EVALUATION OF REAGENTS AND METHODS FOR GENOME EDITING IN POTATO (*SOLANUM TUBEROSUM* L.)

By

Nathaniel Martin Butler

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Plant Breeding, Genetics and Biotechnology—Crop and Soil Sciences—
Doctor of Philosophy

2015

ABSTRACT

EVALUATION OF REAGENTS AND METHODS FOR GENOME EDITING IN POTATO (*SOLANUM TUBEROSUM* L.)

By

Nathaniel Martin Butler

Genome editing using sequence-specific nucleases (SSNs) is rapidly becoming a standard tool for genetic engineering in crop species. The implementation of zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and CRISPR/Cas (clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated systems (Cas)) for inducing double-strand breaks enables targeting of virtually any sequence for genetic modification. Targeted mutagenesis via nonhomologous end-joining (NHEJ) and gene targeting via homologous recombination (HR) have been demonstrated in a number of plant species but reports have been limited in vegetatively propagated crops, such as potato (*Solanum tuberosum* Group Tuberosum L.)

The aim of this dissertation was to develop reagents and methods for genome editing in potato. This was accomplished by demonstrating TALEN and CRISPR/Cas reagents targeting the potato *ACETOLACTATE SYNTHASE1* (*ALS1*) gene were successful in inducing targeted mutations in reporter and endogenous gene targets. Targeted mutations using CRISPR/Cas were capable of both clonal and germline transmission, making CRISPR/Cas the preferred reagent for this application. TALEN and CRISPR/Cas reagents were also used in combination with a geminivirus expression vector for gene targeting experiments to incorporate point mutations within

the *ALS1* locus. Transformed events modified by both TALEN and CRISPR/Cas reagents in the geminivirus expression vector carried gene targeting modifications that supported reduced herbicide susceptibility phenotypes. Gene targeting modification detection and reduced herbicide susceptibility phenotypes were enhanced by regenerating lines under high selection. The evaluated reagents and methods in this dissertation provide a frame work for genome editing in potato and other vegetatively propagated crops and have important implications for basic research and agriculture.

I’d like to dedicate this dissertation to my dear wife, Emily Ertel who has supported me throughout my graduate education and has been my highest inspiration.

ACKNOWLEDGEMENTS

I'd like to first and foremost acknowledge Dan Voytas and members of the Voytas lab at the University of Minnesota who made this work possible. Our interactions at numerous meetings and my training visits were central to the development of my understanding of genome editing and made me feel like a member of the lab. I'd also like to thank members of the Douches lab who provided support in the lab, field, greenhouse and in life. I'd also like to thank members of the Buell lab who kept me thinking about potato on a genome level. I'd also like to thank members of my committee, Jim F. Hancock and Guo-Qing Song for keeping me grounded in traditional breeding and the foundations of tissue-culture. I'd also like to thank Dan Buhr and Chris Thornburn for assisting me in the lab and greenhouse, and Swathi Nadakuduti for carrying on genome editing work in the Douches lab.

Last but certainly not least, I would like to thank my advisor, Dave Douches and committee member, C. Robin Buell who are in my eyes, the father and mother of modern potato genetics. I will carry with me the knowledge and insight provided by Dave and Robin throughout my career, and I look forward to continuing to learn from them into the future.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

A	adenine
Agrobacterium	*Agrobacterium tumefaciens*
ALS	*acetolactate synthase*
barley	*Hordeum vulgare*
BeYDV	*bean yellow dwarf virus*
bp	base pairs
BT	*Bacillus thuringiensis*
C	cytosine
canola	*Brassica rapa*
cas	*CRISPR-associated systems*
CoDA	context-dependent assembly
Corn, maize	*Zea mays*
cotton	*Gossypium ssp.*
CPB	Colorado potato beetle
CRISPR/Cas	clustered regularly interspaced short palindromic repeats/CRISPR-associated systems
crRNA	CRISPR RNA
D	aspartic acid
DM	doubled-monoploid
DNA	deoxyribonucleic acid
DSB	double-strand break
EMS	ethyl methanesulfonate

EPA	Environmental Protection Agency
FDA	Food and Drug Administration
G	guanine
GM	genetically modified
GT	gene targeting
H	histidine
HR	homologous recombination
kb	kilobase pairs
late blight	*Phytophthora infestans*
MCR	mutagenic chain reactions
N	asparagine
NHEJ	nonhomologous end-joining
OPEN	Oligomerized Pool ENgineering
PAM	protospacer adjacent motif
petunia	*Petunia hybrida*
PLRV	*potato leafroll virus*
potato	*Solanum tuberosum* Group Tuberosum L.)
QSR	quick serve restaurants
Rep	Replicase
Rep/RepA	Replicase/replicase A
rice	*Oryza sativa*
RIL	first recombinant inbred line
RNA	ribonucleic acid
RGEN	RNA-guided endonuclease
RNAi	RNA interference

RVD	repeat variable diresidue
S genes	susceptibility genes
sgRNA	single-guide RNA
SSN	sequence-specific nuclease
soybean	*Glycine max*
sugar beets	*Beta vulgaris*
T	thymine
T-DNA	transfer DNA
TAL	transcription activator-like
TALEN	transcription activator-like effector nuclease
tobacco	*Nicotiana tabacum*
tomato	*Solanum lycopersicum*
tracrRNA	*trans*-activating crRNA
TRV	*tobacco rattle virus*
USDA	US Department of Agriculture
USDA-APHIS	US Department of Agriculture-Animal and Plant Health Inspection Service
VIGS	virus-induced gene silencing
VInv	*vacuolar invertase*
wheat	*Triticum aestivum*
ZFN	Zinc finger nuclease

CHAPTER 1: INTRODUCTION

The limitations of conventional genetic engineering

The first genetically engineered crops

Genetic engineering of plants has been developing since the 1983 breakthrough when the first foreign DNA was delivered to plant cells and regenerated using *Agrobacterium tumefaciens* (*Agrobacterium*) (Barton *et al.*, 1983; Herrera-Estrella *et al.*, 1983). The new technology promised to revolutionize agriculture by increasing yields and accelerating the progress of crop trait development beyond what was possible with random mutagenesis and traditional breeding (Borlaug, 2000). This so-called second green revolution has led to genetically modified (GM) crops being grown on 447 million acres in 28 different countries in 2013, with worldwide acreage increasing more than 100-fold since 1996 (James, 2014). Nevertheless, the focus of first generation GM crops on inputs traits such as herbicide and pest resistance and the use of “unnatural” DNA have created an insurgency of public concern revolved around GM technology (Maeseele, 2013; Wohlers, 2013). This public controversy has limited the potential of GM crops worldwide and the success of vegetable and fruit GM crops in the US (Van Montagu, 2011).

The first commercially available GM crop, the Flavr Savr tomato was introduced in 1994 by the Californian company, Calgene. The Flavr Savr tomato had great potential for market success by providing output traits geared towards both processors and consumers. Before its release, analysts predicted the GM tomato would have an annual

market value of at least $500 million and would be used to supply major processors such as Campbell's (Thomas, 2013). However, a combination of aggressive anti-GM protests in Europe and the US and delays with the US Food and Drug Administration (FDA) led Campbell's and other major processors to publically announce they would not be accepting the GM tomato. The lack of processor support and uncertainty of the public about GM technology led to the GM tomato being pulled from the US market in 1997. This series of events was repeated with other early GM vegetable and fruit crops, such as the GM potato and ultimately resulted in the industry's focus on input traits in major commodity crops, such as corn (*Zea mays*) and soybeans (*Glycine max*).

The first GM potato was developed by the Monsanto subsidiary, NatureMark under the NewLeaf trademark. Virus resistance to the *potato leafroll virus* (*PLRV*) and insect resistance to the Colorado Potato Beetle (CPB) were the focus input traits for the NewLeaf potato (Perlak *et al.*, 1993; Kaniewski and Thomas, 2004). In 1991, the company had developed four varieties of CPB-resistant potato using the Atlantic, Russet Burbank, Snowden, and Superior cultivars. The NewLeaf potatoes were rapidly adopted by growers and expanded in acreage from 1,800 in 1995 to 55,000 acres in 1998 (Toevs *et al.*, 2011). Upon the initial success of the NewLeaf potatoes, NatureMark introduced the NewLeaf Plus potato in the Russet Burbank cultivar that had resistance to both CPB and *PLRV*, and were developing other traits such as resistance to late blight (*Phytophthora infestans*) and tubers with increased starch content (Kaniewski and Thomas, 2004). However, public announcements made by major quick serve restaurants (QSR), such as McDonald's in response to public concern about GM technology led to the NewLeaf potatoes being pulled from the market in 2002 (Toevs *et*

al., 2011). The failure of the GM tomato and potato initiated a transition in GM crop development which led to the first generation of commercially successful GM crops.

The first generation of successful GM crops were developed in soybean, corn, cotton (*Gossypium ssp.),* canola (*Brassica rapa*) and sugar beets (*Beta vulgaris*) with input traits for tolerance to the herbicide, Roundup® (Roundup Ready) and for insect resistance using insecticidal Cry proteins from the bacterium, *Bacillus thuringiensis* (BT). Since 1996, production of first generation Roundup Ready and BT GM crops has steady been increasing and as of 2013, accounts for 93%, 85% and 82% of soybeans, corn and cotton, respectively, produced in the US. First generation GM crop production growth is also being seen worldwide, particularly in developing countries, where 75%, 82%, 32% and 31% of the soybeans, cotton, corn and canola, respectively, were GM in 2011 (James, 2014).

The GM debate

The rapid adoption and market success of the first generation GM crops is largely due to the focus on critical inputs traits, which reduced production costs and improved yield, and the development of crops which are not consumed directly by the public. This later point has recently been coming back as a public issue as more information about GM technology and widespread use of GM crops is being made available to the public (Maeseele, 2013). One response from the public concerning GM technology is to have mandatory labeling of foods made from GM crops (Wohlers, 2013). Such legislature has only been proposed in a select number of states but emphasizes the public's uncertainty of GM technology.

Public concern for GM technology stems from different aspects of genetic engineering and the public's perception of nature (Wohlers, 2013). The idea of having food that comes from "natural" or organic sources is becoming more important to public consumers and GM is being considered "unnatural" since it is based on taking DNA from one species and transferring it to another unrelated species through transgenesis (Hunter, 2014). This notion of GM crops being "unnatural" is supported by the US Department of Agriculture (USDA) definitions for organic foods which classifies any GM crops as being non-organic (USDA, 2012). The rapid growth of the "natural" and organic food industry speaks to the public's interest in their food, where it comes from, and how it was made.

All native genetic engineering

To address this issue, companies such as Simplot Plant Sciences and Okanagan Specialty Foods have developed genetically engineered crops using "all native" rather than transgenic DNA. The use of "all native" DNA differs from first generation transgenics in that it uses DNA from closely related species which could otherwise be transferred to a new variety using traditional breeding (Hunter, 2014). This notion has been stretched into the assumption that "all native" DNA exists in nature when in the cases of both the Simplot Innate™ potato and Okanagan Arctic® apple, RNA-interference (RNAi) mechanisms are used to reduce native gene function (Waltz, 2015a; Waltz, 2015b). Although the development of such varieties through the same RNAi mechanisms is possible with traditional breeding, it remains unlikely. Nevertheless, public acceptance of "all native" genetically engineered crops has already

been demonstrated in public surveys and by regulatory agencies and are currently on the market (Toevs *et al.*, 2011; Hunter, 2014).

Another public concern of GM technology is related to the target traits of first generation GM crops (Maeseele, 2013; Wohlers, 2013). The input traits of first generation GM crops benefit farmers directly and consumers indirectly by providing lower food prices. However, the growth of the "natural" and organic food industry suggests this benefit to consumers is being outweighed by the public's uncertainty of GM technology.

To address this issue, both the Simplot Innate™ potato and Okanagan Arctic® apple provide output traits that benefit consumers directly (Waltz, 2015a; Waltz, 2015b). This approach is most effective in food crops, such as vegetables and fruits that are purchased and consumed directly by the public. Hence, both the Innate™ potato and Arctic® apple have reduced browning traits that improves table-life and creates new uses of fresh apples and potatoes. Furthermore, the reduced browning trait also benefits producers by allowing longer storability and reduced cosmetic damage during processing. Nevertheless, in a case of history repeating itself, McDonald's and apple growers have not initially accepted the new products and the fate of the Innate™ potato and Arctic® apple lie in the hands of the public.

The second generation of GM crops

Conventional genetic engineering in crop species using transgenes or "all native" DNA and *Agrobacterium* is subject to regulation by the US Department of Agriculture-Animal and Plant Health Inspection Service (USDA-APHIS), Food and Drug

Administration (FDA) and US Environmental Protection Agency (EPA) (Wohlers, 2013). The development of a GM crop and de-regulation process allowing its commercial release has been estimated to cost approximately $136 million and take over 10 years (McDougall *et al.*, 2011). This costly investment is another major contributing factor to the industry's focus on input traits and high production crops for the development of first generation GM crops. Nevertheless, these costs have kept small biotechnology companies from developing GM crops and competing with large companies, such as Monsanto and Syngenta (Christou, 2013).

A sea change in biotechnology is being seen in the development of a second generation of GM crops that is outpacing the already outdated agency regulations for genetic engineering. Current regulations are triggered with the use of "plant pests" such as *Agrobacterium* and plant pathogen DNA to develop GM crops (Wohlers, 2013). The advancement of alternative DNA delivery methods, such as particle bombardment and "all native" DNA has allowed companies to develop genetically engineered crops that avoid agency regulation (Waltz, 2012). Companies can gain information about if their crop will be regulated or not by submitting "Am I regulated?" letters to the USDA (www.aphis.usda.gov/). An increasing number of letters are being submitted to the USDA from large and small biotechnology companies for this purpose and most are being granted "non-regulated" status (Camacho *et al.*, 2014). This side-stepping of agency regulations is allowing small biotechnology companies to get a foothold in the GM market but may be preventing informative safety studies and vital checkpoints (Tuteja *et al.*, 2012).

Genome editing for genetic engineering

Genome editing using sequence-specific nucleases (SSNs) provides an alternative to conventional genetic engineering and could have major advantages over existing technologies in terms of efficacy and safety (Puchta and Fauser, 2013; Voytas and Gao, 2014). In contrast to conventional genetic engineering, genome editing uses existing DNA in a given crop species to confer new traits and does not rely on the persistent function of transgenes. This is an important distinction from transgenic or "all native" genetic engineering that relies on random integration and persistent function of introduced DNA within the genome.

Genome editing instead relies on the transient action of SSNs and/or repair templates within plant cells and the regeneration of modified lines. SSNs can be designed to target a specific region of the genome to guide integration of new DNA (Weinthal *et al.*, 2013a), make directed modifications to DNA (Baltes *et al.*, 2014), make targeted mutations in DNA (Curtin *et al.*, 2011), or make large rearrangements (Qi *et al.*, 2013a). Theoretically, the non-random nature of genome editing allows more stability across GM crops engineered with a certain trait and is safer than conventional genetic engineering by reducing so-called "off-target" effects (Andersen *et al.*, 2015). However, data concerning the effects of genome editing is limited and methodology and effective reagents are needed for generating genome edited lines in crop species.

SSNs for genome editing

Genome editing is enabled by SSNs that create breaks in DNA and harness DNA repair pathways to modify DNA in living cells (Figure 1). The field of genome editing has been developing ever since the discovery of the DNA double helix and has been a long anticipated goal for molecular biologists and geneticists (Watson and Crick, 1953; Doudna and Charpentier, 2014). Early experiments in yeast and mammalian cells demonstrated that oligonucleotides and other small molecules could bind DNA in a sequence-specific manner and cleave DNA upon replication or treatment with certain chemicals. The induction of DNA damage at these specific regions of the genome forced living cells to repair the break or perish (Cho *et al.*, 1995; Broitman *et al.*, 1987). These early approaches are still used today by biotech companies for genome editing but have largely been replaced with SSN technology (Lusser *et al.*, 2012).

Homing endonucleases

The first SSNs used for genome editing, called homing endonucleases or meganucleases were first characterized in yeast and subsequently identified in other microbes (Netter *et al.*, 1974; Belfort and Perlman, 1995) (Figure 2A). The process of "homing" involves microbial group I and group II self-splicing introns which are capable of duplicating into recipient alleles that lack the intronic sequence and are inherited in a dominant manner (Jacquier and Dujon, 1985). The group I introns first described from yeast duplicate by encoding the homing endonuclease, I-SceI within the intronic sequence. I-SceI was shown to be capable of forming a DNA double-strand break within the target allele which facilitated repair by homologous recombination and incorporation

of the intronic sequence and the I-SceI coding sequence. The self-replicating nature of homing endonucleases and group I and group II introns inspired the development of so-called mutagenic chain reactions (MCR) or gene drives that are actively in development using homing endonucleases and other SSNs to control animal populations (Gantz and Bier, 2015; Chan *et al.*, 2013).

The discovery of homing endonucleases in other microbes led to the classification of at least five different families based on biological host range and conserved physical structure (Stoddard, 2005). Analysis of the physical structures of representative members of each of these families revealed each family has unique catalytic cores that most likely originate from ancestral nucleases. The most studied homing endonuclease families, the GIY-YIG, PD-(D/E)xK, and LAGLIDADG homing endonucleases originate from phage, bacterial and eukaryotic/archaeal hosts, respectively. The other two families contain an HNH nuclease active site and are structurally divergent from each other and the other families, originating from phage and protists. The diversity in homing endonucleases and their DNA binding domains supports great variability in DNA specificities, affecting target sequence quality, length and binding affinity (Stoddard, 2011).

The structural diversity of homing endonucleases is reflected in the utilization of different DNA binding and cleavage mechanisms that vary across homing endonucleases (Stoddard, 2011). This is due to the holistic involvement of homing endonuclease structure for DNA binding and subsequent cleavage. For example, the LAGLIDADG homing endonucleases use antiparallel β-sheets that run the length of its structure to interact with major grooves of target DNA binding sites (Jurica *et al.*, 1998).

The interactions that result in the homing endonuclease's affinity for DNA and cleavage involve both sequence-specific and non-specific contracts (Scalley-Kim *et al.*, 2007). The intricate interactions homing endonucleases require for DNA binding and cleavage differ greatly across homing endonucleases and target sequences and depend on particular environmental conditions (pH, ionic state, etc.). This holistic nature of homing endonucleases requires extensive screening of natural homing endonucleases or protein engineering and testing in order to identify homing endonucleases capable of targeting a desired sequence and host species (Stoddard, 2011). The cost associated with homing endonuclease development has restricted the application of this SSN for genome editing to a select number of labs and biotechnology companies but has paved the way for other SSN technologies.

Zinc finger nucleases

Zinc finger nucleases (ZFNs) provided the first breakthrough in SSN technology as being both versatile and effective reagents for genome editing in plants (Carroll, 2011) (Figure 2B). Unlike homing endonucleases, all ZFNs are synthetic, relying on the DNA binding properties of zinc fingers and the catalytic function of the *Fok*I nuclease. The development of ZFNs was made possible by the discovery that the natural type IIS restriction enzyme, *Fok*I had separable DNA binding and cleavage domains (Li *et al.*, 1992). This provided a unique opportunity to engineer vertically any DNA binding protein into a nuclease (Kim and Chandrasegaran, 1994). The well characterized Cys_2His_2 zinc fingers from the diverse family of zinc finger transcription factors proved to be a useful conjugate DNA recognition domain (Pavletich and Pabo, 1991). Each approximately 30 amino acid zinc finger is capable of binding a single atom of zinc and

three base pairs (bp) of DNA. The modular fashion in which zinc fingers could be assembled allowed versatility in target DNA sequences and subsequent ZFNs (Kim *et al.*, 1996).

Early studies using ZFNs revealed severe context dependence, such that certain zinc fingers worked better in combination with others, and toxicity using certain ZFNs (Pruett-Miller *et al.*, 2008; Ramirez *et al.*, 2008). The observed toxicity associated with ZFNs was attributed to excessive off-targeting due to the limited number of zinc fingers that can be assembled in a single ZFN, which is typically two to six zinc fingers (Cornu *et al.*, 2008) . To overcome this issue, the *Fok*I nuclease domain was separated into N- and C-termini and fused to separate ZFNs which target adjacent target sequences on opposite strands separated by a 5 to 7 bp "spacer" (Miller *et al.*, 2007; Szczepek *et al.*, 2007). This strategy required both ZFNs to simultaneously bind a target sequence before cleavage could occur and increased target site specificity to 12 to 36 bp. This approach was highly successful useful due to the homodimeric characteristics of *Fok*I which was previously discovered (Bitinaite *et al.*, 1998). Overcoming off-targeting issues was critical for the success of ZFNs for genome editing but another issue of context dependence was limiting the development of new ZFNs.

The effects of context dependence are rooted in steric interactions between adjacent DNA binding motifs and a conjugate DNA target (Cornu *et al.*, 2008). The ability to modularly assemble zinc fingers suggested context dependence may be less of an issue for ZFNs compared to homing endonucleases (Grizot *et al.*, 2011). However, modularly assembled ZFNs had major context dependence issues and

alternative approaches were needed to design ZFNs (Joung *et al.*, 2010; Ramirez *et al.*, 2008).

In response to the issue of context dependence, a brute force approach was developed for assembling ZFNs called Oligomerized Pool ENgineering (OPEN) (Maeder *et al.*, 2008). OPEN relied on pre-selected pools of zinc finger proteins containing three zinc fingers targeting different three bp subsites to be tested against target DNA sequences in a bacterial two-hybrid (B2H), and selected based on binding affinity. This approach proved to be very effective for assembling efficient ZFNs but was labor intensive, required special expertise and was not conducive for broad adoption. A more broadly accepted approach for assembling ZFNs, called Context-Dependent Assembly (CoDA) was developed by identifying N- and C-terminal three zinc fingers arrays that were previously determined to function with a common middle finger (Sander *et al.*, 2011). Hence, combinations of 319 N-terminal fingers and 344 C-terminal fingers can be assembled with 18 fixed middle-fingers to target a given sequence. However, some flexibility in target DNA sequence is needed since only certain combinations of fingers are functional. This limitation of ZFN assembly and cell toxicity has restricted the broad adoption of ZFNs for genome editing but efficient ZFNs, such as Zif268 have been effective reagents for gene therapy and have provided an important gold standard for subsequent SSN platforms (Ellis *et al.*, 2013).

Transcription activator-like effector nucleases

The elucidation of the molecular code underlying transcription activator-like (TAL) effector DNA binding and the development of transcription activator-like effector

nucleases (TALENs) marked a significant event in molecular biology history (Figure 2C). Never before was it possible to modularly design a protein to target a specified DNA sequence one base at a time. Such a one-to-one molecular code would allow enzymes, ranging from nucleases to methylases to be targeted to specified regions of the genome for applications ranging from genome editing to making directed epigenetic modifications to whole chromosome rearrangements. Furthermore, synthetic transcription factors could be engineered to act as master regulators of complex gene networks and control quantitative traits. The implications of the TAL effector code is far reaching and originates from nature's greatest molecular biologists—pathogens.

Xanthomonas ssp. is a gram-negative bacterial pathogen that is the causal agent of bacterial blight in a range of host plant species (Kay and Bonas, 2009). *Xanthomonas* infects host plants by injecting effector proteins, called TAL effectors into the plant cell using a type III secretion system. Once in the plant cell, TAL effectors are localized to the nucleus where they act as transcription activators, primarily upregulating disease susceptibility genes (*S* genes).

TAL effectors bind DNA targets using a central protein domain comprising of tandem 33 to 34 amino acid repeats (Doyle *et al.*, 2013). The nearly identical repeats have discrete variation primarily in amino acids 12 and 13, known as repeat variable diresidues (RVDs). The RVDs were found computationally and experimentally to confer specificity to specific DNA nucleotides, allowing each repeat of the DNA binding domain to bind a single nucleotide (Moscou and Bogdanove, 2009; Boch *et al.*, 2009). These findings were based on computational associations made of RVDs with conjugate target DNA sequences and function of modularly assembled DNA binding domains. The TAL

effector modularity exceeded previously described DNA binding proteins, such as zinc fingers by supporting one-to-one base recognition and could be used to engineer synthetic nucleases similar to ZFNs by fusing the DNA binding domain to *Fok*I, creating TALENs (Li *et al.*, 2011a; Christian *et al.*, 2010).

TALENs gained immediate attention from the scientific community as being effective and highly versatile nucleases for genome editing (Baker, 2011). Free-access software was quickly developed to aid TALEN users to design TAL binding domains based on the context dependence seen with natural TAL effectors (Doyle *et al.*, 2012). These target sequence requirements were at first thought to improve TAL effector DNA binding efficiency if not being essential (Moscou and Bogdanove, 2009; Cermak *et al.*, 2011). However, upon further experimentation, it was determined that some context dependence seen in nature was not necessarily affecting TAL effector DNA binding (Mussolino *et al.*, 2011; Li *et al.*, 2011b). For example, options for not having the nucleotide thymine and guanine at the first or last positions in the DNA target site, respectively was found to be not essential. This resulted in the release of different versions of the TAL effector design software that allowed options to be turned on or off to increase or decrease target site stringency, respectively at the user's discretion (Doyle *et al.*, 2012).

Consideration of context dependence within TAL effector DNA target sites was even further complicated by the use of different TAL effector architectures and different spacer lengths (Miller *et al.*, 2011; Mussolino *et al.*, 2011; Bedell *et al.*, 2012; Cermak *et al.*, 2011; Mahfouz *et al.*, 2011; Li *et al.*, 2011b). The development of TAL effector architectures or scaffolds originates from the optimization of TAL effectors for use as

TALENs by truncating N- and C-termini of the TAL effector amino acid sequence flanking the central DNA binding domain. This resulted in different DNA binding activities depending on the number of RVDs and length of spacer between each TALEN. Nevertheless, robust studies testing different TAL effector truncations identified versatile architectures that have proven to be effective across multiple species in different kingdoms (Bedell *et al.*, 2012; Miller *et al.*, 2011; Zhang *et al.*, 2013). Once efficient TAL effector architectures were established, extensive studies were conducted on the nature of the RVD-nucleotide interaction and exploration of artificial RVDs.

The solved crystal structure of the TAL effector in complex with double-strand DNA illustrated the eloquence of the interaction and allowed for development of new RVDs (Mak *et al.*, 2012; Deng *et al.*, 2012). In nature, over 20 RVDs exist, each with variable affinities for each of the four nucleotides (Sun and Zhao, 2013). The most widely used RVDs—NI, NG, HD and NN primarily recognize A, T, C, and G/A, respectively. The crystal structures revealed that the TAL effector binds the major groove of target double-strand DNA as a right-handed superhelix. Each repeat forms a two-helix structure that presents its RVD within a loop in proximity to its target nucleotide. The first amino acid of the RVD—amino acid 12 is either histidine (H) or asparagine (N) and does not contact the target nucleotide directly. Instead, the side-chain of H or N forms a hydrogen bond with the backbone carbonyl oxygen of amino acid 8, alanine of each repeat, stabilizing conformation of the RVD-loop.

The second amino acid of the RVD—amino acid 13 is involved in sequence-specific contacts with the target nucleotide using different interactions. HD and NN RVDs utilize a hydrogen bond between the side chain of aspartic acid (D) and

asparagine (N) and the amine group or N7 nitrogen base of the opposing cytosine and guanine base, respectively. NG and NI RVDs utilize van der Waal contacts with the α carbon backbone of glycine and aliphatic side chain of isoleucine with the methyl group or C8 and N8 of the opposing thymine or adenine base, respectively. With this information, new synthetic RVDs have been developed which confer new affinities to target nucleotides (Yang *et al.*, 2014; Juillerat *et al.*, 2015). Interesting, one of these new RVDs, RV, confers specificity to all four nucleotides and can be used as a "wild card" RVD. However, the difficulty of assembling TAL effector repeats and the ability to target only one target sequence per TALEN pair has limited the use of TALENs for genome editing.

CRISPR/Cas

Leading up to the discovery of CRISPR/Cas (clustered regularly interspaced short palindromic repeats/CRISPR-associated systems), all SSN platforms were based on engineering new DNA binding proteins to bind and cleave novel target DNA (Figure 2D). These SSN platforms require a new nuclease to be engineered for each DNA target that can vary in DNA binding and nuclease efficiency (Juillerat *et al.*, 2014; Gaj *et al.*, 2013). However, the discovery of CRISPR/Cas revealed another class of nucleases that utilizes RNA to recognize and target DNA sequences, called RNA-guided endonucleases (Wiedenheft *et al.*, 2012). The characterization of CRISPR/Cas and its development for genome editing has become a paradigm in genetic engineering and has raised important ethical questions concerning genome editing (Doudna and Charpentier, 2014). The major factors contributing to the success of CRISPR/Cas is its ability to target multiple DNA targets using a single nuclease (Cas9) in a process called

multiplexing, and the overwhelming robust efficacy of Cas9 across all biological kingdoms. The power of CRISPR/Cas has made synthetic biology an inevitability rather than possibility, and actions are already being taken to address ethical issues.

The CRISPR/Cas story begins in 1987 by Nakata and colleagues studying the *iap* gene in *Escherichia coli* (Ishino *et al.*, 1987; Hsu *et al.*, 2014). The researchers reported a set of 29 bp repeats downstream of the *iap* gene that were regularly interspaced by five intervening 32 bp nonrepetitive spacer sequences. This initial finding was later validated by the sequencing of other microbial genomes and the discovery of so-called clustered repeat sequences (Mojica *et al.*, 2000). Clustered repeat sequences were identified in over 40% of sequenced bacteria and 90% of archaea, which led to the term CRISPR being coined in 2002 (Jansen *et al.*, 2002). Around the same time, clusters of CRISPR-associated (*cas*) genes were identified adjacent to the repeat sequences that were conserved with CRISPR loci, providing a basis for classifying three different types of CRISPR systems: type I-III (Haft *et al.*, 2005). Two of these types—type I and III were found to contain multiple *cas* genes that formed complexes with CRISPR RNA (crRNA) to allow recognition and cleavage of target DNA and RNA (Brouns *et al.*, 2008; Hale *et al.*, 2009). The other type—type II only included a few *cas* genes but the function of these *cas* genes remained unknown.

The first breakthrough in the characterization of CRISPR/Cas was made when the spacer sequences separating the CRISPR repeats were found to originate from plasmids and phage DNA (Bolotin *et al.*, 2005). This information was intriguing since other studies showed that CRISPR loci were capable of being transcribed and conferred immunity in archaea carrying spacers corresponding to the viral DNA (Tang *et al.*, 2002;

Mojica *et al.*, 2005). The hypothesis that type II CRISPR spacers direct Cas enzymes to cleave viral DNA through the transcribed spacers as an immune memory and defense system was later confirmed in 2007 in the bacterial strain, *Streptococcus thermophiles* by researchers at the dairy company Danisco (Barrangou *et al.*, 2007).

From there, CRISPR research quickly picked up pace. In just a year, the type II Cas9 nuclease DNA targeting mechanism was revealed by characterization of the essential protospacer adjacent motif (PAM) that directs Cas9 to target DNA and prevents self-cleaving of CRISPR loci (Deveau *et al.*, 2008). This information also suggested that Cas9 was the only Cas protein involved in target DNA cleavage in type II systems and could be used independently of other Cas proteins for biotechnological applications (Garneau *et al.*, 2010).

The final component of the type II CRISPR/Cas mechanism lied with the RNA which guided Cas9 to its DNA targets. Charpentier and colleagues revealed this portion of the mechanism by showing a noncoding *trans*-activating crRNA (tracrRNA) hybridized to the previously characterized crRNA to form a RNA-guiding complex with Cas9 (Deltcheva *et al.*, 2011). This three-component system was later simplified by combining the crRNA and tracrRNA into a single-guide RNA (sgRNA) that could be expressed with Cas9 to cleave target DNA (Jinek *et al.*, 2012; Gasiunas *et al.*, 2012). In 2013, in a pair of studies using the type II CRISPR/Cas system from *Steptococcus thermophiles* and *Streptococcus pyogenes*, genome editing using CRISPR/Cas was first performed in mammalian cells, providing the basis for future development of CRISPR/Cas for genome editing in plants, animals and microbes using strategies never before possible (Mali *et al.*, 2013; Cong *et al.*, 2013).

The Cas9 nuclease has two catalytic domains for DNA cleavage—RuvC and HNH. Each catalytic domain is responsible for nicking a strand of a target double-strand DNA and creating a blunt-ended double-strand break (Nishimasu *et al.*, 2014; Jinek *et al.*, 2014). This is in contrast to the double-strand breaks made by the *Fok*I nuclease which create staggered ends (Li *et al.*, 1992). Cas9 alone with a sgRNA can target a 20-bp DNA sequence for cleavage using homology at the 5' end of the sgRNA with the target sequence, and a PAM sequence at the 3' end of the target sequence (Jinek *et al.*, 2012). Different Cas9 orthologs have different PAM sequences, such as the 5'-NGG and 5'- NGGNG PAM sequences from *Stretrococcus pyogenes* and *Streptococcus thermophilius* CRISPR 3, respectively (Esvelt *et al.*, 2013). Once Cas9 is bound to a PAM, the Cas9-sgRNA complex will begin DNA strand separation using an unknown mechanism and energy source (Anders *et al.*, 2014; Jinek *et al.*, 2014). Cleavage occurs once the sgRNA-DNA heteroduplex is formed and a conformation change in HNH and RvC domains is made. This mechanism has been shown to be very effective for cleaving DNA targets in different biological backgrounds, but is prone to off-targeting due to the utilization of Watson-Crick base-pair recognition of target DNA and the ability of Cas9 to bind all available PAMs.

