Inland Norway University of Applied Sciences

Faculty of Applied Ecology, Agricultural Sciences and Biotechnology

Arghavan Armin

Master Thesis

Design and construction of cloning vectors to improve CRISPR-Cas9 gene editing efficiency in potato

Masters in Applied Experimental Biotechnology

2023 - 2024

Consent to lending by University College Library	YES ☒	NO ☐
Consent to accessibility in digital archive Brage	YES ☒	NO ☐

Acknowledgements

The completion of my thesis would not have been feasible without the assistance and encouragement of many individuals who helped me through this journey. Upon completion of this research, it gives me great pleasure to thank everyone who helped make this study possible.

Firstly, I would like to acknowledge the Biotechnology Department at Hamar Campus at Inland Norway University of Applied Sciences for providing me with the opportunity to pursue my education, and Frøydis Deinboll Myromslien, head of the Biotechnology Department for her valuable insight and support.

I would like to express my gratitude to my main supervisor, Professor Robert Wilson, for his continuous support, inspiration, and advice throughout this project, along with my co-supervisors, Professor Wenche Johansen and Diana Katherine Castillo Avila, for their insight and unconditional support during the project.

Additionally, I am grateful to the lab engineers, Wenche Kristiansen, Hanne Greaker, and Anne Bergljot Falck-Ytter, for their support and instruction regarding laboratory procedures and instruments.

In the end, I would like to acknowledge the entire personnel at INN, Hamar Campus, for their assistance in seeing this project through to completion.

Abbreviations

A: adenine nucleobase

AS1: asparagine synthetase 1

bp: base pair

C: cytosine nucleobase

Cas9: CRISPR-associated protein 9

CRE: cis-regulatory elements

CRISPR: clustered regularly interspaced short palindromic repeats

DNA: deoxyribonucleic acid

DSB: double-strand break

DW: dry weight

E. coli: *Escherichia coli*

EDTA: ethylenediaminetetraacetic acid

FW: fresh weight

G: guanine nucleobase

gDNA: genomic DNA

HDR: homology-directed DNA repair

LB – Ampicillin: Luria-Bertani medium containing ampicillin

NHEJ: nonhomologous end joining

NTC: non-template control

PAM: protospacer adjacent motif

Pol II: RNA polymerase II

Pol III: RNA polymerase III

RNA: ribonucleic acid

RNP: ribonucleoprotein

RT: room temperature

RT-qPCR: reverse transcription-quantitative polymerase chain reaction

SGA: steroidal glycoalkaloid

sgRNA: single guide RNA

snRNA: small nuclear RNA

S.O.C medium: super optimal broth with catabolite repression

StSSR2: sterol side chain reductase 2

T: thymine nucleobase

TBP: TATA-binding protein

TF: transcription factor

TFIID: multi-subunit transcription factor II D

tracrRNA: trans-activating CRISPR RNA

TSS: transcription start site

USE: upstream sequence element

UMGMA: unweighted pair group method with arithmetic mean

UV: ultraviolet

VInv: vacuolar invertase

Table of Contents

Abstract

Potatoes are beneficial for ensuring global food security and CRISPR-Cas9 gene editing technology can be used to improve numerous potato traits. This technology employs a single guide RNA (sgRNA) to direct the genome editing process mediated by the Cas9 endonuclease. Precise genome editing is achievable through careful planning and evaluation of sgRNA expression to increase the efficiency of CRISPR-Cas9 gene editing.

This project aimed to design and construct different pChimera cloning vectors harboring StU6 promoter upstream of a sgRNA scaffold using In-Fusion cloning.

Firstly, the StU6 promoter was amplified in potato Désirée and Asterix cultivars and cloned into an empty vector and *E. coli* cells were transformed. Then, the plasmid was isolated from the putative positive colonies, sequenced and distinct allelic variations corresponding to four promoter sequences were found in potato Désirée and Asterix cultivars.

The phylogenetic analysis of the StU6 promoters distinguished a high degree of sequence identity in Désirée and Asterix clones, and seven clones from Désirée were chosen for In-Fusion cloning. DNA fragments of seven StU6 promoters and pChimera were PCR amplified to retrieve insert and vector fragments, respectively. In-Fusion cloning was performed to clone the insert fragment into a linearized vector to create the final cloning vector containing endogenous potato StU6 promoter, and Stellar Competent Cells were transformed. Colony PCR was performed to verify putative positive colonies, but no amplification was observed.

Overall, suboptimal PCR conditions or the use of a too high insert-to-vector ratio for the In-Fusion cloning reaction may potentially be the contributing factors to the lack of amplification.

Keywords: In-Fusion cloning, CRISPR-Cas 9 gene editing, StU6 promoter, sgRNA expression, PCR, RT-qPCR, colony PCR, pChimera, potato, protoplast, Désirée, Asterix, sanger sequencing, phylogenetic analysis.

1. Introduction

1.1 Background of Study

The rate at which the world's population is growing is alarming and by 2050, it is predicted to surpass 10 billion people (Thatcher et al., 2018). As a result, food demand is projected to increase from +35% to +56% between 2010 and 2050, and the population at the risk of hunger will increase from -91% to +8% (Van Dijk et al., 2021).

This concern can be addressed by implementing potatoes as an important component of a wholesome and balanced diet in combination with other vegetables and whole-grain foods (Devaux et al., 2020). Due to its high nutritional value, producibility in all climate regions, and high production volume, potatoes can play a crucial role in ensuring global food security for the growing population (Haverkort et al., 2009).

1.2 The Role of Potato in Global Food Security

Potato (*Solanum tubersom L.*) is the world’s fourth major crop following maize, rice and wheat, and it is often regarded as a staple food in many parts of Europe and the Americas. The number of edible potato varieties (cultivars) grown worldwide exceeds 4800 being diverse in genetics, from diploid (2n = 2x = 24) to pentaploid (2n = 5x = 60) varieties, whereas common potato varieties are tetraploid (2n = 4x = 48) with a basic chromosome number of 12 (Watanabe, 2015).

Approximately 19 million hectares of cropland are being used for the cultivation of potatoes worldwide, with 378 million tons produced per year (Faostat, 2017). The potato is cultivated as a summer crop in the tropical highlands of Central

and South America as well as in lowland temperate regions of the world, a winter crop in lowland subtropical regions, and a spring and autumn crop in the Mediterranean. In some regions of the world, including the equatorial highlands of South America and East Africa, as well as certain regions of China and Brazil, it can be grown all year round (Bradshaw, 2019).

According to the Norwegian Agriculture Industry, 367,400 metric tons of potatoes were produced in Norway in 2022 with 40% of the total national production in the Hedmark region (*Norway: Production Volume of Potatoes 2022 | Statista*, n.d.).

The health benefits of potatoes appear promising due to the abundance of carbohydrates, resistant starch, quality proteins, vitamins C, B6 and potassium (Camire et al., 2009). Although potatoes are typically thought of as a source of carbohydrates, they contain more protein than most cereals, including rice and wheat, at 10% dry weight (DW) and 2% fresh weight (FW). Although they are a good source of lysine, their nutritional value is limited by low levels of sulphur amino acids, methionine and cysteine (Bártová et al., 2015; Friedman, 1996).

Yellow-fleshed potatoes also contain high levels of lutein, zeaxanthin, and other antioxidants that may help protect against degenerative diseases and age-related illnesses (Burgos, Muñoa, et al., 2013), while the purple and red-fleshed cultivars are an important source of anthocyanin (Burgos, Amoros, et al., 2013). Despite the low concentrations of iron and zinc in potatoes compared with cereals and vegetables, their bioavailability is greater due to the higher levels of ascorbic acid and low levels of phytic acid, which are iron absorption enhancer and inhibitor, respectively (Fairweather-Tait, 1983).

1.3 CRISPR-Cas9 Gene Editing Technology

For most life forms, the capacity to resist viral predation is a crucial component of survival. This includes bacteria, which have evolved in a variety of hostile environmental conditions despite competition for limited resources and bacteriophage infection. The capacity of bacteria to adaptably increase their genetic inventory, dynamically regulate genome homeostasis and impede viruses with a variety of defence mechanisms accounts for their long-term viability (Barrangou, 2015).

Clustered regularly interspaced short palindromic repeats (CRISPRs) are nucleic acid-based adaptive immunity widely found in bacteria and archaea, which were first discovered in the genome of *Escherichia coli* in 1987 by Jennifer Doudna and Emmanuelle Charpentier (Wiedenheft et al., 2012). Repetitive sequences with a repeat–spacer–repeat pattern were discovered in phylogenetically divergent bacterial and archaeal genomes and later recognized to be identical to the DNA fragments of pathogenic bacteriophages that had previously infected the prokaryote. The idea that CRISPRs offer a genetic memory of infection was inspired by this observation (Bolotin et al., 2005; Mojica et al., 2005). The underlying mechanisms of CRISPR were subsequently discovered by Francisco Mojica, who gave it its name and suggested its numerous applications as a widespread genome editing tool (Mojica et al., 2005).

Cas9 (CRISPR-associated protein 9) endonuclease plays a crucial part in prokaryotes' antiviral defence mechanism by performing cuts in double-stranded DNA. Single guide RNA (sgRNA) is a short RNA sequence that serves as a template for the endonuclease to recognize CRISPR sequences and cleave particular DNA strands complementary to the spacer sequence (Mali et al., 2013).

CRISPR/Cas9-mediated double-strand break (DSB) repair mechanism is employed by the cell to repair the break; either by nonhomologous end joining (NHEJ) or homology-directed DNA repair (HDR) pathways (Tang et al., 2019). While random deletions and insertions are created by the NHEJ pathway, the HDR pathway implements precise insertions, base substitutions across DSB sites or two DSBs, and other modifications by using homologous donor DNA sequences from sister chromatids or foreign DNA (Tang et al., 2019).

The objective of genome editing experiments is to convert a targeted DNA sequence into a new, desired DNA sequence that harbours the mutation of interest depending on the application (Anzalone et al., 2020). These might include (i) point mutations, or the conversion of DNA base pairs; (ii) deletion; (iii) insertion; or (iv) a combination of the aforementioned modifications (including replacement of DNA base pairs). Each of these modifications is mediated by a distinct class of CRISPR–Cas editing agents (Anzalone et al., 2020).

In principle, CRISPR-Cas9 can perform DSBs at any locus that is 3 base pairs (bp) upstream of an NGG protospacer adjacent motif (PAM), where N is any nucleobase; adenine (A), cytosine (C), guanine (G) or thymine (T). (Liu et al., 2021). The 20-nt guide sequence of the sgRNA and the existence of a PAM sequence next to the target region in the genome are believed to be the factors that closely regulate the targeting specificity of Cas9 (Fu et al., 2013).

However, significant safety concerns are raised by the possibility that the DNA repair process induced by DSBs might cause unintentional mutations at homologous off-target loci in addition to causing alterations at target sites (Fu et al., 2013; Liu et al., 2021). These off-target activities can be significantly reduced by employing high-fidelity Cas9 variants, selecting a better target sequence, or

the application of inhibitors that quickly stop endonuclease activity (He et al., 2019; Rose et al., 2020).

1.4 Genome Editing in Potato

One of the main long-term objectives of plant breeding efforts has been the improvement of crop's nutritional value such as essential amino acids composition (Chakraborty et al., 2000; Goo et al., 2013; Raina & Datta, 1992), increased plant resistance to pathogens (Kieu et al., 2021), as well as reduction of the toxic compounds (Sawai et al., 2014).

From a health perspective, the contents of hazardous secondary metabolites in potato tubers known as steroidal glycoalkaloids (SGAs) should not exceed more than 20 mg/100 g of the FW. Alpha-chaconine and alpha-solanine are the two main SGAs found in farmed potato tubers (*Solanum tuberosum*) (Krits et al., 2007). Significant reduction in SGA contents in potatoes has been achieved to ensure food safety and palatability via sterol side chain reductase 2 gene (StSSR2) knockout that is involved in SGA metabolism (Zheng et al., 2021).

Additionally, acrylamide is a carcinogenic compound produced in potato tubers from the reduction of glucose and fructose by plant invertases at high temperatures (Vinci et al., 2012). To decrease the accumulation of reducing sugars and the synthesis of acrylamide in potato tubers, the genes encoding vacuolar invertase (VInv) and asparagine synthetase 1 (AS1) were silenced in potato Atlantic and Désirée cultivars (Halford et al., 2022; Ly et al., 2023).

Many potato traits can be improved through the implementation of Cas9 endonuclease and sgRNA to cleave DNA complementary to a specified target area while the cellular repair mechanism offers opportunities for gene editing at the cleavage site (Jinek et al., 2012). NHEJ is the most common repair outcome, potentially leading to gene disruptions through indels. Although less common,

HDR permits precise genetic alterations when an appropriate template is available.

In order to prevent the integration of foreign DNA into the potato genome, it is advised from a scientific and regulatory standpoint to use a DNA-free genome editing technique that delivers CRISPR-Cas9 ribonucleoproteins (RNPs) to potato protoplasts (Andersson et al., 2018). RNPs have been reported to exhibit greater specificity and faster onset of action due to their ability to function without requiring intracellular transcription and translation (Kim et al., 2014). This method offers a viable substitute for conventional DNA delivery methods by producing transgene-free targeted genome editing, as well as reducing off-target mutations (Cho et al., 2013; Liang et al., 2017; Svitashev et al., 2016).

Compared to diploid potato cultivars, tetraploid cultivars provide a substantially higher yield (Wang et al., 2022). However, it is challenging to maintain the desirable features of potato clones through mutagenesis because of the high heterozygosity in the tetraploid potato genome. Although challenging, full allelic potato gene editing has been achieved by applying endogenous potato StU6 promoter to drive the expression of CRISPR cassette (Johansen et al., 2019).

1.5 U6 Promoter

Promoters are essential regulatory elements found upstream of the transcription start site (TSS) playing a key role in regulating gene transcription through the employment of transcription factors (TFs) and RNA polymerases (Kor et al., 2023). The efficiency of plant gene editing can be significantly improved by the selection of the appropriate promoter. Therefore, studying the U6 promoter is crucial since it is the most optimal promoter for the expression of sgRNA for CRISPR-Cas9 gene editing (Johansen et al., 2019; Kor et al., 2023).

The relationship between the promoter structure and the selection of RNA polymerase for transcription has already been investigated in the literature (Dahlberg & Lund, 1988). While RNA polymerase II (pol II) is involved in the transcription of most coding U-snRNA (small nuclear ribonucleoprotein) class's genes such as U1, U2 and U5, RNA polymerase III (pol III) transcribes other non-coding snRNA genes, including U6 in plants (Waibel & Filipowicz, 1990).

The U6 promoter drives the expression of the U6 snRNA gene in plants and contains a TATA box or Goldberg–Hogness box region located 25-35 bp upstream of the TSS (Guerineau & Waugh, 1993). This conserved motive is present in archaea and eukaryotes and distinguished by a recurring pattern of A and T bases with a consensus of TATAWAW, where W is either A or T (Kanai et al., 2023).

In eukaryotes, the first stage of transcription initiation occurs at the TATA box and is known as preinitiation complex formation. The multi-subunit transcription factor II D (TFIID) attaches to the minor groove of the TATA box at its TATA-binding protein (TBP) subunit, which triggers the formation of the preinitiation complex, illustrated in Figure 1 (Starr & Hawley, 1991; White, 2011).

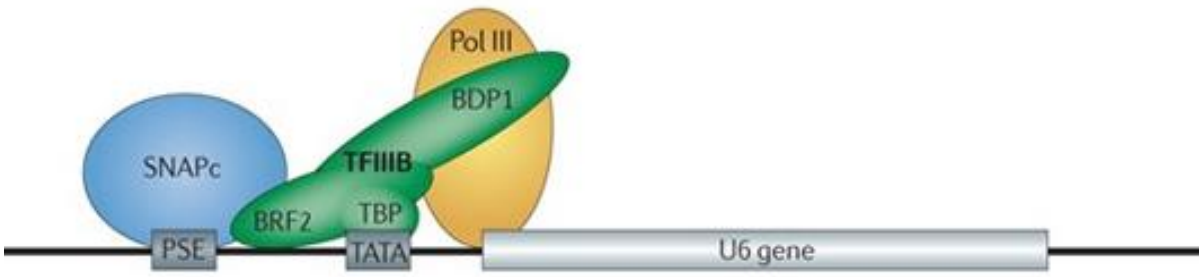

Figure 1 - The promoter structure and basal transcription machinery of RNA polymerase III-transcribed genes demonstrating TATA box region, TBP, TFIID and RNA polymerase III (White, 2011).

Additionally, most plant promoters contain an Upstream Sequence Element (USE) 30 bp upstream of the TATA box region, which increases the expression of a neighbouring gene (Kanai et al., 2023). Both these motives are cis-regulatory elements (CREs) present in plant promoter sequences and are necessary for the regulation of gene transcription. Figure 2 presents the USE and TATA box motives upstream of the TSS (+1) of the U6 snRNA gene in *Arabidopsis thaliana* and *Ricinus communis*.

