



**Inland Norway  
University of  
Applied Sciences**

Faculty of Applied Ecology, Agricultural Sciences and Biotechnology

**Neda Kianpour**

**Master thesis**

**Identification of optimal sgRNA  
candidates for mutagenesis of the  
ALS1 gene in *Ipomoea batatas* using  
in-vitro cleavage assay and  
CRISPR/Cas9 technology**

**Masters in applied experimental biotechnology**

**2022-2023**



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

## Acknowledgement

I would like to express my deepest gratitude and appreciation to my supervisor, Professor Wenche Johansen, for her unwavering support, guidance, and encouragement throughout my research. Her expertise, insights, and constructive feedback have been invaluable to the completion of this work. I would also like to thank the faculty members in the Biotechnology department, whose courses and seminars have equipped me with the knowledge and skills necessary to undertake this research. I am grateful to the faculty of Applied Ecology, Agricultural Sciences and Biotechnology for providing me with a stimulating academic environment, state-of-the-art facilities, and financial support. Special thanks go to my colleagues and friends, who have provided me with a sense of community and moral support during the ups and downs of the research process. Last but not least, I would like to express my heartfelt gratitude to my soltaneh ghalb, my husband, Jonas Haukedal, my parents in Iran, as well as my new family in Norway for their unwavering love, encouragement, patient, and support throughout my academic journey. Thank you all for contributing to the successful completion of my master's thesis.

## Abbreviations

ALS1; Aceto Lactate Synthase

bp; base pair

Cas; CRISPR-associated proteins

CDS; Coding Sequence

CRISPR; Clustered Regularly Interspaced Short Palindromic Repeats

CRISPR-Cas; Clustered Regularly Interspaced Short Palindromic Repeats-CRISPR-associated proteins

crRNA; CRISPR RNA

dCas9; deactivated / dead Cas9

DNA; Deoxyribo Nucleic Acid

dNTP; deoxyNucleoside Tri Phosphate

DSB; Double-Strand Break

*E. coli*; *Escherichia coli*

EDTA; Ethylene Diamine Tetraacetic Acid

EtOH; Ethanol

gRNA; guide RNA

HDR; Homology Directed Repair

HNH; an endonuclease domain named for characteristic Histidine and Asparagine residues

IVT; In-vitro Transcription

mRNA; messenger RNA

NaOAc; Sodium Acetate

NBT; New Breeding Technique

nCas9; nickase Cas9

NHEJ; Non-Homologous End Joining

NLS; Nuclear Localization Signal

PAM; Protospacer-Adjacent Motif

PCR; Polymerase Chain Reaction

PDS; phytoene desaturase

pegRNA; prime editing guide RNA

pre-crRNA; precursor-CRISPR RNA

RNA; Ribo Nucleic Acid

RNP; Ribo Nucleo Protein

rpm; revolutions per minute

RuvC; an endonuclease domain named for an *E. coli* protein involved in DNA repair

sgRNA; single guide RNA

SpCas9; *Streptococcus pyogenes* Cas9

SSN; Sequence-Specific Nuclease

TALEN; Transcription Activator-Like Effector Nuclease

TracrRNA; Trans-activating CRISPR RNA

ZFN; Zinc Finger Nuclease

---

# Table of Contents

<b>ABSTRACT.....</b>	<b>10</b>
<b>1. INTRODUCTION .....</b>	<b>11</b>
1.1 BACKGROUND OF THE STUDY .....	11
1.2 SWEET POTATO, A VITAL FOOD STAPLE .....	12
1.3 ALS1 GENE IN SWEET POTATO .....	13
1.4 TRADITIONAL PLANT MODIFICATION .....	13
1.5 GENOME EDITING IN SWEET POTATO .....	14
1.6 CRISPR/CAS-BASED GENOME EDITING .....	14
1.7 IMPACT OF THE STUDY.....	19
<b>2. AIM OF THE STUDY .....</b>	<b>20</b>
2.1 SPECIFIC OBJECTIVES.....	20
<b>3. MATERIAL AND METHODS .....</b>	<b>21</b>
3.1 EXPERIMENTAL DESIGN METHODOLOGY IN RESEARCH.....	22
3.2 PLASMIDS DNA MATERIALS.....	22
3.3 SANGER SEQUENCING OF ALS1 GENE IN FIVE DISTINCT PLASMIDS DNA .....	22
3.3.1 <i>Cycle sequencing reaction of ALS1 gene in five distinct plasmids DNA</i> .....	22
3.3.2 <i>Purification of the extension products</i> .....	23
3.3.3 <i>Capillary electrophoresis</i> .....	23
3.3.4 <i>Data analysis of sanger sequencing data using CLC software</i> .....	23
3.4 INSERTING OF THE ALS1 INSERT FRAGMENTS .....	24
3.4.1 <i>Transformation of E. coli TOP10 cells with ALS1 insert fragments</i> .....	24
3.4.2 <i>Selection of the transformed cells</i> .....	24
3.4.3 <i>Culturing of putative positive transformants</i> .....	24
3.4.4 <i>Plasmids DNA isolation</i> .....	25

3.4.5	<i>Restriction digest analysis of putative positive plasmids DNA</i> .....	25
3.5	IN-VITRO CLEAVAGE ASSAY .....	25
3.5.1	<i>sgRNA designing by using CRISPOR software</i> .....	25
3.5.2	<i>In-vitro transcription (IVT) reaction of sgRNA</i> .....	26
3.5.3	<i>In-vitro cleavage assay</i> .....	26
3.6	PRODUCING IN-VITRO RECOMBINANT CAS9 BY IN-FUSION CLONING.....	27
3.6.1	<i>Construct expression vector for in-vitro recombinant Cas9 production using CLC software</i> .....	27
3.6.2	<i>Primer designing for in-fusion cloning</i> .....	28
3.6.3	<i>PCR amplifications of insert (CDS of Cas9 and NLS) and vector (pET302_NT)</i> .....	28
3.6.4	<i>Isolation and purification of PCR products (gel extraction)</i> .....	29
3.6.5	<i>In-fusion cloning reaction</i> .....	29
3.6.6	<i>Transformation of the Stellar competent cells with in-fusion cloning mixture</i> .....	29
3.6.7	<i>Selection of the transformed cells</i> .....	30
3.6.8	<i>Verification by colony PCR</i> .....	30
3.6.9	<i>Culturing putative positive pET302_His_Cas9_NLS transformation</i> .....	31
3.6.10	<i>Plasmid DNA isolation</i> .....	31
3.6.11	<i>Restriction digest analysis of putative positive pET302_His_Cas9_NLS</i> .....	31
<b>4.</b>	<b>RESULTS</b> .....	<b>32</b>
4.1	ALS1 GENE VERIFICATION .....	32
4.1.1	<i>Sanger sequencing technique was performed to determine the sequence of insert (ALS1 gene)</i> .....	32
4.1.2	<i>Sequence alignment data for the confirmation of the existence of the insert (ALS1 gene)</i> .....	32
4.1.3	<i>Restriction digestion was performed to analyse the putative positive colonies containing plasmids DNA</i> .....	33
4.2	IN-VITRO CLEAVAGE ASSAY .....	34
4.2.1	<i>Five sgRNA designed for plasmids DNA targets (PCR_IbALS1_S1.1 and PCR_IbALS1_S4.1)</i> .....	34



---

4.2.2	<i>In-vitro cleavage assay was performed to identify suitable sgRNA for PCR_IbALS1_S1.1</i>	35
4.2.3	<i>In-vitro cleavage assay was performed to identify suitable sgRNA for PCR_IbALS1_S4.1</i>	37
4.3	PRODUCING IN-VITRO RECOMBINANT CAS9 BY IN-FUSION CLONING .....	39
4.3.1	<i>PCR was performed to amplify the insert (CDS of Cas9 and NLS) and the vector (pET302_NT)</i>	39
4.3.2	<i>Gel extraction was performed to isolate and purify PCR products.....</i>	40
4.3.3	<i>In-fusion cloning reaction was performed to clone the insert (CDS of Cas9 and NLS) into the vector (pET302_NT).....</i>	41
4.3.4	<i>Colony PCR was performed to verify the putative positive pET302_His_Cas9_NLS</i>	42
4.3.5	<i>Restriction digest was performed to analyze the putative positive pET302_His_Cas9_NLS</i>	43
<b>5.</b>	<b>DISCUSSION.....</b>	<b>45</b>
5.1	SANGER SEQUENCING AND PCR OF ALS1 GENE IN FIVE DISTINCT PLASMIDS.....	45
5.2	PCR-BASED PRIMER DESIGN AND DNA AMPLIFICATION .....	47
5.3	SGRNA DESIGNING .....	47
5.4	EVALUATING SGRNA EFFICIENCY BY IN-VITRO CLEAVAGE ASSAYS .....	49
5.5	IN-FUSION CLONING-BASED IN-VITRO PRODUCTION AND CHARACTERIZATION OF RECOMBINANT CAS9 PROTEIN 50	
<b>6.</b>	<b>CONCLUSION .....</b>	<b>52</b>
<b>7.</b>	<b>REFERENCES .....</b>	<b>53</b>
<b>A.</b>	<b>APPENDIX.....</b>	<b>60</b>

## Abstract

Plant genome editing, a transformative technology in agriculture and biotechnology, faces challenges in efficiently delivering CRISPR/Cas components. Current methods, including *Agrobacterium*-mediated delivery, particle bombardment, and PEG-mediated protoplast transformation, have limitations due to random plasmid DNA integration and unpredictable genetic expression. As a solution, there's a growing interest in developing gene-edited lines with reduced foreign genetic integration risk using DNA-free delivery methods like RNP complexes (pre-assembled ribonucleoprotein) or transient plasmid DNA expression. Hence, this study aimed to create tools for CRISPR/Cas9-based gene editing in *Ipomoea batatas* (sweet potato) by targeting the ALS1 gene as a reliable null-mutant phenotype. The focus was designing and evaluating optimal sgRNA candidates for CRISPR/Cas9 RNP-based genome editing. Additionally, the performance characteristics of an engineered recombinant Cas9 with those of a commercially available Cas9 were aimed to be compared. Firstly, the selection of target plasmids DNA, PCR\_IbALS1\_S1.1 and PCR\_IbALS1\_S4.1, represented the specific genomic region of interest (ALS1 gene) in *Ipomoea batatas* among four plasmids DNA. Then, five distinct sgRNAs were designed based on coding sequences and evaluated through in-vitro cleavage assays. The findings demonstrated effective cleavage by SpCas9 guided by a single sgRNA (FP\_IVT\_T7\_Spa2\_S1.1) in the complementary DNA sequence of plasmid DNA PCR\_IbALS1\_S1.1. Efforts were made to create in-vitro recombinant Cas9 protein through expression vector construction, primer designing, in-fusion cloning, and PCR. However, this endeavour yielded no product, primarily attributed to the substantial size of the insert fragment (CDS of Cas9 and NLS ~ 9849 bp). Thus, this study has laid the foundation and provided fundamental insights for future in-vitro ALS1 CRISPR/Cas9-based genome editing using RNP complexes in *Ipomoea batatas*, despite the challenges encountered and the outcome.

**Keywords:** CRISPR/Cas9 based genome editing, plant genome editing, transfection efficiency, ALS1 gene, in-vitro cleavage assay, and CRISPR/Cas9 DNA-free delivery.

# 1. Introduction

## 1.1 Background of the study

*Ipomoea batatas* (sweet potato) is one of the most important food crops in the world, especially in developing countries. It is a rich source of carbohydrates, vitamins, minerals, and antioxidants and has multiple uses as food, feed, and industrial raw material. However, sweet potato production and quality are affected by various stresses, such as pests, diseases, drought, salinity, and temperature. Therefore, improving sweet potato traits through genetic engineering is a promising strategy to enhance its productivity and resilience. (Luo, et al., 2023)

One of the critical challenges in the genetic engineering of sweet potato is the efficient delivery of foreign DNA into its genome. Sweet potato is a highly heterozygous and hexaploid crop, which makes it difficult to achieve stable and precise integration of transgenes. Moreover, conventional methods, such as *Agrobacterium*-mediated transformation or particle bombardment, may raise biosafety and regulatory concerns due to the random insertion of foreign DNA and the potential for gene flow. (Nawiri, et al., 2017)

To overcome these limitations, a novel approach based on CRISPR/Cas genome editing has emerged as a powerful tool for targeted modification of plant genomes. CRISPR/Cas is a system that consists of a Cas nuclease and a gRNA (guide RNA) that directs the Cas nuclease to a specific DNA sequence for cleavage. The resulting DSB (double-strand break) can then be repaired by either NHEJ (non-homologous end joining) or HDR (homology-directed repair), leading to gene knockout or gene replacement, respectively. CRISPR/Cas can be delivered into plant cells as plasmid DNA or RNP complexes (ribonucleoprotein). The latter method has the advantage of reducing the risk of foreign DNA integration and increasing the efficiency and specificity of genome editing. (Wang, et al., 2018)

Several studies have reported the successful application of CRISPR/Cas-based genome editing in sweet potatoes for various purposes, such as improving disease resistance, modifying starch biosynthesis, and creating herbicide tolerance. (Hou, et al., 2023) (Wang, et al., 2018) However, most of these studies have used DNA-based delivery methods, which may still need to be improved regarding biosafety and regulation. Therefore, developing and optimizing DNA-free delivery methods for sweet potatoes is necessary.

One of the critical steps in DNA-free delivery is the design and evaluation of optimal sgRNAs that can effectively guide the Cas nuclease to the target site. The choice of sgRNA can affect the efficiency and specificity of genome editing and the occurrence of off-target effects. (Wang, et al., 2018) Therefore, it is essential to use reliable methods to select and test sgRNAs before applying them in plant cells. Moreover, genome editing techniques require the selection of suitable endogenous genes that can serve as reliable and visible markers for evaluating the editing efficiency and specificity. (Ly, et al., 2022) In plants, two genes widely used for this purpose are ALS (aceto lactate synthase) and PDS (phytoene desaturase), which are involved in herbicide resistance and carotenoid biosynthesis. These genes have also been successfully applied for optimizing CRISPR/Cas approaches in sweet potatoes, a significant crop with a complex genome. (Chincinska, et al., 2022).

Moreover, different forms of Cas9 may have other performance characteristics, such as activity, specificity, stability, and delivery efficiency. One of the factors that can affect the performance of Cas9 is its origin and production method. Cas9 can be either engineered recombinant or commercially available. (Lee, et al., 2018) Comparing the performance characteristics of engineered recombinant and commercially available can help to optimize the design and application of CRISPR/Cas9-based genome editing system. (Konstantakos, et al., 2022) Hence, this study aimed to create tools for CRISPR/Cas9-based genome editing in sweet potatoes by targeting the ALS1 gene and comparing the performance features of a recombinant Cas9 that is engineered versus a Cas9 that is commercially bought.

## 1.2 Sweet potato, a vital food staple

The sweet potato, scientifically known as *Ipomoea batatas* (L.) Lam, part of the Convolvulaceae family, stands out as a prominent dicotyledonous plant. (Bovell-Benjamin, A. C., 2006) Convolvulaceae has only one known natural hexaploid species, which is sweet potato. The sweet potato is the only species of *Ipomoea* with significant economic importance. (Srisuwan, et al., 2006) Its importance as a primary root crop on a global scale is noteworthy, holding the seventh position in terms of annual production. (Simmonds, N., 1993) Cultivated in over 100 countries, the sweet potato is a valuable food source, animal feed and industrial raw material. To meet the increasing food demands of the world's growing population, enhancing yield production and ensuring stability is essential, especially under challenging conditions. (Zhang, et al., 2011) To overcome these challenges, genetic improvement of sweet potatoes is needed. One of the promising methods for genetic improvement is CRISPR/Cas9-based genome editing, which can introduce precise and targeted modifications in the DNA of sweet potato. (Watanabe, K., 2015) However, to apply

CRISPR/Cas9-based genome editing in sweet potato, it is necessary to evaluate its efficacy and specificity, which are the ability to edit the desired gene and avoid off-target effects. One of the ways to assess the efficacy and specificity of CRISPR/Cas9-based genome editing is to use a reliable target gene that can be easily detected and measured. (Butler, et al., 2015) (Nekrasov, et al., 2013) ALS1 gene in sweet potatoes is one of the best candidates to be used as a reliable target for testing the efficacy and specificity of CRISPR/Cas9-based genome editing.

### 1.3 ALS1 gene in sweet potato

ALS1 is a gene that encodes acetolactate synthase, an enzyme involved in the first step of the biosynthesis of branched-chain amino acids such as valine, leucine, and isoleucine. It is often used as a target gene for testing gene editing techniques or approaches introducing point mutations in ALS that do not entirely abolish its enzymatic activity. (Shimizu, et al., 2002) Moreover, the ALS1 gene is suitable for its application as a reporter/selection target in sweet potato, as it fulfills the following criteria: (1) they allow the assessment of the efficiency of new genome editing techniques; (2) they facilitate the identification of mutated events (e.g., co-transformation with the CRISPR/Cas constructs targeting ALS and selected endogenous gene/genes); (3) they enable the selection of genome-edited events; and (4) they support the generation of non-transgenic genome-edited plant lines. Therefore, the ALS1 gene can be a reliable target for testing the efficacy and specificity of CRISPR/Cas9-based genome editing in sweet potatoes. (Chincinska, et al., 2022)

### 1.4 Traditional plant modification

Humans have modified plants for thousands of years to improve their traits and characteristics, such as yield, quality, resistance, and adaptation. Traditional plant modification methods include selective breeding, crossbreeding, mutagenesis, and somatic hybridization. These methods rely on the natural variation and recombination of plant genes and the induction of random mutations or fusion of different plant cells. Traditional plant modification methods have been essential for the development of agriculture and food security, as well as for the enhancement of biodiversity and environmental sustainability. (Hamadani, et al., 2021) However, these methods have limitations and drawbacks, such as low efficiency, lack of precision, limited gene pool, and ecological risks. Therefore, modern biotechnology techniques, such as genetic engineering and genome editing, have been developed to overcome these challenges and accelerate plant modification. (Kumar, et al., 2020)

## 1.5 Genome editing in sweet potato

Sweet potato is an important crop providing food security and income for millions worldwide. However, sweet potato improvement is challenging due to its complex genetic structure, long breeding cycle, and susceptibility to biotic and abiotic stresses. To overcome these limitations, NBTs (new breeding techniques) have emerged as promising tools for enhancing sweet potato traits and quality. (Aglawe, et al., 2018) Among these techniques, SSNs (sequence-specific nucleases) are widely used for creating targeted and precise modifications in the sweet potato genome without introducing foreign DNA or leaving any transgenic traces. This allows the generation of improved sweet potato varieties that are free of transgenes and suitable for processing. Among the SSNs, the most widely used ones are ZFNs (zinc-finger nucleases), TALENs (transcription activator-like effector nucleases), and CRISPR/Cas (clustered regularly interspaced short palindromic repeats/CRISPR-associated proteins), to precisely modify the genetic information encoded within a sweet potato's DNA. All the mentioned tools introduce DSBs (double-stranded breaks) at specific target sites within the DNA molecule, leading to the activation of the DNA repair mechanism in NHEJ and HDR. NHEJ repairs the DNA by directly rejoining the broken ends, which can result in small insertions or deletions (indels) at the target site, leading to gene knockout or random mutagenesis. (Shamshirgaran, et al., 2022) However, HDR repairs the DNA using a donor template, enabling gene knock-in, gene correction, or targeted mutagenesis. Unlike CRISPR/Cas, TALEN and ZFN are based on the fusion of a DNA-binding domain and a FokI endonuclease domain, which can be time-consuming, costly, and technically challenging. CRISPR/Cas-based genome editing is more straightforward, cheaper, and more efficient than TALEN and ZFN, as it can be designed for any genomic target and multiplexed by adding multiple gRNAs. (Gaj, et al., 2013)

## 1.6 CRISPR/Cas-based genome editing

The CRISPR/Cas technology is a revolutionary genome editing tool that enables precise modifications within the DNA of various organisms, including plants. It consists of two main components: the gRNA (guide RNA), which directs the Cas protein to the specific target site in the genome, and the Cas protein (such as Cas9), which acts as a pair of molecular scissors to cut the DNA at the desired location. This system has gained immense popularity due to its ease of use, high precision, and versatility in altering specific genes, facilitating various applications in research, agriculture, medicine, and other fields. (Asmamaw, M., & Zawdie, B., 2021). Hence, the

CRISPR/Cas system uses RNA to guide and cleave DNA, while TALENs and ZFNs use proteins to bind and cut DNA.

In 1987, Japanese researchers made a significant observation regarding CRISPR, which stands for Clustered Regularly Interspaced Short Palindromic Repeats. They noticed repetitive DNA sequences in the *iap* gene of *E. coli* (*Escherichia coli*) bacteria, and this discovery marked the first hints of CRISPR. (Ishino Y., et al., 1987) The CRISPR/Cas system is an adaptive immune system in prokaryotes, defending against invasive genetic elements like viruses and plasmids. It accomplishes this through a combination of CRISPR arrays, Cas proteins, and small RNA molecules. The CRISPR array consists of repetitive direct repeats, complemented by *tracr*RNAs (trans-activating CRISPR RNA). These direct repeats are interspersed with unique spacer sequences specific to each target. Integrated spacers allow the expression of gRNAs or crRNAs (CRISPR RNA), which play a crucial role in recognizing and destroying foreign DNA. Additionally, the leader sequence located before the CRISPR array and containing a promoter plays a vital role in spacer acquisition. (Barrangou, R., & Marraffini, L. A., 2014)

The CRISPR-mediated immunity steps can be divided into three stages: adaptation, crRNA maturation, and targeting. During adaptation, a prokaryote acquires a new spacer from foreign DNA and integrates it into the CRISPR array. This involves the recognition of a specific motif called the PAM (protospacer adjacent motif) by a Cas protein. The Cas protein cleaves the foreign DNA and inserts a fragment between existing repeats in the CRISPR array, creating a new spacer that recalls past infections. (Hille, F., & Charpentier, E., 2016) In the crRNA maturation stage, the CRISPR array is transcribed into a long RNA molecule called pre-crRNA. Cas proteins and other factors then process the pre-crRNA into short RNA molecules called crRNAs. These crRNAs consist of a spacer and repeat sequences complementary to a *tracr*RNA. The crRNAs combine with the *tracr*RNA and the Cas nuclease to form a complex ready for targeting. (Makarova, K. S., & Koonin, E. V., 2015) Finally, in the targeting stage, the crRNA-*tracr*RNA-Cas nuclease complex scans the genome for the target DNA site that matches the spacer and PAM sequences. Once the target DNA site is found, the crRNA base pairs with it and guides the Cas nuclease to cleave it. This leads to the degradation or modification of the target DNA, effectively preventing its expression or replication. (Karvelis, et al., 2013) This CRISPR-mediated adaptive immune leads to interference and degradation of foreign genetic material. (Nidhi, et al., 2021)

Recent studies indicate that the CRISPR/Cas system can be divided into two classes based on the structure of effector proteins: Class 1 and Class 2. (Makarova, et al., 2011) The effector protein is the critical component for the gene-editing function. It is a protein that interacts with the gRNA to locate the

target DNA sequence and perform gene editing or regulation. Class 1 uses multi-subunit effector complexes for target interference. This class includes types I, III, and IV, which rely on multiple Cas proteins to perform their functions. (Koonin, E. V., & Makarova, K. S., 2019) Class 2 involves a single effector protein for target interference. They are structured simpler than Class 1 and include types II, V, and VI. Class 2 systems are often used in genome editing applications due to the simplicity of their effector proteins, like the well-known Cas9 in type II CRISPR systems. (Makarova, K. S., Wolf, et al., 2015)

CRISPR/Cas9 is a powerful tool in genome editing, which has two components: sgRNA and Cas9 protein. It uses a sgRNA to direct the Cas9 protein to specific DNA sequences, allowing it to cut the DNA at precise locations. This cutting action enables genetic code modification by inserting, deleting, or replacing particular DNA sequences.

The first component, sgRNA, consists of two main components: the crRNA and the tracrRNA combined into a single synthetic RNA molecule. First, the crRNA contains a customizable sequence called spacer sequence, 20 nucleotides long, that match the target DNA sequence the CRISPR system aims to edit (protospacer). (Uniyal, et al., 2019) The recognition of the target DNA sequence depends heavily on the crRNA sequence's 3' end, also called the "seed region." Therefore, based on structural analysis, the functional sgRNAs could be distinguished from the non-functional ones by the accessibility of the last three bases in the seed region. (Wong, et al., 2015) This sequence guides the Cas protein to recognize the specific gene location that needs modification. Hence, CRISPR/Cas9-based genome editing is based on the designing and sequence of these 20 nucleotides of the crRNA that should be matched with the target (protospacer). This specific sequence can be designed online by using several online sgRNA design tools. Second, the tracrRNA provides stability to the RNA molecule and assists in binding with the Cas protein. It is a scaffold that helps in combining the crRNA with the Cas protein, aiding in the formation of the Cas9-sgRNA complex, which locates and edits the targeted DNA sequence. The combination of these two segments (crRNA and tracrRNA) into a sgRNA forms the functional guide for the Cas protein, enabling precise gene editing at the desired location in the genome.

The second component, Cas9, is a crucial component of the CRISPR/Cas9-based genome editing system and has two main parts: the target recognition lobe and the nuclease lobe. The target recognition lobe is responsible for binding to the gRNA and the target DNA. The nuclease lobe contains two domains: the HNH and Ruc-V-like. These domains are the ones that cut the DNA strands. The HNH domain cuts the complementary strand of crRNA, and the Ruc-V-like domain cuts the non-complementary strand. (Asmamaw, M., & Zawdie, B., 2021) The nuclease lobe also has a



carboxyl-terminal domain that interacts with the PAM sequence. Cas9 has multiple naturally occurring variants derived from various bacterial species. Based on their origin, structure, and function, there are at least six significant classes of Cas9 variants. (Bhattacharya, S., & Satpati, P., 2022) Among Cas9 variants, SpCas9 is the original and most commonly used variant of the Cas9 enzymes. It is derived from the bacteria *Streptococcus pyogenes*. SpCas9 is known for recognizing a specific DNA sequence called the PAM, which in the case of SpCas9 is 5'-NGG-3'. Once it identifies the PAM sequence, SpCas9 can cleave both strands of the DNA, allowing for precise gene editing. SpCas9 has two domains: a RuvC domain and a HNH domain, responsible for cleaving the non-target and target strands of DNA, respectively. (Kleinstiver, et al., 2015)

SpCas9 has two forms, recombinant Cas9 and commercial Cas9, which can be used for genome editing with CRISPR/Cas-based genome editing. However, they have some differences in their origin, production, and properties. Recombinant Cas9 protein is a form of Cas9 produced by expressing the Cas9 gene in a host organism, such as *E. coli*, and then purifying the protein from the cell lysate. Recombinant Cas9 can be modified by introducing mutations or tags to alter its activity, specificity, or affinity. Recombinant Cas9 protein can form RNP complexes with guide RNAs, which can be delivered into cells. (Hu, et al., 2018) However, commercial Cas9 is a term that refers to Cas9 products that biotechnology companies sell for research or therapeutic purposes. Commercial Cas9 can be recombinant Cas9 protein, Cas9 mRNA or plasmid DNA. Commercial Cas9 products are usually optimized for high purity, stability, and efficiency. Commercial Cas9 products may also have different features, such as Cas9 variants, fluorescent labels, or delivery vehicles. (Paix, et al., 2018)

Then, the sgRNA-tracrRNA-SpCas9 complex scans the genome for the target DNA site that has a PAM. Once the target DNA site is found, the sgRNA base pairs with it and activates the two domains of Cas9 to cleave it. This results in a DSB in the target DNA 3–4 bp upstream of the PAM, which the cell's mechanisms can repair in two ways, as mentioned before: NHEJ and HDR. (Puchta, H., 2017). NHEJ is often error-prone, as it can introduce insertions or deletions (indels) at the repair site. These indels can cause frameshift mutations or disrupt gene function, while HDR is a DNA repair mechanism that uses a homologous template to repair the DNA damage precisely. HDR can insert, delete, or replace specific sequences at the repair site. (Steinert, et al., 2016)

Introducing foreign DNA fragments through HDR using CRISPR/Cas9 can lead to the creation of GMOs (genetically modified organisms). In many regions, including the EU and several other countries globally, such genetic modifications in organisms trigger regulations and guidelines set for GMOs. This involves thorough risk assessment, stringent authorization processes, and specific legislative controls before these genetically modified organisms can be cultivated or utilized

commercially. (Ahmad, et al., 2016) However, recent CRISPR/Cas9 technology advancements have focused on developing DNA-free genome editing approaches. Techniques such as RNP delivery, base editing, and prime editing offer potential solutions to avoid the regulatory challenges associated with GMOs, as these methods avoid incorporating exogenous genetic material into the edited organism's genome. (Zhang, et al., 2021)

Moreover, CRISPR/Cas components can be delivered by two methods: DNA-free delivery and DNA-based methods. The DNA-free delivery method involves directly delivering a Cas9 protein and gRNA complex to the cells without using plasmids or viruses. In contrast, the DNA-based delivery method uses plasmids or viruses that carry the genes encoding the Cas9 protein and the gRNA. (Ma, et al., 2020) The plasmids or viruses are then transferred into the cells by various techniques, such as electroporation, microinjection, or biolistics. The cells then express the Cas9 protein and the gRNA from the DNA and perform the genome editing. Furthermore, the RNP complex can act faster than the DNA-based method, as it does not require transcription and translation of the Cas9 and gRNA. (Zhang, et al., 2021) It also reduces the off-target effects of genome editing, as the Cas9-gRNA complex is degraded over time and does not persist in the cell.

Several modified forms of the Cas9 protein have been developed to improve or alter its function for DNA-free genome editing applications, of which nCas9 (Cas9 nickase) and dCas9 (dead Cas9) are the most widely used and studied.

nCas9 is accomplished by introducing precise point mutations in one of the DNA cleavage domains of Cas9. These mutations effectively hinder the cutting capability of either the RuvC or HNH domain, forming a single-strand nick on the designated DNA. Prime editing uses only nCas9 fused to a reverse transcriptase enzyme. (Eid, et al., 2018) The nCas9 guides a pegRNA (prime editing guide RNA) to the target site, containing both the targeting and edited sequences. The reverse transcriptase copies the edited sequence from the pegRNA into the nicked strand, creating a flap of mismatched DNA. The flap is then resolved by cellular mechanisms, resulting in precise and controlled alterations to the DNA, such as insertions, deletions, or substitutions. This method can make a wide range of edits without inducing double-strand breaks. (Lee, et al., 2023)

By introducing targeted point mutations, like D10A and H840A, in Cas9, we can effectively turn off certain functional domains of the Cas9 protein. This modification transforms the Cas9 protein into a variant known as dCas9, which loses its ability to cleave DNA but retains its capacity to bind to desired DNA sequences selectively. (Piatek, et al., 2015) To achieve this, dCas9 is combined with a gRNA to recognize and associate with specific DNA sequences. dCas9 is used in base editing and

prime editing, two techniques for changing single nucleotides in the DNA without creating DSB. (Brezgin, et al., 2019)

Base editing uses either nCas9 or dCas9 and a base editor enzyme, such as cytidine deaminase or adenine deaminase. The nCas9 or dCas9 guides the base editor to the target site, where the enzyme converts one base into another, such as C to T or A to G. This method can directly install point mutations in the DNA without inducing a double-strand break. (Eid, et al., 2018)

## 1.7 Impact of the study

This study aims to develop a tool for CRISPR/Cas9-based gene editing in *Ipomoea batatas* (sweet potato) by targeting the ALS1 gene, which can produce a reliable null-mutant phenotype. Moreover, this study aims to compare the performance characteristics of a commercially available Cas9 with those of a recombinant Cas9 that is engineered in this study. Therefore, this study involves designing and evaluating optimal sgRNA candidates for CRISPR/Cas9 RNP-based genome editing and creating in-vitro recombinant Cas9 protein through expression vector construction, primer designing, in-fusion cloning, and PCR. However, due to the time constraints of the project, this study will only test some plasmid DNA. Nevertheless, the results of this study are expected to provide a foundation and fundamental insights for future in-vitro ALS1 CRISPR/Cas9-based genome editing using RNP complexes in *Ipomoea batatas* and in-vitro recombinant Cas9 production.

## 2. Aim of the study

The aim of this study was to create tools for CRISPR/Cas-based gene editing in sweet potato by targeting the ALS1 gene as a reliable null-mutant phenotype. Furthermore, to compare the performance characteristics of an engineered recombinant Cas9 with those of a commercially available Cas9. The ALS1 gene encodes acetolactate synthase (ALS), which is an enzyme involved in branched-chain amino acid biosynthesis. The ALS1 gene is frequently used as a target gene for testing gene editing techniques or approaches that introduce point mutations in ALS without entirely abolishing its enzymatic activity. Additionally, the ALS1 gene is suitable for use as a reporter/selection target in sweet potatoes. Therefore, ALS1 is a suitable gene for testing the efficacy and specificity of CRISPR/Cas genome editing in sweet potato.

### 2.1 Specific Objectives

- To design and evaluate optimal sgRNAs for CRISPR/Cas9 RNP-based genome editing of ALS1 in sweet potato.
- To compare the performance characteristics of an engineered recombinant Cas9 with those of a commercially available Cas9.