In order to confront the issue of off-targeting, the Cas9 nuclease has been modified by mutagenizing either the RuvC or HNH domains creating a nickase, or both domains creating a dead Cas9 (dCas9) (Ran *et al.*, 2013; Tsai *et al.*, 2014). The Cas9 nickase provides advantages over wild-type Cas9 nucleases by increasing target recognition from 20 to 40-bp by supplying two sgRNA simultaneously that target opposite strands of adjacent target DNA. This allows a double-strand break to form only

when Cas9 nickases simultaneously bind each target site and nick each stand. Using a strategy similar to the one used for ZFNs and TALENs, dCas9 can be fused to *Fok*I and used to target adjacent target DNA, requiring heterodimerization of FokI for cleavage to occur. Like zinc fingers and TAL effectors, dCas9 can also be fused to other enzymes and extends these benefits by using one optimized protein for multiple DNA targets (Mali *et al.*, 2013; Hsu *et al.*, 2014). Modifications have also been made to the sgRNA that have reduced off-targeting involving sgRNA target site truncation to 17 to 18-bp (Fu *et al.*, 2014). This counterintuitive approach along with the others involving Cas9 have been effective in reducing off-targeting rates, but off-targeting remains a topic of on-going research and a significant consideration for safety.

In the last thirty years, major advances in SSN technology have made genome editing efficient and available to virtually any lab (Kim and Kim, 2014). The robust efficacy of CRISPR/Cas in particular has been startling to even researchers in the field and has raised some important ethical issues about using genome editing for synthetic biology (Cyranoski, 2015; Ledford, 2015; Lanphier *et al.*, 2015). This issue has recently come to a head with the first demonstrations of human embryo modifications using CRISPR/Cas and the potential to take these experiments further to developing germline modified humans (Liang *et al.*, 2015). CRISPR/Cas is also being developed for mutagenic chain reactions (MCR) formally referred to as gene drives in animals, such as *Drosophila melanogaster* where CRISPR/Cas reagents are perpetuated in the germline of animals using mechanisms similar to group I and group II self-splicing introns (Gantz and Bier, 2015; Port *et al.*, 2015). This allows genes critical to procreation, such as gene controlling sex expression in insects to be modified to control

animal populations. The steady progression towards germline embryo manipulation, gene drives and the uncertainty of the off-target effects warrants further study of plants and animals modified using genome editing technologies.

Genome editing in plants using SSNs

Targeted mutagenesis using SSNs

Targeted mutagenesis or targeted gene knock-outs can be accomplished in living plant cells by inducing double-strand breaks and facilitating formation of insertions and deletions at the break site via nonhomologous end-joining (NHEJ) (Gorbunova and Levy, 1999; Puchta, 2005) (Figure 1A). Certain radiation and chemicals, such as ethyl methanesulfonate (EMS) are capable of mutagenizing DNA in a random manner but cause severe off-targeting (Wang *et al.*, 2012; Tsai *et al.*, 2011). As an alternative to random mutagenesis, SSNs can be used to direct DNA damage to the coding or promoter sequences of target genes and minimize off-targeting effects (Voytas and Gao, 2014). The reduction of off-targeting and improved on-targeting mutagenesis allows for more efficient recovery and identification of modified cells that can be regenerated in plants as modified events.

The first successful SSN-mediated targeted mutagenesis experiments in plants were done using ZFNs (Carroll, 2011). Although homing endonucleases, such as I-SceI and I-CreI were available years before ZFNs, their limited efficiency in plant cells for targeted mutagenesis prevented their early success, but have been demonstrated more recently in Arabidopsis *(Arabidopsis thaliana*) and maize (Gao *et al.*, 2010; Antunes *et*

al., 2012). Nevertheless, ZFNs were demonstrated to be effective in Arabidopsis, tobacco (*Nicotiana tabacum*), maize and soybean (Osakabe *et al.*, 2011; Townsend *et al.*, 2009; Curtin *et al.*, 2011; Lloyd *et al.*, 2005) (Table 1). Strategies for delivering ZFNs included induced ZFN expression in seedlings in which ZFN reagents had been integrated within the genome and to protoplast and cell cultures (Osakabe *et al.*, 2011; Lloyd *et al.*, 2005). Induction of ZFNs was used to reduce toxicity associated with constitutive ZFN expression in stable lines and facilitate high ZFN expression in plant tissues (Zhang *et al.*, 2010). Protoplast and cell culture-mediated delivery was particularly effective for preliminary evaluation of ZFN reagents that were ultimately used for other genome editing applications (Townsend *et al.*, 2009; Wright *et al.*, 2005). Consequently, these transient transformation methods were less effective for recovering regenerated plants and their use demonstrated the difficulty of generating and detecting ZFN-mediated targeted mutations in plants.

The approaches required to generate and detect ZFN-mediated targeted mutations in plants suggested insufficient ZFN expression was limiting the formation of NHEJ-mediated targeted mutations. This observation was supported by the fact that many more copies of ZFN constructs can be delivered to individual cells in protoplast and cell cultures as opposed to *Agrobacterium*-mediated delivery (Alonso *et al.*, 2003; Freeman *et al.*, 1984). To improve ZFN delivery, an approach was developed employing the *Tobacco Rattle Virus* (*TRV*) (Marton *et al.*, 2010). *TRV* is a RNA-virus that was previously developed for Virus-Induced Gene Silencing (VIGS) and delivery of RNAi reagents (Sha *et al.*, 2013; Brigneti *et al.*, 2004). The delivery of ZFNs and sgRNA for CRISPR/Cas using *TRV* was effective in inducing heritable targeted mutations in a

reporter construct within tobacco and petunia (*Petunia hybrida*) lines and endogenous loci of constitutively expressing Cas9 tobacco lines, respectively (Ali *et al.*, 2015). The use of *TRV* remains the most efficient method of transiently delivering ZFNs for the purpose of generating targeted mutations in plant tissues capable of regeneration. However, these *TRV*-mediated targeted mutations and others using conventional transformation must be fixed into the germline by selfing or cross.

An alternative approach to recovering plants with targeted mutations is to direct integration of exogenous DNA to a double-strand break (Tzfira *et al.*, 2003; Chilton and Que, 2003; Salomon and Puchta, 1998). Previously described methods rely on the efficiency of SSNs to make double-strand breaks and induce NHEJ targeted mutations in tissues capable of regeneration (Puchta, 2005). However, directed integration of transfer DNA (T-DNA) from *Agrobacterium* can also be used to “knock-in” new sequence that can be selected for and “knock-out” endogenous gene function or “stack” new transgenes at a specified locus, so called “trait stacking” (Ainley *et al.*, 2013; D’Halluin *et al.*, 2013). This approach is particularly useful for the biotech industry and GM crop development where released GM crop lines have an already characterized locus. Variations of this approach include creating SSN target sites flanking the exogenous DNA, creating staggered ends for insertion, and removing integrated transgenes (Weinthal *et al.*, 2013; Petolino *et al.*, 2010). Targeted gene insertion and “trait stacking” has been demonstrated in a number of plant species but still relies on the introduction of transgenic sequence.

TALENs provided the first breakthrough for targeted mutagenesis in plants with an eloquent study in rice (*Oryza sativa*) (Li *et al.*, 2012). Elucidation of the molecular

code underlying TAL effector binding not only allowed for new TAL effectors to be constructed, but also it allowed gene targets of endogenous TAL effectors to be predicted (Noël *et al.*, 2013). One endogenous TAL effector, AvrXa7 was known to be a major virulence factor of *Xanthamonas oryzae* causing bacterial blight by targeting a discrete promoter region of *Os11N3* (or *OsSWEET14)* susceptibility gene in rice (Kay and Bonas, 2009). With this information, the researchers constructed new TALENs that targeted *OsSWEET14* promoter region to introduce targeted mutations and disrupt binding of the endogenous AvrXa7 TAL effector. Targeted mutations in the *OsSWEET14* promoter were discovered in both callus lines and T1 progeny of transformed rice. These mutations ranged from mono- to bi-allelic and were capable of conferring resistance to *X. oryae*. This was the first study in which targeted mutations of an endogenous target were recovered in primary transformed events without extensive screening or inserting new exogenous sequence for selection while supporting a clear phenotype. Following this study, a handful of other examples of TALEN-mediated mutagenesis were reported in primary monocots such as wheat (*Triticum aestivum*), barley (*Hordeum vulgare*), rice and maize (Wendt *et al.*, 2013; Shan *et al.*, 2013a; Liang *et al.*, 2014; Char *et al.*, 2015; Wang *et al.*, 2014). However, these examples were greatly outweighed by later reports using CRISPR/Cas.

The immediate success of CRISPR/Cas for targeted mutagenesis in plants can be explained by a number of factors but ultimately lies with the robust nature of CRISPR/Cas (Belhaj *et al.*, 2013). The ease of sgRNA design and the use of a single Cas9 nuclease allowed labs to quickly adopt the technology and test CRISPR/Cas reagents in their species of choice. This resulted in a “tsunami” of studies using

CRISPR/Cas in plants starting with the first reports in 2013 (Nekrasov *et al.*, 2013; Li *et al.*, 2013; Shan *et al.*, 2013b).

The first plants to be modified using CRISPR/Cas were Arabidopsis, *Nicotiana benthamiana* and rice. These first reports demonstrated CRISPR/Cas could be used to form targeted mutations in regenerated plants with high efficiency and make modifications that support a clear phenotype, with an example in rice (Shan *et al.*, 2013b). These initial studies were followed up on with extensive investigations into the multigenerational inheritance of CRISPR/Cas-mediated targeted mutations using Arabidopsis (Feng *et al.*, 2014). These studies revealed that most primary events with Cas9 integrated had chimeric targeted mutations in somatic tissues and required to be taken to further generations to fix specific mutations. The issue of chimerism was not new to genome editing and demonstrated the efficacy of CRISPR/Cas reagents in primary plant materials and the need to "segregate out" Cas9 to fix mutations (Marton *et al.*, 2010; Zhang *et al.*, 2010; Curtin *et al.*, 2011).

The approach of generating and fixing mutations generated by SSNs in later generations has become a standard practice for genome editing in plants, and was most effectively demonstrated in a landmark study in wheat (Wang *et al.*, 2014). In this study, TALENs were used to target the *MLO* genes of hexaploid wheat (*Triticum aestivum* L., 2*n* = 42, AABBDD) which confer susceptibility to powdery mildew. Although natural loss-of-function *MLO* alleles exist in the A and B genomes, no known source of loss-of-function *MLO* alleles had been discovered in the D genome. By targeting the MLO genes of all three genomes, mono- and bi-allelic loss-of-function alleles were created in all three genomes and were fixed to homozygosity through selfing. Complete

MLO knock-outs showed remarkable resistance to powdery mildew which echoed the previous described study using TALENs for bacterial blight in rice (Li *et al.*, 2012). The success of these studies led to reports of CRISPR/Cas in other crop species, including tomato (*Solanum lycopersicum*) and maize (Brooks *et al.*, 2014; Liang *et al.*, 2014).

Homologous recombination using SSNs

The ultimate goal of genome editing is to make directed modifications at specified locations within the genome. Historically, this has been accomplished in microbes such as yeast and *E. coli* using homologous recombination (Smith, 1989; Orr-Weaver *et al.*, 1981). Homologous recombination or gene targeting is a fundamental biological process with roles in both meiosis and somatic cells for use in high fidelity DNA repair (Schuermann *et al.*, 2005). Gene targeting for uses in genome editing relies on the utilization of an exogenous DNA template for DNA repair (Puchta *et al.*, 1996; Puchta *et al.*, 1993; Salomon and Puchta, 1998) (Figure 1B). For gene targeting to occur, homologous sequence on either side of a desired modification or new sequence of a repair template is used in an intricate DNA repair pathway to incorporate the new sequence (Pacher *et al.*, 2007; Siebert and Puchta, 2002). In yeast, gene targeting occurs at relatively high frequencies without the need to induce DNA damage at target sites (Wu and Lichten, 1995). However, in higher eukaryotes, such as plants, gene targeting occurs at much lower frequencies at a rate of approximately one in 10^5 to 10^7 transfected cells (Wright *et al.*, 2005). The use of SSNs to induce DNA damage at target sites has shown to be an effective strategy in plants and animals to improve gene targeting efficiencies (Carroll, 2004; Puchta, 2005).

The first reports of gene targeting in plants using SSNs predate reports of targeted mutagenesis due to the ability to select for new sequence incorporated within the target site (Puchta *et al.*, 1993; Salomon and Puchta, 1998; Puchta *et al.*, 1996). Once the first SSNs became available, Hohn and colleagues conducted gene targeting experiments using the homing endonuclease, I-SceI to target a broken reporter integrated within tobacco for gene targeting (Puchta *et al.*, 1996). Once a double-strand break was formed within the reporter, a repair template was provided which was capable of repairing the broken reporter and conferring resistance to selection. Resistant lines were screened for modifications and gene targeting frequencies were found to range from 0.6% to 18.8% of the evaluated lines. The observed frequencies were higher than previous attempts which only repair templates were provided, supporting later studies (Puchta and Fauser, 2013; Paszkowski *et al.*, 1988). These early experiments demonstrated gene targeting could be improved in plants using SSNs but the number of available SSNs were lacking at the time.

The application of ZFNs for gene targeting proved to be more useful than previously described attempts to use ZFNs for targeted mutagenesis in plants (Petolino, 2015). Reports from Arabidopsis, tobacco and maize were successful in performing gene targeting using ZFNs in plant protoplasts and regenerated events (Shukla *et al.*, 2009; Townsend *et al.*, 2009; Qi *et al.*, 2013b) (Table 1). One study, in particular was effective for both generating gene targeting regenerated events and for quantifying gene targeting frequencies in plant cells (Wright *et al.*, 2005). In this study, Wright and colleagues used a reporter in which the “GUS:NptII” protein fusion coding sequence was broken by a 600 bp deletion and a target sequence for the ZFN, Zif268. The

reporter was capable of being repaired with a template which included the missing sequence and homology on either side of the deletion. Once gene targeting had occurred, gene targeting events could be detected using GUS activity and regenerated using kanamycin selection. The reporter was integrated within lines of tobacco and used for protoplast transformation. Gene targeting frequencies were improved three-fold when both repair template and ZFN were delivered compared to only the repair template. This improved efficiency resulted in gene targeting in approximately one in 10^3 cells. Although an improvement, these efficiencies are only a modest improvement from previous attempts using only repair templates and underscore the importance of repair template delivery for efficient gene targeting.

The mechanism of homologous recombination requires a repair template is available at the time DNA damage is formed (Schuermann *et al.*, 2005). In somatic cells, the repair template is typically provided by a sister chromatid which shares extensive homology with the broken chromosome (Vu *et al.*, 2014). However, in gene targeting, repair template homology with the broken chromosome is limited to what can be delivered efficiently to a cell. In plants, repair templates can be efficiently delivered using protoplasts and cell cultures (Paszkowski *et al.*, 1988). Furthermore, the limited repair template size can be compensated for by using protoplasts and delivering a high volume of repair template DNA to the cell (Freeman *et al.*, 1984). This approach has proven to be effective for enhancing rates of gene targeting and improving detection in plants cells (Wright *et al.*, 2005; Townsend *et al.*, 2009). However, protoplasts can only be isolated and regenerated from a select number of plant species (Davey *et al.*, 2005).

To address this issue, Voytas and colleagues have used a plant geminivirus to deliver gene targeting reagents to tobacco cells (Baltes *et al.*, 2014). The *Bean Yellow Dwarf Virus* (*BeYDV*) used for the study replicates through a single-strand DNA intermediate and exists as a double-strand DNA virus in host plant cells (Liu *et al.*, 1998; Liu *et al.*, 1997). The relatively few essential viral elements made it possible to construct an *Agrobacterium* T-DNA that was capable of delivering the geminivirus to plant cells and allow the virus to replicate to a high-copy number within the plant nucleus. In order to compare gene targeting efficiencies, the same GUS:NptII tobacco reporter lines used by Wright and colleagues was used for *Agrobacterium*-mediated transformation using the geminivirus and a conventional T-DNA reference carrying the same Zif268 SSN and repair template. Using this system, it was found that the geminivirus was capable of increasing gene targeting frequencies by 80-fold compared to the conventional T-DNA. Efficient gene targeting was found to be dependent on both SSN activity and geminivirus replication of the repair template, and was capable of creating modified regenerated events. This exciting new approach combines the benefits of *Agrobacterium*-mediated transformation with the enhanced gene targeting efficiencies seen with protoplasts and has great potential for genome editing in crop species.

Genome editing in potato

Potato (*Solanum tuberosum* Group Tuberosum L.) provides an ideal opportunity for developing methology and reagents for genome editing in a crop species. Being

among the first crops to be genetically engineered, potato is highly amendable to tissue culture and can be regenerated from leaf and stem tissues with high efficiency (Banerjee *et al.*, 2006; Chakravarty *et al.*, 2007). *Agrobacterium*-mediated transformation has been established in a number of potato cultivars but other delivery methods have also been successful (Barampuram and Zhang, 2011).

A key advantage of using potato for evaluating methods of genome editing is the ability to perform both clonal propagation and germline crosses with diploid and tetraploid genotypes (Hanneman, 1989). Most crop species are polyploids and require vigorous breeding efforts to combine complex traits into a single cultivar (Comai, 2005). Potato shares this limitation as most US cultivars are tetraploid (Hirsch *et al.*, 2013). However, a wealth of diploid germplasm exists in potato, such as the doubled-monoploid (DM) used to sequence the potato genome that can be used for plant transformation (PGSC, 2011; Paz and Veilleux, 1999; Hanneman, 1989). Hence, genetically engineered diploid and tetraploid potato genotypes can be clonally propagated to study the phenotypic effects of genome editing and also crossed to determine the heritability of modifications.

The release of the potato genome has accelerated potato research and has provided vital tools for functional genomics in potato (PGSC, 2011). The potato genome website allows users to search the latest genome assembly for genes of interest and identify sequences that share identity or similarity with a given target (Hirsch *et al.*, 2014; Sharma *et al.*, 2013). The ability to search the potato genome for potentially redundant sequence or gene copies provides vital information for designing genome editing reagents (Curtin *et al.*, 2012; Voytas and Gao, 2014; Liu *et al.*, 2013). For

example, if a given gene was being targeted for mutagenesis and complete knock-out, all gene copies would need to be identified and mutagenized within a single regenerated event. This would involve identifying a single or multiple target sites that are conserved across each gene copy that could be used as a target site(s) for designing SSN reagents (Curtin *et al.*, 2011). This is a common issue in plants given the occurrence of highly duplicated genomes (Cui *et al.*, 2006). CRISPR/Cas provides a key advantage for genome editing in plant species by being capable of targeting multiple target sites using a single nuclease (Cong *et al.*, 2013). Nevertheless, examples of targeted mutagenesis and other forms of genome editing are limited in potato and other crop species and warrant further development.

Dissertation organization and objectives

The remaining dissertation is organized into two research chapters and a conclusion chapter. The objectives of chapter 2 entitled, “Generation and inheritance of targeted mutations in potato (*Solanum tuberosum* L.) using the CRISPR/Cas system,” are as follows:

1. Design and assemble CRISPR reagents targeting the potato *ALS1* gene.
2. Generate T_0 events expressing *ALS1* CRISPR reagents, screen events and characterize targeted mutations.
3. Determine inheritance of CRISPR targeted mutations through clonal and germline generations.

The objectives of chapter 3 entitled, “Geminivirus-mediated genome editing in potato (*Solanum tuberosum* L.) using sequence-specific nucleases,” are as follows:

1. Design and assemble TALENs and CRISPR reagents targeting the Zif268 target site and TALENs targeting the potato *ALS1* gene.
2. Compare ZFN, TALEN and CRISPR nuclease efficiencies in potato cells using a single-strand annealing assay (SSA).
3. Compare conventional T-DNA and geminivirus-mediated gene targeting efficiencies in potato cells using the GUPTII reporter system.
4. Design and assemble potato *ALS1* repair template for gene targeting that include modifications conferring herbicide and kanamycin resistance.
5. Generate T_0 events transformed with *ALS1* CRISPR and TALEN gene targeting reagents, screen events and characterize modifications.
6. Assess T_0 events for herbicide resistance phenotype using an herbicide spray assay.

The conclusion chapter 4 summaries the findings of the dissertation with other examples of genome editing in potato and provides prospects for future genome editing research in potato.

APPENDICES

Appendix A: Tables

Table 1. List of plant species modified using genome editing. Plant species (Species), *Arabidopsis thaliana* (Arabidopsis), *Oryza sativa* (Rice), *Nicotiana tabacum* (Tobacco) and Zea mays (Maize) are listed with examples of genome editing using meganuclease (Meganuc.), zinc finger nuclease (ZFN), transcription activator-like effector nuclease (TALEN), and clustered regularly interspaced short palindromic repeats/CRISPR-associated systems (CRISPR) sequence-specific nucleases (SSN platform). Demonstration of nonhomologous end-joining (NHEJ) and gene targeting (GT) is indicated (Repair type) with NHEJ efficiencies provided for each reference (Reference). The transformation method (Transformation method) and the demonstration of germ-line mutations (Germline) used for NHEJ is also provided.

Species	SSN platform	Repair type	Efficiency (NHEJ)	Transformation method (NHEJ)	Germ-line	Reference
Arabidopsis	Meganuc.	NHEJ	34%	*Agrobacterium*	Yes	Antunes *et al.*, 2012
Arabidopsis	ZFN	NHEJ	7.9%	*Agrobacterium*	Yes	Lloyd *et al.*, 2005
Arabidopsis	TALEN	NHEJ	N/A	Protoplasts	No	Cermak *et al.*, 2011
Arabidopsis	CRISPR	NHEJ, GT	5.6%	Protoplasts	No	Li *et al.*, 2013
Rice	TALEN	NHEJ	48%	*Agrobacterium*	Yes	Li *et al.*, 2012
Rice	CRISPR	NHEJ	50%	*Agrobacterium*	Yes	Mao *et al.*, 2013
Tobacco	Meganuc.	GT	0.68%	*Agrobacterium*	Yes	Fauser *et al.*, 2012
Tobacco	ZFN	NHEJ, GT	0.39%	Protoplasts	Yes	Townsend *et al.*, 2009
Tobacco	TALEN	NHEJ, GT	32%	Protoplasts	No	Zhang *et al.*, 2013
Tobacco	CRISPR	NHEJ, GT	38%	Protoplasts	No	Li *et al.*, 2013
Maize	Meganuc.	NHEJ, GT	3%	*Agrobacterium*	Yes	Gao *et al.*, 2010
Maize	ZFN	NHEJ, GT	0.05%	Cell culture	Yes	Shukla *et al.*, 2009
Maize	TALEN	NHEJ	10%	*Agrobacterium*	Yes	Char *et al.*, 2015
Maize	CRISPR	NHEJ	13%	Protoplast	No	Liang *et al.*, 2014

Appendix B: Figures

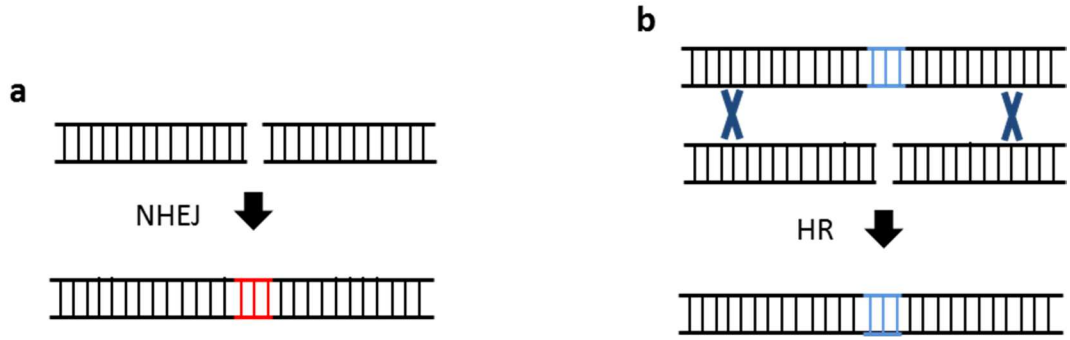

Figure 1. DNA repair pathways for repairing double-strand breaks. Induction of double-strand breaks may result in **a**) error prone nonhomologous end-joining (NHEJ) and in the introduction of random mutations at the break site (red segment) or **b**) homologous recombination (HR) and the introduction of new or modified sequence (blue segment) in the presence of a repair template.

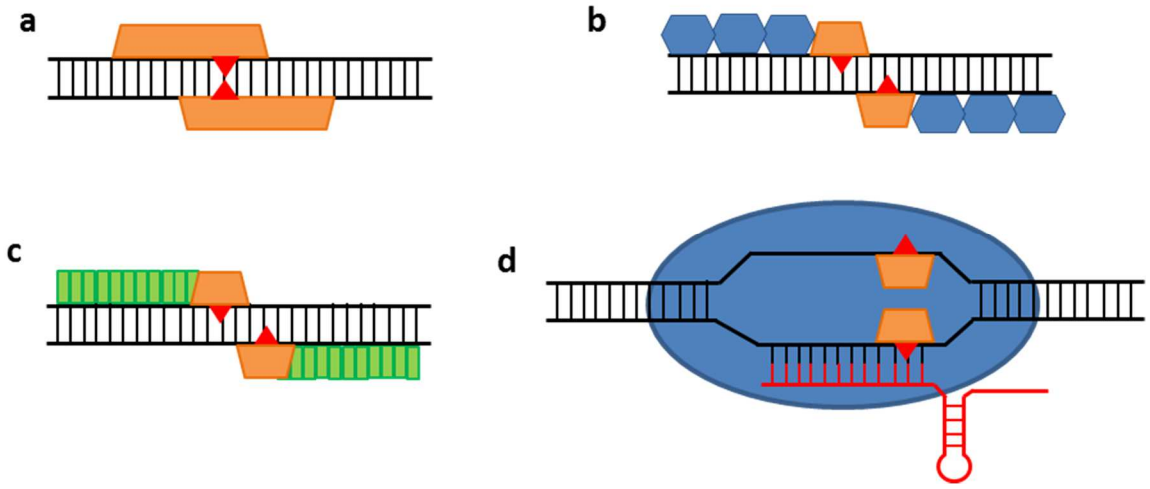

Figure 2. **Schematic of sequence-specific nuclease platforms used for genome editing**: **a**) homing endonucleases (meganuclease), **b**) zinc finger nucleases (ZFNs), **c**) transcription activator-like effector nucleases (TALENs), and **d**) clustered regularly interspaced short palindromic repeats/CRISPR-associated system (CRISPR/Cas). Red triangles indicate cleavage positions which create double-strand breaks and the red strand in d) represents a single-guide RNA. All other colored shapes represent DNA binding proteins.

REFERENCES

REFERENCES

Ainley, W.M., Sastry-Dent, L., Welter, M.E., et al. (2013) Trait stacking via targeted genome editing. *Plant Biotechnol. J.*, **11**, 1126–34. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23953646.

Ali, Z., Abul-faraj, A., Li, L., et al. (2015) Efficient virus-mediated genome editing in plants using the CRISPR/Cas9 system. *Mol. Plant*. Available at: http://dx.doi.org/10.1016/j.molp.2015.02.011.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., et al. (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science*, **301**, 653–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12893945.

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. Available at: http://dx.doi.org/10.1038/nature13579.

Andersen, M.M., Landes, X., Xiang, W., et al. (2015) Feasibility of new breeding techniques for organic farming. *Trends Plant Sci.*, **20**, 426–434. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1360138515001120.

Antunes, M.S., Smith, J.J., Jantz, D. and Medford, J.I. (2012) Targeted DNA excision in *Arabidopsis* by a re-engineered homing endonuclease. *BMC Biotechnol.*, **12**, 1–12. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3536558&tool=pmcentrez&rendertype=abstract.

Baker, M. (2011) Method of the Year 2011. *Nat. Methods*, **9**, 1–1. Available at: https://www.youtube.com/watch?v=zDkUFzZoQAs.

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. Available at: http://www.plantcell.org/cgi/doi/10.1105/tpc.113.119792.

Banerjee, A.K., Prat, S. and Hannapel, D.J. (2006) Efficient production of transgenic potato (*S. tuberosum* L. *ssp.* andigena) plants via *Agrobacterium tumefaciens*-mediated transformation. *Plant Sci.*, **170**, 732–738. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0168945205004127.

Barampuram, S. and Zhang, Z.J. (2011) Recent advances in plant transformation J. A. Birchler, ed. *Methods Mol. Biol.*, **701**, 1–35. Available at: http://www.springerlink.com/index/10.1007/978-1-61737-957-4.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A. and Horvath, P. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science.*, **315**, 1709–1712. Available at: http://www.sciencemag.org/cgi/doi/10.1126/science.1138140.

Barton, K.A., Binns, A.N., Matzke, A.J. and Chilton, M.D. (1983) Regeneration of intact tobacco plants containing full length copies of genetically engineered T-DNA, and transmission of T-DNA to R1 progeny. *Cell*, **32**, 1033–43. Available at: http://www.ncbi.nlm.nih.gov/pubmed/6301678.

Bedell, V.M., Wang, Y., Campbell, J.M., et al. (2012) *In vivo* genome editing using a high-efficiency TALEN system. *Nature*. Available at: http://www.nature.com/doifinder/10.1038/nature11537.

Belfort, M. and Perlman, P.S. (1995) Mechanisms of intron mobility. *J. Biol. Chem.*, **270**, 30237–30240.

Belhaj, K., Chaparro-Garcia, A., Kamoun, S. and Nekrasov, V. (2013) Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods*, **9**, 39. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3852272&tool=pmcentrez&rendertype=abstract.

Bitinaite, J., Wah, D.A., Aggarwal, A.K. and Schildkraut, I. (1998) FokI dimerization is required for DNA cleavage. *PNAS*, **95**, 10570–10575.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A. and Bonas, U. (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*, **326**, 1509–1512. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19933107.

Bolotin, A., Quinquis, B., Sorokin, A. and Dusko Ehrlich, S. (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, **151**, 2551–2561.

Borlaug, N.E. (2000) Ending world hunger. The promise of biotechnology and the threat of antiscience zealotry. *Plant Physiol.*, **124**, 487–490.

Brigneti, G., Martín-Hernández, A.M., Jin, H., Chen, J., Baulcombe, D.C., Baker, B. and Jones, J.D.G. (2004) Virus-induced gene silencing in Solanum species. *Plant J.*, **39**, 264–272. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15225290.

Broitman, S.L., Im, D.D. and Fresco, J.R. (1987) Formation of the triple-stranded polynucleotide helix, poly(A.A.U). *PNAS*, **84**, 5120–5124.

Brooks, C., Nekrasov, V., Lippman, Z.B. and Eck, J. Van (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.

Brouns, S.J.J., Jore, M.M., Lundgren, M., et al. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*, **321**, 960–964. Available at: http://www.sciencemag.org/cgi/doi/10.1126/science.1159689.

Camacho, A., Deynze, A. Van, Chi-Ham, C. and Bennett, A.B. (2014) Genetically engineered crops that fly under the US regulatory radar. *Nat. Biotechnol.*, **32**, 1087–1091. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25380439.

Carroll, D. (2011) Genome engineering with zinc-finger nucleases. *Genetics*, **188**, 773–782.

Carroll, D. (2004) Using nucleases to stimulate homologous recombination. In *Genetic Recombination*. New Jersey: Humana Press, pp. 195–208. Available at: http://link.springer.com/10.1385/1-59259-761-0:195.

Cermak, T., Doyle, E.L., Christian, M., et al. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.*, **39**, e82. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3130291&tool=pmcentrez&rendertype=abstract.

Chakravarty, B., Wang-Pruski, G., Gustafson, V. and Regan, S. (2007) Genetic transformation in potato: approaches and strategies. *Am. J. Potato Res.*, **84**, 301–311.

Chan, Y.S., Huen, D.S., Glauert, R., Whiteway, E. and Russell, S. (2013) Optimising homing endonuclease gene drive performance in a semi-refractory species: the *Drosophila melanogaster* experience. *PLoS One*, **8**, e54130.

Char, S.N., Unger-Wallace, E., Frame, B., Briggs, S.A., Main, M., Spalding, M.H., Vollbrecht, E., Wang, K. and Yang, B. (2015) Heritable site-specific mutagenesis using TALENs in maize. *Plant Biotechnol. J.*, n/a–n/a. Available at: http://doi.wiley.com/10.1111/pbi.12344.

Chilton, M. and Que, Q. (2003) Targeted integration of T-DNA into the tobacco genome at double-stranded breaks: new insights on the mechanism of T-DNA integration. *Plant Physiol.*, **133**, 956–965. Available at: http://www.plantphysiol.org/content/133/3/956.short.

Cho, J., Parks, M.E. and Dervan, P.B. (1995) Cyclic polyamides for recognition in the minor groove of DNA. *PNAS*, **92**, 10389–10392.

Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J. and Voytas, D.F. (2010) Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*, **186**, 757–761. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2942870&tool=pmcentrez&rendertype=abstract.

Christou, P. (2013) Plant genetic engineering and agricultural biotechnology 1983-2013. *Trends Biotechnol.*, **31**, 125–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23375945.

Clasen, B.M., Stoddard, T.J., Luo, S., et al. (2015) Improving cold storage and processing traits in potato through targeted gene knockout. *Plant Biotechnol. J.*, n/a–n/a. Available at: http://doi.wiley.com/10.1111/pbi.12370.

Comai, L. (2005) The advantages and disadvantages of being polyploid. *Nat. Rev. Genet.*, **6**, 836–846. Available at: http://www.nature.com/doifinder/10.1038/nrg1711.

Cong, L., Ran, F.A., Cox, D., et al. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–23. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23287718.

Cornu, T.I., Thibodeau-Beganny, S., Guhl, E., Alwin, S., Eichtinger, M., Joung, J.K. and Cathomen, T. (2008) DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Mol. Ther.*, **16**, 352–358.

