CRISPR-CAS9 TARGETED MUTAGENESIS OF GREEN FLUORESCENT PROTEIN TRANSGENE AND A NATIVE, GRANULE BOUND STARCH SYNTHASE, GENE IN POTATO

A Dissertation

by

STEPHANY ELEANA TOINGA VILLAFUERTE

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Keerti S. Rathore
Committee Members,	Maria Isabel Vales
	Carol A. Loopstra
	Clint Magill
	Michael Thomson
Intercollegiate Faculty Chair,	Dirk B. Hays

May 2021

Major Subject: Molecular and Environmental Plant Sciences

ABSTRACT

Potato (*Solanum tuberosum* L.) was genetically modified using the CRISPR-Cas9 system, first to evaluate its efficacy in knocking out a genome-integrated transgene and then to alter its starch profile. The CRISPR system, which acts as an adaptive immune system in *Streptococcus pyogenes*, is being used to engineer eukaryotic genomes in a precise manner. It has proven to be a simple, yet efficient, tool to generate targeted mutations. Potato is an important crop that serves as a staple food for 1.3 billion people and is also a source of starch for industrial, non-food purposes. We first used transgenic potato plants that contained either a single-copy green fluorescence protein (*gfp*) gene insert or four different *gfp* gene inserts to evaluate the efficacy of the CRISPR-Cas9 system to disrupt a target gene integrated in the genome of this crop. Three sgRNAs were examined for targeted knock out of this gene. In combination with Cas9, all three sgRNAs successfully knocked out the gene, however their efficiencies differed greatly. The nature of mutations was analyzed from various knockout events. Once the technology and protocols were established, the second phase of the project involved disruption of the granule-bound starch synthase (*gbssI*) gene using two different sgRNAs. The objective was to modify starch profile by eliminating amylose in the tuber. Lugol-iodine staining of the tuber showed amylose reduction in various edited events and these results were confirmed using both the perchloric acid and enzymatic method. One event out of six examined showed a complete knock out of all *gbss* alleles and total elimination of amylose from the tuber. Viscosity profiles of the tuber starch from different events were determined

using Rapid Visco Analyzer (RVA). The results further confirmed the alterations in starch composition in the tubers from these events.

DEDICATION

To God, for being that perfect father who saved me and who never abandons me, despite my mistakes.

To my mother, Ghen Villafuerte, for being at all times the biggest support in my life and being an example of hard work in all aspects of life.

To my sisters, Samantha and Megan, for being my first and best friends, for being great sources of unconditional support.

To my father, Pedro Toinga, for showing his support in my personal and professional life.

ACKNOWLEDGEMENTS

I would like to give my sincere gratitude to my advisor, Dr. Keerti Rathore, for welcoming me in his laboratory, for his constant guidance and support throughout my research project. I am also tankful to the former and current members of Dr. Rathore's laboratory, Ms. LeAnne Campbell, Dr. Devendra Pandeya and Dr. Madhusudhana Janga, for all their help and support, and for making the laboratory a comfortable space to work every day.

I am thankful to the members of my committee, Dr. Isabel Vales, Dr. Carol Loopstra, Dr. Clint Magill and Dr. Michael Thomson, for their constructive advice and suggestions on my research proposal and dissertation.

My acknowledgements go to Texas A&M University, especially faculty and staff involve with the Molecular and Environmental Plant Sciences Program, for being always welcoming, a great guidance and help for every student part of the program.

The plant material for the first section of this project was provided by the Texas A&M Potato Breeding Program, Department of Horticultural Sciences, Texas A&M University (Dr. Isabel Vales). Also, the measurements of amylose content by perchloric acid method and enzymatic method (Megazyme kit) were performed with help and guidance from current and former members of this laboratory: Douglas Scheuring, Jeewan Pandey, Sanjeev Gautam and Angel Chappell.

Access to the Rapid Visco Analyzer to determine starch viscosity profiles was provided by the Cereal Quality Laboratory, Department of Nutrition and Food Science,

Texas A&M University (Dr. Joseph Awika and Mr. Suleiman Althawab). Furthermore, I would like to acknowledge Dr. Alan Pepper for his guidance in the use of CLC genomics program.

I will always be grateful for all the support and encouragement from my family, my mother Ghen, my sisters Samantha and Megan, my father Pedro my grandparents Olga and Mario and my aunt Chelita. Additionally, I am thankful to all my friends I have met throughout my PhD program from Texas A&M University and Church group. Thanks to God being with me every step of the way and putting the right people in my path.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by dissertation committee consisting of Professor Keerti Rathore (advisor) of the Department of Soil and Crops Sciences, and as committee members: Professor Maria Isabel Vales of the Department of Horticultural Sciences, Professor Carol A. Loopstra of the Department of Ecosystem Science and Management, Professor Clint Magill of the Department of Plant Pathology and Microbiology and Professor Michael Thomson of the Department of Soil and Crops Sciences.

Funding Sources

Graduate study was supported by the College of Agriculture & Life Sciences – Grand Challenges Program and the Molecular & Environmental Plant Sciences Program, Texas A&M University.

NOMENCLATURE

AB	Antibiotics
BD	Breakdown
Cas9	CRISPR associated protein 9
CIM	Callus Induction Medium
cP	Centipoise
CRISPR	Clustered regularly interspaced short palindromic repeats
DSB	Double-stranded break
FV	Final Viscosity
gbss/GBSS	Granule Bound Starch Synthase
gfp/GFP	Green Fluorescence Protein
HR (HDR)	Homologous recombination (Homology-Directed Repair)
hyg	Hygromycin resistance gene
LSM	Linsmaier and Skoog Medium
MS	Murashige and Skoog
NHEJ	Non-homologous end-joining
nptII	Neomycin phosphotransferase II (kanamycin-resistance gene)
PAM	Protospacer Adjacent Motif
PCR	Polymerase Chain Reaction
PIM	Preinduction medium
PT	Pasting temperature

PV	Peak Viscosity
RIM	Root Induction Medium
RVA	Rapid Visco Analyzer
SB	Setback
SIM	Shoot Induction Medium
w/v	weight/volume
YEP	Yeast Extract Peptone

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

1.1. Potato

Potato (*Solanum tuberosum* L.) is an important crop that serves as a staple food for around 1.3 billion people worldwide. Although consumed mainly as a staple food in the developing countries, there is also an increasing demand in the developed countries in recent years, especially for the colorful varieties (imparted by the presence of xanthophylls), driven not only by the appearance but also by the antioxidant activity. Potato is the third most important food crop in the world after rice and wheat (Kathleen *et al*., 2011, Stokstad, 2019).

Potato's place of origin is the Andes mountains of South America; the crop was domesticated in the region of Southern Peru, Northern Bolivia, around Lake Titicaca. It belongs to the family *Solanaceae* which consists of more than 4,000 native varieties of potato; a majority of them localized in the Andes Mountains. Not all of these varieties are used for commercial purposes; in reality only, a few varieties have been selected for consumption. However, wild relatives of potato can serve as a rich source of useful traits (CIP, Watanabe, 2015).

Tubers, the modified stems that are the edible part of this crop, are rich in polysaccharides and secondary metabolites that are required not only for the growth and development of the plant, but also have health benefits when consumed as a part of the human diet. Potato serves as a primary source of carbohydrates in the diets of many people. The tubers are generally composed of around 80% water and 20% solids; however, these

proportions vary among different varieties. The solid portion consists of starch, protein (mainly patatin), vitamins (vitamin C, B1, B3 and B6), flavanols, polyphenols, phenolic acid, riboflavin, folate, potassium, iron, etc. (Navarre *et al*., 2009).

Cultivated potato crops are commonly autotetraploid with 48 chromosomes (2n=4x=48 with a basic chromosome number of 12), but there are also diploid species and even triploid and pentaploid species. Generally, potato is propagated vegetatively, which means that a new plant is obtained from a tuber or piece of tuber, referred to as a "seed" by the potato community. However, potato is capable of generating "true seed", that develops within the fruit (also called a berry) which is similar in shape to a green tomato. True seeds are usually produced during a breeding process and are not used for mass propagation (Watanabe, 2015).

Compared to other crops, production of new commercial varieties of potatoes is not as frequent as the breeding process is very slow and complicated, taking up to 15 years. This is because potato, mainly grown for food, is an autotetraploid crop that is highly heterozygous, vegetatively propagated, and often gametophytic self-incompatible.
Even in the case where they are self-compatible, the progeny suffers from inbreeding depression. Something that breeders have to keep in mind during the crossing of established varieties with wild genotypes, with desirable traits, is that the next generation might lose the established traits and express unwanted traits from the wild type. This is why improvement of a well-established variety is uncommon (Watanabe, 2015, Stokstad, 2019).

Potato is grown worldwide, with China being the biggest producer of potatoes according to FAO data for 2018. The amount produced by China and India (the second biggest producer) represents about one third of the overall world's production. In 2018, worldwide production was over 368 million metric tons, including 422.9 million cwt produced in the United states (USDA-NASS, 2020). Idaho was the biggest producer (130.9 million cwt), with Washington being the second big producer (103.1 million cwt), and Texas producing 6.9 million cwt (USDA-NASS, 2020).

According to FAO estimates, two thirds of potatoes that are harvested worldwide are used as food. In this regard, the tubers are used as either fresh for cooking or they are processed before consumption. For example, frozen potatoes are used for making French fries and chips. In addition, potatoes are used for industrial, non-food uses mainly because of their high starch content. Some 5% to 15% of the tubers produced are utilized as seed for the next planting cycle. Some farmers tend to save their own seed for planting while others prefer to buy disease-free, certified seed (FAO, 2008).

1.2. Starch

Starch is a glucose polymer that serves as the main carbohydrate reserve (long-term storage in amyloplasts). It is synthesized in the leaf chloroplasts during daytime for short-term storage as transient starch consisting of small granules. It is also degraded in photosynthetic tissue during the day and night-time and serves as an energy source for various biological processes. In heterotrophic tissues such as tubers and seeds, it is synthesized in the amyloplasts from sucrose translocated from the source tissues for long-

term storage as large sized granules. The amount of starch that a plant produces and stores depends on environmental factors such as temperature and water stress. Potato starch is also an important source of carbohydrates in human diets, in addition to what is obtained from fruits, cereals, such as corn, wheat and rice, as well as other root and tuber crops such as cassava, yams, and sweet potato. Compared to the seeds (where the starch is accompanied with protein and lipid), the starch in tubers or roots tends to form larger granules and is accompanied with lower protein and lipid content (Whistler & Daniel, 2000, Jobling, 2004, Grommers & van der Krogt, 2009, Balet *et al*., 2019, Martin & Smith, 1995, Van Harsselaar *et al*., 2017, Buchanan *et al*., 2015).

Starch is mainly composed of two types of polysaccharides: amylose and amylopectin; depending on their ratio, the physical and chemical properties of the starch can vary, such as gelatinization, swelling capacity, solubility in water, texture and recrystallisation. For example, when starch is partially hydrolyzed, the resultant products are glucose, maltose and dextrins. The dextrins being used as thickeners in food industrycome from the amylopectin molecule present in the starch. The enzymes involved in the synthesis of both amylose and amylopectin are starch synthases (Denyer *et al*., 2001, Liu, 2017, Jobling, 2004, Fredriksson *et al*., 1998, Karlsson *et al*., 2007, Lemos *et al*., 2019, Buchanan *et al*., 2015).

A representation of the enzymes involved in the process is shown in Figure 1.1.

1.2.1. Amylose

Amylose, a linear polymer composed of α-D-glucose units joined by α-1,4-glycosidic bonds (Figure 1.2), constitutes 20 to 30% of the starch present in a plant, with its relative level depending on the crop and the variety (Jobling, 2004, Lemos *et al*., 2019). This molecule is formed by the enzymatic reaction on ADP-glucose catalyzed by granule bound starch synthase (GBSS) (Figure 1.1). One of the structural characteristics of this molecule is that it has a high capacity to bind iodine in a solution, resulting in a blue coloration (Denyer *et al*., 2001).

In terms of health benefits for humans, higher levels of amylose are desired. This is because a high-amylose starch is more resistant to enzymatic digestion in the stomach resulting in slower sugar absorption in the intestine. The term used for such starch is resistant starch. Thus, for patients suffering from hyperglycemia, higher amylose diets are beneficial for their health because of their lower glycemic index. However, for industrial applications, the opposite is true, i.e., lower levels or even complete absence of amylose in the starch is desirable (Noakes *et al*., 1996, Jobling, 2004).

1.2.2. Amylopectin

Amylopectin, a highly branched polysaccharide composed of glucose units linked primarily by α-1,4-glycosidic bonds (occasionally α-1,6-glycosidic bonds) forming high-density branches having a higher molecular weight and more organized form than amylose, constitutes 70 to 80% of the starch in most crops (Figure 1.2). It is water soluble and stable under storage (the level of stability varies depending on the source). Crops that

show a phenotype with higher amylopectin are known as “waxy” which can also refer to a starch that lacks the presence of amylose or with levels less than 10% (Svegmark *et al*., 2002, Jobling, 2004, Balet *et al*., 2019). When performing iodine-based staining tests, amylopectin shows up as a reddish-brown coloration because its iodine-binding capacity is much lower compared to amylose (Denyer *et al*., 2001).

Starches from “waxy” crops are used mainly for industrial purposes; the reason is that it gelatinizes easily, forms clear pastes, and acts as a stabilizer, thickener and emulsifier in different food products. Because of these characteristics, amylopectin is also more stable under freeze-thaw cycles compared to the other types of starches (Jobling, 2004). Other uses for high-amylopectin starch, besides the food industry are: adhesive for paper and wood, textile and pharmaceuticals and as a biodegradable substitute for polystyrene for manufacturing disposable plates (FAO, 2008).

1.3. CRISPR-Cas9 System

Clustered regularly interspaced short palindromic repeats (CRISPR) and associated protein (Cas9), present in *Streptococcus pyogenes*, act as an adaptive immune system. It has been harnessed to create targeted mutations in the genomes of Eukaryotic organisms. In comparison to other site-directed nuclease systems, it has proved to be simpler, more cost-effective and efficient. The CRISPR-Cas9 system relies on two components: a guide RNA and a nuclease. These two acts together to introduce targeted double-stranded break (DSBs) in the DNA (Cong *et al*., 2013, Jiang & Doudna, 2017).

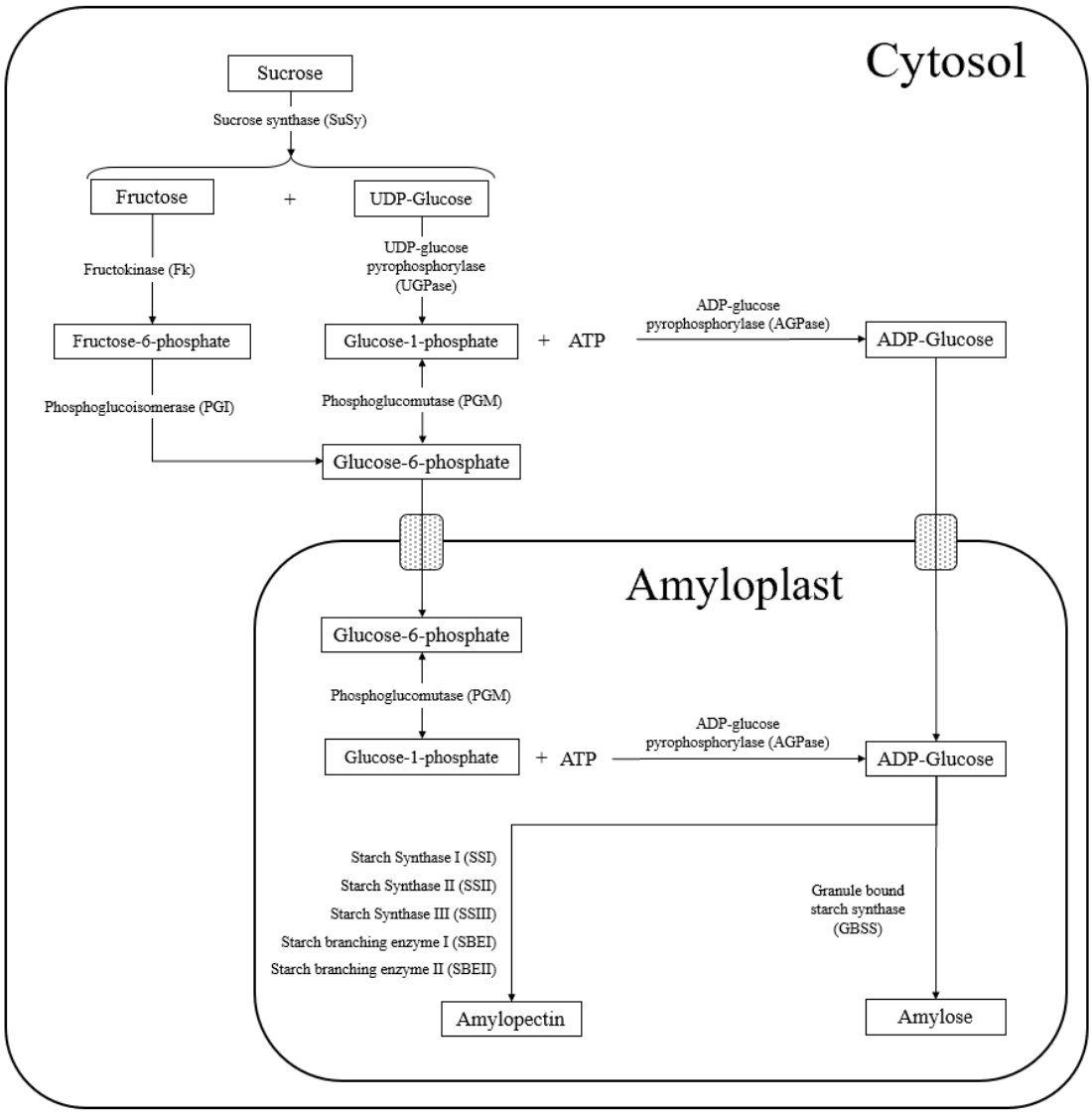


Figure 1.1. Starch biosynthesis pathway adapted from: Beckles (2010), Nazarian-Firouzabadi and Visser (2017) and Van Harsselaar *et al*. (2017).

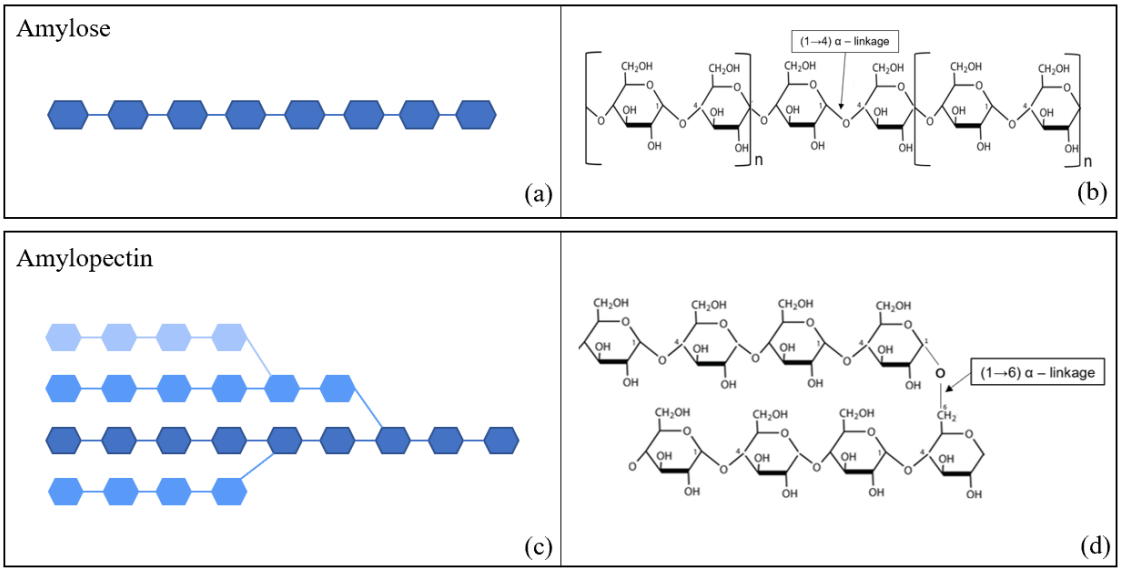


Figure 1.2 (a) Amylose is a linear chain of α-D-glucose units. (b) The amylose molecules are linked together by α-1,4-glycosidic bonds. (c) Amylopectin is a branched chain of α-D-glucose units. (d) The amylopectin branching points are linked through α-1,6-linkages (Adapted from: Bennett "General, Organic and Biochemistry).