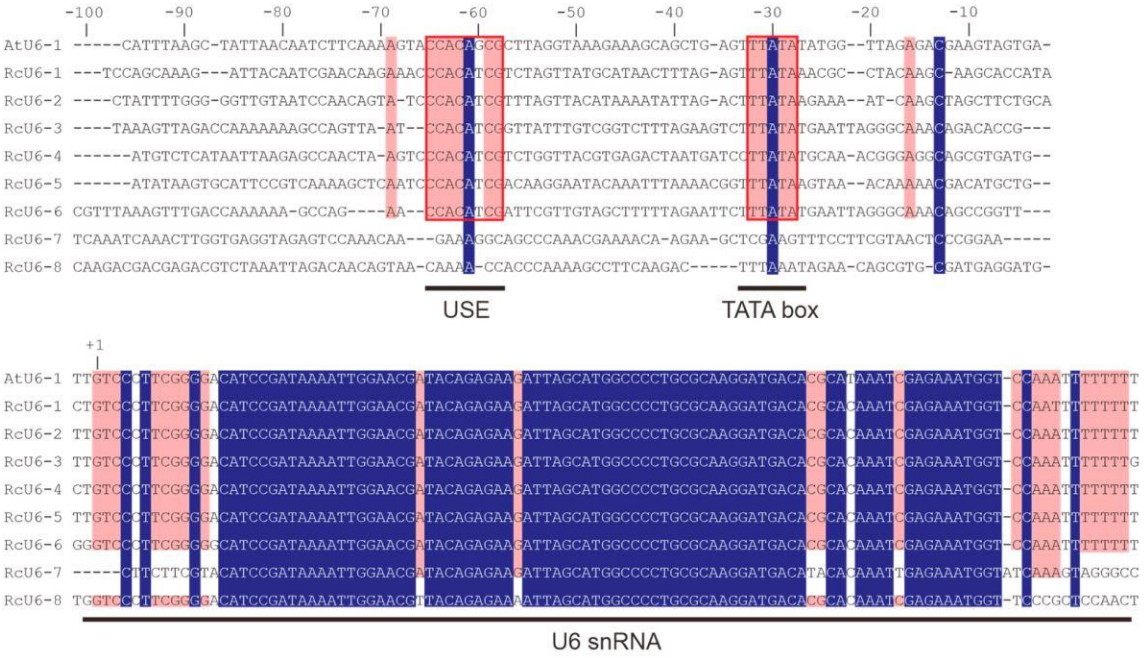

Figure 2 – The USE, TATA box, TSS and U6 snRNA gene in *Arabidopsis thaliana* (AtU6-1) and *Ricinus communis* (RcU61-8) (Kanai et al., 2023).

1.6 In-Fusion Cloning

In-Fusion cloning can be implemented for fast, efficient, and directional cloning of one or more DNA fragments into any vector at any location. The key

component of the In-Fusion cloning method is the application of the In-Fusion enzyme premix, which efficiently and precisely fuses DNA fragments (such as PCR-generated inserts and linearized vectors) by recognizing 15-bp overlaps at their ends. Therefore, no restriction digestion, phosphatase treatment, or ligation is required. Using specific primers especially designed for the amplification of the target sequences, these 15-bp overhangs can be generated by a PCR reaction. The schematic overview of the In-Fusion cloning method from In-Fusion® HD Cloning Kit User Manual (Takara Bio USA, Inc) is presented in Figure 3.

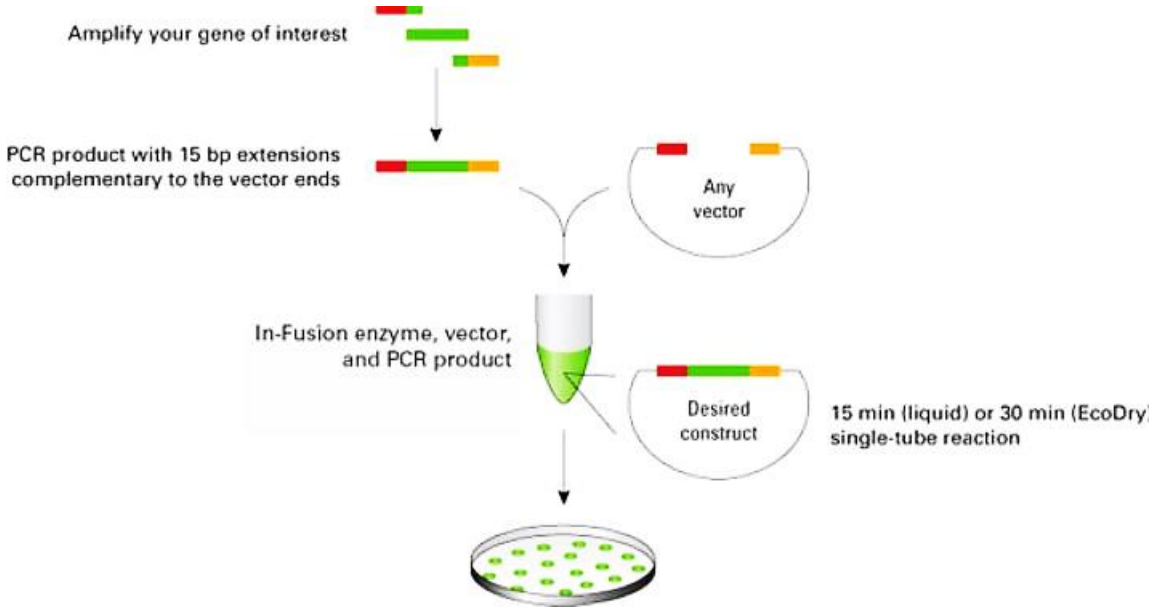

Figure 3 - In-Fusion cloning method for cloning of amplified insert fragment into a linearized vector with 15-bp complementary overhangs in a single tube reaction, In-Fusion® HD Cloning Kit User Manual (Takara Bio USA, Inc).

pChimera vector (2909 bp) is conventionally used for plant gene editing, which contains AtU6-26 promoter from *Arabidopsis thaliana* and sgRNA scaffold sequence, as well as ampicillin resistance gene as a selection marker in its backbone. The features of this plasmid are demonstrated in Figure 4.

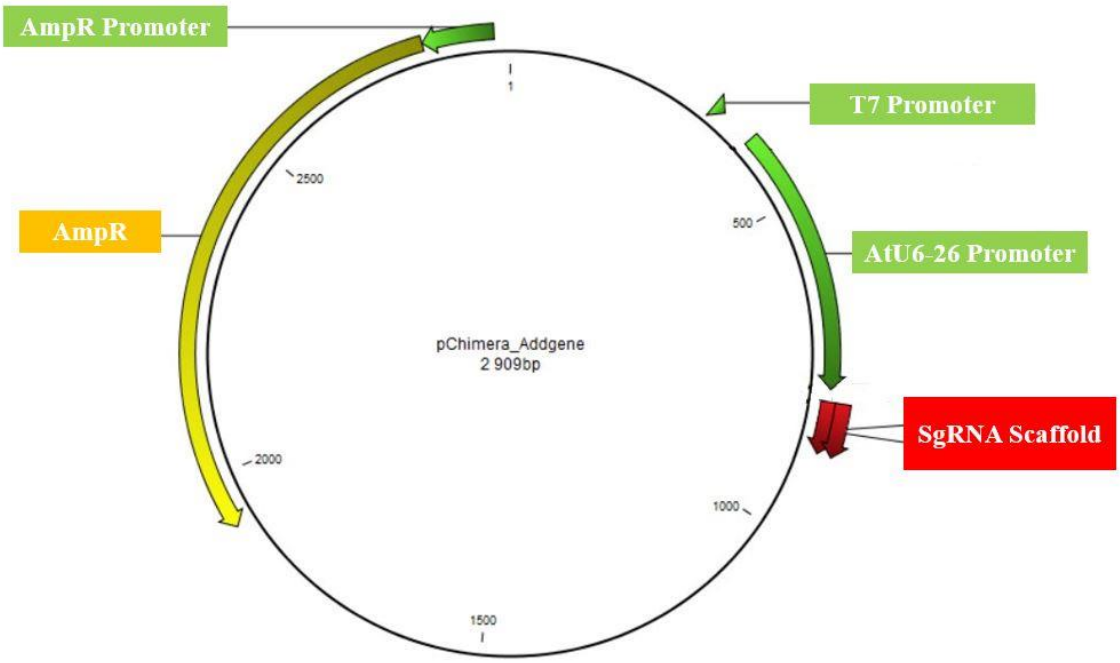

Figure 4 – The map of pChimera cloning vector (2909 bp) demonstrating T7 promoter, AtU6-26 promoter, sgRNA scaffold, ampicillin resistance (AmpR) gene and AmpR gene promoter in CLC Main Work Bench.

To construct the desirable cloning vector, the AtU6-26 promoter can be replaced by a potato promoter (StU6) using In-Fusion cloning. For this purpose, specific primers can be designed to amplify the StU6 promoter and pChimera vector to create PCR products with 15-bp complementary overhangs.

1.7 Impact of this Study

An efficient method of converting genetic knowledge into improved crop varieties is through genome editing. The promising future of potato gene editing and the numerous applications of CRISPR-Cas9 technology in this field has grown exponentially during the past decades and requires vigilant experimental efforts.

This experiment helps us to choose the proper promoter for the expression of the sgRNA scaffold in potato protoplast via the design and construction of cloning vectors that harbour endogenous potato StU6 promoter. After transfection of potato protoplast with several vectors, RT-qPCR can be used to compare the transcription level of sgRNA and ultimately to determine the efficacy of each StU6 promoter for CRISPR-Cas9 gene editing.

1.8 Aim and Objectives

The overall aim of this study was to design and construct different cloning vectors containing endogenous potato StU6 promoter using In-Fusion cloning to drive the expression of sgRNA in potato protoplasts in order to improve the efficiency of CRISPR Cas9 gene editing.

1.8.1 Specific Objectives

1. PCR Amplification of StU6 promoter from potato gDNA

2. Cloning of the StU6 promoter into Zero Blunt® vector and bacterial transformation

3. Plasmid isolation and sequencing to verify positive clones

4. Phylogenetic analysis to select individual clones

5. Restriction digest analysis to confirm the identity of pChimera

6. PCR amplification of StU6 promoter and pChimera

7. In-Fusion cloning, bacterial transformation and colony PCR

2. Material and Methods

2.1 Study Design and Area

A research project's ability to run consistently and effectively is greatly impacted by the careful design of the experimental procedure. Since an experimental design is essential to secure the goals and objectives of the project, a schematic presentation of the experimental design in presented in Figure 5.

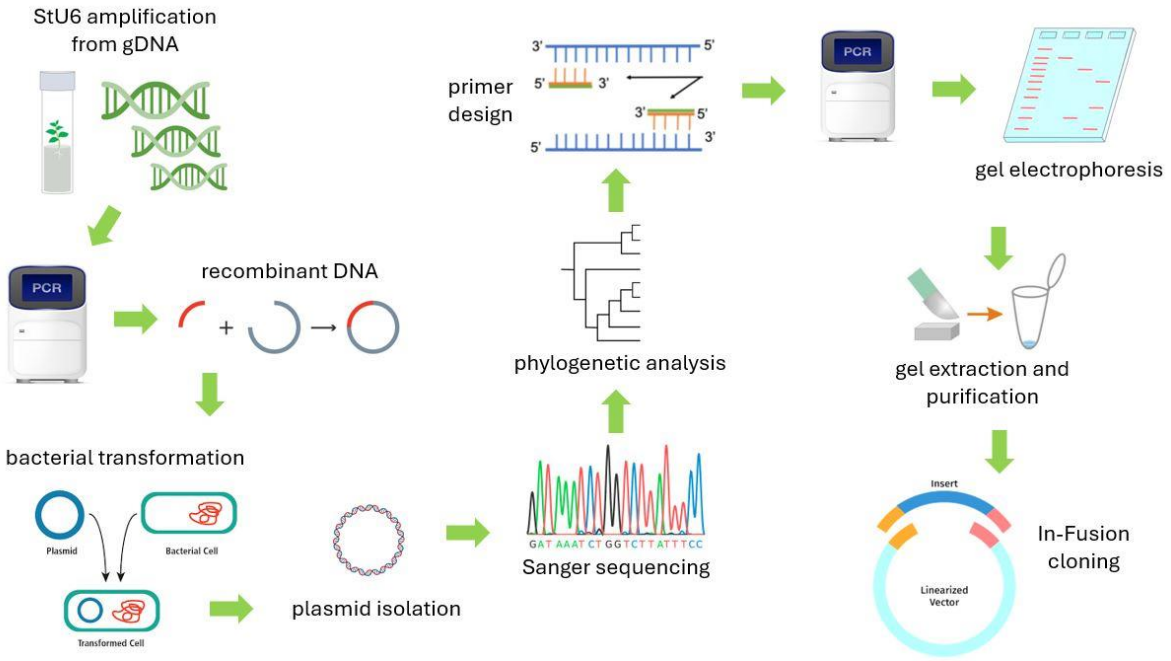

Figure 5. Schematic representation of the experimental design.

This study was carried out in the Biohus laboratory facility of the Inland Norway University of Applied Sciences, Hamar Campus. To accomplish the goals and objectives of this investigation, the workflow shown in Figure 6 was followed.

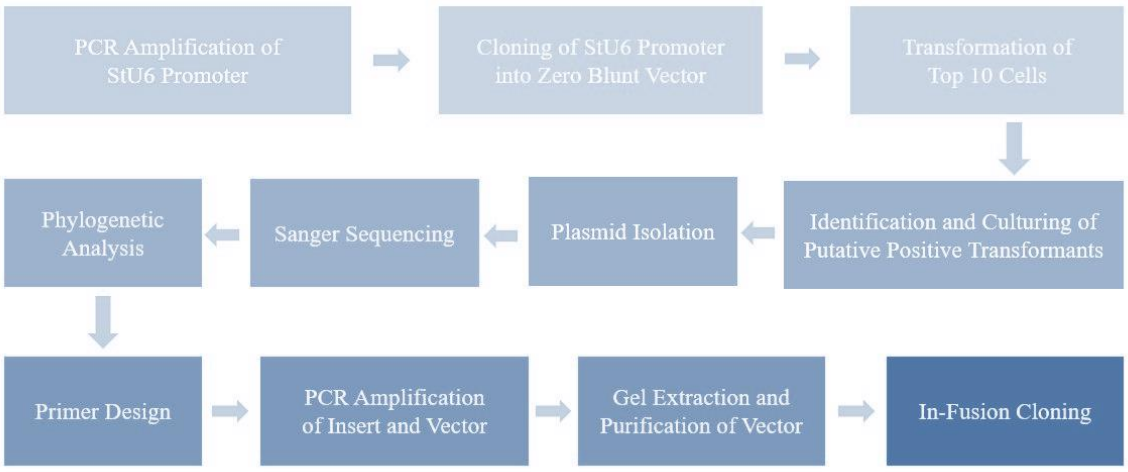

Figure 6 – Experimental workflow for design and construction of cloning vectors for transfection of potato protoplasts using In-Fusion cloning.

2.2 PCR Amplification of the StU6 Promoter

In order to retrieve the StU6 promoter from potato Désirée and Asterix cultivars gDNA, 4 sets of forward primers were used for the amplification of StU6-1, StU6-2, StU6-3 and StU6-4, as well as a common reverse primer for all of them, reported in Table 1 (Johansen et al., 2019).

Table 1 – The forward and reverse primer sequences used for the amplification of StU6 promoter from potato gDNA.

Promoter	**Forward Primer (5'-3')**	**Common Reverse Primer (5'-3')**
StU6-1	AGCAAGATGCAATGTATCAACTCA	GCCATGCTAATCTTCTCTGTATCG
StU6-2	ACCACTTAAACTGAGAACAGTCAA	
StU6-3	TTCACTTAGTTCAGTTGCATTATGTC	
StU6-4	GATAAATTCTTAAAGTTGAGTAACC	

The PCR reaction was performed according to the user protocol for Phusion High-Fidelity PCR Master Mix (F531L, Thermo Scientific, USA), and the PCR components and cycles are reported in Tables A.1 and A.2 in the Appendix, respectively. 1.5% agarose gel electrophoresis was used to visualize StU6 promoter bands in Désirée and Asterix potato cultivars, presented in Figure 7.

2.3 Cloning of StU6 Promoter and Transformation

After the PCR amplification of Désirée and Asterix StU6 promoters, 1 µl of 4 amplicons (StU6-1, StU6-2, StU6-3 and StU6-4) were cloned into pCR® Blunt vector according to the manufacturer's protocol for Zero Blunt® PCR Cloning kit (Invitrogen, USA). The details of the ligation reaction are presented in Table A.3 in the Appendix.

Thirty-three µl OneShot Top 10 Chemically Competent Cells (Invitrogen, USA) were used for StU6-1, StU6-2 and StU6-3, while 50 µl cells were used for StU6-4. The cells were thawed on ice and transformed using 3 µl of the ligation tubes according to the producer's protocol with minor modifications. The heat shock was performed at 42 ˚C for 45 seconds and the cells were transferred on ice immediately.

Two hundred and fifty µl of S.O.C. (super optimal broth with catabolite repression) medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 20 mM glucose) was added into each tube and incubated on shaking incubator for 1 hour. Then, the cells were plated out on LB plates supplemented with Ampicillin to a final concentration of 100 µg/ml to grow overnight at 37 ˚C.

2.4 Plasmid Isolation and Quantification

Two putative positive colonies from the transformation reaction of StU6-1, StU6-2 and StU6-3 and 20 colonies from StU6-4 in both Désirée and Asterix cultivars were selected and transferred into sterile plastic culture tubes containing 3 ml of LB broth medium with 100 µg/ml of Ampicillin. The tubes were then placed into a shaking incubator (225 rpm) for 18 hours at 37 °C.