The expected outcomes of this study were:

- To identify one or more sgRNAs that can efficiently cleave the ALS1 gene in-vitro.
- To produce an engineered recombinant Cas9 protein that can be used for RNP-based genome editing in sweet potato.
- To provide fundamental insights for future in-vitro ALS1 genome editing using RNP complexes in sweet potato.

### 3. Material and Methods

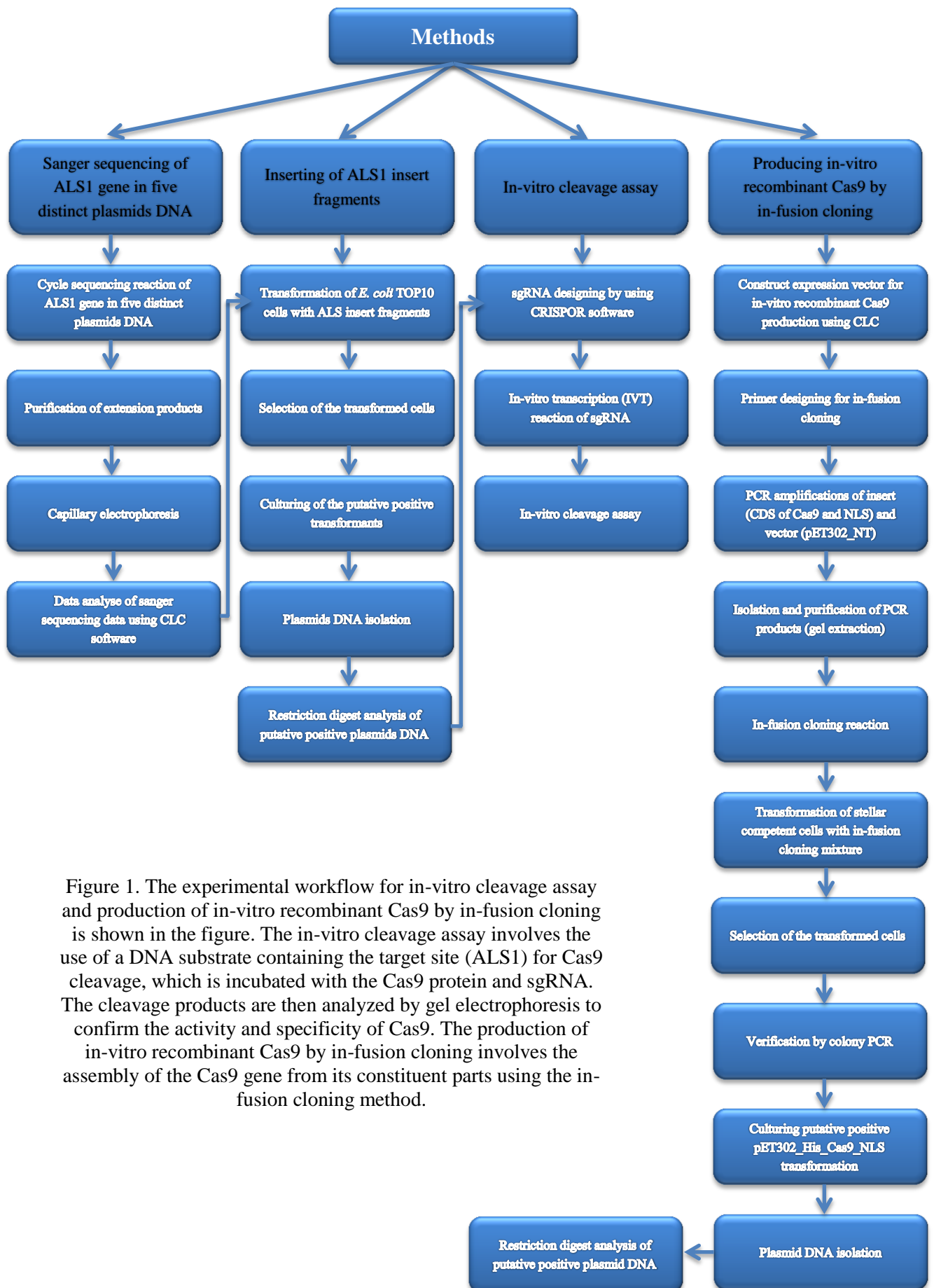


Figure 1. The experimental workflow for in-vitro cleavage assay and production of in-vitro recombinant Cas9 by in-fusion cloning is shown in the figure. The in-vitro cleavage assay involves the use of a DNA substrate containing the target site (ALS1) for Cas9 cleavage, which is incubated with the Cas9 protein and sgRNA. The cleavage products are then analyzed by gel electrophoresis to confirm the activity and specificity of Cas9. The production of in-vitro recombinant Cas9 by in-fusion cloning involves the assembly of the Cas9 gene from its constituent parts using the in-fusion cloning method.

### 3.1 Experimental design methodology in research

The experimental study design was implemented for this project, and the research was carried out within the laboratory facilities housed in the biohus at Inland Norway University of Applied Sciences in Hamar, Norway. In pursuit of the study's specified aims and objectives, a meticulously structured experimental design, depicted in figure 1, was employed.

### 3.2 Plasmids DNA materials

Five distinct plasmids DNA containing the ALS1 gene cloned in Zero Blunt Vector (3.5 kb) were provided by Professor Wenche Johansen, my main supervisor, from *Ipomoea batatas* cultivars. These plasmids were named PCR\_IbALS\_S1.1, PCR\_IbALS\_S1.2, PCR\_IbALS\_S4.1, PCR\_IbALS\_S4.2, and PCR\_IbALS\_S4.3, respectively. The plasmids were then used for further experimentation.

### 3.3 Sanger sequencing of ALS1 gene in five distinct plasmids DNA

#### 3.3.1 Cycle sequencing reaction of ALS1 gene in five distinct plasmids DNA

The cycle sequencing reaction was performed to amplify the ALS1 gene in five distinct plasmids DNA. One reaction mixture contained 0.5  $\mu\text{L}$  of plasmid DNA [400 ng/ $\mu\text{L}$ ], 1.6  $\mu\text{L}$  of each forward [M13 Forward (-20)] and reverse [M13 Reverse] primers [2  $\mu\text{M}$ ], 2  $\mu\text{L}$  of BigDye buffer [5X], 0.5  $\mu\text{L}$  of BigDye Terminator v.3.1, and 5.4  $\mu\text{L}$  of nuclease-free water. Two separate master mixes were prepared for the cycle sequencing reaction as table 10 (see Appendix A.8 Tables). Accordingly, 0.5  $\mu\text{L}$  of each plasmid DNA [400 ng/ $\mu\text{L}$ ] was added to each labeled microcentrifuge tubes based on plasmids DNA names. The cycle sequencing reaction was then performed as per Table 1.

Table 1: Cycle sequencing condition for amplifying the ALS1 gene in five distinct plasmids DNA.

Step	Temperature	Time	Number of cycles
Denaturation	95°C	30 seconds	30
Annealing	55°C	30 seconds	
Extension	72°C	2 minutes	

### 3.3.2 Purification of the extension products

After the cycle sequencing reaction of the ALS1 gene, the extension products were purified using the ethanol precipitation method. Upon completion of the cycle sequencing program, the following components were applied to each ten labeled microcentrifuge tubes (each tube contains 10  $\mu$ L of mixture reaction): 2  $\mu$ L of EDTA [125 mM], 2  $\mu$ L of NaOAc [3 M] [pH 5.2], 52  $\mu$ L of 96% EtOH, and 10  $\mu$ L of nuclease-free water. The microcentrifuge tubes were then centrifuged at max speed (14.8 rpm) in a microfuge for 30 minutes at 4°C. The supernatant formed in each microcentrifuge tube was removed carefully by pipetting. Following this, the remaining pellets per each microcentrifuge tube were washed with 70  $\mu$ L of 70% EtOH to remove residual salt and other contaminants. In a subsequent step, all ten microcentrifuge tubes were centrifuged at max speed (14.8 rpm) in a microfuge for 10 minutes. The remaining supernatant in each microcentrifuge tube was removed carefully by pipetting, followed by air-drying of all ten samples for 30 minutes at room temperature until all the ethanol had evaporated. Afterward, all ten pellets were resuspended in 10  $\mu$ L deionized formamide.

### 3.3.3 Capillary electrophoresis

All ten purified samples were then applied to adjoining wells of a 96-well micro-titer plate (from A1 to B2 and from C2 to H2; only water was used to adjust the plate). As the last step, the 96-well micro-titer plate was placed into the ABI 3130xl Genetic Analyser (Applied Biosystems, USA) for capillary electrophoresis.

### 3.3.4 Data analysis of sanger sequencing data using CLC software

Data analysis of Sanger sequencing was performed using the software CLC Main Workbench [Version 7.9.3] provided with the capillary electrophoresis instrument to analyse the sequence data, generate the chromatograms, and then compare the sequence data to a reference sequence (Sweetpotato\_CM008339.1) and perform a BLAST search to identify the sequence of the target fragment. CLC Main Workbench [Version 7.9.3] was used to analyse DNA sequencing data in the following manner:

**Trim sequences:** Raw sequencing data (ten sequences from Sanger sequencing data) were imported into CLC Main Workbench [Version 7.9.3] and analyzed. Sequencing reads were trimmed to remove low-quality regions and adapter sequences.

**Alignment and assembly:** The trimmed reads were aligned to a reference genome (Sweetpotato\_CM008339.1) using the Map Reads to Reference tool in CLC Main Workbench

[Version 7.9.3]. The conflicting regions in the sequence alignment or assembly were identified. The alignment or assembly for errors such as misaligned reads, low-quality regions, or gaps was inspected. The conflicting regions were edited manually. The revised alignment or assembly for new conflicts or errors was checked. These steps were repeated until the alignment or assembly was free of conflicts and errors. Eventually, the edited alignment or group to a reference genome or sequence data was compared to verify accuracy and completeness.

Validation: Selected variants were validated by BLAST, resulting in four distinct plasmids DNA with four different insert fragments. The validation showed that one of the plasmids (PCR\_IbALS1\_S4.2) did not contain the ALS1 gene.

## 3.4 Inserting of the ALS1 insert fragments

### 3.4.1 Transformation of *E. coli* TOP10 cells with ALS1 insert fragments

The heat shock method was used to transform *E. coli* TOP10 cells with ALS1 insert fragments. Briefly, 50  $\mu\text{L}$  of pre-melted *E. coli* TOP10 cells were mixed with 2.5  $\mu\text{L}$  of each purified extension product and incubated on ice for 30 minutes. The mixture was then heat-shocked at 42°C for 45 seconds, followed by immediate transfer to ice for 2 minutes. Then, 448  $\mu\text{L}$  of pre-heated SOC medium was added, and the mixture was incubated at 37°C for 1 hour with shaking (225 rpm).

### 3.4.2 Selection of the transformed cells

Following incubation, 10  $\mu\text{L}$ , 50  $\mu\text{L}$ , and 100  $\mu\text{L}$  of the transformation mixture were spread onto different pre-heated LB agar plates containing kanamycin [50  $\mu\text{g}/\text{mL}$ ]. Subsequently, the transformation residue was centrifuged at 6000 rpm for 5 min in a microfuge (Thermo Scientific™, USA). Then, the supernatant was carefully discarded by pipetting, and 100  $\mu\text{L}$  of pre-heated SOC medium was added to the pellet. The content of this tube was cultured onto a pre-heated LB agar plate containing kanamycin [50  $\mu\text{g}/\text{mL}$ ]. Finally, all plates were inverted and incubated at 37°C overnight.

### 3.4.3 Culturing of putative positive transformants

After incubation at 37°C overnight, bacterial colonies were screened for antibiotic resistance. Putative positive transformants were then selected, named, and cultured in LB broth medium containing



kanamycin [50 µg/mL]. The culture was incubated overnight at 37°C with shaking (225 rpm) for 18 hours.

#### **3.4.4 Plasmids DNA isolation**

All steps of plasmid DNA isolation were performed following the manufacturer's instructions provided with the PureYield™ Plasmid Miniprep System (Promega). The purity and yield of the plasmids DNA were assessed by measuring the absorbance at 260 nm and 280 nm using a NanoDrop (Thermo Scientific™, USA).

#### **3.4.5 Restriction digest analysis of putative positive plasmids DNA**

A restriction digest reaction mixture was prepared by combining approximately 400 ng of isolated plasmid DNA, 1 µL of EcoRV restriction enzyme, 1 µL of NEB Buffer 3.1 [10X], and nuclease-free water to a final volume of 10 µL. The reaction mixture was mixed gently by pipetting up and down several times. Afterward, the reaction mixture was incubated at 37°C for 1 hour. Restriction digest products were analyzed by running them on a 1% agarose gel at 90 volts for 1 hour. The DNA fragments were visualized under UV light by placing the gel on a UV illuminator. The results of the restriction digest analysis were interpreted by comparing the observed DNA fragment sizes to the expected fragment sizes based on the known sequence and restriction enzyme sites of the plasmid DNA.

### **3.5 In-vitro cleavage assay**

#### **3.5.1 sgRNA designing by using CRISPOR software**

sgRNAs used for this experiment were designed using the free online tool, CRISPOR software, in the following manner. The CRISPOR website (<http://crispor.tefor.net/>) was searched, and the appropriate organism and Cas9 enzyme for our target sequence were selected. The target sequence of interest was entered into the tool, ensuring that it met the criteria for specificity and lack of off-target regions. The PAM site for the selected Cas9 enzyme was identified. The CRISPOR tool generated sgRNA sequences complementary to the target sequence and included them in the PAM site. Eventually, five sgRNA sequences (20 bp spacer sequences) were selected with the highest predicted specificity and lowest predicted off-target effects for use in the cleavage experiment. Two sgRNAs were selected for PCR\_IbALS1\_S1.1, and three sgRNAs were selected for PCR\_IbALS1\_S4.1.

Table 2: Designed sgRNAs by using CRISPOR software for targeting PCR\_IbALS1\_S1.1 and PCR\_IbALS1\_S4.1 plasmids DNA.

Plasmid DNA name	sgRNA name	sgRNA sequence (20 nt)
PCR_IbALS1_S1.1	FP_IVT_T7_Spa1_S1.1	CTCACCAGATTGGTGGCGCC
PCR_IbALS1_S1.1	FP_IVT_T7_Spa2_S1.1	AACGCCTCACGAACAATGCG
PCR_IbALS1_S4.1	FP_IVT_T7_Spa1_S4.1	TCGGTACCGATCATCCGCCG
PCR_IbALS1_S4.1	FP_IVT_T7_Spa2_S4.1	AACAGCTGGTTGTCCCCGAT
PCR_IbALS1_S4.1	FP_IVT_T7_Spa3_S4.1	ACCTGAATGGCATACTGAGG

### 3.5.2 In-vitro transcription (IVT) reaction of sgRNA

The in-vitro transcription (IVT) reaction of sgRNA was performed following the manufacturer's instructions provided with the Guide-it™ sgRNA in-vitro transcription and Screening System. The transcribed sgRNA was then analyzed by 1% agarose gel electrophoresis and NanoDrop (Thermo Scientific™, USA) to confirm the size and purity, respectively.

Table 3: Forward primer sequences were designed by using Guide-it™ sgRNA in-vitro transcription and Screening System.

Plasmid DNA target	sgRNA name	Forward primer sequence (58 bp) [from 5' to 3']
PCR_IbALS1_S1. 1	FP_IVT_T7_Spa1_S1.1	CCTCTAATACGACTCACTATAGGCTCACCAGATTGGTGGCGCCGTTTAAGAGCTATGC
PCR_IbALS1_S1. 1	FP_IVT_T7_Spa2_S1.1	CCTCTAATACGACTCACTATAGGAACGCCTCACGAACAATGCGGTTTAAGAGCTATGC
PCR_IbALS1_S4. 1	FP_IVT_T7_Spa1_S4.1	CCTCTAATACGACTCACTATAGGTCGGTACCGATCATCCGCCGTTTAAGAGCTATGC
PCR_IbALS1_S4. 1	FP_IVT_T7_Spa2_S4.1	CCTCTAATACGACTCACTATAGGAACAGCTGGTTGTCCCCGATGTTTAAGAGCTATGC
PCR_IbALS1_S4. 1	FP_IVT_T7_Spa3_S4.1	CCTCTAATACGACTCACTATAGGACCTGAATGGCATACTGAGGGTTTAAGAGCTATGC

### 3.5.3 In-vitro cleavage assay

The in-vitro cleavage assay was performed by incubating the RNP (SpCas9-sgRNA) complex with the amplified plasmid DNA for 1 hour at 37°C, following the tables provided in Appendix A.8, tables 11 and 12. Afterward, to stop the operation of the Cas9, samples were set at 65°C for 15 minutes. Then, the cleaved plasmids DNA were digested and linearized with the restriction enzyme (*MluI* for PCR\_IbALS1\_S1.1 and *NcoI* for PCR\_IbALS1\_S4.1) and incubated for an additional hour at 37°C. It should be mentioned that two different concentrations of Cas9 were used to check the effectiveness of this protein in the process of DNA cleavage for PCR\_IbALS1\_S4.1. In the first series of the

experiment, only the concentration of 30 nM was used, and in the second series of the investigation, twice the concentration of 60 nM was used. Still, the rest of the experiment components did not change. Then, the samples were loaded onto a 1% agarose gel and run at 90 volts for 1 hour. The DNA fragments were visualized under UV light by placing the gel on a UV-illuminator. The cleavage patterns of the treated sample were compared with the untreated sample to confirm the presence of Cas9 cleavage products. Eventually, the results were analyzed to determine the efficiency and specificity of Cas9 cleavage.

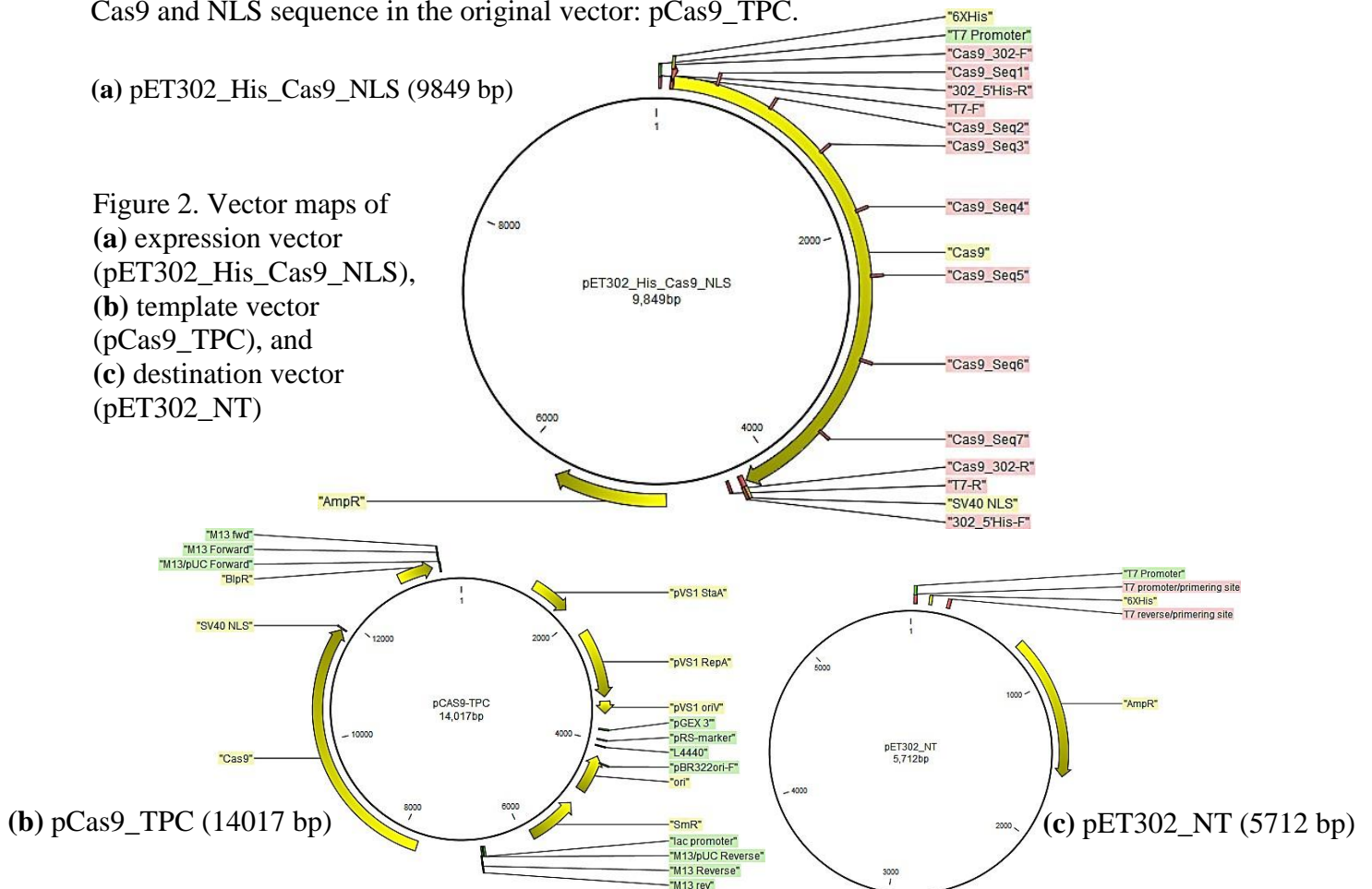
## 3.6 Producing in-vitro recombinant Cas9 by in-fusion cloning

### 3.6.1 Construct expression vector for in-vitro recombinant Cas9 production using CLC software

The pET302-NT vector (5712 bp) (see Appendix A.3.2.2) was used as a destination vector. The pCas9\_TPC vector (14017 bp) (see Appendix A.3.2.1) was used as a DNA template containing the insert to be amplified (CDS of Cas9 and NLS). The expression vector construction (pET302\_His\_Cas9-NLS) (9849 bp) (see Appendix A.3.2.3) was achieved by fusing the insert fragment (4137 bp) from the pCas9\_TPC template, which included 4104 bp corresponding to the CDS of Cas9 and 21 bp compared to the CDS of NLS, with the remaining 12 bp compared to the Cas9 and NLS sequence in the original vector: pCas9\_TPC.

(a) pET302\_His\_Cas9-NLS (9849 bp)

Figure 2. Vector maps of  
(a) expression vector (pET302\_His\_Cas9-NLS),  
(b) template vector (pCas9\_TPC), and  
(c) destination vector (pET302\_NT)



### 3.6.2 Primer designing for in-fusion cloning

Primers with overlapping sequences for in-fusion cloning were designed using the In-Fusion Cloning Primer Design Tool *v1.0* powered by Teselagen. Two oligonucleotide primers were selected to anneal to the 5' and 3' ends of the DNA fragments to be fused. The primers were designed to have overlapping sequences of at least 20-30 bases to allow for efficient fusion of the fragments. The regions of overlap between the fragments were identified. The forward primer for the first fragment was designed to anneal to the 5' end of the fragment and include the overlapping sequence. The reverse primer for the second fragment was designed to anneal to the 3' end of the fragment and include the overlapping sequence. The melting temperature  $T^M$  of the primers was verified to ensure that the  $T_m$  of the forward and reverse primers were similar and in the range of 55-65°C. This will provide efficient amplification during PCR.

Table 4: Designed primers for in-fusion cloning using the In-Fusion Cloning Primer Design Tool *v1.0* powered by Teselagen.

Oligo Name	Sequence	Length (bp)	GC %	$T_m$ (°C)
(Destination vector) Forward primer 302_5'His_F	GTGAATTCGCTCGAGATCG	19	53	58.6
(Destination vector) Reverse primer 302_5'His_R	GTGATGATGATGATGATGCATATG	24	38	58
(Insert) Forward primer Cas9_302-F	CATCATCATCATCACGATAAGAAGTACTCTATCGGACTCG	40	40	66.6
(Insert) Reverse primer Cas9_302-R	CTCGAGCGAATTCCTCAAACCTTCTTCTTCTTAGG	39	46	69.1

### 3.6.3 PCR amplifications of insert (CDS of Cas9 and NLS) and vector (pET302\_NT)

PCR amplification of the insert (CDS of Cas9 and NLS) was performed using primers specific to the insert sequence. The forward (Cas9\_302\_F) and reverse (Cas9\_302\_R) primers were designed to have 40 and 39 nucleotides in length, respectively, and a melting temperature  $T^M$  of 66.6°C and 69.1°C, respectively. To control the experimental conditions, a control sample was also needed, which did not have a DNA template (pCas9\_TPC). For the control sample, the rest of the experimental steps, including PCR and the use of agarose gel to examine the sample, were performed in the same way as the experimental sample. (see Appendix A.8, table 14) Once the PCR had been completed, the products were analyzed by running them on a 1% agarose gel. The gel was run at 90

volts for 1 hour. The DNA fragments were visualized under UV light by placing the gel on a UV illuminator. The gel was analyzed for the presence of the expected product size.

PCR amplification of the vector (pET302\_NT) was performed using primers specific to the destination vector. The forward (302\_5'His\_F) and reverse (302\_5'His\_R) primers were designed to have 19 and 24 nucleotides in length, respectively, and a melting temperature <sup>TM</sup> of 58.6°C and 58°C, respectively. To control the experimental conditions, a control sample was also needed, which did not have a DNA template (pET302\_NT). For the control sample, the rest of the experimental steps, including PCR and the use of agarose gel to examine the sample, were performed in the same way as the experimental sample. (see Appendix A.8, table 15) Once the PCR had been completed, the products were analyzed by running them on a 1% agarose gel. The gel was run at 90 volts for 1 hour. The DNA fragments were visualized under UV light by placing the gel on a UV illuminator. The gel was analyzed for the presence of the expected product size.

### **3.6.4 Isolation and purification of PCR products (gel extraction)**

All steps of isolation and purification of PCR products (gel extraction) were performed following the manufacturer's instructions provided with the NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel). The bands of interest (two separate bands, one for insert fragment, and one for vector fragment) were excised from the agarose gel using a clean and sharp scalpel (avoiding any nearby bands or smears), and NucleoSpin® Gel and PCR Clean-up kit was used to isolate and purify PCR products. Afterward, a 1% agarose gel was prepared to analyze for the presence of the expected product size. The gel was run at 90 volts for 1 hour. The DNA fragments were visualized under UV light by placing the gel on a UV illuminator.

### **3.6.5 In-fusion cloning reaction**

The purified PCR products were combined with the linearized destination vector (pET302\_NT) and the in-fusion cloning enzyme mix (5X In-fusion HD Enzyme Mix) according to the manufacturer's instructions. The reaction was incubated at 50°C for 15 minutes.

### **3.6.6 Transformation of the Stellar competent cells with in-fusion cloning mixture**

The in-fusion cloning reaction was applied to Stellar competent cells and transformed using heat shock. One round bottom tube was labeled, and then 50 µL of pre-melted Stellar capable cells were mixed with 2.5 µL of in-fusion cloning mixture and incubated on ice for 30 minutes. The mixture was

then heat-shocked at 42°C for 45 seconds and immediately transferred to ice for 2 minutes. Then, 448 µL of pre-heated SOC medium was added, and the mixture was incubated at 37°C for 1 hour with shaking (225 rpm).

### 3.6.7 Selection of the transformed cells

After incubation, the transformation mixture was spread onto pre-heated LB agar plates containing ampicillin [100 µg/mL]. The transformation residue was then centrifuged at 6000 rpm for 5 min in a microfuge (Thermo Scientific™, USA). The supernatant was carefully discarded by pipetting, and the pellet was resuspended in 100 µL of pre-heated SOC medium. The content of this tube was cultured onto a pre-heated LB agar plate containing ampicillin [100 µg/mL]. Finally, all plates were incubated inverted at 37°C overnight.

### 3.6.8 Verification by colony PCR

After overnight culture of transformed cells, a master mix was prepared according to the table below to perform colony PCR.

Table 5: Master mixes preparation for verification by colony PCR.

Components	Master mix
Buffer B1 [1X]	10 µL
Forward primer (Cas9_302_F) [0.2 µM]	4 µL
Reverse primer (Cas9_302_R) [0.2 µM]	4 µL
MgCl <sub>2</sub> [1.5 mM]	12 µL
HOT firepol [2.5 U/µL]	5 µL
dNTPs [200 µM]	4 µL
Nuclease-free water	151 µL

After overnight culture of transformed cells, a sterile pipette tip was used to pick a bacterial colony of interest from each LB agar plate. The colony was then transferred to a PCR strip containing 20 µL of master mix. PCR was performed using a thermocycler with the following conditions:

Table 6: PCR reaction for colony PCR was performed as table below.

Step	Temperature	Time	Number of cycles
Initial denaturation	95°C	15 minutes	1
Denaturation	95°C	15 seconds	35
Annealing	64°C	15 seconds	
Extension	72°C	4 minutes	
Final extension	72°C	10 minutes	1
Hold	4°C	-	-

The PCR products were analyzed by running them on a 1% agarose gel and visualizing the bands under UV light. The gel was run at 90 volts for 1 hour. The DNA fragments were visualized under UV light by placing the gel on a UV-illuminator. The gel was then analyzed for the presence of the expected product size.

### **3.6.9 Culturing putative positive pET302\_His\_Cas9\_NLS transformation**

After overnight incubation of transformed cells, a sterile pipette tip was used to pick a bacterial colony of interest from each LB agar plate. The colony was then transferred to a sterile plastic culture tube (15 mL) containing 4  $\mu$ L of LB broth medium containing ampicillin [100  $\mu$ g/mL]. The culture was incubated overnight at 37°C with shaking (225 rpm) for 18 hours.

### **3.6.10 Plasmid DNA isolation**

The plasmid DNA isolation procedure was performed according to the manufacturer's instructions provided with the PureYield™ Plasmid Miniprep System (Promega). The purity and yield of the plasmid DNA were assessed by measuring the absorbance at 260 nm and 280 nm using a NanoDrop spectrophotometer (Thermo Scientific™, USA).

### **3.6.11 Restriction digest analysis of putative positive pET302\_His\_Cas9\_NLS**

A restriction digest reaction mixture was prepared by combining approximately 200 ng of isolated plasmid (pET302\_His\_Cas9\_NLS), 1  $\mu$ L of *Xho*I restriction enzyme, and 1  $\mu$ L of rcut smart buffer. The reaction mixture was gently mixed by pipetting up and down several times and incubated at 37°C for 1 hour. The restriction digest products were then analyzed by running them on a 1% agarose gel. The gel was run at 90 volts for 1 hour, and the DNA fragments were visualized under UV light by placing the gel on a UV-illuminator. The observed DNA fragment sizes were compared to the expected fragment sizes based on the known sequence and restriction enzyme sites of the plasmid DNA.

## 4. Results

### 4.1 ALS1 gene verification

#### 4.1.1 Sanger sequencing technique was performed to determine the sequence of insert (ALS1 gene)

Sanger sequencing was employed to determine the nucleotide sequence of the ALS1 gene inserts within five different plasmids: PCR\_IbALS\_S1.1, PCR\_IbALS\_S1.2, PCR\_IbALS\_S4.1, PCR\_IbALS\_S4.2, and PCR\_IbALS\_S4.3. The obtained raw sequence data underwent cleaning and trimming processes. Subsequently, the sequences related to ten amplicons for gDNA were assembled to generate the ALS1 consensus sequence. Then, the obtained sequencing chromatograms underwent thorough analysis to ascertain the accurate nucleotide sequence within the ALS1 gene inserts.

The Sanger sequencing chromatograms of all plasmids revealed clear and high-quality sequences, confirming the precise determination of the insert's nucleotide sequence. The Sanger sequencing method was conducted using specific equipment and primer sequences following established protocols to ensure accurate and reliable results.

#### 4.1.2 Sequence alignment data for the confirmation of the existence of the insert (ALS1 gene)

The consensus sequences of each individual were aligned with their respective reference sequences (Sweetpotato\_CM008339.1, as mentioned in Appendix A.4). This alignment allowed for the analysis of the actual gene length or the length of mRNA CDS within the consensus sequences. The obtained results were duly noted and documented.

The sequence alignment revealed a substantial degree of similarity between the sequences, strongly indicating the presence of the ALS1 gene within four of the plasmids (Zero\_Blunt\_pCR\_S1.1, Zero\_Blunt\_pCR\_S1.2, Zero\_Blunt\_pCR\_S4.1, and Zero\_Blunt\_pCR\_S4.3). Additionally, the sequences from these four plasmids were translated into corresponding amino acid sequences, and alignment based on these amino acid sequences (refer to Appendix A.6 and A.7) further confirmed the existence of the ALS1 gene.

However, one plasmid (Zero\_Blunt\_pCR\_S4.2) did not confirm the existence of the ALS1 gene. Possible factors contributing to this discrepancy may include the absence of the ALS1 gene in



that specific plasmid, sequencing artifacts, or other technical limitations during the sequence analysis process.

Table 7: The insert was amplified by PCR using specific primers and cloned into the zero blunt vector (see Appendix A.5.2). The genomic consensus sequence was aligned with the reference genomic sequence (Sweetpotato\_CM008339.1, see Appendix A.4) To analyse the gene structure of the ALS1 gene. The length of the coding genomic sequence was determined for each plasmid sample. The coding genomic sequence was 836 bp for Zero\_Blunt\_pCR\_S1.1 and Zero\_Blunt\_pCR\_S1.2 (sequence details in Appendix A.5.1), while it was 980 bp for Zero\_Blunt\_pCR\_S4.1 and 977 bp for Zero\_Blunt\_pCR\_S4.3.