As is true for other organisms, plants are constantly exposed to different biological and environmental conditions that can cause damage or breaks in their DNA. For this reason, cells possess repair mechanisms that are triggered in order to maintain genome stability. If the DNA damage is too extensive, cell death pathways are activated. Depending on the damage, there are multiple routes to repair the DNA. If the DNA has DSBs, the cell will use either the more predominant, non-homologous end-joining (NHEJ) repair mechanism or, less frequently, the homologous recombination (HR) mechanism. CRISPR-Cas9 system takes advantage of these repair pathways to introduce mutations (indels or replacements) at the target sites (Figure 1.3) (Qi, 2019, Jiang & Doudna, 2017, Puchta, 2005).

The NHEJ pathway is simpler and the change in the sequence at the DSB site is minimal, mainly in the form of short indels (insertions or deletions). This pathway is more common compared to the HR pathway. During HR, the DNA gets repaired using an available template, for example a homologous sister chromatid. While this repair pathway is more accurate, it occurs at much lower frequency (Que *et al*., 2019).

In this project, CRISPR-Cas9 system was used to create DSBs and trigger the NHEJ pathway, thus introducing indels in the target genes (non-native and native) present in the potato genome and rendering them non-functional. This was achieved by designing a single guide RNA (sgRNA) that consists of tracrRNA:crRNA linked to a ~20 base-long 'guide' that is complementary to one of the strands of the target DNA. An important requirement is that the target DNA strand (protospacer) has to be preceded by a protospacer adjacent motif (PAM site), which for Cas9 consists of "NGG" (Figure 1.3). The sgRNA can be supplied directly in the form of RNA or can be encoded from DNA. Guided by such an sgRNA, the Cas9 nuclease will create a double-stranded break at the DNA target site triggering NHEJ (Cong *et al*., 2013).

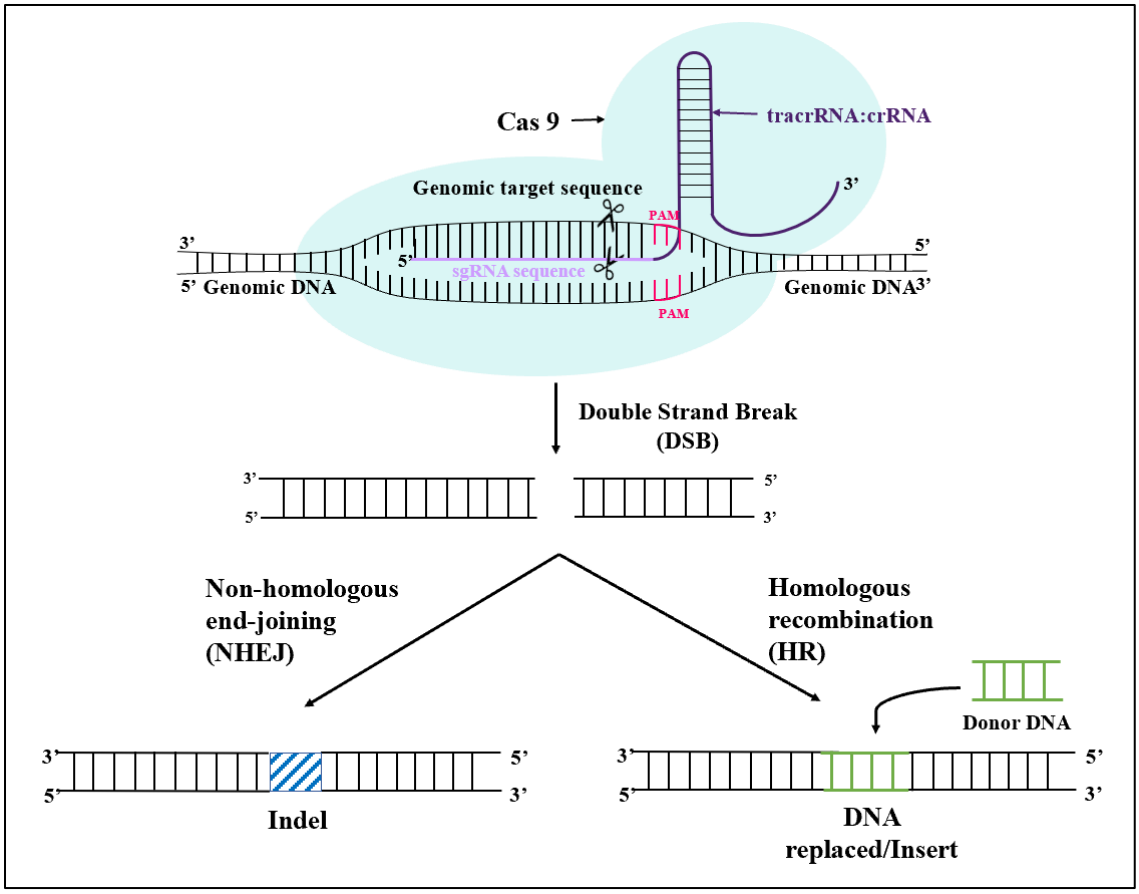


Figure 1.3 Schematic representation of CRISPR-Cas9 system and repair mechanisms. NHEJ: Non-homologous end-joining; HR: Homologous recombination.

CRISPR-Cas9 system has been previously used to introduce modifications in the genome of potato, targeting genes such as: S-RNase (S-locus RNase) gene, which is involved in the gametophytic self-incompatibility of the plant (Enciso-Rodriguez *et al*., 2019) and StPPO2 gene which is part of the family of PPO (Polyphenol Oxidases) genes present in potato that are responsible of the enzymatic browning of the tuber (González *et al*., 2020). A group from Swedish University of Agricultural Sciences (Andersson *et al*., 2017) targeted the potato *gbss* gene in order to change the starch composition through

transient transfection of protoplasts followed by regeneration. The protoplast system has an advantage in that if the editing occurs soon after transfection, the resulting plant will not be chimeric. However, there are some downsides to the use of protoplasts, including the expense and complexity of the procedure of isolating the protoplasts and later to regenerate the plants following extensive cell/tissue culture steps. Moreover, the edited plants likely contain somaclonal variations due to a much longer culture period (Andersson *et al*., 2017). Therefore, in this study, I utilized the *Agrobacterium* method for the transformation of explants and obtained transgenic lines via organogenesis, a simpler alternative system. Note that the genome-integrated CRISPR-Cas9 system in the plant is expected to remain functional throughout the life of a plant and its progeny, generating mutations at the remaining unmutated sites in the cells. However, the CRISPR components can be segregated out in the following generations.

1.4. Objectives

The objectives of this project were:

- To establish a potato transformation protocol that is efficient, and results in the generation of plants with stable integration of the transgene resulting in a desired trait and also capable of forming tubers. My plan is to use green fluorescence protein (*gfp*) as the transgene under this objective.

- Test the efficacy of the CRISPR-Cas9 system to generate mutations in a non-native gene residing in the potato genome. The *gfp* expressing plants generated during the first part of this project are to be used to achieve this objective.

- Once the potato transformation protocol is established and the efficiency of the CRISPR-Cas9 system is evaluated, use it to target a native gene in the potato genome. Specifically, my aim is to target *gbss* gene for mutations to generate amylose-free potato events.

2. MATERIALS AND METHODS

2.1. Transformation protocol

2.1.1. Establishment of potato transformation protocol

To establish which genotype to use for transformation experiments, we evaluated a transformation/regeneration protocol on three potato varieties: Russet Norkotah, White LaSoda (white-skin clonal variant of the red skin potato clone Red LaSoda) and Texas Yukon Gold (a clonal variant of Yukon Gold with fewer eyes), which were provided by the Texas A&M Potato Breeding Program. The protocol published by Chetty *et al*. (2015) was used with some modifications.

2.1.1.1. *Agrobacterium tumefaciens* culture preparation

Agrobacterium tumefaciens (LBA4404) carrying the binary vector, pBINmGFP5-ER construct (Sunilkumar *et al*., 2002), that contains a kanamycin-resistance gene (*nptII*) cassette and also a green fluorescence protein (*gfp*) expression cassette, was used to evaluate transformation/regeneration ability of the three varieties (Figure 2.1). Seven days prior to transformation, bacterial cells were streaked on agar-solidified Yeast Extract Peptone (YEP) medium containing two antibiotics: kanamycin (100 mg/L) and rifampicin (10 mg/L).

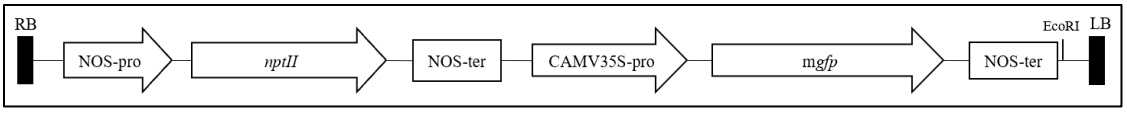


Figure 2.1 T-DNA region of the transformation construct harboring the green fluorescence protein (m*gfp*) gene under the control of the CaMV35S promoter, and kanamycin resistance gene (*nptII*) under the control of the NOS promoter (NOS-pro) for selection.

Following bacterial growth on a plate, 2 mL of YEP medium [supplemented with kanamycin (100 mg/L) and rifampicin [10 mg/L)] was inoculated with a single colony and grown overnight on a shaker at 28 °C at 200 rpm. The next day, two of these cultures were combined in a 15 mL disposable tube and centrifuged at 2800 rpm for 20 min at 4 °C. The supernatant was discarded and using 1 mL of preinduction medium [PIM; (Rathore *et al*., 2015)], the bacterial pellet was resuspended and then brought to 10 mL final volume with PIM, supplemented with 20 µL of acetosyringone stock (10 mg/mL stock). The bacterial suspension was transferred to a 125 mL flask and incubated overnight at 28 °C on a shaker (200 rpm).

Prior to transformation, the culture was adjusted to an OD_{600} of 0.6 with PIM and by adding acetosyringone at the rate of 2 µL (10 mg/mL stock) per mL of culture.

2.1.1.2. **Explant preparation**

Transformation explants were obtained from *in vitro* cultured plants of each variety, which were grown on Linsmaier and Skoog Medium [LSM; (Linsmaier & Skoog, 1965)] at 20 °C and 16h light/8h dark cycle. Internodes proved to be better explants for transformation compared to the leaves. Five mm-long internode segments obtained from

4-week-old plants were therefore used for transformation. Two days prior to transformation, the explants were cut and placed on Callus Induction Medium (CIM; Table 2-1), over Whatman filter paper, and incubated at 25 °C, with a 16h light/8h dark photoperiod.

2.1.1.3. **Transformation**

The explants were incubated with an *Agrobacterium* suspension in a Petri dish on a shaker at at 2 rpm (to avoid damage to the explants) for 20 min. Following inoculation, the explants were blotted dry, to avoid overgrowth of *Agrobacterium*, and were transferred on top of a Whatman filter paper, placed over CIM supplemented with acetosyringone (1 mL/L of 10 mg/mL stock solution) in a Petri plate for two days under conditions described earlier. Following this cocultivation, the explants were transferred to fresh CIM medium supplemented with carbenicillin (500 mg/L), cefotaxime (250 mg/mL) and kanamycin (100 mg/L), without the filter paper.

Explants were subcultured every two weeks until the shoot primordia appeared on the calli that grew from each cut end of the explant. Such cultures were then transferred to a Shoot Induction Medium (SIM). When the new shoots had elongated to at least two cm in length or showed root formation, these were transferred to a Root Induction Medium (RIM). Both SIM and RIM media were supplemented with the same antibiotics (carbenicillin-500 mg/L, cefotaxime-250 mg/mL and kanamycin-100 mg/L). The components of each media are provided in Table 2-1.

Once the regenerated plantlets developed roots, these were transferred to jars containing media and grown under conditions used previously for potato propagation (Linsmaier and Skoog Medium; 20 °C and under 16h light/8h dark photoperiod). For microtuber formation, the medium components were changed to Murashige and Skoog salts, Murashige and Skoog vitamins (Murashige & Skoog, 1962) and 90 g/L of sucrose (without any hormones).

Table 2-1 Media composition used during transformation/regeneration of potato. Callus induction medium (CIM), shooting induction medium (SIM) and rooting induction medium (RIM)

Components	**Callus induction medium**	**Shooting induction medium**	**Rooting induction medium**
MS Salts	4.33 g/L	4.33 g/L	4.33 g/L
MS Vitamins	10 mL/L (100x stock)	10 mL/L (100x stock)	10 mL/L (100x stock)
Sucrose	20 g/L	20 g/L	20 g/L
Myoinositol	100 mg/L	100 mg/L	100 mg/L
NAA	2 mL/L (0.1 mg/mL stock)	200 µL/L (0.1 mg/mL stock)	-
Zeatin Riboside	2.5 mL (1mg/mL stock)	2 mL/L (1mg/mL stock)	-
GA3	200 µL (0.1 mg/mL stock)	200 µL (0.1 mg/mL stock)	-
Agar-Noble	8 g/L	8 g/L	8 g/L
pH	5.8	5.8	5.8

These experiments were conducted to evaluate transformation/regeneration capability of each variety. It was possible to obtain transgenic plants of the variety Yukon Gold only using the protocol described above.

2.1.2. Examination of *gfp* expressing events

To monitor *gfp* expression, each callus event was observed for fluorescence resulting from exposure to UV light under a Zeiss M2 BIO Fluorescence Combination Zoom Stereo/Compound microscope. Fluorescence was visualized using either a 500 nm Long Pass Filter or a 525 nm Narrow Band Pass Filter that blocks red-colored fluorescence from plastids and the photographs were taken with a Zeiss AxioCam color digital camera coupled to the microscope. The shoots that showed green fluorescence resulting from *gfp* expression were selected and transferred to root induction media and propagated through to subsequent stages.

Following regeneration of plants expressing the *gfp* gene, we examined the transgenic events to ascertain transgene copy number via Southern blot analysis before proceeding with next steps. Genomic DNA was obtained from leaves using the Plant Isolate DNA Extraction Kit from Alfa Aesar and Southern blot analysis was performed following restriction enzyme digestion of the DNA with EcoRI and using standard protocols. The regenerated plants expressing the *gfp* gene were grown to tuberization stage, both *in vitro* and in soil.

The objective of this step was to establish a reliable and efficient transformation protocol starting with the *Agrobacterium*-mediated infection of potato explants to regeneration of plants that can survive transfer to soil and grow to produce tubers.

2.2. Examination of the efficacy of CRISPR-Cas9 system to generate mutations in potato

2.2.1. Design of gRNAs to target the non-native, *gfp* transgene present in the transgenic potato genome

In order to evaluate the efficacy of CRISPR-Cas9 system, we first targeted a gene whose mutation results in a phenotype that can be easily and rapidly detected in a non-destructive manner. Thus, we targeted the *gfp* transgene in GFP expressing events that were obtained from the experiments conducted to establish a workable transformation protocol. Two events were selected for this work, one with a single copy of the *gfp* transgene integrated in its genome and the second one with four integrated copies of the *gfp* transgene. Note that since the *nptII* gene and kanamycin were used to select the *gfp* expressing events, the hygromycin-resistance gene and hygromycin selection were used for introducing the Cas9-gRNA construct to knock out *gfp*.

In order to assemble the binary vectors needed for the experiment, the following plasmid vectors were used: pTC212, pTC241 and pCGS751. These vectors were kindly provided by Dr. Daniel Voytas, University of Minnesota. The same three sequences were targeted as those by Janga *et al*. (2017) who had used sgRNA Scorer 2.0 as a web-based

tool to select the guide RNAs to target sequences within the coding region of *gfp* gene. The schematic representation of the three target sites is shown in Figure 2.2.

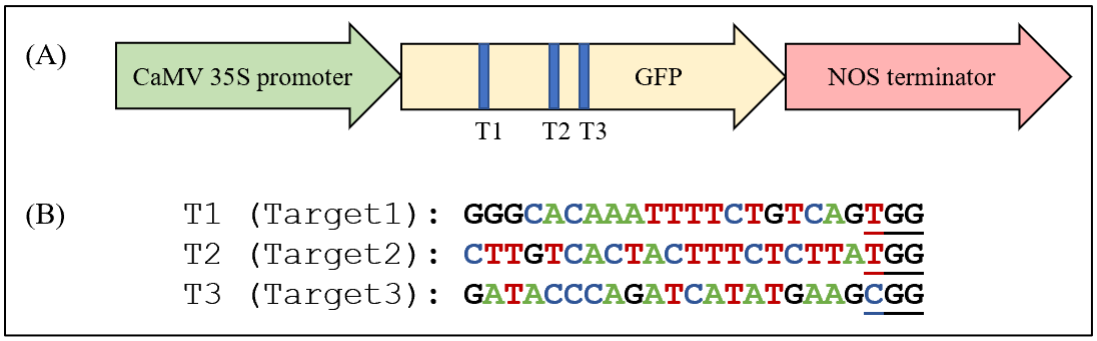


Figure 2.2 (A) Schematic representation of the target sites in the *gfp* coding region. (B) Sequences of the targets selected. Underlined nucleotides represent the PAM (protospacer adjacent motif) sequence. Adapted from Janga *et al*, 2017.

Using the Golden Gate cloning method and Esp31 as the restriction enzyme, each pair of oligonucleotides were annealed (Table 2-2), phosphorylated and introduced into pTC241 (Thermo Fisher Scientific, USA) following the method described in (Cermak *et al*., 2015, Cong *et al*., 2013) with slight modifications. The plasmid vector pTC212 contains the Cas9 gene that was codon optimized for *Arabidopsis*. (Baltes *et al*., 2014)

Table 2-2 Primers corresponding to the target sequences, with the respective overhangs that will allow the oligos to be incorporated into the transformation vector.

Primer	**Sequence 5' to 3'**
Target1-For	gattGGGCACAAATTTTCTGTCAG
Target1-Rev	aaacCTGACAGAAAATTTGTGCCC
Target2-For	gattgCTTGTCACTACTTTCTCTTA
Target2-Rev	aaacTAAGAGAAAGTAGTGACAAGC
Target3-For	gattGATACCCAGATCATATGAAG
Target3-Rev	aaacCTTCATATGATCTGGGTATC

These two DNA sequences, pTC241 and pTC212, were cloned into the binary vector pCGS751 by the Golden Gate cloning method using AarI restriction enzyme (Thermo Fisher Scientific, USA) (Cermak *et al*., 2015, Cong *et al*., 2013). A schematic representation of the final transformation vector is shown in Figure 2.3

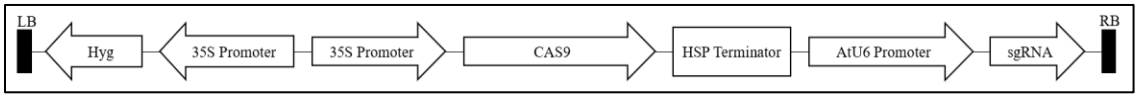


Figure 2.3 T-DNA region of genetic construct a harboring Hygromycin resistance (*Hyg*) gene under the control of the CaMV 35S promoter, CRISPR associated protein 9 (CAS9) under the control of the CaMV 35S promoter and the single guide RNA (sgRNA) expressing sequence under the control of the *Arabidopsis* U6 promoter (AtU6 promoter).

The objective of this step was to become familiar with the gRNA prediction programs and to assess the accuracy of the predicted scores for each gRNA.