Cui, L., Wall, P.K., Leebens-Mack, J.H., et al. (2006) Widespread genome duplications throughout the history of flowering plants. *Genome Res.*, **16**, 738–749.

Curtin, S.J., Voytas, D.F. and Stupar, R.M. (2012) Genome engineering of crops with designer nucleases. *Plant Genome J.*, **5**, 42–50. Available at: https://www.crops.org/publications/tpg/abstracts/5/2/42.

Curtin, S.J., Zhang, F., Sander, J.D., et al. (2011) Targeted mutagenesis of duplicated genes in soybean with zinc-finger nucleases. *Plant Physiol.*, **156**, 466–473. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3177250&tool=pmcentrez&rendertype=abstract.

Cyranoski, D. (2015) Ethics of embryo editing divides scientists. *Nature*, **519**, 272–272. Available at: http://www.nature.com/doifinder/10.1038/519272a.

D'Halluin, K., Vanderstraeten, C., Hulle, J. Van, et al. (2013) Targeted molecular trait stacking in cotton through targeted double-strand break induction. *Plant Biotechnol. J.*, **11**, 933–941.

Davey, M.R., Anthony, P., Power, J.B. and Lowe, K.C. (2005) Plant protoplasts: Status and biotechnological perspectives. *Biotechnol. Adv.*, **23**, 131–171.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J. and Charpentier, E. (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, **471**, 602–607.

Deng, D., Yan, C., Pan, X., Mahfouz, M., Wang, J., Zhu, J.-K., Shi, Y. and Yan, N. (2012) Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science*, **335**, 720–723. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22223738.

Deveau, H., Barrangou, R., Garneau, J.E., Labonté, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P. and Moineau, S. (2008) Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.*, **190**, 1390–1400.

Doudna, J.A. and Charpentier, E. (2014) The new frontier of genome engineering with CRISPR-Cas9. *Science*, **346**, 1077–1093.

Doyle, E.L., Booher, N.J., Standage, D.S., Voytas, D.F., Brendel, V.P., Vandyk, J.K. and Bogdanove, A.J. (2012) TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic Acids Res.*, **40**, 117–122.

Doyle, E.L., Stoddard, B.L., Voytas, D.F. and Bogdanove, A.J. (2013) TAL effectors: highly adaptable phytobacterial virulence factors and readily engineered DNA-targeting proteins. *Trends Cell Biol.*, **23**, 390–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23707478.

Ellis, B.L., Hirsch, M.L., Porter, S.N., Samulski, R.J. and Porteus, M.H. (2013) Zinc-finger nuclease-mediated gene correction using single AAV vector transduction and enhancement by Food and Drug Administration-approved drugs. *Gene Ther.*, **20**, 35–42. Available at: http://dx.doi.org/10.1038/gt.2011.211.

Esvelt, K.M., Mali, P., Braff, J.L., Moosburner, M., Yaung, S.J. and Church, G.M. (2013) Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods*, **10**, 1116–23. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24076762.

Feng, Z., Mao, Y., Xu, N., et al. (2014) Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in *Arabidopsis*. *PNAS*, **111**, 4632–4637. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1400822111.

Freeman, J.P., Draper, J., Davey, M.R., Cocking, E.C., Gartland, K.M. a, Harding, K. and Pental, D. (1984) A comparison of methods for plasmid delivery into plant protoplasts. *Plant Cell Physiol.*, **25**, 1353–1365.

Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M. and Joung, J.K. (2014) Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.*, **32**, 279–284. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24463574.

Gaj, T., Mercer, A.C., Sirk, S.J., Smith, H.L. and Barbas, C.F. (2013) A comprehensive approach to zinc-finger recombinase customization enables genomic targeting in human cells. *Nucleic Acids Res.*, **41**, 3937–3946. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23393187.

Gantz, V.M. and Bier, E. (2015) The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. *Science*, **348**, 442–444. Available at: http://www.sciencemag.org/cgi/doi/10.1126/science.aaa5945.

Gao, H., Smith, J., Yang, M., et al. (2010) Heritable targeted mutagenesis in maize using a designed endonuclease. *Plant J.*, **61**, 176–187. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19811621.

Garneau, J.E., Dupuis, M.-È., Villion, M., et al. (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, **468**, 67–71.

Gasiunas, G., Barrangou, R., Horvath, P. and Siksnys, V. (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *PNAS*, **109**, E2579–E2586.

Gorbunova, V. and Levy, A.A. (1999) How plants make ends meet: DNA double-strand break repair. *Trends Plant Sci.*, **4**, 263–269.

Grizot, S., Duclert, A., Thomas, S., Duchateau, P. and Pâques, F. (2011) Context dependence between subdomains in the DNA binding interface of the I-CreI homing endonuclease. *Nucleic Acids Res.*, **39**, 6124–36. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3152339&tool=pmcentrez&rendertype=abstract.

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**, 0474–0483.

Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., Wells, L., Terns, R.M. and Terns, M.P. (2009) RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell*, **139**, 945–956. Available at: http://dx.doi.org/10.1016/j.cell.2009.07.040.

Hanneman, R. (1989) The potato germplasm resource. *Am. Potato J.*, **66**, 655–667. Available at: http://link.springer.com/article/10.1007/BF02853985.

Herrera-Estrella, L., Depicker, A., Montagu, M. Van and Schell, J. (1983) Expression of chimaeric genes transferred into plant cells using a Ti-plasmid-derived vector. *Nature*, **303**, 209–13. Available at: http://adsabs.harvard.edu/abs/1983Natur.303..209H.

Hirsch, C.D., Hamilton, J.P., Childs, K.L., et al. (2014) Spud DB: a resource for mining sequences, genotypes, and phenotypes to accelerate potato breeding. *Plant Genome*, **7**. Available at: https://www.crops.org/publications/tpg/abstracts/7/1/plantgenome2013.12.0042.

Hirsch, C.N., Hirsch, C.D., Felcher, K., et al. (2013) Retrospective view of North American potato (*Solanum tuberosum* L.) breeding in the 20th and 21st centuries. *G3*, **3**, 1003–1013. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23589519.

Hsu, P.D., Lander, E.S. and Zhang, F. (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, **157**, 1262–1278. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0092867414006047.

Hunter, P. (2014) “Genetically Modified Lite” placates public but not activists: new technologies to manipulate plant genomes could help to overcome public concerns about GM crops. *EMBO Rep.*, **15**. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24451237.

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.

Jacquier, A. and Dujon, B. (1985) An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell*, **41**, 383–394.

James, C. (2014) *Global status of commercialized biotech/GM Crops*, Available at: http://www.isaaa.org/resources/publications/briefs/41/default.asp.

Jansen, R., Embden, J.D. Van, Gaastra, W. and Schouls, L.M. (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.*, **43**, 1565–1575.

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. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22745249.

Jinek, M., Jiang, F., Taylor, D.W., et al. (2014) Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science*, **343**, 1247997. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24505130.

Joung, J.K., Voytas, D.F. and Cathomen, T. (2010) Reply to "Successful genome editing with modularly assembled zinc finger nucleases." *Nat. Methods*, **7**, 91–92.

Juillerat, A., Dubois, G., Valton, J., et al. (2014) Comprehensive analysis of the specificity of transcription activator-like effector nucleases. *Nucleic Acids Res.*, **42**, 5390–402. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4005648&tool=pmcentrez&rendertype=abstract.

Juillerat, A., Pessereau, C., Dubois, G., et al. (2015) Optimized tuning of TALEN specificity using non-conventional RVDs. *Sci. Rep.*, **5**, 8150. Available at: http://www.nature.com/doifinder/10.1038/srep08150.

Jurica, M.S., Monnat, R.J. and Stoddard, B.L. (1998) DNA recognition and cleavage by the LAGLIDADG homing endonuclease I-CreI. *Mol. Cell*, **2**, 469–476.

Kaniewski, W.K. and Thomas, P.E. (2004) The potato story. *AgBioForum*, **7**, 41–46.

Kay, S. and Bonas, U. (2009) How Xanthomonas type III effectors manipulate the host plant. *Curr. Opin. Microbiol.*, **12**, 37–43. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19168386.

Kim, H. and Kim, J.-S. (2014) A guide to genome engineering with programmable nucleases. *Nat. Rev. Genet.*, **15**, 321–34. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24690881.

Kim, Y.G., Cha, J. and Chandrasegaran, S. (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *PNAS*, **93**, 1156–1160. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=40048&tool=pmcentrez&rendertype=abstract.

Kim, Y.G. and Chandrasegaran, S. (1994) Chimeric restriction endonuclease. *PNAS*, **91**, 883–887.

Lanphier, E., Urnov, F., Haecker, S.E., Werner, M. and Smolenski, J. (2015) Don't edit the human germ line. *Nature*, **519**, 410–411. Available at: http://www.nature.com/doifinder/10.1038/519410a.

Ledford, H. (2015) CRISPR, the disruptor. *Nature*, **522**, 20–24. Available at: http://www.nature.com/doifinder/10.1038/522020a.

Li, J.F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M. and Sheen, J. (2013) Multiplex and homologous recombination–mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.*, **31**, 688–691. Available at: http://www.nature.com/doifinder/10.1038/nbt.2654.

Li, L., Wu, L.P. and Chandrasegaran, S. (1992) Functional domains in Fok I restriction endonuclease. *PNAS*, **89**, 4275–4279.

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. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22565958.

Li, T., Huang, S., Jiang, W.Z., Wright, D., Spalding, M.H., Weeks, D.P. and Yang, B. (2011a) TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res.*, **39**, 359–372. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3017587&tool=pmcentrez&rendertype=abstract.

Li, T., Huang, S., Zhao, X., Wright, D.A., Carpenter, S., Spalding, M.H., Weeks, D.P. and Yang, B. (2011b) Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. *Nucleic Acids Res.*, **39**, 6315–6325. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3152341&tool=pmcentrez&rendertype=abstract.

Liang, P., Xu, Y., Zhang, X., et al. (2015) CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes. *Protein Cell*. Available at: http://link.springer.com/10.1007/s13238-015-0153-5.

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. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24576457.

Liu, L., Liu, L., Davies, J.W., Davies, J.W., Stanley, J. and Stanley, J. (1998) Mutational analysis of *bean yellow dwarf virus*, a geminivirus of the genus Mastrevirus that is adapted to dicotyledonous plants. *J. Gen. Virol.*, **79**, 2265–2274.

Liu, L., Tonder, T. Van, Pietersen, G., Davies, J.W. and Stanley, J. (1997) Molecular characterization of a subgroup I geminivirus from a legume in South Africa. *J. Gen. Virol.*, **78**, 2113–2117.

Liu, W., Yuan, J.S. and Stewart, C.N. (2013) Advanced genetic tools for plant biotechnology. *Nat. Rev. Genet.*, **14**, 781–93. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24105275.

Lloyd, A., Plaisier, C.L., Carroll, D. and Drews, G.N. (2005) Targeted mutagenesis using zinc-finger nucleases in *Arabidopsis*. *PNAS*, **102**, 2232–7. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=548540&tool=pmcentrez&rendertype=abstract.

Lusser, M., Parisi, C., Plan, D. and Rodríguez-Cerezo, E. (2012) Deployment of new biotechnologies in plant breeding. *Nat. Biotechnol.*, **30**, 231–239. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22398616.

Maeder, M.L., Thibodeau-beganny, S., Osiak, A., et al. (2008) Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol. Ther.*, **31**, 294–301.

Maeseele, P. (2013) Risk conflicts, critical discourse analysis and media discourses on GM crops and food. *Journalism*. Available at: http://jou.sagepub.com/cgi/doi/10.1177/1464884913511568.

Mahfouz, M.M., Li, L., Shamimuzzaman, M., Wibowo, A., Fang, X. and Zhu, J.-K. (2011) De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *PNAS*, **108**, 2623–2628. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3038751&tool=pmcentrez&rendertype=abstract.

Mak, A.N.-S., Bradley, P., Cernadas, R.A., Bogdanove, A.J. and Stoddard, B.L. (2012) The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science*, **335**, 716–719. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22223736.

Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L. and Church, G.M. (2013) Cas9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.*, **31**, 833–838. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23907171.

Marton, I., Zuker, A., Shklarman, E., Zeevi, V., Tovkach, A., Roffe, S., Ovadis, M., Tzfira, T. and Vainstein, A. (2010) Nontransgenic genome modification in plant cells. *Plant Physiol.*, **154**, 1079–1087. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2971589&tool=pmcentrez&rendertype=abstract.

McDougall, P., Centre, V.B. and Kingdom, U. (2011) *The cost and time involved in the discovery, development and authorisation of a new plant biotechnology derived trait. A consultancy study for Crop Life International*.

Miller, J.C., Holmes, M.C., Wang, J., et al. (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.*, **25**, 778–785. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17603475.

Miller, J.C., Tan, S., Qiao, G., et al. (2011) A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.*, **29**, 143–148. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21179091.

Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J. and Soria, E. (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.*, **60**, 174–182.

Mojica, F.J.M., Díez-Villaseñor, C., Soria, E. and Juez, G. (2000) Biological significance of a family of regularly spaced repeats in the genomes of *Archaea*, *Bacteria* and mitochondria. *Mol. Microbiol.*, **36**, 244–246.

Moscou, M.J. and Bogdanove, A.J. (2009) A simple cipher governs DNA recognition by TAL effectors. *Science*, **326**, 1501.

Mussolino, C., Morbitzer, R., Lütge, F., Dannemann, N., Lahaye, T. and Cathomen, T. (2011) A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.*, **39**, 9283–93. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3241638&tool=pmcentrez&rendertype=abstract.

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. Available at: http://www.nature.com/doifinder/10.1038/nbt.2655.

Netter, P., Petrochilo, E., Slonlmski, P.P., Bolotin-fukuhara, M., Coen, D. and Ginitique, C. De (1974) Mitochondrial genetics VII. Allelism and mapping studies of ribosomal mutants resistant to chloramphenicol, erythromycin and spiramycin in *S. cerevisiae*. *Genetics*, **78**, 1063–1100.

Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F. and Nureki, O. (2014) Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*, **156**, 935–949. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24529477.

Noël, L.D., Denancé, N. and Szurek, B. (2013) Predicting promoters targeted by TAL effectors in plant genomes: from dream to reality. *Front. Plant Sci.*, **4**, 333. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3759964&tool=pmcentrez&rendertype=abstract.

Orr-Weaver, T.L., Szostak, J.W. and Rothstein, R.J. (1981) Yeast transformation: a model system for the study of recombination. *PNAS*, **78**, 6354–6358.

Osakabe, K., Osakabe, Y. and Toki, S. (2011) Site-directed mutagenesis in *Arabidopsis* using custom-designed zinc finger nucleases. *PNAS*, **108**, 433–433. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1018365108.

Pacher, M., Schmidt-Puchta, W. and Puchta, H. (2007) Two unlinked double-strand breaks can induce reciprocal exchanges in plant genomes via homologous recombination and nonhomologous end joining. *Genetics*, **175**, 21–9. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1775016&tool=pmcentrez&rendertype=abstract.

Paszkowski, J., Baur, M., Bogucki, a and Potrykus, I. (1988) Gene targeting in plants. *EMBO J.*, **7**, 4021–4026.

Pavletich, N.P. and Pabo, C. (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1Å. *Science*, **252**, 809–817.

Paz, M. and Veilleux, R. (1999) Influence of culture medium and *in vitro* conditions on shoot regeneration in *Solanum phureja* monoploids and fertility of regenerated doubled monoploids. *Plant Breed.*, **118**, 53–57. Available at: http://onlinelibrary.wiley.com/doi/10.1046/j.1439-0523.1999.118001053.x/abstract.

Perlak, F.J., Stone, T.B., Muskopf, Y.M., et al. (1993) Genetically improved potatoes: protection from damage by Colorado potato beetles. *Plant Mol. Biol.*, **22**, 313–321.

Petolino, J.F. (2015) Genome editing in plants via designed zinc finger nucleases. *Vitr. Cell. Dev. Biol. - Plant*, **51**, 1–8. Available at: http://link.springer.com/10.1007/s11627-015-9663-3.

Petolino, J.F., Worden, A., Curlee, K., Connell, J., Strange Moynahan, T.L., Larsen, C. and Russell, S. (2010) Zinc finger nuclease-mediated transgene deletion. *Plant Mol. Biol.*, **73**, 617–28. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20454835.

PGSC (2011) Genome sequence and analysis of the tuber crop potato. *Nature*, **475**, 189–195. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21743474.

Port, F., Muschalik, N. and Bullock, S.L. (2015) Systematic evaluation of *Drosophila* CRISPR tools reveals safe and robust alternatives to autonomous gene drives in basic research. *G3*, **5**, 1493–1502. Available at: http://g3journal.org/cgi/doi/10.1534/g3.115.019083.

Pruett-Miller, S.M., Connelly, J.P., Maeder, M.L., Joung, J.K. and Porteus, M.H. (2008) Comparison of zinc finger nucleases for use in gene targeting in mammalian cells. *Mol. Ther.*, **16**, 707–717.

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

Puchta, H., Dujon, B. and Hohn, B. (1993) Homologous recombination in plant cells is enhanced by *in vivo* induction of double strand breaks into DNA by a site-specific endonuclease. *Nucleic Acids Res.*, **21**, 5034–5040. Available at: http://nar.oxfordjournals.org/content/21/22/5034.short.

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. *PNAS*, **93**, 5055–5060. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=39405&tool=pmcentrez&rendertype=abstract.

Puchta, H. and Fauser, F. (2013) Synthetic nucleases for genome engineering in plants: prospects for a bright future. *Plant J.*, 1–15. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24112784.

Qi, Y., Li, X., Zhang, Y., et al. (2013a) Targeted deletion and inversion of tandemly arrayed genes in *Arabidopsis thaliana* using zinc finger nucleases. *G3*, **3**, 1707–15. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23979943.

Qi, Y., Zhang, Y., Zhang, F., Baller, J.A., Cleland, S.C., Ryu, Y., Starker, C.G. and Voytas, D.F. (2013b) Increasing frequencies of site-specific mutagenesis and gene targeting in *Arabidopsis* by manipulating DNA repair pathways. *Genome Res.*, **23**, 547–554. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23282329.

Ramirez, C.L., Foley, J.E., Wright, D. a, et al. (2008) Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat. Methods*, **5**, 374–375.

Ran, F.A., Hsu, P.D., Lin, C.Y., et al. (2013) Double nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity. *Cell*, **154**, 1380–1389. Available at: http://dx.doi.org/10.1016/j.cell.2013.08.021.

Salomon, S. and Puchta, H. (1998) Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *EMBO J.*, **17**, 6086–95. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1170935&tool=pmcentrez&rendertype=abstract.

Sander, J.D., Dahlborg, E.J., Goodwin, M.J., et al. (2011) Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat. Methods*, **8**, 67–9. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3018472&tool=pmcentrez&rendertype=abstract.

Scalley-Kim, M., McConnell-Smith, A. and Stoddard, B.L. (2007) Coevolution of a homing endonuclease and its host target sequence. *J. Mol. Biol.*, **372**, 1305–1319.

Schuermann, D., Molinier, J., Fritsch, O. and Hohn, B. (2005) The dual nature of homologous recombination in plants. *Trends Genet.*, **21**, 172–181.

Sha, A., Zhao, J., Yin, K., Tang, Y., Wang, Y., Wei, X., Hong, Y. and Liu, Y. (2013) Virus-based microRNA silencing in plants. *Plant Physiol.* Available at: http://www.ncbi.nlm.nih.gov/pubmed/24296072.

Shan, Q., Wang, Y., Chen, K., et al. (2013a) Rapid and efficient gene modification in rice and *Brachypodium* using TALENs. *Mol. Plant*, **6**, 1365–1368. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1674205214609254.

Shan, Q., Wang, Y., Li, J., et al. (2013b) Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.*, **31**, 686–688. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23929337.

Sharma, S.K., Bolser, D., Boer, J. de, et al. (2013) Construction of reference chromosome-scale pseudomolecules for potato: integrating the potato genome with genetic and physical maps. *G3*, **3**, 2031–47. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3815063&tool=pmcentrez&rendertype=abstract.

Shukla, V., Doyon, Y. and Miller, J. (2009) Precise genome modification in the crop species Zea mays using zinc-finger nucleases. *Nature*, **459**. Available at: http://www.nature.com/nature/journal/vaop/ncurrent/full/nature07992.html.

Siebert, R. and Puchta, H. (2002) Efficient repair of genomic double-strand breaks by homologous recombination between directly repeated sequences in the plant genome. *Plant Cell*, **14**, 1121–1131.

Smith, G.R. (1989) Homologous recombination in *E. coli*: multiple pathways for multiple reasons. *Cell*, **58**, 807–809.

Stoddard, B. (2011) Homing endonucleases: from microbial genetic invaders to reagents for targeted DNA modification. *Structure*, **19**, 7–15. Available at: http://www.sciencedirect.com/science/article/pii/S0969212610004399.

Stoddard, B.L. (2005) Homing endonuclease structure and function. *Q. Rev. Biophys.*, **38**, 49–95.

Sun, N. and Zhao, H. (2013) Transcription activator-like effector nucleases (TALENs): a highly efficient and versatile tool for genome editing. *Biotechnol. Bioeng.*, **110**, 1811–21. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23508559.

Szczepek, M., Brondani, V., Büchel, J., Serrano, L., Segal, D.J. and Cathomen, T. (2007) Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat. Biotechnol.*, **25**, 786–793.

Tang, T.-H., Bachellerie, J.-P., Rozhdestvensky, T., Bortolin, M.-L., Huber, H., Drungowski, M., Elge, T., Brosius, J. and Hüttenhofer, A. (2002) Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*. *PNAS*, **99**, 7536–7541.

Thomas, M. (2013) Fields of gold. *Nature*, **497**, 5–6.

Toevs, E.A., Guenthner, J.F., Johnson, A.J., McIntosh, C.S. and Thornton, M.K. (2011) An industry perspective of all-native and transgenic potatoes. *AgBioForum*, **14**, 14–19.

Townsend, J.A., Wright, D.A., Winfrey, R.J., Fu, F., Maeder, M.L., Joung, J.K. and Voytas, D.F. (2009) High-frequency modification of plant genes using engineered zinc-finger nucleases. *Nature*, **459**, 442–5. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2743854&tool=pmcentrez&rendertype=abstract.

Tsai, H., Howell, T., Nitcher, R., et al. (2011) Discovery of rare mutations in populations: TILLING by sequencing. *Plant Physiol.*, **156**, 1257–68. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3135940&tool=pmcentrez&rendertype=abstract.

Tsai, S.Q., Wyvekens, N., Khayter, C., Foden, J.A., Thapar, V., Reyon, D., Goodwin, M.J., Aryee, M.J. and Joung, J.K. (2014) Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.*, **32**, 569–576. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24770325.

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

Tzfira, T., Frankman, L., Vaidya, M. and Citovsky, V. (2003) Site-specific integration of *Agrobacterium tumefaciens* T-DNA via double-stranded intermediates. *Plant Physiol.*, **133**, 1011–1023. Available at: http://www.plantphysiol.org/content/133/3/1011.short.

USDA (2012) *USDA Organic 101*.

Van Montagu, M. (2011) It is a long way to GM agriculture. *Annu. Rev. Plant Biol.*, **62**, 1–23. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21314429.

Voytas, D.F. and Gao, C. (2014) Precision genome engineering and agriculture: opportunities and regulatory challenges. *PLoS Biol.*, **12**, e1001877. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4051594&tool=pmcentrez&rendertype=abstract.

Vu, G.T.H., Cao, H.X., Watanabe, K., Hensel, G., Blattner, F.R., Kumlehn, J. and Schubert, I. (2014) Repair of site-specific DNA double-strand breaks in barley occurs via diverse pathways primarily involving the sister chromatid. *Plant Cell*, **26**, 2156–2167. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24876253.

Waltz, E. (2015a) Nonbrowning GM apple cleared for market. *Nat. Biotechnol.*, **33**, 326–327. Available at: http://www.nature.com/doifinder/10.1038/nbt0415-326c.

Waltz, E. (2012) Tiptoeing around transgenics. *Nat. Biotechnol.*, **30**, 215–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22398610.

Waltz, E. (2015b) USDA approves next-generation GM potato. *Nat. Biotechnol.*, **33**, 12–13.

Wang, T.L., Uauy, C., Robson, F. and Till, B. (2012) TILLING in extremis. *Plant Biotechnol. J.*, **10**, 761–72. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22651686.

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C. and Qiu, J.-L. (2014) Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol.*, **32**, 947–951. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25038773.

Watson, J.D. and Crick, F.H.C. (1953) Molecular structure of nucleic acids. *Nature*, **171**, 737–738. Available at: http://www.nature.com/physics/looking-back/crick/\nhttp://www.ncbi.nlm.nih.gov/pubmed/13054692.

Weinthal, D.M., Taylor, R.A. and Tzfira, T. (2013) Nonhomologous end joining-mediated gene replacement in plant cells. *Plant Physiol.*, **162**, 390–400. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23509176.

Wendt, T., Holm, P.B., Starker, C.G., Christian, M., Voytas, D.F., Brinch-Pedersen, H. and Holme, I.B. (2013) TAL effector nucleases induce mutations at a pre-selected location in the genome of primary barley transformants. *Plant Mol. Biol.*, **83**, 279–285. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23689819.

Wiedenheft, B., Sternberg, S.H. and Doudna, J.A. (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature*, **482**, 331–338.

Wohlers, A.E. (2013) Labeling of genetically modified food. *Polit. Life Sci.*, **32**, 58–72. Available at: http://www.bioone.org/doi/abs/10.2990/32_1_58.

Wright, D. a, Townsend, J.A., Winfrey, R.J., Irwin, P.A., Rajagopal, J., Lonosky, P.M., Hall, B.D., Jondle, M.D. and Voytas, D.F. (2005) High-frequency homologous recombination in plants mediated by zinc-finger nucleases. *Plant J.*, **44**, 693–705. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16262717.

Wu, T.C. and Lichten, M. (1995) Factors that affect the location and frequency of meiosis-induced double- strand breaks in *Saccharomyces cerevisiae*. *Genetics*, **140**, 55–66.

Yang, J., Zhang, Y., Yuan, P., et al. (2014) Complete decoding of TAL effectors for DNA recognition. *Cell Res.*, **24**, 628–631. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4011339&tool=pmcentrez&rendertype=abstract.

Zhang, F., Maeder, M.L., Unger-Wallace, E., et al. (2010) High frequency targeted mutagenesis in *Arabidopsis thaliana* using zinc finger nucleases. *PNAS*, **107**, 12028–12033. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2900673&tool=pmcentrez&rendertype=abstract.

Zhang, Y., Zhang, F., Li, X., Baller, J.A., Qi, Y., Starker, C.G., Bogdanove, A.J. and Voytas, D.F. (2013) Transcription activator-like effector nucleases enable efficient plant genome engineering. *Plant Physiol.*, **161**, 20–27. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3532252&tool=pmcentrez&rendertype=abstract.

CHAPTER 2: GENERATION AND INHERITANCE OF TARGETED MUTATIONS IN POTATO (*SOLANUM TUBEROSUM* L.) USING THE CRISPR/CAS SYSTEM

A paper to be submitted to PLoS ONE

Nathaniel M. Butler[1], Paul A. Atkins[2], Daniel F. Voytas[2], David S. Douches[1]

[1]Department of Plant, Soils and Microbial Sciences, Michigan State University, East Lansing, Michigan 48824, USA

[2]Department of Genetics, Cell Biology and Development and Center for Genome Engineering, University of Minnesota, Minneapolis, Minnesota 55455, USA

Abstract

Genome editing using sequence-specific nucleases (SSNs) offers an alternative approach to conventional genetic engineering and an opportunity to extend the benefits of genetic engineering in agriculture. Currently available SSN platforms, such as zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and CRISPR/Cas (clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated systems (Cas)) have been used in a range of plant species for targeted mutagenesis via nonhomologous end-joining (NHEJ) are just beginning to be explored in crops such as potato (*Solanum tuberosum* Group Tuberosum L.). In this study, CRISPR/Cas reagents expressing one of two single-guide RNA (gRNA) targeting the potato *ACETOLACTATE SYNTHASE1* (*StALS1*) gene were tested for inducing targeted mutations in callus and stable events of diploid and

tetraploid potato using *Agrobacterium*-mediated transformation with either a conventional 35S or a modified geminivirus T-DNA. The percentage of primary events with targeted mutations ranged from 3-60% per transformation and from 0-29% above an expected threshold based on the number of *ALS* alleles. Primary events with mutation frequencies above the expected threshold were used for targeted mutation cloning and inheritance studies using clonal propagation and crosses or selfing. Four of the nine primary events used for targeted mutation cloning had more than one type of mutation, and eight primary events contained targeted mutations that were maintained across clonal generations. Somatic mutations were most evident in the diploid background with three of the four candidates having more than two mutation types at a single *ALS* locus. Conversely, in the tetraploid background, four of the five candidates carried only one mutation type and in two of these cases, carried the same mutation at two different *ALS* loci. Single mutations were inherited through the germline of both diploid and tetraploid primary events with transmission percentages ranging from 87-100%. This demonstration of CRISPR/Cas in potato extends the range of plant species modified using CRISPR/Cas and provides a framework for future studies.

Introduction

Genome editing using sequence-specific nucleases (SSNs) is rapidly becoming a powerful tool for genetic engineering in crop species. Genetic engineering has played an important role in the development of modern agriculture and has contributed significantly to improvements in crop yield, quality and disease resistance (Klümper and

Qaim, 2014). Conventional genetic engineering relies on the action of trans-, intra-, or cisgenes to confer novel traits (Hunter, 2014). In contrast, genome editing relies on the action of SSNs to induce double-strand breaks (DSBs) at specific genomic sites and employing DNA repair pathways to induce target mutations through nonhomologous end-joining (NHEJ) or incorporate new sequence through homologous recombination (HR) (Curtin *et al.*, 2012). Modifications are typically unlinked to integrated SSN reagents and can be segregated out of progeny using selfing or crossing to non-transgenic parents. Founding SSN platforms, such as meganucleases that are based on natural endonucleases, have limited sequence specificity and are costly to engineer, limiting their applications for genome editing (Takeuchi *et al.*, 2011; Stoddard, 2011). Subsequent SSN platforms, including zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) are synthetic endonucleases employing a customizable DNA binding domain and the FokI nuclease (Miller *et al.*, 2011; Miller *et al.*, 2007). The fusion of these domains provide flexible sequence specificity and both ZFNs and TALENs have demonstrated efficacy in a range of crop species (Li *et al.*, 2012; Shukla *et al.*, 2009; Wendt *et al.*, 2013; Curtin *et al.*, 2011).

CRISPR/Cas (clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated systems (Cas)) represents an alternative class of SSNs that are RNA-guided endonucleases (RGENs) and feature robust activity and simple design (Belhaj *et al.*, 2013). In contrast to other SSN platforms, RGENs consist of a common nuclease and specific guide RNA to direct nuclease binding and cleavage of target DNA. The type II CRISPR/Cas system from *Streptococcus pyrogenes* used for genome editing employs a common nuclease, Cas9 and a CRISPR RNA (crRNA) and

trans-activation CRISPR RNA (tracrRNA) duplex as a specific guide RNA to target protospacer-adjacent motif (PAM)-containing DNA (Jinek *et al.*, 2012). Alone, Cas9 will bind transiently to PAM-containing DNA but requires involvement of the crRNA:tracrRNA duplex for high fidelity binding and cleavage (Gasiunas *et al.*, 2012; Sternberg *et al.*, 2014). For simplicity, the crRNA and tracrRNA have been fused into a single-guide RNA (sgRNA) which can be designed to target a specific sequence by modulating the first 20 nucleotides of the sgRNA to match the complementary strand of a 'protospacer' target DNA site (Jinek *et al.*, 2012). Co-expression Cas9 with one or more sgRNA provides a two-component system capable of targeting multiple loci for modification (Cong *et al.*, 2013).

Genetic engineering in agriculture is at an important crossroads. Increasing pressure from the public to develop 'safer' biotechnology and advances in SSN and next-generation sequencing (NGS) technology have provided an opportunity to advance genome editing methodology and usher in a new generation of genetic engineering (Palmgren *et al.*, 2014). Most crops, like potato (*Solanum tuberosum* Group Tuberosum L.) are amendable to plant transformation but lack sufficient genetic resources to validate genetic studies and assess gene function. Random mutagenesis using ethyl methanesulfonate (EMS), radiation, or T-DNA integration requires generating large mutant collections and extensive screening to identify informative mutants. CRISPR/Cas and other SSN platforms allow mutagenesis of target genes and direct assessment of gene function.

This report demonstrates the use of CRISPR/Cas for targeted mutagenesis in both diploid and tetraploid potato. CRISPR/Cas reagents targeting the potato

ACETOLACTATE SYNTHASE1 (*StALS1*) gene were expressed in leaf explants via *Agrobacterium tumefaciens* (*Agrobacterium*) using a conventional 35S T-DNA expression vector (Curtis and Grossniklaus, 2003) or a modified geminivirus T-DNA expression vector (Baltes *et al.*, 2014). Both sgRNAs and T-DNAs tested were capable of generating targeted mutations in stable events. Single targeted mutations in primary events were capable of being carried through clonal generations and the germline as Cas9-free progeny.

Materials and Methods

Plant materials

The tetraploid *S. tuberosum* cultivar "Désirée" (Désirée) and a diploid self-incompatible breeding line, MSX914-10 (X914-10) were used in the study. X914-10 was produced from a cross between the doubled-monoploid (DM) *S. tuberosum* Group Phureja line used to construct the potato reference genome (Xu *et al.*, 2011) and 84SD22, a heterozygous *S. tuberosum* x *S. chacoense* hybrid breeding line and has high transformation efficiency (Felcher *et al.*, 2012). Désirée is a red-skinned variety with high transformation efficiency (Verhoeven and Dijkhuis, 1988). Three to four-week-old tissue culture plants used for *Agrobacterium* transformation were grown in Magenta® boxes (Phytotech, Shawnee Mission, KS) on light racks set to 16-h-light/8-h-night photoperiod at 22C. Eight to ten-week-old soil-grown plants used in crosses or selfing were grown in greenhouses under the same photoperiod as tissue culture plants.