Plasmid DNA was isolated from 3 ml cultures using PureYield™ Plasmid Miniprep System (A1222, Promega, USA) according to the manufacturer's instructions. Only minor modifications were done where water was used for pellet resuspension prior to cell lysis and centrifugation was performed at the maximum speed of 14.8 rpm. After elution, a Qubit 4 Fluorometer was used to quantify the concentration of plasmid DNA in each sample (reported in Table 5).

2.5 Sanger Sequencing

2 plasmid samples from the Désirée cultivar amplicons were obtained for StU6-1, StU6-2 and StU6-3, as well as 20 samples for StU6-4. Additionally, 2 plasmid samples from the Asterix cultivar were obtained for StU6-1, StU6-2 and 20 samples for StU6-4. The samples were analysed by Sanger sequencing (Eurofins Genomics, Ebersberg, Germany).

The sequencing data associated with StU6-1, StU6-2, StU6-3 and StU6-4 in Désirée and StU6-1, StU6-2 and StU6-4 in Asterix were trimmed and analysed using QIAGEN CLC Main Workbench (CLCMWB) (version 7.9.3, QIAGEN Aarhus AS) and are reported in the Appendix (page 75-78).

A sequence pattern of 5'-GTCCCTTCGGGGA-3' was used to locate the snRNA gene and the first G was annotated as TSS. To confirm the presence of potato

promoter sequence in Désirée and Asterix clones, TATA box and USE conserved motives were studied in CLCMWB. The TTATAT consensus sequence was used to identify TATA box regions in the clones, while CCACATCG or CCTCATCG consensus sequences were used for USE identification.

The StU6 promoter sequences were retrieved from the NCBI nucleotide sequence database using accession numbers (Z17290.1, Z17292.1, Z17293.1 and Z17301.1 for StU6-1, StU6-2, StU6-3 and StU6-4, respectively) and are presented in FASTA format in Table 2.

Table 2 - StU6 promoter sequences retrieved from NCBI nucleotide sequence database in FASTA format, as well as the accession number and length (bp).

Promoter	**Accession Number**	**Length (bp)**	**FASTA Sequence**
StU6-1	Z17290.1	421	>Z17290.1 S.tuberosum DNA for U6 small nuclear RNA promoter region AAATGGTACAAGTTGAATATGGGGGCAAATCTGGACTCTAGGCTT AGTTGGGCTCTATGTGCATATAAAAGCAAGAGCAAAAACCTGTAG CTAGGTCCAGGCCCATGCCTTTGGTAAAACTCAATGTGCTAATTCT CCCTCATCGTCTGCAGAGAGAAGCCTCGCTGTGTTTATATAATTGA ACAGTAACATGCATGCTT
StU6-2	Z17292.1	433	>Z17292.1 S.tuberosum DNA for U6 small nuclear RNA promoter region GATCCAGGCCCATGCAGTTGAAAAATACTCAACTAGAAAGCTATTT TCCTCACATCGGCTAAAGAAAGCTTCTTTGTTTTATTTATATTGCGT AACATTAACATCTATAATT
StU6-3	Z17293.1	372	>Z17293.1 S.tuberosum DNA for U6 small nuclear RNA promoter region GATCCAGGCCCGCAAAAGAAACCCAACAAGCAAATTATCCCTCAT CGAATGCATAAAGCTTCTTTGTCTAGTTTATATGGCGGAATATTAAC ATGTGTGCTT
StU6-4	Z17301.1	318 321 374	>Z17301.1 S.tuberosum gene for U6 small nuclear RNA AAACAAGCGCAAAAAGGAGTCCAGGCCCGTGTTAGCGTGAAGAC TCAACCAGCGATTTCTCCCTCATCGGTTGCACAGAAAAGCTGTGTG TTGTTTATATGGCGAAACCTAACAGTCTGACTTGTCCCTTCGGGGA CATCCGATAAAATTGGAACGATACAGAGAAGATTAGCATGGCCCCT GCGCAAGGATGACACGCACAAATCGAGAAATGGTCCAAATTTTTTT TTGCCATTTTTTCCGAGCTCCATTTTCAAATTTTTTGGGGGTTTTGA AGTCGCCTA

Additionally, the promoter sequence of the *Arabidopsis thaliana* U6-1 snRNA gene was retrieved from NCBI using the accession number (X52527.1) and is presented in Table 3.

Table 3 - The accession number, length (bp) and FASTA sequence associated with the *Arabidopsis thaliana* U6-1 snRNA gene from NCBI.

Name	**Accession Number**	**Length (bp)**	**FASTA Sequence**
Arabidopsis thaliana U6-1 snRNA gene	X52527.1	795	>X52527.1 Arabidopsis thaliana U6-1 snRNA gene AGAAATCTCAAAATTCCGGCAGAACAATTTTGAATCTCGATCCGTAGAAACGAG ACGGTCATTGTTTTAGTTCCACCACGATTATATTTGAAATTTACGCTGAGTGTG AGTGAGACTTGCATAAGAAAATAAAATCTTTAGTTGGGAAAAAATTCAATAATA TAAATGGGCTTGAGAAGGAAGCGAGGGATAGGCCTTTTTCTAAAATAGGCCCAT TTAAGCTATTAACAATCTTCAAAAGTACCACATCGCTTAGGTAAAGAAAGCAGC TGAGTTTATATATGGTTAGAGACGAAGTAGTGATTGTCCCTTCGGGGACATCCG ATAAAATTGGAACGATACAGAGAAGATTAGCATGGCCCCTGCGCAAGGATGACA CGCATAAATCGAGAAATGGTCCAAATTTTTTTTGGCAAAAATTTTCAGATTTTT TCTTCATCTGTAGATTTCTGGGTTTTTTTTTCCGTTTCGGTGAATCATAAGTGA AGTTTTGGATGCAAATCTGCGCGAAAAAAGTTGGACCTGCAATGAGCTTATTTA GATAGCTAAGACAAAGTGATTGGTCCGTTGTTTCAGTTCTGATTGTCAGAGAGT TTGTTTCGAGACGGCGACACCAATGCGTTTTGTTAACCAGATTTCGGGTAAGAA ATGTATCGAGAGTTTGTTTCGAGACGGCTACATCATTTTCTTATGAAGGGTGAA ATTAGATAGACCAAAGATTGAAACACAACATTTCTTTCACAAAAATATAATAAA CTTGATAGCATTTAGGATCAGCTACTCTCACTAATCAGT

Each clone was furthered searched in NCBI using BLASTn RefSeq Genome Database (refseq_genome) associated with *Solanum tuberosum* (taxied 4113). The best hits with sequence ID, score, e-value, identity and gaps in Désirée and Asterix clones are presented in the Appendix in Tables A.13 and A.14, respectively.

Additionally, the sequences from 26 Désirée and 24 Asterix clones were further aligned in CLCMWB and the clones with 100% identity were excluded (data not reported). Based on this similarity assessment, the Asterix clones were further aligned with Désirée clones and excluded from the cloning experiment since they were 100% identical to the Désirée clones (data not reported).

Furthermore, seven representative StU6 clones from Desiree (StU6-1 D1, StU6-2 D1, StU6-2 D2, StU6-3 D1, StU6-4 D1, StU6-4 D6, StU6-4 D14) were aligned with the *Arabidopsis thaliana* U6-1 snRNA gene (X52527.1) and StU6 promoter sequences from NCBI (Z17290.1, Z17292.1, Z17293.1 and Z17301.1) in CLCMWB (Figure 8).

2.6 Phylogenetic Analysis

Seven representative clones from Désirée (StU6-1 D1, StU6-2 D1, StU6-2 D2, StU6-3 D1, StU6-4 D1, StU6-4 D6, StU6-4 D14) and 5 from Asterix (StU6-1 A1, StU6-1 A2, StU6-2 A2, StU6-4 A1, StU6-4 A2) were used for the phylogenetic analysis.

The promoter sequences retrieved from NCBI (Z17290.1, Z17292.1, Z17293.1 and Z17301.1) were then aligned with the representative Désirée and Asterix clones in CLCMWB to study their similarities and a phylogenetic tree was constructed based on the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) method (reported in the results, Figure 9).

2.7 Restriction Digest Analysis of the Cloning Vector

pChimera (Plasmid #61476) was ordered from AddGene and a NanoDrop spectrophotometer (Thermo Scientific, USA) was used for quality and quantity evaluation using 2 µl of plasmid DNA (reported in Table 8).

To validate the identity of pChimera in the tube, a restriction digest analysis was performed using PvuII restriction enzymes (New England Biolabs). According to the protocol available on the New England Biolabs website for PvuII

restriction digest analysis enzyme, 5.2 µl of the plasmid (0.19 µg/µl) was digested by PvuII enzyme and NEBuffer r3.1 (New England Biolabs). The reaction details are presented in the Appendix in Table A.4.

After 15 minutes of incubation at 37 ˚C, 4 µl of the sample was diluted in 6 µl of Nuclease-free water. Then, 2 µl of 6X Gel Loading Dye Purple (New England Biolabs) was added and the bands were visualized by 1% TAE agarose gel electrophoresis (reported in the results, Figure 10).

2.8 Primer Design for the Amplification of Insert and Vector

The online Takara Bio primer design tool was used to design primers specific for the amplification of insert (StU6-1 D1, StU6-2 D1, StU6-2 D2, StU6-3 D1, StU6-4 D1, StU6-4 D6, StU6-4 D14) and vector (pChimera) for InFusion cloning. The forward and reverse primer sequences for each PCR reaction are presented in Table 4.

Table 4 - The forward and reverse primer sequences for the PCR amplification of insert fragments (StU6-1 D1, StU6-2 D1, StU6-2 D2, StU6-3 D1, StU6-4 D1, StU6-4 D6, StU6-4 D14) and vector (pChimera).

Template	Direction	Primer Name	Sequence (5'-3')
StU6-1 D1	Forward	1D13Ins1F	CCAGAGCTCCCTAGGAGTTTAACATATTTACAAATTGACA
	Reverse	1D14Ins1R	CTTCTCGAAGACCCCCAAGCATACATGTTACTGTTCA
StU6-2 D1	Forward	2D13Ins1F	CCAGAGCTCCCTAGGTACTACTTGCGCGTGCAT
	Reverse	2D14Ins1R	CTTCTCGAAGACCCCCAATTATAGATGTTAATGTTACGCA
StU6-2 D2	Forward	2D23Ins1F	CCAGAGCTCCCTAGGTGAAAAAATAATACTAAATTTTTGA
	Reverse	2D24Ins1R	CTTCTCGAAGACCCCCAATTATAGATGTTAATGTTACGC
StU6-3 D1	Forward	3D13Ins1F	CCAGAGCTCCCTAGGTCAGTTAATACAGAAAGAAAAAATC
	Reverse	3D14Ins1R	CTTCTCGAAGACCCCCAAGCATACATGTTAATATTCCG
StU6-4 D1	Forward	4D13Ins1F	CCAGAGCTCCCTAGGTGTCTTTATACACCCCTACCTAG
	Reverse	4D14Ins1R	CTTCTCGAAGACCCCCGAGCTAAACTGTTAGGTTTCGC
StU6-4 D6	Forward	4D63Ins1F	CCAGAGCTCCCTAGGTTATGTCTTTATACAACCCTACC
	Reverse	4D64Ins1R	CTTCTCGAAGACCCCCAAGTCAGACTGTTAGGTTTCG
StU6-4 D14	Forward	4D143Ins1F	CCAGAGCTCCCTAGGATAGGTAAATTCTGAAATATGAGGT
	Reverse	4D144Ins1R	CTTCTCGAAGACCCCCGAACTGGACTGTTAGGTTTCGC
pChimera	Forward	AllDallallDVF	GGGGTCTTCGAGAAGACCTG
	Reverse	AllDallallDVR	CCTAGGGAGCTCTGGACAAGAC

The components used for the PCR reaction, as well as the PCR cycles are reported in the Appendix in Tables A.5, A.6 and A.7. 1.5 % agarose gel electrophoresis was performed to visualize the PCR products (reported in the results, Figures 11 and 12).

A NanoDrop spectrophotometer (Thermo Scientific, USA) was used to evaluate the concentration of insert fragments using 1 µl of the PCR reaction and 1 µl of the elution buffer for blanking the instrument from NucleoSpin® and PCR clean-up kit (Macherey-Nagel™, Takara Bio, USA, Inc.) (Table 9).

2.9 Gel Extraction and Purification of Vector

In order to extract and purify the vector, fresh 0.8% agarose gel was excised around the vector borders from lanes 2 and 4 (Figure 13). Two gel slices containing 50 and 25 ng/µl linearized pChimera were retrieved and purified using the NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel™, Takara Bio, USA, Inc.) protocol with minor modifications. Incubation with NTI solution was performed at 50 ˚C for 8 minutes and elution buffer was pre-heated at 70 ˚C for 5 minutes to increase efficiency.

Elution was performed using 30 µl of the elution buffer to purify the plasmid DNA from samples 1 and 2 and labelled as primary elution. After incubation at RT for 5 minutes, the NucleoSpin® Gel and PCR Clean-up collection tubes were placed into new Eppendorf tubes and secondary elution was achieved for each sample using 30 µl of the elution buffer. After another 5-minute incubation at RT, 1.5 µl of the primary and secondary elution from each sample were mixed to bring to the total volume of 3 µl and labelled as an aliquot mixture.

After elution, a NanoDrop spectrophotometer (Thermo Scientific, USA) was used to evaluate the quantity and purity of the extracted DNA fragment in each tube (reported in the results, Table 10). For this purpose, 1 µl of the linearized plasmid DNA was used for NanoDrop quantification and 1 µl of the elution buffer from the purification step was used for blanking the instrument.

Gel electrophoresis was further performed using 1% agarose gel to verify the presence of vector in the primary and secondary elution solution, as well as the aliquot mixture from both samples (Figure 14).

2.10 In-Fusion Cloning

The cloning reaction was performed according to the In-Fusion® HD Cloning Kit user manual using 1 and 2 µl of insert (StU6-1 D1) and vector, respectively. A negative control reaction was also included using 1 µl of PCR H_2O instead of insert. The reaction details are presented in Table A.8 in the Appendix.

Fifty µl of Stellar Competent Cells were transformed by heat shock at 42 ˚C for 45 seconds using 2.5 µl of the In-Fusion and negative control reactions, and grown on a prewarmed S.O.C. medium (37 ˚C) in a shaking incubator at 37 ˚C for 1 hour. Agar plates containing LB medium and 100 mg/ml ampicillin were used to spread 5, 100 and 200 µl of the In-Fusion reaction, as well as the negative control. For the 5 µl plates, an additional amount of prewarmed S.O.C. medium was added to facilitate spreading. The plates were then transferred to an incubator to grow at 37 ˚C overnight.

The next day the bacterial growth was inspected (presented in Figure 15) and 17 putative positive colonies from 5 µl StU6-1 D1 plate were selected for colony PCR using HOT FIREPol® DNA polymerase. An area of the plate with no colony formation was also included as NTC for the PCR reaction. The reaction details and PCR cycles associated with colony PCR are presented in the Appendix in Tables A.9 and A.10, respectively. 5 µl of the PCR reaction was used to visualize bands by 1.5% agarose gel electrophoresis illustrated in Figure 16.

3. Results

3.1 PCR Amplification of the StU6 Promoter

PCR amplification of the StU6 promoter from Désirée and Asterix cultivars generated a single fragment amplicon for StU6-1 (421 bp), StU6-2 (433 bp), StU6-3 (372 bp) and 3 fragments for StU6-4 (318, 321 and 374 bp) on 1.5% agarose gel electrophoresis (Figure 7).

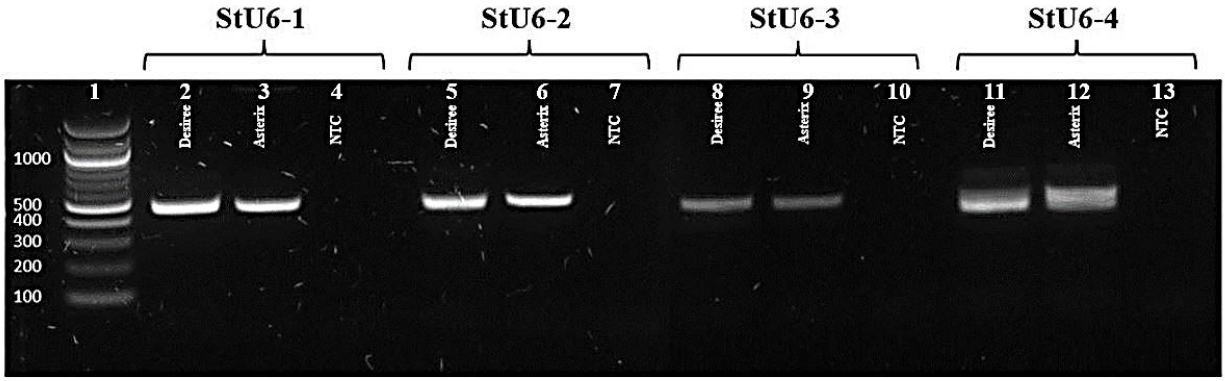

Figure 7 - PCR amplification of StU6 promoter in Désirée and Asterix potato cultivars. Lane 1: 100 bp DNA ladder, lane 2: StU6-1 in Désirée (421 bp), lane 3, StU6-1 in Asterix (421 bp), lane 4: NTC for StU6-1, lane 5: StU6-2 in Désirée (433 bp), lane 6: StU6-1 in Asterix (433 bp), lane 7: NTC for StU6-2, lane 8: StU6-3 in Désirée (372 bp), lane 9, StU6-3 in Asterix (372 bp), lane 10: NTC for StU6-3, lane 11: StU6-4 in Désirée (3 bands of 318, 321 and 374 bp), lane 12: StU6-4 in Asterix (3 bands of 318, 321 and 374 bp), lane 13: NTC for StU6-4.