2.2.2. Qualitative analysis of knockout events

Following transformation with the Cas9/sgRNA constructs, the explants were observed via fluorescence microscopy every week to examine the knockout phenotype (absence of GFP fluorescence). At the time of transferring new shoots to the root induction media, only the shoots that showed the *gfp*-knockout phenotype were transferred. From this group of regenerated plantlets, ten events were selected randomly to move forward to the next phase of the experiment. Once the selected events were transferred to jars, it was not possible to utilize a fluorescence microscope to observe the plantlets. Instead, a handheld UV light source NIGHTSEA BlueStarTM flashlight in combination with VG2

Filter Glasses were used to observe the fluorescence status of the plants and also to choose 1-2 leaves for DNA isolation for sequencing purposes.

2.2.3. Molecular characterization of mutations

To ascertain the nature of mutations present in the knockout events, PCR primers were designed as specified in Janga *et al*. (2017), flanking the target site and the amplicons were sequenced and aligned to the original sequence. This step allowed us to determine the nature of mutations present in each event. Because we used a single-*gfp*-copy event and also a four-*gfp*-copies event, the sequencing method for each was different.

In case of the single-copy event, Sanger sequencing was used with a set of primers shown in Table 2-3 and the sequences obtained were analyzed and separated using the web-based tool CRISPR-ID.

For the 4-*gfp*-copies event, next generation sequencing was used and primers with tags were designed and the sequences were analyzed using CLC and CRISPResso. (Table 2-4) Use CRISPR-Cas9 system to knock out the *gbss* gene to create amylose-free potato.

Table 2-3 Primers used to PCR amplify the *gfp* fragment where the three target sites are located.

Primer	**Sequence 5' to 3'**
CAS9.GFP.M-F	GATGTGATATCTCCACTGACGTAAG
CAS9.GFP.M-R	TGATAATGATCAGCGAGTTGCACG

Table 2-4 Primers with tags used to PCR amplify the *gfp* fragment containing the target sites.

Targets	Event	Forward	Reverse
Target 1	1	tctagcGGACTCTAGAGGATCCAAGGAG	cagtcaGATCCTGTTGACGAGGGTGTC
	2	tctagcGGACTCTAGAGGATCCAAGGAG	acgtcaGATCCTGTTGACGAGGGTGTC
	3	gagtcaGGACTCTAGAGGATCCAAGGAG	cgactgGATCCTGTTGACGAGGGTGTC
	4	gagtcaGGACTCTAGAGGATCCAAGGAG	tgatagGATCCTGTTGACGAGGGTGTC
	5	gagtcaGGACTCTAGAGGATCCAAGGAG	gtcacgGATCCTGTTGACGAGGGTGTC
	6	gagtcaGGACTCTAGAGGATCCAAGGAG	atgatgGATCCTGTTGACGAGGGTGTC
	7	gagtcaGGACTCTAGAGGATCCAAGGAG	cagtcaGATCCTGTTGACGAGGGTGTC
	8	gagtcaGGACTCTAGAGGATCCAAGGAG	acgtcaGATCCTGTTGACGAGGGTGTC
Target 2	1	gtactcGGACTCTAGAGGATCCAAGGAG	cgactgGATCCTGTTGACGAGGGTGTC
	2	gtactcGGACTCTAGAGGATCCAAGGAG	tgatagGATCCTGTTGACGAGGGTGTC
	3	gtactcGGACTCTAGAGGATCCAAGGAG	gtcacgGATCCTGTTGACGAGGGTGTC
	4	gtactcGGACTCTAGAGGATCCAAGGAG	atgatgGATCCTGTTGACGAGGGTGTC
	5	gtactcGGACTCTAGAGGATCCAAGGAG	cagtcaGATCCTGTTGACGAGGGTGTC
	6	gtactcGGACTCTAGAGGATCCAAGGAG	acgtcaGATCCTGTTGACGAGGGTGTC
	7	tctagcGGACTCTAGAGGATCCAAGGAG	cgactgGATCCTGTTGACGAGGGTGTC
	8	tctagcGGACTCTAGAGGATCCAAGGAG	tgatagGATCCTGTTGACGAGGGTGTC
	9	tctagcGGACTCTAGAGGATCCAAGGAG	gtcacgGATCCTGTTGACGAGGGTGTC
	10	tctagcGGACTCTAGAGGATCCAAGGAG	atgatgGATCCTGTTGACGAGGGTGTC
Target 3	1	gctagtGGACTCTAGAGGATCCAAGGAG	gtcacgGATCCTGTTGACGAGGGTGTC
	2	gctagtGGACTCTAGAGGATCCAAGGAG	atgatgGATCCTGTTGACGAGGGTGTC
	3	gctagtGGACTCTAGAGGATCCAAGGAG	cagtcaGATCCTGTTGACGAGGGTGTC
	4	gctagtGGACTCTAGAGGATCCAAGGAG	acgtcaGATCCTGTTGACGAGGGTGTC
	5	atgctaGGACTCTAGAGGATCCAAGGAG	cgactgGATCCTGTTGACGAGGGTGTC
	6	atgctaGGACTCTAGAGGATCCAAGGAG	tgatagGATCCTGTTGACGAGGGTGTC
	7	atgctaGGACTCTAGAGGATCCAAGGAG	gtcacgGATCCTGTTGACGAGGGTGTC
	8	atgctaGGACTCTAGAGGATCCAAGGAG	atgatgGATCCTGTTGACGAGGGTGTC
	9	atgctaGGACTCTAGAGGATCCAAGGAG	cagtcaGATCCTGTTGACGAGGGTGTC
	10	atgctaGGACTCTAGAGGATCCAAGGAG	acgtcaGATCCTGTTGACGAGGGTGTC

2.2.4. Design of sgRNAs to target a native gene in potato

The sequence of the target gene, encoding Granule Bound Starch Synthase (GBSSI), was obtained from the database Spud DB-Potato Genomics Resource at Michigan State University and primers were designed based on the sequence (Table 2-5).

Table 2-5 Primers used to PCR amplify *gbss* fragment where target sites are located.

Primer	**Sequence 5' to 3'**
GBSS-GENE-FOR	ATGGCAAGCATCACAGCTTCAC
GBSS-GENE-REV	CATTGTCCAGATAATCTAGTCCAGC

Because the sequence present in the database is not from the commercial variety used in this project (Yukon Gold), the first step was to amplify the *gbss*I gene from Yukon Gold genomic DNA and verify the sequence for the target sectors in order to design the gRNAs. Once the sequence of the gene was confirmed, gRNAs were designed with the aid of web-based tools: sgRNA Scorer 2.0, WU-CRSPR and CRISPR-P2 2.0.

In order to increase the chances of obtaining a knockout mutation, two different gRNAs were designed to target two different sites within the coding region. The binary vector used for this set of experiments was assembled following the same steps that were used to construct the vector used to knock out the *gfp* gene. The oligonucleotide pairs that were annealed, phosphorylated and introduced into pTC241 are presented in the Table 2-6, and the schematic representation of the two target sites within the gene can be seen in Figure 2.4.

Table 2-6 Primers corresponding to the target sequences, with the respective overhangs that will allow the oligos to be incorporated into the transformation vector.

Primer	**Sequence 5' to 3'**
Target1-For	aaacTTGTTTGTGGAAAGGGAATG
Target1-Rev	gattgCATTCCCTTTCCACAAACAA
Target2-For	gattgAATCTTCCTGATGAATTCAG
Target2-Rev	aaacCTGAATTCATCAGGAAGATT

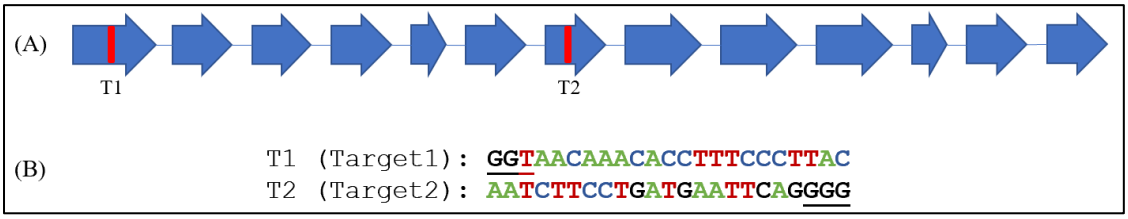


Figure 2.4 (A) Schematic representation of the target sites in the *gbss*I gene. (B) Sequences of the selected targets. Underlined nucleotides represent the PAM (protospacer adjacent motif) sequence.

Once the oligos were incorporated into pTC241, this and pTC212 (which contains the Cas9), were cloned into the binary vector pCGS751 by the Golden Gate cloning method using AarI restriction enzyme (Thermo Fisher Scientific, USA) (Cermak *et al*., 2015, Cong *et al*., 2013). Each gRNA was incorporated into a different construct and these were transferred separately into *Agrobacterium* cells. The schematic representation of the construct is shown in Figure 2.5.

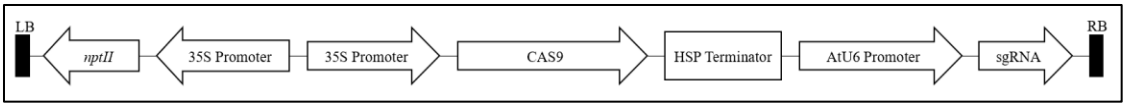


Figure 2.5 T-DNA region of gene construct harboring kanamycin resistance (*nptII*) gene under the control of CaMV 35S promoter, CRISPR associated protein 9 (CAS9) under the control of CaMV 35S promoter and the single guide RNA (sgRNA) expressing sequence under the control of *Arabidopsis* U6 promoter (AtU6 promoter).

2.2.5. Qualitative analysis of knockout events

Because the desired phenotype could only be assessed at the tuberization stage, all the regenerated, independent events obtained following transformation were grown until the *in vitro* tuberization stage. The microtubers growing on each individual event were tested via the histochemical Lugol-Iodine method. The expected knockout, higher amylopectin phenotype was visible as red-colored starch staining instead of the normal blue coloration.

Three individual events for each target, with and without the desired phenotype, were moved to soil in order to obtain plants that produce normal-size tubers. These tubers were also tested via the histochemical Lugol-Iodine method to confirm the phenotype seen previously in microtubers. Once we had qualitative results from the tubers grown in soil, the remaining tuber, specifically the eyes, were used as seeds to grow plants that were then used for the molecular characterization and the quantitative starch composition/property analyses.

2.2.6. Molecular characterization of mutations

Once we obtained the events that showed the desired phenotype, the next step was to analyze the nature of mutations that were present in the transformed events. The PCR primers designed flanking each of the target sites are shown in Table 2-7.

Table 2-7 Primers used to PCR amplify the *gbss* fragment where each target site is located.

Primer	**Sequence 5' to 3'**
GBSS-Target 1-F	ATGGCAAGCATCACAGCTTCAC
GBSS-Target 1-R	CATTGTCCAGATAATCTAGTCCAGC
GBSS-Target 2-F	TCACTCGATTGCACGTTACC
GBSS-Target 2-R	GTAAAGGTTTTGCGTCCATGACC

Respective PCR amplicons were cloned into linearized pMiniTTM vector using the PCR Cloning kit (New England BioLabs) followed by its transformation into DH5α *E. coli.* Twenty bacterial colonies for each target were used to extract DNA that were then sequenced via Sanger sequencing. These sequences were aligned to the original gene sequence.

This step allowed us to determine the nature of mutations and aided in correlating the presence of mutations with the phenotype observed.

2.2.7. Specific gravity measurements and quantitative analysis of starch composition in knockout events

The harvested tubers were first weighed in the air and then in water in order to calculate the specific gravity values for each event using the following equation:

$$Specific\ gravity = \frac{Weight\ in\ air}{weight\ in\ air - weight\ in\ water}$$

The freeze dried samples of tubers from each event were obtained following the method described by Fajardo *et al*. (2013). This involved cutting the tubers into small cubes, storing them in a freezer overnight at -80 °C, and then subjecting them to freeze drying. After 4 to 5 days of freeze drying, the samples were ground and stored at -80 °C.

The percentage of dry matter content in the tubers from various events was also calculated after measuring the weight of the samples before and after freeze drying.

Two different methodologies were used to measure the percentage of amylose present in the tubers from each event. The first method used perchloric acid and the second one was based on an enzymatic reaction.

2.2.7.1. **First Method: Use of perchloric acid solution**

First, an amylose/amylopectin standard curve was generated using different proportions of each compound. From each biological sample, 20 – 30 mg was mixed with 500 µL of 45% (w/v) perchloric acid solution in a 50 ml plastic tube in a plastic tube ???? and incubated at room temperature for four minutes. Then, 16 mL of ultra-pure water was added and the contents were mixed by vortexing. This solution in the tube was kept on the lab bench for ten minutes before transferring 40 µL to a microtiter plate. After adding 50 µL of iodine solution (2 g KI and 1g I_2 mixed in 1 L of ultra-pure water) to each well, the

absorbance was read at 550 nm and 620 nm in SpectraMax® 190 Absorbance Plate Reader, Molecular Devices.

The absorbance values from solutions with different amylose/amylopectin ratio obtained to generate a standard curve (Figure 2.6) are specified in Table 2-8.

Table 2-8 Absorbance values at 550 nm and 620 nm for solutions with different ratios of amylose/amylopectin, used to generate a standard curve.

% of Amylose (y)	% of Amylopectin	550nm	620nm	620/550 (x)
0	100	0.1047	0.077	0.735435
10	90	0.1048	0.0864	0.824427
20	80	0.1064	0.0967	0.908835
30	70	0.1086	0.1069	0.984346
40	60	0.1052	0.1114	1.058935
50	50	0.1113	0.1295	1.163522
60	40	0.1098	0.1302	1.185792
70	30	0.1137	0.1451	1.276165
80	20	0.1166	0.1547	1.326758
90	10	0.1225	0.1698	1.386122
100	100	0.1205	0.1734	1.439004
Blank	Blank	0	0	

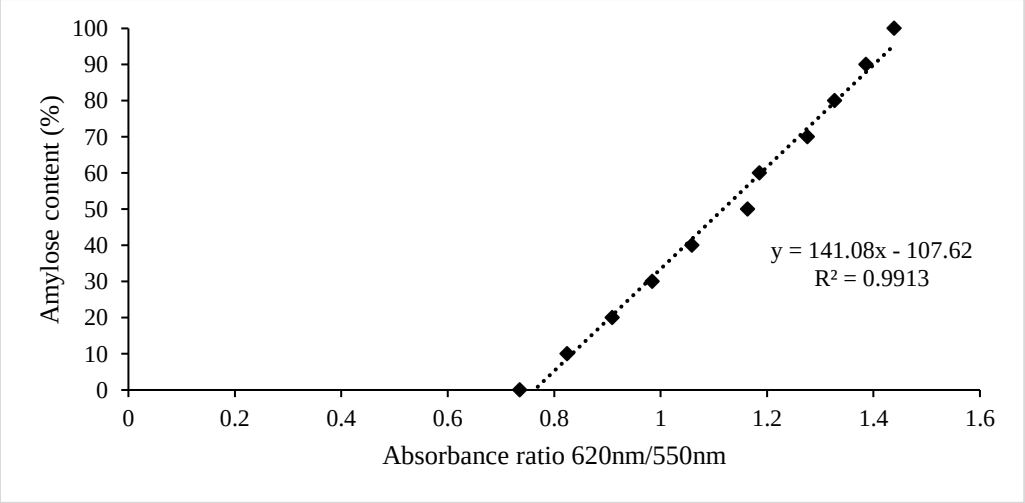


Figure 2.6 Standard curve for different amylose/amylopectin mixtures.

After linearization, the following equation (with an R2 of 0.9913) was obtained (Figure 2.6):

$$y=141.08x-107.6$$

Where y is the percentage of amylose present in the sample and x is calculated by dividing the OD620 with OD550 for each sample.

2.2.7.2. **Second Method: Use of commercial Amylose/Amylopectin kit (Megazyme, Bray, Ireland)**

To conduct this assay, 20-25 mg of each biological sample was used. The assay is based on the separation of amylopectin from amylose using Concanavalin A. This lectin binds and precipitates amylopectin leaving amylose molecules in the supernatant that can then be measured colorimetrically with glucose/peroxidase. The manufacturer's instructions were followed for each step to prepare the samples, and then 40 µL from the final solution was transferred to a microtiter plate, and the absorbance was read at 510 nm.

Once measurements were taken for both: a set of biological samples treated with the kit and a set of the biological samples that were not treated with it, the following formula was applied to obtain the percentage of amylose (w/w) present in each sample:

$$Amylose\ (\%) = \frac{Absorbance\ (Con\ A\ Supernatant)}{Absorbance\ (Total\ Starch\ Aliquot)} x66.8$$

2.2.8. Starch Viscosity Measurements

Viscosity was measured using Rapid Visco Analyzer (RVA, PerkinElmer). Again, freeze-dried samples were used for this assay. One g of ground flour sample was used for

this assay. The moisture values for various samples were fairly consistent. Distilled water was added to each one gram of sample to bring final suspension weight to 28 g. The program used to calculate viscosity was a standard program supplied by the equipment manufacturer.

3. RESULTS

3.1. Transformation protocol

3.1.1. Establishment of potato transformation protocol

The protocol published by Chetty *et al*. (2015) with some modifications was used for potato transformation. Out of the three varieties that were evaluated, only one variety, Yukon Gold, was able to regenerate into plantlets; the other two varieties, Russet Norkotah and White LaSoda did not progress beyond the callus stage. (Table 3-1, Figure 3.1)

Steps related to the transformation/regeneration protocol, starting from the preparation of the explants (pre-inoculation) to obtaining tubers in the soil, are shown in Figure 3.2 The entire process can take at least nine months and sometimes longer.

Regenerated plantlets were incubated in dark on Murashige and Skoog (MS) medium, MS vitamins and sucrose (90 g/L) to obtain *in vitro* tuberization. Various stages from transformation to microtuber formation are shown in Figure 3.3.

Table 3-1 Comparison among Russet Norkotah, White LaSoda and Texas Yukon Gold varieties, related to their ability to form callus and regenerate following transformation.

Variety	Callus formation (%)	Callus transformed (%)	Regeneration from transformed callus (%)
Russet Norkotah	100	92.86	0
White LaSoda	100	44	0
Texas Yukon Gold	96.15	68.00	64.71

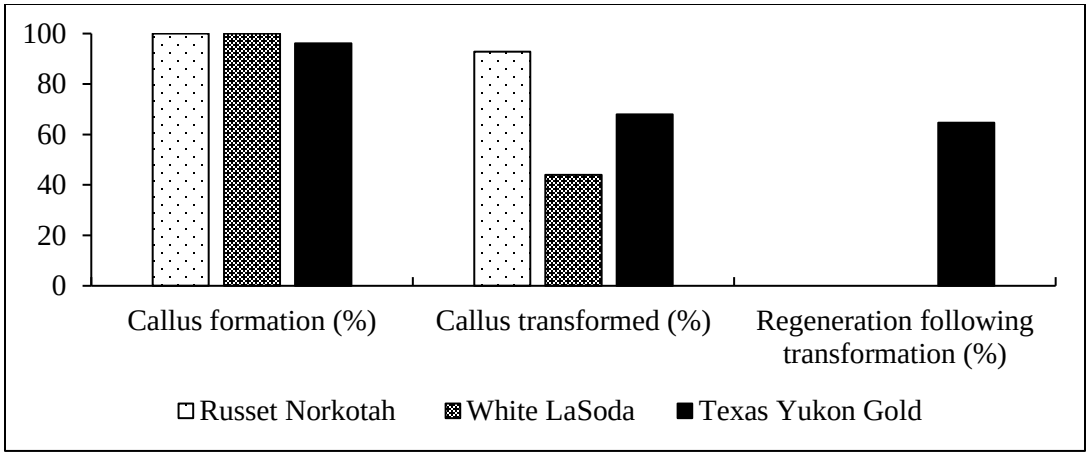


Figure 3.1 Comparison among three varieties (Russet Norkotah, White LaSoda and Texas Yukon Gold) related to their ability to form callus and regenerate following transformation.