Fruit was harvested three weeks following fruit set. An inbred diploid line, M6 (Jansky *et al.*, 2014) was used in crosses with X914-10 events while Désirée events were selfed.

CRISPR/Cas reagent preparation

The *Streptococcus pyogenes* Cas9 gene was codon-optimized for *Arabidopsis* and synthesized (GenScript, Piscataway, NJ) as previously described (Baltes *et al*., 2014). Cas9 and individual sgRNA driven by the *Arabidopsis* U6 RNA pol III promoter (Waibel and Filipowicz, 1990) were cloned into Gateway-compatible binary vectors pMDC32 (35S; (Curtis and Grossniklaus, 2003)) and pLSL (LSL; (Baltes *et al*., 2014)). The Rep and RepA coding sequences from the *Bean Yellow Dwarf Virus* (*BeYDV*) were cloned into the pMDC32 vector for co-expression with pLSL reagents (Baltes *et al*., 2014).

Agrobacterium-mediated transformation

Agrobacterium-mediated transformation of potato leaf explants was conducted as previously described (Paz and Veilleux, 1999). Approximately 20-40 hygromycin-resistant events rooting in 5 mg/L hygromycin B (Life Technologies, Grand Island, NY) were sampled for each transformation. Callus was sampled by excising wounded surfaces of leaf explants that included both callus and non-callus tissues. Sampled callus from three to four leaf explants were combined for genomic DNA extractions.

PCR and restriction enzyme digestion assays

Genomic DNA was extracted from callus and leaf tissues using the DNeasy Plant Mini kit (Qiagen, Valencia, CA). For T-DNA PCRs, primers 5'-

CCTGTCGTGCCAGCTGC-3' and 5'-TGTTGAGAACTCTCGACGTCCTGC-3' were used for LSL T-DNA, and primers 5'- CGAGCTCCACCGCGG-3' and 5'-CCTCCTTAGACGTTGCAGTC-3' for Rep T-DNA. Primary PCR amplicons of *StALS* loci were generated using primers 5'-GGTTGACATTGATGGTGAC-3' and 5'-GCCTAGAACTAGTTATGTAG-3' with 100 ng genomic DNA and Phusion High-Fidelity DNA Polymerase (NEB, Ipsich, MA). Primary amplicons were purified using the QIAquick PCR purification kit (Qiagen, Valencia, CA) and digested overnight with *Alo*I (Life Technologies, Grand Island, NY) or *Bsl*I (NEB, Ipswich, MA) using recommended conditions. Resistant bands were purified from 2% agarose gels using QIAquick Gel Extraction kit (Qiagen, Valencia, CA) and subcloned using the Topo TA Cloning kit (Life Technologies, Grand Island, NY) for Sanger sequencing at the Michigan State University Research Technology Support Facility (RTSF).

Results and Discussion

The *StALS1* gene was chosen as a target locus for designing two sgRNAs (gRNA746 and gRNA751). Each sgRNA target site is separated by 215 base pairs (bp) and is localized to the 3' end of the *StALS1* coding sequence (Figure 3A). A closely related paralog of *StALS1* (PGSC0003DMG400034102), *StALS2* (PGSC0003DMG400007078), is also targeted by gRNA751 and contains a single nucleotide polymorphism (SNP) in the target site of gRNA746 (Figure 3A; lowercase). Both sgRNAs include restriction enzyme sites proximal to the PAM to facilitate detection and cloning of NHEJ mutations at target loci (Figure 3A; underlined).

Agrobacterium-mediated delivery of CRISPR/Cas reagents was used for both transient expression and generation of primary events (Figure 6). Expression of an *Arabidopsis* codon-optimized Cas9 and individual sgRNAs were driven by a doubled cauliflower mosaic virus 35S promoter and *Arabidopsis* U6 promoter, respectively, in either a conventional pCambia T-DNA backbone (35S; (Curtis and Grossniklaus, 2003)) or modified geminivirus backbone (LSL; (Baltes *et al.*, 2014)) (Figure 7). Geminivirus T-DNA constructs were co-transformed with a conventional T-DNA constitutively expressing the replicase (Rep) coding sequence required for geminivirus replicon release and expression of Cas9 (Baltes *et al.*, 2014).

One week following *Agrobacterium* infection, calli was collected and total genomic DNA extracted for target mutation detection using enrichment PCR (Fig. 3B and 8). A conventional enrichment PCR failed to detect targeted mutations in most samples with only slight detection in Désirée calli transformed with gRNA746 in the 35S T-DNA (Figure 8; arrows). A modified enrichment PCR achieved more sensitive detection and was used to determine if the callus samples contained targeted mutations (Tables 2 and 4). Overall, targeted mutations could be detected in calli of both genotypes using either sgRNA in the conventional 35S T-DNA but not the geminivirus LSL T-DNA. The reduction of targeted mutations in calli transformed with the geminivirus LSL T-DNA was also observed in *Agrobacterium*-infiltrated tobacco leaves and supports the use of the geminivirus vector system for promoting HR rather than NHEJ mutagenesis (Baltes *et al.*, 2014).

To determine if targeted mutations detected in calli could also be detected in stable expression lines, transformed calli were regenerated and resulting primary events

were screened for targeted mutations using a restriction enzyme digestion assay (Figure 4A, Tables 2 and 5). Hygromycin selection was used during regeneration and in a rooting assay to generate stable events (Table 2; total events). Total genomic DNA was extracted from leaf tissue of primary events (T_0) and used for PCR amplification and overnight digestion with a restriction enzyme that cleaves within the sgRNA target site (Figure 3A). Digested amplicons were subjected to gel electrophoresis and mutation frequencies estimated using the fraction of resistant band intensity divided by the sum of the resistant and digested band intensities (Table 5). Events with mutation frequencies equal or greater than a 25% and 12.5% threshold for X914-10 and Désirée, respectively were considered mutant events based on expected single allele mutation frequencies across both *StALS* loci (Table 2; # above threshold).

In X914-10, mutant events accounted for 15% (gRNA746) and 3% (gRNA751) of 35S T-DNA lines and 3% (gRNA746) and none (gRNA751) of LSL T-DNA lines. In Désirée, mutant events accounted for 29% (gRNA746) and 3% (gRNA751) of 35S T-DNA lines and none of the LSL expression lines (Table 2). An increase in the number of mutant events from gRNA746 in relation to gRNA751 using the conventional 35S T-DNA (approximately five and ten-fold for X914-10 and Désirée, respectively) may be due to a GG motif at the 3' end of the target sequence of gRNA746 but not gRNA751 that has shown to improve sgRNA efficiency (Figure 3A) (Farboud and Meyer, 2015; Xu *et al.*, 2015). Furthermore, the lack of mutant events from using the geminivirus T-DNA support the results of the transient assay but also could be the result of inefficient co-transformation of the geminivirus and the Rep T-DNAs. To investigate this possibility, primers specific to each T-DNA were used in a PCR assay (Figure 9). Although the Rep

T-DNA could be detected in most events, the LSL T-DNA could only be clearly detected in one event (Figure 9B; lane 13). These results suggest the LSL T-DNA is not properly integrating into the genomes of primary events and could be acting transiently.

In order to characterize mutant alleles in stable expression lines and track them across clonal generations, a subset of nine mutant events derived from the gRNA746 35S T-DNA, four from X914-10 and five from Désirée, were vegetatively propagated in tissue culture (clonal generation 1; CG_1) using shoot tip explants and leaf tissue sampled for targeted mutation cloning. Total genomic DNA from both clonal generations, T_0 and CG_1 were used in restriction enzyme digestion assays to produce resistant bands (Figure 4A). Resistant bands were excised and subcloned for Sanger sequencing. Sequence reads were aligned with wild-type sequence to identify targeted mutations and their corresponding locus (i.e. *StALS1* or -*2*) (Figure 4B and 10).

Insertion-deletion mutations were identified in all nine mutant events ranging from a single bp insertion (X51-28 (+1)) to a 38 bp deletion (X46-27 (-38)) (Figure 4B and 10). Four of the nine mutant events had more than one type of mutation and most likely contain somatic mutations (X46-3, -27, -32 and D46-44). Somatic mutations were most evident in the diploid background with three of the four events having more than two mutation types at a single locus (X46-3, -27 and -32). Conversely, in the tetraploid background, four of the five events carried only one mutation type and in two of these cases, carried the same mutation at both *StALS* loci (Des46-7 and -9). The occurrence of a 4 bp deletion across D46-7, -8 and -9 is most likely an artifact of transformation (i.e. taken from the same callus). Nevertheless, discovery of the same 4 bp deletion mutation at different loci within D46-7 and -9 suggests a preference for this mutation

type and might be explained by microhomology ("TGG") within the gRNA746 target site (Qi *et al.*, 2013). Individual mutations were identified across T_0 and CG_1 generations in eight of the nine mutant events with a preference for mutations identified more frequently in the T_0 generation (X46-3, -27, -32, D46-7, -8, -9, -44 and D51-5). Complete mutagenesis of all *StALS* alleles was not observed in the nine primary events analyzed and is most likely due to *ALS* being an essential gene (Smith *et al.*, 1989). The maintenance of individual mutation types across clonal generations in both genetic backgrounds suggests individual mutations can be carried into future clonal generations.

Inheritance of germline mutations and CRISPR/Cas reagents was also evaluated in progeny of three mutant events (Figure 5, 11-12, Table 3). Tetraploid mutant events, D46-9 and -44 were selfed and diploid event, X46-3 was crossed to a self-compatible diploid pollinator line, M6 (Jansky *et al.*, 2014). Progeny from each population were screened for inheritance of CRISPR/Cas reagents ("Cas9") and Cas9-free progeny were assessed for targeted mutations (Figure 5, 11-12). The percentage of Cas9-free progeny ranged from 19-37% across ploidy types and 87-100% Cas9-free progeny contained targeted mutations (Figure 11-12, Table 3). To determine if targeted mutations detected in progeny were inherited from primary events, two progeny from each population were chosen for mutation cloning. Cas9-containing progeny from X46-3 and D46-9 (X46-3_49 and D46-9_6, respectively) contained new somatic mutations along with mutations from primary events (Figure 5B). Conversely, Cas9-free progeny from all three primary events (X46-3_66, D46-9_7, and D46-44_8) inherited targeted mutations from primary event and new germline mutations with an expected number of

mutant alleles. Interesting, targeted mutations that predominated in tetraploid primary events also predominated in progeny regardless of Cas9 inheritance (ex: D46-44_24 (-3)). This is likely due to opportunity for multiple mutant alleles in the tetraploid background and enrichment of mutant alleles through selfing. Furthermore, the lack of a wild-type band signal from X46-3_66 and the presence of only one mutant type in *StALS1* is intriguing and could possibly be due to an additional mutation allele in *StALS2* that was not identified.

This report in potato and other recent reports in tomato and citrus support the use of CRISPR/Cas for targeted mutagenesis in members of the Solanaceae family and vegetatively propagated plant species (Nekrasov *et al.*, 2013; Brooks *et al.*, 2014; Lor *et al.*, 2014; Jia and Wang, 2014). The ability to make targeted mutations in diploid potato events and introgression of self-compatibility from self-compatible diploid lines, such as M6 provides a never before opportunity to fix targeted mutations and conduct functional genomics in potato (Jansky *et al.*, 2014). Furthermore, the geminivirus T-DNA used in this study has previously been shown to be effective for promoting HR in tobacco but could also potentially be used for transient expression of genome editing reagents (Baltes *et al.*, 2014). This approach would be beneficial in clonally propagated species that cannot be used in genetic crosses to remove genome editing reagents and achieve non-regulated status. Nevertheless, further analysis of LSL T-DNA integration in the genomes of geminivirus-modified events must be conducted.

APPENDICES

Appendix A: Tables

Table 2. Summary of targeted mutation screen of primary events and enrichment PCR results from callus. Diploid (X914-10) and tetraploid (Désirée) genotypes were stably transformed with gRNA746 and gRNA751 CRISPR/Cas reagents in a conventional 35S or geminivirus LSL T-DNA backbone using hygromycin selection (Total events). A restriction enzyme digestion assay and quantification of resistant and digested bands were used to identify events with at least 1% mutation frequencies (# with mutations) and events above a threshold using expected single allele mutation frequencies (# above threshold) (Table 5). Percentages are of total events and modified enrichment PCR results come from Figure 3B and Table 4.

Genotype	gRNA	T-DNA	Total events	# with mutations	% with mutations	# above threshold	% above threshold	Modified Enrichment PCR
X914-10	746	35S	27	15	55%	4	15%	+
X914-10	746	LSL	32	13	41%	1	3%	-
X914-10	751	35S	35	3	9%	1	3%	+
X914-10	751	LSL	39	1	3%	0	0%	-
Désirée	746	35S	35	21	60%	10	29%	+
Désirée	746	LSL	33	12	36%	0	0%	-
Désirée	751	35S	37	4	11%	1	3%	+
Désirée	751	LSL	21	1	5%	0	0%	-

Table 3. Summary of targeted mutation screen of progeny from primary events and inheritance of Cas9. Progeny from diploid (X46-3) and tetraploid (D46-9, D46-44) primary events were screened for inheritance of Cas9 (# of Cas9-free progeny) and Cas9-free progeny were screened for targeted mutations (# of Cas9-free progeny with mutations) (Figures 11 and 12). Mutation transmission percentages are of Cas9-free progeny with targeted mutations and percent of Cas9-free are of the number of progeny screened. Mutations detected are targeted mutations cloned from primary events (F_0) and progeny (F_1).

Primary event	Mutations detected (bp)	# of progeny screened	# of progeny Cas9-free	# of Cas9-free progeny with mutations	Mutation transmission (%)	Cas9-free (%)
X46-3	-2, -3, -4, -5, -6, -11, -12, -17	48	18	16 out of 18	89%	37%
D46-9	+1, -4, -8, -11, -13, -17, -23	31	6	6 out of 6	100%	19%
D46-44	+1, -3, -4, -10, -11	25	8	7 out of 8	87%	32%

Table 4. Enrichment PCR assay band quantification data. Diploid (X914-10) and tetraploid (Désirée) potato leaf explants were transformed with CRISPR/Cas reagents in the conventional 35S T-DNA (35S), geminivirus LSL T-DNA (LSL) or untransformed controls (none). *Alo*I and *Bsl*I restriction enzymes were used for gRNA746 and 751, respectively. ImageJ was used for band quantification and normalization was done by dividing enriched by primary band intensities. Enrichment PCR results were determined as positive (+), negative (-), or non-detectable (ND) if enriched bands have normalized intensities equal or over 0.5, less than 0.5 and equal or more than 0.05, or less then 0.05, respectively.

Genotype	gRNA	T-DNA	Enriched	Primary	Normalized enriched	Enrichment PCR
X914-10	746	35S	10738.05	5024.44	2.14	+
X914-10	751	35S	4640.20	5760.75	0.81	+
X914-10	746	LSL	533.85	7090.53	0.08	-
X914-10	751	LSL	773.92	11611.19	0.07	-
X914-10	746	none	34.12	9487.60	0.00	ND
X914-10	751	none	18.12	3230.42	0.01	ND
Désirée	746	35S	13715.77	1289.65	10.64	+
Désirée	751	35S	763.51	1487.72	0.51	+
Désirée	746	LSL	714.68	6374.63	0.11	-
Désirée	751	LSL	251.78	2969.52	0.08	-
Désirée	746	none	20.92	4384.95	0.00	ND
Désirée	751	none	18.53	1358.15	0.01	ND

Table 5. Restriction enzyme digestion assay band quantification data from primary events expressing CRISPR/Cas reagents.

X914-10	*Resistant band*	*Digested band*	*Digested band + correction*	*Mutation frequency (%)*
(35S)				
X1046-1	102.2	6396.9	8790.9	1.1
X1046-3	**9433.8**	**1471.3**	**2021.9**	**82.3**
X1046-6	1928.2	5567.0	7650.4	20.1
X1046-7	640.0	3793.5	5213.2	10.9
X1046-22	1035.5	4666.2	6412.4	13.9
X1046-24	342.9	3136.5	4310.3	7.4
X1046-25	139.4	5422.8	7452.2	1.8
X1046-26	443.7	4970.7	6830.9	6.1
X1046-27	**8974.4**	**2495.9**	**3429.9**	**72.3**
X1046-28	833.9	5331.4	7326.6	10.2
X1046-30	731.3	1996.6	2743.8	21.0
X1046-31	2895.5	3615.3	4968.3	36.8
X1046-32	**2635.2**	**2306.9**	**3170.2**	**45.4**
X1046-33	122.0	6265.9	8610.8	1.4
X1046-35	95.0	5830.4	8012.3	1.2
(LSL)				
X1046-4	213.7	11426.6	15702.7	1.3
X1046-5	1533.9	11326.8	15565.6	9.0
X1046-6	6985.7	7020.9	9648.3	42.0
X1046-14	1645.2	9430.7	12960.0	11.3
X1046-19	530.2	3407.1	4682.1	10.2
X1046-21	137.3	3780.8	5195.7	2.6
X1046-22	668.3	5143.7	7068.7	8.6
X1046-24	159.1	9790.1	13453.8	1.2
X1046-32	2363.5	8295.2	11399.5	17.2
X1046-33	291.7	10548.0	14495.4	2.0
X1046-36	513.4	9451.7	12988.9	3.8
X1046-37	182.8	10031.3	13785.3	1.3
X1046-38	146.8	8679.1	11927.1	1.2
(35S)				
X1051-15	791.5	5018.4	6297.6	11.2
X1051-18	534.8	3778.1	4741.2	10.1
X1051-28	**8957.0**	**3966.9**	**4978.1**	**64.3**
(LSL)				
X1051-35	54.4	2600.4	3263.2	1.6
(WT)				
X914-10 (46)	4.1	5063.3	6958.2	0.1
Désirée (46)	10.9	12825.1	17624.7	0.1
X914-10 (51)	17.3	4826.7	6057.0	0.3
Désirée (51)	6.1	2780.5	3489.2	0.2

Désirée	*Resistant band*	*Digested band*	*Digested band + correction*	*Mutation frequency (%)*
(35S)				
Des46-6	316.5	4534.8	6231.8	4.8
Des46-7	**7834.7**	**4335.8**	**5958.4**	**56.8**
Des46-8	**8847.1**	**4074.4**	**5599.2**	**61.2**
Des46-9	**8357.9**	**3071.7**	**4221.3**	**66.4**
Des46-10	159.0	4691.1	6446.6	2.4
Des46-13	644.5	5401.9	7423.4	8.0
Des46-14	715.2	5191.2	7134.0	9.1
Des46-17	1151.2	6204.5	8526.4	11.9
Des46-22	4428.6	4547.1	6248.7	41.5
Des46-26	1120.1	5036.9	6921.9	13.9
Des46-42	420.4	5435.1	7469.0	5.3
Des46-43	1767.6	5808.8	7982.6	18.1
Des46-44	**6278.1**	**1152.1**	**1583.2**	**79.9**
Des46-48	78.0	5443.6	7480.8	1.0
Des46-49	1004.6	4865.4	6686.2	13.1
Des46-51	279.7	5645.7	7758.5	3.5
Des46-52	1589.1	2840.6	3903.6	28.9
Des46-53	141.4	4739.7	6513.4	2.1
Des46-55	419.0	4767.3	6551.4	6.0
Des46-56	640.8	4016.7	5519.9	10.4
Des46-62	47553.2	6588.9	9054.7	84.0
(LSL)				
Des46-10	284.6	5670.9	7793.1	3.5
Des46-25	495.8	6048.8	8312.5	5.6
Des46-26	646.6	8210.9	11283.7	5.4
Des46-29	1076.8	6706.1	9215.8	10.5
Des46-30	544.6	5940.9	8164.1	6.3
Des46-31	1441.7	7718.5	10607.1	12.0
Des46-33	390.9	10328.1	14193.2	2.7
Des46-34	624.8	9424.7	12951.7	4.6
Des46-35	450.7	9086.4	12486.9	3.5
Des46-37	1151.5	10029.9	13783.5	7.7
Des46-41	359.9	9613.3	13211.0	2.7
Des46-44	261.3	6937.9	9534.2	2.7
(35S)				
Des51-5	**3623.0**	**3276.4**	**4111.6**	**46.8**
Des51-24	299.8	2938.8	3687.9	7.5
Des51-25	36.2	2335.8	2931.2	1.2
Des51-54	41.0	2611.6	3277.2	1.2
(LSL)				
Des51-19	52.4	1411.3	1771.1	2.9

Table 5 (cont'd). Diploid (X) and tetraploid (D) primary events generated using gRNA746 (46) and gRNA751 (51) CRISPR/Cas reagents were screened using a restriction enzyme digestion assay (Figure 4A). Resistant (448 bp) and digested (326 and 357 bp for gRNA746 and gRNA751, respectively) bands were quantified using ImageJ software. Digested bands were corrected for size by multiplying the digested band intensity by the size ratio of the resistant band by the digested band (Digested band + correction). Targeted mutation frequency percentages were calculated by dividing the resistant band by the sum of both resistant and digested bands and multiplying by 100. Listed events have mutation frequencies over 1% and are organized by T-DNA. Events with targeted mutation frequencies over thresholds for expected single allele mutation frequencies (25 and 12.5% for X914-10 and Désirée, respectively) are shaded and bolded events were used for cloning targeted mutations (Figures 4 and 10). Wild-type (WT) controls are shown using restriction enzyme digestion assays for both gRNA746 and gRNA751.

Appendix B: Figures

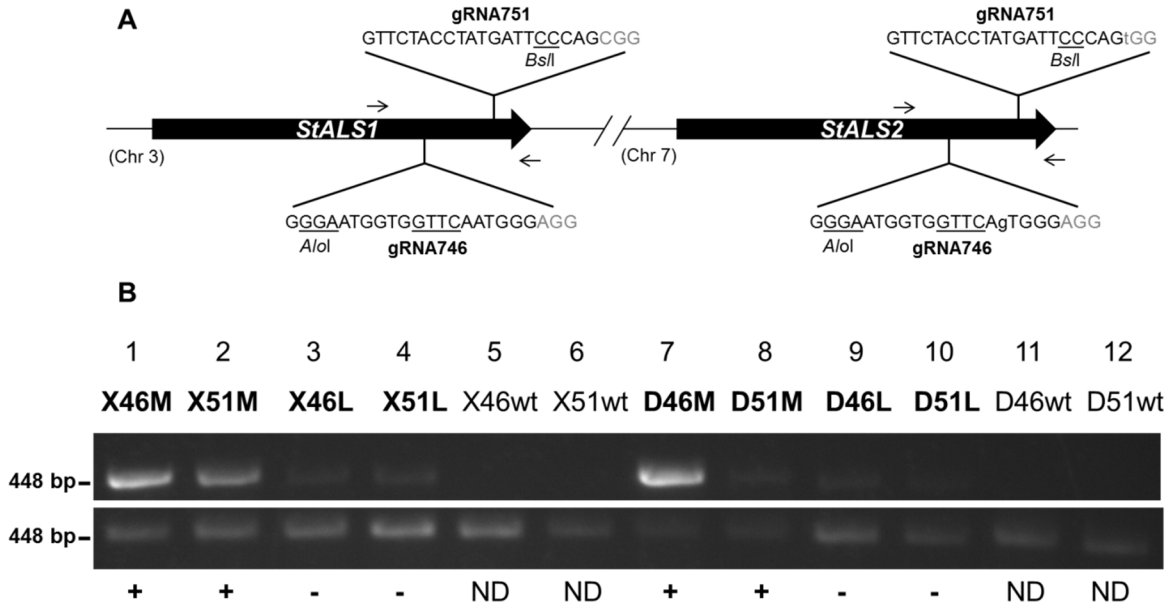

Figure 3. Generation of targeted mutations in callus tissues of potato using CRISPR/Cas reagents. A. Target sites of single-guide RNA within potato *StALS1* and -*2* genes. A single nucleotide polymorphism (lowercase) exists in the gRNA746 target site of *StALS2* but not gRNA751. *Alo*I and *Bsl*I restriction enzyme sites exist in sgRNA target sites of both genes (underlined). Arrows indicate primers used for enrichment PCR and restriction enzyme digestion assays. PAM sequences are in gray. **B.** Modified enrichment PCR assay using potato callus tissue transformed with gRNA746 and gRNA751 CRISPR/Cas reagents. Total genomic DNA was subjected to PCR amplification of the *StALS* target site (bottom image; 448 bp), digested overnight with *Alo*I (lanes 1, 3, 5, 7, 9, 11) or *Bsl*I (lanes 2, 4, 6, 8, 10, 12), and reamplified (top image; 448 bp) to generate an enriched amplicon. Enriched band intensities were normalized by dividing the quantified band intensity of the enriched band by the primary PCR amplicon (Table 4). Positive (+), negative (-) and non-detectable (ND) enriched bands have normalized intensities equal or over 0.5, less than 0.5 and equal or more than 0.05, or less then 0.05, respectively. Diploid (X; lanes 1-6) and tetraploid (D; lanes 7-12) genotypes were tested using both sgRNAs in the conventional 35S (M; lanes 1, 2, 7, 8) and geminivirus LSL (L; lanes 3, 4, 9, 10) T-DNA backbones. Wild-type (wt; lanes 5, 6, 11, 12) genomic DNA was used as negative controls.

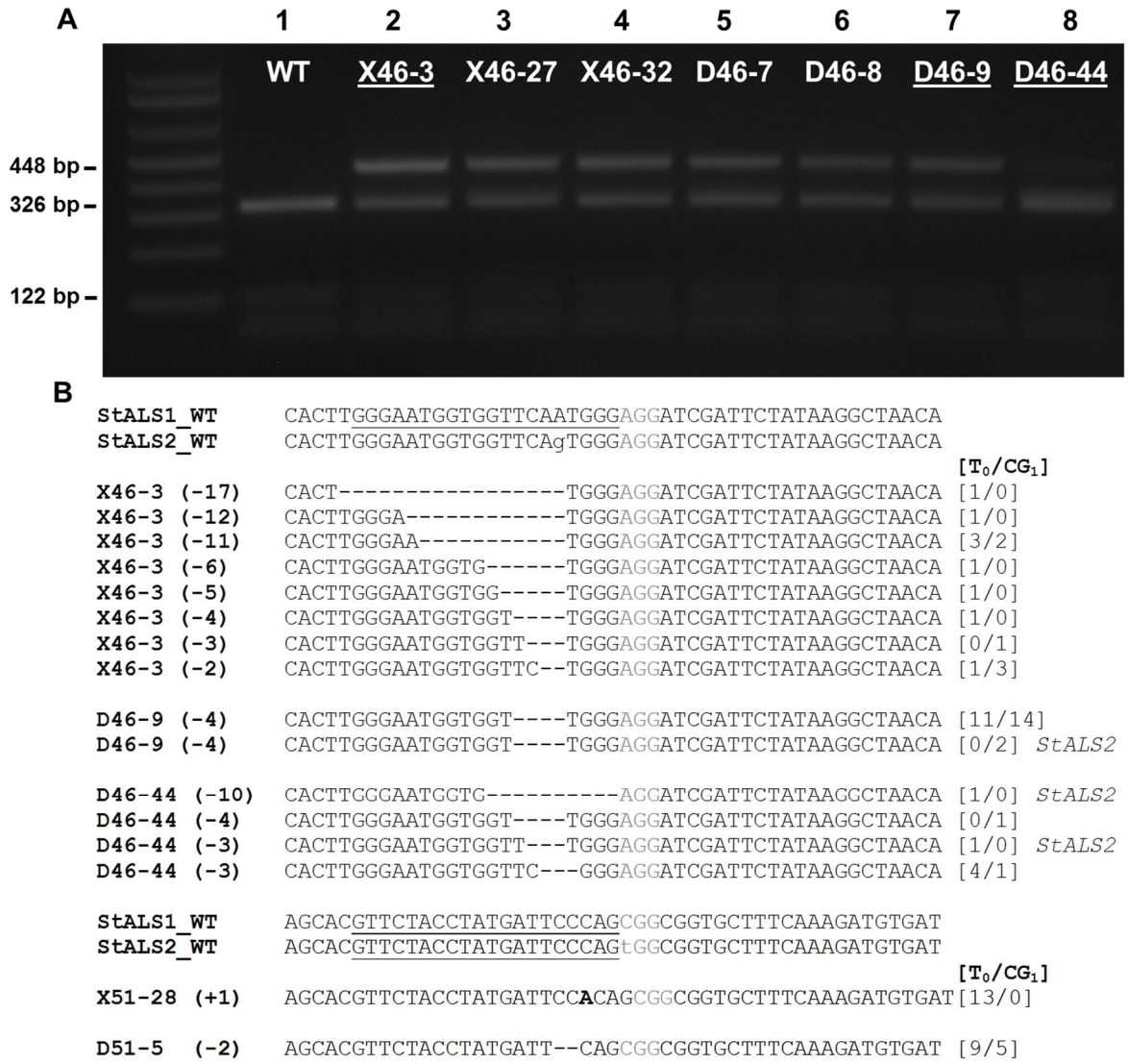

Figure 4. Generation and cloning of targeted mutations in primary events of potato using CRISPR/Cas reagents. A. Restriction enzyme digestion assay of diploid (X; lanes 2-4) and tetraploid (D; lanes 5-8) primary events. Total genomic DNA from regenerated events was subjected to PCR amplification of the *StALS* target site and digested overnight with *Alo*I yielding a 448 bp resistant band and 326 bp and 122 bp digested bands. Wild-type X914-10 (WT; lane 1) and Désirée (Figure 5A) genomic DNA were used as a negative controls. **B.** Cloned targeted mutations in primary events of potato. Diploid (X) and tetraploid (D) events constitutively expressing gRNA746 (46) and gRNA751 (51) CRISPR/Cas reagents were used for cloning. Resistant bands from restriction enzyme digestion assays were excised from 2.0% agarose gels, purified, and subcloned for Sanger sequencing. Sanger reads from each event were aligned to *StALS1* and -*2* wild-type sequence (WT) from each sgRNA target site (gRNA746; top alignments, gRNA751; bottom alignments). The lengths of deletions (-) or insertions (+)

Figure 4 (cont'd). are in parenthesis to the left of each cloned mutation and the number of reads generated in the primary event (T_0) or first clonal generation (CG_1) are in brackets on the right. All targeted mutations were cloned from *StALS1* unless indicated on the right. PAM sequences are in gray.

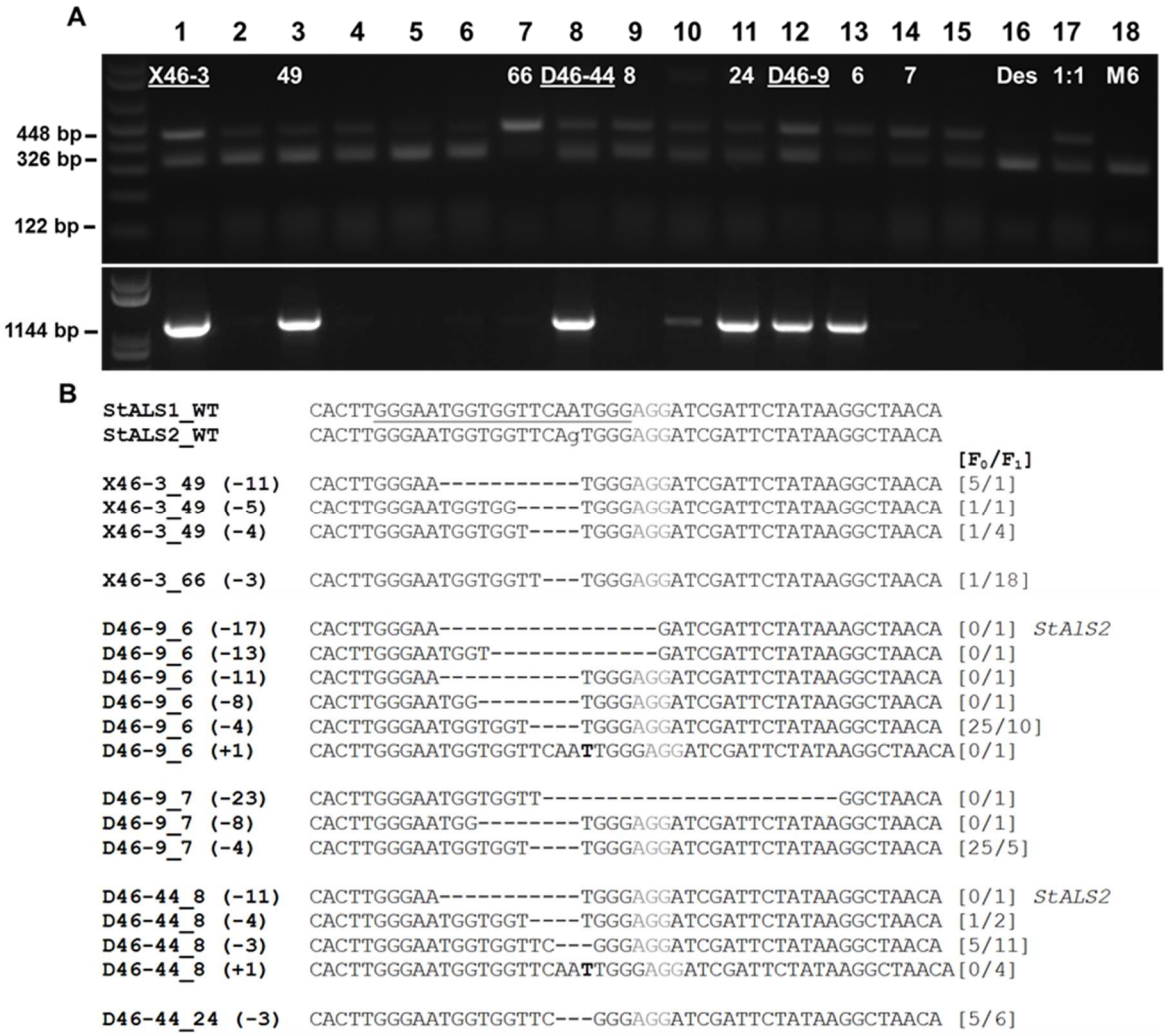

Figure 5. Inheritance of targeted mutations and Cas9 in progeny of primary CRISPR/Cas events. Three primary events with cloned targeted mutations (lanes 1, 8 and 12; underlined) were used to generate genetic populations to assess inheritance of targeted mutations. The diploid event (lane 1; X46-3) was crossed to an inbred diploid line, M6 (lane 18) as the female parent while tetraploid events (lanes 8, 12; D46-44, D46-9) were selfed. Six progeny from the X46-3 population (lanes 2-7) and three progeny from the D46-44 (lanes 9-11) and D46-9 (lanes 13-15) populations were assessed for **A**) targeted mutations using a restriction digestion assay (top gel) and inheritance of Cas9 (bottom gel) and used for **B**) cloning targeted mutations using previously described methods (Figures 4 and 10). The PCR assay used for detecting Cas9 (**A**; bottom gel) produced a 1144 bp amplicon with each lane corresponding to the top gel and is further described in Figure 11. Wild-type Désirée and M6 were used as negative controls (lanes 16 and 18, respectively) and a 1:1 template mixture with wild-type and mutated DNA was used as a positive control (lane 17). The lengths of deletions (-) or insertions (+) of the targeted mutations in progeny (**B**) are in parenthesis

Figure 5 (cont'd). to the left of each cloned mutation and the number of reads generated in the primary event (F_0) or individual progeny (F_1) are in brackets on the right. All targeted mutations were aligned to wild-type sequence and cloned from *StALS1* unless indicated on the right. PAM sequences are in gray.

