi: 10.1093/molbev/mst088

Li, M., Li, X., Zhou, Z., Wu, P., Fang, M., Pan, X., et al. (2016). Reassessment of the four yield-related genes Gn1a, DEP1, GS3, and IPA1 in rice using a CRISPR/Cas9 System. *Front. Plant Sci.* 7, 377. doi: 10.3389/fpls.2016.00377

Li, J., Zhang, H., Si, X., Tian, Y., Chen, K., Liu, J., et al. (2017). Generation of thermosensitive male-sterile maize by targeted knockout of the ZmTMS5 gene. *J. Genet. Genomics* 44, 465–468. doi: 10.1016/j.jgg.2017.02.002

Li, A., Jia, S., Yobi, A., Ge, Z., Sato, S. J., Zhang, C., et al. (2018a). Editing of an Alpha-Kafirin gene family increases, digestibility and protein quality in Sorghum. *Plant Physiol.* 177, 1425–1438. doi: 10.1104/pp.18.00200

Li, T., Yang, X., Yu, Y., Si, X., Zhai, X., Zhang, H., et al. (2018b). Domestication of wild tomato is accelerated by genome editing. *Nat. Biotechnol.* 36, 1160–1163. doi: 10.1038/nbt.4273

Li, T., Liu, B., Spalding, M. H., Weeks, D. P., and Yang, B. (2012). High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat. Biotechnol.* 30, 390–392

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

Liang, Z., Chen, K., Li, T., Zhang, Y., Wang, Y., Zhao, Q., et al. (2017). Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. *Nat. Commun.* 8, 14261. doi: 10.1038/ncomms14261

Liang, Z., Chen, K., Zhang, Y., Liu, J., Yin, K., Qiu, J. L., et al. (2018). Genome editing of bread wheat using biolistic delivery of CRISPR/Cas9 in vitro transcripts or ribonucleoproteins. *Nat. Protoc.* 13, 413–430. doi: 10.1038/nprot.2017.145

Lillestol, R. K., Redder, P., Garrett, R. A., and Brugger, K. (2006). A putative viral defence mechanism in archaeal cells. *Archaea* 2, 59–72. doi: 10.1155/2006/542818

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

Long, C., Amoasii, L., Mireault, A. A., Mcanally, J. R., Li, H., Sanchez-Ortiz, E., et al. (2016). Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* 351, 400–403. doi: 10.1126/science.aad5725

Lor, V. S., Starker, C. G., Voytas, D. F., Weiss, D., and Olszewski, N. E. (2014). Targeted Mutagenesis of the Tomato PROCERA; Gene Using Transcription Activator-Like Effector Nucleases. *Plant Physiology* 166, 1228. doi: 10.1126/science.aad5725

Louwen, R., Staals, R. H., Endtz, H. P., Van Baarlen, P., and Van Der Oost, J. (2014). The role of CRISPR-Cas systems in virulence of pathogenic bacteria. *Microbiol. Mol. Biol. Rev.* 78, 74–88. doi: 10.1128/MMBR.00039-13

Luria, N., Smith, E., Reingold, V., Bekelman, I., Lapidot, M., Levin, I., et al. (2017). A new Israeli tobamovirus isolate infects tomato plants harboring Tm-22 resistance genes. *PloS One* 12, e0170429. doi: 10.1371/journal.pone.0170429

Maher, M. F., Nasti, R. A., Vollbrecht, M., Starker, C. G., Clark, M. D., and Voytas, D. F. (2019). Plant gene editing through de novo induction of meristems. *Nat. Biotechnol.* 38, 84–89. doi: 10.1038/s41587-019-0337-2

Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I., and Koonin, E. V. (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct* 1, 7. doi: 10.1186/1745-6150-1-7

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., Dicarlo, J. E., et al. (2013). RNA-guided human genome engineering via Cas9. *Science* 339, 823–826. doi: 10.1126/science.1232033

Malnoy, M., Viola, R., Jung, M. H., Koo, O. J., Kim, S., Kim, J. S., et al. (2016). DNA-free genetically edited grapevine and apple protoplast using CRISPR/Cas9 Ribonucleoproteins. *Front. Plant Sci.* 7, 1904. doi: 10.3389/fpls.2016.01904

Malzahn, A., Lowder, L., and Qi, Y. (2017). Plant genome editing with TALEN and CRISPR. *Cell Biosci.* 7, 21. doi: 10.1186/s13578-017-0148-4

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

Moffett, H. F., Harms, C. K., Fitzpatrick, K. S., Tooley, M. R., Boonyaratankornkit, J., and Taylor, J. J. (2019). B cells engineered to express pathogen-specific antibodies protect against infection. *Sci. Immunol.* 4, eaax0644.

Moscou, M. J., and Bogdanove, A. J. (2009). A simple cipher governs DNA recognition by TAL effectors. *Science* 326, 1501. doi: 10.1126/science.1178817

Murovec, J., Gucek, K., Bohanec, B., Avbelj, M., and Jerala, R. (2018). DNA-free genome editing of Brassica oleracea and B. rapa Protoplasts using CRISPR-Cas9 Ribonucleoprotein complexes. *Front. Plant Sci.* 9, 1594. doi: 10.3389/fpls.2018.01594

Nagpal, R., Kumar, A., Kumar, M., Behare, P. V., Jain, S., and Yadav, H. (2012). Probiotics, their health benefits and applications for developing healthier foods: a review. *FEMS Microbiol. Lett.* 334, 1–15. doi: 10.1111/j.1574-6968.2012.02593.x

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

Nakade, S., Yamamoto, T., and Sakuma, T. (2017). Cas9, Cpf1 and C2c1/2/3-What's next? *Bioengineered* 8, 265–273. doi: 10.1080/21655979.2017.1282018

Narayanan, N., Beyene, G., Chauhan, R. D., Gaitan-Solis, E., Gehan, J., Butts, P., et al. (2019). Biofortification of field-grown cassava by engineering expression of an iron transporter and ferritin. *Nat. Biotechnol.* 37, 144–151. doi: 10.1038/s41587-018-0002-1

Nature Plants Editorial. (2018). A CRISPR definition of genetic modification. *Nat. Plants* 4, 233. doi: 10.1038/s41477-018-0158-1

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/s41598-017-00578-x

Nelson, C. E., Hakim, C. H., Ousterout, D. G., Thakore, P. I., Moreb, E. A., Castellanos Rivera, R. M., et al. (2016). *In vivo* genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* 351, 403–407. doi: 10.1126/science.aad5143

Nguyen, T. D., Kang, J. H., and Lee, M. S. (2007). Characterization of Lactobacillus plantarum PH04, a potential probiotic bacterium with cholesterol-lowering effects. *Int. J. Food Microbiol.* 113, 358–361. doi: 10.1016/j.ijfoodmicro.2006.08.015

Nicolia, A., Proux-Wera, E., Ahman, I., Onkokesung, N., Andersson, M., Andreasson, E., et al. (2015). Targeted gene mutation in tetraploid potato through transient TALEN expression in protoplasts. *J. Biotechnol.* 204, 17–24

Okada, A., Arndell, T., Borisjuk, N., Sharma, N., Watson-Haigh, N. S., Tucker, E. J., et al. (2019). CRISPR/Cas9-mediated knockout of Ms1 enables the rapid generation of male-sterile hexaploid wheat lines for use in hybrid seed production. *Plant Biotechnol. J.* 17, 1905–1913. doi: 10.1111/pbi.13106

Ortigosa, A., Gimenez-Ibanez, S., Leonhardt, N., and Solano, R. (2018). 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

Osakabe, K., Osakabe, Y., and Toki, S. (2010). Site-directed mutagenesis in *Arabidopsis* using custom-designed zinc finger nucleases. *Proc. Natl. Acad. Sci. U. S. A* 107, 12034–12039. doi: 10.1073/pnas.1000234107

Panavas, T., Serviene, E., Brasher, J., and Nagy, P. D. (2005). Yeast genome-wide screen reveals dissimilar sets of host genes affecting replication of RNA viruses. *Proc. Natl. Acad. Sci. U. S. A.* 102, 7326–7331. doi: 10.1073/pnas.0502604102

Papikian, A., Liu, W., Gallego-Bartolome, J., and Jacobsen, S. E. (2019). Site-specific manipulation of *Arabidopsis* loci using CRISPR-Cas9 SunTag systems. *Nat. Commun.* 10, 729. doi: 10.1038/s41467-019-08736-7

Peng, A., Chen, S., Lei, T., Xu, L., He, Y., Wu, L., et al. (2017). Engineering canker-resistant plants through CRISPR/Cas9-targeted editing of the susceptibility gene CsLOB1 promoter in citrus. *Plant Biotechnol. J.* 15, 1509–1519. doi: 10.1111/pbi.12733

Pineda, M., Lear, A., Collins, J. P., and Kiani, S. (2019). Safe CRISPR: challenges and possible solutions. *Trends Biotechnol.* 37, 389–401. doi: 10.1016/j.tibtech.2018.09.010

Piperno, D. R., and Flannery, K. V. (2001). The earliest archaeological maize (Zea mays L.) from highland Mexico: new accelerator mass spectrometry dates and their implications. *Proc. Natl. Acad. Sci. U. S. A* 98, 2101–2103. doi: 10.1073/pnas.98.4.2101

Puchta, H., Dujon, B., and Hohn, B. (1996). Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. *Proc. Natl. Acad. Sci. U. S. A* 93, 5055–5060. doi: 10.1073/pnas.93.10.5055

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

Pyott, D. E., Sheehan, E., and Molnar, A. (2016). Engineering of CRISPR/Cas9-mediated potyvirus resistance in transgene-free *Arabidopsis* plants. *Mol. Plant Pathol.* 17, 1276–1288. doi: 10.1111/mpp.12417

Rees, H. A., and Liu, D. R. (2018). Base editing: precision chemistry on the genome and transcriptome of living cells. *Nature Reviews Genetics*. 19, 770–788.

Rosinski-Chupin, I., Sauvage, E., Fouet, A., Poyart, C., and Glaser, P. (2019). Conserved and specific features of Streptococcus pyogenes and Streptococcus agalactiae transcriptional landscapes. *BMC Genomics* 20, 236. doi: 10.1186/s12864-019-5613-5

Schwank, G., Koo, B. K., Sasselli, V., Dekkers, J. F., Heo, I., Demircan, T., et al. (2013). Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* 13, 653–658. doi: 10.1016/j.stem.2013.11.002

Settachaimongkon, S., Van Valenberg, H. J., Gazi, I., Nout, M. J., Van Hooijdonk, T. C., Zwietering, M. H., et al. (2016). Influence of Lactobacillus plantarum WCFS1 on post-acidification, metabolite formation and survival of starter bacteria in set-yoghurt. *Food Microbiol.* 59, 14–22. doi: 10.1016/j.fm.2016.04.008

Shah, T., Andleeb, T., Lateef, S., and Noor, M. A. (2018). Genome editing in plants: advancing crop transformation and overview of tools. *Plant Physiol. Biochem.* 131, 12–21. doi: 10.1016/j.plaphy.2018.05.009

Shan, Q., Wang, Y., Chen, K., Liang, Z., Li, J., Zhang, Y., et al. (2013). Rapid and efficient gene modification in rice and Brachypodium using TALENs. *Mol. Plant* 6, 1365–1368.

Shan, Q., Zhang, Y., Chen, K., Zhang, K., and Gao, C. (2015). Creation of fragrant rice by targeted knockout of the OsBADH2 gene using TALEN technology. *Plant Biotechnol. J.* 13, 791–800.

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. doi: 10.1038/nbt.3833

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

Sidira, M., Santarmaki, V., Kiourtzidis, M., Argyri, A. A., Papadopoulou, O. S., Chorianopoulos, N., et al. (2017). Evaluation of immobilized Lactobacillus plantarum 2035 on whey protein as adjunct probiotic culture in yoghurt production. *LWT* 75, 137–146. doi: 10.1016/j.lwt.2016.08.026

Smyth, S. J. (2017). Canadian regulatory perspectives on genome engineered crops. *GM Crops Food* 8, 35–43. doi: 10.1080/21645698.2016.1257468

Sorek, R., Lawrence, C. M., and Wiedenheft, B. (2013). CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu. Rev. Biochem.* 82, 237–266. doi: 10.1146/annurev-biochem-072911-172315

South, P. F., Cavanagh, A. P., Liu, H. W., and Ort, D. R. (2019). Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. *Science* 363, eaat9077. doi: 10.1126/science.aat9077

Soyk, S., Muller, N. A., Park, S. J., Schmalenbach, I., Jiang, K., Hayama, R., et al. (2017). Variation in the flowering gene SELF PRUNING 5G promotes day-neutrality and early yield in tomato. *Nat. Genet.* 49, 162–168. doi: 10.1038/ng.3733

Spencer, J. M., and Zhang, X. (2017). Deep mutational scanning of S. pyogenes Cas9 reveals important functional domains. *Sci. Rep.* 7, 16836. doi: 10.1038/s41598-017-17081-y

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

Subburaj, S., Chung, S. J., Lee, C., Ryu, S. M., Kim, D. H., Kim, J. S., et al. (2016). Site-directed mutagenesis in Petunia x hybrida protoplast system using direct delivery of purified recombinant Cas9 ribonucleoproteins. *Plant Cell Rep.* 35, 1535–1544. doi: 10.1007/s00299-016-1937-7

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

Sun, Y., Jiao, G., Liu, Z., Zhang, X., Li, J., Guo, X., et al. (2017). Generation of high-Amylose rice through CRISPR/Cas9-mediated targeted mutagenesis of starch branching enzymes. *Front. Plant Sci.* 8, 298. doi: 10.3389/fpls.2017.00298

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

Symington, L. S., and Gautier, J. (2011). Double-strand break end resection and repair pathway choice. *Annu. Rev. Genet.* 45, 247–271. doi: 10.1146/annurev-genet-110410-132435

Tabebordbar, M., Zhu, K., Cheng, J. K. W., Chew, W. L., Widrick, J. J., Yan, W. X., et al. (2016). In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* 351, 407–411. doi: 10.1126/science.aad5177

Taylor, R. A., Ryan, S. J., Lippi, C. A., Hall, D. G., Narouei-Khandan, H. A., Rohr, J. R., et al. (2019). Predicting the fundamental thermal niche of crop pests and diseases in a changing world: a case study on citrus greening. *J. Appl. Ecol.* 56, 2057–2068. doi: 10.1111/1365-2664.13455

Toda, E., Koiso, N., Takebayashi, A., Ichikawa, M., Kiba, T., Osakabe, K., et al. (2019). An efficient DNA- and selectable-marker-free genome-editing system using zygotes in rice. *Nat. Plants* 5, 363–368. doi: 10.1038/s41477-019-0386-z

Tripathi, J. N., Ntui, V. O., Ron, M., Muiruri, S. K., Britt, A., and Tripathi, L. (2019). CRISPR/Cas9 editing of endogenous banana streak virus in the B genome of Musa spp. overcomes a major challenge in banana breeding. *Commun. Biol.* 2, 46. doi: 10.1038/s42003-019-0288-7

Tuteja, N., Verma, S., Sahoo, R. K., Raveendar, S., and Reddy, I. N. (2012). Recent advances in development of marker-free transgenic plants: regulation and biosafety concern. *J. Biosci.* 37, 167–197. doi: 10.1007/s12038-012-9187-5

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

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, 402–452. doi: 10.3390/ijms20020402

Vesa, T., Pochart, P., and Marteau, P. (2000). Pharmacokinetics of Lactobacillus plantarum NCIMB 8826, Lactobacillus fermentum KLD, and Lactococcus lactis MG 1363 in the human gastrointestinal tract. *Aliment. Pharmacol. Ther.* 14, 823–828. doi: 10.1046/j.1365-2036.2000.00763.x

Waltz, E. (2016a). CRISPR-edited crops free to enter market, skip regulation. *Nat. Biotechnol.* 34, 582. doi: 10.1038/nbt0616-582

Waltz, E. (2016b). Gene-edited CRISPR mushroom escapes US regulation. *Nature* 532, 293. doi: 10.1038/nature.2016.19754

Wamaitha, M. J., Nigam, D., Maina, S., Stomeo, F., Wangai, A., Njuguna, J. N., et al. (2018). Metagenomic analysis of viruses associated with maize lethal necrosis in Kenya. *Virol. J.* 15, 90. doi: 10.1186/s12985-018-0999-2

Wang, R. L., Stec, A., Hey, J., Lukens, L., and Doebley, J. (1999). The limits of selection during maize domestication. *Nature* 398, 236–239. doi: 10.1038/18435

Wang, F., Wang, C., Liu, P., Lei, C., Hao, W., Gao, Y., et al. (2016a). Enhanced rice blast resistance by CRISPR/Cas9-targeted mutagenesis of the ERF transcription factor gene OsERF922. *PloS One* 11, e0154027. doi: 10.1371/journal.pone.0154027

Wang, G. P., Yu, X. D., Sun, Y. W., Jones, H. D., and Xia, L. Q. (2016b). Generation of marker- and/or backbone-free transgenic wheat plants via Agrobacterium-mediated transformation. *Front. Plant Sci.* 7, 1324. doi: 10.3389/fpls.2016.01324

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.7b02745

Wang, X., Tu, M., Wang, D., Liu, J., Li, Y., Li, Z., et al. (2018). CRISPR/Cas9-mediated efficient targeted mutagenesis in grape in the first generation. *Plant Biotechnol. J.* 16, 844–855. doi: 10.1111/pbi.12832

Whelan, A. I., and Lema, M. A. (2015). Regulatory framework for gene editing and other new breeding techniques (NBTs) in Argentina. *GM Crops Food* 6, 253–265. doi: 10.1080/21645698.2015.1114698

Woo, J. W., Kim, J., Kwon, S. I., Corvalán, 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

Wurtzel, E. T., Vickers, C. E., Hanson, A. D., Millar, A. H., Cooper, M., Voss-Fels, K. P., et al. (2019). Revolutionizing agriculture with synthetic biology. *Nat. Plants.* 5, 1207–1210. doi: 10.1038/s41477-019-0539-0

Wyvekens, N., Topkar, V. V., Khayter, C., Joung, J. K., and Tsai, S. Q. (2015). Dimeric CRISPR RNA-Guided FokI-dCas9 nucleases directed by truncated gRNAs for highly specific genome editing. *Hum. Gene Ther.* 26, 425–431. doi: 10.1089/hum.2015.084

Yang, Y., Wang, L., Bell, P., Mcmenamin, D., He, Z., White, J., et al. (2016). A dual AAV system enables the Cas9-mediated correction of a metabolic liver disease in newborn mice. *Nat. Biotechnol.* 34, 334–338. doi: 10.1038/nbt.3469

Yao, L., Zhang, Y., Liu, C., Liu, Y., Wang, Y., Liang, D., et al. (2018). OsMATL mutation induces haploid seed formation in indica rice. *Nat. Plants* 4, 530–533. doi: 10.1038/s41477-018-0193-y

Ye, L., Wang, J., Beyer, A. I., Teque, F., Cradick, T. J., Qi, Z., et al. (2014). Seamless modification of wild-type induced pluripotent stem cells to the natural CCR5Delta32 mutation confers resistance to HIV infection. *Proc. Natl. Acad. Sci. U. S. A* 111, 9591–9596. doi: 10.1073/pnas.1407473111

Yin, H., Xue, W., Chen, S., Bogorad, R. L., Benedetti, E., Grompe, M., et al. (2014). Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat. Biotechnol.* 32, 551–553. doi: 10.1038/nbt.2884

Yin, H., Song, C. Q., Dorkin, J. R., Zhu, L. J., Li, Y., Wu, Q., et al. (2016). Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo. *Nat. Biotechnol.* 34, 328–333. doi: 10.1038/nbt.3471

Young, J. M., Kuykendall, L. D., Martínez-Romero, E., Kerr, A., and Sawada, H. (2001). A revision of Rhizobium Frank 1889, with an emended description of the genus, and the inclusion of all species of Agrobacterium Conn 1942 and Allorhizobium undicola de Lajudie et al., 1998 as new combinations: Rhizobium radiobacter, R. rhizogenes, R. rubi, R. undicola, and R. vitis. *Int. J. Syst. Evol. Microbiol.* 51, 89–103. doi: 10.1099/00207713-51-1-89

Zago, M., Fornasari, M. E., Carminati, D., Burns, P., Suarez, V., Vinderola, G., et al. (2011). Characterization and probiotic potential of Lactobacillus plantarum strains isolated from cheeses. *Food Microbiol.* 28, 1033–1040. doi: 10.1016/j.fm.2011.02.009

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.

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, H., Zhang, J. S., Lang, Z. B., Botella, J. R., and Zhu, J. K. (2017a). Genome editing-principles and applications for functional genomics research and crop improvement. *Crit. Rev. Plant Sci.* 36, 291–309. doi: 10.1080/07352689.2017.1402989

Zhang, Y., Bai, Y., Wu, G., Zou, S., Chen, Y., Gao, C., et al. (2017b). Simultaneous modification of three homoeologs of TaEDR1 by genome editing enhances powdery mildew resistance in wheat. *Plant J.* 91, 714–724. doi: 10.1111/tpj.13599

Zhang, Y., Li, D., Zhang, D., Zhao, X., Cao, X., Dong, L., et al. (2018a). Analysis of the functions of TaGW2 homoeologs in wheat grain weight and protein content traits. *Plant J.* 94, 857–866. doi: 10.1111/tpj.13903

Zhang, Y., Massel, K., Godwin, I. D., and Gao, C. (2018b). Applications and potential of genome editing in crop improvement. *Genome Biol.* 19, 210. doi: 10.1186/s13059-018-1586-y

Zhang, A., Liu, Y., Wang, F., Li, T., Chen, Z., Kong, D., et al. (2019a). Enhanced rice salinity tolerance via CRISPR/Cas9-targeted mutagenesis of the OsRR22 gene. *Mol. Breed.* 39, 47. doi: 10.1007/s11032-019-0954-y

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

Zhou, J., Peng, Z., Long, J., Sosso, D., Liu, B., Eom, J. S., et al. (2015). Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *Plant J.* 82, 632–643. doi: 10.1111/tpj.12838

Zhou, H., He, M., Li, J., Chen, L., Huang, Z., Zheng, S., et al. (2016). Development of commercial thermo-sensitive Genic male sterile rice accelerates hybrid rice breeding using the CRISPR/Cas9-mediated TMS5 editing system. *Sci. Rep.* 6, 37395. doi: 10.1038/srep37395

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, 1211–1216. doi: 10.1038/nbt.4272

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 05 May 2020
doi: 10.3389/fpls.2020.00535

Complex Trait Loci in Maize Enabled by CRISPR-Cas9 Mediated Gene Insertion

Huirong Gao, Jasdeep Mutti, Joshua K. Young, Meizhu Yang, Megan Schroder, Brian Lenderts, Lijuan Wang, Dave Peterson, Grace St. Clair, Spencer Jones, Lanie Feigenbutz, Wally Marsh, Min Zeng, Susan Wagner, Jeffry Farrell, Kay Snopek, Chris Scelonge†, Xiaoyi Sopko, Jeffry D. Sander, Scott Betts, A. Mark Cigan† and N. Doane Chilcoat*

Research and Development, Corteva Agriscience, Johnston, IA, United States

Modern maize hybrids often contain biotech and native traits. To-date all biotech traits have been randomly inserted in the genome. Consequently, developing hybrids with multiple traits is expensive, time-consuming, and complex. Here we report using CRISPR-Cas9 to generate a complex trait locus (CTL) to facilitate trait stacking. A CTL consists of multiple preselected sites positioned within a small well-characterized chromosomal region where trait genes are inserted. We generated individual lines, each carrying a site-specific insertion landing pad (SSILP) that was targeted to a preselected site and capable of efficiently receiving a transgene via recombinase-mediated cassette exchange. The selected sites supported consistent transgene expression and the SSILP insertion had no effect on grain yield. We demonstrated that two traits residing at different sites within a CTL can be combined via genetic recombination. CTL technology is a major step forward in the development of multi-trait maize hybrids.

Keywords: maize, CRISPR-Cas9, complex trait loci, trait stack, gene target, gene expression, recombinase-mediated cassette exchange, genetic crossing

OPEN ACCESS

Edited by:
Hiroshi Ezura,
University of Tsukuba, Japan

Reviewed by:
Yiping Qi,
University of Maryland, College Park, United States
Anna-Maria Botha-Oberholster,
Stellenbosch University, South Africa

****Correspondence:***
Huirong Gao
huirong.gao@corteva.com

†Present address:
Chris Scelonge,
KWS Gateway Research Center, LLC., St. Louis, MO, United States
A. Mark Cigan,
Genus PLC, DeForest, WI, United States

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *17 January 2020*
Accepted: *08 April 2020*
Published: *05 May 2020*

Citation:
Gao H, Mutti J, Young JK, Yang M, Schroder M, Lenderts B, Wang L, Peterson D, St. Clair G, Jones S, Feigenbutz L, Marsh W, Zeng M, Wagner S, Farrell J, Snopek K, Scelonge C, Sopko X, Sander JD, Betts S, Cigan AM and Chilcoat ND (2020) Complex Trait Loci in Maize Enabled by CRISPR-Cas9 Mediated Gene Insertion. Front. Plant Sci. 11:535. doi: 10.3389/fpls.2020.00535

INTRODUCTION

In the early years of biotech crop cultivation, a single transgene was used to confer resistance to insects or tolerance to herbicides. This benefited growers by reducing yield losses from pests and weeds, while at the same time reducing soil erosion and the use of chemical pesticides (Koziel et al., 1993; Padgette et al., 1995). Over the last ∼30 years some insects have developed resistance to insecticidal proteins used in first generation products and herbicide-resistant weeds have become problematic. The need for new traits that are durable and broad-spectrum has been partially met by the introduction of products that employ multiple transgenes (Que et al., 2010). Additionally, the objectives of transgenic traits have expanded, with research in many areas including disease resistance, drought tolerance, nitrogen use efficiency, and grain quality. Traditionally, multiple traits have been brought together using genetic backcrossing, referred to as trait introgression (TI) (Peng et al., 2014). However, introgression of more than four traits which reside at different part of genome in an inbred line is impractical in developing commercial products (Mumm and Walters, 2001; Petolino and Kumar, 2016; Chen and Ow, 2017). Moreover, backcrossing introgression brings along unintended genome sequences adjacent to the transgene which can lower yield. To overcome these challenges, improved methods for trait assembly are needed.

Several approaches have been taken to improve trait assembly. TI has been enhanced using molecular markers which facilitate selection and reduce the number of backcrosses needed (Peng et al., 2014). Transgenic constructs containing multiple gene expression cassettes have been used, but this approach is inefficient (Cao et al., 2002; Dafny-Yelin and Tzfira, 2007; Que et al., 2010). Sequential transformation methods have been shown to enable stacking traits by inserting a transgene immediately adjacent to an existing biotech trait using recombinase-mediated cassette exchange (RMCE) or zinc finger/homing endonucleases (Ow, 2011; Ainley et al., 2013; D'Halluin et al., 2013; Kumar et al., 2015; Petolino and Kumar, 2016; Srivastava and Thomson, 2016). Although modular trait stacking resolved some of the issues associated with the direct transformation of a large plasmid, it has not been adopted for product development in part due to concern that the newly added transgene may alter expression of the original transgene because of spatial proximity and partly because the process has low efficiency. Both of those molecular stacking approaches are inflexible; should one of the trait genes not be required in a geography or have lost efficacy, it can't be easily separated or replaced by conventional breeding.

To facilitate development of multi-trait products in maize, we have developed a complex trait locus (CTL) approach. A CTL is comprised of multiple gene-targeting sites carefully positioned within a small well-characterized region in the genome. To create a CTL, individual transgenic lines are created that have a site-specific insertion landing pad (SSILP) targeted to a preselected site using the CRISPR-Cas9 system. Each of the SSILP lines are capable of receiving a transgene via high-efficiency RMCE. Genetic crossing then is used to link traits via meiotic recombination. The tightly linked trait genes can then be introgressed into other inbreds as a single locus. The relatively short, but adequate genetic distance between these transgenes also allows removal of a transgene if needed. This CTL approach was not possible previously because targeted gene insertion in crop plants was inefficient before the advent of CRISPR-Cas9 technology. Our results show that the CRISPR-Cas9 system enables robust gene targeting via homology directed repair (HDR) and can be used to establish trait gene landing sites in maize elite inbred lines. We found that transgene expression was consistent across the preselected sites and that inserted transgenes had minimal effects on neighboring endogenous gene expression. As expected, these sites can be genetically linked through traditional crossing. CTLs enable efficient and flexible production of maize hybrids with multiple transgenic traits. Insertion of SSILP at these sites had no impact on yield.

RESULTS

Selection of Chromosomal Location for Complex Trait Loci

We used four criteria to decide where to locate CTLs in the maize genome: (1) regions with conserved haplotype within non-stiff stalk (NSS) and stiff stalk (SS) germplasm pools; (2) regions with low gene density that are not used in forward breeding; (3) regions with high recombination frequency to minimize the donor sequences around the CTL while introgressing the region; (4) regions harboring existing commercially valuable traits. Genomic sequences of >1,000 elite lines were scanned using a 10-cM window to identify regions of 4–5 cM to serve as CTLs. We selected four chromosomal regions to generate CTLs in maize which we refer to as CTL1, 2, 3, and 4, respectively (**Figure 1**). In the four CTLs, the DNA sequence was conserved among 84, 44, 34, and 55% of the SS inbreds and among 56, 72, 73, and 84% of the NSS inbreds, respectively. The regions had a gene density of 20, 21, 4 and 8 genes per cM and an average ratio of physical-to-genetic distance of 0.4, 0.6, 0.2, and 0.2 Mb/cM, respectively, as estimated based on the maize B73 reference genome sequence v2. A stacked insect-resistant and herbicide-tolerant maize event DP-004114-3 (Diehn et al., 2013) resides at 53.5 cM on chromosome 1, which is within CTL1. CTL3 is in a telomeric region of Chr 3, which can facilitate TI with a single crossover.

Selection of CRISPR-Cas9 Target Sites Within CTLs

CRISPR-Cas9 was used to introduce trait genes into preselected sites within each CTL. To minimize regulatory concerns and product development costs, CRISPR-Cas9 target sites (CTS) were selected based on the following criteria: (1) the target site is at least 2 kb away from any known gene; (2) the CTS DNA sequence is unique in the genome and conserved among the targeted inbred lines; (3) the genomic sequences of 200–500 bp flanking the CTS are unique in the genome; and (4) spacing of the CTSs within a CTL would accommodate genetic crossing to recombine traits. A total of 30, 21, 13, and 12 CTSs were selected for CTL1, CTL2, CTL3, and CTL4, respectively (**Figure 1**, **Table 1**, and **Supplementary Table S1**). These sites spanned 4.18 cM (2.5 million base pairs, Mbp) at CTL1, 4.28 cM (3.2 Mbp) at CTL2, 2.35 cM (0.6 Mbp) at CTL3, and 3.04 cM (0.7 Mbp) at CTL4. Most of the target sites were 0.1–3 cM apart, suitable for both genetic stacking and subsequent segregation as a single locus.

Targeting Site-Specific Insertion Landing Pad to Preselected Sites Using CRISPR-Cas9

Although trait genes can be inserted directly at the selected CTS via HDR, efficiencies for CRISPR-Cas9-enabled gene insertion are low (Svitashev et al., 2015; Endo et al., 2016; Shi et al., 2017; Danilo et al., 2018; Hummel et al., 2018). Because many constructs need to be screened and tested in the early stages of product development, it is not practical to directly insert trait gene cassettes on a large scale using HDR at present. Therefore, we adopted a two-step strategy for trait gene insertion (**Figure 2**). First, a SSILP about 3 kb in length was inserted into target sites using CRISPR-Cas9. Trait gene cassettes were then integrated into the SSILP via RMCE in a second transformation of the characterized SSILP transgenic plants (Li et al., 2009; Anand et al., 2019). RMCE utilizes the flippase (FLP) recombinase and FLP recognition targets (FRT) to insert a gene in the target site, and has been shown to be an efficient technique for site-specific

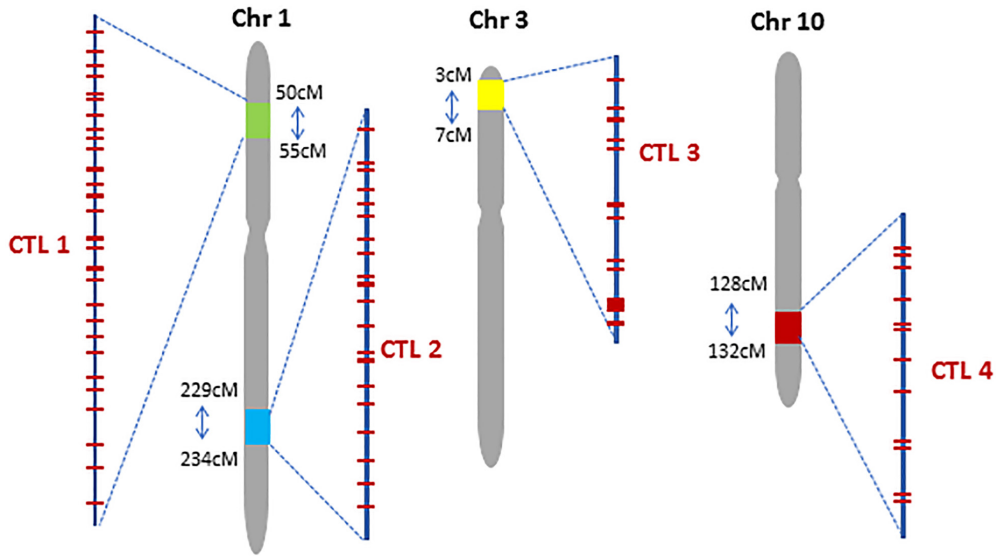


FIGURE 1 | Chromosomal location of four Complex Trait Loci (CTL) in the maize genome. Red bars within each CTL represent preselected CRISPR targeting sites.

transgene insertion (Li et al., 2009; Ow, 2011; Srivastava and Thomson, 2016; Anand et al., 2019).

To insert SSILPs at preselected sites, immature embryos were co-bombarded with four DNA plasmids (**Supplementary Figure S1**) containing: repair template, *Streptococcus pyogenes* Cas9, guide RNA, and the maize morphogenic genes *Baby boom* (*Bbm*) and *WUSCHEL2* (*Wus2*) (Svitashev et al., 2015; Lowe et al., 2016). The DNA repair template consisted of the SSILP flanked by two DNA sequences of approximately 400-bp homologous to the genomic sequences immediately adjacent to the CTS (**Figure 3A**). Two unique sequences, PSA and PSB, flanking the SSILP also were included to facilitate high-throughput PCR screening (**Figure 3A**). The same SSILP sequence was used for all target sites, but the homologous arms varied to match the genomic sequences bordering each CTS. Approximately, 1,000 immature embryos per CTS were used for genotypes PH184C and HC69 while 500 immature embryos were used for PHH5G because it has a higher transformation frequency.

We used junction PCR assays to detect SSILP insertion in T0 plants regenerated from embryogenic calli. In this assay, PCR amplification of the target region was coupled with nested quantitative PCR (qPCR) to detect SSILPs (**Figures 3A,B**). Plants positive for both 5′ and 3′ junctions (hereafter referred to as 2 × HDR events) were further analyzed with overlapping long PCR (**Figures 3A,C**). In PH184C, events with SSILP insertion were identified for 28 out of 30 sites at CTL1, 19 out of 20 sites at CTL2, 13 out of 13 sites at CTL3, and 10 out of 12 sites at CTL4 (**Table 1** and **Supplementary Table S1**). Among the 28 sites at CTL1, the insertion frequency varied from 0.3 to 7.1% (**Table 1**). The site TS13 at CTL4 had the highest insertion frequency with 18% of the events positive for both junctions (**Supplementary Table S1**). A selected subset of CTSs in CTL1 also were targeted for SSILP insertion in the inbreds PHH5G and HC69. Positive events were identified for all sites and the insertion frequency was generally similar to that seen in PH184C. While most of the 2 × HDR T0 plants had mono-allelic insertion at the CTS, we found three events with bi-allelic insertion of SSILP at the CTS.

We identified some 2 × HDR T0 plants that were free of the helper genes (Cas9, gRNA, *Bbm*, and *Wus2*), indicating that transient helper gene expression during transformation can be sufficient to enable homologous gene targeting. However, most of the 2 × HDR plants contained one or more copies of the helper genes. To remove any helper DNA sequences and repair template that might have randomly inserted into the genome, the 2 × HDR T0 plants were crossed to recurrent parent (RP, the wild type of the same inbred line as initially transformed) to produce T1 seeds, and T1 plants were crossed to RP again to generate T2 seeds. PCR assays were used to detect the presence of helper genes. Out of the 89 CTSs with 2 × HDR events, we obtained helper gene-free T1 plants for 67 sites (**Table 2**). The integrity of inserted SSILPs at each CTS and the absence of the helper genes and other plasmid DNA fragments in the genome were further verified using Southern-by-Sequencing (SbS) analysis (Zastrow-Hayes et al., 2015); perfect SSILPs at 57 CTS out of the 67 sites were confirmed by SbS (**Table 2**).

Transgene Expression at CTL Sites

For useful transgenic trait development, genomic sites must be able to support transgene expression. To assess the effect of insertion site on transgene expression, protein expression levels of the Neomycin phosphotransferase II (*NPTII*) gene in SSILPs were measured. In the PH184C lines, the NPTII protein in leaves averaged 38 ppm with relatively low variation among sites, either within or across CTLs (**Figure 4**). Some, or perhaps all the observed variation is not due to insertion site effect since similar expression variation was observed among independent events at a single site, such as TS10-e1 and TS10-e2 at 54.56 cM in PHH5G. There are a small number of sites where gene expression

TABLE 1 | CRISPR-Cas9 mediated insertion of SSILP in preselected sites at CTL1 in PH184C.

CRISPR target site	Genetic location (cM)	Number of shoot regenerated	Number of shoot with target site modified	Target site modification frequency	Number of shoot positive HDR1	Number of shoot positive HDR2	Number of shoot positive 2 × HDR	2 × HDR frequency
TS49	50.87	214	198	93%	2	4	3	1.4%
TS50	50.95	263	218	83%	4	8	8	3%
TS51	51.06	300	280	93%	4	5	11	3.7%
TS41	51.27	356	220	62%	7	7	11	3.1%
TS71^	51.32	979	871	89%	25	13	59	6%
TS72	51.33	309	287	93%	7	4	7	2.3%
TS81	51.45	220	170	77%	3	6	6	2.7%
TS73	51.48	252	191	76%	1	4	7	2.8%
TS14	51.54	293	277	95%	4	3	2	0.7%
TS74	51.61	161	129	80%	1	4	1	0.6%
TS75*^	51.68	899	716	80%	4	6	15	2.6%
TS84*	51.68	366	273	75%	6	7	6	1.6%
TS76	51.69	264	198	75%	3	5	18	6.8%
TS77^	51.72	666	502	75%	8	7	14	2.1%
TS78	51.75	329	188	57%	1	1	7	2.1%
TS19	51.95	217	17	8%	0	1	0	0%
TS85	51.95	217	168	77%	1	3	1	0.5%
TS86	52.54	216	183	85%	1	1	6	3.7%
TS8	52.56	217	205	95%	2	2	9	4.1%
TS43	52.8	179	140	78%	1	5	3	1.7%
TS11	53.15	177	174	98%	2	6	7	4%
TS47	53.21	200	171	86%	1	2	7	3.5%
TS80	53.23	336	330	98%	3	1	4	1.2%
TS52	53.25	222	151	68%	0	2	0	0%
TS87	53.57	302	298	99%	6	2	12	4%
TS88	53.59	370	193	52%	3	2	1	0.3%
TS45^	53.66	616	562	91%	16	10	36	5.8%
TS44	54.16	246	221	90%	0	1	9	3.7%
TS46	54.43	244	220	90%	7	0	5	2%
TS10	54.56	224	209	93%	4	4	4	1.8%

*Mutation of target sites (TS) in the regenerated shoots was detected using a quantitative PCR (qPCR) assay. Insertion events with both the right and left junctions positive (2 × HDR) were identified using nested junction PCR assays. *Two target sites at the same approximate genetic location, physically 6 kb apart from each other. ^Data from two rounds of transformation.*

was different from the norm. For example, the SSILP at TS34 at CTL2 in two different genetic backgrounds expressed NPTII approximately 33% higher than any of the other 15 sites tested (**Figure 4**). Genetic background was found to have a greater influence on the NPTII expression levels than genomic location. For the sites at CTL1, the average NPTII protein content was 37, 69 and 83 ppm in PH184C, HC69, and PHH5G, respectively. Similarly, the CTL2 sites in HC69 had a higher NPTII expression than that in PH184C (**Figure 4**). Since all sites tested support high level transgene expression and the position effect is smaller than the genotype background effect, we believe that these SSILPs are suitable for product development.

Effects of Site-Specific Insertion Landing Pad Insertion on Expression of Neighboring Endogenous Genes

One concern related to the random- or targeted-insertion of a transgene in the plant genome is that the insert may affect the expression of neighboring endogenous genes. As an attempt to minimize interactions of SSILPs with nearby genes, we selected target sites that were at least 2 kb away from endogenous genes. To test if this distance is adequate, RNA sequencing was used to analyze nine PH184C lines, including seven lines with SSILPs inserted at CTL1 and two lines with insertions at CTL2. Because the two CTLs are 175 cM apart, the CTL2 SSILP lines were used as comparator to determine the effect of SSILP insertion on nearby genes within the CTL1 region. In the vicinity of the CTL1 target sites from 49.45 to 55.47 cM, there were 83 endogenous maize genes, of which 69 expressed in the leaf tissues of PH184C (**Supplementary Table S2**). None of these genes showed significant differential expression [absolute value of log2 (fold change) <1; false discovery-corrected $P > 0.05$] in pairwise comparisons between a CTL1 insertion line and the control, except for one gene, Zm00001d027859. The transcript level of this gene was reduced 67% in the line TS10 [log2 (fold change) $= -1.61$ and false discovery-corrected $P = 0.049$; **Figure 5**]. However, this gene was not differentially expressed

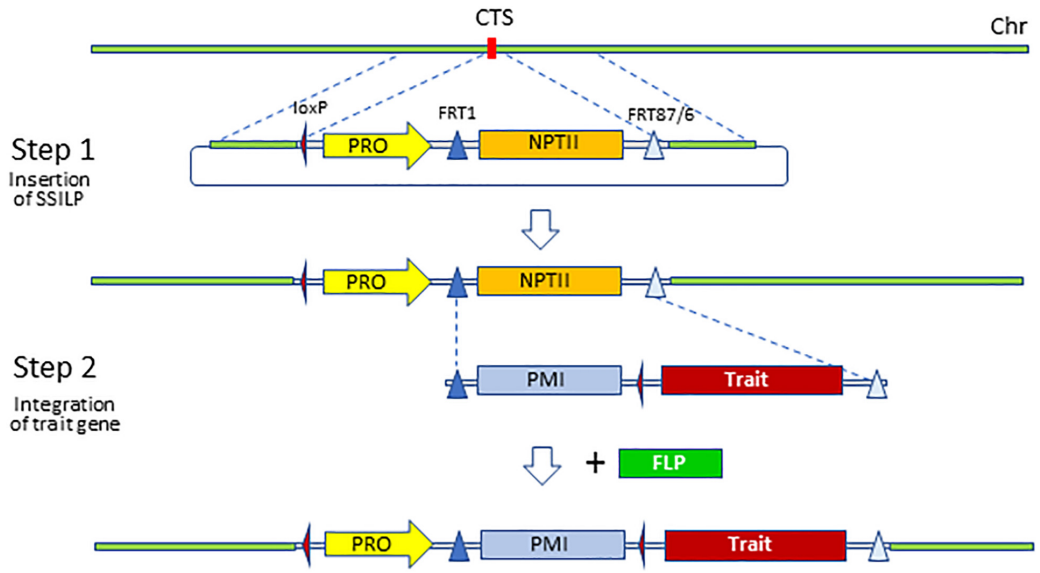


FIGURE 2 | Two-step strategy to integrate trait gene to preselected sites. Site-specific insertion landing pad (SSILP) was inserted to a CRISPR-Cas9 target site (CTS) via homology-directed repair. Lines containing the SSILP can be retransformed with a trait gene which is integrated to the SSILP via RMCE. The PRO in SSILP serves as promoter trap for the selection marker in the trait gene vector. The selectable marker phosphomannose isomerase (PMI) can be removed via CRE-mediated recombination if desired. FLP, flippase recombinase; FRT, flippase recognition target; loxP, side triangle, locus of X-over P1 site; PRO, promoter; NTPII, neomycin phosphotransferase II; Chr, chromosome.

in the other six CTL1 insertion lines relative to the CTL2 comparator. Zm00001d027859 is located at the genetic position 54.58 cM, approximately 5 kb downstream of the SSILP insertion site at 54.56 cM in the line TS10. This spatial proximity is likely responsible for the observed reduction in expression. Overall, these results indicate that the SSILPs generated in this study have little impact on the expression of neighboring endogenous genes.

Integration of Trait Genes to Site-Specific Insertion Landing Pads

Next, we evaluated whether the SSILP lines generated in this study were competent for RMCE. Eleven CTL1 SSILP lines in PH184C were used for trait gene insertion via FLP/FRT-mediated RMCE (**Table 3**, **Figure 2**). A total of 326 donor constructs were tested in seven SSILP lines using particle bombardment. Putative T0 RMCE plants were generated for most of the constructs at the seven sites with an average of 4% recovery (T0 plants/embryos used). Among T0 plants, ∼45% were quality events that had a single copy of the trait gene integrated into the SSILP (**Table 3**). We also tested RMCE using *Agrobacterium*-mediated transformation; the putative T0 RMCE recovery rate ranged from 5.9 to 9.3% in four SSILP lines tested (**Table 3**). These results indicate that all SSILP lines tested are capable of high-efficiency RMCE.

Trait Stacking via Genetic Recombination

To determine if trait genes inserted at SSILPs in two different lines can be linked on the same chromosome through genetic recombination, PH184C inbred lines containing a SSILP at various sites in CTL1 were tested. Previously, we had generated a transgenic PH184C line named M14, which carries the phosphomannose isomerase (*PMI*) gene at 51.54 cM on Chr 1 within CTL1. This line was crossed with homozygous plants from 15 SSILPs (**Table 4**). The F1 progeny from each cross were backcrossed to WT inbred PH184C. BC1 seeds were assayed using qPCR to determine the presence or absence of the *PMI* gene and the *NPTII* gene at the SSILP. Most progeny will have either *PMI* or *NPTII*. However, a small number of progeny are expected to have chromosomal crossover between the two genes, and will have both *PMI* and *NPTII* or neither of them. The genetic distance between the two genes was calculated based on the recombination frequency (i.e., 1% recombinants is equal to 1 cM). For 14 out of 15 crosses, the observed genetic distance between M14 (*PMI*) and SSILP (*NPTII*) closely matched the predicted genetic distance based on the B73 reference genome (**Table 4**). For the cross of M14 with SSILP-TS72, no recombinants were recovered from 1,400 BC1 seeds analyzed. The SSILP-TS72 is located 0.17 cM from the PMI insertion site, the closest SSILP among the 15 sites. A larger BC1 population may need to be screened to identify recombinants for this SSILP. These results demonstrate that the genes integrated into SSILPs in CTL1 can be stacked by genetic crossing.

Presence of Site-Specific Insertion Landing Pad Has No Impact on Grain Yield

To be useful for transgenic product development, CTL insertion sites must be agronomically neutral. To evaluate the impact of SSILP insertion at CRISPR target sites on plant productivity, we

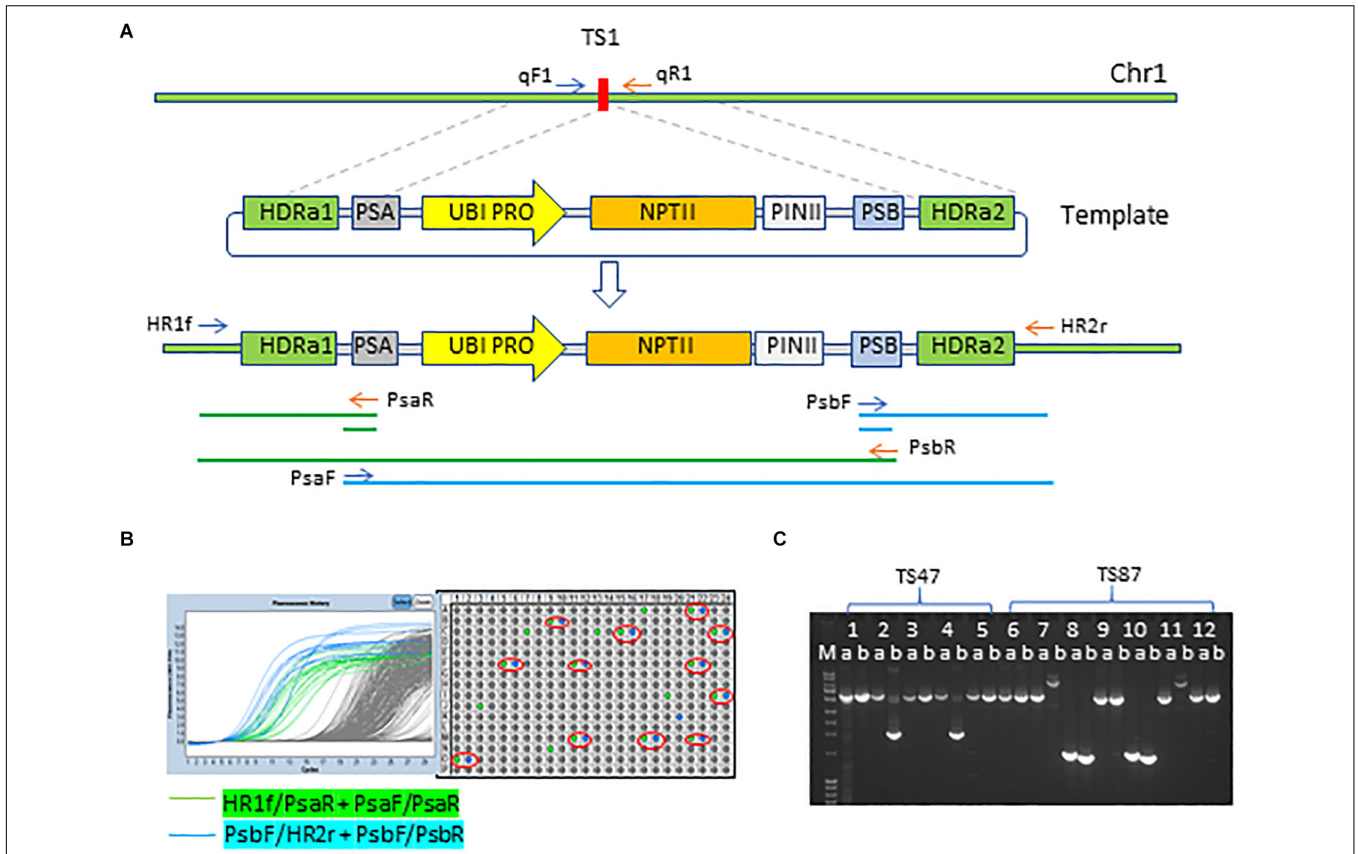


FIGURE 3 | Insertion of SSI landing pad to CRISPR-Cas9 target site. **(A)** Insertion of SSI landing pad (SSILP) into CRISPR-Cas9 target site 1 (TS1) on chromosome 1 (Chr 1). UBI PRO, maize UBIQUITIN1 promoter; NPTII, neomycin phosphotransferase II; PINII, potato proteinase inhibitor II terminator; HDRa, homology-directed repair arm; PSA and PSB, unique sequences to facilitate high-throughput screening for insertion events. HR1f, HR2r, PsaR, PsaF, PsbF, and PsbR are PCR primers and arrows indicate the direction of primers. **(B)** Screening of T0 plants for insertion events with PCR and qPCR. The plants positive for both junctions are indicated in red circles. **(C)** Image of agarose gel showing overlapping long PCR products for two SSILP insertion sites (TS47 and TS87). Lane a, PCR products amplified with primers HR1f and PsbR; Lane b, PCR products amplified with primers PsaF and HR2r; Lane M, molecular-weight size markers. PCR products of the sample 1, 3, 5, 6, 9, and 12 have expected size.

TABLE 2 | Insertion of SSILP in preselected target sites at four complex trait loci (CTL) and recovery of clean plants free of genome-editing helper genes.

CTL	Genotype	Sites targeted	Sites 2 × HDR	Sites 2 × HDR T1 seed	Sites 2 × HDR and null helpers (qPCR)	Sites 2 × HDR perfect and null helpers (SbS)
1	PH184C	36	34	24	19	15
1	HC69	6	6	6	6	3
1	PHH5G	24	24	18	17	17
2	PH184C	20	20	15	11	10
2	HC69	11	10	9	6	5
3	PH184C	13	13	9	4	3
4	PH184C	12	10	8	4	4

Thirty unique sites within CTL1 in the inbred line PH184C were targeted using donor vectors containing FRT1/87. Among those 30 sites, six also were targeted using the donor vector carrying FRT1/6. A subset of 6 and 24 out of the 30 sites were targeted in HC69 and PHH5G, respectively. For CTL2, SSILP was inserted into 20 unique sites in PH184C. Only 11 out of the 21 sites were targeted in HC69. 2 × HDR, both junctions PCR positive; SbS, Southern-by-Sequencing.

ran a multi-location hybrid field trial in the United States. In this trial, lines were evaluated for grain yield and grain moisture. Plant and ear height, and growing degree units required for pollen shed and silking were also measured. SSILP lines at 14 CRISPR target sites in the PH184C background were used as pollen donors (BC1F3 generation) to cross with three elite tester lines to make hybrids hemizygous for the CTL transgenes. As controls for each SSILP line, a null segregant at the BC1F2 generation was used to produce BC1F3 plants as pollen donors for hybrid seed production. Three hybrids for the 14 SSILP lines and corresponding nulls were planted at 12 locations. The yields were similar among the 14 SSILP lines and no difference was

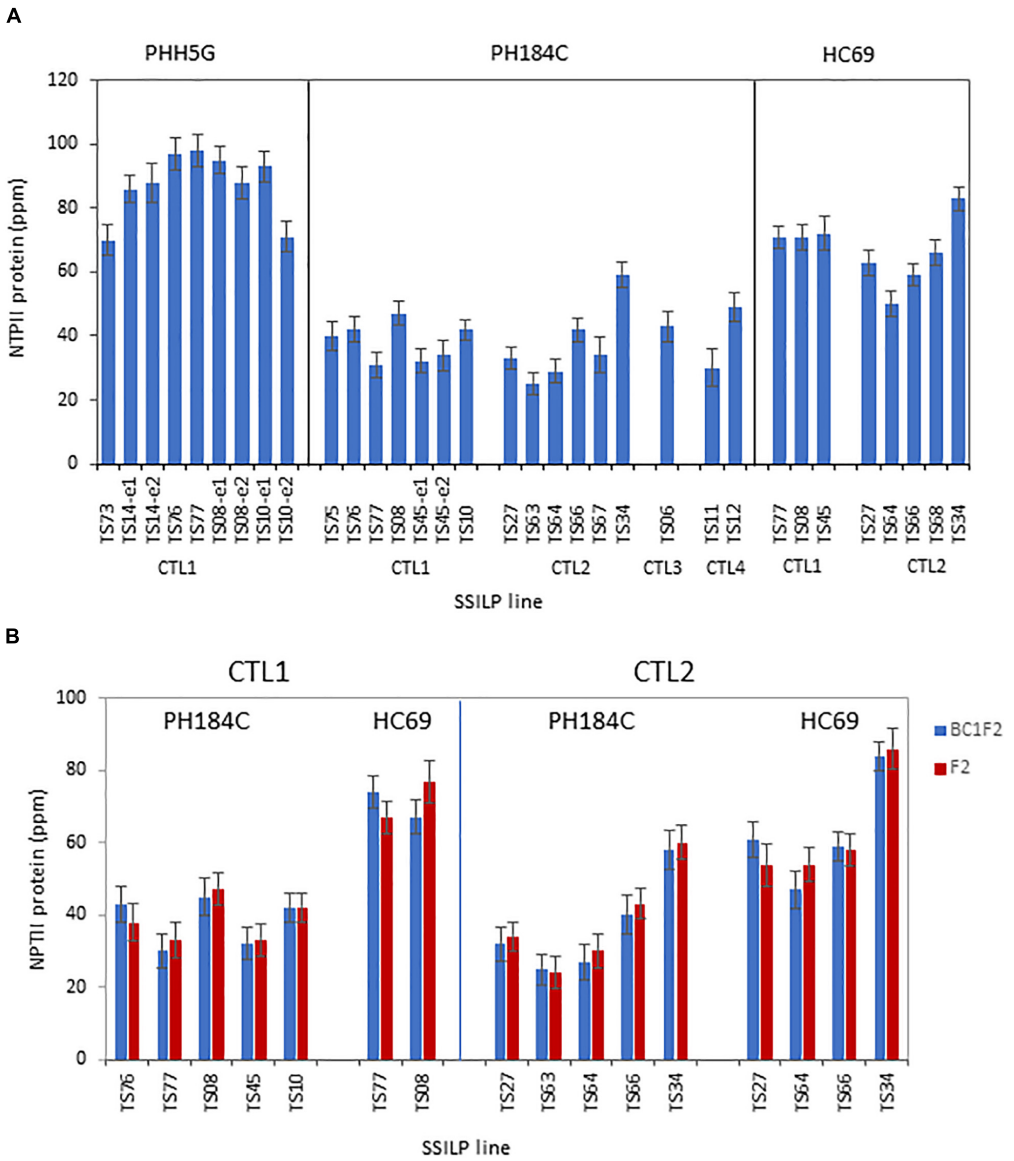


FIGURE 4 | Protein expression of the *NPTII* gene targeted to preselected sites at CTLs. **(A)** NPTII protein content in leaves of 3-week-old plants was measured using ELISA. The plants contain one copy of SSILP and were selected based on PCR genotyping results. One to two independent insertion events per target site were analyzed. Error bars, SE; *N* = 8–32. A total of 33 lines were measured in the study. **(B)** NPTII protein content in leaves of 3-week-old BC1F2 and F2 PH184C or HC69 plants was quantified using ELISA. The plants contain one copy of SSILP and were selected based on PCR genotyping results. One event per target site was analyzed. Error bars, SE; *N* = 7–17. A total of 16 lines were measured in the study.

observed between the hybrid with SSILP and it's corresponding null segregants (**Figure 6**). The trials demonstrated no yield impact attributable to SSILP insertion in all 14 lines. Like yield, other non-yield traits measured did not show a significant difference between the SSILP hybrids and nulls.

DISCUSSION

To construct complex trait loci, we have used CRISPR-Cas9 to generate many independent maize target lines, each containing a single SSILP in one of four preselected genomic regions. DNA sequence analyses confirmed precise insertion of SSILPs via HDR. The *NPTII* gene in the SSILP functioned properly as evidenced by resistance to G418, and we observed minimal variation in its expression level among different insertion sites within a genotype but significant difference between genotypes. Expression of endogenous genes neighboring SSILP insertions was largely unaffected. Trait genes were integrated into the SSILP with high efficiency via RMCE. By crossing the SSILP lines with a pre-established insertion event within the CTL1, we demonstrated that SSILPs could be linked through genetic recombination.

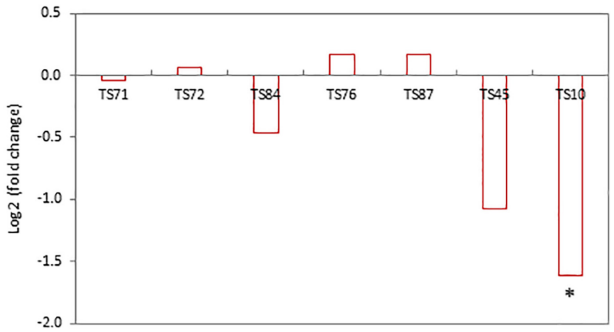


FIGURE 5 | Expression of Zm0001d027859 in PH184C lines with SSILP inserted in the CTL1 region compared to that in the lines with insertion in CTL2. RNA-sequencing was performed to determine gene expression in the leaf tissues of 2-week-old seedlings. The transcript levels of the endogenous genes in the vicinity of CTL1 target sites in seven lines were compared to the CTL2 control. Contrasts are presented as log2 (fold change). Asterisk indicates false discovery-corrected $P < 0.05$.

Complex trait loci have several desirable characteristics. First, they enable high-quality assembly of products with multiple traits by effectively creating a single genetic locus for TI. Second, given the ease of crossing and screening recombinants, traits can be added and removed as needed. Third, because trait genes are inserted in well characterized locations, many of the costs associated with testing novel construct-site combinations are eliminated.

This CTL approach is different from previously reported transgene stacking using RMCE (Ow, 2011; Nandy et al., 2015; Srivastava and Thomson, 2016; Chen and Ow, 2017). CRISPR-Cas-enabled gene targeting allows precise positioning of trait genes within a small, preselected region on a chromosome. In contrast, recombinase-mediated gene insertion is molecular stacking, and relies on a randomly inserted site in the genome. The resulting molecular linkage between two transgenes can't be broken easily in genetic crossing and the closeness of the stacked transgenes raises concerns about expression interactions. Although SSILP/recombinase were used to insert trait genes in a two-step process, they are not essential components for a CTL. A trait gene could be directly inserted into a preselected site within a CTL using CRISPR-Cas9. We used a two-step process because it enables more efficient generation of transgenic lines. By using a pre-established SSILP at preselected sites, many constructs can be easily evaluated with a very small number of events since those insertion sites are well-characterized.

A few insights into plant gene targeting via HDR and gene expression were generated as part of our CTL construction and characterization. This is the largest example to date using CRISPR-Cas9 to promote homology-directed insertion in maize. We recovered targeted insertion events in 93% of tested sites (69 out of 74). As expected, no insertion was obtained without efficient DNA cleavage, as shown in TS19 at CTL1 (**Table 1**), TS77 at CTL2 and TS15 at CTL4 (**Supplementary Table S1**). However, a high mutation frequency did not always result in a high frequency of insertion, for example TS52 at CTL1, TS62-HC69 at CTL2, and TS5 at CTL4 (**Table 1** and **Supplementary Table S1**). Other studies have concluded that several factors, including chromatin structure, DNA sequence of target sites and homology arms can influence the efficiency of CRISPR-Cas9 and HDR (Kuscu et al., 2014; Wu et al., 2014; Liu et al., 2016). We did observe variation across different sites, but overall the CRISPR-Cas system is robust, and we could obtain HDR at most sites. Although CRISPR-Cas9 from *S. pyogenes* (SpyCas9) was used successfully to insert a SSILP at numerous preselected sites in this study, other CRISPR systems like Cas12a (Cpf1), Cas9 orthologs from *Streptococcus thermophilus* (SthCas9) and *Staphylococcus aureus* (SaCas9) could also be

TABLE 3 | Integration of trait gene to SSILP at CTL1 in PH184C.

SSILP location	FRT	Delivery method	Number of constructs	Transformation frequency (percent)	RMCE* ratio (percent)	Quality event rate (percent)
TS50	1/87	PB	326	3.9	38	1.5
TS71	1/87	PB	326	3.8	50	1.4
TS84	1/87	PB	326	3.2	42	1.1
TS76	1/87	PB	326	4.4	53	1.7
TS8	1/87	PB	326	4.3	48	1.8
TS43	1/87	PB	326	4.5	43	1.6
TS45	1/87	PB	326	3.7	45	1.5
TS77	1/6	Agro	74	7.8	58	4.6
TS8	1/6	Agro	74	9.3	52	4.9
TS45	1/6	Agro	74	5.9	55	3.2
TS10	1/6	Agro	74	8.4	53	4.4

*Data are averages of 326 PB constructs or 74 Agro constructs. The frequency of transformation is the number of regenerated T0 plants for each construct divided by the number of embryos used in transformation. The RMCE ratio is the number of RMCE events divided by the number of T0 plants analyzed with PCR. The quality event rate is the number of RMCE events divided by the number of embryos used in transformation. SSILP, site-specific insertion landing pad; FRT, flippase (FLP)/FLP-recombinase targets; PB, particle bombardment; Agro, Agrobacterium-mediated transformation. *RMCE events are characterized by (1) presence of single intact copy of the donor genes (PMI, PAT, trait gene); (2) absence of the marker gene NPTII; (3) presence of FRT1 and FRT87 or FRT6 junctions; and (4) absence of unintended DNA sequence insertion including that derived from vector backbone, Bbm, Wus2, and FLP gene.*

TABLE 4 | Genetic stacking of *NPTII* in SSILPs with *PMI* at the chromosomal location 51.54 cM within CTL1.

Target site	Predicated genetic position (cM)	Predicated distance between M14 and SSILP (cM)	Number of BC1 seeds	Number of recombinants	Observed genetic distance (CM)
TS50	50.95	0.55	2688	15	0.56
TS41	51.27	0.23	2016	5	0.25
TS71	51.32	0.18	2079	3	0.14
TS72	51.33	0.17	1481	0	0.00
TS84	51.68	0.18	1153	4	0.35
TS75	51.68	0.18	1932	7	0.36
TS76	51.69	0.19	1587	2	0.13
TS78	51.75	0.25	1001	4	0.40
TS86	52.54	1.04	1291	9	0.70
TS8	52.56	1.06	1748	15	0.86
TS43	52.80	1.3	3578	36	1.01
TS11	53.15	1.65	1721	22	1.28
TS87	53.57	2.07	2261	53	2.34
TS45	53.66	2.16	1509	36	2.39
TS10	54.56	3.06	1913	55	2.88

Maize PH184C plants homozygous for the NPTII gene, which has been inserted in preselected target sites at CTL1 as part of the site-specific insertion landing pad (SSILP), were crossed with the M14 line which contains the herbicide resistance gene PMI at the location 51.54 cM on Chromosome 1. The resulting F1 progenies were crossed with the wild-type PH184C, producing BC1 seeds. Genotyping of BC1 seeds was performed with qPCR assays. Recombinants, seeds were positive for both NPTII and PMI and null for both NPTII and PMI.

used for construction of CTLs in maize and other crops (Steinert et al., 2015; Begemann et al., 2017; Tang et al., 2017), or combined with SpyCas9 to increase target density.