In case of transformation with the *gfp* gene, the tissues were monitored at each stage under a fluorescence microscope (for the presence of GFP activity. During the shoot induction stage, only the shoots that showed GFP expression were transferred to the next medium for root induction. (Figure 3.4)

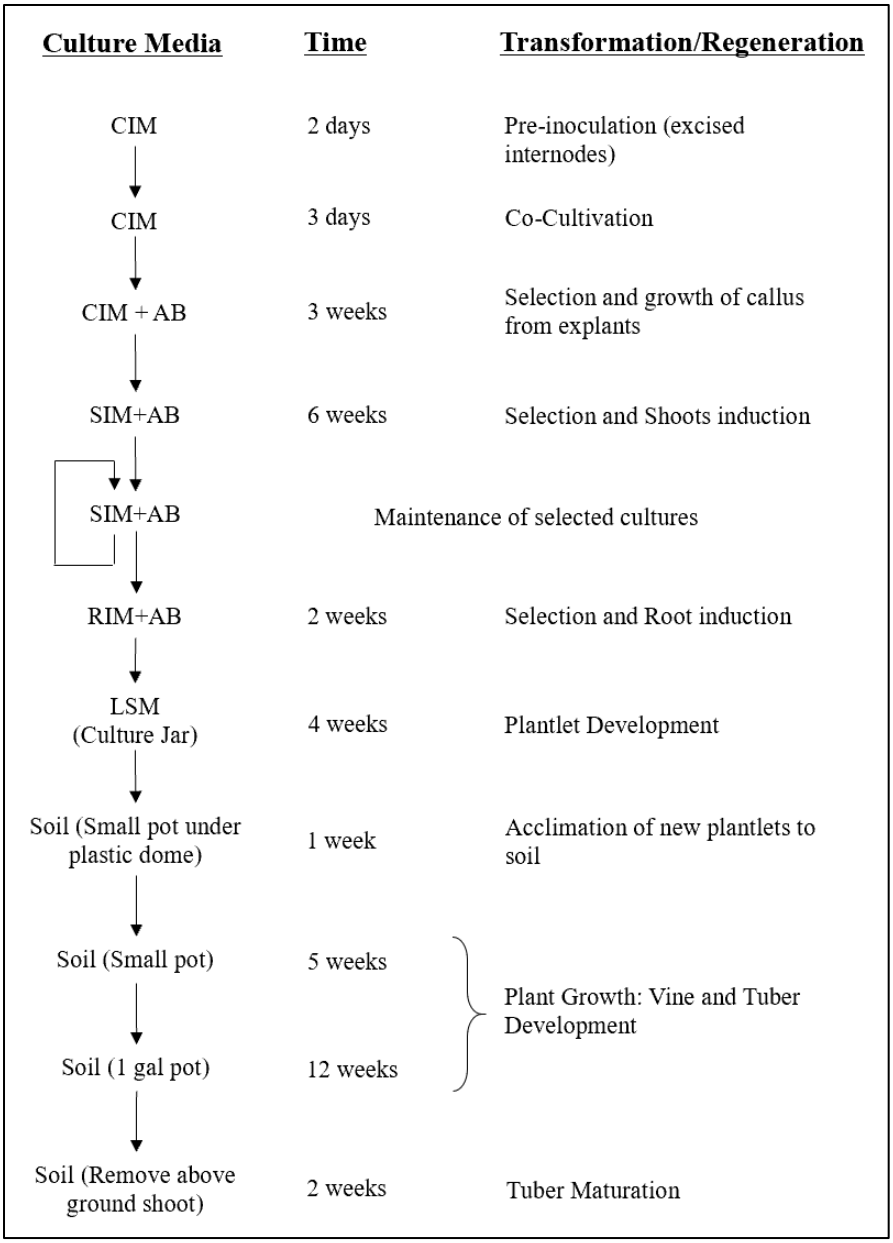


Figure 3.2 Culture steps and media used to transform, select and regenerate potato plants. CIM: Callus Induction Medium; SIM: Shoot Induction Medium; RIM: Root Induction Medium; LSM: Linsmaier and Skoog medium; AB: Antibiotics.

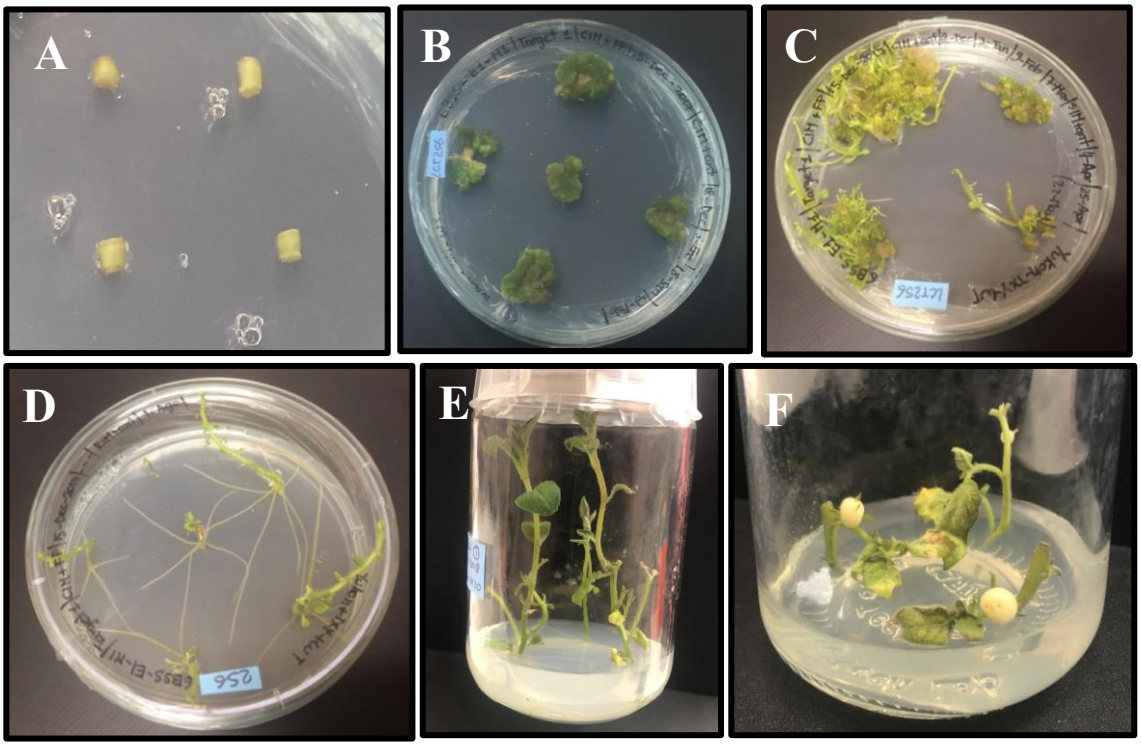


Figure 3.3 Stages of potato transformation/regeneration/microtuber formation. A: Internode explants used for transformation. B: Callus formation stage. C: Shoot formation stage. D: Root formation stage. E: Maintenance/plantlet development stage in a culture jar. F: *In vitro* tuberization stage in a culture jar.

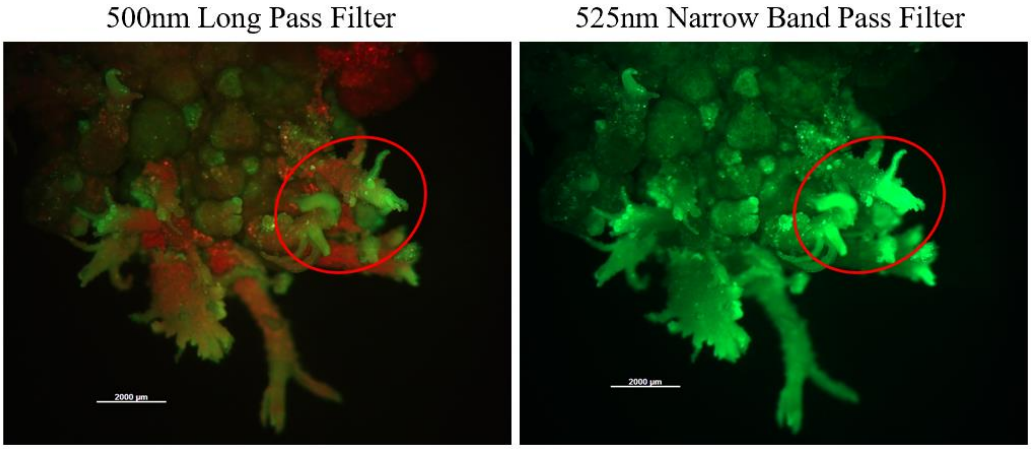


Figure 3.4 Regenerating culture on Shoot-induction medium as seen under 500 nm long pass filter or 525 nm narrow band-pass filter. Shoots showing GFP fluorescence can be seen inside the red circle.

3.1.2. Examination of the GFP events

Ten randomly selected events were evaluated visually for *gfp* expression. These were scored as follows: (+) Low *gfp* expression, (++) Medium *gfp* expression and (+++) High *gfp* expression. Examples of various *gfp* expression levels are shown in Figure 3.5.

Ten events that were scored are shown in Table 3-2. Events 3, 4 and 10 being the ones with the lowest score; Events 1, 5, 8 and 9 having a medium score and Events 2, 6 and 7 being the events with the highest score.

Using standard protocol, Southern blot analysis was conducted following restriction enzyme digestion of the genomic DNA with EcoRI. Two different probes were used, separately, to detect either the *nptII* or the *gfp* gene. This analysis allowed us to ascertain transgene copy number in various *gfp*-expressing events.

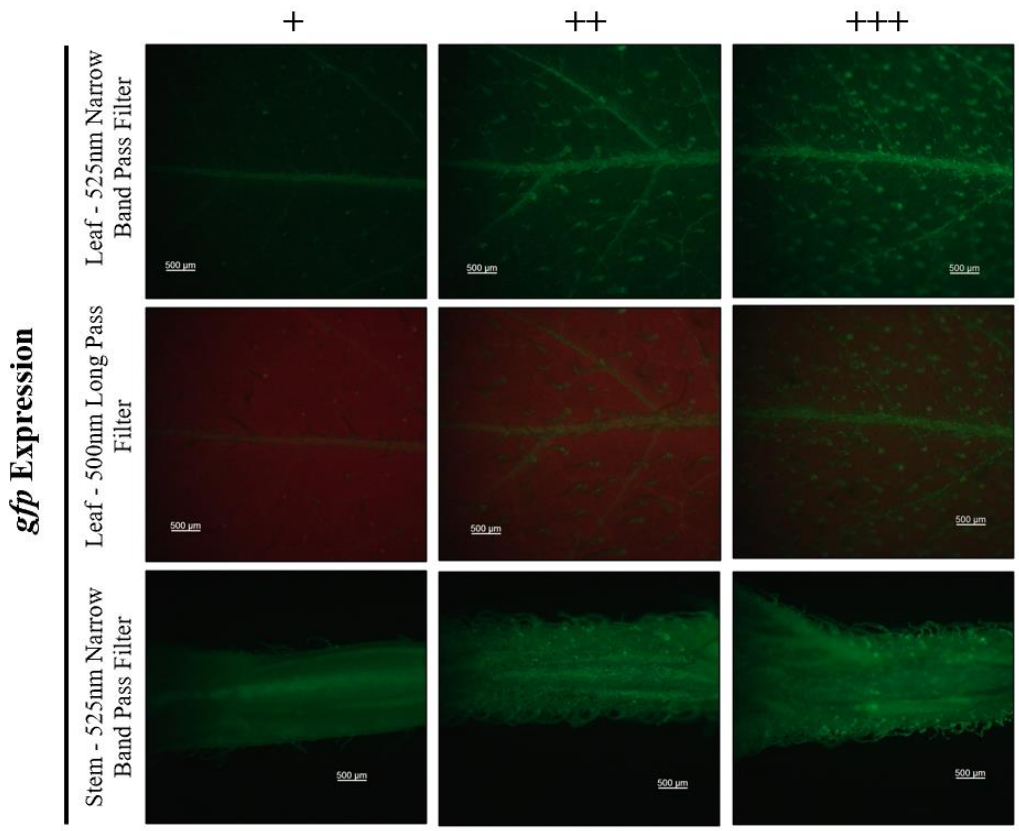


Figure 3.5 Examples of each type of GFP expression score: (+) Low *gfp* expression, (++) Medium *gfp* expression and (+++) High *gfp* expression. The tissues examined were leaf and stem, either under 500 nm long pass filter or 525 nm narrow band-pass filter.

Table 3-2 Visual fluorescence intensity scores for ten *gfp* expressing events. Scores being: (+) Low *gfp* expression, (++) Medium *gfp* expression and (+++) High *gfp* expression.

Event	***gfp* expression**
1	++
2	+++
3	+
4	+
5	++
6	+++
7	+++
8	++
9	++
10	+

With the *nptII* probe, we found six events to have two copies and four events to have single-copy integration. In case of the events showing two copies, events 2, 6, 7 and 8 showed a similar pattern for the double bands, which possibly indicates that these events were siblings. Only four events were probed with the *gfp* gene to confirm if these showed the same *gfp* copy number as the *nptII* gene. Event 6 showed integration of four copies of the *gfp* fragment, while event no. 5, 9 and 10 showed single-copy integration of the *gfp* gene (Figure 3.6).

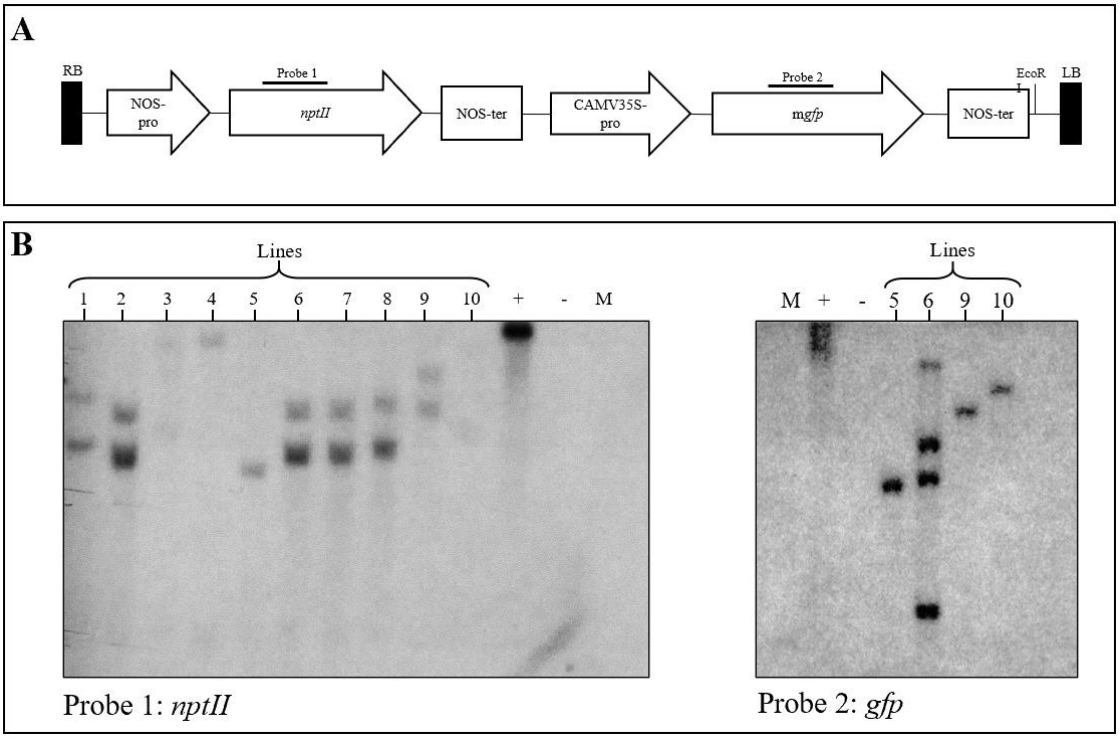


Figure 3.6 Southern blot analysis on *gfp* expressing events. A: Partial map of the transformation construct harboring the green fluorescence protein (*gfp*) and kanamycin resistance (*nptII*) genes. The positions of EcoRI restriction site and the two probes are also shown. B: Blots obtained using the *nptII* probe (left) and the *gfp* probe (right).

It is important to note that in case of Event 6, that showed integration of four copies of the *gfp* gene, not all copies are necessarily functional because transgene expression depends on the site of integration in the genome and the integrity of the entire GFP expression cassette.

The results described above establish that I have a reliable and efficient *Agrobacterium*-mediated transformation and regeneration protocol for potato. Importantly, the regenerated plants can survive transfer to soil and grow to produce tubers.

3.2. Examination of the efficacy of CRISPR-Cas9 system to generate mutations in potato

3.2.1. Retransformation of *gfp*-expressing potato events with gene-editing reagents

Two events, obtained from the experiments described earlier, were selected for this work. One event had a single copy of the *gfp* transgene integrated (Event 5) and the second event had four copies of the *gfp* transgene integrated in its genome (Event 6). Because both the events were obtained using the *nptII* gene for selection, the plant selectable marker used for CRISPR experiments was the hygromycin-resistance gene, *hyg*. The gRNAs tested were the ones designed previously by Janga *et al*. (2017) described in Materials and Methods section.

3.2.2. Qualitative analysis of knockout events

Following transformation with Cas9/sgRNA constructs to knock out the *gfp* gene, the explants were examined via fluorescence microscopy every week to find the knockout phenotype (absence of GFP fluorescence). Starting at week three, calli for gRNA2 and gRNA3 started to show GFP silencing; and at week four, calli for gRNA1 showed GFP silencing. As shown in Figure 3.7, silencing of the *gfp* gene was more noticeable with gRNA3, followed by gRNA2 and gRNA1 being the lowest in terms of showing the expected phenotype. These differences were seen throughout the culture process and were observed in both, single-copy event and four-copy event.

Thirty days after infection, the highest knock out efficiencies were observed with gRNA3 at 59.69% (single *gfp* copy event) and 53.23% (four *gfp* copies event) based on visual observations, followed by 22.96% and 26.25% for gRNA2 (single and four *gfp*-copy, respectively) and 11.73% and 23.13% for gRNA1 (single and four *gfp*-copy, respectively). Thus, gRNA3 proved to be the most efficient while gRNA1 was the least effective in knocking out *gfp* gene function. (Figure 3.8)

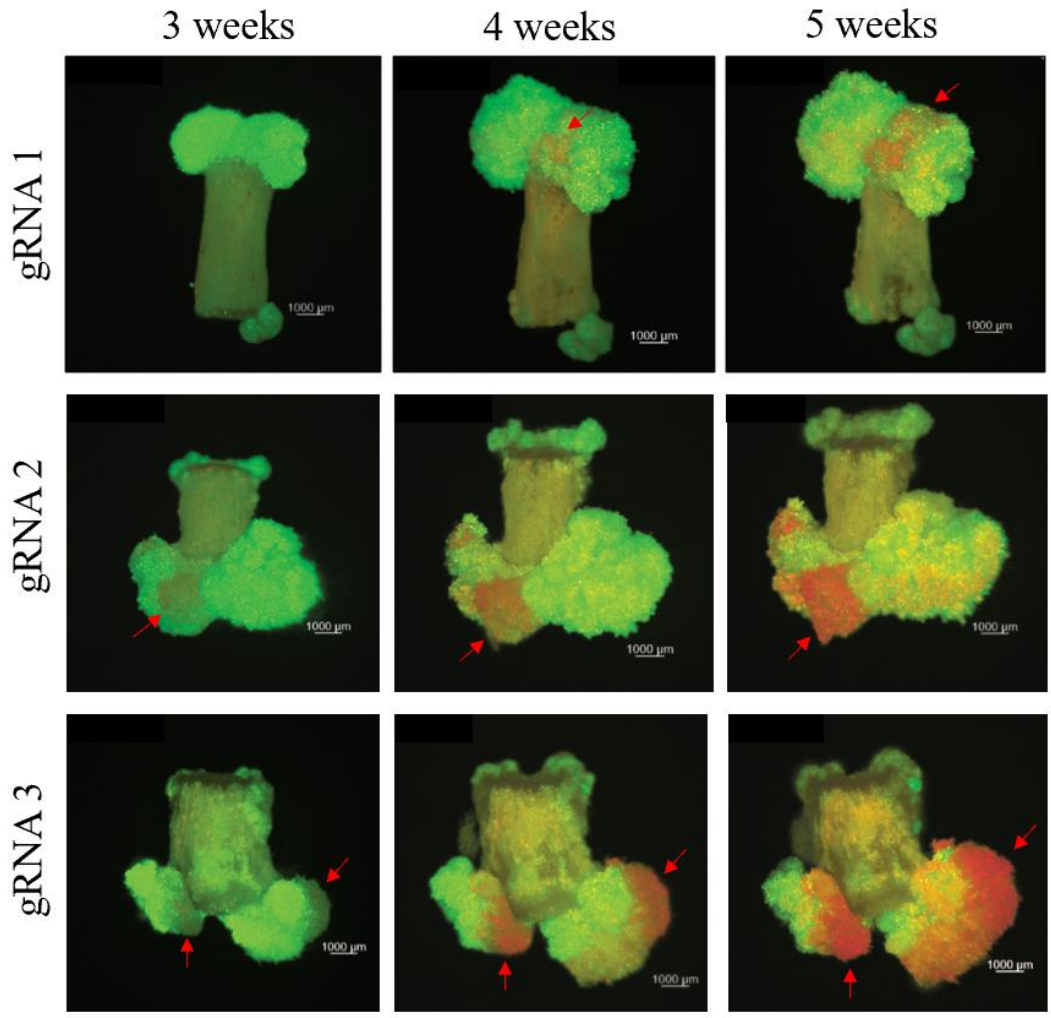


Figure 3.7 Progression of CRISPR-Cas9-mediated mutations in the *gfp* gene in the callus tissue growing at the cut surface of internode explants at 3 weeks, 4 weeks and 5 weeks following transformation. Red arrows show parts of callus where the *gfp* gene has been successfully mutated and became non-functional.