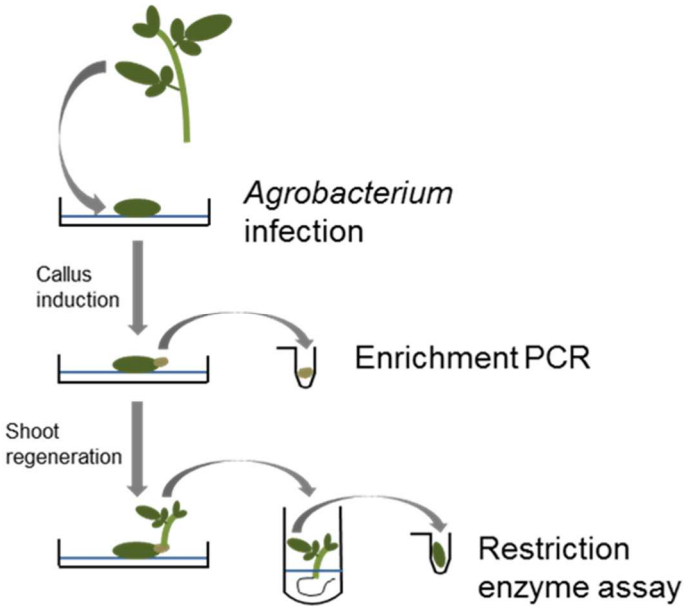

Figure 6. Schematic for delivering CRISPR/Cas reagents to potato leaf explants and detecting targeted mutations in callus and primary events.

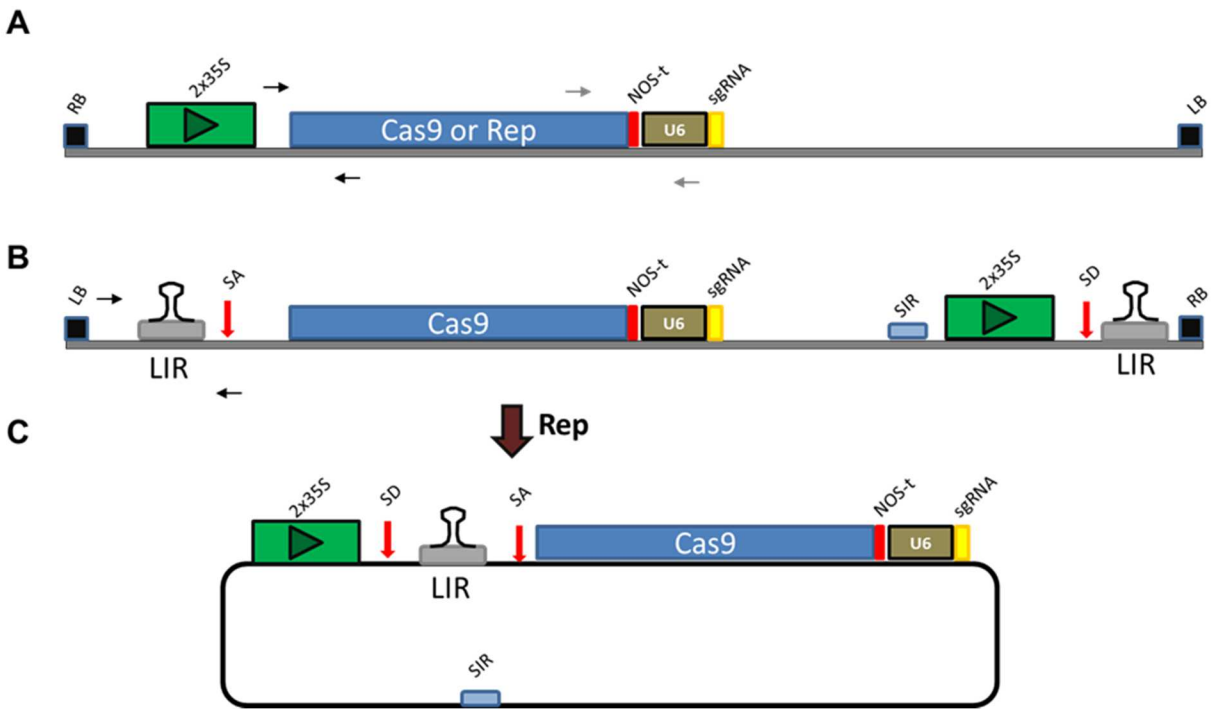

Figure 7. **Binary T-DNA vector constructs used for expressing CRISPR/Cas reagents and PCR primers used for detecting reagents. A.** Conventional 35S T-DNA backbone (pMDC32; (Curtis and Grossniklaus, 2003)) used to express Cas9 and geminivirus replicase (Rep) coding sequences (Baltes *et al*., 2014). Black and gray arrows represent PCR primers used for detecting Rep and Cas9, respectively (Figures 5, 9 and 11). **B.** Geminivirus LSL backbone (pLSL; (Baltes *et al*., 2014)) with *cis*-acting viral elements, long-intergenic region (LIR) and short-intergenic region (SIR) in an L-S-L arrangement with splicing acceptor (SA) and splicing donor (SD) sites flanking the transcribed region. Black arrows represent PCR primers used for detecting the LSL backbone (Figure 9). **C.** Upon co-transformation with Rep, the viral replicon is released and replicated to a high copy number within the plant nucleus. A doubled 35S promoter (2x35S) was used to drive Cas9 and Rep expression with a nopaline synthase transcriptional terminator (NOS-t). Single-guide RNA (sgRNA) expression is driven by an *Arabidopsis* U6 promoter (U6). T-DNAs are delineated by left (LB) and right (RB) borders and contain a selectable hygromycin-resistance marker gene.

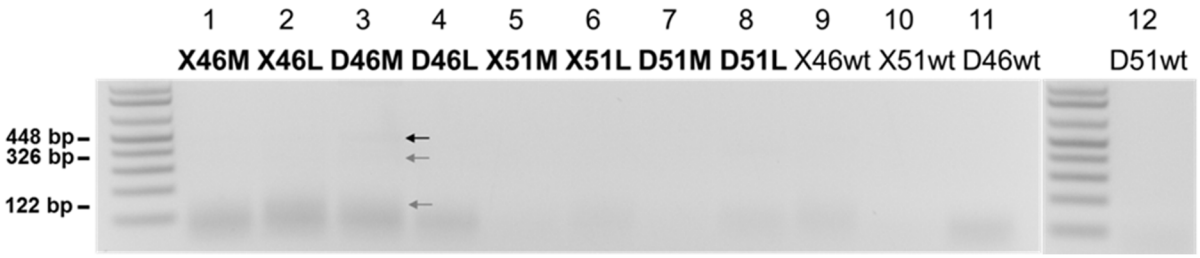

Figure 8. Enrichment PCR assay using potato callus tissue transformed with gRNA746 and gRNA751 CRISPR/Cas reagents. Total genomic DNA was digested overnight with *Alo*I (lanes 1, 2, 3, 4, 9, 11) or *Bsl*I (lanes 5, 6, 7, 8, 10, 12), used for PCR amplification of the *StALS* target site, and redigested overnight to generate an enriched amplicon. For gRNA746, an enriched amplicon of 448 bp (black arrow) and digest products of 326 bp and 122 bp (gray arrows) were generated. Diploid (X; lanes 1-2, 5-6, 9-10) and tetraploid (D; lanes 3-4, 7-8, 11-12) genotypes were tested using both sgRNAs in the conventional 35S (M; lanes 1, 3, 5, 7) and geminivirus LSL (L; lanes 2, 4, 6, 8) T-DNA backbones. Wild-type (wt; lanes 9-12) genomic DNA was used as negative controls.

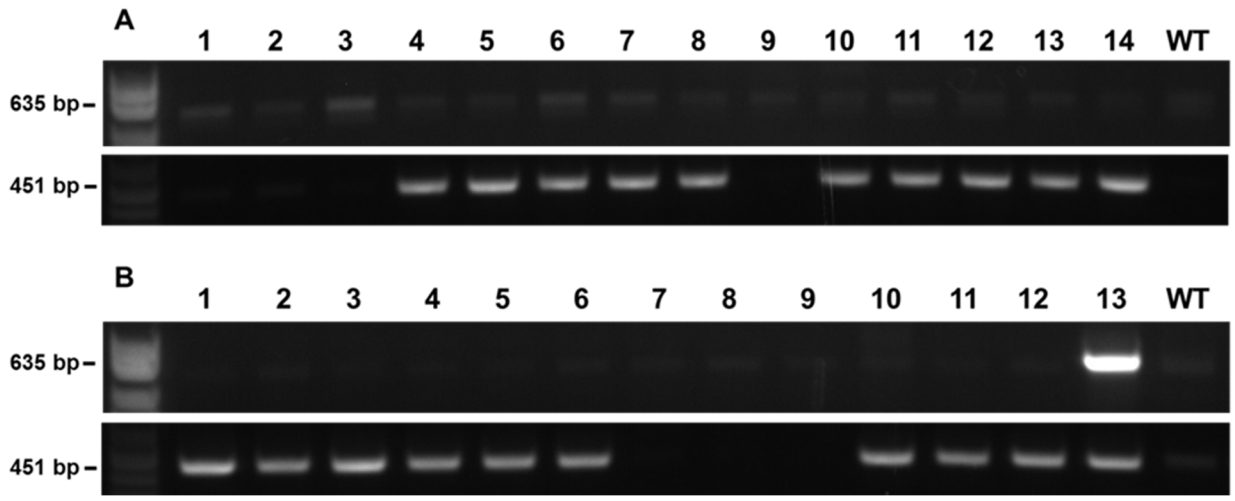

Figure 9. Detection of T-DNA integration in primary events. A PCR assay was used to detect integration of LSL T-DNA and Rep T-DNA in co-transformed events of diploid (**A**; X914-10) and tetraploid (**B**; Désirée) potato. (Table 5; LSL). Primers specific to the LSL T-DNA and Rep T-DNA were used for top and bottom images of each panel, respectively (Figure 7). Expected amplicons were 635 bp and 451 bp in size for LSL and Rep T-DNA, respectively and were generated using Phusion High-Fidelity DNA Polymerase (NEB, Ipsich, MA) and total genomic DNA from primary event leaf tissue. Lane numbering follows the order of events listed in Table 5 (LSL) with lanes 1-13 (X914-10) and lanes 1-12 (Désirée) generated using gRNA746 and lane 14 (X914-10) and lane 13 (Désirée) generated using gRNA751. Wild-type (WT) controls are shown for each genetic background.

Figure 10. Additional targeted mutations in primary events of potato using CRISPR/Cas reagents. Cloned mutations from diploid (X) and tetraploid (D) events constitutively expressing gRNA746 (46) CRISPR/Cas reagents are shown. Sanger reads from each event were aligned to *StALS1* and -*2* wild-type sequence (WT) from the gRNA746 target site. The lengths of deletions (-) or insertions (+) are in parenthesis to the left of each cloned mutation and the number of reads generated in the primary event (T_0) or first clonal generation (CG_1) are in brackets on the right. All targeted mutations were cloned from *StALS1* unless indicated on the right. PAM sequences are in gray.

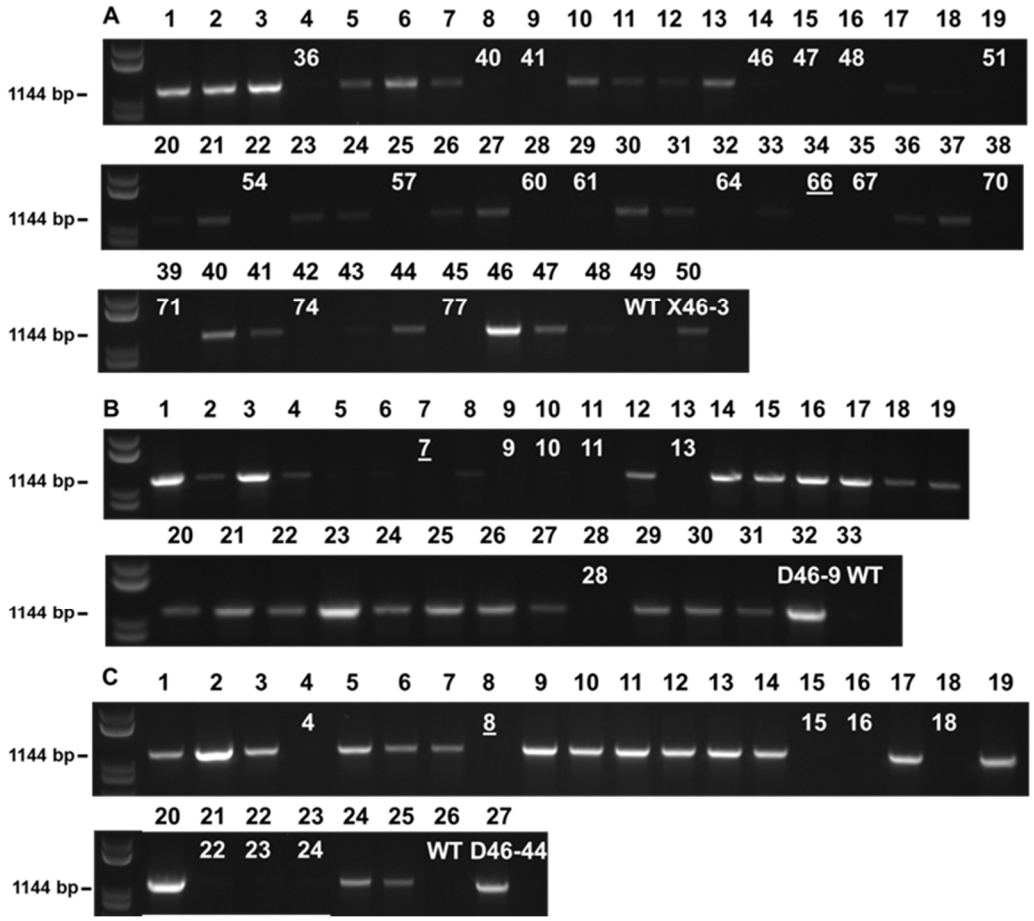

Figure 11. Inheritance of Cas9 in progeny of primary events. A PCR assay was used to detect Cas9 in progeny of diploid (**A**; X46-3) and tetraploid (**B** and **C**; D46-9 and D46-44, respectively) primary events (Figure 5 and Table 3). Primers specific to Cas9 and the *Arabidopsis* U6 promoter were used to generate a 1144 bp expected amplicon (Figure 7; gray arrows). The expected amplicon was generated using GoTaq® Green Master Mix (Promega, Madison, WI) and total genomic DNA from progeny (**A**; lanes 1-48, **B**; lanes 1-31, **C**; lanes 1-25) and primary events (**A**; lane 50, **B**; lane 32, **C**; lane 27). Wild-type (WT) controls are shown for each genetic background and underlined progeny were used for targeted mutation cloning (Figure 5, Table 3).

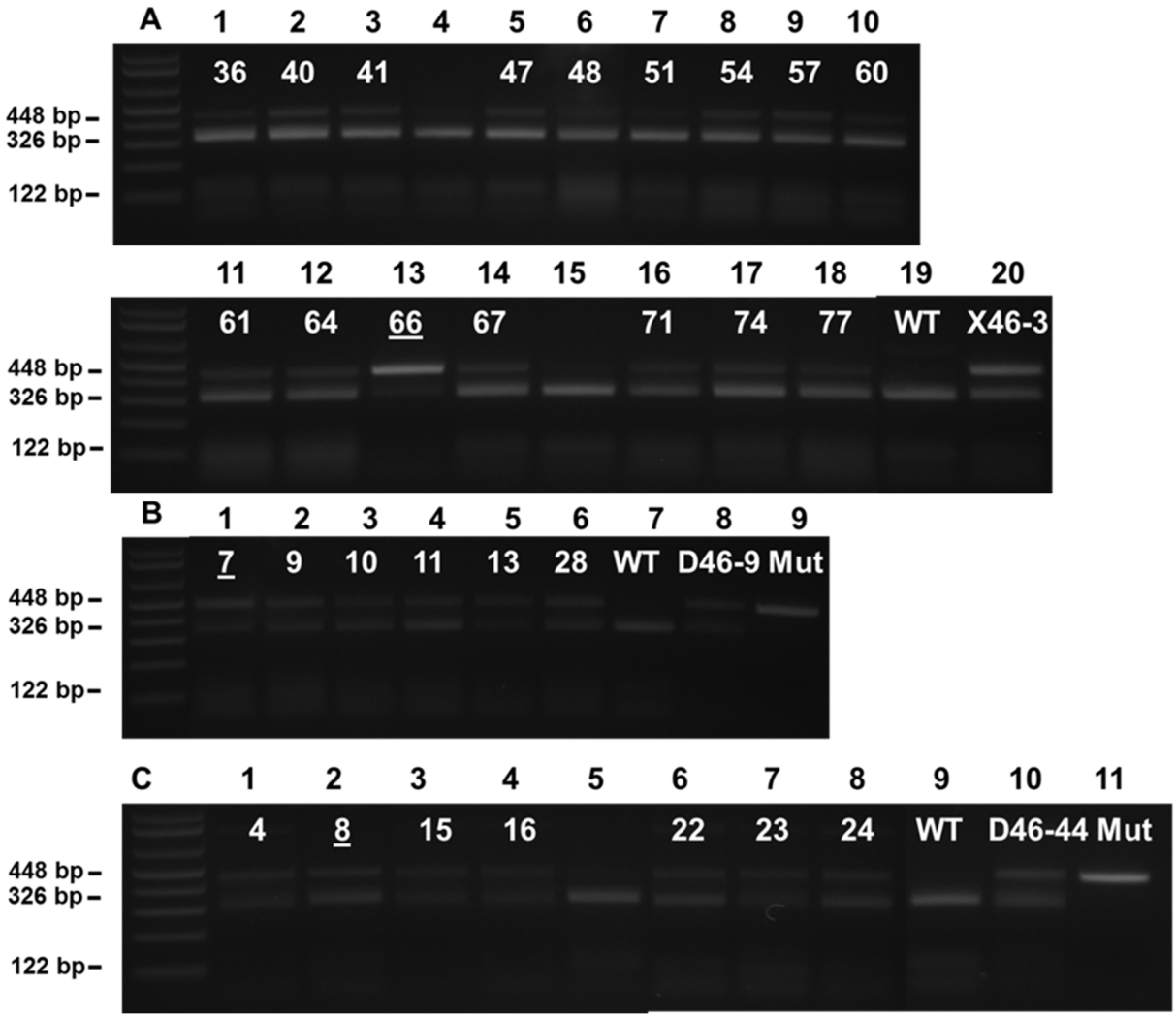

Figure 12. Inheritance of targeted mutations in progeny of primary events. A restriction enzyme digestion assay was used to detect targeted mutations in progeny of diploid (**A**; X46-3) and tetraploid (**B** and **C**; D46-9 and D46-44, respectively) primary events as previously described (Figure 5, Table 3). Primary amplicons were generated from progeny (**A**; lanes 1-18, **B**; lanes 1-6, **C**; lanes 1-8) and primary events (**A**; lane 20, **B**; lane 8, **C**; lane 10). Wild-type (WT) controls are shown for each genetic background. Mutant (Mut) controls were generated using mutant template DNA.

REFERENCES

REFERENCES

Baltes NJ, Gil-Humanes J, Cermak T, Atkins PA, Voytas DF. DNA replicons for plant genome engineering. Plant Cell. 2014;26: 151–163. doi:10.1105/tpc.113.119792

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

Brooks C, Nekrasov V, Lippman ZB, Eck J Van. Efficient gene editing in tomato in the first generation using the clustered regularly interspaced short palindromic repeats/CRISPR-associated9 system. Plant Physiol. 2014; doi:10.1104/pp.114.247577

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013;339: 819–23. doi:10.1126/science.1231143

Curtin SJ, Voytas DF, Stupar RM. Genome engineering of crops with designer nucleases. Plant Genome J. 2012;5: 42–50. doi:10.3835/plantgenome2012.06.0008

Curtin SJ, Zhang F, Sander JD, Haun WJ, Starker C, Baltes NJ, et al. Targeted mutagenesis of duplicated genes in soybean with zinc-finger nucleases. Plant Physiol. 2011;156: 466–473. doi:10.1104/pp.111.172981

Curtis M, Grossniklaus U. A gateway cloning vector set for high-throughput functional analysis of genes in planta. Plant Physiol. 2003;133: 462–469. doi:10.1104/pp.103.027979

Farboud B, Meyer BJ. Dramatic enhancement of genome editing by CRISPR/Cas9 through improved guide RNA design. Genetics. 2015;199: 959–971. doi:10.1534/genetics.115.175166

Felcher KJ, Coombs JJ, Massa AN, Hansey CN, Hamilton JP, Veilleux RE, et al. Integration of two diploid potato linkage maps with the potato genome sequence. PLoS One. 2012;7: e36347. doi:10.1371/journal.pone.0036347

Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. PNAS. 2012;109: E2579–E2586. doi:10.1073/pnas.1208507109

Hunter P. “Genetically Modified Lite” placates public but not activists: new technologies to manipulate plant genomes could help to overcome public concerns about GM crops. EMBO Rep. 2014;15. doi:10.1002/embr.201338365

Jansky SH, Chung YS, Kittipadukal P. M6: A diploid potato inbred line for use in breeding and genetics Research. J Plant Regist. 2014; doi:10.3198/jpr2013.05.0024crg

Jia H, Wang N. Targeted genome editing of sweet orange using Cas9/sgRNA. PLoS One. 2014;9: e93806. doi:10.1371/journal.pone.0093806

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

Klümper W, Qaim M. A meta-analysis of the impacts of genetically modified crops. PLoS One. 2014;9: e111629. doi:10.1371/journal.pone.0111629

Li T, Liu B, Spalding MH, Weeks DP, Yang B. High-efficiency TALEN-based gene editing produces disease-resistant rice. Nat Biotechnol. 2012;30: 390–392. doi:10.1038/nbt.2199

Lor VS, Starker CG, Voytas DF, Weiss D, Olszewski NE. Targeted mutagenesis of the tomato *PROCERA* gene using transcription activator-like effector nucleases. Plant Physiol. 2014; doi:10.1104/pp.114.247593

Miller JC, Holmes MC, Wang J, Guschin DY, Lee Y-L, Rupniewski I, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol. 2007;25: 778–785. doi:10.1038/nbt1319

Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, et al. A TALE nuclease architecture for efficient genome editing. Nat Biotechnol. 2011;29: 143–148. doi:10.1038/nbt.1755

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

Palmgren MG, Edenbrandt AK, Vedel SE, Andersen MM, Landes X, Østerberg JT, et al. Are we ready for back-to-nature crop breeding? Trends Plant Sci. 2014; 1–10. doi:10.1016/j.tplants.2014.11.003

Paz M, Veilleux R. Influence of culture medium and *in vitro* conditions on shoot regeneration in *Solanum phureja* monoploids and fertility of regenerated doubled monoploids. Plant Breed. 1999;118: 53–57. Available: http://onlinelibrary.wiley.com/doi/10.1046/j.1439-0523.1999.118001053.x/abstract

PGSC. Genome sequence and analysis of the tuber crop potato. Nature. 2011;475: 189–195. doi:10.1038/nature10158

Qi Y, Zhang Y, Zhang F, Baller JA, Cleland SC, Ryu Y, et al. Increasing frequencies of site-specific mutagenesis and gene targeting in *Arabidopsis* by manipulating DNA repair pathways. Genome Res. 2013;23: 547–554. doi:10.1101/gr.145557.112

Shukla VK, Doyon Y, Miller JC, DeKelver RC, Moehle EA, Worden SE, et al. Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. Nature. 2009;459: 437–441. doi:10.1038/nature07992

Smith J, Schloss J, Mazur B. Functional expression of plant *acetolactate synthase* genes in *Escherichia coli*. PNAS. 1989;86: 4179–4183. Available: http://www.pnas.org/content/86/11/4179.short

Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature. 2014;507: 62–67. doi:10.1038/nature13011

Stoddard B. Homing endonucleases: from microbial genetic invaders to reagents for targeted DNA modification. Structure. 2011;19: 7–15. doi:10.1016/j.str.2010.12.003

Takeuchi R, Lambert AR, Mak AN-S, Jacoby K, Dickson RJ, Gloor GB, et al. Tapping natural reservoirs of homing endonucleases for targeted gene modification. PNAS. 2011;108: 13077–82. doi:10.1073/pnas.1107719108

Verhoeven HA, Dijkhuis P. Introduction of foreign genes into potato cultivars Bintje and Desiree using an *Agrobacterium tumefaciens* binary vector. Plant Cell Rep. 1988;7: 47–50.

Waibel F, Filipowicz W. *U6* genes of *Arabidopsis* are transcribed by RNA polymerase III but contain the same two upstream promoter elements as RNA polymerase II-transcribed U-snRNA genes. Nucleic Acids Res. 1990;18: 3451–58. Available: http://nar.oxfordjournals.org/content/18/12/3451.short

Wendt T, Holm PB, Starker CG, Christian M, Voytas DF, Brinch-Pedersen H, et al. TAL effector nucleases induce mutations at a pre-selected location in the genome of primary barley transformants. Plant Mol Biol. 2013;83: 279–285. doi:10.1007/s11103-013-0078-4

Xu H, Xiao T, Chen C, Li W, Meyer C, Wu Q, et al. Sequence determinants of improved CRISPR sgRNA design. Genome Res. 2015; gr.191452.115. doi:10.1101/gr.191452.115

CHAPTER 3: GEMINIVIRUS-MEDIATED GENOME EDITING IN POTATO (*SOLANUM TUBEROSUM* L.) USING SEQUENCE-SPECIFIC NUCLEASES

Nathaniel M. Butler[1], Nicholas J. Baltes[2], Daniel F. Voytas[2], David S. Douches[1]

Paper to be submitted to the Plant Biotechnology Journal

[1]Department of Plant, Soils and Microbial Sciences, Michigan State University, East Lansing, Michigan 48824, USA

[2]Department of Genetics, Cell Biology and Development and Center for Genome Engineering, University of Minnesota, Minneapolis, Minnesota 55455, USA

Summary

Genome editing using sequence-specific nucleases (SSNs) is rapidly becoming a standard tool for genetic engineering in crop species. The implementation of zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and CRISPR/Cas (clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated systems (Cas)) for inducing double-strand breaks enables targeting of virtually any sequence for modification. Targeted mutagenesis via nonhomologous end-joining has been demonstrated extensively as being the preferred DNA repair pathway in plants. However, gene targeting via homologous recombination remains more elusive but could be a powerful tool for directed DNA repair. To overcome barriers associated with gene targeting, a geminivirus replicon (GVR) was developed in tobacco and was found to drastically improve gene targeting efficiencies of an integrated reporter. To test the efficacy of GVRs in a crop species, SSNs targeting

the potato *ACETOLACTATE SYNTHASE1* (*ALS1*) gene were developed and incorporated into conventional 35S and GVR T-DNAs. A repair template designed to incorporate herbicide-inhibiting point mutations within the *ALS1* locus was included within these T-DNAs and used in transformation experiments. Transformed events modified with GVR reagents held both point mutations that were capable of supporting a reduced herbicide susceptibility phenotype while events modified from the 35S reagent held no detectable mutations and were similar to wild-type. Regeneration of transformed events led to improved detection of point mutations that supported a stronger phenotype. These results support the use of geminiviruses for delivering genome editing reagents in plant species and an approach to gene targeting in a vegetatively propagated species.

Introduction

Genome editing is rapidly becoming a standard tool for genetic improvement in crop species. Major advancements in sequence-specific nuclease (SSN) technology, such as zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and CRISPR/Cas (clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated systems (Cas)) has enabled application of genome editing in a range of crops species and allowed development of transgene-free genetically engineered crops (Palmgren *et al.*, 2014; Lusser *et al.*, 2012b). Genome editing allows generation of transgene-free genetically modified events by utilizing DNA

repair pathways and modifying target DNA in *trans* without relying on stable integration of genome editing reagents (Curtin *et al.*, 2012).

One method of genome editing involves the induction of DNA double-strand breaks by expressing SSNs *in vivo*. Once double-strand breaks have formed, DNA repair follows two major pathways: nonhomologous end-joining (NHEJ) and homologous recombination. NHEJ repair is often imprecise and can result in the introduction of insertions or deletions at the break site which can be used for targeted mutagenesis of endogenous genes or reporters (Zhang *et al.*, 2013; Li *et al.*, 2012). Homologous recombination is a more precise pathway that utilizes a DNA template for repair. The ability to manipulate an exogenously supplied homologous repair template allows incorporation of new sequence at or near a break site in a process referred to hereafter as gene targeting (Bibikova *et al.*, 2003; Puchta *et al.*, 1996). Gene targeting provides many advantages over targeted mutagenesis but occurs at low frequencies in plant cells (Shukla *et al.*, 2009; Terada *et al.*, 2002; Lee *et al.*, 1990).

Plant viruses have the potential to become powerful tools for genome editing in plants. For decades, mammalian viruses been used for gene therapy in humans by directing high expression of genome editing reagents in pathological tissues (Kotterman and Schaffer, 2015; Giacca and Zacchigna, 2012). More recently, plant viruses such as the *Tobacco Rattle Virus* (*TRV*) have been used with *Agrobacterium tumefaciens* to efficiently deliver RNA interference (RNAi) reagents, ZFNs, and single-guide RNA (sgRNA) used in CRISPR/Cas for genetic modification in Solanaceous species, petunia (*Petunia hybrid*) and tobacco (*Nicotiana tabacum*) (Marton *et al.*, 2010; Sha *et al.*, 2013; Brigneti *et al.*, 2004; Burch-Smith *et al.*, 2006; Ali *et al.*, 2015). The limitations in

carrying capacity of *TRV* and other RNA viruses prevent their use beyond expression of relatively small SSNs and sgRNAs and are unable to efficiently deliver large DNA repair templates.

Geminiviruses may be able to overcome the limitations of RNA viruses by allowing a larger carrying capacity and producing a DNA replicon capable of acting as a repair template for gene targeting (Baltes *et al.*, 2014; Richter *et al.*, 2014). As a DNA virus, geminiviruses such as the *Bean Yellow Dwarf Virus* (*BeYDV*) replicate within the plant nucleus through a double-strand intermediate to a high-copy number using host polymerases (Liu *et al.*, 1998; Liu *et al.*, 1997). The *BeYDV* genome encodes relatively few *cis*-acting elements and requires only one geminivirus *trans*-acting element for replication (Rep/RepA), making it readily amendable to plant transformation. Previous studies have combined the essential *cis*-acting elements of the *BeYDV*, the long intergenic region (LIR), the short intergenic region (SIR) into a LIR-SIR-LIR (LSL) orientation within a T-DNA backbone (pLSL T-DNA) with Rep/RepA delivered on a separate construct to generate geminivirus replicons (GVRs) within tobacco leaf cells (Mor *et al.*, 2003; Zhang and Mason, 2006). More recently, co-delivery of a pLSL T-DNA containing both SSNs and repair template with Rep/RepA was shown to coordinate SSN expression with increased repair template copy number in the plant nucleus and improve the gene targeting efficiency of an integrated reporter in tobacco (Baltes *et al.*, 2014).

In this study, the efficacy of GVRs in concert with SSNs was tested in wild-type and constitutively expressing Rep/RepA mutant genotypes of potato (*Solanum tuberosum* Group Tuberosum L.) by modifying both reporter and endogenous targets.

Transformed events supported reduced herbicide susceptibility phenotypes and gene targeting modifications incorporated using a repair template. Regeneration of transformed events under higher selection for gene targeting modifications resulted in enhanced levels of gene targeting in regenerated events and further reduced susceptibility to herbicide.