3.2 Quantification of Plasmid DNA Containing StU6 Promoter

The concentration of plasmid DNA (ng/µl) in Désirée and Asterix clones evaluated by Qubit are reported in Table 5.

Table 5 - The concentration of plasmid DNA (ng/µl) evaluated by Qubit harbouring StU6 promoter in Désirée and Asterix clones.

Désirée			**Asterix**		
Promoter	**Clone**	**Concentration (ng/µl)**	**Promoter**	**Clone**	**Concentration (ng/µl)**
StU6-1	D1	221	StU6-1	A1	212
	D2	278		A2	439
StU6-2	D1	394	StU6-2	A1	386
	D2	682		A2	435
StU6-3	D1	391	-	-	-
	D2	785		-	-
StU6-4	D1	323	StU6-4	A1	323
	D2	540		A2	312
	D3	337		A3	438
	D4	369		A4	266
	D5	397		A5	264
	D6	315		A6	609
	D7	239		A7	316
	D8	307		A8	565
	D9	448		A9	194
	D10	299		A10	389
	D11	222		A11	202
	D12	348		A12	471
	D13	342		A13	243
	D14	456		A14	347
	D15	362		A15	284
	D16	534		A16	523
	D17	284		A17	230
	D18	342		A18	358
	D19	372		A19	334
	D20	316		A20	315

3.3 Sequence Analysis

The trimmed sequencing data associated with the StU6 promoter sequence in Désirée and Asterix are provided in the Appendix. The length of the promoter (bp), TSS location (bp), as well as the location and sequence associated with the TATA box and USE in Désirée and Asterix clones are presented in Tables 6 and 7, respectively.

Table 6 - The promoter length (bp), TSS location (bp), TATA box and USE location and sequence in Désirée clones.

Cultivar	**Promoter**	**Clone**	**Length (bp)**	**TSS (bp)**	**TATA Box Position (bp)**	**TATA Box Sequence**	**USE Position (bp)**	**USE Sequence**
Désirée	StU6-1	D1	476	422	392-400	TTATATAAT	359-366	CCTCATCG
		D2	476	422	392-400	TTATATAAT	359-366	CCTCATCG
	StU6-2	D1	461	407	277-382	TTATAT	344-351	CCACATCG
		D2	489	435	405-410	TTATAT	372-379	CCACATCG
	StU6-3	D1	427	373	343-348	TTATAT	310-317	CCTCATCG
		D2	427	373	343-348	TTATAT	310-317	CCTCATCG
	StU6-4	D1	376	322	293-298	TTATAT	260-267	CCTCATCG
		D2	376	322	293-298	TTATAT	260-267	CCTCATCG
		D3	376	322	293-298	TTATAT	260-267	CCTCATCG
		D4	429	375	346-351	TTATAT	313-320	CCTCATCG
		D5	376	322	293-298	TTATAT	260-267	CCTCATCG
		D6	373	319	290-295	TTATAT	257-264	CCTCATCG
		D7	376	322	293-298	TTATAT	260-267	CCTCATCG
		D8	376	322	293-298	TTATAT	260-267	CCTCATCG
		D9	376	322	293-298	TTATAT	260-267	CCTCATCG
		D10	376	322	293-298	TTATAT	260-267	CCTCATCG
		D11	376	322	293-298	TTATAT	260-267	CCTCATCG
		D12	373	319	290-295	TTATAT	257-264	CCTCATCG
		D13	429	375	346-351	TTATAT	313-320	CCTCATCG
		D14	429	375	346-351	TTATAT	313-320	CCTCATCG
		D15	376	322	293-298	TTATAT	260-267	CCTCATCG
		D16	376	322	293-298	TTATAT	260-267	CCTCATCG
		D17	376	322	293-298	TTATAT	260-267	CCTCATCG
		D18	376	322	293-298	TTATAT	260-267	CCTCATCG
		D19	376	322	293-298	TTATAT	260-267	CCTCATCG
		D20	376	322	293-298	TTATAT	260-267	CCTCATCG

Table 7 - The promoter length (bp), TSS location (bp), TATA box and USE location and sequence in Asterix clones.

Cultivar	Promoter	Clone	Length (bp)	TSS (bp)	TATA Box Position (bp)	TATA Box Sequence	USE Position (bp)	USE Sequence
Asterix	StU6-1	A1	475	421	391-399	TTATATAAT	358-365	CCTCATCG
		A2	476	422	392-400	TTATATAAT	359-366	CCTCATCG
	StU6-2	A1	488	434	404-409	TTATAT	371-378	CCACATCG
		A2	488	434	404-409	TTATAT	371-378	CCACATCG
	StU6-4	A1	376	322	293-298	TTATAT	260-267	CCTCATCG
		A2	429	375	346-351	TTATAT	313-320	CCTCATCG
		A4	429	375	346-351	TTATAT	313-320	CCTCATCG
		A5	429	375	364-351	TTATAT	313-320	CCTCATCG
		A6	429	375	364-351	TTATAT	313-320	CCTCATCG
		A8	429	375	364-351	TTATAT	313-320	CCTCATCG
		A9	376	322	293-298	TTATAT	260-267	CCTCATCG
		A10	376	322	293-298	TTATAT	260-267	CCTCATCG
		A11	429	375	346-351	TTATAT	313-320	CCTCATCG
		A12	429	375	346-351	TTATAT	313-320	CCTCATCG
		A13	376	322	293-298	TTATAT	260-267	CCTCATCG
		A14	376	322	293-298	TTATAT	260-267	CCTCATCG
		A15	429	375	346-351	TTATAT	313-320	CCTCATCG
		A16	376	322	293-298	TTATAT	260-267	CCTCATCG
		A17	429	375	346-351	TTATAT	313-320	CCTCATCG
		A18	376	322	293-298	TTATAT	260-267	CCTCATCG
		A19	429	375	346-351	TTATAT	313-320	CCTCATCG
		A20	429	373	346-351	TTATAT	313-320	CCTCATCG

Additionally, the distance of the TATA box and USE from each other and TSS associated with Désirée and Asterix clones are reported in the Appendix in Tables A.11 and A.12, respectively. In all StU6-1 clones, the TATA box and USE were located 21 and 55 bp upstream of the TSS, respectively. For StU6-2 and StU6-3 clones (in only Désirée), the TATA box and USE region were identified 24 and 55 bp upstream of the TSS, while for StU6-4, they were located at a 24 and 54 bp distance. The distance of the TATA box and USE region in all four promoters was similar (25 bp) for all clones.

These conserved motives were also observed in the *Arabidopsis thaliana* U6-1 snRNA gene (X52527.1) sequence upstream of the TSS and are demonstrated in Figure 8. TATA box and USE regions were located 23 and 54 bp upstream of the TSS with a 24 bp distance between the two. Additionally, in Z17290.1, Z17292.1 and Z17293.1 sequences, the TATA box and USE regions were 24 and 55 bp upstream of the TSS, respectively. However, this distance was shorter for Z17301.1 (23 and 54 bp for the TATA box and USE, respectively). In all the four promoters, USE was located 25 bp upstream on the TATA box (Figure 8).

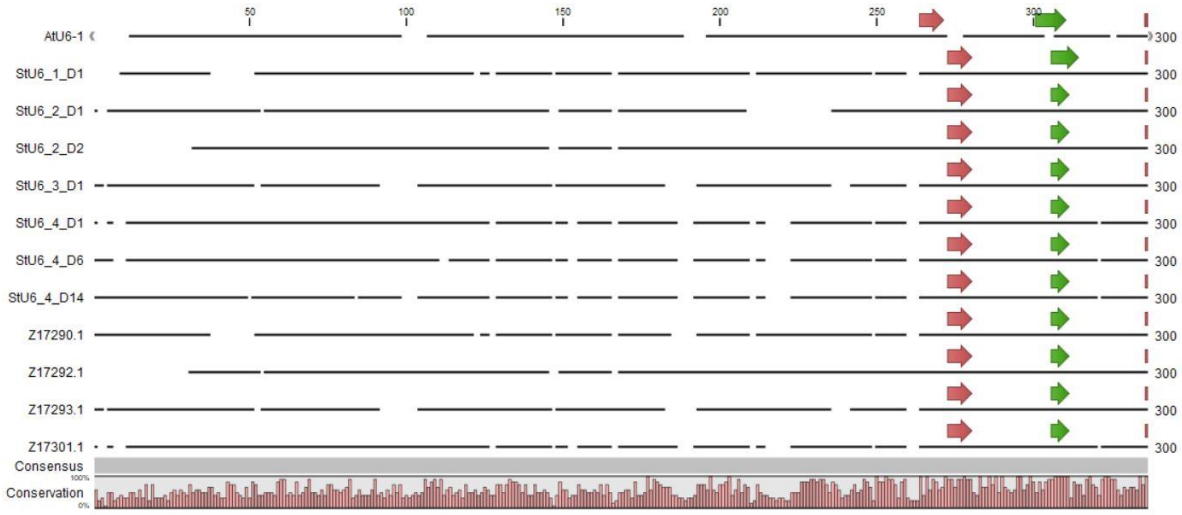

Figure 8 - The sequence alignment of the seven representative StU6 and *Arabidopsis thaliana* U6-1 snRNA gene (AtU6-1) and StU6 promoters (Z17290.1, Z17292.1, Z17293.1 and Z17301.1) from NCBI in CLCMWB demonstrating USE (red arrows), TATA box (green arrows) and TSS (red oblongs) elements.

3.4 Phylogenetic Analysis

The multiple sequence alignments of StU6 promoter sequences as the query in NCBI's genome database showed 100% identity in Désirée StU6-3 D1 and D2, as well as StU6-4 D1, D2, D3, D5, D7, D8, D9, D10, D11, D15, D16, D17, D18 and D19 (Appendix, Tables A.13). In Asterix, StU6-2 A1 and A2, StU6-4 A1,

A9, A10, A13, A14, A16 and A18 clones shared 100 % sequence identity, presented in the Appendix, Table A.14. However, the alignment score and number of gaps were different.

The phylogenetic tree based on the UPGMA algorithm revealed the presence of four different clusters each corresponding to one of the StU6 promoter sequences shown in Figure 9. The first cluster demonstrated a high degree of conservation between StU6-1 D1, StU6-1 A1, StU6-1 A2 and Z17290 sequences, while the second cluster showed similarity among StU6-2 D1, StU6-2 D2, StU6-2 A2 and Z17292. StU6-3 D1 and Z17293 formed a third cluster, whereas StU6-4 D1, StU6-4 D6, StU6-4 D14, StU6-4 A1, StU6-4 A2 and Z17301 formed the fourth cluster.

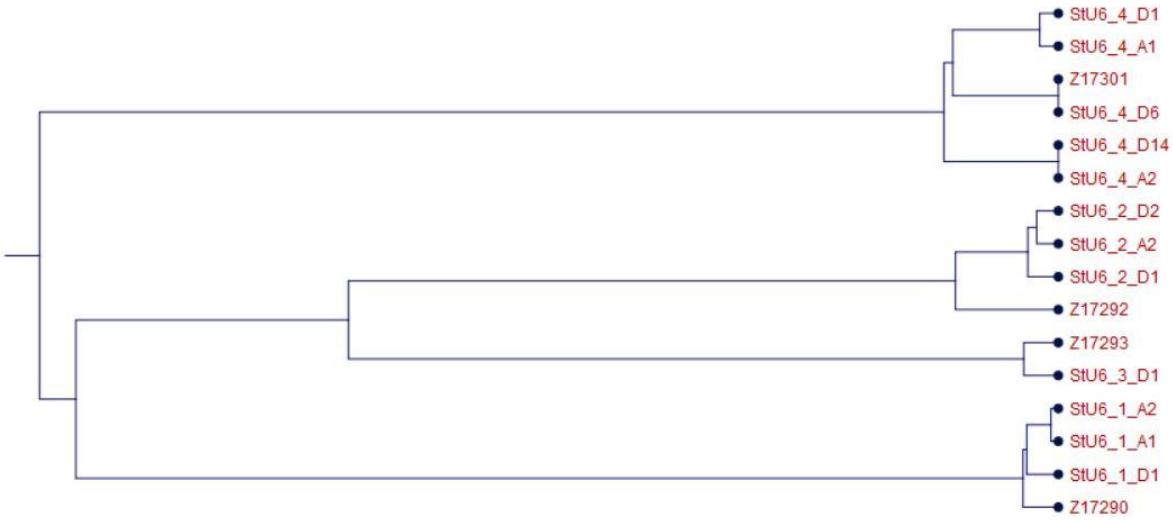

Figure 9 - The phylogenetic tree of 12 representative clones from Désirée and Asterix cultivars and 4 StU6 promoters (Z17290, Z17292, Z17293 and Z17301) demonstrating four distinct clusters based on the UPGMA method.

3.5 Quantification and Restriction Digest Analysis of pChimera

A NanoDrop spectrophotometer (Thermo Scientific, USA) was used for the evaluation of plasmid concentration. The plasmid DNA concentration, as well as 260/280 and 260/230 ratios are presented in Table 8.

Table 8 – The evaluation of pChimera, demonstrating the plasmid concentration, 260/280 and 260/230 ratios by NanoDrop spectrophotometry.

Sample	**Concentration (ng/µl)**	**260/280**	**260/230**
pChimera gDNA	265.9	1.85	2.11

2 fragments of approximately 570, 2300 bp were observed on 1.5% agarose gel as a result of treatment by PvuII restriction enzyme, shown in Figure 10.

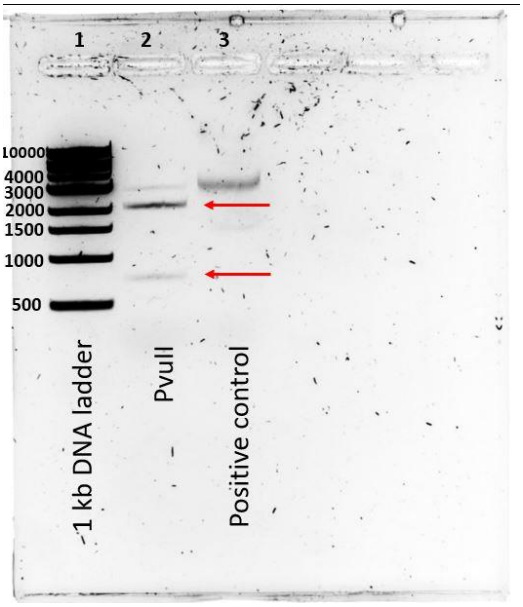

Figure 10 – pChimera restriction digest analysis using PvuII enzyme. Lane 1: 1 kb DNA ladder, lane 2: PvuII restriction digest reaction (2 fragments of 776 and 2133 bp shown by red arrows), lane 3: Uncut plasmid as positive control (1 fragment of approximately 2900 bp).

3.6 PCR Amplification of Insert and Vector

The PCR fragments associated with StU6 promoter (insert) and pChimera (vector) are presented in Figures 11 and 12, respectively.

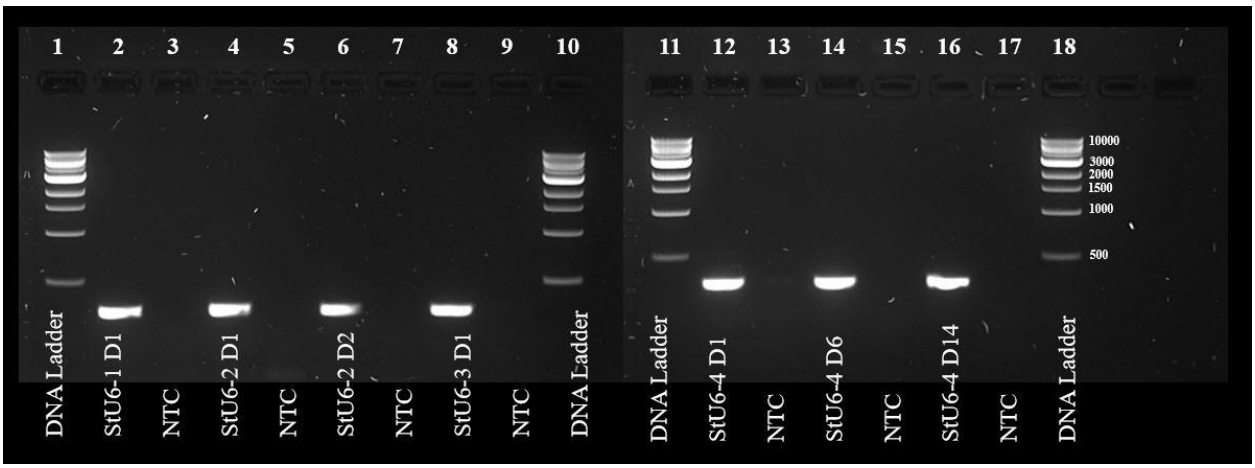

Figure 11 - PCR amplification of insert. Lane 1, 10, 11 and 18: 1 kp DNA ladder, lane 2: StU6-1 D1, lane 3: NTC for StU6-1 D1, lane 4: StU6-2 D1, lane 5: NTC for StU6-2 D1, lane 6: StU6-2 D2, lane 7: NTC for StU6-2 D2, lane 8: StU6-3 D1, lane 9: NTC for StU6-3 D1, lane 12: StU6-4 D1, lane 13: NTC for StU6-4 D1, lane 14: StU6-4 D6, lane 15: NTC for StU6-4 D6, lane 16: StU6-4 D14, lane 17: NTC for StU6-4 D14.