Insert name	Length (bp)	Gene	Product (enzyme)	Organism	Sequence
Zero_Blunt_pCR_S1.1	836	ALS1	Aceto Lactate Synthase	<i>Ipomoea batatas</i>	
Zero_Blunt_pCR_S1.2	836	ALS1	Aceto Lactate Synthase	<i>Ipomoea batatas</i>	See Appendix
Zero_Blunt_pCR_S4.1	980	ALS1	Aceto Lactate Synthase	<i>Ipomoea batatas</i>	A.5.1
Zero_Blunt_pCR_S4.3	977	ALS1	Aceto Lactate Synthase	<i>Ipomoea batatas</i>	

#### 4.1.3 Restriction digestion was performed to analyse the putative positive colonies containing plasmids DNA

The restriction digestion analysis was performed as a pivotal step to verify the presence of the desired DNA fragments within the putative positive colonies containing plasmids DNA (PCR\_IbALS1\_S1.1, PCR\_IbALS1\_S1.2, PCR\_IbALS1\_S4.1, and PCR\_IbALS1\_S4.3). The plasmids were enzymatically digested with the *EcoRV* restriction enzyme, targeting specific regions, followed by the separation of resulting fragments using agarose gel electrophoresis.

During the interpretation of the gel electrophoresis analysis, distinct bands were observed, indicating successful digestions in all four colonies tested. Notably, these observed bands precisely corresponded to the anticipated fragment sizes (4400 bp, 4400 bp, 4300 bp, and 4300 bp for PCR\_IbALS1\_S4.1, PCR\_IbALS1\_S4.3, PCR\_IbALS1\_S1.1, and PCR\_IbALS1\_S1.2, respectively). This alignment was based on the known restriction enzyme recognition sites and the sequences of the inserts, unequivocally confirming the presence of the sought-after DNA fragments within the colonies.

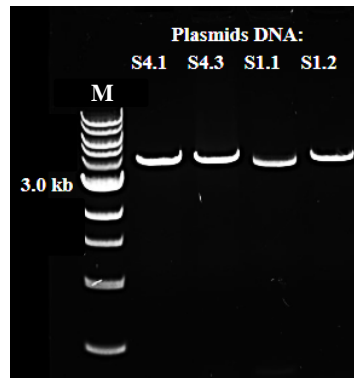


Figure 3. The figure shows the agarose gel electrophoresis of the plasmids digested by *EcoRV* enzyme. The plasmids contain the insert (ALS1 gene) and the vector (zero blunt vector) (see Appendix A.3.1). The gel electrophoresis image reveals the size and presence of the digested plasmid fragments. M denotes the 1 kb DNA ladder used as a DNA marker for size estimation. S4.1, S4.3, S1.1, and S1.2 refer to PCR\_IbALS1\_S4.1, PCR\_IbALS1\_S4.3, PCR\_IbALS1\_S1.1, and PCR\_IbALS1\_S1.2 respectively.

## 4.2 In-vitro cleavage assay

### 4.2.1 Five sgRNA designed for plasmids DNA targets (PCR\_IbALS1\_S1.1 and PCR\_IbALS1\_S4.1)

In this investigation, the CRISPOR software was employed to devise sgRNAs targeting the coding sequence (CDS) of the ALS1 gene integrated into plasmid DNA constructs, namely "PCR\_IbALS1\_S1.1" and "PCR\_IbALS1\_S4.1". A total of five forward primers were formulated to generate five distinct sgRNAs, with two directed towards the cloned ALS1 fragment in plasmid DNA "PCR\_IbALS1\_S1.1" and three aimed at the ALS1 fragment within plasmid DNA "PCR\_IbALS1\_S4.1" (refer to table 8). The ALS1 gene was the focal gene of interest.

CRISPOR facilitated the identification of potential sgRNA sequences based on the provided target gene sequence. The selection criteria prioritized sgRNAs with the highest anticipated efficiency scores and minimal off-target effects. To validate the specificity of the chosen sgRNA, the BLAST tool was employed to scrutinize potential off-target sites within the genome. This scrutiny revealed no substantial sequence similarity between the selected sgRNA and other genomic regions.

To generate DNA templates containing sgRNA encoding sequences, five IVT-forward primers were designed based on five different sgRNA spacer sequences. The IVT-forward primers were synthesized by Invitrogen and used to amplify the sgRNA templates by PCR. The PCR products were then subjected to in-vitro transcription and purification to produce high amounts of sgRNA.

Table 8: Five sgRNA spacer sequences designed for in-vitro transcription sgRNA for plasmid DNAs "PCR\_IbALS1\_S1.1" and "PCR\_IbALS1\_S4.1" with the length of 20 nucleotides. From 5' to 3'.

Plasmid DNA	sgRNA	sgRNA sequence (20 nt)	PAM	Off-target
PCR_IbALS1_S1.1	FP_IVT_T7_Spa1_S1.1	CTCACCAGATTGGTGGCGCC	GGG	3
PCR_IbALS1_S1.1	FP_IVT_T7_Spa2_S1.1	AACGCCTCACGAACAATGCG	AGG	5
PCR_IbALS1_S4.1	FP_IVT_T7_Spa1_S4.1	TCGGTACCGATCATCCGCCG	CGG	3
PCR_IbALS1_S4.1	FP_IVT_T7_Spa2_S4.1	AACAGCTGGTTGTCCCCGAT	TGG	5
PCR_IbALS1_S4.1	FP_IVT_T7_Spa3_S4.1	ACCTGAATGGCATACTGAGG	AGG	7

Table 9: Five IVT-forward primers were designed by using Guide-it™ sgRNA in-vitro transcription and Screening System. (see Appendix A.1.2)

Plasmid DNA target	sgRNA name	IVT-forward primer sequence (58 bp) [from 5' to 3']
PCR_IbALS1_S1.1	FP_IVT_T7_Spa1_S1.1	CCTCTAATACGACTCACTATAGGCTCACCAGATTGGTGGCGCCGTTTAAGAGCTATGC
PCR_IbALS1_S1.1	FP_IVT_T7_Spa2_S1.1	CCTCTAATACGACTCACTATAGGAACGCCTCACGAACAATGCGGTTTAAGAGCTATGC
PCR_IbALS1_S4.1	FP_IVT_T7_Spa1_S4.1	CCTCTAATACGACTCACTATAGGTCGGTACCGATCATCCGCCGGTTTAAGAGCTATGC
PCR_IbALS1_S4.1	FP_IVT_T7_Spa2_S4.1	CCTCTAATACGACTCACTATAGGAACAGCTGGTTGTCCCCGATGTTTAAGAGCTATGC
PCR_IbALS1_S4.1	FP_IVT_T7_Spa3_S4.1	CCTCTAATACGACTCACTATAGGACCTGAATGGCATACTGAGGTTTAAGAGCTATGC

#### 4.2.2 In-vitro cleavage assay was performed to identify suitable sgRNA for PCR\_IbALS1\_S1.1

The in-vitro cleavage assay was performed to evaluate the precision and efficacy of two distinct sgRNAs, namely FP\_IVT\_T7\_Spa1\_S1.1 and FP\_IVT\_T7\_Spa2\_S1.1, on the plasmid PCR\_IbALS1\_S1.1. This assay aimed to gauge the sgRNAs' capability to induce targeted DNA cleavage by RNP complexes (SpCas9 + in-vitro transcribed sgRNA). The interpretation of the results unfolded as follows.

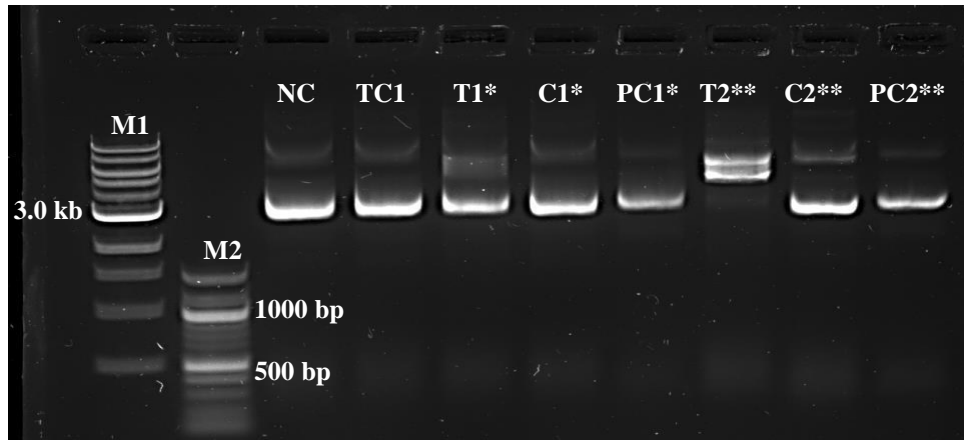


Figure 4. Agarose gel electrophoresis was performed to separate and visualize the DNA fragments obtained from the in-vitro cleavage assay. The plasmid DNA PCR\_IbALS1\_S1.1 was subjected to in-vitro cleavage assay using SpCas9 and two different sgRNAs (FP\_IVT\_T7\_Spa1\_S1.1 and FP\_IVT\_T7\_Spa2\_S1.1). M1 and M2: Refer to the 1 kb and 100 bp DNA ladder utilized as a DNA marker for size determination. NC:

Represents the negative control reaction, which consists of a circular plasmid along with *MluI* restriction enzyme. T: Denotes the treatment reactions, encompassing a circular plasmid combined with SpCas9, sgRNA, and *MluI* restriction enzyme. TC: Signifies the treatment control reaction, comprising a plasmid mixed with SpCas9 enzyme, but lacking sgRNA, along with *MluI* restriction enzyme. C: Represents the sgRNA specific control reactions, involving a plasmid devoid of SpCas9, but containing sgRNA and *MluI* restriction enzyme.

PC: Denotes the positive control, which consists of another circular plasmid, Cas9 and sgRNA with *MluI* restriction. \* and \*\*: Refer to FP\_IVT\_T7\_Spa1\_S1.1 and FP\_IVT\_T7\_Spa2\_S1.1 respectively.

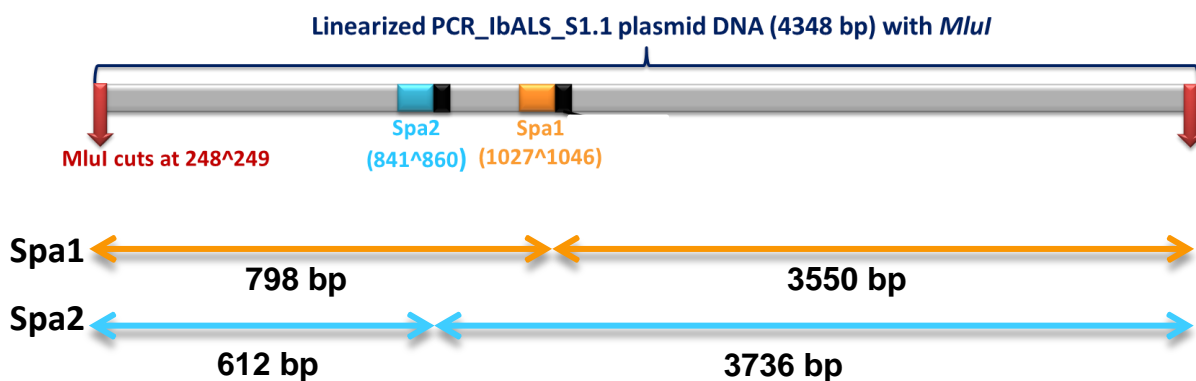


Figure 5. The schematic image presented illustrates the precise cleavage site by the RNP complex on the ALS1 gene. The figure highlights that the SpCas9 enzyme identifies the PAM sequence within the RNP complex, specifically GGG and AGG for Spa1 (FP\_IVT\_T7\_Spa1\_S1.1) and Spa2 (FP\_IVT\_T7\_Spa2\_S1.1), respectively. Both Spa1 and Spa2 bind to their designated locations within the complex, inducing DSBs. These DSBs generate distinct fragments that serve as confirmation of successful sgRNA performance in targeting the ALS1 gene (PCR\_IbALS1\_S1.1).

Upon analyzing lane 3 (T1\*) in figure 4, the in-vitro cleavage assay utilizing FP\_IVT\_T7\_Spa1\_S1.1 sgRNA exhibited no discernible cleavage bands on the gel. Gel analysis showcased an identical pattern to the undigested PCR\_IbALS1\_S1.1 plasmid, indicating a notable absence of cleavage activity. This outcome implies that FP\_IVT\_T7\_Spa1\_S1.1 lacks the efficiency to bind and cleave the target region, resulting in the absence of observable cleavage products.

Conversely, lane 6 (T2\*\*) in figure 4 revealed distinct cleavage bands in the in-vitro cleavage assay using FP\_IVT\_T7\_Spa2\_S1.1 sgRNA, indicating successful DNA cleavage. Gel analysis confirmed the presence of specific cleavage products, aligning with the anticipated fragment sizes (3736 bp and 612 bp) based on the sgRNA target sites within the PCR\_IbALS1\_S1.1 plasmid. This affirmative outcome validates the efficient and specific cleavage activity of FP\_IVT\_T7\_Spa2\_S1.1 (as an influential part of the RNP complex), affirming its capacity to bind and precisely cleave the target region.

The figure also shows the agarose gel electrophoresis of the control reactions, which consist of different combinations of plasmid, SpCas9 and sgRNA with *MluI* digestion. NC denotes the negative control, which has no SpCas9 and no sgRNA. TC denotes the treatment control, which has SpCas9 but no sgRNA. C denotes the sgRNA specific control, which has no SpCas9 but has sgRNA. PC denotes the positive control, which has another plasmid type, SpCas9 and sgRNA. All the control reactions show a single DNA fragment at around 4300 bp, indicating that none of the circular plasmids were cleaved by the RNP complex. This confirms the specificity and activity of the RNP complex for the target sequence of the ALS1 gene.

### **4.2.3 In-vitro cleavage assay was performed to identify suitable sgRNA for PCR\_IbALS1\_S4.1**

The in-vitro cleavage assay was executed to appraise the cleavage efficacy and specificity of three distinct sgRNAs—FP\_IVT\_T7\_Spa1\_S4.1, FP\_IVT\_T7\_Spa2\_S4.1, and FP\_IVT\_T7\_Spa3\_S4.1—targeting plasmid PCR\_IbALS1\_S4.1. The primary aim was to gauge the capacity of these sgRNAs to prompt targeted DNA cleavage. The observations made during the analysis are delineated below.

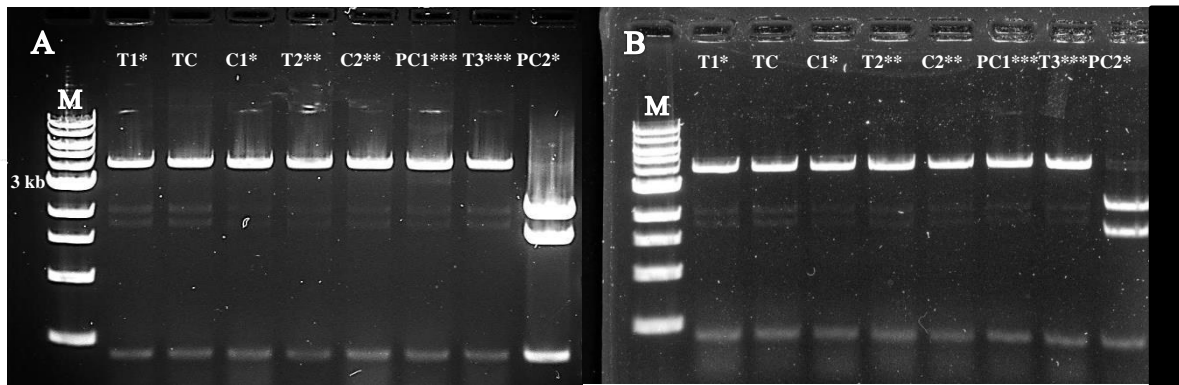


Figure 5. Agarose gel electrophoresis was performed to separate and visualize the DNA fragments obtained from the in-vitro cleavage assay. The plasmid DNA PCR\_IbALS1\_S4.1 was subjected to in-vitro cleavage assay using SpCas9 and three different sgRNAs (FP\_IVT\_T7\_Spa1\_S4.1, FP\_IVT\_T7\_Spa2\_S4.1, and FP\_IVT\_T7\_Spa3\_S4.1). M: Refers to the 1 kb DNA ladder utilized as a DNA marker for size determination. NC: Represents the negative control reaction, which consists of a circular plasmid along with *NcoI* restriction enzyme. T: Denotes the treatment reactions, encompassing a circular plasmid combined with SpCas9, sgRNA, and *NcoI* restriction enzyme. TC: Signifies the treatment control reaction, comprising a plasmid mixed with SpCas9 enzyme, but lacking sgRNA, along with *NcoI* restriction enzyme. C: Represents the sgRNA specific control reactions, involving a plasmid devoid of SpCas9, but containing sgRNA and *NcoI* restriction enzyme. PC: Denotes the positive control, which consists of another circular plasmid, Cas9 and sgRNA with *NcoI* restriction. \*, \*\*, and \*\*\*: Refer to FP\_IVT\_T7\_Spa1\_S4.1, FP\_IVT\_T7\_Spa2\_S4.1, and FP\_IVT\_T7\_Spa3\_S4.1 respectively. The final concentration of SpCas9 in A and B agarose gel electrophoresis figures are 30 nM and 60 nM respectively.

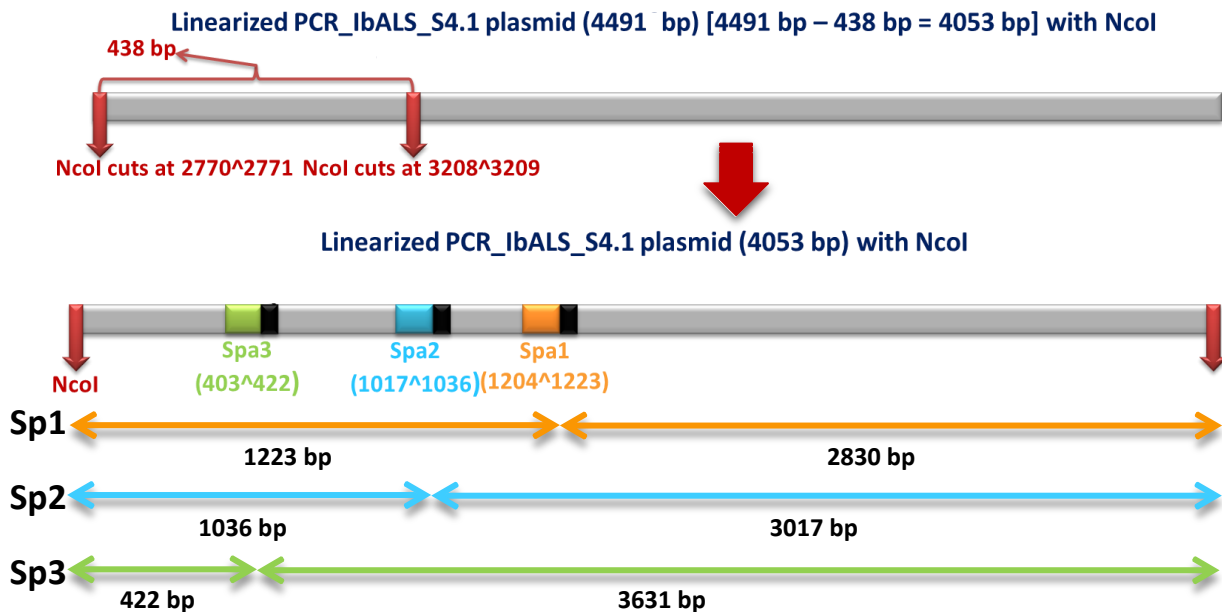


Figure 6. The schematic image presented illustrates the precise cleavage site by the RNP complex on the ALS1 gene. The figure highlights that the SpCas9 enzyme identifies the PAM sequence within the RNP complex, specifically CGG, TGG, and AGG for Spa1 (FP\_IVT\_T7\_Spa1\_S4.1), Spa2 (FP\_IVT\_T7\_Spa2\_S4.1), and Spa3 (FP\_IVT\_T7\_Spa3\_S4.1) respectively. All Spa1, Spa2, and Spa3 bind to their designated locations within the complex, inducing DSBs. These DSBs generate distinct fragments that serve as confirmation of successful sgRNA performance in targeting the ALS1 gene (PCR\_IbALS1\_S4.1).

Upon scrutinizing lanes 1 (T1\*), 4 (T2\*\*), and 7 (T3\*\*\*) in figure 5, the in-vitro cleavage assay employing FP\_IVT\_T7\_Spa1\_S4.1, FP\_IVT\_T7\_Spa2\_S4.1, and FP\_IVT\_T7\_Spa3\_S4.1 sgRNAs did not reveal discernible cleavage bands on the gel. Gel analysis depicted a pattern mirroring the undigested PCR\_IbALS1\_S4.1 plasmid, implying a notable absence of significant cleavage activity. This outcome suggests that FP\_IVT\_T7\_Spa1\_S4.1 fails to efficiently bind and cleave the target region, resulting in the absence of observable cleavage products.

Moreover, the figure displays the agarose gel electrophoresis of the control reactions, which comprise different combinations of plasmid, SpCas9 and sgRNA with *NcoI* digestion. NC indicates the negative control, which lacks SpCas9 and sgRNA. TC indicates the treatment control, which contains SpCas9 but lacks sgRNA. C indicates the sgRNA specific control, which contains sgRNA but lacks SpCas9. PC indicates the positive control, which contains another plasmid type, SpCas9 and sgRNA. All the control reactions except PC2\*, exhibit a single DNA fragment at around 4000 bp, implying that none of the circular plasmids were cleaved by the RNP complex. This verifies the specificity and activity of the RNP complex for the target sequence of the ALS1 gene.

## 4.3 Producing in-vitro recombinant Cas9 by in-fusion cloning

### 4.3.1 PCR was performed to amplify the insert (CDS of Cas9 and NLS) and the vector (pET302\_NT)

The figure depicts the PCR amplification of the insert and the vector from the genomic DNA of sweet potato samples. The insert consists of the CDS of Cas9 and the NLS (pCas9\_TPC), while the vector is pET302\_NT. The gel electrophoresis image shows the size and intensity of the amplified DNA fragments. A DNA marker was loaded in lane ladder (M), which contains fragments from 500 bp to 10 kb. Lane 2 (NCi) and lane 4 (NCv) were negative controls without DNA templates (insert or vector), and no bands were detected in these lanes. This confirms the absence of any nonspecific amplification or contamination during the PCR process. Lane 1 (Ti) and lane 3 (Tv) were experimental lanes with DNA templates (insert and vector), and bands of different sizes and intensities were observed in these lanes. The expected sizes of the insert and the vector were about 4137 bp and 5712 bp, respectively. Lane 1 (Ti) showed a clear and strong band of about 4100 bp, indicating successful and specific amplification of the insert. Lane 3 (Tv) showed a band of about 5700 bp, indicating successful and specific amplification of the vector. The sizes and intensities of the bands in lane 1 (Ti) and lane 3 (Tv) matched the expected values, demonstrating the accuracy and efficiency of the PCR amplification.

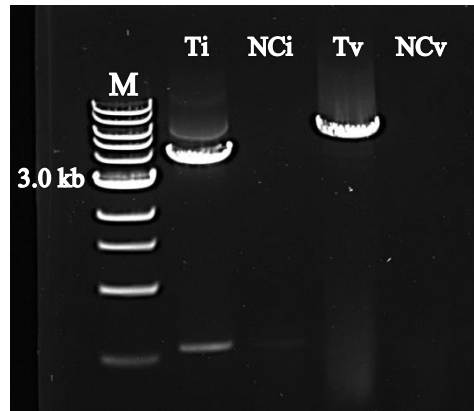


Figure7. The figure shows the PCR amplification of the insert and the vector using two different PCR conditions for each. The insert contains the CDS of Cas9 and the NLS, while the vector is pET302\_NT. M indicates the 1 kb DNA ladder used as a DNA marker for size estimation. Ti indicates the treatment reaction for the insert, which has IFP and IRP and DNAi. NCI indicates the negative control reaction for the insert, which has IFP and IRP but no DNAi. Tv indicates the treatment reaction for the vector, which has VFP and VRP and DNAv. NCv indicates the negative control reaction for the vector, which has VFP and VRP but no DNAv. DNAi refers to insert (pCas9\_TPC); IFP refers to Insert Forward Primer: Cas9\_302\_F; IRP refers to Insert Reverse Primer: Cas9\_302\_R; DNAv refers to pET302\_NT; VFP refers to Vector Forward Primer: 302\_5'His\_F; VRP refers to Vector Reverse Primer: 302\_5'His\_R.

#### 4.3.2 Gel extraction was performed to isolate and purify PCR products

The results obtained from the gel electrophoresis analysis of the gel extraction process for the insert (CDS of Cas9 and NLS) and vector (pET302\_NT) purified PCR products are depicted in Figure 8. During the interpretation of the gel image, several observations were made. Firstly, a DNA marker with known fragment sizes was employed as a reference in lane ladder (M) to estimate the size of the isolated DNA fragments. The marker consisted of fragments ranging from 500 bp to 10 kb.

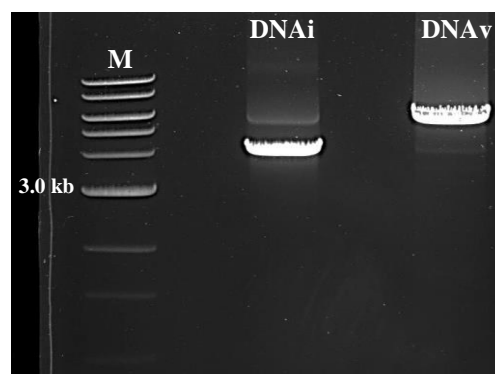


Figure 8. The gel extraction process was carried out to isolate and purify the PCR products, which consisted of the insert (CDS of Cas9 and NLS) and the vector (pET302\_NT). In this context, DNAi refers to the insert (pCas9\_TPC), while DNAv refers to the vector pET302\_NT. M indicates the 1 kb DNA ladder used as a DNA marker for size estimation.



In lane DNA<sub>i</sub>, a distinct and well-defined band of approximately 4100 bp was observed, indicating the successful amplification of the insert (CDS of Cas9 and NLS) PCR product. The observed band aligns precisely with the expected size of the insert, thereby confirming the specific isolation of the target DNA fragment. Moving on to lane DNA<sub>v</sub>, a prominent band of approximately 5700 bp was detected, representing the amplified vector (pET302\_NT) PCR product. This band corresponds exactly to the expected size of the vector, thus validating the successful amplification of the target DNA fragment.

An important observation is that no additional bands or smearing were observed in lanes DNA<sub>i</sub> and DNA<sub>v</sub>, suggesting an absence of contaminants or residual agarose during the gel extraction process. The presence of clear and well-defined bands in both lanes (DNA<sub>i</sub> and DNA<sub>v</sub>) provides strong evidence for the high purity and integrity of the amplified DNA fragments.

#### 4.3.3 In-fusion cloning reaction was performed to clone the insert (CDS of Cas9 and NLS) into the vector (pET302\_NT)

Gel electrophoresis results presented in Figure 9 illustrate the outcome of the in-fusion cloning reaction involving the insert (CDS of Cas9 and NLS) and the vector (pET302\_NT). These results were analyzed to evaluate the success of the cloning process and the presence of the desired DNA constructs (pET302\_His\_Cas9\_NLS). During the interpretation, several observations were made.

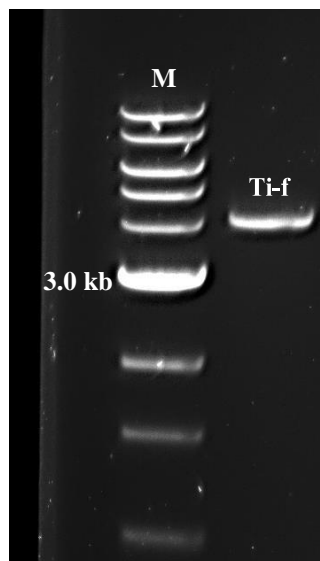


Figure 9. The figure shows the result of the in-fusion cloning reaction, where the insert consisting of the coding sequence (CDS) of Cas9 and NLS was incorporated into the vector pET302\_NT. The DNA ladder (1 kb) labeled as "M" was used as a marker to estimate the size of the DNA fragments. Additionally, the "Ti-f" designation refers to the treatment reaction specifically carried out for the in-fusion cloning process to clone the insert into the vector.

Firstly, a DNA marker containing fragments of known sizes was included in the ladder lane (M) as a reference for estimating the size of the amplified DNA products. The marker encompassed fragments ranging from 500 bp to 10 kb.

However, upon examining the lane corresponding to the in-fusion reaction (Ti-f), no distinct or intense band was observed within the expected size range, which is approximately 10 kb. This indicates an unsuccessful cloning reaction. Furthermore, the absence of a band corresponding to the desired insert (CDS of Cas9 and NLS) and vector (pET302\_NT) constructs (pET302\_His\_Cas9-NLS) suggests that the cloning process did not yield the intended results.

#### 4.3.4 Colony PCR was performed to verify the putative positive pET302\_His\_Cas9-NLS

Figure 10 presents the gel electrophoresis results obtained from the colony PCR verification of the insert (CDS of Cas9 and NLS) into the vector (pET302\_NT), aiming to construct the expression vector (pET302\_His\_Cas9-NLS). Lamentably, the colonies tested did not contain the intended DNA fragments. The gel image was carefully analyzed to confirm the absence of the desired DNA constructs. During interpretation, several observations were made.

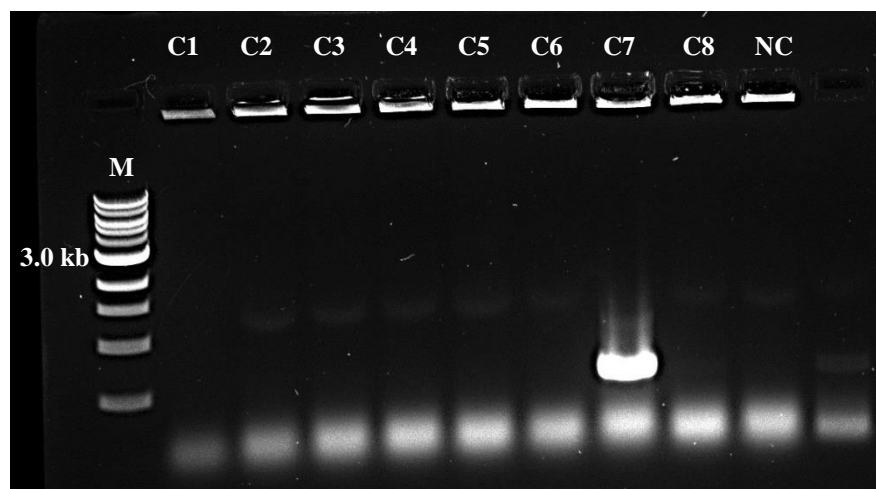


Figure 10. The figure shows the in-fusion cloning verification by colony PCR. The in-fusion cloning was performed to insert the CDS of Cas9 and the NLS into the pET302\_NT vector. The colony PCR was performed to confirm the presence of the desired DNA construct (pET302\_His\_Cas9-NLS) in the colonies. M denotes the 1 kb DNA ladder used as a DNA marker for size estimation. C (1-8) denotes the colony numbers tested by colony PCR. NC denotes the negative control, which has no template DNA.

A DNA marker with known fragment sizes ranging from 500 bp to 10 kb was loaded into the lane labeled "M" as a reference for estimating the size of the amplified DNA products. From lanes 1 to 8 (C1 to C8), no distinct or intense band was observed within the expected size range of approximately

14 kb, indicating unsuccessful colony PCR verification. This absence of a band corresponding to the desired constructs (pET302\_His\_Cas9-NLS) suggests that the colonies tested did not contain the intended DNA fragments.

To validate the experimental procedure and eliminate potential issues, appropriate controls were included. In Lane 9 (NC), representing the negative control (no template DNA), no amplification band was detected. This confirms the absence of contamination or false positives.

#### 4.3.5 Restriction digest was performed to analyze the putative positive pET302\_His\_Cas9-NLS

Results presented in Figure 11 demonstrate the gel electrophoresis analysis conducted after performing the restriction digest on the putative positive pET302\_His\_Cas9-NLS construct. Disastrously, the cloning process did not yield the desired DNA constructs. Furthermore, the primary objective of this analysis was to assess the efficiency of the digestion process and confirm the presence of the anticipated DNA fragments. During the interpretation of the gel image, several observations were made.