Results

GVR delivery and heterologous protein expression

A deconstructed mild strain of the *BeYDV* has recently been used to construct an *Agrobacterium* T-DNA capable of delivering GVRs to plant cells (Baltes *et al.*, 2014). This so called pLSL T-DNA contains two viral *cis*-acting elements required for *BeYDV* replication, the LIR and SIR with a single *trans*-acting element, Rep/RepA delivered on a separate T-DNA (Rep). The L-S-L arrangement of the LIR and SIR elements on the pLSL T-DNA allows for heterologous protein expression driven by a cauliflower mosaic virus 35S promoter (35S) upon replication release of the circular GVR. Concomitant expression of Rep/RepA should facilitate GVR replication release by acting on the stem-loop structure of the LIR and initiating rolling-circle replication (Stenger *et al.*, 1991) (Figure 13A).

To test if GVRs could replicate and express heterologous protein in potato leaf cells, a pLSL T-DNA expressing β-glucuronidase (pLSL-GUS) was used for *Agrobacterium*-mediated transformation of wild-type tetraploid (cv Désirée) potato leaf explants and whole explants were stained for GUS activity (Figure 13A-B). When

concomitantly delivered with Rep, strong GUS staining was seen in cells of wounded areas (Figure 13B; inset). However, in the absence of Rep, little staining was observed. These observations suggest Rep/RepA induces efficient heterologous protein expression and replication of GVRs in potato cells.

To further investigate this possibility, wounded areas were sampled from Désirée and wild-type diploid (X914-10) potato leaf explants transformed with pLSL-GUS for GUS activity quantification or PCR detection of circularized GVRs after 2, 5 and 7 days post inoculation (dpi) in the presence or absence of Rep (Figure 18). Similar to GUS staining experiments, circularized GVRs were most evident in Désirée wounded areas concomitantly transformed with Rep compared with wounded areas without Rep/RepA expression 7 dpi (Figures 13C and 18B). GVR replication in both Désirée and X914-10 was induced 5 dpi consistent with observations made in tobacco (Figure 18) (Zhang and Mason, 2006; Baltes *et al.*, 2014). Following GVR induction, GUS activity in wounded areas of both genotypes increased 7 dpi consistent with GUS staining experiments (Figures 13B and 18). Unexpectedly, GVR circularization and GUS activity could also be detected in wounded areas of both genotypes in the absence of Rep/RepA (Figure 18). No obvious patterns in protein expression or circularization were seen in these treatments and are putatively due to background factors such as host nucleases and transposons (Desai and Shankar, 2003; Dooner and Weil, 2007).

The requirement of Rep/RepA expression for efficient GVR replication and heterologous protein production led us to question if a constitutively expressing Rep/RepA mutant could be developed in potato. To test this, the constitutively expressing Rep T-DNA was used for stable transformation in X914-10 using

hygromycin selection for T-DNA integration. Twenty-eight hygromycin-resistant events were propagated in tissue-culture and used in GVR replication assays to identify an event capable of supporting high GVR replication (Figure 19). Among the events evaluated, event D52 displayed both strong staining (data not shown) and high GVR replication and was chosen for further experimentation. Unlike wild-type X914-10 and Désirée, D52 supported a significant 3-fold induction of GVR replication 7 dpi (vs 5 dpi) ($P<0.02$) and gradual heterologous protein production peaking at 7 dpi, similar to wild-type (Figures 13D and 18) ($P<0.05$) compared to 2 dpi treatments. The difference in GVR replication between wild-type X914-10 and D52 is likely due to variation in Rep/RepA expression and interactions with the pLSL-GUS T-DNA (Stenger *et al.*, 1991). To simply subsequent experiments, we chose to sample wounded areas 7 dpi and focus on X914-10 genotypes in stable gene targeting experiments.

SSN activity and development of SSNs

SSNs have been tested in a number of plant species but data in Solanaceous species, such as potato is limited (Clasen *et al.*, 2015; Lor *et al.*, 2014; Brooks *et al.*, 2014; Sawai *et al.*, 2014; Li *et al.*, 2013). To demonstrate SSN activity and their utility for gene targeting in potato cells, three major SSN platforms, ZFN, TALEN and CRISPR/Cas were tested in a single-strand annealing assay (SSA) (Zhang *et al.*, 2013). To construct the SSA reporter T-DNA (pSSA), the GUS coding sequence was disrupted by a 250 base pair (bp) direct repeat separated by a 60 bp target site for Zif268 (ZFN) and driven by a cauliflower mosaic virus (35S) promoter. Upon formation of a double-strand break within the target site, the SSA direct repeat is used to repair the reporter and reconstitute the GUS coding sequence, allowing GUS expression (Figure 14A). To

test this reporter, leaf explants were transformed with pSSA and stained for GUS activity (Figure 20B). Strong GUS staining was observed in cells of wounded areas concomitantly transformed with ZFN, but not in the absence of ZFN (Figure 20B; inset). In addition to ZFN, both TALEN (TALENz) and CRISPR/Cas (CRISPRz) reagents were designed to target the Zif268 target site and used in the assay (Figure 20A). All three SSN platforms resulted in an approximate 4-fold increase in reporter activity. These results suggest each major SSN platform is active in potato cells and could putatively be used for gene targeting.

The potato *ACETOLACTATE SYNTHASE1* (*ALS1*) gene (PGSC0003DMG400034102) was chosen as an endogenous target for gene targeting. The 3' end of the *ALS1* coding sequence contains two point mutation sites (W563L and S642T) characterized in *Arabidopsis* that confer reduced susceptibility to ALS-inhibiting herbicides (Sathasivan *et al.*, 1991). To target this region, one TALEN and one CRISPR/Cas reagent were designed to target each point mutation site (W563L; TALEN(-), CRISPR(-) and S642T; TALEN, CRISPR). The further downstream S642T mutation was focused on due to its proximity to the end of the *ALS1* coding sequence and the opportunity to incorporate new sequence downstream of *ALS1*. To test the activity of TALEN and CRISPR reagents, a SSA reporter was constructed incorporating a 60 bp S642T target site and W563L reagents were used as negative controls (Figure 14A). TALEN and CRISPR reagents resulted in significant 5 and 2-fold increases, respectively in reporter activity compared to negative controls, demonstrating their activity in potato cells (Figure 14B) ($P<0.05$).

The reporter activity using the TALEN reagents suggested TALENs may be capable of inducing targeted NHEJ mutations in transformed events. To investigate this possibility, X914-10 and D52 genotypes were transformed with pLSL-TALEN and p35S-TALEN T-DNAs, respectively and hygromycin-resistant events were screened for targeted NHEJ mutations (not shown). None of the 26 and 42 events generated from p35S-TALEN or pLSL-TALEN contained NHEJ mutations at or above the 6.25% mutation detection limit. The lack of detectable NHEJ mutations in events transformed with TALEN reagents could be due to the formation of somatic mutations that fall below detection limits (Lor *et al.*, 2014).

GVR-mediated gene targeting

The demonstration of GVR replication and heterologous protein expression along with SSN activity suggested GVRs were capable of delivering gene targeting reagents to potato cells. In order to test gene targeting efficiencies, a previously established reporter (referred here as pGUPTII) was concomitantly transformed with conventional (p35S) and GVR (pLSL) ZFN gene targeting reagents to potato leaf explants and evaluated for GUPTII reporter repair (GUSNPTII) (Figure 15A). The GUPTII reporter was designed with a 600 bp deletion within the GUS:NptII fusion coding sequence and 60 bp Zif268 target sequence in place of the missing sequence. Upon gene targeting, the GUPTII repair template (RT) with the missing 600 bp and flanking homologous sequence is capable of reconstituting the GUS:NptII coding sequence and allowing GUS expression and PCR detection of GUSNptII (Wright *et al.*, 2005).

To determine if gene targeting is enhanced by GVR reagent delivery, pLSL-ZFN/RT and 35S-ZFN/RT reagents were concomitantly delivered with the pGUPTII reporter in the presence of Rep and wounded areas analyzed for GUS activity and GUSNptII detection (Figures 15 and 21). Rep was delivered with both conventional and GVR reagents to normalize Rep/RepA pleiotropic effects (Baltes *et al.*, 2014). GVR delivery of the gene targeting reagents resulted in a significant 8-fold increase in reporter activity over the pGUPTII-only control while no significant increase was observed using the conventional gene targeting reagents (Figure 15B). These results were further supported by GUPTNptII detection (Figure 21B). Interestingly, GVR reagents were capable of significant increases in gene targeting, regardless of Rep (Figure 15B). These results suggest early GVR replication and heterologous protein expression previously discovered (2d and 5d; Figure 18B) were sufficient for enhancing gene targeting efficiencies at 7 dpi (Figure 15B and 21A), even in the absence of Rep.

The results of the GUPTII assay suggest GVRs are capable of delivering gene targeting reagents to potato cells and facilitating efficient gene targeting. However, it is unclear if the GVRs are affecting gene targeting by enhancing SSN protein levels and subsequent double-strand break frequencies, or providing additional DNA template for repair. To test this, pLSL T-DNAs were constructed with ZFN replaced with green fluorescent protein (GFP) (pLSL-GFP/RT) or the GUPTII repair template replaced with heterologous tobacco *ALS* sequence (pLSL-ZFN/RT(-)). Delivery of each GVR with Rep failed to facilitate efficient gene targeting suggesting both SSN and repair template components are necessary but independently no sufficient for GVR-mediated enhancement of gene targeting (Figure 21). These results demonstrate the essential

nature of SSN activity for efficient gene targeting in plant cells and reflect observations made in tobacco (Puchta *et al.*, 1996; Baltes *et al.*, 2014).

GVR-mediated gene targeting of the ALS1 locus

The ability of GVRs to modify the GUPTII reporter in potato and tobacco leaf cells suggests GVRs could also modify an endogenous locus (Baltes *et al.*, 2014). The *ALS1* locus provides an ideal target for gene targeting given the ubiquitous nature of *ALS* expression and the availability of ALS-specific point mutations conferring reduced herbicide susceptibility in a broad range of plant species (Endo and Toki, 2013). SSNs developed to target the S642T point mutation site in the potato *ALS1* gene were cloned into both p35S and pLSL T-DNAs along with an *ALS1* repair template. The *ALS1* repair template, referred here as RT1 was first constructed to contain both W563L and S642T point mutations and silent mutations within TALEN and sgRNA binding sites to prevent SSN off-targeting of the repair template. Left and right homology arms were designed to cover the *ALS1* coding sequence stopping at a non-functional start codon (1.6 kilobases (kb)) and 1 kb downstream of the *ALS1* coding sequence, respectively. A *Bam*HI restriction enzyme site at the end of the *ALS1* coding sequence was used for both screening for gene targeting modifications in transformed events (Figure 22) and constructing a second *ALS1* repair template, RT2 which includes a T2A:NptII translational fusion (Figure 16). RT2 was also used to construct an *ALS1* transgene which incorporates all the modifications in RT1 and -2 but contains a functional *ALS1* coding sequence (*ALSm*) that is driven by a 2.5 kb promoter region upstream of the endogenous *ALS1* coding sequence.

Stable plant transformations were conducted using wild-type X914-10 (primary) for p35S reagents and mutant D52 (secondary) for pLSL reagents to test the efficacy of the gene targeting reagents. D52 was chosen for pLSL transformations to simplify reagent delivery and potentially improve transformation efficiency. Transformations were first conducted with TALEN and RT1 reagents without direct selection for gene targeting modifications, relying on the efficiency of the reagents (Figure 22). Primary transformations were conducted using p35S-TALEN/RT1 and secondary transformations with pLSL-TALEN/RT1. Hygromycin-resistant events were screened for both gene targeting modifications and NHEJ targeted mutations using restriction digestion and T7 endonucleaseI (T7EI) assays, respectively. None of the 72 primary and 78 secondary events contained any detectable gene targeting modifications or NHEJ mutations. These results reflect the previous evaluation of the TALEN reagents for producing NHEJ mutations in stable events (data not shown) and supports findings in tobacco that GVRs do not improve NHEJ frequencies (Baltes *et al.*, 2014). Furthermore, the lack of detectable gene targeting modifications within transformed events suggested direct selection for gene targeting and more sensitive detection methods were needed in the event somatic modifications were being formed (Feng *et al.*, 2014).

Recovery of ALS1 modified events using direct selection

In order to improve gene targeting efficiency in transformed events, new gene targeting reagents were developed incorporating RT2 to allow direct selection for gene targeting modifications and to deliver RT2 independently on GVRs using the D52 background (Figure 16A). The T2A:NptII translational fusion used in RT2 allows for

independent function of both ALSm and NptII proteins in modified cells, supporting resistance to both ALS-inhibiting herbicides and kanamycin, respectively. In addition to the GVR pLSL-TALEN/RT2 reagent, a p35S-TALEN/RT2 reagent was also constructed to compare conventional T-DNA and GVR delivery in secondary transformations. Furthermore, a modified pLSL (pLSLm) T-DNA was constructed that incorporates RT2 but does not include a 35S promoter or SSN reagents. Our reasoning was that by delivering RT2 on a GVR (pLSLm) and SSNs on a separate 35S T-DNAs (TALEN and CRISPR), gene targeting efficiency could be improved by altering the coordination of SSN expression and repair template availability. This approach was supported by findings in tobacco where gene targeting efficiencies were significantly enhanced by delivering a repair template on a GVR compared to SSNs (Baltes *et al.*, 2014).

The new gene targeting reagents were tested in D52 in secondary plant transformations using kanamycin selection (50 mg/L; Kan50) (Table 6). Transformations were carried out using pLSL and p35S reagents by transforming pLSL-TALEN/RT2 and p35S-TALEN/RT2 in two replicate experiments, resulting in 12 Kan50 resistant events (Q lines) and 4 Kan50 resistant events (RR lines). Transformations were carried out using the pLSLm T-DNA by transforming pLSLm and either TALEN (13 Kan50 resistant events; P lines) or CRISPR (8 Kan50 resistant events; O lines) in two replicate experiments. Kan50 resistant lines were screened using PCR detection of gene targeting modifications in leaf tissues (Figure 23). Bands generated from Kan50 resistant lines were faint and could not be clearly seen. To improve detection, band quantification was used to identify lines with band intensities at least two-fold higher than internal controls. Using this criterion, 41.7% and 12.5% of the Q and O lines,

respectively were determined as being positive for gene targeting modifications while none of the RR or P lines were above this threshold. To validate the bands generated in the screen, two representative lines, P31 and Q94 were chosen for gene targeting modification cloning (Figure 16B). PCR products from both representative lines were cloned and sequenced. In both cases, new sequence incorporated by the repair template was identified and linked to the template-locus junction. These results confirm the validity of bands used for evaluating gene targeting events and the likely presence of somatic gene targeting modifications within events.

The confirmation of gene targeting modifications, including W563L and S642T point mutations, and kanamycin resistance of secondary events suggested the events may also display reduced susceptibility to ALS-inhibiting herbicides. To test this, an imidizolinone herbicide was used to spray tissue-culture plants and changes in fresh weight as a percentage of non-sprayed controls were compared to X914-10 wild-type and *ALSm* transgenic line, R31 four weeks post spraying (Figure 16C). R31 maintained approximately 25% positive growth, while both D52 and RR10 (p35S-TALEN/RT2) showed no significant difference from X914-10, ranging from 16 to 27% negative growth. O lines were also similar to X914-10 but showed significant growth in one line (O69). P lines showed growth variability across lines, ranging from 14% negative to 27% positive growth with one line showing significant positive growth (P8). Q lines had less variability than P lines with all three evaluated lines showing significantly improved growth from X914-10 and two lines having positive growth (Q33 and Q71). Overall, five of the nine evaluated events (55%) displayed reduced susceptibility to imidizolinone herbicide and 18% of the kanamycin resistant events had detectable gene targeting

modifications. These results support the use of GVRs for promoting gene targeting modifications in secondary events and gene targeting efficacy of TALEN and CRISR/Cas reagents delivered by GVRs.

Enhancement of gene targeting modification using regeneration

Previous studies have shown production of somatic mutations and chimerism is common among plant events transformed with SSN reagents (Wang *et al.*, 2014b; Feng *et al.*, 2014; Gao *et al.*, 2010). The difficulty to detect gene targeting modifications in secondary events suggested this may also be the case in potato (Figure 23). A recent study in a related species, tomato (*Solanum lycoperscicon*) used multiple rounds of regeneration and SSN induced expression to enhance levels of SSN-mediated modifications in primary events (Lor *et al.*, 2014). To test this approach, three events derived from each p35S, pLSL, and pLSLm reagent was used for regeneration on high kanamyacin selection (100 mg/L; Kan100) in two replicate experiments (Figure 17A). Although no Kan100 resistant lines were recovered from X914-10, D52, or p35S-TALEN/RT2-derived events, a similar number of lines were recovered from events derived from pLSLm+CRISPR (DD lines), pLSLm+TALEN (EE lines) and pLSL-TALEN/RT2 (FF lines) reagents (Table 6). Regenerated events were screened for gene targeting modifications using PCR as previously described (Figure 24). Unlike secondary events, clear gene targeting modification bands were observed in regeneration events at similar frequencies across reagents. To confirm the bands represent gene targeting modifications, two representative regeneration events, EE39 and FF26 were used for cloning and sequencing the modified locus (Figure 17B). Both representative events were confirmed as containing gene targeting modifications which

included NptII template-specific sequence linked to the template-locus junction. These results suggest the regeneration of secondary events under high selection for gene targeting modifications was capable of enhancing the level of gene targeting modifications in regenerated lines.

To investigate the phenotypic effects of enhanced of gene modifications in regenerated lines, three regenerated lines per reagent were subjected to the herbicide spray assay previously described and compared to X914-10, R31, and progenitor secondary events (Figure 17C). In general, regenerated lines showed improvements in reduced herbicide susceptibility compared to progenitor lines. DD lines showed similar growth to its progenitor and other O lines with significant growth in two events (DD5 and DD11). EE lines showed less growth variability compared to its progenitor and other P lines, and displayed positive growth similar to R31, with significant positive growth in two events (EE35 and EE39). FF lines showed the most dramatic improvements in reduced herbicide susceptibility which ranged from 8% to 85% positive growth, with events FF11 and FF26 displaying significant growth improvements over its progenitor. The sustained variation in both gene targeting modifications and reduced herbicide susceptibility phenotypes in secondary and regenerated events reflect the putative somatic nature of these modifications that can be explained by the multicellular origin of shoot organogenesis (Faize *et al.*, 2010; Poethig, 1989; Zhu *et al.*, 2010). These results and previous studies support the use of regeneration to enhance SSN-mediated modification in chimeric events and modified tissues (Chen, 2011; Marton *et al.*, 2010; Ali *et al.*, 2015; Lor *et al.*, 2014).

Discussion

The utility of plant viruses for delivering genome editing reagents is just being realized in studies using TRV and now with geminiviruses (Ali *et al.*, 2015; Marton *et al.*, 2010; Baltes *et al.*, 2014). Geminiviruses not only provide strong heterologous protein expression but because of their DNA genomes, can serve as potent repair templates for gene targeting. Furthermore, the broad host range of geniminiviruses, such as the *Bean Yellow Dwarf Virus* allows a number of different model and crop species to be modified using the same essential viral elements. Together with the rapid development SSN technology positions geminiviruses to be become powerful tools for genome editing.

Efficient geminivirus replication is dependent on the expression of the viral *trans*-acting element, Rep/RepA. Rep/RepA has been shown to have pleiotropic effects in both monocot and dicot species which promotes cell-cycling progression of nondividing cells to enter the S-phase (Ascencio-Ibáñez *et al.*, 2008). This transition provides the virus with the host factors necessary for DNA replication, but also improves rates of regeneration in plant transformation experiments (Gordon-Kamm *et al.*, 2002). The development of a constitutively expressing Rep/RepA expressing potato line made use of both of these benefits of Rep/RepA expression and simplified reagent delivery. However, the ability of the geminivirus to replicate and facilitate increases in gene targeting frequencies independent of Rep/RepA in both potato genotypes questions the necessity of Rep/RepA in certain plant species and may not be necessary in potato.

Geminiviruses have historically been used for producing high levels of heterologous protein in plant tissues (Mor *et al.*, 2003; Zhang and Mason, 2006; Chen

et al., 2011). It would then seem intuitive that geminiviruses could be harnessed to express sequence-specific nucleases and improve rates of nonhomologous end-joining in regenerating plant tissues (Baltes *et al.*, 2014). However, no improvements in NHEJ were observed in transformed tissues or in stable events (Figure 22). Still, SSN activity either driven by a geminivirus or another source was necessary to achieve efficient gene targeting, and replication of the repair template by a geminivirus alone did not result in gene targeting enhancement (Figure 21).

Previous work has suggested geminivirus replication of the repair template limits gene targeting efficiencies and that a repair template delivered alone on a geminivirus can be used to further improve gene targeting (Baltes *et al.*, 2014). This approach both reduces the size of the geminivirus replicon, putatively allowing for more efficient replication and the need to deliver sequence-specific nucleases on the geminivirus which could be lost to recombination or be toxic to plant cells (Maeder *et al.*, 2008). This approach was made possible in potato by use of a Rep/RepA expressing line but did not seem to improve gene targeting efficiencies beyond what was possible with the sequence-specific nuclease expressing geminivirus in stable events (Table 6).

Production of somatic modifications in lines transformed with sequence-specific nucleases has been previously reported in other plants species and using geminivirus delivery (Feng *et al.*, 2014; Richter *et al.*, 2014). Putatively, more efficient reagents will capable of inducing stable germline mutations in primary events (Brooks *et al.*, 2014). However, the drastic differences between frequencies of nonhomologous end-joining and homologous recombination in plant cells makes development of germline homologous recombination modified lines more difficult in primary events (Wright *et al.*,

2005). This is even further complicated in vegetatively propagated species, like potato that cannot be taken through the germline without changing cultivar characteristics (Douches *et al.*, 1996). Nevertheless, difficulties with homologous recombination may be overcome by utilizing so-called in planta genome editing where homologous recombination is allowed to occur during the life of the plant and modifications can be fixed within cell lines (Fauser *et al.*, 2012). This approach proved effective in potato by putting primary events through a subsequent round of regeneration, applying selection, and regenerating modified cells (Figure 17). The variation in detectable modification and phenotype across regenerated lines suggests detected modifications are somatic and require a germline generation to stabilize modifications.

Materials and Methods

Vector construction

pLSL, p35S and Rep T-DNAs used in replication and the GUPTII assays, and Gateway® (Life Technologies, Carlsbad, CA) entry vectors were obtained from Baltes *et* al. 2014. The pGUPTII reporter and pGUSNptII control were modified from pDW1364 and pDW1273 (Wright *et al.*, 2005), respectively using *Hind*III and *Sac*I sites for restriction enzyme cloning into the Gateway-compatible binary vector, pMDC32 (p35S) (Curtis and Grossniklaus, 2003). The pSSA reporter was cloned using PCR amplification of the GUS coding sequence from pBI101 (Jefferson *et al.*, 1987) and restriction enzyme cloning into pENTR™ (Life Technologies, Carlsbad, CA) using *Bam*HI and *Xho*I sites, and subsequently recombined into p35S. *Avr*II and *Nde*I sites

were used to clone the Zif268 target site from pDW1364 and the potato *ALS1* S642T target site into pSSA and pSSA-S642T reporters, respectively. TALENs were constructed using GoldenGate cloning with NΔ152/C63 N- and C-terminal truncations and TALEN coding sequences separated by a T2A translational skipping sequence (Cermak *et al.*, 2011). CRISPR/Cas vectors were constructed with an *Arabidopsis*-optimized *Steptococcus pyogenes* Cas9 and single-guide RNAs were expressed from an *Arabidopsis thaliana* U6 Pol II promoter (Baltes *et al.*, 2014). TALEN and CRISPR/Cas reagents were recombined into pLSL or p35S Gateway-compatible vectors with or without repair templates. The left and right homology arms for the *ALS1* repair template, RT1 were cloned from X914-10 genomic DNA and fused to a gBlock® (Integrated DNA Technologies, Coralville, IA) containing a 369 bp *ALS1* target sequence with W563L, S642T and silent mutations using overlapping PCR. *Xba*I and *Bam*HI sites in RT1 were used to clone a T2A:NptII fusion downstream of the *ALS1* target sequence to construct RT2. The pLSLm T-DNA was modified from pLSL using *Bam*HI restriction enzyme digestion and removal of the cauliflower mosaic virus 35S promoter.

Plant materials

The tetraploid *S. tuberosum* cultivar “Désirée” (Désirée) and diploid breeding line, MSX914-10 (X914-10) were used in the study. X914-10 is a diploid breeding line from the Michigan State University Potato Breeding and Genetics Program produced from a cross between the doubled-monoploid (DM) *S. tuberosum* Group Phureja line used to construct the potato reference genome (PGSC, 2011) and 84SD22, a heterozygous *S. tuberosum* x *S. chacoense* hybrid breeding line (Felcher *et al.*, 2012). Désirée is a red-

skinned tetraploid potato cultivar with high plant transformation efficiency. Tissue culture plants used for *Agrobacterium*-mediated transformation and herbicide spray assays were propagated from shoot tip and axillary bud explants (four per box) and grown in Magenta® boxes (Phytotech, Shawnee Mission, KS) using Murashige and Skoog (MS) medium (Phytotech; product # M519) with 3% sucrose on light racks set to 16-h-light/8-h-night photoperiod at 22C for three to four weeks. Tissue-culture plants in herbicide spray experiments were individually weighed and either not sprayed (no spray) or sprayed to saturation (approximately 5 mL) using an aqueous solution containing Imazamox PESTANAL ® (7 mg/L) (Sigma-Aldrich, St. Louis, MO; product # 34227) and a plastic spray bottle (Sigma-Aldrich, St. Louis, MO; product # BAF116340000).

Agrobacterium-mediated transformation

Plant transformations for reporter assays and stable transformations were conducted using previously described methods (Paz and Veilleux, 1999). Stable transformation experiments used 180 to 200 leaf explants per experiment and reporter assays used 4 leaf explants per biological replication. Leaf explants were prepared from three to four-week old plants and placed on wounded areas induction media for five to seven days prior to inoculation with *Agrobacterium tumefaciens* GV3101. Two days post inoculation, leaf explants were washed with MS medium containing Cefotaxime (250 mg/L) and Timentin (150 mg/L) antibiotics and placed on regeneration media. Hygromycin (5 mg/L) (Sigma-Aldrich, St. Louis, MO; product # 10687) selection was used for primary transformations and reporter assays for T-DNA selection while kanamycin (Sigma-Aldrich, St. Louis, MO; product # K1377) was used for direct

selection of gene targeting modifications. Regenerated events were rooted on MS media containing 3% sucrose and the same selection used in regeneration media.

PCR detection and cloning

The Expand™ Long Template PCR system (Roche, Basel, Switzerland; product # 11681834001) was used to detect circularized geminivirus replicons using 5'-GTTTCACTTCACACATTATTACTG-3' and 5'-TGTTGAGAACTCTCGACGTCCTGC-3' primer sequences (675 bp). For all other PCR, the Phusion® High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA; product # M0531) was used in combination with the following primers: 5'-GCAGCTGGCACGACAGG-3' and 5'-TGTTGAGAACTCTCGACGTCCTGC-3' were used for pLSL T-DNA (592 bp), 5'-GGCAAAGTGTGGGTCAATAATC-3' and 5'-CCAGGAGGAAGCCATTGTTAT-3' for GUSNptII (1,835 bp), 5'-GGTTGACATTGATGGTGAC-3' and 5'-GCCTAGAACTAGTTATGTAG-3' for *ALS1* (448 bp), 5'-GGTTGACATTGATGGTGAC-3' and 5'-CCATTCGGTTATTGCATC-3' for restriction digestion assays (2,354 bp), 5'-CCTCCTTTCACTTCTCACCTTTA-3' and 5'-GACAGGTCGGTCTTGACAAA-3' for screening secondary events (2,241 bp), and 5'- CTGGCTGCTATTGGGCGAAG-3' and 5'- ACTCCTGCATTTCCACCATTA-3' for screening regeneration events (1,815 bp). All PCR were run using 100 ng genomic DNA purified from young, fully emerged leaves of transformed/regenerated events or wounded areas of leaf explants using the DNeasy Plant Kit (Qiagen, Venlo, Limburg; product # 69104). Individual plants were used as biological replications in herbicide spray experiments. Leaf explants were sampled by excising wounded areas using a sterile scalpel blade and sampled from three to four explants for a single biological replication. PCR bands were quantified using ImageJ

software (http://imagej.nih.gov/ij/) and corrected for size by multiplying the target band intensity by the size ratio of the target band by the *ALS1* band. PCR products were cloned using the Topo® TA cloning kit (Life Technologies, Carlsbad, CA; product # 450071).

GUS quantification and statistics

GUS staining was conducted using X-gluc (5-bromo-4-chloro-3-indoyl-β-D-glucuronic acid) and previously described methods (Butler and Hannapel, 2012). GUS activity quantification was conducted using MUG (4-methylumbelliferyl-β-D-glucuronide) following previously described methods (Jefferson *et al.*, 1987). Protein samples were quantified using Protein Assay Dye Reagent (Bio-Rad, Hercules, CA; product # 5000001) using BSA as a standard (New England Biolabs, Ipswich, MA; product # B9000S). Samples for GUS activity quantification were blanked using time zero samples and read on a Synergy H1 Hybrid plate reader (BioTek, Vinooski, VT). P values were generated using a two-tailed Student's t test.

APPENDICES

Appendix A: Tables

Table 6. Summary of gene targeting screens of secondary and regenerated events. Secondary events transformed with conventional (p35S) and GVR (pLSL and pLSLm) T-DNA (T-DNA) carrying TALEN and CRISPR reagents (Fig. 2A) and the *ALS1* repair template containing NptII (RT2) capable of rooting in kanamycin 50mg/L media were selected as resistant events (Total Kan50 resistant events). A PCR assay was used to identify positive secondary events (# PCR positive events (secondary)) and is shown as a percentage of total Kan50 resistant events (% PCR positive events (secondary)). Regenerated secondary events capable of rooting in kanamycin 100mg/L media were chosen as resistant events (Total Kan100 resistant events). A PCR assay was used to identify positive regeneration events (# PCR positive events (regeneration)) and is shown as a percentage of total Kan100 resistant events (% PCR positive events (secondary)).

T-DNA	Total Kan50 resistant events	Kan50 resistant lines	# PCR positive events (secondary)	% PCR positive events (secondary)	Total Kan100 resistant events	Kan100 resistant lines	# PCR positive events (regeneration)	% PCR positive events (regeneration)
p35S-TALEN/RT2	4	RR	0	0%	0	(none)	0	0%
pLSLm+CRISPR	8	O	1	12.5%	31	DD	10	32.2%
pLSLm+TALEN	13	P	0	0%	27	EE	9	33.3%
pLSL-TALEN/RT2	12	Q	5	41.7%	29	FF	10	34.5%

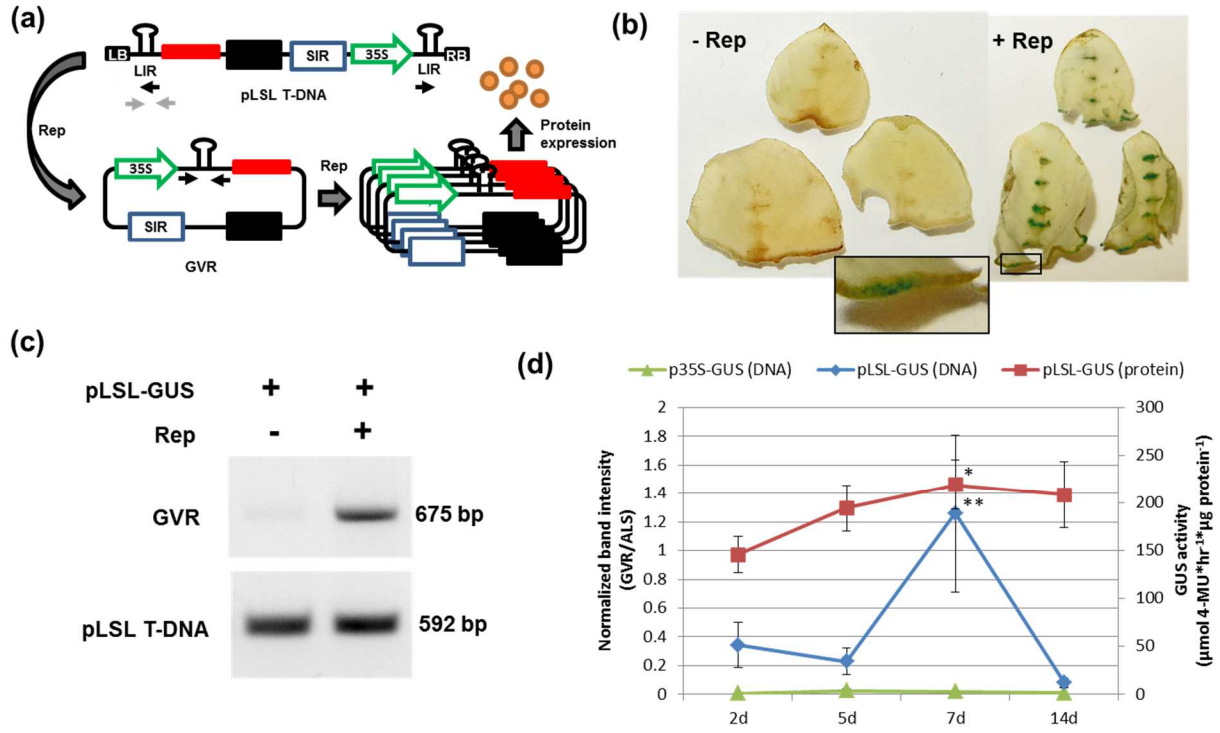

Figure 13. Delivery of the geminivirus replicon (GVR) to potato leaf explants. (a) Schematic of pLSL T-DNA used for *Agrobacterium*-mediated delivery of GVRs to potato leaf tissues. Replicase (Rep) is delivered on a separate p35S T-DNA binary vector (not shown). LB and RB; left and right T-DNA borders, respectively. SIR and LIR; short and long intergenic regions, respectively. 35S; cauliflower mosaic virus promoter. Red and black rectangles; cloning sites for heterologous sequence. Black and light gray arrows; priming sites used for PCR detection of circularized GVRs and pLSL T-DNA, respectively. (b) GUS staining of potato leaf explants transformed with pLSL-GUS. Potato leaf explants were transformed with a pLSL T-DNA that expresses the *uidA* (GUS) coding sequence upon GVR circularization (pLSL-GUS; Baltes *et al.*, 2014). To prepare pLSL-GUS, the GUS coding sequence was cloned into the cloning site closest to the LB (red rectangle; a). Transformations were conducted in the presence (+Rep) or absence (-Rep) of Rep and leaf explants were stained for GUS activity 7 days post inoculation (dpi). Inset is magnification of stained wounded areas from the +Rep treatment (open black rectangle). Images are from Désirée. (c) PCR detection of circularized GVRs in potato leaf explants transformed with pLSL-GUS. Potato leaf explants transformed with pLSL-GUS in the presence (+Rep) or absence (-Rep) of Rep were sampled for PCR detection of circularized GVRs (675 bp) and the pLSL T-DNA (592 bp) using priming sites from panel (a). Images are from Désirée. (d) Time-course

Figure 13 (cont'd). of GVRs in potato leaf explants constitutively expressing Rep. Leaf explants prepared from a mutant potato line constitutively expressing Rep (D52; Figure 19) were transformed with pLSL-GUS and control p35S-GUS T-DNAs and sampled after 2, 5, 7 and 14 dpi. Leaf explant tissues were sampled for PCR detection of circularized GVRs using priming sites from panel (a) (DNA; primary axis) and GUS activity quantification (protein; secondary axis). PCR products were run on 1% agarose gels (d) for band intensity quantification and band intensities were normalized to the potato *ALS1* gene. Error bars represent standard deviations from three biological replications. *P < 0.05 and **P < 0.02; 2d.