Figure 12 - PCR amplification of vector. Lane 1 and 4: 1 kb DNA ladder, lane 2: High intensity band of pChimera (2520 bp), lane 3: NTC.

A NanoDrop spectrophotometer (Thermo Scientific, USA) was used for the evaluation of insert fragments after the PCR reaction. DNA concentration (ng/µl), 260/280 and 260/230 ratios, as well as yield (ng) are presented in Table 9. The yield (ng) was calculated according to the below formula to determine the accurate amount of insert needed for the In-Fusion cloning reaction.

$$\text{Yield (ng)} = \text{Concentration (ng/µl)} \times \text{Volume (µl)}$$

Table 9 – The quantification of insert fragments by NanoDrop spectrophotometry, demonstrating the DNA concentration (ng/µl), 260/280 and 260/230 ratios, and yield (µg).

Sample	**Volume (µl)**	**Concentration (ng/µl)**	**260/280**	**260/230**	**Yield (µg)**
StU6-1 D1	25	498.02	1.51	0.83	12.45
StU6-2 D1	25	354.85	1.39	0.65	8.87
StU6-2 D2	25	478.43	1.47	0.67	11.96
StU6-3 D1	25	525.36	1.52	0.85	13.13
StU6-4 D1	25	392.27	1.46	0.73	9.80
StU6-4 D6	25	370.57	1.46	0.71	9.26
StU6-4 D14	25	550.88	1.59	0.89	13.77

3.7 Gel Extraction and Purification of Vector

Figure 13 illustrates high-intensity bands associated with linearized pChimera (25 and 50 ng/µl) after PCR amplification on 0.8% agarose gel.

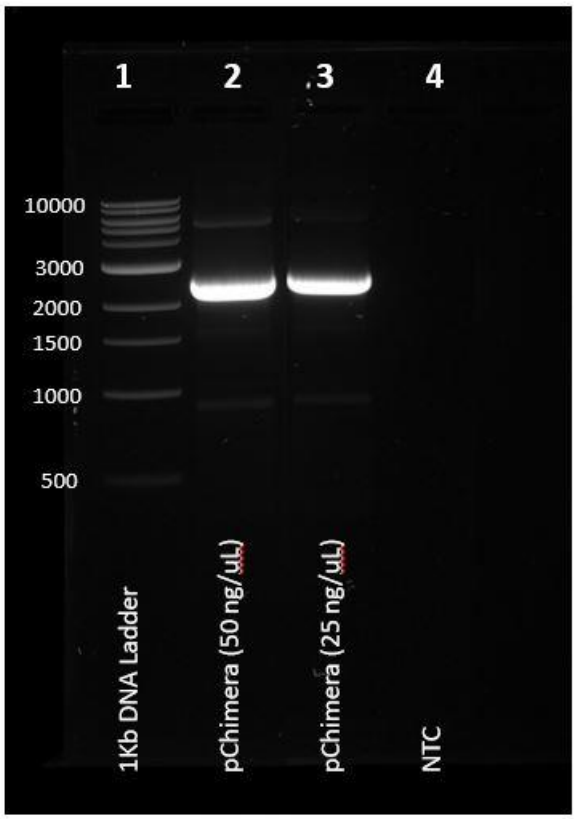

Figure 13 - Linearized vector on 0.8% agarose gel. Lane 1: 1 kb DNA Ladder, lane 2: 50 ng/µl pChimera (2520 bp), lane 3: 25 ng/µl pChimera (2520 bp), lane 5: NTC.

The quantity and purity of the extracted vector after the elution step were evaluated by NanoDrop spectrophotometer (Thermo Scientific, USA) and are presented in Table 10. The yield was highest in the primary elution of sample 1 (843.3 ng) and lowest in the aliquot mixture of sample 1 primary and secondary elution (49.62 ng).

Table 10 – The quantification of linearized pChimera vector by NanoDrop spectrophotometry, demonstrating the plasmid concentration (ng/µl), 260/280 and 260/230 ratios, and yield (ng) in primary and secondary elution, as well as the aliquot mixture of primary and secondary elution in sample 1 and 2.

Linearized pChimera	**Volume (µl)**	**Concentration (ng/µl)**	**260/280**	**260/230**	**Yield (ng)**
Sample 1 - Primary Elution	30	28.11	2.11	1.55	843.3
Sample 2 - Primary Elution	30	21.42	2.25	1.84	642.6
Sample 1 - Secondary Elution	30	15.48	2.17	0.19	464.4
Sample 2 - Secondary Elution	30	6.73	3.01	0.38	201.9
Sample 1 – Aliquot Mixture	3	19.54	2.33	0.40	49.62
Sample 2 – Aliquot Mixture	3	17.02	2.44	0.80	51.06

Figure 14 demonstrates the presence of vector in the elution solution by 1% agarose gel electrophoresis. A 2520 bp band was observed in all the lanes associated with the purified vector. The intensity of the bands corresponded to the DNA concentration and was highest in the primary elution, followed by the aliquot mixture and secondary elution.

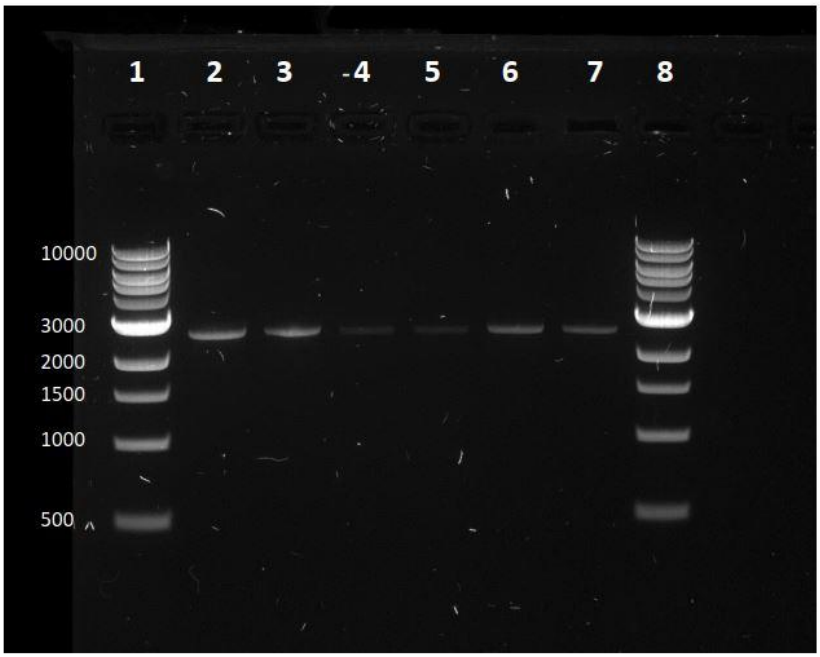

Figure 14 - Verification of the presence of linearized pChimera vector (2520 bp) in the primary and secondary elution buffer, as well as aliquot mixture. Lane 1 and 8: 1 kb DNA Ladder, lane 2: pChimera in primary elution from sample 1, lane 3: pChimera in primary elution from sample 2, lane 4: pChimera in secondary elution from sample 1, lane 5: pChimera in secondary elution from sample 2, lane 6: pChimera in the aliquot mixture from sample 1, lane 7: pChimera in the aliquot mixture from sample 2.

3.8 In-Fusion Cloning

The colony formation of the StU6-1 D1 promoter and the negative control in 5, 100 and 200 µl plates are presented in Figure 15. In all the plates, the number of bacterial colonies was almost double in StU6-1 D1 promoter plates compared to the negative controls. Moreover, the number of colonies corresponded to the volume of cells in both StU6-1 D1 and negative control plates.

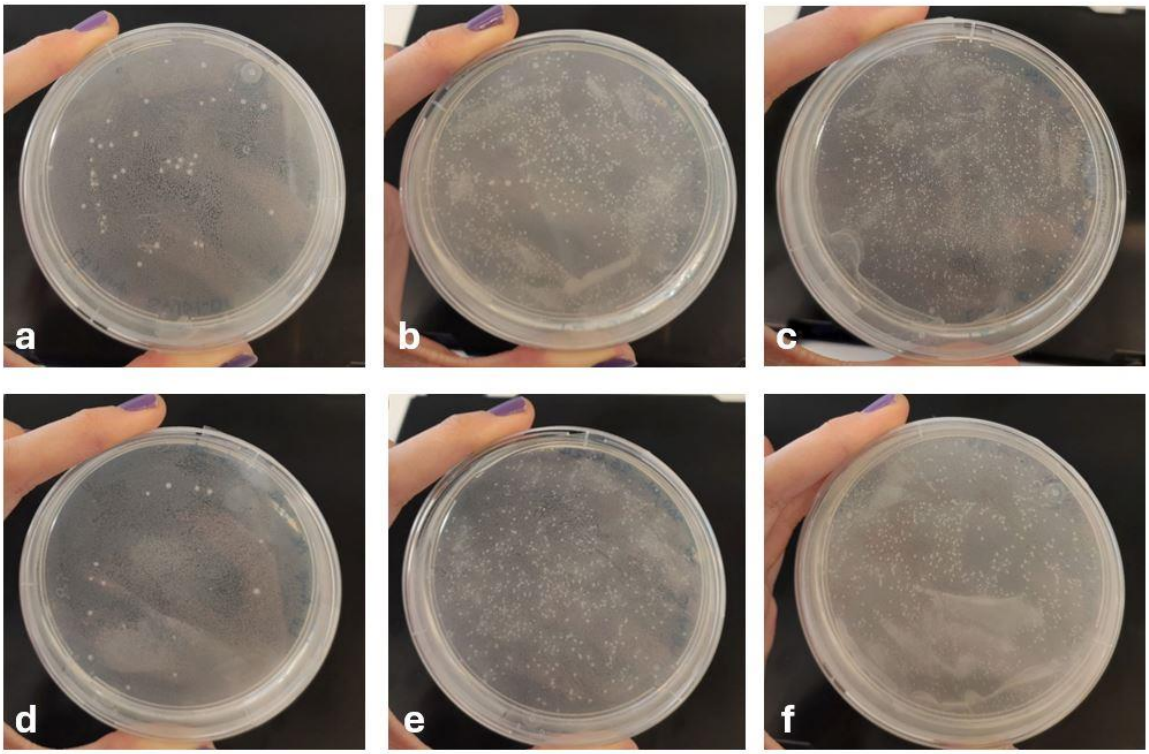

Figure 15 - Colony formation on LB agar with ampicillin. a) 5 µl StU6-1 D1, b) 100 µl StU6-1 D1, c) 200 µl StU6-1 D1, d) 5 µl negative control, e) 100 µl negative control, f) 200 µl negative control.

Figure 16 presents the colony PCR of 17 putative positive colonies from 5 µl StU6-1 D1 plates, as well as negative control. No amplification was observed in either lane.

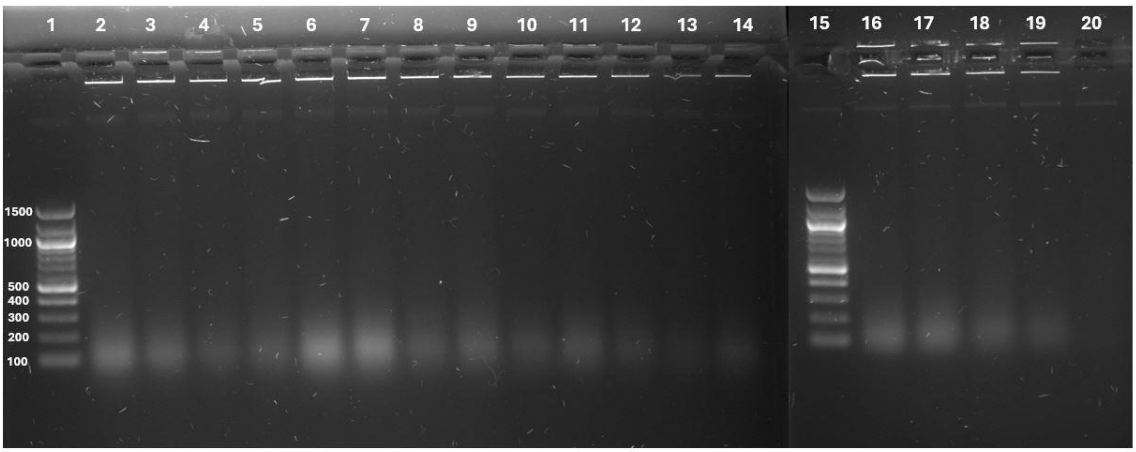

Figure 16 - Colony PCR of putative positive colonies from 5 µl StU6-1 D1 plates. Lane 1 and 15: 100 bp DNA ladder, lane 2-19: StU6-1 D1 template DNA, lane 20: NTC.

4. Discussion

4.1 PCR Amplification of the StU6 Promoter

The PCR amplification of the StU6 promoter generated one amplicon for StU6-1, StU6-2 and StU6-3 (421, 433 and 372 bp, respectively) in Désirée and Asterix cultivars, while 3 amplicons were observed for StU6-4 (318, 321 and 374 bp) in both cultivars (shown in Figure 7). This size polymorphism can be explained by the complex allelic variations in tetraploid potatoes and was previously reported in the literature, where Johansen et. al. used 4 different primer sets for the amplification of the StU6 promoter (Z17290, Z17292, Z17293, Z17301) in potato Désirée and Wotan cultivars to address this polymorphism (Johansen et al., 2019).

4.2 Cloning of the StU6 Promoter, Bacterial Transformation, and Plasmid Isolation

For StU6-4, a higher amount of OneShot Top 10 Chemically Competent Cells was used for the cloning into Zero Blunt® vector (50 µl) and 20 putative positive colonies were isolated, as opposed to the other reactions where 33 µl of cells were used for transformation and 2 colonies were isolated. This has been done to increase the chance of identifying allelic variations reported previously in StU6-4, where 3 different variants of this promoter were retrieved from Désirée and Wotan cultivars (Johansen et al., 2019).

After bacterial transformation, we were able to retrieve 2 plasmid samples from the StU6-3 clones in Désirée, but not any from Asterix. Since the same growth conditions were used for all clones, variability in bacterial culture growth conditions cannot provide a satisfactory explanation for this disparity. Instead, this can be due to low plasmid yield in Asterix StU6-3 A1 and A2 clones as a

result of incomplete lysis of bacterial cells, inefficient binding of plasmid DNA to the purification matrix, or degradation of plasmid DNA during the isolation process.

The concentration of the plasmid DNA containing StU6 promoter ranged from 194 to 785 ng/µl in Désirée and Asterix clones indicating good yield and an efficient method of plasmid isolation using PureYield™ Plasmid Miniprep System (A1222, Promega, USA).

4.3 Sequence Analysis

In all Désirée and Asterix clones, TATA box and USE conserved motives were identified upstream of the TSS associated with StU6 promoter sequences at NCBI showing a high degree of sequence identity. The length of the promoter sequence was similar in StU6-1 and StU6-3 clones from Désirée (476 and 427 bp, respectively), but different in StU6-2 (461 and 489) and StU6-4 clones (373, 376 and 429 bp) shown in Table 6. This variation in promoter length was also present among StU6-1 (475 and 476 bp) and StU6-4 (376 and 429 bp) Asterix clones, but not for StU6-2 (488 bp), reported in Table 7.

These length differences are associated with indels (insertions and deletions) in single nucleotides in polyploid potato cultivars that can affect paralogous and homologous loci. While homologous alleles share a common ancestry, retain similar sequences and perform related functions, paralogous alleles are within the same species that arose from gene duplication events. These duplicated genes can be found either on separate chromosomes (segmental duplication) or on the same chromosome (tandem duplication) (Fitch, 1970; Holland, 1999; Patterson, 1988).

Paralogous genes can also exhibit functional divergence, which results in the acquisition of new activities or the specialization of already-existing ones, despite sharing similar sequences and performing comparable functions. This can lead to the expansion of gene families and the generation of genetic diversity within a species (Li et al., 2019).

Additionally, the location of the TSS, TATA box and USE varied among the clones. Figure 17 presents the forward and reverse primers, USE, TATA box and TSS (+1) associated with the Désirée StU6-4 D1 clone in the CLCMWB as an example.

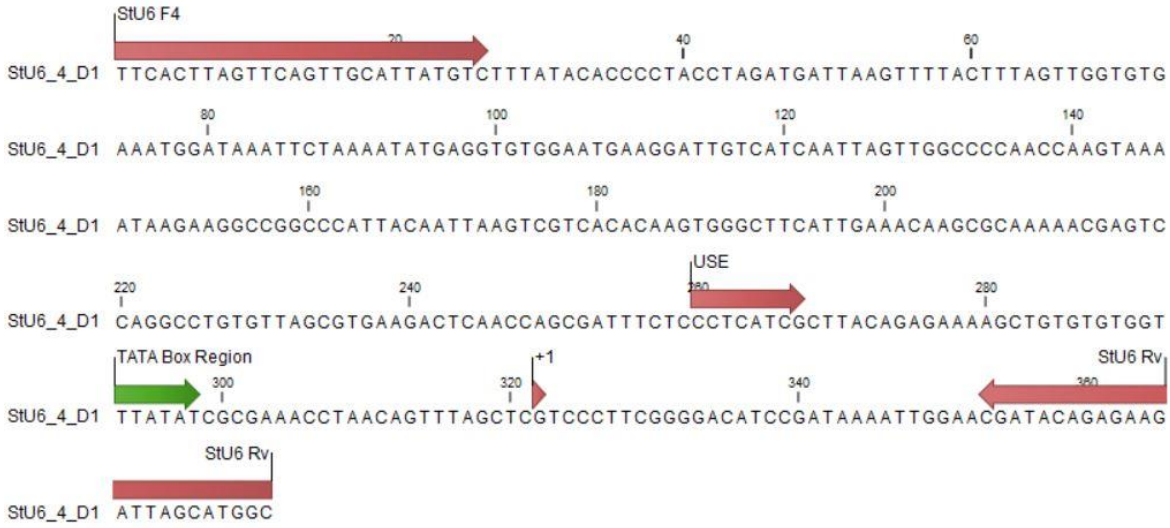

Figure 17 – The forward and reverse primers (StU6 F4 and StU6 Rv, respectively), USE, TATA box and TSS (+1) in Désirée StU6-4 D1 in CLCMWB.