To obtain a usable SSILP, HDR must take place at the target site. In this study, we were able to obtain 2 × HDR events with high frequency. We also found events with HDR occurring only at one end of the SSILP insert while the other end likely was repaired through NHEJ because the junction PCR was negative. These 1 × HDR events often had truncations, insertion of other plasmid fragments or rearrangement of the template DNA. We selected events for 2 × HDR and against 1 × HDR using junction PCR, long PCR and sequencing, and the insert sequence integrity of the selected 2 × HDR events was verified by SbS analysis. The sites were mostly transformed once. Why some sites were more efficient than other sites SSILP insertion requires further study.

Four plasmids were co-bombarded to insert a SSILP to CRISPR target sites. We adopted this strategy, instead of using an all-in-one plasmid, because vectors containing one or two gene expression cassettes are easy to construct and it allows changes in components and plasmid ratio in transformation. Although using individual plasmids may increase the possibility of random plasmid insertion at different locations in the genome, this work and previous studies have shown co-bombarded plasmids tend to insert at the same location in the genome, especially when the morphogenic genes are used in transformation (Gao et al., 2020). We obtained helper-gene free plants for the majority of the SSILP lines generated. The random insertion of co-bombarded plasmids was not a significant limiting factor in populating CTLs with SSILP.

Selected SSILP lines were tested for reception of trait genes using many constructs. RMCE frequency with *Agrobacterium*-mediated transformation was higher than that in bombardment. These results likely are due to different FRT sites used in the site-specific integration. In *Agrobacterium*-mediated transformation, the SSILP and donor vectors had the FRT1/6 pair while the FRT1/87 pair was used in the SSILP and donor in bombardment. It has been reported that the FRT1/6 had a lower excision rate than FRT1/87 when the FLP recombinase was present (Anand et al., 2019). There is a 1 nt difference between FRT1 and FRT87 in the spacer region, but FRT1 and FRT6 differ by 3 nts. FRT cross-reactivity was found to be negatively correlated with RMCE frequency, and the relationship was more pronounced in *Agrobacterium*-mediated transformation. The RMCE frequency was not significantly different between FRT1/87 and FRT1/6 when using bombardment (Li et al., 2009; Anand et al., 2019).

We found that the plants containing SSILP had normal growth and development in the greenhouse and in the field. The SSILP insertion sites were preselected at least 2 kb away from any known gene. RNA-seq analysis demonstrated that insertion of SSILP had no significant effect on expression of nearby endogenous genes. The SSILP plants were generated by back-crossing with wildtype twice. Our work on CRISPR-waxy corn product development (Gao et al., 2020) and several other studies using CRISPR-Cas9 has found very limited or no off-target cutting in plants (Tang et al., 2018; Hahn and Nekrasov, 2019; Li et al., 2019; Young et al., 2019). A few albino seedlings were found among the selfed BC1F2 plants in two events. However, this albino phenotype was not caused by SSILP insertion, was seen in null segregants as well, and is a common occurrence in maize genetics. We selected SSILP from events without abnormal plants. Occasional off-type plants also are seen in traditional genetic crosses or transgenic regenerants; we observed nothing unusual in this CRISPR-Cas9 mediated gene insertion work. Importantly, yield test showed that SSILPs were agronomically neutral when compared to null segregants.

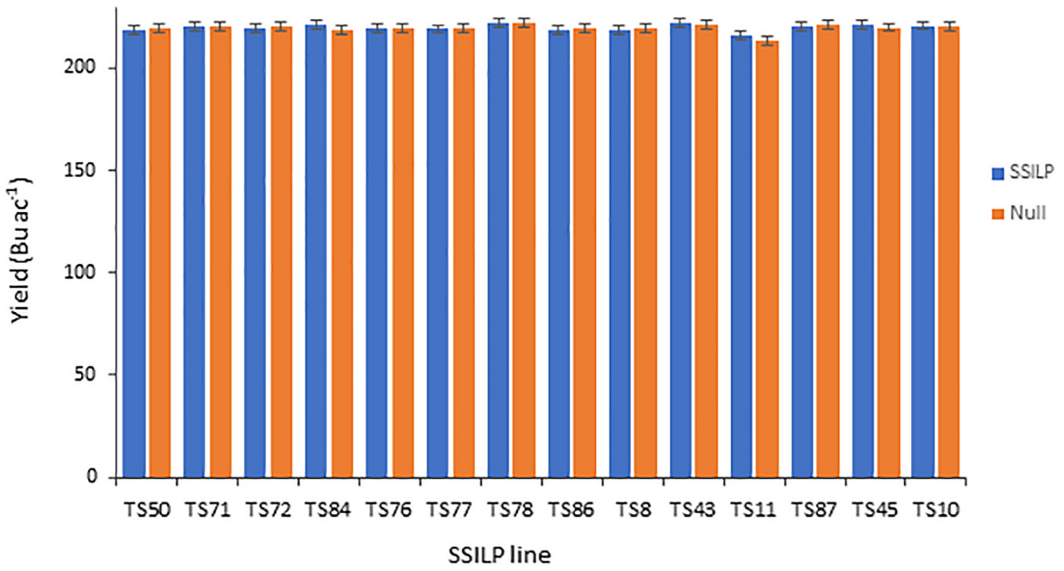


FIGURE 6 | Grain yield of maize hybrids containing SSILP and corresponding nulls. Yield trials were carried out in the United States corn-belt in 2017. Each bar represents data from three hybrids at 12 locations. All analyses were implemented using ASReml with output of the model presented as best linear unbiased prediction (BLUP). The yield of the SSILP hybrids were not significantly different from their corresponding nulls ($P > 0.05$, two-tailed test).

Most sites supported similar level transgene expression within an inbred line, a result that is consistent with earlier reports (Chawla et al., 2006; Nanto et al., 2009; Betts et al., 2019). It has been proposed that transgenic event recovery is dependent upon the ability of the selectable marker or screenable marker to be expressed. Because of this, events in regions of the genome where silencing occur will not be recovered (Francis and Spiker, 2004). While it is possible such repressive locations exist in the maize genome, in this report most of the sites tested supported gene expression, suggesting that at least in the chromosome regions studied here, consistent transgene expression at preselected sites is the norm. It is noteworthy that expression at identical sites was significantly different across different genetic backgrounds, PH184C < HC69 < PHH5G. Given the ability to target a unique SSI landing pad at multiple genetically identical target sites across different maize elite inbreds, this is the first report on examining how differences in genetic background influence gene expression without the linkage-related complications associated with backcrossing.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

Three Pioneer® inbred lines PH184C, HC69, and PHH5G were used in these experiments. M14 is a transgenic PH184C line generated via meganuclease mediated HDR, carrying the selectable marker gene *PMI*. SSILP lines used in transformation for trait gene integration were heterozygous plants from crossing of homozygous BC1F3 or later generations with wild type PH184C. Embryo donor plants and transgenic plants were grown in greenhouses as previously described (Shi et al., 2017).

CRISPR-Cas9 Target Sites Selection

The genetic and physical location of CTSs on chromosomes were calculated based on marker prediction using the B73 reference genome. The DNA sequences from the chromosomal region of the four CTLs were scanned using proprietary bioinformatic tools for unique sequence regions at least 2 kb away from any native gene, then potential CTSs were identified by first locating a suitable PAM for *S. pyogenes* Cas9, NGG and then extracting the sequence between 17 and 24 bp 5′ of the PAM for use as the spacer in the sgRNA. The off-targeting cutting potential of Cas9 based on the selected CTSs was evaluated by searching the B73 reference genome and transformation inbred lines for closely matching targets using Bowtie 2 and PSI-BLAST (Altschul et al., 1997; Langmead and Salzberg, 2012). Only CTSs different from other genomic locations by at least two mismatches in the target site seed region were selected (1–10 bp 5′ of the PAM).

Plasmid Construction and Maize Transformation

The single guide RNA gene consists of a maize U6 polymerase III promoter, a CRISPR RNA, a trans-activating CRISPR RNA and a terminator (**Supplementary Figure S1**). The Cas9 expression cassette contains the maize UBIQUITIN1 promoter (UBI PRO), *S. pyogenes* Cas9 endonuclease and potato protease inhibitor II terminator (PINII). The Cas9 DNA sequence was maize codon optimized and the potato ST-LS1 intron and the nuclear localization signals from the SV40 were added for appropriate expression and nuclear targeting in maize, as previously described (Svitashev et al., 2015). Constructs were assembled using chemically synthesized DNA fragments with standard DNA techniques. *NPTII* served as a transformation selection marker. To improve regeneration of plants, morphogenic regulators *Bbm*

(also known as ovule development protein 2 or *ODP2*) and *Wus2* were expressed under control of the maize UBI1 promoter and In2-2 promoter, respectively (**Supplementary Figure S1**), and the plasmids were constructed as described previously (Lowe et al., 2016).

Embryos from PHH5G line used in SSILP insertion contained a pre-integrated T-DNA of *Bbm* and *Wus2* to enhance transformation. Biolistic-mediated transformation of maize immature embryos was performed as previously described (Svitashev et al., 2015). Briefly, gold particles, 0.6 μm in diameter, were washed with 100% ethanol and sterile distilled water. The plasmid DNA purified with QIAprep Spin Miniprep (Qiagen, Germany) and mixture of Cas9-gRNA/donor template/BBM/WUS2 = 5/5/2.5/2.5 was precipitated on the washed gold particles using a water-soluble cationic lipid TransIT-2020 (Mirus). Fifty microliters of gold particles (water solution of 10 mg/mL) and 1 μL of TransIT-2020 water solution were added to the premixed DNA, mixed gently. DNA-coated gold particles were then centrifuged at 8,000 *g* for 1 min. The pellet was rinsed with 100 μL of 100% ethanol and re-suspended by a brief sonication. Immediately after sonication, DNA-coated gold particles were loaded onto the center of a macro-carrier (10 μL of each) and allowed to air dry. Immature embryos 9–11 days after pollination were bombarded using a PDS-1000 Helium Gun (Bio-Rad) with a rupture pressure of 425 psi. Post-bombardment culture, selection, and plant regeneration were carried out as described (Svitashev et al., 2015).

To integrate a trait gene to the SSILP, plasmids used in *Agrobacterium*-mediated transformation contain six expression cassettes: the trait gene and phosphinothricin acetyltransferase (*PAT*) gene in the donor, the transformation selection marker *PMI* (promoter-less), the transformation enhancer *Bbm* and *WUS2*, and FLP recombinase for RMCE (**Supplementary Figure S2**). The coding sequences, promoters and terminators as well as FRT and loxP were PCR-amplified or chemically synthesized, verified by DNA sequencing and assembled in a Gateway-modified derivative of pSB11. The plasmids then were co-integrated into the super binary pSB1 vector in *Agrobacterium tumefaciens* strain LBA4404 by electroporation. *Agrobacterium*-mediated transformation of maize immature embryos was performed as described previously (Lowe et al., 2016). For biolistic-mediated transformation, individual plasmids of UBI:WUS2, UBI:BBM, and UBI:FLP were co-delivered with the donor plasmid containing *PMI* and the trait gene flanked by FRT1 and FRT87 sites to immature embryos as described above.

DNA Extraction and Genotyping by PCR

Genomic DNA was extracted from leaves as described previously (Shi et al., 2017). A qPCR assay was used to estimate the copy number of each CTS. Shoots with no modification contain two copies of the wild-type CTS, shoots with CTS modification either due to NHEJ or SSILP insertion in one of the two homologous chromosomes has one intact copy, while modification in both chromosomes would reduce the copy number to zero. qPCR was performed using Qiagen QuantiTect Multiplex PCR Master Mix (Qiagen, Germany) primers and probe specific for each CTS (**Supplementary Table S3**). Junction PCR assays were used to detect SSILP insertion at each CTS. In this assay, to increase screening throughput, PCR amplification of the target region was coupled with a nested qPCR to detect SSILPs. PCR was performed using 2x Extract-N-amp PCR Ready Mix (Cat# E3004, Sigma) or 2× Phusion Flash High-fidelity PCR Master Mix (Cas#F548L, Thermo Fisher Scientific). For nested PCR used in screening SSILP insertion events, the first PCR was carried out in 5 μL of reaction mixtures for 20 cycles. Fifteen μL of the reaction mixture containing 2×TagMan Master Mix (LGC Cat# KBS-1001-001) and primers then were added, and the second PCR was performed using LightCycler 480 (Roche Life Science) for 30 cycles. Data were analyzed using the Endpoint Genotyping Software (Roche Life Science). Long PCR was performed using Extensor Master Mix (Cat# AB-0792, Thermo Fisher Scientific). HR1f and HR2r primers which varied among CTS, primers and probes are listed in **Supplementary Table S3**. To detect trait gene integration at the SSILP via FLP mediated RMCE, qPCR was performed (**Supplementary Table S4**). The qPCR for identifying recombinants from crossing of M14 and the SSILP lines was performed using Qiagen QuantiTect Multiplex PCR Master Mix (Qiagen, Germany). Primers and probes are listed in **Supplementary Table S4**.

Detection of Plasmid DNA in Plants

Presence of plasmid DNA in the genome of T0 or T1 plants was determined by qPCR and SbS. QPCR was performed using Qiagen QuantiTect Multiplex PCR Master Mix (Qiagen, Germany) with primers and probes listed in **Supplementary Table S4**. SbS was performed as described by Zastrow-Hayes et al. (2015). A capture-probe library was created to cover the four plasmids used in transformation (**Supplementary Figure S1**). Illumina whole-genome sequencing libraries were constructed from DNA derived from plants. Hybridizations and sequencing were carried out as described (Zastrow-Hayes et al., 2015).

Quantification of Neomycin Phosphotransferase II Proteins by ELISA

Plants were grown in greenhouse in 4 × 8 cell flats in a randomized complete block design. Leaf punches were taken from the third leaves of 3-week-old greenhouse grown heterozygous plants. The leaf samples were ground in 500 μL PBST with two metal beads using a Geno/Grinder at 1,650 rpm for 60 s followed by centrifugation at 4°C (3,889 *g*) for 10 min. Total protein was quantified using the Bradford Protein Assay Kit (Bio-Rad) per the manufacturer's instructions. The NPTII protein was measured using a polyclonal antibody-based sandwich ELISA assay as previously described (Schmidt and Alarcon, 2011). The NPTII specific polyclonal antibody assay standard curve was linear from 2 to 40 ng/mL. Samples were assayed in duplicate with comparison of interpolations across varying sample dilution. Controls (negative and known low and high positive) were included on each assay plate. Results were reported as parts per million based on total protein.

Endogenous Gene Expression Analysis by RNA Sequencing

Plants were grown in greenhouse in 4 × 8 cell flats in a randomized complete block design. Leaf samples for the SSILP

lines and controls were taken from 2-week-old seedlings in four biological replicates. Each biological replicate consists of 3–4 individual plants. Total RNAs were extracted using the Qiagen RNeasy kit for total RNA isolation (Qiagen, Germantown, MD, United States). Sequencing libraries from the resulting total RNAs were prepared using the TruSeq mRNA-Seq kit according to the manufacturer's instructions (Illumina, San Diego, CA, United States) and sequenced on the Illumina HiSeq 2500 system with Illumina TruSeq SBS version 3 reagents. Reads were trimmed based on quality scores, filtered and mapped to the PH184C reference transcripts using the aligner software Bowtie2 v2.3.4.1 (Langmead and Salzberg, 2012). The PH184C reference sequences were established using following steps: First, the PH184C CTL1 scaffold was mapped to B73 RefGen v4.59 using Minimap v2.14-r883 to determine the general location of CTL1 in the genome (Li, 2016; Jiao et al., 2017). Next, the B73 gene sequences from this region were extracted from RefGen v4.59 and mapped to the PH184C CTL1 scaffold using GMAP v2018-07-04 (Wu and Watanabe, 2005). Then the PH184C mapped gene sequences on the CTL1 scaffold, which could be slightly different from the B73 sequences, were extracted and served as the reference transcripts for RNA-seq analysis (Portwood et al., 2019).

Statistical Analysis of RNA-Seq Data

RSEM v1.2.28 was used to estimate transcript abundance (Li and Dewey, 2011). Initial exploratory analysis of the aligned sequences showed high variability due to plant location in the greenhouse, which contributed significantly to the overall variability in the sequencing data. Spatial affects were captured using surrogate variable analysis (svaseq package in R) (Leek, 2014). Using the DESeq2 package in R, differential analysis was performed by first selecting likelihood-ratio test models to test for overall significance of the SSILP site variable (Love et al., 2014). The full model included the SSILP site information and surrogate variables found in the surrogate variable analysis, while the null model included only the surrogate variables to account for spatial variability. The Wald test then was performed for pairwise comparisons of each SSILP site against the CTL2 control. The differential expression is defined as the false discovery-corrected $P < 0.05$ and the absolute value of log2 (fold change) > 1.

Hybrid Seed Production and Yield Test

Regenerated T0 plants containing a SSILP were backcrossed to respective wild type inbred line for two successive generations followed by selfing to develop BC1F2 seeds. Plants homozygous for the SSILP or null segregant which did not contain the SSILP were selfed to produce BC1F3 seeds. BC1F3 plants were planted in the field, and individual plants were selected for uniformity. Selected BC1F3 plants were crossed to three elite tester female lines to produce hybrid seeds for yield testing. F1 hybrid seeds were therefore either null or hemizygous for a SSILP.

Hybrid yield testing was conducted at 12 locations in the United States. The experimental design was two-row plots nested by tester and a set of SSILP lines and corresponding null segregants as controls. Fertilizer was applied, and weeds and pests were controlled according to local practices. Small plot combines were used to collect grain mass and grain moisture data. Grain yield was calculated by normalizing all entries to 15% moisture.

A mixed model framework was used to perform the multi-location data analyses (Betts et al., 2019). Analysis was implemented using ASReml (VSN International Ltd.), and the values are best linear unbiased prediction (BLUP) (Gilmour et al., 2009). The SSILP hybrids were compared to corresponding null-segregant controls, and statistical differences were determined at $P < 0.05$.

Material Availability

Novel biological materials described in this publication may be available to the academic community and other not-for-profit institutions solely for non-commercial research purposes upon acceptance and signing of a material transfer agreement between the author's institution and the requestor. In some cases, such materials may originally contain genetic elements described in the manuscript that were obtained from a third party(s), and the authors may not be able to provide materials including third party genetic elements to the requestor because of certain third-party contractual restrictions placed on the author's institution. In such cases, the requester will be required to obtain such materials directly from the third party. The authors and authors' institution do not make any express or implied permission(s) to the requester to make, use, sell, offer for sale, or import third party proprietary materials. Obtaining any such permission(s) will be the sole responsibility of the requestor. To protect Corteva Agriscience's proprietary germplasm, such germplasm will not be made available except at the discretion of Corteva Agriscience and then only in accordance with all applicable governmental regulations.

DATA AVAILABILITY STATEMENT

All data supporting the findings of this study are available in the article/**Supplementary Material**. RNA sequencing data are available in the SRA database in the BioProject PRJNA612862. Plasmids accession number are MT221180, MT221181, MT221182, and MT221179 correspond plasmid 1, 2, 3, and 4 on **Supplementary Figure S1**.

AUTHOR CONTRIBUTIONS

HG designed, coordinated experiments. JY and HG selected CTS. JM selected CTL regions. JM and HG conducted stacks. MY, MS, and BL performed molecular analysis on SSILPs. LW constructed DNA vectors. DP, GS, SJ, LF, WM, MZ, SW performed plant transformation. JF and KS conducted sampling and plant care. NC designed RNA-seq experiment. CS and HG led gene expression studies. XS performed statistical design and analysis for gene expression. JS led the trait gene insertion experiments. JM conducted field trial. SB, AC, NC, and HG designed SSILP structure. HG and NC wrote the manuscript. JM, JY, XS, JS, SB, and AC edited the manuscript.

FUNDING

There is no other funding. Research was done in Corteva Agriscience.

ACKNOWLEDGMENTS

We thank Laura Church and team for greenhouse plant care. Mary Betty, Dennis O'Neil, and their teams for nucleic acid analysis. Kelli Van Waus and team for protein analysis. Gina Zastrow-Hayes and team for sequencing. Terry Hu and team for their help in plant transformation. Mauricio La Rota, and John Van Hemert for assistance with RNA-seq data analysis. Sutirtha Basu, Dave Stucker, Maria Fedorova, Julian Chaky, and Michael Lassner for providing project support. We are grateful to Todd Jones and Jinrui Shi for critical review of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00535/full#supplementary-material

REFERENCES

Ainley, W. M., Sastry-Dent, L., Welter, M. E., Murray, M. G., Zeitler, B., Amora, C. D. R., et al. (2013). Trait Stacking via targeted genome editing. *Plant Biotechnol. J.* 11, 1126–1134.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402.

Anand, A., Wu, E., Li, Z., TeRonde, S., Arling, M., Lenderts, B., et al. (2019). High efficiency Agrobacterium-mediated site-specific gene integration in maize utilizing the FLP-FRT recombination system. *Plant Biotechnol. J.* 17, 1636–1645. doi: 10.1111/pbi.13089

Begemann, M. B., Gray, B. N., January, E., Gordon, G. C., He, Y., Liu, H., et al. (2017). Precise insertion and guided editing of higher plant genomes using Cpf1 CRISPR nucleases. *Sci. Rep.* 7:11606. doi: 10.1038/s41598-017-11760-6

Betts, S., Basu, S., Bolar, J., Booth, R., Chang, S., Cigan, A. M., et al. (2019). Uniform expression and relatively small position effects characterize sister transformants in maize and soybean. *Front. Plant Sci.* 10:1209. doi: 10.3389/fpls.2019.01209

Cao, J., Zhao, J. Z., Tang, D., Shelton, M., and Earle, D. (2002). Broccoli plants with pyramided cry1Ac and cry1C Bt genes control diamondback moths resistant to Cry1A and Cry1C proteins. *Theor. Appl. Genet.* 105, 258–264. doi: 10.1007/s00122-002-0942-0

Chawla, R., Ariza-Nieto, M., Wilson, A. J., Moore, S. K., and Srivastava, V. (2006). Transgene expression produced by biolistic-mediated, site-specific gene integration is consistently inherited by the subsequent generations. *Plant Biotechnol. J.* 4, 209–218. doi: 10.1111/j.1467-7652.2005.00173.x

Chen, W., and Ow, D. W. (2017). Precise, flexible and affordable gene stacking for crop improvement. *Bioengineered* 8, 451–456. doi: 10.1080/21655979.2016.1276679

Dafny-Yelin, M., and Tzfira, T. (2007). Delivery of multiple transgenes to plant cells. *Plant Pysiol.* 145, 1118–1128. doi: 10.1104/pp.107.106104

Danilo, B., Perrot, L., Botton, E., Nogué, F., and Mazier, M. (2018). The DFR locus: a smart landing pad for targeted transgene insertion in tomato. *PLoS One* 13:e0208395. doi: 10.1371/journal.pone.0208395

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

Diehn, S., Lu, A. L., Nowatzki, T. M., Register, J. C. III, Scelonge, C. J., Zhong, C. X., et al. (2013). *Maize Event DP-00411-3 and Methods for Detection thereof.* US Patent US8,575,434 B2.

Endo, M., Mikami, M., and Toki, S. (2016). Biallelic gene targeting in rice. *Plant Physiol.* 170, 667–677. doi: 10.1104/pp.15.01663

Francis, K. E., and Spiker, S. (2004). Identification of Arabidopsis thaliana transformants without selection reveals a high occurrence of silenced T-DNA integrations. *Plant J.* 41, 464–477. doi: 10.1111/j.1365-313x.2004.02312.x

Gao, H., Gadlage, M. J., Lafitte, H. R., Lenderts, B., Yang, M., Schroder, M., et al. (2020). Superior field performance of waxy corn engineered using CRISPR–Cas9. *Nat. Biotechnol.* [Epub ahead of print]. doi: 10.1038/s41587-020-0444-0

Gilmour, A. R., Gogel, B. J., Cullis, B. R., and Thompson, R. (2009). *ASReml User Guide 3.0*.

Hahn, F., 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

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

Jiao, Y., Peluso, P., Shi, J., Liang, T., Stitzer, M. C., Wang, B., et al. (2017). Improved maize reference genome with single-molecule technologies. *Nature* 546, 524–527. doi: 10.1038/nature22971

Koziel, M. G., Beland, G. L., Bowman, C., Carozzi, N. B., Crenshaw, R., Crossland, L., et al. (1993). Field performance of elite transgenic maize plants expressing an insecticidal protein derived from *Bacillus thuringiensis*. *Nat. Biotechnol.* 11, 194–200. doi: 10.1038/nbt0293-194

Kumar, S., AlAbed, D., Worden, A., Novak, S., Wu, H., Ausmus, C., et al. (2015). A modular gene targeting system for sequential transgene stacking in plants. *J. Biotchnol.* 207, 12–20. doi: 10.1016/j.jbiotec.2015.04.006

Kuscu, C., Arslan, S., Singh, R., Thorpe, J., and Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat. Biotechnol.* 32, 677–683. doi: 10.1038/nbt.2916

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Leek, J. T. (2014). svaseq: removing batch effects and other unwanted noise from sequencing data. *Nucleic Acids Res.* 42, 12961–12972. doi: 10.1093/nar/gku864

Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323. doi: 10.1186/1471-2105-12-323

Li, H. (2016). Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics* 32, 2103–2110. doi: 10.1093/bioinformatics/btw152

Li, J., Manghwar, H., Sun, L., Wang, P., Wang, G., Sheng, H., et al. (2019). Whole genome sequencing reveals rare off-target mutations and considerable inherent genetic or/and somaclonal variations in CRISPR/Cas9-edited cotton plants. *Plant Biotechnol. J.* 17, 858–868. doi: 10.1111/pbi.13020

Li, Z., Xing, A., Moon, B. P., McCardell, R. P., Mills, K., and Falco, S. C. (2009). Site-specific integration of transgenes in soybean via recombinase-mediated DNA cassette exchange. *Plant Physiol.* 151, 1087–1095. doi: 10.1104/pp.109.137612

Liu, X., Homma, A., Sayadi, J., Yang, S., Ohashi, J., and Takumi, T. (2016). Sequence features associated with the cleavage efficiency of CRISPR/Cas9 system. *Sci. Rep.* 6:19675. doi: 10.1038/srep19675

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-558

Lowe, K., Wu, E., Wang, N., Hoerster, G., Hastings, C., Cho, M. J., et al. (2016). Morphogenic regulators Baby boom and Wuschel improve monocot transformation. *Plant Cell* 28, 1998–2015. doi: 10.1105/tpc.16.00124

Mumm, R. H., and Walters, D. S. (2001). Quality control in the development of transgenic crop seed products. *Crop Sci.* 41, 1381–1389. doi: 10.2135/cropsci2001.4151381x

Nandy, S., Zhao, S., Pathak, B. P., Manoharan, M., and Srivastava, V. (2015). Gene stacking in plant cell using recombinases for gene integration and nucleases for marker gene deletion. *BMC Biotechnol.* 15:93. doi: 10.1186/s12896-015-0212-2

Nanto, K., Sato, K., Katayama, Y., and Ebinuma, H. (2009). Expression of a transgene exchanged by the recombinase-mediated cassette exchange (RMCE) method in plants. *Plant Cell Rep*. 28, 777–785. doi: 10.1007/s00299-009-0683-5

Ow, D. W. (2011). Recombinase-mediated gene stacking as a transformation operating system. *J. Integr. Plant Biol.* 53, 512–519. doi: 10.1111/j.1744-7909.2011.01061.x

Padgette, S. R., Kolacz, K. H., Delannay, X., Re, D. B., LaVallee, B. J., Tinius, C. N., et al. (1995). Development, identification, and characterization of a glyphosate-tolerant soybean line. *Crop Sci.* 35, 1451–1461. doi: 10.2135/cropsci1995.0011183x003500050032x

Peng, T., Sun, X., and Mumm, R. H. (2014). Optimized breeding strategies for multiple trait integration: I. Minimizing linkage drag in single event introgression. *Mol. Breed.* 33, 89–104.

Petolino, J. F., and Kumar, S. (2016). Transgenic trait deployment using designed nucleases. *Plant Biotechnol. J.* 14, 503–509. doi: 10.1111/pbi.12457

Portwood, J. L., Woodhouse, M. R., Cannon, E. K., Gardiner, J. M., Harper, L. C., Schaeffer, M. L., et al. (2019). MaizeGDB 2018: the maize multi-genome genetics and genomics database. *Nucleic Acids Res.* 47, D1146–D1154. doi: 10.1093/nar/gky1046

Que, Q., Chilton, M. D., de Fontes, C. M., He, C., Nuccio, M., Zhu, T., et al. (2010). Trait stacking in transgenic crops: challenges and opportunities. *GM Crops* 1, 220–229. doi: 10.4161/gmcr.1.4.13439

Schmidt, J., and Alarcon, C. (2011). *Immunoassay Method Validation: Immunoassays in Agricultural Biotechnology Shan/Immunoassays*, Chap. 6. Hoboken, NJ: John Wiley & Sons, Inc, 115–138.

Shi, J., Gao, H., Wang, H., Lafitte, R., Archibald, R. L., Yang, M., et al. (2017). ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. *Plant Biotechnol. J.* 15, 207–216. doi: 10.1111/pbi.12603

Srivastava, V., and Thomson, J. (2016). Gene stacking by recombinases. *Plant Biotechnol. J.* 14, 471–482. doi: 10.1111/pbi.12459

Steinert, J., Schiml, S., Fauser, F., and Puchta, H. (2015). Highly efficient heritable plant genome engineering using Cas9 orthologues from *Streptococcus thermophilus* and *Staphylococcus aureus*. *Plant J.* 84, 1295–1305. doi: 10.1111/tpj.13078

Svitashev, S., Young, J. K., Schwartz, C., Gao, H., 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

Tang, X., Liu, G., Zhou, J., Ren, Q., You, Q., Tian, L., et al. (2018). Large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. *Genome Biol.* 19:84. doi: 10.1186/s13059-018-1458-5

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* 3:17103. doi: 10.1038/nplants.2017.103

Wu, T. D., and Watanabe, C. K. (2005). GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* 21, 1859–1875. doi: 10.1093/bioinformatics/bti310

Wu, X., Scott, D. A., Kriz, A. J., Chiu, A. C., Hsu, P. D., Dadon, D. B., et al. (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat. Biotechnol.* 32, 670–676. doi: 10.1038/nbt.2889

Young, J., Zastrow-Hayes, G., Deschamps, S., Svitashev, S., Zaremba, M., Acharya, A., et al. (2019). CRISPR-Cas9 editing in maize: systematic evaluation of Off-target activity and its relevance in crop improvement. *Sci. Rep.* 9:6729. doi: 10.1038/s41598-019-43141-6

Zastrow-Hayes, G. M., Lin, H., Sigmund, A. L., Hoffman, J. L., Alarcon, C. M., Hayes, K. R., et al. (2015). Southern-by-sequencing: a robust screening approach for molecular characterization of genetically modified crops. *Plant Genome* 8, 1–15. doi: 10.3835/plantgenome2014.08.0037

Conflict of Interest: All authors were employees of Corteva Agriscience at the time of their contributions to any experiments works. Patents related to this work have been filed.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 21 May 2020
doi: 10.3389/fpls.2020.00621

The Improvement of CRISPR-Cas9 System With Ubiquitin-Associated Domain Fusion for Efficient Plant Genome Editing

*Xuelian Zheng, Caiyan Qi, Lijia Yang, Quan Quan, Binglin Liu, Zhaohui Zhong, Xu Tang, Tingting Fan, Jianping Zhou and Yong Zhang**

Department of Biotechnology, School of Life Sciences and Technology, Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu, China

Genome editing technology represented by CRISPR-Cas9 had been widely used in many biological fields such as gene function analysis, gene therapy, and crop improvement. However, in the face of the complexity of the eukaryotic genome, the CRISPR-Cas9 genome editing tools have shown an unstable editing efficiency with large variability at different target sites. It was important to further improve the editing efficiency of the CRISPR-Cas9 system among the whole genome. In this study, based on the previous single transcription unit genome editing system (STU-SpCas9), using the ubiquitin-associated domain (UBA) to enhance the stability of Cas9 protein, we constructed three Cas9-UBA fusion systems (SpCas9-SD01, SpCas9-SD02, and SpCas9-SD03). Four different target sites of rice OsPDS, OsDEP1 and OsROC5 genes were chosen to evaluate the genome editing efficiency in rice protoplasts and stable transformed rice plants. The results showed that the fusion of UBA domains did not affect the cleavage mode of Cas9 protein, and effectively increase the editing efficiency of STU-SpCas9 at the target sites. This new CRISPR-Cas9-UBA system provided a new strategy and tool for improving the genome editing efficiency of CRISPR-Cas9 in plants.

Keywords: CRISPR-Cas9, UBA domain, rice, genome editing efficiency, high efficiency

OPEN ACCESS

Edited by:
Hiroshi Ezura,
University of Tsukuba, Japan

Reviewed by:
Kaijun Zhao,
Chinese Academy of Agricultural Sciences, China
Tao Sun,
Genentech, Inc., United States

****Correspondence:***
Yong Zhang
zhangyong916@uestc.edu.cn

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *04 January 2020*
Accepted: *22 April 2020*
Published: *21 May 2020*

Citation:
Zheng X, Qi C, Yang L, Quan Q, Liu B, Zhong Z, Tang X, Fan T, Zhou J and Zhang Y (2020) The Improvement of CRISPR-Cas9 System With Ubiquitin-Associated Domain Fusion for Efficient Plant Genome Editing. Front. Plant Sci. 11:621. doi: 10.3389/fpls.2020.00621

INTRODUCTION

The CRISPR-Cas9 system has been the most widely used genome editing technology for gene function analysis, gene therapy, and crop improvement in eukaryotic species because of its simple construction, high efficiency, and low cost (Li et al., 2013; Nekrasov et al., 2013; Shan et al., 2013; Liang et al., 2017; Lowder et al., 2018; Zhou et al., 2019). Numerous CRISPR-Cas9 tools have been developed to achieve targeted mutagenesis, base editing, precise editing by homology-directed repair (HDR) and transcriptional regulation in plants (Chen et al., 2019; Zhang et al., 2019). However, the CRISPR-Cas9 system still has some shortcomings such as off-target effects, ineffectiveness at some genomic sites, considerably variable editing efficiency, etc. The editing frequencies (insertion/deletion, indel) at different target sites are quite variable. For example in rice, the efficiency of some target sites was as high as 90–100%, while the others were less than 1% (Liu et al., 2017, 2019; Ding et al., 2019; Mao et al., 2019). Therefore, it is desirable to develop an ideal CRISPR-Cas9 system with sustained high activity at whole genome target sites.

The CRISPR-Cas9 editing efficiency is mainly dependent on two aspects: the genomic target sites and the CRISPR-Cas9 system itself. There have been some studies to alter the accessibility of genomic target site to improve editing efficiency *in vivo*, such as the proxy-CRISPR, CRISPR-chrom and Cas9-TV (Chen et al., 2017; Hille et al., 2018; Ding et al., 2019; Liu et al., 2019). There also have been many attempts to enhance the activity of CRISPR-Cas9 system to improve editing efficiency, which included expanding Cas9 nucleases and their variants, optimizing the structure of Cas9 for highly active protein, Cas9 codon optimization and improvement of sgRNA design, the dual Pol II promoter systems for high transcription level of Cas9 and sgRNAs, using strong promoters to enhance the transcription level of Cas9 gene in the expression system, etc. (Fauser et al., 2014; Cermak et al., 2015, 2017; Kleinstiver et al., 2015; Lowder et al., 2016; Tang et al., 2016; Casini et al., 2018; Hu et al., 2018; Chen et al., 2019; Mao et al., 2019; Zhang et al., 2019; Zhong et al., 2019). It is well known that the intracellular content of proteins is not only related to the expression level, but also to the degradation rate. Except for the traditional methods to improve the expression level through the transcription and translation of Cas9 protein, we consider whether there are other solutions to improve the intracellular content of Cas9 protein, such as inhibition of degradation.

Most proteins were degraded via the ubiquitin/26S proteasome pathway, which involved the participation of a large number of specific proteins and enzymes (Vierstra, 2009). A class of UBA (ubiquitin-associated) domains has been found in many proteins involved in protein degradation processes. In general, the UBA domains contained about 45 amino acid residues that were conserved from yeast to higher eukaryotes, the most conserved residues were generally non-polar, indicating that the UBA domain was unlikely to be directly involved in phosphorylation or ubiquitination (Hofmann and Bucher, 1996; Bertolaet et al., 2001). The yeast ubiquitin receptors Rad23, Dsk2, and Ddi1 delivered polyubiquitylated protein substrates to the proteasome for destruction. These receptors contained an N-terminal ubiquitin-like (UBL) domain that was recognized by the 26S proteasome and one or more C-terminal UBA domains that bound to substrates. However, these receptor proteins themselves escaped degradation and had long half-lives because of their C-terminal UBA domains. The C-terminal UBA domain of Rad23, Dsk2 and Ddi1 acted as an intrinsic stabilization signal that protected these receptors from proteasomal degradation by inhibiting multi-ubiquitin chain assemble or preventing the generation of initiation sites for degradation, which are required for proper engagement of the inherent unfolding machinery of the proteasome (Chen et al., 2001; Raasi and Pickart, 2003; Heessen et al., 2005; Raasi et al., 2005; Farmer et al., 2010; Heinen et al., 2011).

A few UBA chimeric protein studies have identified that the UBA domain can be used to enhance the stability of the target protein, prolong its half-life, and successfully improve the activity of the target protein. The UBA2 domain used to increase the stability of a destabilized GFP reporter protein in yeast (Heessen et al., 2005). The UBA1 and UBA2 domains of Arabidopsis ubiquitin receptor protein RAD23a functioned as a portable stable signal that extends the half-life of two unstable transcription factors HFR1 and PIF3 in Arabidopsis (Jang et al., 2012). The UBA domain of Arabidopsis DDI1 protein also increased the half-life of the unstable protein JAZ10.1 associated with jasmonic acid signaling (Jang et al., 2012). Based on these results, we have hypothesized that fusion modification of Cas9 protein with UBA domain might also increase the half-life of Cas9 protein and enhance Cas9 activity. Therefore, we tested this concept by fusing the Cas9 protein with three different Arabidopsis UBA domains (UBA1, UBA2, and UBA3) and developing the novel Cas9-UBA editing systems based on our previous reported STU-Cas9 system (Tang et al., 2016, 2019). The new Cas9-UBA genome editing systems effectively enhanced the activity of Cas9 protein and improve the editing efficiency of multiple target sites of OsPDS, OsDEP1, and OsROC5 genes in rice. These new systems can significantly improve the editing efficiency of target gene sites, and provides an alternative method for plant gene targeted mutagenesis and crop genome editing breeding.

RESULTS

All Three Cas9-UBA Fusion Proteins Can Play Editing Activities in Rice Protoplasts

To investigate whether Cas9 proteins fused with different UBA stable domains (SD) had editing activities, the three Cas9-UBA systems were used and compared to the STU-Cas9 (pGEL028) system (**Figure 1A**). With these four Cas9 systems (SpCas9, SpCas9-SD01, SpCas9-SD02, SpCas9-SD03), we targeted four sites (OsPDS-sgRNA01, OsPDS-sgRNA02, OsDEP1-sgRNA02, and OsROC5-sgRNA01, **Figure 1B**) in the rice genome. The resulting 16 constructs were used for transient transformation of rice protoplasts (**Supplementary Table S2**). All four Cas9 systems showed significant mutagenesis in four target sites resulting from error-prone NHEJ, as revealed by restriction fragment length polymorphism (RFLP) analysis (**Figure 1C**). The frequency of mutation was measured by high-throughput deep sequencing of PCR amplicons as a sum of insertion and deletions at the target sites. Mutation frequencies induced by the four Cas9 systems ranged from 20 to 30% across the four targets (**Figure 2A**). The STU-SpCas9 and three SpCas9-SD systems showed similar editing efficiencies and no significant differences were observed. This indicated that the fusion of different UBA SD domains does not affect the nuclease digestion activity of the Cas9 protein. The further analysis of NHEJ mutations in all samples showed no significant differences in the deletions profiles at target sites in the different Cas9-SD systems. Most of mutations produced by the four Cas9 systems were deletions, and the most frequent deletion positions were at 3 to 5 bp upstream of the PAM site (**Figure 2B**, **Supplementary Figure S1A**). The deletion majority ranged from 1 to 3 bp in size, and 1 bp deletions were the most predominant deletion type (**Figure 2C**, **Supplementary Figure S1B**). These results indicated that the fusion of the UBA domain does not influence the cleavage mode of Cas9 protein, and they were completely consistent with our previous report that the NHEJ repair outcomes are largely dictated by the sequence

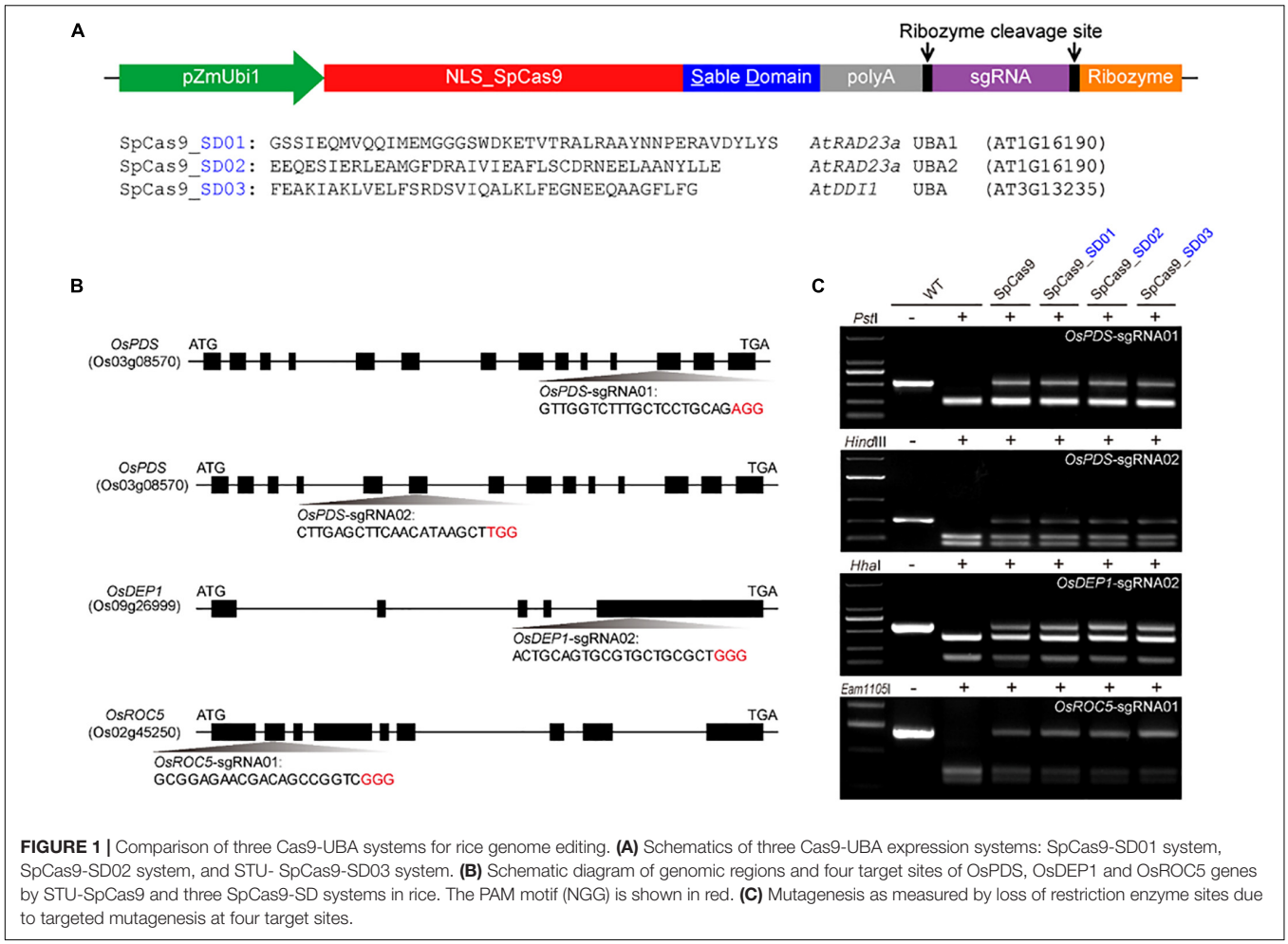


FIGURE 1 | Comparison of three Cas9-UBA systems for rice genome editing. **(A)** Schematics of three Cas9-UBA expression systems: SpCas9-SD01 system, SpCas9-SD02 system, and STU- SpCas9-SD03 system. **(B)** Schematic diagram of genomic regions and four target sites of OsPDS, OsDEP1 and OsROC5 genes by STU-SpCas9 and three SpCas9-SD systems in rice. The PAM motif (NGG) is shown in red. **(C)** Mutagenesis as measured by loss of restriction enzyme sites due to targeted mutagenesis at four target sites.

composition of the target sites but not the expression systems (Tang et al., 2016, 2019).

Evaluation of STU-SpCas9 and Three SpCas9-SD Systems in Rice Transgenic Plants

We further compared the editing efficiency of STU-SpCas9 and three SpCas9-SD systems with different UBA SD domains in stable transgenic rice plants. Eight constructs, respectively, targeting two rice genes (OsPDS-sgRNA01 and OsDEP1-sgRNA02) were transformed into rice calli mediated by Agrobacterium. Analysis of individual T0 transgenic plants from independent calli revealed that the two target sites had high efficiency mutation (**Table 1**). For OsPDS-sgRNA01 site constructs, more than 15 transgenic plants were obtained for RFLP analysis and genotyping with Sanger sequencing (**Figure 3A**, **Supplementary Figure S2A**). With the control STU-SpCas9 system, 12 out of 16 (75.0%) T0 plants were mutated and 9 plants (56.3%) carried bi-allelic mutations. For SpCas9-SD01 system, 12 out of 15 (80.0%) were mutated and 11 plants (73.3%) carried bi-allelic mutations. For SpCas9-SD02 system, 16 out of 18 (88.9%) were mutated of 13 plants (72.2%) carried biallelic mutations. For SpCas9-SD03 system, 14 out of 16 (87.5%) were mutated and 14 plants (87.5%) carried biallelic mutations. The albino phenotype among all biallelic mutants was observed since the *OsPDS* gene knockout (**Figure 3C**). The mutation rate and bi-allelic mutation efficiency of all three SpCas9-SD systems were higher than those of the STU-SpCas9 system. It suggested that UBA SD domains improved the editing efficiency of Cas9.

RFLP analysis and genotyping results similar to above were also observed in OsDEP1-sgRNA02 site constructs (**Figure 3B**, **Supplementary Figure S2B**). For the STU-SpCas9 control system, 11 out of 16 (68.8%) T0 plants were mutated and 8 plants (50.0%) carried bi-allelic mutations. For SpCas9-SD01 system, 14 out of 17 (82.4%) were mutated and 6 plants (35.3%) carried bi-allelic mutations. For SpCas9-SD02 system, 15 out of 17 (88.2%) were mutated of six plants (35.3%) carried biallelic mutations. For SpCas9-SD03 system, 15 out of 18 (83.3%) were mutated and 11 plants (61.1%) carried biallelic mutations (**Table 1**). Knockout of OsDEP1 results in semi-dwarf, panicle erectness, a reduced panicle length, an increased number of grains per panicle and a consequent increase in grain yield (Huang et al., 2009; Zhou et al., 2009). We compared the trait measures such as plant height, panicle length, grain number per panicle, 1000-grain weight, the grain length and width in the wildtype

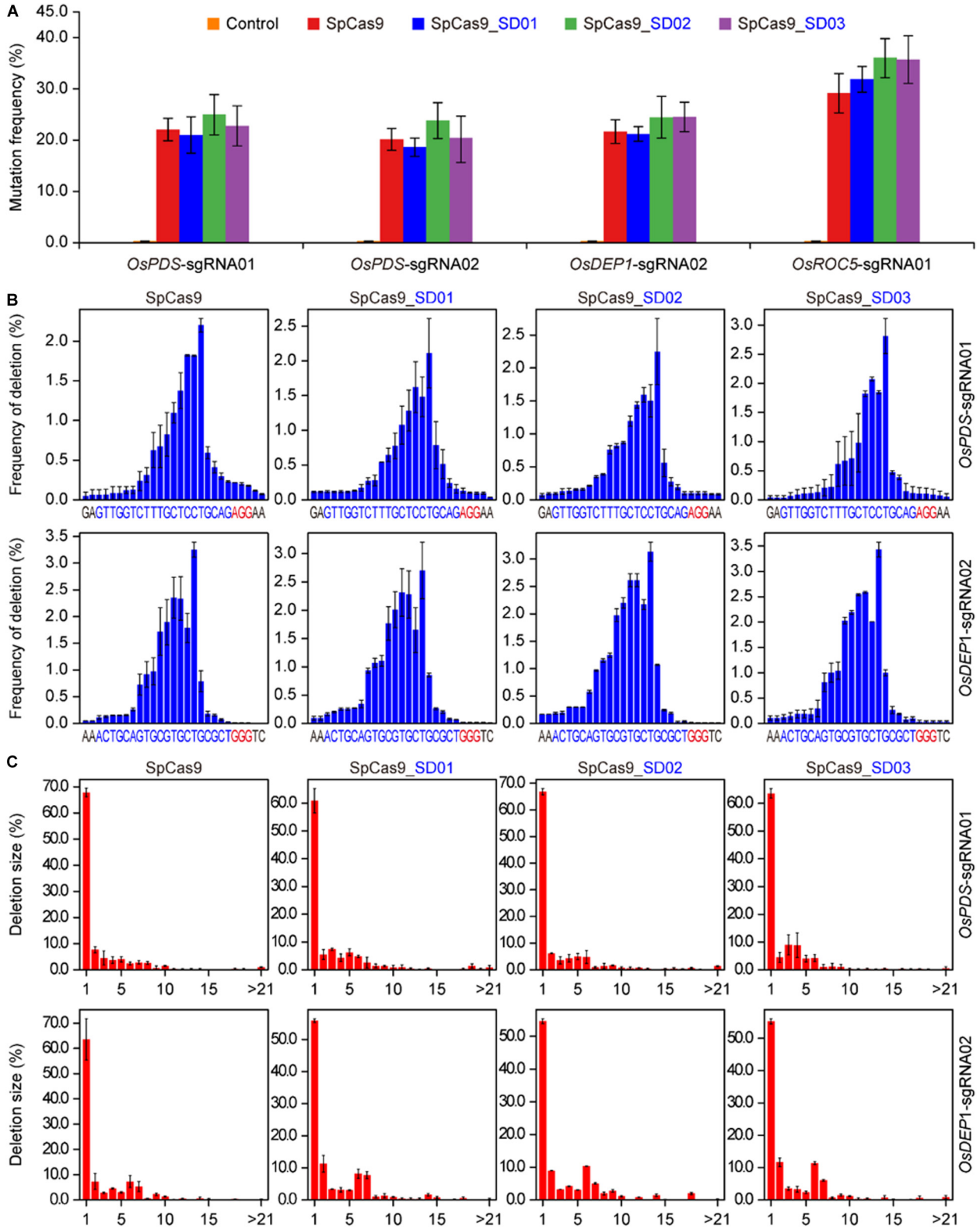


FIGURE 2 | High-throughput sequencing analysis of three Cas9-UBA systems in rice cells. **(A)** Comparison of mutation frequencies of four STU-SpCas9 systems at four different target sites. **(B)** Comparison of positional deletion frequencies at OsPDS-sgRNA01 and OsDEP1-sgRNA02 target sites. The PAM sites are highlighted in red. **(C)** Comparison of deletion of different sizes at OsPDS-sgRNA01 and OsDEP1-sgRNA02 target sites. Each plant represents the same target site, while each column represents the same SpCas9 expression strategy. Error bars represent standard deviations of three biological replicates (n = 3).

TABLE 1 | Mutation rates of four Cas9 systems with OsPDS-sgRNA01 and OsDEP1-sgRNA02 in rice transgenic plants.

Constructs	Targeted locus	Tested T0 plants*	Mutated T0 plants (number; ratio*)	Biallelic mutation plants (number; ratio**)
SpCas9	*OsPDS*-sgRNA01	16	12; 75.0%	9; 56.3%
SpCas9_SD01		15	12; 80.0%	11; 73.3%
SpCas9-SD02		18	16; 88.9%	13; 72.2%
SpCas9_SD03		16	14; 87.5%	14; 87.5%
SpCas9	*OsDEP1-sgRNA02*	16	11; 68.8%	8; 50.0%
SpCas9_SD01		17	14; 82.4%	6; 35.3%
SpCas9-SD02		17	15; 88.2%	6; 35.3%
SpCas9_SD03		18	15; 83.3%	11; 61.1%

**All the tested T0 plants were transgenic positive plants confirmed by PCR. **Mutation rates in stable transgenic T0 plants. Each plant was genotyped by Sanger sequencing of PCR amplicons.*

and all four Cas9 systems. All typical *DEP1* loss-of-function traits were observed in all biallelic mutants of the four Cas9 systems (**Figure 3D**, **Supplementary Figure S3**). But there was no significant difference among the phenotype of the STU-SpCas9 system and three STU-SpCas9 systems. These results indicated that the STU-SpCas9 systems are more efficient and robust than the STU-SpCas9 system for targeted mutagenesis in stable tranformation rice. Hence, the results from stable transgenic plants are consistent with those from rice protoplasts.

DISCUSSION

At present, CRISPR-Cas9-based genome editing has become the most widely used tool in various organisms. Numerous CRISPR-Cas9 tools and platforms have been developed, including the expanded nucleases and their variants, expression systems for multiplexing, precise genome editing with HDR or base editing, epigenome editing and transcriptome regulation (Liang et al., 2017; Lowder et al., 2018; Chen et al., 2019; Mao et al., 2019; Zhang et al., 2019; Zhou et al., 2019). Most of studies focus on increasing the editing efficiency of CRISPR-Cas9 system, especially the precise editing efficiency; expanding the genomic target ranges, achieving the sustained whole-genome editing; reducing off-target effect and improving specificity (Chen et al., 2019; Mao et al., 2019; Zhang et al., 2019). As mentioned before, improving the genome editing efficiency of CRISPR-Cas9 is mainly from two aspects: improving CRISPR-Cas9 system activity and improving the accessibility of genomic target sites. It is well known that protein expression level can be increased by enhancing the transcription and translation. Many previous studies are about how to increasing the expression level of Cas9 in cells. In this study we attempt to improve the genome editing efficiency of Cas9 from another perspective. The cleavage activity of Cas9 protein is closely related to its intracellular content, and the intracellular content of each protein is precisely regulated by cell. The balance of protein content is not only related to the expression level (input), but also to the degradation of the protein (output). It could be an effective strategy to increase the intracellular protein content by regulating the "input" and "output" of a protein at the same time. Most previous reports only focus on increasing Cas9 protein expression level (increasing input), but ignore the inhibition of Cas9 protein degradation (reducing output). Therefore, we consider to further enhance the stability of Cas9 protein and extend its half-life while maintaining high-level expression of Cas9 protein. This could increase the intracellular content of Cas9 protein, and improve the genome editing efficiency of Cas9 efficiently, stably and indiscriminately in the whole genome.

In a previous report, we demonstrated three promising STU-Cas9 systems for high-efficient plant genome editing, which were based on ribozyme, Csy4 and tRNA, respectively (Tang et al., 2016, 2019). STU-Cas9-RZ, STU-Cas9-Csy4, and STU-Cas9-tRNA systems all had similar editing efficiency in single and multiple gene target sites. Therefore, in this study, we constructed three STU-SpCas9-UBA systems (SpCas9-SD01, SpCas9-SD02, and SpCas9-SD03) based on STU-Cas9-RZ system to ensure the high-level expression of Cas9 protein (increased input). The UBA domain is a motif of about 40 amino acids, which competitively inhibited the degradation of the proteasome by binding to a substrate polyubiquitin chain (Raasi and Pickart, 2003; Raasi et al., 2005). The chimeric GFP reporter protein fused with the UBA domain at the C-terminus showed a significant increase in fluorescence levels in yeast (Heessen et al., 2005). A previous study in Arabidopsis showed that fusion of UBA1, UBA2, and UBA from AtRAD23a and AtDDI1 with endogenous transcription factors HFR1, PIF3 and intracellular jasmonic acid signaling protein JAZ10.1 significantly enhanced the stability of these proteins, and enhanced the functions of these proteins correspondingly (Jang et al., 2012). Therefore, we had chimeric three UBA domains to Cas9 protein separately to construct SpCas9-SD01, SpCas9-SD02, and SpCas9-SD03 fusion protein to enhance the stability of Cas9 and reduce its degradation at the same time (reduced output).

The original DNA sequences of UBA1, UBA2, and UBA domains were fused to Cas9 protein, respectively. The cleavage activity and genome editing efficiency of three fusion Cas9 proteins at four target sites were inspected in rice protoplasts. The results showed that the Cas9 proteins fused with the three original UBA domain sequences not only did not improve the genome editing efficiency of Cas9 protein, but decreased the editing activity at most of the target sites (data not shown). Considering the differences between rice and Arabidopsis, the rice gene expression system and protein degradation system were different from those of Arabidopsis. We had codon-optimized all three UBA domains (SD01, SD02, and SD03)

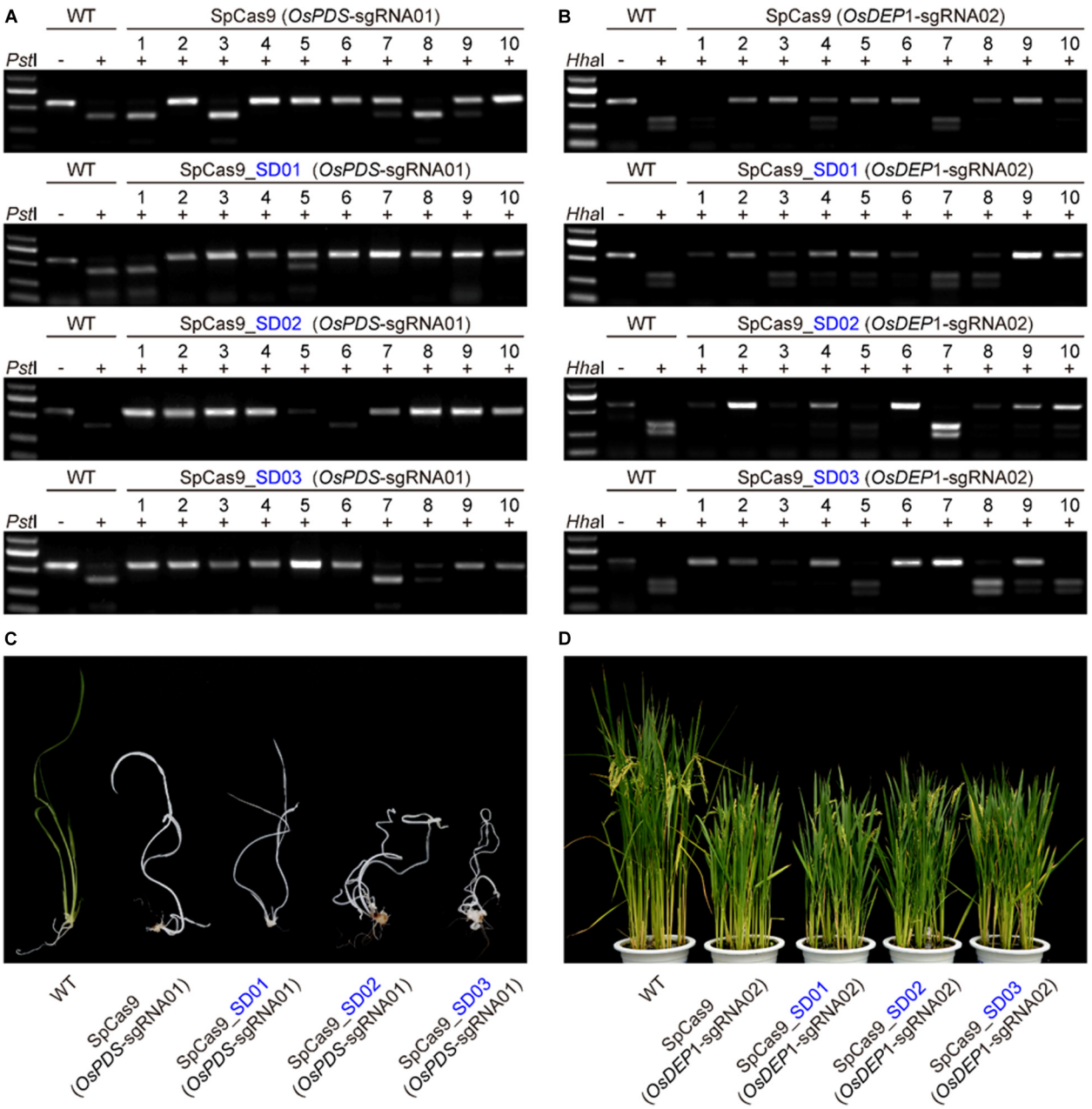


FIGURE 3 | Evaluation of three Cas9-UBA systems editing frequency with rice stable transformation. **(A)** RFLP analysis of independent T0 plants of OsPDS-sgRNA01 site in different SpCas9-SD systems. **(B)** RFLP analysis of independent T0 plants of OsDEP1-sgRNA02 site in different SpCas9-SD systems. **(C)** Phenotype of wild type and targeted mutagenesis plants in different SpCas9-SD systems at the OsPDS-sgRNA01 target site. **(D)** Phenotype of wild type and targeted mutagenesis plants in different SpCas9-SD systems at the OsDEP1-sgRNA02 target site.

according to the preference of rice codon usage. Three SpCas9-SD systems fused with codon-optimized SD domains showed similar cleavage activity in rice protoplast, which indicated that codon optimization was an effective method. The subsequent rice stable transformation results also showed that the codon-optimized SD domains successfully improved the genome editing efficiency of Cas9 protein at two endogenous gene target sites. Moreover, compared with SpCas9-SD01 and SpCas9-SD03, SpCas9-SD02 significantly improved the genome editing more efficiently at all target sites. This was consistent with previous reports that UBA2 has better stability and can more effectively extend protein half-life (Jang et al., 2012).

The UBA domain was a small motif that did not affect the function of chimeric proteins when increase the stability of targeted proteins. In our study, in addition to improving the editing efficiency, the C-terminal fused UBA domain did not

affect the cleavage function of Cas9 protein. Three endogenous genes OsPDS, OsDEP1, and OsROC5 related to rice albino, semi-dwarf and panicle erectness, and curling leaf phenotype were selected. Genotyping analysis according to high-thought sequencing and Sanger sequencing in protoplasts and stable transformed plants showed that the cleavage mode of three SpCas9-SD systems were consistent with STU-Cas9 control in all four target sites. Almost all cleavages occurred 3, 4 nt near the PAM, and 1 bp indel mutation was the most common mutation type (**Figure 2**, **Supplementary Figure S2**). Our previous work revealed that the editing mode was not related to the expression system, mainly related to the characteristics around the target site (Tang et al., 2016, 2019), which was consistent in this study. On the other hand, phenotyping analysis of mutant plants in OsPDS-sgRNA01 and OsDEP1-sgRNA02 showed that three SpCas9-SD systems were not different from the STU-Cas9 control. At OsPDS-sgRNA01, all the mutant plants obtained through four Cas9 systems showed albino phenotype. And the phenotype of mutant plants at OsDEP1-sgRNA02 were semi-dwarf, panicle erectness, consistent with previous reports (Huang et al., 2009; Zhou et al., 2009). The fusion of UBA domain improved the editing efficiency of CRISPR-Cas9 system, which was generally considered to lead to the increase of off-target effect. However, in plant species, a large-scale whole genome sequencing was performed to detect cleavage of off-target sites by Cas9 nuclease in stably transgenic Arabidopsis, rice and cotton, revealing that Cas9 activity was highly specific with very low-level off-targeting. And that potential low-level off-target effects could be avoided by designing highly specific sgRNA, or be excluded by outcrossing to different varieties which is typical during commercial seed multiplication (Feng et al., 2014; Li et al., 2018; Tang et al., 2018; Chen et al., 2019). Thus, whether the SpCas9-SD systems will result in potential very low-level off-target should be considered and evaluated in follow-up study and agricultural application.

MATERIALS AND METHODS

Plant Materials and Growth Condition

Oryza sativa L. japonica cultivar Nipponbare was used in this study. Plants were grown in soil in growth chambers at 28°C and 60% relative humidity under a long-day setting (16 h under the light and 8 h in the dark). For protoplast preparation, the sterilized seeds were placed in the 1/2 MS solid medium for 11 days in a dark chamber at 28°C. For the rice stable transformation, the sterilized seeds were placed in the N6-D solid medium to induce the callus for 7 days under light at 32°C.

Construction of the Vectors

According to the UBA1 and UBA2 amino acid sequences of Arabidopsis AtRAD23a gene (GenBank: AT1G16190) and the UBA amino acid sequence of AtDDI1 gene (GenBank: AT3G13235), UBA-SD01, UBA-SD02, and UBA-SD03 DNA fragments were synthesized by rice codon preference optimization (**Figure 1A** and **Supplementary Table S1**).

We used Cas9 expression backbone vector pGEL028 (ZmUbi1-Cas9-NLS-polyA-RZ site-ccdB-gRNA-RZ) (Tang et al., 2016, 2019) for this study. To make CRISPR-Cas9-UBA gene edit system backbone vector pGEL113, pGEL115, pGEL116, the three Cas9-UBA fusion cassettes, SpCas9-SD01, SpCas9-SD02, SpCas9-SD03 were cloned into pGEL028 using fusion PCR and ligase, all primers are listed in **Supplementary Table S1**. For pGEL113 construction, DNA fragments 1 was obtained from PCR products using primers pTX72-Cas9-F, UBA1-A from template pGEL028, DNA fragments 2 was obtained from PCR products using primers UBA1-B and UBA1-C from UBA1 domain,and DNA fragments 3 was obtained from PCR products using primers UBA1-D and pTX72-ccdB-R template pGEL028, respectively. Fragments 1 and 2 were fused to make fragment 3 using primers pTX72-Cas9-F and pTX72-ccdB-R to get fragment 4. Finally, fragment 4 was cut by *Bsa*I and then cloned into *Bsa*I-digested pGEL028. The pGEL115, pGEL116 vectors were obtained through a similar method.