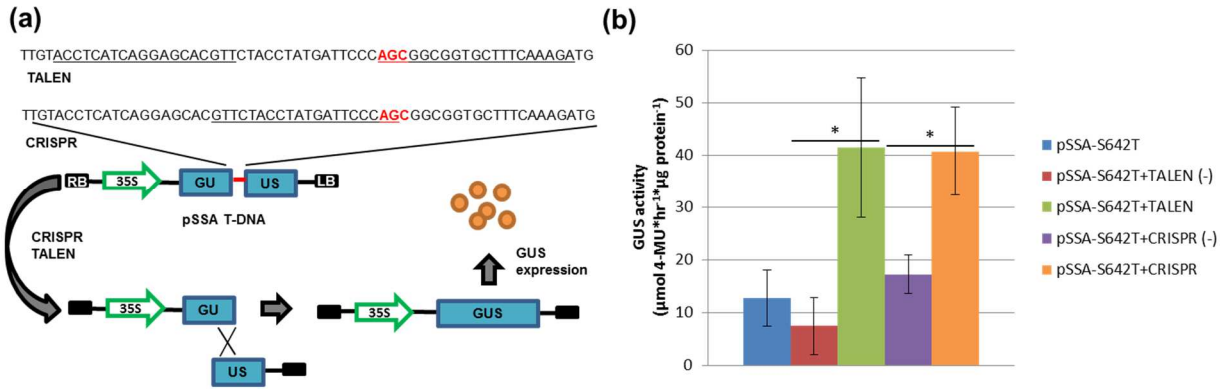

Figure 14. Sequence-specific nuclease (SSN) activity in potato leaf explants. (a) Single-strand annealing assay (SSA) incorporating the S642T target site from the potato *ALS1* gene (red line; sequence) delivered on a T-DNA (pSSA-S642T). The SSA reporter cassette was constructed with the GUS coding sequence (GUS) disrupted by a 60 bp S642T target sequence and a 250 base pair (bp) direct repeat of the GUS coding sequence. Binding sites for the p35S-TALEN (TALEN) and -CRISPR/Cas (CRISPR) reagents targeting S642T are underlined with the S642T codon in red. Upon formation of a double-strand break, the direct repeat is used for DNA repair and reconstitution of the GUS coding sequence and GUS expression. LB and RB; left and right T-DNA borders, respectively. 35S; cauliflower mosaic virus promoter. (b) GUS activity quantification of potato leaf explants transformed with pSSA-S642T and SSN reagents. Leaf explants prepared from wild-type potato plants were co-transformed with the pSSA-S642T reporter, TALEN and CRISPR SSN reagents targeting the S642T target site, or negative control p35S-TALEN (TALEN(-)) and -CRISPR/Cas (CRISPR(-)) reagents targeting a heterogeneous potato *ALS1* target site. Leaf explant tissues were sampled for GUS activity quantification 7 days post inoculation (dpi) and error bars represent standard deviations from three biological replications. *P < 0.05; negative control. Data is from X914-10.

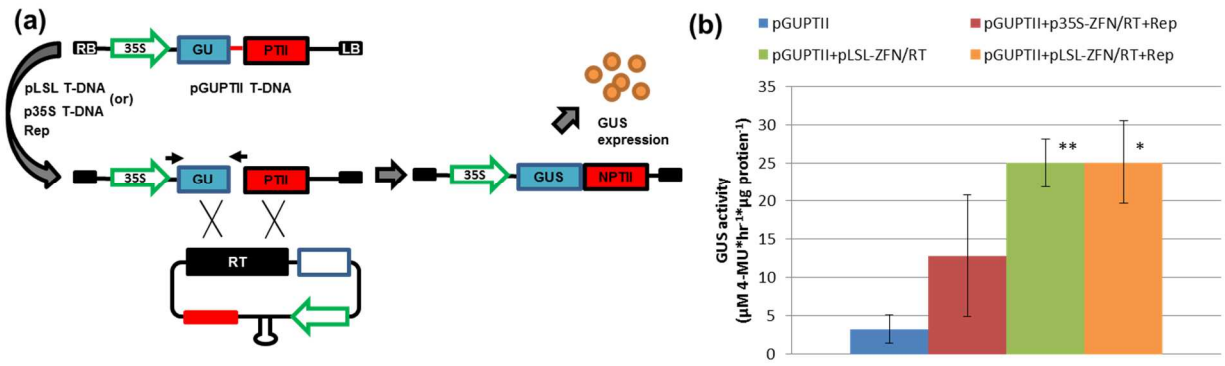

Figure 15. Gene targeting efficiency in potato leaf explants. (a) GUPTII reporter assay incorporating the Zif268 target site delivered on a T-DNA (pGUPTII). The GUPTII reporter cassette was constructed with the GUS:NptII translational fusion coding sequence (GUSNptII) disrupted by a 600 base pair (bp) deletion and a 60 bp Zif268 target site (red line; Figure 20). pLSL and p35S T-DNAs incorporating the Zif268 ZFN (ZFN) and a repair template (RT) incorporating the 600 bp deletion and flanking sequence homologous to the GUPTII reporter were transformed with pGUPTII in the presence (+Rep) or absence of Rep (Baltes *et al.*, 2014). Gene targeting-mediated correction of the pGUPTII reporter results in the reconstitution of the GUSNptII coding sequence and GUS expression. LB and RB; left and right T-DNA borders, respectively. 35S; cauliflower mosaic virus promoter. Black arrows; priming sites used for PCR detection of repaired pGUPTII reporter (GUSNptII). (b) GUS activity quantification of potato leaf explants transformed with pGUPTII and gene targeting reagents. Leaf explants prepared from wild-type potato plants were transformed with the pGUPTII reporter and p35S and pLSL-ZFN T-DNA gene targeting reagents (p35S-ZFN/RT and pLSL-ZFN/RT) in the presence (+Rep) or absence of Rep. Leaf explant tissues were sampled for GUS activity quantification 7 days post inoculation (dpi) and error bars represent standard deviations from three biological replications. *P < 0.05 and **P < 0.02; pGUPTII. Data is from Désirée.

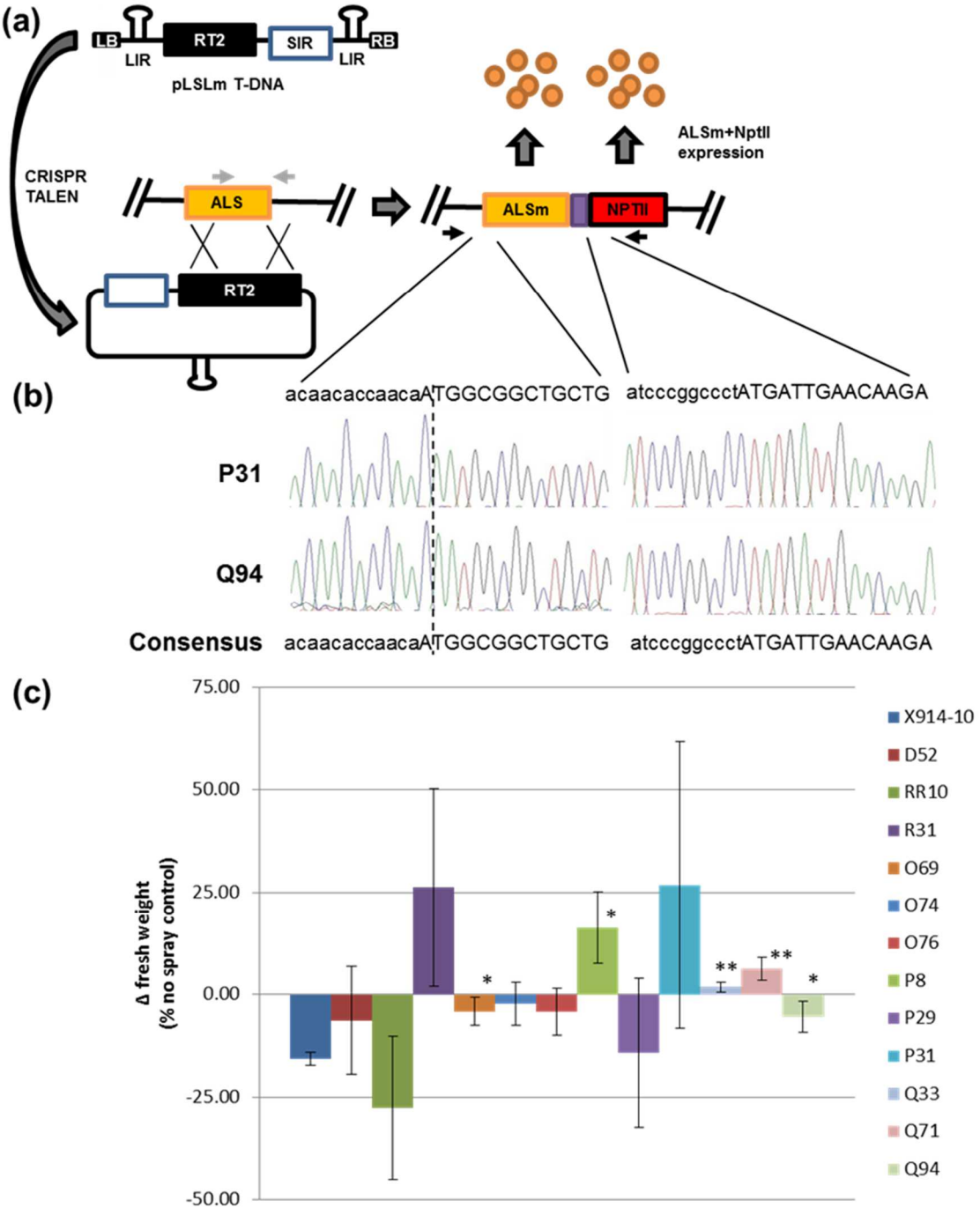

Figure 16. GVR-mediated gene targeting of the potato *ALS1* gene and herbicide susceptibility in secondary events. (a) Gene targeting modification of the potato *ALS1* gene using GVRs with (pLSL; Figure 13A) or without (pLSLm) sequence-specific nucleases (SSNs). Gene targeting modification of the *ALS1* gene with RT2 results in the incorporation of W563L (not shown) and S642T (Figure 14A) point mutations and T2A:NptII within the endogenous locus, and allows modified *ALS1* (*ALSm*) and NptII

Figure 16 (cont'd). (NptII) protein expression. LB and RB; left and right T-DNA borders, respectively. SIR and LIR; short and long intergenic regions, respectively. Black arrows; priming sites used for PCR detection (Figure 23) and cloning (b) of the modified *ALS1* locus. Light gray arrows; priming sites used for PCR detection of the endogenous *ALS1* (Figures 13C, 19, 21, 22) and gene targeting modification digest assays (Figure 22). (b) Cloned gene targeting modifications of the *ALS1* gene in secondary events. Primers specific to the gene targeting modified *ALS1* locus were used for PCR and to clone the locus-template junction (left sequences), both W563L and S642T mutations (not shown), and incorporated T2A:NptII (right sequences). The dotted line represents the locus-template junction where sequences to the left are locus-specific and the sequences to the right are template-specific. Uppercase sequence on the left and right represent the coding sequences for *ALSm* and NptII, respectively. Sanger sequencing traces from clones originating from P31 (top traces) and Q94 (bottom traces) primary events are shown with consensus sequences for both clones (bottom sequence). P31 and Q94 primary events were generated using pLSLm+TALEN and pLSL-TALEN/RT reagents, respectively. (c) Herbicide susceptibility in secondary events. Tissue culture plants were used in an herbicide spray assay for determining herbicide susceptibility in wild-type (X914-10), primary (D, R lines), and secondary events (RR, O, P, Q lines). Primary events were generated by transforming X914-10 with Rep (D52) or the *ALS1* transgene (R31) and applying hygromycin selection. Secondary events were generated by transforming D52 with p35S-TALEN/RT2 (RR10), pLSLm+CRISPR (O69, O74, O76), pLSLm+TALEN (P8, P29, P31), or pLSL-TALEN/RT2 (Q33, Q71, Q94) and applying 50 mg/L kanamycin (Kan50) selection. Change in fresh weight (Δ fresh weight) was calculated as a percentage of the no spray controls for each line. Error bars represent standard deviations from three biological replications. *P < 0.05 and **P < 0.02; X914-10.

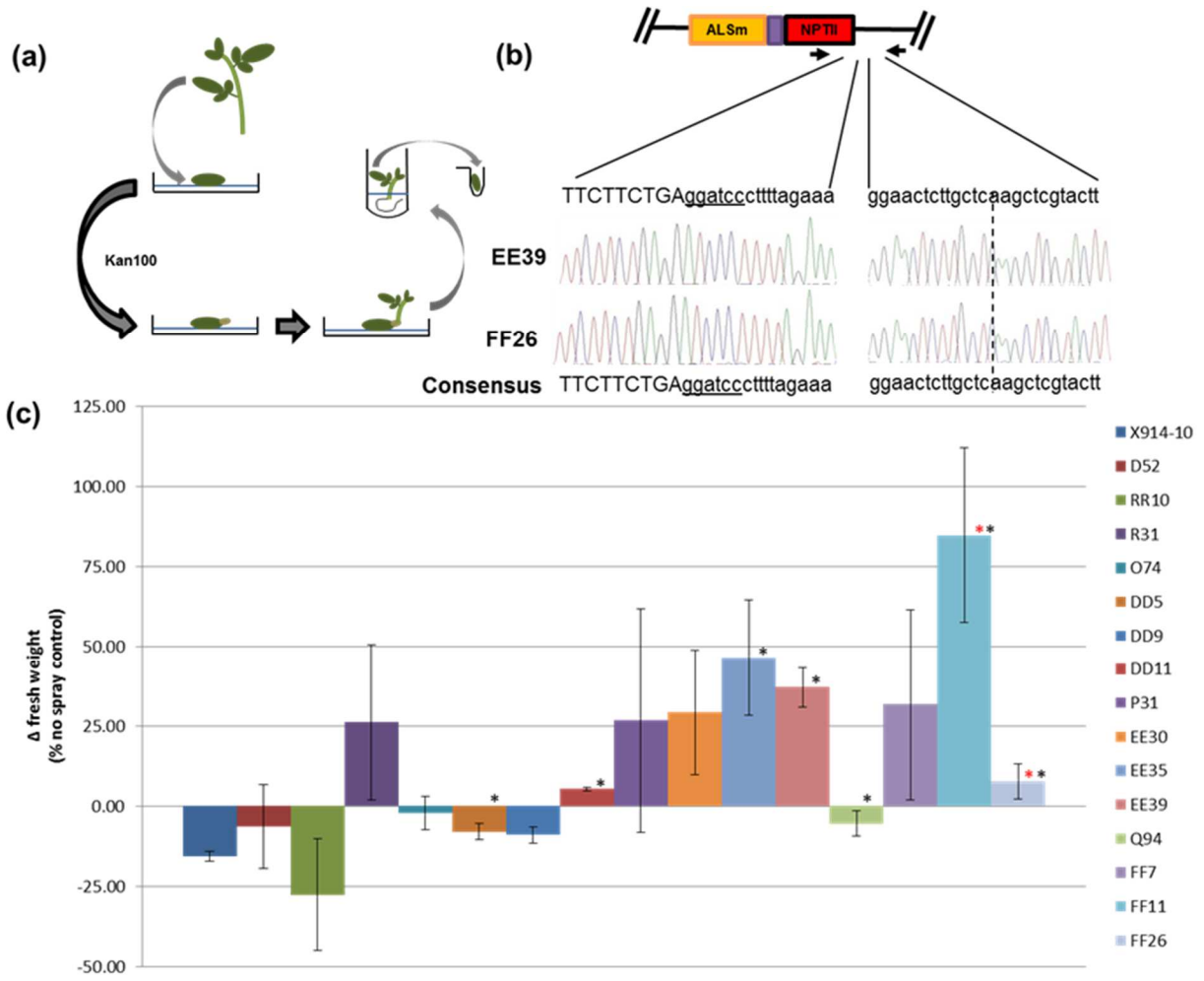

Figure 17. Regeneration of secondary events on high kanamycin selection and herbicide susceptibility in regenerated events. (a) Regeneration of secondary events on high kanamycin selection. Secondary events from D52 were used for regeneration on kanamycin 100mg/L (Kan100) selection media. Primary events capable of rooting in Kan100 where used for cloning the gene targeting modified *ALS1* locus using leaf tissue (b) and in herbicide susceptibility experiments (c). (b) Cloned gene targeting modifications of the *ALS1* gene in regenerated events. Primers specific to the gene targeting modified *ALS1* locus (black arrows) were used for PCR (Figure 24) and to clone the template-locus junction (right sequences) and incorporated *NptII* (left sequences). The dotted line represents the template-locus junction where sequences to the left are template-specific and the sequences to the right are locus-specific. Uppercase and underlined sequence on the left represent the coding sequence for *NptII* and incorporated *Bam*H1 site, respectively. Sanger sequencing traces from clones originating from EE39 (top traces) and FF26 (bottom traces) regenerated events are shown with consensus sequences for both clones (bottom sequence). EE39 and FF26 events were regenerated from P31 and Q94 secondary events, respectively.

Figure 17 (cont'd). (c) Herbicide susceptibility in regenerated events. Four-week old potato plants were used in an herbicide spray assay for determining herbicide susceptibility in wild-type (X914-10), secondary (D52, RR10, R31, O74, P31, Q94), and regenerated events (DD, EE, FF lines). Primary and secondary events were generated using methods described in Figure 16. Regenerated events were produced using methods from (b) and originate from O74 (DD5, DD9, DD11), P31 (EE30, EE35, EE39), and Q94 (FF7, FF11, FF26) secondary events. Herbicide spray experiments were carried out as described in Figure 16, and X914-10 and secondary event data comes from Figure 16. Error bars represent standard deviations from three biological replications. *P < 0.05; X914-10. *P < 0.05; Q94.

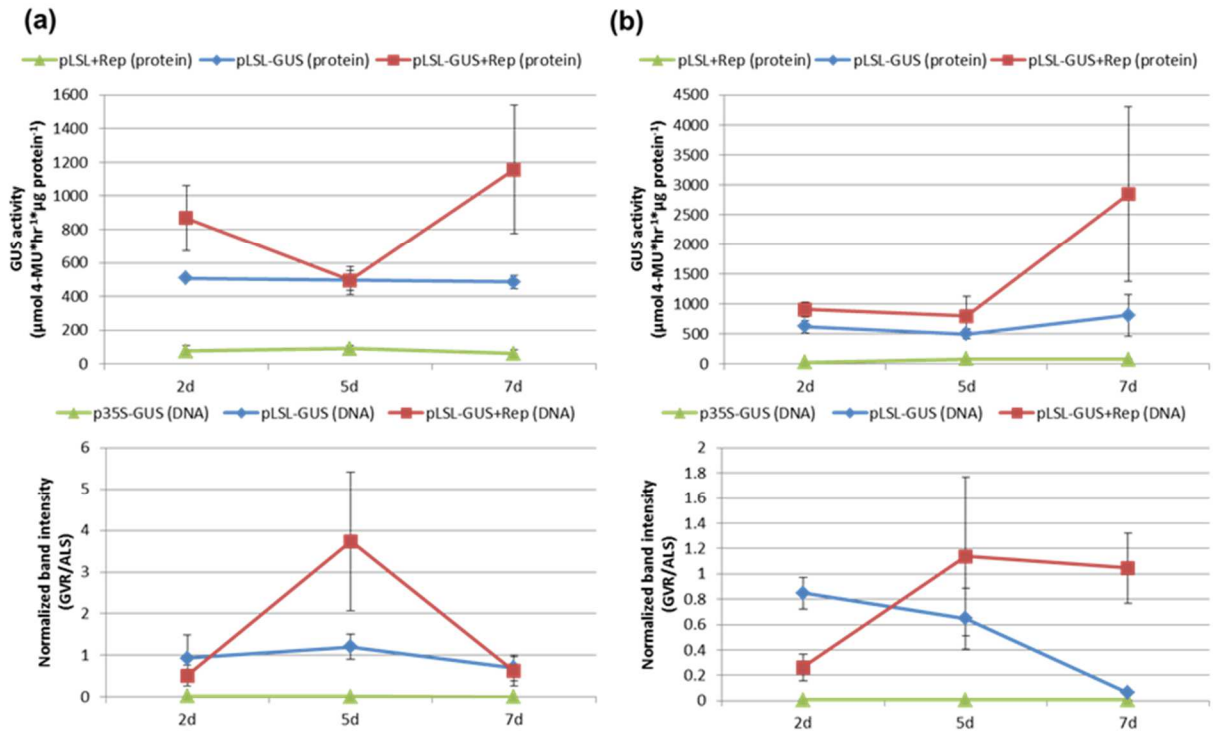

Figure 18. Time-course of GVRs (DNA) and GVR-mediated protein expression (protein) in wild-type potato leaf explants. Leaf explants prepared from wild-type X914-10 (a) and Désirée (b) were transformed with pLSL-GUS and control pLSL T-DNAs in the presence (+Rep) or absence of Rep and sampled after 2, 5 and 7 dpi. Leaf explant tissues were sampled for GUS activity quantification (top graphs) and PCR detection of circularized GVRs (bottom graphs) (Figure 13). Error bars represent standard deviations from three biological replications.

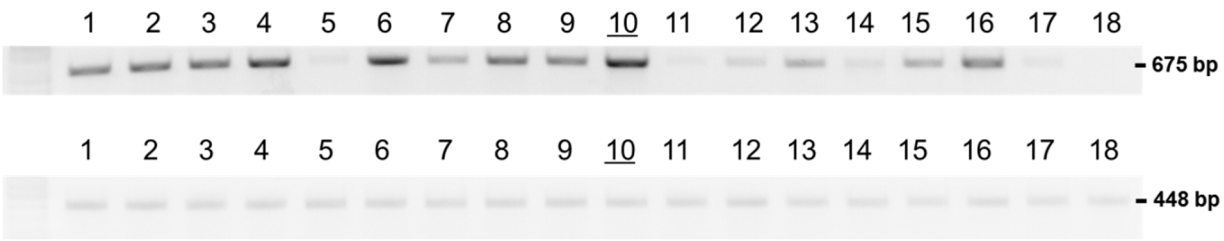

Figure 19. GVR screen of primary potato events transformed with Rep. Leaf explants from hygromycin-resistant primary events transformed with Rep (#1-17) and X914-10 wild-type (#18) were transformed with pLSL-GUS and leaf explant tissues were collected for PCR detection of circularized GVRs (top gel; Figure 13) and *ALS1* (bottom gel). Primary event, D52 (#10; underlined) was chosen for use in the study.

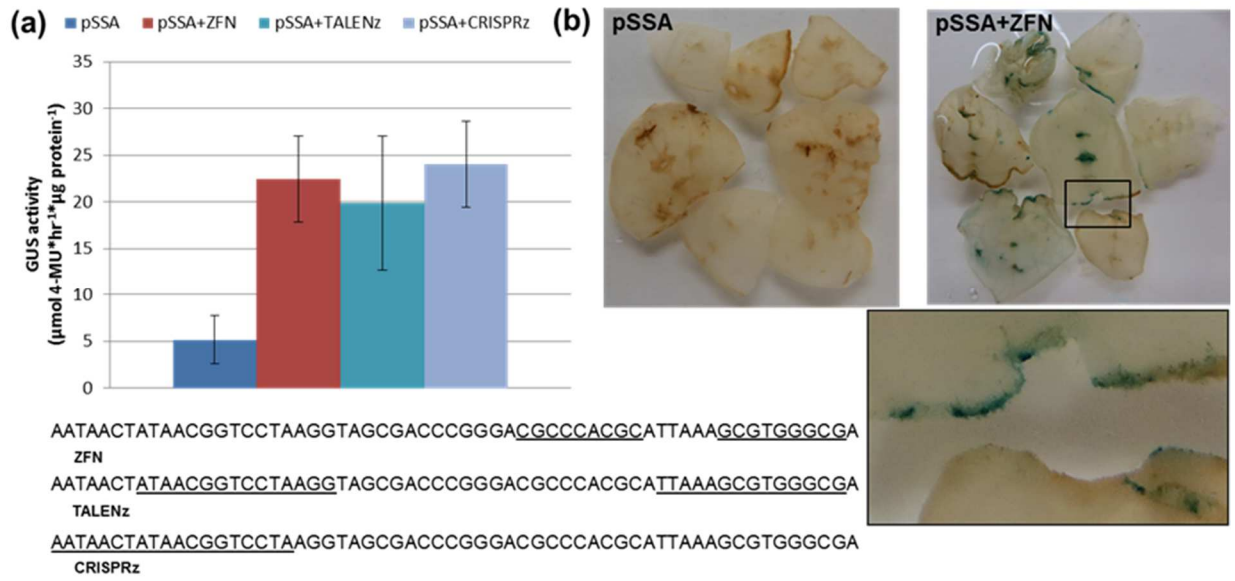

Figure 20. Single-strand annealing assay (SSA) comparing ZFN, TALEN and CRISPR/Cas SSN activity in potato leaf explants. (a) A SSA reporter incorporating the Zif268 target site from the pGUPTII T-DNA (Wright *et al.*, 2005) was delivered on a T-DNA (pSSA) and transformed with SSN reagents. Binding sites for the p35S-Zif268 (ZFN), -TALEN (TALENz), and -CRISPR/Cas (CRISPRz) reagents targeting the Zif268 target site are underlined (sequence). The SSA reporter cassette was constructed as described in Figure 14 with a 60 base pair (bp) Zif268 target sequence separating the direct repeat. Leaf explant tissues were sampled for GUS activity quantification 7 days post inoculation (dpi) and error bars represent standard deviations from three biological replications. (b) GUS staining of potato leaf explants transformed with pSSA. Potato leaf explants were transformed with pSSA in the presence (+ZFN) or absence of ZFN and were stained for GUS activity 7 days post inoculation (dpi). Inset is magnification of stained wounded areas tissue from +ZFN treatment (open black rectangle).

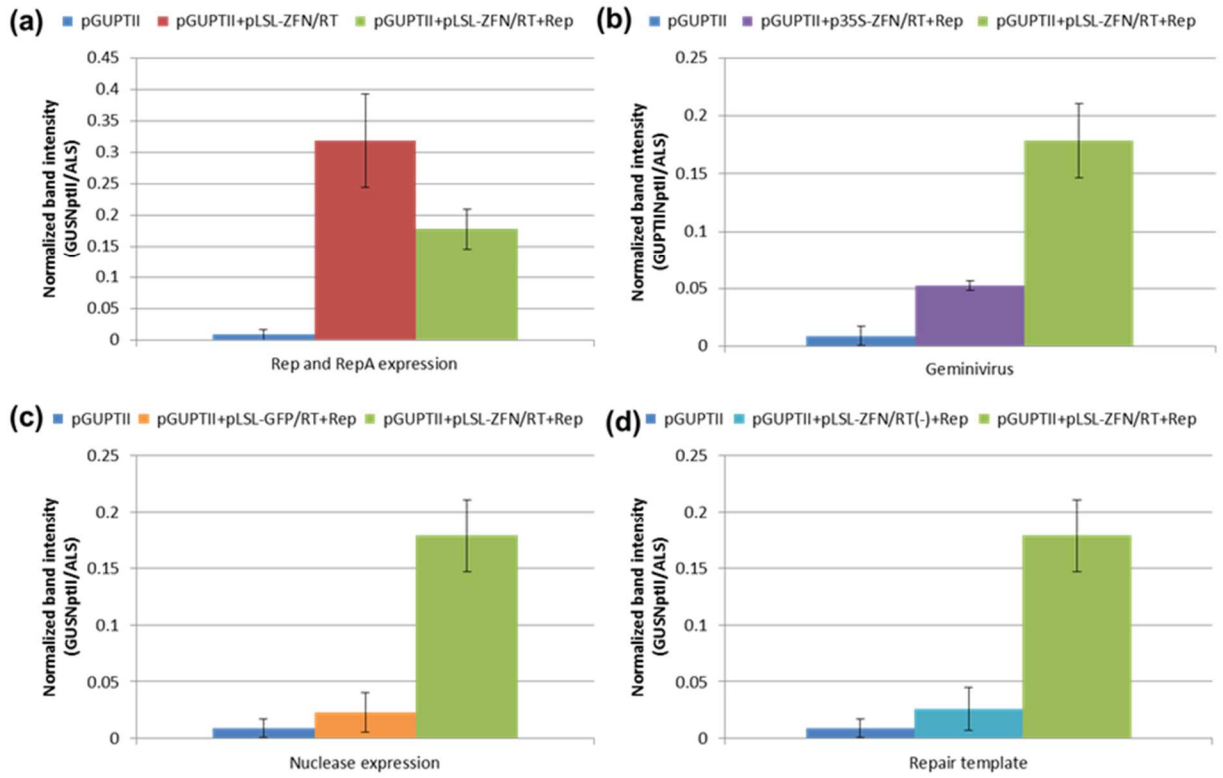

Figure 21. PCR detection of pGUPTII reporter repair in potato leaf explants transformed with gene targeting reagents. Leaf explants prepared from wild-type potato plants were transformed with the pGUPTII reporter and pGUPTII gene targeting reagents (pLSL-ZFN/RT (a) and p35S-ZFN/RT (b)) in the presence (+Rep) or absence of Rep (Figure 15). Control experiments were conducted using pLSL-GFP/RT (c) and pLSL-ZFN/RT(-) (d) T-DNAs where the Zif268 was replaced by green fluorescent protein (GFP) and the pGUPTII repair template (RT) was replaced with a heterologous tobacco *ALS* RT (RT-), respectively. Leaf explant tissues were sampled for PCR detection 7 days post inoculation (dpi) using priming sites shown in Figure 15 (black arrows) and band quantification was carried out as described in Figure 13. Error bars represent standard deviations from three biological replications. Data is from Désirée.