The distance of the TATA box and USE regions from the TSS are reported in the Appendix in Tables A.11 and A.12 for Désirée and Asterix clones, respectively. In all the 48 clones studied from Désirée and Asterix, the TATA box was located 25 bp downstream of the USE. The distance of the TATA box

region from the TSS varied from 21-24 bp, while this gap was 54-55 bp for USE. These results were similar in the *Arabidopsis thaliana* U6-1 snRNA gene and StU6 promoter sequences from NCBI (Figure 8).

In a study by Kanai et. al in castor (*Ricinus communis*), the same interval between the USE and TATA box region (25 bp) was reported, but in their study, the TATA box was located 28 bp upstream of the TSS demonstrated in Figure 2 (Kanai et al., 2023).

The consensus TATA box sequence shows high similarity but different lengths among the clones in our study. While the TTATAT sequence was found in most clones, the TTATATAAT sequence was observed in StU6-1 Désirée and Asterix clones. Furthermore, the consensus USE sequence of CCTCATCG was identified for StU6-1, StU6-3 and StU6-4 clones, while StU6-2 Désirée and Asterix clones had T to A base transversion mutation.

This point mutation occurs on the DNA strand when a single or two-ring purine (A or G) is changed to a single-ring pyrimidine (T or C), or vice versa. Although transversion might occur spontaneously during cellular exogenous oxidative stress, ionizing radiation or certain alkylating chemicals may catalyse this conversion (Friedberg et al., 2005; Kino & Sugiyama, 2001).

The conserved TATA box and USE sequences have been previously investigated in other plant promoters, suggesting an important functional role of these elements in plant gene expression (Kanai et al., 2023). Therefore, we studied the location and sequence of the TATA box and USE motives to confirm the presence of the promoter in Désirée and Asterix clones.

4.4 Phylogenetic Analysis

The phylogenetic analysis of the StU6 promoter was performed using the UPGMA algorithm and revealed a high degree of similarity between Désirée and Asterix clones. However, the differences in the promoter sequence and length of the four promoters lead to the formation of four distinguished clusters on the phylogenetic tree shown in Figure 9. Each of these clusters corresponded to one of the StU6 promoters and served as a confirmation for the promoter identity in Désirée and Asterix clones.

The phylogenetic tree demonstrates the differences and similarities between different organisms during evolution (Sharma et al., 2018). One of the commonly used methods of phylogenetic analysis is the UPGMA algorithm, which is based on clustering the two nearest clusters, and then computing the joint pair's distance by taking the average (Bhambri & Gupta, 2012). This method is relatively fast and sensitive and performs well when the variation of branch length's coefficient is low (Higgins & Sharp, 1989; Tateno et al., 1982).

However, this algorithm directly scales the branch lengths and clusters them based on genetic distances. Therefore, it is most appropriate for datasets that evolved at reasonably stable rates (Liesebach & Sinkó, 2008). Moreover, this method is sensitive to input order and inaccurate tree construction can result from its assumption of a constant rate of evolution, which is not necessarily the case, particularly for datasets where there are notable changes in the rates of evolution among lineages (Podani, 1997).

Overall, while UPGMA has its advantages in terms of simplicity and speed, its limitation lies in handling complex evolutionary scenarios. Moreover, its sensitivity to data input makes it less suitable for certain phylogenetic analyses compared to more sophisticated methods.

4.5 Quantification and Restriction Digest Analysis of pChimera

The resonance structures of pyrimidine and purine bases present in the structure of nucleic acids are responsible for absorption peaks in the UV spectrum's 260–280 nm (Wilfinger et al., 1997). Therefore, the purity of DNA fragments can be determined by comparing the absorbance maxima (260 nm) to the absorbance at 280 nm. It is commonly acknowledged that a 260/280 ratio of approximately 1.8 indicates good-quality DNA and a lower ratio can arise from the presence of residual phenol, guanidine, or other reagents present in the tube (Brian, 2015).

A secondary indicator of nucleic acid purity is the 260/230 values, which are frequently greater than the corresponding 260/280 values. Expected 260/230 values for DNA are typically between 2.0 - 2.2 and a lower ratio indicates the presence of contaminants that absorb light at 230 nm, such as proteins, phenols, guanidine, EDTA, Tris buffer and carbohydrates (Brian, 2015).

In this regard, the high purity of pChimera plasmid was elucidated by a 260/280 ratio of 1.85 and a 260/230 ratio of 2.11, respectively (Table 8). Furthermore, the identity of pChimera was confirmed by the visualization of 776 and 2133 bp fragments on agarose gel after treatment with PvuII restriction enzyme, which performs two cuts on the plasmid backbone at 222 and 998 bp locations.

However, a low-intensity band of 2909 kb was also observed which shows the restriction digest experiment was not 100% efficient and a small proportion of the plasmids were not cut (Lane 2, Figure 10).

Several bands can be produced by uncut circular plasmid DNA on gel electrophoresis, which correlates to distinct plasmid forms, such as the supercoiled and the nicked (relaxed circular) form (Swanson, n.d.). The supercoiled form is the most compact form of the plasmid, which moves fastest through the gel, appearing as the lowest band, while the nicked form occurs when

one strand of the plasmid is nicked, relaxing the supercoiling. It migrates more slowly than the supercoiled form, appearing as a higher band.

Although we expected not to see any bands in lane 3 for the uncut pChimera as a positive control (Figure 10), we observed a 2909 bp band, which was likely associated with the nicked form of the relaxed circular plasmid DNA. A faint band of approximately 3000 bp was also observed in the same lane associated with the supercoiled form of plasmid DNA.

4.6 Insert and Vector Amplification

The presence of a single fragment amplicon on gel electrophoresis after PCR reaction indicated the presence of insert and vector demonstrated in Figures 11 and 12. However, a 100 bp DNA ladder was not available in the laboratory to determine the fragment size and a 1 kb DNA ladder was used instead. Since we used specific primers for the amplification of the insert, the single band observed on the gel was believed to be the insert.

The concentration of the seven StU6 promoters after PCR amplification ranged from 354.85 to 550.88 ng/µl with low 260/280 and 260/230 ratios (1.39 - 1.59 and 0.65 – 0.89, respectively) and are shown in Table 9. These high values can be explained by the presence of PCR reagents in the tube that was directly used for NanoDrop quantification.

The insert yield was considerably high ranging from 8.87 µg at the lowest and 13.77 µg at the highest for StU6-2 D1 and StU6-4 D14, respectively. Since specific primers were used for the amplification of each promoter, the size differences among them cannot provide a good explanation for this difference.

However, it can be explained by the efficiency of the PCR reaction to generate more amplicons for the latter.

4.7 Vector Purification and Quantification

An important feature of pChimera vector is the presence of an AtU6 promoter (389 bp) adjacent to the sgRNA scaffold shown in Figure 4, which makes it applicable for use in many CRISPR-Cas9 gene editing experiments (Shimatani et al., 2017, 2019).

In this study, the AtU6 promoter sequence was replaced by endogenous potato StU6 promoter by PCR. For this purpose, specific primers were designed to amplify pChimera excluding the AtU6 promoter sequence, with 15-bp overhangs complementary to the insert sequence. However, both the original vector and the PCR amplified pChimera contain an ampicillin resistance gene as a selection marker and grow on ampicillin plates, which made it necessary for an additional purification step to separate them.

Therefore, aliquots of the PCR reaction were first visualized on 0.8% agarose gel recommended by the kit manual (NucleoSpin® Gel and PCR Clean-up, Macherey-Nagel™, Takara Bio, USA, Inc.) and the corresponding bands containing the linearized vector were sliced and further subjected to purification (Figure 13). This allowed for a precise separation of the linearized vector from the original vector present in the tube reaction.

The concentration of the purified vector was higher in the primary elution (28.11 and 21.42 ng/µl for samples 1 and 2, respectively) compared to the secondary elution (15.48 and 6.73 ng/µl for samples 1 and 2, respectively) reported in Table 10. The purity of the vector was high in all the samples ranging from 2.11 to 3.01

associated with sample 1 primary elution and sample 2 secondary elution, respectively.

The yield was the highest in the primary elution of sample 1 (843.3 ng) followed by sample 2 primary elution (642.6 ng). Therefore, the first was used for the In-Fusion cloning reaction to secure enough vector DNA for the cloning reaction. The single band fragment observed in gel electrophoresis after the elution further served as confirmation for the presence of vector (Figure 14).

4.8 In-Fusion Cloning

Bacterial colony formation on antibiotic agar after In-Fusion cloning is a commonly used method for the selection and isolation of putative-positive bacterial cells. pChimera vector contains an ampicillin resistance gene, which allows transformed cells to survive in the presence of this antibiotic.

However, it was not a clear confirmation that the cloning reaction was successful since no positive control reaction was included. A positive control sample for the colony PCR was not possible since we lacked the primer pair for the pChimera vector backbone that flanks the insertion site. For this purpose, a colony of the pChimera-containing strain on a selection plate would have been required to rule out setup issues with the PCR reaction.

Furthermore, no bands were observed from the colony PCR of the putative positive colonies which could be due to either a failed cloning reaction or technical issues during the colony PCR (Figure 15). Generally, the use of a suboptimal insert-to-vector ratio, ineffective primer design or low-quality DNA templates can be the reasons for In-Fusion reaction failure. Improper concentrations of insert DNA can lead to inefficient or incomplete incorporation

of the insert into the vector. This can increase the likelihood of self-ligation of the vector fragments with blunt ends due to nuclease activity in the reaction, leading to the formation of background colonies containing unmodified vectors. These transformed cells do not have the insert but are still able to grow on ampicillin plates since the vector contains an ampicillin resistance gene in its backbone.

The ratio of insert and vector used for the cloning reaction was 8:1 (498 and 56.2 ng/μl, respectively). However, the evaluation of the insert concentration directly from the PCR reaction was not reliable because the reaction mixture contained other contaminants leading to the overestimation of DNA concentration. The presence of dNTPs can potentially result in a false increase in DNA concentration because they absorb light at the wavelengths typically used in spectrophotometry (260 nm). Additionally, secondary structures within the DNA sample, such as hairpins or primer dimers, might potentially affect the accuracy of DNA concentration measurements by NanoDrop.

While evaluating DNA concentration directly from a PCR reaction can provide a quick estimate of DNA yield, it is generally recommended to purify the PCR product before quantification to remove contaminants and ensure accurate measurements.

Moreover, the quality and purity of the DNA fragments are crucial for successful cloning and impurities can hinder the cloning reaction. For the In-Fusion cloning reaction, we used the insert fragment directly from the PCR tube that also contained Phusion HF buffer (containing $MgCl_2$) and dNTPs, which in high concentrations might interfere with In-Fusion enzyme activity.

Additionally, low transformation efficiency can lead to a failure to obtain colonies on selective agar plates even if the cloning reaction itself is successful. This can be due to factors such as improper handling of competent cells,

inefficient uptake of DNA by the cells, or poor growth conditions post-transformation. However, this was not the case in our study since the transformed cells grew on ampicillin agar in both StU6-1 D1 and negative control reactions. Moreover, the number of transformed cells was almost twice from the StU6-1 D1 cloning reaction compared to the negative control which could reflect inaccurate pipetting.

Ultimately, the mutations created by error by DNA polymerase during the PCR reaction can affect the outcome of the cloning reaction. This can create incompatible insert and vector ends that might prevent efficient ligation. However, it was unlikely in our experiment because we used high-fidelity Phusion Hot Start II DNA Polymerase with a low error rate (4.4×10^{-7} in Phusion HF buffer).

Technical errors during colony PCR can arise from ineffective annealing of the primers to the template DNA. Since the same primer set employed for the amplification of StU6-1 D1 was used at this step (1D13Ins1F and 1D14Ins1R, reported in Table 4) and generated PCR products earlier (shown in Figure 11), the mismatch between the primers and the target DNA is highly unlikely to explain unsuccessful amplification.

While the implementation of inadequate bacterial colonies can result in low PCR amplification efficiency, excessive amounts can prevent amplification. However, since we used 17 individual colonies as the template for colony PCR and neither of them generated bands, a suboptimal number of bacterial cells used for colony PCR was also unlikely to explain no band amplification.

In total, the use of a too high insert: vector ratio for the In-Fusion cloning reaction or suboptimal PCR conditions can be a potential explanation for why were did not get any band amplification. For future research, it is suggested that the cloning reaction be repeated using an adjusted amount of insert and vector (2:1

or 3:1 according to the user manual). However, it is necessary to first purify insert fragments from the PCR reagents in order to achieve a clear estimation of insert concentration. After confirmation by colony PCR, plasmid isolation from positive colonies should be performed to purify the vector for protoplast transfection. Additionally, restriction digestion analysis of the final construct and sequencing can be utilized for the final verification of the construct.

5. Conclusion

CRISPR-Cas9 technology can potentially be used as a tool for improving many potato traits through precision genome editing. This technology involves using a CRISPR-associated endonuclease, Cas9, in conjunction with a single guide RNA (sgRNA) complementary to a specified target area to perform double-stranded cleavage on DNA strands.

In the CRISPR-Cas9 system, the sgRNA scaffold is usually engineered as a single RNA molecule that combines the targeting sequence with the tracrRNA scaffold. The sgRNA molecule can be synthesized in vitro or generated using expression plasmids in cells. Once formed, the sgRNA-Cas9 complex binds to the target DNA sequence through base-pairing between the targeting sequence and the complementary DNA sequence. The Cas9 protein then induces a DSB at the target site, which can be repaired by cellular DNA repair mechanisms to introduce specific genetic modifications, such as insertions, deletions, or substitutions.

The StU6 promoter is a crucial regulatory component located upstream of the TSS and plays an important role in the transcription of the snRNA gene. Therefore, determining the efficiency of endogenous potato StU6 promoter for gene expression is essential to improve CRISPR-Cas gene editing output.

This project aimed to create constructs utilizing a U6 promoter that has been PCR amplified from potato (*solanum tuberosum*) Désirée and Asterix cultivars and to clone them into a plasmid vector (pChimera) upstream of a non-coding sgRNA scaffold sequence. pChimera is conventionally used for plant gene editing and already contains the AtU6 promoter from *Arabidopsis thaliana* upstream of the sgRNA scaffold. Removal of the AtU6 promoter sequence from pChimera by PCR facilitates the cloning of each of the StU6 promoters by In-Fusion cloning. In-Fusion cloning was used to make the cloning vector using PCR-amplified

StU6 promoter and linearized pChimera (excluding AtU6 promoter sequence) as insert and vector, respectively. The selection and isolation of putative positive bacterial cells is often accomplished by observing the growth of bacterial colonies on ampicillin agar subsequent to In-Fusion cloning. The ampicillin resistance gene included in the pChimera backbone enables transformed cells to endure the presence of ampicillin.

Nevertheless, the absence of a positive control reaction made it difficult to conclude with certainty that the cloning reaction was successful. In addition, no bands were seen from the colony PCR of the presumed positive colonies, which could indicate a cloning reaction failure or PCR technical difficulties.

Therefore, it is strongly suggested that the In-Fusion cloning reaction be repeated using an adjusted ratio of insert to vector. Additionally, it is advised to purify insert fragments from the PCR reaction and quantify DNA concentration prior to the cloning reaction.

The future objective of this project is to study the expression of sgRNA by RT-PCR to determine which homologous StU6 promoter from potato will provide the strongest level of transcription. However, additional restriction digestion analysis and sanger sequencing of the final construct are needed to ensure that we have the desired construct for protoplast transfection.

6. References

Andersson, M., Turesson, H., Olsson, N., Fält, A., Ohlsson, P., Gonzalez, M. N., Samuelsson, M., & Hofvander, P. (2018). Genome editing in potato via CRISPR-Cas9 ribonucleoprotein delivery. *Physiologia Plantarum*, *164*(4), 378–384. https://doi.org/10.1111/ppl.12731

Anzalone, A. V., Koblan, L. W., & Liu, D. R. (2020). Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. *Nature Biotechnology*, *38*(7), 824–844.

Barrangou, R. (2015). The roles of CRISPR–Cas systems in adaptive immunity and beyond. *Current Opinion in Immunology*, *32*, 36–41.

Bártová, V., Bárta, J., Brabcová, A., Zdráhal, Z., & Horáčková, V. (2015). Amino acid composition and nutritional value of four cultivated South American potato species. *Journal of Food Composition and Analysis*, *40*, 78–85.

Bhambri, P., & Gupta, O. P. (2012). A novel method for the design of phylogenetic tree. *Int J IT Eng Appl Sci Res*, *1*(1), 24–28.