Four endogenous genes target site, OsPDS-sgRNA01, OsPDS-sgRNA02, OsDEP1-sgRNA02, and OsROC5-sgRNA01 were selected to construct expression vectors based on pGEL113, pGEL115, pGEL116 and pGEL028, respectively. sgRNAs were synthesized as duplexed oligonucleotides (**Supplementary Table S1**). Oligos were annealed and cloned into *Bsa*I linearized backbone vectors. All the vectors used in this study were provided as **Supplementary Table S2**.

Rice Protoplast Transformation and Agrobacteria-Mediated Transformation

Rice protoplast transformation was performed as described previously (Zhang et al., 2013; Tang et al., 2016; Zhong et al., 2018, 2019), Each protoplast transformation experiment was performed in three biological replicates. Rice stable transformation was carried out as previously published protocol (Zheng et al., 2016; Zhong et al., 2018; Tang et al., 2019; Zhou et al., 2019).

Detection of Targeted Gene Mutations

Restriction fragment length polymorphism (RFLP) and single-strand conformational polymorphism (SSCP) methods were used for mutation detection and analysis in protoplasts and T0 transgenic plants (Tang et al., 2016; Zheng et al., 2016). The genomic DNA was used for PCR with KOD FX DNA polymerase (TOYOBO) using primers OsPDS-F1/R1 for the OsPDS-sgRNA01 target site, primers OsPDS-F2/R2 for the OsPDS-sgRNA02 target site, OsDEP1-F/R for the OsDEP1-sgRNA02 target site, OsROC5-F/R for the OsROC5-sgRNA01 target site. OsPDS-sgRNA01 target site PCR amplicons were digested with *Pst*I, OsPDS-sgRNA02 target site PCR amplicons were digested with *Hin*dIII, OsDEP1-sgRNA02 target site PCR amplicons were digested with *Hha*I and OsROC5-sgRNA01 target site PCR amplicons were digested with *Eam*1105I. Editing in regenerated mutation T0 plants was confirmed by Sanger sequencing of PCR amplicons. The genotyping and the mutation

ratio comparison in stable transgenic T0 plants were all based on sequencing results. All primers and restriction enzymes were listed in **Supplementary Table S1**.

High-Throughput Sequencing Analysis

High-throughput sequencing analysis was carried out as published previously for detection and quantification of mutations for the protoplast DNA (Tang et al., 2017; Zhong et al., 2018). Genome regions of targeted sites were PCR-amplified using primers listed in **Supplementary Table S1**. Purified DNA samples were quantified and were sequenced using Illumina Hiseq 2,500 platform. Data processing was analyzed by CRISPRMatch (You et al., 2018). The mean averages and standard deviations of three biologically independent replicates were calculated.

Measurement of Rice Yield Related Traits

The T1 generations originated from T0 mutation plants were used to test agronomic traits. The rice yield-related traits were measured according to previous method (Zhou et al., 2019). Three individual plants were used for data collection for each genotype. Each sample was tested for three times. The data were analyzed by Excel and SPSS.13 for calculation and significant differences analysis.

DATA AVAILABILITY STATEMENT

The sequencing data generated in this study can be found in NCBI using accession number PRJNA602039.

AUTHOR CONTRIBUTIONS

XZ and YZ conceived and designed the experiments. LY and XT generated all Vectors. CQ, LY, and ZZ performed rice protoplasts experiments and analyzed the HTS data. CQ, LY, BL, and QQ performed stable transgenic the rice experiments and analyzed the plants. CQ, LY, TF, and JZ performed the rice seeds analysis and other experiments. YZ and XZ analyzed the data and wrote the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

This work was supported by the National Science and Technology Major Project of China (2019ZX08010003-001-002) and National Natural Science Foundation of China (31960423, 31771486) to XZ and YZ.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00621/full#supplementary-material

FIGURE S1 | Comprehensive analysis of three SpCas9-SD systems in rice protoplasts. **(A)** Comparison of positional deletion frequencies at two target sites OsPDS-sgRNA02 and OsROC5-sgRNA01. The PAM sites are highlighted in red. **(B)** Comparison of deletion sizes at two target sites. Each plant represents the same target site, while each column represents the same SpCas9 expression strategy. Error bars represent standard deviations of three biological replicates ($n = 3$).

FIGURE S2 | Sanger sequencing reveals the genotype of mutant plants at OsPDS-sgRNA01 and OsDEP1-sgRNA02 sites in different SpCas9-SD systems. The PAM sequence is in red and the target sequence is in blue.

FIGURE S3 | Targeted mutations of OsDEP1-sgRNA02 site led to diverse yield performance in different SpCas9-SD systems. **(A)** Comparison of panicle morphology in different SpCas9-SD systems. **(B)** Comparison of seed length and width in different SpCas9-SD systems. **(C)** Comparison of plant height ($n = 7$), panicle length, grain number per panicle and 1000-grain weight among all SpCas9-SD systems. Bar graphs show average mutation frequency from more than three biologically independent replicates with error bars representing standard deviations ($n \geq 3$).

TABLE S1 | Oligos used in this study.

TABLE S2 | Constructs used in this study.

REFERENCES

Bertolaet, B. L., Clarke, D. J., Wolff, M., Watson, M. H., Henze, M., Divita, G., et al. (2001). UBA domains of DNA damage-inducible proteins interact with ubiquitin. *Nat. Struct. Biol.* 8, 417–422. doi: 10.1038/87575

Casini, A., Olivieri, M., Petris, G., Montagna, C., Reginato, G., Maule, G., et al. (2018). A highly specific SpCas9 variant is identified by in vivo screening in yeast. *Nat. Biotechnol.* 36, 265–271. doi: 10.1038/nbt.4066

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

Cermak, T., Curtin, S. J., Gil-Humanes, J., Cegan, R., Kono, T. J. Y., Konecna, E., et al. (2017). A multipurpose toolkit to enable advanced genome engineering in plants. *Plant. Cell* 29, 1196–1217. doi: 10.1105/tpc.16.00922

Chen, F., Ding, X., Feng, Y., Seebeck, T., Jiang, Y., and Davis, G. D. (2017). Targeted activation of diverse CRISPR-Cas systems for mammalian genome editing via proximal CRISPR targeting. *Nat. Commun.* 8:14958. doi: 10.1038/ncomms14958

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

Chen, L., Shinde, U., Ortolan, T. G., and Madura, K. (2001). Ubiquitin-associated (UBA) domains in Rad23 bind ubiquitin and promote inhibition of multi-ubiquitin chain assembly. *EMBO Rep.* 2, 933–938. doi: 10.1093/embo-reports/kve203

Ding, X., Seebeck, T., Feng, Y., Jiang, Y., Davis, D. G., and Chen, F. (2019). Improving CRISPR-Cas9 Genome Editing Efficiency by Fusion with Chromatin-Modulating Peptides. *CRISPR J.* 2, 51–63. doi: 10.1089/crispr.2018.0036

Farmer, L. M., Book, A. J., Lee, K. H., Lin, Y. L., Fu, H. Y., and Vierstra, R. D. (2010). The RAD23 family provides an essential connection between the 26S proteasome and ubiquitylated proteins in arabidopsis. *Plant. Cell* 22, 124–142. doi: 10.1105/tpc.109.072660

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

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. *Proc. Natl. Acad. Sci. USA*. 111, 4632–4637. doi: 10.1073/pnas.1400822111

Heessen, S., Masucci, M. G., and Dantuma, N. P. (2005). The UBA2 domain functions as an intrinsic stabilization signal that protects Rad23 from proteasomal degradation. *Mol. Cell* 18, 225–235. doi: 10.1016/j.molcel.2005.03.015

Heinen, C., Acs, K., Hoogstraten, D., and Dantuma, N. P. (2011). C-terminal UBA domains protect ubiquitin receptors by preventing initiation of protein degradation. *Nat. Commun.* 2:191. doi: 10.1038/ncomms1179

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

Hofmann, K., and Bucher, P. (1996). The UBA domain: a sequence motif present in multiple enzyme classes of the ubiquitination pathway. *Trends. Biochem. Sci.* 21, 172. doi: 10.1016/0968-0004(96)30015-7

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

Huang, X. Z., Qian, Q., Liu, Z. B., Sun, H. Y., He, S. Y., Luo, D., et al. (2009). Natural variation at the DEP1 locus enhances grain yield in rice. *Nat. Genet.* 41, 494–497. doi: 10.1038/ng.352

Jang, I. C., Niu, Q. W., Deng, S. L., Zhao, P. Z., and Chua, N. H. (2012). Enhancing protein stability with retained biological function in transgenic plants. *Plant. J.* 72, 345–354. doi: 10.1111/j.1365-313X.2012.05060.x

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

Li, J., Manghwar, H., Sun, L., Wang, P., Wang, G., Sheng, H., et al. (2018). Whole genome sequencing reveals rare off-target mutations and considerable inherent genetic or/and somaclonal variations in CRISPR/Cas9-edited cotton plants. *Plant Biotechnol. J.* 17, 858–868. doi: 10.1111/pbi.13020

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

Liang, Z., Chen, K. L., Li, T. D., Zhang, Y., Wang, Y. P., Zhao, Q., et al. (2017). Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. *Nat. Commun.* 8:14261. doi: 10.1038/ncomms14261

Liu, G., Yin, K., Zhang, Q., Gao, C., and Qiu, J. L. (2019). Modulating chromatin accessibility by transactivation and targeting proximal dsgRNAs enhances Cas9 editing efficiency in vivo. *Genome Bio.* 20:145. doi: 10.1186/s13059-019-1762-8

Liu, X., Wu, S., Xu, J., Sui, C., and Wei, J. (2017). Application of crispr/cas9 in plant biology. *Acta Pharm. Sin. B.* 7, 292–302. doi: 10.1016/j.apsb.2017.01.002

Lowder, L., Malzahn, A., and Qi, Y. (2016). Rapid evolution of manifold CRISPR systems for plant genome editing. *Front. Plant. Sci.* 7:1683. doi: 10.3389/fpls.2016.01683

Lowder, L. G., Zhou, J., Zhang, Y., Malzahn, A., Zhong, Z., Hsieh, T. F., et al. (2018). Robust transcriptional activation in plants using multiplexed CRISPR-Act2.0 and mTALE-act systems. *Mol. Plant* 11, 245–256. doi: 10.1016/j.molp.2017.11.010

Mao, Y., Botella, J. R., Liu, Y., and Zhu, J. K. (2019). Gene editing in plants: progress and challenges. *Natl. Sci. Rev.* 3, 421–437. doi: 10.1093/nsr/nwz005

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

Raasi, S., and Pickart, C. M. (2003). Rad23 ubiquitin-associated domains (UBA) inhibit 26 S proteasome-catalyzed proteolysis by sequestering lysine 48-linked polyubiquitin chains. *J. Biol. Chem.* 278, 8951–8959. doi: 10.1074/jbc.M212841200

Raasi, S., Varadan, R., Fushman, D., and Pickart, C. M. (2005). Diverse polyubiquitin interaction properties of ubiquitin-associated domains. *Nat. Struct. Mol. Biol.* 12, 708–714. doi: 10.1038/nsmb962

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

Tang, X., Liu, G., Zhou, J., Ren, Q., You, Q., Tian, L., et al. (2018). A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. *Genome Biol.* 19:84. doi: 10.1186/s13059-018-1458-5

Tang, X., Lowder, L. G., Zhang, T., Malzahn, A. A., Zheng, X. L., Voytas, D. F., et al. (2017). A CRISPR-Cpf1 system for efficient genome editing and transcriptional repression in plants. *Nat. Plants* 3:17103. doi: 10.1038/nplants.2017.103

Tang, X., Ren, Q., Yang, L., Bao, Y., Zhong, Z., He, Y., et al. (2019). 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

Tang, X., Zheng, X., Qi, Y., Zhang, D., Cheng, Y., Tang, A., et al. (2016). A single transcript CRISPR-Cas9 system for efficient genome editing in plants. *Mol. Plant.* 9, 1088–1091. doi: 10.1016/j.molp.2016.05.001

Vierstra, R. D. (2009). The ubiquitin-26S proteasome system at the nexus of plant biology. *Nat. Rev. Mol. Cell. Biol.* 10, 385–397. doi: 10.1038/nrm2688

You, Q., Zhong, Z., Ren, Q., Hassan, F., Zhang, Y., and Zhang, T. (2018). CRISPRMatch: an automatic calculation and visualization tool for high-throughput CRISPR genome-editing data analysis. *Int. J. Biol. Sci.* 14, 858–862. doi: 10.7150/ijbs.24581

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., 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

Zheng, X., Yang, S., Zhang, D., Zhong, Z., Tang, X., Deng, K., et al. (2016). Effective screen of CRISPR/Cas9-induced mutants in rice by single-strand conformation polymorphism. *Plant Cell Rep.* 35, 1545–1554. doi: 10.1007/s00299-016-1967-1

Zhong, Z., Sretenovic, S., Ren, Q., Yang, L., Bao, Y., Qi, C., et al. (2019). Improving Plant Genome Editing with High-Fidelity xCas9 and Non-canonical PAM-Targeting Cas9-NG. *Mol. Plant* 12, 1027–1036. doi: 10.1016/j.molp.2019.03.011

Zhong, Z., Zhang, Y., You, Q., Tang, X., Ren, Q., Liu, S., et al. (2018). Plant genome editing using FnCpf1 and LbCpf1 nucleases at redefined and altered PAM sites. *Mol. Plant* 11, 999–1002. doi: 10.1016/j.molp.2018.03.008

Zhou, J., Xin, X., He, Y., Chen, H., Li, Q., Tang, X., et al. (2019). Multiplex QTL editing of grain-related genes improves yield in elite rice varieties. *Plant Cell Rep.* 38, 475–485. doi: 10.1007/s00299-018-2340-3

Zhou, Y., Zhu, J., Li, Z., Yi, C., Liu, J., Zhang, H., et al. (2009). Deletion in a quantitative trait gene qPE9-1 associated with panicle erectness improves plant architecture during rice domestication. *Genetics* 183, 315–324. doi: 10.1534/genetics.109.102681

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 12 June 2020
doi: 10.3389/fpls.2020.00575

Precise CRISPR-Cas9 Mediated Genome Editing in Super Basmati Rice for Resistance Against Bacterial Blight by Targeting the Major Susceptibility Gene

*Kashaf Zafar[1,2], Muhammad Zuhaib Khan[1], Imran Amin[1], Zahid Mukhtar[1], Sumera Yasmin[3], Muhammad Arif[1], Khansa Ejaz[3] and Shahid Mansoor[1]**

[1] Agricultural Biotechnology Division, National Institute for Biotechnology and Genetic Engineering, Constituent College of Pakistan Institute of Engineering and Applied Sciences, Faisalabad, Pakistan, [2] Department of Biotechnology, Balochistan University of Information Technology, Engineering and Management Sciences, Quetta, Pakistan, [3] Soil and Environmental Biotechnology Division, National Institute for Biotechnology and Genetic Engineering, Constituent College of Pakistan Institute of Engineering and Applied Sciences, Faisalabad, Pakistan

OPEN ACCESS

Edited by:
Vladimir Nekrasov, Rothamsted Research, United Kingdom

Reviewed by:
Bing Yang, University of Missouri, United States
Jana Streubel, Leibniz University Hannover, Germany

***Correspondence:**
Shahid Mansoor
shahidmansoor7@gmail.com

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *03 February 2020*
Accepted: *17 April 2020*
Published: *12 June 2020*

Citation:
Zafar K, Khan MZ, Amin I, Mukhtar Z, Yasmin S, Arif M, Ejaz K and Mansoor S (2020) Precise CRISPR-Cas9 Mediated Genome Editing in Super Basmati Rice for Resistance Against Bacterial Blight by Targeting the Major Susceptibility Gene. Front. Plant Sci. 11:575. doi: 10.3389/fpls.2020.00575

Basmati rice is famous around the world for its flavor, aroma, and long grain. Its demand is increasing worldwide, especially in Asia. However, its production is threatened by various problems faced in the fields, resulting in major crop losses. One of the major problems is bacterial blight caused by *Xanthomonas oryzae pv. oryzae (Xoo)*. *Xoo* hijacks the host machinery by activating the susceptibility genes (*OsSWEET* family genes), using its endogenous transcription activator like effectors (TALEs). TALEs have effector binding elements (EBEs) in the promoter region of the *OsSWEET* genes. Out of six well-known TALEs found to have EBEs in Clade III *SWEET* genes, four are present in *OsSWEET14* gene's promoter region. Thus, targeting the promoter of *OsSWEET14* is very important for creating broad-spectrum resistance. To engineer resistance against bacterial blight, we established CRISPR-Cas9 mediated genome editing in Super Basmati rice by targeting 4 EBEs present in the promoter of *OsSWEET14*. We were able to obtain four different Super Basmati lines (SB-E1, SB-E2, SB-E3, and SB-E4) having edited EBEs of three TALEs (*AvrXa7*, *PthXo3,* and *TalF*). The edited lines were then evaluated in triplicate for resistance against bacterial blight by choosing one of the locally isolated virulent *Xoo* strains with *AvrXa7* and infecting Super Basmati. The lines with deletions in EBE of *AvrXa7* showed resistance against the *Xoo* strain. Thus, it was confirmed that edited EBEs provide resistance against their respective TALEs present in *Xoo* strains. In this study up to 9% editing efficiency was obtained. Our findings showed that CRISPR-Cas9 can be harnessed to generate resistance against bacterial blight in indigenous varieties, against locally prevalent *Xoo* strains.

Keywords: genome editing, CRISPR-Cas9, rice improvement, bacterial blight, *Xanthomonas oryzae*

INTRODUCTION

In plants, different genome editing strategies have been exploited including zinc finger nucleases, transcription activator-like effector nucleases (TALEs), and clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR associated (Cas) nucleases. Among these tools, the RNA guided CRISPR-Cas9 system has become the method of choice for genome editing because of its simplicity, ease of performing, and versatility. This system exploits the complementary base pairing mechanism of DNA to guide site-specific Cas9 endonuclease to the target site. The guide RNA (gRNA) screens the template, recognizes the specific complementary target sequence, and signals to Cas9 to introduce a double stranded break (DSB) at the target site. A triplet of nucleotides (NGG) at the 3′ end of the target site, also known as protospacer-associated motif (PAM), is essential for Cas9 to introduce a double stranded break (DSB) 3 bp upstream of the PAM sequence. These DSBs are repaired either via imprecise non-homologous end joining (NHEJ) or template directed precise homology directed repair (HDR) (Barrangou et al., 2007; Cong et al., 2013; Mali et al., 2013; Hsu et al., 2014; Wright et al., 2016). Until today, this technology has been successfully used to engineer resistance against various pathogens and for agronomic trait enhancement.

Rice is a staple food crop which belongs to the family Poaceae, genus *Oryza*. It has been cultivated for more than 10,000 years and is the second most commonly cultivated cereal in the world (Sasaki, 2001; Pazuki and Sohani, 2013). For Asian countries, it contributes 50–80% of daily calories and serves more than 90% of the population (Khush, 2005; Zeigler and Barclay, 2008). Among different varieties of rice, Basmati rice is famous around the globe for its flavor, aroma, and long grain. Due to food insecurity, rice demand is increasing every day (Shobarani et al., 2010). However, Basmati rice production is threatened by various problems in the field resulting in major crop losses. One of the grave problems is bacterial blight caused by *Xanthomonas oryzae pv. Oryzae (Xoo)*. It is the most destructive and deadly bacterial disease and can cause up to 75% crop loss. *Xoo* hijacks the host machinery by activating the susceptibility genes (*SWEET* family genes), using its endogenous transcription activator like effectors (TALE). TALEs have their effector binding elements (EBEs) in the promoter region of *OsSWEET* genes. These effectors divert sugars from the plant cell to fulfill the pathogen's nutritional needs (Chen et al., 2012). Most of the geographically distinct *Xoo* strains target *OsSWEET14,* which encodes the sucrose-efflux transporter family (Chen et al., 2010). There are six known TALEs which target promoter regions of *OsSWEET* genes, and the EBEs for four different TALEs, *PthXo3, AvrXa7*, *TalC,* and *TalF* (previously known as Tal5) are present in *OsSWEET14* (Antony et al., 2010; Yu et al., 2011; Streubel et al., 2013; Oliva et al., 2019) (**Figure 1A**).

Rice has been used extensively as a model for performing genome editing studies via CRISPR-Cas9 due to its diploid nature (Jiang et al., 2013; Hu et al., 2016; Abe et al., 2018). There are some prior studies where genome editing has been done to develop resistance against bacterial blight using different genome editing platforms including CRISPR-Cas9 (Blanvillain-Baufumé et al., 2017; Oliva et al., 2019; Xu et al., 2019). Customization of the CRISPR-Cas9 tool offers an avenue to target and mutate TALE binding elements in the promoter region of *OsSWEET14* to resist the spread of bacterial blight (Jiang et al., 2013). Different types of natural mutations have been reported in *OsSWEET14* genes in different rice cultivars which provide immunity against *Xoo* strains (Hutin et al., 2015b). However, no natural mutations have been reported in Super Basmati rice (Zaka et al., 2018). Here, we have employed CRISPR-Cas9 technology to engineer resistance against bacterial blight in Super Basmati rice for the first time by targeting the promoter region of the *OsSWEET14* gene (**Figure 1B**).

MATERIALS AND METHODS

gRNAs Design and Construct Development

The genome sequence of Super Basmati rice was not available in nucleotide databases. To know the exact sequence of the *OsSWEET14* gene promoter and to design gRNAs, different primer sets (*Os*P-F: 5′ ATTGGCACTTTCTGTCATGCATG 3′ and *Os*P-R: 5′ GCAAGATCTTGATTAACTAGCTAGC 3′) were first designed on the *OsSWEET14* gene promoter based on the sequences of other rice varieties available in the database. The genomic DNA of Super Basmati rice was extracted using the cetyl trimethylammonium bromide (CTAB) method (Stewart, 1993) followed by polymerase chain reaction (PCR). An amplicon of 457 bp in length was cloned into a pTZ57 R/T (Thermo Fisher Scientific, United States) vector and sent for sequencing. All gRNAs were designed manually to target their respective EBEs (**Table 1**). The overhangs of GGCA and AAAC were added a 5′ end of forward and reverse gRNA respectively for cloning at the *Bsa*I site (**Table 2**). The gRNAs were screened for potential off-targets using the online tool Cas-OFFinder[1]. Analysis revealed that there were no off-targets. The gRNAs were synthesized, annealed, and cloned at the *Bsa*I site in pRGEB32 (Addgene plasmid # 63142) under the rice *U3* promoter (Xie et al., 2015). The pRGEB32 already has rice codon optimized Cas9 expressed under the *Ubi* promoter. The constructs with the Cas9 gene and gRNA cassette (shown in **Figure 2**) were confirmed by PCR and restriction followed by Sanger sequencing. The Cas9 gene was confirmed by PCR using specific primers (Cas9-F 5′ AGCATCGGCCTGGACATCGGC 3′ and Cas9 R-5′ CCGGAACTTGATCATGTGGG 3′). The full length Cas9 was also confirmed in constructs using primer set 3-full Ca9 (**Supplementary Table S1**).

Plant Material

Super Basmati (Pakistan's indigenous rice variety) was used to establish genome editing against bacterial blight. The rice variety IR24 was used as a susceptible control. Rice seeds were obtained from DNA Markers and the Applied Genomics Lab of National Institute for Biotechnology and Genetic Engineering (NIBGE), Faisalabad, Pakistan. The seeds were manually de-husked, and surface sterilized with 70% ethanol followed by dipping in 50%

[1]http://www.rgenome.net/cas-offinder/

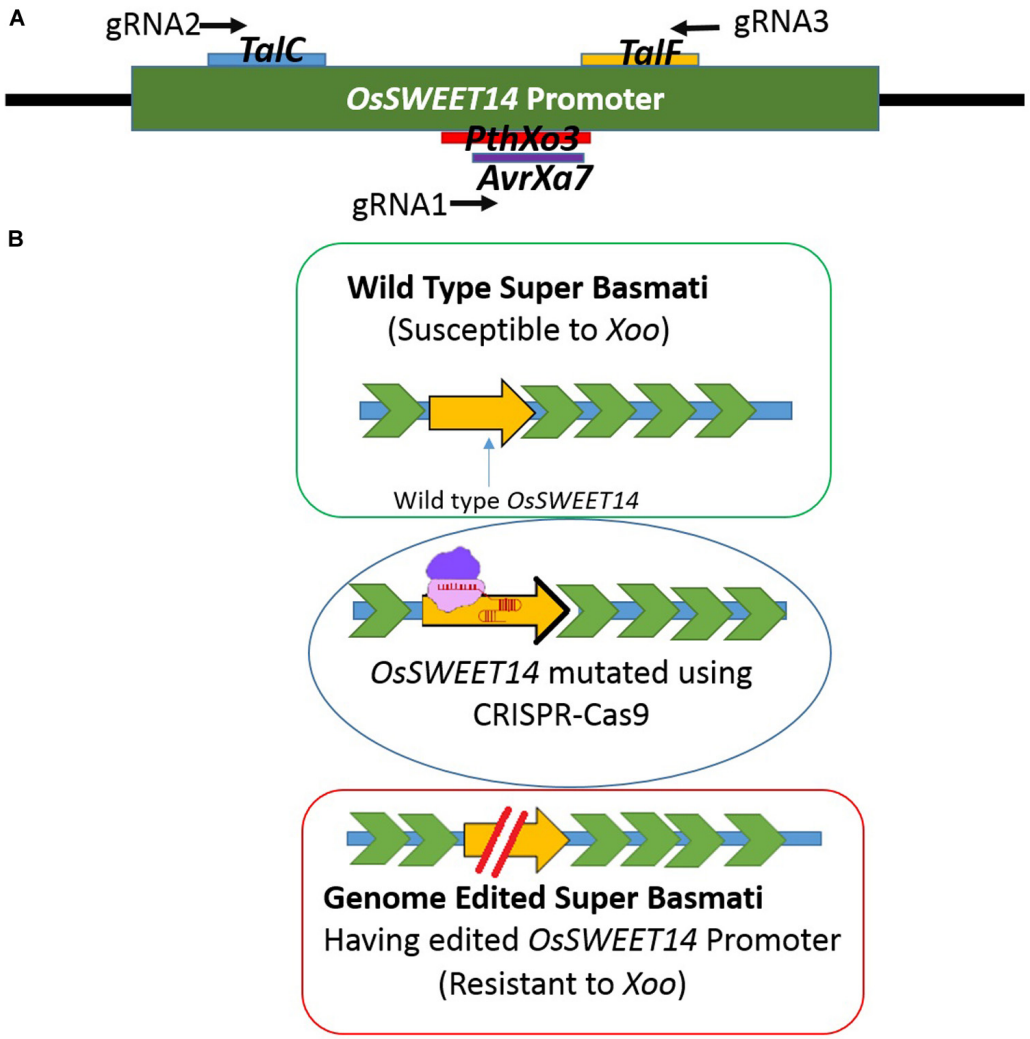


FIGURE 1 | (A) Schematic figure of the promoter showing EBEs location on the promoter against which three gRNAs were designed. gRNA1 was designed to target *PthXo3* and *AvrXa7*, gRNA2 for *TalC* and gRNA3 targeted *TalF*. **(B)** Schematic figure of developing resistance by editing the *OsSWEET14* gene. The intact promoter was susceptible to *Xoo* whereas the EBE edited promoter can be resistant to *Xoo*.

(v/v) commercial bleach (having 5.25% sodium hypochlorite). After washing with distilled water, seeds were placed on MS media (Murashige and Skoog, 1962) supplemented with vitamins and 2,4-Dichlorophenoxyacetic acid (2,4-D) for callus formation.

Rice Transformation and Growth Conditions

The 28 days old embryogenic calli were selected and placed on an osmotic medium with 0.25 M mannitol. The constructs with gRNA and Cas9 were confirmed again and coated on 1 μm gold particles (Bio-Rad, United States) for biolistic transformation. The rice calli were transformed through bombardment by a gene gun (PDS-1000/Bio-Rad, United States) following optimized protocol for biolistic transformation of Super Basmati rice (Mukhtar and Hasnain, 2018). For each gRNA approximately 4,000 calli were transformed. The transformed calli were placed at 28–30°C for 24 h in a dark room under sterile conditions. After 24 h, the transformed calli were shifted to the selection medium with 50 mg/L hygromycin. After 15 days they were again shifted to fresh selection medium. After completing 30 days on the selection medium, the surviving calli were shifted to pre-regeneration and then to regeneration media. The regenerated plantlets were shifted to rooting media. After the development of roots, they were transferred to soil and kept in the greenhouse under controlled conditions. After 15 days the plants were shifted to large pots and allowed to grow to maturity and the seeds were then harvested.

Confirmation of Transgenic Rice

To confirm the presence of Cas9 and gRNA cassette, DNA was extracted from all plants that were developed through tissue culture, using the CTAB method (Stewart, 1993). The

TABLE 1 | Sequences of EBEs present in the *OsSWEET14* promoter region.

TALE EBE	Sequence	References
AvrXa7	ATAAACCCCCTCCAACCAGGTGCTAA	Antony et al., 2010
PthXo3	ATATAAACCCCCTCCAACCAGGTGCTAAG	Antony et al., 2010
TalF	TAAGCTCATCAAGCCTTCA	Streubel et al., 2013
TalC	CATGCATGTCAGCAGCTGGTCAT	Yu et al., 2011

TABLE 2 | Sequences of gRNAs to target EBEs of the *OsSWEET14* promoter region.

gRNA	Sequence	Target
*Os*PgRNA1	F: 5′-GGCA G ATATAAACCCCCTCCAACC-3′ R: 5′-AAAC GGTTGGAGGGGGTTTATAT C-3′	*AvrXa7, PthXo3*
*Os*PgRNA2	F: 5′-GGCA GGGCATGCATGTCAGCAGC-3′ R: 5′-AAAC GCTGCTGACATGCATGCCC-3′	*TalC*
*Os*PgRNA3	F: 5′-GGCA G TGAGTTTGCTTTGCTTGAA-3′ R: 5′-AAAC TTCAAGCAAAGCAAACTCA C-3′	*TalF* (Previously *Tal5*)

Red bases show the overhangs added to make gRNAs compatible for cloning in pRGEB32.

construct was confirmed by PCR using forward primer at vector backbone (OsPRGEB32 F 5′ GGTGCTACCAGCAAATGC TGGAAGCCG3′) and reverse primer designed on gRNA (*Os*PgRNA1-R: 5′-AAACGGTTGGAGGGGGTTTATATC-3′, *Os*PgRNA2-R: 5′-AAAC GCTGCTGACATGCATGCCC-3′ and *Os*PgRNA3-R: 5′-AAAC TTCAAGCAAAGCAAACTCAC-3′ for gRNA 1, 2, and 3 constructs, respectively). The 273 bp amplicon confirmed the presence of construct (**Supplementary Figure S2**). Gene specific primers were used to confirm the presence of Cas9 (**Table 1**- primer set 2 (partial Cas9) and 3 (full Cas9). Promoter fragments of the *OsSWEET14* gene were amplified using primers (**Supplementary Table S1**- primer set 1) from all the plants and PCR products were sent for Sanger sequencing.

T7 Endonuclease Assay

For the T7 endonuclease assay, the genomic DNA was extracted from T_0 plants and the flanking target region was amplified using primer set 1 (**Supplementary Table S1**- *Os*P-F and *Os*P-R) to amplify the promoter (**Supplementary Table S1**). The purified PCR amplicons were denatured and renatured then subjected to T7 endonuclease I (NEB, M0302) following the manufacturer's instructions. The products were resolved on 2% agarose gel and then stained with ethidium bromide.

Screening Against Bacterial Blight Resistance

The screening of edited lines against bacterial blight was performed under containment glasshouse conditions. The 1st round of screening was performed on T_0 plants. Two plants from each line were selected for inoculation. T_1 seeds were then collected from primary transformants. Seeds from each line were sown in separate pots containing soil (Excluding SB-E1 because seed filling was disturbed for this line). When the plants reached the three leaf stage, they were arranged in different batches for screening against bacterial blight. The experiment was conducted in three batches. There were six pots in each batch. Each pot had three plants of wild type Super Basmati, SB-E2, SB-E3, SB-E4, IR-4 (susceptible check), and a negative control. So, the edited plants were arranged along with their negative, susceptible, and wild type controls. There were three pots for each line, so a total of nine plants were inoculated for each line. The wild type Super Basmati has no editing, and the IR-24 line is susceptible to almost all kinds of *Xoo* strains, and was therefore used as a susceptible control. The negative control is the plant for which scissors were dipped in distilled water, without any *Xoo* strain, to check the lesion length introduced due to scissor injury. All of these lines were screened for bacterial blight resistance using a local virulent *Xoo* strain. This strain was selected on the basis that one of the four EBEs of *OsSWEET14* is *AvrXa7* which was present in the strain. The inoculum was prepared as described by Tu et al. (2000). All the lines were maintained in a uniform environment in the glasshouse at 30°C and 85% relative humidity. For leaf clip inoculation, 45 days old plants were used. For each plant three to four leaves were uniformly inoculated with scissors dipped in bacterial suspension (Kauffman, 1973). At 14 days post-inoculation (dpi) the plant's responses were observed and data was recorded by measuring the full length of the leaf and lesion length in centimeters (cm). The percentage disease leaf area (%DLA) was then calculated using the below formula.

$$\%\ \text{DLA (cm)} = [\text{Lesion Length (cm)}/\ \text{Full leaf length (cm)}]^{*}\ 100$$

RNA Extraction, cDNA Synthesis and RT-qPCR

The 30 days old rice seedlings were inoculated with scissors dipped in *Xoo* suspension. The total RNA was extracted after 24 h post-inoculation (hpi) from rice leaves using TRIzol reagent following the manufacturer's instructions (Invitrogen, United States). The isolated RNA was treated with DNaseI as per the manufacturer's instructions (Thermo Fisher Scientific, United States). The complementary DNA (cDNA) was synthesized from RNA using a RevertAid first strand cDNA synthesis kit (Thermo Fisher Scientific, United States) following instructions given by the manufacturer. The induction of *OsSWEET14* by *Xoo* was determined in wild type as well as in edited plants by RT-qPCR. The *OsSWEET14* transcripts were amplified using *OsSWEET14*-specific primers OsSWEET14-RT-F/OsSWEET14-RT-*R* (**Supplementary Table S1**). A volume of 25 μL reaction mixture was made using 12.5 μL SYBR Green Real-Time PCR Master Mix (Thermo Fisher Scientific, United States), 0.1 pmole of forward and reverse primers, 2.5 μL cDNA (∼25 ng), and 9.5 μL water. The conditions were optimized and finally the reaction was performed using a Bio-Rad iQ5 thermal cycler (Bio-Rad, United States). The expression of sucrose phosphate synthase (SPS) was used as an internal control and was amplified using primers SPS-F/SPS-R (**Supplementary Table S1**). The quantification results were analyzed using the $2^{-\Delta\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). Each reaction was performed in triplicate.

FIGURE 2 | Schematic diagram of the construct used for genome editing. All three gRNAs were cloned in the same way. The construct was expressing gRNA under the *OsU3* promoter. The rice codon-optimized Cas9 was expressing under the *Ubi* promoter.

Phenotyping

The plants were observed visually during growth. After harvesting the seeds from edited plants, the germination of the seeds was checked. The seeds were dehusked, sterilized and placed in six-well plates in water and the rate of germination was recorded. This was done for all the lines except SB-E1 because we were unable to obtain seeds from this line. The root and shoot lengths were measured by growing three to four seeds from each line vertically on square plates with ½ MS media (no sucrose was added). After 1 week, the root and shoot lengths were again measured and compared with the wild type.

RESULTS

Development of Constructs

The sequence of the *OsSWEET14* gene promoter was amplified using the primer set *Os*P-F/*Os*P-R (**Supplementary Table S1**). This primer set was designed on the sequences of other rice varieties available in the database. Upon amplification a PCR, a product of 457 bp, was amplified which was cloned and sequenced. The sequenced region of the *OsSWEET14* promoter from Super Basmati rice was submitted to Genbank and is available online under Accession No. MK791135.1. Three different gRNAs (**Table 2**) were designed to target EBEs (**Table 1**) of *AvrXa7*, *PthXo3*, *TalF,* and *TalC* in the promoter region. Due to overlapping EBEs for *AvrXa7* and *PthXo3*, the first gRNA (gRNA1) was designed to target both EBEs simultaneously. Similarly, gRNA2 and gRNA3 were designed to target *TalC* and *TalF*, EBEs, respectively.

The constructs containing the Cas9 gene and gRNA cassette were initially confirmed by PCR. Gene specific primers were used to confirm the Cas9 gene with an amplicon length of 490 bp (**Supplementary Figure S3A**). In the final constructs, full-length Cas9 (4100 bp) was confirmed (**Supplementary Figure S3B**). The gRNA cassette was confirmed using the forward primer (OsPRGEB32 F GGTGCTACCAGCAAATGCTGGAAGCCG) and reverse primer of the respective gRNA (**Table 2**). The 273 bp product confirmed the presence of gRNA along with its scaffold in the pRGEB32 vector (**Supplementary Figure S1A**). The presence of gRNAs in constructs was also confirmed by restriction analysis. gRNA insertion in pRGEB32 results in disruption of the *Bsa*I site. Due to this disruption, restriction with *Bsa*I and *Hind*III enzymes did not release 400 bp fragments in positive clones. Thus, the absence of 400 bp confirmed that all the gRNAs were successfully cloned (**Supplementary Figures S1B,C**). Finally, all constructs were confirmed by Sanger sequencing (**Supplementary Figure S1D**).

Transgenic Plants Development

The process of transgenic plant development and screening against bacterial blight is shown in **Supplementary Figures S4B,C**. Cas9 expressing transgenic rice lines were developed under the *Ubi* promoter. All the three gRNAs (designated as gRNA1, gRNA2, and gRNA3) were expressed under rice U3 promoter in a pRGEB32 vector with rice codon-optimized Cas9. A total of 48 transgenic lines were obtained for gRNA1, 35 lines for gRNA2, and 51 lines for gRNA3. Both edited and wild type lines were maintained at 30°C in a containment glasshouse.

Mutation Detection

To determine whether CRISPR-Cas9 was able to cleave and generate DSBs at the target site and repaired it via NHEJ, total genomic DNA was isolated using the CTAB method. Target regions were amplified using PCR and subjected to T7 Endonuclease I (T7EI) mutation detection analysis which degrades single-stranded regions of non-complementarity resulting from indels; these regions are then detected by sequencing. The 457 bp target region from T_0 was digested with T7E1 and the expected bands (200 and 250 bp approximately) were observed in all five edited lines (**Figure 3C**). The wild type plants did not show these bands and thus confirmed editing in five lines. Moreover, to further confirm the editing, a 457 bp promoter fragment of the *OsSWEET14* gene from all transgenic plants was sent for Sanger sequencing. We were successful in obtaining two transgenic lines with a mutated promoter fragment for gRNA1 (targeting *AvrXa7* and *PthXo3*). Out of these two lines, one line had 24 bp deletion while the second line showed 4 bp deletion at the target site (**Figure 3A**). We were unable to obtain transgenic lines for gRNA2. All those lines that were edited with gRNA2 had intact EBE of *TalC*. Out of 51 transgenic lines for gRNA3, we were able to obtain three lines with 4, 5, and 18 bp deletions at the target site. Thus, in two of the three lines (having 4 and 18 bp deletions), we successfully disrupted the EBE of *TalF*. In the third line, 5 bp deletion was not able to disrupt the EBE of *TalF* (**Figure 3B**). These results showed that CRISPR-Cas9 was able to do editing at the target site while no such activity was observed in the control plants. The nature of mutations was also observed in edited lines. We obtained only one biallelic mutation in SB-E1 while the rest of the mutations were mono-allelic (**Supplementary Table S2**). However, the edits were transmitted successfully to their progeny.

Bacterial Blight Resistance Assays

The edited lines were challenged with a prevalent *Xoo* strain via the leaf clip inoculation method. The strain has *AvrXa7*

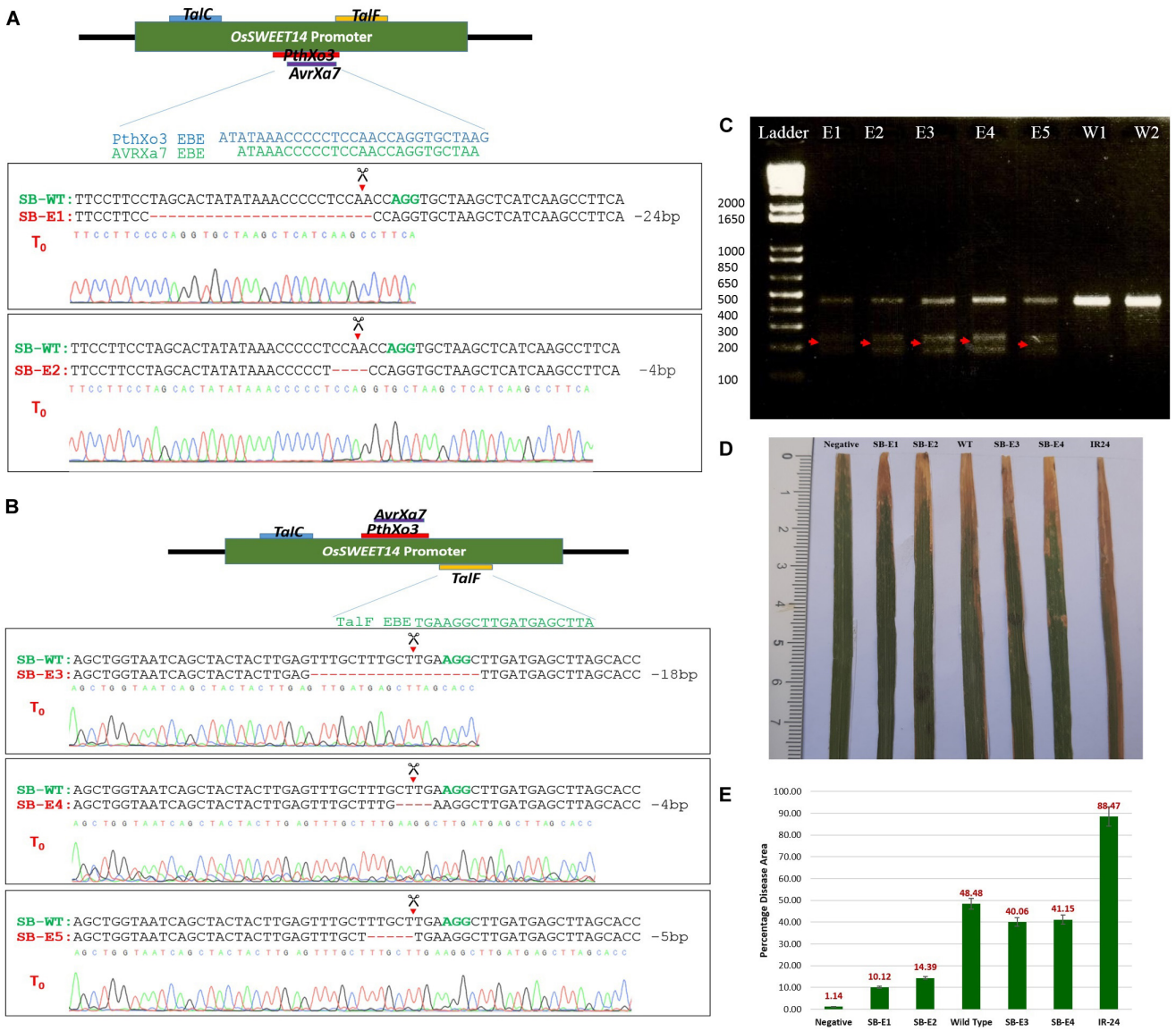


FIGURE 3 | (A) Sanger sequencing of *AvrXa7* and *PthXo3* EBE edited plants. The plants were named as SB-WT (Super Basmati wild type), SB-E1 (Super Basmati edited 1) and SB-E2 (Super Basmati edited 2). Green colored bases AGG represents the PAM sequence and red dashes symbolize the deletion at the target site. SB-E1 and SB-E2 have 24 and 4 bp deletion at the target site, respectively. **(B)** Sanger sequencing of *TalF* EBE edited plants. The plants were named SB-WT (Super Basmati wild type), SB-E3 (Super Basmati edited 3), SB-E4 (Super Basmati edited 4), and SB-E5 (Super Basmati edited 5). SB-E3 and SB-E4 have 18 and 4 bp deletion at the target site, respectively, disrupting *TalF* EBE, whereas SB-E5 had 5 bp deletion at the target site but *TalF* EBE remained intact. **(C)** T7 Endonuclease assay of T_0 edited and wild type plants. Red arrows showed the bands from edited plants that are absent in wild type plants (E1 = SB-E1, E2 = SB-E2, E3 = SB-E3, E4 = SB-E4, E5 = SB-E5, W1 and W2 = Wild Type Super Basmati). **(D)** Lesions induced by *Xoo* after 14 days post-inoculation on edited and control plants. SB-E1 and SB-E2 has reduced lesion length as compared to control plants. Negative = Plant was inoculated with scissors dipped in water. All the other lines were inoculated with scissors dipped in *Xoo* suspension. **(E)** Average of Percentage disease leaf area (% DLA) of the edited and control plants. The lowest DLA 10.12% was observed in the case of SB-E1 plants where 24 bp deletion was detected. The highest DLA 88.47% was present in IR-24 which was used as a susceptible check.

TALE. The presence of *AvrXa7* was not only confirmed by sequencing but the same strain was screened against a rice cultivar (IRBB7) which showed resistance against this strain. IRBB7 has Xa7 gene, which provides resistance against *Xoo* isolates with TALE *AvrXa7* (Yang et al., 2000). This strain was tested to induce the *OsSWEET14* gene. It was able to induce *OsSWEET14*. Therefore, this strain was selected for inoculating edited lines. The wild type Super Basmati rice was used as the control and IR24 (IRRI Line) as a susceptible check. All the inoculated lines were kept at 30°C in a greenhouse with 85% humidity. Two plants from the control and T_0 edited lines, whereas nine plants from control and

T_1 edited lines were inoculated. The percentage disease leaf area (%DLA) for three inoculated leaves from each plant was calculated (**Figure 3E** and **Supplementary Tables S3, S4**). After 14 days post-inoculation, plants were observed for bacterial blight infection. There were significant differences in the rate of infection among edited and control lines. The disease area among edited and control lines were compared in terms of %DLA covered by infection. In *AvrXa7*/*PthXo3* edited lines, the incidence of infection was very low as compared to the wild type Super Basmati and IR24 lines (**Figure 3D**). The lowest DLA (10.12%) covered by infection was seen in the case of Super Basmati edited line 1 (SB-E1) in which 24 bp deletion in overlapping EBEs of *AvrXa7* and *PthXo3* occurred. The other variant, SB-E2, showed 14.4% DLA, which was also lower than the control plants. SB-E3 and SB-E4, with disrupted *TalF* EBE and intact *AvrXa7*/*PthXo3* EBE, showed 40 and 41% DLA, respectively, when inoculated with the same strain (possessing only *AvrXa7*). 48 and 88% disease incidence were observed for wild type Super Basmati and IR24, respectively. This confirmed our hypothesis that editing EBEs in the promoter of the *OsSWEET* genes provides resistance against corresponding TALEs.

Relative *OsSWEET14* Induction by *Xoo* Strain

Xoo strains encode different TALEs to activate endogenous *SWEET* genes for a successful infection. The *OsSWEET14* activation by the *Xoo* strain in wild type Super Basmati Rice and edited lines were determined using real-time quantitative PCR. Expression of *OsSWEET14* was induced by inoculating both wild type and edited lines by the *Xoo* strain. Five lines [(1) Non-inoculated wild type Super Basmati, (2) Inoculated wild type Super Basmati, (3) Inoculated edited line SB-E2, (4) Inoculated SB-E3, and (5) Inoculated SB-E4] were selected to compare *OsSWEET14* activation by *Xoo*. Upon infiltration with the *Xoo* strain, the expression of *OsSWEET14* was considerably high in WT-SB as compared to non-inoculated WT-SB (**Figure 4**). The edited line (SB-E2) carrying a mutation for *AvrXa7,* showed very low expression of *OsSWEET14.* Whereas in the SB-E3 and SB-E4 lines (carrying edited *TalF* but intact *AvrXa7)*, the strain was still able to induce *OsSWEET14*. This clearly indicates that the *Xoo* strain was able to induce *OsSWEET14* expression to establish successful infection in Super Basmati rice. In addition, the induction of *OsSWEET14* expression in SB-E3 and SB-E4 edited lines clearly showed that the *Xoo* strain contained *AvrXa7* TALE to infect Super Basmati rice. As SB-E2 edited line has mutated EBE for *AvrXa7* and the strain was unable to induce *OsSWEET14* expression. As a result, relative expression of *OsSWEET14* was significantly lower in the SB-E2 edited line as compared to other inoculated lines.

Phenotyping

The plants were carefully observed for any phenotypic changes. All the edited plants showed a normal phenotype and were fertile just like the wild type plants. The harvested seeds also showed a normal phenotype. The germination was checked for seeds from all the edited lines (Except SB-E1) and seeds showed normal germination (**Supplementary Figure S5**). The shoot and root lengths were also normal for the edited lines (**Figure 5**).

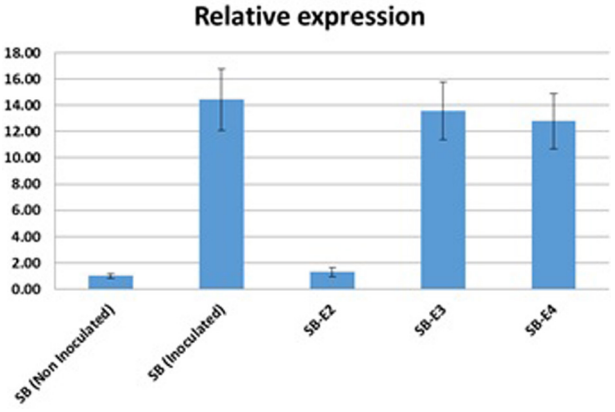


FIGURE 4 | *OsSWEET14* Induction by *Xoo* strain: Relative mRNA levels ($2^{-\Delta\Delta Ct}$) of *OsSWEET14* in leaves of wild type and edited rice lines were compared. qRT-PCR was conducted in wild type Super Basmati and edited lines (SB-E2, SB-E3, SB-E4). Samples were harvested after 24 h of Inoculation with a locally virulent *Xoo* strain (mean ± s.e.m., $n = 3$ leaf samples from biological replicates) with expression normalized to rice *SPS* levels; repeated independently three times with comparable results.

FIGURE 5 | Root length and shoot length of wild type and edited plants. Edited plants show normal root and shoot length just like wild type plants (two representative plants are shown in figure out of three).

DISCUSSION

In a global scenario, bacterial blight is a destructive disease of rice, and in Asia specifically, it severely damages crop yield. Traditionally, various approaches have been used to develop

resistance in rice cultivars which include the introgression of resistance genes (R genes), host-derived resistance, and loss of susceptibility through mutation in recessive R genes (Dou and Zhou, 2012; Kim et al., 2015; Busungu et al., 2016; Li et al., 2019). Despite the effectiveness of R genes, the constant introduction of R-genes in rice breeding programs may result in the emergence of new *Xoo* strains, which can overcome R gene-mediated resistance (Ji et al., 2016). With the advent of new technologies like genome editing, various strategies have been employed, e.g., TALENs and CRISPR-Cas9 to develop disease resistance (Khan et al., 2014). Here in this study we used CRISPR-Cas9 technology to develop disease resistance in Basmati rice.

One approach which can be used to remove disease predisposition is to mutate the susceptible portion of the genome (Iyer-Pascuzzi and McCouch, 2007; Hutin et al., 2015a; Kourelis and van der Hoorn, 2018; Zaidi et al., 2018). The idea to mutate the bacterial blight susceptible region of the rice genome sparked from the presence of three naturally occurring recessive R genes, i.e., *xa13,* x*a25,* and *xa41 (t).* These are mutated alleles of *OsSWEET11*, *OsSWEET13,* and *OsSWEET14,* respectively, which offers resistance against their respective TALEs present in the *Xoo* strain (Chu et al., 2006; Yang et al., 2006; Yuan et al., 2009; Hutin et al., 2015b). It has been reported that *Xa41*(t) (having 18 bp deletion) offers broad-spectrum resistance against a large collection of *Xoo* strains (Hutin et al., 2015b). Recently it was reported that some *Xoo* strains still cause infection in rice despite R genes (Carpenter et al., 2018; Doucouré et al., 2018). These failures have prompted plant genome engineers to develop resistance against this rapidly evolving pathogen by mutating susceptibility genes. TALENs have been employed in the past, but extensive protein engineering makes developing broad-spectrum immunity against bacterial blight difficult (Li et al., 2012). The prokaryotic immune system (CRISPR-Cas9) was customized and used for editing susceptibility genes in rice (Jiang et al., 2013). In the current study this customized tool was used for editing the *OsSWEET14* gene's promoter of Basmati, to obtain resistance against bacterial blight.

Basmati rice is famous for its long grain and aroma and none of the natural mutations are reported in *OsSWEET14* for Super Basmati rice. For the *OsSWEET13* gene, one deletion and substitutions are naturally present in the promoter which disrupts all the EBE variants of *PthXo2* TALE, i.e., *PthXo2A*, *PthXo2B,* and *PthXo2C* (Zaka et al., 2018; Oliva et al., 2019). It was reported that in the majority of cases, *Xoo* strains induce the expression of *OsSWETT14* for the onset of infection because it has the EBEs of four different TALEs. *OsSWEET14* was targeted previously as a single target as well as in combination with other *OsSWEET* genes to create broad-spectrum resistance in different rice cultivars (Blanvillain-Baufumé et al., 2017; Oliva et al., 2019; Xu et al., 2019). They have developed different lines with edited EBEs of the *OsSWEET14* gene's promoter. Each EBE edited line provided resistance against specific TALE. In the present study, the *OsSWEET14* gene's promoter was selected to modify the respective EBEs of four TALEs via CRISPR-Cas9. We have successfully mutated EBEs of three TALEs (*AvrXa7, PthXo3,* and *TalF*). The edited rice lines (SB-E1, SB-E2, SB-E3, and SB-E4) have modified alleles of *OsSWEET14*. Among these four edited lines, SB-E1 and SB-E2 did not respond to *AvrXa7/PthXo3* to establish a successful infection. The new germplasm created in this way showed resistance against a locally virulent *Xoo* strain. These results indicated that CRISPR-Cas9 technology can also be employed on Basmati rice to create resistance against bacterial blight.

Blanvillain-Baufumé et al. (2017) created the allele library of *OsSWEET14* for developing resistance against bacterial blight. The *OsSWEET14* gene's promoter was targeted using TALENS, and they observed 51% editing efficiency. In the current study the EBEs of the *OsSWEET14* gene's promoter were targeted using CRISPR-Cas9 and we were able to obtain only 9% editing efficiency. All the editing events were in close proximity to the predicted cleavage site by Cas9 which is three base pairs upstream of the PAM sequence. This indicates the specificity of this editing system. However, we observed low editing efficiency as compared to the previous report of editing the *OsSWEET14* gene's promoter. The reason for low editing efficiency could be the difference in the editing tool (TALENs) and rice variety as they have used a non-Basmati background. The editing efficiency can be improved by using Cas9 fusion with chromatin-modulating peptides (CMPs), derived from high mobility group proteins. This fusion exhibited many folds improved activity (Ding et al., 2019). Thus, such type of improved systems can be tested in elite varieties, like Super Basmati, to check improvement in editing efficiency. Blanvillain-Baufumé et al. (2017) obtained mixed events (insertion/deletion) at the target sites including the insertion of 22 bp and deletions of up-to 51 bp, whereas another study by Jiang et al. (2013) reported only deletions at the target site. In our case, we were only able to obtain deletions at the target sites, because the repairing of DSB (introduced by CRISPR-Cas9) was done through a random NHEJ process. Therefore, obtaining deletions, insertions, or point mutations at the target site is solely dependent on the host repair machinery.

The diversity of TALEs present in different *Xoo* strains was studied previously by analyzing *Xoo* genome sequence data. Most of the Asian strains had *AvrXa7* (*OsSWEET14*) and *PthXo2* (*SWEET13*) (Oliva et al., 2019). Sequencing data of the most prevalent *Xoo* strain infecting Basmati showed that this strain had *AvrXa7* for infecting Basmati. The induction of *OsSWEET14* was also checked by the tested *Xoo* strain and RT-qPCR results confirmed *OsSWEET14* induction by the tested *Xoo* strain. So, targeting the promoter of *OsSWEET14* in Super Basmati is very effective in generating broad-spectrum resistance. However, this strategy can be modified depending on the TALEs present in *Xoo* strains and their respective EBEs in rice cultivars. More work is still required, however, in order to deal with bacterial blight by characterizing maximum *Xoo* isolates. Some *Xoo* TALEs also target *OsSWEET13* and *OsSWEET11* for the onset of infection. But the isolates which we have characterized were mostly targeting *OsSWEET14*, which agrees with studies reported by Blanvillain-Baufumé et al. (2017).

In a previous report it was shown that Asian *Xoo* strains did not have *TalF* (Oliva et al., 2019). Our study also supports their finding because our locally isolated strain did also not possess

TalF. The selected *Xoo* strain was inoculated on all the edited lines and disease development was recorded. The symptoms were reduced on *AvrXa7* edited lines. When *TalF*-edited lines were challenged with the same *Xoo* strain, the symptoms were not reduced. This was further confirmed by RT-qPCR of the control and edited lines. The relative expression of *OsSWEET14* was similar in wild type SB and *TalF* edited lines (SB-E3, SB-E4), whereas the strain was unable to induce *OsSWEET14* in the *AvrXa7* edited line (SB-E2). This showed that the strain has *AvrXa7* to infect Super Basmati. These findings confirmed our hypothesis that targeting EBEs in the promoter region only provides specific resistance to their corresponding TALE present in the *Xoo* strain.

Mutations in susceptibility genes can also have side effects on normal plant physiology (van Schie and Takken, 2014). In a previous study, rice plants with a *OsSWEET14* TDNA insertion mutant had smaller seeds as compared to wild type plants, although it showed resistance against *Xoo* strains (Antony et al., 2010). In contrast, *OsSWEET14* EBE edited rice plants were no longer susceptible to *Xoo* and showed normal growth (Li et al., 2012). In the current study the edited lines were visually observed during the growth period and after harvest. All the edited lines showed no detectable growth defects in greenhouse conditions except for SB-E1. In SB-E1 seed filling was disturbed. The reason could be attributed to a large deletion in the promoter region. This also indicates the role of *OsSWEET* genes in seed filling (Sosso et al., 2015; Yang et al., 2018; Hu et al., 2019). However, further investigation is required in this regard. Thus, in comparison to former studies, we came across both types of results, and out of four EBE mutated lines we observed abnormality in only one edited line.

There are some previous reports of CRISPR-Cas9 mediated genome editing of rice to develop resistance against bacterial blight. Editing in the promoter fragment of *OsSWEET* genes was reported using CRISPR-Cas9 to develop resistance against bacterial blight (Oliva et al., 2019; Xu et al., 2019). But the majority of previous genome editing was performed in rice cultivar kitake (japonica) and cannot be used for breeding programs of Basmati rice. Recently, genome editing via CRISPR-Cas9 was performed to mutate EBEs of *OsSWEET11*, *OsSWEET13,* and *OsSWEET14* to create resistance in Indica cultivar (IR-64 and Ciherang-Sub1). These cultivars can be used by breeders of Asia and Africa in breeding programs but there was no report on creating resistance in the Basmati cultivars. Therefore, we performed genome editing in the elite Super Basmati rice cultivar to be incorporated into breeding programs. The present study was designed to establish CRISPR-Cas9 mediated genome editing in Basmati rice. To the best of our knowledge, this is the first report of genome editing in Basmati rice using the CRISPR-Cas9 toolbox.

This study shows the potential of CRISPR-Cas9 based genome editing in elite Super Basmati rice. However, there is still more work needed to deal with bacterial blight, by characterizing the maximum of local *Xoo* isolates, because some *Xoo* TALEs also target *OsSWEET13* and *OsSWEET11*. So, to create broad-spectrum resistance against all the native *Xoo* strains, multiplexing can also be done in Basmati. Further, a study can be planned to establish multiplex genome editing against bacterial blight by simultaneously targeting multiple EBEs in the promoter regions of *OsSWEET* genes, which will ultimately result in reducing bacterial blight. This approach of dealing with bacterial blight by targeting the promoter of *OsSWEET* genes will not prevent the adaptation of pathogens. The strength of this method also depends on the ability of local *Xoo* strains to acclimate to recessive resistance alleles. To develop durable resistance against pathogens, it is better to create major changes in EBEs of the *OsSWEET* gene's promoter. Combining these edited alleles with locally effective resistance genes can be a more effective way to reduce disease pressure. In conclusion, by understanding TALE interaction with EBEs of *OsSWEET14* genes, we were able to create resistance against a corresponding *Xoo* strain.

Genome editing can have off-target effects. The gRNAs used in this study were analyzed for off-targeting against the rice genome available in the database. All the gRNAs did not have any off-targets. Furthermore, if any off-target mutations are still present as a result of genome editing or tissue culturing, they will be eliminated during crossing. Full genome sequencing will be required to confirm any off-targeting. Finally, our initial data shows the potential of CRISPR-Cas9 based genome editing in elite Basmati cultivar. However, the establishment of a transgene-free genome editing protocol for targeting multiple EBEs simultaneously is still needed in Super Basmati rice to create broad-spectrum resistance against a large collection of *Xoo* strains. Such transgene-free genome-edited elite lines created in this way can be added in breeding programs.

CONCLUSION

Our results show that targeting EBEs of respective TALEs, employing the CRISPR-Cas9 approach, can provide highly selective and promising immunity against bacterial blight.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

SM planned the work. KZ, MK, and IA planned the experiments. KZ performed all the experiments and wrote the manuscript. MK and IA edited and finalized the manuscript. ZM helped establishing tissue culture of Super Basmati. SY and KE performed the resistance assays against *Xoo*. MA provided the seeds for this study and give critical suggestions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00575/full#supplementary-material

REFERENCES

Abe, K., Araki, E., Suzuki, Y., Toki, S., and Saika, H. (2018). Production of high oleic/low linoleic rice by genome editing. *Plant Physiol. Biochem.* 131, 58–62. doi: 10.1016/j.plaphy.2018.04.033

Antony, G., Zhou, J., Huang, S., Li, T., Liu, B., White, F., et al. (2010). Rice *xa13* recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene *Os-11N3*. *Plant Cell* 22, 3864–3876. doi: 10.1105/tpc.110.078964

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

Blanvillain-Baufumé, S., Reschke, M., Solé, M., Auguy, F., Doucoure, H., Szurek, B., et al. (2017). Targeted promoter editing for rice resistance to *Xanthomonas oryzae pv. oryzae* reveals differential activities for *SWEET 14*−inducing TAL effectors. *Plant Biotechnol. J.* 15, 306–317.

Busungu, C., Taura, S., Sakagami, J.-I., and Ichitani, K. (2016). Identification and linkage analysis of a new rice bacterial blight resistance gene from XM14, a mutant line from IR24. *Breed. Sci.* 66, 636–645. doi: 10.1270/jsbbs.16062

Carpenter, S. C., Mishra, P., Ghoshal, C., Dash, P. K., Wang, L., Midha, S., et al. (2018). A strain of an emerging Indian *pv. oryzae* pathotype defeats the rice bacterial blight resistance gene without inducing a clade III gene and is nearly identical to a recent Thai isolate. *Front. Microbiol.* 9:2703. doi: 10.3389/fmicb.2018.02703

Chen, L.-Q., Hou, B.-H., Lalonde, S., Takanaga, H., Hartung, M. L., Qu, X.-Q., et al. (2010). Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* 468, 527–532. doi: 10.1038/nature09606

Chen, L.-Q., Qu, X.-Q., Hou, B.-H., Sosso, D., Osorio, S., Fernie, A. R., et al. (2012). Sucrose efflux mediated by *SWEET* proteins as a key step for phloem transport. *Science* 335, 207–211. doi: 10.1126/science.1213351

Chu, Z., Yuan, M., Yao, J., Ge, X., Yuan, B., Xu, C., et al. (2006). Promoter mutations of an essential gene for pollen development result in disease resistance in rice. *Genes Dev.* 20, 1250–1255. doi: 10.1101/gad.1416306

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

Ding, X., Seebeck, T., Feng, Y., Jiang, Y., Davis, G. D., and Chen, F. (2019). Improving CRISPR-Cas9 genome editing efficiency by fusion with chromatin-modulating peptides. *CRISPR J.* 2, 51–63. doi: 10.1089/crispr.2018.0036

Dou, D., and Zhou, J.-M. (2012). Phytopathogen effectors subverting host immunity: different foes, similar battleground. *Cell Host Microbe* 12, 484–495. doi: 10.1016/j.chom.2012.09.003

Doucouré, H., Pérez-Quintero, A. L., Reshetnyak, G., Tekete, C., Auguy, F., Thomas, E., et al. (2018). Functional and genome sequence-driven characterization of TAL effector gene repertoires reveals novel variants with altered specificities in closely related Malian *Xanthomonas oryzae pv. oryzae* strains. *Front. Microbiol.* 9:1657. doi: 10.3389/fmicb.2018.01657

Hsu, P. D., Lander, E. S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157, 1262–1278. doi: 10.1016/j.cell.2014.05.010

Hu, B., Wu, H., Huang, W., Song, J., Zhou, Y., and Lin, Y. (2019). *SWEET* gene family in *Medicago truncatula*: genome-wide identification, expression and substrate specificity analysis. *Plants* 8:338. doi: 10.3390/plants8090338

Hu, X., Wang, C., Fu, Y., Liu, Q., Jiao, X., and Wang, K. (2016). Expanding the range of CRISPR/Cas9 genome editing in rice. *Mol. Plant* 9, 943–945. doi: 10.1016/j.molp.2016.03.003

Hutin, M., Pérez-Quintero, A. L., Lopez, C., and Szurek, B. (2015a). MorTAL Kombat: the story of defense against TAL effectors through loss-of-susceptibility. *Front. Plant Sci.* 6:535. doi: 10.3389/fpls.2015.00535

Hutin, M., Sabot, F., Ghesquière, A., Koebnik, R., and Szurek, B. (2015b). A knowledge−based molecular screen uncovers a broad−spectrum *OsSWEET14* resistance allele to bacterial blight from wild rice. *Plant J.* 84, 694–703. doi: 10.1111/tpj.13042

Iyer-Pascuzzi, A. S., and McCouch, S. R. (2007). Recessive resistance genes and the *Oryza sativa-Xanthomonas oryzae* pv. oryzae pathosystem. *Mol. Plant Microbe Interact.* 20, 731–739.