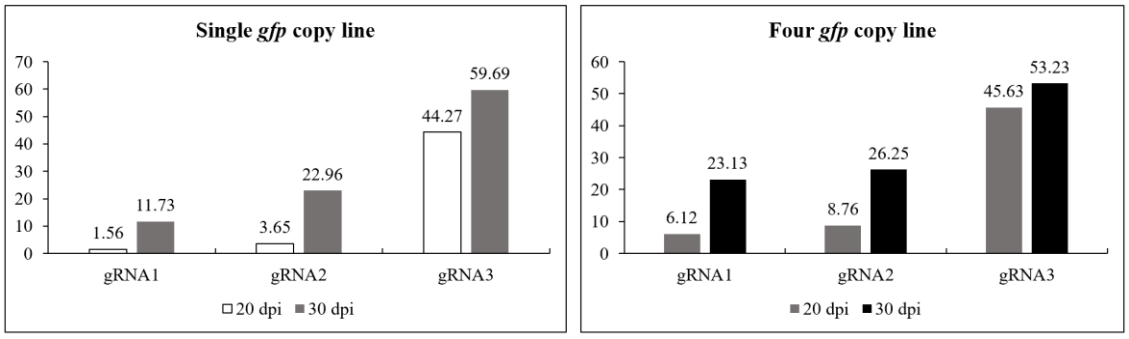


Figure 3.8 *gfp*-fluorescence suppression observed in cultures at 20- and 30-days post infection in response to three different gRNA constructs. Calli were observed visually under a fluorescence microscope.

Approximately, nine weeks after infection, only the shoots that showed *gfp*-knockout phenotype were transferred to the root induction media (Figure 3.9) to regenerate complete plantlets. Multiple events from each gRNA tested were selected randomly to move forward to the next phase of the experiment and to perform molecular characterization.

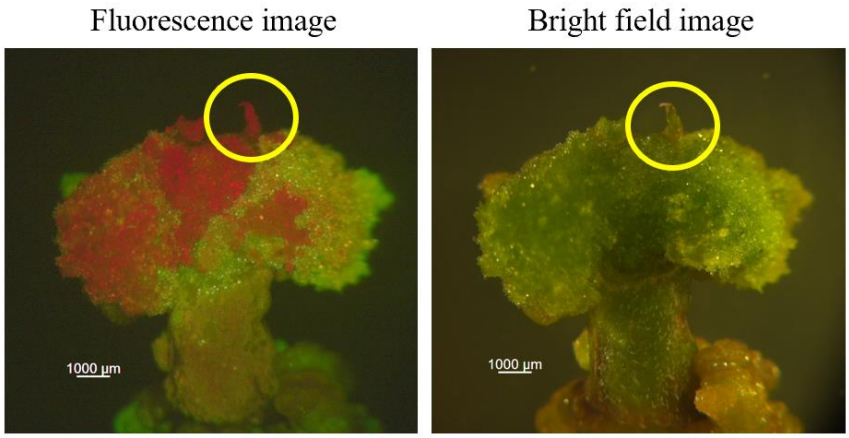


Figure 3.9 *gfp* silencing in cultures on shoot induction medium, 9 weeks after infection. Knockout shoot primordium circled in yellow.

3.2.3. Molecular characterization of mutations

In order to ascertain the nature of mutations in both single-copy and four-copy events, the target portion of the genomic DNA was sequenced using different techniques. DNA from the event that had a single copy of the *gfp* transgene integration was PCR amplified and sequenced by Sanger sequencing. The event with four *gfp* copies integrated was sequenced using Next generation sequencing.

3.2.3.1. Nature of mutations in single-copy *gfp* transgene integration event

Sanger sequencing results for single-copy *gfp* knockout events showed a variety of mutations, including SNPs, small (1-4 bp) indels and large deletions (10 – 73 bp). Various types of mutations observed are shown in Figure 3.10. Interestingly, gRNA1 and gRNA2 resulted in a wider variety of mutations compared to gRNA3. Whereas gRNA3 resulted in small deletions that were similar in nature among different events.

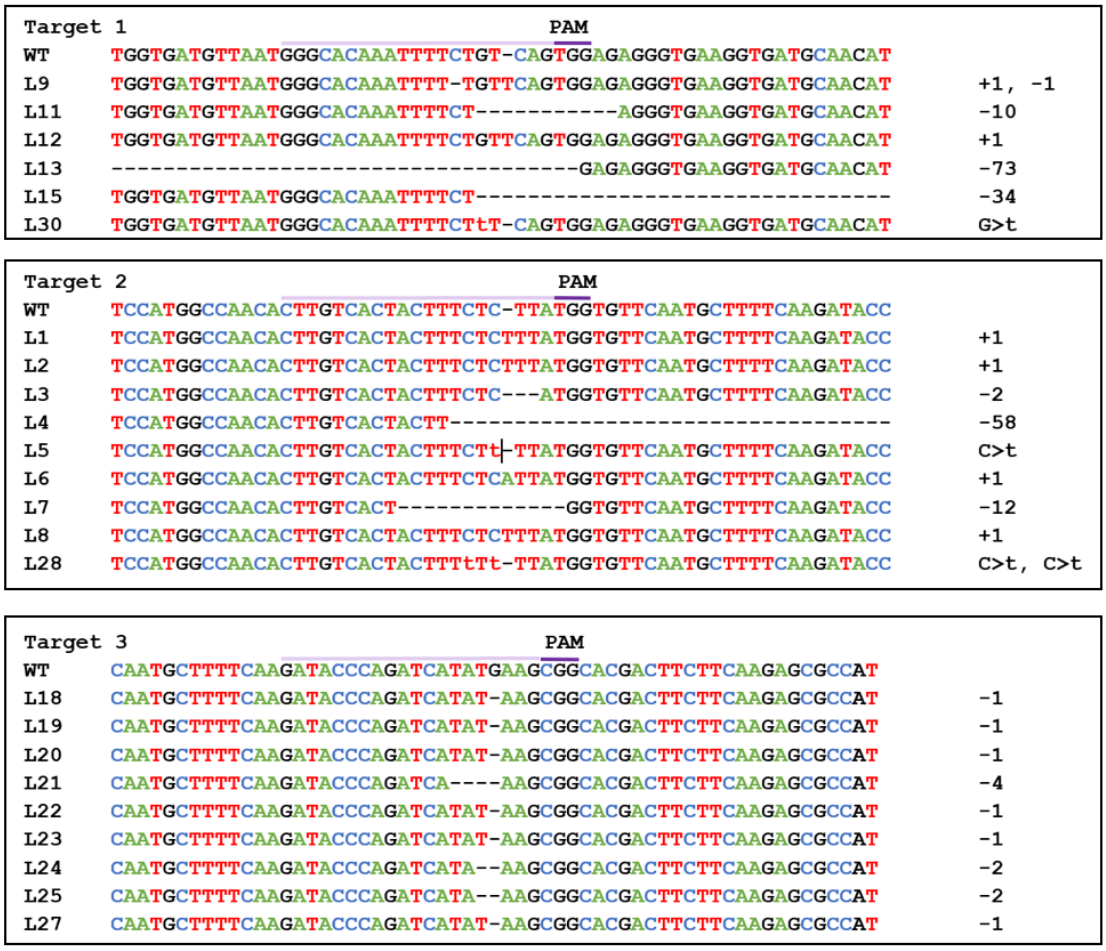
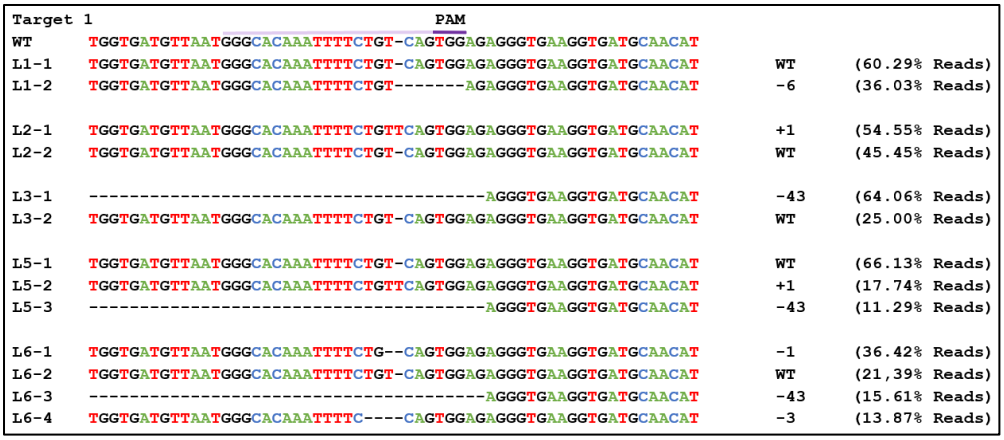


Figure 3.10 Nature of mutations obtained by sequencing the PCR products that were amplified from the genomic DNA from leaves of various silenced events. The target sequence in the wild-type (WT) *gfp* is shown with an overline, with PAM indicated with darker overline. L indicates the knockout event number for each target.

3.2.3.2. **Nature of mutations in four-copy *gfp* transgene event:**

For the 4-*gfp*-copies knockout events, next generation sequencing was used. Primers with tags were designed and the sequences were analyzed by CLC and CRISPResso. These programs allowed us to obtain reads for each sequence, reads below 10% were not taken into account. The nature of mutations found for each target site is shown in Figure 3.11.

Although it is not possible to know whether all four *gfp* copies of the gene are functional, if four different mutations are obtained in any of the complete knockout event, it would indicate that even for a tetraploid species, such as potato, it would be possible to knock out all four alleles of a target native gene. The sequencing results for the three target regions are shown in Figure 3.11. In case of Targets 1 and 2, sequences showing one or two types of mutations and wild-type sequences were observed. On the other hand, for Target 3, most of the events showed 1 – 4 different types of mutations, with the exception of event L4 that had one wild-type sequence. Interestingly for this target, three events showed only one type of mutation (same 1 bp deletion). Another interesting feature of the results shown in Figure 3.11 is that for all three targets, multiple events showed exactly the same type of large deletion. For example, for Target 1, events L3, L5 and L6 show the same 43 bp deletion, for Target 2, events L2, L3, L4 and L5 show the same 94 bp deletion, and for Target 3, events L5, L6, L7 and L9 show the same 78 bp deletion. The results showed clearly that it is possible to achieve mutations at four target sites in the potato genome, thus raising the expectation that all four alleles of a native target gene in the tetraploid potato can be knocked out using the CRISPR system.



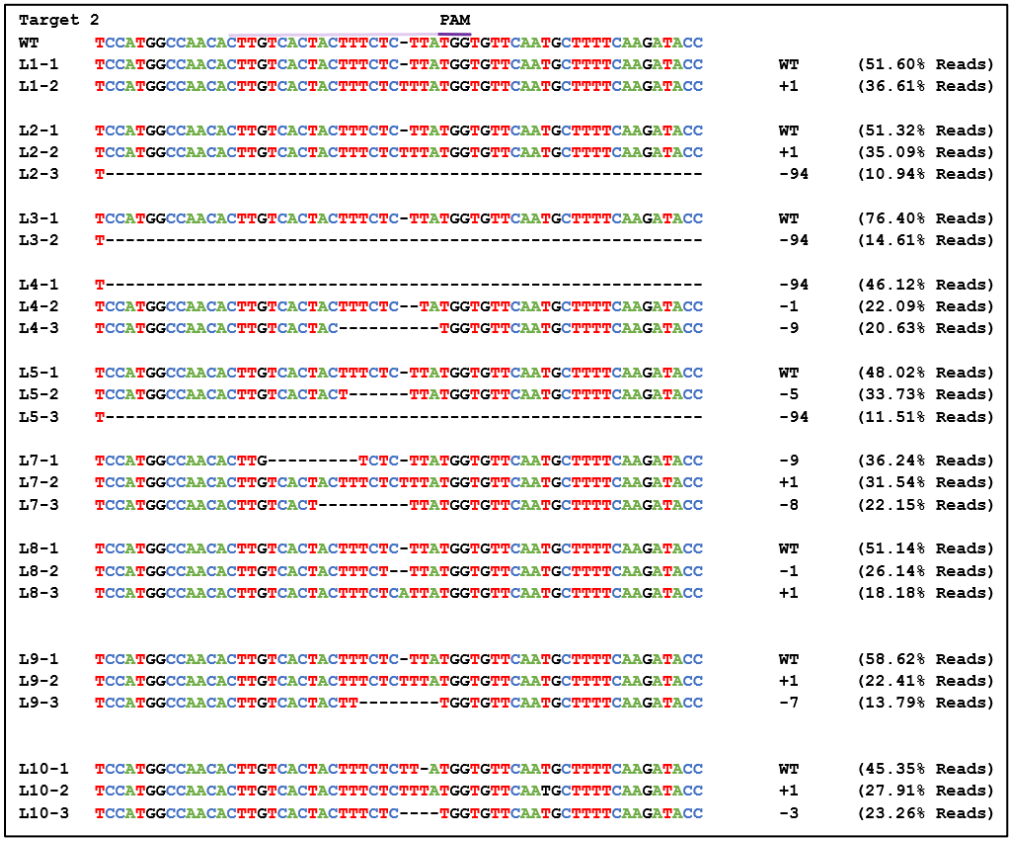


Figure 3.11 Nature of mutations obtained by the next generation sequencing of PCR products that were amplified from the genomic DNA of leaves from different events regenerated following transformation with CRISPR constructs. Overlined: target sequence in the wild-type *gfp* gene; Dark purple line: PAM sequence; WT: wild-type sequence; L: event number and the different sequenced/type of mutation found.

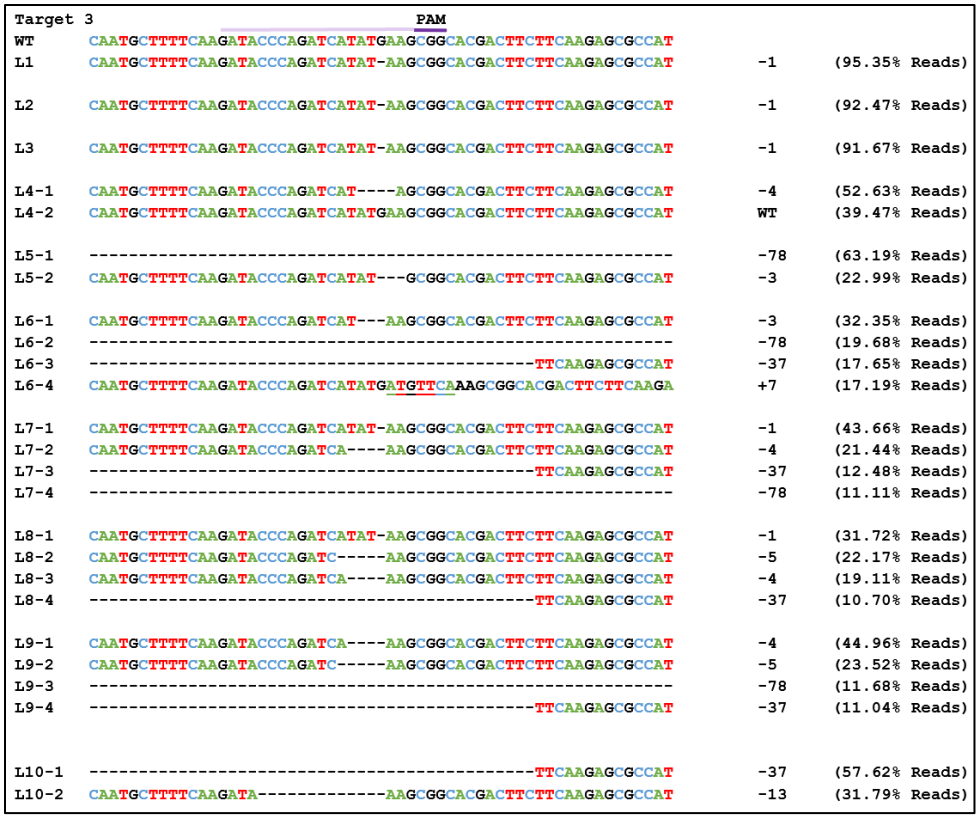


Figure 3.11 Continued Nature of mutations obtained by the next generation sequencing of PCR products that were amplified from the genomic DNA of leaves from different events regenerated following transformation with CRISPR constructs. Overlined: target sequence in the wild-type *gfp* gene; Dark purple overline: PAM sequence; WT: wild-type sequence; L: event number and the different sequence/type of mutation found.

3.3. Using the CRISPR-Cas9 system to knock out the native *gbss* gene to create amylose-free potato

3.3.1. Design of gRNAs to target the native, *gbss* gene in potato

In vitro-grown, wild-type Yukon Gold plants were used to knock out the *gbss* gene in this potato variety (four copies in its tetraploid genome). Results obtained from experiments conducted with four-copy *gfp* transgene plants had indicated that the

CRISPR-Cas9 system is capable of mutating all four copies of a gene in the potato genome. The *nptII* gene, that allowed us to perform transformant selection on kanamycin, was used to introduce the CRISPR-Cas9 system targeting the *gbss* genes in potato. This selection system was preferred compared to *hyg*/Hygromycin because of it facilitates higher transformant recovery. The gRNAs targeting the *gbss* gene were designed using the web-based tools CRISPR-P2 2.0., WU-CRSPR and sgRNA Scorer 2.0 (Table 3-3).

Table 3-3 Scores for each target site in each web-based tool used.

Target	gRNA	CRISPR-P	WU-CRSPR	sgRNA Scorer
Target 1	CATTCCCTTTCCACAAACAATGG	0.1827	60	-0.0459624
Target 2	AATCTTCCTGATGAATTCAGGGG	0.7243	75	-0.85140405

3.3.2. Qualitative analysis of knockout events

The microtubers growing on each individual event were examined via the histochemical Lugol-Iodine method. Examples of the phenotype are shown in Figure 3.12. In all, microtubers from 37 individual events, wherein Target site 1 was targeted for mutation, were evaluated. None of these showed the expected complete knockout phenotype (red-colored starch in the entire tuber) indicating that one or more alleles remained unmutated. However, one event (T1-L27) showed a few, small, red-colored patches randomly distributed throughout the mostly blue-colored tuber section (Figure 3.13). Three Target 1 events (T1-L1, T1-L27 and T1-L32) were selected for growth in soil and more extensive characterization.