Bolotin, A., Quinquis, B., Sorokin, A., & Ehrlich, S. D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, *151*(8), 2551–2561. https://doi.org/10.1099/mic.0.28048-0

Bradshaw, J. E. (2019). Improving the Nutritional Value of Potatoes by Conventional Breeding and Genetic Modification. In A. M. I. Qureshi, Z. A. Dar, & S. H. Wani (Eds.), *Quality Breeding in Field Crops* (pp. 41–84). Springer International Publishing. https://doi.org/10.1007/978-3-030-04609-5_3

Brian, M. (2015). Assessment of Nucleic Acid Purity. *Wilmington, MA, USA: Thermo Fisher Scientific*.

Burgos, G., Amoros, W., Muñoa, L., Sosa, P., Cayhualla, E., Sanchez, C., Díaz, C., & Bonierbale, M. (2013). Total phenolic, total anthocyanin and phenolic acid concentrations and antioxidant activity of purple-fleshed potatoes as affected by boiling. *Journal of Food Composition and Analysis*, *30*(1), 6–12.

Burgos, G., Muñoa, L., Sosa, P., Bonierbale, M., Zum Felde, T., & Díaz, C. (2013). In vitro Bioaccessibility of Lutein and Zeaxanthin of Yellow Fleshed Boiled Potatoes. *Plant Foods for Human Nutrition*, *68*(4), 385–390. https://doi.org/10.1007/s11130-013-0381-x

Camire, M. E., Kubow, S., & Donnelly, D. J. (2009). Potatoes and Human Health. *Critical Reviews in Food Science and Nutrition*, *49*(10), 823–840. https://doi.org/10.1080/10408390903041996

Chakraborty, S., Chakraborty, N., & Datta, A. (2000). Increased nutritive value of transgenic potato by expressing a nonallergenic seed albumin gene from *Amaranthus hypochondriacus*. *Proceedings of the National Academy of Sciences*, *97*(7), 3724–3729. https://doi.org/10.1073/pnas.97.7.3724

Cho, S. W., Lee, J., Carroll, D., Kim, J.-S., & Lee, J. (2013). Heritable gene knockout in Caenorhabditis elegans by direct injection of Cas9–sgRNA ribonucleoproteins. *Genetics*, *195*(3), 1177–1180.

Dahlberg, J. E., & Lund, E. (1988). *In Birnstiel, ML (ed.) Structure and Function ofMajor and Minor Small Nuclear Ribonucleoprotein Particles*. Springer Verlag, Berlin.

Devaux, A., Goffart, J.-P., Petsakos, A., Kromann, P., Gatto, M., Okello, J., Suarez, V., & Hareau, G. (2020). Global food security, contributions from sustainable potato agri-food systems. *The Potato Crop: Its Agricultural, Nutritional and Social Contribution to Humankind*, 3–35.

Fairweather-Tait, S. J. (1983). Studies on the availability of iron in potatoes. *British Journal of Nutrition*, *50*(1), 15–23.

Faostat, F. A. O. (2017). Available online: Http://www. fao. org/faostat/en/# data. *QC (Accessed on January 2018)*.

Fitch, W. M. (1970). Distinguishing homologous from analogous proteins. *Systematic Zoology*, *19*(2), 99–113.

Friedberg, E. C., Walker, G. C., Siede, W., & Wood, R. D. (2005). *DNA repair and mutagenesis*. American Society for Microbiology Press. https://books.google.com/books?hl=ru&lr=&id=VAKsBAAAQBAJ&oi=fnd&pg=PT41&dq=E.C.+Friedberg,+G.C.+Walker,+W.+Siede,+DNA+Repair+and+Mutagenesis,+ASM,+Washington,+DC,+1995.+&ots=2wCp_SMZUi&sig=VK_x1bK_4AWHMyRBd5JYJNJu4nM

Friedman, M. (1996). Nutritional Value of Proteins from Different Food Sources. A Review. *Journal of Agricultural and Food Chemistry*, *44*(1), 6–29. https://doi.org/10.1021/jf9400167

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., & Sander, J. D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature Biotechnology*, *31*(9), 822–826.

Goo, Y.-M., Kim, T.-W., Lee, M.-K., & Lee, S.-W. (2013). Accumulation of PrLeg, a perilla legumin protein in potato tuber results in enhanced level of sulphur-containing amino acids. *Comptes Rendus. Biologies*, *336*(9), 433–439.

Guerineau, F., & Waugh, R. (1993). The U6 small nuclear RNA gene family of potato. *Plant Molecular Biology*, *22*(5), 807–818. https://doi.org/10.1007/BF00027367

Halford, N. G., Raffan, S., & Oddy, J. (2022). Progress towards the production of potatoes and cereals with low acrylamide-forming potential. *Current Opinion in Food Science*, *47*, 100887.

Haverkort, A. J., Struik, P. C., Visser, R. G. F., & Jacobsen, E. (2009). Applied Biotechnology to Combat Late Blight in Potato Caused by Phytophthora Infestans. *Potato Research*, *52*(3), 249–264. https://doi.org/10.1007/s11540-009-9136-3

He, X., Wang, Y., Yang, F., Wang, B., Xie, H., Gu, L., Zhao, T., Liu, X., Zhang, D., & Ren, Q. (2019). Boosting activity of high-fidelity CRISPR/Cas9 variants using a tRNAGln-processing system in human cells. *Journal of Biological Chemistry*, *294*(23), 9308–9315.

Higgins, D. G., & Sharp, P. M. (1989). Fast and sensitive multiple sequence alignments on a microcomputer. *Bioinformatics*, *5*(2), 151–153.

Holland, P. W. (1999). Gene duplication: Past, present and future. *Seminars in Cell & Developmental Biology*, *10*(5), 541–547. https://www.sciencedirect.com/science/article/pii/S108495219990335X

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, *337*(6096), 816–821. https://doi.org/10.1126/science.1225829

Johansen, I. E., Liu, Y., Jørgensen, B., Bennett, E. P., Andreasson, E., Nielsen, K. L., Blennow, A., & Petersen, B. L. (2019). High efficacy full allelic CRISPR/Cas9 gene editing in tetraploid potato. *Scientific Reports*, *9*(1), 17715.

Kanai, M., Hikino, K., & Mano, S. (2023). Cloning and Functional Verification of Endogenous U6 Promoters for the Establishment of Efficient CRISPR/Cas9-Based Genome Editing in Castor (Ricinus communis). *Genes*, *14*(7), 1327.

Kieu, N. P., Lenman, M., Wang, E. S., Petersen, B. L., & Andreasson, E. (2021). Mutations introduced in susceptibility genes through CRISPR/Cas9 genome editing confer increased late blight resistance in potatoes. *Scientific Reports*, *11*(1), 4487.

Kim, S., Kim, D., Cho, S. W., Kim, J., & Kim, J.-S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Research*, *24*(6), 1012–1019.

Kino, K., & Sugiyama, H. (2001). Possible cause of G· C→ C· G transversion mutation by guanine oxidation product, imidazolone. *Chemistry & Biology*, *8*(4), 369–378.

Kor, S. D., Chowdhury, N., Keot, A. K., Yogendra, K., Chikkaputtaiah, C., & Sudhakar Reddy, P. (2023). RNA Pol III promoters—Key players in precisely targeted plant genome editing. *Frontiers in Genetics*, *13*, 989199.

Krits, P., Fogelman, E., & Ginzberg, I. (2007). Potato steroidal glycoalkaloid levels and the expression of key isoprenoid metabolic genes. *Planta*, *227*(1), 143–150. https://doi.org/10.1007/s00425-007-0602-3

Li, Y., Zhu, J., Wu, L., Shao, Y., Wu, Y., & Mao, C. (2019). Functional divergence of PIN1 paralogous genes in rice. *Plant and Cell Physiology*, *60*(12), 2720–2732.

Liang, Z., Chen, K., Li, T., Zhang, Y., Wang, Y., Zhao, Q., Liu, J., Zhang, H., Liu, C., & Ran, Y. (2017). Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. *Nature Communications*, *8*(1), 14261.

Liesebach, H., & Sinkó, Z. (2008). A contribution to the systematics of the genus Tilia with respect to some hybrids by RAPD analysis. *Dendrobiology*, *59*, 13–22.

Liu, M., Zhang, W., Xin, C., Yin, J., Shang, Y., Ai, C., Li, J., Meng, F.-L., & Hu, J. (2021). Global detection of DNA repair outcomes induced by CRISPR–Cas9. *Nucleic Acids Research*, *49*(15), 8732–8742.

Ly, D. N. P., Iqbal, S., Fosu-Nyarko, J., Milroy, S., & Jones, M. G. (2023). Multiplex CRISPR-Cas9 gene-editing can deliver potato cultivars with reduced browning and acrylamide. *Plants*, *12*(2), 379.

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., & Church, G. M. (2013). RNA-Guided Human Genome Engineering via Cas9. *Science*, *339*(6121), 823–826. https://doi.org/10.1126/science.1232033

Mojica, F. J. M., D�ez-Villase�or, C., Garc�a-Mart�nez, J., & Soria, E. (2005). Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *Journal of Molecular Evolution*, *60*(2), 174–182. https://doi.org/10.1007/s00239-004-0046-3

Norway: Production volume of potatoes 2022 | Statista. (n.d.). Retrieved May 1, 2024, from https://www.statista.com/statistics/644073/annual-production-volume-of-potatoes-in-norway/

Patterson, C. (1988). Homology in classical and molecular biology. *Molecular Biology and Evolution*, *5*(6), 603–625.

Podani, J. (1997). On the sensitivity of ordination and classification methods to variation in the input order of data. *Journal of Vegetation Science*, *8*(1), 153–156. https://doi.org/10.2307/3237253

Raina, A., & Datta, A. (1992). Molecular cloning of a gene encoding a seed-specific protein with nutritionally balanced amino acid composition from Amaranthus. *Proceedings of the National Academy of Sciences*, *89*(24), 11774–11778. https://doi.org/10.1073/pnas.89.24.11774

Rose, J. C., Popp, N. A., Richardson, C. D., Stephany, J. J., Mathieu, J., Wei, C. T., Corn, J. E., Maly, D. J., & Fowler, D. M. (2020). Suppression of unwanted CRISPR-Cas9

editing by co-administration of catalytically inactivating truncated guide RNAs. *Nature Communications*, *11*(1), 2697.

Sawai, S., Ohyama, K., Yasumoto, S., Seki, H., Sakuma, T., Yamamoto, T., Takebayashi, Y., Kojima, M., Sakakibara, H., & Aoki, T. (2014). Sterol side chain reductase 2 is a key enzyme in the biosynthesis of cholesterol, the common precursor of toxic steroidal glycoalkaloids in potato. *The Plant Cell*, *26*(9), 3763–3774.

Sharma, A., Jaloree, S., & Thakur, R. S. (2018). Review of Clustering Methods: Toward Phylogenetic Tree Constructions. In B. Tiwari, V. Tiwari, K. C. Das, D. K. Mishra, & J. C. Bansal (Eds.), *Proceedings of International Conference on Recent Advancement on Computer and Communication* (Vol. 34, pp. 475–480). Springer Singapore. https://doi.org/10.1007/978-981-10-8198-9_50

Shimatani, Z., Ariizumi, T., Fujikura, U., Kondo, A., Ezura, H., & Nishida, K. (2019). Targeted Base Editing with CRISPR-Deaminase in Tomato. In Y. Qi (Ed.), *Plant Genome Editing with CRISPR Systems* (Vol. 1917, pp. 297–307). Springer New York. https://doi.org/10.1007/978-1-4939-8991-1_22

Shimatani, Z., Kashojiya, S., Takayama, M., Terada, R., Arazoe, T., Ishii, H., Teramura, H., Yamamoto, T., Komatsu, H., & Miura, K. (2017). Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. *Nature Biotechnology*, *35*(5), 441–443.

Starr, D. B., & Hawley, D. K. (1991). TFIID binds in the minor groove of the TATA box. *Cell*, *67*(6), 1231–1240. https://doi.org/10.1016/0092-8674(91)90299-e

Svitashev, S., Schwartz, C., Lenderts, B., Young, J. K., & Mark Cigan, A. (2016). Genome editing in maize directed by CRISPR–Cas9 ribonucleoprotein complexes. *Nature Communications*, *7*(1), 1–7.

Swanson, L. (n.d.). *Plasmids 101: How to Verify Your Plasmid Using a Restriction Digest Analysis*. Retrieved May 19, 2024, from https://blog.addgene.org/plasmids-101-how-to-verify-your-plasmid

Tang, X.-D., Gao, F., Liu, M.-J., Fan, Q.-L., Chen, D.-K., & Ma, W.-T. (2019). Methods for enhancing clustered regularly interspaced short palindromic repeats/Cas9-mediated homology-directed repair efficiency. *Frontiers in Genetics*, *10*, 551.

Tateno, Y., Nei, M., & Tajima, F. (1982). Accuracy of estimated phylogenetic trees from molecular data: I. Distantly Related Species. *Journal of Molecular Evolution*, *18*(6), 387–404. https://doi.org/10.1007/BF01840887

Thatcher, A., Waterson, P., Todd, A., & Moray, N. (2018). State of Science: Ergonomics and global issues. *Ergonomics*, *61*(2), 197–213. https://doi.org/10.1080/00140139.2017.1398845

Van Dijk, M., Morley, T., Rau, M. L., & Saghai, Y. (2021). A meta-analysis of projected global food demand and population at risk of hunger for the period 2010–2050. *Nature Food*, *2*(7), 494–501.

Vinci, R. M., Mestdagh, F., & De Meulenaer, B. (2012). Acrylamide formation in fried potato products–Present and future, a critical review on mitigation strategies. *Food Chemistry*, *133*(4), 1138–1154.

Waibel, F., & Filipowicz, W. (1990). U6 snRNA genes of Arabidopsis are transcribed by RNA polymerase III but contain the same two upstream promoter elements as RNA polymerase ll-transcribed U-snRNA genes. *Nucleic Acids Research*, *18*(12), 3451–3458.

Wang, F., Xia, Z., Zou, M., Zhao, L., Jiang, S., Zhou, Y., Zhang, C., Ma, Y., Bao, Y., Sun, H., Wang, W., & Wang, J. (2022). The autotetraploid potato genome provides insights into

highly heterozygous species. *Plant Biotechnology Journal*, *20*(10), 1996–2005. https://doi.org/10.1111/pbi.13883

Watanabe, K. (2015). Potato genetics, genomics, and applications. *Breeding Science*, *65*(1), 53–68.

White, R. J. (2011). Transcription by RNA polymerase III: More complex than we thought. *Nature Reviews Genetics*, *12*(7), 459–463.

Wiedenheft, B., Sternberg, S. H., & Doudna, J. A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature*, *482*(7385), 331–338.

Wilfinger, W. W., Mackey, K., & Chomczynski, P. (1997). Effect of pH and Ionic Strength on the Spectrophotometric Assessment of Nucleic Acid Purity. *BioTechniques*, *22*(3), 474–481. https://doi.org/10.2144/97223st01

Zheng, Z., Ye, G., Zhou, Y., Pu, X., Su, W., & Wang, J. (2021). Editing sterol side chain reductase 2 gene (*StSSR2*) via CRISPR/Cas9 reduces the total steroidal glycoalkaloids in potato. *All Life*, *14*(1), 401–413. https://doi.org/10.1080/26895293.2021.1925358

Appendix

Table A.1 - The components for the PCR amplification of the StU6 promoter from potato Désirée and Asterix genomic DNA.

Components	**Stock Concentration**	**Final Concentration**	**1 rx (µl)**
Phusion HF Buffer	5X	1X	10
dNTPs	10 mM	200 µM	1
Forward Primer	10 µM	1 µM	2.5
Reverse Primer	10 µM	1 µM	2.5
Template DNA	10 ng/µL	10 ng/µL	2
Phusion Hot Start II DNA Polymerase	2 U/µL	0.02 U/µL	0.5
Nuclease-Free H_2O			31.5
Final Volume			**50 µl**

Table A.2 - PCR cycles for the amplification of the StU6 promoter from potato Désirée and Asterix genomic DNA.

Stages	**Cycles**	**Temperature**	**Time**	**Number of Cycles**
1	Initial denaturation	98 °C	30 seconds	1
2	Denaturation Annealing Extension	98 °C 55 °C 72 °C	10 seconds 20 seconds 10 seconds	35
3	Final Extension	72 °C 4 °C	5 minutes ∞	1

Table A.3 - The components and volumes associated with the ligation reaction of PCR amplified StU6 fragments into PCR Zero Blunt vector according to the Zero Blunt® PCR Cloning kit protocol.

Components	**Volume (µl)**
Blunt PCR Product	1
pCR Zero Blunt Vector	1
5X Express Link T4 DNA Ligase Buffer (5 U/µl)	2
Express Link T4 DNA Ligase (5 U/µl)	1
PCR Grade Water	5
Total Volume	**10 µl**

Table A.4 - The reaction mixture for the restriction digestion of pChimera.

Components	**Volume (µl)**
gDNA (0.19 ng/µl)	5.2
10X NEBuffer r3.1	5
PvuII Restriction Enzyme	1
Nuclease-Free H_2O	38.8
Total volume	**50 µl**

Table A.5 - The components for the PCR amplification of insert and vector.

Components	**Stock Concentration**	**Final Concentration**	**1 rx (µl)**
Phusion HF Buffer	5X	1X	10
dNTPs	10 mM	200 µM	1
Forward Primer	10 µM	1 µM	2.5
Reverse Primer	10 µM	1 µM	2.5
Template DNA			1
Phusion Hot Start II DNA Polymerase	2 U/µL	0.02 U/µL	0.5
Nuclease-Free H_2O			32.5
Final Volume			**50 µl**

Table A.6 - PCR cycles and conditions for the amplification of insert.