Ji, Z., Ji, C., Liu, B., Zou, L., Chen, G., and Yang, B. (2016). Interfering TAL effectors of *Xanthomonas oryzae* neutralize R-gene-mediated plant disease resistance. *Nat. Commun.* 7:13435. doi: 10.1038/ncomms13435

Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B., and Weeks, D. P. (2013). Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in *Arabidopsis*, tobacco, sorghum and rice. *Nucleic Acids Res.* 41:e188. doi: 10.1093/nar/gkt780

Kauffman, H. (1973). An improved technique for evaluating resistance of rice varieties to *Xanthomonas oryzae*. *Plant Dis. Rep.* 57, 537–541.

Khan, M. A., Naeem, M., and Iqbal, M. (2014). Breeding approaches for bacterial leaf blight resistance in rice (*Oryza sativa L.*), current status and future directions. *Eur. J. Plant Pathol.* 139, 27–37. doi: 10.1007/s10658-014-0377-x

Khush, G. S. (2005). What it will take to feed 5.0 billion rice consumers in 2030. *Plant Mol. Biol.* 59, 1–6. doi: 10.1007/s11103-005-2159-5

Kim, S.-M., Suh, J.-P., Qin, Y., Noh, T.-H., Reinke, R. F., and Jena, K. K. (2015). Identification and fine-mapping of a new resistance gene, *Xa40*, conferring resistance to bacterial blight races in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 128, 1933–1943. doi: 10.1007/s00122-015-2557-2

Kourelis, J., and van der Hoorn, R. A. (2018). Defended to the nines: 25 years of resistance gene cloning identifies nine mechanisms for R protein function. *Plant Cell* 30, 285–299. doi: 10.1105/tpc.17.00579

Li, C., Li, W., Zhou, Z., Chen, H., Xie, C., and Lin, Y. (2019). A new rice breeding method: CRISPR/Cas9 system editing of the *Xa13* promoter to cultivate transgene−free bacterial blight−resistant rice. *Plant Biotechnol. J.* 18, 313–315. doi: 10.1111/pbi.13217

Li, T., Liu, B., Spalding, M. H., Weeks, D. P., and Yang, B. (2012). High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat. Biotechnol.* 30, 390–392. doi: 10.1038/nbt.2199

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., Dicarlo, J. E., et al. (2013). RNA-guided human genome engineering via Cas9. *Science* 339, 823–826. doi: 10.1126/science.1232033

Mukhtar, Z., and Hasnain, S. (2018). Optimization of particle bombardment conditions for rice (*Oryza sativa* L.) transformation. *Pak. J. Agric. Sci.* 55, 271–278. doi: 10.21162/pakjas/18.3678

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

Oliva, R., Ji, C., Atienza-Grande, G., Huguet-Tapia, J. C., Perez-Quintero, A., Li, T., et al. (2019). Broad-spectrum resistance to bacterial blight in rice using genome editing. *Nat. Biotechnol.* 37, 1344–1350. doi: 10.1038/s41587-019-0267-z

Pazuki, A., and Sohani, M. M. (2013). Phenotypic evaluation of scutellum-derived calluses in 'Indica'rice cultivars. *Acta Agric. Slov.* 101, 239–247.

Sasaki, T. (2001). Rice genome analysis to understand the rice plant as an assembly of genetic codes. *Photosynth. Res.* 70, 119–127.

Shobarani, N., Prasad, G., Prasad, A., Sailaja, B., Muthuraman, P., Numeera, S., et al. (2010). Rice Almanac–India. *DRR Tech. Bull.* 5, 6–7.

Sosso, D., Luo, D., Li, Q.-B., Sasse, J., Yang, J., Gendrot, G., et al. (2015). Seed filling in domesticated maize and rice depends on *SWEET*-mediated hexose transport. *Nat. Genet.* 47, 1489–1493. doi: 10.1038/ng.3422

Stewart, C. N. Jr. (1993). A rapid CTAB DNA isolation technique useful for RAPD fingerprinting and other PCR applications. *Biotechniques* 14, 748–750.

Streubel, J., Pesce, C., Hutin, M., Koebnik, R., Boch, J., and Szurek, B. (2013). Five phylogenetically close rice *SWEET* genes confer TAL effector−mediated susceptibility to *Xanthomonas oryzae pv. oryzae*. *New Phytol.* 200, 808–819. doi: 10.1111/nph.12411

Tu, J., Datta, K., Khush, G., Zhang, Q., and Datta, S. (2000). Field performance of *Xa21* transgenic indica rice (*Oryza sativa L.)*, IR72. *Theor. Appl. Genet.* 101, 15–20. doi: 10.1007/s001220051443

van Schie, C. C., and Takken, F. L. (2014). Susceptibility genes 101: how to be a good host. *Annu. Rev. Phytopathol.* 52, 551–581. doi: 10.1146/annurev-phyto-102313-045854

Wright, A. V., Nuñez, J. K., and Doudna, J. A. (2016). Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. *Cell* 164, 29–44. doi: 10.1016/j.cell.2015.12.035

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

Xu, Z., Xu, X., Gong, Q., Li, Z., Li, Y., Wang, S., et al. (2019). Engineering broad-spectrum bacterial blight resistance by simultaneously disrupting variable TALE-binding elements of multiple susceptibility genes in rice. *Mol. Plant* 12, 1434–1446. doi: 10.1016/j.molp.2019.08.006

Yang, B., Sugio, A., and White, F. F. (2006). *Os8N3* is a host disease-susceptibility gene for bacterial blight of rice. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10503–10508. doi: 10.1073/pnas.0604088103

Yang, B., Zhu, W., Johnson, L. B., and White, F. F. (2000). The virulence factor *AvrXa7* of *Xanthomonas oryzae pv. oryzae* is a type III secretion pathway-dependent nuclear-localized double-stranded DNA-binding protein. *Proc. Natl. Acad. Sci. U.S.A.* 97, 9807–9812. doi: 10.1073/pnas.170286897

Yang, J., Luo, D., Yang, B., Frommer, W. B., and Eom, J. S. (2018). *SWEET11* and 15 as key players in seed filling in rice. *New Phytol.* 218, 604–615. doi: 10.1111/nph.15004

Yu, Y., Streubel, J., Balzergue, S., Champion, A., Boch, J., Koebnik, R., et al. (2011). Colonization of rice leaf blades by an African strain of *Xanthomonas oryzae pv. oryzae* depends on a new TAL effector that induces the rice nodulin-3 *Os11N3* gene. *Mol. Plant Microbe Interact.* 24, 1102–1113. doi: 10.1094/MPMI-11-10-0254

Yuan, M., Chu, Z., Li, X., Xu, C., and Wang, S. (2009). Pathogen-induced expressional loss of function is the key factor in race-specific bacterial resistance conferred by a recessive R gene *xa13* in rice. *Plant Cell Physiol.* 50, 947–955. doi: 10.1093/pcp/pcp046

Zaidi, S. S.-E.-A., Mukhtar, M. S., and Mansoor, S. (2018). Genome editing: targeting susceptibility genes for plant disease resistance. *Trends Biotechnol.* 36, 898–906. doi: 10.1016/j.tibtech.2018.04.005

Zaka, A., Grande, G., Coronejo, T., Quibod, I. L., Chen, C.-W., Chang, S.-J., et al. (2018). Natural variations in the promoter of *OsSWEET13* and *OsSWEET14* expand the range of resistance against *Xanthomonas oryzae pv. oryzae*. *PLoS One* 13:e0203711. doi: 10.1371/journal.pone.0203711

Zeigler, R. S., and Barclay, A. (2008). The relevance of rice. *Rice* 1, 3–10. doi: 10.1007/s12284-008-9001-z

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 16 June 2020
doi: 10.3389/fpls.2020.00800

CRISPR/Cas9-Based Gene Editing Using Egg Cell-Specific Promoters in Arabidopsis and Soybean

Na Zheng[1,2], Ting Li[3], Jaime D. Dittman[4], Jianbin Su[2], Riqing Li[2], Walter Gassmann[2], Deliang Peng[1], Steven A. Whitham[4], Shiming Liu[1]* and Bing Yang[2,5]**

[1] State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China, [2] Division of Plant Sciences, Christopher S. Bond Life Sciences Center, and Interdisciplinary Plant Group, University of Missouri, Columbia, MO, United States, [3] The Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, United States, [4] Department of Plant Pathology and Microbiology, Iowa State University, Ames, IA, United States, [5] Donald Danforth Plant Science Center, St. Louis, MO, United States

CRISPR/Cas9-based systems are efficient genome editing tools in a variety of plant species including soybean. Most of the gene edits in soybean plants are somatic and non-transmissible when Cas9 is expressed under control of constitutive promoters. Tremendous effort, therefore, must be spent to identify the inheritable edits occurring at lower frequencies in plants of successive generations. Here, we report the development and validation of genome editing systems in soybean and Arabidopsis based on Cas9 driven under four different egg-cell specific promoters. A soybean ubiquitin gene promoter driving expression of green fluorescent protein (GFP) is incorporated in the CRISPR/Cas9 constructs for visually selecting transgenic plants and transgene-evicted edited lines. In Arabidopsis, the four systems all produced a collection of mutations in the T2 generation at frequencies ranging from 8.3 to 42.9%, with egg cell-specific promoter AtEC1.2e1.1p being the highest. In soybean, function of the gRNAs and Cas9 expressed under control of the CaMV double 35S promoter (2x35S) in soybean hairy roots was tested prior to making stable transgenic plants. The 2x35S:Cas9 constructs yielded a high somatic mutation frequency in soybean hairy roots. In stable transgenic soybean T1 plants, AtEC1.2e1.1p:Cas9 yielded a mutation rate of 26.8%, while Cas9 expression driven by the other three egg cell-specific promoters did not produce any detected mutations. Furthermore, the mutations were inheritable in the T2 generation. Our study provides CRISPR gene-editing platforms to generate inheritable mutants of Arabidopsis and soybean without the complication of somatic mutagenesis, which can be used to characterize genes of interest in Arabidopsis and soybean.

Keywords: CRISPR/Cas9, gene editing, egg cell-specific promoter, Arabidopsis, *Glycine max*, soybean

OPEN ACCESS

Edited by:
Sandeep Kumar,
Corteva Agriscience™, United States

Reviewed by:
Yanfei Mao,
Shanghai Institutes for Biological Sciences (CAS), China
Qi-Jun Chen,
China Agricultural University, China

***Correspondence:**
Steven A. Whitham
swhitham@iastate.edu
Shiming Liu
smliuhn@yahoo.com
Bing Yang
yangbi@missouri.edu

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *11 April 2020*
Accepted: *19 May 2020*
Published: *16 June 2020*

Citation:
Zheng N, Li T, Dittman JD, Su J, Li R, Gassmann W, Peng D, Whitham SA, Liu S and Yang B (2020) CRISPR/Cas9-Based Gene Editing Using Egg Cell-Specific Promoters in Arabidopsis and Soybean. Front. Plant Sci. 11:800. doi: 10.3389/fpls.2020.00800

INTRODUCTION

With the advent of genome editing technologies, such as zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeat (CRISPR) and CRISPR-associated protein (Cas) (CRISPR/Cas) in particular, targeted mutagenesis and precise base changes in genomes of interest can be achieved in ways that

were unimaginable 10 years ago (Zhang et al., 2018). These engineered nucleases can generate double-stranded DNA breaks (DSBs) at pre-chosen genomic loci, and repairs to the DSBs *in vivo* lead to site-specific genetic alterations. Two main pathways are used to repair DSBs *in vivo*: error-prone non-homologous end joining (NHEJ) and error-free homology-directed repair (HDR) in the presence of template DNA. The former is the predominant event that introduces insertions/deletions (indels) that range from one to hundreds of base pairs (Voytas, 2013).

The type II CRISPR/Cas9 system from *Streptococcus pyogenes* is the first described and the most popular CRISPR system for genome editing. CRISPR/Cas9 consists of two components, the Cas9 nuclease and a chimeric single guide RNA (gRNA) derived from the fusion of a crRNA (CRISPR RNA) and a trans-activating crRNA preceded by a spacer (or guide) sequence of 18–20 nucleotides complementary to the target DNA (or protospacer). The Cas9 protein cleaves the target DNA to cause DSBs predominantly located 3 bp upstream of the protospacer adjacent motif (PAM) sequence (5′-NGG-3′). Due to the ease of assembly and high frequency of inducing mutations, the CRISPR/Cas9 system is widely used for gene editing in various organisms, including yeast, mouse, fish, human cells, and plant species (Jinek et al., 2012; Cong et al., 2013; Hwang et al., 2013; Jiang et al., 2013; Li et al., 2013; Mali et al., 2013; Nekrasov et al., 2013; Shan et al., 2013).

Soybean (*Glycine max*) is one of the most economically important crops for food, vegetable oil and animal feed. As a paleopolyploid species, gene function studies in soybean are frequently complicated by genetic redundancy, in addition to low frequencies of genetic transformation. Nevertheless, CRISPR/Cas9 genome editing systems have been successfully utilized in soybean since the first demonstrations in 2015. It has been applied to create targeted mutations in hairy roots, somatic embryos and stable transgenic plants (Jacobs et al., 2015; Li et al., 2015, 2019; Michno et al., 2015; Sun et al., 2015; Du et al., 2016; Chilcoat et al., 2017; Cai et al., 2018, 2020; Kanazashi et al., 2018; Al Amin et al., 2019; Bao et al., 2019; Campbell et al., 2019; Cheng et al., 2019; Do et al., 2019; Zhang et al., 2019; Michno et al., 2020; Wang et al., 2020; Wu et al., 2020). Various promoters have been deployed for expression of Cas9. For example, the constitutive cauliflower mosaic virus (CaMV) 35S promoter (35S) is the most used, especially in hairy roots (Jacobs et al., 2015). The soybean SCREAM M4 promoter (pM4) (Bai et al., 2020), parsley ubiquitin promoter (Kanazashi et al., 2018), and translation elongation factor 1 alpha 2 (*EF1A2*) promoter (Li et al., 2015) have also been used to constitutively activate Cas9 expression.

Compared with the constitutive promoters, germline specific promoters for Cas9 expression can improve the frequency and heritability of mutations significantly in Arabidopsis. Expression of Cas9 under control of the *DD45* (egg cell and early embryo), *Yao* (shoot apical and root meristem-active), tomato *Lat52* (pollen) and *EC* (egg cells, embryo) promoters can increase the frequency of heritable edits in the T2 generation. These promoters also reduce the rate of somatic mutations (Wang et al., 2015; Yan et al., 2015; Mao et al., 2016). The lower frequency of chimerism reduces the need to screen large numbers of individual plants and conduct multiple generational analyses to acquire the desired mutants in Arabidopsis. The efficiency of *Agrobacterium*-mediated soybean transformation is very low, labor-intensive and time-consuming, so a high-efficiency CRISPR/Cas9 system based on germline specific promoters may reduce the chimerism and thus, the workload of characterizing edited plants.

In this work, we present easy-to-use binary vector systems with Cas9 driven by egg cell-specific promoters (ECp) for efficient site-specific mutagenesis in Arabidopsis and soybean based on *Agrobacterium*-mediated transformation. In the system, a GFP marker can also be used to identify transgenic and transgene-free plants. We validated the ECp-Cas9 systems and *Agrobacterium*-mediated protocol by targeting two genes each in Arabidopsis and soybean. Our results showed that egg cell-specific promoters can induce mutations of endogenous genes in Arabidopsis and soybean, and multiple, independent mutations can be obtained from the progeny of individual single lines. We confirmed that the continuous presence of the Cas9/gRNA construct in transgenic plants can cause mutagenesis of target genes of interest in subsequent generations.

RESULTS

Construction of an ECp-Cas9/gRNA System for Genome Editing in Arabidopsis and Soybean

The cloning strategy for building the series of gRNA vectors for insertion into the destination ECp-Cas9/gRNA binary vector suitable for *Agrobacterium*-mediated transformation of Arabidopsis and soybean is shown in **Figure 1**. The gRNA vectors (pCRgRNA1 to pCRgRNA6) were designed and constructed to generate individual gRNA units each consisting of a small nuclear RNA (snRNA, U6 or U3) promoter and a gRNA followed by a poly-T terminator. The multiple cloning sites between the U6 or U3 and the gRNA scaffold sequence contain two *Bsm*BI sites which facilitate an insertion of a double-stranded DNA fragment with two unique overhangs generated after annealing two complementary oligonucleotides. The inserted sequence in each gRNA vector forms the spacer or guide sequence of a gRNA gene designed specifically to target the genomic locus of interest (**Figure 1A**). The six gRNA cassettes were designed and constructed such that, after digestion with *Bsa*I, each unit has overhanging ends compatible to the ends of the adjacent units, which are assembled into the intermediate pENTR_ccdB vector using Golden Gate assembly (**Supplementary Table S1**). In order to have the flexibility of making constructs containing fewer than 6 gRNA units, two additional vectors (pCRgRNA2T and pCRgRNA4T) were designed and constructed to enable 2- and 4-gRNA units (**Supplementary Table S1**). Therefore, the vectors can be used to construct gRNA modules consisting of 2, 4, or 6 gRNAs (**Figure 1B**). In this study, we used the 2-gRNA version to target two different genes at once for mutagenesis (**Figure 2**).

For egg cell-specific expression of Cas9, promoters of four egg cell-specific genes were fused to the Cas9 coding sequence to establish the four CRISPR/Cas9 systems used in this study

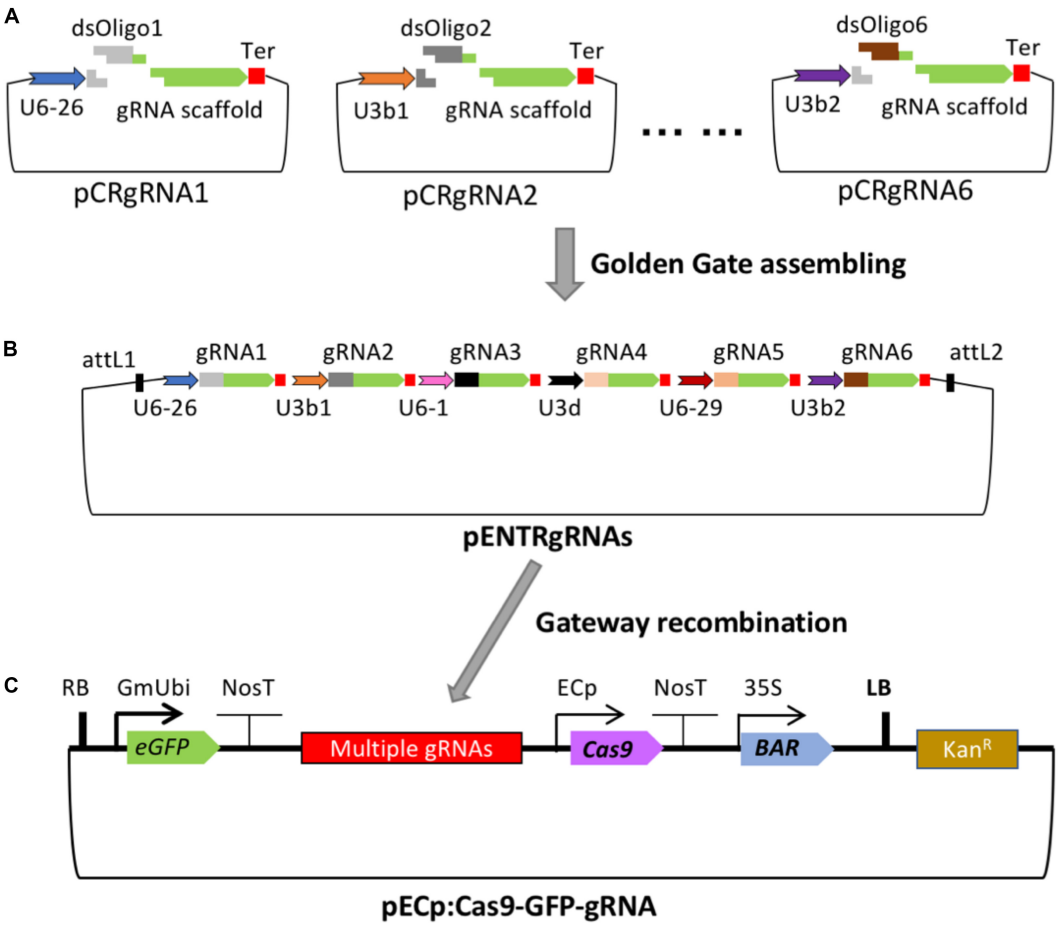


FIGURE 1 | Cloning strategy for adding gRNA modules to the T-DNA vectors carrying Cas9, eGFP, and Bar. **(A)** The pCRgRNA vectors (with two *Bsm*BI and two *Bsa*I recognition sites) have different *Arabidopsis* U6 or U3 promoters to drive expression of each gRNA. A common transcription terminator (Ter) follows each gRNA scaffold. Each gRNA vector can be digested with *Bsm*BI for the insertion of double-stranded oligonucleotides as the guide sequence of a specific gRNA (dsOligo#). **(B)** Up to six unique gRNA expression cassettes can be assembled into an intermediate construct, pENTR4-sgRNAs. **(C)** The gRNA module flanked by the *att*L1 and *att*L2 sites is mobilized into the binary vector P1300-ECp-Cas9-GFP-ccdB by Gateway recombination. The resulting Cas9/sgRNA binary construct is used for *Agrobacterium*-mediated transformation. The features of the plasmids are not drawn to scale.

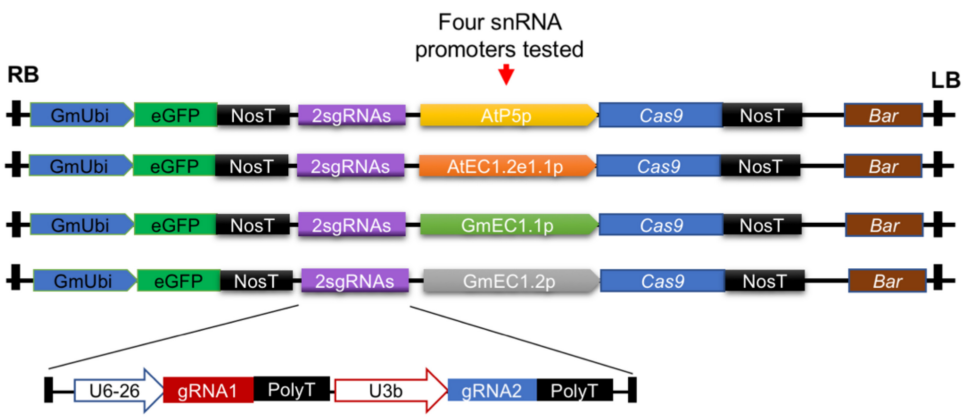


FIGURE 2 | Design of constructs to test ability of four different egg cell-specific promoters to activate expression of Cas9 in Arabidopsis and soybean. In these studies, a gRNA module expressing two different gRNAs (2sgRNAs) targeting unique loci in Arabidopsis and soybean was used. GmUbi, soybean ubiquitin promoter; eGFP, enhanced green fluorescent protein; NosT, nopaline synthase terminator; Bar, bialaphos resistance.

(**Figure 2**). This strategy was used to allow us to identify the promoter that worked best for genome editing in both Arabidopsis and soybean. The promoter from the AT1G71470 locus of Arabidopsis, referred to as AtP5p, was used for expression of the Arabidopsis codon-optimized Cas9. Another Arabidopsis promoter, referred to as AtEC1.2e1.1p, was adapted from the fusion of AtEC1.1 and AtEC1.2 *cis*-regulatory elements (Wang et al., 2015). Additionally, two soybean promoters [GmEC1.1p and GmEC1.2p from loci encoding *G. max* egg cell-secreted protein 1.1 (LOC100801164) and *G. max* egg cell-secreted protein 1.2 (LOC102670289), respectively] were identified and used for Cas9 expression. The Cas9 expression plasmids each contained a cassette of the *ccdB* gene flanked by the Gateway recombination sequences *att*L1 and *att*L2 (**Supplementary Figure S1**). The gRNA modules from pENTR-gRNAs could be mobilized to the individual destination vectors pGW-ECp:Cas9-GmUbi:GFP through Gateway recombination, resulting in a single binary vector pECp:Cas9/gRNA for *Agrobacterium*-mediated plant transformation with the *bar* gene driven by the CaMV 35S promoter as the transformation selection marker (**Figures 1C**, **2**).

The guide sequences of gRNA genes were selected and designed based on the *Arabidopsis thaliana* ecotype Wassilewskija-2 (WS-2) or soybean cultivar Williams 82 genome sequences using the CRISPR Genome Analysis Tool (Brazelton et al., 2015)[1]. The corresponding target genomic regions were PCR-amplified and confirmed by Sanger sequencing prior to gRNA design. All pCRgRNA constructs were confirmed for sequence accuracy at the insertion sites and the flanking regions by Sanger sequencing. The binary plasmids were mobilized into *Agrobacterium* strain GV3101 or EHA105 for transformation of Arabidopsis and soybean, respectively.

The overall strategy to design and apply a CRISPR/Cas9 construct for targeted mutagenesis in Arabidopsis and soybean is illustrated in **Figure 3**. The main steps include designing and constructing individual gRNA gene units, assembling the gRNA units into a guide RNA cassette, transferring the gRNAs into ECp:Cas9 vectors, transforming *Agrobacterium*, and performing *Agrobacterium*-mediated plant transformation. Basta resistance and GFP fluorescence were used to screen and select for the transgenic or transgene-free plants of the T0, T1, and T2 generations. The final step involves genotyping edited plants through PCR-amplification of the targeted regions followed by restriction enzyme digestion analysis (PCR-RE) or T7 endonuclease I (T7E1) assay and further confirmation of edits by Sanger sequencing (**Figure 3**).

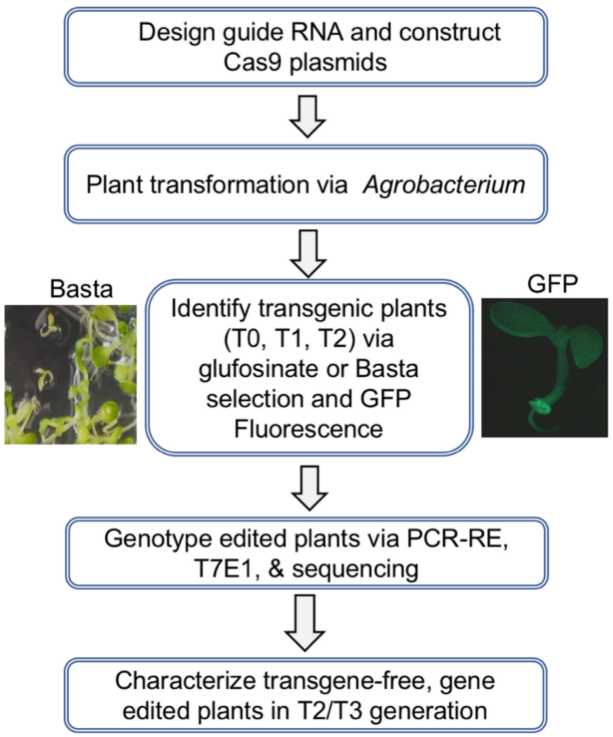


FIGURE 3 | Overall strategy to generate stable transgenic Arabidopsis and soybean lines and identify progeny plants carrying site-specific mutations.

Targeted Mutagenesis of *AtRPS4* and *AtRPS4B* in Arabidopsis

To investigate whether the ECp-Cas9/gRNA systems could induce site-specific mutations, we first tested the ability of each to induce mutations in Arabidopsis. The Arabidopsis genes encoding the TIR-NB-LRR proteins AtRPS4 and AtRPS4B were chosen as the targets. These resistance proteins activate effector-triggered host immunity upon recognizing two bacterial effectors, AvrRps4 from *Pseudomonas syringae* and PopP2 from *Ralstonia solanacearum* (Saucet et al., 2015). The two gRNA genes (gRPS4 and gRPS4B, one gRNA for each target gene) were designed and constructed into one gRNA cassette using pCRgRNA1 and pCRgRNA2T as intermediate cloning vectors. The Cas9 cleavage site (three nucleotides upstream of the NGG PAM) of each target gene overlapped with a restriction enzyme recognition site (*Xba*I for *AtRPS4* and *Bgl*II for *AtRPS4B*), which facilitated genotyping using the PCR-RE approach (**Figure 4A**).

The same gRNA cassette was mobilized individually into pGW-ECp:Cas9-GmUbi:GFP recipient vectors, resulting in four different ECp:Cas9/gRNA constructs that were individually introduced into Arabidopsis through *Agrobacterium*-mediated transformation using the floral-dipping method. The Bar and GFP markers were used to select the transgene positive plants. Compared with non-transgenic Arabidopsis WS-2 plants, the transgenic plants grew well on the MS solid medium supplemented with Basta and showed strong green fluorescence under fluorescence microscopy (**Figures 4B–D**).

We first established a way to identify true transgenic plants by combining Basta resistance and GFP fluorescence presence to screen the T1 and T2 generations. Several transgenic T1 plants obtained from four different ECp:Cas9/gRNA constructs were subjected to Basta, GFP fluorescence and Cas9/gRNA-PCR screening (**Supplementary Table S2**). Compared to Basta screening, the GFP fluorescence was more consistent with the presence of the gRNA genes and Cas9 as further confirmed

[1]http://cbc.gdcb.iastate.edu/cgat/

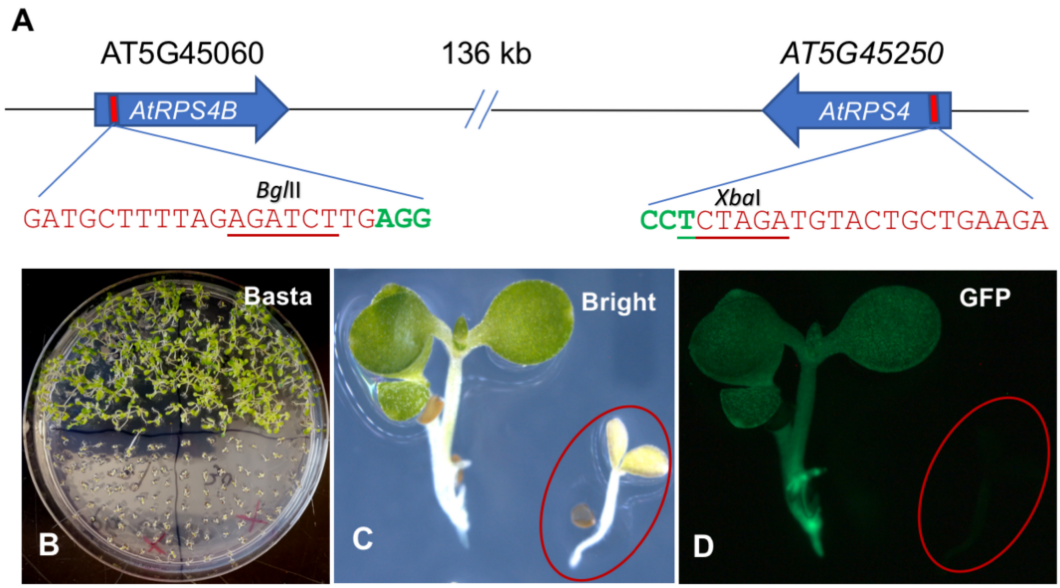


FIGURE 4 | Generation of Arabidopsis plants carrying pECp-Cas9/2sgRNA constructs targeting *AtRPS4* and *AtRPS4B*. **(A)** Schematic representation of *AtRPS4* and *AtRPS4B* and their gRNA target sites. Nucleotides in red and green indicate gRNA target sites and the PAM sequence, respectively. Underlined nucleotides represent the restriction enzyme site in the target gene used to detect mutations. **(B)** Plants grown in MS solid medium with Basta. **(C,D)** Transgenic plants under the bright and fluorescent (GFP) light, respectively. circles indicate the location of a seedling not obviously visible under fluorescent light in **(D)**.

by the PCR approach (**Supplementary Table S2**). These results indicated that transgenic plants could be best identified by using the two methods together.

We next examined the extent of gene editing in the T1 populations by PCR-amplifying the relevant genomic regions and digesting the amplicons with the restriction enzymes *Xba*I for *AtRPS4* and *Bgl*II for *AtRPS4B* (PCR-RE approach) followed by Sanger sequencing some of the randomly selected amplicons. For AtEC1.2e1.1p, among 20 transgenic T1 lines selected, 3 lines contained mutations at one target locus (*AtRPS4*) (**Supplementary Table S2**). Two mutations were heterozygous, while one was bi-allelic. The frequency of mutations for *AtRPS4* was 15%, suggesting that the T-DNA transferred into zygotes of Arabidopsis flowers started to function at an early stage. No mutation was detected at the *AtRPS4B* target locus, suggesting a low efficiency of the guide-RNA for *AtRPS4B*. However, no mutation was detected at the two target loci with the three constructs containing other three ECp driving Cas9 expression. These results indicate that site-specific mutants can be obtained in the T1 generation with one of four egg cell promoters.

Since ECp:Cas9/gRNA should continue to function in egg cells, zygotes and the early stage embryos in the reproductive stage of transgenic plants in the T1 or successive generations, we focused on detecting mutations at the target loci in the T2 generation derived from each CRISPR construct. We counted T2 plants that survived Basta selection in MS medium and with GFP fluorescence and calculated the segregation ratio from individual T1 lines (**Supplementary Table S3**). The T2 plants from several T1 lines with a 3:1 segregation ratio for each Cas9/gRNA construct were selected for genotyping using the PCR-RE approach (**Figure 5**). For AtEC1.2e1.1p, the mutation frequencies ranged from 18.2 to 63.6% for *AtRPS4* and from 0 to 18.2% for *AtRPS4B* (**Table 1**). Genotyping and sequencing analysis of T2 plants showed that AtP5p, GmEC1.1p and GmEC1.2p induced site-specific mutations at 12.8, 25.2, and 0%, respectively, for *AtRPS4*, but induced no mutation at the *AtRPS4B* target locus (**Table 1**). The AtEC1.2e1.1p:Cas9 construct induced the highest number of mutations at both loci in comparison with the other egg cell-specific promoters. These results demonstrated that AtEC1.2e1.1p was the most efficient promoter compared to AtP5p, GmEC1.1p and GmEC1.2p for gene editing in Arabidopsis.

We also determined the exact identities of edits in individual T2 plants from the single T1 lines. Sanger sequencing was performed for seven plants from line #18 of AtEC1.2e1.1p:Cas9 construct, and five plants contained single mutations for *AtRPS4* and two contained double mutations for *AtRPS4* and *AtRPS4B*. A variety of mutations (heterozygous, homozygous and bi-allelic) occurred in the *AtRPS4* single mutants (**Table 2**). The results demonstrate that multiple independent mutations can be obtained from just a single T1 line, which could be useful for generating allelic series.

The GFP fluorescence and PCR-RE approach were also used to select transgene-free mutants in the T3 generation (**Figure 6**). Through genetic segregation, non-GFP (transgene-free) T3 plants could be recovered that carried mutations in either *AtRPS4*, *AtRPS4B* or both genes. These results suggest that it is feasible to use visible GFP fluorescence to remove the transgenic plants, then use the PCR-RE approach to identify the mutants from the previous generation through genetic segregation. This strategy can be used to obtain transgene-free mutants from either the T2 or T3 generation.

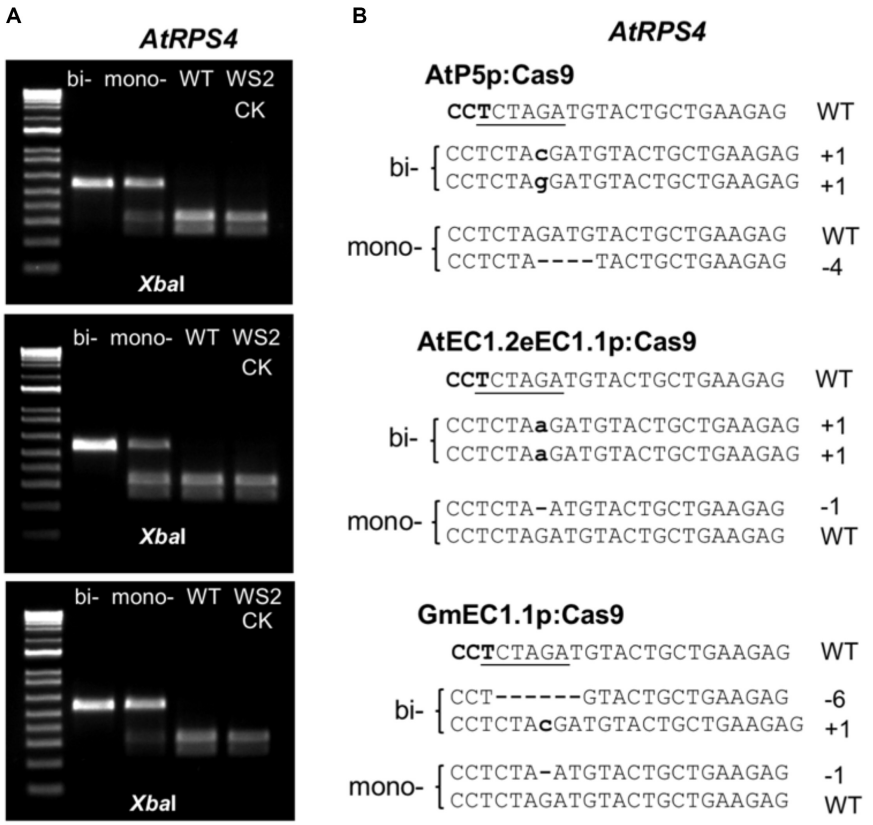


FIGURE 5 | Genotyping Arabidopsis T2 progeny with the PCR-RE and sequencing approaches. **(A)** Gel images of PCR-amplicons from the segregating T2 progeny along with parent WS2 digested with *Xba*I. **(B)** Sequencing results of PCR amplicons derived from the plants as shown in **(A)**. bi-, biallelic mutants; mono-, monoallelic mutants; WT, wild type segregants; WS2 CK, parent control.

TABLE 1 | Efficiency of targeted mutagenesis by Cas9 driven by four different promoters in T2 generation of Arabidopsis.

Promoter for Cas9	T1 Line	*AtRPS4* number of plants	Editing efficiency	*AtRPS4B* number of plants	Editing efficiency	Total T2 plants
AtEC1.2e1.1p	#18	WT(4)/Heter-(3)/Homo-(2)/Bi-(2)	63.6%	WT(9)/Heter-(1)/Bi-(1)	18.2%	11
	#21	WT(4)/Heter-(3)/Bi-(2)	55.6%	WT(9)	0%	9
	#26	WT(8)/Heter-(1)/Bi-(1)	20%	WT(9)/Heter-(1)	10.0%	10
	#42	WT(9)/Bi-(2)	18.2%	WT(10)/Heter-(1)	9.0%	11
	#46	WT(3)/Heter-(1)/Bi-(3)	57.1%	WT(7)	0%	7
AtP5p	#5	WT(15)	0%	WT(15)	0%	15
	#11	WT(8)/Heter-(2)	20.0%	WT(10)	0%	10
	#20	WT(7)/Heter-(3)/Bi-(1)	36.4%	WT(11)	0%	11
	#27	WT(12)	0%	WT(12)	0%	12
	#41	WT(12)/Heter-(1)	7.8%	WT(13)	0%	13
GmEC1.1p	#10	WT(9)/Heter-(3)	25.0%	WT(12)	0%	12
	#36	WT(13)	0%	WT(13)	0%	13
	#42	WT(5)/Heter-(4)/Bi-(3)	63.6%	WT(12)	0%	12
	#52	WT(7)/Heter-(1)	12.5%	WT(8)	0%	8
	#72	WT(9)/Heter-(2)/Bi-(1)	25.0%	WT(12)	0%	12
GmEC1.2p	#2	WT(27)	0%	WT(27)	0%	27
	#3	WT(24)	0%	WT(24)	0%	24
	#4	WT(4)	0%	WT(4)	0%	4
	#5	WT(19)	0%	WT(19)	0%	19

Hetero-, heterozygous mutants; Bi-, biallelic mutants; WT, wildtype segregants.

TABLE 2 | The representative mutations induced by AtEC1.2e1.1p:Cas9 in T2 plants of #18.

	AtRPS4		*AtRPS4B*	
WS2	**CCT**CTAGATGTACTGCTGAAGAG	WT	GTGATGCTTTTAGAGATCTTG**AGG**	WT
AtEC #18-1	CCTCTA**g**GATGTACTGCTGAAGAG	+1	GTGATGCTTTTAGAGATCTTGAGG	WT
	CCTCTA**g**GATGTACTGCTGAAGAG	+1	GTGATGCTTTTAGAGATCTTGAGG	WT
AtEC #18-2	CCTCTA–TGTACTGCTGAAGAG	−2	GTGATGCTTTTAGAGATCTTGAGG	WT
	CCTCTAGATGTACTGCTGAAGAG	WT	GTGATGCTTTTAGAGATCTTGAGG	WT
AtEC #18-3	CCTCTA-ATGTACTGCTGAAGAG	−1	GTGATGCTTTTAGAGATCTTGAGG	WT
	CCTCTA-ATGTACTGCTGAAGAG	−1	GTGATGCTTTTAGAGATCTTGAGG	WT
AtEC #18-4	CCTCTA**a**GATGTACTGCTGAAGAG	+1	GTGATGCTTTTAGAGATCTTGAGG	WT
	CCTCTAGATGTACTGCTGAAGAG	WT	GTGATGCTTTTAGAGATCTTGAGG	WT
AtEC #18-5	CCTCTA**a**GATGTACTGCTGAAGAG	+1	GTGATGCTTTTAGAGATCTTGAGG	WT
	CCTCTA-ATGTACTGCTGAAGAG	−1	GTGATGCTTTTAGAGATCTTGAGG	WT
AtEC #18-6	CCTCTA-ATGTACTGCTGAAGAG	−1	GTGATGCTTTTAGAGAT–TGAGG	−2
	CCTCTAGATGTACTGCTGAAGAG	WT	GTGATGCTTTTAGAGATC**t**TTGAGG	+1
AtEC #18-7	CCTCTA**c**GATGTACTGCTGAAGAG	+1	GTGATGCTTTTAGAGATCTTGAGG	WT
	CCTCTA**g**GATGTACTGCTGAAGAG	+1	GTGATGCTTTTAGAGAT-TTGAGG	−1

The green letters are PAM sequences of Cas9 target sites. Lower letters are insertions and dashed lines are deletions induced by Cas9.

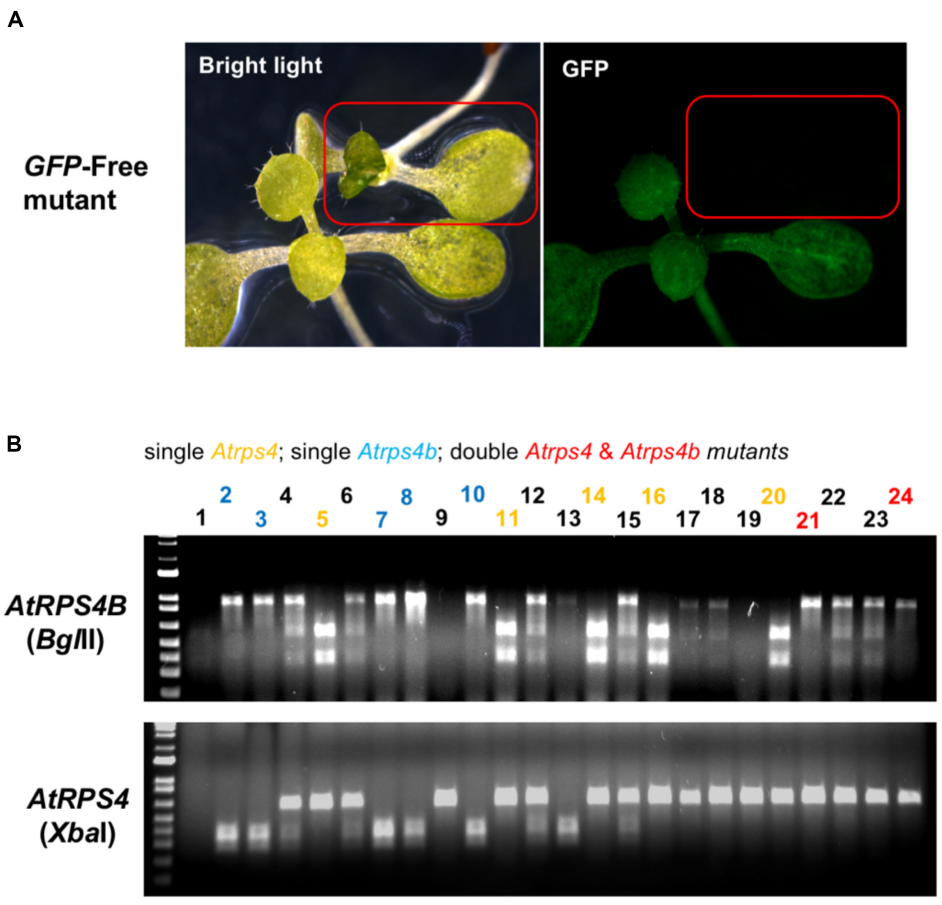


FIGURE 6 | Selection of transgene-free, mutant Arabidopsis plants by absence of GFP fluorescence and PCR-RE genotyping. **(A)** T3 T-DNA-free edited plants can be selected under the fluorescence microscope. **(B)** Gel electrophoresis images of PCR-amplicons digested with restriction enzymes (*Bgl*II for *AtRPS4B* and *Xba*I for *AtRPS4*). Homozygous single and double mutants were obtained in the T3 generation.

ECp:Cas9/gRNA Constructs Induce Mutagenesis of Two *GmAGO7* Genes in Soybean

To test whether our ECp-Cas9/gRNA systems could introduce mutations at genomic target loci in soybean, we chose *GmAGO7a* (Glyma.01G053100) and *GmAGO7b* (Glyma.02G111600) for targeted mutagenesis. In plants, ARGONAUTE (AGO) proteins are associated with small RNA (sRNA) mediated repression of gene expression through either direct cleavage or other mechanisms, such as target destabilization or translational repression (Carbonell et al., 2012; Meister, 2013). ARGONAUTE7 (AGO7), a key regulator in the trans-acting small interfering RNAs (ta-siRNA) pathway, plays a conserved role in controlling leaf pattern among species. In Arabidopsis, the *ago7* mutants display increased leaf length and downward-curled leaf margin due to accelerated juvenile-to-adult transition; however, the mutants did not show obvious defects in leaf polarity (Hunter et al., 2003, 2006; Fahlgren et al., 2006). Overexpressing *SlAGO7* in tomato exhibited pleiotropic phenotypes, including improved axillary bud formation, altered leaf morphology and inflorescence architecture, and increased fruit yield (Lin et al., 2016). Loss-of-function of *Mtago7* resulted in lobed leaf margins and more widely spaced lateral organs in *Medicago truncatula* (Zhou et al., 2013). There are two *AGO7* paralogous genes (*GmAGO7a* and *GmAGO7b*) in the annotated Williams 82 genome[2].

We first verified the efficiency of the CRISPR-Cas9 system and activity of the gRNA in the soybean hairy root system. The CaMV double 35S (2x35S) promoter was used to drive expression of Cas9, and two gRNAs were designed to target a unique sequence in each gene. Cassettes expressing two different pairs of gRNAs targeting *GmAGO7a* and *GmAGO7b* – gAGO7a1 and gAGO7b2 (named gAGO7a1/b2) and gAGO7a2 and gAGO7b1 (named gAGO7a2/b1) – were constructed and each was recombined into the Cas9 binary vector, resulting in constructs that simultaneously targeted *GmAGO7a* and *GmAGO7b* (**Figure 7**).

The resulting two constructs were individually introduced into *Agrobacterium rhizogenes* strain K599 to induce hairy roots from infected soybean cotyledons. Hairy roots from individual cotyledons were collected and pooled for genomic DNA extraction. Individual DNA samples were used to PCR-amplify the relevant regions using the gene-specific primers for *GmAGO7a* and *GmAGO7b*. The T7 endonuclease I (T7E1) analysis, an assay that involves using T7E1 to digest the PCR amplicons of target genomic loci (Guschin et al., 2010), revealed a range of 80 to 100% mutation frequencies for gAGO7a1 and gAGO7a2, about 20% for gAGO7b1 and no detected mutation for gAGO7b2 (**Supplementary Figure S2**). The results indicated that the gRNAs (gAGO7a1 and gAGO7a2) for *GmAGO7a* and their promoters (U6-26 and U3b) were highly active, but the gAGO7b1 and gAGO7b2 gRNAs were less active.

To validate the ability of ECp:Cas9/gRNA to induce mutagenesis and examine the correlation of gRNA activities between hairy roots and stable transgenic plants, we used five constructs expressing Cas9 under different egg cell-specific promoters and gRNAs (gAGO7a2/b1 and gAGO7a1/b2) (pAtP5p:Cas9-gAGO7a1/b2, pAtP5p:Cas9-gAGO7a2/b1, pAtEC1.2e1.1p:Cas9-gAGO7a2/b1, pGmEC1.1p:Cas9-gAGO7a2/b1, and pGmEC1.2p:Cas9-gAGO7a2/b1) for stable soybean transformation. These five constructs were introduced individually into the Williams 82 cultivar of soybean by the *A. tumefaciens*-mediated transformation method. The GFP fluorescence was used to further screen the transgenic plants from the regenerated plants that survived from the herbicide glufosinate (**Figures 8A–D**). In addition, PCR-amplifications of Cas9 and gRNA genes were used to identify the transgenic plants (**Figure 8E**). Plants were screened by resistance to glufosinate (*bar*), GFP fluorescence and PCR-amplification, and 2 and 3 independent T0 transgenic plants were obtained for AtP5p:Cas9-gAGO7a1/b2 and AtP5p:Cas9-gAGO7a2/b1, respectively (**Table 3**). T0 plants were grown to maturity so that we could determine the mutation efficiency in the T1 generation. The three T0 AtP5p:Cas9+gGmAGO7a2/b1 lines (ST397-1, -4, and -5), produced transgenic progeny, but none that were tested carried a mutation at either gRNA target sites. As expected, the T0 AtP5p:Cas9-gAGO7a2/b1 plants (ST397-2 and -3) that survived Bar selection but were negative for GFP fluorescence and the PCR tests for the Cas9 and gRNA transgenes produced no transgenic progeny and carried no mutations at the *GmAGO7a* or *GmAGO7b* target sites (**Table 3**).

For the AtEC1.2e1.1p, GmEC1.1p and GmEC1.2p constructs, 13, 2, and 1 independent T0 transgenic lines were positive for glufosinate screen (*bar*) and the PCR amplification, respectively (**Table 3**). Out of the 13 T0 lines from the AtEC1.2e1.1p:Cas9-gAGO7a2/b1 construct, 10 lines produced transgenic T1 progeny, and of these, five lines produced T1 progeny some of which carried mutations in *GmAGO7a*, but none had mutations in *GmAGO7b*. Among the 97 T1 progeny plants from the five lines, 26 plants contained mutations for *GmAGO7a* a mutation efficiency of 26.8% based on the T7E1 assay (**Table 3**). For other non-transgenic plants, we also checked for occurrence of the mutation at both target loci, and as expected no mutation was found in these plants. For the GmEC1.1p:Cas9-gAGO7a2/b1, two transgenic lines (ST411-12, -18) produced transgenic progeny, but no mutation was detected at either target site. Unfortunately, no seed was produced from the only transgenic line derived from GmEC1.2p:Cas9-gAGO7a2/b1. These results indicate that the AtEC1.2e1.1p promoter is likely the most efficient for gene editing in soybean and that the frequencies of mutations induced by the gRNA were consistent with the hairy root system.

We expected that Cas9 would be activated in egg cells, zygotes or early embryos beyond the T1 generation. Therefore, we further examined the prevalence of mutations in T2 plants produced from T1 lines that were derived from AtEC1.2e1.1p:Cas9-gAGO7a2/b1 and carried the Cas9 and gRNA transgenes. A total of 31 T1 lines produced from 6 T0 lines were selected for further analysis in the T2 generation. Twenty-two T1 lines produced T2 plants that carried inheritable transgenes and mutations, while seven T1 lines (ST410-14-19, ST410-25-22,

[2]phytozome.jgi.doe.gov

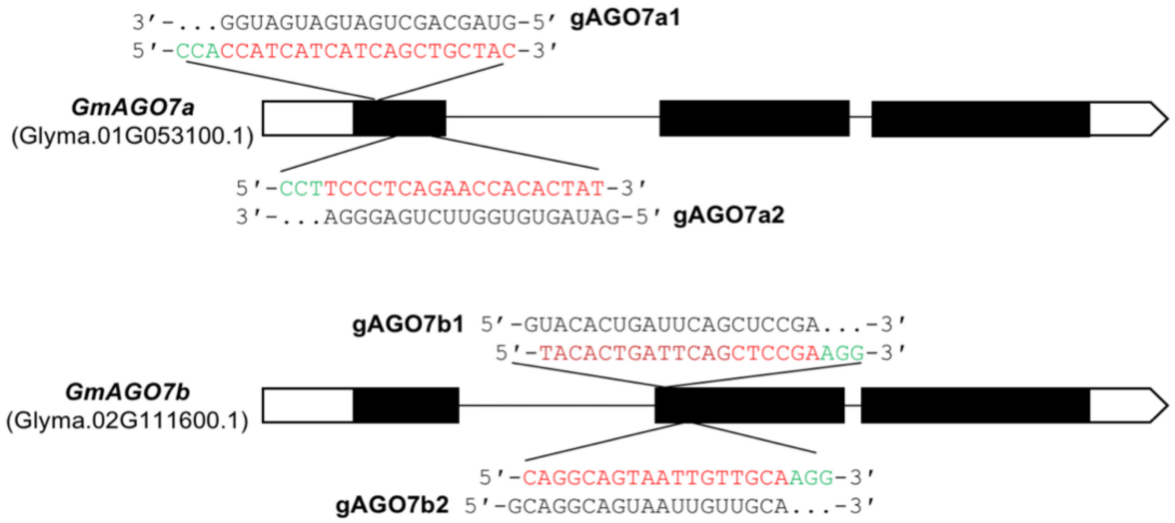


FIGURE 7 | gRNA target sites in *GmAGO7a* and *GmAGO7b*. The structure of the *GmAGO7a* and *GmAGO7b* genes are represented by exons (black bars), introns (black lines) and UTRs (white bars). The sequences of the target sites are shown in red, PAM sequences in green and gRNA sequences in black. A total of four gRNAs were constructed into two Cas9 plasmids, one containing gAGO7a1 and gAGO7b2 (gAGO7a1/b2) and another containing gAGO7a2 and gAGO7b1 (gAGO7a2/b1).

ST410-36-2, ST410-36-4, ST410-36-9, ST410-36-11, ST410-44-6) did not produce any seed (**Table 4**). A total of 145 out of 389 T2 plants tested positive for mutations in *GmAGO7a* by using the T7E1 assay. Sanger sequencing the PCR-amplicons from some of the T7E1 positive plants revealed site-specific mutations in *GmAGO7a* (**Figure 9**). Most mutants were heterozygous, but 11 were homozygous from T1 lines ST410-4-5, ST410-4-14, ST410-25-4, ST410-44-5) (**Table 4**). We found only one mutant for *GmAGO7b* among 389 T2 plants (ST410-14-3-7) (**Table 4**). We were unable to correlate a mutant phenotype with homozygous mutations in *GmAGO7a*, suggesting that there is a redundant function provided by *GmAGO7b*. The results indicate that Cas9 and gRNAs are still active in egg cells, zygotes and early embryos in the T1 and T2 generations. For construct GmEC1.1p:Cas9-gAGO7a2/b1, eight T1 lines were selected for analyses of transgene presence and occurrence of mutations in the T2 generation. No mutation was detected in either *GmAGO7a* or *GmAGO7b* among 140 T2 plants (**Table 4**). These results demonstrated that the AtEC1.2e1.1p promoter is the best promoter for egg cell-specific Cas9 expression and genome editing in Arabidopsis and soybean.

DISCUSSION

In the present study, we built a CRISPR/Cas9 platform for genome editing in soybean with the aid of proof-of-concept experiments in Arabidopsis. The system depends on the conditional expression of Cas9 at the reproductive stage by using an egg cell-specific promoter. The use of tissue-specific promoters is expected to have multiple benefits that includes reducing the potential toxicity associated with Cas9 expressed under the otherwise strong and constitutive promoters. In addition, the expression of Cas9 in germ cells (egg cells, zygotes, and early embryos) results in heritable edits and reduces somatic mutations in plants derived from organogenesis, which has proven to be a feasible approach particularly in Arabidopsis (Wang et al., 2015; Yan et al., 2015; Mao et al., 2016). Our CRISPR system also features a visual marker (GFP) under a soybean ubiquitin gene promoter that can be used to avoid growing and maintaining escapes from soybean transformation projects. The system can be used for multiplex genome editing as up to 6 gRNA cassettes can be integrated into a single construct. Finally, we validated the system for site-specific mutagenesis by targeting two genes each in soybean and Arabidopsis up to the T2 and T3 generations, respectively.

Genome editing in soybean still faces several challenges although numerous studies have demonstrated its feasibility. First, efficient and reliable transformation technology in soybean has lagged behind some other crop species. Genetic transformation of soybean is still genotype dependent, and only a very limited number of genotypes are transformable (Yamada et al., 2012; Altpeter et al., 2016). The most widely and routinely used transformation platforms depend on organogenesis, a process that is based on cotyledonary-nodes as explants and *Agrobacterium*-mediated DNA delivery, to regenerate whole plants (Paz et al., 2006). This method often leads to a high occurrence of escape and chimeric plants, and it makes soybean transformation labor-intensive and expensive. The number of genuine transgenic plants is usually low from each transformation project. We observed evidence for escapes and chimerism during analyses of T0 and T1 plants in which some T0 plants that were Bar positive surviving glufosinate selection, but then tested negative for GFP and/or the Cas9 and gRNA transgenes (**Table 3**). In addition, some T0 lines that tested positive for Bar, GFP, and the Cas9 and gRNA transgenes were

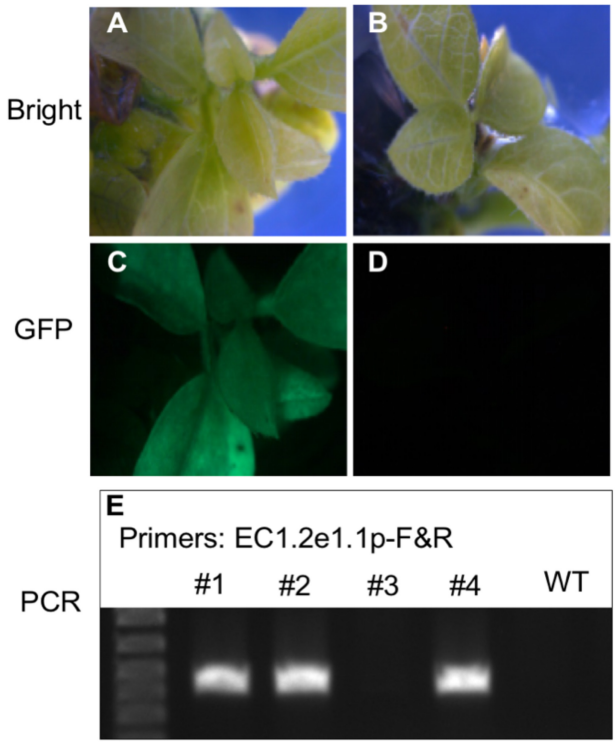


FIGURE 8 | Detection of the transgenic soybean plants using GFP fluorescence and PCR-based genotyping. **(A–D)** Use of GFP fluorescence to detect transgenic soybean plants. **(A,C)** Images of a transgenic plant expressing GFP under bright and fluorescent light, respectively. **(B,D)** Images of a negative control plant under bright and fluorescent light. **(E)** Gel electrophoresis image of the PCR amplicons produced using T-DNA-specific primers targeting the promoter region.

likely chimeric, because their progeny neither inherited the transgenes nor carried mutations in the target genes (**Table 3**). Our soybean CRISPR/Cas9 system, therefore, is ideal, because it allows us to select the Cas9/gRNA positive plants that can be used to generate many progenies that each contain independent and inheritable mutated alleles.

We tested four different EC promoters and found only one (AtEC1.2e1.1p) was capable of inducing mutations in progenies of T1 and T2 generations in soybean. There are possible explanations for the failure of the other three promoters. First, Cas9 was expressed but not as levels sufficient to cause DSBs within the context of chromatin during the reproductive stage of soybean plants. It is worth to point out that two promoters (AtP5p and GmEC1.1p) are expected to be active, because they did enable Cas9 to induce mutagenesis in Arabidopsis. However, the third promoter (GmEC1.2p) may not be functional, because Cas9 failed to induce mutagenesis in both Arabidopsis and soybean. Future work is needed to determine whether Cas9 mRNA and protein expression is activated by the three promoters. Second, the guide RNAs used in the present study may not be the most active ones. More guide RNAs remain to be designed and tested. Finally, the number of T0 transgenic soybean plants derived from each of the three constructs was small, which is a common limitation of soybean transformation. We obtained only eight, four and four T0 plants from AtP5p:Cas9, GmEC1.1p:Cas9, and Gm1.2p:Cas9, respectively, while 14 T0 plants were generated from AtEC1.2e1.1p:Cas9. The low number of T0 plants prevents us from making definitive conclusions about the three promoters that did not induce mutations in soybean. Future work to increase the efficiency of soybean transformation will make it possible to generate larger numbers of T0 plants with constructs carrying more gRNA and Cas9 configurations.

In the present study, we tested the correlation of gRNA activities among hairy roots and stable transgenic plants. Hairy roots yielded very high frequencies of mutagenesis for two of four gRNAs (gAGO7a1 and gAGO7a2) and relatively low or no activity by another two (gAGO7b1 and gAGO7b2). Similarly, the frequency of mutations induced by gAGO7a2 is much higher than gAGO7b1 in the stable transgenic plants. The findings suggest that the hairy root system is a quick way to test and select highly active gRNA before proceeding to stable transformation.

The second challenge facing genome editing in soybean is the genome size and complexity. Soybean, e.g., Williams 82, is a palaeopolyploid and contains a genome of ∼1.15 gigabases (Schmutz et al., 2010), about eight times as larger than the *A. thaliana* genome (135 megabases) (Initiative, 2000). Prior to site-specific cleavage of the target DNA, Cas9 and gRNA first recognize the PAM sequence and then the PAM-proximal region through the complementarity between the gRNA and target strand (Sternberg et al., 2014). The implication of this mode of recognition is that the genome editing efficiency is expected to be negatively correlated to the genome size and complexity. Our data from Arabidopsis and soybean supports this notion. For instance, three of the egg cell-specific promoters driving Cas9 could produce mutants in Arabidopsis, while only one of three was able to do so in soybean at much lower percentage. Alternatively, the expression levels of Cas9 and gRNA may be higher in Arabidopsis than in soybean.

It is highly likely the editing efficiency would be correlated with the abundance of Cas9/gRNA ribonucleoprotein complexes. Based on this, to further improve our soybean CRISPR/Cas9 system, several strategies and approaches can be explored. First, we are interested in approaches to increase the abundance of Cas9 in egg cells, zygotes and early embryos of the T0 or T1 generation transgenic plants. For example, it was shown that zCas9 (maize codon-optimized Cas9) coding sequence followed by the pea ribulose-1,5-bisphosphate carboxylase small subunit (*rbcS*) E9 gene terminator (rbcS E9t) induced edits at higher rates than zCas9 followed by the *Agrobacterium Nos* gene terminator in Arabidopsis (Wang et al., 2015). The constructs described in the present study all carried Cas9 followed by the *Nos* terminator. In addition, the incorporation of introns in gene coding regions especially at the 5′ end was shown to enhance gene expression (Le Hir et al., 2003). Therefore, use of a different terminator (e.g., pea rbcS E9 or soybean rbcS1 terminator) along with soybean codon-optimized Cas9+intron could increase expression level of Cas9 transcripts expressed under control of the EC promoters. It may also be possible to use soybean U6 promoters for expressing gRNAs to improve the

TABLE 3 | Genotyping analysis of T0 and T1 transgenic soybean plants from five CRISPR constructs.

		T0				T1	
CRISPR Construct	**Line**	***Bar***	***GFP***	***Cas9***	**gRNA**	**Transgenic vs. (Total plants)**	**# of mutant in *GmAGO7a***
AtP5p:Cas9 (gAGO7a1/b2)	ST398-4	+	+	+	+	0(19)	None
	ST398-5	+	+	+	+	0(27)	None
AtP5p:Cas9 (gAGO7a2/b1)	ST397-1	+	+	+	+	11(26)	None
	ST397-2	+	–	–	–	0(27)	None
	ST397-3	+	–	–	–	0(25)	None
	ST397-4	+	+	+	+	11(27)	None
	ST397-5	+	+	+	+	17(27)	None
	ST397-10	+	–	–	–	T0 died	
AtEC1.2e1.1p:Cas9 (gAGO7a2/b1)	ST410-4	+	ND	+	+	6(15)	2
	ST410-7	+	ND	+	+	1(18)	None
	ST410-8	+	ND	+	+	5(15)	None
	ST410-9	+	ND	-	-	T0 died	
	ST410-11	+	ND	+	+	0(17)	None
	ST410-13	+	ND	+	+	2(2)	None
	ST410-14	+	ND	+	+	21(23)	7
	ST410-18	+	ND	+	+	3(3)	None
	ST410-25	+	ND	+	+	30(36)	4
	ST410-32	+	ND	+	+	11(15)	None
	ST410-35	+	ND	+	+	0(14)	None
	ST410-36	+	ND	+	+	12(12)	7
	ST410-40	+	ND	+	+	0(17)	None
	ST410-44	+	ND	+	+	6(6)	6
GmEC1.1p:Cas9 (gAGO7a2/b1)	ST411-6	+	ND	–	–	0(6)	None
	ST411-12	+	ND	+	+	8(18)	None
	ST411-17	+	ND	–	–	0(9)	None
	ST411-18	+	ND	+	+	4(11)	None
GmEC1.2p:Cas9 (gAGO7a2/b1)	ST412-20	+	ND	+	+	T0 died	
	ST412-29	+	ND	–	–	T0 died	
	ST412-27	+	ND	–	–	T0 died	
	ST412-31	+	ND	–	–	0(14)	None

ND, no GFP detection was performed.

genome editing efficiency (Di et al., 2019). Such modifications to increase the levels of activated Cas9/gRNA complexes in germline cells are expected to increase the frequency of targeted mutagenesis without sacrificing germline-specific expression.