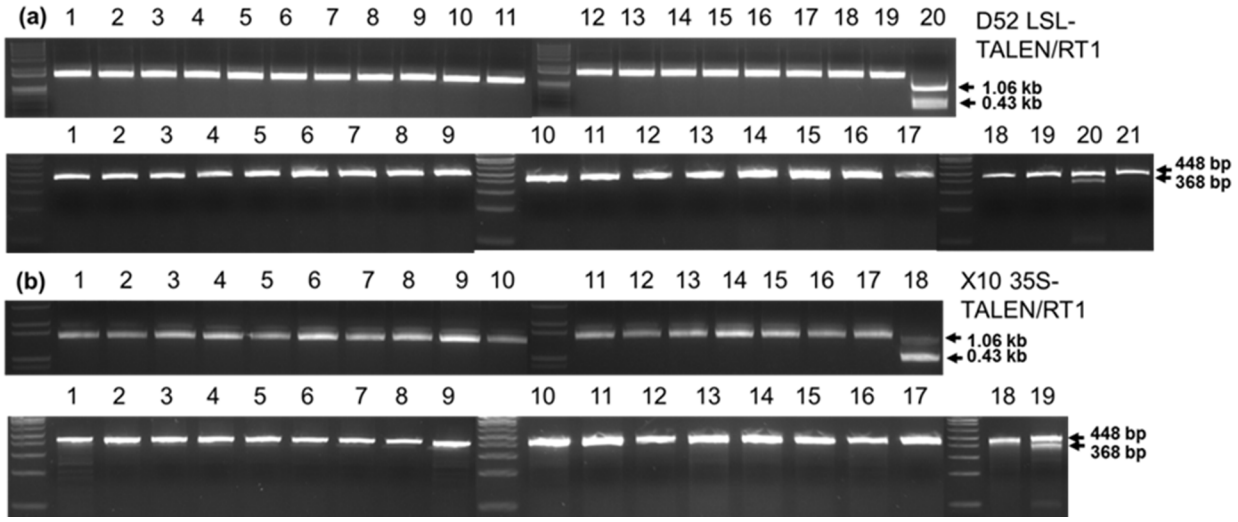

Figure 22. Screen for gene targeting modifications and nonhomologous end-joining (NHEJ) mutations in *ALS1* of primary and secondary events without using direct selection for gene targeting. The *ALS1* repair template (RT1) used for transformation was constructed without T2A:NptII (not shown), and transformations were conducted using hygromycin for T-DNA selection. Secondary events (lanes 1-19) were transformed with D52 using pLSL-TALEN/RT1 T-DNA (a), and primary events (lanes 1-17) were transformed with X914-10 using p35S-TALEN/R2 T-DNA (b) and screened for gene targeting modifications (upper gels) or NHEJ mutations (lower gels) in the *ALS1* gene. Gene targeting screening was conducted using a digest assay using priming sites from Figure 16A (light gray arrows) and BamHI restriction enzyme digestion of the PCR amplicon purified using the QIAquick PCR purification kit (Qiagen, Lenlo, Limburg; product # 28104) (Figure 17B). Wild-type amplicons will remain undigested (1.49 kb) while gene targeting modified amplicons will form 1.06 kb and 0.43 kb digest bands (lanes 20 (a) and 18 (b); black arrows). NHEJ screening was conducted using a T7EI assay using priming sites from Figure 16A and T7 endonuclease I (NEB, Ipwich, MA; product # M0302) digestion (Huang *et al.*, 2012). Wild-type amplicons (448 bp) will remain undigested (lanes 21 (a) and 18 (b)) while amplicons with NHEJ mutations will form a 368 bp band (lanes 20 (a) and 19 (b)).

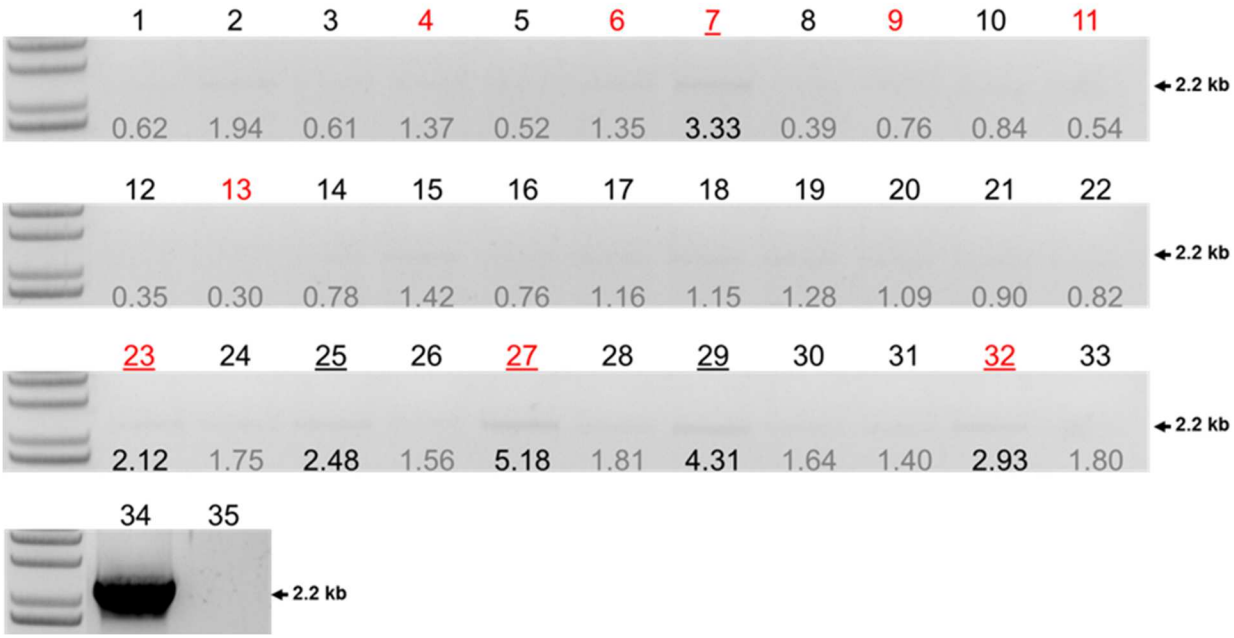

Figure 23. Screen for gene targeting modifications in *ALS1* of secondary events using direct selection for gene targeting. Secondary transformations of D52 were conducted using pLSLm+CRISPR (lanes 1-8), pLSLm+TALEN (lanes 9-21) and pLSL-TALEN/RT (lanes 22-33) with direct selection for gene targeting using kanamycin 50 mg/L. PCR screening was conducted using priming sites specific to the modified *ALS1* locus (Figure 16A; black arrows) with an expected amplicon size of 2.2 kb (lane 34; positive control). PCR products were run on 1.0% agarose gels and band intensities were quantified and normalized to *ALS1* as described in Figure 13 for each secondary event (values). Secondary events were considered positive if they had band intensities equal or more than two-fold *ALS1* (black values, underlined) and nine were chosen for further analysis (red; Figures 16 and 17). D52 was used as a negative control (lane 35).

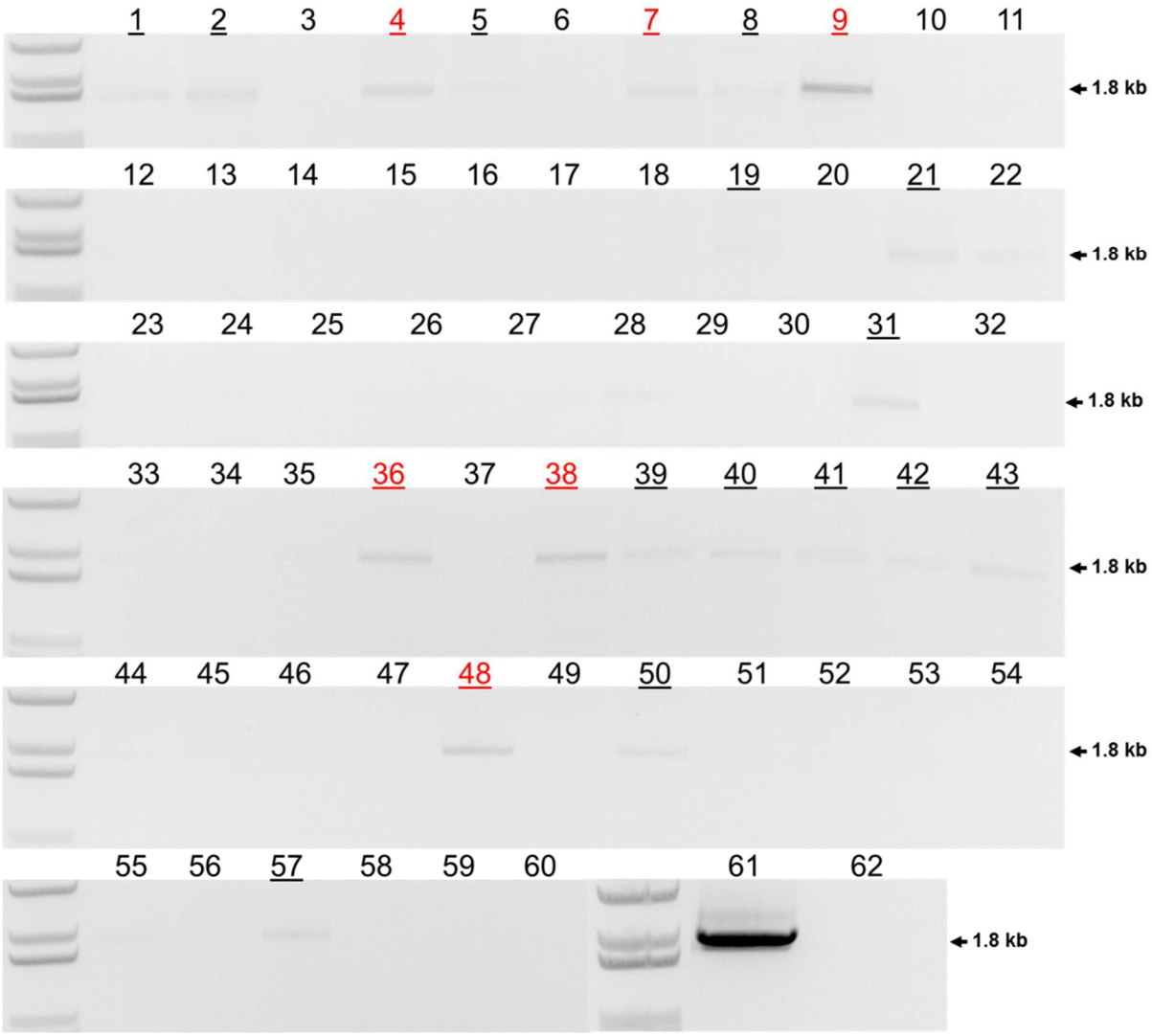

Figure 24. Screen for gene targeting modifications in *ALS1* of regenerated events using high direct selection for gene targeting. Regeneration was conducted using O74 (lanes 1-31), P31 (not shown), and Q94 (lanes 32-60) secondary events with high direct selection for gene targeting using kanamycin 100 mg/L (Figure 17). PCR screening was conducted using priming sites specific to the modified *ALS1* locus (Figure 17B; black arrows) with an expected amplicon size of 1.8 kb (lane 61; positive control). Regenerated events were considered positive if they had visible bands (underlined) and nine were chosen for further analysis (red; Figures 17). D52 was used as a negative control (lane 62).

REFERENCES

REFERENCES

Ali, Z., Abul-faraj, A., Li, L., et al. (2015) Efficient virus-mediated genome editing in plants using the CRISPR/Cas9 system. *Mol. Plant*. Available at: http://dx.doi.org/10.1016/j.molp.2015.02.011.

Ascencio-Ibáñez, J.T., Sozzani, R., Lee, T.-J., Chu, T.-M., Wolfinger, R.D., Cella, R. and Hanley-Bowdoin, L. (2008) Global analysis of *Arabidopsis* gene expression uncovers a complex array of changes impacting pathogen response and cell cycle during geminivirus infection. *Plant Physiol.*, **148**, 436–454.

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. Available at: http://www.plantcell.org/cgi/doi/10.1105/tpc.113.119792.

Bibikova, M., Beumer, K., Trautman, J.K. and Carroll, D. (2003) Enhancing gene targeting with designed zinc finger nucleases. *Science*, **300**, 764. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16169774.

Brigneti, G., Martín-Hernández, A.M., Jin, H., Chen, J., Baulcombe, D.C., Baker, B. and Jones, J.D.G. (2004) Virus-induced gene silencing in Solanum species. *Plant J.*, **39**, 264–272. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15225290.

Brooks, C., Nekrasov, V., Lippman, Z.B. and Eck, J. Van (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.

Burch-Smith, T.M., Schiff, M., Liu, Y. and Dinesh-Kumar, S.P. (2006) Efficient virus-induced gene silencing in *Arabidopsis*. *Plant Physiol.*, **142**, 21–27. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1557620&tool=pmcentrez&rendertype=abstract.

Butler, N.M. and Hannapel, D.J. (2012) Promoter activity of polypyrimidine tract-binding protein genes of potato responds to environmental cues. *Planta*, **236**, 1747–1755. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22868575.

Cermak, T., Doyle, E.L., Christian, M., et al. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.*, **39**, e82. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3130291&tool=pmcentrez&rendertype=abstract.

Chen, G.Q. (2011) Effective reduction of chimeric tissue in transgenics for the stable genetic transformation of *Lesquerella fendleri*. *HortScience*, **46**, 86–90.

Chen, Q., He, J., Phoolcharoen, W. and Mason, H.S. (2011) Geminiviral vectors based on *bean yellow dwarf virus* for production of vaccine antigens and monoclonal antibodies in plants. *Hum. Vaccin.*, **7**, 331–338. Available at: http://www.landesbioscience.com/journals/vaccines/article/14262/.

Clasen, B.M., Stoddard, T.J., Luo, S., et al. (2015) Improving cold storage and processing traits in potato through targeted gene knockout. *Plant Biotechnol. J.*, n/a–n/a. Available at: http://doi.wiley.com/10.1111/pbi.12370.

Curtin, S.J., Voytas, D.F. and Stupar, R.M. (2012) Genome engineering of crops with designer nucleases. *Plant Genome J.*, **5**, 42–50. Available at: https://www.crops.org/publications/tpg/abstracts/5/2/42.

Curtis, M. and Grossniklaus, U. (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol.*, **133**, 462–469. Available at: http://www.plantphysiol.org/content/133/2/462.short.

Desai, N.A. and Shankar, V. (2003) Single-strand-specific nucleases. *FEMS Microbiol. Rev.*, **26**, 457–491. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12586391.

Dooner, H.K. and Weil, C.F. (2007) Give-and-take: interactions between DNA transposons and their host plant genomes. *Curr. Opin. Genet. Dev.*, **17**, 486–492. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0959437X0700161X.

Douches, D.S., Maas, D., Jastrzebski, K. and Chase, R.W. (1996) Assessment of potato breeding progress in the USA over the last century. *Crop Sci.*, **36**, 1544. Available at: https://www.crops.org/publications/cs/abstracts/36/6/CS0360061544.

Endo, M. and Toki, S. (2013) Creation of herbicide-tolerant crops by gene targeting. *J. Pestic. Sci.*, **38**, 49–59. Available at: http://jlc.jst.go.jp/DN/JST.JSTAGE/jpestics/D12-073?lang=en&from=CrossRef&type=abstract.

Faize, M., Faize, L. and Burgos, L. (2010) Using quantitative real-time PCR to detect chimeras in transgenic tobacco and apricot and to monitor their dissociation. *BMC Biotechnol.*, **10**, 53. Available at: http://www.biomedcentral.com/1472-6750/10/53.

Fauser, F., Roth, N., Pacher, M., Ilg, G., Sánchez-Fernández, R., Biesgen, C. and Puchta, H. (2012) In planta gene targeting. *PNAS.*, **109**, 7535–40. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3358861&tool=pmcentrez&rendertype=abstract.

Felcher, K.J., Coombs, J.J., Massa, A.N., Hansey, C.N., Hamilton, J.P., Veilleux, R.E., Buell, C.R. and Douches, D.S. (2012) Integration of two diploid potato linkage maps with the potato genome sequence. *PLoS One*, **7**, e36347. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3338666&tool=pmcentrez&rendertype=abstract.

Feng, Z., Mao, Y., Xu, N., et al. (2014) Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in *Arabidopsis*. *PNAS*, **111**, 4632–4637. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1400822111.

Gao, H., Smith, J., Yang, M., et al. (2010) Heritable targeted mutagenesis in maize using a designed endonuclease. *Plant J.*, **61**, 176–187. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19811621.

Giacca, M. and Zacchigna, S. (2012) Virus-mediated gene delivery for human gene therapy. *J. Control. Release*, **161**, 377–388. Available at: http://dx.doi.org/10.1016/j.jconrel.2012.04.008.

Gordon-Kamm, W., Dilkes, B.P., Lowe, K., et al. (2002) Stimulation of the cell cycle and maize transformation by disruption of the plant retinoblastoma pathway. *Science.*, **99**, 11975–11980.

Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W. (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.*, **6**, 3901–3907.

Kotterman, M.A. and Schaffer, D. V (2015) Engineering adeno-associated viruses for clinical gene therapy. *Nat. Rev. Genet.*, **15**, 445–451.

Lee, K.Y., Lund, P., Lowe, K. and Dunsmuir, P. (1990) Homologous recombination in plant cells after *Agrobacterium*-mediated transformation. *Plant Cell*, **2**, 415–425.

Li, J.F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M. and Sheen, J. (2013) Multiplex and homologous recombination–mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.*, **31**, 688–691. Available at: http://www.nature.com/doifinder/10.1038/nbt.2654.

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. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22565958.

Liu, L., Liu, L., Davies, J.W., Davies, J.W., Stanley, J. and Stanley, J. (1998) Mutational analysis of *bean yellow dwarf virus*, a geminivirus of the genus *Mastrevirus* that is adapted to dicotyledonous plants. *J. Gen. Virol.*, **79**, 2265–2274.

Liu, L., Tonder, T. Van, Pietersen, G., Davies, J.W. and Stanley, J. (1997) Molecular characterization of a subgroup I geminivirus from a legume in South Africa. *J. Gen. Virol.*, **78**, 2113–2117.

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 Physiol.,* **166**, 1288-1291.

Lusser, M., Parisi, C., Plan, D. and Rodríguez-Cerezo, E. (2012) Deployment of new biotechnologies in plant breeding. *Nat. Biotechnol.*, **30**, 231–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22398616.

Maeder, M.L., Thibodeau-beganny, S., Osiak, A., et al. (2008) Rapid “open-source” engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol. Ther.*, **31**, 294–301.

Marton, I., Zuker, A., Shklarman, E., Zeevi, V., Tovkach, A., Roffe, S., Ovadis, M., Tzfira, T. and Vainstein, A. (2010) Nontransgenic genome modification in plant cells. *Plant Physiol.*, **154**, 1079–1087. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2971589&tool=pmcentrez&rendertype=abstract.

Mor, T.S., Moon, Y.-S., Palmer, K.E. and Mason, H.S. (2003) Geminivirus vectors for high-level expression of foreign proteins in plant cells. *Biotechnol. Bioeng.*, **81**, 430–437. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12491528.

Palmgren, M.G., Edenbrandt, A.K., Vedel, S.E., et al. (2014) Are we ready for back-to-nature crop breeding? *Trends Plant Sci.*, 1–10. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1360138514002908.

Paz, M. and Veilleux, R. (1999) Influence of culture medium and *in vitro* conditions on shoot regeneration in *Solanum phureja* monoploids and fertility of regenerated doubled monoploids. *Plant Breed.*, **118**, 53–57. Available at: http://onlinelibrary.wiley.com/doi/10.1046/j.1439-0523.1999.118001053.x/abstract.

Poethig, S. (1989) Genetic mosaics and cell lineage analysis in plants. *Trends Genet.*, **5**, 273–277. Available at: http://linkinghub.elsevier.com/retrieve/pii/0168952589901017.

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. *PNAS*, **93**, 5055–5060. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=39405&tool=pmcentrez&rendertype=abstract.

Richter, K.S., Kleinow, T. and Jeske, H. (2014) Somatic homologous recombination in plants is promoted by a geminivirus in a tissue-selective manner. *Virology*, **452-453**, 287–96. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24606706.

Sathasivan, K., Haughn, G.W. and Murai, N. (1991) Molecular basis of imidazolinone herbicide resistance in *Arabidopsis thaliana* var Columbia. *Plant Physiol.*, **97**, 1044–1050.

Sawai, S., Ohyama, K., Yasumoto, S., 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. Available at: http://www.plantcell.org/cgi/doi/10.1105/tpc.114.130096.

Sha, A., Zhao, J., Yin, K., Tang, Y., Wang, Y., Wei, X., Hong, Y. and Liu, Y. (2013) Virus-based microRNA silencing in plants. *Plant Physiol.* Available at: http://www.ncbi.nlm.nih.gov/pubmed/24296072.

Shukla, V.K., Doyon, Y., Miller, J.C., et al. (2009) Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature*, **459**, 437–441. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19404259.

Stenger, D.C., Revington, G.N., Stevenson, M.C. and Bisaro, D.M. (1991) Replicational release of geminivirus genomes from tandemly repeated copies: evidence for rolling-circle replication of a plant viral DNA. *PNAS*, **88**, 8029–8033.

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

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C. and Qiu, J.-L. (2014) Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol.*, **32**, 947–951. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25038773.

Wright, D.A, Townsend, J.A, Winfrey, R.J., Irwin, P.A, Rajagopal, J., Lonosky, P.M., Hall, B.D., Jondle, M.D. and Voytas, D.F. (2005) High-frequency homologous recombination in plants mediated by zinc-finger nucleases. *Plant J.*, **44**, 693–705. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16262717.

Zhang, X. and Mason, H. (2006) *Bean Yellow Dwarf Virus* replicons for high-level transgene expression in transgenic plants and cell cultures. *Biotechnol. Bioeng.*, **93**, 271–279. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16187337.

Zhang, Y., Zhang, F., Li, X., Baller, J.A., Qi, Y., Starker, C.G., Bogdanove, A.J. and Voytas, D.F. (2013) Transcription activator-like effector nucleases enable efficient plant genome engineering. *Plant Physiol.*, **161**, 20–27. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3532252&tool=pmcentrez&rendertype=abstract.

Zhu, X.-Y., Chai, S.-J., Chen, L.-P., Zhang, M.-F. and Yu, J.-Q. (2010) Induction and origin of adventitious roots from chimeras of *Brassica juncea* and *Brassica oleracea*. *Plant Cell, Tissue Organ Cult.*, **101**, 287–294. Available at: http://link.springer.com/10.1007/s11240-010-9686-8.

CHAPTER 4: CONCLUSION

A new era of genetic engineering

Plant genetic engineering has entered a new era. The recent developments in sequence-specific nuclease (SSN) technology are enabling efficient modification of plant DNA and are forcing regulatory agencies to rethink current regulations on genetically modified (GM) crops. Original research leading to what we now know as genome editing was fueled by a human desire to better understand the function of DNA in a living cell and how such a complex molecule is capable of maintaining its integrity over millions and millions of generations. This passion for understanding the homeostasis of nature has ironically created tools in which synthetic organisms are just within our reach (Räsänen *et al.*, 2015). We have begun to explore these new tools in model plant species such as Arabidopsis, rice and tobacco but we have much to learn about its potential to reshape agriculture and accelerate crop improvement.

Genome editing in potato and other crop species is just beginning. Founding SSN platforms, such as homing endonucleases and ZFNs are costly to develop and have limited efficacy in plant cells, limiting their adoption. However, the development of TALENs provided more efficient reagents that could be designed to most DNA targets. The development of TALENs led to landmark studies for targeted mutagenesis in rice, wheat and also—potato.

Early reports of genome editing in potato

A long sought after trait for potato improvement has been reduced sugar accumulation in potatoes after long-term storage, so called cold-induced sweetening (Dale and Bradshaw, 2003). Cold storage is used to reduce sprouting and extend storage life of potatoes before they are used for processing. High-temperature processing, such as frying will result in browning and accumulation of acrylamide in French fries and chips if reducing sugars have accumulate in the stored potatoes. Previous work has shown that RNAi-mediated "knock-down" of the vacuolar invertase gene (*Vlnv*) can significantly reduce cold-induced sweetening (Wu *et al.*, 2011; Bhaskar *et al.*, 2010; Zhu *et al.*, 2014). However, the use of transgenic or "all-native" DNA requires these varieties to be de-regulated before commercial use.

Researchers at the company, Calyxt (formally Cellectis Plant Sciences) used protoplast transformation to deliver TALENs targeting the *Vlnv* gene in the tetraploid cultivar, Ranger Russet (Clasen *et al.*, 2015). Protoplast transformation was used to both improve the efficiency of TALEN-mediated targeted mutagenesis and to provide a transient method to deliver the TALENs without relying on integration. This later point is particularly important in clonally propagated species, such as tetraploid potato that cannot be selfed or used in a genetic cross to remove the integrated TALEN reagents without altering the characteristics of the cultivar. However, the use of protoplasts extended the time required to regenerate whole plants to approximately 90 days compared to the 40 days needed for *Agrobacterium*-mediated transformation. Furthermore, only 18 events contained targeted mutations out of the over 600 lines regenerated (3%) and only two events contained a complete "knock-out" and were

TALEN-free (0.33%). The limited targeted mutagenesis efficiency reported in this study questions the efficacy of TALENs for targeted mutagenesis in potato and suggests other SSN platforms may be more productive.

SSN-mediated targeted mutagenesis in potato

The targeted mutagenesis work reported in this thesis was performed using ZFNs, TALENs and CRISPR/Cas in both diploid and tetraploid potato. Preliminary assessments of ZFN, TALEN and CRISPR/Cas activity in diploid potato cells demonstrated that each platform was capable of creating targeted mutations within a reporter. With this information, TALENs and CRISPR/Cas reagents were designed to target the endogenous *ACETOLACTATE SYNTHASE1* (*ALS1*) gene and were tested in a reporter assay, in callus tissues and in regenerated events. Two sets of TALENs and two sgRNA were designed to test each SSN platform and were expressed from conventional 35S or modified geminivirus T-DNA. Similar to frequencies seen in previous studies using TALENs in potato, none of the 278 and 483 total regenerated events contained detectable targeted mutations in the diploid and tetraploid backgrounds, respectively (Clasen *et al.*, 2015; Sawai *et al.*, 2014). However, the TALEN set used later for gene targeting experiments was capable of inducing targeted mutations in a reporter assay. This is most likely due to the transient nature of the reporter in this assay, and the reduced opportunity for the reporter target site to be methylated which has been shown to block TALEN binding (Joung and Sander, 2012).

CRISPR/Cas proved to be more effective than TALENs for targeted mutagenesis in potato. Both sgRNA tested were capable of forming detectable targeted mutations in callus tissues of diploid and tetraploid potato. These results were further validated using a reporter assay in which one of the two sgRNA showed significant activity. Confirmation of CRISPR/Cas efficacy for targeted mutagenesis in potato came from experiments with regenerated primary events. Targeted mutagenesis efficiency ranged from 3% to 55% and 5% to 60% in diploid and tetraploid backgrounds, respectively in primary events. These mutations ranged from single bp insertions (typically A) to 38 bp deletions, similar to reports in other plant species (Feng *et al.*, 2014; Brooks *et al.*, 2014). Interestingly, some primary events held single mutations that were inherited in later generations. These results differ from those reported in Arabidopsis where chimerism dominated in primary events. Nevertheless, new mutation alleles were discovered in Cas9-free progeny due to persistent Cas9 activity during gametogenesis. Overall, the high efficiency of targeted mutagenesis in primary events and high mutation transmission sets CRISPR/Cas apart from existing SSN platforms for targeted mutagenesis in potato.

Geminivirus-mediated gene targeting in potato

The demonstration of successful targeted mutagenesis in potato led to development of new reagents for gene targeting in diploid potato. Diploid potato was the focus of the gene targeting experiments to simplify detection of modifications and for potential applications in diploid functional genomics. A novel approach to gene targeting

was tested that employed the use of a geminivirus for delivering the repair template required for gene targeting (Baltes *et al.*, 2014). Previous studies in tobacco demonstrated the ability of this geminivirus to enhance rates of gene targeting by targeting an integrated reporter. However, it was unclear if this geminivirus developed in tobacco would be capable of replicating and improving gene targeting frequencies in other plant species. Preliminary assessments of the geminivirus in replication assays demonstrated that indeed the geminivirus was capable of replicating in potato cells and that replication was efficiently induced by the viral protein, Rep/RepA. However, replication was also observed in the absence of Rep/RepA, putting into question the necessity of this element for geminivirus replication in potato.

The ability of the geminivirus to replicate in potato cells suggested it may also be capable of improving gene targeting rates. This was tested using an established reporter assay and the ZFN, Zif268. These experiments demonstrated that gene targeting is indeed enhanced using the geminivirus in contrast to a conventional *Agrobacterium* delivery. Furthermore, the enhancement in gene targeting seemed to be independent of Rep/RepA expression, further supporting previous observations of geminivirus replication in the absence of Rep/RepA. The improvements in gene targeting using the geminivirus were clearly demonstrated but it was unclear if SSN expression was necessary. To address this, a geminivirus was tested that carried only the repair template and was delivered in the absence of a SSN. The lack of detectable gene targeting demonstrated the importance of SSNs for efficient gene targeting even when additional repair template was made available.

Previous work in tobacco suggested the engineered geminivirus was most effective if the repair template was delivered using the geminivirus, and SSNs could be either delivered on the geminivirus with the repair template or on a separate construct. Both approaches were tested in potato using the *ALS1* TALENs and CRISPR/Cas reagents, with the CRISPR/Cas reagents only tested on a separate construct. This later approach was further facilitated by the development of a Rep/RepA expressing line that did not require transformation of Rep/RepA along with the geminivirus. Evaluation of transformed events revealed gene targeting modifications originating from the repair template and a reduced herbicide susceptibility phenotype conferred by the modifications. Although modifications could be cloned from these events, they could not be easily detected and were most likely somatic in nature. To overcome this issue, individual events were subjected to a second round of regeneration with high selection for gene targeting modification. This approach allowed for direct detection of gene targeting modifications that supported a stronger herbicide susceptibility phenotype. However, the variability in detectable modifications and phenotype across regenerated lines suggested that the modifications remained somatic.

A paradigm in potato breeding

Potato has traditionally been bred at the tetraploid level and vegetatively propagated. However, recent interest in diploid breeding has been picking up pace in both public and private sectors due to the difficulty of fixing important traits, problems with clones accumulating virus, and the re-discovery of self-compatible diploid

germplasm, such as the *S. chacoense* line M6 (Jansky *et al.*, 2014; De Jong and Rowe, 1971; Birhman and Hosaka, 2000; Lindhout *et al.*, 2011). M6 was recently crossed to the potato line used to sequence the potato genome, DM to create the first recombinant inbred lines (RILs) in potato (Endelman and Jansky, 2015). This first RIL population and others in development are valuable resources for mapping important genes and have paved the way for functional genomics in potato.

Self-compatible, inbred germplasm, such as the first RIL population could be used as a power tool for genome editing in potato. Self-compatible, inbred lines capable of being transformed can be modified using genome editing reagents and subsequent modifications can be fixed by selfing. This approach has been shown to be very effective in other self-compatible crops, such as rice and wheat would useful for creating homozygous modifications and segregating out genome editing reagents while maintaining the characteristics of the progenitor line (Li *et al.*, 2012; Wang *et al.*, 2014). Only then can the full potential of genome editing in potato for functional genomics and GM crop development be fully realized.

REFERENCES

REFERENCES

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

Bhaskar, P.B., Wu, L., Busse, J.S., Whitty, B.R., Hamernik, A.J., Jansky, S.H., Buell, C.R., Bethke, P.C. and Jiang, J. (2010) Suppression of the vacuolar invertase gene prevents cold-induced sweetening in potato. *Plant Physiol.*, **154**, 939–48. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2948980&tool=pmcentrez&rendertype=abstract.

Birhman, R.K. and Hosaka, K. (2000) Production of inbred progenies of diploid potatoes using an *S-locus inhibitor* (*Sli*) gene, and their characterization. *Genome*, **43**, 495–502. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10902714.

Brooks, C., Nekrasov, V., Lippman, Z.B. and Eck, J. Van (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.

Clasen, B.M., Stoddard, T.J., Luo, S., et al. (2015) Improving cold storage and processing traits in potato through targeted gene knockout. *Plant Biotechnol. J.*, n/a–n/a. Available at: http://doi.wiley.com/10.1111/pbi.12370.

Dale, M.F.B. and Bradshaw, J.E. (2003) Progress in improving processing attributes in potato. *Trends Plant Sci.*, **8**, 310–312.

De Jong, H. and Rowe, P.R. (1971) Inbreeding in cultivated diploid potatoes. *Potato Res.*, **14**, 74–83. Available at: http://link.springer.com/10.1007/BF02355931.

Endelman, J.B., Jansky, S.H. (2015) Genotyping-by-sequencing of a diploid potato F2 population. Plant and Animal Genome XXIII, San Diego, CA

Feng, Z., Mao, Y., Xu, N., et al. (2014) Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in *Arabidopsis*. *PNAS*, **111**, 4632–4637. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1400822111.

Jansky, S.H., Chung, Y.S. and Kittipadukal, P. (2014) M6: A diploid potato inbred line for use in breeding and genetics research. *J. Plant Regist.* Available at: https://www.crops.org/publications/jpr/abstracts/0/0/jpr2013.05.0024crg.

Joung, J.K. and Sander, J.D. (2012) TALENs: a widely applicable technology for targeted genome editing. *Nat. Rev. Mol. Cell Biol.*, **14**, 49–55. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23169466.

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. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22565958.

Lindhout, P., Meijer, D., Schotte, T., Hutten, R.C.B., Visser, R.G.F. and Eck, H.J. van (2011) Towards F1 hybrid seed potato breeding. *Potato Res.*, **54**, 301–312. Available at: http://link.springer.com/10.1007/s11540-011-9196-z.

Räsänen, E., Pulkkinen, O., Virtanen, T., Zollner, M. and Hennig, H. (2015) Fluctuations of Hi-Hat timing and dynamics in a virtuoso drum track of a popular music recording. *PLoS One*, **10**, e0127902. Available at: http://dx.plos.org/10.1371/journal.pone.0127902.

Sawai, S., Ohyama, K., Yasumoto, S., 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. Available at: http://www.plantcell.org/cgi/doi/10.1105/tpc.114.130096.

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C. and Qiu, J.-L. (2014) Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol.*, **32**, 947–951. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25038773.

Wu, L., Bhaskar, P.B., Busse, J.S., Zhang, R., Bethke, P.C. and Jiang, J. (2011) Developing cold-chipping potato varieties by silencing the vacuolar invertase gene. *Crop Sci.*, **51**, 981. Available at: https://www.crops.org/publications/cs/abstracts/51/3/981.

Zhu, X., Richael, C., Chamberlain, P., Busse, J.S., Bussan, A.J., Jiang, J. and Bethke, P.C. (2014) Vacuolar invertase gene silencing in potato (*Solanum tuberosum* L.) improves processing quality by decreasing the frequency of sugar-end defects. *PLoS One*, **9**, e93381. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3973568&tool=pmcentrez&rendertype=abstract.