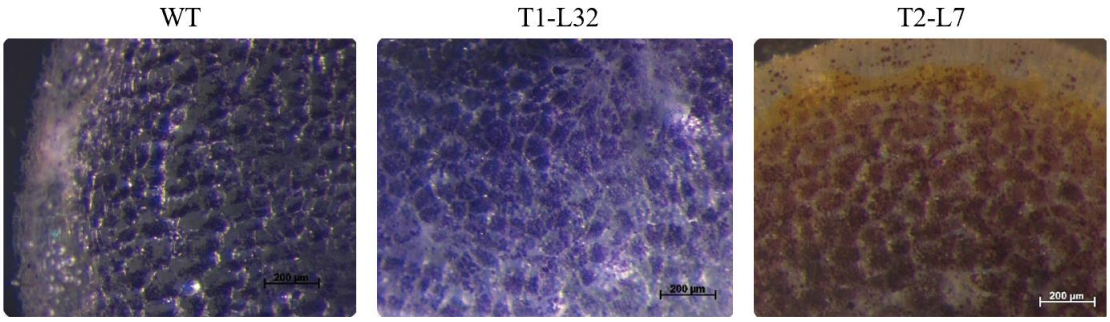


Figure 3.12 Examples of the different phenotypes detected via the histochemical Lugol-Iodine method in microtubers. Control: wild-type tuber; T1-L32: wherein Target site 1 was targeted for mutation, but does not show the complete knockout phenotype; T2-L7: wherein Target site 2 was targeted for mutation showing complete knockout phenotype (reddish-brown coloration).

In case where Target site 2 was the mutation target, microtubers from 32 individual events were examined for starch staining. Only one event (T2-L7) showed the red-colored phenotype throughout the tuber (Figure 3.12), thus indicating a complete knockout resulting from mutations in all four *gbss* alleles present in the potato genome and absence of amylose. Staining of tubers from soil-grown plant of this event confirmed these results (Figure 3.13). Tubers from two other events (T2-L2 and T2-L8) showed a few red-colored patches within a largely blue-colored tuber section (Figure 3.13). These three Target 2 events (T2-L2, T2-L7 and T2-L8) were selected for more extensive characterization.

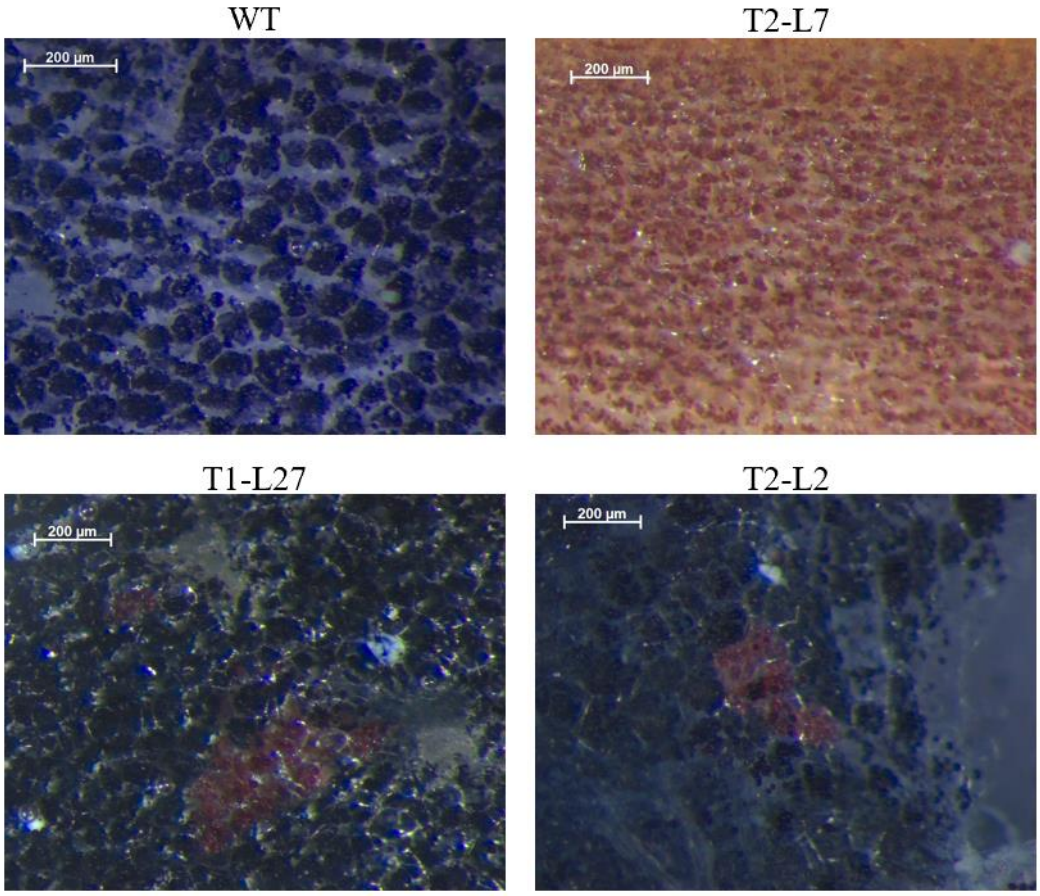


Figure 3.13 Events with different phenotypes detected via the histochemical Lugol-Iodine method in tubers collected from soil-grown plants. Control: wild-type tuber; T2-L7: wherein Target site 2 was targeted for mutation showing complete knockout phenotype (reddish-brown coloration). T1-L27: wherein Target site 1 was targeted for mutation showing a small sector of complete knockout phenotype surrounded by largely blue-colored tuber section. T2-L2: wherein Target site 2 was targeted for mutation showing both a small sector of complete knockout phenotype surrounded by largely blue-colored tuber section.

3.3.3. Molecular characterization of mutations

PCR amplicons from each *gbss* knockout event were cloned into linearized pMiniTTM vector using the PCR Cloning kit (New England BioLabs) followed by its transformation into DH5α *E. coli*. Plasmid DNA was extracted from twenty colonies for

each target and sequenced via Sanger sequencing. The sequences with "clean" chromatograms were aligned to the original gene sequence.

Indels were seen in each event, with most sequences showing deletions. However, two events had one bp insertion when Target site 1 was targeted for mutation (Events T1-L1 and T1-L32). In each of the three events, wherein Target site 1 was subjected to mutation, at least one allele was still intact. As reported earlier, even with one unmutated, functional allele of *gbss* gene, there is little reduction in the levels of amylose in potato tuber (Andersson *et al*., 2017). This is a likely explanation for the largely blue colored phenotype in these three events. (Figures 3.12, 3.13 and 3.14)

When Target site 2 was subjected to mutation, we observed a clear knockout phenotype only in event T2-L7. Three different types of mutations were observed in this event (Figure 3.14). The fact that this event shows amylose-free phenotype and only three types of mutations suggests that two alleles likely have homozygous mutations. For the other two events where Target site 2 was subjected to mutation, Event T2-L2 showed three types of mutations, whereas event T2-L8 showed only two types of mutations. While the sequencing results showing a lack of wild-type sequence would suggest that these events had one or two homozygous, biallelic mutations, these events did not show the phenotype of a of a complete knockout event. An explanation for this discrepancy is presented in Discussion.

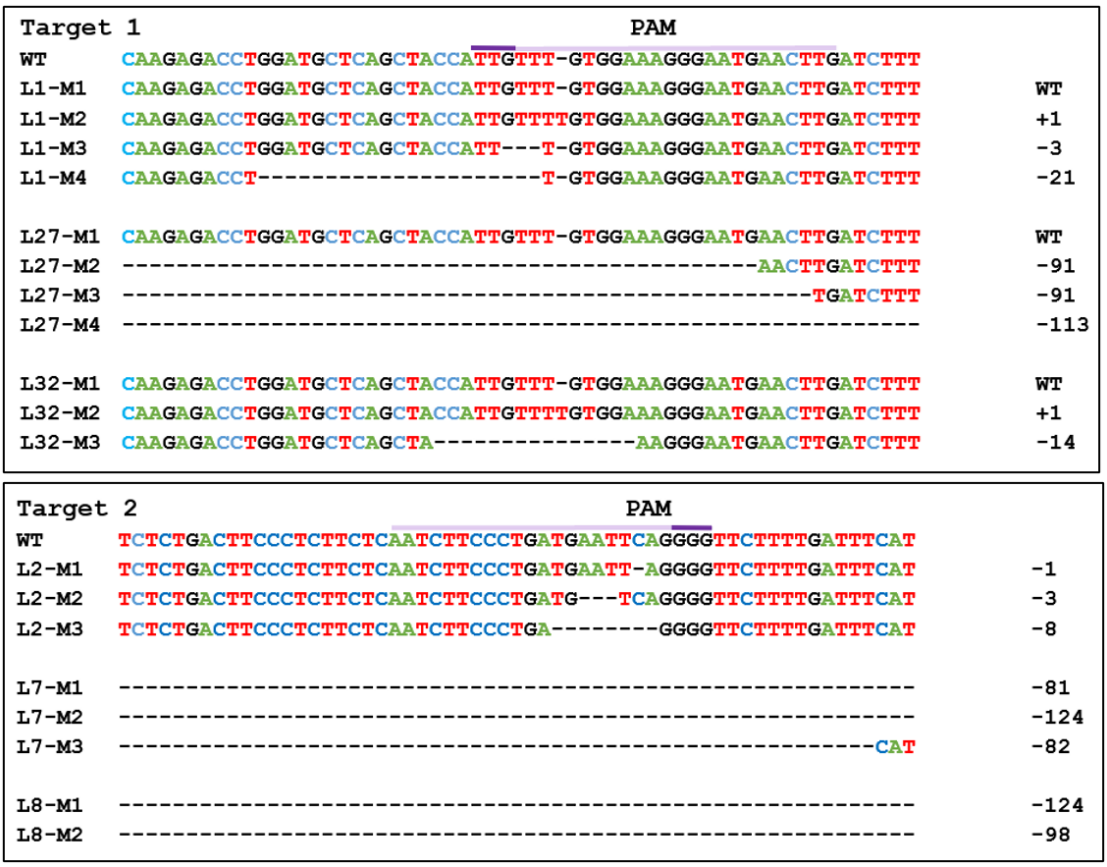


Figure 3.14 Nature of mutations obtained by Sanger sequencing of clones obtained from PCR products that were amplified from the genomic DNA of leaves from three events per target site that were characterized extensively. Overlined: target sequence in the native *gbss* gene; Dark, purple line: PAM sequence; WT: wild-type sequence; L: event number and the different sequences/type of mutations found.

3.3.4. Specific gravity measurements and quantitative analysis of starch composition in knockout events

The specific gravity of tubers from each event was calculated and the values are presented in Figure 3.15.

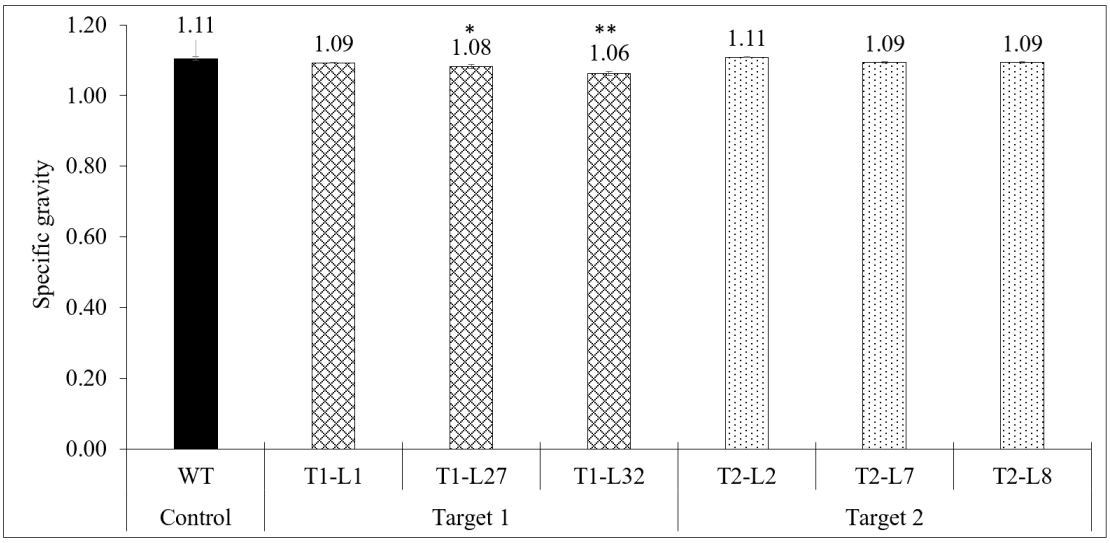


Figure 3.15 Specific gravity values in wild-type (WT), Yukon Gold and six different knockout events. Data represent mean ±SE, *P<0.05, **P<0.01; n=2-3.

Specific gravity measurements allowed us to compare each event with the unmodified control and thus determine if the genetic modification altered the dry matter content in the tubers. The complete knockout event (T2-L7) was not significantly different from the WT control, with specific gravity values of 1.09 and 1.11, respectively. However, tubers from the event T1-L27 with some reduction in amylose content and T1-L32 with no reduction in amylose content had specific gravity values of 1.08 and 1.06, respectively.

The percentage of dry matter was calculated for tubers from each event and the values are presented in Figure 3.16.

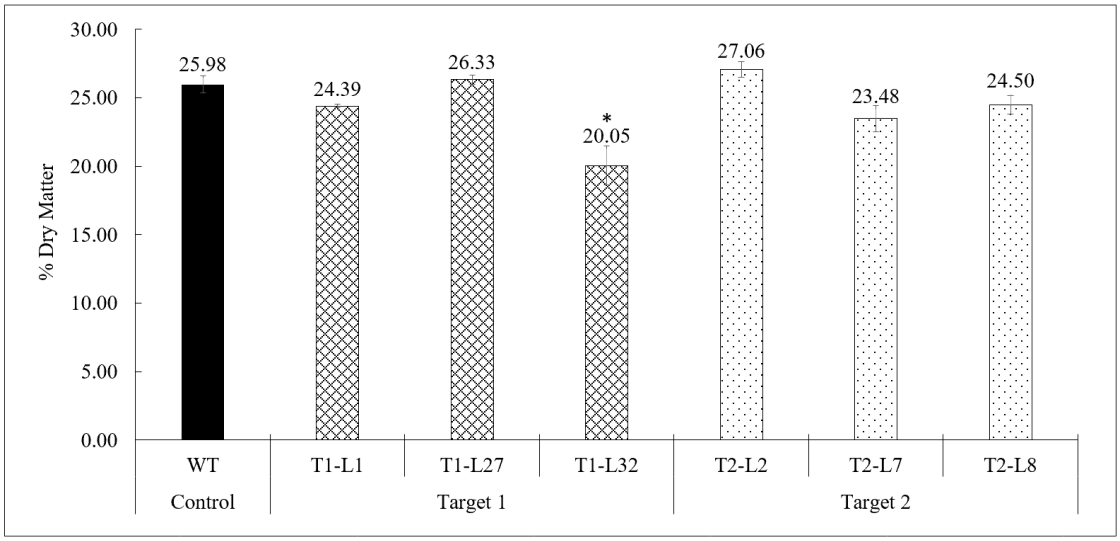


Figure 3.16 Dry matter values in wild-type (WT), Yukon Gold and six different knockout events. Data represent mean ±SE, *P<0.05; n=2-3.

A quantitative analysis of amylose content (%) was conducted in the tubers from wild-type and all six events (three per target site) with two different methodologies. This allowed us to correlate phenotype observed by histochemical assays and the mutations that were present in the *gbss* copies.

3.3.4.1. **Measurement of amylose content using the perchloric acid method**

This method allowed us to compare starch composition in the tubers of control (wild-type) with those from mutant events. The wild-type control tubers have an amylose content of 32.40%. Tubers from events T1-L1 and T1-L32 (Target site 1) and event T2-L8 (Target site 2) have percent amylose content similar to that of wild-type tubers. Event T2-L7, the complete knockout event showed absence of any amylose in the tuber. Event T1-L27 (Target site 1 and 22.2% amylose) and Event T2-L2 (Target site 2 and 13.93%

amylose) showed lower amylose content in the tuber although not enough to be considered complete knockout events. (Figure 3.17). As explained in the Discussion, the chimeric nature of these edited events can account for this anomaly.

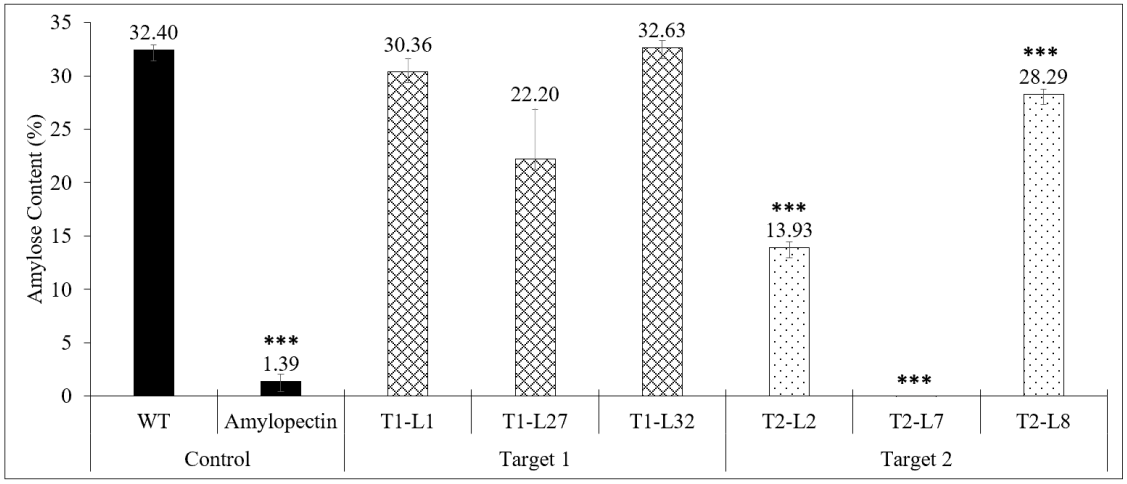


Figure 3.17 Colorimetric determination of amylose content using the perchloric acid method. Amylose values in wild-type (WT), Yukon Gold and six different knockout events. 100% amylopectin solution was also used as a control. Data represent mean ±SE, ***P<0.001; n=3-6.