MATERIALS AND METHODS

Plant Materials

Arabidopsis thaliana (ecotypes Wassilewskija-2, Ws-2) plants were grown vertically on half-strength Murashige and Skoog (1/2 MS) plates at pH 5.6–5.8 (adjusted with 1 N KOH), supplemented with 0.85% (*w*/*v*) agar and 1% (*w*/*v*) sucrose. All plants were grown at 22°C under long-day conditions (16 h light/8 h dark).

The soybean cultivar Williams 82 was used for hairy root transformation and whole plant stable transformation. Soybean seeds were surface-sterilized for 14–16 h with chlorine gas. Seeds were germinated on 1/4 Gamborg's solid medium under long-day conditions (16-h light/8-h dark, 28°C/24°C) in growth chambers. After 5 days, healthy plants were selected for hairy root transformation. Soybean plants were grown in clay pots containing Pro-mix Bx Biofungicide potting mix supplemented with Osmocote slow-releasing fertilizer (14-14-14) and were grown in greenhouse to maturity under a photoperiod of 16 h light and 8 h dark at 28–24°C.

Vector Construction

Different ECp were used to drive Cas9 expression in this study. The Arabidopsis AtP5 promoter (919 bp, Chr1:26929821. . .26930831 in TAIR 10) fused to Cas9 (pCambia:AtP5p:Cas9) was constructed in pCAMBIA3300. This vector was modified by inserting the cassette of *att*R1-ccdB-*att*R2, which is used to insert gRNA modules by Gateway recombination, resulting in the destination vector, pGW-AtP5p:Cas9. The AtEC1.2e1.1p ECp (1,362 bp, Chr2:9282423. . .9283302 + Chr1:28810535. . .28811064 in TAIR 10) was assembled through overlapping PCR amplification using Arabidopsis genomic DNA and primers. To identify soybean

TABLE 4 | Genotyping analysis of CRISPR/Cas9-induced mutations in T1 and T2 soybean plants.

CRISPR construct	T0 Line	T1		T2		
		No.	Mutation (*GmAGO7a*)	Mutant plants (*GmAGO7a*)	Mutant plants (*GmAGO7b*)	Total plants
AtEC1.2e1.1p:Cas9 (gAGO7a2/b1)	ST410-4	#3	–	0	0	21
		#4	–	0	0	18
		#5	+	14 (3 Homo-)	0	20
		#6	–	0	0	19
		#14	+	11 (2 Homo-)	0	20
		#15	–	0	0	15
	ST410-14	#2	+	7	0	16
		#3	+	0	1	15
		#6	+	9	0	15
		#19	+	ND	ND	ND
	ST410-25	#3	+	7	0	16
		#4	+	9 (2 Homo-)	0	14
		#8	+	8	0	13
		#22	+	ND	ND	ND
	ST410-32	#1	–	0	0	17
		#2	–	0	0	20
		#5	–	0	0	17
		#12	–	0	0	18
	ST410-36	#1	+	1	0	14
		#2	+	ND	ND	ND
		#3	+	4	0	16
		#4	–	ND	ND	ND
		#6	+	9	0	19
		#9	+	ND	ND	ND
		#11	+	ND	ND	ND
	ST410-44	#1	+	8	0	8
		#2	+	9	0	9
		#3	+	10	0	10
		#4	+	19	0	19
		#5	+	20 (4 Homo-)	0	20
		#6	+	ND	ND	ND
Sum	**6**	**31**	**22**	**145**	**1**	**389**
GmEC1.1p:Cas9 (gAGO7a2/b1)	ST411-12	#2	–	0	0	19
		#5	–	0	0	17
		#17	–	0	0	18
		#18	–	0	0	18
	ST411-18	#6	–	0	0	18
		#12	–	0	0	18
		#14	–	0	0	18
		#20	–	0	0	14
Sum	**2**	**8**	**0**	**0**	**0**	**140**

ND, no seeds or DNA samples were collected.

ECp, the Arabidopsis egg cell secreted protein 1.1 (AT1G76750) was used as a query to BLAST search against the soybean reference genome to retrieve seven egg cell secreted protein sequences and their promoter sequences. Promoters of GmEC1.1 (1,446 bp, Chr20:40598179. . .40599767 in Wm82.a2v1) and GmEC1.2 (1,334 bp, Chr06:18296898. . .18298365 in Wm82.a2v1) were PCR-amplified using soybean genomic DNA as a template and gene-specific primers. GmEC1.1 and GmEC1.2 each represent one member of their clades among the seven GmECs (**Supplementary Figure S3**). To construct a Cas9 expression cassette with the 2x35S promoter, the 2x35S promoter was PCR-amplified from the pTF101 plasmid. All the promoters were swapped with the AtP5p in pCambia:AtP5p:Cas9 at *Nco*I and *Spe*I through the Gibson cloning method, resulting in constructs containing each of the four different promoters for expression of Cas9. The GmUbi-GFP-NosT cassette was then inserted into each vector through restriction and ligation at the *Hind*III site using the standard molecular cloning

FIGURE 9 | Representative mutations induced by egg cell-specific expression of Cas9 in soybean plants derived from five independent T0 lines. Red letters and underlined letters in Williams 82 indicate the respective PAM and gRNA target sequences. Dashed lines and lowercase letters denote deletions and insertions, respectively, at the *GmAGO7a* target site.

methods (Ausubel et al., 1998), resulting in the GFP version of the plasmids. The primer information is provided in **Supplementary Table S4**.

For the construction of gRNA genes, the intermediate vectors pENTR4-ccdB which was modified by inserting the cassette of *att*L1-ccdB-*att*L2. Briefly, each gRNA vector (pCRgRNA) has a unique promoter and poly-T terminator. The specific gRNA spacer sequence was inserted at the two *Bsm*BI restrict sites. To construct a specific gRNA gene, two 20 – 24 nt complementary oligonucleotides were annealed to produce a double-stranded DNA oligonucleotide (dsOligo). To make a 2-gRNAs cassette, the first dsOligo was designed with a 5′ overhang of ATTG in the sense strand and a 5′ of overhang AAAC in the antisense strand; and the second dsOligo was designed to contain a 5′ overhang of GTCA in the sense strand and a 5′ of overhang AAAC in the antisense strand. All oligonucleotides were synthesized and purchased from Integrated DNA Technology (Coralville, IA, United States). The individual gRNA cassettes were assembled into the pENTR4-ccdB vector by the Golden Gate assembly method using *Bsa*I. After sequencing the guide RNA regions, the gRNA cassette was finally mobilized to four different pGW-ECp:Cas9-GFP-ccdB constructs by using Gateway LR Clonase (Thermo Fisher Scientific, Waltham, MA, United States).

Escherichia coli strain DH5α and DB3.1 were used for molecular cloning of Cas9/gRNA constructs. *Agrobacterium tumefaciens* strain GV3101 and EHA105 were used for Arabidopsis and soybean transformation, respectively. *A. rhizogenes* strain K599 was used for the soybean hairy root transformation. *E. coli* cells were grown in Luria–Bertani (LB) medium supplemented with appropriate antibiotics at 37°C with a standard culture technique, while *Agrobacterium* strains were grown at 28°C in Luria–Bertani (LB) medium with appropriate antibiotics (Ausubel et al., 1998).

Hairy Root Transformation

Agrobacterium rhizogenes strain K599 containing the Cas9/gRNA binary constructs was used for the soybean hairy root induction. Soybean cotyledons of cultivar Williams 82 were inoculated with the transformed K599 strain using a previously described protocol (Kereszt et al., 2007) with slight modifications. Bacterial cells were scraped from the plates and suspended in 1 mL sterile water. The blades of sterilized scissors were immersed into the bacterial suspension, then used to cut off the 1/4 of the cotyledon that was attached to the stem. The cut cotyledons were placed on a stack of one or two sterile pre-wet paper towels (with 1/4 Gamborg's liquid medium plus 200 μg/ml Timentin) in an ice-cream box. The cotyledons were kept in a growth chamber with a photoperiod of 16-h light/8-h dark at 28°C/24°C. After cultivation for 2∼3 weeks, hairy roots were collected for further analysis. Hairy roots induced by K599 lacking CRISPR construct were used as control.

Transformation of Arabidopsis and Soybean

Arabidopsis transformation was performed by using the floral dipping protocol as described (Pike et al., 2019). *Agrobacterium*-mediated transformation of a fixed number of half seed explants of Williams 82 genotype was performed at the Iowa State University Plant Transformation Facility as described (Paz et al., 2004). The plants were grown in greenhouses with a 16-h day at 28°C and an 8-h night at 24°C. The transgenic plants were further confirmed by GFP fluorescence and PCR assays on the Cas9 and gRNA transgenes.

Molecular Characterization of CRISPR Plants of Arabidopsis and Soybean

Genomic DNA samples were extracted from Arabidopsis leaves using the CTAB (cetyltrimethyl ammonium bromide) method (Murray and Thompson, 1980). Soybean genomic DNA was extracted from newly expanding primary leaves of T1 and T2 seedlings using the QIAamp Fast DNA Tissue Kit

(Qiagen, Germantown, MD, United States). Genomic DNA was used for PCR-amplification of relevant regions with specific primers flanking the target sites (**Supplementary Table S4**). PCR reaction conditions were optimized for each primer pair and are available upon request. PCR amplicons were assessed for mutations using the T7 endonuclease I (T7E1) assay or restrict enzyme digestion and Sanger sequencing. For the T7E1 assay, PCR-amplicons obtained from the transgenic tissues were mixed with the respective amplicon derived from wild type, denatured (95°C for 5 min) and reannealed (ramp down to 25°C at 5°C/min), then subjected to T7E1. For PCR-RE, the PCR-amplicons were used directly for digestion with appropriate restriction enzymes. For sequencing, the PCR amplicons derived from the T7E1-positive samples were treated with ExoSAP-IT (Affymetrix, Santa Clara, CA, United States) and subsequently evaluated by the Sanger sequencing method by the University of Missouri-Columbia DNA Core Facility[3]. The sequencing chromatograms were carefully examined for exact patterns that might indicate mono-allelic or bi-allelic mutations.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

[3] https://dnacore.missouri.edu

AUTHOR CONTRIBUTIONS

BY, SW, and SL designed the research. NZ, TL, JD, JS, and RL performed the experiments. DP and WG analyzed the data. NZ and BY wrote the manuscript with input from all other authors.

FUNDING

Funding was provided by the Iowa Soybean Association (SW and BY), the National Natural Science Foundation of China (31972248 to SL), and the China Scholarship Council (NZ).

ACKNOWLEDGMENTS

We thank Diane Luth at the Iowa State University Plant Transformation Facility for soybean transformation. We thank the Iowa State University Crop Bioengineering Center for publication subvention.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00800/full#supplementary-material

REFERENCES

Al Amin, N., Ahmad, N., Wu, X., Pu, T., Ma, Y., Du, X., et al. (2019). CRISPR-Cas9 mediated targeted disruption of FAD2-2 microsomal omega-6 desaturase in soybean (*Glycine max*.L). *BMC Biotechnol.* 19:9. doi: 10.1186/s12896-019-0501-2

Altpeter, F., Springer, L. E., Bartley, A. E., Blechl, T. P., Brutnell, V., Citovsky, L. J., et al. (2016). Advancing crop transformation in the era of genome editing. *Plant Cell* 28, 1510–1520. doi: 10.1105/tpc.16.00196

Ausubel, F., Brent, R., Kingston, D., Moore, J., Seidman, T., and Struhl, K. (1998). *Current Protocols in Molecular Biology*. New York, NY: John Wiley and Sons.

Bai, M., Yuan, H., Kuang, P., Gong, S., Li, Z., Zhang, B., et al. (2020). Generation of a multiplex mutagenesis population via pooled CRISPR-Cas9 in soya bean. *Plant Biotechnol. J.* 18, 721–731. doi: 10.1111/pbi.13239

Bao, A., Chen, L., Chen, S., Chen, Q., Hao, W., Guo, D., et al. (2019). CRISPR/Cas9-mediated targeted mutagenesis of GmSPL9 genes alters plant architecture in soybean. *BMC Plant Biol.* 19:131. doi: 10.1186/s12870-019-1746-6

Brazelton, V. A., Zarecor, D. A., Wright, Y., Wang, J., Liu, K., Chen, B., et al. (2015). A quick guide to CRISPR sgRNA design tools. *GM Crops Food* 6, 266–276. doi: 10.1080/21645698.2015.1137690

Cai, Y., Chen, X., Liu, C., Guo, S., Sun, C., Wu, B., et al. (2018). CRISPR/Cas9-mediated targeted mutagenesis of GmFT2a delays flowering time in soya bean. *Plant Biotechnol. J.* 16, 176–185. doi: 10.1111/pbi.12758

Cai, Y., Wang, L., Chen, T., Wu, L., Liu, S., Sun, C., et al. (2020). Mutagenesis of GmFT2a and GmFT5a mediated by CRISPR/Cas9 contributes for expanding the regional adaptability of soybean. *Plant Biotechnol. J.* 18, 298–309. doi: 10.1111/pbi.13199

Campbell, B. W., Hoyle, B., Bucciarelli, A. O., Stec, D. A., Samac, W. A., Parrott, A., et al. (2019). Functional analysis and development of a CRISPR/Cas9 allelic series for a CPR5 ortholog necessary for proper growth of soybean trichomes. *Sci. Rep.* 9:14757. doi: 10.1038/s41598-019-51240-7

Carbonell, A., Fahlgren, H., Garcia-Ruiz, K. B., Gilbert, T. A., Montgomery, T., Nguyen, J. T., et al. (2012). Functional analysis of three *Arabidopsis* ARGONAUTES using slicer-defective mutants. *Plant Cell* 24, 3613–3629. doi: 10.1105/tpc.112.099945

Cheng, Q., Dong, T., Su, T., Li, Z., Gan, H., Nan, S., et al. (2019). CRISPR/Cas9-mediated targeted mutagenesis of GmLHY genes alters plant height and internode length in soybean. *BMC Plant Biol.* 19:562. doi: 10.1186/s12870-019-2145-8

Chilcoat, D., Liu, Z. B., and Sander, J. (2017). Use of CRISPR/Cas9 for crop improvement in maize and soybean. *Prog. Mol. Biol. Transl. Sci.* 149, 27–46. doi: 10.1016/bs.pmbts.2017.04.005

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

Di, Y. H., Sun, Z., Hu, Q. Y., Jiang, G. H., Song, B., Zhang, S. S., et al. (2019). Enhancing the CRISPR/Cas9 system based on multiple GmU6 promoters in soybean. *Biochem. Biophys. Res. Commun.* 519, 819–823. doi: 10.1016/j.bbrc.2019.09.074

Do, P. T., Nguyen, H. T., Bui, L. T. N., Tran, G., Stacey, J. D., Gillman, Z. J., et al. (2019). Demonstration of highly efficient dual gRNA CRISPR/Cas9 editing of the homeologous GmFAD2-1A and GmFAD2-1B genes to yield a high oleic, low linoleic and α-linolenic acid phenotype in soybean. *BMC Plant Biol.* 19:311. doi: 10.1186/s12870-019-1906-8

Du, H., Zeng, M., Zhao, X., Cui, Q., Wang, H., Yang, H., et al. (2016). Efficient targeted mutagenesis in soybean by TALENs and CRISPR/Cas9. *J. Biotechnol.* 217, 90–97. doi: 10.1016/j.jbiotec.2015.11.005

Fahlgren, N., Montgomery, M. D., Howell, E., Allen, S. K., Dvorak, A. L., Alexander, L., et al. (2006). Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA affects developmental timing and patterning in *Arabidopsis*. *Curr. Biol.* 16, 939–944. doi: 10.1016/j.cub.2006.03.065

Guschin, D. Y., Waite, G. E., Katibah, J. C., Miller, M. C., Holmes, K., and Rebar, E. J. (2010). A rapid and general assay for monitoring endogenous gene modification. *Methods Mol. Biol.* 649, 247–256. doi: 10.1007/978-1-60761-753-2_15

Hunter, C., Sun, H., and Poethig, R. S. (2003). The *Arabidopsis* heterochronic gene ZIPPY is an ARGONAUTE family member. *Curr. Biol.* 13, 1734–1739. doi: 10.1016/j.cub.2003.09.004

Hunter, C., Willmann, G., Wu, M., Yoshikawa, M., Gutiérrez-Nava, L., and Poethig, S. R. (2006). Trans-acting siRNA-mediated repression of ETTIN and ARF4 regulates heteroblasty in *Arabidopsis*. *Development* 133, 2973–2981. doi: 10.1242/dev.02491

Hwang, W., Fu, D., Reyon, M., Maeder, S., Tsai, J., Sander, R., et al. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 227–229. doi: 10.1038/nbt.2501

Initiative, A. G. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* 408, 796–815. doi: 10.1038/35048692

Jacobs, T., LaFayette, R., Schmitz, P., and Parrott, W. (2015). Targeted genome modifications in soybean with CRISPR/Cas9. *BMC Biotechnol.* 15:16. doi: 10.1186/s12896-015-0131-2

Jiang, W., Zhou, H., Bi, M., Fromm, B., Yang, N., and Weeks, D. P. (2013). Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Res.* 41:e188. doi: 10.1093/nar/gkt780

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

Kanazashi, Y., Hirose, I., Takahashi, M., Mikami, M., Endo, S., Hirose, S., et al. (2018). Simultaneous site-directed mutagenesis of duplicated loci in soybean using a single guide RNA. *Plant Cell Rep.* 37, 553–563. doi: 10.1007/s00299-018-2251-3

Kereszt, A., Li, A., Indrasumunar, C. D., Nguyen, S., Nontachaiyapoom, M., Kinkema, M., et al. (2007). Agrobacterium rhizogenes-mediated transformation of soybean to study root biology. *Nat. Protoc.* 2, 948–952. doi: 10.1038/nprot.2007.141

Le Hir, H., Nott, A., and Moore, R. (2003). How introns influence and enhance eukaryotic gene expression. *Trends Biochem. Sci.* 28, 215–220. doi: 10.1016/S0968-0004(03)00052-5

Li, C., Nguyen, J., Liu, W., Fu, C., Chen, K., Yu, R., et al. (2019). Mutagenesis of seed storage protein genes in soybean using CRISPR/Cas9. *BMC Res. Notes* 12:176. doi: 10.1186/s13104-019-4207-2

Li, J. F., Norville, J., Aach, M., McCormack, D., Zhang, J., Bush, G. M., 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, Z., Liu, A., Xing, B., Moon, J. P., Koellhoffer, L., Huang, R. T., et al. (2015). Cas9-guide RNA directed genome editing in soybean. *Plant Physiol.* 169, 960–970. doi: 10.1104/pp.15.00783

Lin, D., Xiang, Z., Xian, Y., and Li, Z. (2016). Ectopic expression of SlAGO7 alters leaf pattern and inflorescence architecture and increases fruit yield in tomato. *Physiol. Plant.* 157, 490–506. doi: 10.1111/ppl.12425

Mali, P., Yang, K. M., Esvelt, J., Aach, M., Guell, J. E., DiCarlo, J. E., et al. (2013). RNA-guided human genome engineering via Cas9. *Science* 339, 823–826.

Mao, Y., Zhang, Z., Feng, P., Wei, H., Zhang, J., Botella, P., et al. (2016). Development of germ-line-specific CRISPR-Cas9 systems to improve the production of heritable gene modifications in *Arabidopsis*. *Plant Biotechnol. J.* 14, 519–532. doi: 10.1111/pbi.12468

Meister, G. (2013). Argonaute proteins: functional insights and emerging roles. *Nat. Rev. Genet.* 14, 447–459. doi: 10.1038/nrg3462

Michno, J. M., Virdi, A. O., Stec, J., Liu, X., Wang, Y., Xiong, R., et al. (2020). Integration, abundance, and transmission of mutations and transgenes in a series of CRISPR/Cas9 soybean lines. *BMC Biotechnol.* 20:10. doi: 10.1186/s12896-020-00604-3

Michno, J. M., Wang, J., Liu, S. J., Curtin, T. J., Kono, T., and Stupar, R. M. (2015). CRISPR/Cas mutagenesis of soybean and *Medicago truncatula* using a new web-tool and a modified Cas9 enzyme. *GM Crops Food* 6, 243–252. doi: 10.1080/21645698.2015.1106063

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

Nekrasov, V., Staskawicz, D., Weigel, J. D., Jones, B., 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

Paz, M., Shou, Z., Guo, Z., Banerjee, A., and Wang, K. (2004). Assessment of conditions affecting Agrobacterium-mediated soybean transformation using the cotyledonary node explant. *Euphytica* 136, 167–179.

Paz, M. M., Martinez, A. B., Kalvig, T. M., Fonger, M., and Wang, K. (2006). Improved cotyledonary node method using an alternative explant derived from mature seed for efficient Agrobacterium-mediated soybean transformation. *Plant Cell Rep.* 25, 206–213. doi: 10.1007/s00299-005-0048-7

Pike, S., Gassmann, W., and Su, J. (2019). Generating transgenic *Arabidopsis* plants for functional analysis of pathogen effectors and corresponding R Proteins. *Methods Mol. Biol.* 1991, 199–206. doi: 10.1007/978-1-4939-9458-8_18

Saucet, S. B., Ma, P., Sarris, O., Furzer, K. H., Sohn, H., and Jones, J. D. (2015). Two linked pairs of *Arabidopsis* TNL resistance genes independently confer recognition of bacterial effector AvrRps4. *Nat. Commun.* 6:6338. doi: 10.1038/ncomms7338

Schmutz, J., Cannon, J., Schlueter, J., Ma, T., Mitros, W., Nelson, D. L., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463, 178–183.

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

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

Sun, X., Hu, R., Chen, Q., Jiang, G., Song, H., Zhang, R., et al. (2015). Targeted mutagenesis in soybean using the CRISPR-Cas9 system. *Sci. Rep.* 5:10342. doi: 10.1038/srep10342

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, L., Sun, T., Wu, L., Liu, X., Sun, Y., Cai, J., et al. (2020). Natural variation and CRISPR/Cas9-mediated mutation in GmPRR37 affect photoperiodic flowering and contribute to regional adaptation of soybean. *Plant Biotechnol. J.* doi: 10.1111/pbi.13346 [Epub ahead of print].

Wang, Z., Xing, L., Dong, H., Zhang, C., Han, X., and Chen, Q. (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, N., Lu, P., Wang, Q., Zhang, J., Zhang, J., and Wang, N. (2020). Construction and Analysis of GmFAD1-1A and GmFAD2-2A soybean fatty acid desaturase mutants based on CRISPR/Cas9 technology. *Int. J. Mol. Sci.* 21:1104. doi: 10.3390/ijms21031104

Yamada, T., Takagi, K., and Ishimoto, M. (2012). Recent advances in soybean transformation and their application to molecular breeding and genomic analysis. *Breed. Sci.* 61, 480–494. doi: 10.1270/jsbbs.61.480

Yan, L., Wei, Y., Wu, R., Hu, H., Li, W., Yang, R., et al. (2015). High-efficiency genome editing in *Arabidopsis* using YAO promoter-driven CRISPR/Cas9 system. *Mol. Plant* 8, 1820–1823. doi: 10.1016/j.molp.2015.10.004

Zhang, P., Du, J., Wang, Y., Pu, C., Yang, R., Yan, H., et al. (2019). Multiplex CRISPR/Cas9-mediated metabolic engineering increases soya bean isoflavone content and resistance to soya bean mosaic virus. *Plant Biotechnol. J.* doi: 10.1111/pbi.13302 Epub ahead of print].

Zhang, Y., Massel, I. D., Godwin, K., and Gao, C. (2018). Applications and potential of genome editing in crop improvement. *Genome Biol.* 19:210. doi: 10.1186/s13059-019-1622-6

Zhou, C., Han, C., Fu, J., Wen, X., Cheng, J., Nakashima, J., et al. (2013). The trans-acting short interfering RNA3 pathway and no apical meristem antagonistically regulate leaf margin development and lateral organ separation, as revealed by analysis of an argonaute7/lobed leaflet1 mutant in *Medicago truncatula*. *Plant Cell* 25, 4845–4862. doi: 10.1105/tpc.113.117788

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 08 July 2020
doi: 10.3389/fpls.2020.01005

Similar Seed Composition Phenotypes Are Observed From CRISPR-Generated In-Frame and Knockout Alleles of a Soybean *KASI* Ortholog

*Kamaldeep S. Virdi, Madison Spencer, Adrian O. Stec, Yer Xiong, Ryan Merry, Gary J. Muehlbauer and Robert M. Stupar**

Department of Agronomy & Plant Genetics, University of Minnesota, Saint Paul, MN, United States

OPEN ACCESS

Edited by:
Linda Ann Rymarquis, Bayer Crop Science, United States

Reviewed by:
Yanfei Mao, Shanghai Institutes for Biological Sciences (CAS), China
Karen A. Hudson, Agricultural Research Service, United States

***Correspondence:**
Robert M. Stupar
stup0004@umn.edu

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *09 April 2020*
Accepted: *19 June 2020*
Published: *08 July 2020*

Citation:
Virdi KS, Spencer M, Stec AO, Xiong Y, Merry R, Muehlbauer GJ and Stupar RM (2020) Similar Seed Composition Phenotypes Are Observed From CRISPR-Generated In-Frame and Knockout Alleles of a Soybean KASI Ortholog. Front. Plant Sci. 11:1005. doi: 10.3389/fpls.2020.01005

The β-ketoacyl-[acyl carrier protein] synthase 1 (*KASI*) gene has been shown in model plant systems to be critical for the conversion of sucrose to oil. A previous study characterized the morphological and seed composition phenotypes associated with a reciprocal chromosomal translocation that disrupted one of the *KASI* genes in soybean. The principle findings of this work included a wrinkled seed phenotype, an increase in seed sucrose, a decrease in seed oil, and a low frequency of transmission of the translocation. However, it remained unclear which, if any, of these phenotypes were directly caused by the loss of *KASI* gene function, as opposed to the chromosomal translocation or other associated factors. In this study, CRISPR/Cas9 mutagenesis was used to generate multiple knockout alleles for this gene, and also one in-frame allele. These soybean plants were evaluated for morphology, seed composition traits, and genetic transmission. Our results indicate that the CRISPR/Cas9 mutants exhibited the same phenotypes as the chromosomal translocation mutant, validating that the observed phenotypes are caused by the loss of gene function. Furthermore, the plants harboring homozygous in-frame mutations exhibited similar phenotypes compared to the plants harboring homozygous knockout mutations. This result indicates that the amino acids lost in the in-frame mutant are essential for proper gene function. In-frame edits for this gene may need to target less essential and/or evolutionarily conserved domains in order to generate novel seed composition phenotypes.

Keywords: soybean, *KASI*, sucrose, oil, CRISPR, Cas9, seed, mutant

INTRODUCTION

Seed composition traits are critical for soybean end uses. The protein fraction is important for food uses and livestock feed and the oil fraction is useful for food, fuel, and industrial applications. Furthermore, the carbohydrate fraction can impact the end use of the bean, particularly in the development of varieties for human consumption. Therefore, a better understanding of the genes

that govern the seed composition components will be useful to increase the breeding efficiency of desirable traits for commodity and specialty markets.

Recent work from our group identified a fast neutron-induced chromosomal translocation that co-segregated with an increased seed sucrose and reduced oil phenotype (Dobbels et al., 2017). This locus was defined by a reciprocal translocation between chromosomes 8 and 13, which disrupted an internal exon of a β-ketoacyl-[acyl carrier protein] synthase 1 (*GmKASI*) ortholog (soybean gene model Glyma.08G084300). The seeds homozygous for the translocation also exhibited a wrinkled phenotype, consistent with previous observations for mutants of this gene in *Arabidopsis thaliana* (Wu and Xue, 2010). Furthermore, the translocated *KASI* allele was observed to transmit and segregate at a frequency far below Mendelian expectations. This work left three major questions unresolved: (1) Could the seed composition function of the soybean *KasI* gene be validated using CRISPR mutagenesis, and would the phenotype be different from the fast neutron line? (2) Given the extreme nature of the seed composition phenotype in the fast neutron line, would it be possible to generate an intermediate (i.e., less severe) phenotype by generating an in-frame mutant allele for the *KASI* ortholog? (3) Was the reduced transmission of the *KASI* mutant a consequence of the knocked-out *kasI* allele, or did it result from abberant meiosis of the translocated chromosomes?

This study addresses these three questions. We used CRISPR/Cas9 gene editing methods to generate both an in-frame and knockout alleles of the soybean *KASI* ortholog. We observed both the segregation patterns of these alleles, and monitored the seed composition phenotypes of the segregating families. All alleles generated in this study did not have any chromosomal abnormalities at the *KASI* site, but rather exhibited relatively small nucleotide deletions and insertions, as are typical of CRISPR/Cas9 edited sites. We found that these mutant lines also exhibited reduced transmission of the *kasI* alleles, for both in-frame and knockout alleles. Furthermore, we observed similar seed phenotype profiles for the in-frame and the knockout alleles. One of our goals of this study was to generate in-frame mutants with intermediate phenotypes and higher (i.e., normal) transmission levels. However, it appears that the location and amino acids deleted from the in-frame mutant were critical for normal *KASI* function, and thus mimicked the knockout allele phenotypes.

MATERIALS AND METHODS

CRISPR/Cas9 Design and Assembly and Soybean Whole Plant Transformation

Whole protein sequences of soybean KASI and the nearest Arabidopsis ortholog were compared to identify evolutionarily conserved domains that could be targeted for mutagenesis. The protein sequences were obtained from the Phytozome and TAIR websites, respectively. Both proteins were aligned with T-Coffee (https://www.ebi.ac.uk/Tools/msa/tcoffee/) and the aligned fasta file was visualized with BoxShade (https://embnet.vital-it.ch/software/BOX_form.html) software.

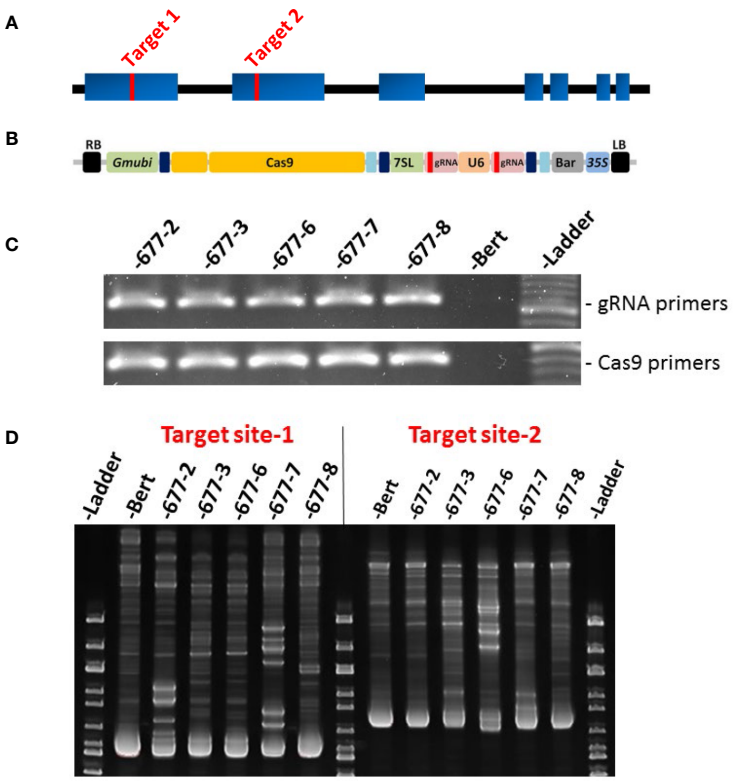
Target sites for gRNAs were identified using the CRISPR-P 2.0 website (http://crispr.hzau.edu.cn/CRISPR2/; Liu et al., 2017). Oligos were synthesized from Integrated DNA Technologies (https://www.idtdna.com/pages). All CRISPR/Cas9 reagents were then assembled as described in Curtin et al. (2018). The final construct was transformed into the disarmed *Agrobacterium rhizogenes* strain 18r12 (Veena and Taylor, 2007). Whole plant soybean transformation was performed in the genetic background of the cultivar "Bert" (Orf and Kennedy, 1992), sub-line Bert-MN-01, using previously published methods (Paz et al., 2006; Liu et al., 2019). CRISPR/Cas9 reagents for whole plant transformation were assembled as described in Curtin et al. (2018). This included the Cas9-encoding sequence, two distinct gRNAs, and sequence encoding the glufosinate selectable marker. These components were driven by *Gmubi, U6, 7sL,* and *35S* promoters, respectively (**Figures 1A, B**).

Polymerase Chain Reaction Targeted Amplicon for Heteroduplex Assays, CAPS Assays, and Sanger Sequencing for Detecting CRISPR/Cas9 Edits

The genomic regions spanning gRNA target sites were amplified by polymerase chain reaction (PCR) using the HotStarTaq Plus master mix (Qiagen, Hilden, Germany) according to the manufacturer's instructions. New mutations at the targeted sites were identified using either heteroduplex of Cleaved Amplified Polymorphic Sequences (CAPS) analyses. Heteroduplex assays were performed as previously reported (Zhu et al., 2014). For CAPS assays, targeted PCR amplicons were digested with BsaWI (New England Biolabs, Ipswich, MA) overnight at 60°C as per the manufacture's guideline. Digested products were run on agarose electrophoresis gels (1.3%). The presence of digestion-resistant PCR amplicons indicated CRISPR/Cas9 induced edits/mutations had occurred. Two approaches were used for Sanger sequencing. First, targeted PCR amplicons were directly sequenced and.abi files containing sequence information were then analyzed by Inference of CRISPR Edits (ICE) software (https://ice.synthego.com; Hsiau et al., 2018) to identify mutations. Alternatively, PCR amplicons were sub-cloned into the Topo TA cloning vector (Thermo Fisher Scientific, Waltham, MA) as per the manufacturer's instructions and individual positive clones were sequenced. Each sequence file (.abi) was visualized in ABI sequence scanner software. DNA sequence files were aligned with MultAlin software (http://multalin.toulouse.inra.fr/multalin/; Corpet, 1988).

Whole Genome Sequencing and Bioinformatics

Selected plants were resequenced to confirm new mutations and the presence/absence of the transgene sequences. As both T-DNA and plasmid backbone sequences were integrated into the genome of the main T_0 plant of interest (see Results section),



E

Target site 1		Target site 2	
677-2		**677-2**	
50% WT	TTGTGATCACCGGAATGGGCC	89% WT	TTGATGATAGCCGGTGGGAC
17% Δ3-bp	TTGTGATCAC---AATGGGCC	11% Δ1-bp	TTGATGATAGC-GGTGGGAC
13% Δ1-bp	TTGTGATCACCG-AATGGGCC	**677-3**	
10% Δ3-bp	TTGTGATCA---GAATGGGCC	57% WT	TTGATGATAGCCGGTGGGAC
10% Δ3-bp	TTGTGATCACCG---TGGGCC	27% +1-bp	TTGATGATAGCTCGGTGGGAC
677-3		16% Δ1-bp	TTGATGATAGC-GGTGGGAC
65% WT	TTGTGATCACCGGAATGGGCC	**677-6**	
35% Δ1-bp	TTGTGATCACCG-AATGGGCC	38% +1-bp	TTGATGATAGCTCGGTGGGAC
677-6		32% Δ5-bp	TTGATG-----CGGTGGGAC
55% Δ1-bp	TTGTGATCACCG-AATGGGCC	18% WT	TTGATGATAGCCGGTGGGAC
45% WT	TTGTGATCACCGGAATGGGCC	10% Δ5-bp	TTGATGATAG-----GGGAC
677-7		**677-7**	
63% WT	TTGTGATCACCGGAATGGGCC	65% WT	TTGATGATAGCCGGTGGGAC
20% Δ5-bp	TTGTGATCACCG-----GGCC	18% Δ1-bp	TTGATGATAGC-GGTGGGAC
17% Δ2-bp	TTGTGATCACCG--ATGGGCC	17% +1-bp	TTGATGATAGCTCGGTGGGAC
677-8		**677-8**	
90% WT	TTGTGATCACCGGAATGGGCC	100% WT	TTGATGATAGCCGGTGGGAC
10% Δ1-bp	TTGTGATCACCG-AATGGGCC		

FIGURE 1 | CRISPR/Cas9 targeting of soybean GmKASI and identification of mutations in T0M0 plants **(A)** Predicted gene model of *GmKASI* (Glyma.08g084300). Exons are depicted as blue rectangles. Red bars in exon 1 and exon 2 indicate the target sites for gene editing. **(B)** The T-DNA of the assembled CRISPR/Cas9 construct carrying two gRNAs, each targeting the respective sites shown in **(A)**. Cas9, gRNA for target-1, gRNA for target-2, and BASTA (glufosinate) selectable marker were driven by *Glycine* max ubiquitin (*Gmubi*), Arabidopsis ubiquitin (*U6*), *7sL*, and *35S* promoters, respectively (image adopted from Curtin et al 2018). **(C)** T-DNA presence in five independent T0M0 events (677-2, 677-3, 677-6, 677-7, 677-8) detected by PCR. Two sets of primer pairs were used, specific to the gRNA and Cas9 regions of the T-DNA. **(D)** Heteroduplex assays showed novel amplicon bands in the transformed plants compared to the 'Bert' wild-type control, indicating the presence of *GmKASI* edited alleles. **(E)** Sanger sequence analysis for all T_0M_0 events revealed various *GmKASI* edited alleles at both target sites. The genomic region spanning each target site was PCR amplified, Sanger sequenced, and analyzed with ICE software. Each T_0M_0 event showed a different proportion (indicated by %) of edited alleles (Δ indicates deletions; + indicates insertions). The underlined bases are the Protospacer Adjacent Motif (PAM) site at each target site.

henceforth these sequences will be distinguish using the terms "T-DNA" and "backbone," rather than the term "transgene."

DNA from young leaves at the second trifoliate stage was extracted with the DNeasy plant kit (Qiagen). Sequencing was performed at University of Minnesota Genomics Center to approximately 20x coverage per genotype. The T-DNA and plasmid backbone insertion sites were identified using the method of Michno et al. (2020) using the bash script TransGeneMap (https://github.com/MeeshCompBio/Soybean_Scripts) with the following modifications. Initial quality was assessed using Fastqc version 0.11.7 (Andrews, 2010). Trimmomatic version 0.33 (Bolger et al., 2014) was used for adapter removal, keeping a minimum read length of 40 bp and quality cutoff set to a phred score of 20. Read mapping to the soybean genome was conducted using bwa version 0.7.17 (Li and Durbin, 2010) following the same parameters as Michno et al. (2020). To locate the sites of integration in the genome, a FASTA file was generated using the entire sequence of the plasmid (including both the T-DNA and backbone) for read mapping, which allowed us to identify integration of the backbone as well as the T-DNA sequences. Orphan reads were mapped back to the soybean reference genome (Wm82.a2.v1) using bowtie2 version 2.3.4.1.

To discern the zygosity state of the vector backbone insertion, reads from sequenced plants were aligned to the soybean reference genome using the bash script Fastq2ReadmapGmaxV2 (https://github.com/MeeshCompBio/Soybean_Scripts) with the following modifications. Initial quality was again assessed using Fastqc version 0.11.7, adapter removal was done using cutadapt version 1.18 (Martin, 2011), and bwa version 0.7.17 was used for alignment. IGV version 2.3.97 (Robinson et al., 2011) was used to visually screen for the position and state of the vector backbone. Plants with reads aligning across the insertion site with reduced read depth at the site compared to the surrounding region were determined to be heterozygous. Plants with no reads aligning across the insertion site or mate pairs spanning the site (indicating a large insertion) were determined to be homozygous for the vector backbone insertion.

Plant Materials, Growth, and Morphological Analysis

All plant materials were grown under similar growth conditions in a single greenhouse. The greenhouse temperature was maintained between 21–23°C. Natural light was supplemented with 600 watt high pressure sodium (HPS) lamps using a photoperiod day length of 14 h. All CRISPR/Cas9 edited and control plants were planted at the same time in a propagation mix growth medium (Sungro brand). Leaf samples for DNA extraction were harvested at the second trifoliate stage. Plants from different CRISPR/Cas9 families and control genotypes were randomly arranged. Plants were fertilized every 2 weeks using 400 ppm Jack's water-soluble 20-3-19 fertilizer until they reached the R7 stage. Plants were watered after every 2 d until the R7 stage and then reduced to once per week until the R8 stage. Individual plants were manually harvested and threshed to maintain genetic purity. Threshed seeds were kept in packets in the greenhouse for 1 week to promote drying. Transmissible mutations were identified from the progeny derived from three different T_0 plants (see *Results*). A series of segregating mutant alleles was identified in the progeny of the T_0 plant WPT677-3. Five of these families, each segregating for a distinct combination of alleles for the Glyma.08G084300 gene, were phenotyped in downstream seed composition analysis (see *Near Infrared Scan for seed Composition*), along with a control family that was homozygous for the wild-type allele (**Table 1**). Furthermore, seedling growth rates were measured for two homozygous mutant lines compared to homozygous wild-type siblings; ten seedlings were grown and measured from each genotype.

Near Infrared Scan for Seed Composition

Two growouts were performed for seed composition analysis of the segregating families derived from WPT677-3 (**Table 1**). Within each of these two experiments, three to 17 biological replicates (the median number of plants among the mutant families was eight) were measured for each mutant family (detailed information on the number of plants in each family per experiment are provide in **Supplementary Tables S1** and **S2**). As these families were segregating, each plant was genotyped and grouped into the appropriate mutant class (homozygous mutant, heterozygous mutant, or homozygous wild-type) for seed composition analysis. Approximately 20 g of whole soybean seeds from individual plants were ground to a fine powder using a water cooled Foss KN195 Knifetec rotary grinder. The seeds were exposed to three consecutive pulses of grinding for 10 s each (30 s total) while rocking the grinder to ensure that all the seed material was ground. After completion, the material was removed from the grinder and immediately placed in a bag and immediately vacuum sealed until near infrared (NIR) scanning was performed. NIR scans and calculations of predicted values for each seed composition trait were computed as previously described (Dobbels et al., 2017).

Statistical Analysis

All statistical analyses were conducted in R 3.5.2 (R Core Team, 2018, https://www.r-project.org/) and figures were produced with ggplot2_3.2.1 package (Wickham, 2016).

Data Availability

Sequence data have been deposited into the NCBI Short Read Archive under the Bioproject identifier PRJNA640373.

RESULTS

CRISPR/Cas9 Induced Targeted Mutagenesis of the Soybean *KASI* (*GmKASI*)

The predicted soybean gene model for Glyma.08G084300, henceforth referred to as *GmKASI*, indicates seven exons (https://soybase.org/; **Figure 1A**). To determine the guide RNA (gRNA) target sites for CRISPR/Cas9 gene editing, we first

TABLE 1 | Status of targeted mutations, T-DNA presence/absence, and plasmid backbone presence/absence in a sub-set of CRISPR/Cas9 confirmed to carry mutations in *GmKASI*.

Plant	Gen.	Glyma.08G084300	T-DNA			Plasmid backbone		
		Edits (site1.site2)	Status	Chr	Integration	Status	Chr	Integration
WPT677-3	T_0M_0	Various	Heteroyg.	5	5839464. 5839707	Heterozyg.	8	17781297. 17781301
WPT677-3-35	T_1M_1	+10/+107.WT/WT[a]	Absent	N/A	N/A	Heterozyg.	8	17781297. 17781301
WPT677-3-43	T_1M_1	WT/Δ1.WT/Δ1[b]	Absent	N/A	N/A	Heterozyg.	8	17781297. 17781301
WPT677-3-44	T_1M_1	Δ1/+1.WT/+1[a]	Absent	N/A	N/A	Heterozyg.	8	17781297. 17781301
WPT677-3-47	T_1M_1	WT/WT.WT/WT[c]	Absent	N/A	N/A	Heterozyg.	8	17781297. 17781301
WPT677-3-48	T_1M_1	WT/Δ6.WT/+1[b]	Absent	N/A	N/A	Homozyg.	8	17781297. 17781301
WPT677-3-22-04	T_2M_2	WT/Δ6.WT/+1[b]	Absent	N/A	N/A	Heterozyg.	8	17781297. 17781301
WPT677-3-22-10[e]	T_2M_2	WT/Δ6.WT/WT[d]	Absent	N/A	N/A	Untested		

[a]*Biallelic knockout genotype.*
[b]*Heterozygous knockout genotype.*
[c]*Homozygous wild-type genotype.*
[d]*Heterozygous in-frame genotype.*
[e]*Genotype WPT677-3-22-10 was not subjected to WGS. The edited alleles were determined by Sanger sequencing and the absence of the T-DNA sequence was based on PCR analysis (see Results section).*
Δ = deletion.
+ = insertion.

compared the protein sequences of *GmKASI* and its Arabidopsis ortholog. The genes are highly conserved, with 86.5% identity at the amino acid level (**Supplementary Figure S1**). Based on the conserved regions of the protein, two gRNA target sites were selected, one in exon 1 (target site-1) and second in exon 2 (target site-2; **Figure 1A**).

The resulting construct (**Figure 1B**) was transformed into soybean, resulting in five T_0 plants: WPT677-2, WPT677-3, WPT677-6, WPT677-7, and WPT677-8. PCR assays with primers specific to gRNA and Cas9 regions of the T-DNA detected the presence of transgene sequences in all T_0 plants (**Figure 1C**). A heteroduplex assay was conducted for both target sites to test for the presence of mutations. Heteroduplex assays involve melting and subsequent renaturation of the target site PCR amplicons prior to electrophoresis. If the homologous gene copies have different sequences due to new mutations, then the renatured DNA will have some imperfect double-stranded complexes, resulting in slower migration and thus novel bands observed on the gel. The heteroduplex assays for the T_0 plants all showed novel amplicon bands, indicating new mutations occurred in the T_0 plants (**Figure 1D**). Therefore, this generation was renamed as T_0M_0. Sanger sequencing of the PCR products from each target region confirmed various edited alleles of *GmKASI* in these plants (**Figure 1E**). To further confirm the mutations, sub-cloned target site-1 amplicons from plants WPT677-2, WPT677-3, and WPT677-6 were sequenced (**Supplementary Figure S2**). These assays confirmed editing at the site, though at a relatively low frequency. CAPS assays at target site-1 revealed digestion-resistant bands in all T_0M_0 plants, further confirming mutated alleles of *GmKASI* (**Supplementary Figure S3**).

Generation of Stable and Heritable Knockout and In-Frame Mutant Alleles of *GmKASI* in Soybean

CAPS assays and Sanger sequencing were used to screen the inheritance of mutations in subsequent generations. CAPS assays on T_1M_1 families of WPT677-2 and WPT677-6 plants showed that the targeted mutations were successfully transmitted, as both families segregated for wild type and mutant alleles (**Supplementary Figure S4**). All T_1M_1 plants from both families also inherited T-DNA sequences, as detected by PCR. This finding indicates that multiple unlinked copies of T-DNA were likely integrated in the genome during transformation. However, the WPT677-7 and WPT677-8 families did not inherit the T-DNA nor mutations in the T_1M_1 generation (**Supplementary Figure S5**), suggesting the T-DNA did not stably integrate into the genome and all mutations observed in the T_0M_0 generation occurred in somatic cells.

Both mutations and T-DNA sequences were inherited in the T_1M_1 generation of plant WPT677-3 (**Supplementary Figure S6**; **Supplementary Table S3**). In this family, the mutations and T-DNA sequences appeared to segregate independently in this generation. A nearly 3:1 segregation ratio was observed for the T-DNA sequences (32 plants carrying T-DNA: 13 plants not carrying T-DNA). This finding suggested the T-DNA insertion in the T_0M_0 plant was likely a single copy event. The WPT677-3 lineage was subjected to further analyses in the T_1M_1 and T_2M_2 generations to identify individuals with heritable and stable targeted mutations in the absence of the CRISPR/Cas9 T-DNA. We selected two homozygous mutant plants (WPT677-3-35, WPT677-3-44), three heterozygous plants (WPT677-3-43, WPT677-3-48, WPT677-3-22-10), and one homozygous wild

type plant (WPT677-3-47). A PCR based assay did not detect the presence of the T-DNA in any of these plants.

Detailed Sanger sequence analysis of plants WPT677-3-35, WPT677-3-44, WPT677-3-43, WPT677-3-48, and WPT677-3-22-10 revealed various edited alleles of *GmKASI* at both target sites (**Figure 2**). We performed Whole Genome Sequencing (WGS) of the T_0M_0 plant (WPT677-3), five T_1M_1 plants (WPT677-3-35, WPT677-3-43, WPT677-3-44, WPT677-3-47, WPT677-3-48), and one T_2M_2 plant (WPT677-3-22-04). At the time of the WGS experiment, the WPT677-3-22-10 plant was not available so we instead sequenced WPT 677-3-22-04, a sibling that did not carry the T-DNA sequences. No mutations were detected in WPT677-3-47, which served as the unedited control in these experiments. Notably, novel mutations at both target sites in the T_1M_1 generation were detected which were not present in the parental T_0M_0 plant (WPT677-3). One possible explanation is that heritable mutations continued to occur after the WPT677-3 leaves were sampled for the initial analysis. We detected two plants carrying biallelic mutations (WPT-677-3-35 and WPT677-3-44), two plants carrying heterozygous knock out alleles (677-3-43 and 677-3-48), and one plant carrying a heterozygous in-frame allele (WPT677-3-22-10) (**Table 1**). We detected edits at both target sites in WPT677-3-22-04, resulting in another knock out allele (**Table 1**); we did not follow this line for phenotypic analysis. All of the edited alleles were stably inherited to the T_2M_2 generation of WPT677-3-35, WPT677-3-44, WPT677-3-43, WPT677-3-44, WPT677-3-48, and the T_3M_3 generation of WPT677-3-22-10 (**Supplementary Tables S1** and **S2**).

Whole genome analysis confirmed that only the T_0M_0 plant contained the T-DNA sequences, which were integrated on chromosome 5 (**Table 1**). All of the selected progeny plants were confirmed to lack the T-DNA, indicating that this locus segregated away in these plants and the targeted sites are thus stable (**Supplementary Figure S6**). However, the WGS analyses detected the presence of the vector backbone (**Supplementary Figure S7**) sequences at a specific locus on chromosome 8 in all plants of these plants (**Table 1**). Therefore, the generational nomenclature for these plants remained in the T_xM_x form. Resequencing analysis indicated that the entire vector backbone, roughly 10 kb of DNA, was included in this insertion. When examined in greater detail, it was determined that the vector backbone insertion was present in the heterozygous state in the T_0M_0 plant WPT677-3 (**Table 1**). It appears that the backbone insertion segregated in the subsequent generation, exhibiting both heterozygous (4 plants) and homozygous (1 plant) progeny among the five sequenced T_1M_1 plants (**Table 1**). As the target gene (Glyma.08G084300) and the backbone insertion are located 11.4 Mb from one another on chromosome 8, it is possible that some of the edits are genetically linked to this backbone insertion. However, this is unlikely to be true for all edits, as the respective edits and the backbone insertion may be located on homologous chromosomes in some cases. The one plant identified as homozygous for the backbone insertion (WPT677-3-48) exhibited a heterozygous edited allele (**Table 1**), indicating that this particular mutation was not perfectly linked to the backbone. Presumably, some WPT677-3 descendants segregated out the backbone insertion while maintaining edited alleles of Glyma.08G084300, however no such lines were selected for WGS in this study.

GmKASI has a paralog copy (Glyma.05g129600) in the soybean genome. RNA expression of both *GmKASI* and its paralog copy have similar transcript expression profiles across all tissues (**Supplementary Figure S8A**; data extracted from Severin et al., 2010). This suggests that the paralog copy may have functional redundancy with *GmKASI*. We sought to investigate if our gRNAs exhibited any off-target mutagenesis to this paralog copy. The two gRNAs for target site-1 and target site-2 in *GmKASI* have three and two mismatches, respectively,

Alleles	Plant ID	Target site-1	Target site-2
WT..WT	Bert	TTGTGATCACCGGAATGGGCC	TTGATGATAGCCGGTGGGAC
WT..WT	677-3-47	TTGTGATCACCGGAATGGGCC	TTGATGATAGCCGGTGGGAC
Δ1..WT	677-3-44	TTGTGATCACCG-AATGGGCC	TTGATGATAGCCGGTGGGAC
+1..+1	677-3-44	TTGTGATCACCGAGAATGGGCC	TTGATGATAGCTCGGTGGGAC
+10..WT	677-3-35	TTGTGATCACCGC..AGAATGGGCC	TTGATGATAGCCGGTGGGAC
+107..WT	677-3-35	TTGTGATCACCGGC..TAATGGGCC	TTGATGATAGCCGGTGGGAC
Δ6..WT	677-3-22	TTGTGA------GAATGGGCC	TTGATGATAGCCGGTGGGAC
Δ1..Δ1	677-3-43	TTGTGATCACCG-AATGGGCC	TTGATGATAGC-GGTGGGAC
Δ6..+1	677-3-48	TTGTGATCA------TGGGCC	TTGATGATAGCACGGTGGGAC

FIGURE 2 | Isolation of multiple stable alleles of *GmKASI* inherited in T_1M_1 plants Eight independent *GmKASI* alleles with a combination of mutations at target site 1 and target site 2 were inherited in the T_1M_1 progeny of the 677-3 T_0M_0 plant. Each target site was amplified by Polymerase Chain Reaction and Sanger sequenced. Plant 677-3-47 carried homozygous wild-type allele. Plants 677-3-35 and 677-3-44 were biallelic, carrying two heterozygous mutant alleles, and are thus shown twice (Δ indicates deletions; + indicates insertions; red letters indicate inserted bases). The ".." in the plant 677-3-35 alleles indicates additional inserted bases not shown. 677-2-22, 677-3-43, and 677-3-48 each carried one mutant allele. The *GmKASI* alleles are depicted as *mutation attarget site 1 .. mutation* at target site 2. The underlined bases are the Protospacer Adjacent Motif (PAM) site at each target site.

when compared to the paralog copy (**Supplementary Figure S8B**). We sequenced PCR products of the paralog copy for both gRNA sites in plants WPT677-3-22, WPT677-35, WPT677-3-43, WPT677-3-44, WPT677-3-47, and WPT677-3-48. No evidence of new mutations were detected at either site in any of the assessed plants (**Supplementary Figure S8C**).

Both Knockout and In-Frame Edited Alleles of GmKASI Altered Seed Morphology and Seed Composition Traits in Soybean

We evaluated lines with edited alleles of *GmKASI*, wild-type segregants, and nontransformed cv. Bert for morphological and seed composition phenotypes based on two growouts in the greenhouse. Plants carrying homozygous mutant alleles displayed a range of slow growth compared to homozygous wild type siblings at the seedling stage (**Supplementary Figure S9**) and maintained slower growth rates at later stages (**Figure 3A**). Homozygous knock out and in-frame mutant plants both showed wrinkled and shriveled seed phenotypes (**Figure 3B**; **Supplementary Figure S10**) consistent with the fast neutron mutants described by Dobbels et al. (2017). Similar seed phenotypes were also confirmed in T_2M_2 seeds harvested from descendants of a separate CRISPR lineage derived from the T_0M_0 plant WPT677-6 (data not shown). The progeny from two WPT677-3 mutant lineages (WPT677-3-43-69 and WPT677-3-48-06) appeared to have many seeds with relatively normal shapes compared to the other mutants. However, both WPT677-3-43-69 and WPT677-3-48-06 carried heterozygous wild-type alleles at site 1 and site 2, therefore many of their seed progeny would be expected to carry a combination of wild type alleles at these two sites, rendering such seeds developmentally normal.

We analyzed seed composition traits from lineages of the WPT677-3 edited families of *GmKASI*, including wild-type segregants, and the nontransformed cv. Bert as controls, in two independent NIR experiments. To confirm the genotypes of individual plants, we genotyped each individual in each replicate. Detailed information about the genotype of individual plants from each CRISPR/Cas9 family are provided in **Supplementary Tables S1** and **S2**. In the first experiment, we evaluated six CRISPR/Cas9 families carrying one wild type, one in-frame, and six knock out alleles of *GmKASI*. WPT677-3-35 and WPT677-3-44 are biallelic mutants that produced families with all mutant plants. WPT677-3-43 (knock out), WPT677-3-48 (knock out) and WPT677-3-22-10 (in-frame) were heterozygous plants that produced families consisting of segregating wild type, heterozygous, and homozygous mutant individuals. All the plants carrying homozygous mutant alleles displayed an increase in seed sucrose content and a decrease in total seed oil content (**Figure 4A**). The individuals with homozygous in-frame alleles showed nearly identical seed phenotypes as the homozygous knockout individuals. On average, the homozygous mutant plants from all the families exhibited significantly higher sucrose content (10.36% on dry matter basis *vs.* a wild type value of 7.03%) and lower oil

FIGURE 3 | Both knockout and in-frame edited alleles of *GmKASI* displayed seed phenotypes**(A)** Representative plants of segregating T_2M_2 progeny from the heterozygous parental 677-3-43 plant. Plants carrying the *GmKASI* edited allele in the homozygous state displayed slow growth and semi-dwarf phenotypes. The genotype of each T_2M_2 plant is depicted as "*edit at target site 1 .. edit at target site 2*." **(B)** Sample of seeds from homozygous wild-type and homozygous mutant plants segregating from three different heterozygous plants; 677-3-43, 677-3-48, and 677-3-22-10. Wild-type segregants are shown on the left, while homozygous mutant segregants are shown on the right. Plants 677-3-43 and 677-3-48 carried knockout edited alleles, while plant 677-3-22-10 carried an in-frame edited allele. Homozygous mutant seeds from homozygous knockout and in-frame alleles showed similar wrinkled and cracked seed morphologies.

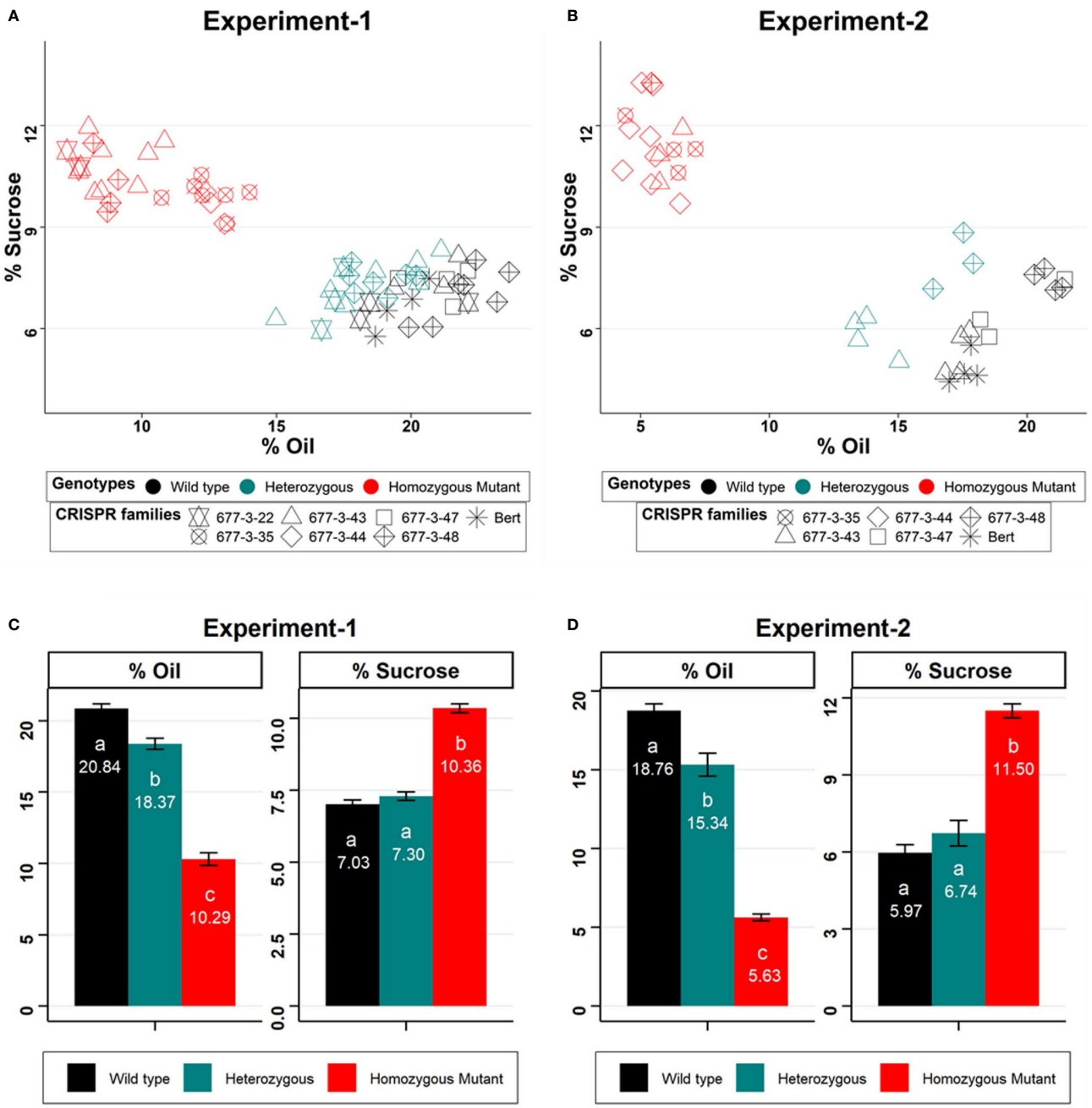


FIGURE 4 | *GmKASI* regulates sucrose and oil content in soybean seeds **(A, B)** Scatter biplot of percent total oil and sucrose content for multiple CRISPR families (denoted by different symbols) based on NIR prediction. Data are shown from two independent greenhouse experiments. The genotypes are shown by different colors and refer to the parent plant that produced the seed. The parents were either homozygous for a mutation, heterozygous, or homozygous wild type. Three to 17 biological replicates were measured for each family. **(C, D)** Barplot of mean oil and sucrose content for each of the seed progeny from wild-type, heterozygous, and mutant genotypes in the two experiments. Homozygous mutants from all families in both experiments showed an increase in sucrose and a decrease in total oil content. Oil and sucrose measurements are based on the percent of seed, on a dry matter basis. One-way ANOVA and post-hoc Tukey-HSD (*alpha* = 0.01) tests were conducted. Genotypes with significantly different means are indicated by different letters (a, b, c). Error bars represent the mean ± standard error for each genotype class.

content (10.29% on dry matter basis *vs.* a wild type value of 20.84%) compared to homozygous wild type individuals in the first replicate (**Figure 4C**). The second experiment of this study showed similar results, although the extent of alteration in sucrose and oil contents were greater than in the first experiment (**Figures 4B, D**). Homozygous mutant plants showed an increase of sucrose content from 5.97% to 11.50% and a reduction of oil content from 18.76% to 5.63% as compared

to wild type. Among other seed composition traits, homozygous mutant plants also showed a significant increase in linolenic fatty acid (*experiment 1*: from 8.55% to 15.48% of total fat; *experiment 2:* from 7.51% to 16.99% of total fat) and a decrease in linoleic fatty acid (*experiment 1*: from 49.42% to 30.57% of total fat; *experiment 2:* from 52.56% to 44.56% of total fat) compared to wild type (**Supplementary Figures S11** and **S12**).

Non-Mendelian Inheritance of *GmKASI* Edited Alleles in Soybean

A low recovery of homozygous mutant segregants was observed in all three independent heterozygous lineages (progeny from WPT677-3-43, WPT677-3-48, and WPT677-3-22-10). In all cases, segregating progeny showed non-Mendelian inheritance with lower than expected transmission of the mutant alleles of *GmKASI* (**Table 2**). The segregation ratio significantly deviated from the expected 1:2:1 ratio in all three families (chi^2 p value <0.001), with a particularly low recovery of homozygous mutant individuals.

DISCUSSION

This study set out to resolve three standing questions that resulted from previous work on *GmKasI* (Dobbels et al., 2017). First, would CRISPR mutagenesis validate the seed composition function of the *GmKASI* gene (Glyma.08G084300), and would the phenotype be different from the fast neutron line? This question was clearly addressed in the current study, as the seed composition phenotypes observed among a range of different CRISPR mutant lines were very similar to those observed in the previously published fast neutron translocation mutant of GmKasI (Dobbels et al., 2017). Compared to wild type seeds, the homozygous mutant seeds exhibited a wrinkled surface phenotype, an increase in percentage sucrose, a decrease in oil, and a redistribution of some fatty acid levels. Therefore, the CRISPR lines validated the function of *GmKasI* in determining the seed morphology and seed composition phenotypes. It is important to note that the phenotypic analysis of the WPT677-3 family may have been complicated by the presence of a vector backbone insertion in many of these plants. However, there is ample evidence indicating that the vector backbone insertion did not influence the mutant phenotypes observed in the WPT677-3 mutant descendants. First, plant WPT677-3-47, which carried wild-type alleles for *GmKASI* and was heterozygous for the backbone insertion (**Table 1**), only produced wild type (not wrinkled) seeds. Second, T_2M_2 seeds harvested from descendants of the WPT677-6 family showed numerous wrinkled individuals, while presumably not carrying the backbone insertion, as they were derived from an independent T_0 event. These findings further confirm that the mutant phenotypes in this study are caused by the mutations at *GmKasI*.

Progeny from heterozygous mutant individuals typically displayed an intermediate mean phenotype for the seed composition traits (**Figure 4**). However, this is likely an outcome of the need to pool multiple individuals in each NIR run. We hypothesize that the progeny from the heterozygous lines consists of a combination of homozygous wild-type, heterozygous, and homozygous mutant individuals. Each individual seed likely displayed the composition traits of either the wild type group (presumably this includes both homozygous wild type and heterozygous individuals) or the extreme mutant phenotype (for homozygous mutant individuals). However, the pooled combination shows an intermediate phenotype. The traits tend to be more similar to the wild type composition presumably because the transmission of the homozygous mutant type is relatively low compared to the other groups (**Table 2**). It is also noteworthy that differences were observed in the severity of the seed composition traits between the first and second greenhouse experiments. While the experiments were grown using the same families and in the same greenhouse, they were grown at different times. These differences between experiments suggest that seed composition traits are influenced by microenvironment differences, even in greenhouse conditions.