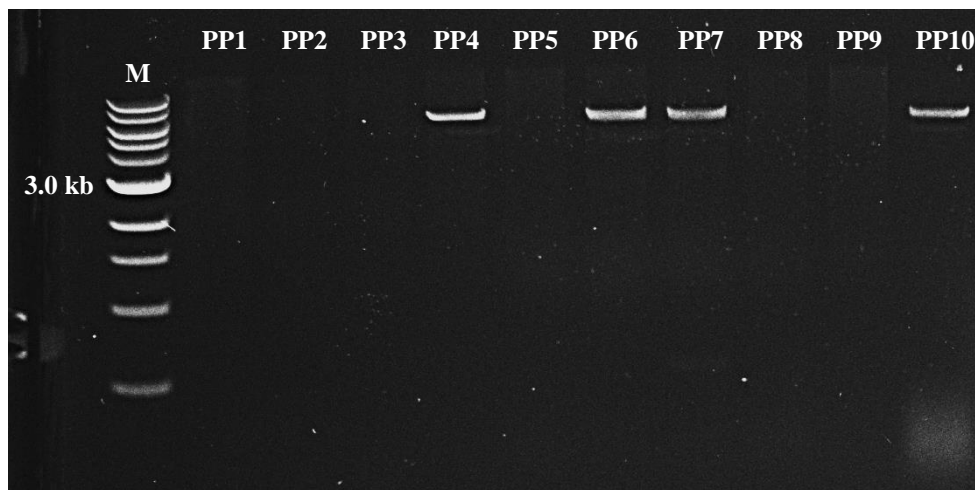


Figure 11. The figure demonstrates the restriction digestion analysis of the putative positive pET302\_His\_Cas9-NLS construct. The restriction digest was performed to confirm the presence and orientation of the insert (CDS of Cas9 and the NLS) in the vector (pET302\_NT). The gel electrophoresis image reveals the size and presence of the digested DNA fragments. M denotes the 1 kb DNA ladder used as a DNA marker for size estimation. PP1 to PP10 denote the putative positive pET302\_His\_Cas9-NLS that may contain the desired DNA construct.

Firstly, a DNA marker with known fragment sizes was loaded into the ladder lane (M), serving as a reference for estimating the size of the digested DNA fragments. This marker encompassed fragments ranging from 500 base pairs to 10 kilobases.

However, upon examining lanes 1 to 10 (PP1 to PP10), it was evident that no distinct or intense bands were observed within the expected size range (~ 9849 bp). This indicates an undesirable outcome of the restriction digest. The absence of bands corresponding to the anticipated digested fragments strongly suggests that the putative positive pET302\_His\_Cas9-NLS construct did not undergo the expected enzymatic cleavage.

## 5. Discussion

Sweet potato is a crop that needs genetic improvement to overcome various challenges. However, it is hard to insert foreign DNA into its genome by traditional methods, which may also cause biosafety and regulation issues. (Abdallah, et al., 2022) CRISPR/Cas is a new method that can edit the genome of sweet potato by using a Cas enzyme and a gRNA. The Cas enzyme can cut the DNA at a specific place guided by the gRNA, and the cell can repair the cut by deleting or inserting a gene. (Zhang, D., et al., 2020) Many studies have used CRISPR/Cas to edit the genome of sweet potato for different goals. (Hou, X., et al., 2023) (Xiao, Y., et al., 2023) (Tussipkan, D., & Manabayeva, S. A. (2021). But most of them have used DNA-based delivery methods, which may still have some problems in terms of biosafety and regulation. So, it is important to develop and optimize DNA-free delivery methods for sweet potato, such as RNP-based genome editing. (Tsanova, et al., 2021)

This study focused on establishing tools for CRISPR/Cas9 RNP-mediated gene editing of the ALS1 gene in *Ipomoea batatas* (sweet potato). The ALS1 gene encodes acetolactate synthase, an enzyme that is involved in the biosynthesis of branched-chain amino acids and confers herbicide resistance. The ALS1 gene is a reliable target for gene editing because it has a single-copy, conserved sequence among different sweet potato varieties, and it can cause phenotypic alterations, such as chlorosis and dwarfism, upon mutation. (Sedeek, et al., 2019) The study designed and evaluated five sgRNAs that can direct the SpCas9 nuclease to cleave the ALS1 gene. The study also attempted to produce recombinant Cas9 protein by in-fusion cloning and compare its performance with commercial Cas9. The findings revealed that one sgRNA (FP\_IVT\_T7\_Spa2\_S1.1) was capable of cleaving the ALS1 gene in a plasmid DNA (PCR\_IbALS1\_S1.1) that contained the ALS1 sequence. However, the study encountered difficulties in producing recombinant Cas9 protein due to the large size of the insert fragment (CDS of Cas9 and NLS ~ 9849 bp). Therefore, this study has contributed to the development and optimization of CRISPR/Cas9 RNP-based gene editing of sweet potato, but it also highlighted some challenges and limitations that need to be addressed in future research.

### 5.1 Sanger sequencing and PCR of ALS1 gene in five distinct plasmids

Sanger sequencing and PCR are two complementary techniques for analyzing DNA sequences. Sanger sequencing is a method for determining the nucleotide sequence of a DNA fragment by using chain-terminating dideoxynucleotides and electrophoresis. (Crossley, et al., 2020) PCR is a method for amplifying a specific DNA fragment by using primers, DNA polymerase, and thermal cycling.

In this study, sanger sequencing and PCR were used to identify and characterize the ALS1 gene in five distinct plasmids. The results of this study demonstrate the successful application of sanger sequencing and BLAST search to confirm the presence and identity of the ALS1 gene insert in four out of five plasmids. These methods are reliable and accurate for determining the nucleotide and amino acid sequences of the insert, as well as for comparing them with the reference sweet potato gene and protein sequences. The results also show that the ALS1 gene insert was cloned and maintained without any mutations or alterations in four plasmids, indicating the high fidelity and stability of the cloning and transformation processes. However, one plasmid (PCR\_IbALS1\_S4.2) did not contain the ALS1 gene insert, as evidenced by the lack of similarity and alignment with the reference sequence. This could be due to several reasons, such as:

- The plasmid DNA was contaminated or degraded during the extraction or purification steps, resulting in the loss of the insert or the generation of false signals.
- The plasmid DNA was not properly transformed into the host cells, resulting in the failure of the insert integration or the selection of the empty vector.
- The plasmid DNA was subjected to recombination or rearrangement events during the replication or maintenance in the host cells, resulting in the deletion or modification of the insert.

These potential causes could be further investigated by performing additional experiments, such as:

- Repeating the plasmid DNA extraction and purification steps, and verifying the quality and quantity of the DNA by using spectrophotometry or gel electrophoresis.
- Repeating the plasmid DNA transformation and selection steps, and verifying the presence and expression of the insert by using PCR or RT-PCR.
- Performing restriction enzyme digestion or Southern blot analysis on the plasmid DNA, and verifying the size and location of the insert by using gel electrophoresis or hybridization.

These experiments could help to identify and resolve the problem of the missing ALS1 gene insert in the plasmid Zero\_Blunt\_pCR\_S4.2, and to ensure the consistency and validity of the results. The confirmation of the ALS1 gene insert in four plasmids is an important step for the development and optimization of CRISPR/Cas9 RNP-based gene editing in sweet potato. The ALS1 gene is a suitable target for gene editing because it has four advantages:

- It is a single and stable gene that does not vary among different sweet potato varieties.
- It is a widely expressed gene that can be detected and measured in different parts of the plant and at different growth stages.

- It is a herbicide-responsive gene that can be used to select the edited plants by using herbicides, such as sulfonylurea and imidazolinone, that affect its function.
- It is a phenotype-associated gene that can be used to evaluate the editing efficiency by seeing the changes in the plant appearance, such as yellowing and stunting, which happen from its mutation.

By using the confirmed plasmids as templates, the next step of this study is to design and synthesize the sgRNAs that can guide the Cas9 nuclease to cleave the ALS1 gene.

## 5.2 PCR-based primer design and DNA amplification

PCR is a powerful technique for amplifying specific DNA sequences from a complex mixture of nucleic acids. However, the success and accuracy of PCR depend largely on the design and quality of the primers used in the reaction. Primers are short synthetic oligonucleotides that anneal to the complementary regions of the target DNA and serve as the starting point for DNA polymerase extension. Therefore, primers should be designed to have optimal properties such as length, melting temperature, GC content, specificity, and lack of secondary structures or self-complementarity. These factors can affect the efficiency, specificity, and fidelity of PCR amplification.

In this study, M13 Forward (-20) and M13 Reverse primers were used for cycle sequencing reaction of ALS1 gene in five distinct plasmids DNA (see Appendix A.1.1). A single band of the expected size (~ 4.3 kb) was observed for each plasmid, indicating that the insert DNA (ALS1 gene) was successfully cloned into the zero blunt vector (see Appendix A.3.1). M13 Forward (-20) and M13 Reverse are oligonucleotides that are complementary to the sequences flanking the multiple cloning site of the M13 bacteriophage-derived cloning. They are commonly used in DNA sequencing reactions to amplify and read the insert DNA that is cloned into the vector. They are also useful for screening and characterizing the insert DNA by PCR amplification and agarose gel electrophoresis. This experiment-based study demonstrates the importance of using M13 Forward (-20) and M13 Reverse primers for cloning, sequencing, and analysing the insert DNA in the M13-derived vector. These findings show that these primers are reliable, specific, and efficient for PCR and sequencing reactions. A protocol for sanger sequencing and agarose gel electrophoresis using these primers were provided, which can be used for rapid screening and verification of the insert DNA.

## 5.3 sgRNA designing

sgRNA designing is a crucial step for CRISPR/Cas9 gene editing, as it determines the specificity and efficiency of the genome editing process. sgRNA is a single-stranded RNA molecule that consists of a 20-nucleotide target sequence (crRNA) and a Cas9-binding sequence (tracrRNA). The target

sequence is complementary to a region of the genomic DNA that contains a PAM, which is essential for Cas9 recognition and cleavage. The tracrRNA sequence is conserved and forms a duplex with the crRNA sequence, facilitating the interaction with Cas9.

In this study, five sgRNA designed for plasmids DNA targets (PCR\_IbALS\_S1.1 and PCR\_IbALS\_S4.1) for the editing of the ALS1 gene were designed and evaluated. Various online tools and databases were used to select the best candidates for sgRNA design. But only seven of them had the sequence data for potato genome, which is essential for designing sgRNAs for CRISPR/Cas9-mediated genome editing in potato. These tools have different features and benefits, such as detecting PAM sequences, predicting off-target effects, optimizing sgRNA properties, and providing sgRNA synthesis services. (Hajiahmadi, et al., 2019) Finally, CRISPOR software was chosen as the best candidates because of the following reasons: Firstly, it supports more than 150 genomes, including many non-conventional model organisms, and allows users to request new genomes or provide their own custom genomes. (Concordet, et al., 2018) Secondly, it ranks the guide RNAs according to different scores that evaluate potential off-targets in the genome of interest and predict on-target activity. Thirdly, it also recommends the best scores for a particular assay based on a published comparison. Fourthly, it provides a comprehensive solution from selection, cloning and expression of guide RNA as well as providing primers needed for testing guide activity and potential off-targets. Finally, it also offers batch design for genome-wide CRISPR and saturation screens. (Concordet, et al., 2018)

Five sgRNAs designed for plasmids DNA targets (PCR\_IbALS\_S1.1 and PCR\_IbALS\_S4.1) using CDS from mentioned plasmids by following guidelines for sgRNA design, such as choosing a target sequence that is close to the mutation site, high GC content, a G nucleotide at the first position and an A or T nucleotide at the 17th position, and minimal similarity to other genomic regions (refers to table 9).

The performance of the sgRNAs was tested in-vitro, using different methods and parameters, such as PCR, sanger sequencing, and agarose gel electrophoresis. The editing efficiency, specificity, and fidelity of the sgRNAs were measured, and compared them with each other and with the control sgRNAs.

The results showed that only a single sgRNA (FP\_IVT\_T7\_Spa2\_S1.1) performed better than others in terms of editing the ALS1 gene. The results showed that the sgRNAs with higher specificity, lower off-target activity, and higher compatibility with the Cas9 variant had higher editing efficiency and fidelity. This also could enhance the specificity and efficiency of the gene editing in the living system.



## 5.4 Evaluating sgRNA efficiency by in-vitro cleavage assays

The effectiveness of the purified sgRNAs was assessed through an in-vitro cleavage assay. This analysis aimed to evaluate the efficiency of the sgRNAs in facilitating targeted DNA cleavage. Certainly, the efficacy of the Cas9 protein in cleaving the target DNA is profoundly influenced by the precision and design of the sgRNAs directing its action. (Guo, et al., 2023) The significance of meticulously designed single guide RNA (sgRNA) in in-vitro cleavage assays lies in its pivotal role in achieving accurate, efficient, and specific gene editing outcomes. The effectiveness of CRISPR/Cas9-mediated genome editing heavily relies on the precision of the sgRNA sequence, targeting the desired genomic locus. (Tycko, et al., 2016) Comparing various research papers, it's evident that optimal sgRNA design significantly impacts cleavage efficiency and off-target effects. Studies by Liu, Q., et al. (2020) and Chen S. J. (2019) emphasize the importance of minimizing off-target effects through rational sgRNA design, utilizing algorithms that consider sequence complementarity, potential off-target sites, and nucleotide modifications. In contrast, Kocak, D. D.'s work (2019) stresses the influence of secondary structures in sgRNA design, proposing modified nucleotides and structural adjustments to enhance cleavage efficiency. This comparison underscores the criticality of tailored sgRNA design in in-vitro cleavage assays, influencing the accuracy and success of CRISPR-based genome editing.

The results of the in-vitro cleavage assay showed that a single sgRNA (FP\_IVT\_T7\_Spa2\_S1.1) targeting the PCR\_IbALS1\_S1.1 plasmid were able to induce significant DNA cleavage and produce DSB, while none of the three sgRNAs (FP\_IVT\_T7\_Spa1\_S4.1, FP\_IVT\_T7\_Spa2\_S4.1, and FP\_IVT\_T7\_Spa3\_S4.1) targeting the PCR\_IbALS1\_S4.1 plasmid were able to induce significant DNA cleavage. This indicates that these sgRNAs are not suitable for editing the ALS1 gene in the target organism. There are several possible reasons for the failure of these sgRNAs to cleave the target DNA.

One reason could be the low quality or quantity of the sgRNA or the SpCas9 protein used in the assay. The sgRNA and the Cas9 protein are essential components of the CRISPR-Cas9 system, and their optimal concentration and purity are required for efficient and specific DNA cleavage. If the sgRNA or the Cas9 protein were degraded, diluted, or contaminated, they could reduce the cleavage activity or increase the off-target effects. Therefore, it is important to verify the quality and quantity of the sgRNA and the Cas9 protein before performing the assay, and to optimize the reaction conditions accordingly.

Another reason could be the mismatch or the secondary structure of the sgRNA or the target DNA. The sgRNA and the target DNA need to form a stable and complementary base pairing for the Cas9 protein to recognize and cleave the target site. If there is a mismatch between the sgRNA and the

target DNA, or if the sgRNA or the target DNA has a secondary structure that prevents the base pairing, the cleavage efficiency could be reduced or abolished. Therefore, it is important to design the sgRNA with minimal mismatch and secondary structure, and to check the sequence and the structure of the target DNA before performing the assay.

A third reason could be the accessibility or the methylation of the target DNA. The target DNA needs to be accessible and unmethylated for the Cas9 protein to bind and cleave it. If the target DNA is located in a region that is tightly packed or wrapped around histones, or if the target DNA is methylated at the protospacer adjacent motif (PAM) site or the target site, the cleavage activity could be inhibited or blocked. Therefore, it is important to select the target DNA from a region that is accessible and unmethylated, and to treat the target DNA with appropriate enzymes or chemicals before performing the assay.

## 5.5 In-fusion cloning-based in-vitro production and characterization of recombinant Cas9 protein

The aim of this study was to construct and express the pET302\_His\_Cas9-NLS vector, which contains the coding sequence of Cas9 and a nuclear localization signal (NLS) fused to a His-tag. To achieve this goal, several molecular cloning steps, such as in-fusion cloning, colony PCR verification, and restriction digest analysis were performed. However, our results showed that none of these steps were successful, and that we did not obtain the desired DNA constructs.

The in-fusion cloning reaction was performed to insert the CDS of Cas9 and NLS into the pET302\_NT vector, which has a His-tag and a multiple cloning site. We expected to obtain a 10 kb DNA fragment corresponding to the pET302\_His\_Cas9-NLS vector. However, the gel electrophoresis results presented in figure 9 did not show any band of this size in the in-fusion reaction lane (Ti-f). This indicates that the in-fusion cloning reaction did not work, and that the insert and the vector did not join together.

The colony PCR verification was performed to screen the *E. coli* colonies that were transformed with the putative pET302\_His\_Cas9-NLS vector. We expected to amplify a 14 kb DNA fragment corresponding to the pET302\_His\_Cas9-NLS vector. However, the gel electrophoresis results presented in figure 10 did not show any band of this size in any of the colony lanes (C1 to C8). This indicates that the colony PCR verification did not work, and that none of the colonies contained the pET302\_His\_Cas9-NLS vector.

The restriction digest analysis was performed to confirm the presence and orientation of the insert (CDS of Cas9 and NLS) in the putative pET302\_His\_Cas9-NLS vector. We expected to obtain a

9849 bp DNA fragment corresponding to the digested pET302\_His\_Cas9-NLS vector. However, the gel electrophoresis results presented in figure 11 did not show any band of this size in any of the putative positive lanes (PP1 to PP10). This indicates that the restriction digest analysis did not work, and that the putative positive pET302\_His\_Cas9-NLS vector did not undergo the expected enzymatic cleavage.

There are several possible reasons for the failure of the molecular cloning steps in this study.

For the in-fusion cloning reaction, one factor could be the quality and quantity of the insert and vector DNA. If the DNA is not pure, concentrated, or linearized properly, it may affect the efficiency and specificity of the in-fusion reaction. Another factor could be the ratio and amount of the insert and vector DNA. If the ratio is not optimal, it may result in self-ligation of the vector or multiple insertions of the insert. A third factor could be the incubation time and temperature of the in-fusion reaction. (Sleight, et al., 2010) If the time is too short or the temperature is too low, it may not allow enough time for the insert and vector to join together.

For the colony PCR verification, one factor could be the quality and quantity of the colony DNA. If the colony DNA is not sufficient, intact, or representative of the plasmid DNA, it may affect the amplification and detection of the target DNA. Another factor could be the design and concentration of the primers. If the primers are not specific, annealing, or optimal for the target DNA, it may result in nonspecific or no amplification. (Royle, K. E., & Polizzi, K., 2017) A third factor could be the cycling conditions and parameters of the PCR reaction. If the conditions are not suitable for the target DNA, it may affect the yield and quality of the PCR products.

For the restriction digest analysis, one factor could be the quality and quantity of the plasmid DNA. If the plasmid DNA is not pure, concentrated, or supercoiled, it may affect the digestion and separation of the DNA fragments. Another factor could be the choice and concentration of the restriction enzymes. If the enzymes are not compatible, active, or optimal for the target DNA, it may result in incomplete or no digestion. A third factor could be the incubation time and temperature of the digestion reaction. (Molloy, et al., 2004) If the time is too short or the temperature is too high, it may not allow enough time for the enzymes to cleave the target DNA.

The suggestions for further study or research are to repeat the cloning steps with improved and optimized factors, such as using a gel extraction kit, a high-fidelity DNA polymerase, a high-efficiency competent *E. coli* strain, and a colony hybridization assay. Moreover, Cas9 concentration adjustments (in some exceptional cases, concentration reduction), proteinase K digestion, and RNase incorporation for the production process and to perform sequencing, expression, and functional analysis on the pET302\_His\_Cas9-NLS vector and to compare the results with the control vector.

## 6. Conclusion

Firstly, sanger sequencing and BLAST searches confirmed the presence and fidelity of the ALS1 gene insert in four plasmids (PCR\_IbALS\_S1.1, PCR\_IbALS\_S1.2, PCR\_IbALS\_S4.1, and PCR\_IbALS\_S4.3) with 100% similarity to the reference sequence (Sweetpotato\_CM008339). This underscores the accuracy of the cloning and transformation processes, ensuring the integrity of the ALS1 gene without mutations. Conversely, one plasmid (PCR\_IbALS1\_S14.2) exhibited no ALS1 gene insert. Subsequently, two plasmids (PCR\_IbALS\_S1.1 and PCR\_IbALS\_S4.1) were chosen for further analysis through in-vitro cleavage assay.

Secondly, five sgRNAs were meticulously designed and assessed for editing efficacy in ALS1 gene plasmids PCR\_IbALS\_S1.1 and PCR\_IbALS\_S4.1. CRISPOR software was employed for optimal candidate selection. FP\_IVT\_T7\_Spa2\_S1.1 induced significant DNA cleavage and DSBs in PCR\_IbALS1\_S1.1, showcasing pronounced efficacy. In contrast, the three sgRNAs targeting PCR\_IbALS1\_S4.1 exhibited no DNA cleavage capacity.

Finally, constructing the pET302\_His\_Cas9-NLS vector for in-vitro recombinant Cas9 production involved integrating Cas9 and NLS coding sequences into the pET302\_NT vector through in-fusion cloning. However, gel electrophoresis, colony PCR, and restriction digest validation yielded unexpected results, indicating the unequivocal failure of the cloning process and unattained vector. This encapsulates the primary findings and challenges, setting the stage for the subsequent study section.

Thus, this study suggests investigating multiple sgRNAs tailored to specific targets, a minimum of five, optimizing the use of various Cas9 proteins in optimal ratios for precise genomic modifications, and refining in-vitro and in-vivo cleavage assays. Moreover, Cas9 concentration adjustments, proteinase K digestion and RNase incorporation to degrade away the protein in order to visualize the cleaved bands on a gel are suggested for the process of optimal recombinant Cas9 protein production. These recommendations contribute to advancing CRISPR-Cas9 technology in genome editing. Furthermore, exploring alternative insert configurations for in-vitro Cas9 production is advised to improve bacterial transformation efficiency and streamline synthesis. These recommendations collectively contribute to advancing CRISPR-Cas9 technology in genome editing.

## 7. References

- Abdallah, N. A., Prakash, C. S., & McHughen, A. G. (2015). Genome editing for crop improvement: Challenges and opportunities. *GM crops & food*, 6(4), 183–205. <https://doi.org/10.1080/21645698.2015.1129937>
- Aglawe, S. B., Barbadikar, K. M., Mangrauthia, S. K., & Madhav, M. S. (2018). New breeding technique "genome editing" for crop improvement: applications, potentials and challenges. *Biotech*, 8(8), 336. <https://doi.org/10.1007/s13205-018-1355-3>
- Ahmad, A., Munawar, N., Khan, Z., Qusmani, A. T., Khan, S. H., Jamil, A., Ashraf, S., Ghouri, M. Z., Aslam, S., Mubarik, M. S., Munir, A., Sultan, Q., Abd-Elsalam, K. A., & Qari, S. H. (2021). An Outlook on Global Regulatory Landscape for Genome-Edited Crops. *International journal of molecular sciences*, 22(21), 11753. <https://doi.org/10.3390/ijms222111753>
- Asmamaw, M., & Zawdie, B. (2021). Mechanism and applications of CRISPR/Cas-9-mediated genome editing. *Biologics: Targets & Therapy*, 15, 353–361. doi:10.2147/BTT.S326422
- Barrangou, R., & Marraffini, L. A. (2014). CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Molecular cell*, 54(2), 234–244. <https://doi.org/10.1016/j.molcel.2014.03.011>
- Bhattacharya, S., & Satpati, P. (2022). Insights into the Mechanism of CRISPR/Cas9-Based Genome Editing from Molecular Dynamics Simulations. *ACS omega*, 8(2), 1817–1837. <https://doi.org/10.1021/acsomega.2c05583>
- Bovell-Benjamin, A. C. (2006). Sweet Potato: A Review of its Past, Present, and Future Role in Human Nutrition. *Advances in Food and Nutrition Research*, 52, 1-59. [https://doi.org/10.1016/S1043-4526\(06\)52001-7](https://doi.org/10.1016/S1043-4526(06)52001-7)
- Brezgin, S., Kostyusheva, A., Kostyushev, D., & Chulanov, V. (2019). Dead Cas Systems: Types, Principles, and Applications. *International journal of molecular sciences*, 20(23), 6041. <https://doi.org/10.3390/ijms20236041>
- Butler, N. M., Atkins, P. A., Voytas, D. F., & Douches, D. S. (2015). Generation and Inheritance of Targeted Mutations in Potato (*Solanum tuberosum* L.) Using the CRISPR/Cas System. *PLoS one*, 10(12), e0144591. <https://doi.org/10.1371/journal.pone.0144591>
- Chen S. J. (2019). Minimizing off-target effects in CRISPR-Cas9 genome editing. *Cell biology and toxicology*, 35(5), 399–401. <https://doi.org/10.1007/s10565-019-09486>

- Chincinska, I. A., Miklaszewska, M., & Sołtys-Kalina, D. (2022). Recent advances and challenges in potato improvement using CRISPR/Cas genome editing. *Planta*, 257(1), 25. <https://doi.org/10.1007/s00425-022-04054-3>
- Concordet, J. P., & Haeussler, M. (2018). CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic acids research*, 46(W1), W242–W245. <https://doi.org/10.1093/nar/gky354>
- Crossley, B. M., Bai, J., Glaser, A., Maes, R., Porter, E., Killian, M. L., Clement, T., & Toohey-Kurth, K. (2020). Guidelines for Sanger sequencing and molecular assay monitoring. *Journal of veterinary diagnostic investigation : official publication of the American Association of Veterinary Laboratory Diagnosticians, Inc*, 32(6), 767–775. <https://doi.org/10.1177/1040638720905833>
- Eid, A., Alshareef, S., & Mahfouz, M. M. (2018). CRISPR base editors: genome editing without double-stranded breaks. *The Biochemical journal*, 475(11), 1955–1964. <https://doi.org/10.1042/BCJ20170793>
- Gaj, T., Gersbach, C. A., & Barbas, C. F., 3rd (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in biotechnology*, 31(7), 397–405. <https://doi.org/10.1016/j.tibtech.2013.04.004>
- Guo, C., Ma, X., Gao, F., & Guo, Y. (2023). Off-target effects in CRISPR/Cas9 gene editing. *Frontiers in bioengineering and biotechnology*, 11, 1143157. <https://doi.org/10.3389/fbioe.2023.1143157>
- Hajiahmadi, Z., Movahedi, A., Wei, H., Li, D., Orooji, Y., Ruan, H., & Zhuge, Q. (2019). Strategies to Increase On-Target and Reduce Off-Target Effects of the CRISPR/Cas9 System in Plants. *International journal of molecular sciences*, 20(15), 3719. <https://doi.org/10.3390/ijms20153719>
- Hamadani, H., Rashid, S. M., Parrah, J. D., Khan, A. A., Dar, K. A., Ganie, A. A., ... Ali, A. (2021). Traditional farming practices and its consequences. In *Microbiota and Biofertilizers, Vol 2* (pp. 119–128). doi:10.1007/978-3-030-61010-4\_6
- Hille, F., & Charpentier, E. (2016). CRISPR-Cas: biology, mechanisms and relevance. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 371(1707), 20150496. <https://doi.org/10.1098/rstb.2015.0496>
- Hou, X., Guo, X., Zhang, Y., & Zhang, Q. (2023). CRISPR/Cas genome editing system and its application in potato. *Frontiers in genetics*, 14, 1017388. <https://doi.org/10.3389/fgene.2023.1017388>

- Hu, J. H., Miller, S. M., Geurts, M. H., Tang, W., Chen, L., Sun, N., Zeina, C. M., Gao, X., Rees, H. A., Lin, Z., & Liu, D. R. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature*, *556*(7699), 57–63. <https://doi.org/10.1038/nature26155>
- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakata, A. (1987). Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *Journal of bacteriology*, *169*(12), 5429-5433.
- Karvelis, T., Gasiunas, G., Miksys, A., Barrangou, R., Horvath, P., & Siksnys, V. (2013). crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*. *RNA biology*, *10*(5), 841–851. <https://doi.org/10.4161/rna.24203>
- Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z., & Joung, J. K. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*, *529*(7587), 490–495. <https://doi.org/10.1038/nature16526>
- Kocak, D. D., Josephs, E. A., Bhandarkar, V., Adkar, S. S., Kwon, J. B., & Gersbach, C. A. (2019). Increasing the specificity of CRISPR systems with engineered RNA secondary structures. *Nature biotechnology*, *37*(6), 657-666.
- Konstantakos, V., Nentidis, A., Krithara, A., & Paliouras, G. (2022). CRISPR-Cas9 gRNA efficiency prediction: an overview of predictive tools and the role of deep learning. *Nucleic Acids Research*, *50*(7), 3616–3637. doi:10.1093/nar/gkac192
- Koonin, E. V., & Makarova, K. S. (2019). Origins and evolution of CRISPR-Cas systems. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *374*(1772), 20180087. <https://doi.org/10.1098/rstb.2018.0087>
- Kumar, K., Gambhir, G., Dass, A., Tripathi, A. K., Singh, A., Jha, A. K., ... Rakshit, S. (2020). Genetically modified crops: current status and future prospects. *Planta*, *251*(4), 91. doi:10.1007/s00425-020-03372-8
- Lee, J. K., Jeong, E., Lee, J., Jung, M., Shin, E., Kim, Y.-H., ... Kim, J.-S. (2018). Directed evolution of CRISPR-Cas9 to increase its specificity. *Nature Communications*, *9*(1). doi:10.1038/s41467-018-05477-x
- Lee, J., Lim, K., Kim, A., Mok, Y. G., Chung, E., Cho, S.-I., ... Kim, J.-S. (2023). Prime editing with genuine Cas9 nickases minimizes unwanted indels. *Nature Communications*, *14*(1), 1786. doi:10.1038/s41467-023-37507-8

- Liu, Q., Cheng, X., Liu, G., Li, B., & Liu, X. (2020). Deep learning improves the ability of sgRNA off-target propensity prediction. *BMC bioinformatics*, 21(1), 1-15.
- Luo, Z., Yao, Z., Yang, Y., Wang, Z., Zou, H., Zhang, X., ... Huang, L. (2023). Genetic fingerprint construction and genetic diversity analysis of sweet potato (*Ipomoea batatas*) germplasm resources. *BMC Plant Biology*, 23(1), 355. doi:10.1186/s12870-023-04329-1
- Ly, D. N., Iqbal, S., Milroy, S., & Jones, M. G. (2022). Multiplex CRISPR-Cas9 Gene-Editing Can Deliver Potato Cultivars with Reduced Browning and Acrylamide. *Plants*, 12(2), 379. <https://doi.org/10.3390/plants12020379>
- Ma, X., Zhang, X., Liu, H., & Li, Z. (2020). Highly efficient DNA-free plant genome editing using virally delivered CRISPR-Cas9. *Nature plants*, 6(7), 773–779. <https://doi.org/10.1038/s41477-020-0704-5>
- Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F. J., Wolf, Y. I., Yakunin, A. F., van der Oost, J., & Koonin, E. V. (2011). Evolution and classification of the CRISPR-Cas systems. *Nature reviews. Microbiology*, 9(6), 467–477. <https://doi.org/10.1038/nrmicro2577>
- Makarova, K. S., & Koonin, E. V. (2015). Annotation and Classification of CRISPR-Cas Systems. *Methods in molecular biology (Clifton, N.J.)*, 1311, 47–75. [https://doi.org/10.1007/978-1-4939-2687-9\\_4](https://doi.org/10.1007/978-1-4939-2687-9_4)
- Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J., Charpentier, E., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J., Terns, R. M., Terns, M. P., White, M. F., Yakunin, A. F., Garrett, R. A., van der Oost, J., Backofen, R., ... Koonin, E. V. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nature reviews. Microbiology*, 13(11), 722–736. <https://doi.org/10.1038/nrmicro3569>
- Molloy, M. J., Hall, V. S., Bailey, S. I., Griffin, K. J., Faulkner, J., & Uden, M. (2004). Effective and robust plasmid topology analysis and the subsequent characterization of the plasmid isoforms thereby observed. *Nucleic acids research*, 32(16), e129. <https://doi.org/10.1093/nar/gnh124>
- Nawiri, S. O., Oduor, R. O., & Jalemba, A. M. (2017). Genetic engineering of sweet potatoes (*Ipomoea batatas*) using isopentenyl transferase gene for enhanced drought tolerance. *Asian Journal of Agriculture*, 1(02), 85–99. doi:10.13057/asianjagric/g010206



- Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J. D., & Kamoun, S. (2013). Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nature biotechnology*, *31*(8), 691–693. <https://doi.org/10.1038/nbt.2655>
- Nidhi, S., Anand, U., Oleksak, P., Tripathi, P., Lal, J. A., Thomas, G., Kuca, K., & Tripathi, V. (2021). Novel CRISPR-Cas Systems: An Updated Review of the Current Achievements, Applications, and Future Research Perspectives. *International journal of molecular sciences*, *22*(7), 3327. <https://doi.org/10.3390/ijms22073327>
- Paix, A., Folkmann, A., & Seydoux, G. (2017). Precision genome editing using CRISPR-Cas9 and linear repair templates in *C. elegans*. *Methods (San Diego, Calif.)*, *121-122*, 86–93. <https://doi.org/10.1016/j.ymeth.2017.03.023>
- Piatek, A., Ali, Z., Baazim, H., Li, L., Abulfaraj, A., Al-Shareef, S., Aouida, M., & Mahfouz, M. M. (2015). RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. *Plant biotechnology journal*, *13*(4), 578–589. <https://doi.org/10.1111/pbi.12284>
- Puchta, H. (2017). Applying CRISPR/Cas for genome engineering in plants: the best is yet to come. *Current Opinion in Plant Biology*, *36*, 1–8. doi:10.1016/j.pbi.2016.11.011
- Royle, K. E., & Polizzi, K. (2017). A streamlined cloning workflow minimising the time-to-strain pipeline for *Pichia pastoris*. *Scientific reports*, *7*(1), 15817. <https://doi.org/10.1038/s41598-017-16172-0>
- Sedeek, K. E. M., Mahas, A., & Mahfouz, M. (2019). Plant Genome Engineering for Targeted Improvement of Crop Traits. *Frontiers in plant science*, *10*, 114. <https://doi.org/10.3389/fpls.2019.00114>
- Shamshirgaran, Y., Liu, J., Sumer, H., Verma, P. J., & Taheri-Ghahfarokhi, A. (2022). Tools for efficient genome editing; ZFN, TALEN, and CRISPR. *Methods in Molecular Biology (Clifton, N.J.)*, *2495*, 29–46. doi:10.1007/978-1-0716-2301-5\_2
- Shimizu, T., Nakayama, I., Nagayama, K., Miyazawa, T., & Nezu, Y. (2002). Acetolactate Synthase Inhibitors.
- Simmonds, N. (1993). *Sweet Potato: An Untapped Food Resource*. By J. A. Woolfe. Cambridge: Cambridge University Press (1992), pp. 643, hardback £65.00, US\$130.00. ISBN 0-521-40295-6. *Experimental Agriculture*, *29*(1), 132-132. doi:10.1017/S0014479700020512
- Sleight, S. C., Bartley, B. A., Lieviant, J. A., & Sauro, H. M. (2010). In-Fusion BioBrick assembly and re-engineering. *Nucleic acids research*, *38*(8), 2624–2636. <https://doi.org/10.1093/nar/gkq179>