3.3.4.2. **Measurement of amylose content using a commercial Amylose/Amylopectin kit (Megazyme)**

This method provided somewhat similar results as those using the perchloric acid method. With this method the wild-type control tubers gave an amylose content value of 22.09%. However, a solution of 100% amylopectin without any amylose gave a value of 7.18%. (Figure 3.18). Events T1-L1 and T1-L32 (Target site 1) and event T2-L8 (Target site 2) showed percent amylose values similar to that of control tubers at 21.51%, 23.46% and 21.44%, respectively. The knockout T2-L7 event showed percent amylose value of

4.41% in the tuber that is comparable to the solution that contained only amylopectin (7.18%). Event T1-L27 (Target site 1 and 14.33% amylose) and event T2-L2 (Target site 2 and 10.03% amylose) also showed lower percentage amylose content in the tubers, although not enough to be considered complete knockout events. Thus, while percent amylose content values are different with the commercial kit, the overall trend is similar to that obtained with the perchloric acid methodology. (Figure 3.18)

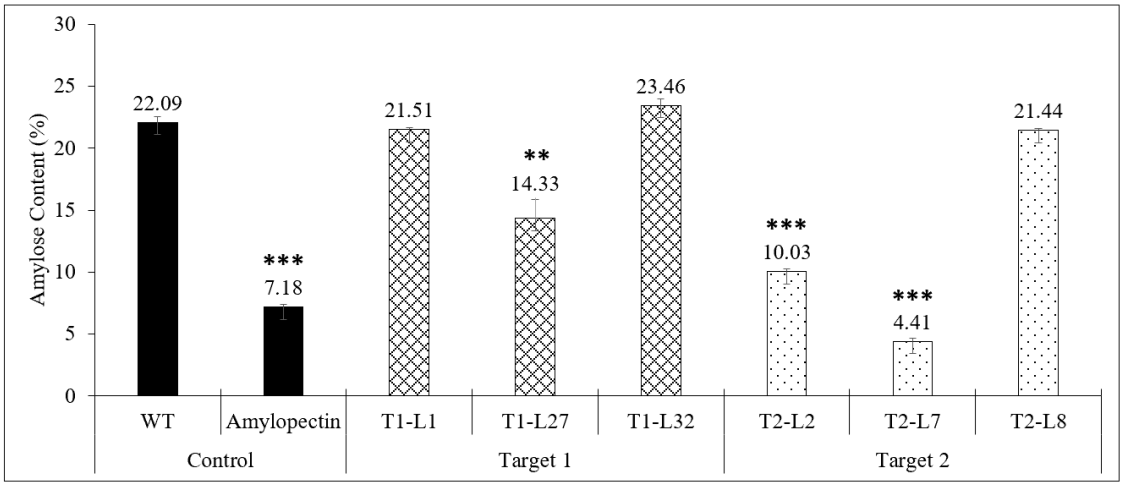


Figure 3.18 Colorimetric determination of amylose content using the Amylose/Amylopectin kit (Megazyme). Amylose values in wild-type (WT), Yukon Gold and six different knockout events. 100% amylopectin solution was also used as a control. Data represent mean ±SE, **P<0.01, ***P<0.001; n=3-6.

3.3.5. Viscosity

Viscosity was measured using a Rapid Visco Analyzer (RVA; PerkinElmer). The same three replicates for each target site were used for RVA analysis as used for earlier assays. The readouts from the RVA include viscosity profiles in relation to temperature

profile for each sample. The complete knockout event, T2-L7, showed a high final viscosity (581.00 cP) as compared to the control sample (297.33 cP). Viscosity values for starches with higher levels of amylopectin are expected to be greater and the RVA results on various samples from the transformed events bear this out (Figure 3.19). These results again confirm those obtained from histochemical and biochemical assays.

RVA also provides another set of data that relate to pasting properties of the starch samples as shown in Table 3-4. These properties depend on the amylose/amylopectin ratio in the starch sample and informs the best use for a given type of starch. The pasting property values for event T2-L7 sample are particularly apparent compared to rest of the samples. The Breakdown value, that indicates paste stability, for event T2-L7 (153.00) is significantly higher compared to the control value (53.00) (Kumar & Khatkar, 2017).

The setback viscosity is a measure of the gelling ability or retrogradation ability of starches. (Kumar & Khatkar, 2017) and indicates recrystallization of amylose molecules in the gel. This value for event T2-L7 is 8 while the control is at 33.33 and rest of the edited events show values higher than 30.00, again reflecting the lack of amylose in event T2-L7. The pasting temperatures for events T2-L2 and T2-L7 are slightly higher than the WT (64.41), being 69.40 and 68.57 respectively.

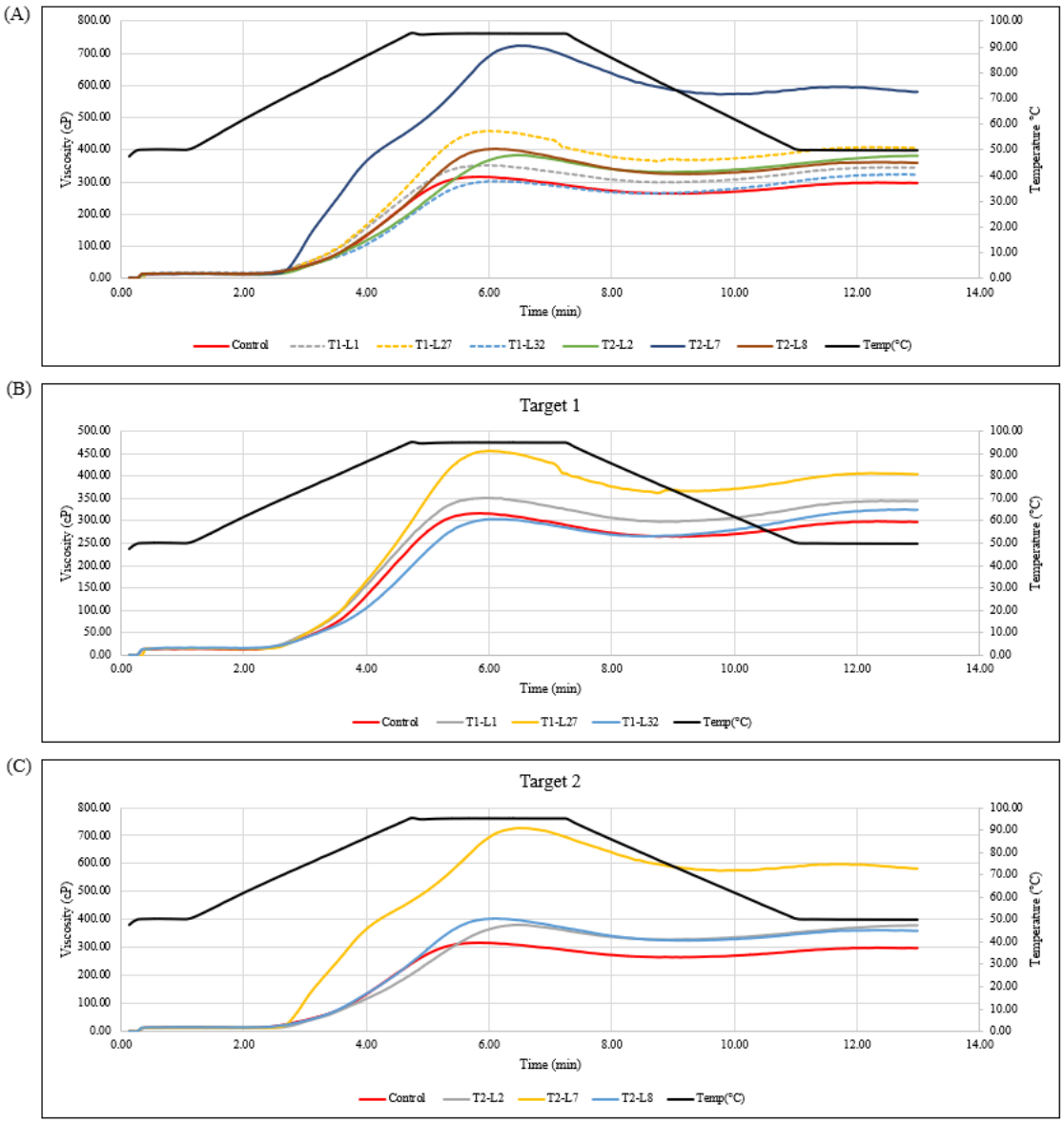


Figure 3.19 Viscosity profiles and temperature profile as analyzed by Rapid Visco Analyzer (RVA). (A) RVA-profiles for control (unedited, wild-type) and edited events for Target 1 and Target 2. (B) RVA-profiles for control (unedited, wild-type) and edited events for Target 1. (C) RVA-profiles for control (unedited, wild-type) and edited events for Target 2.

Table 3-4 Pasting properties of samples from: Control (wild-type) and transformed events.

Sample	Peak Viscosity (cP)	Trough Viscosity (cP)	Breakdown (cP)	Final Viscosity (cP)	Setback (cP)	Peak Time (min)	Pasting Temperature (°C)
Control	317.00	264.00	53.00	297.33	33.33	5.78	64.41
T1-L1	352.50	297.50	55.00	343.50	46.00	5.93	64.41
T1-L27	457.33	359.00	98.33	405.00	46.00	5.96	66.93
T1-L32	307.67	265.00	42.67	324.67	59.67	6.18	66.05
T1-L2	386.00	329.33	56.67	380.67	51.33	6.42	69.40
T2-L7	726.00	573.00	153.00	581.00	8.00	6.49	68.57
T2-L8	402.67	323.67	79.00	358.33	34.67	6.09	66.05

4. DISCUSSION

The CRISPR-Cas9 system is one of the mechanisms that bacteria use to protect themselves from viral attacks; it serves as an adaptative immune mechanism that is capable of generating memory of the past and reacting to the invading genomes (Dy *et al*., 2014, Seed, 2015). This bacterial defense system has been harnessed to introduce mutations at predetermined genome targets and generate desired genetic alterations in Eukaryotes. Early applications in plants have shown the CRISPR system to be a simple, efficient and versatile tool for genome editing. It is not only cost-effective, but also functions successfully in a variety of plant species, including model species as well as the economically important crops (Montecillo *et al*., 2020, Cong *et al*., 2013, Jiang & Doudna, 2017).

The CRISPR-Cas9 system generates indels by causing DSB in the target sequence. The system utilizes Cas9 as the endonuclease and sgRNA containing ~20 nucleotide-long sequence that matches the target DNA site as a guide. The DSB at the target site triggers a repair mechanism in the cell, the most common one being the non-homologous end-joining (NHEJ). Nucleotide insertions or deletions (indels) are generated at the break site causing a frame shift in the gene sequence, thus altering the amino acid sequence and loss of gene function (knockout). The other DNA repair mechanism, called homology-dependent repair (HDR) or homologous recombination (HR), is rare and requires a DNA template. This is used to create more precise, desired mutations such as nucleotide

substitution or insertion of a predetermined DNA sequence (Janga *et al*., 2017, Voytas & Gao, 2014, Puchta, 2005, Jiang & Doudna, 2017).

Until recently, introduction of novel traits in crop plants has relied on breeding methods that allow us to bring in the desired genes from wild relatives, mutants (created by chemical or radiation mutagenesis) or transgenic events. Traditional breeding techniques can take several years or decades to introduce a stable and desirable genetic alteration in many crop plants. CRISPR-Cas9 system allows us to generate mutations in a more precise, rapid and efficient manner and can alter all the alleles in the genome simultaneously. Such gene alterations have proved to be stable and heritable and the system also allows us to perform multiplex genome editing. Importantly, in many countries, gene edited traits require lesser regulatory scrutiny compared to traditional transgenic traits. One of the requirements for introducing mutations through the CRISPR system is prior knowledge of the target gene sequence. The advantages described above have made the CRISPR system a more attractive tool compared to other gene editing technologies to introduce novel traits into commercial crops (Chen *et al*., 2019, Voytas & Gao, 2014, Montecillo *et al*., 2020).

Potato, being an important food crop, makes a good target for introducing new, desirable traits. As stated earlier, the development of new commercial varieties of potatoes by plant breeding is infrequent. In general, the plant breeding process is slow for most crops; it is even slower for potato, partly because it is a vegetatively propagated crop. The reason for this practice is that potato is an autotetraploid crop that is highly heterozygous and often gametophytic self-incompatible. Even in the case where plants are self-

compatible, the progeny suffers from inbreeding depression. Breeders have to be mindful that if commercial varieties are crossed with wild genotypes, there is a high possibility that the next generation might lose the established traits and express unwanted traits obtained from the wild relatives (Stokstad, 2019, Watanabe, 2015).

Potato is a relatively easy species to transform, being one of the first crops to be successfully transformed by *Agrobacterium* method as demonstrated by a study that utilized *A. rhizogenes* to determine the transformation efficiency (Ooms *et al*., 1986). As is the case with most plants, the transformation efficiency of potato depends on the variety and the type of explant used. In the current study, the explant used for transformation were internodes. As shown by Bruce and Shoup Rupp (2019), this explant has shown the greatest transformation efficiency. During the establishment of potato transformation protocol in the current study, leaves were also evaluated in addition to the internodes. However, following cocultivation of leaf segments with *Agrobacterium*, the tissue showed severe damage and no callus growth. All three varieties that were examined using the protocol described in Materials and Methods proved amenable to *Agrobacterium*-mediated transformation. However, only one variety (Yukon Gold) was able to go through various stages of regeneration to yield transgenic plants. Our results support general observations by others in that *Agrobacterium*-mediated transformation and regeneration depend on the plant genotype used.

In this study, *Agrobacterium*-mediated transformation was used to introduce the CRISPR reagents that created mutations in a non-native, integrated transgene and a native gene in potato (*Solanum tuberosum*).

To knock out a non-native gene, previously transformed, *gfp*-expressing plants were used. The objective was to target the *gfp* gene in two different events: one with a single copy of the *gfp* and another one with four copies of the *gfp* gene. In order to test the efficiency of the CRISPR-Cas9 system in potato, the same three sgRNAs used by Janga *et al*. (2017) were deployed to target the *gfp* gene in this crop. The knockout efficiency of the system depends on multiple factors related to the sgRNA, for example, percentage of GC content and the secondary structures. Expression levels of Cas9 and sgRNA also have an effect on the efficiency of the system (Baysal *et al*., 2016, Liang *et al*., 2016, Ma *et al*., 2015).

In this study to examine the efficiency of each sgRNA, gRNA3 proved to be the most efficient (59.69% - single *gfp* copy event and 53.23% - four *gfp* copies event), followed by gRNA 2 (22.96% - single *gfp* copy event and 26.25% - four *gfp* copies event) while gRNA1 was the least efficient (11.73% - single *gfp* copy event and 23.13% - four *gfp* copies event). These results for sgRNA1 and sgRNA2 efficiencies, while not consistent with the predicted scores, are in line with those observed in cotton by Janga *et al*. (2017). In the cotton study, sgRNA1, sgRNA2 and sgRNA3 sequences were based on scores obtained from sgRNA Scorer 1.0. The observed mutation efficiency for sgRNA 3 (predicted score: 91.2) was highest in the cotton study and thus in compliance with the predicted score. However, the expected and observed knockout efficiencies showed discrepancies for the other two sgRNAs. sgRNA 1, with medium score (predicted score: 51.2), proved to be the least efficient while the sgRNA2 with the lowest score (predicted

score: 2.2) showed medium efficiency (Janga *et al*., 2017). Such discrepancy prompted these investigators to examine the prediction score for *gfp* using another web tool, WU-CRISPR (Wong *et al*., 2015). This program considers various characteristics of sgRNA such as the GC content, secondary structure, contiguous stretch of the same nucleotides, and free accessibility of the seed region for target recognition (i.e., avoiding the use of U at position 19 and use of C or U at position 20 in the guide sequence). Thirteen different possible targets were predicted by this tool. Only the sgRNA3, but not sgRNA1 and sgRNA2, was found to be among these 13 predicted targets (Janga *et al*., 2017). Such discrepancies between the predicted and observed sgRNA scores have also been reported by Baysal *et al*. (2016) in rice and Wang *et al*. (2016) in wheat. Our results, taken together with those from these previous studies suggest that the efficiency of the CRISPR-Cas9 system depends on sequence of the gRNA and not on the plant species in which it is being used. Also, it is better to use more than one prediction program to design the sgRNAs to select the most optimum one.

Sequence analyses were performed to determine the nature of mutations present in the *gfp* knockout and *gbss* knockout events. Individual target sites were PCR-amplified and sequenced to determine variations compared to the wild-type sequence.

In the *gfp* knockout experiment, wherein the one-*gfp* copy event was targeted by the CRISPR, Sanger sequencing of PCR amplicons showed mutations (indels and substitutions) in all the regenerated events. In case of four-*gfp* copy events, next generation sequencing was used to identify the mutations. Among these, several events showed 1 - 3 types of mutations along with unmutated (wild-type) copy. A few events had 2 – 4 copies

with mutations without any unmutated (wild-type) copies. Great care was taken to select shoots that showed complete silencing of the *gfp* gene. This fact taken together with the presence of a wild-type sequence suggests that not all four copies of the *gfp* gene were functional. However, there were some events that showed four different types of mutations indicating that it was possible to target four different copies of a gene with the CRISPR system in potato genome. Some events had 2 - 3 mutated copies but absence of a wild-type copy. This suggests the possibility of homozygous mutations in these events as previously observed in cotton by Janga *et al*. (2017), in rice by Ma *et al*. (2015), and in tomato by Pan *et al*. (2016).

The regeneration method used also determines the presence of chimeric tissue in the recovered plant. If the mode of regeneration is via somatic embryogenesis, typically the embryo originates from a single cell thus avoiding the problem of producing a chimeric plant, cotton being a good example (Rathore *et al*., 2015). It is important to note that in cotton, somatic embryos arise after a long culture period in the form of undifferentiated callus, typically taking six to ten months (Janga *et al*., 2017). This allows Cas 9 enough time to mutate all the target alleles in a given callus cell. Thus, the embryo originating from such a cell is likely to have all intended targets mutated in all the cells. Another mode of regeneration is via shoot organogenesis that involves shoot formation from multiple callus cells, as is the case with potato. In this case, the new plantlet will likely originate from both mutated and non-mutated cells. Even the various mutated cells may differ in the type of mutations they carry (Zhang *et al*., 2020). This offers a likely

explanation for the discrepancy observed between the molecular results and the phenotype in some of the *gbss* knockout events (described below).

In case of the *gbss* gene, four different types of variations would indicate different mutations in all four alleles, while a smaller number of mutations would indicate the presence of homozygous mutations or lack of mutation in one or more alleles. Three Events (T1-L1, T1-L27 and T1-L32), wherein Target 1 was subjected to mutation, show at least one wild-type copy of the *gbss* gene. As shown by Andersson *et al.* (2017), even with a single functional copy of the *gbss* gene, the tuber produces a significant amount of amylose, thus retaining the blue color of the starch with the Lugol-iodine staining. In agreement with this study, tubers from T1-L1 and T1-L32 showed blue colored staining. A tuber from Event T1-L27 stained largely blue, but showed some red-stained patches indicating that the fourth *gbss* allele had undergone mutation in these tuber cells (Figure 3.13). As mentioned earlier, the CRISPR components integrated in the genome of a plant remain functional and it is possible that these red-stained cells arose as result of late mutations in the tuber cells. When target 2 was subjected to mutation, 2 – 3 types of mutations were observed in three different events. None of these events had a wild-type, unmutated copy of the *gbss* gene. Only event T2-L7 showed complete lack of amylose and the red-stained starch phenotype. Thus, this event shows an exact match between the genotype and phenotype. Absence of a wild-type copy and the fact that only three types of mutations are seen in this event, suggest that it contains one biallelic, homozygous mutation. Tubers from event T2-L8 showed blue-colored staining throughout, while T2-L2 tubers were stained largely blue with a few red-colored patches (Figure 3.13). As

mentioned earlier, the chimeric nature of the potato plants that regenerates via shoot organogenesis may explain these confounding observations. No wild-type copy of the *gbss* gene was observed in these two events, suggesting that all four alleles in their leaf cells were mutated. But the tubers that arose from the base of the stem may still carry one unmutated copy of the *gbss* gene, responsible for blue staining in the tuber. The chimeric nature of the potato plant and continued activity of the Cas9 to target the unmutated copy is evident from the tuber staining results for event T2-L2 that shows a few red-stained patches surrounded by largely blue-stained starch. Such persistent activity of the genome-integrated CRISPR components resulting in a chimeric plant (in terms of different types of mutations in different sectors of the plant) has been shown in cotton (Janga *et al*., 2019), *Medicago truncatula* (Zhang *et al*., 2020) and *Arabidopsis* (Ma *et al*., 2015).

The levels of amylose in the mutated events were also measured by a perchloric acid method and an enzymatic method (Megazyme). This allowed us to correlate percentage of amylose present in the tuber with the histochemical data. Events that showed a blue-colored phenotype had a percentage of amylose similar to the control (32.40%), the event that showed a knockout phenotype showed complete absence of amylose in the tuber (0%). Interestingly the two events (T1-L27 and T2-L2) that had some red-colored patches surrounded by largely blue-colored staining in the tuber sections had lower percentages of amylose 22.20% and 13.93% respectively. This could be attributed to the presence of a small proportion of tuber cells that had all four alleles of the *gbss* gene knocked out.