The second question of interest addressed the possibility that an in-frame mutation of *GmKasI* may provide a less extreme (i.e., intermediate) seed composition phenotype compared to the knockout mutations. This idea was previously demonstrated in a series of soybean trichome mutants—the in-frame alleles of a *CPR5* gene ortholog demonstrated an intermediate phenotype compared to the wild type and knockout mutants (Campbell et al., 2019). This "weak allele" concept is a promising avenue for using gene editing as a means to develop agriculturally useful phenotypes in cases where full knockout alleles exhibit phenotypes that are too extreme and/or cause secondary undesirable phenotypes. However, the data from our in-frame mutant lines did not support this outcome in the present study. The mutant individuals carrying the homozygous in-frame allele also exhibited very strong alterations in seed sucrose and oil content, comparable to the knockout lines. Thus, an intermediate (i.e., less severe) phenotype was not observed in the in-frame mutant. However, it is notable that the in-frame

TABLE 2 | Non-Mendelian inheritance of *GmkasI* (Glyma.08G084300) mutant alleles.

Plant-ID	Genotype	Planted	Germination	WT	Het	Mut	Chi^2 value	Chi^2 p-value
Bert	WT control	29	100.00%	29	0	0	NA	NA
677-3-22-10	Heterozygous	69	89.86%	27	32	3	18.645	0.00009
677-3-43	Heterozygous	162	96.91%	53	86	18	17.038	0.0002
677-3-48	Heterozygous	89	88.76%	28	45	6	13.785	0.00102

"WT," "Het," and "Mut" are respectively the wild-type, heterozygous, and homozygous mutant segregants derived from a heterozygous parental plant.

mutation was in a highly evolutionarily conserved region (**Supplementary Figure S1**). It may have been expected that perturbations to amino acids in such a conserved region may have strong phenotypic consequences. Perhaps in-frame mutations in a less conserved domain of *GmKasI* would yield intermediate phenotypes.

The third question of interest addressed the low transmission of soybean *kasI* mutant alleles. While reduced transmission of the *kasI* mutant was observed in the fast neutron line, this particular mutation was linked to a translocation event that may have also disrupted meiosis. However, the progeny of the CRISPR lines in the current study heterozygous for the mutations also exhibited low transmission of the mutated alleles, while showing no indications of chromosomal abnormalities. Thus, we conclude that the low transmission of *GmkasI* mutations is a consequence of the mutations *per se*. While we do not know the mechanism for the reduced transmission, it is possible that post-fertilization defects during embryogenesis or seed development may lead to lethality in some of the homozygous mutant individuals.

Taken together, these results from this study confirm that *GmKASI* is an important gene for regulating the sucrose to oil biosynthesis pathway in soybean. This is a function that has also been demonstrated in Arabidopsis (Wu and Xue, 2010). This observation has now been made across three different mutant classes of soybean, including a translocation line, CRISPR knockout alleles, and a CRISPR in-frame allele.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA640373.

AUTHOR CONTRIBUTIONS

All authors contributed to the article and approved the submitted version. KV, GM, and RS conceived and designed the experiments. KV conducted the CRISPR experiments and analyzed the data. MS contributed to genotyping and greenhouse experiments. AS contributed to Sanger sequencing and heteroduplex assays. YX generated transformed plants. RM conducted whole genome sequencing bioinformatics. KV and RS drafted the manuscript.

FUNDING

This work was supported in part by the Minnesota Soybean Research and Promotion Council (projects #150-4140-18-02 and #10-15-47-19-174-7527), and by an endowed chair of molecular genetics to crop improvement.

ACKNOWLEDGMENTS

The authors are grateful to Shaun J. Curtin for providing CRISPR/Cas9 reagents and Jean-Michel Michno for assistance with data management. The authors would also like to thank Austin Dobbels for helpful discussions related to the soybean *kasI* FN mutant.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.01005/full#supplementary-material

REFERENCES

Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data*. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Campbell, B. W., Hoyle, J. W., Bucciarelli, B., Stec, A. O., Samac, D. A., Parrott, W. A., et al. (2019). Functional analysis and development of a CRISPR/Cas9 allelic series for a CPR5 ortholog necessary for proper growth of soybean trichomes. *Sci. Rep.* 9, 14757. doi: 10.1038/s41598-019-51240-7

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* 16, 10881–10890. doi: 10.1093/nar/16.22.10881

Curtin, S. J., Xiong, Y., Michno, J. M., Campbell, B. W., Stec, A. O., Čermák, T., et al. (2018). CRISPR/Cas9 and TALENs generate heritable mutations for genes involved in small RNA processing of Glycine max and Medicago truncatula. *Plant Biotechnol. J.* 16, 1125–1137. doi: 10.1111/pbi.12857

Dobbels, A. A., Michno, J. M., Campbell, B. W., Virdi, K. S., Stec, A. O., Muehlbauer, G. J., et al. (2017). An Induced Chromosomal Translocation in Soybean Disrupts a KASI Ortholog and Is Associated with a High-Sucrose and Low-Oil Seed Phenotype. *G3 (Bethesda)* 7, 1215–1223. doi: 10.1534/g3.116.038596

Hsiau, T., Conant, D., Rossi, N., Maures, T., Waite, K., Yang, J., et al. (2018). Inference of CRISPR edits from Sanger trace data. *bioRxiv*. doi: 10.1101/251082

Li, H., and Durbin, R. (2010). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

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, J., Gunapati, S., Mihelich, N. T., Stec, A. O., Michno, J. M., and Stupar, R. M. (2019). Genome Editing in Soybean with CRISPR/Cas9. *Methods Mol. Biol.* 1917, 217–234. doi: 10.1007/978-1-4939-8991-1_16

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 10–12. doi: 10.14806/ej.17.1.200

Michno, J. M., Virdi, K., Stec, A. O., Liu, J., Wang, X., Xiong, Y., et al. (2020). Integration, abundance, and transmission of mutations and transgenes in a series of CRISPR/Cas9 soybean lines. *BMC Biotechnol.* 20, 10. doi: 10.1186/s12896-020-00604-3

Orf, J. H., and Kennedy, B. W. (1992). Registration of "Bert" soybean. *Crop Sci.* 32, 830. doi: 10.2135/cropsci1992.0011183X003200030058x

Paz, M. M., Martinez, J. C., Kalvig, A. B., Fonger, T. M., and Wang, K. (2006). Improved cotyledonary node method using an alternative explant derived from mature seed for efficient Agrobacterium-mediated soybean transformation. *Plant Cell Rep.* 25, 206–213. doi: 10.1007/s00299-005-0048-7

R Core Team (2018). *R: A language and environment for statistical computing* (Vienna, Austria: R Foundation for Statistical Computing).

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., et al. (2011). Integrative genomics viewer. *Nat. Biotechnol.* 29, 24–26. doi: 10.1038/nbt.1754

Severin, A. J., Woody, J. L., Bolon, Y. T., Joseph, B., Diers, B. W., and Farmer, A. D. (2010). RNA-Seq Atlas of Glycine max: a guide to the soybean transcriptome. *BMC Plant Biol.* 10, 160. doi: 10.1186/1471-2229-10-160

Veena, V., and Taylor, C. G. (2007). Agrobacterium rhizogenes: recent developments and promising applications. *Vitro Cell Dev-Pl.* 43, 383–403. doi: 10.1007/s11627-007-9096-8

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

Wu, G. Z., and Xue, H. W. (2010). Arabidopsis β-ketoacyl-[acyl carrier protein] synthase I is crucial for fatty acid synthesis and plays a role in chloroplast division and embryo development. *Plant Cell* 22, 3726–3744. doi: 10.1105/tpc.110.075564

Zhu, X., Xu, Y., Yu, S., Lu, L., Ding, M., Cheng, J., et al. (2014). An efficient genotyping method for genome-modified animals and human cells generated with CRISPR/Cas9 system. *Sci. Rep.* 4, 6420. doi: 10.1038/srep06420

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 05 August 2020
doi: 10.3389/fpls.2020.00925

Production of Herbicide-Sensitive Strain to Prevent Volunteer Rice Infestation Using a CRISPR-Cas9 Cytidine Deaminase Fusion

Akira Komatsu[1]*, *Miki Ohtake*[1], *Zenpei Shimatani*[2] and *Keiji Nishida*[2]

[1] Institute of Agrobiological Sciences, National Agriculture and Food Research Organization (NARO), Tsukuba, Japan,
[2] Graduate School of Science, Technology and Innovation, Kobe University, Kobe, Japan

OPEN ACCESS

Edited by:
Sandeep Kumar,
Corteva Agriscience™, United States

Reviewed by:
Yiping Qi,
University of Maryland, United States
Terry R. Wright,
Corteva Agriscience™, United States

***Correspondence:**
Akira Komatsu
akomatsu@affrc.go.jp

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *31 March 2020*
Accepted: *05 June 2020*
Published: *05 August 2020*

Citation:
Komatsu A, Ohtake M, Shimatani Z and Nishida K (2020) Production of Herbicide-Sensitive Strain to Prevent Volunteer Rice Infestation Using a CRISPR-Cas9 Cytidine Deaminase Fusion. Front. Plant Sci. 11:925. doi: 10.3389/fpls.2020.00925

When cultivated rice seed fall into fields, they may overwinter and spontaneously germinate the next spring. Such germinated plants are termed "volunteer rice." Volunteer grains originating from feed rice varieties may differ in certain traits, such as quality and taste, as compared with those of rice cultivated for human consumption, which may reduce the overall quality of the final harvested grain. Many rice varieties show resistance to benzobicyclon (BBC), a beta-triketone herbicide (bTH) that inhibits 4-hydroxyphenylpyruvate dioxygenase (HPPD). Recently, the rice gene *HIS1* (*HPPD INHIBITOR SENSITIVE 1*) conferring resistance to BBC and other bTHs was identified. In this study, to suppress the occurrence of volunteer rice infestation, we attempted to generate a BBC-sensitive rice strain *via* the knockout of the *HIS1* gene using genome editing techniques. The production of a *his1* knockout line was carried out by the start-codon substitution or stop-codon creation using CRISPR-Cas9 cytidine deaminase fusion, which is useful as a novel amino acid sequence is not generated due to the shifting of the reading frame. The mutation frequencies of independent transgenic plants were 3.6, 13.5, 13.8, and 21.2% at four gRNAs for start-codon substitution and three stop-codon creations. The *his1* knockout lines were conferred with sensitivity to BBC, re-confirming by genome editing that this is indeed the gene responsible for BBC resistance/sensitivity. The *his1* knockout lines also exhibited a sensitive phenotype to other bTHs, including sulcotrione, mesotrione, tembotrione, and tefuryltrione, compared with the wild-type variety 'Nipponbare.' These results demonstrate the potential of herbicide-sensitive rice produced by genome editing technology as a material to control volunteer feed rice using pre-labeled herbicides for varieties consumed by humans.

Keywords: activation-induced cytidine deaminase (AID), *Oryza sativa* L., targeted nucleotide substitution, *HIS1*, volunteer rice, benzobicyclon, mesotrione, sulcotrione

INTRODUCTION

Genome editing technology has been attracting attention as a means to change and modify the target domains within genes/genomes in diverse species. Recently, the CRISPR-Cas9 system has been developed, which employs a Cas9 endonuclease and guide RNA complex, and has exhibited very high efficiency for target gene editing in various species (Cong et al., 2013; Feng et al., 2013; Jiang et al., 2013; Mali et al., 2013; Nekrasov et al., 2013; Shan et al., 2013). In addition, efforts are in progress to develop a novel genome editing technique that does not involve DNA double-strand breaks. Cytosine base editor (CBE) is one such technology, which employs activation-induced cytidine deaminase (AID) that catalyzes a deamination reaction and couples with molecules harboring DNA sequence recognition ability, thereby modifying the genome sequence *via* nucleotide substitution in domains containing specific DNA sequences. AID is an enzyme that converts cytosine within DNA/RNA into uracil through a deamination reaction. The resulting U-G DNA mismatches trigger nucleotide substitutions (C to T or G to A) through DNA replication and repair pathways. AID prefers single-strand DNA within RNA/DNA hybrid transcripts (R loops) as its substrate. By making complete use of nuclease-impaired CRISPR-Cas9 and AID-related deaminases, researchers have recently developed a BE using rat APOBEC1 (Komor et al., 2016) and Target-AID system using sea lamprey-derived PmCDA1 to demonstrate the direct installation of point mutations (Nishida et al., 2016). The mutation spectrum of Target-AID is highly specific; the system preferentially induces point mutations at cytosine bases within a 5-bp window surrounding the 18 bp upstream of the protospacer adjacent motif (PAM) sequence on the non-complementary strand to gRNA. The capability of Target-AID to avoid cutting genomic double-stranded DNA reduces off-target effects relating to CRISPR-Cas9 nuclease activity and the incidence of cellular cytotoxicity (Nishida et al., 2016). Target-AID system was also applied to higher plants to obtain desirable phenotypes in rice, tomatoes, wheat, and maize (Shimatani et al., 2017; Zong et al., 2017). Simultaneous editing of multiplex traits was also demonstrated in rice (Shimatani et al., 2018). Recently, rationally engineered Cas9 requiring NG-PAM sequences (Cas9-NG) was developed and shown to be compatible with Target-AID (Nishimasu et al., 2018), which was then applied to rice (Endo et al., 2019). Recently, adenine base editors (ABEs) were developed as a system for mediate the conversion of A to G or T to C in genomic DNA (Gaudelli et al., 2017). These systems, together with cytosine base editors by cytidine deaminase, enable introduction of transitions for four patterns (C to T, G to A, A to G, and T to C) at the target site in the genome, expanding the capabilities of base editing. In this manner, the base editing technology continues to evolve and is expected to be capable of achieving breeding goals more efficiently.

Rice is one of the most important crops in the world, with more than half of the global population relying upon it as a staple food (Sasaki, 2008). It is also essential as livestock feed. Worldwide, especially in advanced or emerging countries where the consumption of livestock meat is already high or rapidly increasing, domestic cultivation, and utilization of feed crops is key to maintaining self-sufficiency. Currently, a multitude of high-yield rice varieties have been cultivated specifically for livestock through agricultural breeding efforts, for instance, in Japan (Sakai et al., 2003; Kato, 2008). As the Japanese government encourages the cultivation and harvest of feed rice, farmers have formulated annual production plans by gauging the optimum ratio of edible rice (i.e., human consumption) to feed rice (i.e., livestock consumption) varieties. Under such circumstances, when seeds of cultivated rice migrate onto a paddy field, some of them may overwinter and germinate spontaneously on the following spring. Such germinated seedlings are called "volunteer rice" (Singh et al., 2017). The occurrence of volunteer rice originating from a feed rice variety results in the mixing of feed rice varieties within the edible rice crops to be grown the next year. Edible and feed varieties of rice considerably differ in terms of quality and taste. Hence, such a mixing may cause decline in the quality and lowering its sell prices of harvested rice grains intended for human consumption. To avoid quality loss by such a crop contamination, farmers have a tendency to intentionally cultivate low-yield edible rice varieties as feed rather than using higher-yield varieties bred specifically for livestock consumption.

Beta-triketone herbicides (bTHs) are 4-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitors widely used in agriculture. Benzobicyclon (BBC) is a bTH developed for weed control in paddy fields and is effective against paddy weeds resistant to other types of herbicide, including sulfonylureas. BBC is a prodrug, with its hydrolysate, BBC-OH, is incorporated into plant, acting as an HPPD inhibitor, and mediating plant bleaching (Sekino et al., 2008). Recent studies have revealed that although many Japanese rice varieties show resistance to BBC, certain varieties, including some *Indica* cultivars, remain sensitive to BBC (Maeda et al., 2019). These findings contribute to the identification of rice genes that determine resistance and sensitivity relative to BBC. Specifically, the main quantitative trait locus (QTL) for BBC sensitivity situated on the second chromosome was identified through a QTL analysis of the BC_1F_2 population and chromosomes derived from BBC-sensitive and BBC-resistant rice varieties. As this QTL behaves as a single recessive trait, it was identified as a gene locus determining BBC sensitivity/resistance, with the corresponding wild-type gene being named *HIS1* (Maeda et al., 2019). In addition, an analysis of BBC-sensitive rice cultivar revealed that BBC-OH was absorbed into the roots and transferred to the foliar tissue, but it was not detected in the leaf tissues of the *HIS1*-transformed line, suggesting that HIS1 functions as an enzyme for BBC-OH degradation in plants (Maeda et al., 2019).

In this study, to preclude rice varieties intended for either human or livestock consumption from being mixed together due to field infestation by volunteer rice, we attempted to produce a BBC-sensitive *his1* rice line *via* the knockout of the *HIS1* gene through nucleotide substitution using a CRISPR-Cas9 cytidine deaminase fusion.

MATERIALS AND METHODS

Plasmid Vector Construction

The vectors used for the evaluation of herbicide resistance within rice were constructed as follows. First, a rice-optimized Target-AID vector (**Figure 1A**) was constructed. Rice-optimized *Streptococcus pyogenes* Cas9 (*SpCas9*) vector pZH_MMCas9 and gRNA construction vector pZK_OsU6-gRNA were kindly provided by Dr. Masaki Endo (Mikami et al., 2015), and a RuvC nuclease domain-deficient D10A mutation (nickase, *nCas9*) was introduced *via* PCR and Gibson Assembly methods (New England Biolabs, USA). A plant codon-optimized *PmCDA1* coding sequence was synthesized and inserted following *nCas9* using the same linker peptide as previously described (Nishida et al., 2016). This pZK_OsU6-sgRNA was modified to contain a *Mlu*I site after the *Spe*I site for ease of multiplexing. A target sgRNA sequence was inserted between the OsU6 promoter and sgRNA scaffold using PCR.

Transformation and Regeneration

BBC-resistant *japonica* cultivar (*Oryza sativa* L cv. Nipponbare) was used in this study. The procedures for *Agrobacterium*-mediated transformation were performed as described previously (Toki, 1997; Toki et al., 2006). After co-cultivation of *Agrobacterium* carrying the Target-AID vector with rice scutellum-derived calli (pre-cultured for 5 days) for 3 days, infected calli were transferred to fresh callus induction medium (Toki et al., 2006) containing 50 mg/L hygromycin B (Wako Pure Chemicals) and 400 mg/L carbenicillin disodium salt (Nakarai, Kyoto, Japan) to remove residual *Agrobacterium*. At 21 days following hygromycin selection, proliferating calli were transferred to a fresh pre-regeneration medium containing 40 mg/L hygromycin and 200 mg/L carbenicillin disodium salt. After 8 days of culture, the calli were transferred to a fresh regeneration medium containing 30 mg/L hygromycin B and cultured for 2weeks. The regenerated plants were further cultivated in a greenhouse.

Sequence Analysis

PCR products used for the restriction fragment length polymorphism analysis were also cloned into pCR-Blunt II-TOPO (Invitrogen) and subjected to a sequencing analysis using an ABI 3130 sequencer (Applied Biosystems).

Herbicide Susceptibility Tests for Genome-Edited Lines *In Vitro*

The herbicide susceptibility of genome-edited lines was measured within test tubes (diameter, 2.5 cm; height, 15 cm) containing 10 ml of a Murashige–Skoog (MS) solid medium with herbicide. Five dehusked mature seeds of homozygous *his1* T_3 or wild-type (control) rice plants were surface sterilized *via* two treatments with 4% sodium hypochlorite for 20 min followed by five rinses with sterilized water. The seeds were then immersed in sterilized water for 2 days at 30°C, after which germinated seeds were transferred to tubes containing the solid MS medium composed of half-strength MS salts and agar (1 g/L) containing herbicide; BBC 0.1 μM, BBC-OH 0.1 μM, and 0.3 μM; Mesotrione (MST); Sulcotrione (SLT); Tembotrione (TMT) 0.05 μM and 0.1 μM; and Tefuryltrione (TFT) 0.1 μM and 0.3 μM, cultured at 27°C for 7 to 14 days with 16 h of light (40 μmol m^{-2} s^{-1}) daily. BBC and BBC-OH were obtained from SDS Biotech (Tokyo, Japan), while MST, SLT, TMT, and TFT were obtained from Fujifilm Wako (Tokyo, Japan).

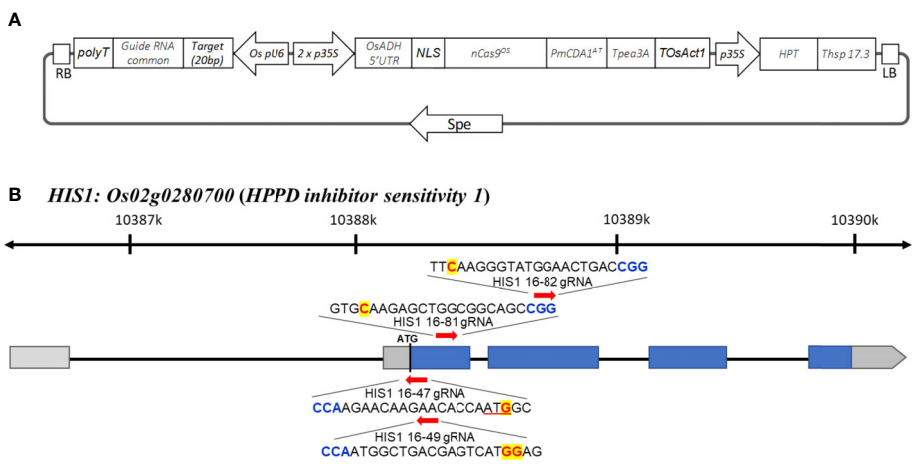


FIGURE 1 | Target-AID vector and targeted sites of rice *HIS1* gene. **(A)** Schematic illustration for the Target-AID vector in this study. The engineered Cas9 nuclease *nCas9^{Os}-PmCDA1^{At}* is under the control of the doubled *cauliflower mosaic virus 35S* promoter, and transcription was terminated with the *Pea3A* terminator from *Pisum sativum*. The *U6* promoter of rice (*OsU6*) was used to drive the expression of gRNAs. **(B)** Schematic map of the gene structure of *HIS1* and the gRNA target sites. The blue and gray boxes indicate translated and untranslated regions (UTRs) on exon(s), respectively; arrows show the positions and direction of gRNA on *HIS1* gene for Target-AID. The PAM motif (NGG) is shown in blue. The target nucleotides for base substitution are shown in red.

Evaluation of BBC and Other bTH Sensitivities of Genome-Edited Lines

The germination of Nipponbare (control) and genome-edited rice seeds was induced with incubation at 30°C for 2 to 3 days with 16-h light (40 μmol m^{-2} s^{-1}) and 8-h dark photoperiods. Approximately 300 ml of soil specific for rice planting (Bonsol Baido, Sumitomo Kagaku, Tokyo, Japan) were mixed with water and deposited into plastic containers. Then, the germinated rice seeds were placed upon the soil surface. Seedlings were grown at 30°C for 7 to 10 days in a greenhouse, with the water level being maintained coincident with the soil surface. After the first leaves expanded, either water only or water containing herbicide (BBC: 0, 0.37, 0.75, 1.49, 2.24, and 4.48 μM; MST: 0.13 μM; SLT: 0.14 μM) was added to each cup so that the surface of the soil was 30 to 40 mm below that of the liquid. The liquid level was then maintained by occasional watering. Seedling growth was checked after 14 days.

RESULTS AND DISCUSSION

Production of an *his1* Knockout Genome-Edited Line by Target-AID

The Target-AID system was used to produce a *his1* knockout rice line (*his1* line). At four points on the *HIS1* gene, 20 bp gRNA sequences were designed (**Figure 1B**). One of such gRNA sequences (16-47 gRNA) introduced a mutation in the start-codon, with a G within the start-codon (ATG) on the second exon being substituted by A. Consequently, the starting methionine is converted to isoleucine (**Figure S1A**). This makes translation initiation impossible; therefore, we predicted that a knockout phenotype would emerge. The gRNA sequences at the three remaining points were designed using AID to create nonsense mutations: the first (16-49 gRNA) converts TGG-encoding tryptophan at the second exon into a stop-codon (TAA, TGA, or TAG; **Figure S1B**), the second (16-81 gRNA) is also situated on the second exon and converts CAA encoding a glutamine residue into a TAA stop-codon (**Figure S1C**), the final gRNA sequence (16-82 gRNA) is situated on the third exon and similarly converts CAA encoding a glutamine residue into a TAA stop-codon (**Figure S1D**). We used the Target-AID system carrying these gRNAs to edit the targets with the aim to facilitate the production of knockout plants. We investigated the efficiency of mutation with the 16-47 gRNA for the start-codon substitution and the other three gRNA sequences (16-49, 16-81, and 16-82 gRNAs) for the stop-codon creation. Following antibiotic hygromycin screening, 83, 52, 65, and 33 regenerated plants were obtained, respectively. For each mutation site, a sequence analysis was performed. Consequently, the mutation efficiency, including insertion, deletion, and substitution, was 12.0% for the 16-47 gRNA, 46.2% for the 16-49 gRNA, 36.9% for the 16-81 gRNA, and 36.4% for the 16-82 gRNA. In addition, the efficiency of indel integration was 9.6, 30.7, 7.7, and 9.1%, whereas the efficiency of the substitution alone was 3.6, 15.4, 29.2, and 27.3%. Finally, the number of lines achieving the targeted substitution rates was 3 (3.6%), 7 (13.5%), 9 (13.8%), and 7 (21.2%) (**Table 1**, **Figure 2**). Therefore, efficiency was the lowest for gRNA initiating the start-codon substitution. In addition, biallelic mutations with only base substitution were only found in one line (1.9%) with 16-49 gRNA.

As the efficiency of the start-codon substitution was lower than that of the stop-codon creation, in the latter case, as there are multiple sites for which gRNA can be designed within the exon domain, a more efficient substitution target site could be selected. However, for a start-codon substitution, only one predetermined site can be used to design gRNA, and it may not always be an optimal sequence.

The efficiency of the indel integration or point mutations using Target-AID in this study was similar to that previously reported for the *OsFTIP1e* gene (*nCas9^{Os}-PmCDA1^{At}*) in rice (Shimatani et al., 2017). In our previous report, we discovered that UGI suppresses indel formation and improves the targeted nucleotide substitution efficiency in mammalian cells (Nishida et al., 2016). However, no clear effect of UGI on indel frequency was confirmed in plant cells (Nishida and Shimatani, unpublished). Zhong et al. (2019) reported that the efficiencies of substitution mutations using Cas9-NG (D10A)-PmCDA1-UGI were 30 and 45%, respectively, which exceeds our maximum efficiency (29.2%). Indel frequency also tended to be slightly lower, suggesting the effect of UGI. Furthermore, a study by Tang et al. (2019) reported that the production efficiencies of base-substituted plants using STU-nCas-PmCDA1-UGI-tRNA were 38.9 and 68.8%, which are both higher than our reported production efficiency. However, the frequency of the indel lines were 16.7 and 25.0%, respectively, and some test plots were considerably higher than our results. These results suggest that the use of UGI does not necessarily lead to increased substitution

TABLE 1 | Mutation frequency at *HIS1* targets using Target-AID.

Variety	gRNA	Regenerated plants of Hyg^R	Regenerated plants with indel mutants	Regenerated plants with substitution mutants	Regenerated plants with start codon mutation or creating stop codon
Nipponbare	*HIS1* 16-47 (Exon2)	83	8 (9.6%)	3 (3.6%)	3 (3.6%)
	HIS1 16-49 (Exon2)	52	16 (30.7%)	8 (15.4%)	7 (13.5%)
	HIS1 16-81 (Exon2)	65	5 (7.7%)	19 (29.2%)	9 (13.8%)
	HIS1 16-82 (Exon3)	33	3 (9.1%)	9 (27.3%)	7 (21.2%)

The transformed and hygromycin-resistant calli were analyzed by sequencing to identify mutation types. Percentage of T_0 plants found with mutations in the target sequence.

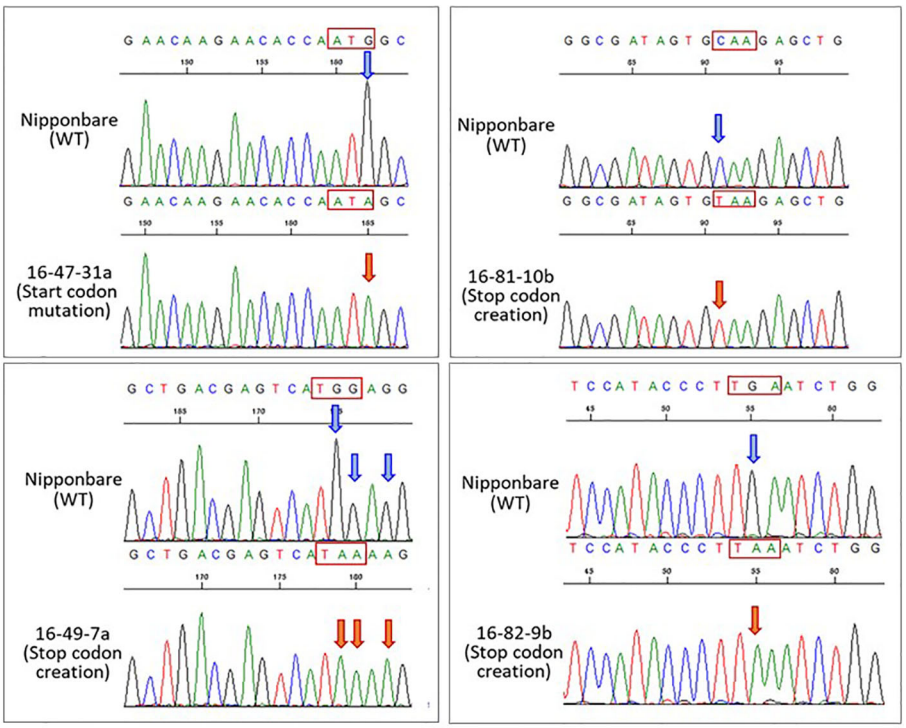


FIGURE 2 | Sequencing chromatograms on nucleotide substitution regions in four *his1* lines. Sequencing chromatograms showing the nucleotide substitutions by Target-AID in 16-47, 16-49, 16-81, and 16-82 gRNAs in the *HIS1* gene of a representative T_2 plant. Blue arrowheads indicate the wild-type nucleotide. Orange arrowheads indicate nucleotide substitutions. The nucleotide sequences in the red box indicate the target codon.

efficiency by reducing indel frequency. Finally, the fact that multiple lines or strains containing start-codon substitutions or stop-codon creation mutations were produced for all four gRNA sequences demonstrates that the mutation efficiency is at practical levels in rice. In cases of start-codon substitutions or stop-codon creation mutations using Target-AID, no new amino acid sequence(s) caused by novel reading frames *via* frame shift are occurring. This case may be considered another advantage for approaches using Target-AID and other CBEs.

Confirmation of Off-Target Mutations

The Target-AID system is used for nucleotide substitutions aimed at specific amino acid residue substitutions. As a result, there are often situations in which the use of gRNA with off-target candidates cannot be avoided. In this study, we investigated how many mismatches with gRNA sequences cause off-targeting.

A family of genes similar to the *HIS1* gene (*HIS1*-like gene: *HSL*) exists in the rice genome on the sixth chromosome in the form of five clusters. From the OsHSL proteins predicted in previous reports, Os06g0176700 (*HSL1A*) and Os06g0178700 (*HSL1B*) are the most similar to *HIS1* (87% sequence identity) (Maeda et al., 2019). We investigated the presence/absence of off-target mutations, including these genes. As a result, we confirmed an off-target mutation (11-base insertion) of *HSL1B* in one out of the three plants within the genome-edited line (16-47) targeting the start-codon sequence (**Table S1**). The 11-bp insertion site of *HSL1B* was predicted to be a region where cytidine deaminase would perform base substitution, and it was presumed that the 11 bp sequence before this region was inserted as a template. As a result, this region became a tandem repeat sequence due to the 11 bp insertion sequence (**Figure S2**).

Conversely, we observed no off-target mutations in the genome-edited lines for the creation of stop codons (**Tables S2–S4**). In the genome-edited plants generated with the 16-49 gRNA, no off-target mutations were observed despite the off-target candidate genes on the sixth chromosome having 100% homology with *HIS1*, except for the PAM sequences (**Table S2**). This result indicates that although the CRISPR-Cas9 system can target any genomic region, depending upon the gRNA design, the specificity of this system depends on the PAM located directly under the target sequence. These target sites must lie immediately 5′ of a PAM sequence that matches the canonical 5′-NGG form, although recognition at sites containing alternate PAM sequences (e.g., 5′-NAG) has also been reported, albeit at less efficient rates (Jinek et al., 2012; Jiang et al., 2013; Pattanayak et al., 2013).

As the *HIS1* gene targeted in this study exhibited a high level of homology with the *HSL* family, it seemed to be a case in which single-base substitution of *HIS1* is problematic. The data

showed that only one plant had the off-target mutation (11-base insertion) on *HSL1B* in a start-codon substitution (gRNA: 16-47), with no other alternative codon. Thus, the creation of a base substitution in a specific codon sequence was found to infer more disadvantageous conditions for mutation efficiency and off-target mutations compared with obtaining indel mutant plants using CRISPR-Cas9. However, in this study, the acquisition of the target base substitution was demonstrated in all target regions and was without mutation(s) for any candidate genes at the off-target sites. Even if off-target mutations occur rarely, they can be detected and eliminated during the selection process in crop breeding. This finding indicates that the base substitution within the target region of the rice genome has been extremely difficult for conventional breeding, but it has been made possible using Target-AID. This development can lead to more efficient and precise breeding in the future.

Evaluation of BBC and Other β-Triketone Sensitivities *In Vitro*

We created knockout rice lines by introducing mutations into *HIS1 via* base editing followed by the evaluation of their BBC sensitivity. For these analyses, the BBC-resistant cultivar 'Nipponbare' was used as an original variety and homozygous T_3 plants were used for the genome-edited line. The 16-47-31a (single-base substitution) and 16-47-32b (12 base deletions) lines carried mutations in the start-codon "ATG" composed of the 16-47 gRNA among the four gRNA sequences designed for *HIS1* (**Table 2**). In contrast to the wild-type 'Nipponbare' exhibiting BBC resistance, these lines apparently acquired sensitivity to 0.1

TABLE 2 | Mutation pattern of genome-edited lines used for evaluation of BBC and other β-triketone sensitivity.

gRNA	*HIS1* genome-edited line	Mutation pattern	
16-47 gRNA	Wild type	GAT**CCA**AGAACAAGAACACCA**ATG**G	
	16-47-20b	GAT**CCA**AGAACAAGAACACC**GAACACC**AATGG	7-base insertion (before ATG)
	16-47-31a	GAT**CCA**AGAACAAGAACACCAAT**A**G	1-base substitution (ATG → ATA)
	16-47-32b	GAT**CCA**AGAACAA----------CTG	12-base deletion (including ATG)
16-49 gRNA	Wild type	AACA**CCAATG**GCTGACGAGTCAT**GG**AGGG	
	16-49-1a	AACA**CC**----------GG	21-base deletion (including ATG)
	16-49-7a	AACA**CCAATG**GCTGACGAGTCAT**AA**A**A**GG	3-base substitution (stop codon creation)
	16-49-14b	AACA**CCAATG**GC----------CGG	12-base deletion (including ATG)
	16-49-28a	AACA**CCAATG**GCTGACGAGTCAT**CA**A**A**GG	3-base substitution (Trp Arg → Ser Lys)
16-81 gRNA	Wild type	GGCGATAGTG**C**AAGAGCTGGCGGCAGC**CGG**	
	16-81-4b	GGCGATAGTG**G**AAGAGCTGGCGGCAGC**CGG**	1-base substitution (Gln → Glu)
	16-81-9b	GGCGATAGTGCAAGAGC-------AGC**CGG**	7-base deletion
	16-81-10b	GGCGATAGTG**T**AAGAGCTGGCGGCAGC**CGG**	1-base substitution (stop codon creation)
16-82 gRNA	Wild type	TTCCAGATT**C**AAGGGTATGGAACTGAC**CGG**	
	16-82-4b	TTCCAGATT**G**AAGGGTATGGAACTGAC**CGG**	1-base substitution (Gln → Glu)
	16-82-9b	TTCCAGATT**T**AAGGGTATGGAACTGAC**CGG**	1-base substitution (stop codon creation)
	16-82-9c	TTCCAGATTCAAGG----------AC**CGG**	11-base deletion

The target nucleotide for base substitution is shown in red. The PAM motif (NGG) is shown in a light blue highlight. Mutation of nucleotide substitution is shown in a green highlight. Mutation by insertion is shown in blue. Mutations by deletion are shown in ----. The yellow highlights show the start-codon.

TABLE 3 | Evaluation of BBC and other β-triketone sensitivity *in vitro*.

gRNA	*HIS1* genome-edited line	Mutation pattern	BBC 0.1μM	BBC-OH 0.1μM	BBC-OH 0.3μM	Mesotrione 0.05μM	Mesotrione 0.1μM	Sulcotrione 0.05μM	Sulcotrione 0.1μM	Tembotrione 0.05μM	Tembotrione 0.1μM	Tefuryltrione 0.1μM	Tefuryltrione 0.3μM
16-47 gRNA	16-47-20b	7-base insertion (before ATG)	R	R	R	R	R	RS	R	RS	RS	RS	R
	16-47-31a	1-base substitution (ATG → ATA)	S	S	S	SS	SS	S	SS	S	SS	S	S
	16-47-32b	12-base deletion (including ATG)	S	S	S	SS	SS	S	SS	S	SS	S	S
16-49 gRNA	16-49-1a	21-base deletion (including ATG)	S	S	S	SS	SS	S	SS	S	SS	S	S
	16-49-7a	3-base substitution (stop codon creation)	S	S	S	SS	SS	S	SS	S	SS	S	S
	16-49-14b	12-base deletion (including ATG)	S	S	S	SS	SS	S	SS	S	SS	S	S
	16-49-28a	3-base substitution (Trp Arg → Ser Lys)	RS	RS	RS	RS	S	RS	S	RS	S	RS	RS
16-81 gRNA	16-81-4b	1-base substitution (Gln → Glu)	S	RS	S	R	R	RS	S	RS	S	RS	S
	16-81-9b	7-base deletion	S	S	S	SS	SS	S	SS	S	SS	S	S
	16-81-10b	1-base substitution (stop codon creation)	S	S	S	SS	SS	S	SS	S	SS	S	S
16-82 gRNA	16-82-4b	1-base substitution (Gln → Glu)	R	R	R	R	R	R	R	R	R	R	R
	16-82-9b	1-base substitution (stop codon creation)	S	S	S	SS	SS	S	SS	S	SS	S	S
	16-82-9c	11-base deletion	S	S	S	SS	SS	S	SS	S	SS	S	S
Nipponbare		Wild type	R	R	R	R	R	R	R	R	R	R	R

Herbicide susceptibility of genome editing lines was tested in test tubes. SS: Leaves completely whitened (Sensitivity). S: Whitening, light green (Sensitivity). RS: Green (Middle of sensitivity and resistance). R: Dark green (Resistance).

μM BBC and its hydrolysate, BBC-OH within a concentration ranging from 0.1 to 0.3 μM (**Table 3**, **Figures 3A, B**). Meanwhile, the 16-47-20b line with a seven-base insertion mutation before the start-codon did not exhibit BBC sensitivity (**Table 3**).

The base substitution lines generated by Target-AID; 16-49-7a, 16-81-10b, and 16-82-9b were made to create a stop-codon on the second or third exon. In a 0.1-μM BBC and a BBC-OH concentration ranging from 0.1 to 0.3 μM, these lines were markedly more sensitive as compared with the wild-type (**Table 3**, **Figures 3C–E**). Similarly, the deletion line on the second or third exon; 16-49-1a, 16-49-14b, 16-81-9b, and 16-82-9c also displayed BBC and BBC-OH sensitivity. There was no difference in susceptibility to BBC and its hydroxide BBC-OH among sensitive strains edited by genome editing. This supports the results of Sekino et al. (2008) and Maeda et al. (2019), in which BBC is a prodrug and the BBC-OH, is incorporated into the plant and functions as an HPPD inhibitor, mediating plant bleaching.

By contrast, in the 16-49-28a line using the 16-49 gRNA and the 16-82-4b line using the 16-82 gRNA, a stop-codon was not created because a substitution mutation to different nucleotide from the purpose in the target region (C to G substitution). Therefore, their resistance to BBC was maintained (**Table 3**, **Figure 4**). Previous studies have shown that C to G substitutions using Target-AID occurred relatively frequently behind to C to T substitutions in yeast (Nishida et al., 2016). In addition, C to G substitution was also observed in the *OsFTIP1e* gene of rice (Shimatani et al., 2018). The use of UGI may improve efficiency while restricting C to T nucleotide substitution (Komor et al., 2016; Nishida et al., 2016; Endo et al., 2019). We are also currently examining the effect of UGI on nucleotide substitution in rice.

All lines sensitive to BBC were also sensitive to bTHs, including mesotrione (MST), sulcotrione (SLT), tembotrione (TMT), and tefuryltrione (TFT) (**Table 3**, **Figure 3**). The lines that underwent a base substitution to a different nucleotide from the target retained resistance to bTHs. Unlike other bTHs, BBC-sensitive lines were sensitive to higher concentration(s) of TFT. A previous study showed that transformed lines expressing *HSL1* did not manifest BBC resistance, however they showed increased resistance to TFT (Maeda et al., 2019). The mechanism of action of TFT may be slightly different from other bTHs.

These findings reconfirmed the previously reported result that *HIS1* is the primary gene determining BBC sensitivity/

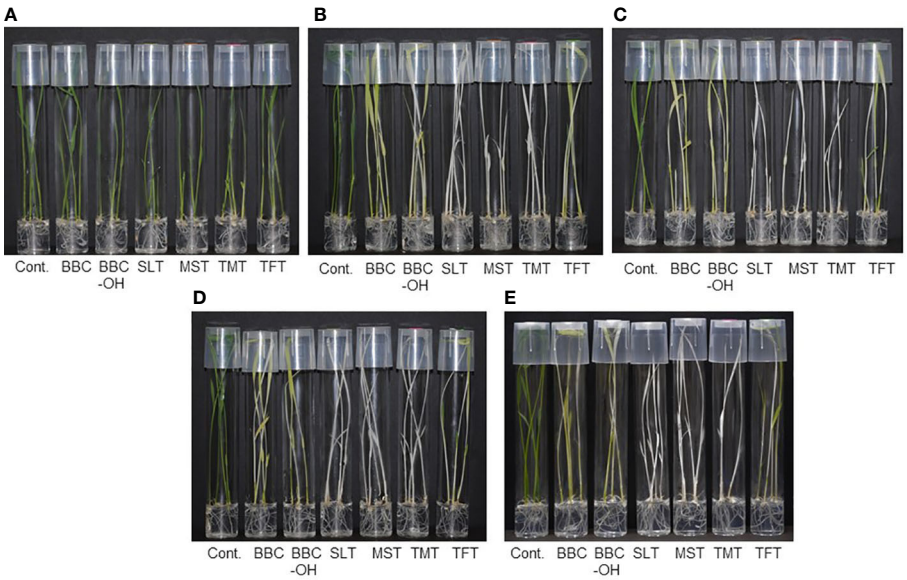


FIGURE 3 | Herbicide susceptibility of *HIS1* genome-edited rice lines. Wild-type Nipponbare **(A)**, *his1* knockout lines by nucleotide substitution, *HIS1*-16-47-31a **(B)**, *HIS1*-16-49-7a **(C)**, *HIS1*-16-81-10b **(D)**, and *HIS1*-16-82-9b **(E)**. Homozygous T_3 rice seeds were germinated and grown on Murashige–Skoog (MS) solid medium in the absence or presence of BBC (0.3 μM), BBC-OH (0.3 μM), MST (0.1 μM), SLT (0.1 μM), TMT (0.1 μM), or TFT (0.3 μM).

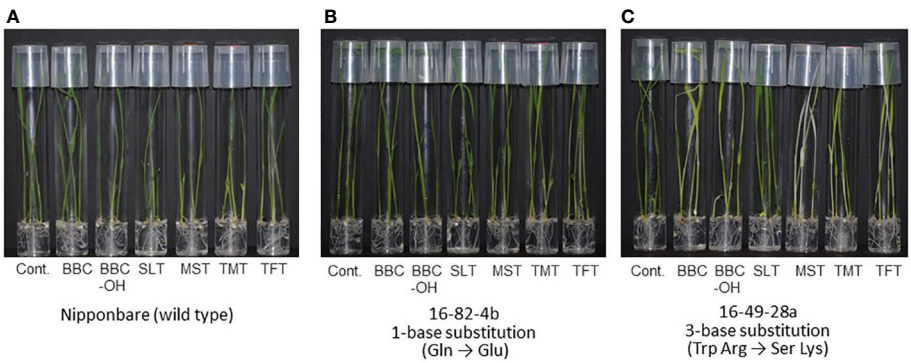


FIGURE 4 | Herbicide susceptibility of *HIS1* genome-edited rice lines. Herbicide susceptibility of wild-type Nipponbare **(A)**, single-base (1 aa) substitution line (16-82-4b) using Nipponbare with *HIS1*-16-82 as gRNA **(B)**, and three-base (2 aa) substitution line (16-49-28a) using Nipponbare with *HIS1*-16-49 as gRNA **(C)**. Homozygous T_3 rice seeds were germinated and grown on MS solid medium in the absence or presence of BBC (0.3 μM), BBC-OH (0.3 μM), MST (0.1 μM), SLT (0.1 μM), TMT (0.1 μM), or TFT (0.3 μM).

resistance while using knockout lines created by genome editing (Maeda et al., 2019). In addition, these results thus suggested the possible application of *his1* and genome editing to the breeding of crops sensitive to multiple bTHs.

Evaluation of BBC and Other β-Triketone Sensitivity in a Greenhouse

In a temperature-controlled greenhouse, we evaluated sensitivity to BBC and other β-triketones using the base-edited lines, including a T_3 *his*1 homozygous line with a start-codon substitution and three lines with a stop-codon creation. Consequently, as with *in vitro* testing, we confirmed withering due to BBC sensitivity in the *his1* base substitution line with varying BBC concentrations (**Figure 5**). Meanwhile, in a test pot at 0.37 μM BBC concentration, no deaths were observed, although growth was suppressed. These base substitution lines also exhibited sensitivity to other bTHs, MST (0.13μM), and SLT (0.14 μM) (**Figure 6**). This finding indicates that genome editing can produce *his1* knockout rice varieties in a short period of time, including varieties and strains that acquired sensitivity to BBC and other bTHs.

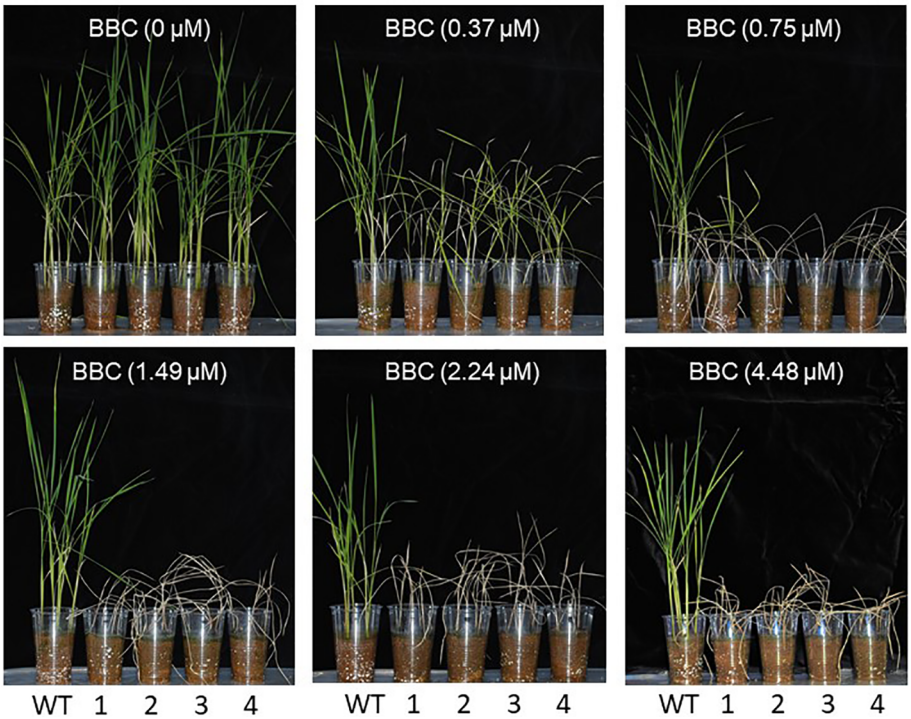


FIGURE 5 | Concentration-dependent effects of BBC on the cultivar Nipponbare (WT, *HIS1*) and *his1* homozygous mutant lines (T_3). 16-47-31a (1; start-codon mutant), 16-49-7a (2; stop-codon creation), 16-81-10b (3; stop-codon creation), and 16-82-9b (4; stop-codon creation) were grown in a controlled environment.

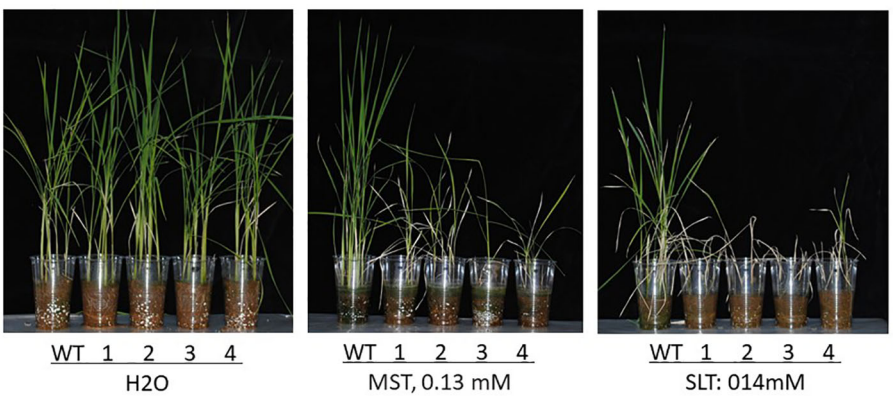


FIGURE 6 | Effects of bTHs (MST, 0.13 mM; SLT, 0.14 mM) on Nipponbare (WT, *HIS1*) and the *his1* homozygous mutant lines (T_3). 16-47-31a (1; start-codon mutant), 16-49-7a (2; stop-codon creation), 16-81-10b (3; stop-codon creation), and 16-82-9b (4; stop-codon creation) were grown in a controlled environment.

These results indicate that varieties and strains having acquired sensitivity to BBC and other bTHs can be produced quickly by *his1* knockout using genome editing techniques, including Target-AID. We also performed a sensitivity evaluation using commercially available herbicides containing either BBC or MST. Consequently, plant death was observed in all *HIS1* genome-edited lines used in the experiment (data not shown). In addition, phenotypes within the temperature-controlled greenhouse for base substitution lines used for evaluation of BBC and other β-triketone sensitivities were not abnormal compared with the wild-type (**Figure 7**). This desirable result suggests that the *his1* rice line is not drastically affected by knockout of *HIS1* (e.g., the pleiotropic effect of reduced expression). Conversely, the original role of his1 in rice plants has not been elucidated yet. In the future, a more robust assessment of

FIGURE 7 | T_3 plants of homozygous *his1* lines. Nipponbare (WT, *HIS1*), the *his1* homozygous mutant lines (T_3) 16-47-31a (1; start-codon mutation), 16-49-7a (2; stop-codon creation), 16-81-10b (3; stop-codon creation), and 16-82-9b (4; stop-codon creation) were grown in a controlled environment.

the agronomic impacts of pinpoint his1 knockouts versus WT lines in the same genome background under field conditions should be conducted. Such *his1* lines can quickly respond to the agricultural goal of suppressing volunteer rice field infestation. Given these results, a further evaluation of BBC sensitivity by cultivation in outdoor paddy fields should be conducted.

In this report, *his1* knockout lines were created using base-editing upon the BBC-resistant variety 'Nipponbare' with almost all lines acquiring BBC sensitivity. At present, we are attempting to produce additional *his1* knockout lines *via* Target-AID using a BBC-resistant practical variety specifically for livestock feed (data not shown). One of the strategic advantages of suppressing the occurrence of volunteer rice by using BBC and other bTHs is that most *japonica* rice varieties grown for human consumption are resistant to these herbicides. In addition, pinpoint gene editing will be advantageous because it can avoid the linkage of vicinity agronomically problematic traits (linkage drags) caused by conventional crossbreeding. In the future, the production of feed rice varieties with the control of volunteer rice infestation will lead to further cultivation and utilization of high-yield feed rice varieties.

CONCLUSION

To reduce the occurrence of volunteer rice field infestation, we created rice strains sensitive to the herbicide BBC *via* genome editing. A strategy for start-codon substitution and stop-codon creation by the targeted-AID system was used to create a *his1* knockout lines. As a result, *his1* knockout lines with nucleotide substitution showed sensitivity to BBC and other β-triketones. These strains are expected to contribute to the suppression of volunteer rice following spraying of BBC and other β-triketones.

AUTHOR'S NOTE

All authors declare that this study adheres to standard biosecurity and institutional safety procedures.

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

AK and KN designed the study, and KN and ZS designed gRNAs and constructed the Target-AID vector. AK and MO carried out the transgenic plant generation, DNA sequence analysis, and evaluation of the herbicide-sensitive strains. AK and KN wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by a grant from the Bio-oriented Technology Research Advancement Institution, NARO (Research Program on Development of Innovative Technology, 29009B).

ACKNOWLEDGMENTS

We thank Dr. H. Yoshida and Dr. S Hirose (NARO) for the advice about evaluation of the herbicide-sensitive strains, and Dr. M. Akasaka and Dr. H. Kato (NARO) for the advice about definition of volunteer rice or feed rice, and Dr. M. Endo and Dr. S. Toki (NARO) for kindly providing the gRNA-Cas9 plant expression vectors. Further we thank M. Hori, A. Yoshida, C. Isahaya and A. Ohuchi for general experimental technical support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00925/full#supplementary-material

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

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 (1), 14–17. doi: 10.1038/s41477-018-0321-8

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

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

Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L. A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* 31, 233–239. doi: 10.1038/nbt.2508

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

Kato, H. (2008). Development of Rice Varieties for Whole Crop Silage (WCS) in Japan.

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

Maeda, H., Murata, K., Sakuma, N., Takei, S., Yamazaki, A., Karim, M. R., et al. (2019). A rice gene that confers broad-spectrum resistance to β-triketone herbicides. *Science* 365, 393–396. doi: 10.1126/science.aax0379

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., et al. (2013). RNA-guided human genome engineering via Cas9. *Science* 339, 823–826. doi: 10.1126/science.1232033

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

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

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

Pattanayak, V., Lin, S., Guilinger, J. P., Ma, E., Doudna, J. A., and Liu, D. R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* 31, 839–843. doi: 10.1038/nbt.2673

Sakai, M., Iida, S., Maeda, H., Sunohara, Y., Nemoto, H., and Imbe, T. (2003). New rice varieties for whole crop silage use in Japan. *Breed. Sci.* 53, 271–275. doi: 10.1270/jsbbs.53.271

Sasaki, T. (2008). From the editor's desk. *Rice* 1, 1–2. doi: 10.1007/s12284-008-9010-y

Sekino, K., Koyanagi, H., Ikuta, E., and Yamada, Y. (2008). Herbicidal activity of a new paddy bleaching herbicide, benzobicyclon. *J. Pestic. Sci.* 33, 364–370. doi: 10.1584/jpestics.G08-11

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

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 deaminasefusion. *Nat. Biotechnol.* 35, 441–443. doi: 10.1038/nbt.3833

Shimatani, Z., Fujikura, U., Ishii, H., Matsui, Y., Suzuki, M., Ueke, Y., et al. (2018). Inheritance of co-edited genes by CRISPR-based targeted nucleotide substitutions in rice. *Plant Physiol. Biochem.* 131, 78–83. doi: 10.1016/j.plaphy.2018.04.028

Singh, V., Burgos, N. R., Singh, S., Gealy, D. R., Gbur, E. E., and Caicedo, A. L. (2017). Impact of volunteer rice infestation on yield and grain quality of rice. *Pest Manag. Sci.* 73 (3), 604–615. doi: 10.1002/ps.4343

Tang, X., Ren, Q., Yang, L., Bao, Y., Zhong, Z., He, Y., et al. (2019). Single transcript unit CRISPR 2.0 systems for robust Cas9 and Cas12a mediated plant genome editing. *Plant Biotechnol. J.* 7, 1431–1445. doi: 10.1111/pbi.13068

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

Toki, S. (1997). Rapid and efficient Agrobacterium-mediated transformation in rice. *Plant Mol. Biol. Rep.* 15, 16–21. doi: 10.1007/BF02772109

Zhong, Z., Sretenovic, S., Ren, Q., Yang, L., Bao, Y., Qi, C., et al. (2019). Improving plant genome editing with high-fidelity xCas9 and non-canonical 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 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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 22 October 2020
doi: 10.3389/fpls.2020.535764

Analyzing Twitter Conversation on Genome-Edited Foods and Their Labeling in Japan

Yutaka Tabei[1], Sachiko Shimura[1,2,3], Yeondae Kwon[2], Shizu Itaka[2,4] and Nobuko Fukino[1]*

[1] Strategic Planning Headquarters, National Agriculture and Food Research Organization, Tsukuba, Japan, [2] Research Center for Agricultural Information Technology, National Agriculture and Food Research Organization, Tokyo, Japan, [3] Institute of Agrobiological Sciences, National Agriculture and Food Research Organization, Tsukuba, Japan, [4] Faculty of Science and Technology, Tokyo University of Science, Noda, Japan

In recent years, the research and development of genome editing technology have been progressing rapidly, and the commercial use of genome-edited soybean started in the United States in 2019. A preceding study's results found that there is public concern with regard to the safety of high-tech foods, such as genetically modified foods and genome-edited foods. Twitter, one of the most popular social networks, allows users to post their opinions instantaneously, making it an extremely useful tool to collect what people are actually saying online in a timely manner. Therefore, it was used for collecting data on the users' concerns with and expectations of high-tech foods. This study collected and analyzed Twitter data on genome-edited foods and their labeling from May 25 to October 15 in 2019. Of 14,066 unique user IDs, 94.9% posted 5 or less tweets, whereas 64.8% tweeted only once, indicating that the majority of users who tweeted on this issue are not as intense, as they posted tweets consistently. After a process of refining, there were 28,722 tweets, of which 2,536 tweets (8.8%) were original, 326 (1.1%) were replies, and 25,860 (90%) were retweets. The numbers of tweets increased in response to government announcements and news content in the media. A total of six prominent peaks were detected during the investigation period, proving that Twitter could serve as a tool for monitoring degree of users' interests in real time. The co-occurrence network of original and reply tweets provided different words from various tweets that appeared with a certain frequency. However, the network derived from all tweets seemed to concentrate on words from specific tweets with negative overtones. As a result of sentiment analysis, 54.5% to 62.8% tweets were negative about genome-edited food and the labeling policy of the Consumer Affairs Agency, respectively, indicating a strong demand for mandatory labeling. These findings are expected to contribute to the communication strategy of genome-edited foods toward social implementation by government officers and science communicators.

Keywords: genome editing, SNS, co-occurrence network, sentiment analysis, public acceptance

OPEN ACCESS

Edited by:
Hiroshi Ezura,
University of Tsukuba, Japan

Reviewed by:
Katsumi Sakata,
Maebashi Institute of Technology, Japan
Ralf Alexander Wilhelm,
Julius Kühn-Institut, Germany

****Correspondence:***
Yutaka Tabei
tabei@affrc.go.jp

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *17 February 2020*
Accepted: *29 September 2020*
Published: *22 October 2020*

Citation:
Tabei Y, Shimura S, Kwon Y, Itaka S and Fukino N (2020) Analyzing Twitter Conversation on Genome-Edited Foods and Their Labeling in Japan. Front. Plant Sci. 11:535764. doi: 10.3389/fpls.2020.535764

INTRODUCTION

Agricultural production is facing new challenges due to the increasing world population, global climate change, and change in consumers' attitudes. To respond to these changes, new breeding technology, such as genome editing, is highly anticipated (Lusser et al., 2011; Yin et al., 2017; Ghogare et al., 2019). Genome editing is an innovative technology that may accelerate breeding by pinpointing and changing specific gene(s) and nucleotide(s) related to yield, biotic and abiotic stress tolerance, nutritional components, growth, and other factors (Dale et al., 2017; Abe et al., 2019; Gomez et al., 2019; Romero and Gatica-Arias, 2019). Genome editing has been used to modify many crops, fish, livestock, and other living organisms; the first commercialized genome-edited crop was the Calyno^{TM} high-oleic soybean[1], which has been cultivated in the United States since 2019. In Japan, the Cross-ministerial Strategic Innovation Promotion Program (SIP) of the Cabinet Office, which started in 2014, and other projects financed by the government have been the main supporters of the development of agricultural and animal products through genome editing technology. Examples include tomatoes rich in gamma-aminobutyric acid (GABA) (Nonaka et al., 2017), potatoes with significantly reduced natural toxins (solanine and chaconine) (Sawai et al., 2014), wheat with altered dormancy (Abe et al., 2019), red sea bream (Kishimoto et al., 2018) and Japanese pufferfish (Kuroyanagi et al., 2018), which have been modified to grow rapidly with more edible meat.

The social implementation of genome-edited products requires three main conditions: (i) the government's handling policy, (ii) intellectual property rights, and (iii) public acceptance (Tabei, 2019a). Japan has just established basic handling systems of genome-edited organisms and foods. This means that if the editing process involves only changes in the genetic code within the range of natural mutation, and no foreign DNA sequence exists in the edited organism's genome, the derived genome-edited food (hereinafter "genome-edited food") may be exempt from regulation as genetically modified (GM) foods under the Food Sanitation Law (Tabei, 2019b). Notification, not safety assessment, is required for genome-edited foods before commercialization. Intellectual property rights cannot be discussed here because it largely depends on individual conditions. Public acceptance is considered essential for its implementation (Araki and Ishii, 2015), and a major premise here is product development that benefits both producers and consumers. A preceding study's results found that there is some public concern with regard to the safety of high-tech foods produced using new breeding techniques (Malyska et al., 2016). For genome-edited crops, close communication of risks and benefits has been proposed for future social integration (Ishii and Araki, 2016). To achieve this, understanding the opinions, concerns, and expectations of consumers is considered important.

Generally, questionnaire surveys have been carried out to identify consumer interests and concerns. This method has the great advantage of garnering more information through answers to detailed questions, along with the information about the respondent's background. On the contrary, this method comes with some shortfalls. First, it is difficult to collect the ideal quantity of opinions in a timely manner and therefore analyze it in response to changing social situations. Second, it has been suggested that surveys do not capture the conversational or hierarchical nature of public opinion formation and that the operationalization of survey questions leads to a narrow definition of public opinion (McGregor, 2019). Recently, social media networks, such as Twitter, have demonstrated to be major drivers of news dissemination and public discourse. It provides a vast amount of semi-structured data in nearly real time and gives a direct access to contents of conversations (Müller et al., 2019). Whittingham et al. (2020) analyzed Twitter posts that discussed genetically modified organisms (GMOs) and found that personality (individual differences in one's tendency to show consistent patterns of thinking, emotion, and behavior) and values (learned beliefs about one's preferred way of action or existence) significantly affected GMO risk perception. Twitter discourse regarding clustered regularly interspaced short palindromic repeats (CRISPR), a genome editing technology, was recently investigated using semantic network analysis (Calabrese et al., 2019) and sentiment analysis (Müller et al., 2019). Therefore, we surmised that Twitter analysis could be applied to collect fresh voice on genome-edited foods because it would be possible to quickly collect information on the aspects that people are concerned about or interested in with regard to changes in the situation surrounding genome-edited foods.

2019 was a milestone year for the regulation of genome-edited products in Japan. The Ministry of Environment and five other ministries released policies for handling genome-edited organisms (Tabei, 2019b). As for labeling of genome-edited foods, on June 20, in a subcommittee meeting, the Cabinet Office's Consumer Committee implied that mandatory labeling for genome-edited products would be difficult (Cabinet Office's Consumer Committee, 2019). On September 19, the Ministry of Health, Labour and Welfare (MHLW) published a notification on genome-edited foods, announcing that it would be implemented from October 1 onward (Ministry of Health, Labour and Welfare, 2019). At the same time, the Consumer Affairs Agency (CAA) revealed their policy on voluntary labeling of genome-edited foods. There was also much discussion about the labeling of GM crops. Consumers demanded clear labeling to guarantee transparency, and in 2001, the Ministry of Agriculture, Forestry and Fisheries (MAFF) and the MHLW issued a labeling policy that lead to mandatory labeling of GM foods if the transgene can be detected (ISAAA, 2006). With regard to the commercialization of genome-edited foods, labeling is likely be a prevalent discussion. So, it seemed important to evaluate how the general public expressed their opinions and responded before and after the labeling policy was announced. Therefore, we decided to narrow our Twitter analysis to tweets related to the labeling of genome-edited foods. In this study, we collected Japanese Twitter data and investigated the relation between government announcements and news published by media. Furthermore, the appearance of words in tweets was investigated using co-occurrence networks, and emotions were determined using sentiment analysis. From these results, the

[1]https://calyxt.com/first-commercial-sale-of-calyxt-high-oleic-soybean-oil-on-the-u-s-market/

use of Twitter analysis for knowing users' opinions with regard to genome-edited foods and their labeling and its potential contribution to the communication strategy toward social implementation was discussed.

MATERIALS AND METHODS

Collecting Twitter Data

To explore public concern regarding genome-edited foods and their labeling, Twitter data were collected from a cloud software, Mieruka Engine® (Plus Alpha Consulting Co., Ltd., Tokyo, Japan), from May 25 to October 15, 2019, with the search string "genome editing (which included 'genome-edited' in Japanese) AND labeling" in Japanese. We performed Mieruka Engine® searches using API provided by the NTT DATA Corporation (Tokyo, Japan), which has the resale rights of Twitter data in Japan. The use of materials for data analysis is permitted by the Copyright Act of Japan.