- Srisuwan, S., Sihachakr, D., & Siljak-Yakovlev, S. (2006). The origin and evolution of sweet potato (*Ipomoea batatas* Lam.) and its wild relatives through the cytogenetic approaches. *Plant Science: An International Journal of Experimental Plant Biology*, 171(3), 424–433. doi:10.1016/j.plantsci.2006.05.007
- Steinert, J., Schiml, S., & Puchta, H. (2016). Homology-based double-strand break-induced genome engineering in plants. *Plant Cell Reports*, 35(7), 1429–1438. doi:10.1007/s00299-016-1981-3
- Tsanova, T., Stefanova, L., Topalova, L., Atanasov, A., & Pantchev, I. (2021). DNA-free gene editing in plants: a brief overview. *Biotechnology, Biotechnological Equipment*, 35(1), 131–138. doi:10.1080/13102818.2020.1858159
- Tussipkan, D., & Manabayeva, S. A. (2021). Employing CRISPR/Cas technology for the improvement of potato and other tuber crops. *Frontiers in Plant Science*, 12, 747476.
- Tycko, J., Myer, V. E., & Hsu, P. D. (2016). Methods for Optimizing CRISPR-Cas9 Genome Editing Specificity. *Molecular cell*, 63(3), 355–370. <https://doi.org/10.1016/j.molcel.2016.07.004>
- Uniyal, A. P., Mansotra, K., Yadav, S. K., & Kumar, V. (2019). An overview of designing and selection of sgRNAs for precise genome editing by the CRISPR-Cas9 system in plants. *3 Biotech*, 9(6), 223. <https://doi.org/10.1007/s13205-019-1760-2>
- Wang, H., Wu, Y., Zhang, Y., Yang, J., Fan, W., Zhang, H., Zhao, S., Yuan, L., & Zhang, P. (2018). CRISPR/Cas9-Based Mutagenesis of Starch Biosynthetic Genes in Sweet Potato (*Ipomoea Batatas*) for the Improvement of Starch Quality. *International Journal of Molecular Sciences*, 20(19), 4702. <https://doi.org/10.3390/ijms20194702>
- Watanabe, K. (2015). Potato genetics, genomics, and applications. *Breeding Science*, 65(1), 53–68. doi:10.1270/jsbbs.65.53
- Wong, N., Liu, W., & Wang, X. (2015). WU-CRISPR: characteristics of functional guide RNAs for the CRISPR/Cas9 system. *Genome biology*, 16, 218. <https://doi.org/10.1186/s13059-015-0784-0>
- Xiao, Y., Zhu, M., & Gao, S. (2022). Genetic and Genomic Research on Sweet Potato for Sustainable Food and Nutritional Security. *Genes*, 13(10), 1833. <https://doi.org/10.3390/genes13101833>
- Zhang, D., Zhang, Z., Unver, T., & Zhang, B. (2020). CRISPR/Cas: A powerful tool for gene function study and crop improvement. *Journal of advanced research*, 29, 207–221. <https://doi.org/10.1016/j.jare.2020.10.003>

Zhang, L., Zhao, H., Gan, M., Jin, Y., Gao, X., Chen, Q., ... Wang, Z. (2011). Application of simultaneous saccharification and fermentation (SSF) from viscosity reducing of raw sweet potato for bioethanol production at laboratory, pilot and industrial scales. *Bioresource Technology*, *102*(6), 4573–4579. doi:10.1016/j.biortech.2010.12.115

Zhang, S., Shen, J., Li, D., & Cheng, Y. (2021). Strategies in the delivery of Cas9 ribonucleoprotein for CRISPR/Cas9 genome editing. *Theranostics*, *11*(2), 614–648. <https://doi.org/10.7150/thno.47007>

## A. Appendix

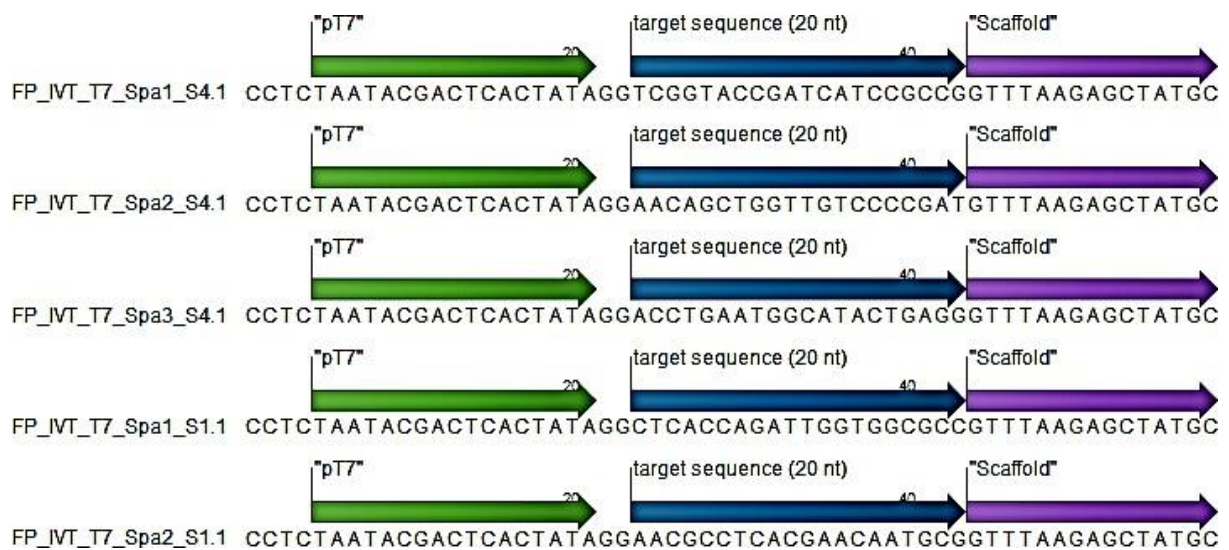
### A.1 Primer sequences

#### A.1.1 Primers used for cycle sequencing reaction of ALS1 gene in five distinct plasmids DNA

M13 Forward (-20) primer: 5' GTAAAACGACGGCCAG 3'

M13 Reverse primer: 5' CAGGAAACAGCTATGAC 3'

#### A.1.2 Primers used for in-vitro transcription (IVT) reaction of sgRNA



#### A.1.3 Primers used for in-fusion cloning

Oligo Name	Sequence
(Destination vector) Forward primer 302_5'His_F	GTGAATTCGCTCGAGATCG
(Destination vector) Reverse primer 302_5'His_R	GTGATGATGATGATGATGCATATG
(Insert) Forward primer Cas9_302-F	CATCATCATCATCACGATAAGAAGTAC TCTATCGGACTCG
(Insert) Reverse primer Cas9_302-R	CTCGAGCGAATTCACTCAAACCTTCCTC TTCTTCTTAGG

---

## A.2 Plasmids DNA sequences

### A.2.1 Plasmids DNA used for cycle sequencing reaction of ALS1 gene in five distinct plasmids DNA and *In-vitro* cleavage assay (\*)

- PCR\_IbALS1\_S1.1 [4348 bp]\*

AGCGCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCA  
GCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT  
AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGG  
CTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTCACACAGGAAACAGCT  
ATGACCATGATTACGCCAAGCTATTTAGGTGACGCGTTAGAATACTCAAGCTAT  
GCATCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAGTGTGCTG  
GAATTCAGGCCTCACGCGCTCCAAGATGATCAGGAATGTCTCCCGCGTCACGA  
GCAGGGCGGGCGTGTTCGCCGCCGAGGGCTACGCGCGGGCCACGGGCTTCCCCG  
GCGTCTGCATCGCCACCTCCGGCCCCGGCGCCACCAATCTGGTGAGCGGACTCG  
CCGACGCGCTTCTGGATAGCTGTCCCATTGTGGCCATTACTGGACAGGTGCCTC  
GGAGGATGATTGGGACAGACGCTTTCAGGAAACCCCAATCGTTGAGGTAACTC  
GATCGATTACGAAGCATAATTATTTAGTTTTAGACGTAGAGGACATCCCTCGCA  
TTGTTTCGTGAGGCGTTTTTCTGGCCAAATCGGGCCGTCCTGGCCCCGGTTTTGAT  
TGATGTTCCCAAGGATATACAGCAACAGCTGGTTATCCCCAATTGGGATCAGCC  
TATGAGATTGCCTGGATACTTATCTCGTTTTGCCAAAACCCCCAAGTGATATGTTG  
TTGGAGCAAATAGTGAGGCTAATTTCTGAGTCCAAGAAGCCTGTCTTTATGTG  
GGGGGTGGGTCTTTACAGTCAAGTGAGGAGTTGAGGCGATTTGTGGAGCTAACT  
GGGATTCCTGTGGCTAGTACTTTAATGGGGCTTGGATCATATCCTTCCTCGGATG  
AACTTTCACTTCAAATGCTGGGAATGCATGGGACTGTGTATGCTAATTAATCTGT  
TGACAAGAGTGATCTGTTGCTTGCATTTGGGGTGAGGTTTCGACGACCGTGTGAC  
TGGAAGCTGGAGGCTTTTGCTAGTCGGGCGAAGATTGTTACATTGATATTGA  
TTCGGCGGAGATTGGGAAGAATAAGCAGCCCCACTGAATTCTGCAGATATCCAT  
CACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCTATAGTG  
AGTCGTATTACAATTCCTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACC  
CTGGCGTTACCCAATTAATCGCCTTGCAGCACATCCCCCTTCGCCAGCTGGCG  
TAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTATA  
CGTACGGCAGTTTAAGGTTTACACCTATAAAAGAGAGAGCCGTTATCGTCTGTT

TGTGGATGTACAGAGTGATATTATTGACACGCCGGGGCGACGGATGGTGATCCC  
CCTGGCCAGTGCACGTCTGCTGTCAGATAAAGTCTCCCGTGAACCTTACCCGGT  
GGTGCATATCGGGGATGAAAGCTGGCGCATGATGACCACCGATATGGCCAGTG  
TGCCGGTCTCCGTTATCGGGGAAGAAGTGGCTGATCTCAGCCACCGCGAAAATG  
ACATCAAAAACGCCATTAACCTGATGTTCTGGGGAATATAAATGTCAGGCATGA  
GATTATCAAAAAGGATCTTCACCTAGATCCTTTTCACGTAGAAAGCCAGTCCGC  
AGAAACGGTGCTGACCCCGGATGAATGTCAGCTACTGGGCTATCTGGACAAGG  
GAAAACGCAAGCGCAAAGAGAAAGCAGGTAGCTTGCAGTGGGCTTACATGGCG  
ATAGCTAGACTGGGCGGTTTTATGGACAGCAAGCGAACCGGAATTGCCAGCTG  
GGGCGCCCTCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTTCT  
TGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGCTCTGATCAAGAGACAGGA  
TGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCC  
GCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGC  
TCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCA  
AGACCGACCTGTCCGGTGCCCTGAATGAACTGCAAGACGAGGCAGCGCGGCTA  
TCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACT  
GAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCT  
GTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCG  
GCGGCTGCATACGCTTGATCCGGCTACCTGCCATTCGACCACCAAGCGAAACA  
TCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGA  
TCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCCGCCAGGCTCA  
AGGCGAGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCT  
TGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCC  
GGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTG  
CTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCG  
CCGCTCCCGATTGCGAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTG  
AATTATTAACGCTTACAATTTCTGATGCGGTATTTTCTCCTTACGCATCTGTGC  
GGTATTTACACCGCATCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCC  
TATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAA  
CCCTGATAAATGCTTCAATAATAGCACGTGAGGAGGGCCACCATGGCCAAGTTG  
ACCAGTGCCGTTCCGGTGCTACCGCGCGCGACGTGCGCGGAGCGGTTCGAGTTC  
TGGACCGACCGGCTCGGGTTCTCCCGGGACTTCGTGGAGGACGACTTCGCCGGT  
GTGGTCCGGGACGACGTGACCCTGTTTCATCAGCGCGGTCCAGGACCAGGTGGTG  
CCGGACAACACCCTGGCCTGGGTGTGGGTGCGCGGCCTGGACGAGCTGTACGCC

GAGTGGTCGGAGGTCGTGTCCACGAACTTCCGGGACGCCTCCGGGCGGCCATG  
ACCGAGATCGGCGAGCAGCCGTGGGGGCGGGAGTTCGCCCTGCGCGACCCGGC  
CGGCAACTGCGTGCACCTTCGTGGCCGAGGAGCAGGACTGACACGTGCTAAAAC  
TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGAC  
CAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAA  
GATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAA  
ACAAAAAACACCGCTACCAGCGGTGGTTTGTGGCCGATCAAGAGCTACCA  
ACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTT  
CTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCT  
ACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAG  
TCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGG  
TCGGGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTA  
CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCG  
AAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGA  
GCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCCG  
GTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGG  
AGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGC  
TGGCCTTTTGTCTACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCG  
TATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCG  
CAGCGAGTCAGTGAGCGAGGAAGCGGAAG

- **PCR\_IbALS1\_S1.2 [4346 bp]**

ATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCT  
ATAGTGAGTCGTATTACAATCACTGGCCGTCGTTTTACAACGTCGTGACTGGG  
AAAACCCTGGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTTCGCCA  
GCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCA  
GCCTATACGTACGGCAGTTTAAGGTTTACACCTATAAAAGAGAGAGCCGTTATC  
GTCTGTTTGTGGATGTACAGAGTGATATTATTGACACGCCGGGGCGACGGATGG  
TGATCCCCCTGGCCAGTGCACGTCTGCTGTCAGATAAAGTCTCCCGTGAACCTTA  
CCCGGTGGTGCATATCGGGGATGAAAGCTGGCGCATGATGACCACCGATATGG  
CCAGTGTGCCGTCTCCGTTATCGGGGAAGAAGTGGCTGATCTCAGCCACCGCG  
AAAATGACATCAAAAACGCCATTAACCTGATGTTCTGGGGAATATAAATGTCAG  
GCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTCACGTAGAAAGCCA

GTCCGCAGAAACGGTGCTGACCCCGGATGAATGTCAGCTACTGGGCTATCTGGA  
CAAGGGAAAACGCAAGCGCAAAGAGAAAGCAGGTAGCTTGCAGTGGGCTTACA  
TGGCGATAGCTAGACTGGGCGGTTTTATGGACAGCAAGCGAACCGGAATTGCC  
AGCTGGGGCGCCCTCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGG  
CTTTCTTGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGCTCTGATCAAGAGA  
CAGGATGAGGATCGTTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTC  
CGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATC  
GGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGGCGCCCGGTTCTT  
TTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAAGACGAGGCAGCG  
CGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTT  
GTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGA  
TCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCA  
ATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCATTCGACCACCAAGCG  
AAACATCGCATCGAGCGAGCACGTA CTCCGGATGGAAGCCGGTCTTGTCGATCAG  
GATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCCGCCAG  
GCTCAAGGCGAGCATGCCCGACGGCGAGGATCTCGTTCGTGACCCATGGCGATG  
CCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTG  
TGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGA  
TATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGG  
TATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTC  
TTCTGAATTATTAACGCTTACAATTTCTGATGCGGTATTTTCTCCTTACGCATCT  
GTGCGGTATTTACACCCGCATCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAA  
CCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACA  
ATAACCCTGATAAATGCTTCAATAATAGCACGTGAGGAGGGCCACCATGGCCA  
AGTTGACCAGTGCCGTTCCGGTGCTCACCGCGCGACGTCGCCGGAGCGGTTCG  
AGTTCTGGACCGACCGGCTCGGGTTCTCCCGGGACTTCGTGGAGGACGACTTCG  
CCGGTGTGGTCCGGGACGACGTGACCCTGTTTCATCAGCGCGGTCCAGGACCAGG  
TGGTGCCGGACAACACCCTGGCCTGGGTGTGGGTGCGCGGCCTGGACGAGCTGT  
ACGCCGAGTGGTCGGAGGTCGTGTCCACGAACTTCCGGGACGCCTCCGGGCCGG  
CCATGACCGAGATCGGCGAGCAGCCGTGGGGGCGGGAGTTCGCCCTGCGCGAC  
CCGGCCGGCAACTGCGTGC ACTTCGTGGCCGAGGAGCAGGACTGACACGTGCT  
AAA ACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTC  
ATGACCAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTA  
GAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCT



---

TGCAAACAAAAAACCACCGCTACCAGCGGTGGTTTGTGGCCGGATCAAGAGC  
TACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATA  
CTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACC  
GCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT  
AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAAGGCGCAG  
CGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGAC  
CTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCC  
CGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTTCGGAACAGGA  
GAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTC  
GGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGC  
GGAGCCTATGGAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTT  
GCTGGCCTTTTGCTCACATGTTCTTTCTGCGTTATCCCCTGATTCTGTGGATAAC  
CGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAG  
CGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCAATACGCAAACCGCC  
TCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGA  
CTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATT  
GGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTG  
AGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCT  
ATTTAGGTGACGCGTTAGAATACTCAAGCTATGCATCAAGCTTGGTACCGAGCT  
CGGATCCACTAGTAACGGCCGCCAGTGTGCTGGAATTCAGGCCTCACGCGCTCC  
AAGATGATGGAATGTCCTCCCGCGTCACGAGCAGGGCGGCGTGTTCCXCCGCCGA  
GGGCTACGCGCGGGCCACGGGCTTCCCCGGCGTCTGCATCGCCACTTCCGGCCC  
CGGCGCCACCAACCTGGTGAGCGGACTCGCTGACGCGCTTCTGGATAGCTGTCC  
CATTGTGGCCATTACTGGACAGGTGCCTCGGAGGATGATTGGGACAGACGCTTT  
CCAGGAAACCCCAATCGTTGAGGTAAGTTCGATCGATTACGAAGCATAATTATT  
AGTTTTAGACGTAGAGGACATCCCTCGCATTGTTTCGTGAGGCGTTTTTCTGGCC  
AAATCGGGCCGTCTGGCCCGGTTTTGATTGATGTTCCCAAGGATATACAGCAA  
CAATTGGTTATCCCCAATTGGGATCAGCCTATGAGATTGCCTGGATACTTATCTC  
GTTTGCCAAAACCCCAAGTGATATGTTGTTGGAGCAAATAGTGAGGCTAATTT  
CTGAGTCAAAGAAGCCTGTCCTTTATGTGGGGGGTGGGTCTTTACAGTCAAGTG  
AGGAGTTGAGGCGATTTGTGGAGCTAACTGGGATTCCTGTGGCTAGTACTTTAA  
TGGGGCTTGGATCATATCCTTCTTCGGATGAACTTCACTTCAAATGCTGGGAAT  
GCATGGGACTGTGTATGCTAATTAATCTGTTGACAAGAGTGATCTGTTGCTTGCA  
TTTGGGGTGAGGTTTCGATGACCGTGTGACTGGGAAGCTGGAGGCTTTTGCTAGT

CGGGCGAAGATTGTTACATTGATATTGATTCCGCGGAGATTGGGAAGAATAAG  
CAGCCCCACTGAATTCTGCAGAT

- **PCR\_IbALS1\_S4.1 [4491 bp]\***

AGCGCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCA  
GCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT  
AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGG  
CTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTCACACAGGAAACAGCT  
ATGACCATGATTACGCCAAGCTATTTAGGTGACGCGTTAGAATACTCAAGCTAT  
GCATCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCCGCCAGTGTGCTG  
GAATTCAGGCTCCGGCCCCGGCGCCACCAATCTCGTCAGTGGCCTCGCCGATGC  
GCTTCTAGATAGCTGTCCGATTGTGCGCCATTACTGGACAAGTGCCGCGGGCGGAT  
GATCGGTACCGATGCCTTCCAGGAAACTCCTATCGTTGAGGTAACGCGTTCGAT  
CACAAAGCATAATTATCTGGTTCTGAATGTAGAGGATATCCCTCGCATTGTCCG  
CGAGGCGTTTTTCTCGCGAAATCGGGTCGTCCTGGCCCCGGTTTTGATTGACGTT  
CCTAAAGATATACAGCAACAGCTGGTTGTCCCCGATTGGGATCAGCCTATGAGA  
TTGCCTGGATACTTGTCTCGCTTGCCTAAACCACCAAACGAAATGCTCTTGGAG  
CAAATAATTAGGCTGATTTCTGAGTCCAAAAGCCAGTTCTGTATGTTGGGGGA  
GGGAGTTTACATTCAAGTGAGGAGTTGAGGCGTTTCGTAGAGCTTACAGGGATT  
CCCGTGGCCAGTACTTTAATGGGGCTTGGATCTTATCCTTGCTCAGATGAGCTCT  
CACTTCAAATGCTGGGAATGCACGGGACTGTGCACGCTAATTATGCTGTTGATA  
AGAGTGATCTTTTGCTTGCCTTTGGAGTTAGGTTTGATGACCGTGTGACTGGTAA  
ATTAGAGGCTTTTGCTAGTCGAGCAAAGATTGTTACATTGATATTGATTCTGCT  
GAGATTGGGAAGAATAAGCAGCCCCATGTATCTGTTTGTGGGGATTTAAAGTTG  
GCTCTGCAGGGGATAAATTCAATGTTGGAGGAGAGGATGAGTGGTAAGCTTAA  
GCTGGATTTCTCGGGTTGGAGACAGGAGGTGTTGGAGCAAAGGCAAAGTTCC  
CCTTGAGTTACAAGACGTTTGGGGATGCTATTCCTCCTCAGTATGCCATTCAGGT  
TCTTGATGAGTTGACTAATGGAAATGCTATAATTACTACTGGCGTTGGGCAGCA  
CCAGATGTGGGCTGCCTGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCG  
AGCATGCATCTAGAGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCAC  
TGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACCTA  
ATCGCCTTGCAGCACATCCCCCTTCGCCAGCTGGCGTAATAGCGAAGAGGCCC  
GCACCGATCGCCCTTCCAACAGTTGCGCAGCCTATACGTACGGCAGTTTAAGG

TTTACACCTATAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTACAGAGTG  
ATATTATTGACACGCCGGGGCGACGGATGGTGATCCCCCTGGCCAGTGCACGTC  
TGCTGTCAGATAAAGTCTCCCGTGAAC TTTACCCGGTGGTGCATATCGGGGATG  
AAAGCTGGCGCATGATGACCACCGATATGGCCAGTGTGCCGGTCTCCGTTATCG  
GGGAAGAAGTGGCTGATCTCAGCCACCGCGAAAATGACATCAAAAACGCCATT  
AACCTGATGTTCTGGGGAATATAAATGTCAGGCATGAGATTATCAAAAAGGATC  
TTCACCTAGATCCTTTTTCACGTAGAAAGCCAGTCCGCAGAAACGGTGCTGACCC  
CGGATGAATGTCAGCTACTGGGCTATCTGGACAAGGGAAAACGCAAGCGCAA  
GAGAAAGCAGGTAGCTTGCAGTGGGCTTACATGGCGATAGCTAGACTGGGCGG  
TTTTATGGACAGCAAGCGAACCGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGG  
TTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTTCTTGCCGCCAAGGATCTGAT  
GGCGCAGGGGATCAAGCTCTGATCAAGAGACAGGATGAGGATCGTTTCGCATG  
ATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTA  
TTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTC  
CGGCTGTCAGCGCAGGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGT  
GCCCTGAATGAACTGCAAGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGAC  
GGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTG  
GCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCT  
GCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGGCGGCTGCATACGCTTGAT  
CCGGCTACCTGCCATTTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGT  
ACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCA  
GGGGCTCGCGCCAGCCGAACTGTTCCGCCAGGCTCAAGGCGAGCATGCCCGACG  
GCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGG  
AAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACC  
GCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCG  
AATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTTCGCAGCG  
CATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAATTATTAACGCTTACAAT  
TTCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATCA  
GGTGGCACTTTTCGGGGAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAA  
TACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAAT  
AATAGCACGTGAGGAGGGCCACCATGGCCAAGTTGACCAGTGCCGTTCCGGTG  
CTCACCGCGCGCGACGTCGCCGGAGCGGTTCGAGTTCTGGACCGACCGGCTCGGG  
TTCTCCCGGGACTTCGTGGAGGACGACTTCGCCGGTGTGGTCCGGGACGACGTG  
ACCCTGTTTCATCAGCGCGGTCCAGGACCAGGTGGTGCCGGACAACACCCTGGCC

TGGGTGTGGGTGCGCGGCCTGGACGAGCTGTACGCCGAGTGGTTCGGAGGTCGT  
GTCCACGAAC TTCGGGACGCCTCCGGGCCGCCATGACCGAGATCGGGCAGC  
AGCCGTGGGGGCGGGAGTTCGCCCTGCGCGACCCGGCCGGCAACTGCGTGCAC  
TTCGTGGCCGAGGAGCAGGACTGACACGTGCTAAACTTCATTTTTAATTTAAA  
AGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAATCCCTTAACGT  
GAGTTTTCGTTCCTACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCT  
TGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCACCGC  
TACCAGCGGTGGTTTGTGGCCGATCAAGAGCTACCAACTCTTTTTCCGAAGGT  
AACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTA  
GTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCT  
AATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTT  
GGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTTCGGGCTGAACGGGGG  
GTTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACC  
TACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGAC  
AGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCC  
AGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCTGGGTTTCGCCACCTCTGACTT  
GAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCC  
AGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGT  
TCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTG  
AGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCG  
AGGAAGCGGAAG

- **PCR\_IbALS1\_S4.3 [4270 bp]**

AGCGCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCA  
GCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT  
AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGG  
CTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTCACACAGGAAACAGCT  
ATGACCATGATTACGCCAAGCTATTTAGGTGACGCGTTAGAATACTCAAGCTAT  
GCATCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAGTGTGCTG  
GAATTCAGGCTCCGGCCCCGGCGCCACCAATCTGGTGAGCGGACTCGCTGACGC  
GCTTCTGGATAGCTGTCCATTGTGGCCATTACTGGACAGGTGCCTCGGAGGAT  
GATTGGGACAGACGCTTTCAGGAAACCCCAATCGTTGAGGTAACCTCGATCGAT  
TACGAAGCATAATTATTTAGTTTTAGACGTAGAGGACATCCCTCGCATTGTTTCGT

---

GAGGCGTTTTTCCTGGCCAAATCGGGCCGTCCTGGCCC GGTTTTGATTGATGTTCC  
CCAAGGATATACAGCAACAATTGGTTATCCCCAATTGGGATCAGCCTATGAGAT  
TGCTGGATACTTATCTCGTTTGCCAAAACCCCCAAGTGATATGTTGTTGGAGCA  
AATAGTGAGGCTAATTTCTGAGTCAAAGAAGCCTGTCCTTTATGTGGGGGGTGG  
GTCTTTACAGTCAAGTGAGGAGTTGAGGCGATTTGTGGAGCTAACTGGGATTCC  
TGTGGCTAGTACTTTAATGGGGCTTGGATCATATCCTTCTTCGGATGAACTTTCA  
CTTCAAATGCTGGGAATGCATGGGACTGTGTATGCTAATTACTCTGTTGACAAG  
AGTGATCTGTTGCTTGCATTTGGGGTGAGGTTTCGATGACCGTGTGACTGGGAAG  
CTGGAGGCTTTTGCTAGTCGGGCGAAGATTGTTACATTGATATTGATTCCGCG  
GAGATTGGGAAGAATAAGCAGCCCCATGTATCGATTTGTGCGGATTTGAAGTTG  
GCCCTGCACTGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGC  
ATCTAGAGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGT  
CGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCA ACTTAATCGCCT  
TGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGA  
TCGCCCTTCCCAACAGTTGCGCAGCCTATACGTACGGCAGTTTAAGGTTTACAC  
CTATAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTACAGAGTGATATTAT  
TGACACGCCGGGGCGACGGATGGTGATCCCCCTGGCCAGTGCACGTCTGCTGTC  
AGATAAAGTCTCCCGTGA ACTTTACCCGGTGGTG CATATCGGGGATGAAAGCTG  
GCGCATGATGACCACCGATATGGCCAGTGTGCCGGTCTCCGTTATCGGGGAAGA  
AGTGGCTGATCTCAGCCACCGCGAAAATGACATCAAAAACGCCATTAACCTGAT  
GTTCTGGGGAATATAAATGTCAGGCATGAGATTATCAAAAAGGATCTTCACCTA  
GATCCTTTTCACGTAGAAAGCCAGTCCGCAGAAACGGTGCTGACCCCGGATGAA  
TGTCAGCTACTGGGCTATCTGGACAAGGGAAAACGCAAGCGCAAAGAGAAAGC  
AGGTAGCTTGCAGTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTTATGGA  
CAGCAAGCGAACCGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTTGGGAAG  
CCCTGCAAAGTAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGGCGCAGG  
GGATCAAGCTCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACA  
AGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTA  
TACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTC  
AGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAA  
TGA ACTGCAAGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTC  
CTTGCGCAGCTGTGCTCGACGTTG TCACTGAAGCGGGAAGGGACTGGCTGCTAT  
TGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGA  
AAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTA

CCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGA  
TGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTC  
GCGCCAGCCGAACTGTTTCGCCAGGCTCAAGGCGAGCATGCCCGACGGCGAGGA  
TCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGG  
CCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCA  
GGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGC  
TGACCGCTTCTCGTGCTTTACGGTATCGCCGCTCCCGATTTCGCAGCGCATCGCC  
TTCTATCGCCTTCTTGACGAGTTCTTCTGAATTATTAACGCTTACAATTTCTGAT  
GCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATCAGGTGGCAC  
TTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCA  
AATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATAGCAC  
GTGAGGAGGGCCACCATGGCCAAGTTGACCAGTGCCGTTCCGGTGCTCACCGCG  
CGCGACGTCGCCGGAGCGGTTCGAGTTCTGGACCGACCGGCTCGGGTTCTCCCG  
GACTTCGTGGAGGACGACTTCGCCGGTGTGGTCCGGGACGACGTGACCCTGTTC  
ATCAGCGCGGTCCAGGACCAGGTGGTGCCGGACAACACCCTGGCCTGGGTGTG  
GGTGCGCGGCCTGGACGAGCTGTACGCCGAGTGGTTCGGAGGTCGTGTCCACGA  
ACTTCCGGGACGCCTCCGGGCCGGCCATGACCGAGATCGGCGAGCAGCCGTGG  
GGGCGGGAGTTCGCCCTGCGCGACCCGGCCGGCAACTGCGTGC ACTTCGTGGCC  
GAGGAGCAGGACTGACACGTGCTAAACTTCATTTTTAATTTAAAAGGATCTAG  
GTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTAACGTGAGTTTTCGT  
TCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTT  
TTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCACCGCTACCAGCGG  
TGTTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTA ACTGGCTT  
CAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCA  
CCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTA  
CCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGA  
CGATAGTTACCGGATAAGGCGCAGCGGTTCGGGCTGAACGGGGGGTTTCGTGCAC  
ACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTG  
AGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCG  
GTAAGCGGCAGGGTTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAA  
ACGCCTGGTATCTTTATAGTCCTGTTCGGGTTTCGCCACCTCTGACTTGAGCGTCG  
ATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAACGCCAGCAACGC  
GGCCTTTTTACGGTTCTTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTCTG  
CGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATAC

---

CGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGG  
AAG

## A.3 Vectors sequences and maps

### A.3.1 Vector used for transformation of *E. coli* TOP10 cells with ALS1 insert fragments

- Zero Blunt Vector sequence [3512 bp]