Specific gravity values represent indirect measure of the dry matter, i.e., it indicates the level of solids compared to the amount of water present in the tubers, with higher values being more desirable for cooking purposes (Yuan *et al*., 2019, Taylor, 2019). Events T1-L27 and T1-L32 showed significantly lower specific gravity values. However, no clear correlation was observed between amylose reduction and specific gravity as indicated by the fact that the tubers from complete knockout event, T2-L7, with no measurable amylose had specific gravity value similar to the tubers from the non-edited plants. These results suggest that the relative levels of amylose and amylopectin do not affect the moisture content of the tuber.

Viscosity profiles were determined with the aid of RVA in the *gbss* knockout events. Disruption of all four copies of the *gbss* gene is expected to eliminate the amylose component from the starch. Such an alteration confers different characteristics to the starch, including swelling and pasting properties, that will determine its use in the industry (Juhász & Salgó, 2008).

The decrease or elimination of amylose content is expected to increase the peak viscosity (PV). This can be seen clearly in event T2-L7 that has PV of 726 cP compared to WT with a PV of 317cP. Such differences based on starch composition have been reported by other investigators. In cassava, Zhao *et al*. (2011) showed that wild-type control had a PV of 885 cP, while the waxy transgenic event B9 had a value of 1094 cP. In another study comparing mutant waxy cassava with the wild-type, Toae *et al*. (2019) observed the same trend. Similarly, in cereals such as maize, Sánchez *et al*. (2010) found

that the wild-type control had a PV of 176 cP, while the waxy variety had a PV value of 973 cP.

The setback value largely depends on the level of amylose present in the starch. This value relates to the cooling stage (or setback) during which retrogradation or realignment of amylose and linear parts of amylopectin occurs. The setback value is greater with higher levels of amylose and vice versa. In the current study, this value for the WT was 33.33, while it was 8.00 for the complete knockout event T2-L7. Such a correlation was also observed in a rice study, where waxy varieties were measured (Allahgholipour *et al*., 2006).

The breakdown value of the starch is higher with a reduction in the amylose content. This value in the WT was 53 cP, while it was 98.33 cP for event T1-L27 and 153 cP for the complete knockout event T2-L7. Such a correlation was also observed in cassava [(Zhao *et al*., 2011) and (Toae *et al*., 2019)] and in potato (Sánchez *et al*., 2010) where waxy phenotype showed a higher breakdown value.

In the current study, pasting temperature values did not differ among the mutated events and the WT. This value for WT was 64.41 °C and for the complete knockout event T2-L7 it was 68.57°C. In a study by Sánchez *et al*. (2010) that compared a waxy potato with WT potato, the pasting temperatures were 65.9°C and 65.2°C, respectively. The cassava investigation by Toae *et al*. (2019) also showed a lack of difference in the pasting temperatures among different mutant events and the wild-type.

In the study conducted by Andersson *et al*. (2017), 25% of the regenerated plants were stunted and eventually died. They attributed this to the accumulation of somaclonal variations during the long culture period needed to regenerate plants following protoplast transformation. Other weaknesses of the protoplast transformation method are the requirement for expensive enzymes and complexity of the entire isolation and regeneration procedures. Thus, the edited plants derived from this method are more likely to contain somaclonal variations due to much longer culture period. Such genetic variations are common when using protoplasts for potato transformation (Shepard *et al*., 1980).

Even though the protoplast transformation system has these weaknesses, there are some advantages to using this technique. The main benefit is that it allows direct delivery of Cas9 enzyme/sgRNA (ribonucleoproteins, RNP). It is also possible to deliver RNP via gene gun or biolistic bombardment. This involves the use of gold or tungsten particles coated with Cas9/sgRNA RNPs for delivery to cells (Sandhya *et al*., 2020). The major strength of these alternative methods is the possible avoidance of transgene integration in the cellular genome, which allows the edited plants to be labeled as "GMO free". The *Agrobacterium*-mediated editing, in general, involves stable integration of the DNA sequences encoding Cas9 and sgRNA and thus can be labeled as GMOs. One disadvantage of the *Agrobacterium* and gene gun method is that the selection process is more difficult, thus lowering the editing efficiency (Sandhya *et al*., 2020). In comparison to the protoplast and gene gun, the *Agrobacterium* system is less expensive and can be performed without highly specialized equipment and chemicals.

In this study I have demonstrated the ability of the CRISPR-Cas9 system to generate mutations within the genome of a commercial variety of potato, Yukon Gold. A reliable and efficient transformation protocol for potato was first established by using the *gfp* gene. In the first phase of this work to evaluate the CRISPR system, the results showed the possibility of knocking out not only a single copy of the *gfp* transgene, but also four copies of this gene. The selection of sgRNA is an important consideration for an efficient CRISPR-Cas9 system. The results obtained in the current study suggest the importance of utilizing more than one gRNA prediction programs to find the best possible option. In the second phase of this work, the native, *gbss* gene was targeted for mutations in order to eliminate amylose in the starch of the tubers. Sequence analyses of the mutations showed insertions, deletions, single nucleotide alterations, homozygous mutations, biallelic mutations and also non-mutated copies. Shoot organogenesis, as the mode of regeneration in potato, likely results in a chimeric T_0 plant. This fact became evident from the results (both histochemical assays and starch composition analyses) obtained for the *gbss* knockout events (T1-L27 and T2-L2). A single complete knockout event (T2-L7) obtained in this investigation shows that while shoot organogenesis mode of regeneration complicates the use of the CRISPR system in potato, it is possible to obtain a complete knockout of a target gene in this tetraploid species. The amylose-free knockout event, T2-L7, should find industrial applications in the future.

5. CONCLUSIONS

1. *Agrobacterium*-mediated transformation of potato is efficient; however, the regeneration process is highly genotype dependent. Out of the three varieties examined, Russet Norkotah, White LaSoda and Texas Yukon Gold, it was possible to recover transgenic plants from Texas Yukon Gold only.
2. The green fluorescent protein gene (*gfp*) expresses well in potato and is a useful reporter gene for this crop.
3. CRISPR-Cas9 reagents introduced via *Agrobacterium*-mediated transformation functioned efficiently in potato. It was possible to knock out as many as four copies of the *gfp* transgene in the potato genome using this system.
4. Mutations in the *gfp* gene included insertions, deletions, single nucleotide alterations and homozygous mutations; however, non-mutated copies were also present in some events.
5. It was possible to knock out the native, *gbss*, gene with the CRISPR-Cas9 system to generate amylose-free (waxy) potato events.
6. It was possible to observe alterations related to starch composition in *in vitro* generated microtubers before the edited events are transferred to soil and grown to obtain normal-sized tubers.
7. Mutations in the *gbss* gene included insertions, deletions, single nucleotide alterations, homozygous mutations, biallelic mutations, however, non-mutated copies were also present in some events.

8. Shoot organogenesis mode of regeneration in potato results in chimeric plants, both with regard to the type as well as the extent of mutations.
9. Two important considerations to take into account for gene editing in potato are: 1. tetraploid genomes of most commercial varieties, and 2. the chimeric nature of the edited events.
10. While it is possible to obtain complete knockout of all the alleles of a native target gene, screening of several tens of edited events may be necessary, especially for a trait that is not easy to identify during the early stages of regeneration.

REFERENCES

Allahgholipour M, Ali AJ, Alinia F, Nagamine T, Kojima Y, 2006. Relationship between rice grain amylose and pasting properties for breeding better quality rice varieties. *Plant Breeding* **125**, 6.

Andersson M, Turesson H, Nicolia A, Fält A-S, Samuelsson M, Hofvander P, 2017. Efficient targeted multiallelic mutagenesis in tetraploid potato (Solanum tuberosum) by transient CRISPR-Cas9 expression in protoplasts. *Plant Cell Reports* **36**, 117-28.

Balet S, Guelpa A, Fox G, Manley M, 2019. Rapid Visco Analyser (RVA) as a Tool for Measuring Starch-Related Physiochemical Properties in Cereals: a Review. *Food Analytical Methods* **12**, 2344-60.

Baltes NJ, Gil-Humanes J, Cermak T, Atkins PA, Voytas DF, 2014. DNA Replicons for Plant Genome Engineering. *The Plant cell* **26**, 151.

Baysal C, Bortesi L, Zhu C, Farré G, Schillberg S, Christou P, 2016. CRISPR/Cas9 activity in the rice OsBEIIb gene does not induce off-target effects in the closely related paralog OsBEIIa. *Molecular Breeding* **36**, 108.

Beckles DM, 2010. Use of Biotechnology to Engineer Starch in Cereals. In. *Encyclopedia of Biotechnology in Agriculture and Food.* .

Bruce MA, Shoup Rupp JL, 2019. Agrobacterium-Mediated Transformation of Solanum tuberosum L., Potato. In: Kumar S, Barone P, Smith M, eds. *Transgenic Plants: Methods and Protocols.* New York, NY: Springer New York, 203-23.

Buchanan BB, Gruissem W, Jones RL, 2015. *Biochemistry and Molecular Biology of Plants*. Wiley.

Cermak T, Starker CG, Voytas DF, 2015. Efficient Design and Assembly of Custom TALENs Using the Golden Gate Platform. In. *Chromosomal Mutagenesis. Methods in Molecular Biology (Methods and Protocols).* Humana Press. (1239.)

Chen K, Wang Y, Zhang R, Zhang H, Gao C, 2019. CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture. *Annual Review of Plant Biology* **70**, 667-97.

Chetty VJ, Narváez-Vásquez J, Orozco-Cárdenas ML, 2015. Potato (Solanum tuberosum L.). In: Wang K, ed. *Agrobacterium Protocols: Volume 2*. New York, NY: Springer New York, 85-96.

Cip. Potato facts and figures. In.: International Potato Center

Cong L, Ran FA, Cox D, *et al.*, 2013. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**, 819-23.

Denyer KaY, Johnson P, Zeeman S, Smith AM, 2001. The control of amylose synthesis. *Journal of Plant Physiology* **158**, 479-87.

Dy RL, Richter C, Salmond GPC, Fineran PC, 2014. Remarkable Mechanisms in Microbes to Resist Phage Infections. *Annual Review of Virology* **1**, 307-31.

Enciso-Rodriguez F, Manrique-Carpintero NC, Nadakuduti SS, Buell CR, Zarka D, Douches D, 2019. Overcoming Self-Incompatibility in Diploid Potato Using CRISPR-Cas9. *Frontiers in Plant Science* **10**.

Fajardo D, Jayanty SS, Jansky SH, 2013. Rapid high throughput amylose determination in freeze dried potato tuber samples. *Journal of visualized experiments : JoVE*, 50407.

Fao, 2008. International Year of the Potato - Uses of Potato. In.

Fredriksson H, Silverio J, Andersson R, Eliasson AC, Åman P, 1998. The influence of amylose and amylopectin characteristics on gelatinization and retrogradation properties of different starches. *Carbohydrate Polymers* **35**, 119-34.

González MN, Massa GA, Andersson M, *et al.*, 2020. Reduced Enzymatic Browning in Potato Tubers by Specific Editing of a Polyphenol Oxidase Gene via Ribonucleoprotein Complexes Delivery of the CRISPR/Cas9 System. *Frontiers in Plant Science* **10**.

Grommers HE, Van Der Krogt DA, 2009. Chapter 11 - Potato Starch: Production, Modifications and Uses. In: Bemiller J, Whistler R, eds. *Starch (Third Edition)*. San Diego: Academic Press, 511-39.

Janga MR, Campbell LM, Rathore KS, 2017. CRISPR/Cas9-mediated targeted mutagenesis in upland cotton (Gossypium hirsutum L.). *Plant Molecular Biology* **94**, 349-60.

Janga MR, Pandeya D, Campbell LM, *et al.*, 2019. Genes regulating gland development in the cotton plant. *Plant biotechnology journal* **17**, 1142-53.

Jiang F, Doudna JA, 2017. CRISPR–Cas9 Structures and Mechanisms. *Annual Review of Biophysics* **46**, 505-29.

Jobling S, 2004. Improving starch for food and industrial applications. *Current Opinion in Plant Biology* **7**, 210-8.

Juhász R, Salgó A, 2008. Pasting Behavior of Amylose, Amylopectin and Their Mixtures as Determined by RVA Curves and First Derivatives. *Starch - Stärke* **60**, 70-8.

Karlsson ME, Leeman AM, Björck IME, Eliasson A-C, 2007. Some physical and nutritional characteristics of genetically modified potatoes varying in amylose/amylopectin ratios. *Food Chemistry* **100**, 136-46.

Kathleen GH, Beverly AC, David R, Bryan TV, 2011. Inheritance of Carotenoid Content in Tetraploid × Diploid Potato Crosses. *Journal of the American Society for Horticultural Science J. Amer. Soc. Hort. Sci.* **136**, 265-72.

Kumar R, Khatkar BS, 2017. Thermal, pasting and morphological properties of starch granules of wheat (Triticum aestivum L.) varieties. *Journal of food science and technology* **54**, 2403-10.

Lemos PVF, Barbosa LS, Ramos IG, Coelho RE, Druzian JI, 2019. Characterization of amylose and amylopectin fractions separated from potato, banana, corn, and cassava starches. *International Journal of Biological Macromolecules* **132**, 32-42.

Liang G, Zhang H, Lou D, Yu D, 2016. Selection of highly efficient sgRNAs for CRISPR/Cas9-based plant genome editing. *Scientific Reports* **6**, 21451.

Linsmaier E, Skoog F, 1965. Organic growth factor requirements of tobacco tissue culture. *Physiologia Plantarium* **18**, 100–27.

Liu S, 2017. Chapter 2 - An Overview of Biological Basics. In: Liu S, ed. *Bioprocess Engineering (Second Edition).* Elsevier, 21-80.

Ma X, Zhang Q, Zhu Q, *et al.*, 2015. A Robust CRISPR/Cas9 System for Convenient, High-Efficiency Multiplex Genome Editing in Monocot and Dicot Plants. *Molecular Plant* **8**, 1274-84.

Martin C, Smith AM, 1995. Starch biosynthesis. *The Plant cell* **7**, 971-85.

Montecillo JaV, Chu LL, Bae H, 2020. CRISPR-Cas9 System for Plant Genome Editing: Current Approaches and Emerging Developments. *Agronomy* **10**, 1033.

Murashige T, Skoog F, 1962. A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum* **15**, 24.

Navarre DA, Goyer A, Shakya R, 2009. Chapter 14 - Nutritional Value of Potatoes: Vitamin, Phytonutrient, and Mineral Content. In: Singh J, Kaur L, eds. *Advances in Potato Chemistry and Technology*. San Diego: Academic Press, 395-424.

Nazarian-Firouzabadi F, Visser RGF, 2017. Potato starch synthases: Functions and relationships. *Biochemistry and Biophysics Reports* **10**, 7-16.

Noakes M, Clifton PM, Nestel PJ, Le Leu R, Mcintosh G, 1996. Effect of high-amylose starch and oat bran on metabolic variables and bowel function in subjects with hypertriglyceridemia. *The American Journal of Clinical Nutrition* **64**, 944-51.

Ooms G, Bossen ME, Burrell MM, Karp A, 1986. Genetic manipulation in potato withAgrobacterium rhizogenes. *Potato Research* **29**, 367-79.

Pan C, Ye L, Qin L, *et al.*, 2016. CRISPR/Cas9-mediated efficient and heritable targeted mutagenesis in tomato plants in the first and later generations. *Scientific Reports* **6**, 24765-.

Puchta H, 2005. The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *Journal of Experimental Botany* **56**, 1-14.

Qi Y, 2019. *Plant Genome Editing with CRISPR Systems*

Methods and Protocols. Springer Science+Business Media, LLC, part of Springer Nature: Humana Press, New York, NY.

Que Q, Chen Z, Kelliher T, Skibbe D, Dong S, Chilton M-D, 2019. Plant DNA Repair Pathways and Their Applications in Genome Engineering. In: Qi Y, ed. *Plant Genome Editing with CRISPR Systems: Methods and Protocols*. New York, NY: Springer New York, 3-24.

Rathore KS, Campbell LM, Sherwood S, Nunes E, 2015. Cotton (Gossypium hirsutum L.). In: Wang K, ed. *Agrobacterium Protocols.* New York Springer Science+Business Media. (2.)

Sánchez T, Dufour D, Moreno IX, Ceballos H, 2010. Comparison of Pasting and Gel Stabilities of Waxy and Normal Starches from Potato, Maize, and Rice with Those of a Novel Waxy Cassava Starch under Thermal, Chemical, and Mechanical Stress. *Journal of Agricultural and Food Chemistry* **58**, 5093-9.

Sandhya D, Jogam P, Allini VR, Abbagani S, Alok A, 2020. The present and potential future methods for delivering CRISPR/Cas9 components in plants. *Journal of Genetic Engineering and Biotechnology* **18**, 25.

Seed KD, 2015. Battling Phages: How Bacteria Defend against Viral Attack. *PLoS pathogens* **11**, e1004847-e.

Shepard JF, Bidney D, Shahin E, 1980. Potato Protoplasts in Crop Improvement. *Science* **208**, 17.

Stokstad E, 2019. The new potato. *Science* **363**, 574-7.

Sunilkumar G, Mohr L, Lopata-Finch E, Emani C, Rathore KS, 2002. Developmental and tissue-specific expression of CaMV 35S promoter in cotton as revealed by GFP. *Plant Molecular Biology* **50**, 463-79.

Svegmark K, Helmersson K, Nilsson G, Nilsson PO, Andersson R, Svensson E, 2002. Comparison of potato amylopectin starches and potato starches — influence of year and variety. *Carbohydrate Polymers* **47**, 331-40.

Taylor A, 2019. Specific gravity of potato tubers. In. *Agriculture and Food.* Department of Primary Industries and Regional Development Goverment of Western Australia.

Toae R, Sriroth K, Rojanaridpiched C, *et al.*, 2019. Outstanding Characteristics of Thai Non-GM Bred Waxy Cassava Starches Compared with Normal Cassava Starch, Waxy Cereal Starches and Stabilized Cassava Starches. *Plants (Basel, Switzerland)* **8**, 447.

Usda-Nass, 2020. Crop Production 2019 Summary (January 2020). In.: United States Department of Agriculture (USDA)

National Agricultural Statistics Service (NASS), 124.

Van Harsselaar JK, Lorenz J, Senning M, Sonnewald U, Sonnewald S, 2017. Genome-wide analysis of starch metabolism genes in potato (Solanum tuberosum L.). *BMC genomics* **18**, 37-.

Voytas DF, Gao C, 2014. Precision genome engineering and agriculture: opportunities and regulatory challenges. *PLoS biology* **12**, e1001877-e.

Wang W, Akhunova A, Chao S, Akhunov E, 2016. Optimizing multiplex CRISPR/Cas9-based genome editing for wheat. *bioRxiv*, 051342.

Watanabe K, 2015. Potato genetics, genomics, and applications. *Breeding science* **65**, 53-68.

Whistler RL, Daniel JR, 2000. Starch. In. *Kirk-Othmer Encyclopedia of Chemical Technology*.

Wong N, Liu W, Wang X, 2015. WU-CRISPR: characteristics of functional guide RNAs for the CRISPR/Cas9 system. *Genome Biology* **16**, 218.

Yuan J, Bizimungu B, Leblanc D, Lague M, 2019. Effects of Field Selection Parameters and Specific Gravity on Culinary Evaluation Traits in a Potato Breeding Programme. *Potato Research* **62**, 361-77.

Zhang H, Cao Y, Zhang H, *et al.*, 2020. Efficient Generation of CRISPR/Cas9-Mediated Homozygous/Biallelic Medicago truncatula Mutants Using a Hairy Root System. *Frontiers in Plant Science* **11**.

Zhao SS, Dufour D, Sánchez T, Ceballos H, Zhang F, 2011. Development of waxy cassava with different Biological and physico-chemical characteristics of starches for industrial applications. *Biotechnology and Bioengineering* **108**, 10.