Refining Twitter Data

There are three types of tweets: original tweets, reply tweets, and retweets. In our study, retweets were regarded as partially or fully agreeing with the opinions of the original tweeters. Although tweet redundancy increases the chance of opinions influencing many people, we excluded duplicates for an accurate analysis. The collected tweet data were refined according to the following rules: (i) if a Twitter user tweets or retweets the same text multiple times, whether intentional or not, only one of them is analyzed; (ii) if the duplicate tweets were from different people, those texts are all subject to analysis – this happens frequently in retweets; and (iii) if a tag or link is different but the text is the same, it is considered the same tweet.

User Profile Data

Twitter users registered their profiles when creating their accounts; some left them blank. For users without profiles, Mieruka Engine® predicted their profiles when possible. Unique user IDs posted from May 25 to October 15 were collected and counted. Because of system rules, user profiles of tweets from June 15 to October 15 were available in Mieruka Engine®. To verify whether the same users tweeted repeatedly on this issue, unique user IDs were categorized based on tweet counts (including original, reply tweets, and retweets) posted during this period.

Tweet Peak Detection

During data collection, the tweet count (i.e., original, replies, and retweets) increased sharply on several occasions. Prominent peaks were detected with Python (version 3.7.6), using SciPy 1.4.1. The prominence of a peak allows measurements of the degree to which a peak is protruding, depending on its position in relation to other peaks. A prominence cut-off of 300 was set to detect important peaks evenly over the entire period of the study. When peaks were detected, influential events were determined from the content of each tweet.

Analysis of Appearance Pattern of Words in Tweets by Co-occurrence Network

On September 19, the MHLW announced that the notification of genome-edited foods would start on October 1, and the CAA declared that labeling of genome-edited foods should not be mandated. To examine the overall picture of how users discussed genome-edited foods and their labeling in response to government's announcements, we created co-occurrence networks based on tweets posted from September 19 to 22. At first, we analyzed 530 original and reply tweets excluding retweets using the free text-mining software KH Coder (Higuchi, 2016). Then we analyzed all 5,410 tweets by the same method described above. Before analysis, we modified the text data of the tweets as follows: (i) we converted half-width Japanese characters to full-width Japanese letters, full-width numbers and English letters to half-width numbers and English letters, respectively; (ii) we excluded URLs and the string of ASCII characters that are considered punctuation characters; and (iii) we normalized Unicode strings. We also excluded some words as stop words (**Supplementary Table 1**). We determined the degree of association between words using the Jaccard coefficient (Romesburg, 1992).

Sentiment Classification of User's Opinion on Genome-Edited Foods and Labeling

To infer users' opinions on these policies, original and reply tweets posted during the same period in the former section were selected for classification according to sentiment on genome-edited food and its labeling. Oftentimes, the sentiments expressed for "genome-edited food" and "labeling" in tweets differed; therefore, they were counted independently, and text and information provided by the URL link were used to determine the sentiments. Three researchers differentiated Twitter sentiment into three groups: positive, negative, and neutral. Tweet classification was performed using the criteria in **Table 1**. In addition, the sentiment of each tweet in the same data set was also determined using the "positive–negative analysis" function of the Mieruka Engine® software.

RESULTS

Tweet Data and User Profiles

From the 29,299 tweets that were extracted with the search string "genome editing (which included 'genome-edited' in Japanese) AND labeling" in Japanese, 577 were excluded in accordance with the policies laid out in the "Materials and Methods" section. Thus, the dataset consisted of 28,722 tweets of which 2,536 tweets (8.8%) were original tweets, 326 tweets (1.1%) were replies, and 25,860 tweets (90%) were retweets.

To learn about the kind of people who were interested in the labeling of genome-edited foods, we categorized the age and gender of unique user IDs (**Table 2**). Of 14,066 unique user IDs, gender and age profiles were available for 12,016. While user profile accuracy is not guaranteed, 8,817 (73.4%) were male, and

TABLE 1 | Tweet classification criteria for sentiment analysis.

Sentiment	Criterion
Positive	This category includes tweets that accept genome-edited food or government policies, explain the technology scientifically, or elucidate the reason for non-mandatory labeling.
Negative	This category consists of tweets against genome-edited foods and MHLW/CAA policies or those that call for a signature petition of severe regulation. Many of them include the following terms: "scary," "don't want to be distributed," "right not to eat," "don't want to buy," "don't want to eat," "dangerous," etc.
Neutral	This category includes tweets that are neutral on genome-edited food/labeling or are just publicizing government policies, etc.

TABLE 2 | Unique user IDs categorized by age and gender.

Age	Gender				Total	
	Male		Female			
10s	168	(1.9)	110	(3.4)	278	(2.3)
20s	1,270	(14.4)	632	(19.8)	1,902	(15.8)
30s	1,748	(19.8)	1,254	(39.2)	3,002	(25.0)
40s	834	(9.5)	524	(16.4)	1,358	(11.3)
50s and over	4,797	(54.4)	679	(21.2)	5,476	(45.6)
Total	8,817	(100.0)	3,199	(100.0)	12,016	(100.0)

The numbers on the left represent the number of unique user IDs of all tweets posted from June 15 to October 15. The numbers in parentheses indicate percentage.

3,199 (26.6%) were female. As for age distribution, the ratio of users in their 10s and 40s were low, accounting for only 2.3% and 11.3%, respectively, while, users in their 30s and 50s and above were relatively high, at 25% and 45.6%, respectively. In particular, users in their 50s and older were the largest group among male users while users in their 30s were the largest group among female users (**Table 2**).

To verify that the same users tweeted repeatedly in each peak, unique user IDs were categorized based on tweet counts (**Supplementary Figure 1**). The study found that 94.9% of user IDs posted five or less tweets, in particular 64.8% and 17.3% posted tweets only once and twice, respectively, whereas 94.9% posted five or less tweets. Two user IDs were noted to post more than 100 tweets, and the maximum tweet count per user ID was 201.

Changes in Tweet Count and Influential Events

We investigated the time course of daily tweet numbers to reveal the responses to information related to the labeling of genome-edited foods. **Figure 1** shows the change in the daily number of tweets from May 25 to October 15. The tweet count ranged from 1 to 3,426 a day. During this period, we also identified six specific peaks. **Table 3** summarizes the date and tweet count in the peaks as well as influential events (e.g., government announcements, media reports, and so on).

When it was expressed in a subcommittee meeting of the Cabinet Office's Consumer Committee that mandatory labeling for genome-edited foods would be difficult, many users posted tweets in response to news regarding it. This is a key reason for the significant increase in tweet count compared to the period before June 21 and thus formed the peak in **Figure 1**(a). The MHLW then began collecting public comments on the procedure for submitting information on genome-edited foods (Ministry of Health, Labour and Welfare, 2019). Many media networks covered the direction of notification, which was a trigger to form the peak in **Figure 1**(b).

The third peak, **Figure 1**(c), several tweets mentioned a tweet by an opposition politician insisting mandatory labeling and tweets introducing newspaper or website articles, and most tweets were about campaigns to collect signatures demanding the mandatory labeling of genome-edited foods.

The fourth peak, **Figure 1**(d), shows that the tweet count rose immediately after the government announced its handling policies on genome-edited foods on September 19. The MHLW announced that starting October 1, it would be requiring notification, not safety assessment, for genome-edited foods before commercialization. The CAA, meanwhile, announced that the labeling policy for genome-edited foods was non-obligatory but recommended that developers provide as much information to consumers as possible. News about the government's policies were publicized by many media outlets over 2 days from September 19 to 20. The increase in tweet count on September 20 appears to be because of widespread tweeting. Many tweets tried to publicize the fact that genome-edited food labeling is not mandatory, of which a few were positive opinions. Examples include "It is reasonable that labeling of genome-edited food is voluntary, as it is scientifically indistinguishable from existing food," and "It's not good to just look at the word genome editing and post negative messages." However, there were also many negative opinions expressing concern that one could unwittingly buy genome-edited foods.

The fifth peak, **Figure 1**(e), was mainly caused by much news that introduced genome editing technology and reviewed government's policies, which were widely publicized in TV programs, newspapers, and so on. These included a special TV program aired by the Japan Broadcasting Corporation (NHK), editorial articles, and others. There were also many tweets about petitions to collect signatures for mandatory labeling. The tweet count on September 25 was 3,426, the highest number during the analysis period.

The cause of the sharp increase in tweet count on October 7, the peak in **Figure 1**(f), was unclear. No major event seemed to have occurred, but there were various kinds of tweets regarding news about labeling, signature activities that oppose to non-obligatory labeling, and so on. Moreover, a newsletter article posted on a bulletin board system (BBS), the

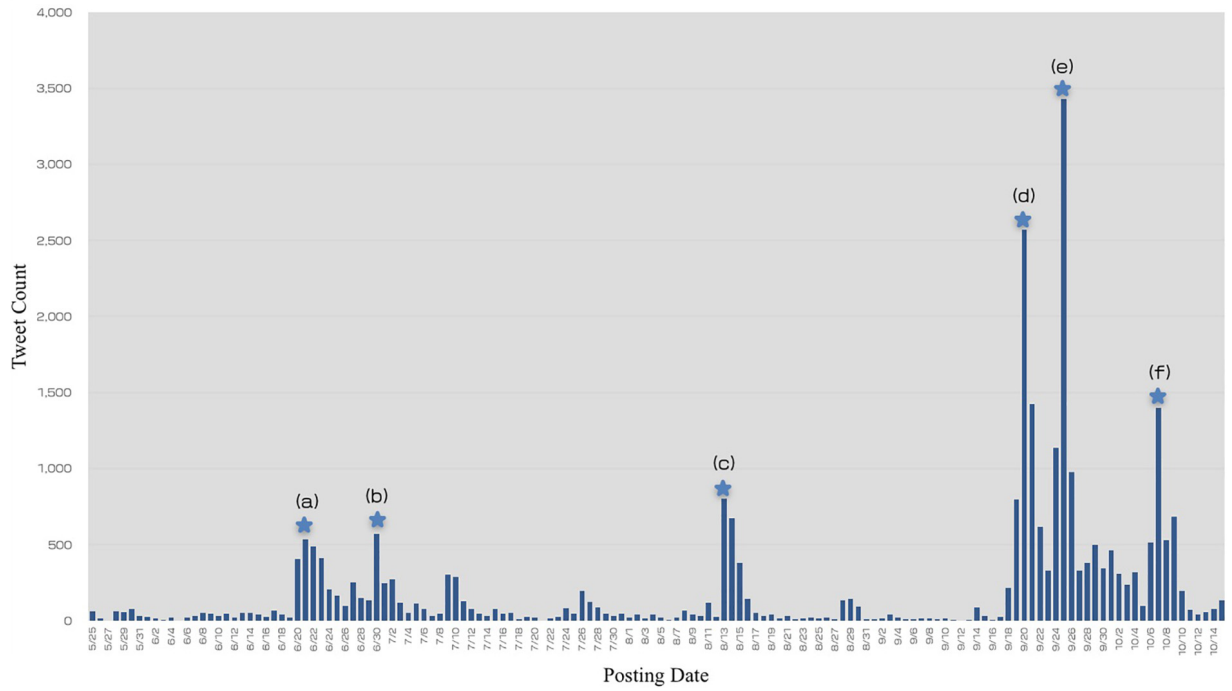


FIGURE 1 | Changes in the daily number of tweets. This figure shows the number of tweets in chronological order from May 15 to October 15. The bars represent the total tweet count including original tweets, reply tweets, and retweets. Prominence of peaks (a–f) were identified with Python (version 3.7.6) using the SciPy 1.4.1.

TABLE 3 | Changes in tweet counts in response to influential events.

Peak[a]	Date	Influential events[b]	No. of all tweets	No. of original and reply	No. of retweets	Prominence[c]
(a)	6/21	A subcommittee meeting of the Cabinet Office's Consumer Committee stated that it is difficult to make the labeling of genome-edited foods mandatory on June 20.	535	54	481	436
(b)	6/30	The Ministry of Health and Welfare (MHLW) began to collect public comments on genome-edited foods from June 27.	569	13	556	566
(c)	8/13	An opposition politician posted a tweet on August 13 insisting mandatory labeling of genome-edited foods.	802	13	789	799
(d)	9/20	The MHLW and the Consumer Affairs Agency announced their policies for handling and labeling genome-edited foods on September 19.	2,569	275	2,294	2,239
(e)	9/25	Many news introducing the basis of genome editing technology and reviewing government's policies were widely publicized in TV programs, newspapers, and so on.	3,426	142	3,284	3,387
(f)	10/7	There seemed to be no major events, but there were various kinds of tweets on news, signature activities, and so on.	1,397	129	1,268	1,297

[a] *The first column represents the peaks in* ***Figure 1****.* [b] *Influential events include government announcements, mass media, and tweets by influential public figures.* [c] *The prominence of each peak is calculated using SciPy 1.4.1, which is module of Python.*

largest in Japan, was tweeted and retweeted. This BBS often discusses opposition to social trends and events that have not spread to society.

Words in Tweets by Co-occurrence Network

Using co-occurrence networks, we analyzed the appearance and relevance of words in original and reply tweets, and all tweets including retweets posted from September 19 to 22 and compared patterns of the two networks.

In the co-occurrence network of original and reply tweets without retweets, seven clusters were observed (subgraphs 1–7; **Figure 2A**), which had the following characteristics: Subgraph 1 consisted of words mainly related to *genome-edited*, *food*, *mandatory labeling*, and *consumer*, which were discussed in the context of the obligatory labeling of genome-edited foods. Furthermore, the word *the Nikkei* (referring to

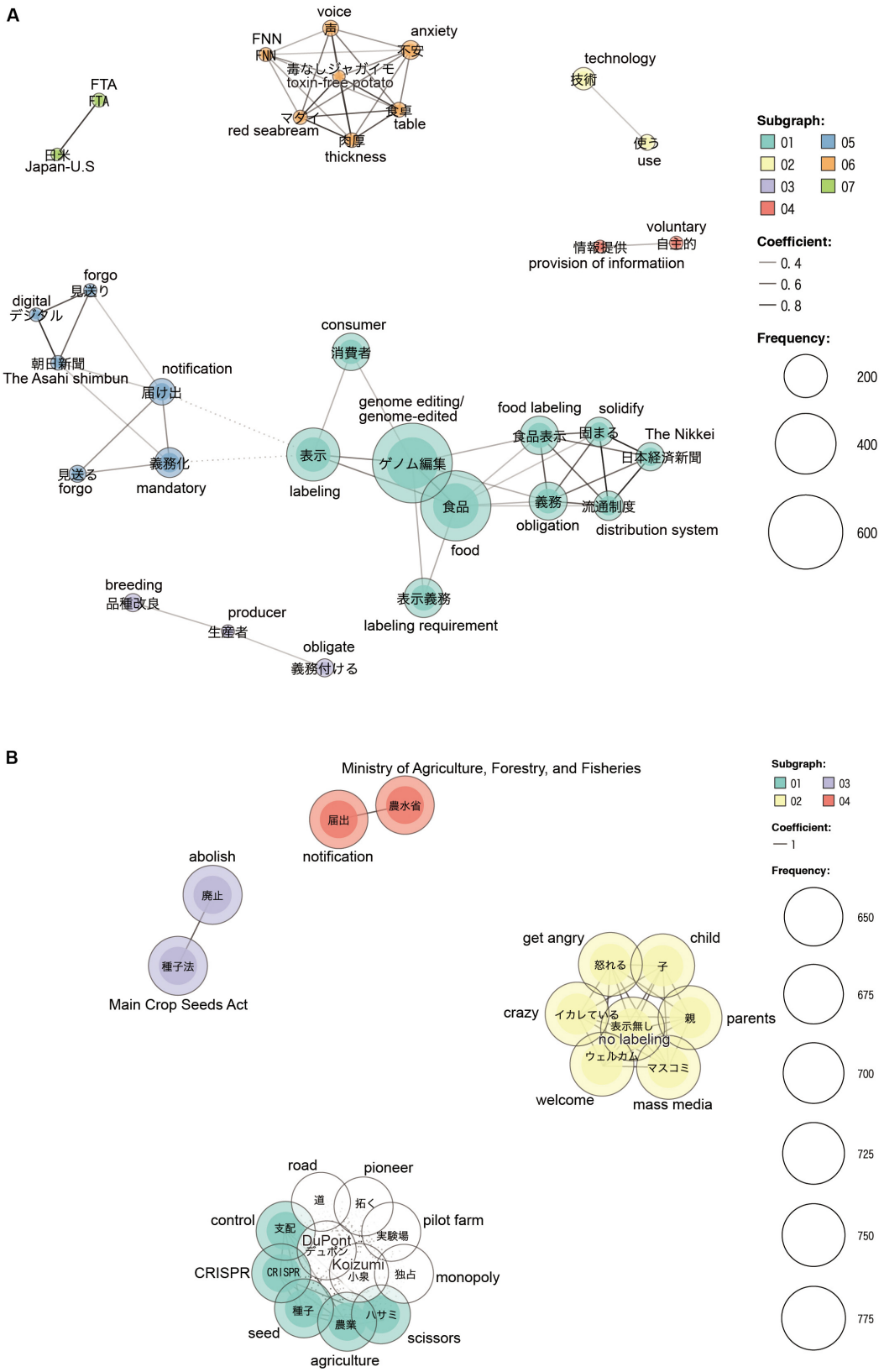


FIGURE 2 | Co-occurrence networks of words in tweets posted from September 19 to 22. **(A)** Co-occurrence network with only original and reply tweets, **(B)** co-occurrence network with all tweets including original tweets, reply tweets, and retweets.

the *Nikkei* newspaper) was strongly related to the words *food labeling, obligation,* and *distribution system*; it was speculated that the content published in the *Nikkei* uses the above words. Subgraph 2 contained two words: *technology* and *use*. Subgraph 3 consisted of three words – *breeding*, *producer*, and *obligate* – which points toward people asking producers for mandatory labeling. Subgraph 4 consisted of two words: *voluntary* and *provision of information*. It shows that many tweeters were concerned about the provision of information (in the notification to the MHLW) being non-mandatory. Subgraph 5 consisted of tweets regarding news of government policies and the commercialization of genome-edited foods, and an article from *Asahi Shimbun* (Asahi newspaper) influenced this subgraph. Subgraph 6 was influenced by tweets regarding studies on potatoes without harmful substances and fast-growing red sea breams, as well as tweets about the resulting public anxiety. Subgraph 7 contained two words: *Japan and United States* and *FTA*. Tweeters were concerned that the United States–Japan FTA might result in the import of undesirable genome-edited foods because of pressure from the United States.

However, in the co-occurrence network of all tweets, two large clusters consisting of 12 and 7 words and two small clusters containing two words were formed. Concerns about food supply and Japanese agriculture formed a cluster (subgraph 1, **Figure 2B**); which included words such as *seed*, *agriculture*, *control*, and *CRISPR*. In addition, these words were also co-occurring with words in subgraph 1, such as *DuPont*, *monopoly*, *pilot farm*, and *Koizumi*, a member of the Japanese Diet, i.e., Japanese parliament, regarded as a key player in abolishing the Major Crop Seeds Act. Representative contentions of the tweets related to this cluster were "CRISPR-Cas9 is a carcinogenic enzyme," "the abolition of Major Crop Seeds Act would cause the control of Japanese agriculture by foreign-affiliated companies," "promotion of local production for local consumption," and so on. These words were associated with the concern that foreign companies would control or monopolize crop seeds. Another cluster (subgraph 2, **Figure 2B**) was formed by the users' anger toward the CAA's policy of not imposing mandatory labeling, and concerns about children's health and food safety. The terms in this cluster were *mass media*, *get angry*, *no labeling*, *parents*, *child*, *crazy*, and *welcome*. The word *welcome* was used in the context of criticizing news which welcomed genome editing technology, and does not pertain to agreement with the CAA policy. Other small clusters contained such terms as *Main Crop Seeds Act*, *abolish*, *the Ministry of Agriculture*, *Forestry and Fisheries*, *notification,* and so on. Tweets with such words seemed opposed to genome editing or to the government's agricultural policies.

The co-occurrence network of original and reply tweets provided different words from various tweets that appeared with a certain frequency, and many of them were derived from news contents. However, the network derived from all tweets seemed to concentrate on words from specific tweets with negative overtones, and the variety of words that appeared in the network was biased, resulting in a simple network diagram.

Sentiment Analysis of Tweets on Genome-Edited Food and Labeling

Five hundred and thirty original and reply tweets posted during 4 days after the government announcements from September 19 to 22 were classified as positive, negative, and neutral based on sentiments on genome-edited food and labeling (**Table 4**). Three researchers performed the classification. For some tweets, Twitter's character limit (140 Japanese characters) and chatty and colloquial style of tweets complicated sentiment judgment. In instances where the classification of a given tweet among the three researchers was inconsistent, the sentiment agreed by two of them was adopted. When there were three different judgments for a tweet, it was classified as neutral. There were 190 (35.8%) and 122 (23.0%) tweets indicating inconsistent judgment among three researchers for genome-edited foods and labeling, respectively.

There were tweets in which users tried to convey scientific knowledge to other users, such as explaining the difference between genome editing and genetic modification, and such tweets were classified as "positive." Tweets that include phrases such as "want to eat" and "want to buy" were also under "positive." Under the "negative" classification were tweets against MHLW/CAA policies, those that call for a signature petition for stricter regulations, and so on. Tweets that mention articles that clearly criticize the government's policies were also classified as "negative," Most "neutral" tweets were those that simply publicized news by mass media networks.

Of the 530 original and reply tweets posted during the period, 289 (54.5%) opposed genome-edited food, and 37 (7.0%) were in favor of it. With regard to the CAA policy, 333 (62.8%) tweets were negative and 27 (5.1%) were positive (**Table 4**). Among the 37 positive tweets about genome-edited food, 24 were in favor of the CAA policy and 8 were against it, for example, "It is better that it is labeled, because I want to buy and eat it." 204 tweets (38.5%) were neutral toward genome-edited food, but 39 opposed the non-mandatory labeling; for instance, "Aside from the discussion on whether genome-edited foods are safe, its labeling is necessary for consumers to have a choice."

For comparison, the sentiment for each tweet was determined using the Mieruka Engine® software, which did not consider the respective sentiments for genome-edited food and labeling policy in a tweet. Of the same 530 tweets mentioned above, 79 tweets (14.9%) were determined to be positive, while 97 (18.3%) were determined to be negative.

DISCUSSION

Genome editing technology is expected to rise in fields such as medicine and agriculture. However, science and technology are not autonomous entities, and research trajectories are largely influenced by public opinion; even if crops produced using new breeding techniques do not fall under GMOs, commercializing them is by no means easy (Ishii and Araki, 2016; Malyska et al., 2016). The public tends to have a vague anxiety about advanced technology that they are not familiar with. Therefore, public acceptance of innovative technology, such as genetic modification or genome editing, is an important requirement for its social

TABLE 4 | Types of sentiments on genome-edited food and its labeling expressed in each tweet posted from September 19 to 22, 2019[a].

		Labeling			
		Positive	**Negative**	**Neutral**	**Total[b]**
Genome-edited food	Positive	24	8	5	37 (7.0)
	Negative	0	286	3	289 (54.5)
	Neutral	3	39	162	204 (38.5)
	Total[b]	27 (5.1)	333 (62.8)	170 (32.1)	530 (100.0)

[a] *Tweet classification was performed by three researchers using the criteria in* ***Table 1***. [b] *The numbers in parentheses indicate percentage.*

implementation. In the 1990s, when GM foods began to be distributed in Japan, there was much debate about labeling to ensure consumers' right to select non-GM foods. As a result, the MAFF and MHLW started the labeling system for GM food in 2001 (ISAAA, 2006; Umeda, 2014). It was assumed that labeling would be a major issue for the social implementation of genome-edited foods; therefore, Twitter analysis relevant to genome-edited foods and their labeling was started.

In this survey, among user IDs whose profiles were available, 71.5% were male, and male users in their 50s and older outnumbered male users in other age groups (**Table 2**). Females in their 30s were the most among all age groups of female users. Generally, Twitter users are relatively evenly distributed across all ages although the number of males in their 40s and females in their 20s tend to be slightly more than other ages[2]. The reason for the deviation in our study was not clear, but users with these profiles were most interested in genome-edited foods and their labeling. A total of 94.9% of user IDs posted five or less tweets, in particular 64.8% tweeted only once. The majority of the users who tweeted about this issue do not seem as intense as they posted tweets consistently.

The study retrieved a total of 29,299 tweets from May 25 to October 15 in 2019 (142 days) using the search string "genome editing (which included 'genome-edited' in Japanese) AND labeling" in Japanese with a maximum tweet number per day of 3,426. To compare the number of tweets on this issue and another much-discussed political issue, tweets on a political issue (i.e., the Tokyo gubernatorial election) were retrieved using the Mieruka Engine® from June 29 to July 29, 2020. The maximum tweet count per day reached 1,343,045 tweets on the election day: June 5. In terms of the Fukushima Daiichi Nuclear Power Plant accident, which was directly related to life-threatening and energy problems, approximately 25 million tweets were posted within a span of 195 days, and the maximum tweet count per day was more than 640,000 (Tsubokura et al., 2018). Comparing these politically and socially hot topics, the issue on labeling of genome-edited food seemed to draw less attention on Twitter. Typically, surveys on public opinion are conducted and targeted several thousands of samples, which are nationally representative. In the present study, online discourse and extent of users' interest on the labeling of genome-edited food were successfully obtained, although bias in Twitter users or limited number of tweets posted on this issue should be considered. To collect opinions from a diverse or targeted range of people, questionnaire surveys should be used together.

To ascertain users' opinions with regard to the labeling of genome-edited foods, Twitter data were collected from May 25 to October 15 using the search string "genome editing (which included 'genome-edited' in Japanese) AND labeling." Throughout this period, the tweet count changed significantly in response to government announcements and news published by the media and other information sources. A spike in the tweet count may be an indication of growing concern among Twitter users (**Figure 1** and **Table 3**). An analysis of the sentiments accompanying original and reply tweets revealed that 289 tweets (54.5%) displayed negative sentiments toward genome-edited foods. Uchiyama et al. (2019) reported the results of an attitude survey on the internet that showed that 43% and 47% of Japanese citizens answered that they did not want to eat genome-edited crops and livestock products, respectively, similar to the results of sentiment analysis in present study. To reveal the reason for this evasion, a more comprehensive analysis of opinions regarding genome-edited foods and the application of genome editing technology to food would be necessary while using a data set not limited to labeling. With regard to the non-mandatory labeling policy of the CAA, 333 tweets (62.8%) opposed it. Sentiments for genome-edited foods and their labeling policy displayed different tendencies. All negative tweets about genome-edited foods were negative or neutral about the labeling policy. However, among tweets that were positive about genome-edited foods, some tweets were negative about the labeling policy. Therefore, the percentage of negative tweets about labeling policy was higher than that for genome-edited foods itself, indicating a strong demand for mandatory labeling, as was the case of GM crops of which labeling was requested by consumers because they had doubts about its safety and wanted to avoid GM foods (Umeda, 2014). These findings highlighted the need to inform people regarding the rationale behind the CAA policy of non-mandatory labeling. A technical reason exists for the non-mandatory labeling of genome-edited foods, because distinguishing between genome-edited foods and those developed using conventional breeding is difficult.

Sometimes, the Twitter character limit (140 characters) and the chatty, colloquial style of tweets complicated sentiment judgment, resulting in approximately one-quarter to one third inconsistent tweets among the three researchers. Sentiments determined by the Mieruka Engine® software tended to be positive more than those determined by researchers in this study.

[2]https://gaiax-socialmedialab.jp/post-30833/

Sentiment analysis of the software was conducted by determining the presence of words in positive and negative categories, and not by the context. This might be a reason for the difference between the two methods. For high-volume and automated processing, the deep learning of text data could potentially allow for the identification of complex sentence structures, taking into account multiple word combinations, and lead to improvement in the accuracy of sentiment analysis.

Retweets accounted for 90% of total tweets in our study; co-occurrence networks were significantly different with or without retweets because of an increase of the proportion of specific words in retweets (**Figure 2**). There are several finding on retweets. Boyd et al. (2010) found that people prefer retweets that contain breaking and timely news. Naveed et al. (2011) found that tweets that address public events or include emoticons that reveal more negative emotions than positive ones are likely to be retweeted; Negative tweets are generally more often retweeted than positive ones regardless of their subject matter (Stieglitz and Dang-Xuan, 2012; Tsugawa and Ohsaki, 2015). These findings are consistent with our results on co-occurrence networks that words from tweets with a negative tone increased when adding retweets to the analysis. Such retweets, which were made up majorly of tweets obtained in our study, might be a product of the increasing disaffection and anxieties of users about the labeling policy. The negative opinions of users, revealed by a co-occurrence network and sentiment analysis, seem to imply the direction of communication strategy of genome-edited food.

The acceptance of GM crops increased among consumers who understood the benefits of GM crops, such as reduced CO_2 emissions (Council for Biotechnology Information Japan, 2017). For some genome-edited crops that are currently being developed in Japan, it is expected that consumers easily understand their benefits; some examples of such crops are high-GABA tomatoes (Nonaka et al., 2017) and potatoes without harmful substances (Sawai et al., 2014). It might be easier to achieve public acceptance and to realize social implementation for such crops. This study indicated the potential of Twitter as a real-time indicator of users' concern. When releasing genome-edited crops into the Japanese market, it is likely that there will be discussions about them pros and cons, and it would be significant to continue analyzing opinions for them.

CONCLUSION

Twitter analysis allowed us to quickly access online discourse in response to government announcements and media reports. Through this study, we were able to identify the information that Twitter users were interested in and also their apprehensions with regard to genome-edited foods and their labeling. These findings may contribute to the communication strategy of genome-edited foods with regard to its social implementation, for example, in case erroneous information spreads, corresponding scientific facts could be provided. However, there are certain disadvantages of Twitter analysis such as the limit on information that can be derived from each tweet due to the character limit (140 Japanese characters), colloquial style, and bias among Twitter users. To achieve a more detailed and accurate analysis, using a combination of a Twitter analysis and other methods, such as questionnaire surveys, would help cover for the deficits of the Twitter analysis. Furthermore, it will be meaningful to analyze consumers' opinions and attitudes at the first commercialization of genome-edited food employing data mining by deep learning, considering many complex factors, such as social situation surrounding genome-edited foods, news content, online discourse or questionnaire survey, and individual profile, among others, in the future.

DATA AVAILABILITY STATEMENT

All dataset except raw Twitter data generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

YT provided the initial research design, discussed the results, and wrote, reviewed, and edited the manuscript. SS wrote the manuscript, performed research design and data analysis, and discussed the results. YK also wrote the manuscript and worked on research design, data collection and analysis, and the discussion. SI discussed the results as well. NF worked on data analysis and discussed the results as well as reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Cross-ministerial Strategic Innovation Promotion Program (SIP), Cabinet Office, Government of Japan, under "Technologies for Smart Bio-industry and Agriculture" (funding agency: Bio-oriented Technology Research Advancement Institution, NARO), and a subsidy from NARO.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Masato Ohta and Mr. Makoto Kasai, who provided valuable opinions and helped us in our data analysis, and Dr. Masahiro Yano, who promoted the research group. The authors would also like to thank Enago (www.enago.jp) for the English language review.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.535764/full#supplementary-material

REFERENCES

Abe, F., Haque, E., Hisano, H., Tanaka, T., Kamiya, Y., Mikami, M., et al. (2019). Genome-edited triple-recessive mutation alters seed dormancy in wheat. *Cell. Rep.* 28, 1362.e4–1369.e4.

Araki, M., and Ishii, T. (2015). Towards social acceptance of plant breeding by genome editing. *Trends Plant Sci.* 20, 145–149. doi: 10.1016/j.tplants.2015.01.010

Boyd, D., Golder, S., and Lotan, G. (2010). "Tweet, tweet, retweet: conversational aspects of retweeting on Twitter," in *Proceedings of the 2010 43rd Hawaii International Conference on System Sciences (HICSS '10)*, New York, NY.

Cabinet Office's Consumer Committee (2019). Available online at: https://www.cao.go.jp/consumer/history/05/kabusoshiki/syokuhinhyouji/bukai/055/shiryou/index.html (accessed February 18, 2020).

Calabrese, C., Ding, J., Millam, B., and Barnett, G. A. (2019). The uproar over gene-edited babies: a semantic network analysis of CRISPR on twitter. *Environ. Commun.* 14, 954–970. doi: 10.1080/17524032.2019.1699135

Council for Biotechnology Information Japan (2017). *A Survey on Consumer Attitude on Genetically Modified Foods*. Available online at: https://cbijapan.com/document/1671/ (accessed February 18, 2020).

Dale, J., James, A., Paul, J. Y., Khanna, H., Smith, M., Peraza-Echeverria, S., et al. (2017). Transgenic cavendish bananas with resistance to fusarium wilt tropical race 4. *Nat. Commun.* 8:1496.

Ghogare, R., Williamson-Benavides, B., Ramírez-Torres, F., and Dhingra, A. (2019). CRISPR-associated nucleases: the dawn of a new age of efficient crop improvement. *Transgen. Res.* 29, 1–35. doi: 10.1007/s11248-019-00181-y

Gomez, M. A., Lin, Z. D., Moll, T., Chauhan, R. D., Hayden, L., Renninger, K., et al. (2019). Simultaneous CRISPR/Cas9-mediated editing of cassava *eIF4E* isoforms *nCBP-1* and *nCBP-2* reduces cassava brown streak disease symptom severity and incidence. *Plant Biotechnol. J.* 17, 421–434. doi: 10.1111/pbi.12987

Higuchi, K. (2016). A two-step approach to quantitative content analysis: KH coder tutorial using Anne of Green Gables (Part I), Ritsumeikan. *Soc. Sci. Rev.* 52, 77–91.

ISAAA (2006). *Labeling GM foods*. Available online at: http://www.isaaa.org/resources/publications/pocketk/7/default.asp (accessed February 18, 2020).

Ishii, T., and Araki, M. (2016). Consumer acceptance of food crops developed by genome editing. *Plant Cell Rep.* 35, 1507–1518. doi: 10.1007/s00299-016-1974-2

Kishimoto, K., Washio, Y., Yoshiura, Y., Toyoda, A., Ueno, T., Fukuyama, H., et al. (2018). Production of a breed of red sea bream *Pagrus major* with an increase of skeletal muscle mass and reduced body length by genome editing with CRISPR/Cas9. *Aquacult. Asia* 495, 415–427. doi: 10.1016/j.aquaculture.2018.05.055

Kuroyanagi, M., Kinoshita, M., and Yoshiura, Y. (2018). Establishment of breeding technology using TILLING method in cultured fish: development of fugu (*Takifugu rubripes*) with a double-muscling phenotype (in Japanese). *Kagaku Seibutsu* 56, 523–525.

Lusser, M., Parisi, C., Rodriguez-Cerezo, E., and Plan, D. (2011). *New Plant Breeding Techniques State-of-the-Art and Prospects for Commercial Development. JRC Technical Report EUR 24760 EN*. Rome: European Commission Joint Research Centre.

Malyska, A., Bolla, R., and Twardowski, T. (2016). The role of public opinion in shaping trajectories of agricultural biotechnology. *Trends Biotechnol.* 34, 530–534. doi: 10.1016/j.tibtech.2016.03.005

McGregor, S. C. (2019). Social media as public opinion: how journalists use social media to represent public opinion. *Journalism* 20, 1070–1086. doi: 10.1177/1464884919845458

Ministry of Health, Labour, and Welfare (2019). *Public Comments for "Food Hygiene Handling Procedures for Food and Additives Derived From Genome Editing Technology (Draft)"*. Tokyo: Ministry of Health, Labour, and Welfare.

Müller, M., Schneider, M., Salathé, M., and Vayena, E. (2019). Combining crowdsourcing and deep learning to assess public opinion on CRISPR-Cas9. *bioRxiv* doi: 10.1101/802454

Naveed, N., Gottron, T., Kunegis, J., and Alhadi, A. C. (2011). "Bad news travel fast: a content-based analysis of interestingness on twitter," in *Proceedings of the 3rd International Web Science Conference*, (New York, NY: Association for Computing Machinery), 1–7.

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

Romero, F. M., and Gatica-Arias, A. (2019). CRISPR/Cas9: development and application in rice breeding. *Rice Sci.* 26, 265–281. doi: 10.1016/j.rsci.2019.08.001

Romesburg, H. C. (1992). *Cluster Analysis for Researchers*. trans, H. Nishida and T. Sato (Tokyo: Uchidarokakuho), 424.

Sawai, S., Ohyama, K., Yasumoto, S., Seki, H., Sakuma, T., Yamamoto, T., et al. (2014). Sterol side chain reductase 2 is a key enzyme in the biosynthesis of cholesterol, the common precursor of toxic steroidal glycoalkaloids in potato. *Plant Cell* 26, 3763–3774. doi: 10.1105/tpc.114.130096

Stieglitz, S., and Dang-Xuan, L. (2012). "Political communication and influence through microblogging – an empirical analysis of sentiment in twitter messages and retweet behavior," in *Proceedings of the 45th Hawaii International Conference on System Sciences*, (Maui, HI).

Tabei, Y. (2019a). "Japanese regulatory frameworks, regulations on gene editing technology and the considerations during the regulatory policy development," in *Proceedings of the International Symposium on the Development Trends of Emerging Precision Breeding*, Taiwan.

Tabei, Y. (2019b). "The regulatory frameworks of genome editing organisms and foods in Japan," in *Proceedings of the KSBS and SABRAO 14th International Conference on Plant Breeding for Sustainable Development*, Gwangju.

Tsubokura, M., Onoue, Y., Torii, H. A., Suda, S., Mori, K., Nishikawa, Y., et al. (2018). Twitter use in scientific communication revealed by visualization of information spreading by influencers within half a year after the Fukushima Daiichi nuclear power plant accident. *PLoS One* 13:e0203594. doi: 10.1371/journal.pone.0203594

Tsugawa, S., and Ohsaki, H. (2015). "Negative messages spread rapidly and widely on social media," in *Proceedings of the 3rd ACM Conference on Online Social Networks*, (New York, NY: ACM), 151–160.

Uchiyama, M., Nagai, A., and Muto, K. (2019). Survey on how the public think about genome editing in foods. *J. Hum. Genet.* 63, 745–748.

Umeda, S. (2014). *Restrictions on Genetically Modified Organisms: Japan*. Available online at: https://www.loc.gov/law/help/restrictions-on-gmos/japan.php#_ftn42 (accessed February 18, 2020).

Whittingham, N., Boecker, A., and Grygorczyk, A. (2020). Personality traits, basic individual values and GMO risk perception of twitter users. *J. Risk Res.* 23, 522–540. doi: 10.1080/13669877.2019.1591491

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

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.

frontiers in Plant Science

REVIEW
published: 02 November 2020
doi: 10.3389/fpls.2020.584151

CRISPR-Cas12a (Cpf1): A Versatile Tool in the Plant Genome Editing Tool Box for Agricultural Advancement

Anindya Bandyopadhyay[1]†, Nagesh Kancharla[1†], Vivek S. Javalkote[1], Santanu Dasgupta[1] and Thomas P. Brutnell[2,3]*

[1] Reliance Industries Ltd., R&D-Synthetic Biology, Navi Mumbai, India, [2] Chinese Academy of Agricultural Sciences, Biotechnology Research Institute, Beijing China, [3] Gateway Biotechnology, Inc., St. Louis, MO, United States

OPEN ACCESS

Edited by:
Vladimir Nekrasov, Rothamsted Research, United Kingdom

Reviewed by:
Bing Yang, University of Missouri, United States
Basavaprabhu L. Patil, Indian Institute of Horticultural Research (ICAR), India
Florian Hahn, University of Oxford, United Kingdom

***Correspondence:**
Anindya Bandyopadhyay
anindya.b@ril.com;
anindya.biotech@yahoo.com

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Biotechnology, a section of the journal Frontiers in Plant Science

Received: *16 July 2020*
Accepted: *28 September 2020*
Published: *02 November 2020*

Citation:
Bandyopadhyay A, Kancharla N, Javalkote VS, Dasgupta S and Brutnell TP (2020) CRISPR-Cas12a (Cpf1): A Versatile Tool in the Plant Genome Editing Tool Box for Agricultural Advancement. Front. Plant Sci. 11:584151. doi: 10.3389/fpls.2020.584151

Global population is predicted to approach 10 billion by 2050, an increase of over 2 billion from today. To meet the demands of growing, geographically and socio-economically diversified nations, we need to diversity and expand agricultural production. This expansion of agricultural productivity will need to occur under increasing biotic, and environmental constraints driven by climate change. Clustered regularly interspaced short palindromic repeats-site directed nucleases (CRISPR-SDN) and similar genome editing technologies will likely be key enablers to meet future agricultural needs. While the application of CRISPR-Cas9 mediated genome editing has led the way, the use of CRISPR-Cas12a is also increasing significantly for genome engineering of plants. The popularity of the CRISPR-Cas12a, the type V (class-II) system, is gaining momentum because of its versatility and simplified features. These include the use of a small guide RNA devoid of trans-activating crispr RNA, targeting of T-rich regions of the genome where Cas9 is not suitable for use, RNA processing capability facilitating simpler multiplexing, and its ability to generate double strand breaks (DSB) with staggered ends. Many monocot and dicot species have been successfully edited using this Cas12a system and further research is ongoing to improve its efficiency in plants, including improving the temperature stability of the Cas12a enzyme, identifying new variants of Cas12a or synthetically producing Cas12a with flexible PAM sequences. In this review we provide a comparative survey of CRISPR-Cas12a and Cas9, and provide a perspective on applications of CRISPR-Cas12 in agriculture.

Keywords: CRISPR, Cas9, Cas12a, NHEJ, base editing, PAM, temperature sensitivity, agriculture

INTRODUCTION

Innovation has always been the driver of agricultural advancement from the earliest days of domestication to today's machine learning-based genomic selection technologies. Although the green revolution provided the caloric increase to sustain the current global populations, this energy intensive form of agriculture is beginning to plateau (Food and Agriculture Organization of the United Nations, 2013). Future global agricultural production will depend increasingly on tools and technologies to improve sustainability and deliver more nutritious crops that will lessen our dependence of animal-based protein and deliver new fiber and plant medicinal products to market.

Unprecedented social and political resistance blocked the broad dissemination of genetically modified crops (GM crops), which has resulted in only a handful of traits being successfully introduced to the market. The precision, ease and low cost of engineered genomes using genome editing technologies promises to greatly reduce the technological and economic restrictions associated with Genetically Modified Organisms (GMOs), but public acceptance is by no means guaranteed (Smart et al., 2017; Callaway, 2018). Nevertheless, plant scientists from industry and academia around the globe have embraced the technology for numerous applications including gene knockouts, fine-tuning gene expression through transcriptional activation/repression, inducing epigenetic changes, multiplex gene editing, and base editing applications in crops. Importantly, the reagents for genome editing can be delivered into the cell without incorporating DNA into the genome (Svitashev et al., 2016; Zhang et al., 2016; Ma et al., 2020) and result in mutations that are identical to those occurring in nature, potentially simplifying the regulatory process associated with traditional GMO crops. Indeed, regulation itself is challenging with many genome editing events as it can be technologically challenging if not impossible to differentiate between a genome-edited change and one that occurs naturally. Consequently, editing by Clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated proteins (Cas) is not only gaining popularity as a trait development tool, but also in achieving legal/regulatory approvals for product development in many countries (Schmidt et al., 2020).

Among clustered regularly interspaced short palindromic repeats (CRISPR) systems, Cas9 and Cas12a (originally identified as Cpf1) (Zetsche et al., 2015) have been most widely utilized and thus are most advanced in application. While the CRISPR-Cas9 system is still the most popular plant genome editing tool, the CRISPR-Cas12a nuclease is gaining broader adoption for multiple applications. One of the major limitations of CRISPR-Cas9 systems is the restriction of edits to regions of high GC content because of a "G" rich PAM sequence requirement (Bernabé-Orts et al., 2019). Although efforts are underway to engineer near "PAMless" Cas9 variants (Walton et al., 2020), only a few applications have been reported in plant systems (Zhong et al., 2019). Cas12a has the capacity to edit "T" rich PAM regions, and generates staggered ends that may promote site-directed integration events. Although the PAM sequence of Cas12a (e.g., TTTV) is longer than Cas9 (e.g., NGG), efforts have also been successful to engineer Cas12a variants with altered PAM specificities (Gao et al., 2017; Tóth et al., 2020). As recently shown in maize, Cas9 was used in combination with FLP recombinase to engineer gene stacks, greatly facilitating breeding efforts to stack traits of interest (Gao et al., 2020). This elegant use of genome editing technologies could be extended further by incorporating Cas12a, enabling a greater selection of target sites and potentially a higher efficiency of gene integration events. As promoters and introns are often AT-rich, the use of Cas12a also affords additional flexibility for engineering efforts (Wolter and Puchta, 2019). Although Cas12a displays a temperature sensitivity that has limited its utility in plant genome editing (Malzahn et al., 2019; Safari et al., 2019; Swarts, 2019), engineered variants have recently been generated with enhanced activities (Schindele and Puchta, 2020). In this review, we provide a comparison of CRISPR Cas9 and CRISPR Cas12a from the perspective of their applications in plant engineering and plant breeding and suggests ways to improve the utility of Cas12a in broadening its agricultural applications. Importantly, our intent is not to identify "the best" nuclease, but rather to highlight the advantages of broadening the molecular biology toolbox to incorporate both Cas9 and Cas12a technologies.

THE CRISPR-Cas SYSTEM

The CRISPR-Cas system evolved as a bacterial immune system to combat the invasion of phages and other mobile genetic elements like plasmids and transposons (Hille et al., 2018). There are three major steps involved in the evolution of CRISPR-Cas systems in bacteria namely:

(i) CRISPR adaptation-integration of foreign invader genomic fragments into a CRISPR array as spacer sequences,
(ii) crRNA biogenesis – the CRISPR array is transcribed into pre-crRNA and processed to mature crRNA's which in turn integrates with the Cas effector proteins to form crRNA effector complexes,
(iii) CRISPR interference – These programmed effector complexes identify and catalyze sequence-specific destruction of foreign invading genomic fragments (Jackson et al., 2017).

Based on the array of cas genes and the nature of the interference complex, the CRISPR-Cas system can be roughly divided into Class I and Class 2 systems which are further divided into six subtypes: Class I, type I, III, and IV are defined by multi-subunit complexes, and Class II, types II, V, and VI are postulated as single subunit effector endonucleases (Makarova et al., 2015; Shmakov et al., 2015). The Class 2 systems might have evolved from Class 1 systems with effector proteins originating from diverse mobile elements (Shmakov et al., 2015). Although Cas9 and Cas12-related proteins are similar in length (∼1100 to ∼1500 amino acids) it is likely that these families evolved independently from distinct transposable element families (Shmakov et al., 2015).

CRISPR-Cas systems have gained much popularity as a revolutionary genome-engineering tool because of their ease of use and multiple genome editing applications in the fields of medicine, agriculture, and animal husbandry. The most popular CRISPR-Cas9, originating from *Streptococcus pyogenes* (spCas9) belongs to the type II CRISPR system and has a protospacer adjacent motif (PAM) requirement of "NGG." Here, we will focus on the development of the Cas12 system that offers distinct advantages for genome engineering.

Cas12a, Class 2-Type V CRISPR System

The discovery and characterization of the Cas12a system was led by researchers at Broad Institute of MIT and Harvard University who identified a series of CRISPR nucleases in *Prevotella* and *Francisella 1* bacteria termed Cas12a (Cpf1) belonging to the

Class 2, type V CRISPR system (Zetsche et al., 2015). Additional effector proteins of this family include Cas12b (C2c1) and Cas12c (C2c3), respectively. Cas12a is an endonuclease which varies in size between 1200 and 1500 amino acids (Shmakov et al., 2015). The PAM sequence requirement for Cas12a is "TTN/TTTN/TTTV". (N = A/T/C/G; V = A/C/G). FnCas12a (from *Francisella novicida*), LbCas12a (from *Lachnospiraceae bacterium*), and AsCas12a (from *Acidaminococcus* sp.) are the three homologs of the Cas12a nucleases (Zetsche et al., 2015) (**Table 1**) which were widely utilized in plant genome editing technologies.

Structural Aspects of Cas12a

CRISPR-Cas12a is a two-component system, consisting of a protein/effector nuclease and a single crRNA. FnCas12a, LbCas12a, and AsCas12a proteins display similar domain organizations and range in size from ∼1300 to ∼1307 amino acids (aa). The crystal structure reveals a bi-lobed organization consisting of an α-helical recognition lobe (REC) and a nuclease lobe (NUC) (Dong et al., 2016; Yamano et al., 2016). The REC lobe consists of two domains Hel-1 and Hel-2, while the NUC lobe is comprised of the RuvC nuclease domain and three supplementary domains: PI, WED, and BH. The RuvC endonuclease domain of Cas12a is subdivided into three discontinuous segments (RuvC I–III), but it lacks the second HNH endonuclease domain and processes its mature crRNA without the utilization of trans-activating crispr RNA (tracrRNA) in comparison with Cas9 proteins (Safari et al., 2019; Wang J. et al., 2020).

The structure of AsCas12a crRNA reveals 20 nt direct repeat (5' handle) sequence and a spacer (guide segment) sequence of 23 nt in length (**Figure 1**). The crucial pseudoknot structure adopted by the direct repeat sequence is essential for the recognition by Cas12a. The pseudoknot structure can be broadly divided into a stem and a loop region. The pseudoknot, starting from -1 to -20 bases, consists of five Watson-Crick base pairs, one noncanonical U–U base pair, one UCUU tetraloop, one reverse Hoogsteen A–U base pair and three 5′-end bases (**Figure 2**). The hydrogen bonds formed within stem and loop regions stabilizes the pseudoknot structure. The bases U (-1), U (-10), U (-16), and A (-18) are conserved across Cas12a homologs indicating formation of similar tetraloop pseudoknot is crucial for the efficiency of endonuclease activity of Cas12a. The guide segment (spacer) sequence is complementary to the target DNA sequence and seed sequences (1–8 bases) are crucial for target specificity of CRISPR-Cas12a system (Dong et al., 2016; Yamano et al., 2016; Li L. et al., 2018; Swarts and Jinek, 2018; Safari et al., 2019). A survey of 16 uncharacterized Cas12a enzymes revealed differences in both PAM recognition and cut site repair mechanism (Zetsche et al., 2019), suggesting that it should be possible to engineer a range of activities into members of the Cas12a family.

crRNA Biogenesis

In contrast to CRISPR-Cas9 systems, type V systems do not require tracrRNA and RNase III for processing of mature crRNA. The transformation of pre-crRNA to mature crRNA (42–44 nt length) is mediated by intrinsic ribonuclease activities of Cas12a domains. The biogenesis of mature crRNA in *F. novicida* starts with the recognition of 27–32 base pair (bp) long spacers located adjacent to 36 bp long repeats by FnCas12a which are expressed as a single transcript (Zetsche et al., 2015). The repeat sequences in the pre-crRNA transcript forms a pseudoknot structure which is readily recognized by Cas12a (Dong et al., 2016; Yamano et al., 2016; Swarts and Jinek, 2018). Pseudoknot binding to divalent cations like Mg^{2+} or Ca^{2+} augments binding of the crRNA to Cas12a. The WED domain of Cas12a catalyzes the processing of the 5′ end of the crRNA but the 3′ end processing mechanism of crRNA is still obscure. A mature crRNA consists of 19–20 nt direct repeat sequence (5′ pseudoknot structure) and 20–24 nt guide or spacer sequence (Swarts and Jinek, 2018; Safari et al., 2019) (**Figure 2**).

Salient Difference Between Cas12a and Cas9

The major differences between Cas9 and Cas12a proteins (see **Figure 3**) include the following:

(i) The PAM requirement for Cas12a is "TTTN" which favors its use in targeting "AT" rich regions in the genome in contrast to the spCas9 system (PAM "NGG") (Zetsche et al., 2015);

(ii) Cas12a cleaves the target DNA strand 18-23 nucleotide (nt) distal of the PAM, leaving staggering ends (5 to 8 nt 5' overhangs) in contrast to blunt ends generated by Cas9 (Zetsche et al., 2015);

(iii) Cas12a processes the mature crRNA into 42 to 44 nt segments in contrast to Cas9 requiring tracrRNA for biogenesis of mature crRNA. This distinct feature of Cas12a makes it advantageous for multiplex gene editing, transcription, epigenetic modulations and base editing (Safari et al., 2019);

(iv) Unlike Cas9, Cas12a contains only one endonuclease domain, RuvC (NUC lobe) for cleavage of target and non-target DNA strands. The cleavage occurs in a sequential manner in which the non-target DNA strand is cleaved first and later the target DNA strand by the RuvC domain (Yamano et al., 2016);

(v) Lower off-target effects have been reported for Cas12a relative to Cas9, that are indistinguishable from spontaneous mutations caused during plant development (Bernabé-Orts et al., 2019). However, it has also been reported that Cas12a and several orthologs are capable cleaving randomized targets *in vitro* that contain up to four mismatches (Murugan et al., 2020). It remains to be seen, however, if this reduced specificity is evidenced in planta;

(vi) One of the major constraints of Cas12a broader adoption in plants is its lower efficiency at low temperatures (Malzahn et al., 2019);

(vii) A modified dCas12a fused to a human apolipoprotein B mRNA editing enzyme did not activate the DNA damage response and increased deamination efficiency and editing specificity relative to a similar dCas9 base editor nickase (Wang X. et al., 2020).

TABLE 1 | List of various CRISPR-Cas12a nucleases and their various applications in crops.

Plant name	Cas12a	Gene codes	Gene targeted	Target Trait	Binary vector	Transformation method	PAM	References
Rice	FnCas12a	*OsDL, OsALS, OsNCED1, OsAO1*	*Drooping leaf; Acetolactate synthase; 9-cis-epoxycarotenoid dioxygenase1; Aldehyde oxidase*	Floral organ identity; Herbicide resistance; Abscisic acid regulation-stress tolerance; caroteniod catabolism and abscisic acid metabolism-stress tolerance	pPZP200	Agrobacterium	TTN	Endo A. et al., 2016
Rice	LbCas12a	*OsEPFL9*	Stomatal developmental gene	Abiotic stress tolerance	pCAMBIA	Agrobacterium	TTTG	Yin et al., 2017
Rice	LbCas12a	*OsPDS, OsBEL*	*Phytoene desaturase*; Bentazon-sensitive-lethal	Carotenoid biosynthetic pathway; Herbicide resistance	pHSN400	Agrobacterium	TTTA,	Tang et al., 2017
Rice	FnCas12a, LbCas12a	*OsBEL, OsRLK*	Bentazon-sensitive-lethal; Receptor like kinases	Herbicide resistance; Biotic stress stimulant response gene	pCambia	Agrobacterium	TTN, TTTG, TTTC	Wang et al., 2017b
Rice	FnCas12a, LbCas12a	*OsCAO1*	*Chlorophyllide-a oxygenase gene*	Photosynthetic efficiency	pUC19	Biolistic Mediated	TTTC	Begemann et al., 2017
Rice	AsCas12a, LbCas12a	*OsPDS, OsDEP1, and OsROC5*	*Phytoene desaturase*; Dense and erect panicle 1; Rice outermost cell-specific gene5	Carotenoid biosynthetic pathway; Regulating carbon-nitrogen metabolism-Yield; leaf rolling controlling; Negatively modulates bulliform cells	pYPQ203 or pYPQ202, pYPQ220, pYPQ230, pYPQ223	Agrobacterium	TTTG	Tang et al., 2017
Rice	FnCas12a	*OsDEP1, OsPDS, and OsEPFL9*	Dense and erect panicle 1; *Phytoene desaturase*; Stomatal developmental gene	Regulating carbon-nitrogen metabolism-Yield; Carotenoid biosynthetic pathway; Abiotic stress tolerance	pYPQ203	Protoplasts transformation	TTTC	Zhong et al., 2018
Rice	LbCas12a	*OsALS*	*Acetolactate synthase*	Herbicide resistance	pCXUN	Particle bombardment	TTTG	Li et al., 2018b
Rice	*FnCas12a, AsCas12a, and LbCas1*	*OsROC5 and OsDEP1*	Rice outermost cell-specific gene5; Dense and erect panicle 1	Leaf rolling controlling; Negatively modulates bulliform cells; Regulating carbon-nitrogen metabolism-Yield	pYPQ141, 210, 230	PEG-mediated protoplasts transfection	TTTC	Malzahn et al., 2019
Rice	LbCas12a	*OsDEP1, OsROC5*	Dense and erect panicle 1; Rice outermost cell-specific gene5	Regulating carbon-nitrogen metabolism-Yield; leaf rolling controlling; Negatively modulates bulliform cells;	STU-Cas12a system	Protoplasts transformation	TTTA	Tang et al., 2019
Rice	LbCas12a	*OsPDS, OsGS3, OsALS, OsNAL*	*Phytoene desaturase*; Grain size; *Acetolactate synthase*; Narrow leaf	Carotenoid biosynthetic pathway; Grain length-yield; Herbicide resistance; grain yield	STU-poly-A vector	Agrobacterium	TTTA, TTTG	Xu et al., 2019
Rice	AsCas12a, LbCas12a	*OsPDS*	*Phytoene desaturase;*	Carotenoid biosynthetic pathway;	pCAMBIA1301	Biolistic Mediated	TTTG	Banakar et al., 2020
Arabidopsis and rice	AsCas12a, LbCas12a	*OsPDS, OsDEP1, and OsROC5*	*Phytoene desaturase;* Dense and erect panicle 1; Rice outermost cell-specific gene5	Carotenoid biosynthetic pathway;	dAsCas12a–SRDX and dLbCas12a–SRDX carrying vector	Floral dip and protoplasts transformation	TTTG	Tang et al., 2017
Arabidopsis	LbCas12a, enLbCas12a, ttLbCas12a	five gene targets	–	–	enAsCas12a and ttLbCas12a carrying vector	Floral dip	TTTC, TTTA	Schindele et al., 2020
Soybean	LbCas12a/ AsCas12a-RNP	*FAD2-1A, FAD2-1B*	*Fattyacid desaturase*	Increase oleicacid levels-Yield improvement	p2GW7	PEG-mediated protoplasts transformation	TTTTA	Kim et al., 2017

(Continued)

TABLE 1 | Continued

Plant name	Cas12a	Gene codes	Gene targeted	Target Trait	Binary vector	Transformation method	PAM	References
Tobacco	FnCas12a	*NtSTF1, NtPDS*	*Phytoene desaturase;* Stenofolia	NtPDS; NtSTF1	pRI201-AN	Agrobacterium	TTN	Endo A. et al., 2016
Tobacco	LbCas12a/ AsCas12a-RNP	*AOC*	*Allen Oxidase Cyclase*	Jasmonic acid biosynthesis metabolism	p2GW7	PEG mediated RNP delivery	TTTN	Kim et al., 2017
Maize	LbCas12a	*Gl2*	Maize Glossy gene2	Epicuticular wax formation-Regulation of fattyacid elongase pathway	pYPQ141, 210, 230	Agrobacterium	TTTG, TTTC	Lee K. et al., 2019
Tomato	LbCas12a	*SlHKT1;2*	Salt tolerance gene 1;2	Salt tolerance-Abiotic stress tolerant	pHRHKT12.1	Agrobacterium	TTTG, TTTA	van Vu et al., 2020
Allotetraploid cotton	LbCas12a	*GhLCA1*	Cloroplastos alterados	Chloroplast developmental gene	pGhRBE3	Agrobacterium	TTTG	Li B. et al., 2019
Citrus	LbCas12a	*CsPDS*	*Phytoene desaturase;*	Carotenoid biosynthetic pathway;	p1380	Agroinfiltration	TTTC	Jia et al., 2019
Wheat	LbCas12a	*TaWaxy* and *TaMTL*	Waxy and Matrilineal	MTL-Haploid induction gene; Waxy-starch synthase gene involved in flour quality	pWMB110-LbCas12a	Agrobacterium	TTTG	Liu et al., 2020

(viii) Intellectual property issues of Cas9 invention rights between the University of California vs Broad Institute of MIT and Harvard are still obscure and disorganized whereas Cas12a patent rights are a single point grant to Broad institute, MIT, and Harvard.

APPLICATION OF CRISPR-Cas12a IN AGRICULTURE

Cas12a editing has been widely utilized in many crops (see **Table 1**) including rice (Endo A. et al., 2016; Begemann et al., 2017; Hu et al., 2017; Tang et al., 2017, 2018; Wang et al., 2017a,b; Yin et al., 2017; Li L. et al., 2018; Li et al., 2019a; Jun et al., 2019; Malzahn et al., 2019; Banakar et al., 2020; Chen et al., 2020; Schindele and Puchta, 2020), wheat (Liu et al., 2020), maize (Lee K. et al., 2019), soybean (Kim et al., 2017), cotton (Li B. et al., 2019), tomato (van Vu et al., 2020), citrus (Jia et al., 2019), tobacco (Endo A. et al., 2016; Endo and Toki, 2019), and the model plant Arabidopsis (Wolter and Puchta, 2019; Schindele and Puchta, 2020). At present, three Cas12a genome editing systems AsCas12a, FnCas12a, and LbCas12a have been demonstrated in plants (Zhong et al., 2018) with varied efficiency.

Rice is one of the most well-studied crops due to its agricultural importance, small genome size, ease of transformation and available genetic resources making it an ideal flagship genome for the grasses (Mishra et al., 2018). These factors have also made it an ideal testing ground for developing genome editing technologies. Codon optimized FnCas12a binary vectors were utilized for targeted mutagenesis in rice (*OsDL, OsALS, OsNCED1, OsAO1*) and tobacco (*NtPDS and NtSTF1*) with average targeted mutation frequencies of 47.2% and 28.2%, respectively (Endo A. et al., 2016). Utilizing the LbCas12a nuclease two endogenous rice genes *OsPDS* and *OsBEL* were targeted with mutation frequencies of 21.4 and 41.2%, respectively (Xu et al., 2017). An independent study that targeted the disruption of *OsPDS* by LbCas12a resulted in a similarly high editing frequency of 32.3% (Banakar et al., 2020). Xu et al., 2017 also demonstrated that pre-crRNAs were more efficient in generating mutants than mature crRNAs in rice. However, the opposite was observed in HEK293T cells (Zetsche et al., 2017). In addition to these proof-of-concept experiments, LbCas12a was also used to create loss-of-function alleles of *OsEPFL9* which regulates stomatal density. These lines increased water use efficiency eight fold in T2 generation plants (Yin et al., 2017).

In vitro experiments conducted with FnCas12a and LbCas12a suggests that the efficiency of Cas12a depends upon the base content of the gene sequence targeted (Wang et al., 2017b). To compare the activity of two nucleases, *Acidaminococcus* sp. BV3L6 (As) and *Lachnospiraceae bacterium* ND2006 (Lb) were used to target six sites in three genes (*OsPDS, OsDEP1*, and *OsROC5*). Mutation frequencies observed ranged from 0.6 to 10% for AsCas12a and 15–25% with LbCas12a across the six targets (Tang et al., 2017). Importantly, whole-genome sequencing analysis of LbCas12a-edited plants revealed zero off-target mutations in the rice genome (Tang et al., 2018).

A potential advantage in using Cas12a in genome editing, is its ability to facilitate site-directed integration due to staggered

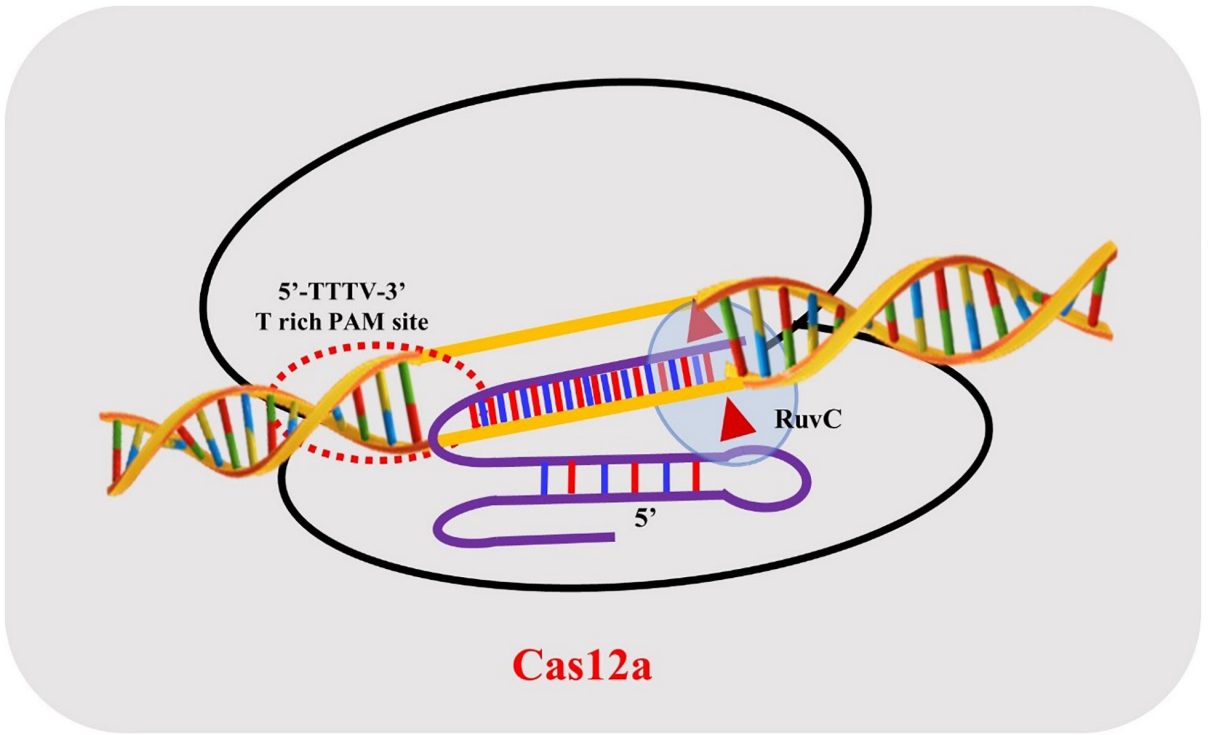


FIGURE 1 | Schematic representation of Cas12a crRNA with the target strand DNA association.