Stages	**Cycles**	**Temperature**	**Time**	**Number of Cycles**
1	Initial denaturation	98 °C	30 seconds	1
2	Denaturation Annealing Extension	98 °C 55 °C 72 °C	10 seconds 20 seconds 10 seconds	35
3	Final Extension	72 °C 4 °C	5 minutes ∞	1

Table A.7 - PCR cycles and conditions for the amplification of vector.

Stages	**Cycles**	**Temperature**	**Time**	**Number of Cycles**
1	Initial denaturation	98 °C	30 seconds	1
2	Denaturation Annealing Extension	98 °C 55 °C 72 °C	10 seconds 20 seconds 45 seconds	35
3	Final Extension	72 °C 4 °C	5 minutes ∞	1

Table A.8 – The components for the In-Fusion cloning reaction.

Components	**Cloning Reaction**	**Negative Control**
Insert	1 µl	-
Linearized Vector	2 µl	2 µl
5X In-Fusion HD Enzyme Premix	2 µl	2 µl
Deionized Water	5 µl	6 µl
Total volume	**10 µl**	**10 µl**

Table A.9 - The reaction details associated with colony PCR of putative positive colonies.

Components	**Stock Concentration**	**Final Concentration**	**1 rx (µl)**
Buffer B1	10 X	1 X	1
$MgCl_2$	100 mM	2 mM	0.2
Forward Primer	10 µM	0.2 µM	0.2
Reverse Primer	10 µM	0.2 µM	0.2
$dNTP_S$	10 mM	0.2 mM	0.2
HOT FIREPol® DNA polymerase	5 U/µl	0.02 U/µl	0.04
Nuclease-Free H_2O			8.16
Final Volume			**10 µl**

Table A.10 - The PCR cycles associated with colony PCR of putative positive colonies.

Stages	**Cycles**	**Temperature**	**Time**	**Number of Cycles**
1	Initial denaturation	95 °C	12 minutes	1
2	Denaturation Annealing Extension	95 °C 55 °C 72 °C	10 seconds 20 seconds 10 seconds	35
3	Final Extension	72 °C 10 °C	5 minutes ∞	1

Table A.11 – The distance of the TATA box and USE from TSS, as well as the distance of the TATA box and USE from each other in Désirée clones.

Cultivar	**Promoter**	**Clones**	**TATA Box Distance from TSS**	**USE Distance from TSS**	**USE Distance from TATA Box**
Désirée	StU6-1	D1	21	55	25
		D2	21	55	25
	StU6-2	D1	24	55	25
		D2	24	55	25
	StU6-3	D1	24	55	25
		D2	24	55	25
	StU6-4	D1	23	54	25
		D2	23	54	25
		D3	23	54	25
		D4	23	54	25
		D5	23	54	25
		D6	23	54	25
		D7	23	54	25
		D8	23	54	25
		D9	23	54	25
		D10	23	54	25
		D11	23	54	25
		D12	23	54	25
		D13	23	54	25
		D14	23	54	25
		D15	23	54	25
		D16	23	54	25
		D17	23	54	25
		D18	23	54	25
		D19	23	54	25
		D20	23	54	25

Table A.12 – The distance of the TATA box and USE from TSS, as well as the distance of the TATA box and USE from each other in Asterix clones.

Cultivar	**Promoter**	**Clones**	**TATA Box Distance from TSS**	**USE Distance from TSS**	**USE Distance from TATA Box**
Asterix	StU6-1	A1	21	55	25
		A2	21	55	25
	StU6-2	A1	24	55	25
		A2	24	55	25
	StU6-4	A1	23	54	25
		A2	23	54	25
		A4	23	54	25
		A5	23	54	25
		A6	23	54	25
		A8	23	54	25
		A9	23	54	25
		A10	23	54	25
		A11	23	54	25
		A12	23	54	25
		A13	23	54	25
		A14	23	54	25
		A15	23	54	25
		A16	23	54	25
		A17	23	54	25
		A18	23	54	25
		A19	23	54	25
		A20	23	54	25

Table A.13 – Multiple sequence alignment of Désirée clones in NCBI's Genome Database (refseq_genome) in *Solanum tuberosum*.

Promoter	Clone	Best Hit	Score	Expect	Identity	Gaps
StU6-1	D1 D2	Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00381 Sequence ID: NW_006239304.1	771 bits (417)	0.0	458/476 (96%)	9/476 (1%)
StU6-2	D1	Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00300 Sequence ID: NW_006239221.1	682 bits (369)	0.0	453/488 (93%)	27/488 (5%)
	D2	Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00300 Sequence ID: NW_006239221.1	863 bits (467)	0.0	482/489 (99%)	1/489 (0%)
StU6-3	D1 D2	Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00047 Sequence ID: NW_006238973.1	789 bits (427)	0.0	427/427 (100%)	0/427 (0%)
StU6-4	D1 D2 D3 D5 D7 D8 D9 D10 D11 D15 D16 D17	Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00105 Sequence ID: NW_006239031.1	695 bits (376)	0.0	376/376 (100%)	0/376 (0%)
	D4 D13 D14 D20	Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00105 Sequence ID: NW_006239031.1	398 bits (215)	4e-109	239/251 (95%)	0/251 (0%)
	D6 D12	Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00105 Sequence ID: NW_006239031.1	520 bits (281)	7e-146	346/377 (92%)	5/377 (1%)
	D18	Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00105 Sequence ID: NW_006239031.1	689 bits (373)	0.0	373/373 (100%)	0/373 (0%)
	D19	Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00105 Sequence ID: NW_006239031.1	693 bits (375)	0.0	375/375 (100%)	0/375 (0%)

Table A.14 – Multiple sequence alignment of Asterix clones in NCBI's Genome Database (refseq_genome) in *Solanum tuberosum.*

Promoter	Clone	Best Hit	Score	Expect	Identity	Gaps
StU6-1	A1	Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00381 Sequence ID: NW_006239304.1	784 bits (424)	0.0	458/473 (97%)	8/473 (1%)
	A2	Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00381 Sequence ID: NW_006239304.1	776 bits (420)	0.0	459/476 (96%)	9/476 (1%)
StU6-2	A1 A2	Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00300 Sequence ID: NW_006239221.1	902 bits (488)	0.0	488/488 (100%)	0/488 (0%)
StU6-4	A1	Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00105 Sequence ID: NW_006239031.1	688 bits (372)	0.0	372/372 (100%)	0/372 (0%)
	A2 A4 A5 A6 A8 A12 A15 A17 A19	Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00105 Sequence ID: NW_006239031.1	398 bits (215)	4e-109	239/251 (95%)	0/251 (0%)
	A9 A10 A13 A14 A16 A18	Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00105 Sequence ID: NW_006239031.1	695 bits (376)	0.0	376/376 (100%)	0/376 (0%)
	A11	Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00105 Sequence ID: NW_006239031.1	459 bits (248)	2e-127	250/251 (99%)	0/251 (0%)
	A20	Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00105 Sequence ID: NW_006239031.1	392 bits (212)	2e-107	238/251 (95%)	0/251 (0%)

Sequencing Data:

The trimmed sequence and length (bp) associated with StU6-1, StU6-2, StU6-3 and StU6-4 in *S. tuberosum* Désirée and Asterix cultivars are presented below.

Désirée:

StU6-1 – D1, D2 (428 bp)

TATGCATGATCCATTTGGTTTATCACCTTGTTTCTAAGTGATTTTTTAATTTTTTTGTGTAAAATTTC
AACGGTTAAACCGAATGGTCGTTTTATAACAGTTTAACATATTTACAAATTGACAACCGTTAAGTC
GAATCAATAGACTTCAAAACCAAACAAACCGATACGGGCACTAATTTCAATAACCAAATGGTACA
AGTTGAATATGGGGGCAAATCTGGACTCTAGGCTTAGTTGGGCTCTATGTGCATGAATGAACATA
AAAGCAAGAGCAAAAACCTGTAGCTAGGTCCAGTCCCATGCCTTTGGAAAAACTCAATGTGCTAA
TTCTCCCTCATCGTCTGCAGAGAGAAGCTTCGCTGTGTTTATATAATTGAACAGTAACATGTATGC
TTGTCCCTTCGGGGACATCCGATAAAATTGGAA

StU6-2 – D1 (413 bp)

TTCCAATTTTATCAGATGTCCCCGAAGGGACAATTATAGATGTTAATGTTACGCAATATAAACAA
AACAAAGAAGCTTTCTCCAGCCGATGTGGGGAAAATAGCTTTCTAATTGAGTATTTTTCAACTACA
TGGGCCTGGACTTAACATTAATGGGCCGACCTTCTTAGTAGAACACACATTGGATTGGTGATCCAT
TGATTTGGCGAAAAACAAGAATAAAATTATCGGAAAAATTTCATAAGTAGTGAGATTAGTCTTCA
TTTTTGTTTGTTTTTTTCAAAATTTAGTATTATTTTTTCATATACATTGTATGCACGCGCAAGTAGT
ATCAATGTATATACACTGATAGCATAAACACAAAAAACTTCGTCTTTTTCGTTGACAACTTAAGTT
TTCCCTCAATGCTATCTA

StU6-2 – D2 (441 bp)

TTCCAATTTTATCGGATGTCCCCGAAGGGACAATTATAGATGTTAATGTTACGCAATATAAACAA
AACAAAGAATCTTTCTTCAGCCGATGTGGGGAAAATAGCTTTCTAGTTGAGTATTTTTCAACTACA
TGGGCCTGGATCTGAATGGAACCCAATTGGGCCTGGACTTAACATTAATGGGCCGACCTTCTTAG
TAGAACACACATTGGATTGGTGATCCATCGATTTGGCGAAAAACAAGAATAAAATTATCGGAAAA
TTTTCATAAGTAGTGAGATTAGTCTTCATTTTTGTTTGTTTTTTTCAAAAATTTAGTATTATTTTTTC
ATATACATTGTATGCACGCGCAAGTAGTATCAATATATATACACTGATAGCATAAACACAAAAAA
CTTCGTCTTTTTCGTTGACAACTTAAGTTTTCCCTCAATGCTATCTA

StU6-3 – D1, D2 (378 bp)

TTCCAATTTTATCGGATGTCCCCGAAGGGACAAGCATACATGTTAATATTCCGCGATATAAACTAG
ACAAAGAAGCTTTATGCATTCGATGAGGGATAATTTGCTTGTTGGGTTTCTTTTGCGGGCCTGGAT
CTTCCATCTGAATGGGCCGGCCCTATCTTATTTCTCTTGCTTCATGGACTTCGATGGAAAAACACA
AAGTAGAACCATGATTTATGAATTATAGGTAATGAGTTATTGGTTTAACTGCTTTTAATAACGGTA
TTGGTACTTTTATTTTATTGGGGTATCGGTTCTATGTTTTGGATTTTTTCTTTCTGTATTAACTGATA
ACTGATAGTAAATTAGTAAGTTATAATTAGAGTTAATATCTCAAAA

StU6-4 – D1, D2, D3, D5, D7, D8, D9, D10, D11, D15, D16, D17, D18, D19, D20 (326 bp)

TTTATACACCCCTACCTAGATGATTAAGTTTTACTTTAGTTGGTGTGAAATGGATAAATTCTAAAA
TATGAGGTGTGGAATGAAGGATTGTCATCAATTAGTTGGCCCCAACCAAGTAAAATAAGAAGGCC
GGCCCATTACAATTAAGTCGTCACACAAGTGGGCTTCATTGAAACAAGCGCAAAAACGAGTCCAG
GCCTGTGTTAGCGTGAAGACTCAACCAGCGATTTCTCCCTCATCGCTTACAGAGAAAAGCTGTGT
GTGGTTTATATCGCGAAACCTAACAGTTTAGCTCGTCCCTTCGGGGACATCCGATAAAATTGGAA

StU6-4 – D4, D13, D14 (379 bp)

TTTATACAACCCTACCTCGATGCTTAAGTTTTACTTTAGTTGGTGTGAAATGGATAAATTCTAAAA
TATGAGGTCTTGTCGGTATGGAATCTACCCTTTTGGAGTGAAGGATCGTGATCAGTTAGTTTGAGT
AAATACTCAAGAAGGCAAGAAGCTGGCCCCAACCAAGTAAAATAAGAAGGCCGGCCCATTACAA
TTAAGTCGTCACACAAGTGGGCTTCATTGAAACAAGCGCAAAAACGAGTCCAGGCCCGTGTTAGC
GTAAAGATTCAACCAGCGATTACTCCCTCATCGGTTACAGAGAAATGCTGTGTGTTGTTTATATGG
CGAAACCTAACAGTCCAGTTCGTCCCTTCGGGGACATCCGATAAAATTGGAA

StU6-4 – D6, D12 (323 bp)

TTTATACAACCCTACCTCGATGCTTAAGTTTTACTTTAGTTGGTGTGAAATGGATAAATTCTAAAA
TATGAGGTCTTGTCGGTATGGAATCTACCCTTTTGGGCCCAACCAAGTAAAATAAGAAGGCCGGC
CCATTACAATTAAGTCATCACACAAGTGGGCTTCATTGAAACAAGCGCAAAAAGGAGTCCAGGCC
CGTGTTAGCGTGAAGACTCAACCAGCGATTTCTCCCTCATCGGTTGCACAGAAAAGCTGTGTGTTG
TTTATATGGCGAAACCTAACAGTCTGACTTGTCCCTTCGGGGACATCCGATAAAATTGGAA

Asterix:

StU6-1 – A1, A2 (476 bp)

AGCAAGATGCAATGTATCAACTCATATGCATGATCCATTTGGTTTATCACCTTGTTTCTAAGTGAT
TTTTAATTTTTTTGTGTAAAATTTCAACGGTTAAACCGAATGGTCGTTTTATAACAGTTTAGCATAT
TTACAAATTGACAACCGTTAAGTCAAATCAATAGACTTCAAAACCAAACAGACCGATACGGGCAC
TAATTTCAATAACCAAATGGTACAAGTTGAATATGGGGGCAAATCTGGACTCTAGGCTTAGTTGG
GCTCTATGTGCATGAATGAACATATAAGCAAGAGCAAAAACCTGTAGCTAGGTCCAGGCCCATGC
CTTTGGAAAAACTCAATGTGCTAATTCTCCCTCATCGTCTGCAGAGAGAAGCTTCGCTGTGTTTAT
ATAATTGAACAGTAACATGTATGCTTGTCCCTTCGGGGACATCCGATAAAATTGGAACGATACAG
AGAAGATTAGCATGCC

StU6-2 – A1, A2 (488 bp)

ACCACTTAAACTGAGAACAGTCAATAGATAGCATTGAGGGAAAACTTAAGTTGTCAACGAAAAA
GACGAAGTTTTTTGTGTTTATGCTATCAGTGTATATACATTGATACTACTTGCGCGTGCATACAAT
GTATATGAAAAAATATTACTAAATTTTGAAAAAAACAAACAAAAATGAAGACTAATCTCACTACT
TATGAAAATTTTCCGATAATTTTATTCTTGTTTTTCGCCAAATCGATGGATCACCAATCCAATGTGT
GTTCTACTAAGAAGGTCGGCCCATTAATATTAAGTCCAGGCCCAATTGGGTTCCATTCAGATCCAG
GCCCATGCAGTTGAAAAATACTCAACTAGAAAGCTATTTTCCCCACATCGGCTGAAGAAAGCTTC
TTTGTTTTATTTATATTGCGTAACATTAACATCTATAATTGTCCCTTCGGGGACATCCGATAAAATT
GGAACGATACAGAGAAGATTAGCATGGC

StU6-4 – A1, A9, A10, A13, A14, A16, A18 (376 bp)

CAGGCTTAGTTCAGTTGCATTATGTCTTTATACACCCCTACCTAGATGATTAAGTTTTACTTTAGTT
GGTGTGAAATGGATAAATTCTAAAATATGAGGTGTGGAATGAAGGATTGTCATCAATTAGTTGGC
CCCAACCAAGTAAAATAAGAAGGCCGGCCCATTACAATTAAGTCGTCACACAAGTGGGCTTCATT
GAAACAAGCGCAAAAACGAGTCCAGGCCTGTGTTAGCGTGAAGACTCAACCAGCGATTTCTCCCT
CATCGCTTACAGAGAAAAGCTGTGTGTGGTTTATATCGCGAAACCTAACAGTTTAGCTCGTCCCTT
CGGGGACATCCGATAAAATTGGAACGATACAGAGAAGATTAGCATGGC

StU6-4 – A2, A4, A5, A6, A8, A11, A12, A15, A17, A19, A20 (429 bp)

TTCACTTAGTTCAGTTGCATTATGTCTTTATACAACTCTACCTCGATGATTAAGTTTTACTTTAGTT
GGTGTGAAATAGGTAAATTCTGAAATATGAGGTCTTGTCGGTATGGAACCTACCTTTTTGGAGTG
AAGGATCGTGATCAGTTAGTTTGAGTAAATACTCAAGAAGGCAAGAAGCTGGCCCCAACCAAGT
AAAATAAGAAGGCCGGCCCATTACAATTAAGTCGTCACACAAGTGGGCTTCATTGAAACAAGCGC

AAAAACGAGTCCAGGCCCGTGTTAGCGTAAAGATTCAACCAGCGATTACTCCCTCATCGGTTACA
GAGAAATGCTGTGTGTTGTTTATATGGCGAAACCTAACAGTCCAGTTCGTCCCTTCGGGGACATCC
GATAAAATTGGAACGATACAGAGAAGATTAGCATGGC