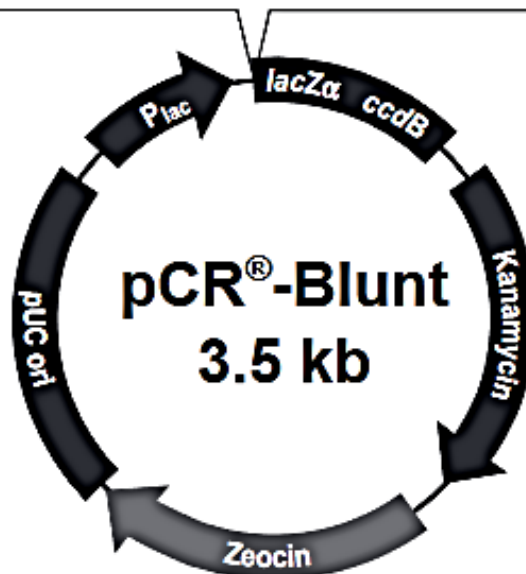
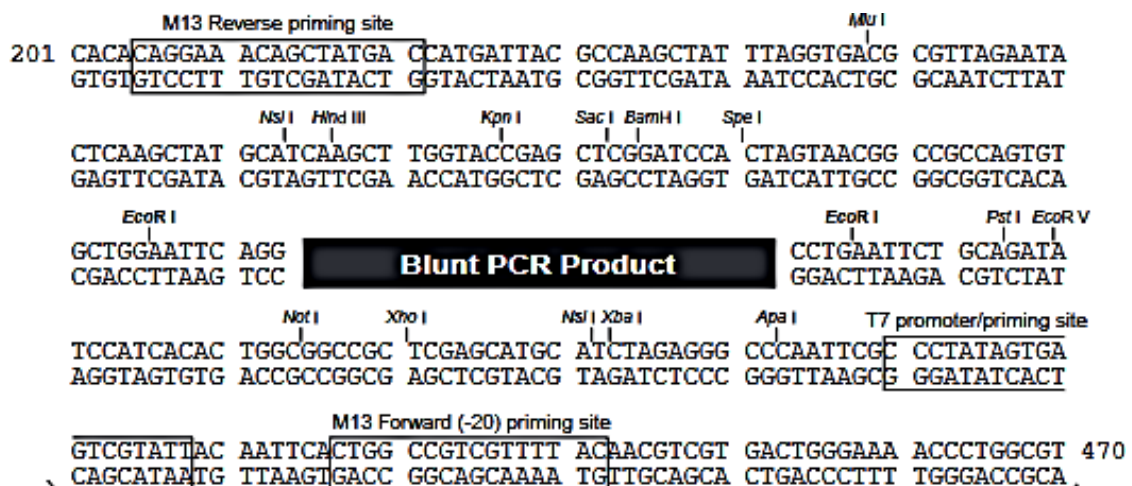
AGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCA  
GCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT  
AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGG  
CTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTCACACAGGAAACAGCT  
ATGACCATGATTACGCCAAGCTATTTAGGTGACGCGTTAGAATACTCAAGCTAT  
GCATCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAGTGTGCTG  
GAATTCAGGCCTGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCAT  
GCATCTAGAGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCAGTGGCC  
GTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGC  
CTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACC  
GATCGCCCTTCCCAACAGTTGCGCAGCCTATACGTACGGCAGTTTAAGGTTTAC  
ACCTATAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTACAGAGTGATATT  
ATTGACACGCCGGGGCGACGGATGGTGATCCCCCTGGCCAGTGCACGTCTGCTG  
TCAGATAAAGTCTCCCGTGAACCTTACCCGGTGGTGCATATCGGGGATGAAAGC  
TGGCGCATGATGACCACCGATATGGCCAGTGTGCCGGTCTCCGTTATCGGGGAA  
GAAGTGGCTGATCTCAGCCACCGCGAAAATGACATCAAAAACGCCATTAACCT  
GATGTTCTGGGGAATATAAATGTCAGGCATGAGATTATCAAAAAGGATCTTCAC  
CTAGATCCTTTTCACGTAGAAAGCCAGTCCGCAGAAACGGTGCTGACCCCGGAT  
GAATGTCAGCTACTGGGCTATCTGGACAAGGGAAAACGCAAGCGCAAAGAGAA  
AGCAGGTAGCTTGCAGTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTTAT  
GGACAGCAAGCGAACCGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTTGGG  
AAGCCCTGCAAAGTAAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGGCGC  
AGGGGATCAAGCTCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGA  
ACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCCGG

CTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCT  
GTCAGCGCAGGGGCGCCCGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCT  
GAATGAACTGCAAGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCG  
TTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGC  
TATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCG  
AGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGG  
CTACCTGCCATTTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTC  
GGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGG  
CTCGCGCCAGCCGAAGTTCGCCAGGCTCAAGGCGAGCATGCCCCGACGGCGA  
GGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAA  
TGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTA  
TCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATG  
GGCTGACCGCTTCCCTCGTGCTTTACGGTATCGCCGCTCCCGATTTCGACGCGCATC  
GCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAATTATTAACGCTTACAATTTCT  
GATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATCAGGTGG  
CACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACAT  
TCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATAG  
CACGTGAGGAGGGCCACCATGGCCAAGTTGACCAGTGCCGTTCCGGTGCTCACC  
GCGCGCGACGTCGCCGGAGCGGTTCGAGTTCTGGACCGACCGGCTCGGGTTCTCC  
CGGGACTTCGTGGAGGACGACTTCGCCGGTGTGGTCCGGGACGACGTGACCCTG  
TTCATCAGCGCGGTCCAGGACCAGGTGGTGCCGGACAACACCCTGGCCTGGGTG  
TGGGTGCGCGGCCTGGACGAGCTGTACGCCGAGTGGTCGGAGGTCGTGTCCACG  
AACTTCCGGGACGCCTCCGGGCCGGCCATGACCGAGATCGGCGAGCAGCCGTG  
GGGGCGGGAGTTCGCCCTGCGCGACCCGGCCGCAACTGCGTGCATTCGTGGC  
CGAGGAGCAGGACTGACACGTGCTAAACTTCATTTTTAATTTAAAAGGATCTA  
GGTGAAGATCCTTTTTGATAATCTCATGACCAAATCCCTTAACGTGAGTTTTCG  
TTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCT  
TTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCACCGCTACCAGCG  
GTGGTTTGTGGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCT  
TCAGCAGAGCGCAGATAACCAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCC  
ACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTT  
ACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAG  
ACGATAGTTACCGGATAAGGCGCAGCGGTCCGGCTGAACGGGGGGTTCGTGCA  
CACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGT



GAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGGCGGACAGGTATCC  
 GGTAAGCGGCAGGGTTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGA  
 AACGCCTGGTATCTTTATAGTCCTGTCTGGGTTTCGCCACCTCTGACTTGAGCGTC  
 GATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAACGCCAGCAAC  
 GCGGCCTTTTTACGGTTCCTGGCCTTTTTGCTGGCCTTTTTGCTCACATGTTCTTTCC  
 TCGTATATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGAT  
 ACCGCTCGCCGACCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGC  
 GGAAG

- Zero Blunt Vector map [3512 bp]



Comments for pCR®-Blunt  
3512 nucleotides

## A.3.2 Vectors used for construct expression vector for in-vitro recombinant Cas9 production using CLC software

### A.3.2.1 pCas9\_TPC [14017 bp]

- pCas9\_TPC sequence [14017 bp]

AGTACTTTGATCCAACCCCTCCGCTGCTATAGTGCAGTCGGCTTCTGACGTTTCAG  
TGCAGCCGTCTTCTGAAAACGACATGTCGCACAAGTCCTAAGTTACGCGACAGG  
CTGCCGCCCTGCCCTTTTCCTGGCGTTTTCTTGTGCGCGTGTTTTAGTCGCATAAA  
GTAGAATACTTGCGACTAGAACCGGAGACATTACGCCATGAACAAGAGCGCCG  
CCGCTGGCCTGCTGGGCTATGCCCGCGTCAGCACCGACGACCAGGACTTGACCA  
ACCAACGGGCCGAACTGCACGCGGCCGGCTGCACCAAGCTGTTTTCCGAGAAG  
ATCACCGGCACCAGGCGCGACCGCCCGGAGCTGGCCAGGATGCTTGACCACCT  
ACGCCCTGGCGACGTTGTGACAGTGACCAGGCTAGACCGCCTGGCCCCGACGAC  
CCGCGACCTACTGGACATTGCCGAGCGCATCCAGGAGGCCGGCGCGGGCCTGC  
GTAGCCTGGCAGAGCCGTGGGCCGACACCACCGCCGGCCGGCCGCATGGTG  
TTGACCGTGTTTCGCCGGCATTGCCGAGTTCGAGCGTTCCTAATCATCGACCGC  
ACCCGGAGCGGGCGCGAGGCCGCCAAGGCCCGAGGCGTGAAGTTTGGCCCCCG  
CCCTACCCTCACCCCGGCACAGATCGCGCACGCCCGCGAGCTGATCGACCAGGA  
AGGCCGCACCGTGAAAGAGGCGGCTGCACTGCTTGGCGTGCATCGCTCGACCCT  
GTACCGCGCACTTGAGCGCAGCGAGGAAGTGACGCCACCGAGGCCAGGCCGGC  
GCGGTGCCTTCCGTGAGGACGCATTGACCGAGGCCGACGCCCTGGCGGCCGCCG  
AGAATGAACGCCAAGAGGAACAAGCATGAAACCGCACCAAGGACGGCCAGGAC  
GAACCGTTTTTTCATTACCGAAGAGATCGAGGCCGAGATGATCGCGGCCGGGTAC  
GTGTTTCGAGCCGCCCGCGCACGTCTCAACCGTGCGGCTGCATGAAATCCTGGCC  
GGTTTGTCTGATGCCAAGCTGGCGGCCTGGCCGGCCAGCTTGGCCGCTGAAGAA  
ACCGAGCGCCGCCGTCTAAAAGGTGATGTGTATTTGAGTAAACAGCTTGCGT  
CATGCGGTGCTGCGTATATGATGCGATGAGTAAATAAACAATAACGCAAGGG  
GAACGCATGAAGGTTATCGCTGTACTIONAACCAGAAAGGCCGGGTCAGGCAAGAC  
GACCATCGCAACCCATCTAGCCCGCGCCCTGCAACTCGCCGGGGCCGATGTTCT  
GTTAGTCGATTCCGATCCCCAGGGCAGTGCCCGCGATTGGGCGGCCCGTGCGGGA  
AGATCAACCGCTAACCGTTGTGCGCATCGACCGCCCGACGATTGACCGCGACGT  
GAAGGCCATCGGCCGGCGCGACTTCGTAGTGATCGACGGAGCGCCCCAGGCCG

---

CGGACTTGGCTGTGTCCGCGATCAAGGCAGCCGACTTCGTGCTGATTCCGGTGC  
AGCCAAGCCCTTACGACATATGGGCCACCGCCGACCTGGTGGAGCTGGTTAAGC  
AGCGCATTGAGGTCACGGATGGAAGGCTACAAGCGGCCTTTGTCGTGTCGCGGG  
CGATCAAAGGCACGCGCATCGGCGGTGAGGTTGCCGAGGCGCTGGCCGGGTAC  
GAGCTGCCATTCTTGAGTCCCGTATCACGCAGCGCGTGAGCTACCCAGGCACT  
GCCGCCGCCGGCACAACCGTTCTTGAATCAGAACCCGAGGGCGACGCTGCCCGC  
GAGGTCCAGGCGCTGGCCGCTGAAATTAATCAAACTCATTTGAGTTAATGAG  
GTAAAGAGAAAATGAGCAAAGCACAAACACGCTAAGTGCCGGCCGTCCGAGC  
GCACGCAGCAGCAAGGCTGCAACGTTGGCCAGCCTGGCAGACACGCCAGCCAT  
GAAGCGGGTCAACTTTCAGTTGCCGGCGGAGGATCACACCAAGCTGAAGATGT  
ACGCGGTACGCCAAGGCAAGACCATTACCGAGCTGCTATCTGAATACATCGCGC  
AGCTACCAGAGTAAATGAGCAAATGAATAAATGAGTAGATGAATTTTAGCGGC  
TAAAGGAGGCGGCATGGAAAATCAAGAACAACCAGGCACCGACGCCGTGGAAT  
GCCCCATGTGTGGAGGAACGGGCGGTTGGCCAGGCGTAAGCGGCTGGGTTGTCT  
GCCGGCCCTGCAATGGCACTGGAACCCCCAAGCCCGAGGAATCGGCGTGACGG  
TCGCAAACCATCCGGCCCGGTACAAATCGGCGCGGCGCTGGGTGATGACCTGGT  
GGAGAAGTTGAAGGCCGCGCAGGCCGCCAGCGGCAACGCATCGAGGCAGAAG  
CACGCCCCGGTGAATCGTGGCAAGCGGCCGCTGATCGAATCCGCAAAGAATCC  
CGGCAACCGCCGGCAGCCGGTGCGCCGTCGATTAGGAAGCCGCCCAAGGGCGA  
CGAGCAACCAGATTTTTTCGTTCCGATGCTCTATGACGTGGGCACCCGCGATAG  
TCGCAGCATCATGGACGTGGCCGTTTTCCGTCTGTCTGAAGCGTGACCGACGAGC  
TGGCGAGGTGATCCGCTACGAGCTTCCAGACGGGCACGTAGAGGTTTCCGCAGG  
GCCGGCCGGCATGGCCAGTGTGTGGGATTACGACCTGGTACTGATGGCGGTTTC  
CCATCTAACCGAATCCATGAACCGATACCGGGAAGGGAAGGGAGACAAGCCCG  
GCCGCGTGTTCCGTCCACACGTTGCGGACGTACTIONCAAGTTCTGCCGGCGAGCCG  
ATGGCGGAAAGCAGAAAGACGACCTGGTAGAAACCTGCATTCGGTTAAACACC  
ACGCACGTTGCCATGCAGCGTACGAAGAAGGCCAAGAACGGCCGCTGGTGAC  
GGTATCCGAGGGTGAAGCCTTGATTAGCCGCTACAAGATCGTAAAGAGCGAAA  
CCGGGCGGCCGGAGTACATCGAGATCGAGCTAGCTGATTGGATGTACCGCGAG  
ATCACAGAAGGCAAGAACCCGGACGTGCTGACGGTTCACCCCGATTACTTTTTG  
ATCGATCCCGGCATCGGCCGTTTTCTCTACCGCCTGGCACGCCGCGCCGCAGGC  
AAGGCAGAAGCCAGATGGTTGTTCAAGACGATCTACGAACGCAGTGGCAGCGC  
CGGAGAGTTCAAGAAGTTCTGTTTACCGTGCGCAAGCTGATCGGGTCAAATGA  
CCTGCCGGAGTACGATTTGAAGGAGGAGGCGGGGCAGGCTGGCCCGATCCTAG

TCATGCGCTACCGCAACCTGATCGAGGGCGAAGCATCCGCCGGTTCCTAATGTA  
CGGAGCAGATGCTAGGGCAAATTGCCCTAGCAGGGGAAAAAGGTGCGAAAAGGT  
CTCTTTCCTGTGGATAGCACGTACATTGGGAACCCAAAGCCGTACATTGGGAAC  
CGGAACCCGTACATTGGGAACCCAAAGCCGTACATTGGGAACCCGGTCACACAT  
GTAAGTGACTGATATAAAAGAGAAAAAAGGCGATTTTTCCGCCTAAACTCTTT  
AAAACCTATTA AAAACTCTTAAAACCCGCCTGGCCTGTGCATAACTGTCTGGCCA  
GCGCACAGCCGAAGAGCTGCAAAAAGCGCCTACCCTTCGGTCGCTGCGCTCCCT  
ACGCCCCGCCGCTTCGCGTCGGCCTATCGCGGCCGCTGGCCGCTCAAAAATGGC  
TGGCCTACGGCCAGGCAATCTACCAGGGCGCGGACAAGCCGCGCCGTGCGCCAC  
TCGACCGCCGGCGCCACATCAAGGCACCCTGCCTCGCGCGTTTTCGGTGATGAC  
GGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAA  
GCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGG  
GTGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATACTG  
GCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTG  
TGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCG  
CTTCCTCGCTCACTGACTCGCTGCGCTCGGTTCGGCTGCGGGCAGCGGTATC  
AGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAG  
GAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGC  
CGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAA  
TCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAACCAGG  
CGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTAC  
CGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCA  
CGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGC  
ACGAACCCCCCGTTCAGCCCGACCCTGCGCCTTATCCGGTAACTATCGTCTTG  
AGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA  
GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGG  
CCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAG  
CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC  
GCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAA  
GGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAC  
GAAAACCTCACGTTAAGGGATTTTGGTTCATGCATGATATATCTCCCAATTTGTGTA  
GGGCTTATTATGCACGCTTAAAAATAATAAAAGCAGACTTGACCTGATAGTTTG  
GCTGTGAGCAATTATGTGCTTAGTGCATCTAATCGCTTGAGTTAACGCCGGCGA  
AGCGGCGTCGGCTTGAACGAATTTCTAGCTAGACATTATTTGCCGACTACCTTG

GTGATCTCGCCTTTCACGTAGTGGACAAATTCTTCCAACCTGATCTGCGCGCGAG  
GCCAAGCGATCTTCTTCTTGTCCAAGATAAAGCCTGTCTAGCTTCAAGTATGACG  
GGCTGATACTGGGCCGGCAGGGCGCTCCATTGCCAGTCGGCAGCGACATCCTTC  
GGCGCGATTTTGCCGGTACTGCGCTGTACCAAATGCGGGACAACGTAAGCACT  
ACATTTGCTCATCGCCAGCCAGTCGGGCGGGCAGTTCCATAGCGTTAAGGTT  
TCATTTAGCGCCTCAAATAGATCCTGTTCAGGAACCGGATCAAAGAGTTCCTCC  
GCCGCTGGACCTACCAAGGCAACGCTATGTTCTTCTTGTCTTTGTCAGCAAGATA  
GCCAGATCAATGTCGATCGTGGCTGGCTCGAAGATACCTGCAAGAATGTCATTG  
CGCTGCCATTCTCAAATTGCAGTTCGCGCTTAGCTGGATAACGCCACGGAATG  
ATGTCGTCGTGCACAACAATGGTGACTTCTACAGCGCGGAGAATCTCGCTCTCT  
CCAGGGGAAGCCGAAGTTTCCAAAAGGTCGTTGATCAAAGCTCGCCGCGTTGTT  
TCATCAAGCCTTACGGTCACCGTAACCAGCAAATCAATATCACTGTGTGGCTTC  
AGGCCGCCATCCACTGCGGAGCCGTACAAATGTACGGCCAGCAACGTCGGTTTCG  
AGATGGCGCTCGATGACGCCAACTACCTCTGATAGTTGAGTCGATACTTCGGCG  
ATCACCGCTTCCCCCATGATGTTTAACTTTGTTTTAGGGCGACTGCCCTGCTGCG  
TAACATCGTTGCTGCTCCATAACATCAAACATCGACCCACGGCGTAACGCGCTT  
GCTGCTTGGATGCCCAGGCATAGACTGTACCCCAAAAAACATGTCATAACAA  
GAAGCCATGAAAACCGCCACTGCGCCGTTACCACCGCTGCGTTCGGTCAAGGTT  
CTGGACCAGTTGCGTGACGGCAGTTACGCTACTTGCATTACAGCTTACGAACCG  
AACGAGGCTTATGTCCACTGGGTTCGTGCCCGAATTGATCACAGGCAGCAACGC  
TCTGTCATCGTTACAATCAACATGCTACCCTCCGCGAGATCATCCGTGTTTCAA  
CCCGGCAGCTTAGTTGCCGTTCTTCCGAATAGCATCGGTAACATGAGCAAAGTC  
TGCCGCCTTACAACGGCTCTCCCGCTGACGCCGTCCCGGACTGATGGGCTGCCT  
GTATCGAGTGGTGATTTTGTGCCGAGCTGCCGGTCGGGGAGCTGTTGGCTGGCT  
GGTGGCAGGATATATTGTGGTGTAACAAATTGACGCTTAGACAACCTTAATAAC  
ACATTGCGGACGTTTTTAATGTACTGAATTAACGCCGAATTGCTCTAGCCAATA  
CGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGAC  
AGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTA  
GCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTG  
TGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACATGATT  
ACGAATTCAAAAATTACGGATATGAATATAGGCATATCCGTATCCGAATTATCC  
GTTTGACAGCTAGCAACGATTGTACAATTGCTTCTTTAAAAAAGGAAGAAAGAA  
AGAAAGAAAAGAATCAACATCAGCGTTAACAACGGCCCCGTTACGGCCCAA  
CGGTCATATAGAGTAACGGCGTTAAGCGTTGAAAGACTCCTATCGAAATACGTA

ACCGCAAACGTGTCATAGTCAGATCCCCTCTTCCTTCACCGCCTCAAACACAAA  
AATAATCTTCTACAGCCTATATATAACAACCCCCCTTCTATCTCTCCTTTCTCAC  
AATTCATCATCTTTCTTTCTCTACCCCAATTTTAAGAAATCCTCTCTTCTCCTCT  
TCATTTTCAAGGTAAATCTCTCTCTCTCTCTCTCTCTCTGTTATTCCTTGTTTTAAT  
TAGGTATGTATTATTGCTAGTTTGTTAATCTGCTTATCTTATGTATGCCTTATGTG  
AATATCTTTATCTTGTTTCATCTCATCCGTTTAGAAGCTATAAAATTTGTTGATTTGA  
CTGTGTATCTACACGTGGTTATGTTTATATCTAATCAGATATGAATTTCTTCATA  
TTGTTGCGTTTGTGTGTACCAATCCGAAATCGTTGATTTTTTTTCATTTAATCGTGT  
AGCTAATTGTACGTATACATATGGATCTACGTATCAATTGTTTCATCTGTTTGTGT  
TTGTATGTATACAGATCTGAAAACATCACTTCTCTCATCTGATTGTGTTGTTACA  
TACATAGATATAGATCTGTTATATCATTTTTTTTTATTAATTGTGTATATATATATG  
TGCATAGATCTGGATTACATGATTGTGATTATTTACATGATTTTGTATTACGT  
ATGTATATATGTAGATCTGGACTTTTTGGAGTTGTTGACTTGATTGTATTTGTGT  
GTGTATATGTGTGTTCTGATCTTGATATGTTATGTATGTGCAGCGAATTCGGCGC  
GCCATGGATAAGAAGTACTCTATCGGACTCGATATCGGAACTAACTCTGTGGGA  
TGGGCTGTGATCACCGATGAGTACAAGGTGCCATCTAAGAAGTTCAAGGTTCTC  
GGAAACACCGATAGGCACTCTATCAAGAAAAACCTTATCGGTGCTCTCCTCTTC  
GATTCTGGTGAAACTGCTGAGGCTACCAGACTCAAGAGAACCGCTAGAAGAAG  
GTACACCAGAAGAAAGAACAGGATCTGCTACCTCCAAGAGATCTTCTCTAACGA  
GATGGCTAAAGTGGATGATTCATTCTTCCACAGGCTCGAAGAGTCATTCCTCGT  
GGAAGAAGATAAGAAGCACGAGAGGCACCCTATCTTCGGAAACATCGTTGATG  
AGGTGGCATAACCACGAGAAGTACCCTACTATCTACCACCTCAGAAAGAAGCTCG  
TTGATTCTACTGATAAGGCTGATCTCAGGCTCATCTACCTCGCTCTCGCTCACAT  
GATCAAGTTCAGAGGACACTTCCTCATCGAGGGTGATCTCAACCCTGATAACTC  
TGATGTGGATAAGTTGTTTCATCCAGCTCGTGCAGACCTACAACCAGCTTTTCGA  
AGAGAACCCTATCAACGCTTCAGGTGTGGATGCTAAGGCTATCCTCTCTGCTAG  
GCTCTCTAAGTCAAGAAGGCTTGAGAACCTCATTGCTCAGCTCCCTGGTGAGAA  
GAAGAACGGACTTTTCGGAAACTTGATCGCTCTCTCTCTCGGACTCACCCCTAA  
CTTCAAGTCTAACTTCGATCTCGCTGAGGATGCAAAGCTCCAGCTCTCAAAGGA  
TACCTACGATGATGATCTCGATAACCTCCTCGCTCAGATCGGAGATCAGTACGC  
TGATTTGTTCTCGCTGCTAAGAACCTCTCTGATGCTATCCTCCTCAGTGATATC  
CTCAGAGTGAACACCGAGATCACCAAGGCTCCACTCTCAGCTTCTATGATCAAG  
AGATACGATGAGCACCACCAGGATCTCACACTTCTCAAGGCTCTTGTTAGACAG  
CAGCTCCCAGAGAAGTACAAAGAGATTTTCTTCGATCAGTCTAAGAACGGATAC

GCTGGTTACATCGATGGTGGTGCATCTCAAGAAGAGTTCTACAAGTTCATCAAG  
CCTATCCTCGAGAAGATGGATGGAACCGAGGAACCTCCTCGTGAAGCTCAATAG  
AGAGGATCTTCTCAGAAAGCAGAGGACCTTCGATAACGGATCTATCCCTCATCA  
GATCCACCTCGGAGAGTTGCACGCTATCCTTAGAAGGCAAGAGGATTTCTACCC  
ATTCCTCAAGGATAACAGGGAAAAGATTGAGAAGATTCTCACCTTCAGAATCCC  
TACTACGTGGGACCTCTCGCTAGAGGAAACTCAAGATTCGCTTGGATGACCAG  
AAAGTCTGAGGAAACCATCACCCCTTGGAACCTCGAAGAGGTGGTGGATAAGG  
GTGCTAGTGCTCAGTCTTTCATCGAGAGGATGACCAACTTCGATAAGAACCTTC  
CAAACGAGAAGGTGCTCCCTAAGCACTCTTTGCTCTACGAGTACTTCACCGTGT  
ACAACGAGTTGACCAAGGTAAAGTACGTGACCGAGGGAATGAGGAAGCCTGCT  
TTTTTGTCAGGTGAGCAAAGAAGGCTATCGTTGATCTCTTGTTCAAGACCAAC  
AGAAAGGTGACCGTGAAGCAGCTCAAAGAGGATTACTTCAAGAAAATCGAGTG  
CTTCGATTCAGTTGAGATTTCTGGTGTGAGGATAGGTTCAACGCATCTCTCGGA  
ACCTACCACGATCTCCTCAAGATCATTAAAGGATAAGGATTTCTTGGATAACGAG  
GAAAACGAGGATATCTTGGAGGATATCGTTCTTACCCTCACCTCTTTGAAGAT  
AGAGAGATGATTGAAGAAAGGCTCAAGACCTACGCTCATCTCTTCGATGATAAG  
GTGATGAAGCAGTTGAAGAGAAGAAGATACACTGGTTGGGGAAGGCTCTCAAG  
AAAGCTCATTAAACGGAATCAGGGATAAGCAGTCTGGAAAGACAATCCTTGATTT  
CCTCAAGTCTGATGGATTCGCTAACAGAACTTCATGCAGCTCATCCACGATGA  
TTCTCTCACCTTTAAAGAGGATATCCAGAAGGCTCAGGTTTCAGGACAGGGTGA  
TAGTCTCCATGAGCATATCGCTAACCTCGCTGGATCTCCTGCAATCAAGAAGGG  
AATCCTCCAGACTGTGAAGGTTGTGGATGAGTTGGTGAAGGTGATGGGAAGGC  
ATAAGCCTGAGAACATCGTGATCGAAATGGCTAGAGAGAACCAGACCACTCAG  
AAGGGACAGAAGAACTCTAGGGAAAGGATGAAGAGGATCGAGGAAGGTATCA  
AAGAGCTTGGATCTCAGATCCTCAAAGAGCACCTGTTGAGAACACTCAGCTCC  
AGAATGAGAAGCTCTACCTCTACTACCTCCAGAACGGAAGGGATATGTATGTGG  
ATCAAGAGTTGGATATCAACAGGCTCTCTGATTACGATGTTGATCATATCGTGC  
CACAGTCATTCTTGAAGGATGATTCTATCGATAACAAGGTGCTCACCAGGTCTG  
ATAAGAACAGGGGTAAGAGTGATAACGTGCCAAGTGAAGAGGTTGTGAAGAAA  
ATGAAGAACTATTGGAGGCAGCTCCTCAACGCTAAGCTCATCACTCAGAGAAA  
GTTTCGATAACTTGACTAAGGCTGAGAGGGGAGGACTCTCTGAATTGGATAAGGC  
AGGATTCATCAAGAGGCAGCTTGTGGAAACCAGGCAGATCACTAAGCACGTTG  
CACAGATCCTCGATTCTAGGATGAACACCAAGTACGATGAGAACGATAAGTTG  
ATCAGGGAAAGTGAAGGTTATCACCTCAAGTCAAAGCTCGTGTCTGATTTCAGA

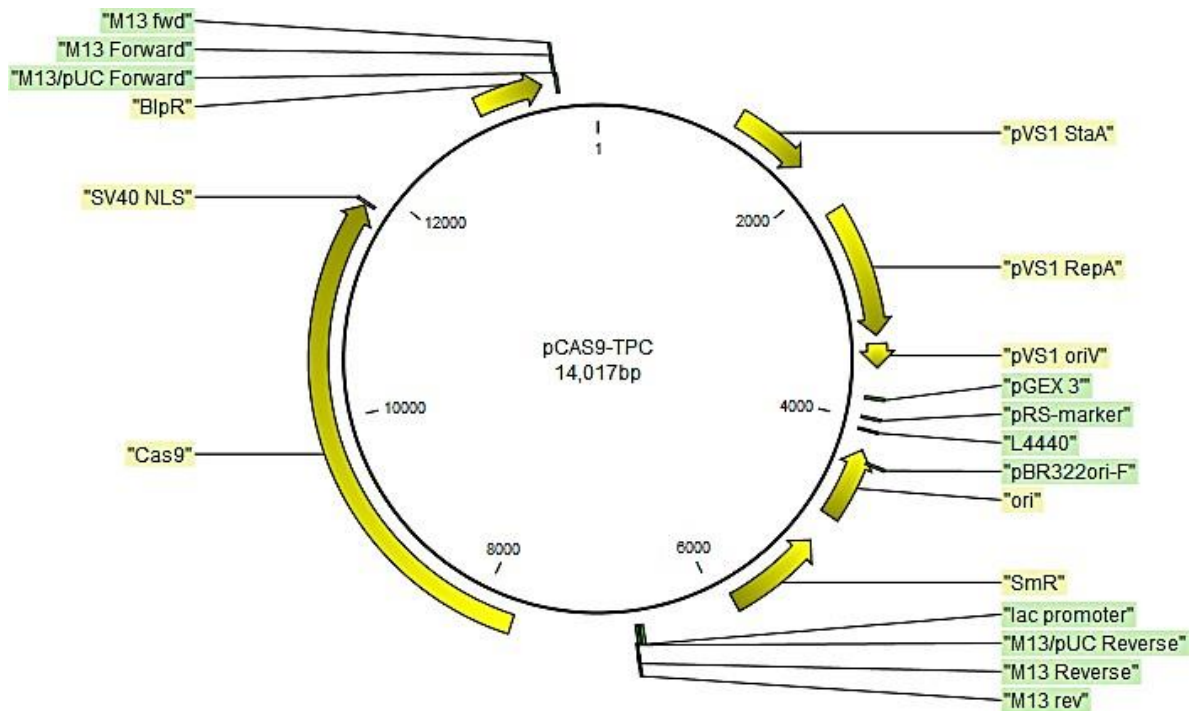
AAGGATTTCCAATTCTACAAGGTGAGGGAAATCAACAACCTACCACCACGCTCAC  
GATGCTTACCTTAACGCTGTTGTTGGAACCGCTCTCATCAAGAAGTATCCTAAG  
CTCGAGTCAGAGTTCGTGTACGGTGATTACAAGGTGTACGATGTGAGGAAGATG  
ATCGCTAAGTCTGAGCAAGAGATCGGAAAGGCTACCGCTAAGTATTTCTTCTAC  
TCTAACATCATGAATTTCTTCAAGACCGAGATTACCCTCGCTAACGGTGAGATC  
AGAAAGAGGCCACTCATCGAGACAAACGGTGAAACAGGTGAGATCGTGTGGGA  
TAAGGGAAGGGATTTGCTACCGTTAGAAAGGTGCTCTCTATGCCACAGGTGAA  
CATCGTTAAGAAAACCGAGGTGCAGACCGGTGGATTCTCTAAAGAGTCTATCCT  
CCCTAAGAGGAACTCTGATAAGCTCATTGCTAGGAAGAAGGATTGGGACCCTA  
AGAAATACGGTGGTTTTGATTCTCCTACCGTGGCTTACTCTGTTCTCGTTGTGGC  
TAAGGTTGAGAAGGGAAAGAGTAAGAAGCTCAAGTCTGTTAAGGAACTTCTCG  
GAATCACTATCATGGAAAGGTCATCTTTCGAGAAGAACCCAATCGATTTCCCTCG  
AGGCTAAGGGATACAAAGAGGTTAAGAAGGATCTCATCATCAAGCTCCCAAAG  
TACTCACTCTTCGAACTCGAGAACGGTAGAAAGAGGATGCTCGCTTCTGCTGGT  
GAGCTTCAAAGGGAAACGAGCTTGCTCTCCCATCTAAGTACGTAACTTTCTTT  
ACCTCGCTTCTCACTACGAGAAGTTGAAGGGATCTCCAGAAGATAACGAGCAG  
AAGCAACTTTTCGTTGAGCAGCACAAAGCACTACTTGGATGAGATCATCGAGCAG  
ATCTCTGAGTTCTCTAAAAGGGTGATCCTCGCTGATGCAAACCTCGATAAGGTG  
TTGTCTGCTTACAACAAGCACAGAGATAAGCCTATCAGGGAACAGGCAGAGAA  
CATCATCCATCTCTTACCCTTACCAACCTCGGTGCTCCTGCTGCTTTCAAGTAC  
TTCGATAACAACCATCGATAGGAAGAGATACACCTCTACCAAAGAAGTGCTCGAT  
GCTACCCTCATCCATCAGTCTATCACTGGACTCTACGAGACTAGGATCGATCTCT  
CACAGCTCGGTGGTGATTCAAGGGCTGATCCTAAGAAGAAGAGGAAGGTTTGA  
GGCGCGCCGAGCTCCAGGCCTCCAGCTTTCGTCCGTATCATCGGTTTCGACAA  
CGTTCGTCAAGTTCAATGCATCAGTTTCATTGCCACACACCAGAATCCTACTAA  
GTTTGAGTATTATGGCATTGGAAAAGCTGTTTTCTTCTATCATTTGTTCTGCTTGT  
AATTTACTGTGTTCTTTCAGTTTTTGTTCGGACATCAAAAATGCAAATGGATGG  
ATAAGAGTTAATAAATGATATGGTCCTTTTGTTCATTCTCAAATTATTATTATCT  
GTTGTTTTTACTTTAATGGGTTGAATTTAAGTAAGAAAGGAACTAACAGTGTGA  
TATTAAGGTGCAATGTTAGACATATAAAACAGTCTTTCACCTCTCTTTGGTTATG  
TCTTGAATTGGTTTGTTCCTTCACTTATCTGTGTAATCAAGTTTACTATGAGTCTA  
TGATCAAGTAATTATGCAATCAAGTTAAGTACAGTATAGGCTTGAGCTCCCTAG  
GCCCCGGCCTGAGGACGCGTCCATGGTTAATTAAGACGTCCGGACCGACTAGTG  
GATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTCTTCGTCAACATGGTGGAG



---

CACGACACGCTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAA  
AGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTC  
CATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGC  
TCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCT  
GCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAA  
AGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCAC  
TGACGTAAGGGATGACGCACAATCAATCCCCTATCCTTCGCAAGACCCTTTTA  
AGGGGAAGTTCATTTCAATTTGGAGAGGACACGCTGAAATCACCAGTCTCTCTG  
TACAAATCNATCTCTCTATAATATTGTGTAAGTAGTTCCCAGATAAGGGAATT  
AGGGTTCTTATAGGGTTTCGCTCAGCTGTTGAGCATATAAGAAACCTTAGTCG  
ATAGATCTGTTGGGGATCTACCATGAGCCCAGAACGACGCCCGGCCGACATCCG  
CCGTGCCACCGAGGCGGACATGCCGGCGGTCTGCACCATCGTCAACCACTACAT  
CGAGACAAGCACGGTCAACTTCCGTACCGAGCCGCAGGAACCGCAGGAGTGA  
CGGACGACCTCGTCCGTCTGCGGGAGCGCTATCCCTGGCTCGTCGCCGAGGTGG  
ACGGCGAGGTCGCCGGCATCGCCTACGCGGGCCCCTGGAAGGCACGCAACGCC  
TACGACTGGACGGCCGAGTCGACCGTGTACGTCTCCCCCGCCACCAGCGGACG  
GGACTGGGCTCCACGCTCTACACCCACCTGCTGAAGTCCCTGGAGGCACAGGGC  
TTCAAGAGCGTGGTCGCTGTCATCGGGCTGCCCAACGACCCGAGCGTGCGCATG  
CACGAGGCGCTCGGATATGCCCCCGCGGCATGCTGCGGGCGGCCGGCTTCAAG  
CACGGGAACCTGGCATGACGTGGGTTTCTGGCAGCTGGACTTCAGCCTGCCGGTA  
CCGCCCCGTCCGGTCCTGCCCGTCACCGAGATCTGATGACCCAACCTTAGTATGT  
ATTTGTATTTGTAAAATACTTCTATCAATAAAAATTTCTAATTCCTAAAACCAAAA  
TCCAGGGGTACCGAACAAGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGACT  
GGGAAAACCTGGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTTCG  
CCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGC  
GCAGCCTGAATGGCGAATGAGCTTGAGCTTGATCAGATTGTCGTTTCCCGCCT  
TCAGTTTAAACTATCAGTGTTTGACAGGATATATTGGCGGGTAAACCTAAGAGA  
AAAGAGCGTTTATTAGAATAACGGATATTTAAAAGGGCGTGAAAAGGTTTATCC  
GTTTCGTCCATTTGTATGTGCATGCCAACCACAGGGTTCCCTCGGGATCAA

- pCas9\_TPC map [14017 bp]



### A.3.2.2 pET302\_NT [5712 bp]

- pET302\_NT sequence [5712 bp]

```

GATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGAT
AACAAATCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACAT
ATGCATCATCATCATCACGTGAATTCGCTCGAGATCGATGATATTCGAGCC
TAGGTATAATCGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGC
TGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGT
CTTGAGGGGTTTTTTGCTGAAAGGAGGA ACTATATCCGGATATCCC GCAAGAGG
CCCGGCAGTACCGGCATAACCAAGCCTATGCCTACAGCATCCAGGGTGACGGTG
CCGAGGATGACGATGAGCGCATTGTTAGATTT CATAACGGTGCCTGACTGCGT
TAGCAATTTAACTGTGATAAACTACCGCATTAAAGCTAGCTTATCGATGATAAG
CTGTCAAACATGAGAATTAATTCTTGAAGACGAAAGGGCCTCGTGATACGCCTA
TTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTT
TTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAA
TATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAA
AAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATCCCTTTTTTGCG
GCATTTTGCCTTCCTGTTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGAT
GCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGC

```

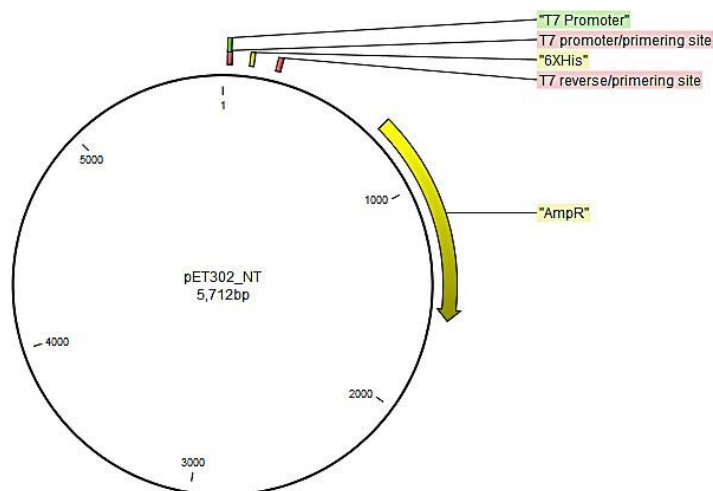
---

GGTAAGATCCTTGAGAGTTTTTCGCCCGAAGAACGTTTTCCAATGATGAGCACT  
TTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGCAAGAG  
CAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTACCA  
GTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGC  
TGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGG  
AGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACCTCG  
CCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTG  
ACACCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCG  
AACTACTTACTCTAGCTTCCCGCAACAATTAATAGACTGGATGGAGGCGGATA  
AAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGA  
TAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCC  
AGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAAC  
TATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGC  
ATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAACT  
TCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTTGATAATCTCATGACC  
AAAATCCCTTAACGTGAGTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAG  
ATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAA  
CAAAAAACCACCGCTACCAGCGGTGGTTTGTGGCCGGATCAAGAGCTACCAA  
CTTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCC  
TTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGA ACTCTGTAGCACCGCCTA  
CATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTC  
GTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTC  
GGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACA  
CCGA ACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAA  
GGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGC  
GCACGAGGGAGCTTCCAGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGT  
TTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGA  
GCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCT  
GGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGT  
ATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGACCCGAACGACCGAGCGC  
AGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTT  
ACGCATCTGTGCGGTATTTACACCGCAATGGTGC ACTCTCAGTACAATCTGCTC  
TGATGCCGCATAGTTAAGCCAGTATACTCCGCTATCGCTACGTGACTGGGTC  
ATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGT

CTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATG  
TGTCAGAGGTTTTTACCCTCATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGC  
TCATCAGCGTGGTTCGTGAAGCGATTCACAGATGTCTGCCTGTTTCATCCGCGTCC  
AGCTCGTTGAGTTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCC  
ATGTTAAGGGCGGTTTTTTTCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGAT  
TTCTGTTTCATGGGGTAATGATACCGATGAAACGAGAGAGGATGCTCACGATAC  
GGGTTACTGATGATGAACATGCCCGTTACTGGAACGTTGTGAGGGTAAACAAC  
TGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAG  
CGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCG  
ATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTT  
TACGAAACACGGAAACCGAAGACCATTTCATGTTGTTGCTCAGGTCGCAGACGTT  
TTGCAGCAGCAGTCGCTTCACGTTTCGCTCGCGTATCGGTGATTCATTCTGCTAAC  
CAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCTCAACGACAGGAGCACGATC  
ATGCGCACCCGTGGCCAGGACCCAACGCTGCCCGAGATGCGCCGCGTGCGGCT  
GCTGGAGATGGCGGACGCGATGGATATGTTCTGCCAAGGGTTGGTTTGCGCATT  
CACAGTTCTCCGCAAGAATTGATTGGCTCCAATTCTTGGAGTGGTGAATCCGTT  
AGCGAGGTGCCCGCGGCTTCCATTCAGGTTCGAGGTGGCCCGGCTCCATGCACCG  
CGACGCAACGCGGGGAGGCAGACAAGGTATAGGGCGGGCGCCTACAATCCATGC  
CAACCCGTTCCATGTGCTCGCCGAGGCGGCATAAATCGCCGTGACGATCAGCGG  
TCCAATGATCGAAGTTAGGCTGGTAAGAGCCGCGAGCGATCCTTGAAGCTGTCC  
CTGATGGTCGTCATCTACCTGCCTGGACAGCATGGCCTGCAACGCGGGCATCCC  
GATGCCGCCGGAAGCGAGAAGAATCATAATGGGGAAGGCCATCCAGCCTCGCG  
TCGCGAACGCCAGCAAGACGTAGCCCAGCGCGTCGGCCGCCATGCCGGCGATA  
ATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGACCAGTGACGAAGGCTTGA  
GCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGC  
GCTCCAGCGAAAGCGGTCTCTCGCCGAAAATGACCCAGAGCGCTGCCGGCACCT  
GTCCTACGAGTTGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTC  
ATGCCCCGCGCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCAT  
CGGTTCGAGATCCCGGTGCCTAATGAGTGAGCTAACTTACATTAATTGCGTTGCG  
CTCACTGCCCGCTTTCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAAT  
CGGCAACGCGCGGGGAGAGGCGGTTTTCGTATTGGGCGCCAGGGTGGTTTTTC  
TTTTACCAGTGAGACGGGCAACAGCTGATTGCCCTTACCAGCCTGGCCCTGAG  
AGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTT  
TGATGGTGGTTAACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATC

CCACTACCGAGATATCCGCACCAACGCGCAGCCCGGACTCGGTAATGGCGCGC  
 ATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATG  
 CCCTCATTACAGCATTTGCATGGTTTGTGAAAACCGGACATGGCACTCCAGTCG  
 CCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTGAGATATTTATGCCAGC  
 CAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCG  
 ATTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCAGTCGCGTACCGTCT  
 TCATGGGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAAAT  
 AACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCC  
 AGCGGATAGTTAATGATCAGCCACTGACGCGTTGCGCGAGAAGATTGTGCACC  
 GCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGACACCACCACGCTG  
 GCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCG  
 TGCAGGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGC  
 CAGTTGTTGTGCCACGCGGTTGGGAATGTAATTCAGCTCCGCCATCGCCGCTTCC  
 ACTTTTTCCCGCGTTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGGGAA  
 ACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGT  
 TTCACATTCACCACCCTGAATTGACTCTCTTCCGGGCGCTATCATGCCATACCGC  
 GAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTATGCG  
 ACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCG  
 CCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACG  
 GGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCG  
 AGCCCGATCTTCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACC  
 TGTGGCGCCGGTGATGCCGGCCACGATGCGTCCGGCGTAGAGGATCGA

- **pET302\_NT map [5712 bp]**



### A.3.2.3 pET302\_His\_Cas9-NLS [9849 bp]

- pET302\_His\_Cas9-NLS sequence [9849 bp]

GATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGAT  
AACAAATCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACAT  
ATGCATCATCATCATCACGATAAGAAGTACTCTATCGGACTCGATATCGGA  
ACTAACTCTGTGGGATGGGCTGTGATCACCGATGAGTACAAGGTGCCATCTAAG  
AAGTTCAAGGTTCTCGGAAACACCGATAGGCACTCTATCAAGAAAACCTTATC  
GGTGCTCTCCTCTTCGATTCTGGTGAAACTGCTGAGGCTACCAGACTCAAGAGA  
ACCGCTAGAAGAAGGTACACCAGAAGAAAGAACAGGATCTGCTACCTCCAAGA  
GATCTTCTCTAACGAGATGGCTAAAGTGGATGATTCATTCTTCCACAGGCTCGA  
AGAGTCATTCCCTCGTGGAAGAAGATAAGAAGCACGAGAGGCACCCTATCTTCG  
GAAACATCGTTGATGAGGTGGCATAACCACGAGAAGTACCCTACTATCTACCACC  
TCAGAAAGAAGCTCGTTGATTCTACTGATAAGGCTGATCTCAGGCTCATCTACC  
TCGCTCTCGCTCACATGATCAAGTTCAGAGGACACTTCCTCATCGAGGGTGATC  
TCAACCCTGATAACTCTGATGTGGATAAGTTGTTTCATCCAGCTCGTGCAGACCT  
ACAACCAGCTTTTCGAAGAGAACCCTATCAACGCTTCAGGTGTGGATGCTAAGG  
CTATCCTCTCTGCTAGGCTCTCTAAGTCAAGAAGGCTTGAGAACCTCATTGCTCA  
GCTCCCTGGTGAGAAGAAGAACGGACTTTTCGGAAACTTGATCGCTCTCTCTCT  
CGGACTCACCCCTAACTTCAAGTCTAACTTCGATCTCGCTGAGGATGCAAAGCT  
CCAGCTCTCAAAGGATACCTACGATGATGATCTCGATAACCTCCTCGCTCAGAT  
CGGAGATCAGTACGCTGATTTGTTCCCTCGCTGCTAAGAACCTCTCTGATGCTATC  
CTCCTCAGTGATATCCTCAGAGTGAACACCGAGATCACCAAGGCTCCACTCTCA  
GCTTCTATGATCAAGAGATACGATGAGCACCACCAGGATCTCACACTTCTCAAG  
GCTCTTGTTAGACAGCAGCTCCCAGAGAAGTACAAAGAGATTTTCTTCGATCAG  
TCTAAGAACGGATACGCTGGTTACATCGATGGTGGTGCATCTCAAGAAGAGTTC  
TACAAGTTCATCAAGCCTATCCTCGAGAAGATGGATGGAACCGAGGAACTCCTC  
GTGAAGCTCAATAGAGAGGATCTTCTCAGAAAGCAGAGGACCTTCGATAACGG  
ATCTATCCCTCATCAGATCCACCTCGGAGAGTTGCACGCTATCCTTAGAAGGCA  
AGAGGATTTCTACCCATTCTCAAGGATAACAGGGAAAAGATTGAGAAGATTCT  
CACCTTCAGAATCCCTTACTACGTGGGACCTCTCGCTAGAGGAAACTCAAGATT  
CGCTTGATGACCAGAAAGTCTGAGGAAACCATCACCCCTTGGAACCTCGAAGA  
GGTGGTGGATAAGGGTGCTAGTGCTCAGTCTTTCATCGAGAGGATGACCAACTT

CGATAAGAACCTTCCAAACGAGAAGGTGCTCCCTAAGCACTCTTTGCTCTACGA  
GTACTTCACCGTGTACAACGAGTTGACCAAGGTTAAGTACGTGACCGAGGGAAT  
GAGGAAGCCTGCTTTTTTGTGTCAGGTGAGCAAAGAAGGCTATCGTTGATCTCTT  
GTTCAAGACCAACAGAAAGGTGACCGTGAAGCAGCTCAAAGAGGATTACTTCA  
AGAAAATCGAGTGCTTCGATTGAGTTGAGATTTCTGGTGTGAGGATAGGTTCA  
ACGCATCTCTCGGAACCTACCACGATCTCCTCAAGATCATTAAAGGATAAGGATT  
TCTTGGATAACGAGGAAAACGAGGATATCTTGGAGGATATCGTTCTTACCCTCA  
CCCTCTTTGAAGATAGAGAGATGATTGAAGAAAGGCTCAAGACCTACGCTCATC  
TCTTCGATGATAAGGTGATGAAGCAGTTGAAGAGAAGAAGATACTACTGGTGG  
GGAAGGCTCTCAAGAAAGCTCATTAAACGGAATCAGGGATAAGCAGTCTGGAAA  
GACAATCCTTGATTTCTCAAGTCTGATGGATTGCTAACAGAACTTCATGCA  
GCTCATCCACGATGATTCTCTCACCTTTAAAGAGGATATCCAGAAGGCTCAGGT  
TTCAGGACAGGGTGATAGTCTCCATGAGCATATCGCTAACCTCGCTGGATCTCC  
TGCAATCAAGAAGGGAATCCTCCAGACTGTGAAGGTTGTGGATGAGTTGGTGA  
AGGTGATGGGAAGGCATAAGCCTGAGAACATCGTGATCGAAATGGCTAGAGAG  
AACCAGACCACTCAGAAGGGACAGAAGAAGTCTAGGGAAAGGATGAAGAGGA  
TCGAGGAAGGTATCAAAGAGCTTGGATCTCAGATCCTCAAAGAGCACCCCTGTTG  
AGAACACTCAGCTCCAGAATGAGAAGCTCTACCTCTACTACCTCCAGAACGGAA  
GGGATATGTATGTGGATCAAGAGTTGGATATCAACAGGCTCTCTGATTACGATG  
TTGATCATATCGTGCCACAGTCATTCTTGAAGGATGATTCTATCGATAACAAGG  
TGCTCACCAGGTCTGATAAGAACAGGGGTAAGAGTGATAACGTGCCAAGTGAA  
GAGGTTGTGAAGAAAATGAAGAACTATTGGAGGCAGCTCCTCAACGCTAAGCT  
CATCACTCAGAGAAAGTTCGATAACTTGACTAAGGCTGAGAGGGGAGGACTCT  
CTGAATTGGATAAGGCAGGATTCATCAAGAGGCAGCTTGTGGAAACCAGGCAG  
ATCACTAAGCACGTTGCACAGATCCTCGATTCTAGGATGAACACCAAGTACGAT  
GAGAACGATAAGTTGATCAGGGAAGTGAAGGTTATCACCTCAAGTCAAAGCT  
CGTGTCTGATTTAGAAAAGGATTTCCAATTCTACAAGGTGAGGGAAATCAACAA  
CTACCACCACGCTCACGATGCTTACCTTAACGCTGTTGTTGGAACCGCTCTCATC  
AAGAAGTATCCTAAGCTCGAGTCAGAGTTCGTGTACGGTGATTACAAGGTGTAC  
GATGTGAGGAAGATGATCGCTAAGTCTGAGCAAGAGATCGGAAAGGCTACCGC  
TAAGTATTTCTTCTACTCTAACATCATGAATTTCTTCAAGACCGAGATTACCCTC  
GCTAACGGTGAGATCAGAAAGAGGCCACTCATCGAGACAAACGGTGAAACAGG  
TGAGATCGTGTGGGATAAGGGAAGGATTTGCTACCGTTAGAAAGGTGCTCTC  
TATGCCACAGGTGAACATCGTTAAGAAAACCGAGGTGCAGACCGGTGGATTCTC

TAAAGAGTCTATCCTCCCTAAGAGGAACTCTGATAAGCTCATTGCTAGGAAGAA  
GGATTGGGACCCTAAGAAATACGGTGGTTTCGATTCTCCTACCGTGGCTTACTCT  
GTTCTCGTTGTGGCTAAGGTTGAGAAGGGAAAGAGTAAGAAGCTCAAGTCTGTT  
AAGGAACTTCTCGGAATCACTATCATGGAAAGGTCATCTTTCGAGAAGAACCCA  
ATCGATTCCTCGAGGCTAAGGGATACAAAGAGGTTAAGAAGGATCTCATCATC  
AAGCTCCCAAAGTACTCACTCTTCGAACTCGAGAACGGTAGAAAGAGGATGCTC  
GCTTCTGCTGGTGAGCTTCAAAGGGAAACGAGCTTGCTCTCCCATCTAAGTAC  
GTTAACTTTCTTTACCTCGCTTCTCACTACGAGAAGTTGAAGGGATCTCCAGAAG  
ATAACGAGCAGAAGCAACTTTTCGTTGAGCAGCACAAGCACTACTTGGATGAG  
ATCATCGAGCAGATCTCTGAGTTCTCTAAAAGGGTGATCCTCGCTGATGCAAAC  
CTCGATAAGGTGTTGTCTGCTTACAACAAGCACAGAGATAAGCCTATCAGGGAA  
CAGGCAGAGAACATCATCCATCTCTTACCCTTACCAACCTCGGTGCTCCTGCTG  
CTTTCAAGTACTTCGATACAACCATCGATAGGAAGAGATACACCTTACCAAAG  
AAGTGCTCGATGCTACCCTCATCCATCAGTCTATCACTGGACTCTACGAGACTA  
GGATCGATCTCTCACAGCTCGGTGGTGATTCAAGGGCTGATCCTAAGAAGAAGA  
GGAAGGTTTGAGTGAATTCGCTCGAGATCGATGATATTCGAGCCTAGGTATAAT  
CGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACC  
GCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAACGGGTCTTGAGGGGT  
TTTTTGCTGAAAGGAGGAACTATATCCGGATATCCCGCAAGAGGCCCGGCAGTA  
CCGGCATAACCAAGCCTATGCCTACAGCATCCAGGGTGACGGTGCCGAGGATG  
ACGATGAGCGCATTGTTAGATTCATACACGGTGCCTGACTGCGTTAGCAATTT  
AACTGTGATAAACTACCGCATTAAAGCTAGCTTATCGATGATAAGCTGTCAAAC  
ATGAGAATTAATTCTTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTTATAG  
GTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGA  
AATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATC  
CGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAG  
AGTATGAGTATTCAACATTTCCGTGTCGCCCTTATCCCTTTTTTGCGGCATTTTG  
CCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGA  
TCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGAT  
CCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTT  
CTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGT  
CGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAA  
AAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACC  
ATGAGTGATAAACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAA



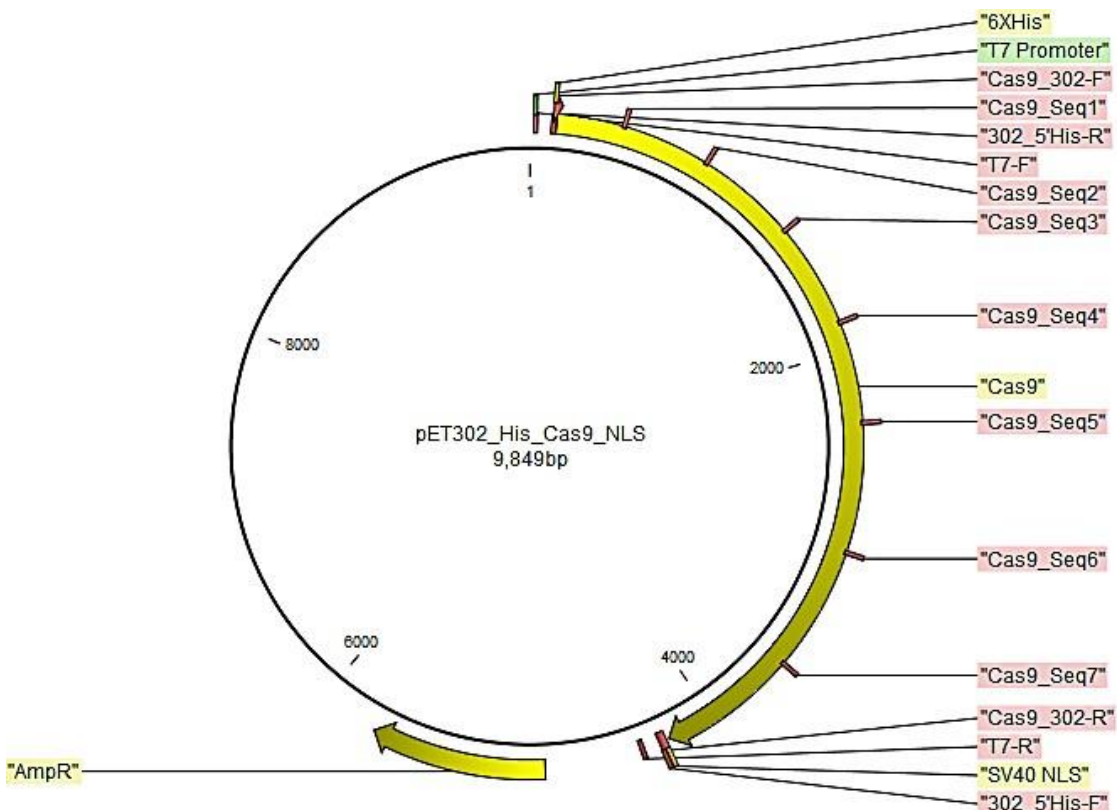
---

GGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACCTCGCCTTGATCG  
TTGGGAACCGGAGCTGAATGAAGCCATAACCAAACGACGAGCGTGACACCACGA  
TGCTGCAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTACTTA  
CTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAG  
GACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGATAAATCTGG  
AGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAA  
GCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGA  
ACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACT  
GTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACCTTCATTTTTAA  
TTTTAAAAGGATCTAGGTGAAGATCCTTTTTTGATAATCTCATGACCAAAATCCCTT  
AACGTGAGTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGAT  
CTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACC  
ACCGCTACCAGCGGTGGTTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCG  
AAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAG  
CCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT  
CTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCG  
GGTTGGACTCAAGACGATAGTTACCGGATAAAGGCGCAGCGGTTCGGGCTGAACG  
GGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAG  
ATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGG  
CGGACAGGTATCCGGTAAGCGGCAGGGTTCGGAACAGGAGAGCGCACGAGGGA  
GCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTTCGGGTTTCGCCACCTC  
TGACTIONGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAA  
AACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTTGCTGGCCTTTTTGCTC  
ACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTT  
GAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGT  
GAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTG  
CGGTATTTACACCGCAATGGTGCACCTCTCAGTACAATCTGCTCTGATGCCGCAT  
AGTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCC  
CCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGC  
ATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTGTCAGAGGTT  
TTCACCGTCATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTG  
GTCGTGAAGCGATTACAGATGTCTGCCTGTTTCATCCGCGTCCAGCTCGTTGAGT  
TTCTCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCG  
GTTTTTTCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTCTGTTCATGG

GGGTAATGATACCGATGAAACGAGAGAGGATGCTCACGATACGGGTTACTGAT  
GATGAACATGCCC GGTTACTGGAACGTTGTGAGGGTAAACA ACTGGCGGTATG  
GATGCGGCGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTA  
ATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCGATGCAGATCC  
GGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACAC  
GGAAACCGAAGACCATTTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGC  
AGTCGCTTCACGTTTCGCTCGCGTATCGGTGATTCATTCTGCTAACCCAGTAAGGCA  
ACCCCGCCAGCCTAGCCGGGTCTCAACGACAGGAGCACGATCATGCGCACCC  
GTGGCCAGGACCCAACGCTGCCCAGATGCGCCGCGTGC GGCTGCTGGAGATG  
GCGGACGCGATGGATATGTTCTGCCAAGGGTTGGTTTGCGCATTACAGTTCTC  
CGCAAGAATTGATTGGCTCCAATTCTTGGAGTGGTGAATCCGTTAGCGAGGTGC  
CGCCGGCTTCCATTCAGGTCGAGGTGGCCCCGGCTCCATGCACCCGCGACGCAACG  
CGGGGAGGCAGACAAGGTATAGGGCGGCGCCTACAATCCATGCCAACCCGTTC  
CATGTGCTCGCCGAGGCGGCATAAATCGCCGTGACGATCAGCGGTCCAATGATC  
GAAGTTAGGCTGGTAAGAGCCGCGAGCGATCCTTGAAGCTGTCCCTGATGGTGC  
TCATCTACCTGCCTGGACAGCATGGCCTGCAACGCGGGCATCCCGATGCCGCCG  
GAAGCGAGAAGAATCATAATGGGGAAGGCCATCCAGCCTCGCGTCGCGAACGC  
CAGCAAGACGTAGCCCAGCGCGTCGGCCGCCATGCCGGCGATAATGGCCTGCTT  
CTCGCCGAAACGTTTGGTGGCGGGACCAGTGACGAAGGCTTGAGCGAGGGCGT  
GCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGCTCCAGCGA  
AAGCGGTCTTCGCCGAAAATGACCCAGAGCGCTGCCGGCACCTGTCCTACGAGT  
TGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCCGCGC  
CCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTTCGAGATC  
CCGGTGCCTAATGAGTGAGCTAACTTACATTAATTGCGTTGCGCTCACTGCCCCG  
CTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCG  
CGGGGAGAGGCGGTTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTACCAGT  
GAGACGGGCAACAGCTGATTGCCCTTACCCGCTGGCCCTGAGAGAGTTGCAGC  
AAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTT  
AACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAG  
ATATCCGCACCAACGCGCAGCCCGACTCGGTAATGGCGCGCATTGCGCCCAGC  
GCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATTACG  
ATTTGCATGGTTTGTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCG  
CTATCGGCTGAATTTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCA  
GACGCGCCGAGACAGA ACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGA

CCCAATGCGACCAGATGCTCCACGCCAGTCGCGTACCGTCTTCATGGGAGAAA  
 ATAATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAAATAACGCCGGAAC  
 ATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGTT  
 AATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACA  
 GGCTTCGACGCCGCTTCGTTCTACCATCGACACCACCACGCTGGCACCCAGTTG  
 ATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCA  
 GACTGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTG  
 CCACGCGGTTGGGAATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCG  
 CGTTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTGATA  
 AGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACATTCAC  
 CACCCTGAATTGACTCTCTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTG  
 CGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTATGCGACTCCTGCATT  
 AGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGGAA  
 TGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCA  
 CCATACCACGCCGAAACAAGCGCTCATGAGCCCCGAAGTGGCGAGCCCGATCTT  
 CCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGG  
 TGATGCCGGCCACGATGCGTCCGGCGTAGAGGATCGA

- **pET302\_His\_Cas9-NLS map [9849 bp]**



## A.4 Reference sequence and its map

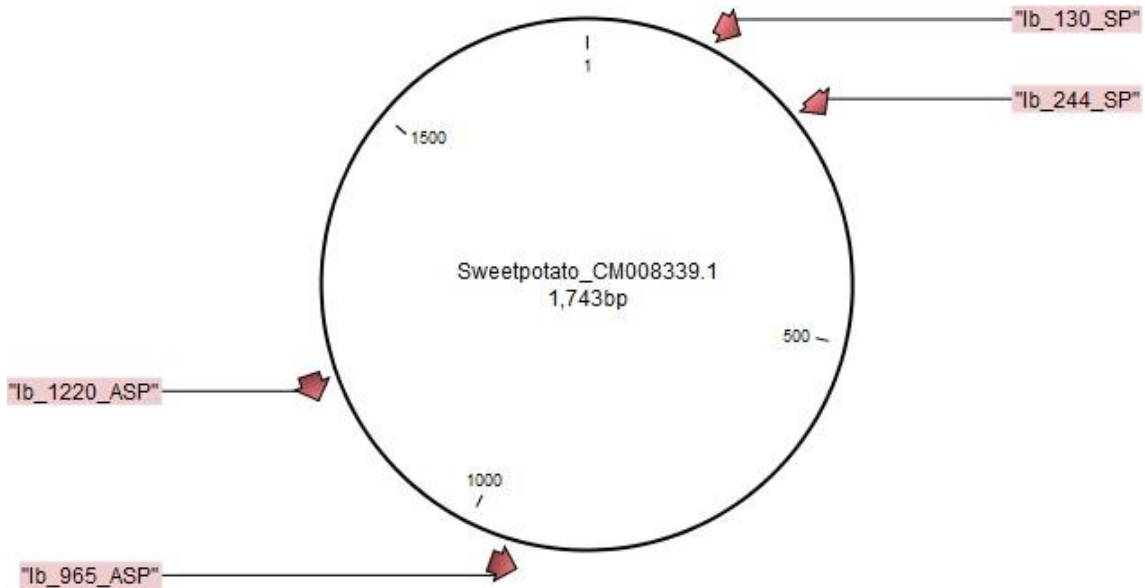
### (Sweetpotato\_CM008339.1)

- Reference sequence (Sweetpotato\_CM008339.1) [1743 bp]