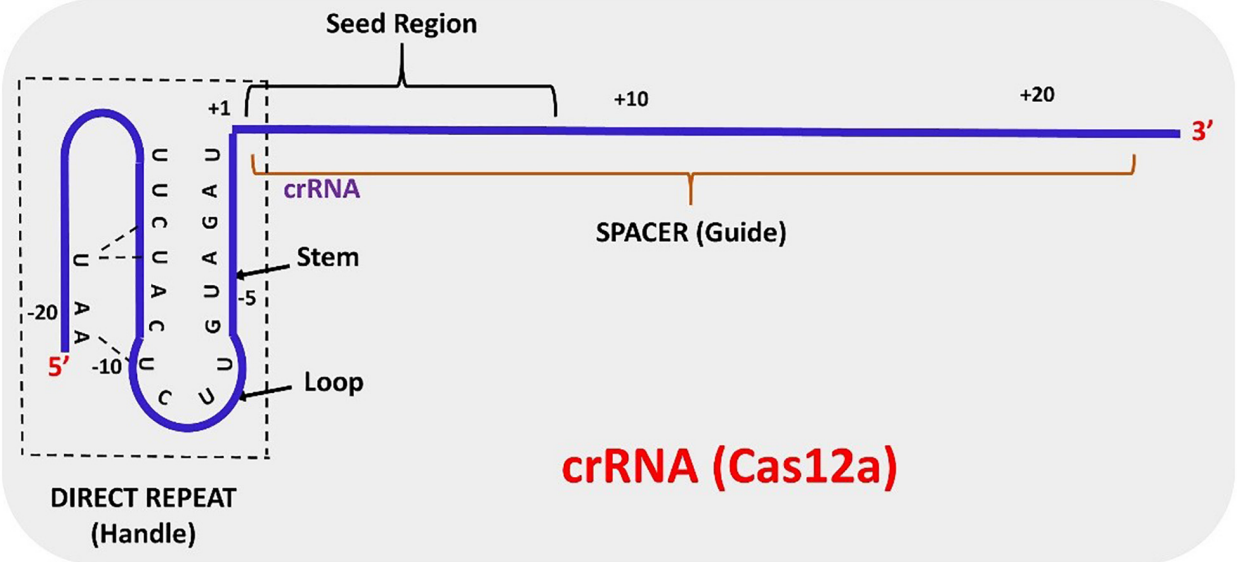


FIGURE 2 | Schematic representation of mature crRNA derived from the maturation of pre-crRNA.

overhangs. The expression of *F. novicida* and *L. bacterium ND2006* nucleases resulted in a high frequency of homology-directed repair (HDR) in rice suggesting a primary advantage of the Cas12a system over Cas9 for targeted gene integration (Begemann et al., 2017). However, the FnCas12a mutation frequency varies with the selection of PAM sequence (e.g. 10% to 35% efficiency with "TTTV" and 5–10% with "TTV") in rice (Zhong et al., 2018), suggesting that site directed targeting of sequences may be highly variable across the genome.

Given the inherent variations in site-specific editing efficiencies, it is challenging to directly compare the mutation efficiencies of Cas9 and Cas12a. Although studies in several

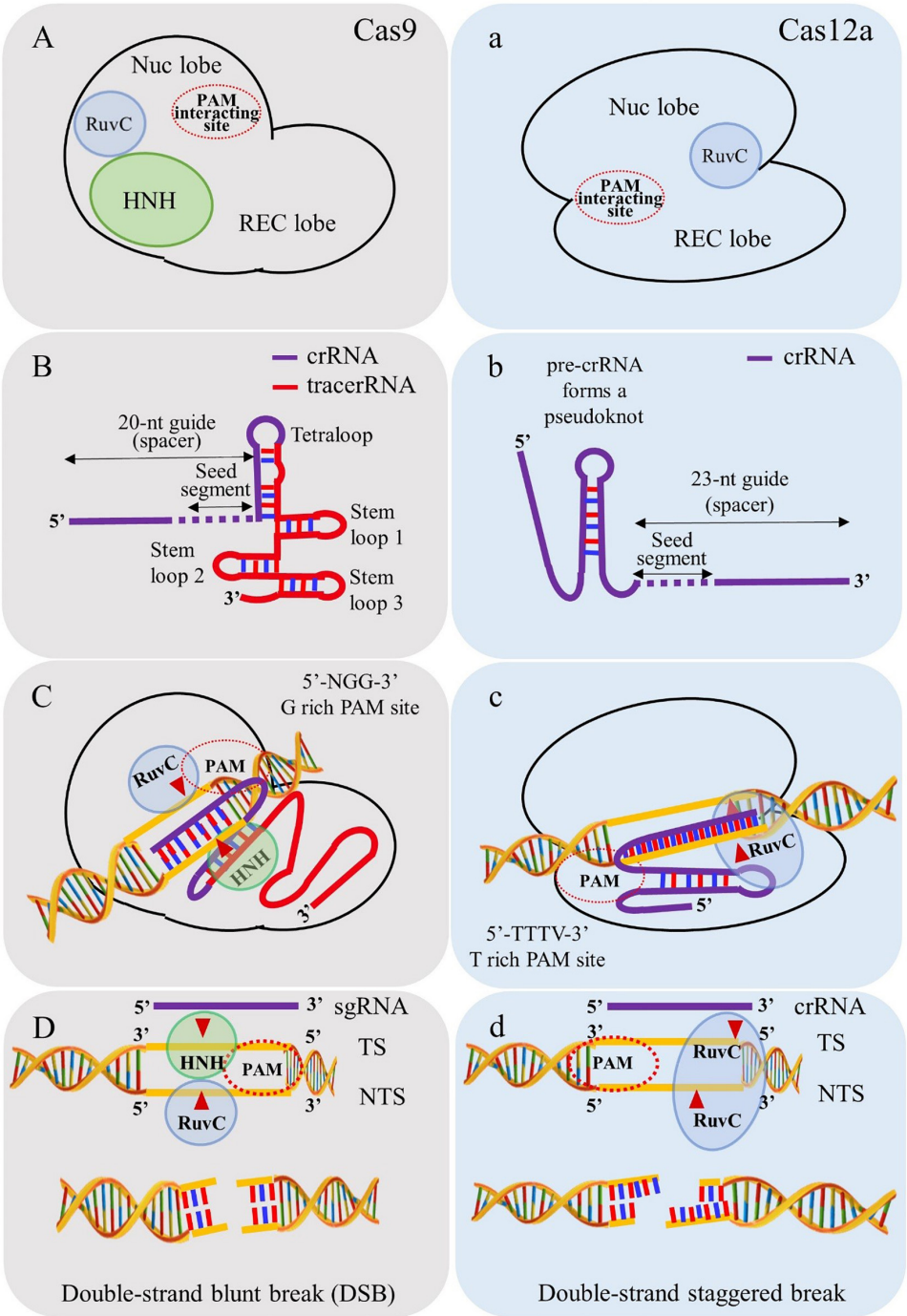


FIGURE 3 | Depiction of salient differences between Cas9 and Cas12a. **(A)** Cas9 contains two endonuclease domains to cleave target strand (TS) and non-target DNA strands (NTS) by HNH and RuvC domains, respectively. **(B)** Cas9 requires tracrRNA for biogenesis of mature crRNA. **(C)** PAM requirement of Cas9 is "NGG" rich regions for cleaving target site. **(D)** Cas9 simultaneously breaks TS and NTS and generates blunt ends. **(a)** Cas12a utilizes single endonuclease domain RuvC for cleaving TS and NTS. **(b)** Cas12a processes its own mature crRNA without intervention of tracrRNA. **(c)** PAM requirements of Cas12a is "TTN/TTTN" favoring "AT" rich regions. **(d)** Cas12a cleaves in a sequential manner in which NTS is cleaved first and followed by TS and generates double strand staggered break (sticky ends).

plant species have suggested lower editing efficiencies associated with Cas12a relative to Cas9 (Lee K. et al., 2019; Malzahn et al., 2019; Liu et al., 2020), Wang and colleagues used Cas9 and Cas12a to target the same loci and in one instance observed a higher efficiency of mutation with Cas9 (Lee K. et al., 2019) and with another target Cas12a was more efficient (Banakar et al., 2020). Various factors which might have attributed to the relative efficiency could be related to the gRNA sequences, epigenetic

modifications of the target site or expression of the endonuclease itself. Despite the variation in editing efficiencies, several groups have utilized Cas12a to edit more recalcitrant genomes including the allotetraploid cotton (Li B. et al., 2019), citrus (Jia et al., 2019), soybean (Kim et al., 2017) and wheat (Liu et al., 2020). Thus, although it appears that Cas12a is generally less efficient, as discussed below, newly developed versions of Cas12a promise to enhance its performance *in planta*.

FUTURE PERSPECTIVE: IMPROVING Cas12a FOR GREATER AND BROADER APPLICATIONS IN AGRICULTURE

Genome editing has opened up tremendous opportunities to improve the pace of agricultural advancement. The EU was one of the first organizations to develop a regulatory framework for genome editing (Friedrichs et al., 2019). They defined three tiers of genome editing. Site-directed nuclease (SDN) 1 events are the result of non-homologous end-joining and result in single nucleotide changes or small indels. SDN2 events result in template-mediated changes of a few nucleotides. In rice, for example, herbicide-resistant mutant lines were developed using template-mediated repair. LbCas12a was used create staggered breaks in the *Acetolactate synthase* (*ALS*) gene in the presence of a template molecule containing the point mutations of interest. Repair through a HDR mechanism (Li et al., 2018a) resulted in the precise introduction of small nucleotide changes. Although herbicide tolerance has been achieved effectively in major crop plants through transgenic approaches, the reduced regulatory barriers associated with CRISPR-mediated edits (especially outside the European Union, e.g., United States and Australia) makes this trait an attractive target for species that have had limited success overcoming regulatory hurdles with transgenic technologies such as rice. CRISPR genome edited crops may be exempted from GMO regulation in several countries on a case-by-case review including the United States, Canada, Australia, Japan, Colombia, Brazil, Argentina, and Chile. Other countries including the United Kingdom, India, Bangladesh, Philippines, and Indonesia are still in the process of developing regulations while the European Union and New Zealand have classified genome-edited crops as GMO's as they consider the process rather than the outcome. Events that insert foreign DNA from another species are likely to trigger the most rigorous regulatory reviews (SDN3) and will likely be considered transgenic in most countries. In the United States, the Environmental Protection Agency (EPA), the United States Department of Agriculture (USDA) and the Food and Drug Administration (FDA) all influence the regulatory path of an engineered plant product. Thus, it is still challenging to predict the time and costs of bringing genome-edited products to market (Schmidt et al., 2020).

Despite the challenges of the current regulatory environment, the scope and scale of genome editing opportunities will likely drive the entire agricultural industry. A few examples that incorporate genome editing technologies include accelerated breeding strategies (Li et al., 2016; Chen et al., 2019), allelic replacements (Ahmar et al., 2020), simultaneous double haploid production and editing (HI-EDIT) (Kelliher et al., 2019; Wang et al., 2019), crop domestication (Van Tassel et al., 2020; Zhang et al., 2020), and gene stacking (Razzaq et al., 2019). Importantly, these technologies are enabling a step change in the pace of crop improvement over conventional breeding and transformation technologies, especially when combined with emerging transformation technologies (e.g., co-expression of maize transcriptional factors, *BABY BOOM* and *WUSCHEL2*) (Lowe et al., 2016), machine learning (e.g., breeding), and imaging technologies. With so many potential opportunities, it is clear that both Cas9 and Cas12a will be utilized extensively in plant breeding in the years ahead. Nevertheless, several technological improvements in Cas12a will help to accelerate its broader adoption and utility.

Although, Cas12a has many advantages there are certain limitations as well which need to be addressed, such as PAM flexibility, to enable its broader application across the genome. Greater extent of modified or engineered versions of Cas12a are needed for single gene or multigene activation or repression. The relatively high temperature dependence of Cas12a is problematic in plant transformation as many crop protocols require low temperature. A Cas12a nickase has not been engineered to date which could facilitate gene integration without DSB, improving the possibility of HDR for allele replacement (**Figure 4**).

Development of Cas12a With Relaxed or "PAM-Less" Requirements

One of the significant limitations of Cas12a in plant genome editing is the relatively long TTTV PAM sequence requirements in eukaryotic genomes (Tóth et al., 2020). The theoretically estimated frequency of the TTTV PAM motif in DNA sequences is 3/256, a considerably more restrictive target space than that of the canonical NGG motif of SpCas9 (16/256). Several groups have recently attempted to engineer alternative PAM site requirements into Cas12a (**Table 2**). Perhaps the most promising is the improved Cas12a variant (*impLbCas12a*) generated by Toth and colleagues (Tóth et al., 2020). After engineering five independent nucleotide changes into Cas12a that had been shown in previous studies to modulate PAM-site selectivity and enzyme cutting efficiency, the *impLbCas12a* enzyme was able to cut at a TNTN consensus sequence with increased activity (Tóth et al., 2020). In addition to engineering relaxed specificities, Chen and colleagues identified two Cas12a variants (*CeCas12a* and *BeCas12a*) with a more stringent PAM site requirement in order to minimize off target events (Chen et al., 2020). This may have applications in engineering synthetic circuits when tight control of target sites is necessary. To continue expanding the repertoire of PAM sites or enhance enzyme activity, technologies such as phage-assisted continuous evolution (PACE) have shown much promise (Esvelt et al., 2011; Komor et al., 2016). In addition, the exploration of Cas12a homologs in diverse bacterial species, such as Lb5Cas12a, BoCas12a, BsCas12a (e.g., Zetsche et al., 2015; Gao et al., 2017) will likely provide both novel insights into Cas12a function and potentially new activities. However, it is important to note that all of these advances have come from studies in non-plant systems,

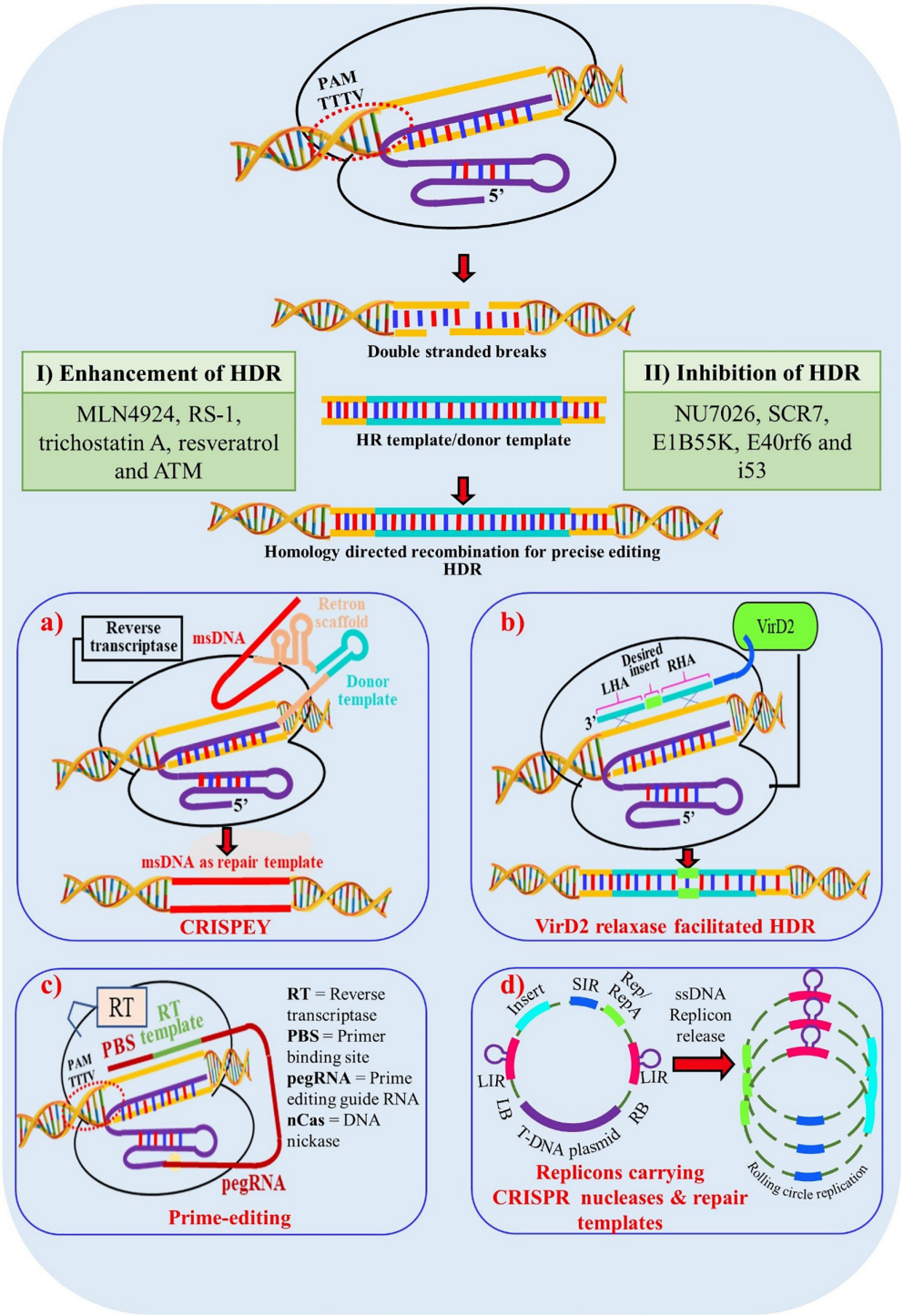


FIGURE 4 | Enhancement of Homology Directed Recombination (HDR) through multiple approaches. **(I)** Addition of chemical components which enhances HDR mechanisms in cells; **(II)** Chemical components which inhibit non homologous end joining (NHEJ) and thus indirectly promote HDR mechanism in cells; **(a)** Enhancement of HDR through CRISPEY (Cas9 Retron precISe Parallel Editing via homologY) method. Utilization of bacterial retron system to generate desired single stranded donor DNAs via multi-copy single-stranded DNA (msDNA); **(b)** Enhancement through VirD2 relaxase gene. A chimeric protein is synthesized with Cas9 protein tethered to the Agrobacterium VirD2 relaxase protein. Cas9 generates a precise DSB and VirD2 relaxse brings the donor template into close proximity to the DSB; **(c)** HDR enhancement through prime-editing for precise genome editing for crop improvement. **(d)** Enhancement through geminiviral replicon system. Utilization of rolling circle mechanism of geminivirus replicon system to generate multiple donor templates in vivo to enhance the success of HDR; msDNA – multi-copy single-stranded DNA; LHA – Left Homologous Arm; RHA – Right Homologous Arm; LB – Left Border; RB – Right Border; LIR – Long Intergenic Regions; SIR – Short Intergenic Regions.

TABLE 2 | List of modified and native Cas12a nucleases with their flexible "PAM" specificities.

CRISPR/Cas12a system	Modification in native CRISPR/Cas12a	PAM specificity	Amino acids Size	References
FnCas12a	Native	TTV,TTTV,KYTV	1300	Zetsche et al., 2015
FnCas12a-RR	N607R/K671R	TYCV, TCTV	1300	Toth et al., 2018; Zhong et al., 2018; Tóth et al., 2020
FnCas12a-RVR	N607R/K613V/N617R	TWTV	1300	Toth et al., 2018; Zhong et al., 2018; Tóth et al., 2020
FnCas12a-RVRR	N607R/K613V/N617R/K671R	TYCV, TCTV,TWTV	1300	Tóth et al., 2020
AsCas12a	Native	TTTV	1307	Zetsche et al., 2015; Kim et al., 2017; Tang et al., 2017; Bernabé-Orts et al., 2019
AsCas12a-RR	S542R/K607R	TYCV,CCCC	1307	Gao et al., 2017
AsCas12a-RVR	S542R/K548V/N552R	TATV	1307	Gao et al., 2017
AsCas12a-RVRR	S542R/K548V/N552R/K607R	TYCV,CCCC,TATV	1307	Tóth et al., 2020
LbCas12a	Native	TTTV	1228	Zetsche et al., 2015; Kim et al., 2017; Tang et al., 2017; Bernabé-Orts et al., 2019; Li et al., 2019b
LbCas12a-RR	G532R/K595R	TYCV,CCCC	1228	Gao et al., 2017; Li et al., 2018a; Zhong et al., 2018; Tóth et al., 2020
LbCas12a-RVR	G532R/K538V/Y542R	TATV	1228	Gao et al., 2017; Li et al., 2018a; Zhong et al., 2018; Tóth et al., 2020
LbCas12a-RVRR	G532R/K538V/Y542R/K595R	TYCV,CCCC,TATV	1228	Tóth et al., 2020
MbCas12a	Native	TTV, TTTV	1373	Zetsche et al., 2015
MbCas12a-RR	N576R/K637R	TYCV, TCTV	1373	Toth et al., 2018
MbCas12a-RVR	N576R/K582V/N586R	TWTV	1373	Toth et al., 2018
MbCas12a-RVRR	N576R/K582V/N586R/K637R	TYCV, TCTV,TWTV	1373	Tóth et al., 2020
enAsCas12a	E174R/S542R/K548R	VTTV,TTTT,TTCN,TATV	1307	Kleinstiver et al., 2019
impLbCas12a	D156R, G532R, K538V, Y542R, k595R	TTTV,TCCV, CCCV, TATC, TACV	1228	Tóth et al., 2020

thus considerable work remains to test these functionalities and develop new variants specifically for plant systems.

Cas12a for Gene Expression Modulation

As mentioned above, Cas9 has been used to manipulate gene expression through the engineering of DNAse dead (dd) enzymes that are directed to specific sites in the genome and similar strategies are now ongoing to engineer Cas12a (**Table 3**). The dual nuclease activity of Cas12a is essential for its ability to create double strand breaks (DSB) in the DNA and is dependent on the RuvC domain (Zhang et al., 2017; Safari et al., 2019). Alteration or mutation of the RuvC domain leads to the formation of a DNAse dead Cas12a (dCas12a) which retains the crRNA processing activity of Cas12a but fails to cleave the DNA (Zetsche et al., 2015). To exploit this feature of Cas12a, Leenay and colleagues created catalytically inactive ddCas12a enzymes to identify the repertoire of PAM sites in an *in-vivo* screen and infer binding specificity based on green fluorescent protein (GFP) readout (Leenay et al., 2017). Transcriptional repression mediated through both the inhibition of elongation and initiation in *E.coli* was achieved when a ddCas12a was targeted to multiple promoter or coding sequences (Zhang et al., 2017). A similar strategy was used to create a ddCas12a variant for *Streptomyces*, an important bacterial species for natural product discovery (Li L. et al., 2018). Additional modifications of Cas12a applied to human (Gao et al., 2018) and *Escherichia coli* (Miao et al., 2019) systems, have increased the activity of DNAse and activation/repression modalities of Cas12a. Although it is possible that similar mutations would also be effective in plant systems, this has yet to be demonstrated. In Arabidopsis, Tang et al. (2017) demonstrated transcriptional repression of *miR159b* utilizing deactivated nuclease domains of dAsCas12a (D908A) and dLbCas12a (D832A). Although, AsCas12a performed better than LbCas12a as a transcriptional repressor in *Arabidopsis thaliana*, species-specific differences are likely to influence binding efficiencies. Thus, further improvement in this area is greatly needed in plants, as LbCas12a is still the most widely utilized editing tool among all of the Cas12a variants, but reports in crop plants are relatively limited.

Cas12a Efficiency Improvement Through Chemical and Engineering Modifications

In addition to engineering the Cas12a protein, several groups have tried to optimize Cas12a cutting efficiencies and reduce off target modifications by modifying the crRNA molecule and Cas12a transcript. Li and colleagues showed in human cell lines by engineering a crRNA molecule containing five 2'-fluoro ribose at the 3' terminus together with an engineered Cas12a mRNA template in which uridine residues were replaced with pseudouridine throughout the entire transcript, cutting efficiencies could be improved. Together, these modifications led to an enhanced cutting efficiency of 300% above the wild-type plasmid template and crRNA controls (Li B. et al., 2017). Extensions of the crRNA at the 5'end also improved the efficiency

TABLE 3 | List of Cas protein activators and repressors and their applications in inducing gene expression in plants and human embryonic cells.

Activator/Repressor	Target Gene/Plant species	References
denAsCas12a–VPR	human cells:S170R, E174R, S542R or K548R	Kleinstiver et al., 2019
enAsBE1.1–1.4	human cells:S170R, E174R, S542R or K548R	Kleinstiver et al., 2019
dAs/LbCas12a-VP64-3xHA-crHDV	Luciferase (luc) gene in HEK293T cells	Gao et al., 2018
dCas9-H3K27 acetyltransferase p300	Flowering locus in Arabidopsis	Lee J. E. et al., 2019
dCas9-H3K9 methyltransferase KRYPTONITE	Flowering locus in Arabidopsis	Lee J. E. et al., 2019
dCas9-VP64	Flowering locus in Arabidopsis	Lee J. E. et al., 2019
dCas9-SRDX (Transcriptional repressor)	Flowering locus in Arabidopsis	Lee J. E. et al., 2019
dCas9-H3K9 methyltransferase G9a	Flowering locus in Arabidopsis	Lee J. E. et al., 2019
dCas9-MS2-VP64	Rice	Lowder et al., 2018
dCas9-mTALE-VP64	Rice	Lowder et al., 2018
dCas9–6TAL–VP128 (dCas9-TV)	Protoplasts of Arabidopsis	Li Z. et al., 2017
dCas9-VP128	Protoplasts of Arabidopsis	Li Z. et al., 2017
dCas9–VP256	Protoplasts of Arabidopsis	Li Z. et al., 2017
dCas9-VPR	Human cells- HEK293T	Chavez et al., 2016
dCas9-SAM	Human cells- HEK293T	Chavez et al., 2016
dCas9-SunTag	Human cells- HEK293T	Chavez et al., 2016

of Cas12a NHEJ and HDR activities (Park et al., 2018). Bin Moon et al., 2018 also demonstrated precise and enhanced indel-generating efficiency of Cas12a, up to 13-fold, with an engineered 3'-uridinylate rich crRNA in human HEK-293T (Human Embryonic Kidney) cells. This may be due in part to the enhanced stability of the molecule especially when Cas12a and crRNA are delivered to the cell as an ribo nucleo protein (RNP) in primary mouse myoblasts (Park et al., 2018).

McMahon et al., 2018 also demonstrated in HEK-293T cells that truncated synthetic RNA's (scrRNA) with chemical modification of nucleotides at 5' and 3' end with PS, 2'-F'5'-O-Me, and substitution with DNA nucleotides were more readily taken up by cells and enhanced its genome editing efficiency of AsCas12a relative to wild-type crRNAs. To identify additional components that may aid editing efficiencies, Ma and colleagues conducted a small molecule library screen and identified VE-822 and AZD-7762 for their ability to enhancing the genome editing efficiency of Cas12a in human pluripotent stem cells (Ma et al., 2018). In summary, chemical modification to the crRNA, Cas12a transcript and the addition of small molecules all were able to improve Cas12a efficiencies in mammalian systems. It remains to be seen, however, if any of these modifications result in similar efficiencies in plant systems. Some of the challenge in introducing chemical modifications could be overcome in plants if they are transformed using biolistics as it is easier to envision an RNP cocktail with small molecules rather than utilizing Agrobacterium transformation.

Improving HDR Efficiency

The DSBs generated by site-specific nucleases (SSNs) are repaired broadly through two repair pathways; non-homologous end joining (NHEJ) or homology directed repair (HDR) methods, generating either random or directed outcomes. In higher organisms and especially in plants the preferred DSB repair mechanism is achieved through NHEJ, where most often small indels are created causing frameshift mutations ultimately creating loss-of-function or "knock-out" alleles. If a DNA template (either single or double stranded) is present when the DNA is cleaved, then DNA repair can be achieved through HDR. This mechanism results in precise gene editing or modifications which is the most preferred way of bringing changes in the plant genome to produce high yielding and disease-resistant varieties in the crops. Unfortunately, the efficiency/frequency of HDR in plants is very low due to numerous factors including the low copy number of donor templates and length of donor templates (Puchta, 2005). Several groups have targeted the DNA repair pathway to engineer higher efficiencies of HDR including the targeted suppression of KU70 and KU80 and the overexpression of RAD54, RAD51, CtIP, CDK1, and Scr7 inhibitor to bypass NHEJ and promote HDR pathways (Shaked et al., 2005; Bozas et al., 2009; Qi et al., 2013; Maruyama et al., 2015; Ye et al., 2018; **Table 4**). These studies reveal that overexpression of HDR-related factors and suppression of NHEJ related factors are promising approaches to homology-directed gene targeting-HGT (Pinder et al., 2015; Robert et al., 2015; Endo M. et al., 2016; Rozov et al., 2019). The utilization of a geminivirus replicon system also enhances donor template delivery and available donor templates enhance homologous recombination (Baltes et al., 2014). However, this system failed to generate a high HDR repair frequency in Arabidopsis, suggesting species specific variation (De Pater et al., 2018; Hahn et al., 2018). A combination of an Agrobacterium VirD2 relaxase gene with Cas9 improved HDR in rice presumably by enhancing the proximity of repair template to the DSBs in rice (Ali et al., 2020). *In planta* gene targeting is another efficient method of HDR which was successfully demonstrated in rice (Sun et al., 2016), maize (Kumar et al., 2016) and Arabidopsis (Fauser et al., 2012; Schiml et al., 2014; Hahn et al., 2018; Wolter and Puchta, 2019). Here, sequence specific nucleases not only generate targeted DSB but also release a homology template from the T-DNA backbone. Another innovative method to increase HDR that has been applied in yeast systems is CRISPEY (Sharon et al., 2018). In this system, a DNA retron is used to tether a template sequence to a gRNA that is then delivered to the genomic target. Although the efficiency of HDR is high in yeast, it remains to be seen if a similar approach will work in plants. Prime editing can be used to introduce precise point mutations and insertions in the plant genome without separate repair templates and with reduced off-targets compared to other genome editing technologies (Anzalone et al., 2019). However, inherent limitations associated with prime editing must still be overcome if it is to be broadly adopted as an efficient, precise

and flexible plant genome editing tool for crop improvement (Lin et al., 2020; Marzec and Hensel, 2020; Xu et al., 2020).

As mentioned above, there are numerous advantages to utilizing an HDR-dependent pathway to engineer alleles of interest. In particular, because a template is provided, this template can be synthesized to contain naturaly occurring or novel alleles of any given locus. However, as shown in **Table 4**, the efficiencies of HDR are quite variable relative to target loci and both Cas12a and Cas9 have been used successfully for HDR. Various approaches and vector construct designs have been used to direct homology-dependent repair pathway utilizing the CRISPR-Cas12a endonuclease system. In one of the first examples of Cas12a-mediated gene replacements in plants, 1 kb of homologous sequence flanking a target sequence was used to insert a selectable marker into the *Chlorophyllide-a oxygenase* (*OsCAO*) locus in rice. Reagents including LbCas12a and FnCas12a plasmids, donor template and the crRNA expression construct were introduced as DNA templates through particle bombardment and insertion events identified (Begemann et al., 2017). Similar frequencies of HDR (4.6–7%) were obtained in the zebrafish model system when reagents were delivered as ribonucleoproteins coupled with donor template DNA. In this example, LbCas12a mediated homologous gene replacement at target loci *slc45a2* (albino) and *tyr* (tyrosinase) at higher efficiency than SpCas9 (Moreno-Mateos et al., 2017). Li et al. (2018a) also utilized RNP delivery but used RNA templates to mediate the

TABLE 4 | List of selected Cas9 and Cas12a homology-directed gene targeting mechanisms and their efficiencies in crops and animals.

Nuclease	HDR Enhancing System	Target Gene	Target Organism/plant	HDR efficiency percentage	References
FnCas12a and LbCas12a	Repairing DNA template flanked by 1000 bp homology DNA fragments/arms	Chlorophyllide-a oxygenase gene (CAO1)	Rice	3-8	Begemann et al., 2017
LbCas12a	Ribonucleo Protein (RNP) Complex	*slc45a2(albino) and tyr (tyrosinase)*	Zebrafish and Xenopus	5-7	Moreno-Mateos et al., 2017
	Synthesis dependent repair with donor repair template coupled with left homologous arm is sufficient for HDR mechanism	Acetolactate synthase gene (ALS)	Rice	0.014	Li et al., 2018a
				0.021	
	Ribozyme based strategy to synthesize crRNA's and DNA repair template	Acetolactate synthase gene (ALS)	Rice	4.6	Li et al., 2019a
				1.7	
	Utilization of homologous recombination enhancers	Acetolactate synthase gene (ALS)	Arabidopsis	1.47	Wolter and Puchta, 2019
	Utilization of viral multi replicon system (de novo engineered geminiviral replicon system) to increase the availability of donor template	Salt-tolerant (SlHKT1;2)	Tomato	4.5 - 9.8	van Vu et al., 2020
CRISPR-SpCas9	SpCas9, sgRNA and single-stranded DNA oligo's (72 base pair) into plant cells	Phytoene desaturase-OsPDS	Rice	6.8	Shan et al., 2013
CRISPR-Cas9	gRNAs targeting Lig4 gene were transformed with Cas9- Lig4 knockout resulted in enhancement of HDR	Acetolactate synthase gene (ALS)	Rice	0.15-1.0	Endo M. et al., 2016
	Geminivirus based vectors to release abundant HDR template	Actin-1 (ACT1)	Rice	6.8 - 19.4	Wang et al., 2017a
		Glutathione S-transferase (GST)		7.7	
CRISPR-SpCas9	DNA donor template contains constitutively expressing PAT gene with 1 kb homologous arm surrounding the target gene	Liguleless1-LIG1	Maize	0.2 - 4.61	Svitashev et al., 2015
		Male fertility genes - Ms26		0.13 - 3.11	
		Male fertility genes - Ms45		0.47 - 1.87	
		Acetolactate synthase gene (ALS1 and ALS2)		1.35 - 2.23	
CRISPR-Cas9-VirD2	Cas9-VirD2 chimeric protein helps in DSB and bringing close proximity of phosphorothioate mediated template DNA through VirD2 relaxase	Acetolactate synthase gene (ALS)	Rice	4.1 - 20.8	Ali et al., 2020
		Histone Deactylase (HDT)		0.2 - 8.7	

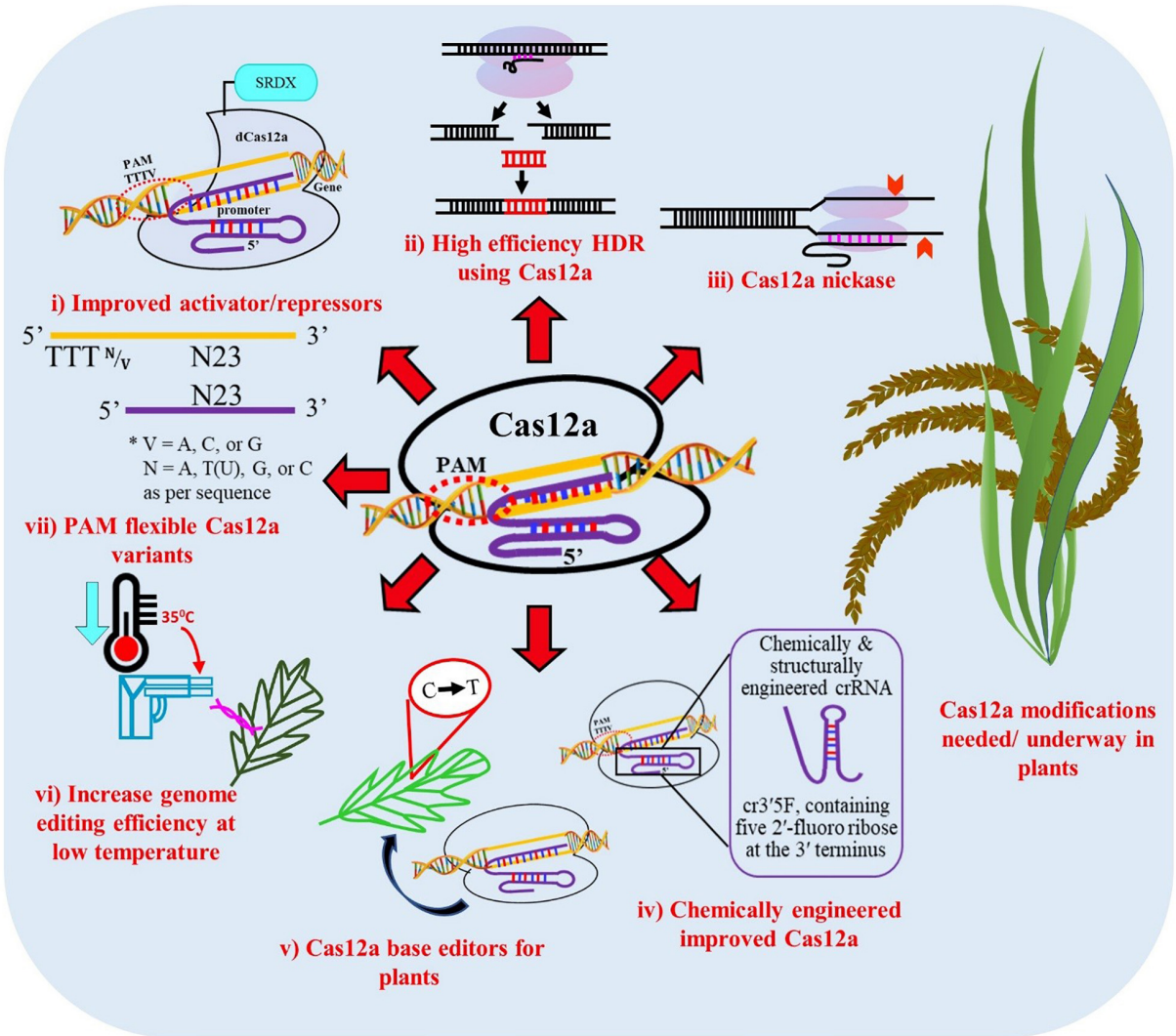


FIGURE 5 | Proposed modifications of Cas12a for improved editing efficiency in plants. **(i)** Improved dCas12a activators and repressors for modifying gene expression, **(ii)** High efficiency gene targeting through homology repair mechanisms, **(iii)** Nickases, **(iv)** Chemically modified and structurally engineered crRNA, **(v)** Base editors for point mutations and indel insertions, **(vi)** Increase genome editing efficiency at low temperature, **(vii)** PAM-flexible Cas12a variants.

HDR of ALS. Further refinements to the Cas12a system including the utilization of ribozymes and silent PAMs for homologous gene replacement in rice and maize have increased the frequency of template mediated repair using LbCas12a (Wolter and Puchta, 2019; Li et al., 2019b). To increase the availability of donor template van Vu et al., 2020 utilized a geminivirus replicon system to introduce a salt tolerance allele of *ANT1* and achieved a higher HDR efficiency rate of 9.8% compared with SpCas9 in tomato. Application of various strategies to insert genes through a homologous repair pathway enables one to edit crops with desired traits at a high frequency which is otherwise not possible with standard transgenic approaches. Further improvement and novel strategies of improving homologous recombination is greatly needed to fulfill important need of allele replacement in higher crops. Multiple techniques need to be tested widely such as combining different Agrobacterium virulence proteins (*Vir* proteins) with Cas12a, recruiting HDR proteins such as RAD group proteins with Cas12a, increasing donor template concentration in the presence of DSB by viral vectors (**Figure 5**).

SUMMARY

The utility of genome editing in plants is clear. From deepening our understanding of fundamental biological processes to engineering synthetic circuits and potentially introducing entirely novel biosynthetic pathways into production hosts, the CRISPR-Cas toolbox truly is revolutionizing plant biology. Here, we have reviewed some of the fundamental differences between the two most widely utilized systems for plant genome engineering and suggest that both Cas9 and Cas12a have unique advantages and disadvantages for genome engineering. As off

target mutations occur at low frequencies with both enzymes (Stupar et al., 2020), it is recommended that seed stocks should not be maintained with active Cas proteins in the genome. In the case of RNP delivery, this is not an issue, but when the reagents are delivered as DNA molecules, it is likely that the frequency of off targets will increase the longer the nuclease remains in the genome and primary targets are exhausted. Once the primary target is mutated and no longer serves as a target for the gRNA, then the probability of the gRNA identifying a new target even with some mismatches likely increases. In practice, and in particularly for plant breeding applications, these potential off-target events would be purged with successive backcrossing and in light of alternative approaches (e.g., chemical mutagenesis), the mutational load associated with CRISPR/Cas will be dramatically lower. Nevertheless, strategies to induce and characterize CRISPR-induced alleles, should incorporate the segregation of the CRISPR transgenes out of the plant genome and ensure that alleles generated are homozygous rather than chimeric/heterozygous.

Future strategies to develop synthetic circuits (Jusiak et al., 2016) or to engineer novel pathways will likely incorporate multiple Cas enzymes that can serve alternatively as repressors or activators of suites of genes (Lowder et al., 2018; Ming et al., 2020). These artificial transcriptional activators or repressors can be guided to specific loci to globally up or down regulate entire suites of genes. It is also easy to envision scenarios where entirely orthologous circuits are introduced and regulated by entirely novel promoter elements. In such a way an entirely new pathway may be introduced and expressed in a developmentally- or environmentally controlled manner.

Although the possibilities of engineering plant systems are exciting, these strategies must also be tempered by the regulatory environment that exists. As with any new technologies the potential benefits will be weighed against the potential risks of the technology. Many in the agricultural industry hope that the development of traits that will directly benefit the consumer will help drive public acceptance of the technology. However, diverse stakeholders and special interest groups who benefit from the fractionation of genome editing technologies into clearly defined buckets (e.g., GMO and non-GMO), will likely oppose the technology no matter how low the risk or big the benefit as we have witnessed with GMO technologies. Thus, it will be critical to establish sound and transparent regulatory frameworks for genome editing technologies and for scientists to not only be good stewards of the technologies but to actively participate in public forums to discuss the technology.

AUTHOR CONTRIBUTIONS

AB and SD conceptualized the idea. NK, TB, and AB wrote, reviewed, and edited the manuscript. VJ and NK drew the images. NK tabulated the tables. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We acknowledge Reliance Industries Ltd., for providing funding for this work. We also would like to thank Chunxia Liu of the Syngenta Beijing Innovation center for help in editing the manuscript. We would also like to thank the Reviewer 3 for many helpful comments and suggestions.

REFERENCES

Ahmar, S., Gill, R. A., Jung, K. H., Faheem, A., Qasim, M. U., Mubeen, M., et al. (2020). Conventional and molecular techniques from simple breeding to speed breeding in crop plants: recent advances and future outlook. *Int. J. Mol. Sci.* 21, 1–24. doi: 10.3390/ijms21072590

Ali, Z., Shami, A., Sedeek, K., Kamel, R., Alhabsi, A., Tehseen, M., et al. (2020). Fusion of the Cas9 endonuclease and the VirD2 relaxase facilitates homology-directed repair for precise genome engineering in rice. *Commun. Biol.* 3:44. doi: 10.1038/s42003-020-0768-9

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

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., Schubert, M., Collingwood, M., Vakulskas, C., Eggenberger, A. L., and Wang, K. (2020). Comparison of CRISPR-Cas9/Cas12a ribonucleoprotein complexes for genome editing efficiency in the rice phytoene desaturase (OsPDS) gene. *Rice* 13:4. doi: 10.1186/s12284-019-0365-z

Begemann, M. B., Gray, B. N., January, E., Gordon, G. C., He, Y., Liu, H., et al. (2017). Precise insertion and guided editing of higher plant genomes using Cpf1 CRISPR nucleases. *Sci. Rep.* 7:11606. doi: 10.1038/s41598-017-11760-6

Bernabé-Orts, J. M., Casas-Rodrigo, I., Minguet, E. G., Landolfi, V., Garcia-Carpintero, V., Gianoglio, S., et al. (2019). Assessment of Cas12a-mediated gene editing efficiency in plants. *Plant Biotechnol. J.* 17, 1971–1984. doi: 10.1111/pbi.13113

Bin Moon, S., Lee, J. M., Kang, J. G., Lee, N. E., Ha, D. I., Kim, D. Y., et al. (2018). Highly efficient genome editing by CRISPR-Cpf1 using CRISPR RNA with a uridinylate-rich 3'-overhang. *Nat. Commun.* 9:3651. doi: 10.1038/s41467-018-06129-w

Bozas, A., Beumer, K. J., Trautman, J. K., and Carroll, D. (2009). Genetic analysis of zinc-finger nuclease-induced gene targeting in *Drosophila*. *Genetics* 182, 641–651. doi: 10.1534/genetics.109.101329

Callaway, E. (2018). CRISPR plants now subject to tough GM laws in European Union. *Nature* 560:16. doi: 10.1038/d41586-018-05814-6

Chavez, A., Tuttle, M., Benjamin, B. W., Ewen-Campen, B., Chari, R., Ter-Ovanesyan, D., et al. (2016). Comparative analysis of Cas9 activators across multiple species. *Nat. Methods* 13, 563–567. doi: 10.1038/nmeth.3871

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

Chen, P., Zhou, J., Wan, Y., Liu, H., Li, Y., Liu, Z., et al. (2020). A Cas12a ortholog with stringent PAM recognition followed by low off-target editing rates for genome editing. *Genome Biol.* 21, 1–13. doi: 10.1186/s13059-020-01989-2

De Pater, S., Klemann, B. J. P. M., and Hooykaas, P. J. J. (2018). True gene-targeting events by CRISPR / Cas-induced DSB repair of the PPO locus with an ectopically integrated repair template. *Sci. Rep.* 8:3338. doi: 10.1038/s41598-018-21697-z

Dong, D., Ren, K., Qiu, X., Zheng, J., Guo, M., Guan, X., et al. (2016). The crystal structure of Cpf1 in complex with CRISPR RNA. *Nature* 532, 522–526. doi: 10.1038/nature17944

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, 1–9. doi: 10.1038/srep38169

Endo, A., and Toki, S. (2019). Targeted mutagenesis using FnCpf1 in tobacco. *Methods Mol. Biol.* 1917, 269–281. doi: 10.1007/978-1-4939-8991-1_20

Endo, M., Mikami, M., and Toki, S. (2016). Biallelic gene targeting in rice 1. *Plant Physiol.* 170, 667–677. doi: 10.1104/pp.15.01663

Esvelt, K. M., Carlson, J. C., and Liu, D. R. (2011). A system for the continuous directed evolution of biomolecules. *Nature* 472, 499–503. doi: 10.1038/nature09929

Fauser, F., Roth, N., Pacher, M., Ilg, G., Sánchez-Fernández, 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

Food and Agriculture Organization of the United Nations (2013). *The State of the World's Land and Water Resources for Food and Agriculture: Managing Systems at Risk*. Abingdon: Routledge, doi: 10.4324/9780203142837

Friedrichs, S., Takasu, Y., Kearns, P., Dagallier, B., Oshima, R., Schofield, J., et al. (2019). An overview of regulatory approaches to genome editing in agriculture. *Biotechnol. Res. Innov.* 3, 208–220. doi: 10.1016/j.biori.2019.07.001

Gao, H., Mutti, J., Young, J. K., Yang, M., Schroder, M., Lenderts, B., et al. (2020). Complex trait loci in maize enabled by crispr-Cas9 mediated gene insertion. *Front. Plant Sci.* 11:535. doi: 10.3389/fpls.2020.00535

Gao, L., Cox, D. B. T., Yan, W. X., Manteiga, J. C., Schneider, M. W., Yamano, T., et al. (2017). Engineered Cpf1 variants with altered PAM specificities increase genome targeting range. *Nat. Biotechnol.* 35, 789–792. doi: 10.1038/nbt.3900

Gao, Z., Herrera-Carrillo, E., and Berkhout, B. (2018). Improvement of the CRISPR-Cpf1 system with ribozyme-processed crRNA. *RNA Biol.* 15, 1458–1467. doi: 10.1080/15476286.2018.1551703

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

Hille, F., Richter, H., Wong, S. P., Bratovič, 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, X., Wang, C., Liu, Q., Fu, Y., and Wang, K. (2017). Targeted mutagenesis in rice using CRISPR-Cpf1 system. *J. Genet. Genomics* 44, 71–73. doi: 10.1016/j.jgg.2016.12.001

Jackson, S. A., McKenzie, R. E., Fagerlund, R. D., Kieper, S. N., Fineran, P. C., and Brouns, S. J. J. (2017). CRISPR-Cas: adapting to change. *Science* 356:eaal5056. doi: 10.1126/science.aal5056

Jia, H., Orboviæ, V., and Wang, N. (2019). CRISPR-LbCas12a-mediated modification of citrus. *Plant Biotechnol. J.* 17, 1928–1937. doi: 10.1111/pbi.13109

Jun, R., Xixun, H., Kejian, W., and Chun, W. (2019). Development and application of CRISPR/Cas system in rice. *Rice Sci.* 26, 69–76. doi: 10.1016/j.rsci.2019.01.001

Jusiak, B., Cleto, S., Perez-piñera, P., and Lu, T. K. (2016). Engineering synthetic gene circuits in living cells with CRISPR technology. *Trends Biotechnol.* 34, 535–547. doi: 10.1016/j.tibtech.2015.12.014

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

Kim, H., Kim, S. T., Ryu, J., Kang, B. C., Kim, J. S., and Kim, S. G. (2017). CRISPR/Cpf1-mediated DNA-free plant genome editing. *Nat. Commun.* 8, 1–7. doi: 10.1038/ncomms14406

Kleinstiver, B. P., Sousa, A. A., Walton, R. T., Tak, Y. E., Hsu, J. Y., Clement, K., et al. (2019). Engineered CRISPR–Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. *Nat. Biotechnol.* 37, 276–282. doi: 10.1038/s41587-018-0011-0

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. doi: 10.1038/nature17946

Kumar, S., Worden, A., Novak, S., Lee, R., and Petolino, J. F. (2016). A trait stacking system via intra-genomic homologous recombination. *Planta* 244, 1157–1166. doi: 10.1007/s00425-016-2595-2

Lee, J. E., Id, M. N., Duro, D. I., and Id, M. S. (2019). CRISPR-based tools for targeted transcriptional and epigenetic regulation in plants. *PLoS One* 14:e0222778. doi: 10.1371/journal.pone.0222778

Lee, K., Zhang, Y., Kleinstiver, B. P., Guo, J. A., Aryee, M. J., Miller, J., et al. (2019). Activities and specificities of CRISPR/Cas9 and Cas12a nucleases for targeted mutagenesis in maize. *Plant Biotechnol. J.* 17, 362–372. doi: 10.1111/pbi.12982

Leenay, R. T., Maksimchuk, K. R., Slotkowski, R. A., Agrawal, R. N., Gomaa, A. A., Briner, A. E., et al. (2017). Identifying and visualizing functional PAM diversity across CRISPR-Cas systems. *Mol. Cell* 62, 137–147. doi: 10.1016/j.molcel.2016.02.031

Li, B., Rui, H., Li, Y., Wang, Q., Alariqi, M., Qin, L., et al. (2019). Robust CRISPR/Cpf1 (Cas12a)-mediated genome editing in allotetraploid cotton (*Gossypium hirsutum*). *Plant Biotechnol. J.* 17, 1862–1864. doi: 10.1111/pbi.13147

Li, L., Wei, K., Zheng, G., Liu, X., Chen, S., Jiang, W., et al. (2018). CRISPR-Cpf1-assisted multiplex genome editing and transcriptional repression in streptomyces. *Appl. Environ. Microbiol.* 84:e00827-18. doi: 10.1128/AEM.00827-818

Li, B., Zhao, W., Luo, X., Zhang, X., Li, C., Zeng, C., et al. (2017). Engineering CRISPR-Cpf1 crRNAs and mRNAs to maximize genome editing efficiency. *Nat Biomed Eng.* 1, 1–21. doi: 10.1038/s41551-017-0066

Li, J., Meng, X., Zong, Y., Chen, K., Zhang, H., Liu, J., et al. (2016). Gene replacements and insertions in rice by intron targeting using CRISPR-Cas9. *Nat. Plants* 2, 1–6. doi: 10.1038/nplants.2016.139

Li, S., Li, J., He, Y., Xu, M., Zhang, J., Du, W., et al. (2019a). Precise gene replacement in rice by RNA transcript-templated homologous recombination. *Nat. Biotechnol.* 37, 445–450. doi: 10.1038/s41587-019-0065-7

Li, S., Li, J., Zhang, J., Du, W., Fu, J., Sutar, S., et al. (2018a). Synthesis-dependent repair of Cpf1-induced double strand DNA breaks enables targeted gene replacement in rice. *J. Exp. Biol.* 69, 4715–4721. doi: 10.1093/jxb/ery245

Li, S., Zhang, X., Wang, W., Guo, X., Wu, Z., Du, W., et al. (2018b). Expanding the scope of CRISPR/Cpf1-mediated genome editing in rice. *Mol. Plant* 11, 995–998. doi: 10.1016/j.molp.2018.03.009

Li, S., Zhang, Y., Xia, L., and Qi, Y. (2019b). CRISPR-Cas12a enables efficient biallelic gene targeting in rice. *Plant Biotechnol. J.* 18, 1351–1353. doi: 10.1111/pbi.13295

Li, Z., Zhang, D., Xiong, X., Yan, B., Xie, W., Sheen, J., et al. (2017). A potent Cas9-derived gene activator for plant and mammalian cells. *Nat. Plants* 3, 930–936. doi: 10.1038/s41477-017-0046-0

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

Liu, H., Wang, K., Jia, Z., Gong, Q., Lin, Z., Du, L., et al. (2020). Efficient induction of haploid plants in wheat by editing of TaMTL using an optimized Agrobacterium-mediated CRISPR system. *J. Exp. Bot.* 71, 1337–1349. doi: 10.1093/jxb/erz529

Lowder, L. G., Zhou, J., Zhang, Y., Malzahn, A., Zhong, Z., Hsieh, T. F., et al. (2018). Robust transcriptional activation in plants using multiplexed CRISPR-Act2.0 and mTALE-act systems. *Mol. Plant* 11, 245–256. doi: 10.1016/j.molp.2017.11.010

Lowe, K., Wu, E., Wang, N., Hoerster, G., Hastings, C., Cho, M. J., et al. (2016). Morphogenic regulators Baby boom and Wuschel improve monocot transformation. *Plant Cell* 28, 1998–2015. doi: 10.1105/tpc.16.00124

Ma, X., Chen, X., Jin, Y., Ge, W., Wang, W., Kong, L., et al. (2018). Small molecules promote CRISPR-Cpf1-mediated genome editing in human pluripotent stem cells. *Nat. Commun.* 9:1303. doi: 10.1038/s41467-018-03760-5

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

Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., et al. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13, 722–736. doi: 10.1038/nrmicro3569

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

Maruyama, T., Dougan, S. K., Truttmann, M., Bilate, A. M., Ingram, J. R., and Ploegh, H. L. (2015). Inhibition of non-homologous end joining increases the efficiency of CRISPR/Cas9-mediated precise [TM: inserted] genome editing. *Nature* 33, 538–542. doi: 10.1038/nbt.3190.Inhibition

Marzec, M., and Hensel, G. (2020). Prime editing: game changer for modifying plant genomes. *Trends Plant Sci.* 25, 722–724. doi: 10.1016/j.tplants.2020.05.008

McMahon, M. A., Prakash, T. P., Cleveland, D. W., Bennett, C. F., and Rahdar, M. (2018). Chemically modified Cpf1-CRISPR RNAs mediate efficient genome editing in mammalian cells. *Mol. Ther.* 26, 1228–1240. doi: 10.1016/j.ymthe.2018.02.031

Miao, C., Zhao, H., Qian, L., and Lou, C. (2019). Systematically investigating the key features of the DNase deactivated Cpf1 for tunable transcription regulation in prokaryotic cells. *Synth. Syst. Biotechnol.* 4, 1–9. doi: 10.1016/j.synbio.2018.11.002

Ming, M., Ren, Q., Pan, C., He, Y., Zhang, Y., Liu, S., et al. (2020). CRISPR–Cas12b enables efficient plant genome engineering. *Nat. Plants* 6, 202–208. doi: 10.1038/s41477-020-0614-6

Mishra, R., Joshi, R. K., and Zhao, K. (2018). Genome editing in rice: recent advances, challenges, and future implications. *Front. Plant Sci.* 9:1361. doi: 10.3389/fpls.2018.01361

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

Murugan, K., Seetharam, A. S., Severin, A. J., and Sashital, D. G. (2020). CRISPR-Cas12a has widespread off-target and dsDNA-nicking effects. *J. Biol. Chem.* 295, 5538–5553. doi: 10.1074/jbc.RA120.012933

Park, H. M., Liu, H., Wu, J., Chong, A., Mackley, V., Fellmann, C., et al. (2018). Extension of the crRNA enhances Cpf1 gene editing in vitro and in vivo. *Nat. Commun.* 9:3313. doi: 10.1038/s41467-018-05641-3

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. (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

Qi, Y., 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

Razzaq, A., Saleem, F., Kanwal, M., Mustafa, G., Yousaf, S., Arshad, H. M. I., et al. (2019). Modern trends in plant genome editing: an inclusive review of the CRISPR/Cas9 Toolbox. *Int. J. Mol. Sci.* 20:4045. doi: 10.3390/ijms20164045

Robert, F., Barbeau, M., Éthier, S., Dostie, J., and Pelletier, J. (2015). Pharmacological inhibition of DNA-PK stimulates Cas9-mediated genome editing. *Genome Med.* 7, 1–11. doi: 10.1186/s13073-015-0215-6

Rozov, S. M., Permyakova, N. V., and Deineko, E. V. (2019). The problem of the low rates of CRISPR/Cas9-mediated knock-ins in plants: approaches and solutions. *Int. J. Mol. Sci.* 20, 1–12. doi: 10.3390/ijms20133371

Safari, F., Zare, K., Negahdaripour, M., Barekati-Mowahed, M., and Ghasemi, Y. (2019). CRISPR Cpf1 proteins: structure, function and implications for genome editing. *Cell Biosci.* 9, 1–21. doi: 10.1186/s13578-019-0298-7

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

Schindele, A., Dorn, A., and Puchta, H. (2020). CRISPR/Cas brings plant biology and breeding into the fast lane. *Curr. Opin. Biotechnol.* 61, 7–14. doi: 10.1016/j.copbio.2019.08.006

Schindele, P., and Puchta, H. (2020). Engineering CRISPR/LbCas12a for highly efficient, temperature-tolerant plant gene editing. *Plant Biotechnol. J.* 18, 1118–1120. doi: 10.1111/pbi.13275

Schmidt, S. M., Belisle, M., and Frommer, W. B. (2020). The evolving landscape around genome editing in agriculture. *EMBO Rep.* 21, 19–22. doi: 10.15252/embr.202050680

Shaked, H., Melamed-bessudo, C., and Levy, A. A. (2005). High-frequency gene targeting in *Arabidopsis* plants expressing the yeast RAD54 gene. *PNAS* 102, 12265–12269. doi: 10.1073/pnas.0502601102

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 8–10. doi: 10.1038/nbt.2652

Sharon, E., Chen, S. A. A., Khosla, N. M., Smith, J. D., Pritchard, J. K., and Fraser, H. B. (2018). Functional genetic variants revealed by massively parallel precise genome editing. *Cell* 175, 544–557.e16. doi: 10.1016/j.cell.2018.08.057

Shmakov, S., Abudayyeh, O. O., Makarova, K. S., Wolf, Y. I., Gootenberg, J. S., Semenova, E., et al. (2015). Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. *Mol. Cell* 60, 385–397. doi: 10.1016/j.molcel.2015.10.008

Smart, R. D., Blum, M., and Wesseler, J. (2017). Trends in approval times for genetically engineered crops in the United States and the European Union. *J. Agric. Econ.* 68, 182–198. doi: 10.1111/1477-9552.12171

Stupar, R. M., Genetics, P., Graham, N., Patil, G. B., Bubeck, D. M., Dobert, R. C., et al. (2020). Plant genome editing and the relevance of off-target changes. *Plant Physiol.* 183, 1453–1471. doi: 10.1104/pp.19.01194

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

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, 1–7. doi: 10.1038/ncomms13274

Svitashev, S., Young, J. K., Schwartz, C., Gao, H., 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

Swarts, D. C. (2019). Making the cut(s): how Cas12a cleaves target and non-target DNA. *Biochem. Soc. Trans.* 47, 1499–1510. doi: 10.1042/BST20190564

Swarts, D. C., and Jinek, M. (2018). Cas9 versus Cas12a/Cpf1: Structure–function comparisons and implications for genome editing. *Wiley Interdiscip. Rev. RNA* 9, 1–19. doi: 10.1002/wrna.1481

Tang, X., Liu, G., Zhou, J., Ren, Q., You, Q., Tian, L., et al. (2018). A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. *Genome Biol.* 19, 1–13. doi: 10.1186/s13059-018-1458-5

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, 1–5. doi: 10.1038/nplants.2017.18

Tang, X., Ren, Q., Yang, L., Bao, Y., Zhong, Z., He, Y., et al. (2019). 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

Toth, E., Czene, B. C., Kulcsar, P. I., Krausz, S. L., Talas, A., Nyeste, A., et al. (2018). Mb- and FnCpf1 nucleases are active in mammalian cells: activities and PAM preferences of four wild-type Cpf1 nucleases and of their altered PAM specificity. *Nucleic Acids Res.* 46, 10272–10285. doi: 10.1093/nar/gky815

Tóth, E., Varga, É, Kulcsár, P. I., Kocsis-Jutka, V., Krausz, S. L., Nyeste, A., et al. (2020). Improved LbCas12a variants with altered PAM specificities further broaden the genome targeting range of Cas12a nucleases. *Nucleic Acids Res.* 48, 3722–3733. doi: 10.1093/nar/gkaa110

Van Tassel, D. L., Tesdell, O., Schlautman, B., Rubin, M. J., DeHaan, L. R., Crews, T. E., et al. (2020). New food crop domestication in the age of gene editing: genetic, agronomic and cultural change remain co-evolutionarily entangled. *Front. Plant Sci.* 11:789. doi: 10.3389/fpls.2020.00789

van Vu, T., 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.* doi: 10.1111/pbi.13373 [Online ahead of print],

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, B., Zhu, L., Zhao, B., Zhao, Y., Xie, Y., Zheng, Z., et al. (2019). Development of a haploid-inducer mediated genome editing system for accelerating maize breeding. *Mol. Plant* 12, 597–602. doi: 10.1016/j.molp.2019.03.006

Wang, J., Zhang, C., and Feng, B. (2020). The rapidly advancing Class 2 CRISPR-Cas technologies: a customizable toolbox for molecular manipulations. *J. Cell. Mol. Med.* 24, 3256–3270. doi: 10.1111/jcmm.15039

Wang, M., Lu, Y., Botella, J. R., Mao, Y., Hua, K., and Zhu, J.-K. (2017a). 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

Wang, M., Mao, Y., Lu, Y., Tao, X., and Zhu, J. K. (2017b). Multiplex gene editing in rice using the CRISPR-Cpf1 system. *Mol. Plant* 10, 1011–1013. doi: 10.1016/j.molp.2017.03.001

Wang, X., Ding, C., Yu, W., Wang, Y., He, S., Yang, B., et al. (2020). Cas12a base editors induce efficient and specific editing with low DNA damage response. *Cell Rep.* 31:107723. doi: 10.1016/j.celrep.2020.107723

Wolter, F., and Puchta, H. (2019). In planta gene targeting can be enhanced by the use of CRISPR/Cas12a. *Plant J.* 100, 1083–1094. doi: 10.1111/tpj.14488

Xu, R., Li, J., Liu, X., Shan, T., Qin, R., and Wei, P. (2020). Development of plant prime-editing systems for precise genome editing. *Plant Commun.* 1:100043. doi: 10.1016/j.xplc.2020.100043

Xu, R., Qin, R., Li, H., Li, D., Li, L., Wei, P., et al. (2017). Generation of targeted mutant rice using a CRISPR-Cpf1 system. *Plant Biotechnol. J.* 15, 713–717. doi: 10.1111/pbi.12669

Xu, R., Qin, R., Li, H., Li, J., Yang, J., and Wei, P. (2019). Enhanced genome editing in rice using single transcript unit CRISPR-LbCpf1 systems. *Plant Biotechnol. J.* 17, 553–555. doi: 10.1111/pbi.13028

Yamano, T., Nishimasu, H., Zetsche, B., Hirano, H., Ian, M., Li, Y., et al. (2016). EIGER detector: application in macromolecular crystallography. *Acta Crystallogr. Sect. D Struct. Biol.* 165, 949–962. doi: 10.1016/j.cell.2016.04.003. Crystal

Ye, L., Wang, C., Hong, L., Sun, N., Chen, D., Chen, S., et al. (2018). Programmable DNA repair with CRISPRa/i enhanced homology-directed repair efficiency with a single Cas9. *Cell Discov.* 4:46. doi: 10.1038/s41421-018-0049-7

Yin, X., Biswal, A. K., Dionora, J., Perdigon, K. M., Balahadia, C. P., Mazumdar, S., et al. (2017). CRISPR-Cas9 and CRISPR-Cpf1 mediated targeting of a stomatal developmental gene EPFL9 in rice. *Plant Cell Rep.* 36, 745–757. doi: 10.1007/s00299-017-2118-z

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

Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E. M., et al. (2017). Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat. Biotechnol.* 35, 31–34. doi: 10.1038/nbt.3737

Zetsche, B., Strecker, J., Abudayyeh, O. O., Gootenberg, J. S., Scott, D. A., and Zhang, F. (2019). A survey of genome editing activity for 16 Cas12a orthologs. *Keio J. Med.* 6, 1–5. doi: 10.2302/kjm.2019-0009-oa

Zhang, X., Wang, J., Wang, J., Cheng, Q., Zheng, X., and Zhao, G. (2017). Multiplex gene regulation by CRISPR-ddCpf1. *Cell Discov.* 3, 1–9. doi: 10.1038/celldisc.2017.18

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, 1–8. doi: 10.1038/ncomms12617

Zhang, Y., Pribil, M., Palmgren, M., and Gao, C. (2020). A CRISPR way for accelerating improvement of food crops. *Nat. Food* 1, 200–205. doi: 10.1038/s43016-020-0051-8

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

Zhong, Z., Zhang, Y., You, Q., Tang, X., Ren, Q., Liu, S., et al. (2018). Plant genome editing using FnCpf1 and LbCpf1 nucleases at redefined and altered PAM sites. *Mol. Plant* 11, 999–1002. doi: 10.1016/j.molp.2018.03.008

Conflict of Interest: TB is employed by Gateway Biotechnology, Inc. The remaining authors are employed by Reliance Industries Ltd.

TB and all other authors ensure that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

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