TTCGTCTCCCGATTTCGCTCCCGACGAGCCGCGGAAGGGGTGCGATGTGCTCGTG  
GAAGCGCTAGAGCGGGAGGGCGTGACGGACGTGTTTCGCGTACCCCGGCGGGCGC  
GTCCATGGAGATTCATCAGGCGCTCACGCGCTCCAAGATGATCAGGAATGTCCT  
CCCGCGTCACGAGCAGGGCGGGCGTGTTCGCCGCCGAGGGCTACGCGCGGGCCA  
CGGGCTTCCCGGCGTCTGCATCGCCACCTCCGGCCCCGGCGCCACCAATCTGG  
TGAGCGGACTCGCCGACGCGCTTCTGGATAGCTGTCCCATTGTGGCCATTACTG  
GACAGGTGCCTCGGAGGATGATTGGGACTGACGCTTTCAGGAAACCCCAATCG  
TTGAGGTAACCTCGATCGATTACGAAGCATAATTATTTAGTTTTAGACGTAGAGG  
ATATCCCTCGCATTGTTTCGTGAGGCGTTTTTCCTGGCCAAATCGGGCCGTCCTGG  
CCCGGTTTTGATTGACGTTCCCAAGGATATACAGCAACAGCTGGTTATCCCCAA  
TTGGGATCAGCCTATGAGATTGCCTGGATACTTATCTCGTTTTGCCAAAACCCCCA  
AGTGATATGTTGTTGGAGCAAATAGTGAGGCTAATTTCTGAGTCCAAGAAGCCT  
GTCCTTTATGTGGGGGGCGGGTCTTTACAGTCAAGTGAGGAGTTGAGGCGATTT  
GTGGAGCTTACTGGGATTCCTGTGGCTAGTACTTTAATGGGGCTTGGATCATATC  
CTTCCTCGGATGAACCTTCACTTCAAATGCTGGGAATGCATGGGACTGTGTATG  
CTAATTACTCTGTTGACAAGAGTGATCTGTTGCTTGCATTTGGGGTGAGGTTCGA  
CGACCGTGTGACTGGGAAGCTGGAGGCTTTTGCTAGTCGGGCGAAGATTGTTCA  
CATTGATATTGATTCCGCGGAGATTGGGAAGAATAAGCAGCCCCATGTGTCGAT  
TTGTGCGGATTTGAAGTTGGCCCTGCAGGGGCTGAATTCGATACTGGAGGAGAG  
GGTAGGTAAGCTGAAGTTGGATTTCTCGGCTTGGAGACAGGAGTTGAATGAGCA  
GAAGGAGAAGTTCCCGTTGGGTTATAAGACGTTTGAGGATGCCATTTCTCCCCA  
GTATGCTATTCAGGTTCTTGATGAGTTGACGAACGGAAATGCTATAATTAGTAC  
TGGTGTGGGGCAGCACCAGATGTGGGCTGCCAGTTTTACAAGTATCGGGAGCC  
CCGCCAGTGGTTGACCTCTGGTGGGTTAGGGGCGATGGGTTTTGGATTGCCAGC  
GGCAATTGGAGCAGCGGTTGGAAGACCGGATGCAGTGGTTGTGGATATCGATG  
GCGATGGTAGTTTCATCATGAATGTGCAGGAATTGGCCACGATTTCGAGTGGAGA  
ATCTCCCGGTTAAAATTCTGCTGTTGAATAACCAGCACTTAGGCATGGTGGTTC  
AATGGGAGGATCGCTTCTACAAGGCAAACCGAGCCCATACTTACCTGGGAAAC  
CCTGCTAACGAGAACCAGATATTCCCGAACATGCTGAAGTTTGCTGAGGCTTGT  
GATGTACCCGCTGCTCGTGTGACCAAGAAGGACGATCTCCGGGCTGCTATTCAG

ACGATGCTTGACACCCCCGGGCCGTACTTGTGGATGTGATTGTGCCACATCAG  
 GAACACGTCCTGCCTATGATCCCTAGCGGCCGAGTTTCAACGATGTGATCACC  
 GAGGGGGATGGCAGA

- Reference sequence map (Sweetpotato\_CM008339.1) [1743 bp]



## A.5 Insert sequences and their primer sets

### A.5.1 Insert sequences

- Zero\_Blunt\_pCR\_S1.1 [836 bp]

CTCACGCGCTCCAAGATGATCAGGAATGTCCTCCCGCGTCACGAGCAGGGCGGC  
 GTGTTCCGCCGCGAGGGCTACGCGCGGGCCACGGGCTTCCCCGGCGTCTGCATC  
 GCCACCTCCGGCCCCGGCGCCACCAATCTGGTGAGCGGACTCGCCGACGCGCTT  
 CTGGATAGCTGTCCATTGTGGCCATTACTGGACAGGTGCCTCGGAGGATGATT  
 GGGACAGACGCTTTCAGGAAACCCCAATCGTTGAGGTAACCTCGATCGATTACG  
 AAGCATAATTATTTAGTTTTAGACGTAGAGGACATCCCTCGCATTGTTTCGTGAG  
 GCGTTTTTCCTGGCCAAATCGGGCCGTCTGGCCCGGTTTTGATTGATGTTCCCA  
 AGGATATACAGCAACAGCTGGTTATCCCCAATTGGGATCAGCCTATGAGATTGC  
 CTGGATACTTATCTCGTTTGCCAAAACCCCAAGTGATATGTTGTTGGAGCAA  
 TAGTGAGGCTAATTTCTGAGTCCAAGAAGCCTGTCTTTATGTGGGGGGTGGGT  
 CTTTACAGTCAAGTGAGGAGTTGAGGCGATTTGTGGAGCTAACTGGGATTCTCG

TGGCTAGTACTTTAATGGGGCTTGGATCATATCCTTCCTCGGATGAACTTTCACT  
TCAAATGCTGGGAATGCATGGGACTGTGTATGCTAATTACTCTGTTGACAAGAG  
TGATCTGTTGCTTGCATTTGGGGTGAGGTTTCGACGACCGTGTGACTGGAAAGCT  
GGAGGCTTTTGCTAGTCGGGCGAAGATTGTTACATTGATATTGATTCCGGCGGA  
GATTGGGAAGAATAAGCAGCCCCA

- **Zero\_Blunt\_pCR\_S1.2 [836 bp]**

CTCACGCGCTCCAAGATGATCAGGAATGTCCTCCCGCGTCACGAGCAGGGCGGC  
GTGTTCCGCCGCCGAGGGCTACGCGCGGGCCACGGGCTTCCCCGGCGTCTGCATC  
GCCACTTCCGGCCCCGGCGCCACCAACCTGGTGAGCGGACTCGCTGACGCGCTT  
CTGGATAGCTGTCCCATTGTGGCCATTACTGGACAGGTGCCTCGGAGGATGATT  
GGGACAGACGCTTTCCAGGAAACCCCAATCGTTGAGGTAACCTCGATCGATTACG  
AAGCATAATTATTTAGTTTTAGACGTAGAGGACATCCCTCGCATTGTTTCGTGAG  
GCGTTTTTCTGGCCAAATCGGGCCGTCCTGGCCCCGGTTTTTGATTGATGTTCCCA  
AGGATATACAGCAACAATTGGTTATCCCCAATTGGGATCAGCCTATGAGATTGC  
CTGGATACTTATCTCGTTTGCCAAAACCCCAAGTGATATGTTGTTGGAGCAA  
TAGTGAGGCTAATTTCTGAGTCAAAGAAGCCTGTCCTTTATGTGGGGGGTGGGT  
CTTTACAGTCAAGTGAGGAGTTGAGGCGATTTGTGGAGCTAACTGGGATTCCTG  
TGGCTAGTACTTTAATGGGGCTTGGATCATATCCTTCTTCGGATGAACTTTCACT  
TCAAATGCTGGGAATGCATGGGACTGTGTATGCTAATTACTCTGTTGACAAGAG  
TGATCTGTTGCTTGCATTTGGGGTGAGGTTTCGATGACCGTGTGACTGGGAAGCT  
GGAGGCTTTTGCTAGTCGGGCGAAGATTGTTACATTGATATTGATTCCGGCGGA  
GATTGGGAAGAATAAGCAGCCCCA

- **Zero\_Blunt\_pCR\_S4.1 [980 bp]**

TCCGGCCCCGGCGCCACCAATCTCGTCAGTGGCCTCGCCGATGCGCTTCTAGAT  
AGCTGTCCGATTGTCGCCATTACTGGACAAGTGCCGCGGCGGATGATCGGTACC  
GATGCCTTCCAGGAACTCCTATCGTTGAGGTAACGCGTTCGATCACAAAGCAT  
AATTATCTGGTTCTGAATGTAGAGGATATCCCTCGCATTGTCCGCGAGGCGTTTT  
TCCTCGCGAAATCGGGTCGTCTTGGCCCCGGTTTTGATTGACGTTCCCTAAAGATAT  
ACAGCAACAGCTGGTTGTCCCCGATTGGGATCAGCCTATGAGATTGCCTGGATA  
CTTGTCTCGCTTGCCTAAACCACCAAACGAAATGCTCTTGGAGCAAATAATTAG  
GCTGATTTCTGAGTCCAAAAGCCAGTTCTGTATGTTGGGGGAGGGAGTTTACA

---

TTCAAGTGAGGAGTTGAGGCGTTTCGTAGAGCTTACAGGGATTCCCGTGGCCAG  
TACTTTAATGGGGCTTGGATCTTATCCTTGCTCAGATGAGCTCTCACTTCAAATG  
CTGGGAATGCACGGGACTGTGCACGCTAATTATGCTGTTGATAAGAGTGATCTT  
TTGCTTGCCTTTGGAGTTAGGTTTGATGACCGTGTGACTGGTAAATTAGAGGCTT  
TTGCTAGTCGAGCAAAGATTGTTACATTGATATTGATTCTGCTGAGATTGGGA  
AGAATAAGCAGCCCCATGTATCTGTTTGTGGGGATTTAAAGTTGGCTCTGCAGG  
GGATAAATTCAATGTTGGAGGAGAGGATGAGTGGTAAGCTTAAGCTGGATTTCT  
CGGGTTGGAGACAGGAGGTGTTGGAGCAAAGGCAAAGTTCCCCTTGAGTTAC  
AAGACGTTTGGGGATGCTATTCCTCCTCAGTATGCCATTCAGGTTCTTGATGAGT  
TGACTAATGGAAATGCTATAATTACTACTGGCGTTGGGCAGCACCAGATGTGGG  
CTGC

- **Zero\_Blunt\_pCR\_S4.3 [977 bp]**

TCCGGCCCCGGCGCCACCAATCTGGTGAGCGGACTCGCTGACGCGCTTCTGGAT  
AGCTGTCCATTGTGGCCATTACTGGACAGGTGCCTCGGAGGATGATTGGGACA  
GACGCTTTCAGGAAACCCCAATCGTTGAGGTAACCTCGATCGATTACGAAGCAT  
AATTATTTAGTTTTAGACGTAGAGGACATCCCTCGCATTGTTTCGTGAGGCGTTTT  
TCCTGGCCAAATCGGGCCGTCCTGGCCCGGTTTTGATTGATGTTCCCAAGGATAT  
ACAGCAACAATTGGTTATCCCCAATTGGGATCAGCCTATGAGATTGCCTGGATA  
CTTATCTCGTTTGCCAAAACCCCAAGTGATATGTTGTTGGAGCAAATAGTGAG  
GCTAATTTCTGAGTCAAAGAAGCCTGTCTTTATGTGGGGGGTGGGTCTTTACA  
GTCAAGTGAGGAGTTGAGGCGATTTGTGGAGCTAACTGGGATTCCTGTGGCTAG  
TACTTTAATGGGGCTTGGATCATATCCTTCTTCGGATGAACTTTCCTTCAAATG  
CTGGGAATGCATGGGACTGTGTATGCTAATTACTCTGTTGACAAGAGTGATCTG  
TTGCTTGCATTTGGGGTGAGGTTTCGATGACCGTGTGACTGGGAAGCTGGAGGCT  
TTTGCTAGTCGGGCGAAGATTGTTACATTGATATTGATTCCGCGGAGATTGGG  
AAGAATAAGCAGCCCCATGTATCGATTTGTGCGGATTTGAAGTTGGCCCTGCAG  
GGGCTGAATTCGATACTGGAGGAGAGGGTAGGTAAGCTGAAGTTGGATTTCTCG  
GCTTGGAGACAGGAGTTGAATGAGCAGAAGGAGAAGTTCCCGTTGGGTTATAA  
GACGTTTGGAGATGCCATTTCTCCCCAGTATGCTATTCAGGTTCTTGATGAGCTG  
ACGAACGGAAATGCTATAATTAGTACTGGTGTGGGCAGCACCAGATGTGGGCT  
GC

## A.5.2 Primer sets used for inserts

### Zero\_Blunt\_pCR\_S1.1

Forward primer [Ib130SP] (24 bp)	CTCACGCGCTCCAAGATGATCAGG
Reverse primer [Ib965ASP] (26 bp)	TGGGGCTGCTTATTCTTCCCAATCTC

### Zero\_Blunt\_pCR\_S1.2

Forward primer [Ib130SP] (24 bp)	CTCACGCGCTCCAAGATGATCAGG
Reverse primer [Ib965ASP] (26 bp)	TGGGGCTGCTTATTCTTCCCAATCTC

### Zero\_Blunt\_pCR\_S4.1

Forward primer [Ib224SP] (21 bp)	TCCGGCCCCGGCGCCACCAAT
Reverse primer [Ib1220ASP] (26 bp)	GCAGCCCACATCTGGTGCTGCCCAAC

### Zero\_Blunt\_pCR\_S4.3

Forward primer [Ib224SP] (21 bp)	TCCGGCCCCGGCGCCACCAAT
Reverse primer [Ib1220ASP] (26 bp)	GCAGCCCACATCTGGTGCTGCCCAAC

## A.6 Protein translation of correct open frame for each insert (ALS1 gene)

- Zero\_Blunt\_pCR\_S1.1 (5' to 3') – [273 aa]

MIRNVLPRHEQGGVFAAEGYARATGFPGVCIATSGPGATNLVSGGLADALLDSCPIV  
 AITGQVPRRMIGTDAFQETPIVEVTRSITKHNYLVLDVEDIPRIVREAFFLAKSGRPGP  
 VLIDVPKDIQQQLVIPNWDQPMRLPGYLSRLPKPPSDMLLEQIVRLISESKKPVLYVG  
 GGSLSSEELRRFVELTGIPVASTLMGLGSSYSSDELSLQMLGMHGTVYANYSVDK  
 SDLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKQP



---

- **Zero\_Blunt\_pCR\_S1.2 (5' to 3') – [273 aa]**

MIRNVLP RHEQGGVFAAEGYARATGFPGVCIATSGPGATNLV SGLADALLDSCPIV  
AITGQVPRRMIGTDAFQETPIVEVTRSITKHNYLVLDVEDIPRIVREAFFLAKSGRPGP  
VLIDVPKDIQQQLVIPNWDQPMRLPGYLSRLPKPPSDMLLEQIVRLISESKKPVLVYG  
GGSLSQSSEELRRFVELTGIPVASTLMGLGSYPSSDELSLQMLGMHGT VYANYSVDK  
SDLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKQP

- **Zero\_Blunt\_pCR\_S4.1 (5' to 3') – [294 aa]**

MIGTDAFQETPIVEVTRSITKHNYLVLNVEDIPRIVREAFFLAKSGRPGPVLIDVPKDI  
QQQLVVPDWDQPMRLPGYLSRLPKPPNEMLLEQIIRLISESKKPVLVYGGGSLHSSE  
ELRRFVELTGIPVASTLMGLGSYPCSDLSLQMLGMHGT VHANYAVDKSDLLAFG  
VRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKQPHVSVCGDLKLALQGINSMLEE  
RMSGKCLKLDFSGWRQEVLEQKAKFPLSYKTFGDAIPPQYAIQVLDEL TNGNAIITG  
VGQHQMWA

- **Zero\_Blunt\_pCR\_S4.3 (5' to 3') – [293 aa]**

MIGTDAFQETPIVEVTRSITKHNYLVLDVEDIPRIVREAFFLAKSGRPGPVLIDVPKDI  
QQQLVIPNWDQPMRLPGYLSRLPKPPSDMLLEQIVRLISESKKPVLVYGGGSLQSSE  
ELRRFVELTGIPVASTLMGLGSYPSSDELSLQMLGMHGT VYANYSVDKSDLLAFG  
VRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKQPHVSI CADLKLALQGLNSILEER  
VGKCLKLDFSAWRQELNEQKEKFPLGYKTFEDAISPQYAIQVLDEL TNGNAIISTGVG  
QHQMWA

## A.7 Amino acid table

Name	Abbr		Molecular Weight	Molecular Formula	Residue Formula	Residue Weight (-H <sub>2</sub> O)
Alanine	Ala	A	89.10	C <sub>3</sub> H <sub>7</sub> NO <sub>2</sub>	C <sub>3</sub> H <sub>5</sub> NO	71.08
Arginine	Arg	R	174.20	C <sub>6</sub> H <sub>14</sub> N <sub>4</sub> O <sub>2</sub>	C <sub>6</sub> H <sub>12</sub> N <sub>4</sub> O	156.19
Asparagine	Asn	N	132.12	C <sub>4</sub> H <sub>8</sub> N <sub>2</sub> O <sub>3</sub>	C <sub>4</sub> H <sub>6</sub> N <sub>2</sub> O <sub>2</sub>	114.11
Aspartic acid	Asp	D	133.11	C <sub>4</sub> H <sub>7</sub> NO <sub>4</sub>	C <sub>4</sub> H <sub>5</sub> NO <sub>3</sub>	115.09
Cysteine	Cys	C	121.16	C <sub>3</sub> H <sub>7</sub> NO <sub>2</sub> S	C <sub>3</sub> H <sub>5</sub> NOS	103.15
Glutamic acid	Glu	E	147.13	C <sub>5</sub> H <sub>9</sub> NO <sub>4</sub>	C <sub>5</sub> H <sub>7</sub> NO <sub>3</sub>	129.12
Glutamine	Gln	Q	146.15	C <sub>5</sub> H <sub>10</sub> N <sub>2</sub> O <sub>3</sub>	C <sub>5</sub> H <sub>8</sub> N <sub>2</sub> O <sub>2</sub>	128.13
Glycine	Gly	G	75.07	C <sub>2</sub> H <sub>5</sub> NO <sub>2</sub>	C <sub>2</sub> H <sub>3</sub> NO	57.05
Histidine	His	H	155.16	C <sub>6</sub> H <sub>9</sub> N <sub>3</sub> O <sub>2</sub>	C <sub>6</sub> H <sub>7</sub> N <sub>3</sub> O	137.14
Hydroxyproline	Hyp	O	131.13	C <sub>5</sub> H <sub>9</sub> NO <sub>3</sub>	C <sub>5</sub> H <sub>7</sub> NO <sub>2</sub>	113.11
Isoleucine	Ile	I	131.18	C <sub>6</sub> H <sub>13</sub> NO <sub>2</sub>	C <sub>6</sub> H <sub>11</sub> NO	113.16
Leucine	Leu	L	131.18	C <sub>6</sub> H <sub>13</sub> NO <sub>2</sub>	C <sub>6</sub> H <sub>11</sub> NO	113.16
Lysine	Lys	K	146.19	C <sub>6</sub> H <sub>14</sub> N <sub>2</sub> O <sub>2</sub>	C <sub>6</sub> H <sub>12</sub> N <sub>2</sub> O	128.18
Methionine	Met	M	149.21	C <sub>5</sub> H <sub>11</sub> NO <sub>2</sub> S	C <sub>5</sub> H <sub>9</sub> NOS	131.20
Phenylalanine	Phe	F	165.19	C <sub>9</sub> H <sub>11</sub> NO <sub>2</sub>	C <sub>9</sub> H <sub>9</sub> NO	147.18
Proline	Pro	P	115.13	C <sub>5</sub> H <sub>9</sub> NO <sub>2</sub>	C <sub>5</sub> H <sub>7</sub> NO	97.12
Pyroglutamic	Glp	U	139.11	C <sub>5</sub> H <sub>7</sub> NO <sub>3</sub>	C <sub>5</sub> H <sub>5</sub> NO <sub>2</sub>	121.09
Serine	Ser	S	105.09	C <sub>3</sub> H <sub>7</sub> NO <sub>3</sub>	C <sub>3</sub> H <sub>5</sub> NO <sub>2</sub>	87.08
Threonine	Thr	T	119.12	C <sub>4</sub> H <sub>9</sub> NO <sub>3</sub>	C <sub>4</sub> H <sub>7</sub> NO <sub>2</sub>	101.11
Tryptophan	Trp	W	204.23	C <sub>11</sub> H <sub>12</sub> N <sub>2</sub> O <sub>2</sub>	C <sub>11</sub> H <sub>10</sub> N <sub>2</sub> O	186.22
Tyrosine	Tyr	Y	181.19	C <sub>9</sub> H <sub>11</sub> NO <sub>3</sub>	C <sub>9</sub> H <sub>9</sub> NO <sub>2</sub>	163.18
Valine	Val	V	117.15	C <sub>5</sub> H <sub>11</sub> NO <sub>2</sub>	C <sub>5</sub> H <sub>9</sub> NO	99.13

## A.8 Tables

Table 10: Master mixes preparation for cycle sequencing reaction of ALS1 gene in five distinct plasmids DNA.

Components	Master mix 1 contains M13 Forward (-20)	Master mix 2 contains M13 Reverse
M13 Forward (-20) [2 µM]	9.6 µL	-
M13 Reverse [2 µM]	-	9.6 µL
BigDye buffer [5X]	12 µL	12 µL
BigDye Terminator v.3.1	3 µL	3 µL
Nuclease-free water	32.4 µL	32.4 µL

Table 11: Sample preparation for PCR\_IbALS1\_S1.1 plasmid DNA.

	+DNA -sgRNA -Cas9	+DNA -sgRNA +Cas9	+DNA +sgRNA +Cas9	+DNA +sgRNA -Cas9	+Con.DNA +sgRNA +Cas9	+DNA +sgRNA +Cas9	+DNA +sgRNA -Cas9	+Con.DNA +sgRNA +Cas9
Sample	1	2	3	4	5	6	7	8
DNA	2 µL [3 nM]	2 µL [3 nM]	2 µL [3 nM]	2 µL [3 nM]	2 µL [3 nM]	2 µL [3 nM]	2 µL [3 nM]	2 µL [3 nM]
sgRNA	-	-	2 µL [30 nM]*	2 µL [30 nM]*	2 µL [30 nM]*	2 µL [30 nM]**	2 µL [30 nM]**	2 µL [30 nM]**
Cas9	-	0.6 µL [30 nM]	0.6 µL [30 nM]	-	0.6 µL [30 nM]	0.6 µL [30 nM]	-	0.6 µL [30 nM]
NEB	2 µL [1X]	2 µL [1X]	2 µL [1X]	2 µL [1X]	2 µL [1X]	2 µL [1X]	2 µL [1X]	2 µL [1X]
H <sub>2</sub> O	16 µL	15.4 µL	13.4 µL	16 µL	13.4 µL	13.4 µL	14 µL	13.4 µL
Final volume	20 µL	20 µL	20 µL	20 µL	20 µL	20 µL	20 µL	20 µL

\*FP\_IVT\_T7\_Spa1\_S1.1 \*\*FP\_IVT\_T7\_Spa2\_S1.1

Table 12: Sample preparation for PCR\_IbALS1\_S4.1 plasmid DNA.

	+DNA +sgRNA +Cas9	+DNA -sgRNA +Cas9	+DNA +sgRNA -Cas9	+DNA +sgRNA +Cas9	+DNA +sgRNA -Cas9	+Con.DNA +sgRNA +Cas9	+DNA +sgRNA +Cas9	+Con.DNA +sgRNA +Cas9
Sample	1	2	3	4	5	6	7	8
DNA	2 µL [3 nM]	2 µL [3 nM]	2 µL [3 nM]	2 µL [3 nM]	2 µL [3 nM]	2 µL [3 nM]	2 µL [3 nM]	2 µL [3 nM]
sgRNA	2 µL [30 nM]*	-	2 µL [30 nM]*	2 µL [30 nM]**	2 µL [30 nM]**	2 µL [30 nM]***	2 µL [30 nM]***	2 µL [30 nM]*
Cas9	0.6 µL [30 nM]	0.6 µL [30 nM]	-	0.6 µL [30 nM]	-	0.6 µL [30 nM]	0.6 µL [30 nM]	0.6 µL [30 nM]
NEB	2 µL [1X]	2 µL [1X]	2 µL [1X]	2 µL [1X]	2 µL [1X]	2 µL [1X]	2 µL [1X]	2 µL [1X]
H <sub>2</sub> O	13.4 µL	15.4 µL	14 µL	13.4 µL	14 µL	13.4 µL	13.4 µL	13.4 µL
Final volume	20 µL	20 µL	20 µL	20 µL	20 µL	20 µL	20 µL	20 µL

\*FP\_IVT\_T7\_Spa1\_S4.1 \*\*FP\_IVT\_T7\_Spa2\_S4.1 \*\*\*FP\_IVT\_T7\_Spa3\_S4.1

Table 13: Experimental and control samples for PCR amplification of insert (CDS of Cas9 and NLS) were prepared as table below.

Components	Experimental	Control
DNA template (pCas9_TPC) [205 ng/ $\mu$ L]	0.5 $\mu$ L	-
Forward primer (Cas9_302_F) [10 $\mu$ M]	2.5 $\mu$ L	2.5 $\mu$ L
Reverse primer (Cas9_302_R) [10 $\mu$ M]	2.5 $\mu$ L	2.5 $\mu$ L
Phusion HF buffer [5X]	10 $\mu$ L	10 $\mu$ L
Phusion DNA polymerase [2 U/ $\mu$ L]	0.5 $\mu$ L	0.5 $\mu$ L
dNTPs [10 mM]	1 $\mu$ L	1 $\mu$ L
Nuclease-free water	33 $\mu$ L	33.5 $\mu$ L

Table 14: PCR cycling condition for PCR amplification of insert (CDS of Cas9 and NLS) was performed as table below.

Step	Temperature	Time	Number of cycles
Initial denaturation	98°C	30 seconds	1
Denaturation	98°C	10 seconds	
Annealing	64°C*	30 seconds	35
Extension	72°C	2 minutes	
Final extension	72°C	7 minutes	1
Hold	4°C	-	-

\*based on  $T_m$  of primers

Table 15: Experimental and control samples for PCR amplification of vector (pET302\_NT) were prepared as table below.

Components	Experimental	Control
DNA template (pET302_NT) [116 ng/ $\mu$ L]	1 $\mu$ L	-
Forward primer (302_5'His_F) [10 $\mu$ M]	2.5 $\mu$ L	2.5 $\mu$ L
Reverse primer (302_5'His_R) [10 $\mu$ M]	2.5 $\mu$ L	2.5 $\mu$ L
Phusion HF buffer [5X]	10 $\mu$ L	10 $\mu$ L
Phusion DNA polymerase [2 U/ $\mu$ L]	0.5 $\mu$ L	0.5 $\mu$ L
dNTPs [10 mM]	1 $\mu$ L	1 $\mu$ L
Nuclease-free water	32.5 $\mu$ L	33.5 $\mu$ L

Table 16: PCR cycling condition for PCR amplification of vector (pET302\_NT) was performed as table below.

Step	Temperature	Time	Number of cycles
Initial denaturation	98°C	30 seconds	1
Denaturation	98°C	10 seconds	
Annealing	60°C*	30 seconds	35
Extension	72°C	3 minutes	
Final extension	72°C	7 minutes	1
Hold	4°C	-	-

\*based on  $T_m$  of primers

## A.9 Materials

### A.9.1 Kits

A list of the kits used in this study can be found in the table below.

Table 17: The kits used for plasmid DNA isolation, in-vitro transcription (IVT) reaction of sgRNA, In-fusion cloning, and Isolation and purification of PCR products (gel extraction).

Kit	Purpose	Supplier
PureYield™ Plasmid Miniprep System	Plasmid DNA isolation	Promega
Guide-it™ sgRNA <i>In-vitro</i> Transcription and Screening System	<i>In-vitro</i> transcription (IVT) reaction of sgRNA	Takara
In-Fusion® HD Cloning	In-fusion cloning	Takara
NucleoSpin® Gel and PCR Clean-up	Isolation and purification of PCR products (gel extraction)	Macherey-Nagel

### A.9.2 Primers

A list of the primers used in this study can be found in the table below. Forward primers used in *In-vitro* transcription (IVT) reaction of sgRNA, were designed by CRISPOR online tool. Primer sets used in in-fusion cloning were designed by in-Fusion Cloning Primer Design Tool *v1.0* powered by Teselagen. All primer sequences are listed in Appendix [A.1](#).

Table 18: Primers used for cycle sequencing, in-vitro transcription (IVT) reaction of sgRNA, and in-fusion cloning.

Primer	Purpose
M13 Forward (-20) M13 Reverse	Cycle sequencing reaction of ALS1 gene in five distinct plasmids DNA
FP_IVT_T7_Spa1_S1.1 FP_IVT_T7_Spa2_S1.1 FP_IVT_T7_Spa1_S4.1 FP_IVT_T7_Spa2_S4.1 FP_IVT_T7_Spa3_S4.1	In-vitro transcription (IVT) reaction of sgRNA
302_5'His_F 302_5'His_R Cas9_302_F* Cas9_302_R*	Primer designing for in-fusion cloning / *Verification by colony PCR for in-fusion cloning

### A.9.3 Plasmids DNA

Five different plasmids DNA containing ALS1 fragments that cloned in Zero Blunt Vector (3.5 kb) provided by my main supervisor, Professor Wenche Johansen, from *Ipomoea batatas* cultivars were used in cycle sequencing reaction of ALS1 gene, and two of them (\*marked in table 19) were used in *in-vitro* cleavage assay. Plasmids DNA sequences are listed in Appendix A.1.2.

Table 19: Plasmids for cycle sequencing reaction of ALS1 gene in five distinct plasmids DNA, and in-vitro cleavage assay.

Plasmid DNA	Purpose
PCR_IbALS1_S1.1* PCR_IbALS1_S1.2 PCR_IbALS1_S4.1* PCR_IbALS1_S4.2 PCR_IbALS1_S4.3	Cycle sequencing reaction of ALS1 gene in five distinct plasmids DNA / * <i>In-vitro</i> cleavage assay

### A.9.4 Vectors

Different vectors were used in this study; Zero Blunt Vector was used for

transformation, pET302\_NT vector was used as a destination vector for constructing an expression vector for *in-vitro* recombinant Cas9 production. CDS of Cas9 and NLS can be found in pCas9\_TPC vector. Vector sequences and maps are listed in Appendix A.3.

Table 10: Vectors used for transformation and Cas9 expression.

Vector	Purpose
Zero Blunt Vector (3512 bp)	Transformation of <i>E. coli</i> TOP10 cells with ALS1 insert fragments
pCas9_TPC (14017 bp) pET302_NT (5712 bp) pET302_His_Cas9_NLS (9849 bp)	Construct expression vector for <i>in-vitro</i> recombinant Cas9 production using CLC software

## A.9.5 Chemicals

A list of the chemicals used in this study can be found in the table below.

Table 11: Different buffers and chemicals were used for cycle sequencing, extension products purification, transformation, plasmid DNA isolation, culturing of putative positive transformants, restriction digest analysis, *in-vitro* cleavage assay, and colony PCR.

Chemical	Purpose
BigDye buffer [5X] BigDye Terminator v.3.1	Cycle sequencing reaction of ALS1 gene in five distinct plasmids DNA
EDTA [125 mM] NaOAc [3 M] [pH 5.2] 96% EtOH 70% EtOH Deionized formamide	Purification of extension products
LB agar medium (pH 7.0) SOC medium	Transformation of <i>E. coli</i> TOP10 cells with ALS1 insert fragments / Transformation of stellar competent cells with in-fusion cloning mixture
LB broth medium (pH 7.0)	Culturing of putative positive transformants
TE buffer	Plasmid DNA isolation
NEB Buffer 3.1 [10X] 1% agarose gel	Restriction digest analysis of putative positive plasmid DNA

SpCas9 [30 nM, 60 nM]	<i>In-vitro</i> cleavage assay
NEB buffer [1X]	
Buffer B1 [1X]	Verification by colony PCR for in-fusion cloning
MgCl <sub>2</sub> [1.5 mM]	
dNTPs [200 µL]	
HOT firepol [2.5 U/µL]	
rcut smart buffer	Restriction digest analysis of putative positive plasmid DNA for in-fusion cloning

## A.9.6 Bacteria

This study employed two different bacterial cells for different transformations.

Table 12: Bacteria for transformation.

Bacteria	Purpose	Supplier
<i>E. coli</i> TOP10 cells	Transformation of <i>E. coli</i> TOP10 cells with ALS1 insert fragments	Invitrogen
Stellar component cells	Transformation of stellar competent cells with in-fusion cloning mixture	Invitrogen

## A.9.7 Restriction enzymes

A list of the restriction enzymes used in this study can be found in the table below.

Table 13: Restriction enzymes used for restriction digest analysis of putative positive plasmid DNA and in-vitro cleavage assay.

Enzyme	Purpose	Supplier
<i>EcoRV</i> restriction enzyme	Restriction digest analysis of putative positive plasmids DNA	New England Biolabs
<i>MluI</i> restriction enzyme	<i>In-vitro</i> cleavage assay	New England Biolabs
<i>NcoI</i> restriction enzyme		
<i>XhoI</i> restriction enzyme	Restriction digest analysis of putative positive pET302_His_Cas9-NLS	New England Biolabs



---

## A.9.8 Antibiotics

A list of the antibiotics used in this study can be found in the table below.

Table 14: Antibiotics used for different transformation.

<b>Antibiotic</b>	<b>Purpose</b>
Kanamycin [50 µg/mL]	Transformation of <i>E. coli</i> TOP10 cells with ALS1 insert fragments
Ampicillin [100 µg/mL]	Transformation of stellar competent cells with in-fusion cloning mixture